

THE UNIVERSITY OF CHICAGO

EXPLORING THE GENETIC EPIDEMIOLOGY OF BREAST CANCER AND THE  
MOLECULAR EPIDEMIOLOGY OF ARSENIC TOXICITY

A DISSERTATION SUBMITTED TO  
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES  
AND THE PRITZKER SCHOOL OF MEDICINE  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

INTERDISCIPLINARY SCIENTIST TRAINING PROGRAM:  
PUBLIC HEALTH SCIENCES

BY  
JAMES LU LI

CHICAGO, ILLINOIS

JUNE 2025

Copyright © 2025 by James Li

All rights reserved

*To my partner Yijia Sun, my parents Xinwei Li and Feifei Lu, and my younger brother Michael Li, for their endless support, encouragement, and unwavering belief in me.*

## TABLE OF CONTENTS

|                                                                                                                                                                                               |             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>LIST OF FIGURES .....</b>                                                                                                                                                                  | <b>V</b>    |
| <b>LIST OF TABLES.....</b>                                                                                                                                                                    | <b>IX</b>   |
| <b>ACKNOWLEDGMENTS .....</b>                                                                                                                                                                  | <b>XII</b>  |
| <b>ABSTRACT.....</b>                                                                                                                                                                          | <b>XV</b>   |
| <b>INTRODUCTION.....</b>                                                                                                                                                                      | <b>XVII</b> |
| <b>1 CHAPTER 1: DISSECTING THE GENETIC ARCHITECTURE OF BREAST<br/>CANCER INTRINSIC SUBTYPES THROUGH CROSS-ANCESTRY GENOME- AND<br/>TRANSCRIPTOME-WIDE ASSOCIATION STUDIES .....</b>           | <b>1</b>    |
| 1.1 ABSTRACT .....                                                                                                                                                                            | 1           |
| 1.2 INTRODUCTION.....                                                                                                                                                                         | 2           |
| 1.3 RESULTS .....                                                                                                                                                                             | 3           |
| 1.4 DISCUSSION.....                                                                                                                                                                           | 11          |
| 1.5 METHODS.....                                                                                                                                                                              | 15          |
| 1.6 APPENDIX .....                                                                                                                                                                            | 23          |
| <b>2 CHAPTER 2: IMPROVED POLYGENIC RISK SCORE PREDICTION OF BREAST<br/>CANCER RISK IN WOMEN OF AFRICAN ANCESTRY LEVERAGING<br/>INFORMATION ACROSS ANCESTRIES, SUBTYPES, AND METHODS .....</b> | <b>119</b>  |
| 2.1 ABSTRACT .....                                                                                                                                                                            | 119         |
| 2.2 INTRODUCTION.....                                                                                                                                                                         | 119         |
| 2.3 RESULTS .....                                                                                                                                                                             | 121         |
| 2.4 DISCUSSION.....                                                                                                                                                                           | 129         |
| 2.5 METHODS.....                                                                                                                                                                              | 133         |
| 2.6 APPENDIX .....                                                                                                                                                                            | 143         |
| <b>3 CHAPTER 3: THE IMPACT OF ARSENIC EXPOSURE ON DNA METHYLATION<br/>IN HUMANS: BUILDING AN EPIGENETIC BIOMARKER OF EXPOSURE AND<br/>TOXICITY RISK.....</b>                                  | <b>174</b>  |
| 3.1 ABSTRACT .....                                                                                                                                                                            | 174         |
| 3.2 INTRODUCTION.....                                                                                                                                                                         | 175         |
| 3.3 METHODS.....                                                                                                                                                                              | 176         |
| 3.4 RESULTS .....                                                                                                                                                                             | 182         |
| 3.5 DISCUSSION.....                                                                                                                                                                           | 188         |
| 3.6 APPENDIX .....                                                                                                                                                                            | 192         |
| <b>4 SUMMARY AND FUTURE DIRECTIONS .....</b>                                                                                                                                                  | <b>305</b>  |
| <b>5 REFERENCES.....</b>                                                                                                                                                                      | <b>307</b>  |



## LIST OF FIGURES

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1.1 Ideogram of GWAS loci associated with each intrinsic breast cancer subtype, including both novel and previously reported loci identified in this study .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 23  |
| Figure 1.2 Risk allele frequencies of all lead variants identified across GWAS analyses for each subtype by population. Note: Risk allele frequencies were derived from control samples. One-sided binomial test p-values are shown for the proportion of lead variants with higher risk allele frequencies in African-ancestry compared to European-ancestry populations. ....                                                                                                                                                                                                                                                                                                                                                                                                                                               | 24  |
| Figure 2.1 Workflow for constructing and validating polygenic risk score models. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 143 |
| Figure 2.2 Developing a PRS model for overall breast cancer tuned by proportions of each subtype. Note: The observed proportion of ER+ cases in the training set was 0.67. Gray error bars indicate the performance (95% CI of the covariate-adjusted AUC) of two-subtype proportion-tuned PRS models developed by combining only ER-positive and ER-negative models by various assigned proportions of ER-positive cases. The navy-blue error bar indicates the performance of the three-subtype proportion-tuned PRS model developed by combining ER-positive, ER-negative, and TNBC PRS models by their observed proportions in the validation set. The turquoise error bar indicates the performance of the three-subtype proportion-tuned PRS model only in cases with complete subtype data. ....                       | 144 |
| Figure 2.3 Estimated lifetime and 10-year absolute risks of developing each breast cancer subtype among women with African ancestry by percentiles of the polygenic risk scores.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 145 |
| Figure 2.4 Density plots for the lifetime risk of developing each subtype of breast cancer between African Americans and European Americans. Red indicates the estimated theoretical lifetime risk of the best performing PRS model for each subtype among African Americans and blue indicates the estimated theoretical lifetime risk of the European 313-/330-variant PRSs for each subtype among European Americans. A) Overall breast cancer B) ER-positive breast cancer C) ER-negative breast cancer D) Triple-negative breast cancer. ....                                                                                                                                                                                                                                                                            | 146 |
| Figure 3.1 Epigenome-wide study of the association between urinary arsenic and whole blood DNAm in HEALS. <b>(A)</b> . Manhattan plot showing P-values for the association of urinary arsenic with DNAm in HEALS. Blue dots represent CpG sites where As exposure is associated with increased DNAm. Red dots represent CpGs sites where As exposure is associated with decreased DNAm. Dashed line represents FDR threshold ( $p=8.12E-05$ ). Solid line represents Bonferroni threshold ( $p=6.93E-08$ ). <b>(B)</b> . As-associated CpGs were enriched for in open seas and depleted in CpG islands. <b>(C)</b> . As-associated CpGs were enriched for in enhancer regions, based on chromatin segmentation categories, across several leukocyte cell types and depleted in regions involved in active transcription. .... | 192 |
| Figure 3.2 Regional plots of several regions containing CpGs showing association with urinary arsenic levels. In each panel, P-values (above) and mean DNAm levels (bottom) are shown above the gene map. FDR-significant CpGs are labeled. Genomic annotations were obtained from the UCSC genome browser. <b>(A)</b> . ATP1B3 region. <b>(B)</b> . AC009518.4 region. <b>(C)</b> . PNMA8B region. <b>(D)</b> . CCDC105 region.....                                                                                                                                                                                                                                                                                                                                                                                          | 193 |

Figure 3.3 Association of arsenic metabolism efficiency (reflecting internal dose) with DNAm at arsenic-associated CpGs. The 116 CpGs plotted represent a high confidence set of As-associated CpGs that had a Bonferroni-adjusted  $p < 0.05$  in their association with urinary As and were also identified to associate with arsenic levels in well water ( $FDR < 0.05$ ). **(A)**. Beta values and confidence intervals for the association of measured arsenic metabolism efficiency (DMA%) with DNAm at As-associated CpGs. >97% of the associations observed are consistent with estimates from EWAS. **(B)**. Beta values and confidence intervals for the association between genetically-determined arsenic metabolism efficiency with DNAm at As-associated CpGs. >87% of CpGs show association that is directionally consistent with estimates from EWAS..... 194

Figure 3.4 Validating DNAm-based biomarkers of As exposure. **(A)**. Association between measured (log2-transformed creatinine-adjusted) and predicted urinary As scores in an independent subset in the Health Effects of Arsenic Longitudinal Study (HEALS). **(B)**. Association between measured (log2-transformed creatinine-adjusted) and predicted urinary As scores among participants in the Bangladesh Vitamin E and Selenium Trial (BEST). **(C)**. Association between measured (log2-transformed creatinine-adjusted) and predicted urinary As scores among participants in the Strong Heart Study (SHS). **(D)**. Associations between urinary As scores predicted by the 255-CpG DNAm-based biomarker and arsenical skin lesion status in HEALS. **(E)**. Associations between urinary As scores predicted by the 255-CpG DNAm-based biomarker and survival outcomes in the SHS. Model #1 adjusts for covariates including age, study center, sex, smoking status, body mass index, estimated glomerular filtration rate. Model #2 controls for potential mediators in addition to Model #1 covariates including LDL cholesterol, HDL cholesterol, systolic blood pressure, diabetes status, hypertension history, albuminuria. . 195

|                                                                                                                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure S 1.1 Ideogram of all GWAS loci identified to associate with intrinsic subtypes in this study by ancestry.....                                                                                                        | 30 |
| Figure S 1.2 QQ-plots and inflation factors for the African-ancestry GWAS meta-analysis for each intrinsic subtype. A) Luminal-A-like, B) Luminal-B-like, C) HER2-enriched-like, and D) Triple-negative breast cancer. ....  | 31 |
| Figure S 1.3 QQ-plots and inflation factors for the European-ancestry GWAS meta-analysis for each intrinsic subtype. A) Luminal-A-like, B) Luminal-B-like, C) HER2-enriched-like, and D) Triple-negative breast cancer. .... | 32 |
| Figure S 1.4 QQ-plots and inflation factors for the cross-ancestry GWAS meta-analysis for each intrinsic subtype. A) Luminal-A-like, B) Luminal-B-like, C) HER2-enriched-like, and D) Triple-negative breast cancer. ....    | 33 |
| Figure S 1.5A Regional plot of GWAS associations at the 1q41 locus (lead variant, rs4011894) for Luminal-A-like breast cancer. ....                                                                                          | 34 |
| Figure S 1.6 Predictive performance for gene expression models built in this study for TWAS analyses .....                                                                                                                   | 47 |

|                                                                                                                                                                                                                                                                        |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure S 1.7 Correlation of mean association estimates (Z-scores) for genes identified in both the current study and previous breast tissue TWAS of intrinsic subtypes .....                                                                                           | 48  |
| Figure S 1.8 Ideogram of GWAS loci color-coded by loci without a nearby TWAS signal, loci with a nearby TWAS signal identified in the present study, and loci with a TWAS signal identified in prior TWAS of intrinsic subtypes. ....                                  | 49  |
| Figure S 2.1 Performance of single-ancestry PRS models of overall breast cancer risk in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                               | 149 |
| Figure S 2.2 Performance of single-ancestry PRS models of ER-positive BC in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                                           | 150 |
| Figure S 2.3 Performance of single-ancestry PRS models of ER-negative BC in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                                           | 151 |
| Figure S 2.4 Performance of single-ancestry PRS models of triple-negative breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                            | 152 |
| Figure S 2.5 Performance of cross-ancestry (xAncestry) PRS models of overall breast cancer risk in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                    | 153 |
| Figure S 2.6 Performance of cross-ancestry (xAncestry) PRS models of ER-positive breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                     | 154 |
| Figure S 2.7 Performance of cross-ancestry (xAncestry) PRS models of ER-negative breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                     | 155 |
| Figure S 2.8 Performance of cross-ancestry (xAncestry) PRS models of triple negative breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                 | 156 |
| Figure S 2.9 Performance of ensemble-based PRS models of breast cancer by subtype. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC. ....                                                                                        | 157 |
| Figure S 2.10 Observed odds ratio for each breast cancer subtype by polygenic risk score (PRS) percentile compared to the middle 40-60th percentile of PRS scores. Note: Error bars represent the 95% confidence interval for each covariate-adjusted odds ratio. .... | 158 |
| Figure S 2.11 Density plots for the lifetime absolute risks of developing each subtype of breast cancer by first-degree family history status. ....                                                                                                                    | 159 |
| Figure S 2.12 Density plots for the 10-year absolute risks of developing each subtype of breast cancer by first-degree family history status. ....                                                                                                                     | 160 |
| Figure S 2.13 Scatterplots of principal components 1 and 2. A) Color-coded by proportion of African ancestry B) Color-coded by continent where each participant resides. ....                                                                                          | 161 |

|                                                                                                                                                                                                                                                                                            |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure S 3.1 QQ-plot for the epigenome-wide association study of log2-transformed creatinine-adjusted urinary arsenic. ....                                                                                                                                                                | 197 |
| Figure S 3.2 Consistency in arsenic measurements between total urinary arsenic obtained from spot urine samples (utilized in the primary EWAS) and A) the sum of inorganic and methylated arsenic species and B) arsenic levels in well water.....                                         | 198 |
| Figure S 3.3 Consistency in effect sizes of the 1,177 FDR-significant identified in the primary EWAS between primary EWAS and A) EWAS of the sum of inorganic and methylated As species and B) EWAS of water As. ....                                                                      | 199 |
| Figure S 3.4 Consistency of EWAS associations between urinary arsenic and DNA methylation (beta values) for the 20 CpGs identified in both the previous EWAS of a HEALS subcohort (n=396) and the current EWAS of a larger HEALS subcohort including the previous subcohort (n=1,186)..... | 200 |
| Figure S 3.5 Forest plot of the associations between epigenetic age acceleration predicted by various DNAm clocks and log2-transformed creatinine-adjusted urinary arsenic levels in HEALS.....                                                                                            | 201 |
| Figure S 3.6 Workflow for developing and validating DNAm-based biomarkers of arsenic exposure. ....                                                                                                                                                                                        | 202 |

## LIST OF TABLES

|                                                                                                                                                                                                                                |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1.1 Novel GWAS loci for breast cancer intrinsic subtypes identified in African-ancestry and cross-ancestry meta-analyses (Bonferroni-adjusted p-value<0.05) .....                                                        | 25  |
| Table 1.2 Heterogeneity in associations of novel intrinsic subtype lead variants across other breast cancer subtypes.....                                                                                                      | 26  |
| Table 1.3 Novel TWAS genes for breast cancer intrinsic subtypes identified in cross-ancestry cell type-aware TWAS analyses (Bonferroni-adjusted p-value<0.05) .....                                                            | 27  |
| Table 1.4 TWAS genes with signals that were not explained by GWAS index variants .....                                                                                                                                         | 29  |
| Table 2.1 Predictive performance of the best performing PRS models for each breast cancer phenotype in the validation set. Note: The 95% CI of each metric is presented in parentheses.                                        | 147 |
| Table 2.2 Association of PRS scores and family history with each breast cancer subtype. Note: PRS score is per SD; Family history is first-degree relatedness. The 95% CI of each odds ratio is presented in parentheses. .... | 148 |
| Table 3.1 Participant characteristics at baseline for a subcohort of HEALS participants [median (25th, 75th percentiles) or n (%)] compared to all HEALS participants. ....                                                    | 196 |
|                                                                                                                                                                                                                                |     |
| Table S 1.1 Studies in AABCG included in the African-ancestry GWAS meta-analysis performed in this study. ....                                                                                                                 | 50  |
| Table S 1.2 All lead variants identified for each intrinsic subtype in European-ancestry meta-analysis (Bonferroni-adjusted p-value<0.05) .....                                                                                | 52  |
| Table S 1.3 All lead variants identified for each intrinsic subtype in the African-ancestry meta-analysis (Bonferroni-adjusted p-value<0.05) .....                                                                             | 56  |
| Table S 1.4 All lead variants identified for each intrinsic subtype in cross-ancestry meta-analysis amongst variants present in both ancestry-specific meta-analyses (Bonferroni-adjusted p-value<0.05) .....                  | 57  |
| Table S 1.5 Validation of triple-negative variants identified previously in AABCG using logistic regression. ....                                                                                                              | 61  |
| Table S 1.6 Consistency of novel lead variants across sequencing platforms in AABCG .....                                                                                                                                      | 62  |
| Table S 1.7 Conditionally independent signals nearby novel GWAS loci identified for breast cancer intrinsic subtypes in this study .....                                                                                       | 65  |
| Table S 1.8 Risk allele frequencies for all lead variants identified for each subtype in this study. ....                                                                                                                      | 66  |
| Table S 1.9 All TWAS genes identified to associate with breast cancer intrinsic subtypes across ancestries and cell types (Bonferroni-adjusted p-value<0.05).....                                                              | 74  |
| Table S 1.10 Number of unique TWAS genes identified by subtype and cell type. ....                                                                                                                                             | 81  |
| Table S 1.11 Ancestry-stratified TWAS associations for the 90 genes identified to associate with intrinsic subtypes (Bonferroni-adjusted p-value<0.05) .....                                                                   | 82  |
| Table S 1.12 Genes previously reported in TWAS conducted across all breast cancer phenotypes including overall breast cancer and various subtypes. ....                                                                        | 95  |

|                                                                                                                                                                                                                    |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table S 1.13 TWAS associations for the closest gene by distance to each novel GWAS loci identified for intrinsic subtypes. ....                                                                                    | 117 |
| Table S 1.14 Suggestive TWAS associations linked to novel GWAS loci identified for intrinsic subtypes.....                                                                                                         | 118 |
| Table S 2.1 Studies from AABCG included in development of polygenic risk score models for breast cancer in women of African ancestry. ....                                                                         | 162 |
| Table S 2.2 Performance of the European 330-variant PRS models when modifying coefficients based on the AABCG training set. ....                                                                                   | 163 |
| Table S 2.3 Association of PRS scores and African ancestry with each breast cancer subtype .                                                                                                                       | 164 |
| Table S 2.4 Association of PRS scores and age with each breast cancer phenotype .....                                                                                                                              | 165 |
| Table S 2.5 Lifetime absolute risks for the best performing PRS models for each breast cancer phenotype by family history.....                                                                                     | 166 |
| Table S 2.6 10-year absolute risks at age 40 for the best performing PRS models for each breast cancer phenotype by family history. ....                                                                           | 167 |
| Table S 2.7 Proportion of genome-wide heritability explained by each PRS model for each breast cancer phenotype .....                                                                                              | 168 |
| Table S 2.8 External validation of PRS models for overall breast cancer .....                                                                                                                                      | 169 |
| Table S 2.9 External validation of PRS models for breast cancer subtypes among additional women of African ancestry from WFBC, BioVU, and BEST. ....                                                               | 171 |
| Table S 3.1 P-value matrices for the correlations between the first 20 surrogate variables (SVs) and inferred cell types (generated using EpiDISH) for the EWASs of urinary arsenic and arsenic in well water..... | 203 |
| Table S 3.2 CpGs associated with log2-transformed creatinine-adjusted urinary arsenic (FDR < 0.05) .....                                                                                                           | 204 |
| Table S 3.3 List of CpGs previously identified in whole blood to associate with urinary arsenic levels in adults after accounting for multiple testing.....                                                        | 235 |
| Table S 3.4 Inflation factors for the EWAS approaches for assessing the association between DNAm and log2-transformed creatinine-adjusted urinary arsenic levels.....                                              | 238 |
| Table S 3.5 Enrichment of As-associated CpGs in DNaseI hypersensitivity sites and EWASs of human health and disease.....                                                                                           | 239 |
| Table S 3.6 Hits for As-associated CpGs in EWASs of other health conditions and diseases in the MRC-IEU EWAS Catalog .....                                                                                         | 240 |
| Table S 3.7 Gene set and pathway enrichment analysis .....                                                                                                                                                         | 270 |
| Table S 3.8 Enrichment of As-associated CpGs in regions in relation to UCSC CpG islands for As-associated CpGs located in regions with multiple vs. one As-associated CpG .....                                    | 290 |
| Table S 3.9 CpGs involved in arsenic and cell type interactions (AxCT) .....                                                                                                                                       | 291 |
| Table S 3.10 List of CpGs and effect sizes for the 255-CpG and 117-CpG biomarkers of urinary arsenic levels developed using the Illumina EPIC and 450k array, respectively. ....                                   | 293 |

|                                                                                                                                                                                                         |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table S 3.11 Participant characteristics of the two study populations used for the external validation of the DNAm-based biomarkers of Arsenic exposure [median (25th, 75th percentiles) or n (%)]..... | 304 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|

## ACKNOWLEDGMENTS

I am filled with immense gratitude even as I just begin to think about everyone who has made this journey possible for me. It is humbling to think about how I hadn't even taken a biology class in high school yet here I am today, having received this opportunity to graduate with a degree in genetic epidemiology. My ambition towards becoming an independent researcher and physician scientist is really a privilege that I do not take for granted – none of this would have been possible without the encouragement, belief, and patience that you all have given me.

I would like to first and foremost thank my advisors Dr. Dezheng Huo and Dr. Brandon Pierce. I am so lucky to have had this opportunity to be mentored by both of you and cannot express enough how much I appreciate all of the effort you put into me from weekly meetings to lengthy discussions. I feel incredibly blessed that you allowed me the freedom to propose and pursue ideas that I was interested in – someday I still can't believe I actually got the chance to explore all these ideas. Your mentorship has truly helped shape me not only as a scientist and independent researcher, but also as a person, and I feel so grateful that you both agreed to take me on as your student.

I would also like to thank my committee members Dr. Lin Chen and Dr. Haoyu Zhang. Dr. Chen, I am beyond grateful for all of your guidance during my graduate studies and the energy you spent teaching me to think critically about statistical approaches to take in my projects; I hope to maintain this rigor you have cultivated in me as I continue in my career. Haoyu, I cannot thank you enough for all the time you devoted towards mentoring me in our bi-weekly meetings and the pre-confluence project; I am immensely grateful for your guidance, giving me such a neat



opportunity to learn from you, and also for training me to rigorously approach and explore questions about the genetic architecture of human diseases. I would also like to thank Dr. Guimin Gao for his mentorship and guidance on utilizing TWAS approaches to identify susceptibility genes for human diseases.

Additionally, I would not be here without the amazing mentorship and training I received as an undergraduate, and I wanted to give a huge shout out to Professor Peter White – who was my undergraduate PI, Dr. Jeff Gaither, and Benjamin Kelly for all of their support as my mentors both in lab and in life. I would also like to thank Dr. Biao Ding who was my PI during my freshman year of undergrad as well as Dr. Jian Wu who at the time was my graduate student mentor. To Dr. Biao Ding, I had promised I would do an MD/PhD, and now I'm halfway there! I am so spoiled to have had all of you as my mentors, and I hope to continue demonstrating that the time and effort you put into mentoring me is something I deeply cherish.

I would also like to thank the MSTP office for all of the support they've provided me throughout both my med and grad school years and wanted to express my gratitude to Dr. Marcus Clark, Alison Anastasio, Haley Burkhardt, Nija Givens, Kristin McCann, and Marisa Davis. I would like to thank the Cancer Health Equity Training Program for the opportunities they have given me to present my research and to hear unique perspectives throughout the community, and I wanted to thank Dr. Eileen Dolan, Rosie Huggins, and Michelle Domecki for all they do to drive this program.

I am also so grateful to have amazing friends who have conveniently also been my lab mates. I wanted to give a huge thank you to Niyati Jain, Bei Wang, Lizeth Tamayo, Lin Tong, Julian McClellan, and Fangyuan Zhao. Niyati and Bei, I am honored to be in both of your cohorts, and it has truly been a blast doing this PhD with you guys. Julian, thank you not only for being a bro I could talk to about life outside of lab, but also for all of your help in developing my computational skills. Lizeth, thank you for accepting me into the lab so graciously and for being my role model. Fangyuan, thank you for being my first contact in PHS and for welcoming me into the lab. Lin, thank you for being so incredibly nice and supportive to me throughout my time in grad school and for helping me navigate all my projects.

I would also like to thank my brothers in all but blood, Jonathan Huang and Jason Cheung. Both of you have seen me in my lowest of lows and highest of highs. I feel incredibly fortunate to have met both of you and got the chance to spend these years in Chicago with you. Thank you for always being there for me. I would also like to thank Lingyang Zhang, Zoey Zhou, Niu Qun, Tongtong, Zhen Ni, Jingwen Zhang, and Youzi. You all have been family to me and I cannot express enough how immensely appreciative I am for all your support. I am grateful beyond.

Lastly, I would like to thank my partner, Yijia Sun, and my parents and my younger brother for their endless love and support of me to pursue my dreams. To my parents and younger brother, thank you for putting up with me and supporting me unconditionally throughout the years. To Yijia, I am truly the luckiest person to have you by my side during this whole journey throughout graduate school and life. I could not have done this without you.

## ABSTRACT

Despite improvements in medical treatments, the global burden of cancer has steadily increased over time. Breast cancer remains the most frequently diagnosed malignancy in women worldwide. Disparities in the incidence and mortality of breast cancer tend to be pronounced among women of African ancestry, driven in part by a higher prevalence of aggressive subtypes and a general lack of effective screening tools. To improve the assessment of cancer risk in diverse populations, as well as our understanding of the underlying etiology of these diseases, it is paramount to study both genetic and environmental factors that contribute to the development of cancer. This dissertation examines two distinct but complementary research foci: the genetic epidemiology of breast cancer and the molecular epidemiology of environmental carcinogen exposure. In Chapter 1, we characterize the genetic architecture of breast cancer by conducting a cross-ancestry genome-wide association study of intrinsic subtypes, along with a cell type-aware transcriptome-wide association study, to identify susceptibility loci and genes for each subtype. Following this, in Chapter 2, we develop and validate polygenic risk score (PRS) models for several breast cancer phenotypes including overall, estrogen receptor-positive, estrogen receptor-negative, and triple-negative breast cancer among African-ancestry women, leveraging information across ancestries, subtypes, and PRS methodologies to improve predictive performance. Chapter 3 then shifts to assessing the molecular epidemiology of arsenic (As) exposure, a well-established carcinogen affecting millions of individuals worldwide. Utilizing DNA methylation (DNAm) data from Bangladeshi adults who are chronically exposed to a wide range of As levels, we identify CpG sites associated with As exposure and then apply Mendelian Randomization to demonstrate that genetically predicted As metabolism efficiency causally impacts DNAm at these sites. We then develop a DNAm-based biomarker of As exposure that

robustly predicts exposures across multiple diverse populations. This biomarker also accurately predicts As-associated toxicities including arsenical skin lesions with high sensitivity and specificity, and associates with overall mortality even in an external population in the United States with lower levels of As exposure. Together, these three chapters demonstrate that leveraging both genetic and epigenetic data can improve our ability to characterize disease risk and improve risk-stratification at the population-level.

## INTRODUCTION

Cancer currently ranks as the second leading cause of death in the United States, following heart disease, and is the leading cause of death among individuals younger than 85 years of age.<sup>1</sup> As recent as 2022, there were an estimated 20 million new cancer cases and approximately 9.7 million cancer-related deaths worldwide.<sup>2</sup> Cancers are known to impact nearly all organ systems in the human body and are characterized by uncontrolled proliferation and spread of abnormal cells, driven by genetic alterations.<sup>3</sup> These genetic alterations can arise from several factors including inherited mutations, exposure to carcinogens such as radiation, environmental chemicals, and infectious organisms, or spontaneous mutations.<sup>4</sup> Given the broad range of biological and environmental factors that contribute to cancer risk, it is crucial that not only genetic factors but also the molecular consequences of environmental exposures are explored to enhance our understanding of cancer etiology, and ultimately, improve prevention efforts.

Breast cancer is the most frequently occurring malignancy among women globally, impacting over 2 million individuals every year.<sup>5</sup> This incidence rate of this cancer is estimated to be 40% higher in African American women compared to that in European American women.<sup>6,7</sup> Multiple factors contribute to this disparity, including unequal access to healthcare, more frequent diagnoses at later stages, and a higher prevalence of aggressive breast cancer subtypes – such as triple-negative and estrogen receptor-negative breast cancers – which are linked to worse survival outcomes and shorter disease latency.<sup>8,9</sup> This unequal cancer burden is especially noticeable among women in Sub-Saharan Africa,<sup>10</sup> where breast cancer mortality ranks among the highest worldwide.<sup>11</sup> As such, it is imperative to enhance the

accuracy of existing risk-assessment tools to more effectively stratify individuals for targeted screening, ultimately to facilitate earlier detection and improved outcomes.

Due to the heterogeneity in clinical features observed in breast cancer patients, the expression of hormone and HER2 receptors have been commonly used to classify breast cancers into four receptor-defined intrinsic subtypes including Luminal-A-like, Luminal-B-like, HER2-enriched-like, and triple-negative breast cancer.<sup>12,13</sup> To explore the etiology of these subtypes, several genome-wide association studies (GWAS) have been performed in European-ancestry populations to identify susceptibility loci associated with each subtype.<sup>14,15</sup> Furthermore, transcriptome-wide association studies (TWAS), which evaluate the association between genetically predicted gene expression and disease risk have complemented GWAS approaches to identify susceptibility genes for intrinsic subtypes;<sup>14,15</sup> however, similar to previous GWAS, these TWAS have largely been limited to European-ancestry populations.<sup>14,15</sup> To better understand the genetic architecture underlying the intrinsic subtypes in non-European populations – particularly for triple-negative breast cancer – cross-ancestry GWAS and TWAS of intrinsic subtypes are needed.

Additionally, while genetic susceptibility plays a critical role in the development of cancer, environmental exposures also substantially impact cancer risk, with many carcinogens that are abundant in our environment. Among environmental carcinogens, inorganic arsenic (As) is recognizable for its widespread impacts across diverse populations, as over 200 million people globally are exposed to As through drinking water,<sup>16</sup> with over 56 million exposed in Bangladesh.<sup>17</sup> Classified as a group 1 carcinogen by the IARC, As has been shown to increase

the risk of cancers at various body sites,<sup>18,19</sup> as well as several chronic diseases including As-associated skin lesions,<sup>20,21</sup> cardiovascular diseases,<sup>22</sup> and overall mortality<sup>23</sup> in those who are chronically exposed. One proposed mechanism by which As exposure affects human health involves dysregulation of the epigenome, with studies reporting its association with both gene-specific and global DNA methylation (DNAm) changes at cytosine-phosphate-guanine (CpG) sites.<sup>24,25</sup> While these DNAm-associated alterations may reflect levels of As exposure in the human body, only few such sites have been identified which has significantly hindered our ability to characterize the effects of As exposure on the epigenome and to utilize these as reliable indicators of As exposure.<sup>26</sup> The identification and validation of a larger number of As-associated DNAm sites is paramount to developing an epigenetic biomarker with the sensitivity and specificity needed to stratify individuals by their risk of As-associated toxicity and disease.

This dissertation aims to systematically explore the genetic epidemiology of breast cancer and the molecular epidemiology of arsenic toxicity in diverse populations. In Chapter 1, we examine the genetic architecture of breast cancer intrinsic subtypes across populations by conducting a cross-ancestry GWAS in European and African populations that jointly models subtype associations while efficiently accounting for missing tumor data to improve power for GWAS loci. Using the resulting summary statistics, we then perform a cross-ancestry cell type-aware TWAS based on bulk breast tissue, by integrating a cell type-deconvolution approach to further enhance the identification of susceptibility genes linked to each intrinsic subtype.

In Chapter 2, we extend beyond our exploration of the genetic architecture of breast cancer to develop polygenic risk score (PRS) models for four breast cancer phenotypes – overall, estrogen

receptor-positive, estrogen receptor-negative, and triple-negative – using GWAS data from women of African ancestry in AABCG. We employed a variety of PRS methods to construct these PRS models and additionally integrated information across ancestries, subtypes, and methods to further improve predictive performance. We then externally validated these models in the “All of Us” study<sup>27</sup> and three additional cohorts of African-ancestry individuals.

In Chapter 3, we move to explore the molecular epidemiology of As exposure. We generate DNAm data from Bangladeshi adults in the Health Effects of Arsenic Longitudinal Study (HEALS)<sup>23</sup> to identify CpG sites influenced by As exposure. We then perform a Mendelian Randomization analysis to demonstrate genetically predict arsenic metabolism, and by proxy internal dose of As exposure, causally affects DNAm at these CpG sites. Following this, we develop a DNAm-based biomarker of As exposure that robustly predicted exposures across diverse populations. We further demonstrate the utility of this marker to stratify individuals by arsenic skin lesion status in HEALS and even all-cause mortality in the Strong Heart Study,<sup>28</sup> a cohort with lower levels of As exposure.



# **1 CHAPTER 1: DISSECTING THE GENETIC ARCHITECTURE OF BREAST CANCER INTRINSIC SUBTYPES THROUGH CROSS-ANCESTRY GENOME- AND TRANSCRIPTOME-WIDE ASSOCIATION STUDIES**

## **1.1 ABSTRACT**

We conducted a combined cross-ancestry GWAS and TWAS to investigate the genetic architecture of breast cancer intrinsic subtypes across 123,669 cases and 110,277 controls of African and European ancestry. In total, we identified 116 GWAS loci across all intrinsic subtypes, including 13 novel loci for intrinsic subtypes uniquely identified in African-ancestry and cross-ancestry meta-analyses. Most of the novel variants identified for triple-negative breast cancer were driven by their associations in African-ancestry-specific analyses. When examining all lead variants identified for triple-negative breast cancer, we observed higher risk allele frequencies in African compared to European ancestry populations, suggesting that common genetic variants may partly underlie the higher prevalence of this aggressive subtype among women of African ancestry. Furthermore, utilizing a cell type-aware TWAS approach for which we constructed bulk and cell type-specific gene expression models, we identified 90 genes associated with intrinsic subtypes, including 48 genes not previously in TWAS reported to associate with these subtypes. Moreover, combined with genes identified in previous TWAS of intrinsic subtypes, these TWAS genes collectively localize nearby 73% of all GWAS loci implicated for these subtypes. Together, these findings highlight the utility of cross-ancestry analyses in improving our understanding of the genetic architecture underlying intrinsic subtypes, particularly for triple-negative breast cancer.

## 1.2 INTRODUCTION

Breast cancer is the most frequently diagnosed malignancy among women worldwide and breast cancer subtypes are highly heterogenous in their clinical presentations and aggressiveness.<sup>29,30</sup>

The expression of hormone and HER2 receptors are routinely utilized to guide the treatment strategies for patients and can be used to classify breast cancer into intrinsic subtypes which include Luminal-A-like, Luminal-B-like, HER2-enriched-like, and triple-negative breast cancer.<sup>12,13</sup> Previous genome-wide association studies (GWAS) of intrinsic subtypes conducted in primarily European populations have identified numerous loci for these subtypes, especially for Luminal-A-like breast cancer.<sup>14,15</sup> However, these studies are limited in their relevance for non-European populations, particularly women of African ancestry, who are more likely to develop more aggressive and basal-like tumors.<sup>31,32</sup> As such, it is critical that GWAS of intrinsic subtypes are conducted in non-European populations to improve our understanding of the genetic architecture for these subtypes, particularly for subtypes other than luminal-like cancers, as well as to address disparities observed at the population-level for these more aggressive subtypes.

Additionally, transcriptome-wide associations (TWAS) have been utilized to complement GWAS by leveraging genetic variants to identify susceptibility genes from associations between genetically predicted gene expression and disease risk.<sup>33–35</sup> While TWAS for breast cancer intrinsic subtypes have been performed in European ancestry populations,<sup>36,37</sup> the relatively fewer number of available samples involving paired genotyping and RNA-sequencing data from non-European populations has limited the power of TWAS in these populations. As such, only one TWAS of triple-negative breast cancer has been conducted among individuals of African ancestry, which managed to identify only two genes at the 2q14.2 locus associated with this

subtype.<sup>38</sup> These limitations highlight the need to use TWAS modeling approaches that afford more power to detect associations in these non-European datasets, as well as to implement cross-ancestry strategies that leverage well-powered populations to boosting the identification of susceptibility genes for intrinsic subtypes in diverse populations.

In this study, we investigated the genetic architecture of breast cancer intrinsic subtypes across populations by first conducting a cross-ancestry GWAS of intrinsic subtypes that jointly modeled subtype associations through efficient handling of missing tumor data to significantly boost power for detecting genetic associations. Following these GWAS analyses, we then leveraged the resulting summary statistics to perform a cross-ancestry TWAS using bulk breast tissue samples, further incorporating cell type-deconvoluted gene expression profiles to enhance our power for identifying susceptibility genes for the intrinsic subtypes.

### **1.3 RESULTS**

Genetic data for genome-wide analyses in women of African ancestry were obtained from the African Ancestry Breast Cancer Genetics (AABCG) consortium<sup>39</sup> and data for women of European ancestry came from the Breast Cancer Association Consortium (BCAC).<sup>14</sup> Briefly, the AABCG dataset included 17,391 breast cancer cases (5,882 Luminal A-like, 1,296 Luminal B-like, 795 HER2-enriched-like, and 2,760 triple-negative) and 18,800 controls. The BCAC dataset included 106,278 breast cancer cases (34,280 Luminal A-like, 6,502 Luminal B-like, 2,793 HER2-enriched-like, and 7,174 triple-negative) and 91,477 controls.<sup>14,39</sup> Detailed descriptions of these study populations and genotyping methods have been published previously.<sup>14,39</sup> For our transcriptome-wide analyses, we obtained and analyzed paired genotyping and bulk-tissue RNA-

sequencing data from healthy breast tissue samples.<sup>38,40,41</sup> These paired samples consisted of 308 individuals of European and 172 individuals of African ancestry from the Genotype-Tissue Expression (GTEx) project<sup>40</sup> and the Komen Tissue Bank.<sup>41</sup> Sample collection and processing of these paired samples has been previously described.<sup>38,40,41</sup>

### 1.3.1 Identification of GWAS loci for breast cancer intrinsic subtypes

To identify variants associated with each intrinsic subtype, we utilized a top-stage polytomous logistic regression (TOP) model<sup>42</sup> to estimate the odds ratios (ORs) and 95% confidence intervals (CIs) for every variant within each AABCG dataset (**Table S 1.1**), adjusting for age and the top 10 genetic principal components (PCs). In short, TOP utilizes an expectation-maximization algorithm to handle missing tumor characteristics, thereby enhancing the estimation of variant associations by leveraging information across subtypes. For each subtype, we performed a fixed-effects meta-analysis of summary statistics across AABCG datasets to obtain African-ancestry GWAS summary statistics. We subsequently conducted a cross-ancestry meta-analysis, integrating African ancestry summary statistics with those previously generated using TOP in European individuals from the BCAC.

In total, across all meta-analyses, we identified 116 GWAS susceptibility loci for intrinsic subtypes at a Bonferroni threshold of significance ( $p < 1.25 \times 10^{-8}$ ) accounting for four subtypes; these loci included 105 for Luminal A-like (96 European, 9 African, 96 cross-ancestry), 18 for Luminal B-like (15 European, 2 African, 14 cross-ancestry), 4 for HER2-enriched-like (4 European, 3 cross-ancestry), and 18 for triple-negative breast cancer (12 European, 7 African, 15 cross-ancestry) (**Figure 1.1, Figure S 1.1, Table S 1.2-S4**). Of the seven loci identified for triple-

negative breast cancer in the African-ancestry meta-analysis, three had been previously identified using traditional logistic regression in a largely overlapping study of African-ancestry participants in AABCG, generally showing similar effect sizes but greater statistical significance in the present study using TOP (**Table S 1.5**).<sup>39</sup> Overall, we observed no obvious genomic inflation in each meta-analysis with the sample size-adjusted  $\lambda_{1,000}$  ranging from 1.006 to 1.043 in the African-ancestry, 1.004 to 1.007 in the European-ancestry, and 1.002 to 1.009 in the cross-ancestry meta-analysis (**Figure S 1.2, Figure S 1.3, Figure S 1.4**).

### **1.3.2 Novel GWAS loci for intrinsic subtypes identified in African-ancestry and cross-ancestry meta-analyses**

Among these GWAS loci, we identified 13 novel loci for breast cancer intrinsic subtypes in the African-ancestry and cross-ancestry meta-analyses including five lead variants for Luminal-A-like (rs4011894 at 1q41, rs55913768 at 10p15.1, rs7100052 at 10p14, rs16963218 at 18q12.1, rs1105262 at 20q13.32), three for Luminal-B-like (rs145680477 at 10q26.13, rs4762310 at 12q23.1, rs930313 at 2q31.1), and five for triple-negative breast cancer (rs112200610 at 1p36.22, rs142302749 at 1p35.2, rs6590696 at 11q25, rs12710697 at 2p24.1, rs11952226 at 5p15.32) (**Table 1.1, Figures S 1.5A-M**); seven of these novel loci had not been reported in previous GWAS of any breast cancer phenotypes and were at least 2 Mb away from any reported variants (**Table 1.1**). Lead variants at these novel loci showed consistent directionality in their associations across AABCG datasets (**Table S 1.6**). Of these variants, 7 were identified in the cross-ancestry meta-analysis while 6 were uniquely identified in the African-ancestry meta-analysis; these seven variants exhibited consistent directionality in their associations across African and European ancestries. Interestingly, all but one of the newly identified GWAS loci for

triple-negative breast cancer in this study were either exclusively identified in the African ancestry-specific meta-analysis or showed stronger associations in this analysis compared to the European-ancestry meta-analysis (**Table 1.1**). Furthermore, 9 of the 13 lead variants were located in intronic regions of genes including *MARK1*, *FAM208B*, *GAREM1*, *CPXM2*, *CFAP54*, *AC092573.2*, *MIR34AHG*, *OPCML*, and *LINC01376*, while the other 4 were located in intergenic regions. We also examined the associations of these lead variants across other intrinsic subtypes and observed that all the lead variants at novel loci for Luminal-B-like and triple-negative breast cancer had nominally significant heterogeneity tests ( $p < 0.05$ ), while only two of the five novel loci for Luminal-A-like exhibited heterogeneity (**Table 1.2**); for variants with evidence of heterogeneity, the odds ratio was consistently stronger for the subtype in which the variant was originally identified as a lead variant compared to other subtypes. We additionally performed conditional analyses at these novel loci with all variants within  $\pm 2$  Mb of each novel lead variant and identified an additional independent signal ( $p < 1 \times 10^{-5}$ ) at the 10p15.1 locus (rs11259192) for Luminal-A-like and at the 1p36.22 locus (rs2483677) for triple-negative breast cancer (**Table S 1.7**).

### **1.3.3 Population-specific differences in risk allele frequency of lead variants**

We examined the population-specific allele frequencies of risk alleles for all lead variants identified across all GWAS meta-analyses by intrinsic subtype. Notably, risk allele frequencies for lead variants associated with triple-negative breast cancer were higher in African-ancestry populations (both in Africa and North America) compared to European-ancestry populations (**Figure 1.2**, **Table S 1.8**). Similar findings were observed for HER2-enriched-like breast cancer with median risk allele frequencies of 0.37 among African-ancestry individuals in Africa, 0.32

among African-ancestry individuals in North America, and 0.24 among individuals of European ancestry; however, these differences were not statistically significant. In contrast, no statistically significant differences in risk allele frequencies across populations were observed for luminal subtypes ( $p>0.05$ ) (**Figure 1.2, Table S 1.8**).

#### **1.3.4 Identifying susceptibility genes for breast cancer intrinsic subtypes using a cell type-aware TWAS approach**

Of the 26,408 and 24,671 genes that had passed quality control when processing RNA-sequencing data from 308 European and 172 African ancestry breast tissue samples, respectively, we constructed bulk-tissue prediction models using elastic net regression, as well as cell type-specific models using MiXcan, incorporating cell type enrichment scores estimated with the “xCell” software package. In total, we modeled 18,083, 8,900, 8,965, 9,089, and 9,251 genes in bulk-tissue, epithelial cells, adipocytes, endothelial cells, and fibroblasts, respectively, with a five-fold cross-validation predictive performance  $r>0.1$  in European-ancestry samples and 20,203, 9,819, 9,402, 9,621, and 10,148 genes, respectively, in African-ancestry samples (**Figure S 1.6**). When applying these models to our GWAS data, a total of 21,973, 22,549, and 24,707 genes were evaluated across all cell types for European-ancestry, African-ancestry, and cross-ancestry meta-analyses, respectively; hence for declaring genes as Bonferroni-significant in the TWAS, we accounted for the number of genes tested, four subtypes, and four cell types, yielding p-value thresholds of  $1.42 \times 10^{-7}$  for European-ancestry,  $1.39 \times 10^{-7}$  for African-ancestry, and  $1.26 \times 10^{-7}$  for cross-ancestry TWAS analyses.

Overall, we identified 90 genes in 56 loci associated with intrinsic subtypes including 85 genes in 54 loci for Luminal-A-like, 3 genes in 3 loci for Luminal-B-like, 1 gene in 1 locus for HER2-enriched-like, and 6 genes in 3 loci for triple-negative breast cancer (**Table S 1.9**). Of these genes, 87 were identified in the European-ancestry TWAS, 1 in the African-ancestry TWAS, and 59 identified in cross-ancestry TWAS meta-analysis; though only 1 gene in the African-ancestry and 2 genes in the cross-ancestry TWAS were uniquely identified compared to genes identified in the European-ancestry TWAS, 16 of the genes identified in the European-ancestry TWAS exhibited stronger associations in the cross-ancestry meta-analysis (**Table S 1.9**). Furthermore, we observed 64 of the 90 genes were identified in the bulk-tissue TWAS, while the remaining 26 genes were identified uniquely in cell type-level TWAS which identified a total of 17, 25, 22, and 25 genes in epithelial cells, adipocytes, endothelial cells, and fibroblasts, respectively (**Table S 1.10**). Moreover, we observed that when a gene had prediction models in multiple cell types, the cell type-level TWAS z-scores were moderately concordant across cell types for a given gene and subtype, with an intraclass correlation coefficient (ICC) of 0.376 (95% CI: 0.373-0.380). On the other hand, the z-scores for a given gene and cell type were only weakly concordant across subtypes, with an ICC of 0.107 (95% CI: 0.104-0.110). These trends were consistent across both European and African ancestry-specific TWAS analyses (**Table S 1.11**). We additionally observed that many identified genes showed evidence of subtype-specificity when utilizing a stringent p-value threshold ( $p > 0.05$ ) for other subtypes including 17 genes for Luminal-A-like and 2 genes for triple-negative breast cancer (**Table S 1.9**).

Among these 90 genes, 48 genes in 36 loci were newly identified to associate with intrinsic subtypes including 46 genes in 36 loci for Luminal-A-like, 1 gene in 1 locus for Luminal-B-like,



and 3 genes in 2 loci for triple-negative breast cancer (**Table 1.3**), while the remaining 42 genes had been previously linked to intrinsic subtypes in TWAS (**Table S 1.12**); for these previously linked genes that had specifically been identified in breast tissue TWAS, both the direction and strength of their associations with intrinsic subtypes were highly consistent with those observed in the present study (**Figure S 1.7**). Of these newly associated genes with intrinsic subtypes, 10 genes showed evidence for subtype-specificity including several previously reported proto-oncogenes such as *JUNB*,<sup>43</sup> *MNI*,<sup>44</sup> *XRCC5*,<sup>45,46</sup> and *ACTN1*<sup>47</sup> that showed positive associations with Luminal-A-like breast cancer (**Table 1.3**). Additional genes we identified that had TWAS associations with concordant directionality with previously reported roles in tumor biology including tumor suppressors such as *ATMIN* (a known interactor of *ATM*)<sup>48</sup> and *RBL2*<sup>49–51</sup> that had negative associations with Luminal-A-like breast cancer and an oncogene, *TRIM32*,<sup>52</sup> that had a positive association with this subtype. Interestingly, *CDKN2A*, which was previously characterized as a tumor suppressor gene,<sup>53,54</sup> exhibited a positive association with Luminal-A-like breast cancer.

### 1.3.5 TWAS genes for intrinsic subtypes commonly localize near GWAS loci

Nearly all of the 90 TWAS-identified genes were located within 2 Mb of intrinsic subtype GWAS lead variants identified across African-ancestry, European-ancestry, and cross-ancestry meta-analyses (**Table S 1.9**) except two genes, *METTL15P1* and *PA2G4P4*, both of which were pseudogenes identified to associate with Luminal-A-like breast cancer (**Table 1.4**). Among the 88 genes within 2 Mb of nearby GWAS lead variants, we identified 5 genes for Luminal-A-like and 1 gene for triple-negative breast cancer that could not be explained by GWAS signals including a protein coding gene, *APOBEC3B*, that was associated with Luminal-A-like breast

cancer (**Table 1.4**). On the other hand, of the 116 GWAS loci associated with intrinsic subtypes, we identified TWAS genes within 2 Mb of 50 loci; in comparison, all genes identified in prior TWAS of intrinsic subtypes were within 2 Mb of 73 GWAS loci (**Figure S 1.8**). Together, this study and prior work implicate TWAS associations within 2 Mb of 85 loci, representing 73% of all 116 GWAS loci identified for intrinsic subtypes in this study.

### **1.3.6 TWAS signals provide functional insight into novel GWAS loci for intrinsic subtypes**

Of the 9 novel GWAS loci for intrinsic subtypes that mapped to intronic regions, 8 corresponded to genes for which TWAS prediction models could be successfully constructed (**Table 1.1**). Apart from the lincRNA, *LINC01376*, identified for triple-negative breast cancer, TWAS associations for these genes were generally weak across all cell types and ancestries, with p-values either exceeding the nominal significance threshold ( $p > 0.05$ ) or falling just below it (**Table S 1.13**). When instead examining TWAS associations within  $\pm 2$  Mb of lead variants at these GWAS loci, we found that one locus identified in this study – 10p15.1 for Luminal-A-like breast cancer – also contained *GDI2*, a TWAS gene we newly identified to associate with Luminal-A-like breast cancer in both bulk breast tissue and epithelial cells (**Table 1.1, Table 1.3**). The lead variant at this locus, rs55913768, was among variants included in the expression prediction models for *GDI2*. After conditioning on rs55913768, the association between *GDI2* and Luminal-A-like breast cancer was no longer statistically significant (**Table S 1.14**), suggesting that the TWAS signal may be driven by this novel GWAS association at 10p15.1. Furthermore, we expanded our search to include suggestive TWAS associations within  $\pm 2$  Mb of lead variants at these novel GWAS loci; these were defined by  $p < 1 \times 10^{-3}$  in any cell type or ancestry, inclusion of at least one variant in the gene prediction model with  $r^2 > 0.1$  relative to the lead variant, and loss of

significance after conditioning on the lead variant ( $p > 1 \times 10^{-3}$ ). Using these criteria, we identified *Clorf115* as a suggestive gene potentially explaining the GWAS signal at 1q41 for Luminal-A-like breast cancer and found additional support for the involvement of *LINC01376* at 2p24.1 in triple-negative breast cancer (**Table S 1.14**).

## 1.4 DISCUSSION

To our knowledge, this is first cross-ancestry genome- and transcriptome-wide association study of breast cancer intrinsic subtypes to date, involving 123,669 cases and 110,277 controls of European and African ancestry from the BCAC and AABCG consortia. In total, we identified 116 GWAS loci associated with intrinsic subtypes across ancestry-specific and cross-ancestry meta-analyses, including 13 loci not previously reported to associate with intrinsic subtypes. Ancestry-specific and cross-ancestry TWAS identified 90 susceptibility genes at 56 loci associated with intrinsic subtypes, including 48 novel associations, several of which demonstrated subtype-specific effects. Additionally, many GWAS and TWAS associations that have been previously reported replicated in our study with consistent magnitude and directionality.<sup>36,37,39</sup> Combining TWAS-identified genes from this study with those from previous TWAS of intrinsic subtypes, our findings reveal that 73% of GWAS loci are located within 2 Mb of a TWAS locus, indicating substantial overlap between disease-associated variants and mapped susceptibility genes.

While prior studies of intrinsic subtypes have primarily focused on European populations, this study highlights the importance of examining the genetic architecture of these breast cancer subtypes across ancestries to uncover novel susceptibility loci. For example, one GWAS locus

we newly identified to associate with Luminal-A-like breast cancer in the cross-ancestry meta-analysis was at 10p15.1 (rs55913768). rs55913768 was selected as a variant in the expression prediction model for *GDI2*, a gene identified in our TWAS to negatively associate with Luminal-A-like breast cancer in bulk-tissue and epithelial cells; after conditioning on this variant, these associations no longer reached nominal significance, indicating that these TWAS signals were driven by this variant. *GDI2* encodes a regulator of Rab GTPases, which are involved in vesicular trafficking and in epithelial-mesenchymal transition in breast cancer cell lines.<sup>55–57</sup> Decreased *GDI2* protein levels has also been associated with estrogen receptor-positive breast cancer in a Mendelian Randomization study of plasma proteins, where rs55913768 was notably the primary instrument used to infer *GDI2* protein levels.<sup>58</sup> These findings highlights the need for further research into the potential tumor suppressive role of *GDI2* in mammary epithelial cells in luminal breast cancers.

We also identified a novel GWAS susceptibility locus for triple-negative breast cancer at 2p24.1 (rs12710697) in the cross-ancestry meta-analysis. This variant is located within *LINC01376*, which has been implicated as a oncogene in epithelial-derived cancers, having been shown to promote tumorigenesis in nasopharyngeal carcinoma,<sup>59</sup> be overexpressed in endometrial carcinomas,<sup>60</sup> and potentially facilitate epithelial-mesenchymal transition in lung cancer cells.<sup>61</sup> Although *LINC01376* was not Bonferroni-significant in our TWAS, a suggestive TWAS association was identified for triple-negative breast cancer ( $p=5.95 \times 10^{-6}$ ) and was attenuated after conditioning on rs12710697, suggesting this variant may potentially impact triple-negative breast cancer by modulating *LINC01376* expression. Notably, rs12710697 was in LD with rs12710696 (European-ancestry  $r^2=0.77$ , African-ancestry  $r^2=0.68$ ), a known estrogen receptor-

negative risk variant,<sup>62</sup> though the latter has not been identified as a genome-wide significant variant for triple-negative breast cancer in this study nor prior studies. Together, these findings indicate the strength of our cross-ancestry meta-analysis to refine association signals for breast cancer intrinsic subtypes by leveraging population differences in linkage disequilibrium.

Our findings further underscore the utility of performing cross-ancestry analyses to identify population-specific patterns of common variation associated with these intrinsic subtypes. We observed that across lead variants identified for triple-negative and HER2-enriched-like breast cancer in this study, the risk allele frequencies generally tended to be higher in African-ancestry compared to European-ancestry populations; this pattern was not observed for either luminal breast cancer subtype. Given that individuals of African ancestry have a higher reported odds of developing basal-like and HER2-enriched-like cancers compared to individuals of European ancestry,<sup>63</sup> this ancestry-specific distribution of risk alleles may partially help explain population-level differences in the prevalences of these subtypes. Moreover, of the GWAS loci newly identified for triple-negative breast cancer in this study, all but one were either uniquely detected in the African ancestry-specific meta-analysis or exhibited stronger associations compared to the European-ancestry-specific meta-analysis. These observations underscore the utility of cross-ancestry analyses to identify susceptibility loci that may be missed in studies consisting of predominantly European-ancestry participants and to improve our understanding behind disparities in the distribution of breast cancer subtypes at population-level. However, even though many of these novel GWAS loci for intrinsic subtypes were largely driven by associations in the African-ancestry meta-analysis, we generally did not observe corresponding TWAS signals at these loci. This likely reflects the limited power of the African-ancestry TWAS, which

identified only one unique gene compared to the European and cross-ancestry analyses. As such, increasing the size of transcriptomic datasets involving African-ancestry individuals is critical to fully characterize mechanisms underlying the regulation of gene expression at these novel GWAS loci.

Several key strengths in our study design allowed us to identify a substantial number of GWAS and TWAS loci for intrinsic subtypes. First, the utilization of TOP as our primary GWAS model significantly improved our power to detect susceptibility loci as it enabled us to borrow information across subtypes by efficiently handling missing tumor subtype data. For direct comparison, a prior African-ancestry GWAS of triple-negative breast cancer that utilized a largely overlapping sample from AABCG identified only three susceptibility loci using traditional logistic regression;<sup>39</sup> in contrast, our TOP-based analysis involving African-ancestry individuals from AABCG identified seven loci for this subtype, expanding the total number of GWAS loci implicated for triple-negative breast cancer from 13 to 18.<sup>14,15,39</sup> Additionally, our identification of numerous TWAS genes in breast tissue – of which over half had not been previously reported for intrinsic subtypes – is largely attributable to the substantially increased number of female breast tissue samples with paired genotyping and RNA-sequencing used in this study to build expression prediction models compared to prior studies.<sup>36–38</sup> Moreover, our utilization of a cell type-aware TWAS approach enabled us to identify 40% additional genes compared to the 64 detected using conventional bulk-tissue prediction models, suggesting that such approaches can not only improve our understanding of cellular mechanisms underlying TWAS associations but also increase our power to detect susceptibility genes.

In summary, to the best of our knowledge, this is the first and largest cross-ancestry GWAS and TWAS of breast cancer intrinsic subtypes. We identified GWAS loci and TWAS genes that have not been previously reported to associate with the intrinsic subtypes, thereby improving our understanding of the genetic architecture of these subtypes. Further research into validating the loci we identified and increasing the amount of transcriptomic data available in non-European ancestries are both needed to better characterize the regulatory mechanisms underlying these genetic associations.

## **1.5 METHODS**

### **1.5.1 Study population involved in GWAS meta-analyses**

GWAS summary statistics involving European-ancestry women were obtained from previously published data involving 106,278 invasive cases and 91,477 controls that were a part of 82 Breast Cancer Association Consortium (BCAC) studies from over 20 countries.<sup>14</sup> African-ancestry participants were drawn from the African Ancestry Breast Cancer Genetics Consortium (AABCG), including 17,391 breast cancer cases and included 5,882 Luminal-A-like, 1,296 Luminal-B-like, 795 HER2-enriched-like, and 2,760 triple-negative breast cancer cases and 18,800 controls.<sup>39</sup> In total, this African-ancestry genetic data was collated across seven datasets defined by genotyping and sequencing platforms<sup>39</sup> and the participating studies of each dataset are detailed in **Table S 1.1**. All participants had previously provided their informed consent, these participating studies received approval by appropriate committees and/or institution review boards. We conducted variant-level quality control and harmonization procedures for genetic data of African-ancestry participants as described below.

### **1.5.2 Processing genetic data for women of African ancestry**

Samples were excluded based on the following criteria: 1) was not genetically identified as female, 2) had a variant call rate below 95%, 3) the proportion of African ancestry was estimated to be under 5% (using Admixture v1.3.0<sup>64,65</sup> with the 1000 Genomes Project<sup>66</sup> as a reference panel), 4) genetic relatedness ( $\text{Pi-HAT} > 0.45$ ) where only one individual from each related pair was retained, and 5) unusually high levels of heterozygosity. We then filtered variants to have an imputation score  $> 0.3$  and excluded those with allele frequencies differing by more than 0.15 across genotyping and sequencing platforms. We performed an initial filter of variants to have at least a minor allele frequency of 0.01 across all these participants. However, since subtypes other than Luminal-A-like breast had smaller sample sizes, we subsequently filtered for variants that had a minor allele frequency count  $\geq 30$  for each subtype, yielding a minimum observed minor allele frequency of 0.0094 for Luminal-A-like, 0.023 for Luminal-B-like, 0.038 for HER2-enriched-like, and 0.011 for Triple-negative breast cancer. Next, we used PLINKv2.0<sup>67</sup> to generate principal components (PCs) after pruning variants using an  $r^2 < 0.2$  cutoff in 500kb windows.

### **1.5.3 Harmonizing and meta-analyzing GWAS summary statistics across African and European ancestries for intrinsic breast cancer subtypes**

To generate GWAS summary statistics for individuals of African ancestry, we used a two-stage polytomous logistic regression (TOP)<sup>42</sup> framework to evaluate the association between each genetic variant and each intrinsic breast cancer subtype, defined by hormone receptor and HER2 status; these receptor-defined subtypes include Luminal A-like, Luminal B-like, HER2-enriched-like, and triple-negative breast cancer. Analyses were conducted separately within each



genotyping/sequencing platform, adjusting for the top 10 principal components (PCs). For variants that could not be evaluated using TOP regression due to model convergence issues or limited sample sizes, we employed logistic regression or Firth regression as appropriate.<sup>67</sup> Finally, we performed a fixed-effects meta-analysis across genotyping and sequencing platforms to obtain African-ancestry-specific summary statistics using METAL.<sup>68</sup>

European ancestry summary statistics had previously been generated across among BCAC participants using TOP regression for five intrinsic subtypes defined using both receptor status and tumor grade including Luminal-A-like, Luminal-B-like, Luminal-B/HER2-negative-like, HER2-enriched-like, and triple-negative breast cancer.<sup>14</sup> To harmonize these results with the African ancestry subtype-specific summary statistics generated in this study, we performed a fixed-effects meta-analysis of the Luminal A-like and Luminal-B/HER2-negative-like GWAS results, producing aggregated Luminal A-like summary statistics for the European dataset.

We then conducted a fixed-effects meta-analysis combining the African-ancestry and European-ancestry summary statistics to generate cross-ancestry meta-analyzed summary statistics for each of the four receptor-defined intrinsic subtypes. We assessed cross-ancestry heterogeneity in the association between each variant and each subtype using Cochran's Q test.

#### **1.5.4 Identification and classification of novel GWAS risk loci for intrinsic subtypes**

Lead variants were identified in each of the three analyses – the African-ancestry meta-analysis, European-ancestry meta-analysis, and cross-ancestry meta-analysis – using a Bonferroni-corrected significance threshold of  $p < 1.25 \times 10^{-8}$  (i.e., the typical genome-wide significant p-value

threshold of  $5 \times 10^{-8}$  divided by the four intrinsic subtypes) with 4 Mb windows. Variants detected in either the African-ancestry or cross-ancestry meta-analyses were classified as novel loci for intrinsic subtypes if they had not been previously reported in GWAS of intrinsic subtypes<sup>14,15</sup> and were located more than 2 Mb from any known intrinsic subtype-associated variants,<sup>14,15</sup> including those identified in the European-ancestry meta-analysis.

### **1.5.5 Identifying independent signals at novel GWAS risk loci**

To identify additional, independent variants at novel loci for breast cancer intrinsic subtypes, we performed conditional analysis for all variants within  $\pm 2$  Mb of each novel lead variant using the GCTA-COJO<sup>69</sup> software package. In particular, for lead variants identified in the African-ancestry meta-analysis, we ran GCTA-COJO directly with African-ancestry GWAS summary statistics and utilized 5,000 randomly selected AABCG controls to compute conditional beta coefficients. We considered variants with a conditional  $p < 10^{-5}$  as independent signals and subsequently repeated this conditional analysis until no additional independent signals were statistically significant. We then used GCTA-COJO to estimate joint odds ratios for all independent signals at each locus by modeling all independent variants simultaneously. For lead variants identified in the cross-ancestry meta-analysis, we similarly ran GCTA-COJO for all variants within  $\pm 2$  Mb of these lead variants, separately, however with African-ancestry summary statistics utilizing the panel of 5,000 African-ancestry controls and European-ancestry summary statistics utilizing European individuals from the 1000 Genomes Project<sup>66</sup> as a linkage disequilibrium reference panel. Following this, conditional estimates were then meta-analyzed and used to identify additional independent signals using the same  $p < 10^{-5}$  threshold. We similarly

jointly modeled all independent signals at each locus and subsequently meta-analyzed these joint estimates across ancestries.

### **1.5.6 Building bulk-tissue and cell type-aware gene expression prediction models in breast tissue**

We retrieved and processed paired RNA-sequencing and genotyping data that had been previously collected from healthy breast tissue samples from the Genotype-Tissue Expression (GTEx) Project<sup>40</sup> and the Komen Tissue Bank<sup>38,41</sup> from women of European and African ancestry. To process RNA-sequencing data of samples from the Komen Tissue Bank, we first merged FastQ files and utilized the “Salmon” software package<sup>70</sup> to auto-infer strandedness. Next, we utilized “FastQC”<sup>71</sup> to perform read quality control. Following this, we extracted unique molecular identifiers (UMI) sequences from these FastQ reads using “UMI-tools,”<sup>72</sup> trimmed adapters using “Trim Galore,”<sup>73</sup> and removed genome contaminants and ribosomal RNA data utilizing “BBSplit”<sup>74</sup> and “SortMeRNA.”<sup>75</sup> We then performed multiple alignment and quantification using “STAR”<sup>76</sup> and “Salmon.”<sup>70</sup> We sorted and indexed these alignments using “samtools,”<sup>77</sup> performed UMI-based deduplication using the MarkDuplicates function “picard”<sup>78</sup> and then performed transcript assembly and quantification using “StringTie”;<sup>79</sup> we annotated transcripts using GENCODEv26<sup>80</sup> to obtain read counts. Since pre-processed GTEx read counts were readily downloadable from the GTEx portal, we combined the read counts from GTEx and Komen samples, separately for African-ancestry and European-ancestry samples. Following this processing of RNA-sequencing data, we retained paired RNA-sequencing and genotyping samples from a total of 308 individuals of European ancestry (124 from GTEx and 184 from

Komen) and 172 individuals of African ancestry (19 from GTEx and 153 from Komen) for subsequent analyses.

For these combined read count matrices, we retained genes with a transcripts per million value  $\geq 0.1$  in at least 20% of samples and a raw read count  $\geq 6$  in at least 20% of samples. After filtering, 26,472 transcripts were retained for the European ancestry paired samples and 24,715 for the African ancestry paired samples. We then utilized “edgeR”<sup>81</sup> to perform trimmed mean of M values (TMM) normalization which resulted in TMM-normalized transcripts per million. We subsequently performed a rank-based inverse normal transform to normalize values across samples. We then utilized PCAforQTL<sup>82</sup> to generate PCs for this RNA expression data to control for unmeasured potential confounders including batch effects; specifically, the elbow approach was used to select PCs to adjust for, which resulted in 11 PCs for African-ancestry and 13 PCs for European-ancestry.

Next, we built ancestry-specific, bulk-tissue gene expression prediction models using elastic net regression<sup>83</sup> with 5-fold cross-validation for variants within  $\pm 1$  Mb in the regions flanking transcription start and stop sites, retaining only models with an  $r^2 > 0.01$ ; we adjusted for covariates including age, an indicator batch variable (whether a sample came from GTEx vs. Komen), and ancestry-specific PCs generated by PCAforQTL. Additionally, we utilized the “MiXcan” package to build cell type-aware expression prediction models for four common cell types in breast tissue including epithelial cells, adipocytes, endothelial cells, and fibroblasts, adjusting for the same covariates as in the bulk-tissue models. Specifically, we used the “xCell”<sup>84</sup> package to infer cell type enrichment scores for these cell types and then input these cell type

enrichment scores as priors in the “MiXcan”<sup>85</sup> package to construct cell type-level models. In total, we modeled 20,203, 9,819, 9,402, 9,621, and 10,148 genes in bulk-tissue, epithelial cells, adipocytes, endothelial cells, and fibroblasts, respectively, in African-ancestry samples; we modeled 18,083, 8,900, 8,965, 9,089, and 9,251 genes, respectively, in European-ancestry samples.

### **1.5.7 Performing cell type-aware TWAS of breast cancer intrinsic subtypes**

Prior to performing TWAS analyses, we constructed ancestry-specific covariance matrices for variants involved in prediction models using the “bigsnpr” R package<sup>86</sup> to enable application of these models to GWAS summary statistics. We then performed z-score imputation to harmonize variants available in our GWAS summary statistics with model variants.<sup>87</sup> For each intrinsic subtype, we utilized the ancestry-specific GWAS summary statistics we generated along with ancestry-specific expression prediction models to perform a bulk-tissue TWAS, as well as cell type-aware TWAS analyses. Next, we performed sample-size weighted Z-score meta-analyses for each TWAS association across European and African ancestries for their corresponding subtype and cell type (including bulk-tissue). TWAS genes for each ancestry (African, European, and cross-ancestry) were identified as those passing the Bonferroni threshold of significance accounting for four subtypes, four cell types, and the total number of genes tested for each ancestry; these p-value thresholds were  $1.42 \times 10^{-7}$  for European-ancestry TWAS,  $1.39 \times 10^{-7}$  for African-ancestry TWAS, and  $1.26 \times 10^{-7}$  for cross-ancestry TWAS. Among the Bonferroni-significant TWAS genes, we classified genes as having evidence of subtype-specificity using a stringent threshold, classifying genes as subtype-specific if their associations with all other subtypes had p-values exceeding 0.05.

### **1.5.8 Performing TWAS conditioned on nearby GWAS variants**

To identify TWAS signals that were independent of nearby GWAS variants, we performed conditional TWAS (cond-TWAS) analyses for Bonferroni-significant genes where we conditioned prediction model variants on lead variants identified in our GWAS meta-analyses ( $p < 1.25 \times 10^{-8}$ ) within  $\pm 2$  Mb in the regions flanking transcription start and stop sites using GCTA-COJO.<sup>69</sup> For TWAS-identified genes in each ancestry (European and African), we performed GCTA-COJO using ancestry-specific LD reference panels: 5,000 AABCG controls for African ancestry and European individuals from the 1000 Genomes Project for European ancestry;<sup>66</sup> ancestry-specific GWAS lead variants were utilized in this analysis. For genes identified in the cross-ancestry TWAS, we first performed ancestry-specific conditional analyses where African-ancestry TWAS models were conditioned on GWAS lead variants from both African-ancestry and cross-ancestry meta-analyses, and European-ancestry models were conditioned on lead variants from both European-ancestry and cross-ancestry meta-analyses; we then meta-analyzed the resulting conditional TWAS associations across ancestries and assessed whether the combined signal remained conditionally independent. The significance threshold for conditional independence was identical to the Bonferroni threshold of significance used to identify genes in the primary TWAS.

## 1.6 APPENDIX

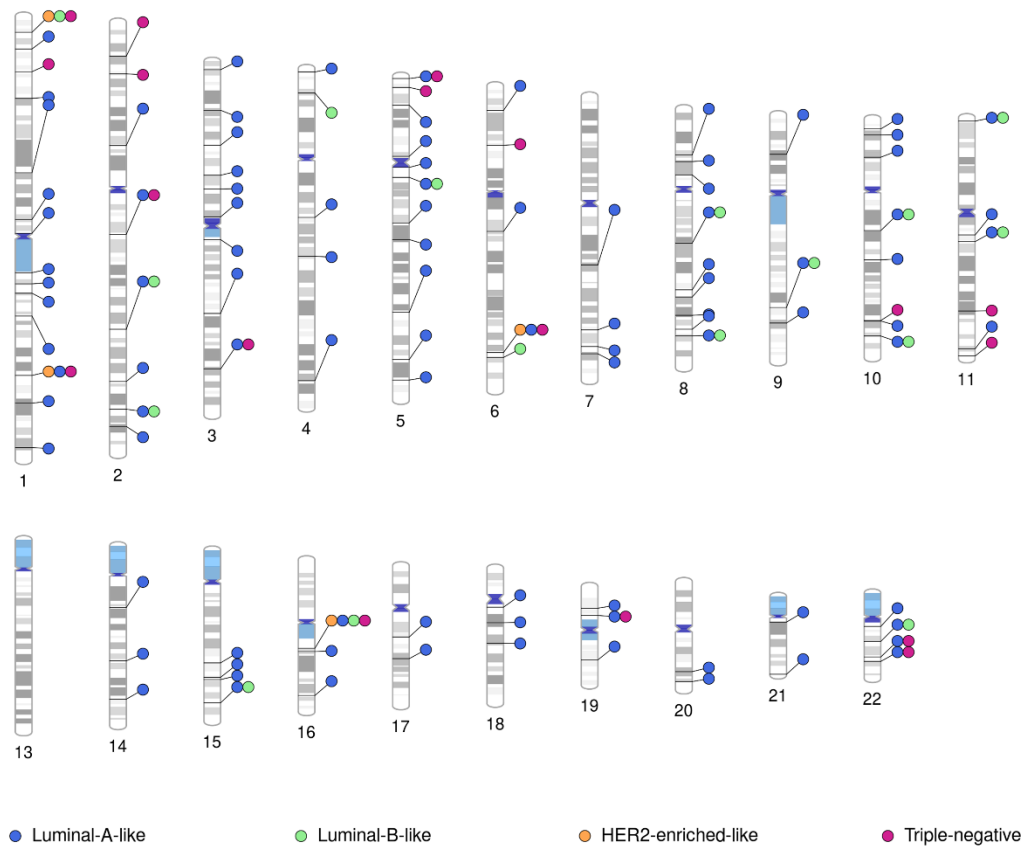


Figure 1.1 Ideogram of GWAS loci associated with each intrinsic breast cancer subtype, including both novel and previously reported loci identified in this study

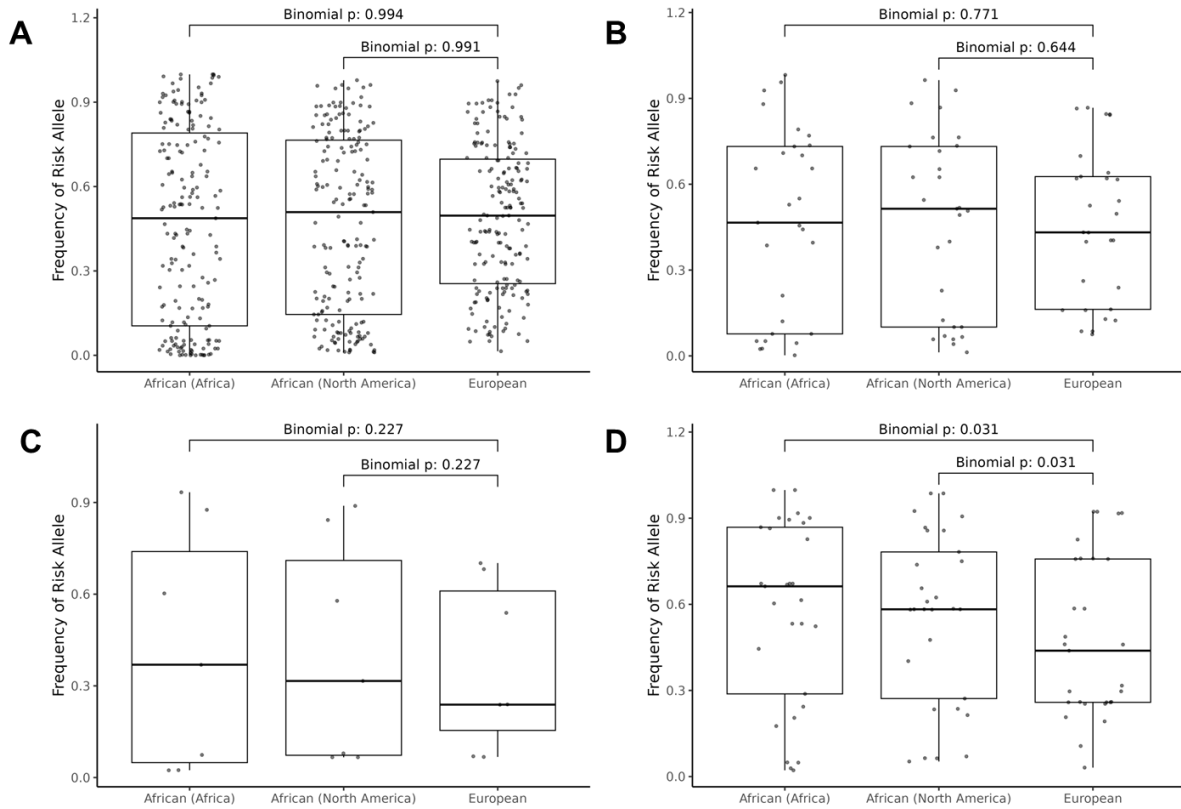


Figure 1.2 Risk allele frequencies of all lead variants identified across GWAS analyses for each subtype by population. Note: Risk allele frequencies were derived from control samples. One-sided binomial test p-values are shown for the proportion of lead variants with higher risk allele frequencies in African-ancestry compared to European-ancestry populations.



Table 1.1 Novel GWAS loci for breast cancer intrinsic subtypes identified in African-ancestry and cross-ancestry meta-analyses (Bonferroni-adjusted p-value<0.05)

| Loci            | rsID               | CHR:POS:REF:ALT  | Alleles <sup>a</sup> | Nearest gene      | Gene region | African-ancestry meta-analysis |          |                  | European-ancestry meta-analysis |          |                  | Cross-ancestry meta-analysis |          |                           |
|-----------------|--------------------|------------------|----------------------|-------------------|-------------|--------------------------------|----------|------------------|---------------------------------|----------|------------------|------------------------------|----------|---------------------------|
|                 |                    |                  |                      |                   |             | OR (95% CI)                    | P-value  | EAF <sup>b</sup> | OR (95% CI)                     | P-value  | EAF <sup>b</sup> | OR (95% CI)                  | P-value  | Het. P-value <sup>c</sup> |
| Luminal-A-like  |                    |                  |                      |                   |             |                                |          |                  |                                 |          |                  |                              |          |                           |
| 1q41            | <b>rs4011894</b>   | 1:220613960:G:A  | A/G                  | <i>MARX1</i>      | Intronic    | 1.15 (1.05, 1.26)              | 0.002842 | 0.04             | 1.05 (1.03, 1.07)               | 9.26E-08 | 0.23             | 1.05 (1.03, 1.07)            | 6.66E-09 | 0.05121                   |
| 10p15.1         | <b>rs55913768</b>  | 10:5707486:G:A   | A/G                  | <i>FAM208B</i>    | Intronic    | 1.08 (1.01, 1.14)              | 0.01753  | 0.10             | 1.04 (1.03, 1.06)               | 5.40E-08 | 0.33             | 1.04 (1.03, 1.06)            | 4.89E-09 | 0.3204                    |
| 10p14           | rs7100052          | 10:9058859:T:G   | G/T                  | -                 | Intergenic  | 1.07 (1.01, 1.13)              | 0.03346  | 0.11             | 1.06 (1.04, 1.09)               | 4.06E-08 | 0.12             | 1.06 (1.04, 1.09)            | 3.99E-09 | 0.8977                    |
| 18q12.1         | rs16963218         | 18:32351771:A:G  | G/A                  | <i>GAREM1</i>     | Intronic    | 0.89 (0.85, 0.93)              | 7.65E-07 | 0.25             | 0.92 (0.88, 0.95)               | 8.59E-06 | 0.04             | 0.91 (0.88, 0.93)            | 4.38E-11 | 0.3724                    |
| 20q13.32        | <b>rs1105262</b>   | 20:59573644:T:G  | G/T                  | -                 | Intergenic  | 0.83 (0.78, 0.88)              | 1.89E-09 | 0.89             | 1.00 (0.98, 1.01)               | 0.580466 | 0.39             | 0.99 (0.97, 1.00)            | 0.05496  | 1.08E-08                  |
| Luminal-B-like  |                    |                  |                      |                   |             |                                |          |                  |                                 |          |                  |                              |          |                           |
| 10q26.13        | <b>rs145680477</b> | 10:123814971:T:C | C/T                  | <i>CPXM2</i>      | Intronic    | 1.72 (1.44, 2.06)              | 3.67E-09 | 0.04             | -                               | -        | -                | -                            | -        | -                         |
| 12q23.1         | rs4762310          | 12:96591322:G:A  | A/G                  | <i>CFAP54</i>     | Intronic    | 0.60 (0.51, 0.71)              | 2.95E-09 | 0.96             | 0.99 (0.94, 1.04)               | 0.571511 | 0.87             | 0.94 (0.90, 0.99)            | 0.02171  | 3.79E-08                  |
| 2q31.1          | rs930313           | 2:173354390:C:G  | G/C                  | <i>AC092573.2</i> | Intronic    | 1.16 (1.06, 1.28)              | 0.001835 | 0.76             | 1.15 (1.09, 1.20)               | 7.19E-08 | 0.84             | 1.15 (1.10, 1.20)            | 5.12E-10 | 0.7753                    |
| Triple-negative |                    |                  |                      |                   |             |                                |          |                  |                                 |          |                  |                              |          |                           |
| 1p36.22         | rs112200610        | 1:9156852:C:T    | T/C                  | <i>MIR34AHG</i>   | Intronic    | 1.28 (1.15, 1.43)              | 5.47E-06 | 0.05             | 1.19 (1.10, 1.30)               | 4.50E-05 | 0.03             | 1.23 (1.15, 1.31)            | 1.77E-09 | 0.2969                    |
| 1p35.2          | <b>rs142302749</b> | 1:31087245:A:G   | G/A                  | -                 | Intergenic  | 1.38 (1.24, 1.55)              | 9.77E-09 | 0.05             | -                               | -        | -                | -                            | -        | -                         |
| 11q25           | <b>rs6590696</b>   | 11:133440533:A:C | C/A                  | <i>OPCML</i>      | Intronic    | 1.22 (1.14, 1.31)              | 8.57E-09 | 0.76             | 1.01 (0.97, 1.05)               | 0.562749 | 0.21             | 1.06 (1.02, 1.09)            | 0.001047 | 1.83E-06                  |
| 2p24.1          | rs12710697         | 2:19121207:C:T   | T/C                  | <i>LINC01376</i>  | Intronic    | 0.90 (0.85, 0.95)              | 0.000108 | 0.37             | 0.93 (0.90, 0.96)               | 1.56E-06 | 0.56             | 0.92 (0.90, 0.95)            | 1.31E-09 | 0.2667                    |
| 5p15.32         | <b>rs11952226</b>  | 5:6120548:T:C    | C/T                  | -                 | Intergenic  | 1.19 (1.13, 1.26)              | 5.52E-10 | 0.32             | -                               | -        | -                | -                            | -        | -                         |

**Footnote:**

Lead variants at loci not previously reported in GWAS of any breast cancer phenotype are in bold.

Bonferroni threshold of significance corresponded to a p-value of  $1.25 \times 10^{-8}$

<sup>a</sup> Effect/Baseline alleles

<sup>b</sup> Effect allele frequency in controls

<sup>c</sup> P-value computed from the heterogeneity test comparing estimates between African and European ancestry meta-analyses

Table 1.2 Heterogeneity in associations of novel intrinsic subtype lead variants across other breast cancer subtypes

| Loci            | rsID        | Alleles <sup>a</sup> | Luminal-A-like   | Luminal-B-like   | HER2-enriched-like | Triple-negative  | Het.<br>P-value <sup>b</sup> |
|-----------------|-------------|----------------------|------------------|------------------|--------------------|------------------|------------------------------|
|                 |             |                      | OR (95% CI)      | OR (95% CI)      | OR (95% CI)        | OR (95% CI)      |                              |
| Luminal-A-like  |             |                      |                  |                  |                    |                  |                              |
| 1q41            | rs4011894   | A/G                  | 1.05 (1.03-1.07) | 1.02 (0.98-1.06) | 1.05 (0.99-1.11)   | 1.00 (0.97-1.03) | 4.83E-02                     |
| 10p15.1         | rs55913768  | A/G                  | 1.04 (1.03-1.06) | 1.05 (1.02-1.09) | 1.04 (0.98-1.09)   | 1.00 (0.97-1.03) | 0.08                         |
| 10p14           | rs7100052   | G/T                  | 1.06 (1.04-1.09) | 1.05 (1.00-1.10) | 1.05 (0.98-1.13)   | 1.03 (0.99-1.07) | 0.59                         |
| 18q12.1         | rs16963218  | G/A                  | 0.91 (0.88-0.93) | 0.87 (0.81-0.93) | 0.89 (0.81-0.97)   | 0.89 (0.84-0.93) | 0.60                         |
| 20q13.32        | rs1105262   | G/T                  | 0.83 (0.78-0.88) | 0.88 (0.77-0.99) | 0.90 (0.77-1.05)   | 1.00 (0.91-1.09) | 5.06E-04                     |
| Luminal-B-like  |             |                      |                  |                  |                    |                  |                              |
| 10q26.13        | rs145680477 | C/T                  | 0.97 (0.87-1.08) | 1.72 (1.44-2.06) | -                  | 1.04 (0.90-1.21) | 1.37E-06                     |
| 12q23.1         | rs4762310   | A/G                  | 1.03 (0.93-1.14) | 0.60 (0.51-0.71) | 0.88 (0.69-1.12)   | 0.90 (0.79-1.03) | 2.13E-06                     |
| 2q31.1          | rs930313    | G/C                  | 1.05 (1.03-1.07) | 1.15 (1.10-1.20) | 1.04 (0.98-1.10)   | 1.00 (0.97-1.03) | 1.78E-05                     |
| Triple-negative |             |                      |                  |                  |                    |                  |                              |
| 1p36.22         | rs112200610 | T/C                  | 1.06 (1.02-1.10) | 0.99 (0.90-1.08) | 1.08 (0.96-1.23)   | 1.23 (1.15-1.31) | 4.06E-04                     |
| 1p35.2          | rs142302749 | G/A                  | 1.05 (0.95-1.15) | 1.19 (0.99-1.42) | 1.03 (0.81-1.32)   | 1.38 (1.24-1.55) | 2.36E-04                     |
| 11q25           | rs6590696   | C/A                  | 1.04 (1.00-1.09) | 1.04 (0.95-1.15) | 1.02 (0.90-1.15)   | 1.22 (1.14-1.31) | 2.28E-04                     |
| 2p24.1          | rs12710697  | T/C                  | 0.97 (0.96-0.98) | 0.97 (0.94-1.00) | 0.97 (0.93-1.02)   | 0.92 (0.90-0.95) | 6.21E-03                     |
| 5p15.32         | rs11952226  | C/T                  | 1.01 (0.96-1.05) | 1.08 (1.00-1.18) | 1.01 (0.91-1.12)   | 1.19 (1.13-1.26) | 1.15E-06                     |

**Footnote:**

Lead variants at loci not previously reported in GWAS of any breast cancer phenotype are in bold.  
Bonferroni threshold of significance corresponded to a p-value of  $1.25 \times 10^{-8}$

<sup>a</sup> Effect/Baseline alleles

<sup>b</sup> P-value computed from the heterogeneity test comparing estimates for each lead variant across subtypes.

Table 1.3 Novel TWAS genes for breast cancer intrinsic subtypes identified in cross-ancestry cell type-aware TWAS analyses (Bonferroni-adjusted p-value<0.05)

| Loci           | Gene Symbol          | Gene Type            | Strongest TWAS Association<br>Z-score (P-value) <sup>a</sup> | Cell Types <sup>b</sup>    | Ancestries <sup>c</sup> | Shows evidence for subtype-specificity <sup>d</sup> |
|----------------|----------------------|----------------------|--------------------------------------------------------------|----------------------------|-------------------------|-----------------------------------------------------|
| Luminal-A-like |                      |                      |                                                              |                            |                         |                                                     |
| 1p12           | <i>HSD3BP4</i>       | Pseudogene           | -6.11 (1.02E-09)                                             | Adip, Bulk                 | E                       | Yes                                                 |
| 1p12           | <i>CH17-472G23.2</i> | Pseudogene           | 7.79 (6.63E-15)                                              | Bulk, Epi, Fib             | E                       | No                                                  |
| 1p11.2         | <i>HIST2H2BA</i>     | Pseudogene           | 12.87 (6.64E-38)                                             | Bulk                       | C, E                    | Yes                                                 |
| 1p11.2         | <i>HIST2H3DP1</i>    | Pseudogene           | 13.97 (2.45E-44)                                             | Bulk, Endo                 | C, E                    | No                                                  |
| 1q22           | <i>MSTO2P</i>        | Pseudogene           | -6.07 (1.26E-09)                                             | Epi                        | E                       | No                                                  |
| 1q22           | <i>RIT1</i>          | Protein Coding       | -5.89 (3.78E-09)                                             | Bulk, Endo, Fib            | C, E                    | No                                                  |
| 2q33.1         | <i>ALS2CR12</i>      | Protein Coding       | 5.41 (6.23E-08)                                              | Fib                        | E                       | No                                                  |
| 2q35           | <i>XRCC5</i>         | Protein Coding       | 6.19 (6.20E-10)                                              | Bulk                       | E                       | Yes                                                 |
| 2q35           | <i>AC098820.3</i>    | Processed Transcript | 8.79 (1.46E-18)                                              | Adip, Bulk                 | E                       | Yes                                                 |
| 3p24.1         | <i>LRRC3B</i>        | Protein Coding       | -13.83 (1.59E-43)                                            | Adip, Bulk, Endo, Epi, Fib | E                       | No                                                  |
| 3p14.1         | <i>PSMD6-AS1</i>     | Antisense Transcript | -5.89 (3.82E-09)                                             | Epi                        | C, E                    | Yes                                                 |
| 3q25.31        | <i>METTL15P1</i>     | Pseudogene           | -5.31 (1.09E-07)                                             | Bulk                       | E                       | No                                                  |
| 3q25.31        | <i>PA2G4P4</i>       | Pseudogene           | -5.47 (4.51E-08)                                             | Epi                        | C, E                    | No                                                  |
| 4q21.23        | <i>SLC25A14P1</i>    | Pseudogene           | -5.72 (1.06E-08)                                             | Bulk, Epi                  | C, E                    | No                                                  |
| 5q11.1         | <i>PARP8</i>         | Protein Coding       | 5.39 (6.90E-08)                                              | Endo                       | E                       | No                                                  |
| 5q11.2         | <i>CTD-2310F14.1</i> | Antisense Transcript | 15.29 (9.22E-53)                                             | Adip, Bulk, Endo           | C, E                    | No                                                  |
| 5q11.2         | <i>ACTBL2</i>        | Protein Coding       | 7.67 (1.67E-14)                                              | Endo                       | E                       | No                                                  |
| 5q14.1         | <i>ATG10-IT1</i>     | sense_intronic       | -5.72 (1.10E-08)                                             | Bulk                       | E                       | No                                                  |
| 7q34           | <i>JHDM1D-AS1</i>    | Antisense Transcript | 6.77 (1.28E-11)                                              | Bulk                       | C, E                    | No                                                  |
| 8p11.23        | <i>ADGRA2</i>        | Protein Coding       | -7.75 (9.30E-15)                                             | Adip                       | C, E                    | No                                                  |
| 8q21.13        | <i>PI15</i>          | Protein Coding       | -7.42 (1.21E-13)                                             | Adip, Endo                 | C, E                    | No                                                  |
| 8q24.13        | <i>RNF139</i>        | Protein Coding       | -5.74 (9.23E-09)                                             | Bulk                       | E                       | No                                                  |
| 8q24.21        | <i>PVT1</i>          | lincRNA              | -6.24 (4.51E-10)                                             | Fib                        | E                       | No                                                  |
| 9p21.3         | <i>RN7SL151P</i>     | Miscellaneous RNA    | -5.67 (1.40E-08)                                             | Bulk                       | E                       | No                                                  |
| 9p21.3         | <i>CDKN2A</i>        | Protein Coding       | 5.49 (3.93E-08)                                              | Adip, Bulk                 | C, E                    | No                                                  |
| 9q33.1         | <i>TRIM32</i>        | Protein Coding       | 6.27 (3.58E-10)                                              | Bulk                       | C, E                    | No                                                  |
| 10p15.1        | <i>GDI2</i>          | Protein Coding       | -5.46 (4.83E-08)                                             | Bulk, Epi                  | C                       | No                                                  |
| 10q25.3        | <i>NHLRC2</i>        | Protein Coding       | 6.28 (3.36E-10)                                              | Endo                       | C, E                    | No                                                  |
| 10q26.13       | <i>RP11-500G22.4</i> | Antisense Transcript | 5.73 (9.99E-09)                                              | Epi, Fib                   | E                       | Yes                                                 |
| 10q26.13       | <i>DMBT1</i>         | Protein Coding       | 13.74 (6.25E-43)                                             | Bulk, Endo                 | C, E                    | No                                                  |
| 11q13.2        | <i>CTSF</i>          | Protein Coding       | 6.41 (1.48E-10)                                              | Adip                       | E                       | No                                                  |
| 12p12.3        | <i>HIST4H4</i>       | Protein Coding       | -5.92 (3.26E-09)                                             | Bulk                       | C, E                    | No                                                  |
| 12p11.22       | <i>MANSC4</i>        | Protein Coding       | 5.66 (1.54E-08)                                              | Bulk                       | E                       | No                                                  |
| 12p11.22       | <i>RP11-425D17.1</i> | Antisense Transcript | 7.77 (7.93E-15)                                              | Bulk, Endo                 | E                       | No                                                  |
| 14q24.1        | <i>ACTN1</i>         | Protein Coding       | 5.80 (6.71E-09)                                              | Bulk                       | C, E                    | Yes                                                 |

Continuation of Table 1.3

|                 |                      |                              |                  |            |      |     |
|-----------------|----------------------|------------------------------|------------------|------------|------|-----|
| 16q12.2         | <i>RBL2</i>          | Protein Coding               | -6.71 (2.01E-11) | Endo       | C, E | No  |
| 16q23.2         | <i>ATMIN</i>         | Protein Coding               | -7.19 (6.33E-13) | Fib        | C, E | No  |
| 16q23.2         | <i>GAN</i>           | Protein Coding               | 8.75 (2.09E-18)  | Bulk       | C, E | No  |
| 18q11.2         | <i>RP11-107K17.2</i> | Pseudogene                   | -5.69 (1.26E-08) | Bulk       | C, E | No  |
| 19p13.13        | <i>JUNB</i>          | Protein Coding               | 5.74 (9.36E-09)  | Bulk       | E    | Yes |
| 19p13.11        | <i>B3GNT3</i>        | Protein Coding               | 6.40 (1.60E-10)  | Adip       | C, E | No  |
| 19q13.31        | <i>RP11-15A1.2</i>   | Antisense Transcript         | 5.34 (9.10E-08)  | Bulk       | E    | No  |
| 21q11.2         | <i>AF127936.7</i>    | lincRNA                      | -6.66 (2.69E-11) | Bulk       | C, E | No  |
| 22q12.1         | <i>MN1</i>           | Protein Coding               | 7.31 (2.62E-13)  | Bulk       | C, E | Yes |
| 22q12.2         | <i>EMID1</i>         | Protein Coding               | -6.05 (1.41E-09) | Fib        | E    | No  |
| 22q13.2         | <i>CHADL</i>         | Protein Coding               | -6.52 (6.86E-11) | Bulk, Endo | C, E | No  |
| Luminal-B-like  |                      |                              |                  |            |      |     |
| 5q11.2          | <i>CTD-2310F14.1</i> | Miscellaneous RNA Transcript | 5.61 (2.03E-08)  | Bulk       | E    | No  |
| Triple-negative |                      |                              |                  |            |      |     |
| 12p11.22        | <i>RP11-425D17.1</i> | Miscellaneous RNA Transcript | 5.49 (4.04E-08)  | Bulk, Endo | E    | No  |
| 19p13.11        | <i>PLVAP</i>         | Protein Coding               | -6.46 (1.07E-10) | Fib        | C, E | Yes |
| 19p13.11        | <i>KCNN1</i>         | Protein Coding               | -5.33 (9.86E-08) | Fib        | E    | No  |

**Footnote:**

Bonferroni threshold of significance corresponded to a p-value of  $1.42 \times 10^{-7}$  for European-ancestry TWAS,  $1.39 \times 10^{-7}$  for African-ancestry TWAS, and  $1.26 \times 10^{-7}$  for cross-ancestry TWAS

<sup>a</sup> Strongest association is defined as the TWAS association with the minimum p-value across all ancestries and cell types

<sup>b</sup> Cell type abbreviations: Bulk (Bulk-tissue), Epi (Epithelial cells), Adip (Adipocytes), Endo (Endothelial cells), Fib (Fibroblasts)

<sup>c</sup> Ancestry abbreviations: E (European-ancestry), A (African-ancestry), C (Cross-ancestry)

<sup>d</sup> Subtype specificity is defined as the minimum p-value of TWAS associations across all ancestries and cell types is above 0.05 for all other subtypes

Table 1.4 TWAS genes with signals that were not explained by GWAS index variants

| Loci                                                                       | Gene Symbol          | Gene Type            | Strongest TWAS Association Z-score (P-value) <sup>a</sup> | Strongest Cond-TWAS Association Z-score (P-value) <sup>b</sup> | Cell Types <sup>c</sup>           | Ancestries <sup>d</sup> | Nearest GWAS Index Variant | Distance to Nearest Index Variant (kb) |
|----------------------------------------------------------------------------|----------------------|----------------------|-----------------------------------------------------------|----------------------------------------------------------------|-----------------------------------|-------------------------|----------------------------|----------------------------------------|
| TWAS genes located > 2 Mb from nearest GWAS lead variant                   |                      |                      |                                                           |                                                                |                                   |                         |                            |                                        |
| Luminal-A-like                                                             |                      |                      |                                                           |                                                                |                                   |                         |                            |                                        |
| 3q25.31                                                                    | <i>METTL15P1</i>     | Pseudogene           | -5.31 (1.09E-07)                                          | -                                                              | Bulk                              | E                       | rs6440006                  | 15290.03                               |
| 3q25.31                                                                    | <i>PA2G4P4</i>       | Pseudogene           | -5.47 (4.51E-08)                                          | -                                                              | Epi                               | C, E                    | rs6440006                  | 15385.7                                |
| Conditionally-significant TWAS genes within +/- 2 Mb of GWAS lead variants |                      |                      |                                                           |                                                                |                                   |                         |                            |                                        |
| Luminal-A-like                                                             |                      |                      |                                                           |                                                                |                                   |                         |                            |                                        |
| 2q35                                                                       | <i>AC098820.3</i>    | Processed Transcript | 8.79 (1.46E-18)                                           | 9.52 (1.73E-21)                                                | Adip, Bulk <sup>e</sup>           | E                       | rs4442975                  | 557.28                                 |
| 12p11.22                                                                   | <i>RP11-967K21.1</i> | Antisense Transcript | 8.20 (2.41E-16)                                           | 7.40 (1.37E-13)                                                | Adip, Bulk, Epi, Fib <sup>e</sup> | C, E                    | rs7297051                  | 141.41                                 |
| 12p11.22                                                                   | <i>RP11-425D17.1</i> | Antisense Transcript | 7.77 (7.93E-15)                                           | 7.26 (3.73E-13)                                                | Bulk, Endo <sup>e</sup>           | E                       | rs7297051                  | 163.74                                 |
| 18q11.2                                                                    | <i>RP11-107K17.2</i> | Pseudogene           | -5.69 (1.26E-08)                                          | -5.26 (1.42E-07)                                               | Bulk <sup>e</sup>                 | E                       | rs170801                   | 571.93                                 |
| 22q13.1                                                                    | <i>APOBEC3B</i>      | Protein Coding       | -7.48 (7.48E-14)                                          | -6.39 (1.67E-10)                                               | Adip, Bulk, Endo <sup>e</sup>     | C, E                    | rs11703601                 | 780.31                                 |
| Triple-negative                                                            |                      |                      |                                                           |                                                                |                                   |                         |                            |                                        |
| 12p11.22                                                                   | <i>RP11-967K21.1</i> | Antisense Transcript | 5.57 (2.53E-08)                                           | 5.47 (4.57E-08)                                                | Epi <sup>e</sup>                  | E                       | rs7297051                  | 141.41                                 |

**Footnote:**

Conditionally-significant TWAS genes were defined by a similar Bonferroni threshold of significance utilized in the primary TWAS which corresponded to a p-value of  $1.42 \times 10^{-7}$  for European-ancestry TWAS,  $1.39 \times 10^{-7}$  for African-ancestry TWAS, and  $1.26 \times 10^{-7}$  for cross-ancestry TWAS

<sup>a</sup> Strongest association is defined as the TWAS association with the minimum p-value across all ancestries and cell types

<sup>b</sup> Strongest conditional TWAS (Cond-TWAS) association is defined as the cond-TWAS association with the minimum p-value across all ancestries and cell types

<sup>c</sup> Cell type abbreviations: Bulk (Bulk-tissue), Epi (Epithelial cells), Adip (Adipocytes), Endo (Endothelial cells), Fib (Fibroblasts)

<sup>d</sup> Ancestry abbreviations: E (European-ancestry), A (African-ancestry), C (Cross-ancestry)

<sup>e</sup> Cell types with Bonferroni-significant Cond-TWAS associations

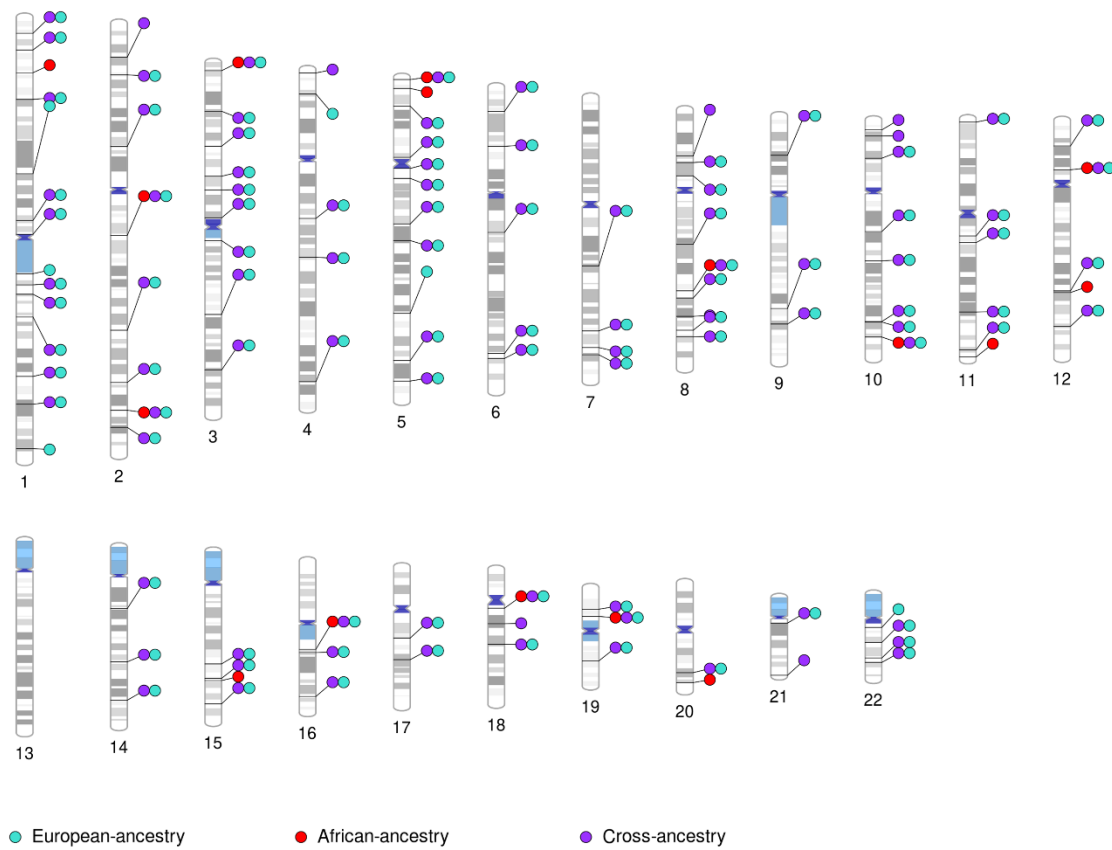


Figure S 1.1 Ideogram of all GWAS loci identified to associate with intrinsic subtypes in this study by ancestry.

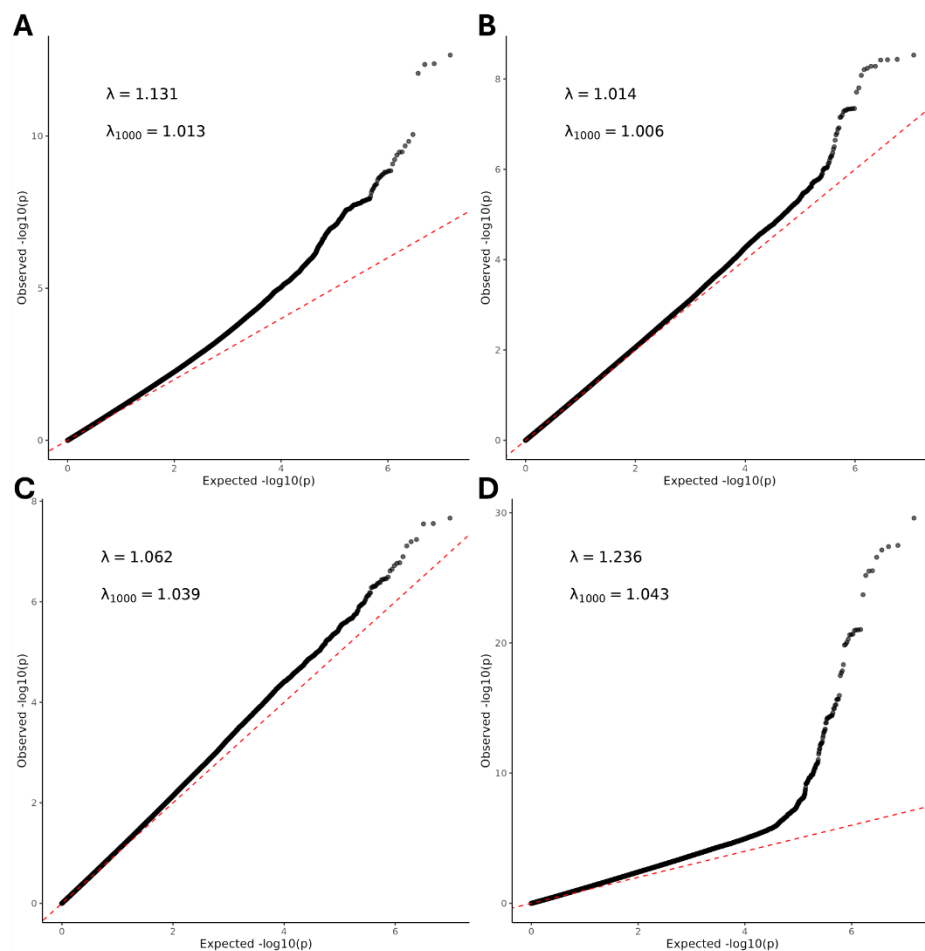


Figure S 1.2 QQ-plots and inflation factors for the African-ancestry GWAS meta-analysis for each intrinsic subtype. A) Luminal-A-like, B) Luminal-B-like, C) HER2-enriched-like, and D) Triple-negative breast cancer.

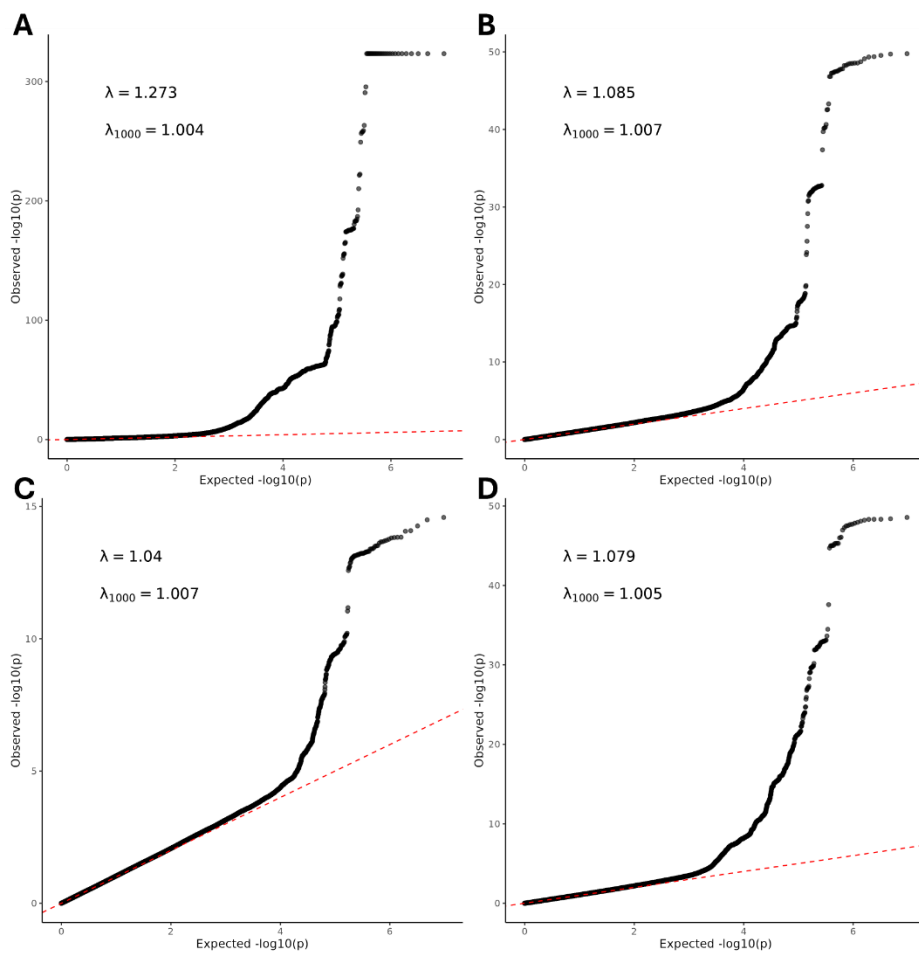


Figure S 1.3 QQ-plots and inflation factors for the European-ancestry GWAS meta-analysis for each intrinsic subtype. A) Luminal-A-like, B) Luminal-B-like, C) HER2-enriched-like, and D) Triple-negative breast cancer.



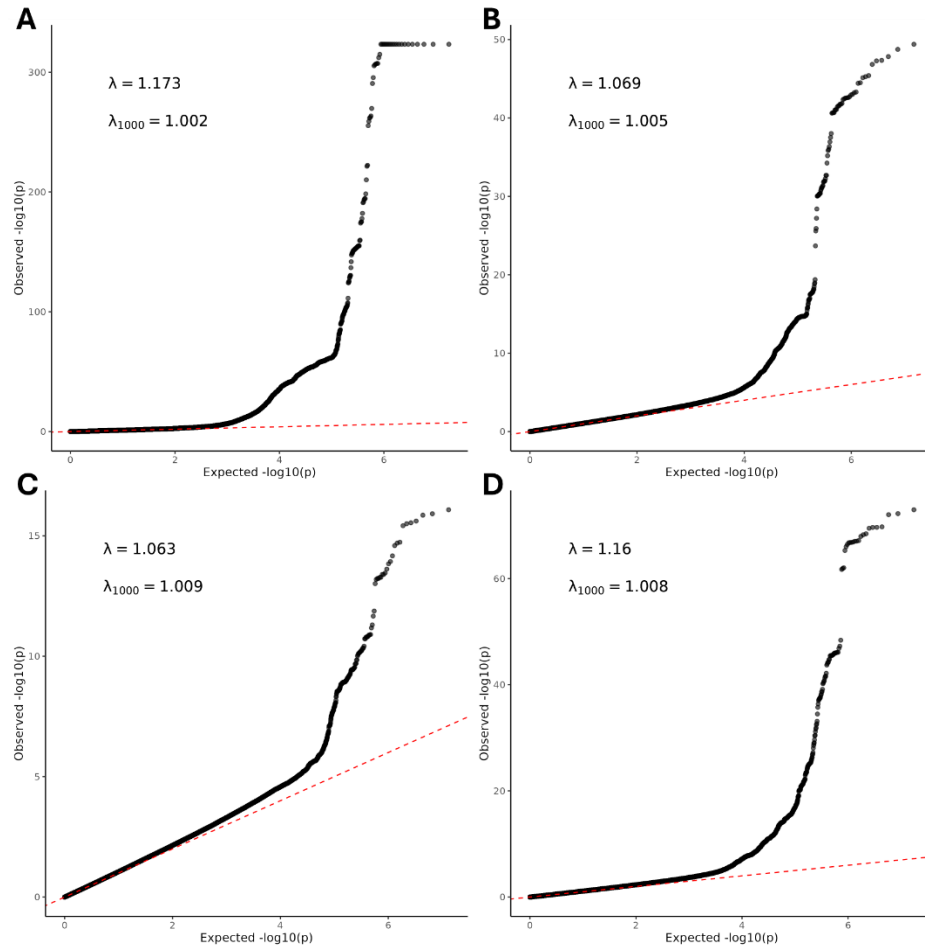


Figure S 1.4 QQ-plots and inflation factors for the cross-ancestry GWAS meta-analysis for each intrinsic subtype. A) Luminal-A-like, B) Luminal-B-like, C) HER2-enriched-like, and D) Triple-negative breast cancer.

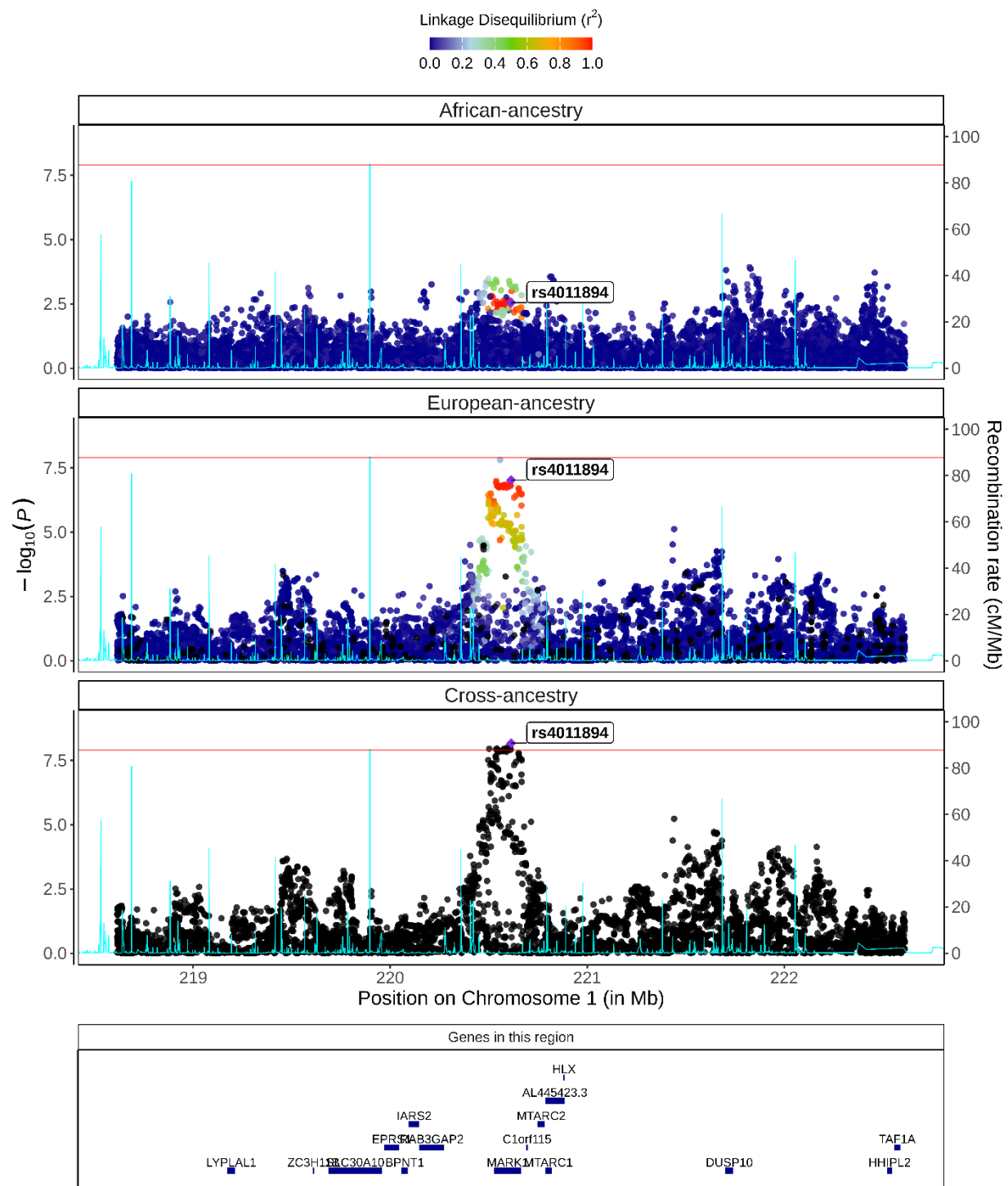


Figure S 1.5A Regional plot of GWAS associations at the 1q41 locus (lead variant, rs4011894) for Luminal-A-like breast cancer.

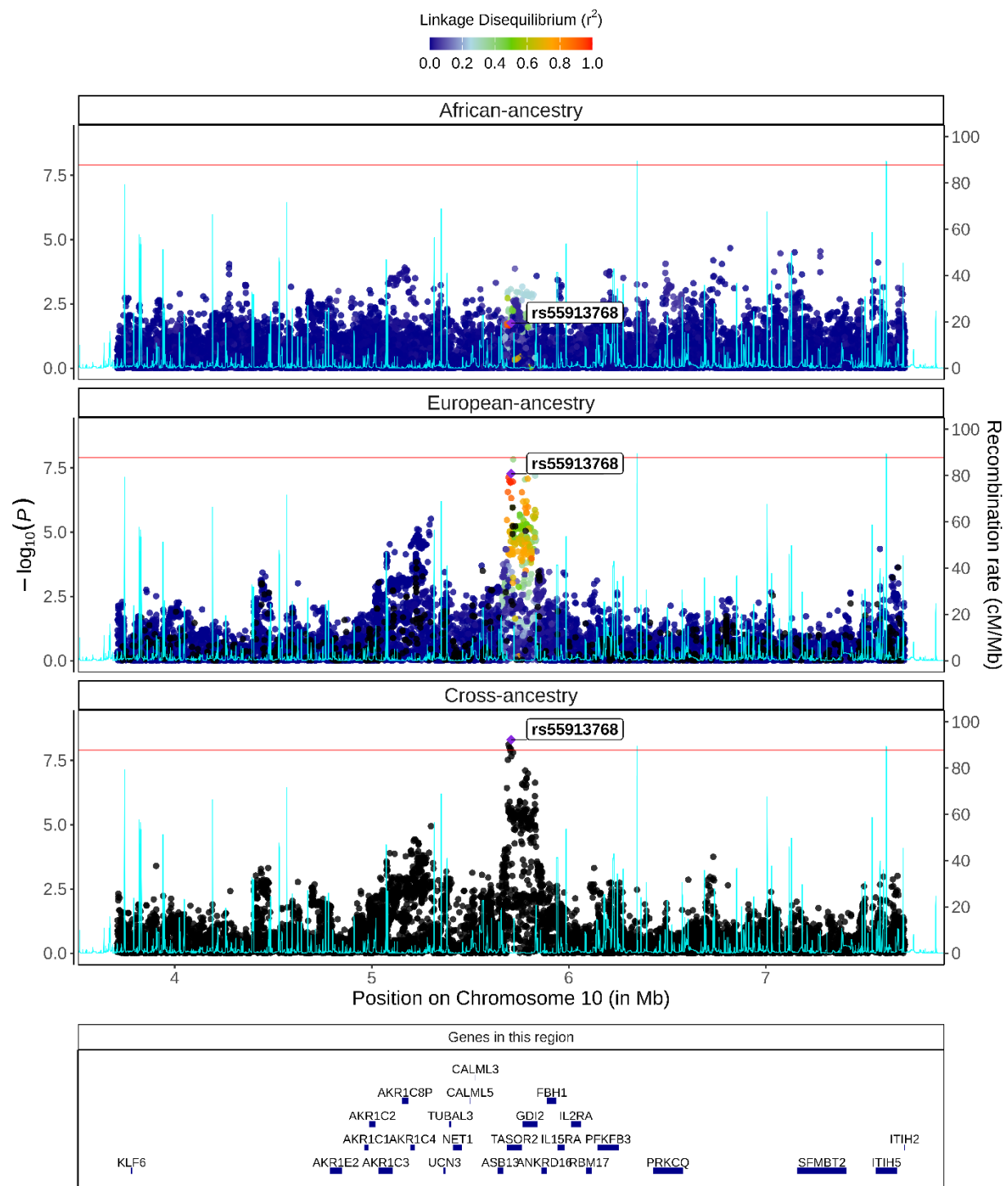


Figure S 1.5B. Regional plot of GWAS associations at the 10p15.1 locus (lead variant, rs55913768) for Luminal-A-like breast cancer.

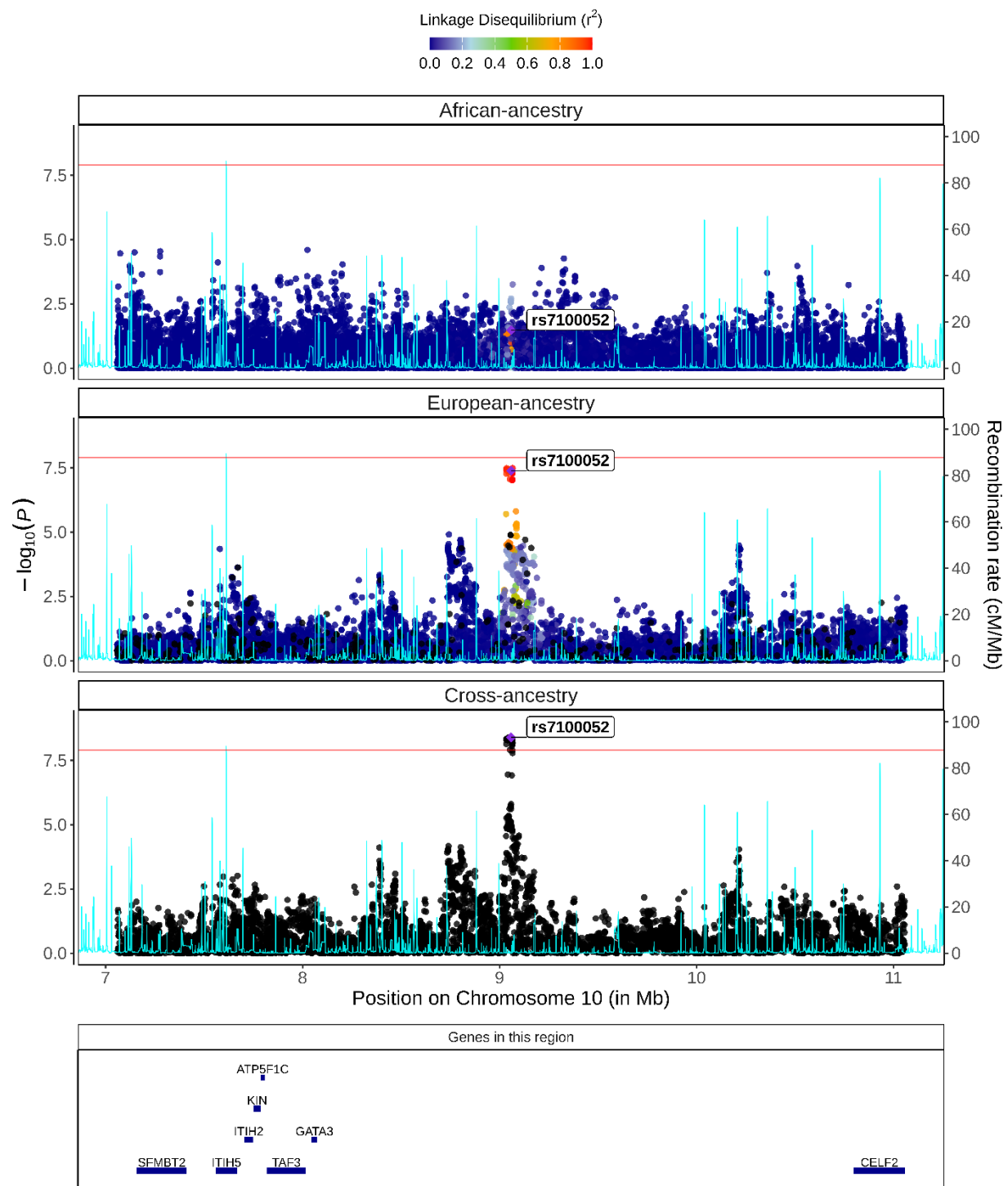


Figure S 1.5C. Regional plot of GWAS associations at the 10p14 locus (lead variant, rs7100052) for Luminal-A-like breast cancer.

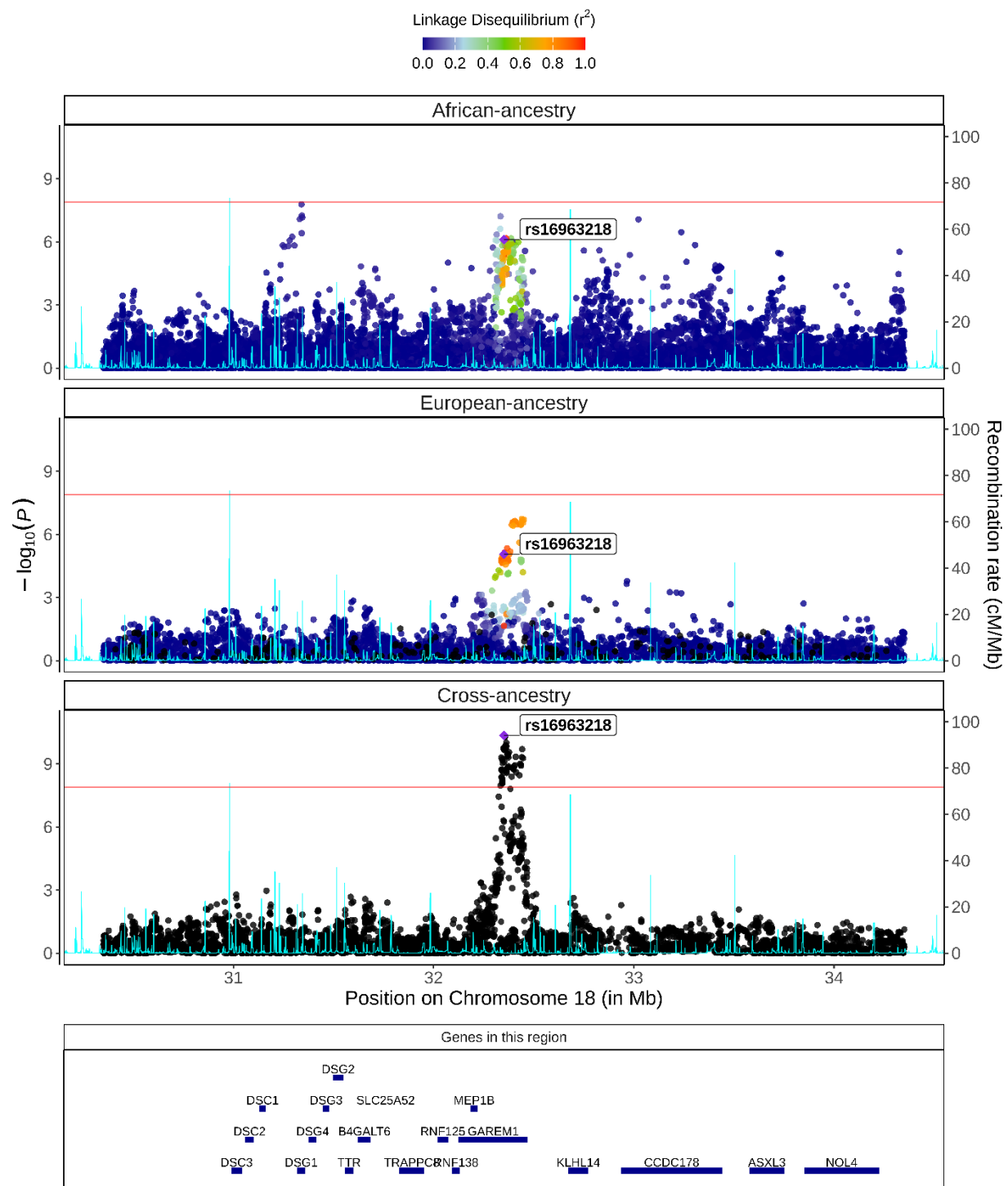


Figure S 1.5D. Regional plot of GWAS associations at the 18q12.1 locus (lead variant, rs16963218) for Luminal-A-like breast cancer.

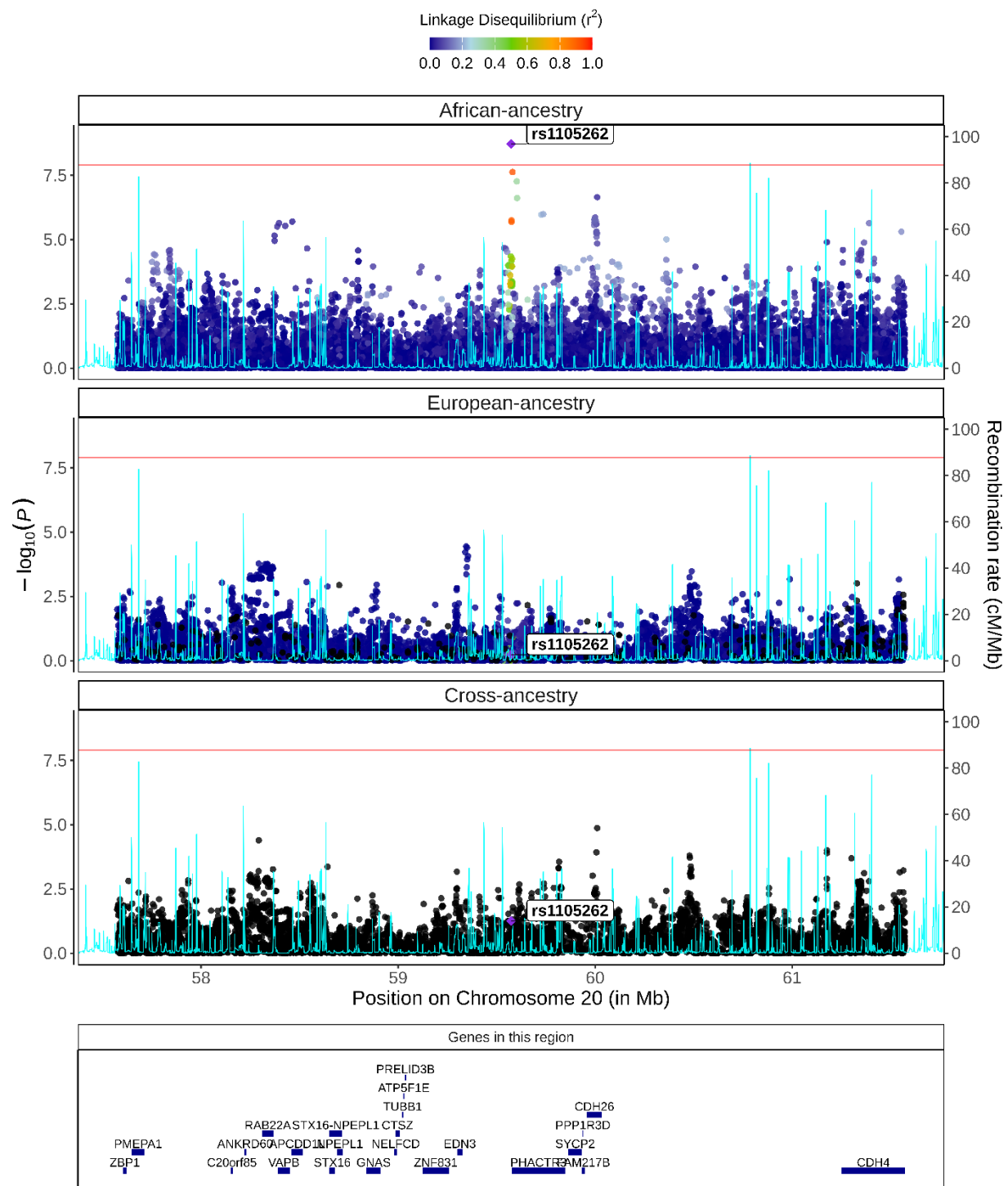


Figure S 1.5E. Regional plot of GWAS associations at the 20q13.32 locus (lead variant, rs1105262) for Luminal-A-like breast cancer.

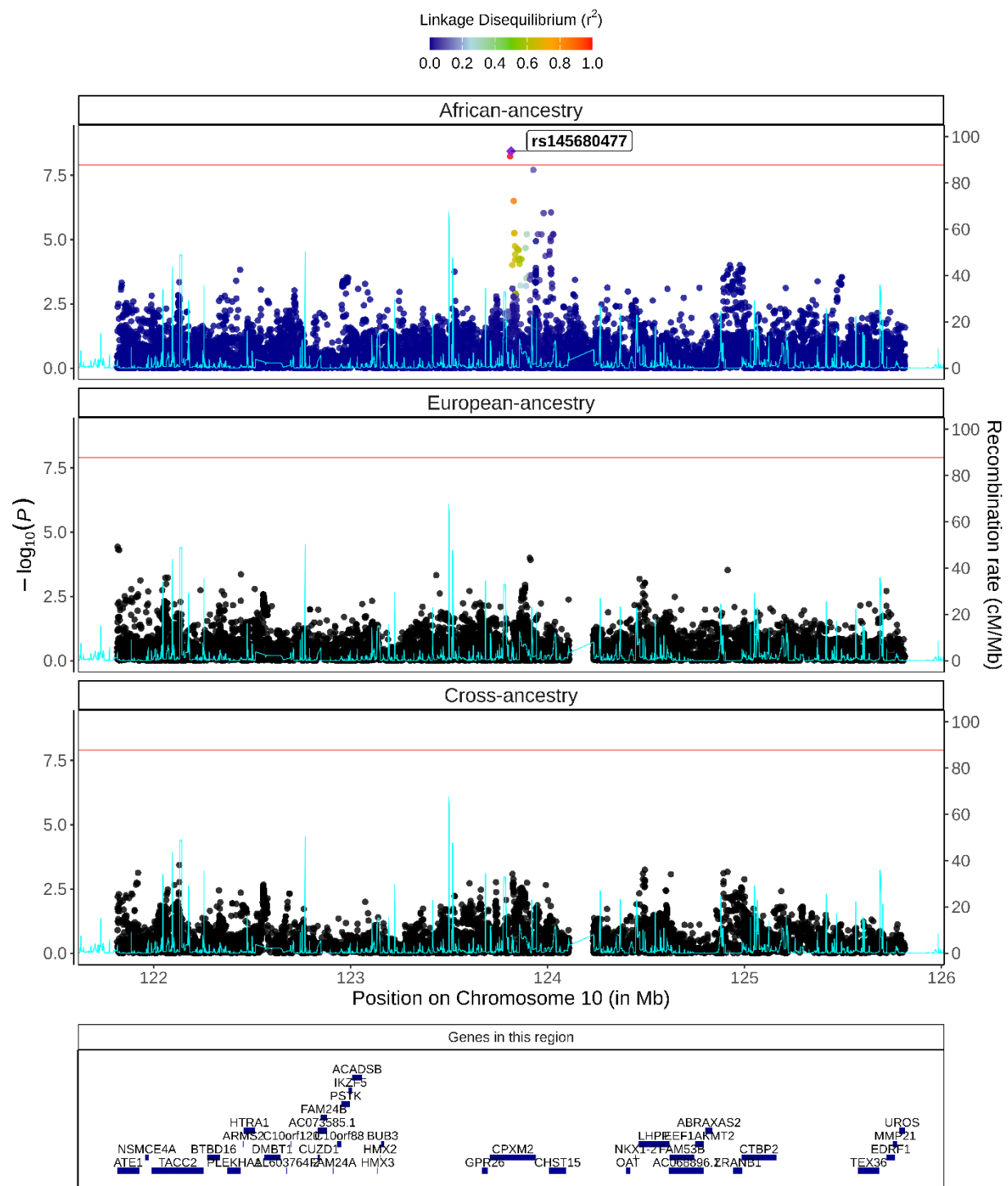


Figure S 1.5F. Regional plot of GWAS associations at the 10q26.13 locus (lead variant, rs145680477) for Luminal-B-like breast cancer.

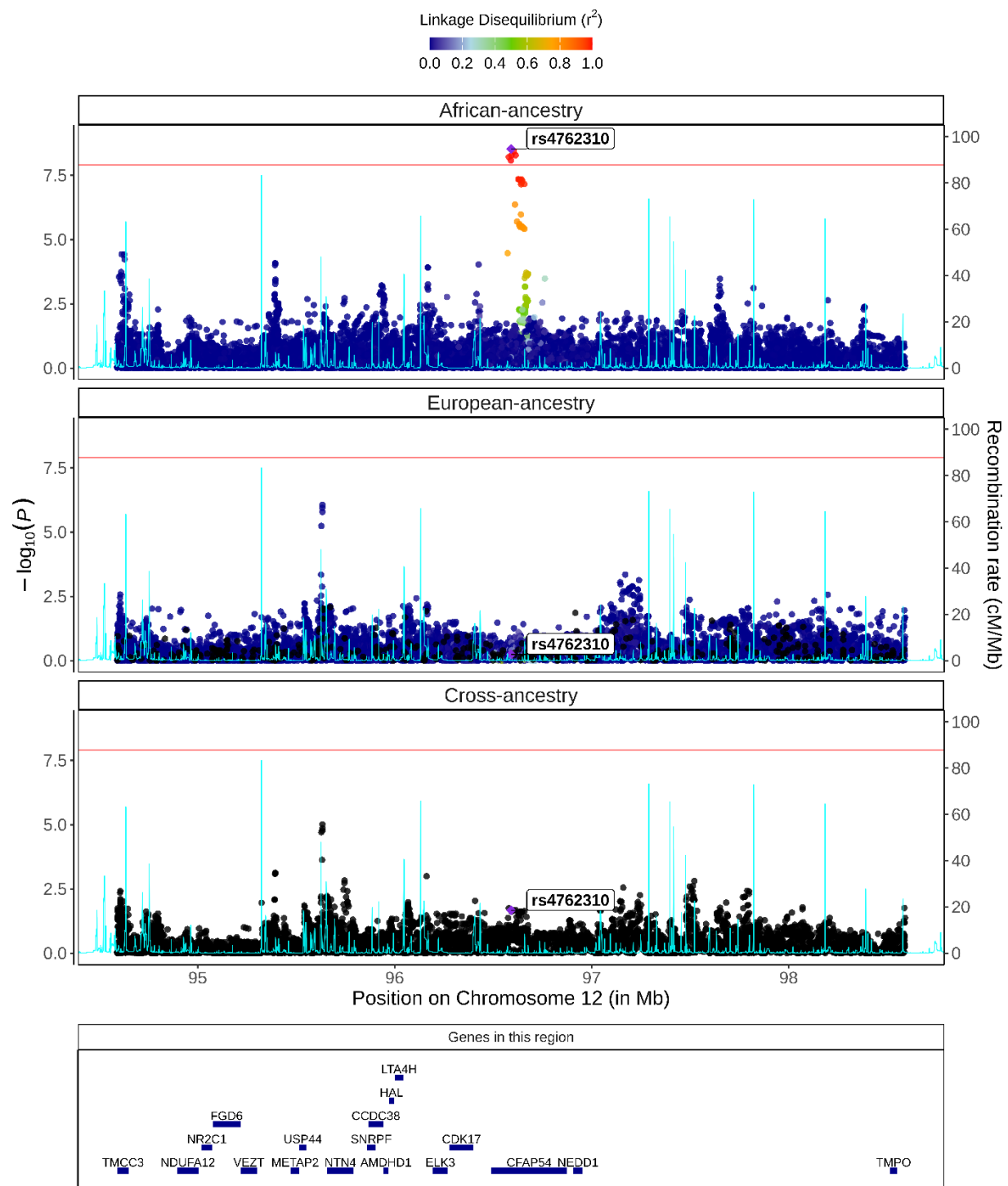


Figure S 1.5G. Regional plot of GWAS associations at the 12q23.1 locus (lead variant, rs4762310) for Luminal-B-like breast cancer.



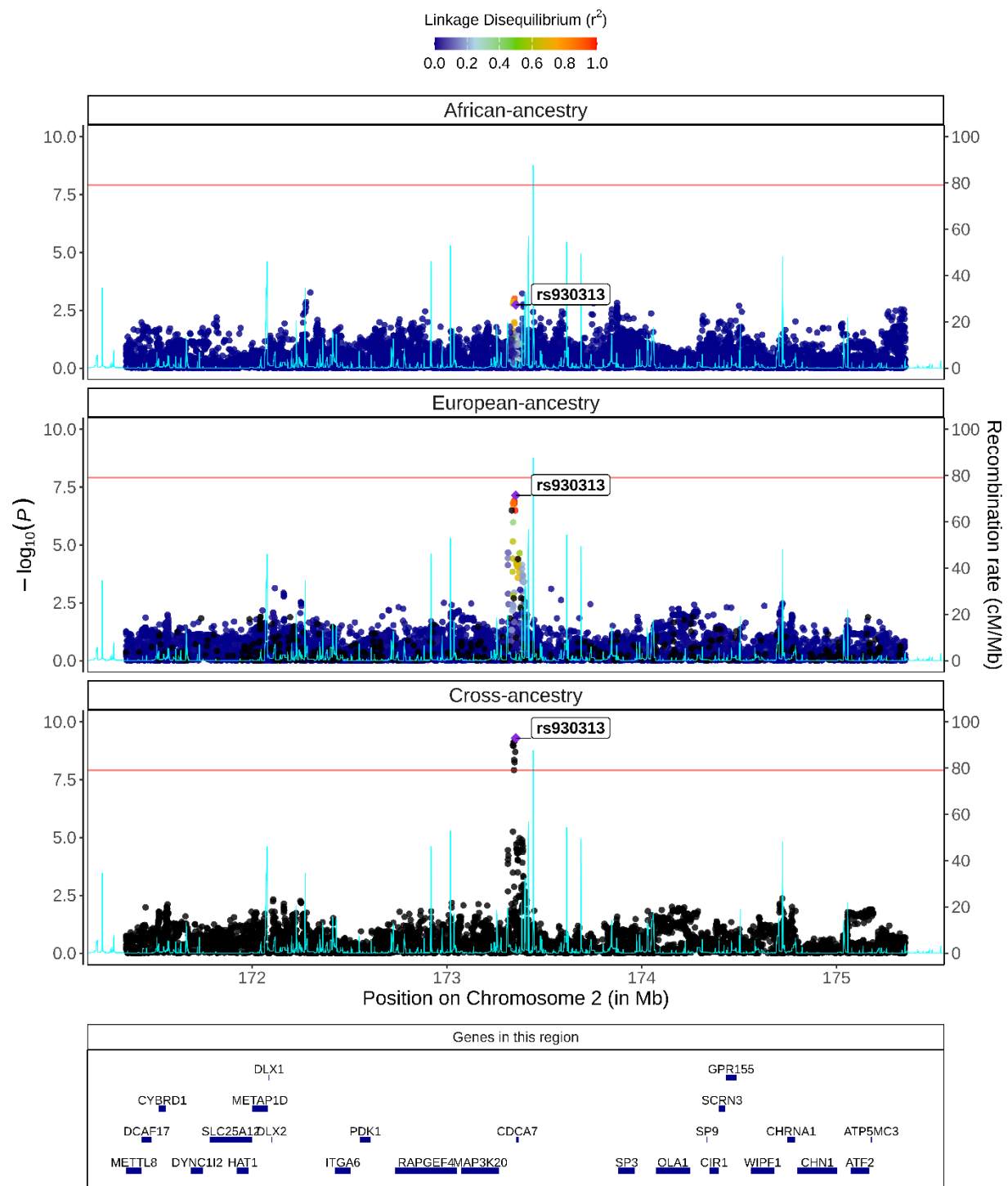


Figure S 1.5H. Regional plot of GWAS associations at the 2q31.1 locus (lead variant, rs930313) for Luminal-B-like breast cancer.

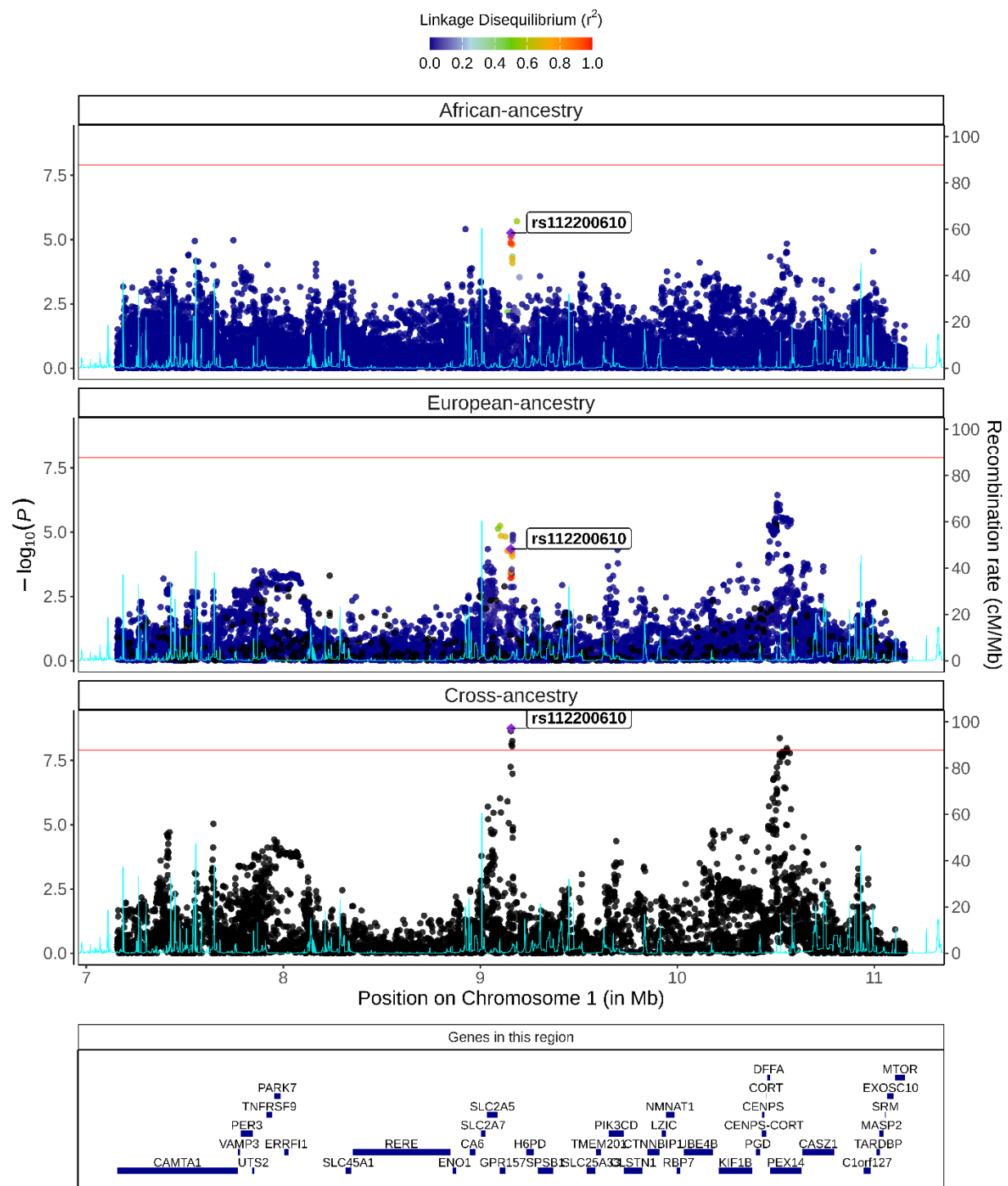


Figure S 1.5I. Regional plot of GWAS associations at the 1p36.22 locus (lead variant, rs112200610) for triple-negative breast cancer.



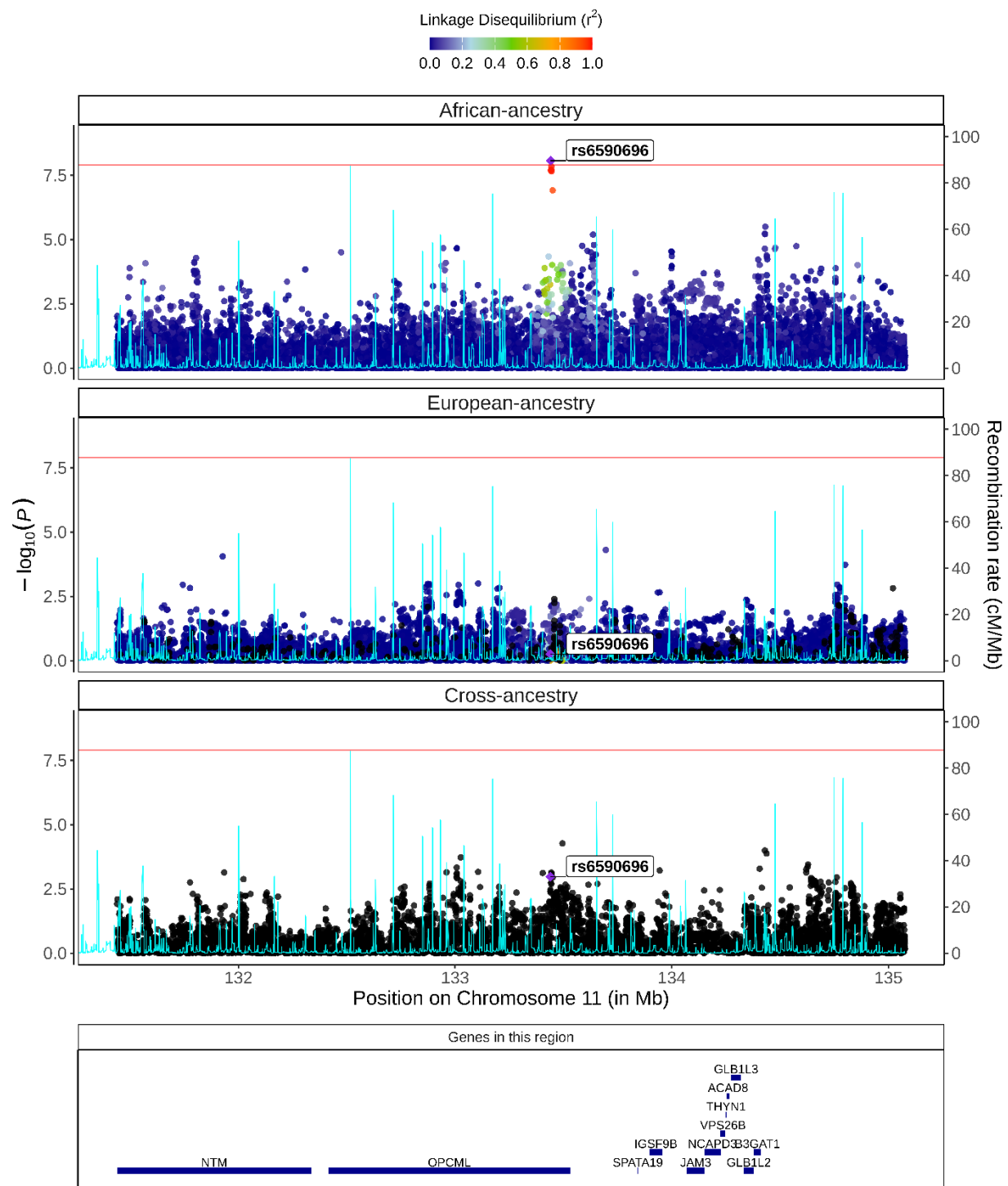


Figure S 1.5K. Regional plot of GWAS associations at the 11q25 locus (lead variant, rs6590696) for triple-negative breast cancer.

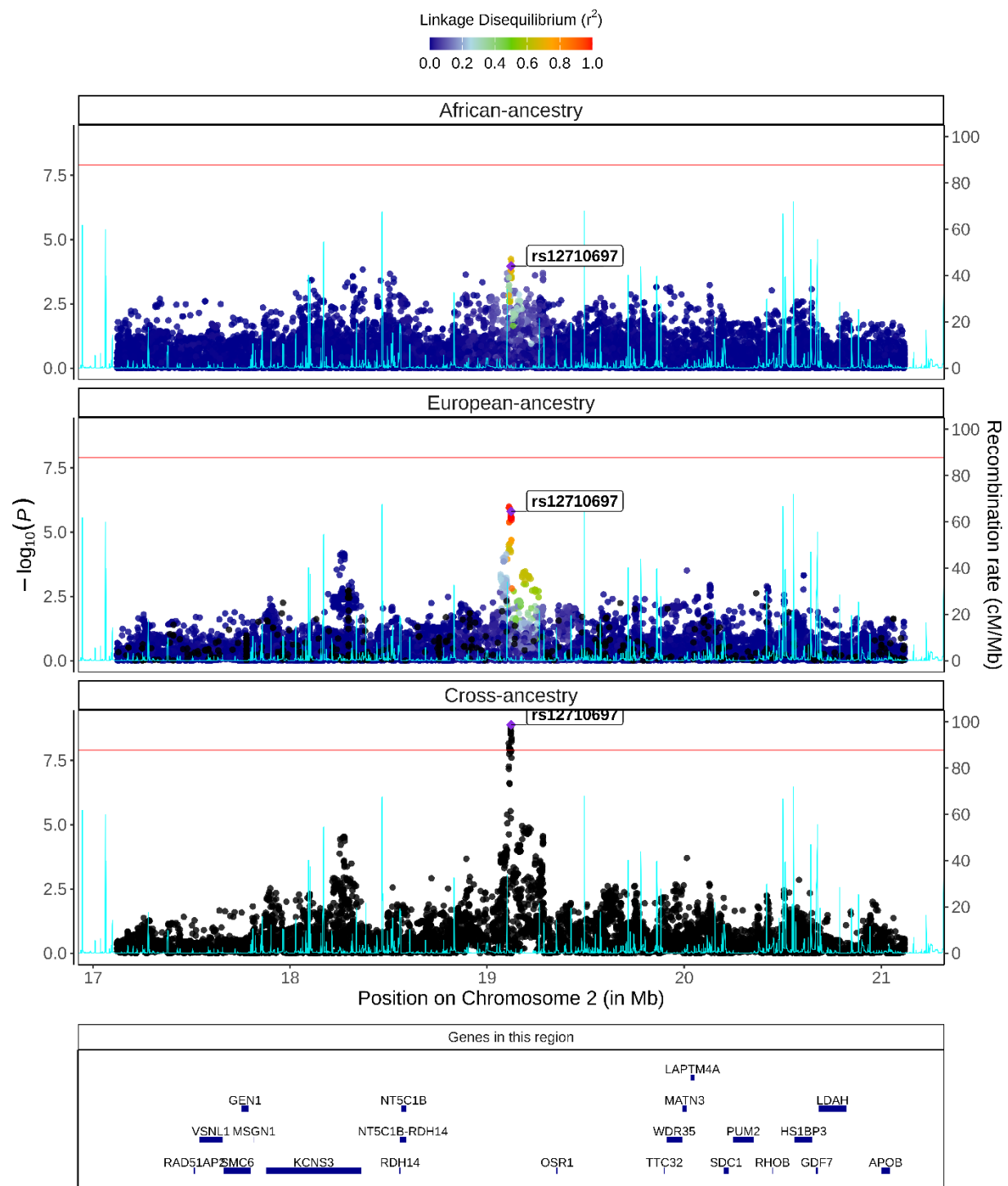


Figure S 1.5L. Regional plot of GWAS associations at the 2p24.1 locus (lead variant, rs12710697) for triple-negative breast cancer.

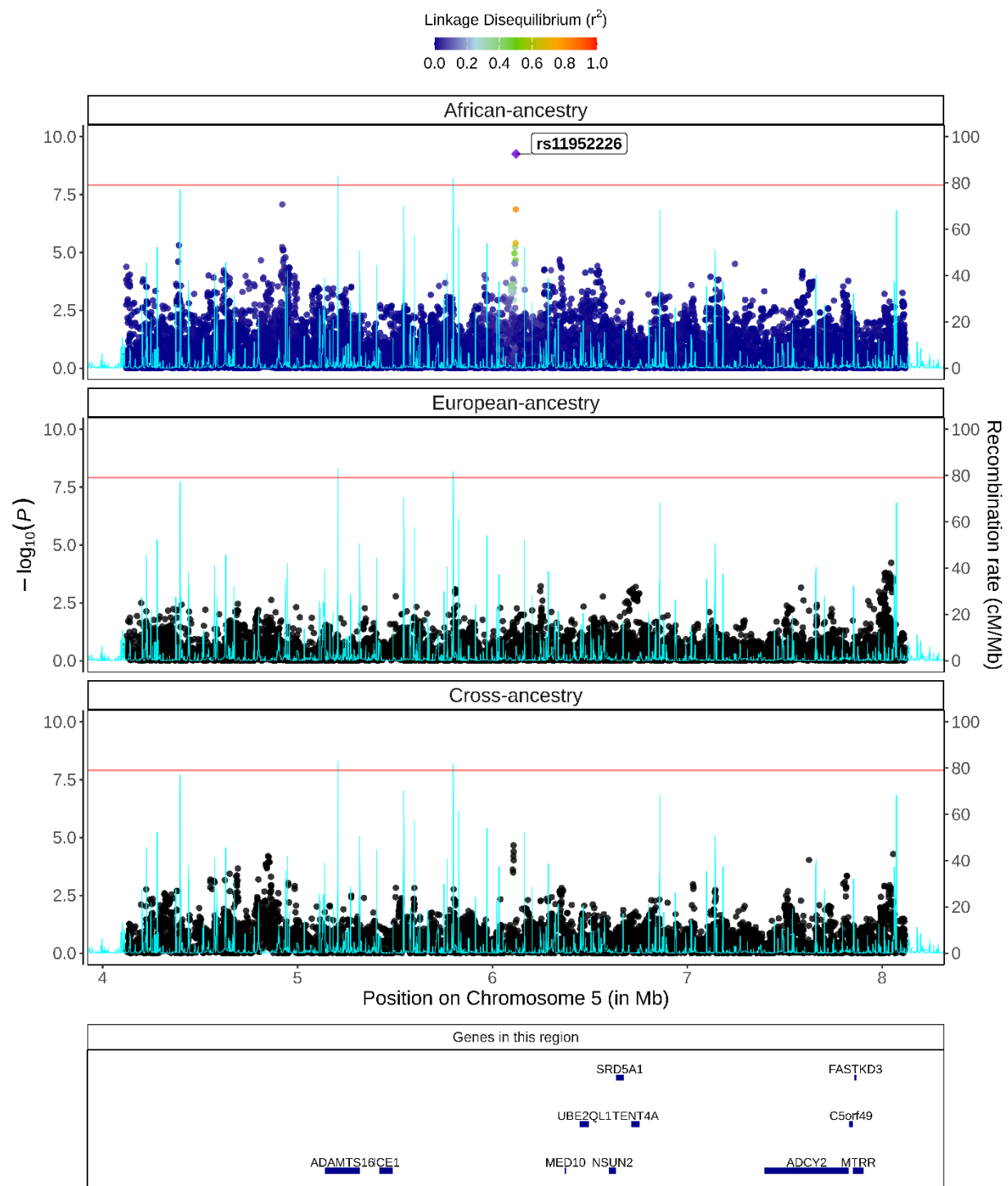


Figure S 1.5M. Regional plot of GWAS associations at the 5p15.32 locus (lead variant, rs11952226) for triple-negative breast cancer.

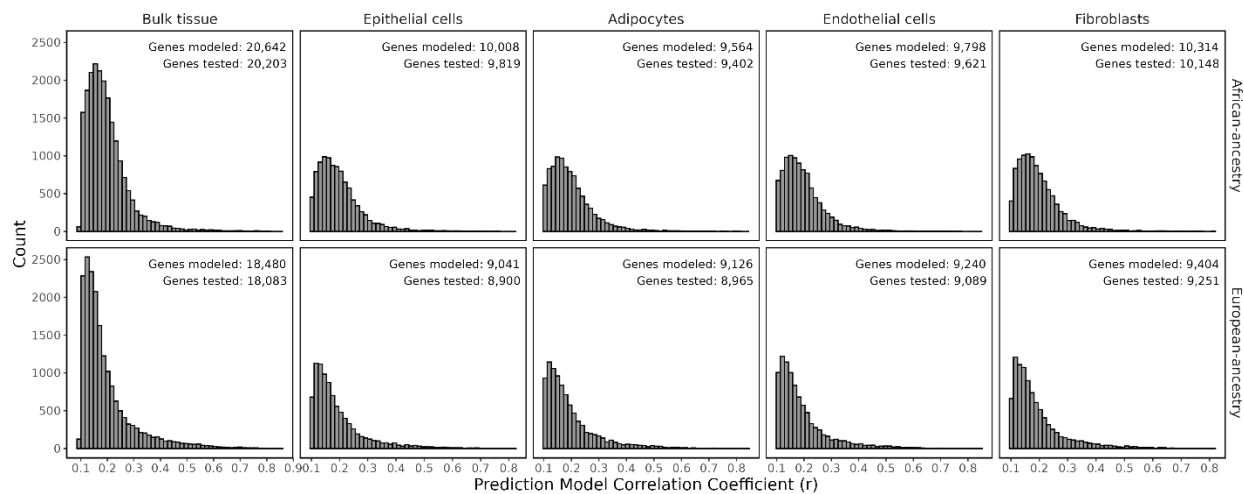


Figure S 1.6 Predictive performance for gene expression models built in this study for TWAS analyses

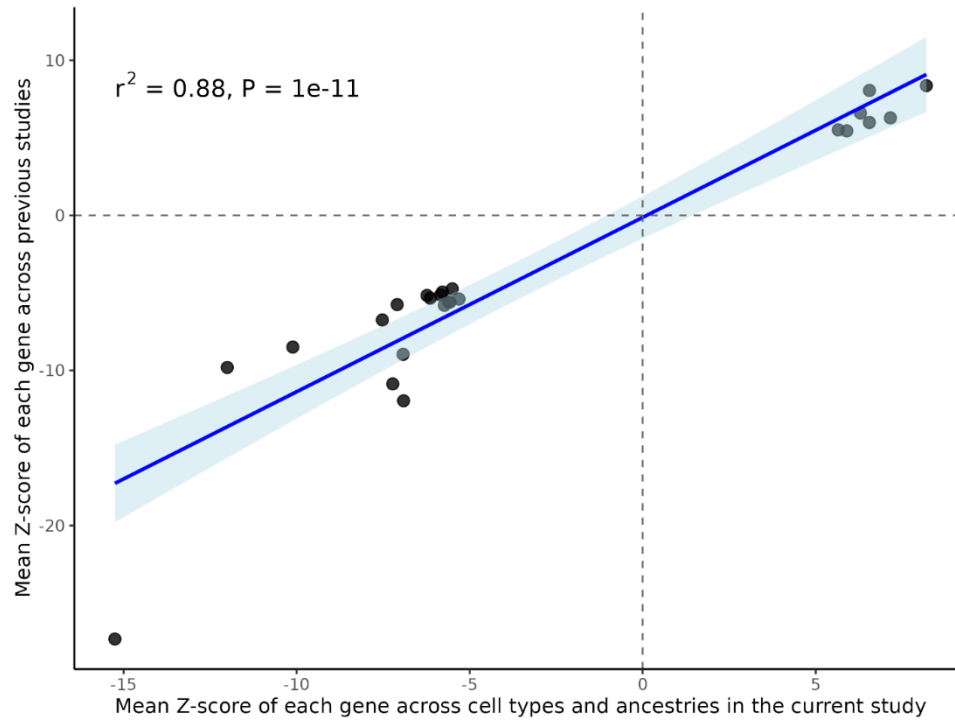


Figure S 1.7 Correlation of mean association estimates (Z-scores) for genes identified in both the current study and previous breast tissue TWAS of intrinsic subtypes



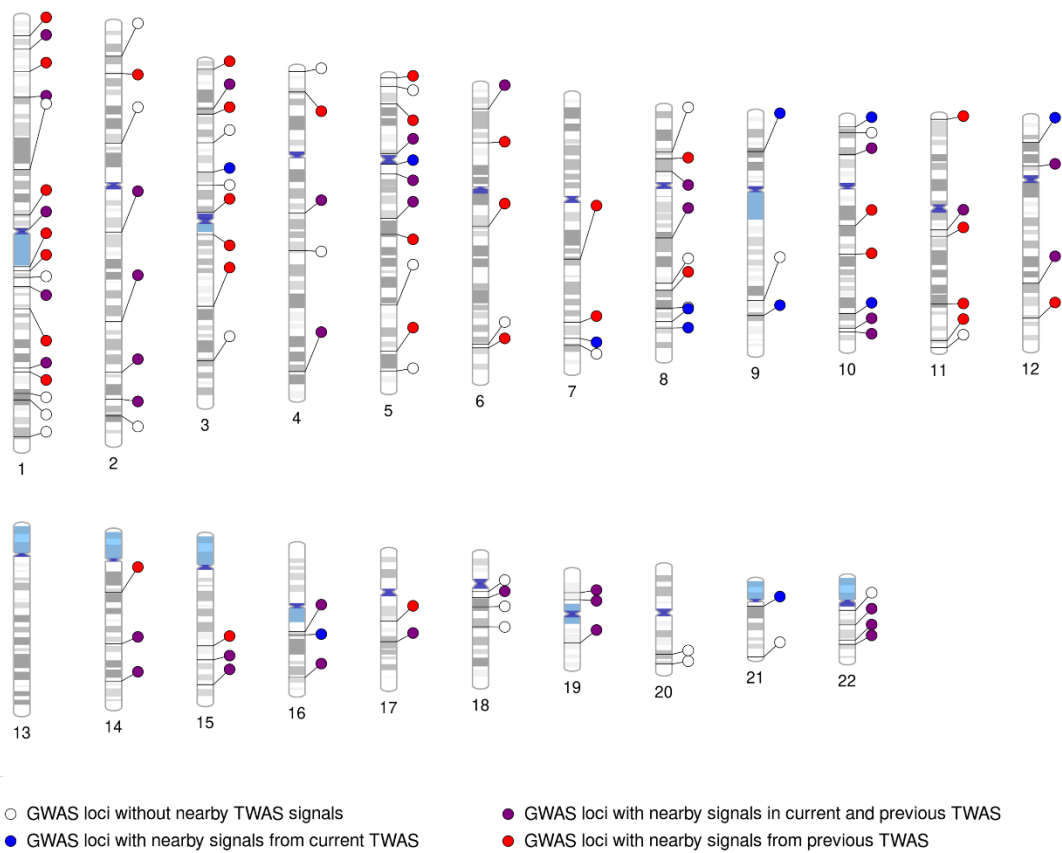


Figure S 1.8 Ideogram of GWAS loci color-coded by loci without a nearby TWAS signal, loci with a nearby TWAS signal identified in the present study, and loci with a TWAS signal identified in prior TWAS of intrinsic subtypes.

Table S 1.1 Studies in AABCG included in the African-ancestry GWAS meta-analysis performed in this study.

| Dataset        | Study   | Cases by subtype |                |                    |                 |                             | Total cases | Controls |
|----------------|---------|------------------|----------------|--------------------|-----------------|-----------------------------|-------------|----------|
|                |         | Luminal-A-like   | Luminal-B-like | HER2-enriched-like | Triple-negative | Missing subtype information |             |          |
| AABC           | CARE    | 85               | 8              | 4                  | 31              | 126                         | 254         | 204      |
|                | CBCS    | 216              | 31             | 34                 | 193             | 140                         | 614         | 570      |
|                | MEC     | 102              | 40             | 12                 | 45              | 379                         | 578         | 888      |
|                | NBHS    | 56               | 24             | 6                  | 20              | 149                         | 255         | 161      |
|                | NC-BCFR | 95               | 35             | 19                 | 40              | 194                         | 383         | 48       |
|                | PLCO    | 6                | 3              | 0                  | 2               | 43                          | 54          | 112      |
|                | SFBCS   | 3                | 3              | 0                  | 2               | 149                         | 157         | 210      |
|                | WCHS    | 27               | 8              | 3                  | 21              | 4                           | 63          | 21       |
| AMBER          | BWHS    | 126              | 20             | 12                 | 42              | 107                         | 307         | 2098     |
|                | CBCS    | 330              | 63             | 35                 | 136             | 38                          | 602         | 1        |
|                | WCHS    | 239              | 61             | 35                 | 88              | 49                          | 472         | 243      |
| BCAC OncoArray | 2SISTER | 19               | 4              | 1                  | 12              | 6                           | 42          | 0        |
|                | CBCS    | 400              | 81             | 40                 | 215             | 119                         | 855         | 46       |
|                | MEC     | 268              | 53             | 25                 | 91              | 168                         | 605         | 607      |
|                | NBHS    | 7                | 1              | 2                  | 8               | 33                          | 51          | 53       |
|                | NC-BCFR | 22               | 3              | 3                  | 16              | 25                          | 69          | 0        |
|                | PLCO    | 8                | 0              | 0                  | 2               | 14                          | 24          | 68       |
|                | SISTER  | 50               | 8              | 3                  | 19              | 50                          | 130         | 163      |
|                | USRT    | 0                | 0              | 0                  | 0               | 26                          | 26          | 38       |
|                | WAABCS  | 9                | 12             | 11                 | 47              | 229                         | 308         | 292      |
| GBHS           | GBHS    | 198              | 70             | 33                 | 111             | 248                         | 660         | 1496     |
| MEGA           | BWHS    | 456              | 75             | 58                 | 191             | 502                         | 1282        | 1879     |
|                | CCPS    | 159              | 34             | 22                 | 69              | 82                          | 366         | 279      |
|                | MDABCS  | 507              | 95             | 58                 | 220             | 414                         | 1294        | 1222     |
|                | MEC     | 616              | 123            | 60                 | 162             | 233                         | 1194        | 914      |
|                | NBCS    | 36               | 28             | 49                 | 82              | 500                         | 695         | 376      |
|                | NBHS    | 30               | 13             | 0                  | 0               | 95                          | 138         | 147      |
|                | NC-BCFR | 75               | 21             | 9                  | 35              | 45                          | 185         | 213      |
|                | NYUWHS  | 7                | 2              | 0                  | 5               | 58                          | 72          | 58       |
|                | SCCS    | 178              | 27             | 15                 | 50              | 438                         | 708         | 678      |
|                | STSBHS  | 354              | 69             | 10                 | 36              | 223                         | 692         | 683      |
|                | WCHS    | 585              | 143            | 87                 | 235             | 276                         | 1326        | 851      |
| ROOT           | BBCS    | 0                | 0              | 0                  | 0               | 94                          | 94          | 102      |
|                | BNCS    | 0                | 0              | 0                  | 0               | 92                          | 92          | 227      |
|                | CCPS    | 106              | 31             | 21                 | 84              | 123                         | 365         | 376      |
|                | NBCS    | 75               | 14             | 22                 | 78              | 513                         | 702         | 602      |

Continuation of Table S 1.1

|              |             |      |      |     |      |      |       |       |
|--------------|-------------|------|------|-----|------|------|-------|-------|
|              | RVGBC       | 0    | 0    | 0   | 0    | 143  | 143   | 254   |
|              | SCCS        | 50   | 3    | 3   | 18   | 52   | 126   | 323   |
| WGS          | GBHS        | 93   | 27   | 23  | 69   | 81   | 293   | 147   |
|              | MEC         | 8    | 2    | 2   | 5    | 194  | 211   | 119   |
|              | NBHS        | 14   | 7    | 13  | 32   | 25   | 91    | 16    |
|              | SCCS        | 123  | 22   | 23  | 77   | 147  | 392   | 2015  |
|              | STSBHS      | 144  | 32   | 42  | 171  | 32   | 421   | 0     |
| All datasets | All studies | 5882 | 1296 | 795 | 2760 | 6658 | 17391 | 18800 |

Table S 1.2 All lead variants identified for each intrinsic subtype in European-ancestry meta-analysis (Bonferroni-adjusted p-value<0.05)

| Loci           | rsID        | CHR:POS:REF:ALT         | Alleles <sup>a</sup> | OR (95% CI)      | P-value   |
|----------------|-------------|-------------------------|----------------------|------------------|-----------|
| Luminal-A-like |             |                         |                      |                  |           |
| 1p36.13        | rs2992756   | 1:18480845:T:C          | C/T                  | 0.94 (0.93-0.95) | 1.53E-16  |
| 1p34.1         | rs1473840   | 1:46135525:A:G          | G/A                  | 1.05 (1.03-1.07) | 4.96E-10  |
| 1p22.3         | rs305450    | 1:87736544:C:T          | T/C                  | 1.05 (1.03-1.07) | 4.75E-09  |
| 1p13.2         | -           | 1:113880354:T:A         | A/T                  | 0.95 (0.93-0.97) | 3.58E-09  |
| 1p11.2         | rs11249433  | 1:121538815:A:G         | G/A                  | 1.15 (1.14-1.17) | 2.01E-84  |
| 1q21.1         | -           | 1:143909199:C:G         | G/C                  | 1.14 (1.12-1.16) | 7.60E-52  |
| 1q21.1         | -           | 1:146125986:G:C         | C/G                  | 1.14 (1.09-1.18) | 8.51E-11  |
| 1q21.2         | rs11205303  | 1:149934520:T:C         | C/T                  | 1.06 (1.04-1.07) | 1.38E-13  |
| 1q22           | -           | 1:155225280:ATTAT:<br>A | A/ATTAT              | 0.95 (0.94-0.96) | 3.89E-12  |
| 1q24.2         | rs4657733   | 1:168117196:G:A         | A/G                  | 0.95 (0.94-0.97) | 2.24E-09  |
| 1q32.1         | rs35383942  | 1:201468704:C:T         | T/C                  | 1.10 (1.07-1.14) | 2.13E-09  |
| 1q41           | rs11117754  | 1:217033363:T:G         | G/T                  | 0.95 (0.93-0.97) | 3.73E-09  |
| 1q43           | rs72755295  | 1:241870961:A:G         | G/A                  | 1.14 (1.09-1.19) | 1.25E-10  |
| 2p13.3         | rs11126226  | 2:69208111:C:T          | T/C                  | 0.94 (0.92-0.96) | 3.93E-09  |
| 2q14.2         | rs12711947  | 2:120486696:T:C         | C/T                  | 1.10 (1.07-1.12) | 4.57E-14  |
| 2q31.1         | rs17726078  | 2:172109838:C:G         | G/C                  | 0.93 (0.92-0.95) | 2.75E-20  |
| 2q33.1         | rs3769821   | 2:201258707:C:T         | T/C                  | 0.95 (0.94-0.97) | 4.74E-11  |
| 2q35           | rs4442975   | 2:217056046:G:T         | T/G                  | 0.85 (0.84-0.86) | 2.81E-108 |
| 2q36.3         | rs13009921  | 2:226404332:T:A         | A/T                  | 1.05 (1.03-1.06) | 9.96E-09  |
| 3p26.1         | rs6762558   | 3:4700567:A:G           | G/A                  | 1.07 (1.05-1.08) | 2.48E-17  |
| 3p24.1         | rs646157    | 3:27310652:T:A          | A/T                  | 1.13 (1.12-1.15) | 1.52E-63  |
| 3p24.1         | rs35263707  | 3:30642605:G:A          | A/G                  | 1.08 (1.06-1.09) | 4.93E-21  |
| 3p21.31        | rs143048713 | 3:46851683:G:GGT        | GGT/G                | 0.91 (0.89-0.93) | 1.04E-14  |
| 3p14.1         | rs3821902   | 3:63956021:T:G          | G/T                  | 1.07 (1.04-1.09) | 1.28E-09  |
| 3p13           | rs6805189   | 3:71482962:T:C          | C/T                  | 0.96 (0.94-0.97) | 1.46E-09  |
| 3p12.1         | rs13066793  | 3:86988393:A:G          | G/A                  | 0.93 (0.90-0.95) | 2.61E-09  |
| 3q12.1         | rs34626154  | 3:100037468:G:A         | A/G                  | 1.06 (1.04-1.08) | 8.59E-12  |
| 3q23           | rs6440006   | 3:141423849:G:A         | A/G                  | 1.07 (1.05-1.08) | 4.64E-18  |
| 3q26.31        | rs78105464  | 3:172560322:GA:G        | G/GA                 | 1.06 (1.04-1.08) | 4.20E-11  |
| 4q21.23        | -           | 4:83448971:TAA:TA       | TA/TAA               | 0.96 (0.94-0.97) | 2.30E-09  |
| 4q24           | rs62331150  | 4:105147856:G:T         | T/G                  | 1.05 (1.03-1.07) | 2.87E-09  |
| 4q34.1         | rs7697216   | 4:174906885:T:C         | C/T                  | 1.14 (1.11-1.17) | 4.79E-29  |
| 5p15.33        | rs2735846   | 5:1299264:C:G           | G/C                  | 0.95 (0.93-0.96) | 1.29E-13  |
| 5p15.1         | rs17611291  | 5:16231085:G:C          | C/G                  | 0.96 (0.94-0.97) | 7.47E-10  |
| 5p12           | rs10941679  | 5:44706396:A:G          | G/A                  | 1.18 (1.16-1.20) | 1.30E-89  |
| 5q11.1         | rs35951924  | 5:50899259:A:AT         | AT/A                 | 0.94 (0.92-0.96) | 1.14E-12  |

Continuation of Table S 1.2

|          |             |                                |                     |                  |           |
|----------|-------------|--------------------------------|---------------------|------------------|-----------|
| 5q11.2   | -           | 5:56757918:AAAAC:A<br>AAACAAAC | AAAACAAA<br>C/AAAAC | 1.22 (1.20-1.25) | 1.09E-96  |
| 5q14.2   | rs4703870   | 5:82196824:G:A                 | A/G                 | 1.06 (1.05-1.08) | 2.23E-13  |
| 5q14.3   | rs11959525  | 5:91408023:A:G                 | G/A                 | 0.91 (0.89-0.93) | 5.35E-18  |
| 5q31.1   | rs6860806   | 5:132304843:A:G                | G/A                 | 1.04 (1.03-1.06) | 8.15E-09  |
| 5q33.3   | rs11135046  | 5:158803005:G:T                | T/G                 | 0.93 (0.92-0.95) | 6.32E-21  |
| 5q35.1   | rs59346687  | 5:170139969:TC:T               | T/TC                | 1.05 (1.03-1.07) | 8.99E-09  |
| 6p23     | rs418053    | 6:13713134:G:C                 | C/G                 | 0.94 (0.92-0.95) | 9.99E-20  |
| 6q14.1   | rs1361549   | 6:81577968:G:A                 | A/G                 | 0.95 (0.93-0.96) | 2.13E-13  |
| 6q25.1   | rs35409891  | 6:149265192:CT:C               | C/CT                | 0.95 (0.93-0.97) | 4.00E-09  |
| 6q25.1   | rs60954078  | 6:151634779:A:G                | G/A                 | 1.15 (1.12-1.18) | 6.70E-23  |
| 7q21.3   | rs13073     | 7:94667457:T:C                 | C/T                 | 0.95 (0.93-0.96) | 8.99E-12  |
| 7q32.3   | rs61729633  | 7:130984153:G:A                | A/G                 | 1.06 (1.04-1.07) | 1.71E-11  |
| 7q34     | rs5887960   | 7:140243902:CT:C               | C/CT                | 1.08 (1.06-1.10) | 3.07E-24  |
| 7q35     | rs720475    | 7:144377836:G:A                | A/G                 | 0.95 (0.93-0.96) | 4.59E-10  |
| 8p12     | rs7463114   | 8:29649578:T:C                 | C/T                 | 0.94 (0.92-0.95) | 1.41E-17  |
| 8p11.23  | rs4286946   | 8:36992428:C:G                 | G/C                 | 0.92 (0.91-0.94) | 6.88E-16  |
| 8q21.13  | rs72658084  | 8:75420821:C:T                 | T/C                 | 1.12 (1.09-1.14) | 4.00E-18  |
| 8q22.3   | rs4734025   | 8:101386664:A:T                | T/A                 | 0.95 (0.93-0.96) | 5.15E-11  |
| 8q23.1   | rs17282623  | 8:105342969:C:G                | G/C                 | 0.92 (0.90-0.94) | 1.49E-13  |
| 8q24.13  | rs73335873  | 8:123597610:G:C                | C/G                 | 1.08 (1.05-1.11) | 5.64E-09  |
| 8q24.21  | rs10096351  | 8:127359926:A:G                | G/A                 | 1.13 (1.11-1.15) | 2.59E-62  |
| 9p21.3   | rs539723051 | 9:21964883:CAAAA:C             | C/CAAAA             | 1.06 (1.04-1.08) | 4.59E-14  |
| 9q31.2   | rs630965    | 9:108123199:C:T                | T/C                 | 1.13 (1.12-1.15) | 1.64E-62  |
| 9q33.1   | rs1895062   | 9:116551207:A:G                | G/A                 | 0.95 (0.94-0.96) | 3.91E-12  |
| 10p12.31 | rs10828248  | 10:21535690:A:G                | G/A                 | 1.09 (1.07-1.10) | 5.19E-29  |
| 10q21.2  | rs10995201  | 10:62540131:A:G                | G/A                 | 0.88 (0.86-0.90) | 3.12E-34  |
| 10q22.3  | rs704010    | 10:79081391:T:C                | C/T                 | 0.91 (0.90-0.93) | 1.19E-32  |
| 10q25.3  | rs521196    | 10:113373449:C:A               | A/C                 | 1.08 (1.06-1.10) | 3.22E-17  |
| 10q26.13 | rs11200014  | 10:121575416:G:A               | A/G                 | 1.34 (1.32-1.36) | 0         |
| 11p15.5  | rs588321    | 11:1875727:G:C                 | C/G                 | 0.92 (0.91-0.94) | 6.17E-26  |
| 11q13.1  | rs578056    | 11:65747403:C:T                | T/C                 | 1.07 (1.05-1.09) | 2.12E-12  |
| 11q13.3  | rs78540526  | 11:69516650:C:T                | T/C                 | 1.45 (1.41-1.48) | 4.79E-181 |
| 11q24.3  | rs11822830  | 11:129591121:A:G               | G/A                 | 1.05 (1.03-1.06) | 1.59E-09  |
| 12p13.1  | rs12422552  | 12:14260997:G:C                | C/G                 | 1.05 (1.03-1.07) | 3.26E-09  |
| 12p11.22 | rs7297051   | 12:28021884:C:T                | T/C                 | 0.88 (0.87-0.90) | 6.17E-47  |
| 12q22    | rs17356907  | 12:95633983:A:G                | G/A                 | 0.91 (0.89-0.92) | 1.14E-32  |
| 12q24.21 | rs2464264   | 12:115397993:G:A               | A/G                 | 0.91 (0.89-0.92) | 7.22E-40  |
| 14q13.3  | rs12891525  | 14:36652242:A:G                | G/A                 | 0.92 (0.90-0.93) | 5.00E-25  |
| 14q24.1  | rs11624333  | 14:68513118:T:C                | C/T                 | 0.90 (0.89-0.92) | 1.54E-33  |
| 14q32.11 | rs941764    | 14:91374725:A:G                | G/A                 | 1.05 (1.04-1.07) | 9.58E-12  |

Continuation of Table S 1.2

|                    |             |                   |       |                  |           |
|--------------------|-------------|-------------------|-------|------------------|-----------|
| 15q22.33           | rs35874463  | 15:67165360:A:G   | G/A   | 1.10 (1.07-1.14) | 3.25E-09  |
| 15q24.2            | rs11072553  | 15:75640436:C:T   | T/C   | 0.95 (0.93-0.96) | 1.34E-09  |
| 15q26.1            | rs150333865 | 15:90985645:G:A   | A/G   | 0.93 (0.91-0.95) | 1.98E-12  |
| 16q12.1            | rs4784227   | 16:52565276:C:T   | T/C   | 1.27 (1.25-1.30) | 1.33E-187 |
| 16q12.2            | -           | 16:54649366:C:CAT | CAT/C | 1.05 (1.04-1.07) | 5.37E-12  |
| 16q23.2            | rs13329835  | 16:80616908:A:G   | G/A   | 1.09 (1.08-1.11) | 1.26E-25  |
| 17q21.2            | rs72826962  | 17:42684371:C:T   | T/C   | 1.24 (1.17-1.33) | 4.08E-11  |
| 17q22              | rs1812715   | 17:55097605:A:G   | G/A   | 1.09 (1.08-1.11) | 1.75E-28  |
| 18q11.2            | rs4800446   | 18:23045177:A:T   | T/A   | 0.96 (0.94-0.97) | 1.87E-09  |
| 18q11.2            | rs170801    | 18:26920935:C:A   | A/C   | 0.92 (0.91-0.94) | 1.94E-21  |
| 18q12.3            | rs9954058   | 18:44831838:G:C   | C/G   | 0.89 (0.86-0.91) | 1.91E-16  |
| 19p13.13           | rs78269692  | 19:13047463:T:C   | C/T   | 1.14 (1.10-1.18) | 1.62E-13  |
| 19p13.11           | rs8107367   | 19:18453895:A:G   | G/A   | 0.93 (0.91-0.94) | 3.99E-23  |
| 19q13.31           | rs12609846  | 19:43779295:C:T   | T/C   | 0.94 (0.92-0.95) | 5.64E-19  |
| 20q13.2            | rs13039563  | 20:53680310:G:A   | A/G   | 1.06 (1.04-1.07) | 2.34E-10  |
| 21q21.1            | rs2403907   | 21:15202135:C:A   | A/C   | 0.91 (0.89-0.92) | 3.97E-35  |
| 22q11.21           | rs9798754   | 22:19778614:C:T   | T/C   | 0.96 (0.94-0.97) | 9.18E-09  |
| 22q12.1            | rs75174485  | 22:28682717:T:G   | G/T   | 1.38 (1.29-1.47) | 6.28E-23  |
| 22q13.1            | rs738321    | 22:38172826:C:G   | G/C   | 0.94 (0.93-0.95) | 3.97E-16  |
| 22q13.2            | rs56051217  | 22:40653398:G:T   | T/G   | 1.13 (1.11-1.16) | 1.04E-25  |
| Luminal-B-like     |             |                   |       |                  |           |
| 1p36.22            | rs2506885   | 1:10520994:A:T    | T/A   | 0.87 (0.84-0.91) | 3.35E-12  |
| 2q35               | rs4442975   | 2:217056046:G:T   | T/G   | 0.89 (0.86-0.93) | 1.43E-10  |
| 4p15.33            | rs12506352  | 4:13578908:A:G    | G/A   | 1.17 (1.11-1.23) | 3.62E-09  |
| 5q11.2             | rs62355901  | 5:56757708:T:C    | C/T   | 1.22 (1.17-1.28) | 3.12E-19  |
| 6q25.2             | -           | 6:152144481:AGT:A | A/AGT | 1.14 (1.10-1.18) | 6.31E-14  |
| 8q21.13            | rs72658084  | 8:75420821:C:T    | T/C   | 1.22 (1.15-1.29) | 7.25E-12  |
| 8q24.21            | rs12550713  | 8:127358703:C:G   | G/C   | 1.11 (1.07-1.15) | 6.16E-09  |
| 9q31.2             | rs648354    | 9:108124826:G:A   | A/G   | 1.12 (1.08-1.16) | 1.22E-09  |
| 10q21.2            | rs10995194  | 10:62528371:G:C   | C/G   | 0.84 (0.79-0.88) | 2.15E-12  |
| 10q26.13           | rs2912779   | 10:121577668:T:C  | C/T   | 0.77 (0.74-0.80) | 1.67E-50  |
| 11p15.5            | rs620315    | 11:1881245:G:A    | A/G   | 0.89 (0.85-0.92) | 1.96E-11  |
| 11q13.3            | rs78540526  | 11:69516650:C:T   | T/C   | 1.23 (1.16-1.31) | 4.80E-11  |
| 15q26.1            | -           | 15:90965985:C:A   | A/C   | 0.85 (0.81-0.90) | 3.68E-09  |
| 16q12.1            | rs9921569   | 16:52538117:T:C   | C/T   | 1.26 (1.21-1.30) | 1.66E-33  |
| 22q12.1            | rs34661406  | 22:28681643:C:CA  | CA/C  | 1.47 (1.33-1.63) | 1.19E-13  |
| HER2-enriched-like |             |                   |       |                  |           |
| 1p36.22            | rs1188526   | 1:10489791:G:C    | C/G   | 1.21 (1.14-1.28) | 6.13E-11  |
| 1q32.1             | rs721559    | 1:202218991:T:C   | C/T   | 0.85 (0.81-0.89) | 2.30E-10  |
| 6q25.1             | rs58343273  | 6:151632045:A:G   | G/A   | 1.39 (1.28-1.51) | 1.88E-14  |

Continuation of Table S 1.2

|                 |             |                  |       |                  |          |
|-----------------|-------------|------------------|-------|------------------|----------|
| 16q12.1         | rs4784227   | 16:52565276:C:T  | T/C   | 1.25 (1.18-1.32) | 2.58E-15 |
| Triple-negative |             |                  |       |                  |          |
| 1q32.1          | rs2169137   | 1:204528785:G:C  | C/G   | 0.85 (0.82-0.88) | 4.81E-22 |
| 2p23.2          | rs138612072 | 2:28823002:CTT:C | C/CTT | 0.82 (0.77-0.87) | 1.61E-11 |
| 2q14.2          | rs11903787  | 2:120330606:G:A  | A/G   | 0.88 (0.84-0.91) | 1.06E-12 |
| 5p15.33         | rs10069690  | 5:1279675:C:T    | T/C   | 1.26 (1.22-1.31) | 8.65E-47 |
| 6p21.32         | rs17215231  | 6:33272092:C:T   | T/C   | 0.82 (0.77-0.87) | 2.37E-10 |
| 6q25.1          | rs60954078  | 6:151634779:A:G  | G/A   | 1.30 (1.23-1.37) | 2.07E-22 |
| 10q25.2         | rs6585202   | 10:113023044:T:C | C/T   | 1.09 (1.06-1.13) | 3.21E-09 |
| 11q22.3         | rs4754322   | 11:108440928:T:C | C/T   | 0.90 (0.87-0.93) | 1.74E-11 |
| 12p11.22        | rs7297051   | 12:28021884:C:T  | T/C   | 0.85 (0.82-0.88) | 3.44E-18 |
| 16q12.1         | rs35668161  | 16:52504913:C:A  | A/C   | 1.11 (1.08-1.15) | 1.30E-10 |
| 19p13.11        | rs67397200  | 19:17290595:C:G  | G/C   | 1.26 (1.23-1.30) | 2.76E-49 |
| 22q13.2         | rs71785733  | 22:40629878:GT:G | G/GT  | 1.16 (1.11-1.21) | 3.37E-10 |

**Footnote:**

Bonferroni threshold of significance corresponded to a p-value of  $1.25 \times 10^{-8}$

<sup>a</sup> Effect/Baseline alleles

Table S 1.3 All lead variants identified for each intrinsic subtype in the African-ancestry meta-analysis (Bonferroni-adjusted p-value<0.05)

| Loci            | rsID        | CHR:POS:REF:ALT    | Alleles <sup>a</sup> | OR (95% CI)      | P-value  |
|-----------------|-------------|--------------------|----------------------|------------------|----------|
| Luminal-A-like  |             |                    |                      |                  |          |
| 2q14.2          | rs7581194   | 2:120497373:A:C    | C/A                  | 1.13 (1.08-1.17) | 9.12E-09 |
| 2q35            | rs16856925  | 2:217096609:A:G    | G/A                  | 0.83 (0.78-0.88) | 8.91E-11 |
| 3p26.1          | rs4684440   | 3:4690838:A:G      | G/A                  | 1.17 (1.11-1.22) | 2.23E-10 |
| 8q22.3          | rs60703653  | 8:102087902:CCCT:C | C/CCCT               | 0.67 (0.59-0.76) | 1.41E-09 |
| 10q26.13        | rs2981578   | 10:121580797:C:T   | T/C                  | 0.84 (0.79-0.89) | 1.82E-09 |
| 15q24.3         | rs11636648  | 15:77043549:C:T    | T/C                  | 1.18 (1.12-1.25) | 4.25E-10 |
| 16q12.1         | rs4784227   | 16:52565276:C:T    | T/C                  | 1.32 (1.22-1.42) | 2.09E-13 |
| 18q11.2         | rs1976128   | 18:24055442:G:A    | A/G                  | 1.16 (1.10-1.22) | 2.53E-09 |
| 20q13.32        | rs1105262   | 20:59573644:T:G    | G/T                  | 0.83 (0.78-0.88) | 1.90E-09 |
| Luminal-B-like  |             |                    |                      |                  |          |
| 10q26.13        | rs145680477 | 10:123814971:T:C   | C/T                  | 1.72 (1.44-2.06) | 3.62E-09 |
| 12q23.1         | rs4762310   | 12:96591322:G:A    | A/G                  | 0.60 (0.51-0.71) | 2.97E-09 |
| Triple-negative |             |                    |                      |                  |          |
| 1p35.2          | rs142302749 | 1:31087245:A:G     | G/A                  | 1.38 (1.24-1.55) | 9.80E-09 |
| 2q14.2          | rs76664032  | 2:118823485:A:G    | G/A                  | 0.81 (0.76-0.87) | 9.95E-09 |
| 5p15.33         | rs10069690  | 5:1279675:C:T      | T/C                  | 1.38 (1.31-1.46) | 2.81E-30 |
| 5p15.32         | rs11952226  | 5:6120548:T:C      | C/T                  | 1.19 (1.13-1.26) | 5.32E-10 |
| 11q25           | rs6590696   | 11:133440533:A:C   | C/A                  | 1.22 (1.14-1.31) | 8.45E-09 |
| 12p11.22        | rs2346760   | 12:27719121:G:C    | C/G                  | 0.54 (0.45-0.66) | 6.47E-10 |
| 19p13.11        | rs11668840  | 19:17288816:T:C    | C/T                  | 0.74 (0.70-0.78) | 3.67E-28 |

**Footnote:**

Bonferroni threshold of significance corresponded to a p-value of  $1.25 \times 10^{-8}$

<sup>a</sup> Effect/Baseline alleles



Table S 1.4 All lead variants identified for each intrinsic subtype in cross-ancestry meta-analysis amongst variants present in both ancestry-specific meta-analyses (Bonferroni-adjusted p-value<0.05)

| Loci           | rsID        | CHR:POS:REF:ALT   | Alleles <sup>a</sup> | OR (95% CI)      | P-value   |
|----------------|-------------|-------------------|----------------------|------------------|-----------|
| Luminal-A-like |             |                   |                      |                  |           |
| 1p36.13        | rs2992756   | 1:18480845:T:C    | C/T                  | 0.95 (0.93-0.96) | 6.95E-16  |
| 1p34.1         | rs1707302   | 1:46135245:A:G    | G/A                  | 1.05 (1.03-1.06) | 6.08E-11  |
| 1p13.2         | rs7513707   | 1:113903258:G:A   | A/G                  | 1.06 (1.04-1.08) | 1.97E-09  |
| 1p11.2         | rs11249433  | 1:121538815:A:G   | G/A                  | 1.15 (1.13-1.17) | 1.44E-84  |
| 1q21.2         | rs11205303  | 1:149934520:T:C   | C/T                  | 1.06 (1.04-1.07) | 7.46E-14  |
| 1q22           | rs4971099   | 1:155183132:A:G   | G/A                  | 1.05 (1.03-1.06) | 1.16E-11  |
| 1q24.2         | rs4657733   | 1:168117196:G:A   | A/G                  | 0.96 (0.94-0.97) | 2.51E-09  |
| 1q41           | rs11572421  | 1:217027238:C:T   | T/C                  | 0.95 (0.93-0.97) | 8.02E-10  |
| 1q41           | rs4011894   | 1:220613960:G:A   | A/G                  | 1.05 (1.03-1.07) | 7.51E-09  |
| 2p13.3         | rs11126226  | 2:69208111:C:T    | T/C                  | 0.94 (0.93-0.96) | 2.87E-09  |
| 2q14.2         | rs12711947  | 2:120486696:T:C   | C/T                  | 1.10 (1.08-1.13) | 1.55E-20  |
| 2q31.1         | rs17726078  | 2:172109838:C:G   | G/C                  | 0.94 (0.93-0.95) | 1.40E-18  |
| 2q33.1         | rs3769821   | 2:201258707:C:T   | T/C                  | 0.95 (0.94-0.96) | 1.27E-12  |
| 2q35           | rs4442975   | 2:217056046:G:T   | T/G                  | 0.86 (0.85-0.87) | 1.49E-113 |
| 2q36.3         | rs2396351   | 2:226410046:C:G   | G/C                  | 1.05 (1.03-1.06) | 2.16E-11  |
| 3p26.1         | rs6762558   | 3:4700567:A:G     | G/A                  | 1.07 (1.06-1.09) | 8.48E-23  |
| 3p24.1         | rs583298    | 3:27320389:T:C    | C/T                  | 1.13 (1.11-1.14) | 2.89E-62  |
| 3p24.1         | rs35263707  | 3:30642605:G:A    | A/G                  | 1.07 (1.06-1.09) | 2.40E-21  |
| 3p21.31        | rs143048713 | 3:46851683:G:GGT  | GGT/G                | 0.91 (0.89-0.93) | 4.39E-16  |
| 3p14.1         | rs3774708   | 3:63895873:A:C    | C/A                  | 1.06 (1.04-1.08) | 2.68E-10  |
| 3p13           | rs11714337  | 3:71533370:G:A    | A/G                  | 0.96 (0.94-0.97) | 4.55E-09  |
| 3p12.1         | rs13066793  | 3:86988393:A:G    | G/A                  | 0.93 (0.91-0.95) | 9.57E-09  |
| 3q12.1         | rs13064704  | 3:100048697:G:A   | A/G                  | 1.06 (1.05-1.08) | 4.97E-13  |
| 3q23           | rs1582874   | 3:141396377:T:C   | C/T                  | 1.06 (1.05-1.07) | 2.77E-16  |
| 3q26.31        | rs78105464  | 3:172560322:GA:G  | G/GA                 | 1.06 (1.04-1.08) | 1.67E-12  |
| 4p16.3         | rs16342     | 4:1983005:T:TAACA | TAACA/T              | 1.05 (1.03-1.07) | 3.33E-09  |
| 4q21.23        | rs6535478   | 4:83472322:C:T    | T/C                  | 0.96 (0.95-0.97) | 3.06E-09  |
| 4q24           | -           | 4:105222285:G:T   | T/G                  | 1.08 (1.05-1.10) | 7.87E-10  |
| 4q34.1         | rs28750347  | 4:174923119:G:A   | A/G                  | 0.89 (0.88-0.91) | 1.12E-28  |
| 5p15.33        | rs2736098   | 5:1293971:C:T     | T/C                  | 0.94 (0.92-0.95) | 4.52E-15  |
| 5p15.1         | rs17611291  | 5:16231085:G:C    | C/G                  | 0.96 (0.94-0.97) | 1.17E-10  |
| 5p12           | rs10941679  | 5:44706396:A:G    | G/A                  | 1.17 (1.15-1.18) | 1.89E-85  |
| 5q11.1         | rs152803    | 5:50916286:C:T    | T/C                  | 0.96 (0.94-0.97) | 4.94E-09  |
| 5q11.2         | rs74345699  | 5:56757652:C:T    | T/C                  | 1.22 (1.20-1.24) | 3.50E-102 |
| 5q14.2         | rs4703870   | 5:82196824:G:A    | A/G                  | 1.06 (1.04-1.08) | 2.15E-12  |
| 5q14.3         | rs35945986  | 5:91430632:CT:C   | C/CT                 | 0.92 (0.91-0.94) | 3.68E-18  |
| 5q33.3         | rs11135046  | 5:158803005:G:T   | T/G                  | 0.93 (0.92-0.95) | 3.99E-22  |

Continuation of Table S 1.4

|          |             |                    |       |                  |           |
|----------|-------------|--------------------|-------|------------------|-----------|
| 5q35.1   | rs7729559   | 5:170141842:C:T    | T/C   | 1.05 (1.03-1.07) | 1.30E-08  |
| 6p23     | rs495633    | 6:13717681:G:A     | A/G   | 0.94 (0.92-0.95) | 5.54E-22  |
| 6q14.1   | rs12525804  | 6:81467674:A:T     | T/A   | 0.95 (0.94-0.97) | 1.11E-12  |
| 6q25.1   | rs2121348   | 6:149274369:T:C    | C/T   | 0.95 (0.93-0.97) | 6.45E-09  |
| 6q25.1   | rs9383590   | 6:151632630:T:C    | C/T   | 1.15 (1.12-1.18) | 8.40E-24  |
| 7q21.3   | rs13073     | 7:94667457:T:C     | C/T   | 0.95 (0.93-0.96) | 2.14E-12  |
| 7q32.3   | rs61729633  | 7:130984153:G:A    | A/G   | 1.06 (1.04-1.08) | 4.07E-13  |
| 7q34     | rs11977670  | 7:140242504:G:A    | A/G   | 1.08 (1.06-1.09) | 8.00E-26  |
| 7q35     | rs720475    | 7:144377836:G:A    | A/G   | 0.95 (0.94-0.97) | 1.35E-09  |
| 8p21.2   | rs13256025  | 8:25974262:C:T     | T/C   | 1.06 (1.04-1.08) | 1.32E-09  |
| 8p12     | rs7463114   | 8:29649578:T:C     | C/T   | 0.93 (0.92-0.95) | 7.89E-21  |
| 8p11.23  | rs13365225  | 8:37000965:A:G     | G/A   | 0.93 (0.92-0.95) | 3.02E-16  |
| 8q21.13  | rs72658084  | 8:75420821:C:T     | T/C   | 1.12 (1.10-1.15) | 2.02E-21  |
| 8q22.3   | rs16867610  | 8:101396319:G:A    | A/G   | 1.06 (1.04-1.08) | 8.52E-12  |
| 8q23.1   | rs2957447   | 8:105345146:A:G    | G/A   | 1.05 (1.04-1.07) | 8.50E-13  |
| 8q23.3   | rs13277568  | 8:115667320:A:G    | G/A   | 0.95 (0.94-0.97) | 3.60E-10  |
| 8q24.13  | rs58847541  | 8:123597926:G:A    | A/G   | 1.06 (1.04-1.08) | 5.42E-10  |
| 8q24.21  | rs13249993  | 8:127338270:G:A    | A/G   | 1.12 (1.10-1.13) | 3.55E-60  |
| 9p21.3   | rs3731257   | 9:21966222:G:A     | A/G   | 0.95 (0.94-0.97) | 1.70E-09  |
| 9q31.2   | rs7862747   | 9:108130619:C:A    | A/C   | 1.13 (1.11-1.14) | 1.44E-62  |
| 9q33.1   | rs1407243   | 9:116552572:C:T    | T/C   | 0.96 (0.94-0.97) | 2.06E-10  |
| 10p15.1  | rs55913768  | 10:5707486:G:A     | A/G   | 1.04 (1.03-1.06) | 5.22E-09  |
| 10p14    | rs7100052   | 10:9058859:T:G     | G/T   | 1.06 (1.04-1.09) | 4.45E-09  |
| 10p12.31 | rs12251016  | 10:21532989:A:T    | T/A   | 1.08 (1.06-1.09) | 1.58E-25  |
| 10q21.2  | rs10995182  | 10:62509005:T:C    | C/T   | 0.88 (0.86-0.90) | 6.86E-33  |
| 10q22.3  | rs704010    | 10:79081391:T:C    | C/T   | 0.91 (0.90-0.93) | 5.35E-35  |
| 10q25.3  | rs521196    | 10:113373449:C:A   | A/C   | 1.07 (1.05-1.09) | 1.17E-17  |
| 10q26.13 | rs11200014  | 10:121575416:G:A   | A/G   | 1.31 (1.30-1.33) | 0.00E+00  |
| 11p15.5  | rs1973765   | 11:1877434:T:C     | C/T   | 0.93 (0.91-0.94) | 2.81E-26  |
| 11q13.1  | rs72922733  | 11:65753275:A:G    | G/A   | 1.07 (1.05-1.08) | 1.35E-12  |
| 11q13.3  | rs78540526  | 11:69516650:C:T    | T/C   | 1.44 (1.41-1.48) | 2.53E-183 |
| 11q24.3  | rs113103254 | 11:129593220:C:CAT | CAT/C | 1.04 (1.03-1.06) | 6.87E-10  |
| 12p13.1  | rs75905550  | 12:14255620:G:C    | C/G   | 1.06 (1.04-1.08) | 8.02E-10  |
| 12p11.22 | rs7297051   | 12:28021884:C:T    | T/C   | 0.89 (0.88-0.90) | 1.56E-44  |
| 12q22    | rs17356907  | 12:95633983:A:G    | G/A   | 0.91 (0.90-0.92) | 7.47E-36  |
| 12q24.21 | rs2454399   | 12:115398031:T:C   | C/T   | 0.91 (0.90-0.92) | 3.78E-41  |
| 14q13.3  | rs12883049  | 14:36662856:A:G    | G/A   | 0.92 (0.90-0.93) | 8.09E-26  |
| 14q24.1  | rs11624333  | 14:68513118:T:C    | C/T   | 0.90 (0.89-0.92) | 5.29E-34  |
| 14q32.11 | rs4904771   | 14:91382854:G:A    | A/G   | 1.07 (1.05-1.09) | 1.14E-11  |
| 15q22.33 | rs35874463  | 15:67165360:A:G    | G/A   | 1.10 (1.07-1.14) | 2.21E-09  |
| 15q24.2  | rs11072553  | 15:75640436:C:T    | T/C   | 0.94 (0.93-0.96) | 2.88E-13  |

Continuation of Table S 1.4

|                    |             |                  |      |                  |           |
|--------------------|-------------|------------------|------|------------------|-----------|
| 15q26.1            | rs12594752  | 15:90988765:C:T  | T/C  | 0.92 (0.90-0.94) | 1.16E-16  |
| 16q12.1            | rs4784227   | 16:52565276:C:T  | T/C  | 1.28 (1.26-1.30) | 2.39E-199 |
| 16q12.2            | rs28539243  | 16:54648152:G:A  | A/G  | 1.05 (1.04-1.07) | 2.31E-12  |
| 16q23.2            | rs7499149   | 16:80614430:A:G  | G/A  | 1.09 (1.07-1.10) | 5.40E-26  |
| 17q21.2            | rs12938957  | 17:42417421:T:C  | C/T  | 1.06 (1.04-1.07) | 3.48E-10  |
| 17q22              | rs1812715   | 17:55097605:A:G  | G/A  | 1.09 (1.07-1.10) | 8.46E-29  |
| 18q11.2            | rs170801    | 18:26920935:C:A  | A/C  | 0.93 (0.91-0.94) | 1.28E-21  |
| 18q12.1            | rs16963218  | 18:32351771:A:G  | G/A  | 0.91 (0.88-0.93) | 4.55E-11  |
| 18q12.3            | rs9954058   | 18:44831838:G:C  | C/G  | 0.89 (0.87-0.91) | 6.92E-21  |
| 19p13.13           | rs138940068 | 19:13142270:T:G  | G/T  | 1.12 (1.08-1.16) | 6.90E-12  |
| 19p13.11           | rs10409505  | 19:18479441:C:G  | G/C  | 0.93 (0.92-0.94) | 3.04E-23  |
| 19q13.31           | rs56681946  | 19:43778879:T:C  | C/T  | 1.07 (1.06-1.09) | 2.39E-21  |
| 20q13.2            | rs13039563  | 20:53680310:G:A  | A/G  | 1.06 (1.04-1.07) | 3.34E-11  |
| 21q21.1            | rs2403907   | 21:15202135:C:A  | A/C  | 0.91 (0.90-0.93) | 5.83E-34  |
| 21q22.3            | rs9808759   | 21:46360308:T:C  | C/T  | 0.92 (0.90-0.95) | 1.14E-08  |
| 22q12.1            | rs9620817   | 22:28712568:A:T  | T/A  | 1.09 (1.07-1.12) | 1.89E-14  |
| 22q13.1            | rs11703601  | 22:38202034:C:T  | T/C  | 0.94 (0.93-0.96) | 1.24E-15  |
| 22q13.2            | rs56051217  | 22:40653398:G:T  | T/G  | 1.13 (1.11-1.16) | 8.38E-27  |
| Luminal-B-like     |             |                  |      |                  |           |
| 1p36.22            | rs596537    | 1:10494652:G:T   | T/G  | 1.13 (1.09-1.17) | 1.28E-12  |
| 2q31.1             | rs930313    | 2:173354390:C:G  | G/C  | 1.15 (1.10-1.20) | 5.49E-10  |
| 2q35               | rs12620095  | 2:217042791:C:G  | G/C  | 0.90 (0.87-0.93) | 2.41E-11  |
| 5q11.2             | rs62355901  | 5:56757708:T:C   | C/T  | 1.21 (1.16-1.26) | 3.76E-19  |
| 6q25.2             | rs34133739  | 6:152113143:A:AC | AC/A | 1.14 (1.10-1.17) | 1.95E-15  |
| 8q21.13            | rs72658084  | 8:75420821:C:T   | T/C  | 1.22 (1.16-1.29) | 1.75E-13  |
| 8q24.21            | rs7006882   | 8:127365421:G:T  | T/G  | 0.91 (0.88-0.94) | 4.04E-09  |
| 9q31.2             | rs525142    | 9:108124254:G:A  | A/G  | 1.10 (1.07-1.14) | 4.40E-09  |
| 10q21.2            | rs10995187  | 10:62513267:G:A  | A/G  | 0.84 (0.80-0.88) | 5.22E-13  |
| 10q26.13           | rs2912779   | 10:121577668:T:C | C/T  | 0.79 (0.77-0.82) | 3.93E-46  |
| 11p15.5            | rs540710    | 11:1926345:C:A   | A/C  | 0.90 (0.87-0.93) | 1.61E-10  |
| 11q13.3            | rs657686    | 11:69517902:A:G  | G/A  | 1.16 (1.11-1.22) | 1.02E-10  |
| 15q26.1            | rs8037137   | 15:90963407:T:C  | C/T  | 0.87 (0.83-0.91) | 2.63E-09  |
| 16q12.1            | rs4784227   | 16:52565276:C:T  | T/C  | 1.27 (1.23-1.32) | 3.71E-38  |
| HER2-enriched-like |             |                  |      |                  |           |
| 1p36.22            | rs2506896   | 1:10543659:T:C   | C/T  | 0.84 (0.80-0.88) | 4.01E-11  |
| 6q25.1             | rs9397436   | 6:151630867:A:G  | G/A  | 1.38 (1.28-1.49) | 8.27E-17  |
| 16q12.1            | rs4784227   | 16:52565276:C:T  | T/C  | 1.25 (1.19-1.32) | 1.34E-16  |
| Triple-negative    |             |                  |      |                  |           |
| 1p36.22            | rs112200610 | 1:9156852:C:T    | T/C  | 1.23 (1.15-1.31) | 1.80E-09  |
| 1q32.1             | rs930947    | 1:204572069:G:A  | A/G  | 0.86 (0.83-0.88) | 1.49E-26  |
| 2p24.1             | rs12710697  | 2:19121207:C:T   | T/C  | 0.92 (0.90-0.95) | 1.13E-09  |

Continuation of Table S 1.4

|          |            |                  |     |                  |          |
|----------|------------|------------------|-----|------------------|----------|
| 2p23.2   | rs57312136 | 2:28922682:C:G   | G/C | 0.83 (0.79-0.87) | 1.63E-13 |
| 2q14.2   | rs11903787 | 2:120330606:G:A  | A/G | 0.90 (0.87-0.93) | 1.73E-11 |
| 3q26.31  | rs2270418  | 3:172523209:T:G  | G/T | 0.89 (0.86-0.92) | 3.51E-10 |
| 5p15.33  | rs10069690 | 5:1279675:C:T    | T/C | 1.29 (1.26-1.33) | 2.88E-73 |
| 6p21.32  | rs17215231 | 6:33272092:C:T   | T/C | 0.83 (0.78-0.88) | 2.85E-10 |
| 6q25.1   | rs7763637  | 6:151628177:G:A  | A/G | 1.15 (1.12-1.19) | 8.98E-26 |
| 10q25.2  | rs7899529  | 10:113019779:G:A | A/G | 1.10 (1.07-1.13) | 3.87E-12 |
| 11q22.3  | rs60639858 | 11:108444159:A:G | G/A | 0.90 (0.88-0.93) | 1.94E-13 |
| 12p11.22 | rs7297051  | 12:28021884:C:T  | T/C | 0.86 (0.83-0.89) | 9.55E-19 |
| 16q12.1  | rs17271951 | 16:52504128:T:C  | C/T | 1.11 (1.08-1.15) | 1.26E-10 |
| 19p13.11 | rs61494113 | 19:17291050:G:A  | A/G | 1.28 (1.25-1.31) | 1.31E-72 |
| 22q13.1  | rs59098711 | 22:40481000:T:C  | C/T | 1.11 (1.07-1.14) | 2.53E-10 |

**Footnote:**

Bonferroni threshold of significance corresponded to a p-value of  $1.25 \times 10^{-8}$

<sup>a</sup> Effect/Baseline alleles

Table S 1.5 Validation of triple-negative variants identified previously in AABCG using logistic regression.

| Loci     | rsID       | Alleles <sup>a</sup> | Previous Study (Logistic Regression) |          | Current Study (TOP) |          |
|----------|------------|----------------------|--------------------------------------|----------|---------------------|----------|
|          |            |                      | OR (95% CI)                          | P-value  | OR (95% CI)         | P-value  |
| 5p15.33  | rs10069690 | T/C                  | 1.38 (1.30-1.47)                     | 1.15E-23 | 1.38 (1.30-1.47)    | 2.61E-30 |
| 19p13.11 | rs12974508 | C/T                  | 1.38 (1.29-1.47)                     | 1.29E-23 | 1.37 (1.28-1.45)    | 4.02E-28 |
| 2q14.2   | rs76664032 | A/G                  | 1.30 (1.20-1.42)                     | 3.64E-10 | 1.23 (1.15-1.32)    | 1.03E-08 |

**Footnote:**

TOP, Two-stage polytomous logistic regression

<sup>a</sup> Effect/Baseline alleles

Table S 1.6 Consistency of novel lead variants across sequencing platforms in AABCG

| Loci           | rsID       | Alleles <sup>a</sup> | Consistency in direction <sup>b</sup> | Sequencing Platform | OR (95% CI)       | P-value  | EAF <sup>c</sup> |
|----------------|------------|----------------------|---------------------------------------|---------------------|-------------------|----------|------------------|
| Luminal-A-like |            |                      |                                       |                     |                   |          |                  |
| 1q41           | rs4011894  | A/G                  | +-+?+++                               | AABC                | 1.09 (0.87, 1.37) | 0.437    | 0.054878         |
|                |            |                      |                                       | AMBER               | 0.89 (0.66, 1.19) | 0.424    | 0.0492655        |
|                |            |                      |                                       | BCAC OncoArray      | 1.26 (0.94, 1.67) | 0.117    | 0.0418619        |
|                |            |                      |                                       | GBHS                | -                 | -        | -                |
|                |            |                      |                                       | MEGA                | 1.18 (1.04, 1.34) | 0.00858  | 0.042844         |
|                |            |                      |                                       | ROOT                | 1.11 (0.71, 1.73) | 0.636    | 0.0273302        |
|                |            |                      |                                       | WGS                 | 1.33 (0.94, 1.89) | 0.111    | 0.0337397        |
| 10p15.1        | rs55913768 | A/G                  | -+++++                                | AABC                | 0.93 (0.78, 1.10) | 0.378    | 0.122038         |
|                |            |                      |                                       | AMBER               | 1.09 (0.91, 1.31) | 0.33     | 0.114156         |
|                |            |                      |                                       | BCAC OncoArray      | 1.23 (1.02, 1.49) | 0.029    | 0.105738         |
|                |            |                      |                                       | GBHS                | 1.23 (0.85, 1.79) | 0.274    | 0.0576691        |
|                |            |                      |                                       | MEGA                | 1.07 (0.98, 1.17) | 0.128    | 0.10644          |
|                |            |                      |                                       | ROOT                | 1.03 (0.79, 1.34) | 0.826    | 0.0875796        |
|                |            |                      |                                       | WGS                 | 1.15 (0.92, 1.42) | 0.212    | 0.0975185        |
| 10p14          | rs7100052  | G/T                  | --+++++                               | AABC                | 0.99 (0.83, 1.18) | 0.931    | 0.106471         |
|                |            |                      |                                       | AMBER               | 0.99 (0.82, 1.19) | 0.897    | 0.108285         |
|                |            |                      |                                       | BCAC OncoArray      | 1.08 (0.89, 1.31) | 0.46     | 0.103929         |
|                |            |                      |                                       | GBHS                | 1.41 (1.04, 1.91) | 0.0276   | 0.0859954        |
|                |            |                      |                                       | MEGA                | 1.08 (0.99, 1.18) | 0.0774   | 0.108495         |
|                |            |                      |                                       | ROOT                | 1.03 (0.82, 1.31) | 0.791    | 0.104945         |
|                |            |                      |                                       | WGS                 | 1.09 (0.88, 1.35) | 0.441    | 0.104702         |
| 18q12.1        | rs16963218 | G/A                  | -----                                 | AABC                | 0.96 (0.85, 1.09) | 0.572    | 0.244201         |
|                |            |                      |                                       | AMBER               | 0.77 (0.67, 0.89) | 0.000326 | 0.248508         |
|                |            |                      |                                       | BCAC OncoArray      | 0.92 (0.80, 1.07) | 0.275    | 0.23934          |
|                |            |                      |                                       | GBHS                | 0.87 (0.70, 1.09) | 0.241    | 0.265151         |
|                |            |                      |                                       | MEGA                | 0.88 (0.83, 0.94) | 0.000117 | 0.254526         |
|                |            |                      |                                       | ROOT                | 0.92 (0.78, 1.09) | 0.343    | 0.26646          |
|                |            |                      |                                       | WGS                 | 0.97 (0.83, 1.13) | 0.681    | 0.262081         |
| 20q13.32       | rs1105262  | G/T                  | ---?---                               | AABC                | 0.79 (0.67, 0.93) | 0.00434  | 0.869891         |
|                |            |                      |                                       | AMBER               | 0.91 (0.75, 1.09) | 0.301    | 0.875986         |
|                |            |                      |                                       | BCAC OncoArray      | 0.74 (0.61, 0.90) | 0.00295  | 0.863667         |
|                |            |                      |                                       | GBHS                | -                 | -        | -                |
|                |            |                      |                                       | MEGA                | 0.82 (0.75, 0.90) | 7.83E-06 | 0.887156         |
|                |            |                      |                                       | ROOT                | 0.91 (0.68, 1.22) | 0.524    | 0.92883          |
|                |            |                      |                                       | WGS                 | 0.89 (0.69, 1.16) | 0.393    | 0.921637         |
| Luminal-B-like |            |                      |                                       |                     |                   |          |                  |
| 2q31.1         | rs930313   | G/C                  | +++++.+                               | AABC                | 1.13 (0.87, 1.48) | 0.367    | 0.758988         |

Continuation of Table S 1.6

|                 |             |     |                |                   |                   |          |           |
|-----------------|-------------|-----|----------------|-------------------|-------------------|----------|-----------|
|                 |             |     | AMBER          | 1.14 (0.85, 1.52) | 0.394             | 0.761328 |           |
|                 |             |     | BCAC OncoArray | 1.40 (1.05, 1.85) | 0.0212            | 0.751365 |           |
|                 |             |     | GBHS           | 1.41 (0.91, 2.17) | 0.12              | 0.728902 |           |
|                 |             |     | MEGA           | 1.14 (1.00, 1.31) | 0.0543            | 0.766831 |           |
|                 |             |     | ROOT           | 0.83 (0.55, 1.27) | 0.398             | 0.752028 |           |
|                 |             |     | WGS            | 1.20 (0.83, 1.73) | 0.343             | 0.763822 |           |
| 10q26.13        | rs145680477 | C/T | +++++++        | AABC              | 2.27 (1.46, 3.53) | 0.000289 | 0.0334027 |
|                 |             |     |                | AMBER             | 2.40 (1.48, 3.88) | 0.000369 | 0.0337597 |
|                 |             |     |                | BCAC OncoArray    | 1.33 (0.74, 2.40) | 0.346    | 0.0351661 |
|                 |             |     |                | GBHS              | 3.36 (1.73, 6.55) | 0.000362 | 0.0373314 |
|                 |             |     |                | MEGA              | 1.38 (1.04, 1.82) | 0.0259   | 0.0361201 |
|                 |             |     |                | ROOT              | 1.98 (0.89, 4.42) | 0.0932   | 0.0360969 |
|                 |             |     |                | WGS               | 1.05 (0.53, 2.09) | 0.882    | 0.0409229 |
| 12q23.1         | rs4762310   | A/G | -----+         | AABC              | 0.55 (0.36, 0.82) | 0.00317  | 0.958171  |
|                 |             |     |                | AMBER             | 0.89 (0.50, 1.60) | 0.705    | 0.95575   |
|                 |             |     |                | BCAC OncoArray    | 0.62 (0.38, 1.02) | 0.0619   | 0.956985  |
|                 |             |     |                | GBHS              | 0.22 (0.11, 0.44) | 1.73E-05 | 0.976741  |
|                 |             |     |                | MEGA              | 0.64 (0.51, 0.81) | 0.000233 | 0.958442  |
|                 |             |     |                | ROOT              | 0.44 (0.20, 0.96) | 0.0402   | 0.965655  |
|                 |             |     |                | WGS               | 1.75 (0.55, 5.60) | 0.345406 | 0.967131  |
| Triple-negative |             |     |                |                   |                   |          |           |
| 1p36.22         | rs112200610 | T/C | +++-.+++       | AABC              | 1.11 (0.83, 1.49) | 0.487    | 0.0532082 |
|                 |             |     |                | AMBER             | 1.12 (0.78, 1.62) | 0.544    | 0.0528895 |
|                 |             |     |                | BCAC OncoArray    | 1.40 (1.04, 1.89) | 0.0288   | 0.056829  |
|                 |             |     |                | GBHS              | 0.64 (0.31, 1.34) | 0.24     | 0.0474359 |
|                 |             |     |                | MEGA              | 1.39 (1.17, 1.65) | 0.000194 | 0.0509678 |
|                 |             |     |                | ROOT              | 1.47 (1.06, 2.04) | 0.0222   | 0.0483264 |
|                 |             |     |                | WGS               | 1.19 (0.89, 1.61) | 0.247    | 0.0555072 |
| 1p35.2          | rs142302749 | G/A | +++++++        | AABC              | 1.11 (0.82, 1.51) | 0.491    | 0.0464181 |
|                 |             |     |                | AMBER             | 1.35 (0.93, 1.95) | 0.113    | 0.0490717 |
|                 |             |     |                | BCAC OncoArray    | 1.56 (1.11, 2.19) | 0.0106   | 0.0416742 |
|                 |             |     |                | GBHS              | 1.21 (0.68, 2.17) | 0.514    | 0.0504668 |
|                 |             |     |                | MEGA              | 1.54 (1.29, 1.83) | 1.57E-06 | 0.0451393 |
|                 |             |     |                | ROOT              | 1.27 (0.89, 1.80) | 0.184    | 0.0481624 |
|                 |             |     |                | WGS               | 1.29 (0.95, 1.76) | 0.106    | 0.0476709 |
| 2p24.1          | rs12710697  | T/C | +-----         | AABC              | 1.07 (0.94, 1.23) | 0.306    | 0.375566  |
|                 |             |     |                | AMBER             | 0.89 (0.74, 1.06) | 0.199    | 0.38386   |
|                 |             |     |                | BCAC OncoArray    | 0.82 (0.70, 0.96) | 0.0114   | 0.375822  |
|                 |             |     |                | GBHS              | 0.85 (0.63, 1.13) | 0.256    | 0.334129  |
|                 |             |     |                | MEGA              | 0.85 (0.78, 0.93) | 0.00044  | 0.374345  |

Continuation of Table S 1.6

|         |            |     |        |                   |                   |          |
|---------|------------|-----|--------|-------------------|-------------------|----------|
|         |            |     | ROOT   | 0.91 (0.75, 1.10) | 0.313             | 0.35324  |
|         |            |     | WGS    | 0.93 (0.79, 1.08) | 0.341             | 0.364824 |
| 5p15.32 | rs11952226 | C/T | ++++++ | AABC              | 1.19 (1.03, 1.37) | 0.0178   |
|         |            |     |        | AMBER             | 1.20 (0.99, 1.44) | 0.059    |
|         |            |     |        | BCAC OncoArray    | 1.26 (1.08, 1.48) | 0.00417  |
|         |            |     |        | GBHS              | 1.22 (0.93, 1.60) | 0.146    |
|         |            |     |        | MEGA              | 1.23 (1.12, 1.34) | 7.54E-06 |
|         |            |     |        | ROOT              | 1.17 (0.98, 1.40) | 0.085    |
|         |            |     |        | WGS               | 1.03 (0.88, 1.21) | 0.688    |
| 11q25   | rs6590696  | C/A | +++++- | AABC              | 1.15 (0.98, 1.34) | 0.0872   |
|         |            |     |        | AMBER             | 1.18 (0.96, 1.45) | 0.122    |
|         |            |     |        | BCAC OncoArray    | 1.14 (0.95, 1.36) | 0.166    |
|         |            |     |        | GBHS              | 1.09 (0.73, 1.61) | 0.676    |
|         |            |     |        | MEGA              | 1.30 (1.17, 1.44) | 4.97E-07 |
|         |            |     |        | ROOT              | 1.29 (1.01, 1.64) | 0.0424   |
|         |            |     |        | WGS               | 0.76 (0.21, 2.71) | 0.676622 |

**Footnote:**<sup>a</sup> Effect/Baseline alleles<sup>b</sup> + corresponds to a positive association, - corresponds to a negative association. Order of platforms for reported directionality of associations: AABC (African American Breast Cancer consortium), AMBER (African American Breast Cancer Epidemiology and Risk Consortium), BCAC OncoArray (African-ancestry women in BCAC), GBHS (Ghana Breast Health Study), MEGA (Multi-ethnic genotyping array), ROOT (GWAS of Breast Cancer in the African Diaspora Consortium), WGS (Whole genome sequencing)<sup>c</sup> Effect allele frequency in controls



Table S 1.7 Conditionally independent signals nearby novel GWAS loci identified for breast cancer intrinsic subtypes in this study

| Locus           | Lead variant | Evaluated variant | Alleles <sup>a</sup> | Marginal analysis |          | Joint analysis <sup>c</sup> |          |
|-----------------|--------------|-------------------|----------------------|-------------------|----------|-----------------------------|----------|
|                 |              |                   |                      | OR (95% CI)       | P-value  | OR (95% CI)                 | P-value  |
| Luminal-A-like  |              |                   |                      |                   |          |                             |          |
| 10p15.1         | rs55913768   | rs55913768        | A/G                  | 1.04 (1.03-1.06)  | 4.89E-09 | 1.05 (1.04-1.07)            | 7.67E-12 |
|                 |              | rs11259192        | A/G                  | 1.04 (1.02-1.06)  | 0.00021  | 1.06 (1.04-1.08)            | 3.98E-07 |
| Triple-negative |              |                   |                      |                   |          |                             |          |
| 1p36.22         | rs112200610  | rs112200610       | T/C                  | 1.23 (1.15, 1.31) | 1.8E-09  | 1.22 (1.14-1.31)            | 3.99E-09 |
|                 |              | rs2483677         | T/C                  | 0.91 (0.88-0.94)  | 4.29E-09 | 0.91 (0.88-0.94)            | 9.82E-09 |

**Footnote:**

Conditionally independent signals were identified at a p-value threshold of  $1 \times 10^{-5}$

<sup>a</sup> Effect/Baseline alleles

<sup>b</sup> Association for both the lead variant and the independent signal included in a joint model

Table S 1.8 Risk allele frequencies for all lead variants identified for each subtype in this study.

| Variant ID<br>(CHR:POS:REF:ALT) | Risk Allele | RAF in<br>African-<br>ancestry<br>(Africa) | RAF in<br>African-<br>ancestry<br>(North<br>America) | RAF in<br>European-<br>ancestry | RAF in<br>African-<br>ancestry<br>(Africa) greater<br>than RAF in<br>European-<br>ancestry | EAF in African-<br>ancestry (North<br>America) greater<br>than RAF in<br>European-ancestry |
|---------------------------------|-------------|--------------------------------------------|------------------------------------------------------|---------------------------------|--------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| Luminal-A-like                  |             |                                            |                                                      |                                 |                                                                                            |                                                                                            |
| 10:121580797:C:T                | C           | 0.9379                                     | 0.8540                                               | 0.4671                          | Yes                                                                                        | Yes                                                                                        |
| 15:77043549:C:T                 | T           | 0.0502                                     | 0.1577                                               | 0.6667                          | No                                                                                         | No                                                                                         |
| 16:52565276:C:T                 | T           | 0.0240                                     | 0.0661                                               | 0.2385                          | No                                                                                         | No                                                                                         |
| 18:24055442:G:A                 | A           | 0.1241                                     | 0.1993                                               | 0.5910                          | No                                                                                         | No                                                                                         |
| 2:217096609:A:G                 | A           | 0.8395                                     | 0.8535                                               | 0.9506                          | No                                                                                         | No                                                                                         |
| 2:120497373:A:C                 | C           | 0.5631                                     | 0.6363                                               | 0.9749                          | No                                                                                         | No                                                                                         |
| 20:59573644:T:G                 | T           | 0.0287                                     | 0.1163                                               | 0.6067                          | No                                                                                         | No                                                                                         |
| 3:4690838:A:G                   | G           | 0.1091                                     | 0.2007                                               | 0.5572                          | No                                                                                         | No                                                                                         |
| 8:102087902:CCT:C               | CCT         | 0.0295                                     | 0.0239                                               | 0.1124                          | No                                                                                         | No                                                                                         |
| 1:121538815:A:G                 | G           | 0.0617                                     | 0.1194                                               | 0.4040                          | No                                                                                         | No                                                                                         |
| 1:143909199:C:G                 | G           | -                                          | -                                                    | 0.3900                          | -                                                                                          | -                                                                                          |
| 1:18480845:T:C                  | T           | 0.5374                                     | 0.5335                                               | 0.4805                          | Yes                                                                                        | Yes                                                                                        |
| 1:149934520:T:C                 | C           | 0.0154                                     | 0.0821                                               | 0.4009                          | No                                                                                         | No                                                                                         |
| 1:155225280:ATTAT:A             | ATTAT       | -                                          | -                                                    | 0.5644                          | -                                                                                          | -                                                                                          |
| 1:146125986:G:C                 | C           | -                                          | -                                                    | 0.0464                          | -                                                                                          | -                                                                                          |
| 1:241870961:A:G                 | G           | -                                          | -                                                    | 0.0312                          | -                                                                                          | -                                                                                          |
| 1:46135525:A:G                  | G           | 0.7258                                     | 0.7310                                               | 0.6648                          | Yes                                                                                        | Yes                                                                                        |
| 1:201468704:C:T                 | T           | -                                          | -                                                    | 0.0561                          | -                                                                                          | -                                                                                          |
| 1:168117196:G:A                 | G           | 0.3426                                     | 0.3864                                               | 0.6629                          | No                                                                                         | No                                                                                         |
| 1:113880354:T:A                 | T           | -                                          | -                                                    | 0.2534                          | -                                                                                          | -                                                                                          |
| 1:217033363:T:G                 | T           | 0.7702                                     | 0.7679                                               | 0.7898                          | No                                                                                         | No                                                                                         |
| 1:87736544:C:T                  | T           | 0.6772                                     | 0.6064                                               | 0.2623                          | Yes                                                                                        | Yes                                                                                        |
| 10:121575416:G:A                | A           | 0.1681                                     | 0.2184                                               | 0.4004                          | No                                                                                         | No                                                                                         |
| 10:62540131:A:G                 | A           | 0.8086                                     | 0.8218                                               | 0.8412                          | No                                                                                         | No                                                                                         |
| 10:79081391:T:C                 | T           | 0.0192                                     | 0.0912                                               | 0.3802                          | No                                                                                         | No                                                                                         |
| 10:21535690:A:G                 | G           | 0.4431                                     | 0.4050                                               | 0.3577                          | Yes                                                                                        | Yes                                                                                        |
| 10:113373449:C:A                | A           | 0.4325                                     | 0.3890                                               | 0.2035                          | Yes                                                                                        | Yes                                                                                        |
| 11:69516650:C:T                 | T           | 0.0021                                     | 0.0125                                               | 0.0753                          | No                                                                                         | No                                                                                         |
| 11:1875727:G:C                  | G           | 0.8275                                     | 0.7921                                               | 0.6075                          | Yes                                                                                        | Yes                                                                                        |
| 11:65747403:C:T                 | T           | 0.2580                                     | 0.2453                                               | 0.1889                          | Yes                                                                                        | Yes                                                                                        |
| 11:129591121:A:G                | G           | 0.8199                                     | 0.7959                                               | 0.5963                          | Yes                                                                                        | Yes                                                                                        |
| 12:28021884:C:T                 | C           | 0.9008                                     | 0.8571                                               | 0.7578                          | Yes                                                                                        | Yes                                                                                        |
| 12:115397993:G:A                | G           | 0.5413                                     | 0.5543                                               | 0.5831                          | No                                                                                         | No                                                                                         |
| 12:95633983:A:G                 | A           | 0.8340                                     | 0.7991                                               | 0.7015                          | Yes                                                                                        | Yes                                                                                        |
| 12:14260997:G:C                 | C           | 0.4327                                     | 0.4126                                               | 0.2595                          | Yes                                                                                        | Yes                                                                                        |
| 14:68513118:T:C                 | T           | 0.9989                                     | 0.9489                                               | 0.7397                          | Yes                                                                                        | Yes                                                                                        |

Continuation of Table S 1.8

|                   |     |        |        |        |     |     |
|-------------------|-----|--------|--------|--------|-----|-----|
| 14:36652242:A:G   | A   | 0.8937 | 0.8604 | 0.7527 | Yes | Yes |
| 14:91374725:A:G   | G   | 0.8058 | 0.7207 | 0.3430 | Yes | Yes |
| 15:90985645:G:A   | G   | 0.7287 | 0.7621 | 0.8651 | No  | No  |
| 15:75640436:C:T   | C   | 0.4279 | 0.4710 | 0.7562 | No  | No  |
| 15:67165360:A:G   | G   | 0.0002 | 0.0104 | 0.0505 | No  | No  |
| 16:52565276:C:T   | T   | 0.0240 | 0.0661 | 0.2385 | No  | No  |
| 16:80616908:A:G   | G   | -      | -      | 0.2239 | -   | -   |
| 16:54649366:C:CAT | CAT | -      | -      | 0.5087 | -   | -   |
| 17:55097605:A:G   | G   | 0.3181 | 0.3920 | 0.6932 | No  | No  |
| 17:42684371:C:T   | T   | -      | -      | 0.0111 | -   | -   |
| 18:26920935:C:A   | C   | 0.7510 | 0.7421 | 0.7186 | Yes | Yes |
| 18:44831838:G:C   | G   | 0.6533 | 0.6910 | 0.9279 | No  | No  |
| 18:23045177:A:T   | A   | 0.9527 | 0.8434 | 0.3625 | Yes | Yes |
| 19:18453895:A:G   | A   | 0.2763 | 0.3144 | 0.6548 | No  | No  |
| 19:43779295:C:T   | C   | 0.5264 | 0.4930 | 0.4342 | Yes | Yes |
| 19:13047463:T:C   | C   | -      | -      | 0.0478 | -   | -   |
| 2:217056046:G:T   | G   | 0.7007 | 0.6527 | 0.4967 | Yes | Yes |
| 2:172109838:C:G   | C   | 0.8027 | 0.7258 | 0.5250 | Yes | Yes |
| 2:120486696:T:C   | C   | 0.6491 | 0.6955 | 0.8976 | No  | No  |
| 2:201258707:C:T   | C   | 0.6366 | 0.5731 | 0.3315 | Yes | Yes |
| 2:69208111:C:T    | C   | 0.8997 | 0.8861 | 0.8482 | Yes | Yes |
| 2:226404332:T:A   | A   | 0.5959 | 0.6091 | 0.6551 | No  | No  |
| 20:53680310:G:A   | A   | 0.0392 | 0.0792 | 0.2389 | No  | No  |
| 21:15202135:C:A   | C   | 0.5620 | 0.5982 | 0.6831 | No  | No  |
| 22:40653398:G:T   | T   | 0.0001 | 0.0184 | 0.1006 | No  | No  |
| 22:28682717:T:G   | G   | 0.2005 | 0.1535 | 0.0145 | Yes | Yes |
| 22:38172826:C:G   | C   | 0.5668 | 0.5780 | 0.6226 | No  | No  |
| 22:19778614:C:T   | C   | 0.5243 | 0.5562 | 0.6251 | No  | No  |
| 3:27310652:T:A    | A   | 0.9900 | 0.8987 | 0.5251 | Yes | Yes |
| 3:30642605:G:A    | A   | 0.3035 | 0.3125 | 0.3251 | No  | No  |
| 3:141423849:G:A   | A   | 0.9178 | 0.8215 | 0.4416 | Yes | Yes |
| 3:4700567:A:G     | G   | 0.2455 | 0.2948 | 0.3882 | No  | No  |
| 3:46851683:G:GGT  | G   | 0.7475 | 0.7693 | 0.8951 | No  | No  |
| 3:100037468:G:A   | A   | -      | -      | 0.2209 | -   | -   |
| 3:172560322:GA:G  | G   | 0.1047 | 0.1247 | 0.2176 | No  | No  |
| 3:63956021:T:G    | G   | 0.0214 | 0.0437 | 0.1361 | No  | No  |
| 3:71482962:T:C    | T   | 0.6247 | 0.6057 | 0.5240 | Yes | Yes |
| 3:86988393:A:G    | A   | 0.9668 | 0.9612 | 0.9067 | Yes | Yes |
| 4:174906885:T:C   | C   | 0.5239 | 0.5889 | 0.8820 | No  | No  |
| 4:83448971:TAA:TA | TAA | -      | -      | 0.4659 | -   | -   |
| 4:105147856:G:T   | T   | 0.0004 | 0.0481 | 0.2268 | No  | No  |

Continuation of Table S 1.8

|                            |           |        |        |        |     |     |
|----------------------------|-----------|--------|--------|--------|-----|-----|
| 5:56757918:AAAAC:AAAACAAAC | AAAACAAAC | -      | -      | 0.1572 | -   | -   |
| 5:44706396:A:G             | G         | 0.1726 | 0.1930 | 0.2500 | No  | No  |
| 5:158803005:G:T            | G         | 0.8661 | 0.7986 | 0.4460 | Yes | Yes |
| 5:91408023:A:G             | A         | 0.5964 | 0.6582 | 0.8405 | No  | No  |
| 5:1299264:C:G              | C         | 0.3652 | 0.3918 | 0.4951 | No  | No  |
| 5:82196824:G:A             | A         | 0.9060 | 0.8845 | 0.7527 | Yes | Yes |
| 5:50899259:A:AT            | A         | -      | -      | 0.6761 | -   | -   |
| 5:16231085:G:C             | G         | 0.8622 | 0.7736 | 0.4395 | Yes | Yes |
| 5:132304843:A:G            | G         | 0.3403 | 0.3882 | 0.5402 | No  | No  |
| 5:170139969:TC:T           | T         | 0.0534 | 0.0961 | 0.2139 | No  | No  |
| 6:151634779:A:G            | G         | -      | -      | 0.0696 | -   | -   |
| 6:13713134:G:C             | G         | 0.4171 | 0.4107 | 0.4309 | No  | No  |
| 6:81577968:G:A             | G         | 0.9954 | 0.9179 | 0.5796 | Yes | Yes |
| 6:149265192:CT:C           | CT        | -      | -      | 0.7046 | -   | -   |
| 7:140243902:CT:C           | C         | -      | -      | 0.5306 | -   | -   |
| 7:94667457:T:C             | T         | 0.0342 | 0.0792 | 0.2713 | No  | No  |
| 7:130984153:G:A            | A         | 0.1172 | 0.1452 | 0.2665 | No  | No  |
| 7:144377836:G:A            | G         | 0.9245 | 0.8884 | 0.7468 | Yes | Yes |
| 8:127359926:A:G            | G         | 0.6060 | 0.6164 | 0.5425 | Yes | Yes |
| 8:75420821:C:T             | T         | 0.0516 | 0.0580 | 0.0860 | No  | No  |
| 8:29649578:T:C             | T         | 0.3868 | 0.3819 | 0.3243 | Yes | Yes |
| 8:36992428:C:G             | C         | 0.9868 | 0.9581 | 0.8250 | Yes | Yes |
| 8:105342969:C:G            | C         | 0.9833 | 0.9525 | 0.8476 | Yes | Yes |
| 8:101386664:A:T            | A         | 0.5284 | 0.5749 | 0.6513 | No  | No  |
| 8:123597610:G:C            | C         | 0.2679 | 0.2089 | 0.0835 | Yes | Yes |
| 9:108123199:C:T            | T         | 0.4663 | 0.5091 | 0.6184 | No  | No  |
| 9:21964883:CAAAA:C         | C         | -      | -      | 0.3184 | -   | -   |
| 9:116551207:A:G            | A         | 0.8930 | 0.8264 | 0.5957 | Yes | Yes |
| 1:121538815:A:G            | G         | 0.0617 | 0.1194 | 0.4040 | No  | No  |
| 1:18480845:T:C             | T         | 0.5374 | 0.5335 | 0.4805 | Yes | Yes |
| 1:149934520:T:C            | C         | 0.0154 | 0.0821 | 0.4009 | No  | No  |
| 1:155183132:A:G            | G         | 0.1935 | 0.2755 | 0.4465 | No  | No  |
| 1:46135245:A:G             | G         | 0.5886 | 0.6143 | 0.6644 | No  | No  |
| 1:217027238:C:T            | C         | 0.7250 | 0.7239 | 0.7901 | No  | No  |
| 1:113903258:G:A            | A         | 0.0049 | 0.0395 | 0.1664 | No  | No  |
| 1:168117196:G:A            | G         | 0.3426 | 0.3864 | 0.6629 | No  | No  |
| 1:220613960:G:A            | A         | 0.0043 | 0.0457 | 0.2281 | No  | No  |
| 10:121575416:G:A           | A         | 0.1681 | 0.2184 | 0.4004 | No  | No  |
| 10:79081391:T:C            | T         | 0.0192 | 0.0912 | 0.3802 | No  | No  |
| 10:62509005:T:C            | T         | 0.9986 | 0.9782 | 0.8671 | Yes | Yes |
| 10:21532989:A:T            | T         | 0.4471 | 0.4062 | 0.3588 | Yes | Yes |

Continuation of Table S 1.8

|                    |     |        |        |        |     |     |
|--------------------|-----|--------|--------|--------|-----|-----|
| 10:113373449:C:A   | A   | 0.4325 | 0.3890 | 0.2035 | Yes | Yes |
| 10:9058859:T:G     | G   | 0.0952 | 0.1072 | 0.1202 | No  | No  |
| 10:5707486:G:A     | A   | 0.0543 | 0.1111 | 0.3301 | No  | No  |
| 11:69516650:C:T    | T   | 0.0021 | 0.0125 | 0.0753 | No  | No  |
| 11:1877434:T:C     | T   | 0.8184 | 0.7727 | 0.6005 | Yes | Yes |
| 11:65753275:A:G    | G   | 0.0401 | 0.0688 | 0.1879 | No  | No  |
| 11:129593220:C:CAT | CAT | 0.8388 | 0.8142 | 0.5830 | Yes | Yes |
| 12:28021884:C:T    | C   | 0.9008 | 0.8571 | 0.7578 | Yes | Yes |
| 12:115398031:T:C   | T   | 0.5362 | 0.5541 | 0.5829 | No  | No  |
| 12:95633983:A:G    | A   | 0.8340 | 0.7991 | 0.7015 | Yes | Yes |
| 12:14255620:G:C    | C   | 0.0014 | 0.0378 | 0.1954 | No  | No  |
| 14:68513118:T:C    | T   | 0.9989 | 0.9489 | 0.7397 | Yes | Yes |
| 14:36662856:A:G    | A   | 0.8144 | 0.8049 | 0.7511 | Yes | Yes |
| 14:91382854:G:A    | A   | 0.0259 | 0.0531 | 0.1690 | No  | No  |
| 15:90988765:C:T    | C   | 0.8435 | 0.8564 | 0.8654 | No  | No  |
| 15:75640436:C:T    | C   | 0.4279 | 0.4710 | 0.7562 | No  | No  |
| 15:67165360:A:G    | G   | 0.0002 | 0.0104 | 0.0505 | No  | No  |
| 16:52565276:C:T    | T   | 0.0240 | 0.0661 | 0.2385 | No  | No  |
| 16:80614430:A:G    | G   | 0.7563 | 0.6418 | 0.2302 | Yes | Yes |
| 16:54648152:G:A    | A   | 0.9083 | 0.8377 | 0.4883 | Yes | Yes |
| 17:55097605:A:G    | G   | 0.3181 | 0.3920 | 0.6932 | No  | No  |
| 17:42417421:T:C    | C   | 0.0074 | 0.0472 | 0.2170 | No  | No  |
| 18:26920935:C:A    | C   | 0.7510 | 0.7421 | 0.7186 | Yes | Yes |
| 18:44831838:G:C    | G   | 0.6533 | 0.6910 | 0.9279 | No  | No  |
| 18:32351771:A:G    | A   | 0.7208 | 0.7500 | 0.9603 | No  | No  |
| 19:18479441:C:G    | C   | 0.1809 | 0.2439 | 0.6472 | No  | No  |
| 19:43778879:T:C    | C   | 0.1971 | 0.2230 | 0.3530 | No  | No  |
| 19:13142270:T:G    | G   | 0.0136 | 0.0201 | 0.0487 | No  | No  |
| 2:217056046:G:T    | G   | 0.7007 | 0.6527 | 0.4967 | Yes | Yes |
| 2:120486696:T:C    | C   | 0.6491 | 0.6955 | 0.8976 | No  | No  |
| 2:172109838:C:G    | C   | 0.8027 | 0.7258 | 0.5250 | Yes | Yes |
| 2:201258707:C:T    | C   | 0.6366 | 0.5731 | 0.3315 | Yes | Yes |
| 2:226410046:C:G    | G   | 0.5318 | 0.5442 | 0.6548 | No  | No  |
| 2:69208111:C:T     | C   | 0.8997 | 0.8861 | 0.8482 | Yes | Yes |
| 20:53680310:G:A    | A   | 0.0392 | 0.0792 | 0.2389 | No  | No  |
| 21:15202135:C:A    | C   | 0.5620 | 0.5982 | 0.6831 | No  | No  |
| 21:46360308:T:C    | T   | 0.0043 | 0.0166 | 0.0729 | No  | No  |
| 22:40653398:G:T    | T   | 0.0001 | 0.0184 | 0.1006 | No  | No  |
| 22:38202034:C:T    | C   | 0.7282 | 0.7482 | 0.7017 | Yes | Yes |
| 22:28712568:A:T    | T   | 0.0543 | 0.0588 | 0.0952 | No  | No  |
| 3:27320389:T:C     | C   | 0.9916 | 0.8989 | 0.5251 | Yes | Yes |

Continuation of Table S 1.8

|                   |       |        |        |        |     |     |
|-------------------|-------|--------|--------|--------|-----|-----|
| 3:4700567:A:G     | G     | 0.2455 | 0.2948 | 0.3882 | No  | No  |
| 3:30642605:G:A    | A     | 0.3035 | 0.3125 | 0.3251 | No  | No  |
| 3:141396377:T:C   | C     | 0.8941 | 0.8116 | 0.4389 | Yes | Yes |
| 3:46851683:G:GGT  | G     | 0.7475 | 0.7693 | 0.8951 | No  | No  |
| 3:100048697:G:A   | A     | 0.0510 | 0.0869 | 0.2256 | No  | No  |
| 3:172560322:GA:G  | G     | 0.1047 | 0.1247 | 0.2176 | No  | No  |
| 3:63895873:A:C    | C     | 0.1788 | 0.1694 | 0.1331 | Yes | Yes |
| 3:71533370:G:A    | G     | 0.9310 | 0.8790 | 0.5794 | Yes | Yes |
| 3:86988393:A:G    | A     | 0.9668 | 0.9612 | 0.9067 | Yes | Yes |
| 4:174923119:G:A   | G     | 0.4873 | 0.5373 | 0.8823 | No  | No  |
| 4:105222285:G:T   | T     | 0.0018 | 0.0172 | 0.0937 | No  | No  |
| 4:83472322:C:T    | C     | 0.0879 | 0.1603 | 0.5008 | No  | No  |
| 4:1983005:T:TAACA | TAACA | 0.1666 | 0.2014 | 0.3598 | No  | No  |
| 5:56757652:C:T    | T     | 0.0376 | 0.0666 | 0.1579 | No  | No  |
| 5:44706396:A:G    | G     | 0.1726 | 0.1930 | 0.2500 | No  | No  |
| 5:158803005:G:T   | G     | 0.8661 | 0.7986 | 0.4460 | Yes | Yes |
| 5:91430632:CT:C   | CT    | 0.5183 | 0.5894 | 0.8377 | No  | No  |
| 5:1293971:C:T     | C     | 0.9431 | 0.8849 | 0.7338 | Yes | Yes |
| 5:82196824:G:A    | A     | 0.9060 | 0.8845 | 0.7527 | Yes | Yes |
| 5:16231085:G:C    | G     | 0.8622 | 0.7736 | 0.4395 | Yes | Yes |
| 5:50916286:C:T    | C     | 0.7596 | 0.7372 | 0.6158 | Yes | Yes |
| 5:170141842:C:T   | T     | 0.0025 | 0.0502 | 0.2224 | No  | No  |
| 6:151632630:T:C   | C     | 0.0133 | 0.0310 | 0.0676 | No  | No  |
| 6:13717681:G:A    | G     | 0.3274 | 0.3432 | 0.4387 | No  | No  |
| 6:81467674:A:T    | A     | 0.6443 | 0.6229 | 0.5666 | Yes | Yes |
| 6:149274369:T:C   | T     | 0.8511 | 0.8313 | 0.7957 | Yes | Yes |
| 7:140242504:G:A   | A     | 0.2414 | 0.2693 | 0.4313 | No  | No  |
| 7:130984153:G:A   | A     | 0.1172 | 0.1452 | 0.2665 | No  | No  |
| 7:94667457:T:C    | T     | 0.0342 | 0.0792 | 0.2713 | No  | No  |
| 7:144377836:G:A   | G     | 0.9245 | 0.8884 | 0.7468 | Yes | Yes |
| 8:127338270:G:A   | A     | 0.5764 | 0.5956 | 0.5402 | Yes | Yes |
| 8:75420821:C:T    | T     | 0.0516 | 0.0580 | 0.0860 | No  | No  |
| 8:29649578:T:C    | T     | 0.3868 | 0.3819 | 0.3243 | Yes | Yes |
| 8:37000965:A:G    | A     | 0.5359 | 0.6016 | 0.8246 | No  | No  |
| 8:105345146:A:G   | G     | 0.1313 | 0.2019 | 0.5320 | No  | No  |
| 8:101396319:G:A   | A     | 0.3828 | 0.3300 | 0.1801 | Yes | Yes |
| 8:115667320:A:G   | A     | 0.1352 | 0.2254 | 0.6284 | No  | No  |
| 8:123597926:G:A   | A     | 0.3597 | 0.2901 | 0.1471 | Yes | Yes |
| 8:25974262:C:T    | T     | 0.0507 | 0.0728 | 0.1980 | No  | No  |
| 9:108130619:C:A   | A     | 0.7779 | 0.7575 | 0.6213 | Yes | Yes |
| 9:116552572:C:T   | C     | 0.5854 | 0.5912 | 0.5907 | No  | Yes |

Continuation of Table S 1.8

|                    |    |        |        |        |     |     |
|--------------------|----|--------|--------|--------|-----|-----|
| 9:21966222:G:A     | G  | 0.9303 | 0.8969 | 0.7445 | Yes | Yes |
| Luminal-B-like     |    |        |        |        |     |     |
| 10:123814971:T:C   | C  | 0.0415 | 0.0351 | -      | -   | -   |
| 12:96591322:G:A    | G  | 0.0255 | 0.0412 | 0.1285 | No  | No  |
| 1:10520994:A:T     | A  | 0.9277 | 0.8830 | 0.6399 | Yes | Yes |
| 10:121577668:T:C   | T  | 0.6550 | 0.6245 | 0.4039 | Yes | Yes |
| 10:62528371:G:C    | G  | 0.8801 | 0.8680 | 0.8443 | Yes | Yes |
| 11:1881245:G:A     | G  | 0.7319 | 0.7159 | 0.6166 | Yes | Yes |
| 11:69516650:C:T    | T  | 0.0021 | 0.0125 | 0.0753 | No  | No  |
| 15:90965985:C:A    | C  | 0.9819 | 0.9638 | 0.8672 | Yes | Yes |
| 16:52538117:T:C    | C  | 0.3958 | 0.3789 | 0.2619 | Yes | Yes |
| 2:217056046:G:T    | G  | 0.7007 | 0.6527 | 0.4967 | Yes | Yes |
| 22:28681643:C:CA   | CA | -      | -      | 0.0220 | -   | -   |
| 4:13578908:A:G     | G  | 0.0447 | 0.0694 | 0.1627 | No  | No  |
| 5:56757708:T:C     | C  | 0.0769 | 0.1008 | 0.1599 | No  | No  |
| 6:152144481:AGT:A  | A  | 0.2106 | 0.2281 | 0.4316 | No  | No  |
| 8:75420821:C:T     | T  | 0.0516 | 0.0580 | 0.0860 | No  | No  |
| 8:127358703:C:G    | G  | 0.5285 | 0.5176 | 0.3989 | Yes | Yes |
| 9:108124826:G:A    | A  | 0.4657 | 0.5143 | 0.6203 | No  | No  |
| 1:10494652:G:T     | T  | 0.4421 | 0.5068 | 0.6987 | No  | No  |
| 10:121577668:T:C   | T  | 0.6550 | 0.6245 | 0.4039 | Yes | Yes |
| 10:62513267:G:A    | G  | 0.9563 | 0.9281 | 0.8452 | Yes | Yes |
| 11:69517902:A:G    | G  | 0.1208 | 0.1246 | 0.1238 | No  | Yes |
| 11:1926345:C:A     | C  | 0.7698 | 0.7341 | 0.6267 | Yes | Yes |
| 15:90963407:T:C    | T  | 0.7089 | 0.7632 | 0.8646 | No  | No  |
| 16:52565276:C:T    | T  | 0.0240 | 0.0661 | 0.2385 | No  | No  |
| 2:217042791:C:G    | C  | 0.7914 | 0.7319 | 0.5416 | Yes | Yes |
| 2:173354390:C:G    | G  | 0.7356 | 0.7637 | 0.8420 | No  | No  |
| 5:56757708:T:C     | C  | 0.0769 | 0.1008 | 0.1599 | No  | No  |
| 6:152113143:A:AC   | AC | 0.5500 | 0.5449 | 0.5252 | Yes | Yes |
| 8:75420821:C:T     | T  | 0.0516 | 0.0580 | 0.0860 | No  | No  |
| 8:127365421:G:T    | G  | 0.3861 | 0.3995 | 0.4309 | No  | No  |
| 9:108124254:G:A    | A  | 0.4554 | 0.4925 | 0.6199 | No  | No  |
| HER2-enriched-like |    |        |        |        |     |     |
| 1:10489791:G:C     | C  | 0.9335 | 0.8892 | 0.7018 | Yes | Yes |
| 1:202218991:T:C    | T  | 0.6025 | 0.5780 | 0.5393 | Yes | Yes |
| 16:52565276:C:T    | T  | 0.0240 | 0.0661 | 0.2385 | No  | No  |
| 6:151632045:A:G    | G  | 0.3692 | 0.3160 | 0.0699 | Yes | Yes |
| 1:10543659:T:C     | T  | 0.8764 | 0.8427 | 0.6815 | Yes | Yes |
| 16:52565276:C:T    | T  | 0.0240 | 0.0661 | 0.2385 | No  | No  |
| 6:151630867:A:G    | G  | 0.0736 | 0.0795 | 0.0674 | Yes | Yes |

Continuation of Table S 1.8

| Triple-negative  |     |        |        |        |     |     |
|------------------|-----|--------|--------|--------|-----|-----|
| 1:31087245:A:G   | G   | 0.0479 | 0.0463 | -      | -   | -   |
| 11:133440533:A:C | C   | 0.8646 | 0.7380 | 0.2068 | Yes | Yes |
| 12:27719121:G:C  | G   | 0.0306 | 0.0130 | -      | -   | -   |
| 19:17288816:T:C  | T   | 0.6145 | 0.5843 | 0.4870 | Yes | Yes |
| 2:118823485:A:G  | A   | 0.7743 | 0.8106 | -      | -   | -   |
| 5:1279675:C:T    | T   | 0.6720 | 0.5826 | 0.2586 | Yes | Yes |
| 5:6120548:T:C    | C   | 0.3760 | 0.3052 | -      | -   | -   |
| 1:204528785:G:C  | G   | 0.1759 | 0.2140 | 0.2599 | No  | No  |
| 10:113023044:T:C | C   | 0.8267 | 0.7500 | 0.4598 | Yes | Yes |
| 11:108440928:T:C | T   | 0.6031 | 0.6094 | 0.5847 | Yes | Yes |
| 12:28021884:C:T  | C   | 0.9008 | 0.8571 | 0.7578 | Yes | Yes |
| 16:52504913:C:A  | A   | 0.0289 | 0.0702 | 0.2534 | No  | No  |
| 19:17290595:C:G  | G   | 0.2880 | 0.2717 | 0.2966 | No  | No  |
| 2:120330606:G:A  | G   | 0.5324 | 0.5814 | 0.7596 | No  | No  |
| 2:28823002:CTT:C | CTT | 0.9174 | 0.9249 | 0.9165 | Yes | Yes |
| 22:40629878:GT:G | G   | 0.0494 | 0.0641 | 0.1064 | No  | No  |
| 5:1279675:C:T    | T   | 0.6720 | 0.5826 | 0.2586 | Yes | Yes |
| 6:151634779:A:G  | G   | -      | -      | 0.0696 | -   | -   |
| 6:33272092:C:T   | C   | 0.9980 | 0.9864 | 0.9225 | Yes | Yes |
| 1:204572069:G:A  | G   | 0.2045 | 0.2359 | 0.2597 | No  | No  |
| 1:9156852:C:T    | T   | 0.0486 | 0.0525 | 0.0312 | Yes | Yes |
| 10:113019779:G:A | A   | 0.8685 | 0.7824 | 0.4606 | Yes | Yes |
| 11:108444159:A:G | A   | 0.6626 | 0.6559 | 0.5854 | Yes | Yes |
| 12:28021884:C:T  | C   | 0.9008 | 0.8571 | 0.7578 | Yes | Yes |
| 16:52504128:T:C  | C   | 0.0219 | 0.0635 | 0.2534 | No  | No  |
| 19:17291050:G:A  | A   | 0.4448 | 0.4022 | 0.2970 | Yes | Yes |
| 2:28922682:C:G   | C   | 0.8946 | 0.9062 | 0.9180 | No  | No  |
| 2:120330606:G:A  | G   | 0.5324 | 0.5814 | 0.7596 | No  | No  |
| 2:19121207:C:T   | C   | 0.6685 | 0.6238 | 0.4385 | Yes | Yes |
| 22:40481000:T:C  | C   | 0.2434 | 0.2340 | 0.1922 | Yes | Yes |
| 3:172523209:T:G  | T   | 0.8835 | 0.8672 | 0.8255 | Yes | Yes |
| 5:1279675:C:T    | T   | 0.6720 | 0.5826 | 0.2586 | Yes | Yes |
| 6:151628177:G:A  | A   | 0.5235 | 0.4760 | 0.3168 | Yes | Yes |
| 6:33272092:C:T   | C   | 0.9980 | 0.9864 | 0.9225 | Yes | Yes |

**Footnote:**

All listed lead variants were identified at a p-value threshold of  $1.25 \times 10^{-8}$

RAF, Risk allele frequency in controls





Table S 1.9 All TWAS genes identified to associate with breast cancer intrinsic subtypes across ancestries and cell types (Bonferroni-adjusted p-value<0.05)

| Loci           | Genomic Coordinates   | Gene Symbol          | Cell Type         | Ancestry          | TWA S Z-score | TWAS P-value | Evidence for subtype-specificity* | Cond-TWAS Z-score | Cond-TWAS P-value | Bonferroni-significant in Cond-TWAS | Nearest GWAS Index Variant | Distance to Nearest Index Variant (kb) |
|----------------|-----------------------|----------------------|-------------------|-------------------|---------------|--------------|-----------------------------------|-------------------|-------------------|-------------------------------------|----------------------------|----------------------------------------|
| Luminal-A-like |                       |                      |                   |                   |               |              |                                   |                   |                   |                                     |                            |                                        |
| 1p36.13        | 1:18480982-18486126   | <i>KLHDC7A</i>       | Adipocytes        | European-ancestry | -7.340        | 2.14E-13     | No                                | 0.060             | 0.95252           | No                                  | 1:18480845:T:C             | 137                                    |
| 1p36.13        | 1:18480982-18486126   | <i>KLHDC7A</i>       | Endothelial cells | European-ancestry | -7.734        | 1.04E-14     | No                                | -0.627            | 0.530343          | No                                  | 1:18480845:T:C             | 137                                    |
| 1p36.13        | 1:18480982-18486126   | <i>KLHDC7A</i>       | Adipocytes        | Cross-ancestry    | -7.263        | 3.80E-13     | No                                | -0.141            | 0.887776          | No                                  | 1:18480845:T:C             | 137                                    |
| 1p36.13        | 1:18480982-18486126   | <i>KLHDC7A</i>       | Endothelial cells | Cross-ancestry    | -7.755        | 8.84E-15     | No                                | -0.955            | 0.339794          | No                                  | 1:18480845:T:C             | 137                                    |
| 1p34.1         | 1:46203334-46221261   | <i>LURAP1</i>        | Breast tissue     | European-ancestry | -5.826        | 5.68E-09     | No                                | -2.443            | 0.014556          | No                                  | 1:46135525:A:G             | 67809                                  |
| 1p34.1         | 1:46203334-46221261   | <i>LURAP1</i>        | Breast tissue     | Cross-ancestry    | -5.744        | 9.25E-09     | No                                | -2.319            | 0.020405          | No                                  | 1:46135525:A:G             | 67809                                  |
| 1p33           | 1:46340177-46365152   | <i>NSUN4</i>         | Adipocytes        | European-ancestry | -5.464        | 4.65E-08     | Yes                               | -3.386            | 0.00071           | No                                  | 1:46135525:A:G             | 204652                                 |
| 1p33           | 1:46340177-46365152   | <i>NSUN4</i>         | Breast tissue     | European-ancestry | -5.539        | 3.04E-08     | Yes                               | -3.285            | 0.00102           | No                                  | 1:46135525:A:G             | 204652                                 |
| 1p12           | 1:119564066-119572067 | <i>HSD3BP4</i>       | Adipocytes        | European-ancestry | -6.106        | 1.02E-09     | Yes                               | -4.179            | 2.93E-05          | No                                  | 1:121538815:A:G            | 1966748                                |
| 1p12           | 1:119564066-119572067 | <i>HSD3BP4</i>       | Breast tissue     | European-ancestry | -5.442        | 5.26E-08     | Yes                               | -3.449            | 0.000563          | No                                  | 1:121538815:A:G            | 1966748                                |
| 1p12           | 1:120267334-120341871 | <i>CH17-472G23.2</i> | Breast tissue     | European-ancestry | 7.488         | 6.97E-14     | No                                | -1.000            | 0.317143          | No                                  | 1:121538815:A:G            | 1196944                                |
| 1p12           | 1:120267334-120341871 | <i>CH17-472G23.2</i> | Epithelial cells  | European-ancestry | 7.791         | 6.63E-15     | No                                | -0.222            | 0.824252          | No                                  | 1:121538815:A:G            | 1196944                                |
| 1p12           | 1:120267334-120341871 | <i>CH17-472G23.2</i> | Fibroblasts       | European-ancestry | 6.973         | 3.11E-12     | No                                | -1.295            | 0.195296          | No                                  | 1:121538815:A:G            | 1196944                                |
| 1p11.2         | 1:120436353-120467739 | <i>NBPFF8</i>        | Endothelial cells | European-ancestry | -8.025        | 1.01E-15     | Yes                               | -1.327            | 0.184446          | No                                  | 1:121538815:A:G            | 1071076                                |
| 1p11.2         | 1:120723923-120841481 | <i>NBPFF26</i>       | Adipocytes        | European-ancestry | 6.935         | 4.08E-12     | No                                | -2.361            | 0.018221          | No                                  | 1:121538815:A:G            | 697334                                 |
| 1p11.2         | 1:120723923-120841481 | <i>NBPFF26</i>       | Breast tissue     | European-ancestry | 8.122         | 4.60E-16     | No                                | -1.857            | 0.063351          | No                                  | 1:121538815:A:G            | 697334                                 |
| 1p11.2         | 1:120723923-120841481 | <i>NBPFF26</i>       | Epithelial cells  | European-ancestry | 8.952         | 3.50E-19     | No                                | -2.374            | 0.017611          | No                                  | 1:121538815:A:G            | 697334                                 |
| 1p11.2         | 1:120723923-120841481 | <i>NBPFF26</i>       | Fibroblasts       | European-ancestry | 6.500         | 8.01E-11     | No                                | -2.132            | 0.032999          | No                                  | 1:121538815:A:G            | 697334                                 |
| 1p11.2         | 1:120723923-120841481 | <i>NBPFF26</i>       | Breast tissue     | Cross-ancestry    | 7.874         | 3.44E-15     | No                                | -1.862            | 0.062629          | No                                  | 1:121538815:A:G            | 697334                                 |
| 1p11.2         | 1:120723923-120841481 | <i>NBPFF26</i>       | Epithelial cells  | Cross-ancestry    | 8.571         | 1.03E-17     | No                                | -2.280            | 0.0226            | No                                  | 1:121538815:A:G            | 697334                                 |
| 1p11.2         | 1:120850819-120850985 | <i>RNVU1-19</i>      | Breast tissue     | European-ancestry | 7.425         | 1.12E-13     | No                                | -1.158            | 0.246788          | No                                  | 1:121538815:A:G            | 687830                                 |
| 1p11.2         | 1:120850819-120850985 | <i>RNVU1-19</i>      | Breast tissue     | Cross-ancestry    | 6.946         | 3.75E-12     | No                                | -1.342            | 0.179566          | No                                  | 1:121538815:A:G            | 687830                                 |
| 1p11.2         | 1:121108210-121117257 | <i>HIST2H2BA</i>     | Breast tissue     | European-ancestry | 12.870        | 6.64E-38     | Yes                               | 0.472             | 0.636571          | No                                  | 1:121538815:A:G            | 421558                                 |
| 1p11.2         | 1:121108210-121117257 | <i>HIST2H2BA</i>     | Breast tissue     | Cross-ancestry    | 12.714        | 4.93E-37     | Yes                               | 0.995             | 0.319499          | No                                  | 1:121538815:A:G            | 421558                                 |
| 1p11.2         | 1:121118195-121118610 | <i>HIST2H3DP1</i>    | Breast tissue     | European-ancestry | 13.968        | 2.45E-44     | No                                | 3.841             | 0.000123          | No                                  | 1:121538815:A:G            | 420205                                 |
| 1p11.2         | 1:121118195-121118610 | <i>HIST2H3DP1</i>    | Endothelial cells | European-ancestry | 13.840        | 1.46E-43     | No                                | 3.390             | 0.000699          | No                                  | 1:121538815:A:G            | 420205                                 |
| 1p11.2         | 1:121118195-121118610 | <i>HIST2H3DP1</i>    | Breast tissue     | Cross-ancestry    | 13.210        | 7.65E-40     | No                                | 3.515             | 0.000439          | No                                  | 1:121538815:A:G            | 420205                                 |
| 1p11.2         | 1:121184810-121392822 | <i>SRGAP2C</i>       | Breast tissue     | European-ancestry | -12.261       | 1.48E-34     | No                                | 0.446             | 0.655798          | No                                  | 1:121538815:A:G            | 145993                                 |
| 1p11.2         | 1:121184810-121392822 | <i>SRGAP2C</i>       | Epithelial cells  | European-ancestry | -12.455       | 1.31E-35     | No                                | 0.434             | 0.663983          | No                                  | 1:121538815:A:G            | 145993                                 |
| 1p11.2         | 1:121184810-121392822 | <i>SRGAP2C</i>       | Breast tissue     | Cross-ancestry    | -11.299       | 1.32E-29     | No                                | 0.545             | 0.58595           | No                                  | 1:121538815:A:G            | 145993                                 |
| 1p11.2         | 1:121396754-121463129 | <i>CH17-437K3.1</i>  | Breast tissue     | European-ancestry | -10.102       | 5.42E-24     | Yes                               | 0.175             | 0.860846          | No                                  | 1:121538815:A:G            | 75686                                  |
| 1p11.2         | 1:121396754-121463129 | <i>CH17-437K3.1</i>  | Breast tissue     | Cross-ancestry    | -10.122       | 4.41E-24     | Yes                               | -0.272            | 0.78535           | No                                  | 1:121538815:A:G            | 75686                                  |

# Continuation of Table S 1.9

|         |                       |                   |                   |                   |         |          |     |        |          |     |                     |          |
|---------|-----------------------|-------------------|-------------------|-------------------|---------|----------|-----|--------|----------|-----|---------------------|----------|
| 1p11.2  | 1:121519112-121571892 | <i>EMBP1</i>      | Breast tissue     | European-ancestry | -8.965  | 3.11E-19 | Yes | -3.221 | 0.001277 | No  | 1:121538815:A:G     | 0        |
| 1p11.2  | 1:121519112-121571892 | <i>EMBP1</i>      | Breast tissue     | Cross-ancestry    | -9.075  | 1.14E-19 | Yes | -3.400 | 0.000674 | No  | 1:121538815:A:G     | 0        |
| 1q22    | 1:155195588-155209051 | <i>THBS3</i>      | Breast tissue     | European-ancestry | 5.601   | 2.13E-08 | No  | -0.053 | 0.958123 | No  | 1:155183132:A:G     | 12456    |
| 1q22    | 1:155195588-155209051 | <i>THBS3</i>      | Adipocytes        | Cross-ancestry    | 5.573   | 2.50E-08 | No  | 1.914  | 0.055677 | No  | 1:155183132:A:G     | 12456    |
| 1q22    | 1:155195588-155209051 | <i>THBS3</i>      | Breast tissue     | Cross-ancestry    | 5.781   | 7.45E-09 | No  | 0.573  | 0.566459 | No  | 1:155183132:A:G     | 12456    |
| 1q22    | 1:155213821-155227422 | <i>GBAP1</i>      | Adipocytes        | European-ancestry | -5.762  | 8.31E-09 | No  | -0.220 | 0.826156 | No  | 1:155225280:ATTAT:A | 0        |
| 1q22    | 1:155213821-155227422 | <i>GBAP1</i>      | Breast tissue     | European-ancestry | -6.130  | 8.81E-10 | No  | -1.051 | 0.2932   | No  | 1:155225280:ATTAT:A | 0        |
| 1q22    | 1:155213821-155227422 | <i>GBAP1</i>      | Fibroblasts       | European-ancestry | -6.203  | 5.52E-10 | No  | -0.685 | 0.49352  | No  | 1:155225280:ATTAT:A | 0        |
| 1q22    | 1:155213821-155227422 | <i>GBAP1</i>      | Adipocytes        | Cross-ancestry    | -5.716  | 1.09E-08 | No  | -0.441 | 0.65918  | No  | 1:155225280:ATTAT:A | 0        |
| 1q22    | 1:155213821-155227422 | <i>GBAP1</i>      | Breast tissue     | Cross-ancestry    | -6.207  | 5.41E-10 | No  | -1.512 | 0.130545 | No  | 1:155225280:ATTAT:A | 0        |
| 1q22    | 1:155745829-155750137 | <i>MSTO2P</i>     | Epithelial cells  | European-ancestry | -6.073  | 1.26E-09 | No  | -2.898 | 0.00376  | No  | 1:155225280:ATTAT:A | 520549   |
| 1q22    | 1:155897808-155911404 | <i>RIT1</i>       | Breast tissue     | European-ancestry | -5.351  | 8.73E-08 | No  | -1.895 | 0.058122 | No  | 1:155225280:ATTAT:A | 672528   |
| 1q22    | 1:155897808-155911404 | <i>RIT1</i>       | Endothelial cells | European-ancestry | -5.845  | 5.08E-09 | No  | -2.905 | 0.003672 | No  | 1:155225280:ATTAT:A | 672528   |
| 1q22    | 1:155897808-155911404 | <i>RIT1</i>       | Fibroblasts       | European-ancestry | -5.275  | 1.33E-07 | No  | -1.856 | 0.063464 | No  | 1:155225280:ATTAT:A | 672528   |
| 1q22    | 1:155897808-155911404 | <i>RIT1</i>       | Endothelial cells | Cross-ancestry    | -5.893  | 3.78E-09 | No  | -2.401 | 0.01636  | No  | 1:155225280:ATTAT:A | 672528   |
| 2q31.1  | 2:171522247-171558133 | <i>CYBRD1</i>     | Endothelial cells | Cross-ancestry    | -5.570  | 2.55E-08 | No  | -4.310 | 1.63E-05 | No  | 2:172109838:C:G     | 551705   |
| 2q33.1  | 2:201233443-201287711 | <i>CASP8</i>      | Breast tissue     | European-ancestry | -5.669  | 1.43E-08 | No  | 0.096  | 0.923912 | No  | 2:201258707:C:T     | 0        |
| 2q33.1  | 2:201233443-201287711 | <i>CASP8</i>      | Endothelial cells | European-ancestry | -6.140  | 8.23E-10 | No  | -0.376 | 0.707116 | No  | 2:201258707:C:T     | 0        |
| 2q33.1  | 2:201233443-201287711 | <i>CASP8</i>      | Fibroblasts       | European-ancestry | -5.644  | 1.66E-08 | No  | 0.038  | 0.97008  | No  | 2:201258707:C:T     | 0        |
| 2q33.1  | 2:201233443-201287711 | <i>CASP8</i>      | Breast tissue     | Cross-ancestry    | -6.316  | 2.69E-10 | No  | -0.027 | 0.978725 | No  | 2:201258707:C:T     | 0        |
| 2q33.1  | 2:201233443-201287711 | <i>CASP8</i>      | Fibroblasts       | Cross-ancestry    | -6.038  | 1.56E-09 | No  | 0.143  | 0.885909 | No  | 2:201258707:C:T     | 0        |
| 2q33.1  | 2:201288271-201357398 | <i>ALS2CR12</i>   | Fibroblasts       | European-ancestry | 5.412   | 6.23E-08 | No  | 0.194  | 0.846315 | No  | 2:201258707:C:T     | 29564    |
| 2q35    | 2:216107464-216206303 | <i>XRCC5</i>      | Breast tissue     | European-ancestry | 6.185   | 6.20E-10 | Yes | 4.066  | 4.78E-05 | No  | 2:217056046:G:T     | 849743   |
| 2q35    | 2:216479030-216498761 | <i>AC098820.3</i> | Adipocytes        | European-ancestry | 6.119   | 9.44E-10 | Yes | 7.441  | 1.00E-13 | Yes | 2:217056046:G:T     | 557285   |
| 2q35    | 2:216479030-216498761 | <i>AC098820.3</i> | Breast tissue     | European-ancestry | 8.792   | 1.46E-18 | Yes | 9.520  | 1.73E-21 | Yes | 2:217056046:G:T     | 557285   |
| 3p24.1  | 3:26622806-26710776   | <i>LRRC3B</i>     | Adipocytes        | European-ancestry | -12.152 | 5.59E-34 | No  | -2.958 | 0.003098 | No  | 3:27310652:T:A      | 599876   |
| 3p24.1  | 3:26622806-26710776   | <i>LRRC3B</i>     | Breast tissue     | European-ancestry | -13.834 | 1.59E-43 | No  | -3.330 | 0.000867 | No  | 3:27310652:T:A      | 599876   |
| 3p24.1  | 3:26622806-26710776   | <i>LRRC3B</i>     | Endothelial cells | European-ancestry | -13.797 | 2.65E-43 | No  | -3.257 | 0.001127 | No  | 3:27310652:T:A      | 599876   |
| 3p24.1  | 3:26622806-26710776   | <i>LRRC3B</i>     | Epithelial cells  | European-ancestry | -7.830  | 4.87E-15 | No  | -2.015 | 0.043947 | No  | 3:27310652:T:A      | 599876   |
| 3p24.1  | 3:26622806-26710776   | <i>LRRC3B</i>     | Fibroblasts       | European-ancestry | -12.236 | 2.00E-34 | No  | -2.857 | 0.00427  | No  | 3:27310652:T:A      | 599876   |
| 3p14.1  | 3:64011964-64016246   | <i>PSMD6-AS1</i>  | Epithelial cells  | European-ancestry | -5.892  | 3.82E-09 | Yes | -1.747 | 0.080629 | No  | 3:63956021:T:G      | 55943    |
| 3p14.1  | 3:64011964-64016246   | <i>PSMD6-AS1</i>  | Epithelial cells  | Cross-ancestry    | -5.586  | 2.32E-08 | Yes | -1.844 | 0.065188 | No  | 3:63956021:T:G      | 55943    |
| 3q25.31 | 3:156713884-156714928 | <i>METTL15P1</i>  | Breast tissue     | European-ancestry | -5.310  | 1.09E-07 | No  | NA     | NA       | NA  | 3:141423849:G:A     | 15290035 |
| 3q25.31 | 3:156809551-156810732 | <i>PA2G4P4</i>    | Epithelial cells  | European-ancestry | -5.470  | 4.51E-08 | No  | NA     | NA       | NA  | 3:141423849:G:A     | 15385702 |
| 3q25.31 | 3:156809551-156810732 | <i>PA2G4P4</i>    | Epithelial cells  | Cross-ancestry    | -5.422  | 5.90E-08 | No  | NA     | NA       | NA  | 3:141423849:G:A     | 15385702 |
| 4q21.23 | 4:83477524-83478424   | <i>SLC25A14P1</i> | Breast tissue     | European-ancestry | -5.688  | 1.29E-08 | No  | -0.525 | 0.599564 | No  | 4:83472322:C:T      | 5202     |
| 4q21.23 | 4:83477524-83478424   | <i>SLC25A14P1</i> | Epithelial cells  | European-ancestry | -5.721  | 1.06E-08 | No  | -0.571 | 0.568269 | No  | 4:83472322:C:T      | 5202     |
| 4q21.23 | 4:83477524-83478424   | <i>SLC25A14P1</i> | Breast tissue     | Cross-ancestry    | -5.361  | 8.27E-08 | No  | -0.458 | 0.646665 | No  | 4:83472322:C:T      | 5202     |
| 4q34.1  | 4:174636914-174829314 | <i>GLRA3</i>      | Breast tissue     | European-ancestry | -7.585  | 3.32E-14 | No  | -3.063 | 0.002193 | No  | 4:174906885:T:C     | 77571    |
| 4q34.1  | 4:174636914-174829314 | <i>GLRA3</i>      | Fibroblasts       | European-ancestry | -6.835  | 8.22E-12 | No  | -2.541 | 0.011062 | No  | 4:174906885:T:C     | 77571    |
| 4q34.1  | 4:174636914-174829314 | <i>GLRA3</i>      | Breast tissue     | Cross-ancestry    | -7.391  | 1.46E-13 | No  | -3.161 | 0.001573 | No  | 4:174906885:T:C     | 77571    |

# Continuation of Table S 1.9

|              |                       |               |                   |                   |            |          |     |        |              |    |                                    |        |
|--------------|-----------------------|---------------|-------------------|-------------------|------------|----------|-----|--------|--------------|----|------------------------------------|--------|
| 4q34.1       | 4:174636914-174829314 | GLRA3         | Fibroblasts       | Cross-ancestry    | -<br>5.881 | 4.09E-09 | No  | -1.745 | 0.0809<br>82 | No | 4:174906885:<br>T:C                | 77571  |
| 5p12         | 5:44744900-44808777   | RP11-53O19.1  | Adipocytes        | European-ancestry | 11.61<br>7 | 3.37E-31 | No  | -0.312 | 0.7554<br>06 | No | 5:44706396:A<br>:G                 | 38504  |
| 5p12         | 5:44744900-44808777   | RP11-53O19.1  | Breast tissue     | European-ancestry | 12.32<br>5 | 6.62E-35 | No  | 0.385  | 0.7000<br>24 | No | 5:44706396:A<br>:G                 | 38504  |
| 5p12         | 5:44744900-44808777   | RP11-53O19.1  | Endothelial cells | European-ancestry | 11.88<br>8 | 1.36E-32 | No  | 0.705  | 0.4810<br>71 | No | 5:44706396:A<br>:G                 | 38504  |
| 5p12         | 5:44744900-44808777   | RP11-53O19.1  | Fibroblasts       | European-ancestry | 10.33<br>6 | 4.85E-25 | No  | 0.852  | 0.3943<br>35 | No | 5:44706396:A<br>:G                 | 38504  |
| 5p12         | 5:44744900-44808777   | RP11-53O19.1  | Breast tissue     | Cross-ancestry    | 12.73<br>4 | 3.85E-37 | No  | 1.849  | 0.0644<br>95 | No | 5:44706396:A<br>:G                 | 38504  |
| 5p12         | 5:44744900-44808777   | RP11-53O19.1  | Fibroblasts       | Cross-ancestry    | 9.151      | 5.62E-20 | No  | 0.605  | 0.5451<br>27 | No | 5:44706396:A<br>:G                 | 38504  |
| 5q11.1       | 5:50665899-50846522   | PARP8         | Endothelial cells | European-ancestry | 5.394      | 6.90E-08 | No  | 3.608  | 0.0003<br>08 | No | 5:50899259:A<br>:AT                | 52737  |
| 5q11.2       | 5:56927874-56929573   | CTD-2310F14.1 | Adipocytes        | European-ancestry | -<br>8.936 | 4.03E-19 | No  | -0.647 | 0.5174<br>66 | No | 5:56757918:A<br>AAAC:AAA<br>ACAAAC | 169956 |
| 5q11.2       | 5:56927874-56929573   | CTD-2310F14.1 | Breast tissue     | European-ancestry | 15.28<br>8 | 9.22E-53 | No  | 2.147  | 0.0317<br>6  | No | 5:56757918:A<br>AAAC:AAA<br>ACAAAC | 169956 |
| 5q11.2       | 5:56927874-56929573   | CTD-2310F14.1 | Endothelial cells | European-ancestry | 8.068      | 7.14E-16 | No  | 1.173  | 0.2407<br>58 | No | 5:56757918:A<br>AAAC:AAA<br>ACAAAC | 169956 |
| 5q11.2       | 5:56927874-56929573   | CTD-2310F14.1 | Adipocytes        | Cross-ancestry    | -<br>7.825 | 5.06E-15 | No  | -0.398 | 0.6908<br>55 | No | 5:56757918:A<br>AAAC:AAA<br>ACAAAC | 169956 |
| 5q11.2       | 5:56927874-56929573   | CTD-2310F14.1 | Breast tissue     | Cross-ancestry    | 13.90<br>7 | 5.75E-44 | No  | 1.769  | 0.0768<br>93 | No | 5:56757918:A<br>AAAC:AAA<br>ACAAAC | 169956 |
| 5q11.2       | 5:57480016-57482809   | ACTBL2        | Endothelial cells | European-ancestry | 7.674      | 1.67E-14 | No  | 1.586  | 0.1128<br>34 | No | 5:56757918:A<br>AAAC:AAA<br>ACAAAC | 722098 |
| 5q14.1-q14.2 | 5:81972025-82276857   | ATG10         | Adipocytes        | European-ancestry | -<br>6.336 | 2.36E-10 | Yes | -0.313 | 0.7539<br>94 | No | 5:82196824:G<br>:A                 | 0      |
| 5q14.1-q14.2 | 5:81972025-82276857   | ATG10         | Breast tissue     | European-ancestry | -<br>6.263 | 3.77E-10 | Yes | -0.287 | 0.7744<br>34 | No | 5:82196824:G<br>:A                 | 0      |
| 5q14.1-q14.2 | 5:81972025-82276857   | ATG10         | Epithelial cells  | European-ancestry | -<br>6.306 | 2.86E-10 | Yes | -0.340 | 0.7341<br>61 | No | 5:82196824:G<br>:A                 | 0      |
| 5q14.1-q14.2 | 5:81972025-82276857   | ATG10         | Fibroblasts       | European-ancestry | -<br>6.368 | 1.92E-10 | Yes | -0.397 | 0.6916<br>48 | No | 5:82196824:G<br>:A                 | 0      |
| 5q14.1-q14.2 | 5:81972025-82276857   | ATG10         | Adipocytes        | Cross-ancestry    | -<br>6.267 | 3.68E-10 | Yes | -0.712 | 0.4767<br>21 | No | 5:82196824:G<br>:A                 | 0      |
| 5q14.1-q14.2 | 5:81972025-82276857   | ATG10         | Breast tissue     | Cross-ancestry    | -<br>5.856 | 4.73E-09 | Yes | -0.378 | 0.7050<br>99 | No | 5:82196824:G<br>:A                 | 0      |
| 5q14.1       | 5:81991995-81992481   | ATG10-IT1     | Breast tissue     | European-ancestry | -<br>5.715 | 1.10E-08 | No  | 0.932  | 0.3513<br>56 | No | 5:82196824:G<br>:A                 | 204343 |
| 5q14.2       | 5:82273358-82278577   | RPS23         | Adipocytes        | European-ancestry | 6.172      | 6.74E-10 | Yes | 0.012  | 0.9903<br>03 | No | 5:82196824:G<br>:A                 | 76534  |
| 5q14.2       | 5:82273358-82278577   | RPS23         | Breast tissue     | European-ancestry | 6.332      | 2.42E-10 | Yes | 0.085  | 0.9324<br>24 | No | 5:82196824:G<br>:A                 | 76534  |
| 5q14.2       | 5:82273358-82278577   | RPS23         | Endothelial cells | European-ancestry | 6.563      | 5.28E-11 | Yes | 0.007  | 0.9948<br>04 | No | 5:82196824:G<br>:A                 | 76534  |
| 5q14.2       | 5:82273358-82278577   | RPS23         | Breast tissue     | Cross-ancestry    | 6.765      | 1.34E-11 | Yes | 1.009  | 0.3129<br>51 | No | 5:82196824:G<br>:A                 | 76534  |
| 5q14.2       | 5:82273358-82278577   | RPS23         | Endothelial cells | Cross-ancestry    | 6.924      | 4.40E-12 | Yes | 0.876  | 0.3809<br>04 | No | 5:82196824:G<br>:A                 | 76534  |
| 5q14.2       | 5:82279462-82386977   | ATP6AP1L      | Breast tissue     | European-ancestry | -<br>5.768 | 8.01E-09 | Yes | 0.309  | 0.7570<br>85 | No | 5:82196824:G<br>:A                 | 82638  |
| 5q14.2       | 5:82279462-82386977   | ATP6AP1L      | Epithelial cells  | European-ancestry | -<br>5.812 | 6.19E-09 | Yes | 0.280  | 0.7792<br>59 | No | 5:82196824:G<br>:A                 | 82638  |
| 5q14.2       | 5:82279462-82386977   | ATP6AP1L      | Breast tissue     | Cross-ancestry    | -<br>5.390 | 7.05E-08 | Yes | 0.199  | 0.8423<br>82 | No | 5:82196824:G<br>:A                 | 82638  |
| 5q14.2       | 5:82279462-82386977   | ATP6AP1L      | Epithelial cells  | Cross-ancestry    | -<br>5.437 | 5.43E-08 | Yes | 0.165  | 0.8688<br>2  | No | 5:82196824:G<br>:A                 | 82638  |
| 6p23         | 6:13574529-13615158   | SIRT5         | Adipocytes        | European-ancestry | 5.525      | 3.29E-08 | No  | 3.395  | 0.0006<br>85 | No | 6:13713134:G<br>:C                 | 97976  |
| 6p23         | 6:13574529-13615158   | SIRT5         | Breast tissue     | European-ancestry | 5.558      | 2.73E-08 | No  | 3.141  | 0.0016<br>86 | No | 6:13713134:G<br>:C                 | 97976  |
| 6p23         | 6:13574529-13615158   | SIRT5         | Fibroblasts       | European-ancestry | 6.272      | 3.56E-10 | No  | 2.972  | 0.0029<br>54 | No | 6:13713134:G<br>:C                 | 97976  |
| 6p23         | 6:13574529-13615158   | SIRT5         | Fibroblasts       | Cross-ancestry    | 6.239      | 4.41E-10 | No  | 3.135  | 0.0017<br>17 | No | 6:13713134:G<br>:C                 | 97976  |
| 7q34         | 7:140177261-140179640 | JHDM1D-AS1    | Breast tissue     | European-ancestry | 6.771      | 1.28E-11 | No  | 1.893  | 0.0583<br>44 | No | 7:140242504:<br>G:A                | 62864  |
| 7q34         | 7:140177261-140179640 | JHDM1D-AS1    | Breast tissue     | Cross-ancestry    | 6.110      | 9.99E-10 | No  | 1.670  | 0.0949<br>91 | No | 7:140242504:<br>G:A                | 62864  |
| 8p11.23      | 8:37784191-37844896   | ADGRA2        | Adipocytes        | European-ancestry | -<br>7.693 | 1.43E-14 | No  | -0.243 | 0.8080<br>83 | No | 8:37000965:A<br>:G                 | 783226 |
| 8p11.23      | 8:37784191-37844896   | ADGRA2        | Adipocytes        | Cross-ancestry    | -<br>7.748 | 9.30E-15 | No  | -0.749 | 0.4540<br>61 | No | 8:37000965:A<br>:G                 | 783226 |
| 8q21.13      | 8:74824537-74855029   | PI15          | Adipocytes        | European-ancestry | -<br>5.575 | 2.48E-08 | No  | -4.025 | 5.70E-05     | No | 8:75420821:C<br>:T                 | 565792 |
| 8q21.13      | 8:74824537-74855029   | PI15          | Endothelial cells | European-ancestry | -<br>7.416 | 1.21E-13 | No  | -4.969 | 6.73E-07     | No | 8:75420821:C<br>:T                 | 565792 |

Continuation of Table S 1.9

|          |                        |                      |                   |                   |            |          |     |        |          |     |                     |        |
|----------|------------------------|----------------------|-------------------|-------------------|------------|----------|-----|--------|----------|-----|---------------------|--------|
| 8q21.13  | 8:74824537-74855029    | <i>PI15</i>          | Endothelial cells | Cross-ancestry    | -<br>7.067 | 1.58E-12 | No  | -4.663 | 3.12E-06 | No  | 8:75420821:C:T      | 565792 |
| 8q24.13  | 8:124474738-124487914  | <i>RNF139</i>        | Breast tissue     | European-ancestry | -<br>5.744 | 9.23E-09 | No  | -1.286 | 0.198576 | No  | 8:123597926:G:A     | 876812 |
| 8q24.21  | 8:127794533-128101253  | <i>PVT1</i>          | Fibroblasts       | European-ancestry | -<br>6.235 | 4.51E-10 | No  | -2.772 | 0.005566 | No  | 8:127359926:A:G     | 434607 |
| 9p21.3   | 9:21699314-21699596    | <i>RN7SL151P</i>     | Breast tissue     | European-ancestry | -<br>5.674 | 1.40E-08 | No  | -3.319 | 0.000902 | No  | 9:21964883:C:AAAA:C | 265287 |
| 9p21.3   | 9:21967753-21995301    | <i>CDKN2A</i>        | Adipocytes        | European-ancestry | 5.433      | 5.53E-08 | No  | 3.261  | 0.001111 | No  | 9:21966222:G:A      | 1531   |
| 9p21.3   | 9:21967753-21995301    | <i>CDKN2A</i>        | Breast tissue     | European-ancestry | 5.324      | 1.01E-07 | No  | 3.328  | 0.000873 | No  | 9:21966222:G:A      | 1531   |
| 9p21.3   | 9:21967753-21995301    | <i>CDKN2A</i>        | Adipocytes        | Cross-ancestry    | 5.494      | 3.93E-08 | No  | 3.448  | 0.000565 | No  | 9:21966222:G:A      | 1531   |
| 9q33.1   | 9:116687302-116701300  | <i>TRIM32</i>        | Breast tissue     | European-ancestry | 6.271      | 3.58E-10 | No  | 0.251  | 0.801435 | No  | 9:116552572:C:T     | 134730 |
| 9q33.1   | 9:116687302-116701300  | <i>TRIM32</i>        | Breast tissue     | Cross-ancestry    | 5.534      | 3.12E-08 | No  | -0.027 | 0.978628 | No  | 9:116552572:C:T     | 134730 |
| 10p15.1  | 10:5765223-5842132     | <i>GDI2</i>          | Breast tissue     | Cross-ancestry    | -<br>5.457 | 4.83E-08 | No  | -1.219 | 0.222753 | No  | 10:5707486:G:A      | 57737  |
| 10p15.1  | 10:5765223-5842132     | <i>GDI2</i>          | Epithelial cells  | Cross-ancestry    | -<br>5.401 | 6.61E-08 | No  | -1.204 | 0.228732 | No  | 10:5707486:G:A      | 57737  |
| 10p12.31 | 10:21513478-21526368   | <i>SKIDA1</i>        | Adipocytes        | European-ancestry | 10.629     | 2.19E-26 | No  | -1.607 | 0.108149 | No  | 10:21532989:A:T     | 6621   |
| 10p12.31 | 10:21513478-21526368   | <i>SKIDA1</i>        | Breast tissue     | European-ancestry | 9.578      | 9.86E-22 | No  | -1.585 | 0.112871 | No  | 10:21532989:A:T     | 6621   |
| 10p12.31 | 10:21513478-21526368   | <i>SKIDA1</i>        | Adipocytes        | Cross-ancestry    | 9.461      | 3.06E-21 | No  | -1.814 | 0.069626 | No  | 10:21532989:A:T     | 6621   |
| 10p12.31 | 10:21513478-21526368   | <i>SKIDA1</i>        | Breast tissue     | Cross-ancestry    | 8.777      | 1.68E-18 | No  | -1.431 | 0.152479 | No  | 10:21532989:A:T     | 6621   |
| 10q25.3  | 10:113854661-113917194 | <i>NHLRC2</i>        | Endothelial cells | European-ancestry | 6.281      | 3.36E-10 | No  | -1.398 | 0.162045 | No  | 10:113373449:C:A    | 481212 |
| 10q25.3  | 10:113854661-113917194 | <i>NHLRC2</i>        | Endothelial cells | Cross-ancestry    | 6.243      | 4.30E-10 | No  | -0.851 | 0.394674 | No  | 10:113373449:C:A    | 481212 |
| 10q26.13 | 10:121928312-121951965 | <i>ATE1-AS1</i>      | Breast tissue     | European-ancestry | 5.695      | 1.24E-08 | No  | -0.434 | 0.664564 | No  | 10:121575416:G:A    | 352896 |
| 10q26.13 | 10:121928312-121951965 | <i>ATE1-AS1</i>      | Fibroblasts       | European-ancestry | 7.727      | 1.11E-14 | No  | -0.201 | 0.841046 | No  | 10:121575416:G:A    | 352896 |
| 10q26.13 | 10:121928312-121951965 | <i>ATE1-AS1</i>      | Breast tissue     | Cross-ancestry    | 5.313      | 1.08E-07 | No  | -0.670 | 0.502855 | No  | 10:121580797:C:T    | 347515 |
| 10q26.13 | 10:121956782-121957098 | <i>RP11-500G22.4</i> | Epithelial cells  | European-ancestry | 5.731      | 9.99E-09 | Yes | -0.757 | 0.44918  | No  | 10:121575416:G:A    | 381366 |
| 10q26.13 | 10:121956782-121957098 | <i>RP11-500G22.4</i> | Fibroblasts       | European-ancestry | 5.521      | 3.36E-08 | Yes | -0.747 | 0.454766 | No  | 10:121575416:G:A    | 381366 |
| 10q26.13 | 10:122560665-122643736 | <i>DMBT1</i>         | Breast tissue     | European-ancestry | 11.012     | 3.34E-28 | No  | -1.054 | 0.291769 | No  | 10:121575416:G:A    | 985249 |
| 10q26.13 | 10:122560665-122643736 | <i>DMBT1</i>         | Endothelial cells | European-ancestry | 13.735     | 6.25E-43 | No  | -0.111 | 0.911978 | No  | 10:121575416:G:A    | 985249 |
| 10q26.13 | 10:122560665-122643736 | <i>DMBT1</i>         | Breast tissue     | Cross-ancestry    | 10.191     | 2.17E-24 | No  | -1.081 | 0.279498 | No  | 10:121580797:C:T    | 979868 |
| 10q26.13 | 10:122560665-122643736 | <i>DMBT1</i>         | Endothelial cells | Cross-ancestry    | 12.802     | 1.59E-37 | No  | -0.294 | 0.768523 | No  | 10:121580797:C:T    | 979868 |
| 11q13.2  | 11:66563463-66568841   | <i>CTSF</i>          | Adipocytes        | European-ancestry | 6.407      | 1.48E-10 | No  | 3.593  | 0.000327 | No  | 11:65753275:A:G     | 810188 |
| 12p12.3  | 12:14767999-14771131   | <i>HIST4H4</i>       | Breast tissue     | European-ancestry | -<br>5.918 | 3.26E-09 | No  | 0.000  | 1        | No  | 12:14260997:G:C     | 507002 |
| 12p12.3  | 12:14767999-14771131   | <i>HIST4H4</i>       | Breast tissue     | Cross-ancestry    | -<br>5.703 | 1.18E-08 | No  | -0.264 | 0.791439 | No  | 12:14260997:G:C     | 507002 |
| 12p11.22 | 12:27762738-27771276   | <i>MANSC4</i>        | Breast tissue     | European-ancestry | 5.657      | 1.54E-08 | No  | 2.258  | 0.023933 | No  | 12:28021884:C:T     | 250608 |
| 12p11.22 | 12:27958084-27972705   | <i>PTHLH</i>         | Epithelial cells  | European-ancestry | 5.361      | 8.30E-08 | No  | 2.277  | 0.022815 | No  | 12:28021884:C:T     | 49179  |
| 12p11.22 | 12:28133249-28581511   | <i>CCDC91</i>        | Adipocytes        | European-ancestry | -<br>5.307 | 1.12E-07 | No  | -3.195 | 0.001399 | No  | 12:28021884:C:T     | 111365 |
| 12p11.22 | 12:28163298-28190738   | <i>RP11-967K21.1</i> | Adipocytes        | European-ancestry | 7.710      | 1.25E-14 | No  | 6.996  | 2.63E-12 | Yes | 12:28021884:C:T     | 141414 |
| 12p11.22 | 12:28163298-28190738   | <i>RP11-967K21.1</i> | Breast tissue     | European-ancestry | 8.200      | 2.41E-16 | No  | 7.233  | 4.74E-13 | Yes | 12:28021884:C:T     | 141414 |
| 12p11.22 | 12:28163298-28190738   | <i>RP11-967K21.1</i> | Epithelial cells  | European-ancestry | 7.948      | 1.90E-15 | No  | 7.399  | 1.37E-13 | Yes | 12:28021884:C:T     | 141414 |
| 12p11.22 | 12:28163298-28190738   | <i>RP11-967K21.1</i> | Fibroblasts       | European-ancestry | 7.523      | 5.35E-14 | No  | 6.946  | 3.76E-12 | Yes | 12:28021884:C:T     | 141414 |
| 12p11.22 | 12:28163298-28190738   | <i>RP11-967K21.1</i> | Adipocytes        | Cross-ancestry    | 7.426      | 1.12E-13 | No  | 6.778  | 1.22E-11 | Yes | 12:28021884:C:T     | 141414 |
| 12p11.22 | 12:28163298-28190738   | <i>RP11-967K21.1</i> | Fibroblasts       | Cross-ancestry    | 7.002      | 2.53E-12 | No  | 6.473  | 9.60E-11 | Yes | 12:28021884:C:T     | 141414 |
| 12p11.22 | 12:28185625-28186190   | <i>RP11-425D17.1</i> | Breast tissue     | European-ancestry | 7.769      | 7.93E-15 | No  | 7.265  | 3.73E-13 | Yes | 12:28021884:C:T     | 163741 |
| 12p11.22 | 12:28185625-28186190   | <i>RP11-425D17.1</i> | Endothelial cells | European-ancestry | 7.319      | 2.50E-13 | No  | 6.822  | 8.98E-12 | Yes | 12:28021884:C:T     | 163741 |
| 12q22    | 12:95657807-95791152   | <i>NTN4</i>          | Breast tissue     | European-ancestry | -<br>6.498 | 8.15E-11 | No  | 0.407  | 0.684222 | No  | 12:95633983:A:G     | 23824  |

Continuation of Table S 1.9

|                 |                      |                      |                   |                   |       |          |     |        |          |     |                  |        |
|-----------------|----------------------|----------------------|-------------------|-------------------|-------|----------|-----|--------|----------|-----|------------------|--------|
| 12q22           | 12:95657807-95791152 | <i>NTN4</i>          | Epithelial cells  | European-ancestry | -     | 1.91E-23 | No  | -0.662 | 0.507925 | No  | 12:95633983: A:G | 23824  |
| 12q22           | 12:95657807-95791152 | <i>NTN4</i>          | Fibroblasts       | European-ancestry | -     | 1.79E-08 | No  | 1.383  | 0.166666 | No  | 12:95633983: A:G | 23824  |
| 12q22           | 12:95657807-95791152 | <i>NTN4</i>          | Breast tissue     | Cross-ancestry    | -     | 1.32E-11 | No  | 0.038  | 0.969457 | No  | 12:95633983: A:G | 23824  |
| 12q22           | 12:95657807-95791152 | <i>NTN4</i>          | Epithelial cells  | Cross-ancestry    | -     | 5.97E-19 | No  | -0.671 | 0.501974 | No  | 12:95633983: A:G | 23824  |
| 12q22           | 12:95657807-95791152 | <i>NTN4</i>          | Fibroblasts       | Cross-ancestry    | -     | 2.50E-08 | No  | 1.283  | 0.199646 | No  | 12:95633983: A:G | 23824  |
| 14q24.1         | 14:67819779-68730218 | <i>RAD51B</i>        | Breast tissue     | European-ancestry | 5.780 | 7.46E-09 | No  | 1.977  | 0.048085 | No  | 14:68513118: T:C | 0      |
| 14q24.1         | 14:67819779-68730218 | <i>RAD51B</i>        | Breast tissue     | Cross-ancestry    | 5.514 | 3.51E-08 | No  | 2.069  | 0.038559 | No  | 14:68513118: T:C | 0      |
| 14q24.1         | 14:68874143-68979440 | <i>ACTN1</i>         | Breast tissue     | European-ancestry | 5.798 | 6.71E-09 | Yes | 5.119  | 3.06E-07 | No  | 14:68513118: T:C | 361025 |
| 14q24.1         | 14:68874143-68979440 | <i>ACTN1</i>         | Breast tissue     | Cross-ancestry    | 5.371 | 7.83E-08 | Yes | 4.749  | 2.04E-06 | No  | 14:68513118: T:C | 361025 |
| 14q32.11-q32.12 | 14:91271323-91417844 | <i>CCDC88C</i>       | Endothelial cells | European-ancestry | 5.518 | 3.43E-08 | No  | 2.160  | 0.030758 | No  | 14:91382854: G:A | 0      |
| 15q24.1         | 15:74836116-74843346 | <i>ULK3</i>          | Adipocytes        | European-ancestry | -     | 4.94E-09 | No  | -4.760 | 1.94E-06 | No  | 15:75640436: C:T | 797090 |
| 15q24.1         | 15:74836116-74843346 | <i>ULK3</i>          | Breast tissue     | European-ancestry | -     | 5.49E-09 | No  | -4.675 | 2.94E-06 | No  | 15:75640436: C:T | 797090 |
| 15q24.2         | 15:75355207-75368630 | <i>MAN2C1</i>        | Breast tissue     | European-ancestry | -     | 1.48E-08 | No  | -0.715 | 0.474507 | No  | 15:75640436: C:T | 271806 |
| 15q24.2         | 15:75355207-75368630 | <i>MAN2C1</i>        | Breast tissue     | Cross-ancestry    | -     | 3.83E-11 | No  | -1.690 | 0.091003 | No  | 15:75640436: C:T | 271806 |
| 15q26.1         | 15:90954870-90963125 | <i>RCCD1</i>         | Breast tissue     | European-ancestry | -     | 1.22E-10 | No  | 0.214  | 0.830524 | No  | 15:90985645: G:A | 22520  |
| 15q26.1         | 15:90954870-90963125 | <i>RCCD1</i>         | Breast tissue     | Cross-ancestry    | -     | 9.75E-11 | No  | -0.162 | 0.871596 | No  | 15:90985645: G:A | 22520  |
| 15q26.1         | 15:90966345-90988624 | <i>PRC1-AS1</i>      | Adipocytes        | Cross-ancestry    | 5.369 | 7.93E-08 | No  | 0.735  | 0.462268 | No  | 15:90985645: G:A | 0      |
| 16q12.1         | 16:52438005-52547802 | <i>TOX3</i>          | Breast tissue     | European-ancestry | -     | 1.83E-53 | No  | -0.465 | 0.641714 | No  | 16:52565276: C:T | 17474  |
| 16q12.1         | 16:52438005-52547802 | <i>TOX3</i>          | Fibroblasts       | European-ancestry | -     | 4.82E-66 | No  | -1.781 | 0.074838 | No  | 16:52565276: C:T | 17474  |
| 16q12.1         | 16:52438005-52547802 | <i>TOX3</i>          | Breast tissue     | Cross-ancestry    | -     | 6.86E-42 | No  | -0.004 | 0.997159 | No  | 16:52565276: C:T | 17474  |
| 16q12.1         | 16:52438005-52547802 | <i>TOX3</i>          | Fibroblasts       | Cross-ancestry    | -     | 3.63E-50 | No  | -1.083 | 0.278772 | No  | 16:52565276: C:T | 17474  |
| 16q12.2         | 16:53433977-53491649 | <i>RBL2</i>          | Endothelial cells | European-ancestry | -     | 2.01E-11 | No  | -3.356 | 0.000792 | No  | 16:52565276: C:T | 868701 |
| 16q12.2         | 16:53433977-53491649 | <i>RBL2</i>          | Endothelial cells | Cross-ancestry    | -     | 6.12E-10 | No  | -3.198 | 0.001384 | No  | 16:52565276: C:T | 868701 |
| 16q23.2         | 16:81035847-81047358 | <i>ATMIN</i>         | Fibroblasts       | European-ancestry | -     | 6.33E-13 | No  | -1.744 | 0.081237 | No  | 16:80616908: A:G | 418939 |
| 16q23.2         | 16:81035847-81047358 | <i>ATMIN</i>         | Fibroblasts       | Cross-ancestry    | -     | 5.90E-12 | No  | -1.959 | 0.0501   | No  | 16:80616908: A:G | 418939 |
| 16q23.2         | 16:81314952-81390884 | <i>GAN</i>           | Breast tissue     | European-ancestry | 8.752 | 2.09E-18 | No  | 1.738  | 0.082139 | No  | 16:80616908: A:G | 698044 |
| 16q23.2         | 16:81314952-81390884 | <i>GAN</i>           | Breast tissue     | Cross-ancestry    | 8.548 | 1.25E-17 | No  | 0.855  | 0.392568 | No  | 16:80616908: A:G | 698044 |
| 17q22           | 17:54899387-54961956 | <i>TOM1L1</i>        | Breast tissue     | European-ancestry | 6.524 | 6.83E-11 | No  | 1.164  | 0.244587 | No  | 17:55097605: A:G | 135649 |
| 17q22           | 17:54951902-54968785 | <i>COX11</i>         | Breast tissue     | European-ancestry | 5.395 | 6.84E-08 | No  | 0.759  | 0.448071 | No  | 17:55097605: A:G | 128820 |
| 17q22           | 17:54951902-54968785 | <i>COX11</i>         | Fibroblasts       | European-ancestry | 5.320 | 1.04E-07 | No  | 0.771  | 0.440848 | No  | 17:55097605: A:G | 128820 |
| 17q22           | 17:54951902-54968785 | <i>COX11</i>         | Breast tissue     | Cross-ancestry    | 5.915 | 3.33E-09 | No  | 1.635  | 0.102052 | No  | 17:55097605: A:G | 128820 |
| 18q11.2         | 18:26348347-26349004 | <i>RP11-107K17.2</i> | Breast tissue     | European-ancestry | -     | 1.26E-08 | No  | -5.263 | 1.42E-07 | Yes | 18:26920935: C:A | 571931 |
| 18q11.2         | 18:26348347-26349004 | <i>RP11-107K17.2</i> | Breast tissue     | Cross-ancestry    | -     | 2.35E-08 | No  | -5.052 | 4.37E-07 | No  | 18:26920935: C:A | 571931 |
| 19p13.13        | 19:12791496-12793315 | <i>JUNB</i>          | Breast tissue     | European-ancestry | 5.742 | 9.36E-09 | Yes | 2.818  | 0.004837 | No  | 19:13047463: T:C | 254148 |
| 19p13.11        | 19:17794828-17813082 | <i>B3GNT3</i>        | Adipocytes        | European-ancestry | 6.395 | 1.60E-10 | No  | 1.885  | 0.059442 | No  | 19:18453895: A:G | 640813 |
| 19p13.11        | 19:17794828-17813082 | <i>B3GNT3</i>        | Adipocytes        | Cross-ancestry    | 5.568 | 2.58E-08 | No  | 1.345  | 0.178683 | No  | 19:18453895: A:G | 640813 |
| 19p13.11        | 19:18391144-18397617 | <i>LRRC25</i>        | Epithelial cells  | European-ancestry | 6.851 | 7.34E-12 | No  | 0.461  | 0.644623 | No  | 19:18453895: A:G | 56278  |
| 19p13.11        | 19:18391144-18397617 | <i>LRRC25</i>        | Fibroblasts       | European-ancestry | 7.215 | 5.41E-13 | No  | 4.077  | 4.57E-05 | No  | 19:18453895: A:G | 56278  |
| 19p13.11        | 19:18391144-18397617 | <i>LRRC25</i>        | Endothelial cells | Cross-ancestry    | 5.584 | 2.36E-08 | No  | 1.019  | 0.308287 | No  | 19:18453895: A:G | 56278  |
| 19p13.11        | 19:18391144-18397617 | <i>LRRC25</i>        | Epithelial cells  | Cross-ancestry    | 5.808 | 6.31E-09 | No  | -0.030 | 0.976436 | No  | 19:18453895: A:G | 56278  |
| 19p13.11        | 19:18391144-18397617 | <i>LRRC25</i>        | Fibroblasts       | Cross-ancestry    | 7.314 | 2.58E-13 | No  | 4.378  | 1.20E-05 | No  | 19:18453895: A:G | 56278  |

Continuation of Table S 1.9

|                    |                       |               |                   |                   |       |          |     |        |          |     |                 |        |
|--------------------|-----------------------|---------------|-------------------|-------------------|-------|----------|-----|--------|----------|-----|-----------------|--------|
| 19q13.31           | 19:43766533-43781257  | KCNN4         | Breast tissue     | European-ancestry | 5.558 | 2.72E-08 | No  | 1.304  | 0.192256 | No  | 19:43779295:C:T | 0      |
| 19q13.31           | 19:43766533-43781257  | KCNN4         | Breast tissue     | Cross-ancestry    | 5.479 | 4.28E-08 | No  | 1.465  | 0.143043 | No  | 19:43778879:T:C | 0      |
| 19q13.31           | 19:43872363-43901385  | ZNF404        | Breast tissue     | European-ancestry | 7.245 | 4.33E-13 | No  | 1.778  | 0.075409 | No  | 19:43779295:C:T | 93068  |
| 19q13.31           | 19:43872363-43901385  | ZNF404        | Breast tissue     | Cross-ancestry    | 7.072 | 1.52E-12 | No  | 1.487  | 0.137079 | No  | 19:43779295:C:T | 93068  |
| 19q13.31           | 19:43891804-43901805  | RP11-15A1.3   | Adipocytes        | European-ancestry | 5.506 | 3.67E-08 | No  | 1.606  | 0.108199 | No  | 19:43779295:C:T | 112509 |
| 19q13.31           | 19:43891804-43901805  | RP11-15A1.3   | Breast tissue     | European-ancestry | 6.783 | 1.17E-11 | No  | 2.656  | 0.0079   | No  | 19:43779295:C:T | 112509 |
| 19q13.31           | 19:43891804-43901805  | RP11-15A1.3   | Epithelial cells  | European-ancestry | 6.678 | 2.41E-11 | No  | 2.791  | 0.005257 | No  | 19:43779295:C:T | 112509 |
| 19q13.31           | 19:43891804-43901805  | RP11-15A1.3   | Fibroblasts       | European-ancestry | 6.927 | 4.30E-12 | No  | 3.316  | 0.000912 | No  | 19:43779295:C:T | 112509 |
| 19q13.31           | 19:43891804-43901805  | RP11-15A1.3   | Breast tissue     | Cross-ancestry    | 5.804 | 6.48E-09 | No  | 1.828  | 0.067486 | No  | 19:43779295:C:T | 112509 |
| 19q13.31           | 19:43891804-43901805  | RP11-15A1.3   | Fibroblasts       | Cross-ancestry    | 6.070 | 1.28E-09 | No  | 2.618  | 0.008856 | No  | 19:43779295:C:T | 112509 |
| 19q13.31           | 19:43902001-43926545  | RP11-15A1.2   | Breast tissue     | European-ancestry | 5.344 | 9.10E-08 | No  | 2.224  | 0.026134 | No  | 19:43779295:C:T | 122706 |
| 21q11.2            | 21:14819699-14918552  | AF127936.7    | Breast tissue     | European-ancestry | -     | 2.69E-11 | No  | -3.126 | 0.001772 | No  | 21:15202135:C:A | 283583 |
| 21q11.2            | 21:14819699-14918552  | AF127936.7    | Breast tissue     | Cross-ancestry    | -     | 1.49E-09 | No  | -2.838 | 0.004541 | No  | 21:15202135:C:A | 283583 |
| 22q12.1            | 22:27748277-27801498  | MN1           | Breast tissue     | European-ancestry | 7.312 | 2.62E-13 | Yes | 0.000  | 1        | No  | 22:28682717:T:G | 881219 |
| 22q12.1            | 22:27748277-27801498  | MN1           | Breast tissue     | Cross-ancestry    | 6.662 | 2.70E-11 | Yes | -0.080 | 0.935914 | No  | 22:28682717:T:G | 881219 |
| 22q12.2            | 22:29205851-29259597  | EMID1         | Fibroblasts       | European-ancestry | -     | 1.41E-09 | No  | -2.313 | 0.02072  | No  | 22:28712568:A:T | 493283 |
| 22q13.1            | 22:38982347-38992804  | APOBEC3B      | Adipocytes        | European-ancestry | -     | 7.48E-14 | Yes | -6.389 | 1.67E-10 | Yes | 22:38202034:C:T | 780313 |
| 22q13.1            | 22:38982347-38992804  | APOBEC3B      | Breast tissue     | European-ancestry | -     | 4.02E-12 | Yes | -6.002 | 1.95E-09 | Yes | 22:38202034:C:T | 780313 |
| 22q13.1            | 22:38982347-38992804  | APOBEC3B      | Endothelial cells | European-ancestry | -     | 1.06E-13 | Yes | -6.359 | 2.02E-10 | Yes | 22:38202034:C:T | 780313 |
| 22q13.1            | 22:38982347-38992804  | APOBEC3B      | Adipocytes        | Cross-ancestry    | -     | 3.04E-11 | Yes | -5.664 | 1.48E-08 | Yes | 22:38202034:C:T | 780313 |
| 22q13.1            | 22:38982347-38992804  | APOBEC3B      | Endothelial cells | Cross-ancestry    | -     | 2.99E-12 | Yes | -5.887 | 3.94E-09 | Yes | 22:38202034:C:T | 780313 |
| 22q13.2            | 22:41229510-41240934  | CHADL         | Breast tissue     | European-ancestry | -     | 6.86E-11 | No  | -2.374 | 0.017597 | No  | 22:40653398:G:T | 576112 |
| 22q13.2            | 22:41229510-41240934  | CHADL         | Endothelial cells | European-ancestry | -     | 8.93E-09 | No  | -2.388 | 0.016943 | No  | 22:40653398:G:T | 576112 |
| 22q13.2            | 22:41229510-41240934  | CHADL         | Breast tissue     | Cross-ancestry    | -     | 4.68E-08 | No  | -1.805 | 0.07102  | No  | 22:40653398:G:T | 576112 |
| Luminal-B-like     |                       |               |                   |                   |       |          |     |        |          |     |                 |        |
| 5q11.2             | 5:56927874-56929573   | CTD-2310F14.1 | Breast tissue     | European-ancestry | 5.609 | 2.03E-08 | No  | 0.034  | 0.972685 | No  | 5:56757708:T:C  | 170166 |
| 15q26.1            | 15:90966345-90988624  | PRC1-AS1      | Breast tissue     | European-ancestry | 5.307 | 1.12E-07 | No  | 0.742  | 0.457977 | No  | 15:90965985:C:A | 360    |
| 16q12.1            | 16:52438005-52547802  | TOX3          | Breast tissue     | European-ancestry | -     | 3.23E-11 | No  | -0.400 | 0.689026 | No  | 16:52538117:T:C | 0      |
| 16q12.1            | 16:52438005-52547802  | TOX3          | Fibroblasts       | European-ancestry | -     | 1.54E-15 | No  | -0.691 | 0.489771 | No  | 16:52538117:T:C | 0      |
| 16q12.1            | 16:52438005-52547802  | TOX3          | Breast tissue     | Cross-ancestry    | -     | 2.64E-08 | No  | -0.014 | 0.988616 | No  | 16:52538117:T:C | 0      |
| 16q12.1            | 16:52438005-52547802  | TOX3          | Fibroblasts       | Cross-ancestry    | -     | 8.39E-14 | No  | -0.954 | 0.340226 | No  | 16:52538117:T:C | 0      |
| HER2-enriched-like |                       |               |                   |                   |       |          |     |        |          |     |                 |        |
| 1q32.1             | 1:202193901-202319781 | LGR6          | Endothelial cells | European-ancestry | -     | 3.50E-09 | No  | 0.021  | 0.983058 | No  | 1:202218991:T:C | 0      |
| 1q32.1             | 1:202193901-202319781 | LGR6          | Fibroblasts       | European-ancestry | -     | 6.47E-08 | No  | 0.162  | 0.871251 | No  | 1:202218991:T:C | 0      |
| Triple-negative    |                       |               |                   |                   |       |          |     |        |          |     |                 |        |
| 2q14.2             | 2:118842171-118847678 | EN1           | Breast tissue     | African-ancestry  | -     | 9.95E-09 | No  | 0.000  | 1        | No  | 2:118823485:A:G | 18686  |
| 12p11.22           | 12:28163298-28190738  | RP11-967K21.1 | Breast tissue     | European-ancestry | 5.566 | 2.61E-08 | No  | 4.990  | 6.03E-07 | No  | 12:28021884:C:T | 141414 |
| 12p11.22           | 12:28163298-28190738  | RP11-967K21.1 | Epithelial cells  | European-ancestry | 5.572 | 2.53E-08 | No  | 5.467  | 4.57E-08 | Yes | 12:28021884:C:T | 141414 |
| 12p11.22           | 12:28185625-28186190  | RP11-425D17.1 | Breast tissue     | European-ancestry | 5.489 | 4.04E-08 | No  | 5.187  | 2.14E-07 | No  | 12:28021884:C:T | 163741 |
| 12p11.22           | 12:28185625-28186190  | RP11-425D17.1 | Endothelial cells | European-ancestry | 5.438 | 5.38E-08 | No  | 5.139  | 2.77E-07 | No  | 12:28021884:C:T | 163741 |
| 19p13.11           | 19:17292131-17310236  | ABHD8         | Breast tissue     | European-ancestry | 8.263 | 1.42E-16 | Yes | 0.667  | 0.504495 | No  | 19:17291050:G:A | 1081   |

Continuation of Table S 1.9

|          |                      |              |               |                   |        |          |     |        |          |    |                 |        |
|----------|----------------------|--------------|---------------|-------------------|--------|----------|-----|--------|----------|----|-----------------|--------|
| 19p13.11 | 19:17292131-17310236 | <i>ABHD8</i> | Breast tissue | Cross-ancestry    | 8.133  | 4.20E-16 | Yes | 1.548  | 0.121632 | No | 19:17291050:G:A | 1081   |
| 19p13.11 | 19:17351448-17377350 | <i>PLVAP</i> | Fibroblasts   | European-ancestry | -6.456 | 1.07E-10 | Yes | -0.402 | 0.688041 | No | 19:17291050:G:A | 60398  |
| 19p13.11 | 19:17351448-17377350 | <i>PLVAP</i> | Fibroblasts   | Cross-ancestry    | -6.018 | 1.77E-09 | Yes | -0.778 | 0.436334 | No | 19:17291050:G:A | 60398  |
| 19p13.11 | 19:17951293-18000080 | <i>KCNN1</i> | Fibroblasts   | European-ancestry | -5.329 | 9.86E-08 | No  | 2.446  | 0.014451 | No | 19:17291050:G:A | 660243 |

**Footnote:**

Bonferroni threshold of significance utilized in both the primary TWAS and conditional (Cond-TWAS) corresponded to a p-value of  $1.42 \times 10^{-7}$  for European-ancestry TWAS,  $1.39 \times 10^{-7}$  for African-ancestry TWAS, and  $1.26 \times 10^{-7}$  for cross-ancestry TWAS

<sup>a</sup> Subtype specificity is defined as the minimum p-value of TWAS associations across all ancestries and cell types is above 0.05 for all other subtypes



Table S 1.10 Number of unique TWAS genes identified by subtype and cell type.

| Subtype            | Bulk tissue | Cell types       |            |                   |             |
|--------------------|-------------|------------------|------------|-------------------|-------------|
|                    |             | Epithelial cells | Adipocytes | Endothelial cells | Fibroblasts |
| Luminal-A-like     | 61          | 17               | 25         | 21                | 22          |
| Luminal-B-like     | 3           | 0                | 0          | 0                 | 1           |
| HER2-enriched-like | 0           | 0                | 0          | 1                 | 1           |
| Triple-negative    | 4           | 1                | 0          | 1                 | 2           |
| All subtypes       | 64          | 17               | 25         | 22                | 25          |

Table S 1.11 Ancestry-stratified TWAS associations for the 90 genes identified to associate with intrinsic subtypes (Bonferroni-adjusted p-value<0.05)

| Gene symbol    | Cell type         | Z-score           |                  | P-value           |                  |
|----------------|-------------------|-------------------|------------------|-------------------|------------------|
|                |                   | European American | African American | European American | African American |
| Luminal-A-like |                   |                   |                  |                   |                  |
| AC098820.3     | Bulk tissue       | 8.79              |                  | 1.46E-18          |                  |
| AC098820.3     | Adipocytes        | 6.12              |                  | 9.44E-10          |                  |
| AC098820.3     | Endothelial cells |                   | 0.46             |                   | 6.45E-01         |
| AC098820.3     | Epithelial cells  | 4.29              |                  | 1.82E-05          |                  |
| AC098820.3     | Fibroblasts       |                   | -0.03            |                   | 9.75E-01         |
| ACTBL2         | Bulk tissue       |                   | 1.65             |                   | 9.80E-02         |
| ACTBL2         | Endothelial cells | 7.67              |                  | 1.67E-14          |                  |
| ACTN1          | Bulk tissue       | 5.80              | 0.08             | 6.71E-09          | 9.35E-01         |
| ACTN1          | Adipocytes        | 3.92              |                  | 8.79E-05          |                  |
| ACTN1          | Endothelial cells | 3.26              |                  | 1.11E-03          |                  |
| ACTN1          | Fibroblasts       | 3.96              |                  | 7.44E-05          |                  |
| ADGRA2         | Bulk tissue       |                   | 0.29             |                   | 7.68E-01         |
| ADGRA2         | Adipocytes        | -7.69             | -1.70            | 1.43E-14          | 8.87E-02         |
| ADGRA2         | Endothelial cells | -0.06             |                  | 9.50E-01          |                  |
| ADGRA2         | Epithelial cells  |                   | 1.22             |                   | 2.21E-01         |
| ADGRA2         | Fibroblasts       | 1.14              | 1.25             | 2.55E-01          | 2.13E-01         |
| AF127936.7     | Bulk tissue       | -6.66             | 0.23             | 2.69E-11          | 8.18E-01         |
| AF127936.7     | Adipocytes        | -2.00             | -1.06            | 4.56E-02          | 2.91E-01         |
| AF127936.7     | Endothelial cells |                   | 0.64             |                   | 5.23E-01         |
| AF127936.7     | Epithelial cells  | -0.83             | 0.96             | 4.05E-01          | 3.35E-01         |
| AF127936.7     | Fibroblasts       | -3.00             |                  | 2.74E-03          |                  |
| ALS2CR12       | Bulk tissue       | 5.10              | -0.44            | 3.33E-07          | 6.63E-01         |
| ALS2CR12       | Adipocytes        |                   | 0.89             |                   | 3.76E-01         |
| ALS2CR12       | Endothelial cells |                   | 0.99             |                   | 3.22E-01         |
| ALS2CR12       | Epithelial cells  | 5.04              | 0.33             | 4.57E-07          | 7.42E-01         |
| ALS2CR12       | Fibroblasts       | 5.41              |                  | 6.23E-08          |                  |
| APOBEC3B       | Bulk tissue       | -6.94             |                  | 4.02E-12          |                  |
| APOBEC3B       | Adipocytes        | -7.48             | 0.62             | 7.48E-14          | 5.33E-01         |
| APOBEC3B       | Endothelial cells | -7.43             | -0.34            | 1.06E-13          | 7.32E-01         |

Continuation of Table S 1.11

|                  |                   |       |       |          |          |
|------------------|-------------------|-------|-------|----------|----------|
| <i>ATE1-AS1</i>  | Bulk tissue       | 5.69  | 0.18  | 1.24E-08 | 8.61E-01 |
| <i>ATE1-AS1</i>  | Endothelial cells |       | 0.93  |          | 3.52E-01 |
| <i>ATE1-AS1</i>  | Epithelial cells  |       | 0.31  |          | 7.55E-01 |
| <i>ATE1-AS1</i>  | Fibroblasts       | 7.73  |       | 1.11E-14 |          |
| <i>ATG10</i>     | Bulk tissue       | -6.26 | -0.23 | 3.77E-10 | 8.20E-01 |
| <i>ATG10</i>     | Adipocytes        | -6.34 | -1.11 | 2.36E-10 | 2.68E-01 |
| <i>ATG10</i>     | Endothelial cells |       | -0.74 |          | 4.62E-01 |
| <i>ATG10</i>     | Epithelial cells  | -6.31 |       | 2.86E-10 |          |
| <i>ATG10</i>     | Fibroblasts       | -6.37 |       | 1.92E-10 |          |
| <i>ATG10-IT1</i> | Bulk tissue       | -5.72 | 0.91  | 1.10E-08 | 3.65E-01 |
| <i>ATMIN</i>     | Bulk tissue       | -0.52 | 0.21  | 6.00E-01 | 8.34E-01 |
| <i>ATMIN</i>     | Adipocytes        | -0.96 | 0.33  | 3.35E-01 | 7.38E-01 |
| <i>ATMIN</i>     | Endothelial cells | -0.34 | 0.15  | 7.33E-01 | 8.84E-01 |
| <i>ATMIN</i>     | Epithelial cells  |       | -1.03 |          | 3.04E-01 |
| <i>ATMIN</i>     | Fibroblasts       | -7.19 | -0.66 | 6.33E-13 | 5.08E-01 |
| <i>ATP6AP1L</i>  | Bulk tissue       | -5.77 | -0.20 | 8.01E-09 | 8.41E-01 |
| <i>ATP6AP1L</i>  | Adipocytes        |       | -0.41 |          | 6.84E-01 |
| <i>ATP6AP1L</i>  | Epithelial cells  | -5.81 | -0.22 | 6.19E-09 | 8.28E-01 |
| <i>B3GNT3</i>    | Bulk tissue       |       | 0.10  |          | 9.22E-01 |
| <i>B3GNT3</i>    | Adipocytes        | 6.40  | -0.82 | 1.60E-10 | 4.10E-01 |
| <i>B3GNT3</i>    | Endothelial cells | 4.54  | -1.02 | 5.51E-06 | 3.07E-01 |
| <i>B3GNT3</i>    | Fibroblasts       |       | -0.09 |          | 9.28E-01 |
| <i>CASP8</i>     | Bulk tissue       | -5.67 | -2.81 | 1.43E-08 | 4.98E-03 |
| <i>CASP8</i>     | Endothelial cells | -6.14 |       | 8.23E-10 |          |
| <i>CASP8</i>     | Fibroblasts       | -5.64 | -2.16 | 1.66E-08 | 3.11E-02 |
| <i>CCDC88C</i>   | Bulk tissue       | 4.49  | 0.06  | 7.04E-06 | 9.50E-01 |
| <i>CCDC88C</i>   | Adipocytes        | 4.74  | 0.27  | 2.17E-06 | 7.86E-01 |
| <i>CCDC88C</i>   | Endothelial cells | 5.52  |       | 3.43E-08 |          |
| <i>CCDC88C</i>   | Epithelial cells  | 3.05  | -0.44 | 2.28E-03 | 6.63E-01 |
| <i>CCDC88C</i>   | Fibroblasts       |       | 0.74  |          | 4.62E-01 |
| <i>CCDC91</i>    | Bulk tissue       | -4.38 | 0.41  | 1.16E-05 | 6.84E-01 |
| <i>CCDC91</i>    | Adipocytes        | -5.31 |       | 1.12E-07 |          |
| <i>CCDC91</i>    | Endothelial cells |       | -0.71 |          | 4.80E-01 |

Continuation of Table S 1.11

|                      |                   |        |       |          |          |
|----------------------|-------------------|--------|-------|----------|----------|
| <i>CCDC91</i>        | Epithelial cells  | -3.20  | 0.21  | 1.37E-03 | 8.37E-01 |
| <i>CCDC91</i>        | Fibroblasts       | -3.61  | -0.68 | 3.01E-04 | 4.98E-01 |
| <i>CDKN2A</i>        | Bulk tissue       | 5.32   | -0.54 | 1.01E-07 | 5.90E-01 |
| <i>CDKN2A</i>        | Adipocytes        | 5.43   | 1.26  | 5.53E-08 | 2.09E-01 |
| <i>CDKN2A</i>        | Endothelial cells | 1.95   |       | 5.16E-02 |          |
| <i>CH17-437K3.1</i>  | Bulk tissue       | -10.10 | -2.10 | 5.42E-24 | 3.56E-02 |
| <i>CH17-437K3.1</i>  | Epithelial cells  |        | -2.10 |          | 3.56E-02 |
| <i>CH17-472G23.2</i> | Bulk tissue       | 7.49   |       | 6.97E-14 |          |
| <i>CH17-472G23.2</i> | Adipocytes        |        | 1.30  |          | 1.93E-01 |
| <i>CH17-472G23.2</i> | Endothelial cells |        | 0.80  |          | 4.22E-01 |
| <i>CH17-472G23.2</i> | Epithelial cells  | 7.79   |       | 6.63E-15 |          |
| <i>CH17-472G23.2</i> | Fibroblasts       | 6.97   |       | 3.11E-12 |          |
| <i>CHADL</i>         | Bulk tissue       | -6.52  | 1.40  | 6.86E-11 | 1.62E-01 |
| <i>CHADL</i>         | Adipocytes        |        | 0.45  |          | 6.53E-01 |
| <i>CHADL</i>         | Endothelial cells | -5.75  |       | 8.93E-09 |          |
| <i>CHADL</i>         | Epithelial cells  | -4.18  |       | 2.86E-05 |          |
| <i>CHADL</i>         | Fibroblasts       | -3.96  |       | 7.54E-05 |          |
| <i>COX11</i>         | Bulk tissue       | 5.40   | 2.43  | 6.84E-08 | 1.52E-02 |
| <i>COX11</i>         | Adipocytes        |        | 2.61  |          | 9.04E-03 |
| <i>COX11</i>         | Fibroblasts       | 5.32   |       | 1.04E-07 |          |
| <i>CTD-2310F14.1</i> | Bulk tissue       | 15.29  | -0.44 | 9.22E-53 | 6.60E-01 |
| <i>CTD-2310F14.1</i> | Adipocytes        | -8.94  | 1.04  | 4.03E-19 | 3.00E-01 |
| <i>CTD-2310F14.1</i> | Endothelial cells | 8.07   |       | 7.14E-16 |          |
| <i>CTSF</i>          | Bulk tissue       | 0.53   | -0.37 | 5.96E-01 | 7.13E-01 |
| <i>CTSF</i>          | Adipocytes        | 6.41   |       | 1.48E-10 |          |
| <i>CTSF</i>          | Epithelial cells  | -1.78  | 1.63  | 7.46E-02 | 1.03E-01 |
| <i>CTSF</i>          | Fibroblasts       | -1.88  | -1.01 | 5.98E-02 | 3.13E-01 |
| <i>CYBRD1</i>        | Bulk tissue       | -4.07  | -0.90 | 4.67E-05 | 3.70E-01 |
| <i>CYBRD1</i>        | Adipocytes        | -4.20  | -2.04 | 2.70E-05 | 4.10E-02 |
| <i>CYBRD1</i>        | Endothelial cells | -5.12  | -2.20 | 3.12E-07 | 2.76E-02 |
| <i>CYBRD1</i>        | Epithelial cells  | -4.28  | -0.94 | 1.90E-05 | 3.50E-01 |
| <i>CYBRD1</i>        | Fibroblasts       | -4.83  | -0.92 | 1.38E-06 | 3.56E-01 |

Continuation of Table S 1.11

|                   |                   |       |       |          |          |
|-------------------|-------------------|-------|-------|----------|----------|
| <i>DMBT1</i>      | Bulk tissue       | 11.01 | 0.13  | 3.34E-28 | 8.98E-01 |
| <i>DMBT1</i>      | Adipocytes        |       | -0.03 |          | 9.73E-01 |
| <i>DMBT1</i>      | Endothelial cells | 13.74 | 0.39  | 6.25E-43 | 6.93E-01 |
| <i>DMBT1</i>      | Fibroblasts       |       | 0.32  |          | 7.49E-01 |
| <i>EMBP1</i>      | Bulk tissue       | -8.96 | -2.10 | 3.11E-19 | 3.56E-02 |
| <i>EMBP1</i>      | Adipocytes        |       | -2.92 |          | 3.46E-03 |
| <i>EMID1</i>      | Bulk tissue       | -5.07 | 0.75  | 3.94E-07 | 4.55E-01 |
| <i>EMID1</i>      | Adipocytes        | -4.81 |       | 1.52E-06 |          |
| <i>EMID1</i>      | Endothelial cells |       | -0.43 |          | 6.68E-01 |
| <i>EMID1</i>      | Epithelial cells  | -3.68 | -0.53 | 2.37E-04 | 5.93E-01 |
| <i>EMID1</i>      | Fibroblasts       | -6.05 |       | 1.41E-09 |          |
| <i>GAN</i>        | Bulk tissue       | 8.75  | 1.25  | 2.09E-18 | 2.10E-01 |
| <i>GAN</i>        | Epithelial cells  |       | 1.33  |          | 1.83E-01 |
| <i>GAN</i>        | Fibroblasts       | 4.35  | -0.50 | 1.33E-05 | 6.17E-01 |
| <i>GBAP1</i>      | Bulk tissue       | -6.13 | -1.44 | 8.81E-10 | 1.49E-01 |
| <i>GBAP1</i>      | Adipocytes        | -5.76 | -1.05 | 8.31E-09 | 2.93E-01 |
| <i>GBAP1</i>      | Endothelial cells |       | -1.61 |          | 1.07E-01 |
| <i>GBAP1</i>      | Fibroblasts       | -6.20 |       | 5.52E-10 |          |
| <i>GDI2</i>       | Bulk tissue       | -4.78 | -2.71 | 1.76E-06 | 6.73E-03 |
| <i>GDI2</i>       | Endothelial cells |       | -2.07 |          | 3.85E-02 |
| <i>GDI2</i>       | Epithelial cells  | -4.75 | -2.63 | 2.03E-06 | 8.44E-03 |
| <i>GLRA3</i>      | Bulk tissue       | -7.59 | -1.04 | 3.32E-14 | 2.97E-01 |
| <i>GLRA3</i>      | Adipocytes        | -0.54 |       | 5.91E-01 |          |
| <i>GLRA3</i>      | Endothelial cells |       | -2.24 |          | 2.51E-02 |
| <i>GLRA3</i>      | Fibroblasts       | -6.83 | 1.06  | 8.22E-12 | 2.89E-01 |
| <i>HIST2H2BA</i>  | Bulk tissue       | 12.87 | 2.21  | 6.64E-38 | 2.70E-02 |
| <i>HIST2H3DP1</i> | Bulk tissue       | 13.97 | 0.89  | 2.45E-44 | 3.73E-01 |
| <i>HIST2H3DP1</i> | Endothelial cells | 13.84 |       | 1.46E-43 |          |
| <i>HIST4H4</i>    | Bulk tissue       | -5.92 | -0.65 | 3.26E-09 | 5.16E-01 |
| <i>HIST4H4</i>    | Adipocytes        | 1.67  | 0.80  | 9.42E-02 | 4.21E-01 |
| <i>HIST4H4</i>    | Endothelial cells |       | -0.06 |          | 9.56E-01 |
| <i>HIST4H4</i>    | Epithelial cells  | -2.04 | 0.02  | 4.14E-02 | 9.88E-01 |

Continuation of Table S 1.11

|                   |                   |        |       |          |          |
|-------------------|-------------------|--------|-------|----------|----------|
| <i>HIST4H4</i>    | Fibroblasts       | 0.90   | 0.07  | 3.67E-01 | 9.41E-01 |
| <i>HSD3BP4</i>    | Bulk tissue       | -5.44  |       | 5.26E-08 |          |
| <i>HSD3BP4</i>    | Adipocytes        | -6.11  |       | 1.02E-09 |          |
| <i>JHDM1D-AS1</i> | Bulk tissue       | 6.77   | -0.32 | 1.28E-11 | 7.46E-01 |
| <i>JHDM1D-AS1</i> | Endothelial cells |        | 0.51  |          | 6.08E-01 |
| <i>JHDM1D-AS1</i> | Epithelial cells  |        | -1.49 |          | 1.36E-01 |
| <i>JUNB</i>       | Bulk tissue       | 5.74   | -0.58 | 9.36E-09 | 5.64E-01 |
| <i>JUNB</i>       | Adipocytes        | -1.54  | -0.17 | 1.24E-01 | 8.63E-01 |
| <i>JUNB</i>       | Endothelial cells | 2.37   |       | 1.79E-02 |          |
| <i>JUNB</i>       | Fibroblasts       |        | -0.24 |          | 8.09E-01 |
| <i>KCNN4</i>      | Bulk tissue       | 5.56   | 0.92  | 2.72E-08 | 3.55E-01 |
| <i>KCNN4</i>      | Adipocytes        |        | 1.00  |          | 3.16E-01 |
| <i>KCNN4</i>      | Fibroblasts       |        | 1.14  |          | 2.53E-01 |
| <i>KLHDC7A</i>    | Bulk tissue       | -3.84  | -1.46 | 1.21E-04 | 1.45E-01 |
| <i>KLHDC7A</i>    | Adipocytes        | -7.34  | -1.29 | 2.14E-13 | 1.96E-01 |
| <i>KLHDC7A</i>    | Endothelial cells | -7.73  | -1.62 | 1.04E-14 | 1.05E-01 |
| <i>KLHDC7A</i>    | Epithelial cells  | -0.87  |       | 3.84E-01 |          |
| <i>KLHDC7A</i>    | Fibroblasts       | -3.25  | -1.12 | 1.14E-03 | 2.62E-01 |
| <i>LRRC25</i>     | Bulk tissue       | 4.50   | 2.05  | 6.74E-06 | 4.08E-02 |
| <i>LRRC25</i>     | Adipocytes        | 3.92   |       | 8.98E-05 |          |
| <i>LRRC25</i>     | Endothelial cells | 5.22   | 1.99  | 1.76E-07 | 4.71E-02 |
| <i>LRRC25</i>     | Epithelial cells  | 6.85   | -1.28 | 7.34E-12 | 1.99E-01 |
| <i>LRRC25</i>     | Fibroblasts       | 7.21   | 1.72  | 5.41E-13 | 8.54E-02 |
| <i>LRRC3B</i>     | Bulk tissue       | -13.83 |       | 1.59E-43 |          |
| <i>LRRC3B</i>     | Adipocytes        | -12.15 |       | 5.59E-34 |          |
| <i>LRRC3B</i>     | Endothelial cells | -13.80 |       | 2.65E-43 |          |
| <i>LRRC3B</i>     | Epithelial cells  | -7.83  |       | 4.87E-15 |          |
| <i>LRRC3B</i>     | Fibroblasts       | -12.24 |       | 2.00E-34 |          |
| <i>LURAP1</i>     | Bulk tissue       | -5.83  | -0.97 | 5.68E-09 | 3.31E-01 |
| <i>LURAP1</i>     | Endothelial cells |        | 0.83  |          | 4.06E-01 |
| <i>LURAP1</i>     | Fibroblasts       | -5.04  |       | 4.58E-07 |          |
| <i>MAN2C1</i>     | Bulk tissue       | -5.66  | -3.58 | 1.48E-08 | 3.47E-04 |

Continuation of Table S 1.11

|                  |                   |       |       |          |          |
|------------------|-------------------|-------|-------|----------|----------|
| <i>MAN2C1</i>    | Endothelial cells |       | -3.61 |          | 3.11E-04 |
| <i>MAN2C1</i>    | Epithelial cells  |       | -3.44 |          | 5.87E-04 |
| <i>MANSC4</i>    | Bulk tissue       | 5.66  |       | 1.54E-08 |          |
| <i>MANSC4</i>    | Endothelial cells | 1.49  |       | 1.36E-01 |          |
| <i>MANSC4</i>    | Epithelial cells  | -1.82 |       | 6.92E-02 |          |
| <i>METTL15P1</i> | Bulk tissue       | -5.31 | 1.25  | 1.09E-07 | 2.12E-01 |
| <i>MN1</i>       | Bulk tissue       | 7.31  | -0.18 | 2.62E-13 | 8.54E-01 |
| <i>MN1</i>       | Adipocytes        |       | 0.94  |          | 3.45E-01 |
| <i>MN1</i>       | Endothelial cells |       | 0.91  |          | 3.63E-01 |
| <i>MN1</i>       | Epithelial cells  |       | -0.38 |          | 7.03E-01 |
| <i>MN1</i>       | Fibroblasts       |       | -0.27 |          | 7.90E-01 |
| <i>MSTO2P</i>    | Bulk tissue       | -5.15 | 1.16  | 2.67E-07 | 2.47E-01 |
| <i>MSTO2P</i>    | Epithelial cells  | -6.07 |       | 1.26E-09 |          |
| <i>MSTO2P</i>    | Fibroblasts       | -5.16 |       | 2.42E-07 |          |
| <i>NBPF26</i>    | Bulk tissue       | 8.12  | 1.01  | 4.60E-16 | 3.11E-01 |
| <i>NBPF26</i>    | Adipocytes        | 6.93  |       | 4.08E-12 |          |
| <i>NBPF26</i>    | Endothelial cells |       | 1.46  |          | 1.45E-01 |
| <i>NBPF26</i>    | Epithelial cells  | 8.95  | 0.84  | 3.50E-19 | 4.01E-01 |
| <i>NBPF26</i>    | Fibroblasts       | 6.50  |       | 8.01E-11 |          |
| <i>NBPF8</i>     | Bulk tissue       | 2.36  | -0.88 | 1.83E-02 | 3.77E-01 |
| <i>NBPF8</i>     | Adipocytes        |       | -0.69 |          | 4.91E-01 |
| <i>NBPF8</i>     | Endothelial cells | -8.03 |       | 1.01E-15 |          |
| <i>NBPF8</i>     | Fibroblasts       |       | 1.62  |          | 1.05E-01 |
| <i>NHLRC2</i>    | Bulk tissue       |       | 0.52  |          | 6.03E-01 |
| <i>NHLRC2</i>    | Adipocytes        |       | 1.73  |          | 8.31E-02 |
| <i>NHLRC2</i>    | Endothelial cells | 6.28  | 1.18  | 3.36E-10 | 2.40E-01 |
| <i>NHLRC2</i>    | Epithelial cells  |       | 0.00  |          | 9.98E-01 |
| <i>NHLRC2</i>    | Fibroblasts       |       | -1.09 |          | 2.78E-01 |
| <i>NSUN4</i>     | Bulk tissue       | -5.54 | 0.95  | 3.04E-08 | 3.45E-01 |
| <i>NSUN4</i>     | Adipocytes        | -5.46 |       | 4.65E-08 |          |
| <i>NSUN4</i>     | Epithelial cells  |       | -0.80 |          | 4.21E-01 |
| <i>NSUN4</i>     | Fibroblasts       |       | 0.79  |          | 4.29E-01 |

Continuation of Table S 1.11

|                  |                   |       |       |          |          |
|------------------|-------------------|-------|-------|----------|----------|
| <i>NTN4</i>      | Bulk tissue       | -6.50 | -2.01 | 8.15E-11 | 4.46E-02 |
| <i>NTN4</i>      | Adipocytes        | -3.87 |       | 1.08E-04 |          |
| <i>NTN4</i>      | Endothelial cells | -2.39 | -0.64 | 1.69E-02 | 5.22E-01 |
| <i>NTN4</i>      | Epithelial cells  | -9.98 | 0.76  | 1.91E-23 | 4.48E-01 |
| <i>NTN4</i>      | Fibroblasts       | -5.63 | -1.00 | 1.79E-08 | 3.20E-01 |
| <i>PA2G4P4</i>   | Bulk tissue       | -5.19 | -0.65 | 2.12E-07 | 5.18E-01 |
| <i>PA2G4P4</i>   | Adipocytes        |       | -0.21 |          | 8.37E-01 |
| <i>PA2G4P4</i>   | Endothelial cells |       | -0.76 |          | 4.49E-01 |
| <i>PA2G4P4</i>   | Epithelial cells  | -5.47 | -0.99 | 4.51E-08 | 3.23E-01 |
| <i>PARP8</i>     | Bulk tissue       | 3.16  | -1.87 | 1.60E-03 | 6.13E-02 |
| <i>PARP8</i>     | Adipocytes        |       | -0.99 |          | 3.21E-01 |
| <i>PARP8</i>     | Endothelial cells | 5.39  | -0.87 | 6.90E-08 | 3.82E-01 |
| <i>PARP8</i>     | Epithelial cells  | 3.34  |       | 8.46E-04 |          |
| <i>PARP8</i>     | Fibroblasts       |       | -0.85 |          | 3.93E-01 |
| <i>PI15</i>      | Bulk tissue       |       | -0.52 |          | 6.02E-01 |
| <i>PI15</i>      | Adipocytes        | -5.57 | -0.27 | 2.48E-08 | 7.88E-01 |
| <i>PI15</i>      | Endothelial cells | -7.42 | -0.61 | 1.21E-13 | 5.42E-01 |
| <i>PI15</i>      | Epithelial cells  |       | -2.00 |          | 4.58E-02 |
| <i>PRC1-AS1</i>  | Bulk tissue       | 4.68  | 1.85  | 2.81E-06 | 6.47E-02 |
| <i>PRC1-AS1</i>  | Adipocytes        | 4.95  | 2.09  | 7.61E-07 | 3.66E-02 |
| <i>PRC1-AS1</i>  | Endothelial cells | 4.95  | 0.91  | 7.61E-07 | 3.64E-01 |
| <i>PRC1-AS1</i>  | Epithelial cells  | -1.13 |       | 2.58E-01 |          |
| <i>PRC1-AS1</i>  | Fibroblasts       | 3.54  | 1.72  | 3.93E-04 | 8.60E-02 |
| <i>PSMD6-AS1</i> | Bulk tissue       |       | 0.98  |          | 3.29E-01 |
| <i>PSMD6-AS1</i> | Epithelial cells  | -5.89 | -0.41 | 3.82E-09 | 6.80E-01 |
| <i>PSMD6-AS1</i> | Fibroblasts       | 0.38  |       | 7.01E-01 |          |
| <i>PTHLH</i>     | Bulk tissue       | 3.71  | 2.34  | 2.10E-04 | 1.95E-02 |
| <i>PTHLH</i>     | Adipocytes        | 0.87  |       | 3.84E-01 |          |
| <i>PTHLH</i>     | Epithelial cells  | 5.36  |       | 8.30E-08 |          |
| <i>PTHLH</i>     | Fibroblasts       |       | -0.24 |          | 8.10E-01 |
| <i>PVT1</i>      | Bulk tissue       |       | -1.73 |          | 8.42E-02 |
| <i>PVT1</i>      | Adipocytes        | 0.56  |       | 5.78E-01 |          |
| <i>PVT1</i>      | Endothelial cells | -3.14 | -1.25 | 1.70E-03 | 2.12E-01 |
| <i>PVT1</i>      | Epithelial cells  |       | 0.52  |          | 6.03E-01 |



Continuation of Table S 1.11

|                      |                   |       |       |          |          |
|----------------------|-------------------|-------|-------|----------|----------|
| <i>PVT1</i>          | Fibroblasts       | -6.24 |       | 4.51E-10 |          |
| <i>RAD51B</i>        | Bulk tissue       | 5.78  | 0.49  | 7.46E-09 | 6.24E-01 |
| <i>RAD51B</i>        | Adipocytes        |       | 0.59  |          | 5.58E-01 |
| <i>RAD51B</i>        | Endothelial cells | 0.06  | -0.68 | 9.51E-01 | 5.00E-01 |
| <i>RAD51B</i>        | Epithelial cells  | 3.88  |       | 1.03E-04 |          |
| <i>RAD51B</i>        | Fibroblasts       |       | -0.85 |          | 3.93E-01 |
| <i>RBL2</i>          | Bulk tissue       | -0.36 | 0.40  | 7.19E-01 | 6.93E-01 |
| <i>RBL2</i>          | Adipocytes        | -2.35 | 1.26  | 1.87E-02 | 2.06E-01 |
| <i>RBL2</i>          | Endothelial cells | -6.71 | -0.03 | 2.01E-11 | 9.75E-01 |
| <i>RBL2</i>          | Fibroblasts       |       | -0.33 |          | 7.42E-01 |
| <i>RCCD1</i>         | Bulk tissue       | -6.44 | -1.39 | 1.22E-10 | 1.63E-01 |
| <i>RCCD1</i>         | Adipocytes        | -4.67 | -2.33 | 3.05E-06 | 1.97E-02 |
| <i>RCCD1</i>         | Endothelial cells | -3.07 |       | 2.14E-03 |          |
| <i>RCCD1</i>         | Epithelial cells  |       | -1.45 |          | 1.48E-01 |
| <i>RCCD1</i>         | Fibroblasts       | -4.11 |       | 4.01E-05 |          |
| <i>RIT1</i>          | Bulk tissue       | -5.35 |       | 8.73E-08 |          |
| <i>RIT1</i>          | Adipocytes        | -5.00 | -1.62 | 5.62E-07 | 1.06E-01 |
| <i>RIT1</i>          | Endothelial cells | -5.84 | -1.31 | 5.08E-09 | 1.90E-01 |
| <i>RIT1</i>          | Epithelial cells  | -4.18 |       | 2.93E-05 |          |
| <i>RIT1</i>          | Fibroblasts       | -5.28 | -0.20 | 1.33E-07 | 8.44E-01 |
| <i>RN7SL151P</i>     | Bulk tissue       | -5.67 | 1.57  | 1.40E-08 | 1.16E-01 |
| <i>RN7SL151P</i>     | Epithelial cells  | -0.48 |       | 6.32E-01 |          |
| <i>RN7SL151P</i>     | Fibroblasts       | -2.27 | -0.04 | 2.34E-02 | 9.68E-01 |
| <i>RNF139</i>        | Bulk tissue       | -5.74 |       | 9.23E-09 |          |
| <i>RNF139</i>        | Endothelial cells | -2.19 |       | 2.88E-02 |          |
| <i>RNF139</i>        | Epithelial cells  | -5.19 |       | 2.14E-07 |          |
| <i>RNF139</i>        | Fibroblasts       | -0.97 |       | 3.31E-01 |          |
| <i>RNVU1-19</i>      | Bulk tissue       | 7.43  | 0.28  | 1.12E-13 | 7.81E-01 |
| <i>RNVU1-19</i>      | Fibroblasts       |       | 2.00  |          | 4.55E-02 |
| <i>RP11-107K17.2</i> | Bulk tissue       | -5.69 | -0.88 | 1.26E-08 | 3.79E-01 |
| <i>RP11-107K17.2</i> | Adipocytes        |       | -1.37 |          | 1.69E-01 |
| <i>RP11-107K17.2</i> | Epithelial cells  |       | -1.59 |          | 1.13E-01 |

Continuation of Table S 1.11

|                      |                   |       |       |          |          |
|----------------------|-------------------|-------|-------|----------|----------|
| <i>RP11-15A1.2</i>   | Bulk tissue       | 5.34  | 0.40  | 9.10E-08 | 6.90E-01 |
| <i>RP11-15A1.2</i>   | Fibroblasts       |       | -0.51 |          | 6.10E-01 |
| <i>RP11-15A1.3</i>   | Bulk tissue       | 6.78  | -1.14 | 1.17E-11 | 2.56E-01 |
| <i>RP11-15A1.3</i>   | Adipocytes        | 5.51  |       | 3.67E-08 |          |
| <i>RP11-15A1.3</i>   | Epithelial cells  | 6.68  |       | 2.41E-11 |          |
| <i>RP11-15A1.3</i>   | Fibroblasts       | 6.93  | -0.79 | 4.30E-12 | 4.27E-01 |
| <i>RP11-425D17.1</i> | Bulk tissue       | 7.77  |       | 7.93E-15 |          |
| <i>RP11-425D17.1</i> | Endothelial cells | 7.32  |       | 2.50E-13 |          |
| <i>RP11-425D17.1</i> | Epithelial cells  |       | -0.68 |          | 4.95E-01 |
| <i>RP11-425D17.1</i> | Fibroblasts       |       | 0.01  |          | 9.90E-01 |
| <i>RP11-500G22.4</i> | Bulk tissue       | 5.00  |       | 5.87E-07 |          |
| <i>RP11-500G22.4</i> | Epithelial cells  | 5.73  |       | 9.99E-09 |          |
| <i>RP11-500G22.4</i> | Fibroblasts       | 5.52  |       | 3.36E-08 |          |
| <i>RP11-53O19.1</i>  | Bulk tissue       | 12.33 | 3.55  | 6.62E-35 | 3.87E-04 |
| <i>RP11-53O19.1</i>  | Adipocytes        | 11.62 |       | 3.37E-31 |          |
| <i>RP11-53O19.1</i>  | Endothelial cells | 11.89 |       | 1.36E-32 |          |
| <i>RP11-53O19.1</i>  | Epithelial cells  | 3.22  |       | 1.29E-03 |          |
| <i>RP11-53O19.1</i>  | Fibroblasts       | 10.34 | -0.94 | 4.85E-25 | 3.47E-01 |
| <i>RP11-967K21.1</i> | Bulk tissue       | 8.20  |       | 2.41E-16 |          |
| <i>RP11-967K21.1</i> | Adipocytes        | 7.71  | 0.83  | 1.25E-14 | 4.04E-01 |
| <i>RP11-967K21.1</i> | Epithelial cells  | 7.95  |       | 1.90E-15 |          |
| <i>RP11-967K21.1</i> | Fibroblasts       | 7.52  | 0.19  | 5.35E-14 | 8.50E-01 |
| <i>RPS23</i>         | Bulk tissue       | 6.33  | 2.40  | 2.42E-10 | 1.66E-02 |
| <i>RPS23</i>         | Adipocytes        | 6.17  |       | 6.74E-10 |          |
| <i>RPS23</i>         | Endothelial cells | 6.56  | 2.26  | 5.28E-11 | 2.40E-02 |
| <i>RPS23</i>         | Fibroblasts       |       | 2.19  |          | 2.87E-02 |
| <i>SIRT5</i>         | Bulk tissue       | 5.56  |       | 2.73E-08 |          |
| <i>SIRT5</i>         | Adipocytes        | 5.53  |       | 3.29E-08 |          |
| <i>SIRT5</i>         | Epithelial cells  | 4.53  |       | 5.79E-06 |          |
| <i>SIRT5</i>         | Fibroblasts       | 6.27  | 1.19  | 3.56E-10 | 2.35E-01 |
| <i>SKIDA1</i>        | Bulk tissue       | 9.58  | -0.11 | 9.86E-22 | 9.11E-01 |
| <i>SKIDA1</i>        | Adipocytes        | 10.63 | -0.84 | 2.19E-26 | 4.01E-01 |

Continuation of Table S 1.11

|                   |                   |        |       |          |          |
|-------------------|-------------------|--------|-------|----------|----------|
| <i>SLC25A14P1</i> | Bulk tissue       | -5.69  | -0.32 | 1.29E-08 | 7.52E-01 |
| <i>SLC25A14P1</i> | Epithelial cells  | -5.72  | 0.19  | 1.06E-08 | 8.48E-01 |
| <i>SRGAP2C</i>    | Bulk tissue       | -12.26 | -0.02 | 1.48E-34 | 9.82E-01 |
| <i>SRGAP2C</i>    | Adipocytes        |        | -0.34 |          | 7.33E-01 |
| <i>SRGAP2C</i>    | Epithelial cells  | -12.46 |       | 1.31E-35 |          |
| <i>THBS3</i>      | Bulk tissue       | 5.60   | 1.60  | 2.13E-08 | 1.10E-01 |
| <i>THBS3</i>      | Adipocytes        | 4.74   | 3.09  | 2.11E-06 | 1.98E-03 |
| <i>THBS3</i>      | Endothelial cells | 4.50   |       | 6.64E-06 |          |
| <i>THBS3</i>      | Epithelial cells  | 4.68   | 1.48  | 2.87E-06 | 1.39E-01 |
| <i>THBS3</i>      | Fibroblasts       | 3.81   | 1.11  | 1.37E-04 | 2.68E-01 |
| <i>TOM1L1</i>     | Bulk tissue       | 6.52   |       | 6.83E-11 |          |
| <i>TOM1L1</i>     | Adipocytes        | 2.70   | -0.16 | 7.00E-03 | 8.75E-01 |
| <i>TOM1L1</i>     | Endothelial cells | 1.37   |       | 1.71E-01 |          |
| <i>TOM1L1</i>     | Epithelial cells  | 4.49   |       | 7.29E-06 |          |
| <i>TOM1L1</i>     | Fibroblasts       | 1.45   |       | 1.46E-01 |          |
| <i>TOX3</i>       | Bulk tissue       | -15.39 | 1.58  | 1.83E-53 | 1.15E-01 |
| <i>TOX3</i>       | Adipocytes        |        | 1.40  |          | 1.62E-01 |
| <i>TOX3</i>       | Endothelial cells | -5.05  | 0.24  | 4.32E-07 | 8.10E-01 |
| <i>TOX3</i>       | Fibroblasts       | -17.17 | 2.34  | 4.82E-66 | 1.91E-02 |
| <i>TRIM32</i>     | Bulk tissue       | 6.27   | -0.62 | 3.58E-10 | 5.37E-01 |
| <i>TRIM32</i>     | Adipocytes        |        | 0.05  |          | 9.63E-01 |
| <i>TRIM32</i>     | Epithelial cells  |        | -1.32 |          | 1.88E-01 |
| <i>TRIM32</i>     | Fibroblasts       |        | -0.73 |          | 4.65E-01 |
| <i>ULK3</i>       | Bulk tissue       | -5.83  | 0.67  | 5.49E-09 | 5.05E-01 |
| <i>ULK3</i>       | Adipocytes        | -5.85  |       | 4.94E-09 |          |
| <i>ULK3</i>       | Endothelial cells |        | 0.29  |          | 7.71E-01 |
| <i>ULK3</i>       | Fibroblasts       |        | 1.50  |          | 1.34E-01 |
| <i>XRCC5</i>      | Bulk tissue       | 6.19   | -1.98 | 6.20E-10 | 4.80E-02 |
| <i>XRCC5</i>      | Endothelial cells |        | -2.12 |          | 3.43E-02 |
| <i>XRCC5</i>      | Epithelial cells  | 3.43   |       | 6.03E-04 |          |
| <i>ZNF404</i>     | Bulk tissue       | 7.24   | 1.03  | 4.33E-13 | 3.04E-01 |
| <i>ZNF404</i>     | Endothelial cells |        | 1.02  |          | 3.07E-01 |
| Luminal-B-like    |                   |        |       |          |          |

Continuation of Table S 1.11

|                      |                   |       |       |          |          |
|----------------------|-------------------|-------|-------|----------|----------|
| <i>CTD-2310F14.1</i> | Bulk tissue       | 5.61  | -0.30 | 2.03E-08 | 7.64E-01 |
| <i>CTD-2310F14.1</i> | Adipocytes        | -3.33 | -0.32 | 8.68E-04 | 7.51E-01 |
| <i>CTD-2310F14.1</i> | Endothelial cells | 3.94  |       | 8.11E-05 |          |
| <i>DMBT1</i>         | Bulk tissue       | 4.10  | -0.06 | 4.10E-05 | 9.50E-01 |
| <i>DMBT1</i>         | Adipocytes        |       | 0.47  |          | 6.36E-01 |
| <i>DMBT1</i>         | Endothelial cells | 4.82  | -0.37 | 1.44E-06 | 7.13E-01 |
| <i>DMBT1</i>         | Fibroblasts       |       | 1.07  |          | 2.86E-01 |
| <i>KLHDC7A</i>       | Bulk tissue       | -2.48 | -0.86 | 1.32E-02 | 3.91E-01 |
| <i>KLHDC7A</i>       | Adipocytes        | -4.20 | -0.72 | 2.61E-05 | 4.69E-01 |
| <i>KLHDC7A</i>       | Endothelial cells | -4.55 | -1.65 | 5.31E-06 | 9.83E-02 |
| <i>KLHDC7A</i>       | Epithelial cells  | -1.00 |       | 3.18E-01 |          |
| <i>KLHDC7A</i>       | Fibroblasts       | -2.88 | -0.62 | 3.95E-03 | 5.38E-01 |
| <i>PRC1-AS1</i>      | Bulk tissue       | 5.31  | 0.46  | 1.12E-07 | 6.46E-01 |
| <i>PRC1-AS1</i>      | Adipocytes        | 4.70  | 0.43  | 2.57E-06 | 6.65E-01 |
| <i>PRC1-AS1</i>      | Endothelial cells | 5.22  | -0.99 | 1.76E-07 | 3.22E-01 |
| <i>PRC1-AS1</i>      | Epithelial cells  | 0.07  |       | 9.47E-01 |          |
| <i>PRC1-AS1</i>      | Fibroblasts       | 4.75  | -0.74 | 2.02E-06 | 4.57E-01 |
| <i>RCCD1</i>         | Bulk tissue       | -4.43 | 0.21  | 9.58E-06 | 8.31E-01 |
| <i>RCCD1</i>         | Adipocytes        | -3.73 | -0.66 | 1.89E-04 | 5.11E-01 |
| <i>RCCD1</i>         | Endothelial cells | -2.85 |       | 4.31E-03 |          |
| <i>RCCD1</i>         | Epithelial cells  |       | 0.56  |          | 5.74E-01 |
| <i>RCCD1</i>         | Fibroblasts       | -2.26 |       | 2.41E-02 |          |
| <i>SLC25A14P1</i>    | Bulk tissue       | -5.13 | -0.01 | 2.97E-07 | 9.94E-01 |
| <i>SLC25A14P1</i>    | Epithelial cells  | -4.94 | -0.17 | 7.96E-07 | 8.68E-01 |
| <i>TOX3</i>          | Bulk tissue       | -6.64 | 1.16  | 3.23E-11 | 2.45E-01 |
| <i>TOX3</i>          | Adipocytes        |       | 1.14  |          | 2.55E-01 |
| <i>TOX3</i>          | Endothelial cells | -3.29 | 0.94  | 1.01E-03 | 3.48E-01 |
| <i>TOX3</i>          | Fibroblasts       | -7.97 | -0.49 | 1.54E-15 | 6.25E-01 |
| HER2-enriched-like   |                   |       |       |          |          |
| <i>LGR6</i>          | Bulk tissue       |       | 1.57  |          | 1.16E-01 |
| <i>LGR6</i>          | Adipocytes        |       | 1.76  |          | 7.82E-02 |
| <i>LGR6</i>          | Endothelial cells | -5.91 | 1.20  | 3.50E-09 | 2.32E-01 |
| <i>LGR6</i>          | Epithelial cells  |       | -0.33 |          | 7.42E-01 |
| <i>LGR6</i>          | Fibroblasts       | -5.41 |       | 6.47E-08 |          |

Continuation of Table S 1.11

|                    |                   |       |       |          |          |
|--------------------|-------------------|-------|-------|----------|----------|
| <i>TOX3</i>        | Bulk tissue       | -4.12 | 0.36  | 3.84E-05 | 7.19E-01 |
| <i>TOX3</i>        | Adipocytes        |       | 0.87  |          | 3.83E-01 |
| <i>TOX3</i>        | Endothelial cells | -2.20 | 0.05  | 2.81E-02 | 9.63E-01 |
| <i>TOX3</i>        | Fibroblasts       | -5.03 | 0.04  | 4.91E-07 | 9.69E-01 |
| Triple-negative    |                   |       |       |          |          |
| <i>ABHD8</i>       | Bulk tissue       | 8.26  | 1.54  | 1.42E-16 | 1.25E-01 |
| <i>ABHD8</i>       | Endothelial cells |       | 1.43  |          | 1.53E-01 |
| <i>CHADL</i>       | Bulk tissue       | -4.56 | 0.07  | 5.14E-06 | 9.47E-01 |
| <i>CHADL</i>       | Adipocytes        |       | 0.88  |          | 3.78E-01 |
| <i>CHADL</i>       | Endothelial cells | -2.79 |       | 5.32E-03 |          |
| <i>CHADL</i>       | Epithelial cells  | -3.63 |       | 2.80E-04 |          |
| <i>CHADL</i>       | Fibroblasts       | -2.66 |       | 7.87E-03 |          |
| <i>EN1</i>         | Bulk tissue       | 2.17  | -5.73 | 3.00E-02 | 9.95E-09 |
| <i>EN1</i>         | Adipocytes        | 0.90  |       | 3.66E-01 |          |
| <i>EN1</i>         | Fibroblasts       | 1.19  |       | 2.33E-01 |          |
| <i>KCNN1</i>       | Bulk tissue       | -1.91 | -0.12 | 5.64E-02 | 9.07E-01 |
| <i>KCNN1</i>       | Adipocytes        | -0.89 |       | 3.72E-01 |          |
| <i>KCNN1</i>       | Endothelial cells |       | -0.74 |          | 4.60E-01 |
| <i>KCNN1</i>       | Fibroblasts       | -5.33 |       | 9.86E-08 |          |
| <i>LGR6</i>        | Bulk tissue       |       | -1.98 |          | 4.72E-02 |
| <i>LGR6</i>        | Adipocytes        |       | -1.14 |          | 2.56E-01 |
| <i>LGR6</i>        | Endothelial cells | -4.54 | -0.12 | 5.54E-06 | 9.07E-01 |
| <i>LGR6</i>        | Epithelial cells  |       | -1.10 |          | 2.70E-01 |
| <i>LGR6</i>        | Fibroblasts       | -4.21 |       | 2.50E-05 |          |
| <i>NTN4</i>        | Bulk tissue       | -3.00 | -1.50 | 2.72E-03 | 1.35E-01 |
| <i>NTN4</i>        | Adipocytes        | -1.06 |       | 2.90E-01 |          |
| <i>NTN4</i>        | Endothelial cells | -0.34 | -0.90 | 7.33E-01 | 3.69E-01 |
| <i>NTN4</i>        | Epithelial cells  | -4.60 | -1.34 | 4.17E-06 | 1.80E-01 |
| <i>NTN4</i>        | Fibroblasts       | -2.26 | -1.41 | 2.38E-02 | 1.60E-01 |
| <i>PLVAP</i>       | Bulk tissue       | 0.59  | 0.24  | 5.53E-01 | 8.14E-01 |
| <i>PLVAP</i>       | Adipocytes        | -4.76 | 4.62  | 1.97E-06 | 3.76E-06 |
| <i>PLVAP</i>       | Endothelial cells | -5.06 |       | 4.10E-07 |          |
| <i>PLVAP</i>       | Fibroblasts       | -6.46 | -0.41 | 1.07E-10 | 6.85E-01 |
| <i>RP11-15A1.3</i> | Bulk tissue       | 4.27  | 1.82  | 1.98E-05 | 6.95E-02 |

Continuation of Table S 1.11

|                      |                   |      |       |          |          |
|----------------------|-------------------|------|-------|----------|----------|
| <i>RP11-15A1.3</i>   | Adipocytes        | 3.80 |       | 1.42E-04 |          |
| <i>RP11-15A1.3</i>   | Epithelial cells  | 4.13 |       | 3.55E-05 |          |
| <i>RP11-15A1.3</i>   | Fibroblasts       | 4.66 | 1.36  | 3.22E-06 | 1.72E-01 |
| <i>RP11-425D17.1</i> | Bulk tissue       | 5.49 |       | 4.04E-08 |          |
| <i>RP11-425D17.1</i> | Endothelial cells | 5.44 |       | 5.38E-08 |          |
| <i>RP11-425D17.1</i> | Epithelial cells  |      | 0.80  |          | 4.25E-01 |
| <i>RP11-425D17.1</i> | Fibroblasts       |      | 1.83  |          | 6.78E-02 |
| <i>RP11-967K21.1</i> | Bulk tissue       | 5.57 |       | 2.61E-08 |          |
| <i>RP11-967K21.1</i> | Adipocytes        | 4.78 | -0.14 | 1.72E-06 | 8.91E-01 |
| <i>RP11-967K21.1</i> | Epithelial cells  | 5.57 |       | 2.53E-08 |          |
| <i>RP11-967K21.1</i> | Fibroblasts       | 5.02 | 0.28  | 5.24E-07 | 7.76E-01 |
| <i>ZNF404</i>        | Bulk tissue       | 4.89 | 0.09  | 9.94E-07 | 9.27E-01 |
| <i>ZNF404</i>        | Endothelial cells |      | 0.54  |          | 5.86E-01 |

**Footnote:**

Bonferroni threshold of significance utilized in both the primary TWAS and conditional (Cond-TWAS) corresponded to a p-value of  $1.42 \times 10^{-7}$  for European-ancestry TWAS,  $1.39 \times 10^{-7}$  for African-ancestry TWAS, and  $1.26 \times 10^{-7}$  for cross-ancestry TWAS

Table S 1.12 Genes previously reported in TWAS conducted across all breast cancer phenotypes including overall breast cancer and various subtypes.

| Gene Symbol             | ENSG ID         | PMID List                                                            | Identified in previous intrinsic subtype TWAS | Identified in current intrinsic subtype TWAS |
|-------------------------|-----------------|----------------------------------------------------------------------|-----------------------------------------------|----------------------------------------------|
| <i>CENPS</i>            | ENSG00000175279 | 36356581, 37164006, 38400758, 38733992                               | Yes                                           | No                                           |
| <i>DFFA</i>             | ENSG00000160049 | 36356581, 37164006, 38400758, 38733992                               | Yes                                           | No                                           |
| <i>PEX14</i>            | ENSG00000142655 | 36356581, 37164006, 38400758, 38733992                               | Yes                                           | No                                           |
| <i>PADI6</i>            | ENSG00000276747 | 34518541                                                             | No                                            | No                                           |
| <i>KLHDC7A</i>          | ENSG00000179023 | 29915430, 34518541, 36356581, 36402776, 37164006, 38400758, 38723630 | Yes                                           | Yes                                          |
| <i>STPG1</i>            | ENSG00000001460 | 34518541                                                             | No                                            | No                                           |
| <i>ZMYND12</i>          | ENSG00000066185 | 34518541                                                             | No                                            | No                                           |
| <i>ZSWIM5</i>           | ENSG00000162415 | 29915430                                                             | No                                            | No                                           |
| <i>MUTYH</i>            | ENSG00000132781 | 34518541, 38733992                                                   | No                                            | No                                           |
| <i>NASP</i>             | ENSG00000132780 | 36402776, 37164006, 38733992                                         | No                                            | No                                           |
| <i>LURAP1</i>           | ENSG00000171357 | 34518541, 38400758                                                   | Yes                                           | Yes                                          |
| <i>NSUN4</i>            | ENSG00000117481 | 34518541, 36356581, 36402776, 37164006, 38400758, 38733992           | Yes                                           | Yes                                          |
| <i>RP4-603I14.3</i>     | ENSG00000225779 | 36356581                                                             | No                                            | No                                           |
| <i>MKNK1-AS1</i>        | ENSG00000269956 | 36356581                                                             | No                                            | No                                           |
| <i>CDC7</i>             | ENSG00000097046 | 34518541                                                             | No                                            | No                                           |
| <i>GSTM4</i>            | ENSG00000168765 | 34518541, 37164006                                                   | No                                            | No                                           |
| <i>GSTM2</i>            | ENSG00000213366 | 30988301, 36356581, 37164006                                         | No                                            | No                                           |
| <i>GSTM1</i>            | ENSG00000134184 | 34518541, 36356581, 37164006, 38733992                               | No                                            | No                                           |
| <i>LRIG2</i>            | ENSG00000198799 | 34518541                                                             | No                                            | No                                           |
| <i>MAGI3</i>            | ENSG00000081026 | 36402776, 38733992                                                   | No                                            | No                                           |
| <i>RSBN1</i>            | ENSG00000081019 | 36356581                                                             | No                                            | No                                           |
| <i>NBPF8</i>            | ENSG00000270231 | 36356581, 38400758                                                   | Yes                                           | Yes                                          |
| <i>PFN1P2</i>           | ENSG00000270392 | 36356581                                                             | No                                            | No                                           |
| <i>H2BP1, HIST2H2BA</i> | ENSG00000223345 | 32115800                                                             | No                                            | Yes                                          |
| <i>H3P4, HIST2H3DP1</i> | ENSG00000213244 | 36356581                                                             | No                                            | Yes                                          |
| <i>FAM72B</i>           | ENSG00000188610 | 29915430, 37164006, 38400758                                         | Yes                                           | No                                           |
| <i>SRGAP2C</i>          | ENSG00000171943 | 36356581, 36690614, 37164006, 38400758, 38733992                     | Yes                                           | Yes                                          |

Continuation of Table S 1.12

|                                            |                 |                                                                  |     |     |
|--------------------------------------------|-----------------|------------------------------------------------------------------|-----|-----|
| <i>LINC02798,</i><br><i>CH17-437K3.1</i>   | ENSG00000227082 | 36356581, 36690614,<br>37164006, 38400758,<br>38723630           | Yes | Yes |
| <i>RP11-343N15.2</i>                       | ENSG00000231429 | 36356581                                                         | No  | No  |
| <i>EMBP1</i>                               | ENSG00000231752 | 36356581, 38723630                                               | Yes | Yes |
| <i>NUDT17</i>                              | ENSG00000186364 | 29915430, 34518541                                               | No  | No  |
| <i>ANKRD34A</i>                            | ENSG00000272031 | 29915430                                                         | No  | No  |
| <i>EFNA1</i>                               | ENSG00000169242 | 36402776                                                         | No  | No  |
| <i>MUC1</i>                                | ENSG00000185499 | 36402776, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>THBS3</i>                               | ENSG00000169231 | 36356581, 36402776,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | Yes |
| <i>GBAP1, GBA1LP</i>                       | ENSG00000160766 | 36356581, 38400758,<br>38723630, 38733992                        | Yes | Yes |
| <i>FAM189B</i>                             | ENSG00000160767 | 36402776, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>CLK2</i>                                | ENSG00000176444 | 36402776, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>ASH1L</i>                               | ENSG00000116539 | 36402776, 37164006,<br>38400758, 38723630                        | Yes | No  |
| <i>ASH1L-AS1</i>                           | ENSG00000235919 | 36356581                                                         | No  | No  |
| <i>MSTO1</i>                               | ENSG00000125459 | 36402776, 37164006,<br>38733992                                  | No  | No  |
| <i>GON4L</i>                               | ENSG00000116580 | 36402776                                                         | No  | No  |
| <i>RIT1</i>                                | ENSG00000143622 | 36356581                                                         | No  | Yes |
| <i>UBQLN4</i>                              | ENSG00000160803 | 36402776                                                         | No  | No  |
| <i>SEMA4A</i>                              | ENSG00000196189 | 30988301, 36356581,<br>38400758, 38723630                        | Yes | No  |
| <i>SLC25A44</i>                            | ENSG00000160785 | 36402776                                                         | No  | No  |
| <i>BGLAP</i>                               | ENSG00000242252 | 36402776                                                         | No  | No  |
| <i>PTPRVP</i>                              | ENSG00000243323 | 36356581                                                         | No  | No  |
| <i>MDM4</i>                                | ENSG00000198625 | 36356581, 38400758                                               | Yes | No  |
| <i>CENPO</i>                               | ENSG00000138092 | 36402776, 37164006                                               | No  | No  |
| <i>DNAJC27-AS1,</i><br><i>DNAJC27-AS1a</i> | ENSG00000224165 | 36356581, 36402776,<br>38733992                                  | No  | No  |
| <i>POMC</i>                                | ENSG00000115138 | 34518541, 38400758                                               | Yes | No  |
| <i>ALK</i>                                 | ENSG00000171094 | 29915430                                                         | No  | No  |
| <i>BCL2L11</i>                             | ENSG00000153094 | 36356581, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>PTPN4</i>                               | ENSG00000088179 | 36402776                                                         | No  | No  |
| <i>AC012363.4</i>                          | ENSG00000224789 | 36356581                                                         | No  | No  |
| <i>MTCO3P43</i>                            | ENSG00000227788 | 36356581                                                         | No  | No  |
| <i>INHBB</i>                               | ENSG00000163083 | 30988301, 37164006,<br>38400758                                  | Yes | No  |
| <i>TFCP2L1</i>                             | ENSG00000115112 | 36402776                                                         | No  | No  |



Continuation of Table S 1.12

|                             |                 |                                                                                                      |     |     |
|-----------------------------|-----------------|------------------------------------------------------------------------------------------------------|-----|-----|
| <i>CYBRD1</i>               | ENSG00000071967 | 34518541, 36356581,<br>36402776, 37164006,<br>38400758, 38723630,<br>38733992                        | Yes | Yes |
| <i>NIF3L1</i>               | ENSG00000196290 | 36402776                                                                                             | No  | No  |
| <i>CASP8</i>                | ENSG00000064012 | 29915430, 34518541,<br>36356581, 36402776,<br>36690614, 37164006,<br>38400758, 38723630,<br>38733992 | Yes | Yes |
| <i>FLACC1,<br/>ALS2CR12</i> | ENSG00000155749 | 29915430, 32115800,<br>36356581, 36402776,<br>37164006, 36690614,<br>38723630                        | No  | Yes |
| <i>TRAK2</i>                | ENSG00000115993 | 36402776                                                                                             | No  | No  |
| <i>PECR</i>                 | ENSG00000115425 | 36402776                                                                                             | No  | No  |
| <i>TNP1</i>                 | ENSG00000118245 | 36356581, 37164006,<br>38400758                                                                      | Yes | No  |
| <i>TNS1</i>                 | ENSG00000079308 | 36402776                                                                                             | No  | No  |
| <i>CTDSP1</i>               | ENSG00000144579 | 36402776                                                                                             | No  | No  |
| <i>SETMAR</i>               | ENSG00000170364 | 36402776                                                                                             | No  | No  |
| <i>AC018816.3</i>           | ENSG00000235978 | 36402776                                                                                             | No  | No  |
| <i>VHL</i>                  | ENSG00000134086 | 34518541                                                                                             | No  | No  |
| <i>LRRC3B</i>               | ENSG00000179796 | 29915430, 36402776                                                                                   | No  | Yes |
| <i>SLC4A7</i>               | ENSG00000033867 | 34518541, 36690614,<br>37164006, 38400758,<br>38723630, 38733992                                     | Yes | No  |
| <i>PRSS46P</i>              | ENSG00000261603 | 29915430, 36356581,<br>37164006, 38733992                                                            | No  | No  |
| <i>PRSS45P</i>              | ENSG00000188086 | 36356581, 37164006,<br>38733992                                                                      | No  | No  |
| <i>MST1</i>                 | ENSG00000173531 | 36402776                                                                                             | No  | No  |
| <i>RNF123</i>               | ENSG00000164068 | 36356581                                                                                             | No  | No  |
| <i>GMPPB</i>                | ENSG00000173540 | 34518541                                                                                             | No  | No  |
| <i>PSMD6-AS2</i>            | ENSG00000239653 | 36356581, 38723630                                                                                   | No  | No  |
| <i>RP11-245J9.5</i>         | ENSG00000271843 | 36356581, 37164006                                                                                   | No  | No  |
| <i>PSMD6</i>                | ENSG00000163636 | 36356581, 37164006,<br>38733992                                                                      | No  | No  |
| <i>VGLL3</i>                | ENSG00000206538 | 36356581, 37164006,<br>38400758                                                                      | Yes | No  |
| <i>CMSS1</i>                | ENSG00000184220 | 36402776, 37164006,<br>38400758, 38733992                                                            | Yes | No  |
| <i>TFG</i>                  | ENSG00000114354 | 36402776                                                                                             | No  | No  |
| <i>ERICH6-AS1</i>           | ENSG00000240137 | 36356581                                                                                             | No  | No  |
| <i>METTL15P1</i>            | ENSG00000174912 | 36356581                                                                                             | No  | Yes |
| <i>LINC00886</i>            | ENSG00000240875 | 30988301, 36356581,<br>37164006, 38400758,<br>38723630                                               | Yes | No  |

Continuation of Table S 1.12

|                      |                 |                                                                               |     |     |
|----------------------|-----------------|-------------------------------------------------------------------------------|-----|-----|
| <i>TNFSF10</i>       | ENSG00000121858 | 30988301                                                                      | No  | No  |
| <i>PIGG</i>          | ENSG00000174227 | 36402776                                                                      | No  | No  |
| <i>DGKQ</i>          | ENSG00000145214 | 34518541                                                                      | No  | No  |
| <i>IDUA</i>          | ENSG00000127415 | 34518541                                                                      | No  | No  |
| <i>RNF212</i>        | ENSG00000178222 | 36402776                                                                      | No  | No  |
| <i>MAEA</i>          | ENSG00000090316 | 32115800                                                                      | No  | No  |
| <i>CNGA1</i>         | ENSG00000198515 | 34518541                                                                      | No  | No  |
| <i>SPATA18</i>       | ENSG00000163071 | 29915430                                                                      | No  | No  |
| <i>THAP9</i>         | ENSG00000168152 | 36356581, 38400758                                                            | Yes | No  |
| <i>SLC10A7</i>       | ENSG00000120519 | 34518541                                                                      | No  | No  |
| <i>EXOC3</i>         | ENSG00000180104 | 36402776, 37164006                                                            | No  | No  |
| <i>NSUN2</i>         | ENSG00000037474 | 36356581                                                                      | No  | No  |
| <i>MARCHF11</i>      | ENSG00000183654 | 34317694, 36356581,<br>37164006, 38400758<br>29915430, 36356581,              | Yes | No  |
| <i>MRPS30-DT</i>     | ENSG00000251141 | 36402776, 38400758,<br>38733992<br>34518541, 36690614,                        | Yes | Yes |
| <i>MRPS30</i>        | ENSG00000112996 | 37164006, 38400758,<br>38723630, 38733992                                     | Yes | No  |
| <i>SLC38A9</i>       | ENSG00000177058 | 36402776                                                                      | No  | No  |
| <i>IL6ST</i>         | ENSG00000134352 | 36402776                                                                      | No  | No  |
| <i>AC022431.3</i>    | ENSG00000234553 | 36356581                                                                      | No  | No  |
| <i>AC008940.1</i>    | ENSG00000228650 | 36356581                                                                      | No  | No  |
| <i>SETD9</i>         | ENSG00000155542 | 34518541, 36402776,<br>36690614, 37164006,<br>38400758, 38733992              | Yes | No  |
| <i>CTD-2310F14.1</i> | ENSG00000271828 | 36356581                                                                      | No  | Yes |
| <i>GUSBP14</i>       | ENSG00000253816 | 36356581                                                                      | No  | No  |
| <i>MSH3</i>          | ENSG00000113318 | 36402776                                                                      | No  | No  |
| <i>CKMT2</i>         | ENSG00000131730 | 34518541                                                                      | No  | No  |
| <i>ATG10</i>         | ENSG00000152348 | 29915430, 34518541,<br>36356581, 37164006,<br>38400758, 38723630,<br>38733992 | Yes | Yes |
| <i>RPS23</i>         | ENSG00000186468 | 34518541, 36356581,<br>37164006, 38400758,<br>38723630                        | Yes | Yes |
| <i>ATP6AP1L</i>      | ENSG00000205464 | 29915430, 34518541,<br>36356581, 37164006,<br>38400758, 38723630,<br>38733992 | Yes | Yes |
| <i>ADGRV1</i>        | ENSG00000164199 | 36690614                                                                      | No  | No  |
| <i>CEP120</i>        | ENSG00000168944 | 34518541                                                                      | No  | No  |
| <i>ADAMTS19</i>      | ENSG00000145808 | 34518541                                                                      | No  | No  |
| <i>P4HA2</i>         | ENSG00000072682 | 36402776, 37164006,<br>38733992                                               | No  | No  |

Continuation of Table S 1.12

|                     |                 |                                                                                         |     |    |
|---------------------|-----------------|-----------------------------------------------------------------------------------------|-----|----|
| <i>MIR3936HG</i>    | ENSG00000233006 | 30988301, 38733992                                                                      | No  | No |
| <i>SLC22A5</i>      | ENSG00000197375 | 36402776, 37164006,<br>38733992                                                         | No  | No |
| <i>IRF1-AS1</i>     | ENSG00000197536 | 36402776, 37164006,<br>38733992                                                         | No  | No |
| <i>IRF1</i>         | ENSG00000125347 | 30988301, 37164006,<br>38733992                                                         | No  | No |
| <i>RAD50</i>        | ENSG00000113522 | 36402776, 38733992                                                                      | No  | No |
| <i>KIF3A</i>        | ENSG00000131437 | 36402776                                                                                | No  | No |
| <i>ZCCHC10</i>      | ENSG00000155329 | 36402776                                                                                | No  | No |
| <i>RP11-32D16.1</i> | ENSG00000254135 | 29915430                                                                                | No  | No |
| <i>LINC02202</i>    | ENSG00000245812 | 36356581                                                                                | No  | No |
| <i>UBLCP1</i>       | ENSG00000164332 | 29915430                                                                                | No  | No |
| <i>CARMIL1</i>      | ENSG00000079691 | 36402776                                                                                | No  | No |
| <i>BTN3A2</i>       | ENSG00000186470 | 29915430, 34518541,<br>36402776                                                         | No  | No |
| <i>H4C12</i>        | ENSG00000273542 | 36356581                                                                                | No  | No |
| <i>ZNF311</i>       | ENSG00000197935 | 36402776, 37164006,<br>38733992                                                         | No  | No |
| <i>UBD</i>          | ENSG00000213886 | 29915430                                                                                | No  | No |
| <i>HLA-F</i>        | ENSG00000204642 | 36356581                                                                                | No  | No |
| <i>CSNK2B</i>       | ENSG00000204435 | 34518541                                                                                | No  | No |
| <i>VPS52</i>        | ENSG00000223501 | 30988301                                                                                | No  | No |
| <i>ZBTB9</i>        | ENSG00000213588 | 34518541                                                                                | No  | No |
| <i>ZNF318</i>       | ENSG00000171467 | 34518541                                                                                | No  | No |
| <i>RP11-250B2.5</i> | ENSG00000260645 | 32115800, 37164006                                                                      | No  | No |
| <i>RPSAP72</i>      | ENSG00000219361 | 36356581                                                                                | No  | No |
| <i>HSF2</i>         | ENSG00000025156 | 34518541, 37164006,<br>38400758                                                         | Yes | No |
| <i>L3MBTL3</i>      | ENSG00000198945 | 29915430, 32115800,<br>36356581, 36690614,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | No |
| <i>RP11-73O6.3</i>  | ENSG00000227678 | 29915430, 32115800                                                                      | No  | No |
| <i>RN7SKP268</i>    | ENSG00000223039 | 36402776                                                                                | No  | No |
| <i>ZBTB2</i>        | ENSG00000181472 | 36402776, 37164006                                                                      | No  | No |
| <i>RMND1</i>        | ENSG00000155906 | 29915430, 38733992                                                                      | No  | No |
| <i>ESR1</i>         | ENSG00000091831 | 36402776, 37164006,<br>38400758                                                         | Yes | No |
| <i>GTF2I</i>        | ENSG00000263001 | 36356581                                                                                | No  | No |
| <i>NCF1</i>         | ENSG00000158517 | 36356581                                                                                | No  | No |
| <i>GTF2IRD2</i>     | ENSG00000196275 | 36356581                                                                                | No  | No |
| <i>SPDYE12P</i>     | ENSG00000184616 | 36356581                                                                                | No  | No |
| <i>CASTOR2</i>      | ENSG00000274070 | 36356581                                                                                | No  | No |
| <i>GTF2IRD2B</i>    | ENSG00000174428 | 30988301                                                                                | No  | No |

Continuation of Table S 1.12

|                      |                 |                                                        |     |     |
|----------------------|-----------------|--------------------------------------------------------|-----|-----|
| <i>GTF2IP1</i>       | ENSG00000277053 | 36356581                                               | No  | No  |
| <i>MTERF1</i>        | ENSG00000127989 | 36402776                                               | No  | No  |
| <i>AKAP9</i>         | ENSG00000127914 | 36402776, 37164006,<br>38733992                        | No  | No  |
| <i>KRIT1</i>         | ENSG00000001631 | 36402776, 38733992                                     | No  | No  |
| <i>BET1</i>          | ENSG00000105829 | 36402776                                               | No  | No  |
| <i>SGCE</i>          | ENSG00000127990 | 36356581, 36402776,<br>37164006, 38400758,<br>38733992 | Yes | No  |
| <i>CLDN15</i>        | ENSG00000106404 | 34518541                                               | No  | No  |
| <i>FIS1</i>          | ENSG00000214253 | 34518541                                               | No  | No  |
| <i>KLHDC10</i>       | ENSG00000128607 | 29915430                                               | No  | No  |
| <i>LINC-PINT</i>     | ENSG00000231721 | 36402776, 38400758,<br>38733992                        | Yes | No  |
| <i>OR2A7</i>         | ENSG00000243896 | 34518541                                               | No  | No  |
| <i>SEPTIN14P8</i>    | ENSG00000255464 | 36356581                                               | No  | No  |
| <i>STC1</i>          | ENSG00000159167 | 34518541, 37164006                                     | No  | No  |
| <i>ZNF703</i>        | ENSG00000183779 | 36690614, 37164006,<br>38400758                        | Yes | No  |
| <i>FAM91A1</i>       | ENSG00000176853 | 36402776                                               | No  | No  |
| <i>MTSS1</i>         | ENSG00000170873 | 36402776                                               | No  | No  |
| <i>MYC</i>           | ENSG00000136997 | 36356581                                               | No  | No  |
| <i>PSIP1P1</i>       | ENSG00000234297 | 29915430                                               | No  | No  |
| <i>MIR31HG</i>       | ENSG00000171889 | 29915430                                               | No  | No  |
| <i>RP11-370B11.3</i> | ENSG00000224549 | 36402776                                               | No  | No  |
| <i>RP11-109D9.4</i>  | ENSG00000261447 | 36402776                                               | No  | No  |
| <i>SMC2</i>          | ENSG00000136824 | 36356581, 37164006,<br>38733992                        | No  | No  |
| <i>CTNNAL1</i>       | ENSG00000119326 | 36402776                                               | No  | No  |
| <i>TMEM245</i>       | ENSG00000106771 | 36690614                                               | No  | No  |
| <i>REXO4</i>         | ENSG00000148300 | 34518541                                               | No  | No  |
| <i>GDI2</i>          | ENSG00000057608 | 32115800, 37164006,<br>38733992                        | No  | Yes |
| <i>NUTM2B-AS1</i>    | ENSG00000225484 | 36402776                                               | No  | No  |
| <i>IFIT2</i>         | ENSG00000119922 | 34518541                                               | No  | No  |
| <i>CASP7</i>         | ENSG00000165806 | 34518541                                               | No  | No  |
| <i>WDR11-AS1</i>     | ENSG00000227165 | 36402776                                               | No  | No  |
| <i>WDR11</i>         | ENSG00000120008 | 36402776                                               | No  | No  |
| <i>RP11-95I16.4</i>  | ENSG00000271670 | 36402776                                               | No  | No  |
| <i>RP11-95I16.6</i>  | ENSG00000278484 | 36402776                                               | No  | No  |
| <i>ATE1</i>          | ENSG00000107669 | 36402776, 38400758,<br>38733992                        | Yes | No  |
| <i>ATE1-AS1</i>      | ENSG00000226864 | 36356581, 38400758,<br>38723630, 38733992              | Yes | Yes |
| <i>RP11-500G22.4</i> | ENSG00000276742 | 36356581                                               | No  | Yes |

Continuation of Table S 1.12

|                       |                 |                                                                  |     |     |
|-----------------------|-----------------|------------------------------------------------------------------|-----|-----|
| <i>NSMCE4A</i>        | ENSG00000107672 | 36402776, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>TACC2</i>          | ENSG00000138162 | 36402776                                                         | No  | No  |
| <i>ARMS2</i>          | ENSG00000254636 | 36402776                                                         | No  | No  |
| <i>DMBT1</i>          | ENSG00000187908 | 36402776                                                         | No  | Yes |
| <i>EEF1AKMT2</i>      | ENSG00000203791 | 36356581                                                         | No  | No  |
| <i>RIC8A</i>          | ENSG00000177963 | 29915430                                                         | No  | No  |
| <i>IFITM3</i>         | ENSG00000142089 | 34518541                                                         | No  | No  |
| <i>AP006621.5</i>     | ENSG00000255284 | 36356581, 37164006,<br>38400758, 38723630,<br>38733992           | Yes | No  |
| <i>CMB9-55F22.1</i>   | ENSG00000279672 | 36356581, 38723630<br>29915430, 36356581,                        | No  | No  |
| <i>AP006621.6</i>     | ENSG00000255142 | 37164006, 38400758,<br>38723630                                  | Yes | No  |
| <i>PIDD1</i>          | ENSG00000177595 | 34518541, 36356581,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | No  |
| <i>RPLP2</i>          | ENSG00000177600 | 29915430, 37164006,<br>38733992                                  | No  | No  |
| <i>CRACR2B</i>        | ENSG00000177685 | 34518541, 38733992                                               | No  | No  |
| <i>CD151</i>          | ENSG00000177697 | 34518541                                                         | No  | No  |
| <i>PRR33</i>          | ENSG00000283787 | 36690614, 37164006,<br>38400758                                  | Yes | No  |
| <i>MRPL23-AS1</i>     | ENSG00000226416 | 32115800, 36356581                                               | No  | No  |
| <i>LINC01219</i>      | ENSG00000232987 | 36356581, 37164006,<br>38400758                                  | Yes | No  |
| <i>ASCL2</i>          | ENSG00000183734 | 34518541                                                         | No  | No  |
| <i>TRPM5</i>          | ENSG00000070985 | 36356581                                                         | No  | No  |
| <i>CMB9-22P13.1</i>   | ENSG00000173727 | 36356581                                                         | No  | No  |
| <i>KRT8P26</i>        | ENSG00000214659 | 36356581                                                         | No  | No  |
| <i>SNX32</i>          | ENSG00000172803 | 29915430, 34518541,<br>36356581, 37164006,<br>38400758, 38733992 | Yes | No  |
| <i>FIBP</i>           | ENSG00000172500 | 36356581, 37164006,<br>38723630                                  | No  | No  |
| <i>SART1</i>          | ENSG00000175467 | 36356581                                                         | No  | No  |
| <i>B4GAT1</i>         | ENSG00000174684 | 29915430                                                         | No  | No  |
| <i>RP11-867G23.10</i> | ENSG00000254510 | 29915430                                                         | No  | No  |
| <i>CTSF</i>           | ENSG00000174080 | 34518541, 37164006                                               | No  | Yes |
| <i>RHOD</i>           | ENSG00000173156 | 36356581                                                         | No  | No  |
| <i>MRPL21</i>         | ENSG00000197345 | 36402776                                                         | No  | No  |
| <i>IGHMBP2</i>        | ENSG00000132740 | 36402776                                                         | No  | No  |
| <i>TPCN2</i>          | ENSG00000162341 | 36402776                                                         | No  | No  |
| <i>RP11-554A11.9</i>  | ENSG00000259799 | 32115800                                                         | No  | No  |
| <i>LTO1</i>           | ENSG00000149716 | 36402776, 37164006,<br>38400758                                  | Yes | No  |

Continuation of Table S 1.12

|                      |                 |                                                                  |     |     |
|----------------------|-----------------|------------------------------------------------------------------|-----|-----|
| <i>PPFIA1</i>        | ENSG00000131626 | 36402776                                                         | No  | No  |
| <i>CTTN</i>          | ENSG00000085733 | 36402776                                                         | No  | No  |
| <i>SIK2</i>          | ENSG00000170145 | 30988301                                                         | No  | No  |
| <i>PPP2R1B</i>       | ENSG00000137713 | 30988301, 37164006,<br>38733992                                  | No  | No  |
| <i>SLC6A12</i>       | ENSG00000111181 | 34518541                                                         | No  | No  |
| <i>NINJ2-AS1</i>     | ENSG00000177406 | 29915430                                                         | No  | No  |
| <i>GRIN2B</i>        | ENSG00000273079 | 36402776                                                         | No  | No  |
| <i>ERP27</i>         | ENSG00000139055 | 34518541                                                         | No  | No  |
| <i>TM7SF3</i>        | ENSG00000064115 | 36402776                                                         | No  | No  |
| <i>PPFIBP1</i>       | ENSG00000110841 | 36402776                                                         | No  | No  |
| <i>RP11-967K21.1</i> | ENSG00000247934 | 36356581, 36402776,<br>38723630                                  | Yes | Yes |
| <i>RP11-425D17.1</i> | ENSG00000278733 | 36402776, 38723630                                               | No  | Yes |
| <i>NUP107</i>        | ENSG00000111581 | 32115800, 36356581                                               | No  | No  |
| <i>PLXNC1</i>        | ENSG00000136040 | 34518541                                                         | No  | No  |
| <i>MLEC</i>          | ENSG00000110917 | 36356581                                                         | No  | No  |
| <i>RILPL1</i>        | ENSG00000188026 | 36402776                                                         | No  | No  |
| <i>DNAJC15</i>       | ENSG00000120675 | 34518541                                                         | No  | No  |
| <i>MIPOL1</i>        | ENSG00000151338 | 36402776                                                         | No  | No  |
| <i>CHURC1</i>        | ENSG00000258289 | 34518541                                                         | No  | No  |
| <i>CTD-2566J3.1</i>  | ENSG00000258837 | 36402776, 38400758,<br>38733992                                  | Yes | No  |
| <i>HMGNI1P3</i>      | ENSG00000258967 | 36356581                                                         | No  | No  |
| <i>GALNT16</i>       | ENSG00000100626 | 29915430                                                         | No  | No  |
| <i>PLEKHD1</i>       | ENSG00000175985 | 29915430                                                         | No  | No  |
| <i>VASH1</i>         | ENSG00000071246 | 36402776                                                         | No  | No  |
| <i>NRDE2</i>         | ENSG00000119720 | 36402776                                                         | No  | No  |
| <i>GPR68</i>         | ENSG00000119714 | 34518541                                                         | No  | No  |
| <i>TRIP11</i>        | ENSG00000100815 | 36402776                                                         | No  | No  |
| <i>ATXN3</i>         | ENSG00000066427 | 36402776                                                         | No  | No  |
| <i>CTD-3051D23.1</i> | ENSG00000258811 | 29915430                                                         | No  | No  |
| <i>SRP14-AS1</i>     | ENSG00000248508 | 36356581                                                         | No  | No  |
| <i>STRCP1</i>        | ENSG00000166763 | 36356581                                                         | No  | No  |
| <i>GLDN</i>          | ENSG00000186417 | 34518541                                                         | No  | No  |
| <i>ULK3</i>          | ENSG00000140474 | 32115800, 37164006,<br>38400758, 38733992                        | Yes | Yes |
| <i>MPI</i>           | ENSG00000178802 | 36356581                                                         | No  | No  |
| <i>MAN2C1</i>        | ENSG00000140400 | 29915430, 36356581,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | Yes |
| <i>RP11-817O13.8</i> | ENSG00000260274 | 30988301, 37164006,<br>38400758, 38723630                        | Yes | No  |
| <i>SIN3A</i>         | ENSG00000169375 | 30988301, 38400758,<br>38733992                                  | Yes | No  |

Continuation of Table S 1.12

|                      |                 |                                                                                                      |     |     |
|----------------------|-----------------|------------------------------------------------------------------------------------------------------|-----|-----|
| <i>CTD-2323K18.1</i> | ENSG00000260269 | 29915430                                                                                             | No  | No  |
| <i>SNUPN</i>         | ENSG00000169371 | 34518541, 37164006,<br>38400758, 38723630                                                            | Yes | No  |
| <i>IMP3</i>          | ENSG00000177971 | 30988301, 37164006,<br>38400758                                                                      | Yes | No  |
| <i>CTD-3065B20.3</i> | ENSG00000259314 | 36356581, 36402776                                                                                   | No  | No  |
| <i>FES</i>           | ENSG00000182511 | 34518541                                                                                             | No  | No  |
| <i>MAN2A2</i>        | ENSG00000196547 | 36356581, 36402776,<br>37164006, 38400758,<br>38733992                                               | Yes | No  |
| <i>HDDC3</i>         | ENSG00000184508 | 36356581, 37164006,<br>38400758, 38733992                                                            | Yes | No  |
| <i>UNC45A</i>        | ENSG00000140553 | 34518541, 38733992                                                                                   | No  | No  |
| <i>RCCD1</i>         | ENSG00000166965 | 28362817, 34317694,<br>36356581, 36402776,<br>36690614, 37164006,<br>38400758, 38723630,<br>38733992 | Yes | Yes |
| <i>PRC1-AS1</i>      | ENSG00000258725 | 36356581, 38400758,<br>38733992                                                                      | Yes | Yes |
| <i>PDXDC1</i>        | ENSG00000179889 | 34518541                                                                                             | No  | No  |
| <i>RP11-297L17.6</i> | ENSG00000234337 | 36356581                                                                                             | No  | No  |
| <i>RP11-467J12.4</i> | ENSG00000277639 | 29915430                                                                                             | No  | No  |
| <i>CHD9</i>          | ENSG00000177200 | 34518541                                                                                             | No  | No  |
| <i>RP11-44F14.6</i>  | ENSG00000279722 | 36402776                                                                                             | No  | No  |
| <i>IRX3</i>          | ENSG00000177508 | 34518541                                                                                             | No  | No  |
| <i>CTD-3032H12.1</i> | ENSG00000259725 | 29915430                                                                                             | No  | No  |
| <i>SLC12A3</i>       | ENSG00000070915 | 34518541                                                                                             | No  | No  |
| <i>ZNF19</i>         | ENSG00000157429 | 36402776, 37164006,<br>38733992                                                                      | No  | No  |
| <i>PHLPP2</i>        | ENSG00000040199 | 36402776                                                                                             | No  | No  |
| <i>DHODH</i>         | ENSG00000102967 | 28362817                                                                                             | No  | No  |
| <i>CDYL2</i>         | ENSG00000166446 | 36690614                                                                                             | No  | No  |
| <i>GCSH</i>          | ENSG00000140905 | 36402776                                                                                             | No  | No  |
| <i>MAP2K4</i>        | ENSG00000065559 | 36356581, 37164006                                                                                   | No  | No  |
| <i>TOM1L2</i>        | ENSG00000175662 | 34518541                                                                                             | No  | No  |
| <i>GOSR1</i>         | ENSG00000108587 | 36402776                                                                                             | No  | No  |
| <i>NAGLU</i>         | ENSG00000108784 | 36356581, 37164006,<br>38400758, 38733992                                                            | Yes | No  |
| <i>HSD17B1P1</i>     | ENSG00000108785 | 32115800                                                                                             | No  | No  |
| <i>TUBG2</i>         | ENSG00000037042 | 34317694, 37164006,<br>38400758, 38733992                                                            | Yes | No  |
| <i>LINC00671</i>     | ENSG00000213373 | 29915430                                                                                             | No  | No  |
| <i>ARHGAP27</i>      | ENSG00000159314 | 34317694, 34518541,<br>36402776, 37164006,<br>38733992                                               | No  | No  |

Continuation of Table S 1.12

|                      |                 |                                                                                         |     |     |
|----------------------|-----------------|-----------------------------------------------------------------------------------------|-----|-----|
| <i>PLEKHM1</i>       | ENSG00000225190 | 34317694, 34518541,<br>36356581, 36402776,<br>37164006, 38723630,<br>38733992           | No  | No  |
| <i>LRRC37A4P</i>     | ENSG00000214425 | 32115800, 34317694,<br>36356581, 38400758,<br>38733992                                  | Yes | No  |
| <i>RP11-798G7.8</i>  | ENSG00000266918 | 34317694, 36356581,<br>36402776                                                         | No  | No  |
| <i>DND1P1</i>        | ENSG00000264070 | 34317694, 36356581                                                                      | No  | No  |
| <i>RP11-707O23.1</i> | ENSG00000280022 | 36356581, 36402776                                                                      | No  | No  |
| <i>MAPK8IP1P2</i>    | ENSG00000263503 | 34317694, 36356581                                                                      | No  | No  |
| <i>RPS26P8</i>       | ENSG00000204652 | 34317694, 36356581                                                                      | No  | No  |
| <i>LINC02210</i>     | ENSG00000204650 | 32115800, 34317694,<br>38723630, 38733992                                               | No  | No  |
| <i>RP11-293E1.1</i>  | ENSG00000265964 | 34317694                                                                                | No  | No  |
| <i>CRHR1</i>         | ENSG00000120088 | 29915430, 34317694,<br>36356581, 36402776,<br>37164006, 38733992                        | No  | No  |
| <i>SPPL2C</i>        | ENSG00000185294 | 34317694                                                                                | No  | No  |
| <i>KANSL1</i>        | ENSG00000120071 | 36356581, 36402776,<br>37164006, 38733992                                               | No  | No  |
| <i>KANSL1-AS1</i>    | ENSG00000214401 | 29915430, 36402776,<br>38723630                                                         | No  | No  |
| <i>MAPK8IP1P1</i>    | ENSG00000262500 | 34317694, 36356581                                                                      | No  | No  |
| <i>RP11-259G18.3</i> | ENSG00000262539 | 34317694, 36356581                                                                      | No  | No  |
| <i>RP11-259G18.1</i> | ENSG00000261575 | 36356581, 38723630                                                                      | No  | No  |
| <i>LRRC37A</i>       | ENSG00000176681 | 29915430, 34317694,<br>34518541, 36402776,<br>37164006                                  | No  | No  |
| <i>LRRC37A2</i>      | ENSG00000238083 | 29915430, 34317694,<br>34518541, 36356581,<br>36402776, 37164006                        | No  | No  |
| <i>ARL17A</i>        | ENSG00000185829 | 34317694, 36356581,<br>36402776, 37164006                                               | No  | No  |
| <i>FAM215B</i>       | ENSG00000232300 | 36356581, 36402776,<br>37164006                                                         | No  | No  |
| <i>NSF</i>           | ENSG00000073969 | 36402776, 37164006,<br>38733992                                                         | No  | No  |
| <i>WNT3</i>          | ENSG00000108379 | 36402776, 37164006                                                                      | No  | No  |
| <i>WNT9B</i>         | ENSG00000158955 | 36402776                                                                                | No  | No  |
| <i>COX11</i>         | ENSG00000166260 | 34518541, 37164006,<br>38400758, 38733992                                               | Yes | Yes |
| <i>RP11-257O5.2</i>  | ENSG00000279059 | 36356581, 36402776,<br>38723630                                                         | Yes | No  |
| <i>STXBP4</i>        | ENSG00000166263 | 29915430, 34518541,<br>36356581, 36690614,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | No  |
| <i>SS18</i>          | ENSG00000141380 | 36402776                                                                                | No  | No  |



Continuation of Table S 1.12

|                      |                 |                                                                                         |     |     |
|----------------------|-----------------|-----------------------------------------------------------------------------------------|-----|-----|
| <i>RP11-349H17.2</i> | ENSG00000275805 | 36356581                                                                                | No  | No  |
| <i>ANKLE1</i>        | ENSG00000160117 | 34317694, 34518541,<br>36356581, 37164006,<br>38400758, 38723630,<br>38733992           | Yes | No  |
| <i>ABHD8</i>         | ENSG00000127220 | 34317694, 36356581,<br>37164006, 38400758,<br>38723630                                  | Yes | Yes |
| <i>IL12RB1</i>       | ENSG00000096996 | 36402776                                                                                | No  | No  |
| <i>LRRC25</i>        | ENSG00000175489 | 34518541, 36356581,<br>36402776, 37164006,<br>38400758, 38723630                        | Yes | Yes |
| <i>SSBP4</i>         | ENSG00000130511 | 36356581, 36402776,<br>37164006, 38400758,<br>38723630, 38733992                        | Yes | No  |
| <i>ELL</i>           | ENSG00000105656 | 34518541, 37164006,<br>38400758, 38733992                                               | Yes | No  |
| <i>FKBP8</i>         | ENSG00000105701 | 36356581                                                                                | No  | No  |
| <i>AC005253.2</i>    | ENSG00000268030 | 36402776                                                                                | No  | No  |
| <i>ARMC6</i>         | ENSG00000105676 | 36402776                                                                                | No  | No  |
| <i>MEF2B</i>         | ENSG00000213999 | 36402776                                                                                | No  | No  |
| <i>HAPLN4</i>        | ENSG00000187664 | 29915430, 36402776,<br>37164006, 38723630                                               | No  | No  |
| <i>SUGP1</i>         | ENSG00000105705 | 36402776, 37164006,<br>38733992                                                         | No  | No  |
| <i>GMIP</i>          | ENSG00000089639 | 36356581                                                                                | No  | No  |
| <i>ATP13A1</i>       | ENSG00000105726 | 36402776, 38733992                                                                      | No  | No  |
| <i>ZNF793-AS1</i>    | ENSG00000266916 | 36356581                                                                                | No  | No  |
| <i>PSG4</i>          | ENSG00000243137 | 36690614                                                                                | No  | No  |
| <i>ZNF575</i>        | ENSG00000176472 | 34518541                                                                                | No  | No  |
| <i>KCNN4</i>         | ENSG00000104783 | 34518541, 37164006,<br>38400758                                                         | Yes | Yes |
| <i>ZNF404</i>        | ENSG00000176222 | 29915430, 32115800,<br>34518541, 36356581,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | Yes |
| <i>RP11-15A1.3</i>   | ENSG00000267058 | 36356581, 37164006,<br>38400758, 38723630                                               | Yes | Yes |
| <i>ZNF155</i>        | ENSG00000204920 | 32115800                                                                                | No  | No  |
| <i>ZNF230-DT</i>     | ENSG00000266921 | 29915430, 32115800                                                                      | No  | No  |
| <i>PPP1R37</i>       | ENSG00000104866 | 36402776                                                                                | No  | No  |
| <i>AC007191.4</i>    | ENSG00000279407 | 36356581                                                                                | No  | No  |
| <i>PLA2G4C</i>       | ENSG00000105499 | 34518541                                                                                | No  | No  |
| <i>TP53INP2</i>      | ENSG00000078804 | 28957356                                                                                | No  | No  |
| <i>CPNE1</i>         | ENSG00000214078 | 32115800, 36356581,<br>38733992                                                         | No  | No  |
| <i>RGS19</i>         | ENSG00000171700 | 36356581, 36402776,<br>37164006                                                         | No  | No  |

Continuation of Table S 1.12

|                      |                 |                                                                  |     |     |
|----------------------|-----------------|------------------------------------------------------------------|-----|-----|
| <i>USP25</i>         | ENSG00000155313 | 36402776                                                         | No  | No  |
| <i>CBS</i>           | ENSG00000160200 | 34518541                                                         | No  | No  |
| <i>YBEY</i>          | ENSG00000182362 | 36356581                                                         | No  | No  |
| <i>HSCB</i>          | ENSG00000100209 | 34317694, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>CTA-292E10.8</i>  | ENSG00000272858 | 36356581, 36402776                                               | No  | No  |
| <i>RHBDD3</i>        | ENSG00000100263 | 36402776                                                         | No  | No  |
| <i>LGALS2</i>        | ENSG00000100079 | 36402776                                                         | No  | No  |
| <i>GGA1</i>          | ENSG00000100083 | 36402776                                                         | No  | No  |
| <i>NOL12</i>         | ENSG00000273899 | 34518541                                                         | No  | No  |
| <i>TRIOBP</i>        | ENSG00000100106 | 34518541                                                         | No  | No  |
| <i>PICK1</i>         | ENSG00000100151 | 36402776, 37164006                                               | No  | No  |
| <i>PLA2G6</i>        | ENSG00000184381 | 36356581, 36402776,<br>37164006, 38400758,<br>38723630, 38733992 | Yes | No  |
| <i>MAFF</i>          | ENSG00000185022 | 36402776, 37164006,<br>38400758, 38733992                        | Yes | No  |
| <i>TMEM184B</i>      | ENSG00000198792 | 36356581, 38400758,<br>38723630, 38733992                        | Yes | No  |
| <i>GTPBP1</i>        | ENSG00000100226 | 36402776                                                         | No  | No  |
| <i>NPTXR</i>         | ENSG00000221890 | 36356581                                                         | No  | No  |
| <i>APOBEC3B</i>      | ENSG00000179750 | 34518541, 36356581,<br>37164006, 38400758,<br>38723630           | Yes | Yes |
| <i>XPNPEP3</i>       | ENSG00000196236 | 36402776                                                         | No  | No  |
| <i>ZC3H7B</i>        | ENSG00000100403 | 36402776                                                         | No  | No  |
| <i>TEF</i>           | ENSG00000167074 | 36402776                                                         | No  | No  |
| <i>ATXN10</i>        | ENSG00000130638 | 36356581                                                         | No  | No  |
| <i>PKDREJ</i>        | ENSG00000130943 | 34518541                                                         | No  | No  |
| <i>CORT</i>          | ENSG00000241563 | 37164006, 38733992                                               | No  | No  |
| <i>CITED4</i>        | ENSG00000179862 | 37164006                                                         | No  | No  |
| <i>AKR1A1</i>        | ENSG00000117448 | 37164006                                                         | No  | No  |
| <i>CCDC17</i>        | ENSG00000159588 | 37164006                                                         | No  | No  |
| <i>PIK3R3</i>        | ENSG00000117461 | 37164006, 38733992                                               | No  | No  |
| <i>CDKN2C</i>        | ENSG00000123080 | 37164006                                                         | No  | No  |
| <i>PHTF1</i>         | ENSG00000116793 | 37164006                                                         | No  | No  |
| <i>DCLRE1B</i>       | ENSG00000118655 | 37164006, 38400758,<br>38733992                                  | Yes | No  |
| <i>SEC22B</i>        | ENSG00000265808 | 37164006, 38400758                                               | Yes | No  |
| <i>NBPF26</i>        | ENSG00000273136 | 37164006, 38400758                                               | Yes | Yes |
| <i>LINC00623</i>     | ENSG00000226067 | 37164006, 38400758,<br>38723630, 38733992                        | Yes | No  |
| <i>FCGR1B</i>        | ENSG00000198019 | 37164006, 38400758                                               | Yes | No  |
| <i>RP11-344P13.4</i> | ENSG00000228826 | 37164006, 38400758                                               | Yes | No  |
| <i>RP11-344P13.6</i> | ENSG00000272583 | 37164006, 38400758                                               | Yes | No  |

Continuation of Table S 1.12

|                                    |                 |                                           |     |    |
|------------------------------------|-----------------|-------------------------------------------|-----|----|
| <i>H3-2</i>                        | ENSG00000203818 | 37164006                                  | No  | No |
| <i>FAM72C</i>                      | ENSG00000263513 | 37164006, 38400758                        | Yes | No |
| <i>SLC50A1,</i><br><i>SLC50A1a</i> | ENSG00000169241 | 37164006, 38733992                        | No  | No |
| <i>TRIM46</i>                      | ENSG00000163462 | 37164006, 38400758                        | Yes | No |
| <i>MTX1</i>                        | ENSG00000173171 | 37164006, 38400758,<br>38723630, 38733992 | Yes | No |
| <i>GBA</i>                         | ENSG00000177628 | 37164006, 38400758,<br>38733992           | Yes | No |
| <i>SCAMP3</i>                      | ENSG00000116521 | 37164006, 38400758,<br>38733992           | Yes | No |
| <i>HCN3</i>                        | ENSG00000143630 | 37164006, 38400758,<br>38733992           | Yes | No |
| <i>PKLR</i>                        | ENSG00000143627 | 37164006, 38400758                        | Yes | No |
| <i>FDPS</i>                        | ENSG00000160752 | 37164006, 38400758,<br>38733992           | Yes | No |
| <i>RUSC1</i>                       | ENSG00000160753 | 37164006, 38400758                        | Yes | No |
| <i>YY1AP1</i>                      | ENSG00000163374 | 37164006, 38733992                        | No  | No |
| <i>DAP3</i>                        | ENSG00000132676 | 37164006, 38733992                        | No  | No |
| <i>SYT11</i>                       | ENSG00000132718 | 37164006, 38733992                        | No  | No |
| <i>GPR161</i>                      | ENSG00000143147 | 37164006, 38400758                        | Yes | No |
| <i>ZC3H11A</i>                     | ENSG00000058673 | 37164006, 38400758                        | Yes | No |
| <i>ZBED6</i>                       | ENSG00000257315 | 37164006                                  | No  | No |
| <i>SNRPE</i>                       | ENSG00000182004 | 37164006                                  | No  | No |
| <i>LINC02767</i>                   | ENSG00000284237 | 37164006, 38733992                        | No  | No |
| <i>GRHL1</i>                       | ENSG00000134317 | 37164006, 38733992                        | No  | No |
| <i>ADCY3</i>                       | ENSG00000138031 | 37164006, 38733992                        | No  | No |
| <i>DNAJC27</i>                     | ENSG00000115137 | 37164006                                  | No  | No |
| <i>EFR3B</i>                       | ENSG00000084710 | 37164006                                  | No  | No |
| <i>RALB</i>                        | ENSG00000144118 | 37164006, 38400758,<br>38733992           | Yes | No |
| <i>AC012363.13</i>                 | ENSG00000235840 | 37164006, 38400758                        | Yes | No |
| <i>DLX2</i>                        | ENSG00000115844 | 37164006, 38400758                        | Yes | No |
| <i>CDCA7</i>                       | ENSG00000144354 | 37164006, 38400758                        | Yes | No |
| <i>BZW1</i>                        | ENSG00000082153 | 37164006, 38733992                        | No  | No |
| <i>CFLAR</i>                       | ENSG00000003402 | 37164006                                  | No  | No |
| <i>DIRC3</i>                       | ENSG00000231672 | 37164006, 38400758                        | Yes | No |
| <i>NEK10</i>                       | ENSG00000163491 | 37164006, 38400758,<br>38733992           | Yes | No |
| <i>TMIE</i>                        | ENSG00000181585 | 37164006                                  | No  | No |
| <i>PTH1R</i>                       | ENSG00000160801 | 37164006                                  | No  | No |
| <i>ATXN7</i>                       | ENSG00000163635 | 37164006                                  | No  | No |
| <i>LINC00994</i>                   | ENSG00000189196 | 37164006, 38733992                        | No  | No |
| <i>FILIP1L</i>                     | ENSG00000168386 | 37164006, 38400758                        | Yes | No |
| <i>TOMM70</i>                      | ENSG00000154174 | 37164006                                  | No  | No |

Continuation of Table S 1.12

|                     |                 |                                 |     |    |
|---------------------|-----------------|---------------------------------|-----|----|
| <i>ZBTB38</i>       | ENSG00000177311 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>RASA2</i>        | ENSG00000155903 | 37164006                        | No  | No |
| <i>TLR10</i>        | ENSG00000174123 | 37164006, 38733992              | No  | No |
| <i>TLR1</i>         | ENSG00000174125 | 37164006, 38733992              | No  | No |
| <i>TLR6</i>         | ENSG00000174130 | 37164006, 38733992              | No  | No |
| <i>FAM114A1</i>     | ENSG00000197712 | 37164006, 38733992              | No  | No |
| <i>ABRAXAS1</i>     | ENSG00000163322 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>PKD2</i>         | ENSG00000118762 | 37164006, 38733992              | No  | No |
| <i>PPM1K</i>        | ENSG00000163644 | 37164006                        | No  | No |
| <i>AHRR</i>         | ENSG00000063438 | 37164006, 38733992              | No  | No |
| <i>TERT</i>         | ENSG00000164362 | 37164006, 38400758              | Yes | No |
| <i>NNT</i>          | ENSG00000112992 | 37164006                        | No  | No |
| <i>RP11-53O19.3</i> | ENSG00000272335 | 37164006, 38400758              | Yes | No |
| <i>MAP3K1</i>       | ENSG00000095015 | 37164006, 38400758              | Yes | No |
| <i>MIER3</i>        | ENSG00000155545 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>ARRDC3-AS1</i>   | ENSG00000281357 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>PDLIM4</i>       | ENSG00000131435 | 37164006, 38733992              | No  | No |
| <i>SLC22A4</i>      | ENSG00000197208 | 37164006, 38733992              | No  | No |
| <i>HSPA4</i>        | ENSG00000170606 | 37164006, 38733992              | No  | No |
| <i>RANBP9</i>       | ENSG00000010017 | 37164006, 38400758              | Yes | No |
| <i>H2AC14</i>       | ENSG00000276368 | 37164006                        | No  | No |
| <i>H2BC15</i>       | ENSG00000233822 | 37164006                        | No  | No |
| <i>H1-5</i>         | ENSG00000184357 | 37164006                        | No  | No |
| <i>ZKSCAN4</i>      | ENSG00000187626 | 37164006                        | No  | No |
| <i>ZKSCAN3</i>      | ENSG00000189298 | 37164006                        | No  | No |
| <i>ZSCAN12</i>      | ENSG00000158691 | 37164006, 38733992              | No  | No |
| <i>TRIM27</i>       | ENSG00000204713 | 37164006, 38733992              | No  | No |
| <i>LINC02829</i>    | ENSG00000229274 | 37164006                        | No  | No |
| <i>FLOT1</i>        | ENSG00000137312 | 37164006                        | No  | No |
| <i>RP11-250B2.3</i> | ENSG00000233967 | 37164006, 38723630              | No  | No |
| <i>RSPO3</i>        | ENSG00000146374 | 37164006                        | No  | No |
| <i>TXLNB</i>        | ENSG00000164440 | 37164006, 38733992              | No  | No |
| <i>CYP51A1</i>      | ENSG00000001630 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>PEG10</i>        | ENSG00000242265 | 37164006, 38400758              | Yes | No |
| <i>BAIAP2L1</i>     | ENSG00000006453 | 37164006                        | No  | No |
| <i>SPDYE3</i>       | ENSG00000214300 | 37164006                        | No  | No |
| <i>PILRB</i>        | ENSG00000121716 | 37164006, 38733992              | No  | No |
| <i>PILRA</i>        | ENSG00000085514 | 37164006, 38733992              | No  | No |

Continuation of Table S 1.12

|                      |                 |                                           |     |     |
|----------------------|-----------------|-------------------------------------------|-----|-----|
| <i>ZCWPW1</i>        | ENSG00000078487 | 37164006, 38733992                        | No  | No  |
| <i>MEPCE</i>         | ENSG00000146834 | 37164006                                  | No  | No  |
| <i>C7orf61</i>       | ENSG00000185955 | 37164006                                  | No  | No  |
| <i>TSC22D4</i>       | ENSG00000166925 | 37164006, 38733992                        | No  | No  |
| <i>NYAP1</i>         | ENSG00000166924 | 37164006, 38733992                        | No  | No  |
| <i>POLR2J</i>        | ENSG00000005075 | 37164006, 38733992                        | No  | No  |
| <i>RP11-486M23.2</i> | ENSG00000253182 | 37164006, 38400758                        | Yes | No  |
| <i>RP11-419C23.1</i> | ENSG00000254038 | 37164006, 38400758                        | Yes | No  |
| <i>HNFG4G</i>        | ENSG00000164749 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>LMX1B</i>         | ENSG00000136944 | 37164006                                  | No  | No  |
| <i>ABO</i>           | ENSG00000175164 | 37164006, 38733992                        | No  | No  |
| <i>MIR1915HG</i>     | ENSG00000204682 | 37164006, 38400758                        | Yes | No  |
| <i>SKIDA1</i>        | ENSG00000180592 | 37164006, 38400758                        | Yes | Yes |
| <i>YME1L1</i>        | ENSG00000136758 | 37164006                                  | No  | No  |
| <i>ADO</i>           | ENSG00000181915 | 37164006                                  | No  | No  |
| <i>FGFR2</i>         | ENSG00000066468 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>GATD1</i>         | ENSG00000177225 | 37164006, 38733992                        | No  | No  |
| <i>CEND1</i>         | ENSG00000184524 | 37164006                                  | No  | No  |
| <i>PANO1</i>         | ENSG00000274897 | 37164006                                  | No  | No  |
| <i>PNPLA2</i>        | ENSG00000177666 | 37164006, 38733992                        | No  | No  |
| <i>CTSD</i>          | ENSG00000117984 | 37164006                                  | No  | No  |
| <i>LSP1</i>          | ENSG00000130592 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>TNNT3</i>         | ENSG00000130595 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>IGF2</i>          | ENSG00000167244 | 37164006, 38400758                        | Yes | No  |
| <i>AP5B1</i>         | ENSG00000254470 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>OVOL1</i>         | ENSG00000172818 | 37164006, 38400758                        | Yes | No  |
| <i>CFL1</i>          | ENSG00000172757 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>MUS81</i>         | ENSG00000172732 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>EFEMP2</i>        | ENSG00000172638 | 37164006, 38400758,<br>38723630, 38733992 | Yes | No  |
| <i>FOSL1</i>         | ENSG00000175592 | 37164006, 38400758                        | Yes | No  |
| <i>DPP3</i>          | ENSG00000254986 | 37164006, 38733992                        | No  | No  |
| <i>RP11-554A11.8</i> | ENSG00000261070 | 37164006                                  | No  | No  |
| <i>RP11-108O10.2</i> | ENSG00000254990 | 37164006, 38733992                        | No  | No  |
| <i>FDXACB1</i>       | ENSG00000255561 | 37164006                                  | No  | No  |
| <i>HSPB2</i>         | ENSG00000170276 | 37164006                                  | No  | No  |
| <i>RP11-237N19.3</i> | ENSG00000281386 | 37164006, 38400758                        | Yes | No  |
| <i>PTHLH</i>         | ENSG00000087494 | 37164006, 38400758                        | Yes | Yes |

Continuation of Table S 1.12

|                      |                 |                                           |     |     |
|----------------------|-----------------|-------------------------------------------|-----|-----|
| <i>CCDC91</i>        | ENSG00000123106 | 37164006, 38400758,<br>38723630, 38733992 | Yes | Yes |
| <i>OVCH1</i>         | ENSG00000187950 | 37164006                                  | No  | No  |
| <i>PRIM1</i>         | ENSG00000198056 | 37164006                                  | No  | No  |
| <i>HSD17B6</i>       | ENSG00000025423 | 37164006                                  | No  | No  |
| <i>NDUFA4L2</i>      | ENSG00000185633 | 37164006                                  | No  | No  |
| <i>R3HDM2</i>        | ENSG00000179912 | 37164006                                  | No  | No  |
| <i>NTN4</i>          | ENSG00000074527 | 37164006, 38400758                        | Yes | Yes |
| <i>RP11-116D17.4</i> | ENSG00000276308 | 37164006, 38400758                        | Yes | No  |
| <i>PAX9</i>          | ENSG00000198807 | 37164006, 38400758                        | Yes | No  |
| <i>IRF2BPL</i>       | ENSG00000119669 | 37164006                                  | No  | No  |
| <i>ADSS1</i>         | ENSG00000185100 | 37164006, 38733992                        | No  | No  |
| <i>SIVA1</i>         | ENSG00000184990 | 37164006                                  | No  | No  |
| <i>EDC3</i>          | ENSG00000179151 | 37164006, 38400758                        | Yes | No  |
| <i>CYP1A1</i>        | ENSG00000140465 | 37164006, 38400758                        | Yes | No  |
| <i>CSK</i>           | ENSG00000103653 | 37164006, 38400758                        | Yes | No  |
| <i>LMAN1L</i>        | ENSG00000140506 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>CPLX3</i>         | ENSG00000213578 | 37164006, 38400758                        | Yes | No  |
| <i>FURIN</i>         | ENSG00000140564 | 37164006                                  | No  | No  |
| <i>PRC1</i>          | ENSG00000198901 | 37164006, 38400758,<br>38733992           | Yes | No  |
| <i>TOX3</i>          | ENSG00000103460 | 37164006, 38400758                        | Yes | Yes |
| <i>RP11-297L17.2</i> | ENSG00000260963 | 37164006, 38400758                        | Yes | No  |
| <i>ZNF23</i>         | ENSG00000167377 | 37164006, 38733992                        | No  | No  |
| <i>TBC1D29P</i>      | ENSG00000266733 | 37164006                                  | No  | No  |
| <i>CRLF3</i>         | ENSG00000176390 | 37164006, 38733992                        | No  | No  |
| <i>CTD-2349P21.9</i> | ENSG00000266490 | 37164006                                  | No  | No  |
| <i>TEFM</i>          | ENSG00000172171 | 37164006                                  | No  | No  |
| <i>HSD17B1</i>       | ENSG00000108786 | 37164006, 38400758                        | Yes | No  |
| <i>MLX</i>           | ENSG00000108788 | 37164006                                  | No  | No  |
| <i>CNTNAP1</i>       | ENSG00000108797 | 37164006, 38400758                        | Yes | No  |
| <i>AOC2</i>          | ENSG00000131480 | 37164006                                  | No  | No  |
| <i>MAPT</i>          | ENSG00000186868 | 37164006, 38733992                        | No  | No  |
| <i>TOM1L1</i>        | ENSG00000141198 | 37164006, 38400758,<br>38733992           | Yes | Yes |
| <i>RPS6KB1</i>       | ENSG00000108443 | 37164006                                  | No  | No  |
| <i>CBX8</i>          | ENSG00000141570 | 37164006, 38723630                        | No  | No  |
| <i>LINC01977</i>     | ENSG00000262772 | 37164006                                  | No  | No  |
| <i>KCTD1</i>         | ENSG00000134504 | 37164006                                  | No  | No  |
| <i>GCDH</i>          | ENSG00000105607 | 37164006, 38733992                        | No  | No  |
| <i>SYCE2</i>         | ENSG00000161860 | 37164006                                  | No  | No  |
| <i>FARSA</i>         | ENSG00000179115 | 37164006, 38400758                        | Yes | No  |

Continuation of Table S 1.12

|                   |                 |                                 |     |    |
|-------------------|-----------------|---------------------------------|-----|----|
| <i>CALR</i>       | ENSG00000179218 | 37164006, 38400758              | Yes | No |
| <i>RAD23A</i>     | ENSG00000179262 | 37164006                        | No  | No |
| <i>LYL1</i>       | ENSG00000104903 | 37164006, 38400758              | Yes | No |
| <i>SAMD1</i>      | ENSG00000141858 | 37164006                        | No  | No |
| <i>MYO9B</i>      | ENSG00000099331 | 37164006, 38400758              | Yes | No |
| <i>BABAM1</i>     | ENSG00000105393 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>MRPL34</i>     | ENSG00000130312 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>ISYNA1</i>     | ENSG00000105655 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>TM6SF2</i>     | ENSG00000213996 | 37164006                        | No  | No |
| <i>MAU2</i>       | ENSG00000129933 | 37164006, 38733992              | No  | No |
| <i>GATAD2A</i>    | ENSG00000167491 | 37164006, 38733992              | No  | No |
| <i>NDUFA13</i>    | ENSG00000186010 | 37164006                        | No  | No |
| <i>YJEFN3</i>     | ENSG00000250067 | 37164006                        | No  | No |
| <i>LYPD5</i>      | ENSG00000159871 | 37164006, 38400758              | Yes | No |
| <i>AC115522.3</i> | ENSG00000268601 | 37164006, 38400758              | Yes | No |
| <i>ZNF283</i>     | ENSG00000167637 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>GIPR</i>       | ENSG00000010310 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>FBXO46</i>     | ENSG00000177051 | 37164006                        | No  | No |
| <i>PHF20</i>      | ENSG00000025293 | 37164006                        | No  | No |
| <i>SCAND1</i>     | ENSG00000171222 | 37164006                        | No  | No |
| <i>CNBD2</i>      | ENSG00000149646 | 37164006                        | No  | No |
| <i>LINC01271</i>  | ENSG00000233077 | 37164006                        | No  | No |
| <i>OPRL1</i>      | ENSG00000125510 | 37164006                        | No  | No |
| <i>TTC28</i>      | ENSG00000100154 | 37164006, 38400758              | Yes | No |
| <i>CHEK2</i>      | ENSG00000183765 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>CCDC117</i>    | ENSG00000159873 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>XBP1</i>       | ENSG00000100219 | 37164006, 38400758              | Yes | No |
| <i>KREMEN1</i>    | ENSG00000183762 | 37164006, 38400758              | Yes | No |
| <i>C22orf23</i>   | ENSG00000128346 | 37164006                        | No  | No |
| <i>BAIAP2L2</i>   | ENSG00000128298 | 37164006, 38400758              | Yes | No |
| <i>CBX6</i>       | ENSG00000183741 | 37164006, 38400758              | Yes | No |
| <i>APOBEC3A</i>   | ENSG00000128383 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>ADSL</i>       | ENSG00000239900 | 37164006, 38400758              | Yes | No |
| <i>SGSM3</i>      | ENSG00000100359 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>MRTFA</i>      | ENSG00000196588 | 37164006, 38400758,<br>38733992 | Yes | No |
| <i>SREBF2</i>     | ENSG00000198911 | 37164006                        | No  | No |

Continuation of Table S 1.12

|                      |                 |                    |     |     |
|----------------------|-----------------|--------------------|-----|-----|
| <i>CCDC163</i>       | ENSG00000280670 | 38733992           | No  | No  |
| <i>CD34</i>          | ENSG00000174059 | 38733992           | No  | No  |
| <i>DNMT3A</i>        | ENSG00000119772 | 38733992           | No  | No  |
| <i>CLK1</i>          | ENSG00000013441 | 38733992           | No  | No  |
| <i>ITPR1</i>         | ENSG00000150995 | 38400758, 38733992 | Yes | No  |
| <i>NELFA</i>         | ENSG00000185049 | 38733992           | No  | No  |
| <i>MRPS18C</i>       | ENSG00000163319 | 38733992           | No  | No  |
| <i>CDKAL1</i>        | ENSG00000145996 | 38733992           | No  | No  |
| <i>BTN2A1</i>        | ENSG00000112763 | 38733992           | No  | No  |
| <i>ABT1</i>          | ENSG00000146109 | 38733992           | No  | No  |
| <i>ZSCAN9</i>        | ENSG00000137185 | 38733992           | No  | No  |
| <i>SYNE1</i>         | ENSG00000131018 | 38733992           | No  | No  |
| <i>BRI3</i>          | ENSG00000164713 | 38733992           | No  | No  |
| <i>MLLT10</i>        | ENSG00000078403 | 38400758, 38733992 | Yes | No  |
| <i>ALG9</i>          | ENSG00000086848 | 38733992           | No  | No  |
| <i>USP28</i>         | ENSG00000048028 | 38733992           | No  | No  |
| <i>RAD51B</i>        | ENSG00000182185 | 38400758, 38733992 | Yes | Yes |
| <i>PPP4R3A</i>       | ENSG00000100796 | 38733992           | No  | No  |
| <i>SMAD3</i>         | ENSG00000166949 | 38400758, 38733992 | Yes | No  |
| <i>SUZ12P1</i>       | ENSG00000264538 | 38733992           | No  | No  |
| <i>ATAD5</i>         | ENSG00000176208 | 38733992           | No  | No  |
| <i>PLEKHH3</i>       | ENSG00000068137 | 38400758, 38733992 | Yes | No  |
| <i>TRMT1</i>         | ENSG00000104907 | 38400758, 38733992 | Yes | No  |
| <i>QPCTL</i>         | ENSG00000011478 | 38733992           | No  | No  |
| <i>BCAS4</i>         | ENSG00000124243 | 38733992           | No  | No  |
| <i>CTA-292E10.6</i>  | ENSG00000226471 | 38733992, 38400758 | Yes | No  |
| <i>AFF1a</i>         | ENSG00000172493 | 38733992           | No  | No  |
| <i>RPL23AP53a</i>    | ENSG00000223508 | 38733992           | No  | No  |
| <i>PCM1a</i>         | ENSG00000078674 | 38733992           | No  | No  |
| <i>SRP54a</i>        | ENSG00000100883 | 38733992           | No  | No  |
| <i>PRORPa</i>        | ENSG00000100890 | 38733992           | No  | No  |
| <i>NEIL1a, NEIL1</i> | ENSG00000140398 | 38733992, 38400758 | Yes | No  |
| <i>EZH1a, EZH1</i>   | ENSG00000108799 | 38733992, 38400758 | Yes | No  |
| <i>PARK7</i>         | ENSG00000116288 | 38733992           | No  | No  |
| <i>ERRFI1</i>        | ENSG00000116285 | 38733992           | No  | No  |
| <i>TESK2</i>         | ENSG00000070759 | 38733992           | No  | No  |
| <i>MAST2</i>         | ENSG00000086015 | 38733992           | No  | No  |
| <i>FAAH</i>          | ENSG00000117480 | 38400758, 38733992 | Yes | No  |
| <i>PRSS44P</i>       | ENSG00000226074 | 38733992           | No  | No  |
| <i>MYL3</i>          | ENSG00000160808 | 38733992           | No  | No  |
| <i>THOC7</i>         | ENSG00000163634 | 38733992           | No  | No  |



Continuation of Table S 1.12

|                        |                 |                    |     |     |
|------------------------|-----------------|--------------------|-----|-----|
| <i>NIT2</i>            | ENSG00000114021 | 38733992           | No  | No  |
| <i>HELQ</i>            | ENSG00000163312 | 38400758, 38733992 | Yes | No  |
| <i>ARRDC3</i>          | ENSG00000113369 | 38400758, 38733992 | Yes | No  |
| <i>BTN3A1</i>          | ENSG00000026950 | 38733992           | No  | No  |
| <i>PGBD1</i>           | ENSG00000137338 | 38733992           | No  | No  |
| <i>ZSCAN23</i>         | ENSG00000187987 | 38733992           | No  | No  |
| <i>ANKIB1</i>          | ENSG00000001629 | 38733992           | No  | No  |
| <i>TECPR1</i>          | ENSG00000205356 | 38733992           | No  | No  |
| <i>TRIM4</i>           | ENSG00000146833 | 38733992           | No  | No  |
| <i>GJC3</i>            | ENSG00000176402 | 38733992           | No  | No  |
| <i>AZGP1</i>           | ENSG00000160862 | 38733992           | No  | No  |
| <i>PMS2P1</i>          | ENSG00000078319 | 38733992           | No  | No  |
| <i>STAG3L5P</i>        | ENSG00000242294 | 38733992           | No  | No  |
| <i>LRWD1</i>           | ENSG00000161036 | 38733992           | No  | No  |
| <i>RNF19A</i>          | ENSG00000034677 | 38733992           | No  | No  |
| <i>FBXO32</i>          | ENSG00000156804 | 38733992           | No  | No  |
| <i>CACFD1</i>          | ENSG00000160325 | 38733992           | No  | No  |
| <i>PIP4K2A</i>         | ENSG00000150867 | 38733992           | No  | No  |
| <i>EGR2</i>            | ENSG00000122877 | 38400758, 38733992 | Yes | No  |
| <i>H19</i>             | ENSG00000130600 | 38400758, 38733992 | Yes | No  |
| <i>CCDC88C</i>         | ENSG00000015133 | 38400758, 38733992 | Yes | Yes |
| <i>PTPN9</i>           | ENSG00000169410 | 38400758, 38733992 | Yes | No  |
| <i>VPS33B</i>          | ENSG00000184056 | 38400758, 38733992 | Yes | No  |
| <i>ZNF18</i>           | ENSG00000154957 | 38733992           | No  | No  |
| <i>ATP6V0A1</i>        | ENSG00000033627 | 38733992           | No  | No  |
| <i>CCR10</i>           | ENSG00000184451 | 38400758, 38733992 | Yes | No  |
| <i>NACC1</i>           | ENSG00000160877 | 38400758, 38733992 | Yes | No  |
| <i>USHBP1</i>          | ENSG00000130307 | 38400758, 38733992 | Yes | No  |
| <i>ZNF45</i>           | ENSG00000124459 | 38400758, 38733992 | Yes | No  |
| <i>CBX7</i>            | ENSG00000100307 | 38400758, 38733992 | Yes | No  |
| <i>FCGR1CP</i>         | ENSG00000265531 | 38400758, 38733992 | Yes | No  |
| <i>RP13-577H12.2</i>   | ENSG00000250431 | 38733992, 38400758 | Yes | No  |
| <i>EBF1</i>            | ENSG00000164330 | 38400758, 38733992 | Yes | No  |
| <i>ENSG00000244349</i> | ENSG00000244349 | 38733992           | No  | No  |
| <i>TENT5A</i>          | ENSG00000112773 | 38400758, 38733992 | Yes | No  |
| <i>ENSG00000226571</i> | ENSG00000226571 | 38733992           | No  | No  |
| <i>FBXL13</i>          | ENSG00000161040 | 38400758, 38733992 | Yes | No  |
| <i>LCN1P1</i>          | ENSG00000119440 | 38733992           | No  | No  |
| <i>ENSG00000273767</i> | ENSG00000273767 | 38400758, 38733992 | Yes | No  |
| <i>AC051649.12</i>     | ENSG00000249086 | 38733992, 38400758 | Yes | No  |

Continuation of Table S 1.12

|                           |                 |                    |     |     |
|---------------------------|-----------------|--------------------|-----|-----|
| <i>CTSW</i>               | ENSG00000172543 | 38733992           | No  | No  |
| <i>SIRT4</i>              | ENSG00000089163 | 38733992           | No  | No  |
| <i>SCAMP2</i>             | ENSG00000140497 | 38400758, 38733992 | Yes | No  |
| <i>FTO</i>                | ENSG00000140718 | 38733992           | No  | No  |
| <i>CHRNA1</i>             | ENSG00000170175 | 38733992           | No  | No  |
| <i>CHST9</i>              | ENSG00000154080 | 38400758, 38733992 | Yes | No  |
| <i>RP11-84A19.4</i>       | ENSG00000269967 | 38400758           | Yes | No  |
| <i>PTP4A2</i>             | ENSG00000184007 | 38400758           | Yes | No  |
| <i>RAD54L</i>             | ENSG00000085999 | 38400758           | Yes | No  |
| <i>HMGCS2</i>             | ENSG00000134240 | 38400758           | Yes | No  |
| <i>PDE4DIP2</i>           | ENSG00000275131 | 38400758           | Yes | No  |
| <i>DLX1</i>               | ENSG00000144355 | 38400758           | Yes | No  |
| <i>TGFBR2</i>             | ENSG00000163513 | 38400758           | Yes | No  |
| <i>ADH5</i>               | ENSG00000197894 | 38400758           | Yes | No  |
| <i>SIRT5</i>              | ENSG00000124523 | 38400758           | Yes | Yes |
| <i>SERINC1</i>            | ENSG00000111897 | 38400758           | Yes | No  |
| <i>PRKRIP1</i>            | ENSG00000128563 | 38400758           | Yes | No  |
| <i>ALKBH4</i>             | ENSG00000160993 | 38400758           | Yes | No  |
| <i>ZFPM2</i>              | ENSG00000169946 | 38400758           | Yes | No  |
| <i>ZMIZ1</i>              | ENSG00000108175 | 38400758           | Yes | No  |
| <i>TLCD5,<br/>TMEM136</i> | ENSG00000181264 | 38400758, 38723630 | Yes | No  |
| <i>ARHGEF12</i>           | ENSG00000196914 | 38400758           | Yes | No  |
| <i>CAVIN1</i>             | ENSG00000177469 | 38400758           | Yes | No  |
| <i>RETREG3</i>            | ENSG00000141699 | 38400758           | Yes | No  |
| <i>TUBG1</i>              | ENSG00000131462 | 38400758           | Yes | No  |
| <i>LGALS1</i>             | ENSG00000100097 | 38400758           | Yes | No  |
| <i>BOD1L1</i>             | ENSG00000038219 | 38400758, 38723630 | Yes | No  |
| <i>LGR6</i>               | ENSG00000133067 | 38400758           | Yes | Yes |
| <i>PIK3C2B</i>            | ENSG00000133056 | 38400758           | Yes | No  |
| <i>WDR43</i>              | ENSG00000163811 | 38400758           | Yes | No  |
| <i>SLC39A7</i>            | ENSG00000112473 | 38400758           | Yes | No  |
| <i>RPS18</i>              | ENSG00000231500 | 38400758, 38723630 | Yes | No  |
| <i>B3GALT4</i>            | ENSG00000235863 | 38400758           | Yes | No  |
| <i>C11orf65</i>           | ENSG00000166323 | 38400758           | Yes | No  |
| <i>USE1</i>               | ENSG00000053501 | 38400758           | Yes | No  |
| <i>OCEL1</i>              | ENSG00000099330 | 38400758, 38723630 | Yes | No  |
| <i>NR2F6</i>              | ENSG00000160113 | 38400758           | Yes | No  |
| <i>DDA1</i>               | ENSG00000130311 | 38400758           | Yes | No  |
| <i>ANO8</i>               | ENSG00000074855 | 38400758           | Yes | No  |
| <i>GTPBP3</i>             | ENSG00000130299 | 38400758           | Yes | No  |
| <i>AC093110.3</i>         | ENSG00000238018 | 38723630           | Yes | No  |

Continuation of Table S 1.12

|                       |                 |          |     |     |
|-----------------------|-----------------|----------|-----|-----|
| <i>ACAP3</i>          | ENSG00000131584 | 38723630 | Yes | No  |
| <i>ACSBG1</i>         | ENSG00000103740 | 38723630 | Yes | No  |
| <i>AGTPBP1</i>        | ENSG00000135049 | 38723630 | No  | No  |
| <i>BANF1</i>          | ENSG00000175334 | 38723630 | No  | No  |
| <i>BLACAT1</i>        | ENSG00000281406 | 38723630 | Yes | No  |
| <i>CH17-118O6.6</i>   | ENSG00000281741 | 38723630 | Yes | No  |
| <i>CTD-3157E16.1</i>  | ENSG00000265519 | 38723630 | No  | No  |
| <i>CYP26B1</i>        | ENSG00000003137 | 38723630 | Yes | No  |
| <i>EEF1A1P16</i>      | ENSG00000213235 | 38723630 | Yes | No  |
| <i>EFR3A</i>          | ENSG00000132294 | 38723630 | Yes | No  |
| <i>FNDCC11</i>        | ENSG00000125531 | 38723630 | Yes | No  |
| <i>GLRA3</i>          | ENSG00000145451 | 38723630 | Yes | Yes |
| <i>HGS</i>            | ENSG00000185359 | 38723630 | Yes | No  |
| <i>IL2RA</i>          | ENSG00000134460 | 38723630 | No  | No  |
| <i>ITGA10</i>         | ENSG00000143127 | 38723630 | Yes | No  |
| <i>KLF7-IT1</i>       | ENSG00000237892 | 38723630 | No  | No  |
| <i>LINC00683</i>      | ENSG00000266256 | 38723630 | No  | No  |
| <i>LINC01266</i>      | ENSG00000224957 | 38723630 | Yes | No  |
| <i>MARCKS</i>         | ENSG00000277443 | 38723630 | No  | No  |
| <i>NPM2</i>           | ENSG00000158806 | 38723630 | Yes | No  |
| <i>NUDT1</i>          | ENSG00000106268 | 38723630 | Yes | No  |
| <i>PPIF</i>           | ENSG00000108179 | 38723630 | Yes | No  |
| <i>RNF20</i>          | ENSG00000155827 | 38723630 | Yes | No  |
| <i>RNVU1-19</i>       | ENSG00000275538 | 38723630 | Yes | Yes |
| <i>RP11-165A20.3</i>  | ENSG00000273107 | 38723630 | No  | No  |
| <i>RP11-303E16.3</i>  | ENSG00000260213 | 38723630 | Yes | No  |
| <i>RP11-307L14.1</i>  | ENSG00000272370 | 38723630 | Yes | No  |
| <i>RP11-416N2.4</i>   | ENSG00000273108 | 38723630 | Yes | No  |
| <i>RP11-459E5.1</i>   | ENSG00000253125 | 38723630 | Yes | No  |
| <i>RP11-467L19.16</i> | ENSG00000278626 | 38723630 | No  | No  |
| <i>RP11-723D22.3</i>  | ENSG00000253829 | 38723630 | Yes | No  |
| <i>RP11-848P1.3</i>   | ENSG00000265743 | 38723630 | Yes | No  |
| <i>RP13-20L14.4</i>   | ENSG00000264812 | 38723630 | No  | No  |
| <i>RP4-621F18.2</i>   | ENSG00000282057 | 38723630 | Yes | No  |
| <i>RPAP2</i>          | ENSG00000122484 | 38723630 | No  | No  |
| <i>RPL3P4</i>         | ENSG00000232573 | 38723630 | Yes | No  |
| <i>RUSC1-AS1</i>      | ENSG00000225855 | 38723630 | Yes | No  |
| <i>TMEM50A</i>        | ENSG00000183726 | 38723630 | No  | No  |
| <i>CTD-3080P12.3</i>  | ENSG00000249201 | 38697998 | No  | No  |
| <i>TET2</i>           | ENSG00000168769 | 38697998 | No  | No  |
| <i>NUP210L</i>        | ENSG00000143552 | 38697998 | No  | No  |

Continuation of Table S 1.12

|                      |                 |          |     |     |
|----------------------|-----------------|----------|-----|-----|
| <i>AC003090.1</i>    | ENSG00000223561 | 38723630 | Yes | No  |
| <i>CTD-2554C21.1</i> | ENSG00000267422 | 38723630 | Yes | No  |
| <i>SLC39A13</i>      | ENSG00000165915 | 38723630 | Yes | No  |
| <i>VPS39</i>         | ENSG00000166887 | 38723630 | Yes | No  |
| <i>RP11-474P2.6</i>  | ENSG00000275481 | 38723630 | Yes | No  |
| <i>EN1</i>           | ENSG00000163064 | 38697998 | Yes | Yes |
| <i>LINC01956</i>     | ENSG00000258910 | 38697998 | Yes | No  |
| <i>BRD9</i>          | ENSG00000028310 | 38697998 | Yes | No  |

Table S 1.13 TWAS associations for the closest gene by distance to each novel GWAS loci identified for intrinsic subtypes.

| <b>Loci</b>     | <b>rsID</b> | <b>Nearby gene</b> | <b>Gene region</b> | <b>Strongest TWAS Association Z-score (P-value)<sup>a</sup></b> |
|-----------------|-------------|--------------------|--------------------|-----------------------------------------------------------------|
| Luminal-A-like  |             |                    |                    |                                                                 |
| 1q41            | rs4011894   | <i>MARK1</i>       | Intronic           | 0.84 (0.402)                                                    |
| 10p15.1         | rs55913768  | <i>FAM208B</i>     | Intronic           | 0.86 (0.389)                                                    |
| 18q12.1         | rs16963218  | <i>GAREM1</i>      | Intronic           | -2.36 (0.018)                                                   |
| Luminal-B-like  |             |                    |                    |                                                                 |
| 10q26.13        | rs145680477 | <i>CPXM2</i>       | Intronic           | 2.05 (0.041)                                                    |
| 12q23.1         | rs4762310   | <i>CFAP54</i>      | Intronic           | -1.65 (0.099)                                                   |
| Triple-negative |             |                    |                    |                                                                 |
| 1p36.22         | rs112200610 | <i>MIR34AHG</i>    | Intronic           | 1.21 (0.227)                                                    |
| 11q25           | rs6590696   | <i>OPCML</i>       | Intronic           | 2.37 (0.018)                                                    |
| 2p24.1          | rs12710697  | <i>LINC01376</i>   | Intronic           | 4.53 (5.95E-06)                                                 |

**Footnote:**

<sup>a</sup> Strongest association is defined as the TWAS association with the minimum p-value across all ancestries and cell types

Table S 1.14 Suggestive TWAS associations linked to novel GWAS loci identified for intrinsic subtypes.

| Loci            | Lead Variant | Gene Symbol      | ENSG ID         | Strongest TWAS Association Z-score (P-value) <sup>a</sup> | Strongest Cond-TWAS Association Z-score (P-value) <sup>b</sup> | Cell Types <sup>c</sup> | Ancestries <sup>d</sup> | Max. r <sup>2</sup> with model variants <sup>e</sup> |
|-----------------|--------------|------------------|-----------------|-----------------------------------------------------------|----------------------------------------------------------------|-------------------------|-------------------------|------------------------------------------------------|
| Luminal-A-like  |              |                  |                 |                                                           |                                                                |                         |                         |                                                      |
| 1q41            | rs4011894    | <i>C1orf115</i>  | ENSG00000162817 | 3.46 (5.41E-04)                                           | -1.08 (0.282)                                                  | Endo                    | E                       | 0.388                                                |
| 10p15.1         | rs55913768   | <i>GDI2</i>      | ENSG00000057608 | -5.46 (4.83E-08)                                          | -1.15 (0.248)                                                  | Bulk, Epi               | C, E                    | 1.000                                                |
| Triple-negative |              |                  |                 |                                                           |                                                                |                         |                         |                                                      |
| 2p24.1          | rs12710697   | <i>LINC01376</i> | ENSG00000236204 | 4.53 (5.95E-06)                                           | 2.79 (5.32E-03)                                                | Bulk, Epi               | C, E                    | 0.226                                                |

**Footnote:**

Suggestive TWAS associations were defined as: minimum p-value across all ancestries and cell types  $<1 \times 10^{-3}$ , GWAS lead variant was within +/- 2Mb of gene transcription start and stop sites, at least one prediction model variant was in LD with the GWAS lead variant ( $r^2 > 0.1$ ), cond-TWAS signal was attenuated ( $p > 1 \times 10^{-3}$ )

<sup>a</sup> Strongest association is defined as the TWAS association with the minimum p-value across all ancestries and cell types

<sup>b</sup> Strongest conditional TWAS (Cond-TWAS) association is defined as the cond-TWAS association with the minimum p-value across all ancestries and cell types

<sup>c</sup> Cell type abbreviations: Bulk (Bulk-tissue), Epi (Epithelial cells), Adip (Adipocytes), Endo (Endothelial cells), Fib (Fibroblasts)

<sup>d</sup> Ancestry abbreviations: E (European-ancestry), A (African-ancestry), C (Cross-ancestry)

<sup>e</sup> The maximum  $r^2$  value between any prediction model variant of the evaluated gene and the GWAS lead variant

## **2 CHAPTER 2: IMPROVED POLYGENIC RISK SCORE PREDICTION OF BREAST CANCER RISK IN WOMEN OF AFRICAN ANCESTRY LEVERAGING INFORMATION ACROSS ANCESTRIES, SUBTYPES, AND METHODS**

### **2.1 ABSTRACT**

Polygenic risk scores (PRSs) demonstrate utility for predicting breast cancer (BC) risk in European-ancestry women but show limited accuracy in African-ancestry women, particularly for aggressive subtypes. Here, we developed PRS models for overall BC, ER (estrogen receptor)-positive, ER-negative, and triple-negative BC (TNBC) in African-ancestry women using genome-wide association study data from the African Ancestry Breast Cancer Genetics consortium (17,391 cases/18,800 controls). PRS models were built using multiple PRS methods and then integrated into models leveraging information across ancestries, subtypes, and methods that significantly improved performance. The best models for overall, ER-positive, ER-negative, and TNBC showed an area under the receiving operating curve (AUC) of 0.612, 0.621, 0.610, and 0.639, respectively, and maintained predictive accuracy in external validation studies with AUCs of 0.612, 0.640, 0.600, and 0.652. We further introduce a parsimonious 162-variant PRS for TNBC with comparable accuracy (0.626) that validated robustly (0.655). These findings demonstrate markedly improved PRS accuracy for BC risk prediction in African-ancestry women. Utilizing these PRS models for screening will help promote more equitable cancer prevention efforts.

### **2.2 INTRODUCTION**

Breast cancer (BC) consists of 32% of all cancers in females,<sup>88</sup> and the mortality rate is estimated to be 40% higher in African American women compared to that in European American women.<sup>6,7</sup> Several factors contribute, in part, to this disparity, including differences in access to healthcare, later stages of diagnosis, and higher rates of aggressive BC subtypes such as triple-negative BC (TNBC) and estrogen receptor negative (ER-) BC, which are both associated with poorer survival outcomes and shorter latency periods.<sup>8,9</sup> This disproportionate burden is even more pronounced in women in Sub-Saharan Africa,<sup>10</sup> where BC mortality rates are among the highest globally.<sup>11</sup> Given the higher prevalence of aggressive BC subtypes in women of African ancestry, risk-stratified screening is particularly critical for this population to improve early detection and outcomes.

Polygenic risk score (PRS) models have emerged as valuable tools for predicting BC risk by summing the effects of common genetic variants identified in genome-wide association studies (GWAS). PRS models developed for women of European ancestry have demonstrated considerable predictive performance, with a 313-variant PRS obtaining an area under receiver operating characteristic curve (AUC) of 0.630, 0.641, and 0.601 for overall, ER-positive (ER+), and ER- BC, respectively.<sup>89</sup> However, these models exhibited lower predictive performance in African ancestry women with AUCs of 0.571, 0.588 and 0.562,<sup>90</sup> respectively. This disparity is largely driven by differences in linkage disequilibrium (LD)<sup>91</sup> and the distribution of common variants between ancestries.<sup>92–96</sup> Attempts to develop PRS models tailored to African American women have similarly struggled to achieve optimal performance, with the highest externally validated AUCs for overall, ER+, and ER- BC being 0.584, 0.595, and 0.576, respectively.<sup>97,98</sup>



these PRS models were developed using relatively modest sample sizes (9,235 cases/10,184 controls) and predominantly utilized the classical forward stepwise selection approach.

The limited predictive accuracy of existing PRS models for women of African ancestry, as well as the poor cross-ancestry portability of PRSs developed in other populations, highlights the need to build improved risk prediction models tailored to this population. In this study, we developed PRS models for four BC phenotypes including overall BC, ER+ BC, ER- BC, and TNBC. These models were constructed using genetic data from a larger number of participants from the African Ancestry Breast Cancer Genetics (AABCG) Consortium<sup>99</sup> and externally validated in the “All of Us” (AoU) study<sup>27</sup> and three additional studies of women of African ancestry. This work extends previous efforts to develop PRS models by incorporating a diverse range of PRS methods, including recently developed multi-ancestry approaches that account for genetic architecture across populations.<sup>100,101</sup>

## 2.3 RESULTS

For four BC phenotypes including overall BC and each of its subtypes (ER+ BC, ER- BC, and TNBC), we developed PRS models using genetic data from the AABCG consortium.<sup>99</sup> The AABCG data, consisting of 17,391 BC cases (9,021 ER+ BC, 4,768 ER- BC, 2,760 TNBC) and 18,800 controls, was randomly split into training (70%), tuning (10%), and validation (20%) sets (**Table S 2.1**). We performed a GWAS in the training set for each phenotype, fine-tuned model parameters in the tuning set, and evaluated model performance in the validation set.

Subsequently, external validation of models was carried out. A schematic overview of our study

is shown in **Figure 2.1**, and methodologic details, including data processing and model development, are described in Methods.

### **2.3.1 Tuning beta coefficients of European PRS models in the training set**

Prior to developing African ancestry-specific PRS models, we evaluated the performance of the previously published European 330-variant models<sup>89,102</sup> in our dataset as a benchmark. For the 281 variants available in the validation set after variant filtering and quality control, we observed limited performance for each phenotype including AUCs of 0.569 (95% CI: 0.556-0.583) for overall BC, 0.583 (95% CI: 0.567-0.599) for ER+ BC, 0.577 (95% CI: 0.555-0.599) for ER- BC, and 0.597 (95% CI: 0.571-0.622) for TNBC (**Table S 2.2**), which were consistent with prior studies in women of African ancestry.<sup>90,97</sup> We observed improved performance when utilizing beta coefficients for these 281 variants obtained from the GWAS performed in our training set, which further increased upon only selecting variants with directionally consistent effect sizes between the original coefficients and those in our training set (**Table S 2.2**). These modified European 330-variant models exhibited AUCs of 0.576 (95% CI: 0.562-0.589) for overall BC, 0.584 (95% CI: 0.567-0.599) for ER+ BC, 0.597 (95% CI: 0.576-0.617) for ER- BC, and 0.616 (95% CI: 0.591-0.642) for TNBC (**Table S 2.2**).

### **2.3.2 PRS models built using existing single-ancestry and cross-ancestry methods**

We developed single-ancestry PRS, including a forward stepwise selection (FSS) approach at different p-value thresholds that utilized individual-level genetic data, and several existing methods that utilized GWAS summary statistics from the training set including PLINK C+T,<sup>103</sup> PRSice-2,<sup>104</sup> Lassosum2,<sup>105</sup> and LDpred2-auto.<sup>106</sup> Among single-ancestry models, the highest

AUCs for each phenotype were 0.560 (95% CI: 0.547-0.573) for overall BC developed using PRSice-2, 0.558 (95% CI: 0.541-0.575) for ER+ BC using Lassosum2, 0.571 (95% CI: 0.550-0.591) for ER- BC using FSS, and 0.607 (95% CI: 0.582-0.633) for TNBC using Lassosum2 (**Figure S 2.1-4**).

We also implemented two cross-ancestry methods including PRS-CSx<sup>101</sup> and CT-SLEB,<sup>100</sup> which directly modeled data from both European and African ancestries; PRS-CSx is limited to HapMap3 variants. Both cross-ancestry methods performed better than single-ancestry methods for overall and ER+ BC, with CT-SLEB having the highest AUCs for overall BC (AUC: 0.592; 95% CI: 0.579-0.605) and ER+ BC (AUC: 0.612; 95% CI: 0.597-0.628) (**Figure S 2.5-6**). For ER- BC and TNBC, while PRS-CSx did not surpass single-ancestry models, CT-SLEB provided higher AUCs: 0.600 (95% CI: 0.580-0.620) for ER- BC and 0.619 (95% CI: 0.595-0.644) for TNBC (**Figure S 2.7-8**).

### **2.3.3 Cross-ancestry PRSs that incorporate European models demonstrate improved performance**

We built hybrid cross-ancestry PRSs, referred to as “xAncestry” models, by linearly combining each model above developed using existing PRS methods (including both single-ancestry and cross-ancestry methods) with the corresponding, modified European 330-variant model for overall BC, ER- BC, ER+ BC, and TNBC within the tuning set. Nearly all xAncestry models showed higher discriminating ability than PRSs developed using existing methods for each phenotype (**Figure S 2.5-8**). Among these xAncestry models, the highest AUCs for each phenotype were 0.595 (95% CI: 0.581-0.608) for overall BC developed using CT-SLEB, 0.617

(95% CI: 0.601-0.632) for ER+ BC using CT-SLEB, 0.609 (95% CI: 0.590-0.630) for ER- BC using PLINK C+T, and 0.639 (95% CI: 0.614-0.666) for TNBC using PRSice-2. Furthermore, we identified several xAncestry models for TNBC had a relatively low number of variants but had comparably high predictive performance including a 162-variant PRS model developed using Lassosum2 (AUC: 0.628, 95% CI: 0.601-0.651) and another 162-variant PRS developed using FSS (AUC: 0.626, 95% CI: 0.600-0.652).

#### **2.3.4 Ensemble-based approaches to borrow information across subtypes and methods**

To further enhance performance, we developed ensemble-based PRS models by incorporating all PRS models generated for each phenotype using either forward stepwise selection or elastic net regression in the tuning set, as well as existing European 313-/330-variant models; these European 330-variant models included models for overall BC, ER+ BC, ER- BC, Luminal-A-like BC, Luminal-B-like BC, HER2-enriched-like BC, Luminal-B/HER2-negative-enriched-like BC, and TNBC.<sup>102</sup> For all phenotypes except TNBC, ensemble-based models achieved a higher AUC than xAncestry models including 0.600 (95% CI: 0.587-0.613) for overall BC developed using forward stepwise selection, 0.621 (95% CI: 0.605-0.636) for ER+ BC using forward stepwise selection, and 0.610 (95% CI: 0.590-0.630) for ER- BC using elastic net regression (**Figure S 2.9**).

#### **2.3.5 Combining PRS models for each subtype by their proportions to build a PRS for overall breast cancer**

We additionally developed a PRS model for overall BC that consisted of the best performing ER+ BC and ER- BC PRSs (ensemble-based models) (**Table 2.1**) that were linearly combined

with weights equivalent to the proportion of each subtype in the training set; the proportion of ER+ cases was 0.67 and the proportion of ER- cases was 0.33. This two-subtype proportion-tuned model exhibited higher predictive performance compared to the ensemble-based model of overall BC (AUC: 0.603, 95% CI: 0.590-0.617) (**Figure 2.2**). We examined PRS models with arbitrarily assigned proportions of ER+ cases and observed that proportion-tuned PRS models were more predictive when the proportion of ER+ cases was closer to the actual proportion in the training set (**Figure 2.2**). We further developed a three-subtype proportion-tuned model by additionally incorporating the best-performing TNBC PRS (xAncestry PRSice-2 model) (**Table 2.1**). The weight of this TNBC PRS was set based on the proportion of cases classified as TNBC (0.19), while the weight of the ER- model was adjusted to account only for ER- cases that were not classified as TNBC (0.14). The performance of this PRS in the validation set was higher (AUC: 0.605, 95% CI: 0.592-0.618) and improved even further when restricting cases to those with complete subtype data (AUC: 0.612, 95% CI: 0.596-0.627) (**Figure 2.2**). Hence, we considered this three-subtype proportion-tuned model as our best performing model of overall BC (**Table 2.1**). For all subsequent analyses of this PRS model of overall BC, we restricted the AABCG validation set to only those with complete subtype data.

### **2.3.6 The association of the best performing PRSs and BC phenotypes**

The best performing PRS model for each phenotype was highly associated with an increased odds of developing that phenotype. The odds ratio (OR) per standard deviation was 1.49 (95% CI: 1.41-1.57) for overall BC, 1.53 (95% CI: 1.44-1.62) for ER+ BC, 1.49 (95% CI: 1.38-1.61) for ER- BC, and 1.65 (95% CI: 1.50-1.82) for TNBC in the validation set (**Table 2.1**). The top 5<sup>th</sup> percentile of PRS scores compared to the middle 40-60<sup>th</sup> percentile showed markedly high odds

ratios, with 2.39 (95% CI: 1.88-3.04) for overall BC, 2.44 (95% CI: 1.89-3.15) for ER+ BC, 2.27 (95% CI: 1.66-3.12) for ER- BC, and 2.70 (95% CI: 1.80-4.05) for TNBC (**Figure S 2.10, Table 2.1**).

### **2.3.7 Geographic and ancestry variations in PRS performance**

We separately assessed the performance of these PRS models in individuals from North American vs. African continents. PRS models for overall BC and ER+ BC showed higher odds ratios among individuals from Africa with statistically significant interactions (**Table S 2.3**); conversely, the odds ratios were higher in individuals from North America for ER- BC and TNBC, though interaction terms were not statistically significant. We also found overall BC and ER+ BC PRS models had suggestively higher odds ratios in individuals with a proportion of African ancestry  $\geq 80\%$  compared to in individuals with  $< 80\%$  African ancestry, while these odds ratios for ER- BC and TNBC PRS models were nearly identical for groups above and below 80% African ancestry (**Table S 2.3**). We did not observe any statistically significant interactions when treating proportion of African ancestry as a continuous variable (**Table S 2.3**). In addition, we did not observe statistically significant interactions between the phenotype-specific PRS models and age at diagnosis (**Table S 2.4**).

### **2.3.8 Absolute lifetime and 10-year risk estimates based on PRS models**

We calculated both the lifetime and 10-year absolute risk for the best performing models for each BC phenotype using age-specific incidence rates in African Americans from the Surveillance, Epidemiology and End Results program (SEER22, 2017-2021) and mortality rates from other causes from the Centers for Disease Control and Prevention, National Center for Health Statistics

(CDC NCHS) WONDER database (2017-2020). For individuals in the top 1<sup>st</sup> percentile, the absolute lifetime risk of BC by 80 years of age was 25.7% for overall BC, 20.9% for ER+ BC, 7.5% for ER- BC, and 7.4% for TNBC, after accounting for competing risks (**Figure 2.3**). In comparison, the absolute risk by 80 years of age for the top percentile of the 313/330-variant PRS in European Americans was higher for overall BC (33.1%) and ER+ BC (31.4%), but lower for ER- BC (4.3%) and TNBC (3.0%) (**Figure 2.4**). Moreover, while the average 10-year risk of overall BC for women 45 years of age (the recommended age for starting annual mammogram screenings)<sup>107</sup> in the United States is 2%, women in the >99<sup>th</sup>, 95-99<sup>th</sup>, and 90-95<sup>th</sup> percentiles of our best performing PRS of reached this 2% risk threshold at much younger ages – 32, 35, and 37 years, respectively. For more aggressive subtypes of BC including ER- BC and TNBC, the ages at which the 10-year risks in the top 1<sup>st</sup> percentile reached 1% were 37 and 38 years of age, respectively.

### **2.3.9 Utilizing family history of breast cancer in combination with PRS scores improves risk stratification**

In the AABCG study, 17.1% of all BC cases (n=2,976) and 8.3% of controls (n=1,554) had first-degree family history of BC. The best performing PRS models were strongly associated with each BC phenotype (**Table 2.2**) in both women with and without family history, though the strength of this association tended to be slightly higher in those with family history. Family history was also associated with risk of each BC phenotype, and we observed slight attenuation in these associations after controlling for PRS scores (**Table 2.2**). Formal tests for interaction between PRS scores and family history were not statistically significant.

To evaluate the combined effect of PRS and family history, we utilized the *iCARE* package<sup>108</sup> to simultaneously model both variables for each phenotype. Though only a small proportion of individuals in our dataset have family history of BC, we observed that those with a family history and high PRS scores had even greater lifetime and 10-year absolute risks at age 40 (**Figure S 2.11 and 12**). For example, the lifetime risks for women in the top 1<sup>st</sup> percentile of PRS scores who also had a positive family history were 53.6%, 45.3%, 15.7%, and 17.2% for overall BC, ER+ BC, ER- BC, and TNBC, respectively (**Table S 2.5**). For 40-year-old women with a positive family history, the 10-year absolute risks for women in the top 1<sup>st</sup> percentile of PRS scores were 11.1%, 8.5%, 2.8%, and 3.1%, respectively; in contrast, for 40-year-old women without a family history, the 10-year risks for the top 1<sup>st</sup> percentile of PRS scores were 3.8%, 2.9%, 1.1%, and 1.1% (**Table S 2.6**).

The SNP-based heritability of each phenotype was computed on the logit scale using LD score regression,<sup>109</sup> with the highest estimate for ER+ BC at 0.716 (SE: 0.187) and the lowest for ER- BC at 0.535 (SE: 0.286) (**Table S 2.7**). Notably, the best performing PRS of TNBC explained a significant proportion of the genome-wide heritability of this phenotype (43.6%); the best-performing PRSs for the other phenotypes explained a smaller but still substantial proportion of their respective genome-wide heritability, ranging from 26.4% to 29.4%.

### **2.3.10 External validation of PRS models in All of Us (AoU) and three additional studies in women of African ancestry**

We externally validated PRS models of overall BC risk in the AoU study, as well as in an independent combined dataset of women of African ancestry from three studies. As a benchmark,



we evaluated the European 330-variant PRS's performance among European American women in AoU and observed an AUC of 0.610 (95% CI: 0.601-0.618) (**Table S 2.8**); on the other hand, the performance of this 330-variant PRS had lower performance in African American women in AoU (AUC: 0.568; 95% CI: 0.550-0.587) and in the independent combined dataset of women of African ancestry 0.563 (95% CI: 0.529-0.595). Among African American women in AoU, the three-subtype proportion-tuned model of overall BC was highly consistent with the performance in the AABCG validation set with an AUC of 0.612 (95% CI: 0.592-0.631); the performance of this model in the independent combined dataset of women of African ancestry was also highly consistent with an AUC of 0.612 (95% CI: 0.581-0.643) (**Table S 2.8**).

We further validated subtype-specific models in the aforementioned independent combined dataset of women of African ancestry. We observed our best performing PRS models for each subtype in the AABCG validation set maintained high predictive accuracy in this external validation including AUCs of 0.640 (95% CI: 0.597-0.682) for ER+ BC, 0.600 (95% CI: 0.547-0.652) for ER- BC, and 0.652 (95% CI: 0.596-0.709) for TNBC (**Table S 2.9**). Of note, the 162-variant xAncestry PRS models for TNBC developed using FSS (0.655, 95% CI: 0.594-0.715) and Lassosum2 (AUC: 0.642, 95%: 0.585-0.701) validated robustly.

## 2.4 DISCUSSION

In this work, we developed and validated PRS models for four breast cancer phenotypes including overall BC, ER+ BC, ER- BC, and TNBC among women of African ancestry by analyzing data from AABCG, the largest genetic dataset of BC in this population. By leveraging information across ancestries, subtypes, and PRS methodologies, we introduced PRS models that

currently have the highest reported predictive performance in women of African ancestry. In our validation set, the AUCs for overall BC, ER+ BC, ER- BC, and TNBC were 0.612, 0.621, 0.610, and 0.639, respectively. These results mark a significant advancement, as they approach predictive performance of European ancestry-based PRS models, which have AUCs of 0.630, 0.641, 0.601, and 0.636, respectively.<sup>89,102</sup> Our models were also robustly validated in external validation studies. Additionally, we introduced two 162-variant PRS models for TNBC that have potential to serve as cost-effective screening tools for this aggressive BC subtype.

The improved performance of our PRS models compared to previous BC PRSs for women of African ancestry can be attributed, in part, to the increased sample size that was nearly double that of earlier studies in this population. By leveraging the largest available genetic dataset of BC patients of African ancestry,<sup>99</sup> we obtained more accurate and reliable effect size estimates for genetic variants, even for historically underpowered subtypes such as TNBC, which had a sample size of around 2,000 cases in the training set. Furthermore, our tuning set was sufficient large to enable hyperparameter optimization for each PRS method, as well as for linear combinations of PRS models. Given that the proportion of European ancestry in African Americans is approximately 20%,<sup>110</sup> we enhanced the predictive power of our models by using multi-ancestry PRS methods, such as CT-SLEB and PRS-CSx, which can jointly model the effects of genetic variants in both African and European populations. These approaches incorporated data from large European datasets, including 94,075 cases/75,017 controls and 133,384 cases/113,789 controls, respectively.<sup>89,102</sup> By accounting for genetic diversity and shared ancestry between African and European populations, these methods increased the statistical power and accuracy of our PRS models in women of African ancestry.

In addition to gaining power in cross-ancestry methods, this study also highlights the importance of integrating information across BC phenotypes and PRS methods to develop PRS models. We observed that ensemble-based approaches exhibited higher discriminating ability for most phenotypes, except for TNBC. While we cannot definitively explain the exact reasons for this improvement, part of the enhanced predictive performance for ER+ BC and ER- BC likely stems from incorporating information from the overall BC risk model, which benefits from a larger sample size and consequently boosts statistical power of these subtype-specific models. In contrast, ensemble-based PRS models for TNBC had lower predictive performance than some of xAncestry models, suggesting that the genetic determinants of TNBC risk may differ from those captured in overall BC models. Additionally, we found that incorporating the best performing models for ER+ BC, ER- BC, and TNBC into the three-subtype proportion-tuned PRS model improved predictive power for overall BC; this finding suggests that accounting for subtype heterogeneity enhances the ability to model genetic risk in overall BC, which is quite diverse in its etiology across subtypes.

The developed PRS models have very good utility in stratifying women by risk of each phenotype for BC screening. The American Cancer Society currently recommends women to start regular annual mammography screenings at 45 years of age,<sup>111</sup> corresponding to an average 10-year risk of about 2% for developing BC. Our findings suggest that women in the top 1st percentile of PRS scores for overall BC risk could benefit from screening as early as age 32, since their 10-year risk would be comparable to the established screening threshold. Notably, compared to the lifetime risks estimated by the European 313-/330-variant PRSs for ER- BC and

TNBC in women of European ancestry, our PRS models predicted noticeably higher risks for these two subtypes in African American women. This suggests that our PRS models may provide better stratification for determining risk and optimal age for initiating intensive screening. Given the lower survival rates associated with TNBC and ER- BC, more intensive screening for African-ancestry women with high PRS scores could reduce the chance of missing aggressive cancers that may develop between traditionally recommended annual mammography screening intervals for women between the ages of 45 and 54, as well as aid in the early detection of cancers that may develop before this recommended age for screening. Consequently, our results support the use of subtype-specific PRS scores for risk stratification. If a 10-year risk threshold of 1% were used as a screening criterion, our findings suggest that women in the top 1<sup>st</sup> percentile of ER- BC and TNBC PRS scores could begin screening as early as 37 and 38 years of age. Remarkably, the lifetime risk for individuals in the top 1<sup>st</sup> percentile of each subtype's PRS was in the same order of magnitude as those previously reported for germline pathogenic variants of moderate penetrance genes, while the lifetime risk for those with both first-degree family history of BC and the top 1<sup>st</sup> percentile of PRS reaches the level of high-penetrance genes.<sup>112</sup> These findings suggest that earlier and more intensive screenings using mammography or magnetic resonance imaging could be conducted for individuals with high PRS scores, similar to current screening recommendations for carriers of high- or moderate-risk gene mutations.<sup>113,114</sup> Furthermore, elevated ER+ BC PRS scores, in combination with other epidemiologic risk factors, may warrant consideration of chemoprevention for women of African ancestry.

In summary, we developed PRS models of overall BC, ER+ BC, ER- BC, and TNBC for women of African ancestry that demonstrate predictive performance comparable to PRS models of BC developed for women of European ancestry. These models effectively stratify individuals by absolute risk of developing each BC subtype and have been robustly validated in multiple external datasets. Our findings have important implications for addressing racial disparities in BC mortality by enabling more precise risk stratification, facilitating earlier and risk-adaptive screenings, and informing tailored intervention strategies. Future validation and clinical evaluation of these models in other populations of women of African ancestry are necessary to assess their applicability and potential impact on reducing disparities in breast cancer outcomes. This study also highlights how integrating information across ancestries, BC subtypes, and PRS methods can enhance PRS model performance, which has broader implications for PRS development of other diseases.

## **2.5 METHODS**

### **2.5.1 Study participants and sources of genetic data**

This study included 17,391 BC cases (9,021 ER+ BC, 4,768 ER- BC, 2,760 TNBC) and 18,800 controls who were all women of African ancestry from the African-Ancestry Breast Cancer Genetic (AABCG) Consortium. The data collection of studies included in AABCG have been previously described and protocols had IRB approval.<sup>99</sup> The genetic data predominantly consisted of 10 primary datasets that included genotyping data on the Multi-Ethnic Genotyping Array (MEGA), whole-genome sequencing (WGS) data, and genotyping data from previous studies/consortia; sample sizes of participants from each genotyping platform and study are detailed in **Table S 2.1**.

We also externally validated models of overall BC risk by computing covariate-adjusted AUCs among all women in the AoU study<sup>115</sup> (6,509 cases and 139,054 controls), those with African ancestry (834 cases and 29,122 controls), European ancestry (4,183 cases and 68,427 controls), and Latina women (682 cases and 26,139 controls). BC subtype information for each participant in AoU were not provided. We additionally performed external validation of PRS models for each of the four BC phenotypes in an independent combined dataset of women of African ancestry (580 cases and 3,011 controls) from three studies of women of African ancestry including The Vanderbilt Biobank (BioVU),<sup>116,117</sup> Black Women: Etiology and Survival of Triple-negative Breast Cancers (BEST) Study,<sup>118</sup> and the Women Firefighters Biomonitoring Collaborative (WFBC).<sup>119</sup>

## 2.5.2 Strategies for PRS development

**Figure 2.1** provides the scheme of our PRS development and validation. The categories of PRS models that we developed are as follows. 1) ***PRS models using existing methods***: This category includes PRS models developed using both single-ancestry and multi-ancestry approaches. These methods include forward stepwise selection (FSS),<sup>89,120</sup> PLINK C+T,<sup>103</sup> PRSice-2,<sup>104</sup> Lassosum2,<sup>105</sup> and LDpred2-auto,<sup>106</sup> as well as more advanced cross-ancestry methods like CT-SLEB<sup>100</sup> and PRS-CSx,<sup>101</sup> which leverage both African ancestry-specific PRSs and European genetic data to improve performance.

2) ***Cross-ancestry (xAncestry) PRS models***: For each phenotype, these models were built by linearly combining each PRS developed using existing methods (including both single- and cross-ancestry methods) with the corresponding European 330-variant PRS model for that

phenotype with modified effect size estimates based on the GWAS of the training set. 3)

**Ensemble-based models:** These models incorporated all developed PRS models, as well as pre-existing European 330-variant PRS models, across all phenotypes using forward stepwise selection or elastic net regression. We aimed to enhance the predictive accuracy of PRS models for each phenotype by leveraging information from all phenotype models. 4) **PRS model of**

**overall BC tuned by proportions of each subtype:** This model of overall BC incorporated the best performing models for each subtype (ER+ BC, ER- BC, and TNBC) as a linear combination where the weight for each subtype PRS model was the proportion of cases with that subtype in the training set. Given that TNBC cases were a subset of ER- cases, the weight given to the ER- BC PRS model was the proportion of ER- cases after subtracting the proportion of TNBC cases.

### 2.5.3 Processing genetic data

All multiallelic variants were converted to biallelic variants within VCF files and subsequently normalized to the hg38 human reference genome using bcftools.<sup>121</sup> Next, we used bcftools to merge genetic data across all datasets. We then extracted the dosage values for each variant from genotyped/imputed datasets (ranging from 0 to 2) with a minor allele frequency (MAF)>0.01 and an imputation information score >0.3; for WGS datasets, we coded the sequenced allele counts similarly to dosages but with discrete values (0, 1, or 2). Following this, we filtered out variants with allele frequencies which differed more than 0.15 across datasets. Next, we retained variants that were missing in less than 10 individuals (0.028% missingness) and subsequently performed mean imputation. Principal components (PCs) were generated in PLINK 2.0 by first pruning for variants with an  $r^2 < 0.2$  within 500kb windows and the association between PCs and proportions of African ancestry inferred by Admixture (v1.3.0)<sup>64</sup> were evaluated (**Figure S 2.13**). We then

randomly subset the genetic dataset into training (70%), tuning (10%), and interval validation sets (20%). To develop PRS models for each BC phenotype, we restricted cases within each of the data subsets to each phenotype.

Due to our collation of datasets that utilized different genotyping and genotyping platforms, many variants that were published as part of the European 313-/330-variant PRSs did not meet this stringent criterion of missingness (228 out of 330 variants met this criterion). To improve our ability to leverage these European models by increasing the variants we could cover, we relaxed this missingness threshold up to 20% for variants in these models, thereby increasing the number of variants to 281 out of 330. Given the large proportion of missingness for these additional 53 variants, we performed imputation of these variants. Variants that were missing in controls were mean imputed based on allele frequencies in controls in the training set. For cases, we imputed genotypes using the method described by Du et al. 2021,<sup>122</sup> using the allele frequency of each variant in controls of training set and the beta coefficient obtained from the GWAS in the training set detailed in the subsection below titled “Generating GWAS summary statistics for each subtype among women of African ancestry.” To prevent any chance of contamination from this imputation of missing data for variants with larger missingness rates in our PRS models, we only utilized these larger missingness variants for evaluating models containing the European 330-variant PRS, but not when building models using existing PRS methods.

#### **2.5.4 Generating GWAS summary statistics for each phenotype among women of African ancestry**



Within the training set, we used PLINKv2.0<sup>67</sup> to perform a GWAS of overall breast cancer status while adjusting for age, genotyping platform, and the top 10 PCs. For each BC subtype (ER+ BC, ER- BC, and TNBC), we performed similar GWAS analyses with the same controls while restricting cases to each subtype.

### **2.5.5 Obtaining European GWAS summary statistics for each subtype**

For overall BC, ER+ BC, and ER- BC, we retrieved summary statistics among women of European ancestry from the Breast Cancer Association Consortium (BCAC) for 122,977 BC cases and 105,974 controls, as well as 18,908 *BRCA1* mutation carriers, of which 9,414 have BC<sup>123,124</sup> (**Figure 2.1**). For TNBC, summary statistics derived from 133,384 BC cases and 113,789 controls, and 18,908 *BRCA1* mutation carriers (of which 9,414 had BC) were utilized.<sup>102</sup>

### **2.5.6 Obtaining European 330-variant PRS models with effect size estimates that were modified based on the GWAS in our training set**

We obtained the European 313-variant PRS models that had been previously constructed for overall BC, ER+ BC, and ER- BC using a forward stepwise selection approach,<sup>89</sup> as well as the previously published European 330-variant PRS of TNBC that additionally included 17 genome-wide significant variants across intrinsic subtypes.<sup>102</sup> For each of the 313-variant PRS models, we obtained effect size estimates for the remaining 17 variants from an established European GWAS meta-analysis of BC;<sup>123</sup> we collectively refer to these models as the European 330-variant PRS models for each phenotype. We additionally built modified versions of these 330-variant PRS models by utilizing coefficients calculated from the GWAS in our training set (see above) and further reduced these 330 variants to only those variants with directionally consistent effect

sizes between the original coefficients and those in the training set of our GWAS. We subsequently utilized these modified 330-variant PRS models to build “xAncestry” and ensemble-based PRS models.

### 2.5.7 Implementing existing PRS approaches

We implemented several existing single-ancestry PRS approaches using only data from AABCG for each BC phenotype (overall BC, ER+ BC, ER- BC, and TNBC). One approach we used was forward stepwise selection (**FSS**) with purpose of identifying parsimonious list of variants.<sup>89,120</sup> In FSS, we utilized individual-level data to perform LD clumping in 5 Mb windows within chromosomes in the training set followed by forward stepwise selection, at several specified p-value thresholds ( $1 \times 10^{-8}$ ,  $1 \times 10^{-7}$ ,  $1 \times 10^{-6}$ ,  $1 \times 10^{-5}$ ,  $1 \times 10^{-4}$ ,  $1 \times 10^{-3}$ , 0.01, 0.05); following this initial clumping and stepwise selection, we performed an additional round of LD clumping and forward stepwise selection on 5 Mb windows that were offset by 2.5 Mb. Only variants with p values smaller than the pre-specified thresholds in the FSS were included. We also implemented several existing methods that utilized GWAS summary statistics from the training set and tuned parameters in the tuning set including PLINK C+T,<sup>103</sup> PRSice-2,<sup>104</sup> Lassosum2,<sup>105</sup> and LDpred2-auto.<sup>106</sup> Briefly, all processed variants were used in **PLINK C+T** and were clumped in 5Mb windows using an  $r^2$  cutoff of 0.1 at p-value thresholds of  $1 \times 10^{-3}$ , 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5. **PRSice-2** was similarly implemented for all processed variants in our dataset using a clumping window of 250kb with an  $r^2$  threshold of 0.1. **LDpred2-auto** and **Lassosum2** were implemented for HapMap3 variants as recommended by the authors.

We additionally implemented two existing cross-ancestry methods, PRS-CSx<sup>125</sup> and CT-SLEB.<sup>126</sup> For each phenotype, we input summary statistics for HapMap3 variants from our training set and previous GWAS in European women, along with LD reference panels generated from the 1000 Genomes Project into **PRS-CSx** to generate African and European ancestry-specific PRS models using a  $\phi$  parameter of 0.01; we then linearly combined these PRSs in the tuning set to obtain a cross-ancestry PRS. We additionally built cross-ancestry PRS models for each phenotype using **CT-SLEB** by inputting summary statistics from our training set with a  $p$ -value  $< 0.3$ , as well as from previous GWAS in European women for the corresponding subtypes. Within CT-SLEB, we performed two-dimensional clumping of African and European ancestry summary statistics using several  $p$ -value thresholds ( $5 \times 10^{-8}$ ,  $5 \times 10^{-7}$ ,  $5 \times 10^{-6}$ ,  $5 \times 10^{-5}$ ,  $5 \times 10^{-4}$ ,  $5 \times 10^{-3}$ ,  $5 \times 10^{-2}$ , 0.1, 0.3) and  $r^2$  parameters (0.01, 0.05, 0.1, 0.2, 0.5, 0.8) prior to applying empirical Bayes estimation, and a subsequent super-learner step to obtain PRS models.

### **2.5.8 Developing cross-ancestry (xAncestry) PRS models by linearly combining each PRS model with the corresponding European model for each phenotype**

To each single- and cross-ancestry model built using aforementioned methods, we then linearly combined the corresponding modified 330-variant PRS model for each phenotype in the tuning set to develop cross-ancestry models that we refer to as “**xAncestry**” models. Though we had built cross-ancestry PRS models using CT-SLEB and PRS-CSx, we still performed this linear combination step since the variants captured by these methods did not necessarily include variants in the parsimonious 330-variant models.

### **2.5.9 Building ensemble-based PRS models**

Aiming to further enhance prediction accuracy of PRS models, we additionally developed **ensemble-based** PRS models for each BC phenotype by incorporating all PRS models we developed across all the phenotypes; these ensemble-based models were designed to leverage the strengths of different modeling approaches and to borrow information across subtypes. We first identified PRS models that were very highly correlated ( $r^2 > 0.95$ ) and only retained the PRS with the most statistically significant association with a given BC phenotype in the tuning set. Next, we applied two variable selection approaches to develop ensemble-based PRSs: 1) forward stepwise selection and 2) elastic net regression. For forward stepwise selection approach, we identified the association between each PRS model and a given BC phenotype using logistic regression, adjusting for age, genotyping platform, and the top 10 PCs in the tuning set and subsequently utilized a p-value threshold of 0.05 for stepwise selection. For elastic net regression approach, we utilized the *glmnet*<sup>127</sup> package in R to perform elastic net regression with 10-fold cross-validation to obtain ensemble-based PRSs for each phenotype.

#### **2.5.10 Developing joint-subtype proportion-tuned PRS models of overall BC**

We first developed a **two-subtype proportion-tuned** PRS model for overall BC that consisted of the best performing ER+ BC and ER- BC PRSs that were linearly combined with weights equivalent to the proportion of each subtype in the training set. We further developed a **three-subtype proportion-tuned** model by additionally incorporating the best-performing TNBC PRS model where the weight of this TNBC PRS was set based as the proportion of cases classified as TNBC, while the weight of the ER- model was adjusted to include only ER- cases that were not classified as TNBC, i.e. ER-/HER2+ subtype.

### **2.5.11 Evaluating PRS models in the validation set**

For each phenotype, we assessed the predictive accuracy of each PRS model in the validation set by computing the covariate-adjusted area under the receiver operating characteristic curve (AUC), adjusting for age, genotyping platform, and the top 10 PCs using the *ROCnReg* R package.<sup>128</sup> We fit logistic regression models to calculate the covariate-adjusted odds ratio per unit increase in standard deviation of PRS. Furthermore, we assessed whether age, first-degree family history of BC, continent, and proportion of inferred African ancestry modified the association between PRS and BC risk by adding interaction terms in separately fit logistic regressions.

### **2.5.12 Calculating absolute risk of models**

We estimated the lifetime and 10-year absolute risks of developing each BC phenotype utilizing a previously described method that adjusts for competing risks including mortality from causes other than BC.<sup>129</sup> Age-specific incidence rates of each BC phenotype in African Americans were obtained from the Surveillance, Epidemiology and End Results program (SEER22, 2017-2021) and mortality rates from other causes were obtained from CDC NCHS's WONDER database (2017-2020). As a reference, we also computed the lifetime absolute risks of developing BC in European Americans using the European 313-/330-variant models by using odds ratios per standard deviation of PRS scores that were previously published in European Americans.<sup>102,130</sup> We additionally utilized the *iCARE* R package<sup>108</sup> to compute absolute risks for combinations of PRSs and first-degree family history of BC, using incidence rates and mortality rates described above and prevalence of family history of BC in our study samples.

### **2.5.13 Examining proportion of genome-wide heritability explained by PRS models**

We utilized LD score regression<sup>109</sup> to compute the genome-wide heritability of each BC subtype on the observed frailty scale using summary statistics obtained from the training set. The LD score reference panel we utilized consisted of 1,953,872 variants genotyped using the MEGA array from a subset of controls in the Roswell Park Comprehensive Cancer Center dataset (n = 2,730), including 851 controls from the Women's Circle of Health Follow-Up Study and 1,879 controls from the Black Women's Health Study.<sup>99</sup> The covariate-adjusted AUC of select PRS models for each subtype was converted into heritability on the observed scale using a previously described method.<sup>131</sup>

## 2.6 APPENDIX

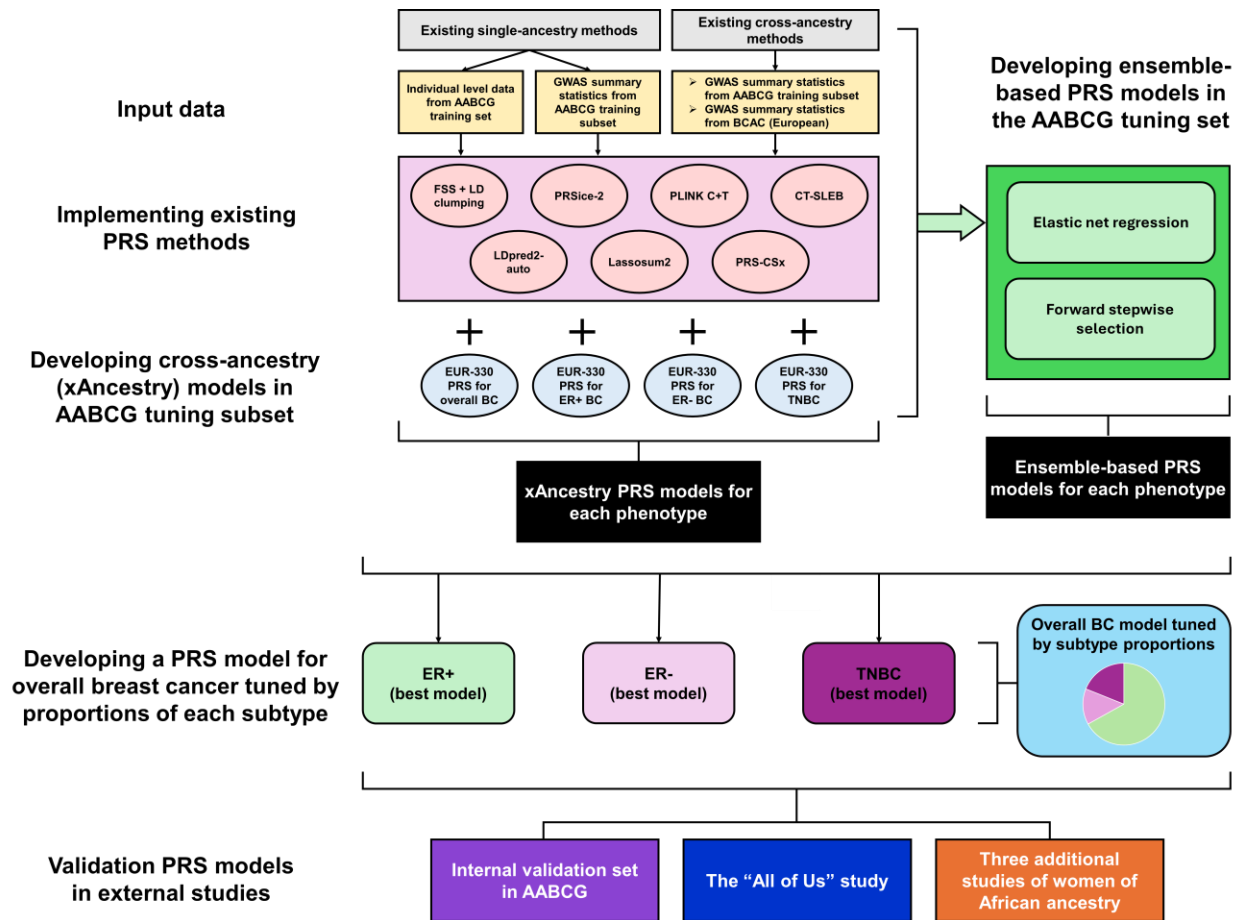


Figure 2.1 Workflow for constructing and validating polygenic risk score models.

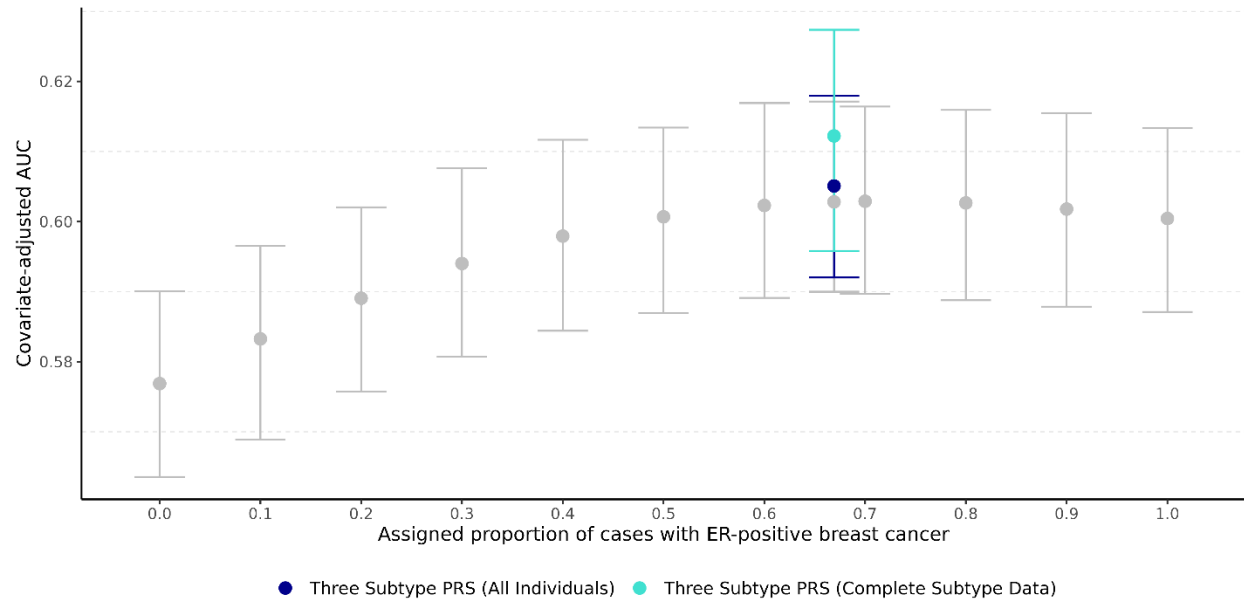


Figure 2.2 Developing a PRS model for overall breast cancer tuned by proportions of each subtype. Note: The observed proportion of ER+ cases in the training set was 0.67. Gray error bars indicate the performance (95% CI of the covariate-adjusted AUC) of two-subtype proportion-tuned PRS models developed by combining only ER-positive and ER-negative models by various assigned proportions of ER-positive cases. The navy-blue error bar indicates the performance of the three-subtype proportion-tuned PRS model developed by combining ER-positive, ER-negative, and TNBC PRS models by their observed proportions in the validation set. The turquoise error bar indicates the performance of the three-subtype proportion-tuned PRS model only in cases with complete subtype data.



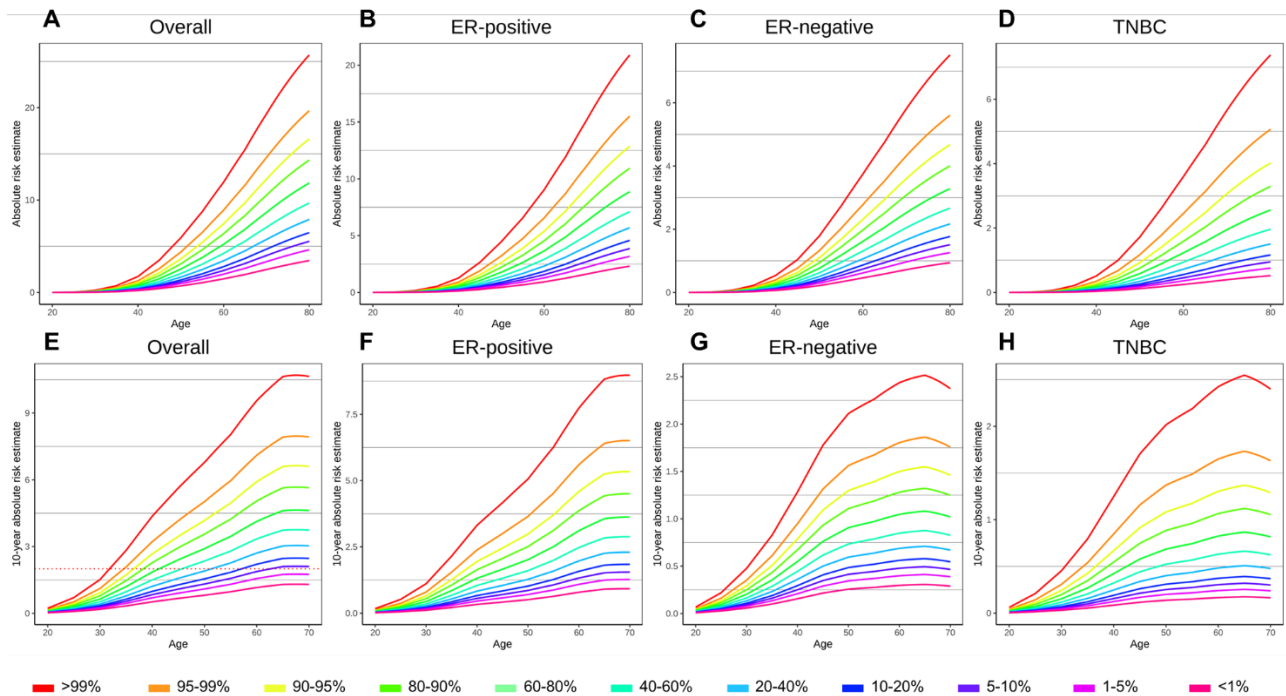


Figure 2.3 Estimated lifetime and 10-year absolute risks of developing each breast cancer subtype among women with African ancestry by percentiles of the polygenic risk scores.

Lifetime absolute risks of developing A) overall breast cancer, B) estrogen receptor (ER)-positive breast cancer, C) ER-negative breast cancer, and D) Triple-negative breast cancer (TNBC). 10-year absolute risk of developing E) overall breast cancer, F) ER-positive breast cancer, G) ER-negative breast cancer, and H) TNBC. The red dotted line in E) corresponds to a 10-year risk of 2% that can be used as a recommendation for screening ages.

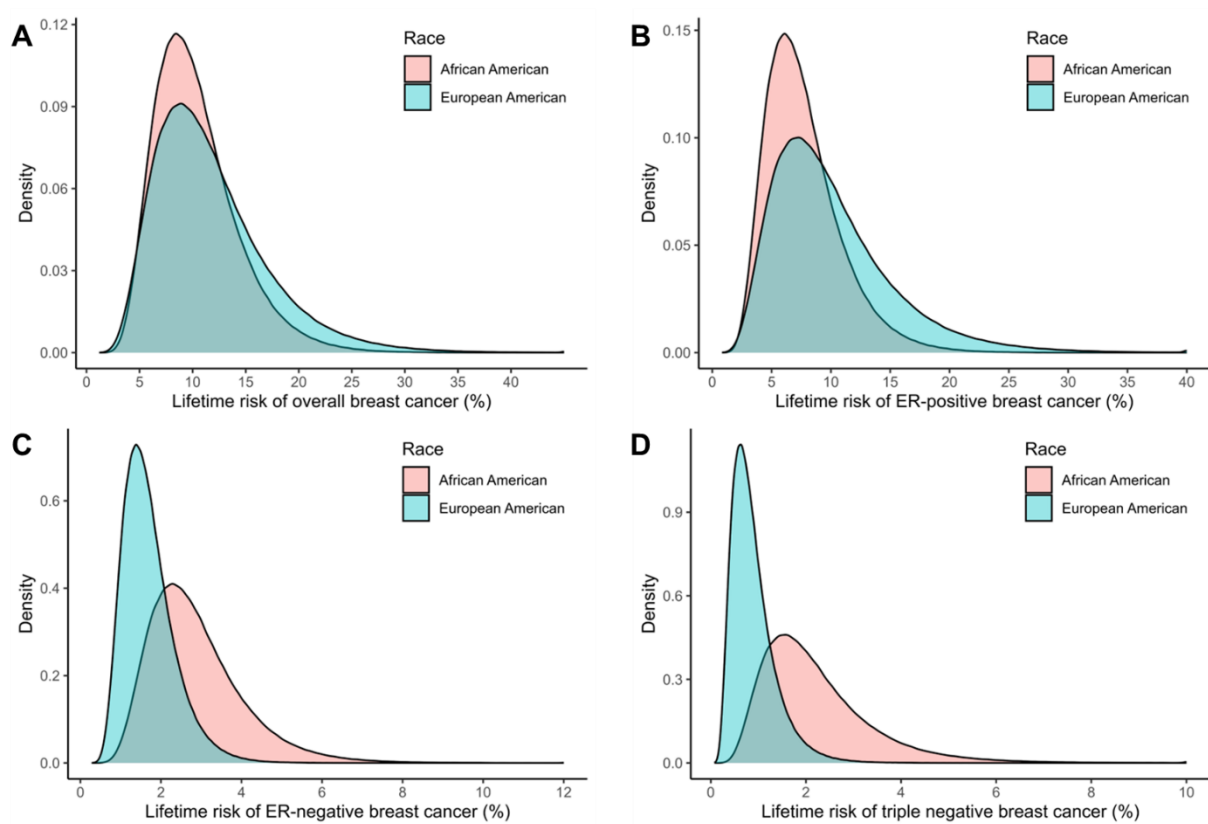


Figure 2.4 Density plots for the lifetime risk of developing each subtype of breast cancer between African Americans and European Americans. Red indicates the estimated theoretical lifetime risk of the best performing PRS model for each subtype among African Americans and blue indicates the estimated theoretical lifetime risk of the European 313-/330-variant PRSs for each subtype among European Americans. A) Overall breast cancer B) ER-positive breast cancer C) ER-negative breast cancer D) Triple-negative breast cancer.

Table 2.1 Predictive performance of the best performing PRS models for each breast cancer phenotype in the validation set. Note: The 95% CI of each metric is presented in parentheses.

| Phenotype                                      | Overall                              | ER-positive                | ER-negative                        | TNBC                       |
|------------------------------------------------|--------------------------------------|----------------------------|------------------------------------|----------------------------|
| Method                                         | Three subtype proportion-tuned model | Ensemble-based model (FSS) | Ensemble-based model (elastic net) | xAncestry model (PRSice-2) |
| Number of SNPs                                 | 2,324,063                            | 592,247                    | 2,007,632                          | 554,959                    |
| AUC                                            | 0.612<br>(0.597, 0.628)              | 0.621<br>(0.605, 0.636)    | 0.610<br>(0.591, 0.630)            | 0.639<br>(0.613, 0.665)    |
| OR per SD                                      | 1.49 (1.41, 1.57)                    | 1.53 (1.44, 1.62)          | 1.49 (1.38, 1.61)                  | 1.65 (1.50, 1.82)          |
| OR of percentile compared to middle 20% of PRS | <1%                                  | 0.52 (0.26, 1.04)          | 0.31 (0.12, 0.76)                  | 0.14 (0.02, 1.09)          |
|                                                | 1-5%                                 | 0.49 (0.33, 0.71)          | 0.44 (0.29, 0.67)                  | 0.32 (0.15, 0.69)          |
|                                                | 5-10%                                | 0.54 (0.39, 0.75)          | 0.61 (0.43, 0.86)                  | 0.35 (0.17, 0.72)          |
|                                                | <10%                                 | 0.52 (0.40, 0.67)          | 0.52 (0.39, 0.69)                  | 0.32 (0.19, 0.55)          |
|                                                | 10-20%                               | 0.66 (0.51, 0.83)          | 0.63 (0.48, 0.82)                  | 0.65 (0.43, 0.99)          |
|                                                | 20-40%                               | 0.81 (0.67, 0.97)          | 0.92 (0.76, 1.12)                  | 0.69 (0.49, 0.96)          |
|                                                | 40-60%                               | 1 (reference)              | 1 (reference)                      | 1 (reference)              |
|                                                | 60-80%                               | 1.22 (1.03, 1.45)          | 1.44 (1.20, 1.74)                  | 1.05 (0.78, 1.43)          |
|                                                | 80-90%                               | 1.65 (1.35, 2.02)          | 1.81 (1.46, 2.25)                  | 2.01 (1.46, 2.77)          |
|                                                | >90%                                 | 2.10 (1.73, 2.55)          | 2.27 (1.85, 2.79)                  | 2.53 (1.84, 3.49)          |
|                                                | 90-95%                               | 1.81 (1.41, 2.32)          | 2.15 (1.67, 2.77)                  | 2.41 (1.63, 3.57)          |
|                                                | >95%                                 | 2.39 (1.88, 3.04)          | 2.44 (1.89, 3.15)                  | 2.70 (1.80, 4.05)          |
|                                                | 95-99%                               | 2.46 (1.90, 3.19)          | 2.46 (1.86, 3.25)                  | 2.70 (1.75, 4.18)          |
|                                                | 99-100%                              | 2.06 (1.30, 3.27)          | 2.30 (1.43, 3.72)                  | 3.04 (1.42, 6.49)          |

Table 2.2 Association of PRS scores and family history with each breast cancer subtype. Note: PRS score is per SD; Family history is first-degree relatedness. The 95% CI of each odds ratio is presented in parentheses.

| Phenotype                                                              | Overall BC Risk   | ER-positive       | ER-negative       | TNBC              |
|------------------------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|
| <b>Association of PRS and cancer risk by family history</b>            |                   |                   |                   |                   |
| PRS unadjusted for family history                                      | 1.49 (1.41, 1.57) | 1.53 (1.44, 1.62) | 1.49 (1.38, 1.61) | 1.65 (1.50, 1.82) |
| PRS in women without family history                                    | 1.49 (1.39, 1.59) | 1.54 (1.43, 1.65) | 1.42 (1.30, 1.56) | 1.60 (1.42, 1.80) |
| PRS in women with family history                                       | 1.61 (1.35, 1.91) | 1.65 (1.37, 1.98) | 1.63 (1.32, 2.01) | 1.94 (1.48, 2.55) |
| <b><i>P for testing interaction between PRS and family history</i></b> | 0.37              | 0.42              | 0.28              | 0.16              |
| <b>Association of family history and cancer risk</b>                   |                   |                   |                   |                   |
| Family history unadjusted for PRS                                      | 3.16 (2.67, 3.74) | 3.08 (2.59, 3.67) | 2.75 (2.20, 3.43) | 2.98 (2.27, 3.89) |
| Family history adjusted for PRS                                        | 3.06 (2.57, 3.64) | 3.01 (2.52, 3.61) | 2.63 (2.10, 3.29) | 2.88 (2.18, 3.78) |

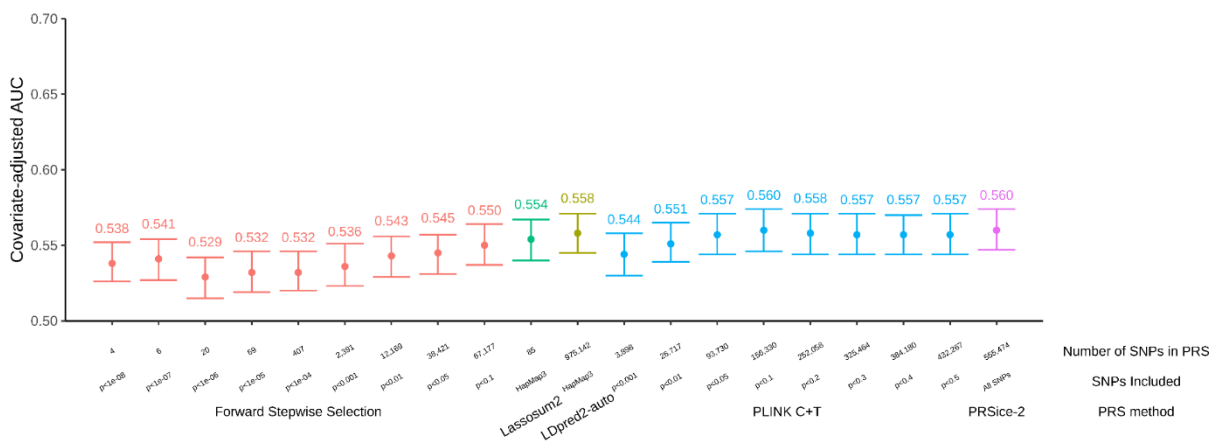


Figure S 2.1 Performance of single-ancestry PRS models of overall breast cancer risk in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

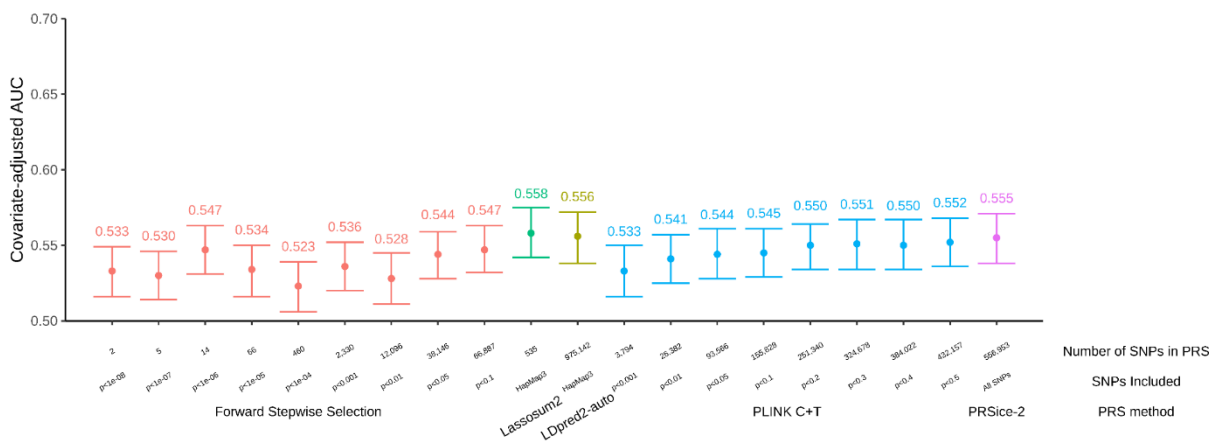


Figure S 2.2 Performance of single-ancestry PRS models of ER-positive BC in the validation set.  
Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

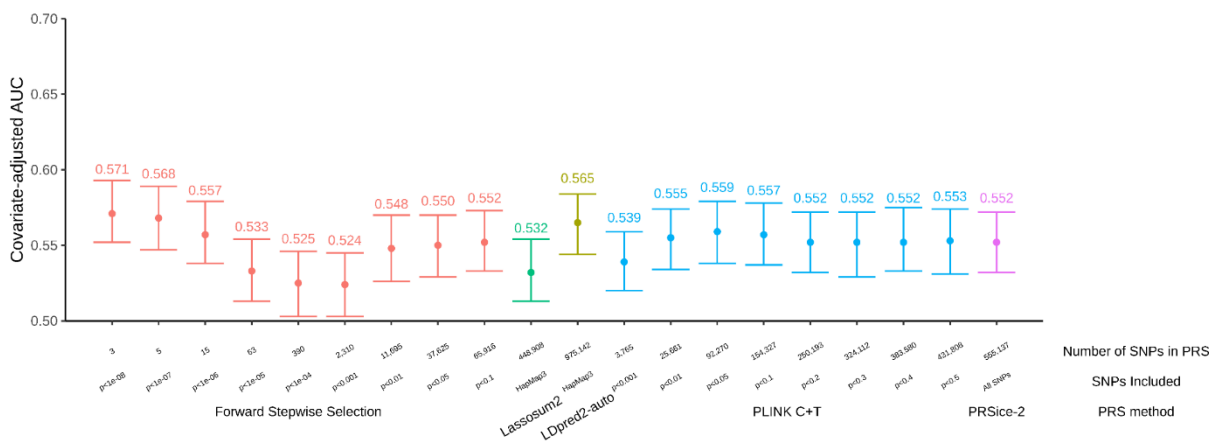


Figure S 2.3 Performance of single-ancestry PRS models of ER-negative BC in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

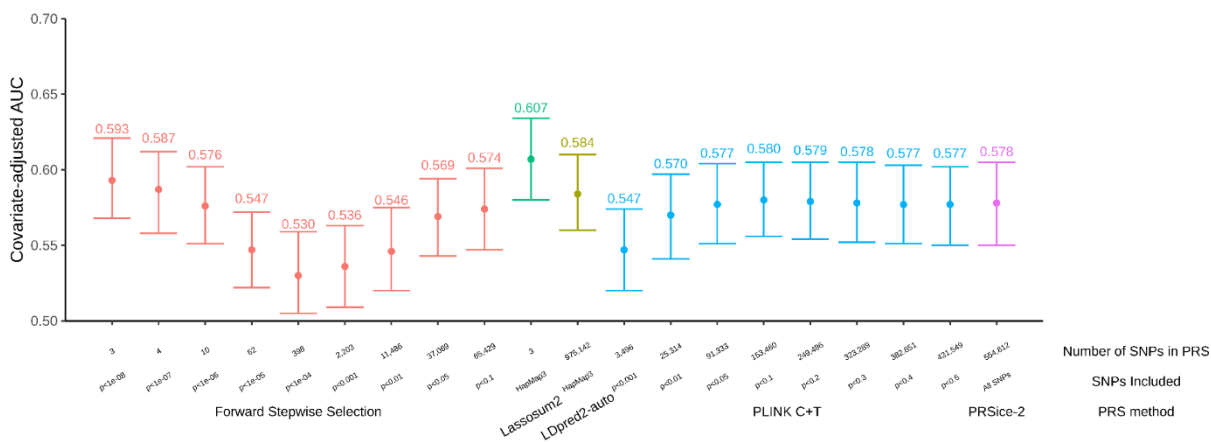


Figure S 2.4 Performance of single-ancestry PRS models of triple-negative breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.



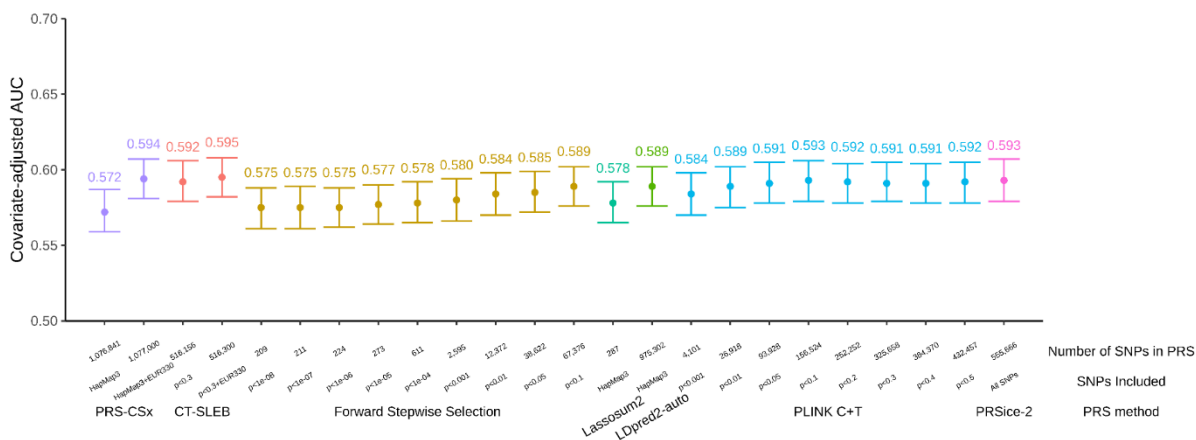


Figure S 2.5 Performance of cross-ancestry (xAncestry) PRS models of overall breast cancer risk in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

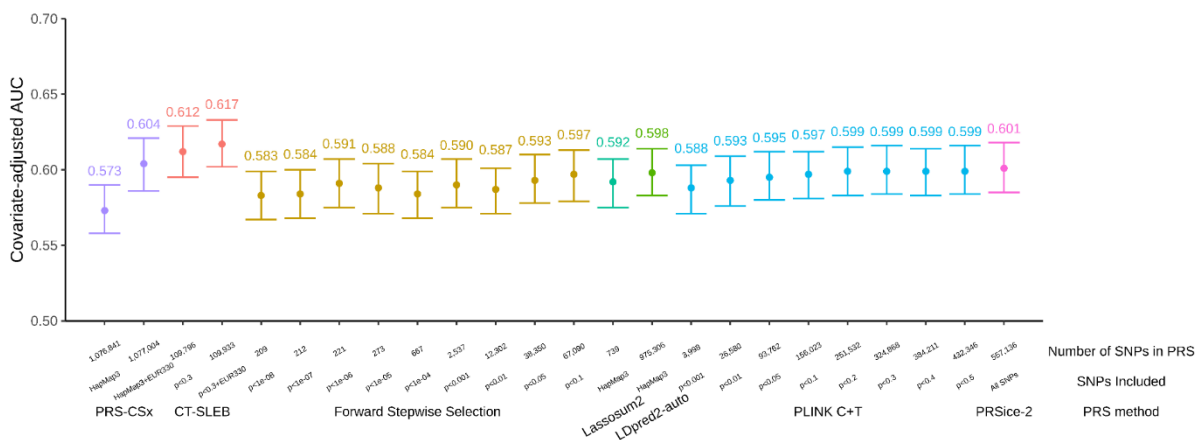


Figure S 2.6 Performance of cross-ancestry (xAncestry) PRS models of ER-positive breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

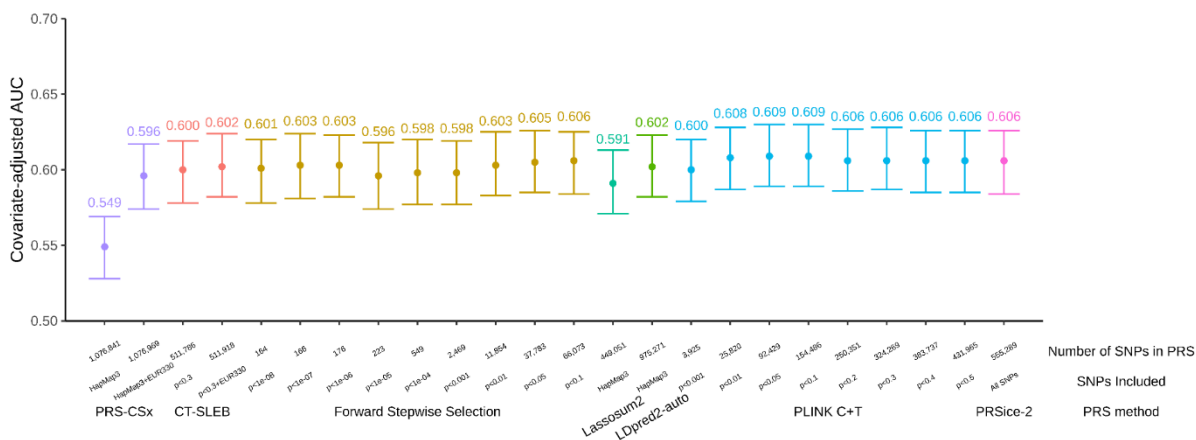


Figure S 2.7 Performance of cross-ancestry (xAncestry) PRS models of ER-negative breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

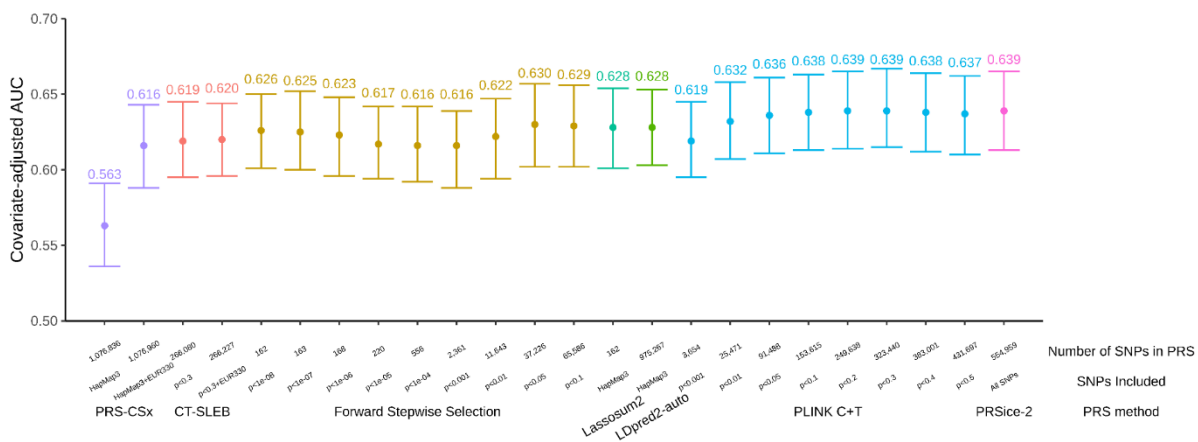


Figure S 2.8 Performance of cross-ancestry (xAncestry) PRS models of triple negative breast cancer in the validation set. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

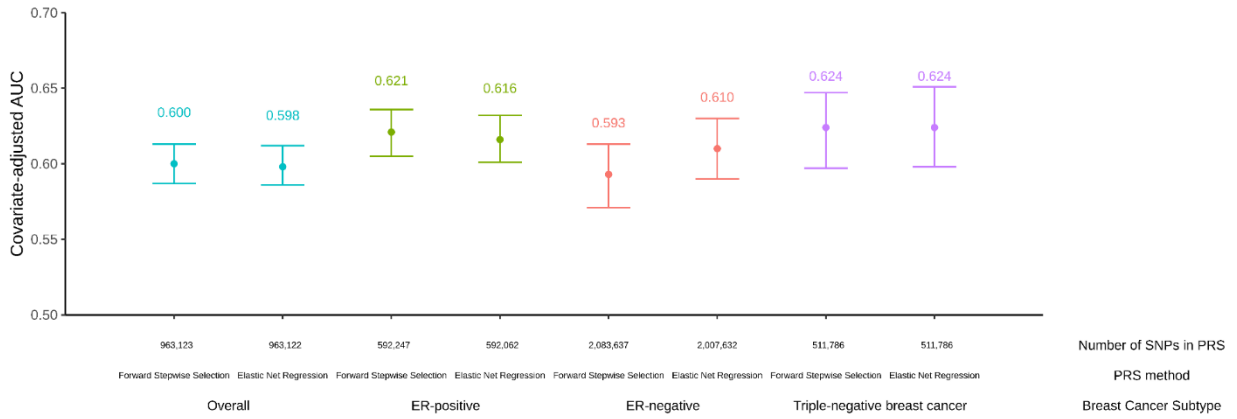


Figure S 2.9 Figure S 2.9. Performance of ensemble-based PRS models of breast cancer by subtype. Note: Error bars represent the 95% confidence interval for each covariate-adjusted AUC.

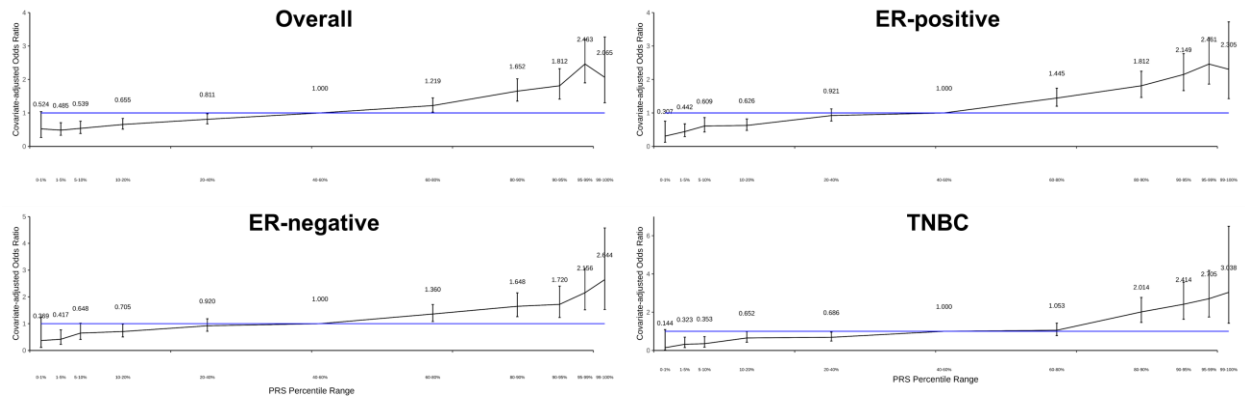


Figure S 2.10 Observed odds ratio for each breast cancer subtype by polygenic risk score (PRS) percentile compared to the middle 40-60th percentile of PRS scores. Note: Error bars represent the 95% confidence interval for each covariate-adjusted odds ratio.

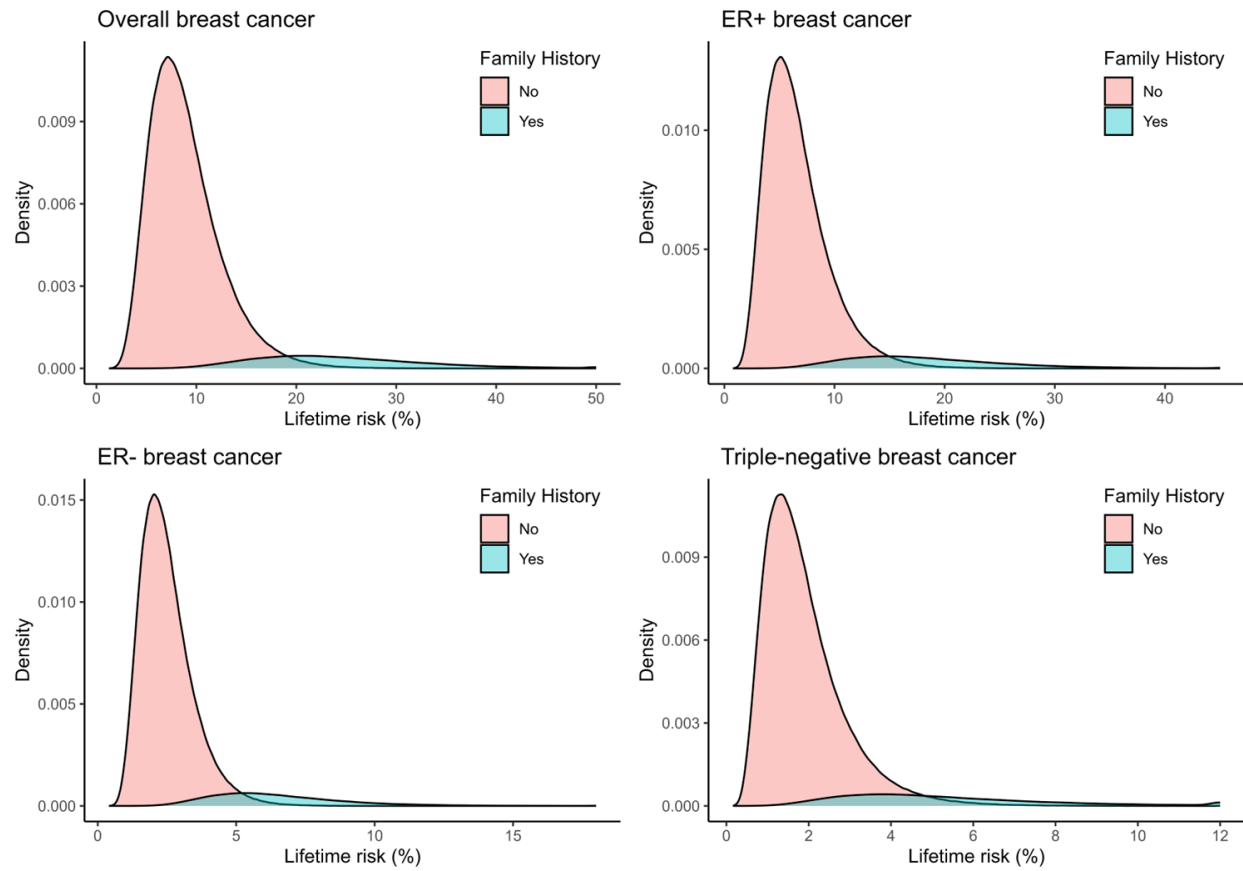


Figure S 2.11 Density plots for the lifetime absolute risks of developing each subtype of breast cancer by first-degree family history status.

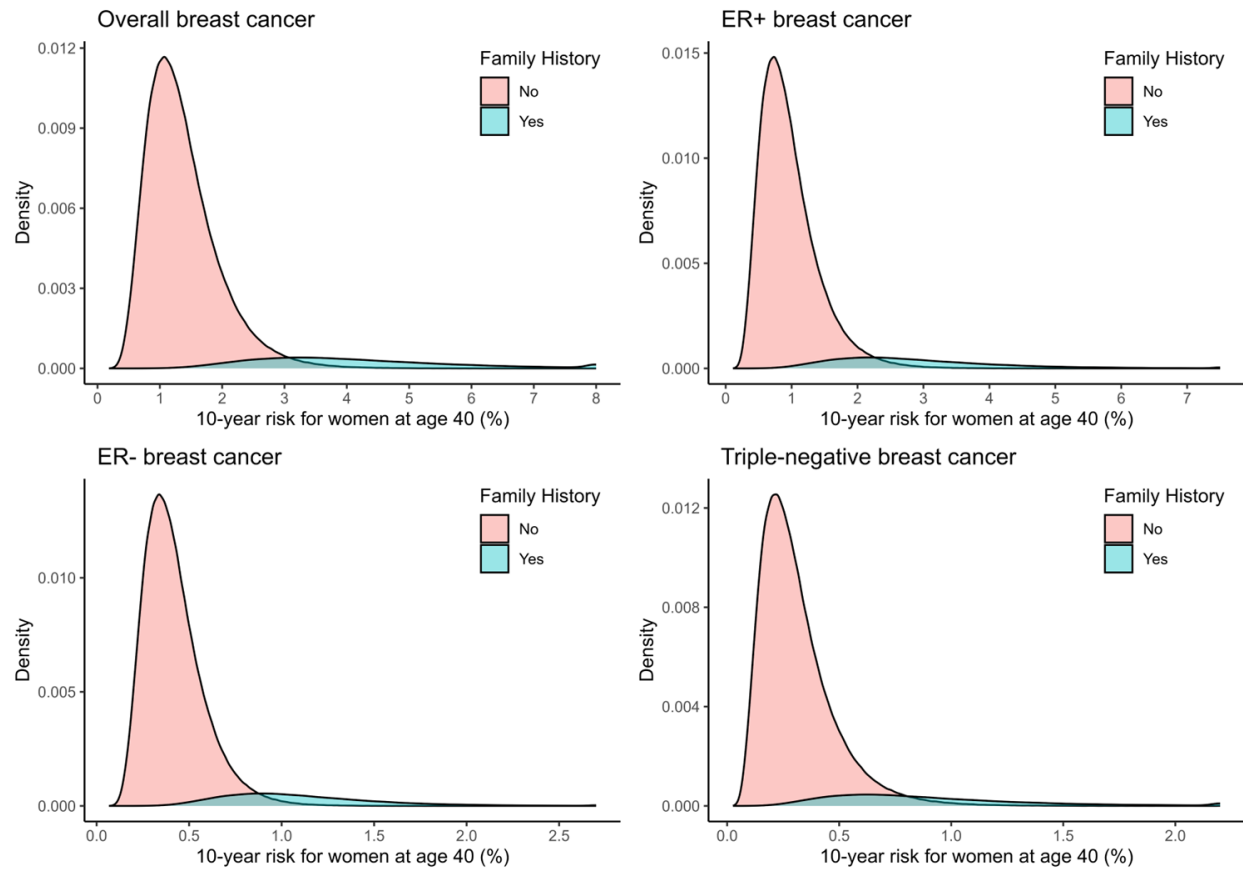


Figure S 2.12 Density plots for the 10-year absolute risks of developing each subtype of breast cancer by first-degree family history status.



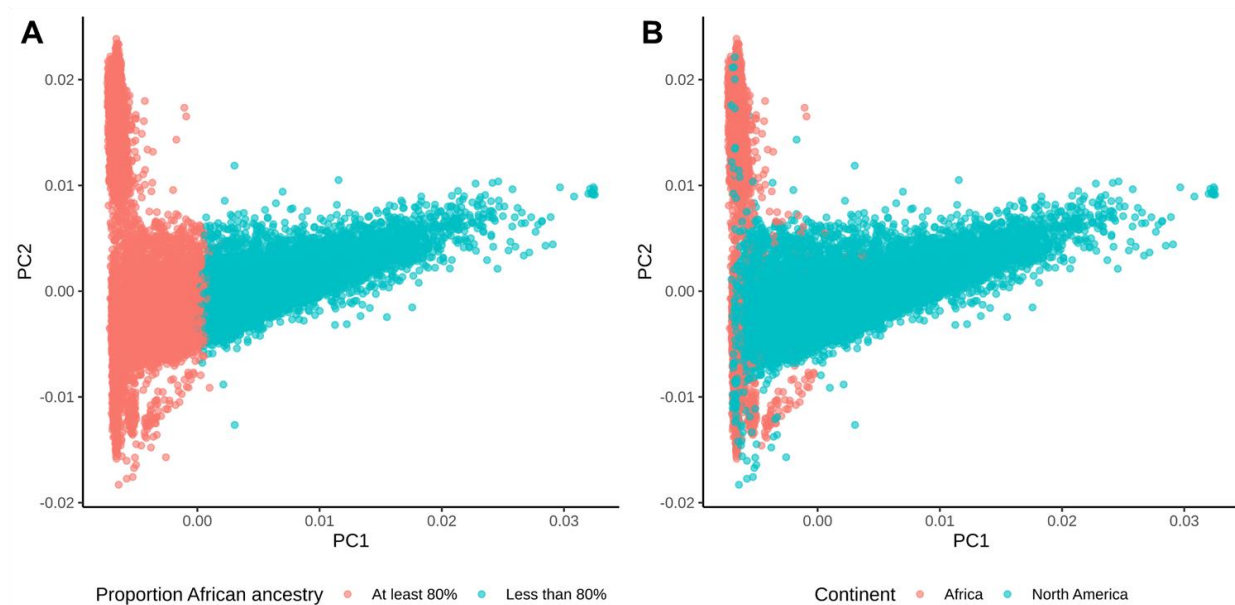


Figure S 2.13 Scatterplots of principal components 1 and 2. A) Color-coded by proportion of African ancestry B) Color-coded by continent where each participant resides.

Table S 2.1 Studies from AACBG included in development of polygenic risk score models for breast cancer in women of African ancestry.

| Genotyping Platform         | Study   | Training Set |             |             |       | Tuning Set |             |             |      | Validation Set |             |             |       |
|-----------------------------|---------|--------------|-------------|-------------|-------|------------|-------------|-------------|------|----------------|-------------|-------------|-------|
|                             |         | Cases        |             | Controls    |       | Cases      |             | Controls    |      | Cases          |             | Controls    |       |
|                             |         | Total        | ER-positive | ER-negative | TNBC  | Total      | ER-positive | ER-negative | TNBC | Total          | ER-positive | ER-negative | TNBC  |
| AABC                        | CARE    | 169          | 84          | 53          | 20    | 146        | 24          | 12          | 10   | 6              | 19          | 61          | 5     |
| AABC                        | CBCS    | 430          | 180         | 220         | 138   | 393        | 65          | 31          | 28   | 19             | 58          | 119         | 36    |
| AABC                        | MEC     | 405          | 227         | 106         | 35    | 623        | 56          | 31          | 11   | 4              | 97          | 117         | 6     |
| AABC                        | NBHS    | 179          | 93          | 23          | 13    | 105        | 18          | 7           | 4    | 3              | 18          | 58          | 4     |
| AABC                        | NC-BCFR | 266          | 153         | 92          | 25    | 31         | 41          | 25          | 11   | 4              | 6           | 76          | 11    |
| AABC                        | PLCO    | 40           | 9           | 6           | 2     | 84         | 6           | 1           | 0    | 0              | 8           | 8           | 0     |
| AABC                        | SFBCS   | 115          | 64          | 37          | 2     | 152        | 13          | 8           | 4    | 0              | 22          | 29          | 0     |
| AABC                        | WCHS    | 45           | 23          | 21          | 16    | 17         | 5           | 3           | 2    | 2              | 2           | 13          | 3     |
| AMBER                       | BWHS    | 221          | 152         | 50          | 36    | 1,479      | 23          | 19          | 1    | 0              | 203         | 63          | 6     |
| AMBER                       | CBCS    | 424          | 277         | 129         | 94    | 1          | 53          | 35          | 18   | 12             | 0           | 125         | 30    |
| AMBER                       | WCHS    | 348          | 244         | 103         | 62    | 177        | 43          | 30          | 13   | 11             | 17          | 81          | 15    |
| BCAC - OncoArray            | 2SISTER | 30           | 20          | 9           | 0     | 0          | 3           | 2           | 1    | 1              | 0           | 9           | 2     |
| BCAC - OncoArray            | CBCS    | 614          | 358         | 203         | 144   | 31         | 75          | 34          | 31   | 24             | 0           | 166         | 47    |
| BCAC - OncoArray            | MEC     | 406          | 281         | 105         | 54    | 404        | 75          | 44          | 28   | 17             | 65          | 124         | 20    |
| BCAC - OncoArray            | NBHS    | 34           | 9           | 8           | 5     | 39         | 7           | 0           | 2    | 1              | 9           | 10          | 2     |
| BCAC - OncoArray            | NC-BCFR | 47           | 24          | 16          | 11    | 0          | 9           | 3           | 5    | 2              | 0           | 13          | 3     |
| BCAC - OncoArray            | PLCO    | 18           | 8           | 2           | 2     | 44         | 4           | 3           | 0    | 0              | 9           | 2           | 0     |
| BCAC - OncoArray            | SISTER  | 83           | 51          | 15          | 12    | 123        | 17          | 9           | 6    | 4              | 12          | 30          | 3     |
| BCAC - OncoArray            | WAABCS  | 226          | 14          | 46          | 31    | 194        | 34          | 1           | 8    | 7              | 32          | 48          | 4     |
| BCAC - OncoArray            | USRT    | 18           | 0           | 0           | 0     | 26         | 3           | 0           | 0    | 0              | 6           | 5           | 0     |
| GBHS                        | GBHS    | 450          | 161         | 153         | 79    | 1,006      | 89          | 27          | 33   | 18             | 161         | 121         | 14    |
| MEGA                        | BWHS    | 899          | 525         | 238         | 129   | 1,308      | 138         | 77          | 35   | 23             | 210         | 245         | 39    |
| MEGA                        | CCPS    | 253          | 168         | 71          | 47    | 193        | 41          | 25          | 13   | 8              | 24          | 72          | 14    |
| MEGA                        | MDABCS  | 899          | 486         | 208         | 148   | 881        | 126         | 65          | 38   | 24             | 104         | 269         | 48    |
| MEGA                        | MEC     | 832          | 564         | 185         | 114   | 636        | 103         | 77          | 19   | 14             | 85          | 259         | 34    |
| MEGA                        | NBHS    | 490          | 43          | 117         | 59    | 272        | 61          | 3           | 12   | 8              | 27          | 144         | 15    |
| MEGA                        | NBHS    | 104          | 44          | 0           | 0     | 107        | 10          | 6           | 0    | 0              | 12          | 24          | 0     |
| MEGA                        | NC-BCFR | 138          | 80          | 42          | 30    | 161        | 18          | 9           | 5    | 3              | 24          | 29          | 2     |
| MEGA                        | NYUWHS  | 53           | 22          | 10          | 5     | 47         | 7           | 4           | 0    | 0              | 3           | 12          | 1     |
| MEGA                        | SCCS    | 497          | 187         | 69          | 36    | 458        | 50          | 22          | 10   | 6              | 78          | 161         | 8     |
| MEGA                        | STSBHS  | 470          | 388         | 50          | 27    | 490        | 65          | 57          | 3    | 0              | 70          | 157         | 9     |
| MEGA                        | WCHS    | 929          | 619         | 264         | 175   | 579        | 149         | 103         | 42   | 25             | 90          | 248         | 35    |
| ROOT                        | BBCS    | 61           | 30          | 29          | 0     | 73         | 13          | 5           | 6    | 0              | 12          | 20          | 0     |
| ROOT                        | BNCs    | 62           | 0           | 0           | 0     | 164        | 13          | 0           | 0    | 0              | 23          | 17          | 0     |
| ROOT                        | CCPS    | 258          | 117         | 91          | 61    | 263        | 38          | 19          | 10   | 8              | 41          | 69          | 15    |
| ROOT                        | NBHS    | 486          | 63          | 81          | 46    | 418        | 75          | 13          | 16   | 9              | 53          | 141         | 23    |
| ROOT                        | RVGBC   | 96           | 17          | 19          | 0     | 184        | 14          | 3           | 1    | 0              | 24          | 33          | 7     |
| ROOT                        | SCCS    | 90           | 54          | 25          | 16    | 219        | 12          | 7           | 3    | 0              | 26          | 24          | 3     |
| WGS                         | GBHS    | 202          | 69          | 90          | 56    | 103        | 24          | 9           | 9    | 5              | 15          | 67          | 8     |
| WGS                         | MEC     | 163          | 97          | 59          | 5     | 84         | 14          | 10          | 3    | 0              | 11          | 34          | 11    |
| WGS                         | NBHS    | 66           | 20          | 46          | 21    | 12         | 10          | 3           | 7    | 4              | 1           | 15          | 6     |
| WGS                         | SCCS    | 267          | 133         | 99          | 51    | 1,433      | 55          | 30          | 18   | 12             | 203         | 70          | 30    |
| WGS                         | STSBHS  | 291          | 117         | 174         | 121   | 0          | 40          | 19          | 21   | 14             | 0           | 90          | 36    |
| Total number of individuals |         | 12,174       | 6,275       | 3,364       | 1,937 | 13,160     | 1,739       | 891         | 487  | 298            | 1,880       | 3,478       | 525   |
|                             |         |              |             |             |       |            |             |             |      |                | 1,855       | 917         | 3,760 |

Table S 2.2 Performance of the European 330-variant PRS models when modifying coefficients based on the AABCG training set.

| <b>Phenotype</b> | <b>Original<br/>coefficients (all<br/>available variants)</b> | <b>Coefficients from<br/>training set GWAS<br/>(all available<br/>variants)</b> | <b>Coefficients from<br/>training set GWAS<br/>(only variants with<br/>directionally<br/>consistent<br/>coefficients as<br/>original)</b> | <b>Number of<br/>variants in this<br/>modified EUR-330<br/>PRS with<br/>directionally<br/>consistent<br/>coefficients</b> |
|------------------|---------------------------------------------------------------|---------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|
| Overall          | 0.569 (0.556, 0.583)                                          | 0.573 (0.560, 0.586)                                                            | 0.576 (0.562, 0.589)                                                                                                                      | 205                                                                                                                       |
| ER-positive      | 0.583 (0.567, 0.599)                                          | 0.579 (0.562, 0.595)                                                            | 0.584 (0.567, 0.599)                                                                                                                      | 207                                                                                                                       |
| ER-negative      | 0.577 (0.555, 0.599)                                          | 0.595 (0.574, 0.616)                                                            | 0.597 (0.576, 0.617)                                                                                                                      | 162                                                                                                                       |
| TNBC             | 0.597 (0.571, 0.622)                                          | 0.605 (0.576, 0.631)                                                            | 0.616 (0.591, 0.642)                                                                                                                      | 160                                                                                                                       |

Table S 2.3 Association of PRS scores and African ancestry with each breast cancer subtype

| Phenotype                                                                                                                        | Overall BC Risk   | ER-positive       | ER-negative       | TNBC              |
|----------------------------------------------------------------------------------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|
| <b>Association of PRS and cancer risk by continent</b>                                                                           |                   |                   |                   |                   |
| PRS in individuals from the African continent                                                                                    | 1.59 (1.31, 1.93) | 1.74 (1.30, 2.34) | 1.33 (1.06, 1.66) | 1.41 (1.06, 1.89) |
| PRS in individuals from the North American continent                                                                             | 1.46 (1.37, 1.54) | 1.50 (1.41, 1.60) | 1.49 (1.38, 1.62) | 1.64 (1.48, 1.82) |
| <i>P for testing interaction between PRS and continent</i>                                                                       | 0.008             | 0.023             | 0.80              | 0.46              |
| <b>Association of PRS and cancer risk by proportion African ancestry</b>                                                         |                   |                   |                   |                   |
| PRS in individuals with $\geq 80\%$ African ancestry                                                                             | 1.52 (1.41, 1.64) | 1.60 (1.47, 1.74) | 1.48 (1.34, 1.64) | 1.64 (1.45, 1.87) |
| PRS in individuals with $< 80\%$ African ancestry                                                                                | 1.45 (1.33, 1.57) | 1.45 (1.33, 1.58) | 1.52 (1.36, 1.71) | 1.65 (1.42, 1.92) |
| <i>P for testing interaction between PRS and proportion African ancestry (<math>\geq 80\%</math> vs. <math>&lt; 80\%</math>)</i> | 0.44              | 0.15              | 0.71              | 0.99              |
| <i>P for testing interaction between PRS and proportion African ancestry (continuous variable)</i>                               | 0.39              | 0.38              | 0.76              | 0.71              |

Footnote: PRS score is per SD

Table S 2.4 Association of PRS scores and age with each breast cancer phenotype

| Phenotype                                                                          | Overall BC Risk   | ER-positive       | ER-negative       | TNBC              |
|------------------------------------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|
| <b>Association of PRS and cancer risk by age group</b>                             |                   |                   |                   |                   |
| <50 years of age                                                                   | 1.51 (1.38, 1.66) | 1.54 (1.38, 1.71) | 1.50 (1.32, 1.69) | 1.73 (1.48, 2.03) |
| 50+ years of age                                                                   | 1.47 (1.37, 1.58) | 1.53 (1.42, 1.64) | 1.49 (1.35, 1.64) | 1.60 (1.42, 1.82) |
| <i>P for testing interaction between PRS and age (&lt;50 vs. 50+ years of age)</i> | 0.52              | 0.79              | 0.98              | 0.67              |
| <i>P for testing interaction between PRS and age (continuous variable)</i>         | 0.95              | 0.83              | 0.92              | 0.25              |

Footnote: PRS score is per SD

Table S 2.5 Lifetime absolute risks for the best performing PRS models for each breast cancer phenotype by family history.

| Phenotype      | Overall breast cancer                |                                | ER-positive                          |                                | ER-negative                          |                                | TNBC                                 |                                |
|----------------|--------------------------------------|--------------------------------|--------------------------------------|--------------------------------|--------------------------------------|--------------------------------|--------------------------------------|--------------------------------|
| PRS Percentile | Positive first-degree family history | No first-degree family history | Positive first-degree family history | No first-degree family history | Positive first-degree family history | No first-degree family history | Positive first-degree family history | No first-degree family history |
| 0-1%           | 8.46%                                | 2.89%                          | 5.63%                                | 1.93%                          | 2.26%                                | 0.87%                          | 1.26%                                | 0.45%                          |
| 1-5%           | 11.37%                               | 3.88%                          | 7.78%                                | 2.66%                          | 3.00%                                | 1.15%                          | 1.85%                                | 0.64%                          |
| 5-10%          | 13.54%                               | 4.67%                          | 9.44%                                | 3.25%                          | 3.55%                                | 1.37%                          | 2.32%                                | 0.81%                          |
| 10-20%         | 15.75%                               | 5.49%                          | 11.16%                               | 3.87%                          | 4.12%                                | 1.59%                          | 2.83%                                | 0.99%                          |
| 20-40%         | 19.07%                               | 6.73%                          | 13.82%                               | 4.84%                          | 4.98%                                | 1.92%                          | 3.66%                                | 1.29%                          |
| 40-60%         | 23.08%                               | 8.30%                          | 17.10%                               | 6.08%                          | 6.04%                                | 2.34%                          | 4.74%                                | 1.68%                          |
| 60-80%         | 27.80%                               | 10.21%                         | 21.11%                               | 7.62%                          | 7.33%                                | 2.85%                          | 6.16%                                | 2.19%                          |
| 80-90%         | 32.99%                               | 12.43%                         | 25.64%                               | 9.46%                          | 8.80%                                | 3.45%                          | 7.89%                                | 2.82%                          |
| 90-95%         | 37.52%                               | 14.48%                         | 29.74%                               | 11.18%                         | 10.15%                               | 3.99%                          | 9.56%                                | 3.44%                          |
| 95-99%         | 43.30%                               | 17.27%                         | 35.13%                               | 13.58%                         | 11.99%                               | 4.75%                          | 11.96%                               | 4.34%                          |
| 99-100%        | 53.63%                               | 22.81%                         | 45.26%                               | 18.50%                         | 15.70%                               | 6.29%                          | 17.16%                               | 6.35%                          |

Table S 2.6 10-year absolute risks at age 40 for the best performing PRS models for each breast cancer phenotype by family history.

| Phenotype      | Overall breast cancer                |                                | ER-positive                          |                                | ER-negative                          |                                | TNBC                                 |                                |
|----------------|--------------------------------------|--------------------------------|--------------------------------------|--------------------------------|--------------------------------------|--------------------------------|--------------------------------------|--------------------------------|
| PRS Percentile | Positive first-degree family history | No first-degree family history | Positive first-degree family history | No first-degree family history | Positive first-degree family history | No first-degree family history | Positive first-degree family history | No first-degree family history |
| 0-1%           | 1.28%                                | 0.43%                          | 0.81%                                | 0.27%                          | 0.38%                                | 0.14%                          | 0.21%                                | 0.07%                          |
| 1-5%           | 1.75%                                | 0.58%                          | 1.14%                                | 0.38%                          | 0.50%                                | 0.19%                          | 0.30%                                | 0.11%                          |
| 5-10%          | 2.11%                                | 0.70%                          | 1.39%                                | 0.46%                          | 0.59%                                | 0.23%                          | 0.38%                                | 0.13%                          |
| 10-20%         | 2.49%                                | 0.82%                          | 1.66%                                | 0.55%                          | 0.69%                                | 0.26%                          | 0.47%                                | 0.16%                          |
| 20-40%         | 3.07%                                | 1.01%                          | 2.08%                                | 0.69%                          | 0.84%                                | 0.32%                          | 0.61%                                | 0.21%                          |
| 40-60%         | 3.80%                                | 1.26%                          | 2.62%                                | 0.88%                          | 1.02%                                | 0.39%                          | 0.79%                                | 0.28%                          |
| 60-80%         | 4.71%                                | 1.56%                          | 3.32%                                | 1.11%                          | 1.25%                                | 0.48%                          | 1.03%                                | 0.36%                          |
| 80-90%         | 5.78%                                | 1.93%                          | 4.14%                                | 1.39%                          | 1.51%                                | 0.58%                          | 1.33%                                | 0.47%                          |
| 90-95%         | 6.79%                                | 2.27%                          | 4.93%                                | 1.66%                          | 1.75%                                | 0.67%                          | 1.63%                                | 0.57%                          |
| 95-99%         | 8.19%                                | 2.75%                          | 6.04%                                | 2.04%                          | 2.09%                                | 0.80%                          | 2.07%                                | 0.72%                          |
| 99-100%        | 11.12%                               | 3.76%                          | 8.45%                                | 2.87%                          | 2.79%                                | 1.07%                          | 3.05%                                | 1.07%                          |

Table S 2.7 Proportion of genome-wide heritability explained by each PRS model for each breast cancer phenotype

| Phenotype                                | Overall       | ER-positive   | ER-negative   | TNBC          |
|------------------------------------------|---------------|---------------|---------------|---------------|
| <b>Genome-wide h2 (observed scale)</b>   | 0.556 (0.136) | 0.716 (0.187) | 0.535 (0.286) | 0.583 (0.387) |
| <b>PRS h2 (observed scale)</b>           | 0.163 (0.022) | 0.189 (0.026) | 0.157 (0.030) | 0.254 (0.049) |
| <b>Proportion of h2 explained by PRS</b> | 29.2%         | 26.4%         | 29.4%         | 43.6%         |

Footnote: Standard errors of heritability estimates are included in parentheses



Table S 2.8 External validation of PRS models for overall breast cancer

| Studies for external validation                          | All of Us                    |                                |                              |                                 | Additional women of African ancestry from WFBC, BioVU, and BEST |
|----------------------------------------------------------|------------------------------|--------------------------------|------------------------------|---------------------------------|-----------------------------------------------------------------|
| Ancestry of validation population                        | African                      | European                       | Hispanic                     | All women                       | African                                                         |
| Number of cases/controls                                 | 834 cases<br>29,122 controls | 4,183 cases<br>68,427 controls | 682 cases<br>26,139 controls | 6,509 cases<br>139,054 controls | 580 cases<br>3,011 controls                                     |
| <b><u>European ancestry PRS (reference)</u></b>          |                              |                                |                              |                                 |                                                                 |
| EUR 330-variant PRS                                      | 0.568 (0.550, 0.587)         | 0.610 (0.601, 0.618)           | 0.574 (0.553, 0.597)         | 0.599 (0.592, 0.606)            | 0.563 (0.529, 0.595)                                            |
| <b><u>Best performing PRS in internal validation</u></b> |                              |                                |                              |                                 |                                                                 |
| Three-subtype<br>Proportion-tuned PRS                    | <b>0.612 (0.592, 0.631)</b>  | 0.607 (0.598, 0.615)           | 0.592 (0.572, 0.611)         | 0.607 (0.601, 0.615)            | <b>0.612 (0.581, 0.643)</b>                                     |
| <b><u>Ensemble-based PRS models</u></b>                  |                              |                                |                              |                                 |                                                                 |
| FSS                                                      | 0.613 (0.594, 0.632)         | 0.602 (0.593, 0.610)           | 0.590 (0.570, 0.610)         | 0.604 (0.596, 0.611)            | 0.611 (0.579, 0.640)                                            |
| Elastic net regression                                   | 0.613 (0.594, 0.632)         | 0.607 (0.598, 0.616)           | 0.591 (0.572, 0.612)         | 0.608 (0.601, 0.615)            | 0.606 (0.576, 0.637)                                            |
| <b><u>xAncestry PRS models</u></b>                       |                              |                                |                              |                                 |                                                                 |
| FSS (1e-8)                                               | 0.572 (0.551, 0.590)         | 0.570 (0.562, 0.579)           | 0.561 (0.540, 0.583)         | 0.571 (0.564, 0.579)            | 0.577 (0.544, 0.608)                                            |
| FSS (1e-7)                                               | 0.570 (0.550, 0.589)         | 0.571 (0.562, 0.579)           | 0.563 (0.542, 0.584)         | 0.572 (0.564, 0.579)            | 0.578 (0.546, 0.608)                                            |
| FSS (1e-6)                                               | 0.574 (0.553, 0.592)         | 0.576 (0.567, 0.584)           | 0.566 (0.545, 0.588)         | 0.576 (0.569, 0.583)            | 0.580 (0.549, 0.611)                                            |
| FSS (1e-5)                                               | 0.577 (0.557, 0.596)         | 0.577 (0.568, 0.585)           | 0.567 (0.546, 0.589)         | 0.578 (0.570, 0.584)            | 0.579 (0.548, 0.611)                                            |
| FSS (1e-4)                                               | 0.573 (0.551, 0.591)         | 0.578 (0.569, 0.586)           | 0.566 (0.545, 0.587)         | 0.578 (0.571, 0.585)            | 0.577 (0.544, 0.607)                                            |
| FSS (0.001)                                              | 0.577 (0.557, 0.595)         | 0.578 (0.569, 0.587)           | 0.564 (0.544, 0.586)         | 0.578 (0.571, 0.586)            | 0.578 (0.546, 0.609)                                            |
| FSS (0.01)                                               | 0.584 (0.563, 0.602)         | 0.579 (0.570, 0.588)           | 0.566 (0.546, 0.588)         | 0.580 (0.573, 0.587)            | 0.574 (0.541, 0.608)                                            |
| FSS (0.05)                                               | 0.587 (0.567, 0.606)         | 0.580 (0.572, 0.589)           | 0.567 (0.546, 0.588)         | 0.582 (0.575, 0.589)            | 0.582 (0.550, 0.612)                                            |
| FSS (0.1)                                                | 0.593 (0.573, 0.612)         | 0.578 (0.569, 0.587)           | 0.564 (0.543, 0.584)         | 0.581 (0.574, 0.588)            | 0.580 (0.549, 0.608)                                            |
| Lassosum2                                                | 0.572 (0.552, 0.591)         | 0.566 (0.557, 0.575)           | 0.557 (0.536, 0.579)         | 0.568 (0.561, 0.575)            | 0.575 (0.545, 0.604)                                            |
| LDpred2-auto                                             | 0.603 (0.583, 0.622)         | 0.576 (0.567, 0.585)           | 0.571 (0.551, 0.592)         | 0.581 (0.575, 0.589)            | 0.605 (0.573, 0.635)                                            |
| PLINK C+T (0.001)                                        | 0.584 (0.564, 0.602)         | 0.577 (0.568, 0.586)           | 0.567 (0.546, 0.588)         | 0.579 (0.572, 0.586)            | 0.580 (0.549, 0.611)                                            |
| PLINK C+T (0.01)                                         | 0.592 (0.572, 0.611)         | 0.579 (0.571, 0.588)           | 0.573 (0.553, 0.594)         | 0.582 (0.575, 0.589)            | 0.590 (0.557, 0.622)                                            |
| PLINK C+T (0.05)                                         | 0.595 (0.575, 0.612)         | 0.578 (0.570, 0.587)           | 0.569 (0.548, 0.588)         | 0.582 (0.574, 0.588)            | 0.601 (0.569, 0.631)                                            |
| PLINK C+T (0.1)                                          | 0.594 (0.574, 0.613)         | 0.579 (0.570, 0.587)           | 0.567 (0.548, 0.588)         | 0.582 (0.574, 0.588)            | 0.601 (0.570, 0.631)                                            |

Continuation of Table S 2.8

|                                                |                      |                      |                      |                      |                      |
|------------------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| PLINK C+T (0.2)                                | 0.595 (0.575, 0.613) | 0.579 (0.570, 0.588) | 0.566 (0.546, 0.586) | 0.582 (0.575, 0.588) | 0.596 (0.565, 0.626) |
| PLINK C+T (0.3)                                | 0.595 (0.575, 0.613) | 0.579 (0.570, 0.588) | 0.567 (0.548, 0.587) | 0.582 (0.575, 0.589) | 0.593 (0.561, 0.622) |
| PLINK C+T (0.4)                                | 0.594 (0.573, 0.612) | 0.579 (0.570, 0.588) | 0.567 (0.547, 0.587) | 0.582 (0.575, 0.588) | 0.593 (0.562, 0.625) |
| PLINK C+T (0.5)                                | 0.594 (0.573, 0.612) | 0.579 (0.570, 0.587) | 0.566 (0.547, 0.587) | 0.581 (0.574, 0.588) | 0.595 (0.564, 0.625) |
| PRSe-2                                         | 0.598 (0.578, 0.615) | 0.580 (0.571, 0.588) | 0.568 (0.548, 0.589) | 0.582 (0.576, 0.589) | 0.595 (0.563, 0.624) |
| PRS-CSx                                        | 0.610 (0.590, 0.629) | 0.608 (0.600, 0.617) | 0.588 (0.568, 0.608) | 0.610 (0.603, 0.617) | 0.619 (0.589, 0.650) |
| CT-SLEB                                        | 0.608 (0.588, 0.627) | 0.608 (0.599, 0.616) | 0.594 (0.575, 0.614) | 0.608 (0.601, 0.615) | 0.599 (0.569, 0.630) |
| <b>PRS models built using existing methods</b> |                      |                      |                      |                      |                      |
| FSS (1e-8)                                     | 0.529 (0.510, 0.549) | 0.510 (0.500, 0.518) | 0.519 (0.497, 0.541) | 0.514 (0.506, 0.521) | 0.546 (0.514, 0.576) |
| FSS (1e-7)                                     | 0.527 (0.506, 0.548) | 0.511 (0.502, 0.521) | 0.524 (0.502, 0.546) | 0.516 (0.509, 0.523) | 0.553 (0.521, 0.586) |
| FSS (1e-6)                                     | 0.536 (0.515, 0.554) | 0.522 (0.513, 0.531) | 0.539 (0.517, 0.561) | 0.528 (0.521, 0.535) | 0.562 (0.530, 0.593) |
| FSS (1e-5)                                     | 0.539 (0.519, 0.558) | 0.517 (0.508, 0.526) | 0.525 (0.501, 0.548) | 0.523 (0.515, 0.530) | 0.536 (0.504, 0.569) |
| FSS (1e-4)                                     | 0.508 (0.489, 0.529) | 0.514 (0.505, 0.523) | 0.519 (0.496, 0.542) | 0.515 (0.509, 0.522) | 0.530 (0.498, 0.561) |
| FSS (0.001)                                    | 0.530 (0.508, 0.551) | 0.510 (0.501, 0.519) | 0.503 (0.481, 0.526) | 0.513 (0.506, 0.520) | 0.529 (0.499, 0.561) |
| FSS (0.01)                                     | 0.539 (0.520, 0.560) | 0.519 (0.509, 0.528) | 0.521 (0.500, 0.542) | 0.522 (0.516, 0.529) | 0.514 (0.483, 0.547) |
| FSS (0.05)                                     | 0.548 (0.528, 0.568) | 0.523 (0.513, 0.532) | 0.517 (0.495, 0.539) | 0.527 (0.520, 0.534) | 0.535 (0.502, 0.566) |
| FSS (0.1)                                      | 0.561 (0.542, 0.580) | 0.517 (0.508, 0.527) | 0.506 (0.484, 0.527) | 0.524 (0.517, 0.531) | 0.529 (0.495, 0.560) |
| Lassosum2                                      | 0.545 (0.525, 0.566) | 0.518 (0.509, 0.527) | 0.519 (0.497, 0.542) | 0.524 (0.517, 0.531) | 0.546 (0.514, 0.579) |
| LDpred2-auto                                   | 0.584 (0.564, 0.603) | 0.528 (0.519, 0.536) | 0.539 (0.517, 0.562) | 0.538 (0.532, 0.546) | 0.583 (0.551, 0.614) |
| PLINK C+T (0.001)                              | 0.543 (0.522, 0.563) | 0.513 (0.504, 0.521) | 0.522 (0.499, 0.545) | 0.521 (0.514, 0.528) | 0.540 (0.509, 0.571) |
| PLINK C+T (0.01)                               | 0.558 (0.538, 0.579) | 0.521 (0.513, 0.530) | 0.537 (0.516, 0.559) | 0.528 (0.521, 0.534) | 0.554 (0.523, 0.585) |
| PLINK C+T (0.05)                               | 0.565 (0.545, 0.585) | 0.516 (0.507, 0.524) | 0.525 (0.504, 0.548) | 0.525 (0.518, 0.532) | 0.572 (0.541, 0.602) |
| PLINK C+T (0.1)                                | 0.563 (0.544, 0.584) | 0.517 (0.508, 0.525) | 0.519 (0.499, 0.542) | 0.524 (0.518, 0.531) | 0.571 (0.542, 0.602) |
| PLINK C+T (0.2)                                | 0.564 (0.544, 0.585) | 0.518 (0.509, 0.527) | 0.518 (0.497, 0.541) | 0.526 (0.519, 0.533) | 0.564 (0.532, 0.598) |
| PLINK C+T (0.3)                                | 0.564 (0.544, 0.585) | 0.519 (0.510, 0.527) | 0.520 (0.499, 0.543) | 0.526 (0.520, 0.533) | 0.558 (0.526, 0.589) |
| PLINK C+T (0.4)                                | 0.563 (0.543, 0.584) | 0.518 (0.509, 0.526) | 0.520 (0.499, 0.543) | 0.525 (0.519, 0.532) | 0.559 (0.528, 0.591) |
| PLINK C+T (0.5)                                | 0.564 (0.543, 0.584) | 0.518 (0.509, 0.526) | 0.519 (0.498, 0.543) | 0.525 (0.519, 0.532) | 0.560 (0.526, 0.592) |
| PRSe-2                                         | 0.568 (0.549, 0.589) | 0.518 (0.509, 0.527) | 0.522 (0.499, 0.543) | 0.526 (0.519, 0.533) | 0.558 (0.526, 0.589) |
| PRS-CSx                                        | 0.594 (0.574, 0.614) | 0.599 (0.589, 0.607) | 0.578 (0.558, 0.600) | 0.599 (0.592, 0.606) | 0.603 (0.569, 0.636) |
| CT-SLEB                                        | 0.606 (0.587, 0.625) | 0.606 (0.598, 0.615) | 0.594 (0.574, 0.613) | 0.606 (0.600, 0.614) | 0.596 (0.565, 0.627) |

Table S 2.9 External validation of PRS models for breast cancer subtypes among additional women of African ancestry from WFBC, BioVU, and BEST.

| Subtype                                         | ER-positive                 | ER-negative                 | TNBC                        |
|-------------------------------------------------|-----------------------------|-----------------------------|-----------------------------|
| Number of cases/controls                        | 281 cases<br>3,011 controls | 155 cases<br>3,011 controls | 101 cases<br>3,011 controls |
| <b><u>European ancestry PRS (reference)</u></b> |                             |                             |                             |
| EUR 330-variant PRS                             | 0.580 (0.534, 0.625)        | 0.567 (0.512, 0.622)        | 0.606 (0.542, 0.671)        |
| <b><u>Ensemble-based PRS models</u></b>         |                             |                             |                             |
| FSS                                             | 0.640 (0.597, 0.682)        | 0.579 (0.526, 0.629)        | 0.627 (0.562, 0.688)        |
| Elastic net regression                          | 0.638 (0.593, 0.680)        | 0.600 (0.547, 0.652)        | 0.627 (0.566, 0.691)        |
| <b><u>xAncestry PRS models</u></b>              |                             |                             |                             |
| FSS (1e-8)                                      | 0.578 (0.533, 0.621)        | 0.590 (0.536, 0.645)        | 0.655 (0.594, 0.715)        |
| FSS (1e-7)                                      | 0.576 (0.532, 0.620)        | 0.592 (0.541, 0.641)        | 0.652 (0.589, 0.712)        |
| FSS (1e-6)                                      | 0.577 (0.535, 0.622)        | 0.583 (0.530, 0.635)        | 0.647 (0.587, 0.702)        |
| FSS (1e-5)                                      | 0.574 (0.530, 0.618)        | 0.576 (0.525, 0.626)        | 0.629 (0.569, 0.686)        |
| FSS (1e-4)                                      | 0.586 (0.541, 0.629)        | 0.584 (0.528, 0.636)        | 0.646 (0.585, 0.708)        |
| FSS (0.001)                                     | 0.593 (0.553, 0.635)        | 0.579 (0.526, 0.635)        | 0.620 (0.554, 0.683)        |
| FSS (0.01)                                      | 0.582 (0.538, 0.625)        | 0.577 (0.522, 0.630)        | 0.608 (0.542, 0.671)        |
| FSS (0.05)                                      | 0.589 (0.548, 0.634)        | 0.581 (0.530, 0.632)        | 0.616 (0.560, 0.671)        |
| FSS (0.1)                                       | 0.587 (0.546, 0.629)        | 0.577 (0.530, 0.630)        | 0.610 (0.555, 0.671)        |
| Lassosum2                                       | 0.584 (0.540, 0.629)        | 0.577 (0.523, 0.631)        | 0.642 (0.585, 0.701)        |
| LDpred2-auto                                    | 0.608 (0.567, 0.649)        | 0.610 (0.556, 0.662)        | 0.648 (0.589, 0.708)        |
| PLINK C+T (0.001)                               | 0.603 (0.556, 0.648)        | 0.586 (0.534, 0.640)        | 0.615 (0.554, 0.677)        |
| PLINK C+T (0.01)                                | 0.603 (0.563, 0.644)        | 0.593 (0.545, 0.647)        | 0.626 (0.565, 0.683)        |
| PLINK C+T (0.05)                                | 0.605 (0.563, 0.648)        | 0.601 (0.547, 0.649)        | 0.633 (0.571, 0.697)        |
| PLINK C+T (0.1)                                 | 0.598 (0.557, 0.640)        | 0.599 (0.546, 0.647)        | 0.634 (0.575, 0.693)        |

Continuation of Table S 2.9

|                                                |                      |                      |                      |
|------------------------------------------------|----------------------|----------------------|----------------------|
| PLINK C+T (0.2)                                | 0.603 (0.560, 0.645) | 0.595 (0.544, 0.645) | 0.649 (0.589, 0.704) |
| PLINK C+T (0.3)                                | 0.600 (0.555, 0.642) | 0.597 (0.546, 0.648) | 0.646 (0.592, 0.704) |
| PLINK C+T (0.4)                                | 0.601 (0.557, 0.643) | 0.597 (0.547, 0.649) | 0.648 (0.590, 0.703) |
| PLINK C+T (0.5)                                | 0.600 (0.554, 0.642) | 0.598 (0.547, 0.653) | 0.648 (0.589, 0.703) |
| PRSice-2                                       | 0.606 (0.563, 0.648) | 0.600 (0.550, 0.652) | 0.652 (0.596, 0.709) |
| PRS-CSx                                        | 0.631 (0.587, 0.674) | 0.599 (0.550, 0.647) | 0.638 (0.578, 0.692) |
| CT-SLEB                                        | 0.635 (0.591, 0.676) | 0.603 (0.546, 0.659) | 0.664 (0.605, 0.717) |
| <b>PRS models built using existing methods</b> |                      |                      |                      |
| FSS (1e-8)                                     | 0.539 (0.493, 0.586) | 0.594 (0.541, 0.646) | 0.621 (0.561, 0.681) |
| FSS (1e-7)                                     | 0.527 (0.481, 0.572) | 0.578 (0.529, 0.629) | 0.615 (0.561, 0.673) |
| FSS (1e-6)                                     | 0.533 (0.488, 0.578) | 0.543 (0.492, 0.595) | 0.607 (0.548, 0.666) |
| FSS (1e-5)                                     | 0.523 (0.478, 0.565) | 0.521 (0.470, 0.571) | 0.553 (0.490, 0.617) |
| FSS (1e-4)                                     | 0.542 (0.500, 0.590) | 0.516 (0.462, 0.570) | 0.577 (0.512, 0.642) |
| FSS (0.001)                                    | 0.560 (0.514, 0.606) | 0.492 (0.442, 0.542) | 0.502 (0.438, 0.572) |
| FSS (0.01)                                     | 0.523 (0.478, 0.565) | 0.504 (0.453, 0.552) | 0.476 (0.417, 0.540) |
| FSS (0.05)                                     | 0.540 (0.496, 0.582) | 0.528 (0.474, 0.581) | 0.518 (0.460, 0.584) |
| FSS (0.1)                                      | 0.537 (0.495, 0.583) | 0.528 (0.474, 0.583) | 0.531 (0.469, 0.596) |
| Lassosum2                                      | 0.554 (0.509, 0.598) | 0.526 (0.474, 0.577) | 0.617 (0.556, 0.680) |
| LDpred2-auto                                   | 0.596 (0.553, 0.635) | 0.588 (0.536, 0.640) | 0.597 (0.534, 0.660) |
| PLINK C+T (0.001)                              | 0.578 (0.534, 0.622) | 0.540 (0.487, 0.590) | 0.503 (0.429, 0.572) |
| PLINK C+T (0.01)                               | 0.580 (0.535, 0.626) | 0.551 (0.501, 0.608) | 0.536 (0.477, 0.600) |
| PLINK C+T (0.05)                               | 0.576 (0.530, 0.618) | 0.564 (0.508, 0.620) | 0.561 (0.494, 0.624) |
| PLINK C+T (0.1)                                | 0.567 (0.525, 0.612) | 0.559 (0.507, 0.614) | 0.566 (0.506, 0.627) |
| PLINK C+T (0.2)                                | 0.575 (0.529, 0.622) | 0.551 (0.497, 0.607) | 0.580 (0.516, 0.640) |
| PLINK C+T (0.3)                                | 0.570 (0.525, 0.614) | 0.556 (0.502, 0.609) | 0.577 (0.514, 0.640) |

|                 |                      |                      |                      |
|-----------------|----------------------|----------------------|----------------------|
| PLINK C+T (0.4) | 0.571 (0.528, 0.614) | 0.556 (0.501, 0.612) | 0.579 (0.519, 0.638) |
| PLINK C+T (0.5) | 0.570 (0.526, 0.614) | 0.557 (0.504, 0.612) | 0.579 (0.518, 0.641) |
| PRSize-2        | 0.582 (0.541, 0.627) | 0.560 (0.502, 0.615) | 0.579 (0.513, 0.636) |
| PRS-CSx         | 0.625 (0.581, 0.666) | 0.562 (0.509, 0.613) | 0.569 (0.509, 0.627) |
| CT-SLEB         | 0.634 (0.593, 0.677) | 0.601 (0.549, 0.656) | 0.663 (0.609, 0.720) |

### **3 CHAPTER 3: THE IMPACT OF ARSENIC EXPOSURE ON DNA METHYLATION IN HUMANS: BUILDING AN EPIGENETIC BIOMARKER OF EXPOSURE AND TOXICITY RISK**

#### **3.1 ABSTRACT**

Inorganic arsenic (As) is a carcinogen, and >200 million people worldwide are exposed to As through drinking water. Long-term As exposure increases the risk of chronic diseases including cancer and cardiovascular disease. One proposed As toxicity mechanism is epigenetic dysregulation, including disruption of DNA methylation (DNAm). While prior work has identified several sites in the epigenome associated with As exposure, our knowledge remains limited. To improve our understanding of the epigenetic effects of As exposure, as well as develop an effective tool to reliably track As exposure effects and toxicity risk, we aimed to discover new As-associated regions and build a robust epigenetic biomarker of As exposure. First, we evaluated the association of As exposure (measured in urine and drinking water) with DNAm (measured in blood cells at >720,000 CpG sites across the genome) among 1,186 adults from the Health Effects of Arsenic Longitudinal Study in Bangladesh, a population with a wide range of As exposure levels. We identified 1,177 CpGs associated with urinary As (FDR<0.05), and further demonstrated that these associations are likely causal using genetic determinants of arsenic metabolism as instrumental variables. As-associated CpGs showed strong enrichment for genomic context and CpG sets related to several disease states previously linked to As exposure in epidemiological studies. We then built a 255-CpG DNAm-based biomarker of As exposure using elastic net regression that was highly predictive of urinary As levels ( $r^2=0.46$ ) and arsenical skin lesion status (AUC=0.69), a common sign of As toxicity. This biomarker was highly

associated with measured urinary As levels in an independent Bangladeshi cohort ( $p=1.3 \times 10^{-19}$ ) and further validated in the Strong Heart Study ( $p=6.7 \times 10^{-6}$ ) where DNAm-predicted As was also associated with overall mortality ( $p=4.6 \times 10^{-4}$ ). We demonstrate the utility of a DNAm-based biomarker of As exposure and toxicity and present guidelines for developing effective epigenetic biomarkers to track environmental chemical exposures.

### 3.2 INTRODUCTION

Chemicals that adversely affect human health are abundant in our environment, but there is currently a lack of effective tools to reliably track these exposures and their effects in the body. Environmental chemicals, including metals and persistent organic pollutants, induce toxicity through various mechanisms including the disruption of DNA methylation (DNAm).<sup>132,133</sup> While such DNAm-associated changes may reflect levels of chemical exposure in the human body, thus far, our ability to fully characterize their effects on DNAm and to utilize them as reliable indicators of exposure has been greatly limited by the relatively small number of sites in the epigenome linked to these chemical exposures.<sup>26,134,135</sup> With the exception of lead,<sup>136</sup> no epigenetic biomarkers currently exist to effectively track chemical exposures in humans with the sensitivity and specificity required to reflect toxicity risk.

In this study, we aimed to demonstrate that exposure to environmental chemicals can be reliably tracked in the human epigenome, even across populations with differing levels of exposures, by building an epigenetic biomarker of Arsenic (As) exposure in Bangladesh, a population chronically exposed to both high and widespread levels of As.<sup>17</sup> Inorganic arsenic (As) is an environmental chemical classified as a group 1 carcinogen by IARC, and long-term exposure

increases the risk of chronic diseases including As-associated skin lesions, cancers,<sup>18,19</sup> cardiovascular diseases,<sup>137</sup> and overall mortality.<sup>23</sup> We generated DNAm data using blood samples from 1,186 adults in the Health Effects of Arsenic Longitudinal Study (HEALS) to identify numerous CpG sites at which As exposure impacts DNAm. We observed strong enrichment of As-associated CpGs in genomic contexts and CpG sets linked to various disease states previously associated with As exposure. We then utilized Mendelian Randomization (an instrumental variable approach) to demonstrate that genetically-predicted As metabolism efficiency causally impacts DNAm at As-associated CpGs. Next, we used our EWAS results to create a DNAm-based biomarker that robustly predicts As exposure across multiple populations. Lastly, we demonstrate that this biomarker of exposure predicts arsenical skin lesion status in HEALS (a hallmark of arsenic toxicity), as well as overall mortality in the Strong Heart Study (SHS), a population with lower As exposure.

### **3.3 METHODS**

#### **3.3.1 Characteristics of the study population**

Health Effects of Arsenic Longitudinal Study (HEALS) is a prospective cohort study aimed at assessing how chronic As exposure from groundwater consumption impacts health outcomes among adults in Araihaazar, Bangladesh.<sup>138</sup> As described previously, participants recruited in 2000-2002 were married couples who resided in Araihaazar for at least 5 years.<sup>139</sup> For the present study, HEALS participants were randomly selected to have their genome-wide DNAm profiled among those randomly selected for genotyping. Informed consent was obtained from all participants.



### 3.3.2 Measuring total As and As speciation in urine

Baseline As levels (prior to As exposure mitigation efforts<sup>140</sup>) were measured from spot urine samples using graphite furnace atomic absorption spectrometry (AAAnalyst 600 spectrometer; Perkin Elmer) at Columbia University's Trace Metal Core Laboratory; the detection limit was 2 ug/L.<sup>141</sup> Additionally, urinary As species were measured using high-performance liquid chromatography and inductively coupling plasma mass spectrometry with dynamic reaction cell technology for detection.<sup>142</sup> The detection limit for arsenic species, including inorganic As III (iAs<sup>III</sup>) and V (iAs<sup>V</sup>), mono-methylated As (MMA), and di-methylated As (DMA), was 0.2 µg/L. To normalize urinary As levels to overall urine volume, total urinary As was divided by urine creatinine measured by a colorimetric Sigma Diagnostics Kit.

### 3.3.3 DNA methylation

The Qiagen FlexiGene DNA kit was used to extract DNA from clotted blood samples, and Zymo EZ-96 DNA Methylation Kits were used for bisulfite conversion. Samples were prepared and analyzed according to Infinium MethylationEPIC (EPIC) array. All samples were collected at baseline but processed in two batches (396 in 2017 and 828 in 2021). For each batch, we removed CpGs with a detection p-value >0.01 in  $\geq 1$  sample, CpGs missing methylation values in >5% of samples, cross-reactive CpGs, CpGs annotating to SNPs or within one base pair, CpGs on sex chromosomes, and probes that were missing in >10% samples;<sup>143</sup> samples with >10% missing probes were also removed. KNN imputation was performed (n=10) for missing beta values followed by BMIQ normalization.<sup>144,145</sup> After filtering, 721,796 CpGs across 1,186 samples were retained for subsequent analyses.

### 3.3.4 Epigenome-wide association analyses of urinary As

Prior to analyses, we log2-transformed creatinine-adjusted urinary As. DNAm beta values were transformed to M-values ( $\log_2[\text{beta}/(1-\text{beta})]$ ). For this study's primary EWAS, we conducted a cross-sectional study to evaluate the association between log2-transformed creatinine-adjusted urinary As measured at baseline and DNAm at each CpG site; we used the *limma*<sup>146</sup> (v3.60.4) package to fit linear regression models, controlling for covariates of methylation batch, age (continuous), sex, smoking status (never [reference], former, current), BMI category (normal [reference], underweight, overweight, obese, and unknown), and surrogate variables (SVs) to capture cell type composition. As a sensitivity analysis, we also performed a reference-based EWAS without SVs, but instead adjusting for cell type composition by utilizing EpiDISH<sup>147</sup> (v2.20.0) to compute immune cell proportions of B-cells, CD4+ T-cells, CD8+ T-cells, NK-cells, Monocytes, Neutrophils, and Eosinophils based on 333 tissue-specific DNase hypersensitive sites (DHS) from the 450K purified blood cell types dataset from the NIH Epigenomics Roadmap<sup>148</sup> that was further improved by Teschendorff et al. 2017.<sup>149</sup> We also conducted an EWAS of the creatinine-adjusted sum of iAs<sup>III</sup>, iAs<sup>V</sup>, MMA, and DMA species, using the subset of HEALS participants with available speciation data. We additionally performed an EWAS using log2-transformed As levels in water from wells that HEALS participants most commonly drank from.

### 3.3.5 Determination of surrogate variables to adjust for

Surrogate variables (SVs) were estimated using the *sva* package to capture the effects of cell type composition and urinary arsenic and covariates were protected for during the generation of these SVs;<sup>150,151</sup> to determine the number of surrogate variables to use, we computed the inflation

factor for the distribution of p-values for each number of SVs and selected the number of SVs that yielded the inflation factor closest to 1. We further checked that these SVs strongly associated with different inferred cell type proportions inferred using the EpiDISH package (**Table S 3.1**). This EWAS approach using surrogate variables has been shown to perform better in genome-wide methylation studies than reference-based cell type-adjusted approaches<sup>152,153</sup> by adjusting for unknown and other accounted for biological and technical artifacts.<sup>150,154</sup> For the primary EWAS of creatinine-adjusted urinary As, as well as sensitivity analyses including the EWAS of the creatinine-adjusted sum of inorganic and methylated As species and the EWAS of arsenic in well water, we adjusted for the first 9 SVs (**Table S 3.1**).

### **3.3.6 Enrichment of As-associated CpGs for genomic context and human diseases**

For As-associated CpGs (FDR<0.05), we assessed enrichment in CpG islands, shelves, and shores based on the UCSC genome browser,<sup>155</sup> as well as KEGG pathways, GO terms, and Hallmark Gene Sets using the “gsameth” and “gometh” functions in the *missMethyl* R package<sup>156</sup> (v1.38.0); these functions evaluate enrichment of each gene set using a hypergeometric test while accounting for the number of CpGs per gene when computing prior probabilities used in the GSEA. We additionally queried As-associated CpGs in the MRC-IEU EWAS catalog<sup>157</sup> to examine overlap with CpGs identified in EWASs of human diseases.

### **3.3.7 Identifying CpGs involved in arsenic and cell type interactions (AxCT)**

To assess interactions between Arsenic exposure and cell-type specific DNAm, we utilized the *CellDMC* method<sup>158</sup> (v2.20.0) to identify CpGs involved in AxCT. *CellDMC* implements a linear

model similar to the model we utilized in our primary EWAS but also includes interaction terms between the exposure variable (in our case, log2-transformed urinary arsenic levels) and each estimated cell type proportion. We adjusted for covariates of sex, age, BMI category, cigarette smoking status, and DNAm batch.

### **3.3.8 Assessing the relationship between arsenic and epigenetic aging**

We utilized the *dnaMethyAge* package<sup>159</sup> to estimate epigenetic age acceleration using 14 DNAm clocks<sup>160–173</sup> including a multi-tissue clock developed by Horvath.<sup>161</sup> We separately regressed each clock's estimates of age acceleration on log2-transformed creatinine-adjusted urinary arsenic, controlling for covariates of sex, cigarette smoking status, batch, and inferred immune cell type proportions.

### **3.3.9 Association of genetically-predicted As metabolism efficiency (reflecting internal dose) with DNA methylation**

In the liver, inorganic As is converted into MMA and then DMA forms, facilitating excretion in urine. The percentage of DMA in urine among all As, MMA, and DMA species (DMA%), is commonly used as a measure of As metabolism efficiency (AME).<sup>174</sup> To evaluate whether As exposure causally impacts DNAm at a high confidence set of 116 As-associated CpGs, we conducted a Mendelian Randomization analysis to assess the association between genetically-predicted DMA%, reflecting internal dose, and DNAm at each As-associated CpG; this high confidence set of CpGs were identified in both EWAS of urinary As (Bonferroni  $p < 0.05$ ) and water As (FDR  $< 0.05$ ). To do this, we performed a two-stage least squares regression for each CpG using two SNPs, rs61735836 and rs4919690 (both with strong and well-established

associations with DMA%)<sup>139,175</sup> as instruments to predict AME, with CpG DNAm as the outcome. We then determined if the direction of the effect of genetically-predicted AME on each CpG was consistent with the direction of association between urinary As and DNAm at each CpG. For example, if increased genetically-predicted AME, and hence lower internal dose of As, is associated with increased DNAm at a specific CpG, then increased urinary As would be expected to be associated with decreased DNAm at that CpG. We tested this directional consistency using binomial tests for consistency across all 116 CpGs. As a sensitivity analysis, we tested the association between measured DMA% and DNAm at these CpGs, controlling for sex, age, cigarette smoking, batch, and inferred cell type proportions.

### **3.3.10 Building a DNAm-based biomarker of As exposure**

We randomly subdivided our 1,186 participants into a training/tuning set (n=1,000) and a testing set (n=186). All 93 skin lesion cases in our subcohort were included in the testing set to evaluate skin lesion outcomes with our DNAm-based biomarker. We performed an EWAS of urinary As in the training/tuning set and selected CpGs at an FDR<0.05. Among these 1,000 participants, we then performed elastic net regression (a combination of lasso and ridge regressions) using these pre-selected CpGs using with 5-fold cross-validation using the *glmnet*<sup>176</sup> package to optimally tune the  $\alpha$  hyperparameter, representing the contribution of lasso and ridge regressions. We applied this model in the testing set to obtain epigenetically-predicted As scores and computed the  $r^2$  between these scores and measured urinary As. Additionally, we evaluated the ability of our DNAm-based biomarker of As to predict baseline skin lesion status in the testing set by computing the covariate-adjusted area under the receiver operating characteristic curve (AUC) using *ROCNReg*,<sup>177</sup> adjusting for age, sex, BMI category, and smoking status. We also

built a DNAm-based biomarker specifically for the Illumina 450K array utilizing the same steps listed above but restricting to CpGs on the 450K array.

### **3.3.11 Validating the DNAm-based biomarker of As exposure in two independent cohorts**

We validated our 450K array DNAm-based biomarker in the Bangladesh Vitamin E and Selenium Trial (BEST), a population-based chemoprevention study of Bangladeshi adults with visible effects of As toxicity;<sup>178</sup> we evaluated the association between measured urinary As levels and epigenetically predicted urinary As scores using the 450K array DNAm-based biomarker (as only 450K data was available in BEST). We validated our EPIC array biomarker in the Strong Heart Study (SHS), a study that has linked low-level As exposure to cardiovascular disease risk in American Indians.<sup>179–182</sup> In SHS, we assessed the correlation between measured and epigenetically-predicted urinary As levels among a subset of participants with DNAm data without cardiovascular disease at baseline (n=2,320). We also assessed the ability of our DNAm-based biomarker to predict overall mortality, and cardiovascular disease (CVD) incidence and mortality using Cox regressions; measured As has been established as a clear biomarker for these survival outcomes in the greater SHS cohort.<sup>181,183</sup> Follow-up was limited to 2009 since water As exposure began to change after the US EPA Final Arsenic Rule was implemented.<sup>184–186</sup> Cox model #1 controlled for potential confounders (age, study center, sex, smoking status, body mass index, estimated glomerular filtration rate). Model #2 additionally controls for potential mediators (Model #1 covariates, LDL cholesterol, HDL cholesterol, systolic blood pressure, diabetes status, hypertension history, albuminuria).

## **3.4 RESULTS**

### 3.4.1 Participant characteristics

The median age of this HEALS subcohort was 38.0 y, and there were slightly more women than men (**Table 3.1**). At baseline, 78% of our participants consumed water from As-contaminated tube wells with As concentrations above the WHO guideline of 10 µg/L, and 55% consumed water with concentrations above the national Bangladesh standard of 50 µg/L. The median urinary As was 204.5 µg/g creatinine.

### 3.4.2 EWAS of urinary As identified many novel and previously reported regions

This study's primary EWAS of log2-transformed creatinine-adjusted urinary As using 721,796 CpGs identified 1,177 CpGs at an FDR<0.05 and 135 CpGs at a Bonferroni threshold ( $p<6.9E-8$ ) (**Figure 3.1A and Table S 3.2**), controlling for covariates and nine surrogate variables (SVs). For most of the 1,177 CpGs (72.3%), higher urinary As levels were associated with decreased methylation (**Table S 3.2**). The majority of these CpGs (98.0%) have not been reported previously (**Table S 3.3**). We observed mild inflation ( $\lambda=1.14$ ) (**Figure S 3.1**) which was lower than the inflation observed for supplementary EWAS approaches, including adjusting for only cell types (estimated using EpiDISH) and covariates ( $\lambda=1.64$ ) and only SVs without covariates ( $\lambda=1.26$ ) (**Table S 3.4**).

We observed high consistency between total urinary As levels obtained from spot urine samples and the sum of inorganic and methylated species and moderate consistency with water As levels (**Figure S 3.2**). Our primary EWAS was highly consistent with the EWAS of the sum of inorganic and methylated As with each EWAS identifying 898 shared FDR-significant CpGs, having nearly identical effect sizes (**Figure S 3.3**). Effect sizes of FDR-significant CpGs were

also highly consistent with the EWAS of water As (**Figure S 3.3**). Additionally, 20 of 34 CpGs previously identified in a subset of 396 HEALS participants ( $\text{FDR} < 0.05$ )<sup>26</sup> were identified in our primary EWAS with highly consistent effect size estimates (**Figure S 3.4**).

### **3.4.3 As-associated CpGs show enrichment based on genomic context and in gene sets related to human health**

As-associated CpGs were enriched in open seas and depleted in CpG islands (**Figure 3.1B**). Utilizing leukocyte data from the NIH Roadmap Epigenomics Consortium<sup>187</sup> as a reference, we observed that As-associated CpGs were depleted in or near active transcription start sites and enriched in inactive chromatin states (weak repressed polycomb sites and quiescent states) and enhancer regions (**Figure 3.1C**). As-associated CpGs showed enrichment in DNaseI hypersensitivity sites ( $p = 2.74 \times 10^{-36}$ ) (**Table S 3.5**).

Most As-associated CpGs (74.8%) had been previously identified in EWASs of health conditions and diseases in the MRC-IEU EWAS catalog,<sup>157</sup> and As-associated CpGs were enriched in the catalog compared to all analyzed CpGs on the EPIC array (Fisher test  $p = 3.51 \times 10^{-27}$ ) (**Table S 3.5, Table S 3.6**). Notably, 440 CpGs were identified in an EWAS of C-reactive protein, 245 CpGs in an EWAS of chronic obstructive pulmonary disease (COPD), and 166 in an EWAS of Type 2 Diabetes (T2D) (**Table S 3.6**). Other phenotypes with overlapping CpGs included aging, eosinophilia, and cancers (kidney, pancreas, lung, colorectal, breast, prostate). Gene set enrichment analyses showed that As-associated CpGs were marginally enriched ( $p < 0.05$ ) for genes that respond to ultraviolet radiation (Hallmark gene set), KEGG pathways related to



kidney/adrenal gland function and T2D, and gene ontology terms related to cell communication, plasma membranes, and blood coagulation (**Table S 3.7**).

### 3.4.4 CpGs with the strongest associations with As were located in regulatory regions

Among As-associated CpGs, the strongest association was between increased As and increased methylation at cg19534475, located in *ATP1B3*, which encodes a subunit of Na<sup>+</sup>/K<sup>+</sup> ATPases that are critical for regulating osmosis and sodium-coupled transport (**Figure 3.2A**). This CpG is located at a DNaseI hypersensitivity site annotated in ENCODE<sup>188</sup> and at a regulatory region in the ReMap Atlas.<sup>189</sup> cg19534475 was previously identified in a subset of HEALS adults,<sup>26</sup> as well as in the SHS, consisting of American Indians exposed to lower levels of As ( $p < 0.05$ ).<sup>179</sup> Additionally, we identified cg14937504 as a novel As-associated CpG in *AC009518.4*, a long noncoding RNA that has been experimentally shown to overexpress during cardiac differentiation<sup>190</sup> (**Figure 3.2B**). Increased As was associated with increased DNAm at cg14937504, and cg14937504 was located next to an H3K27ac mark and on both a DNaseI hypersensitivity site and a ReMap Atlas regulatory region. cg14937504 was previously reported to associate with C-reactive protein levels and COPD.<sup>191,192</sup>

We defined As-associated regions as 5kb windows surrounding each As-associated CpG, with overlapping windows grouped into the same regions. We identified 75 regions containing more than one As-associated CpG, including several previously reported loci such as *ABR* and *SEMA4G* (**Table S 3.2**).<sup>26</sup> For several of these regions, we observed clusters of CpGs showing decreased DNAm in relation to increased As. Strikingly, a cluster of eight CpGs at the TSS of

*PNMA8B* (**Figure 3.2C**) and a cluster of three CpGs at the TSS of *CCDC105* (**Figure 3.2D**) showed decreased DNAm in association with increased As. Both CpG clusters were located at a CpG island/shore, a DNaseI hypersensitivity site, and a ReMap Atlas regulatory region. CpGs in these two clusters had been previously identified to associate with aging,<sup>193</sup> T2D,<sup>192</sup> inflammation (C-reactive protein),<sup>191</sup> and cancers of the kidney,<sup>194</sup> liver,<sup>195</sup> and prostate.<sup>192</sup> Compared to As-associated CpGs in regions with one hit, As-associated in regions with multiple hits were enriched in CpG islands and shores, perhaps due to the higher density of CpGs at islands (**Table S 3.8**).

### 3.4.5 Arsenic-by-cell type interactions (AxCT) analyses identify novel CpGs

To determine if the association of As with DNAm might vary by cell type, we estimated arsenic-cell type interactions (or AxCT) using the *CellDMC* software.<sup>158</sup> In total, we identified 46 CpGs showing evidence of AxCT across B cells, CD8+ T cells, monocytes, and NK cells (**Table S 3.9**). Although no CpGs involved in AxCT were also identified in the primary EWAS, two CpGs involved in AxCT in NK cells annotated to two genes (*PTGDR*, *UACA*) assigned to CpGs identified in our primary EWAS. Interestingly, five CpGs involved in AxCT in monocytes were located in *TYK2*, a gene previously implicated in systemic vasculitides.<sup>196</sup> These findings suggest that considering AxCT interactions can increase power to detect CpGs not detected using conventional EWAS approaches.

### 3.4.6 As exposure shows no clear association with epigenetic aging

We examined the association between urinary As and epigenetic aging, utilizing fourteen established DNAm clocks to estimate epigenetic age acceleration (EAA). Among the clocks we

tested, we observed no associations ( $p < 0.05$ ) between EAA and log2-transformed creatinine-adjusted urinary As levels (**Figure S 3.5**).

### **3.4.7 As metabolism efficiency (AME) causally impacts DNAm at As-associated CpGs**

We observed that direction of association of measured DMA% with DNAm at As-associated CpGs was nearly entirely directionally consistent with the associations between As and DNAm observed in the primary EWAS (**Figure 3.3A**). To assess the causality of the hypothesized effect of internal dose of As on DNAm at these CpGs, we conducted a Mendelian Randomization analysis by utilizing two SNPs that influence AME<sup>139,175</sup> as instruments to estimate each participant's genetically-predicted AME, and by proxy, genetically-predicted internal dose of As. We observed strong directional consistency between the association of genetically-predicted AME with DNAm at As-associated CpGs and the association of As exposure with these CpGs in the primary EWAS (**Figure 3.3B**).

### **3.4.8 DNAm-based biomarkers predict exposure and health risks across populations**

Our final multi-CpG models developed using elastic net regression with 5-fold cross-validation were a 255-CpG model built using EPIC array data and a 117-CpG model built using 450k array data (**Figure S 3.6, Table S 3.10**). In an independent validation set in HEALS, both the 255-CpG model ( $r^2 = 0.46$ ,  $p = 2.73 \times 10^{-26}$ ) and the 117-CpG model ( $r^2 = 0.34$ ,  $p = 4.83 \times 10^{-18}$ ) were strongly predictive of measured urinary As (**Figure 3.4A**). The 255-CpG biomarker of As exposure was clearly predictive of arsenical skin lesion status in the independent validation set in HEALS (AUC=0.69, 95% CI: 0.59-0.77) (**Figure 3.4D**) and comparable to the performance of measured urinary As (AUC=0.65, 95% CI: 0.55-0.74).

We externally validated our 117-CpG model in BEST, another Bangladeshi cohort with high As exposure (**Table S 3.11**), and observed that measured and DNAm-predicted urinary As scores showed clear correlation ( $r^2=0.19$ ,  $p=1.31 \times 10^{-19}$ ) (**Figure 3.4B**). We further validated our 255-CpG model in SHS, a population with much lower As exposure (**Table S 3.11**), and detected clear association between measured and predicted urinary As scores ( $r^2=0.01$ ,  $p=6.03 \times 10^{-6}$ ) (**Figure 3.4C**). Among survival outcomes in SHS, predicted As scores using the 255-CpG biomarker were associated with increased overall mortality, similar to the association of measured As levels (**Figure 3.4E**). The associations between cardiovascular disease (CVD) mortality and both measured and epigenetically predicted As scores were also directionally consistent; however, the direction of associations with CVD incidence differed.

### 3.5 DISCUSSION

In this study, we generated and analyzed genome-wide DNAm data from whole blood samples of 1,186 HEALS participants, a cohort with high As exposure, with the aim of developing an effective epigenetic marker of As exposure and toxicity. We identified 1,177 As-associated CpGs ( $FDR < 0.05$ ), with only 68 identified previously (accounting for multiple testing).<sup>26,179,197,198</sup> As-associated CpGs were depleted in CpG islands and transcription start sites and enriched in several inactive chromatin states, DNaseI hypersensitivity sites, and enhancers, largely consistent with previous studies of the effects of environmental exposures on the epigenome.<sup>26,199</sup> As-associated CpGs were enriched in CpG sets related to kidney function, T2D, inflammation, and cancers (kidney, pancreas, lung, colorectal, breast, prostate), all of which have been linked to As exposure in epidemiological studies.<sup>200–203</sup> While the effect of As exposure on the human

epigenome cannot be studied in a randomized trial, our analysis utilizing SNPs to estimate genetically-predicted AME, reflecting internal dose of As, mimics a randomized study (i.e., “Mendelian randomization”);<sup>204</sup> though As exposure is known to have adverse effects on multiple organs and tissue types,<sup>205</sup> the consistency between the associations observed in our primary EWAS and the effects of genetically-predicted internal dose on As-associated CpGs provides strong evidence that DNAm in blood is causally impacted by As exposure. Lastly, we developed a novel DNAm-based biomarker of As exposure that achieved the highest predictive performance to-date amongst all epigenetic biomarkers developed to track exposure to environmental chemicals, validated across different populations, and associated with health outcomes including arsenical skin lesions and overall mortality.

Our findings highlight several strengths of this study which can serve as guidelines for the future development of robust epigenetic biomarkers including 1) training on a large sample size, 2) including participants with a broad range of exposure levels, 3) ensuring high-quality exposure assessments, including both environmental and biospecimen-derived exposure measures if possible, 4) utilizing individuals with chronic exposure, and 5) replicating biomarkers across multiple external populations. Our ability to identify >1,000 As-associated CpGs as features for our epigenetic biomarker stems from our large sample size (n=1,186) and the wide range of As exposure levels in HEALS. Moreover, our EWAS analyses using total urine As, the sum of inorganic and methylated As urine species, and water As measures produced highly consistent results, suggesting minimal bias due to organic forms of As. This consistency stresses the importance of having high quality exposure assessment for training epigenetic markers. Furthermore, given that all samples were obtained at baseline before any interventions to reduce

As exposure, our As measures likely reflect long-term exposure, thereby contributing to the lasting effects of As on DNAm,<sup>206</sup> and consequently, increased biomarker performance. Additionally, the portability of our blood-based epigenetic biomarker across populations shows that As-associated epigenetic changes in high-exposure populations can be detected in populations with lower exposure and underscores how including individuals with a wide range of exposures when training epigenetic biomarkers can improve their performance across populations. Taken together, these findings demonstrate the relevance of our blood-based biomarker in reliably tracking As exposure.

In addition to tracking exposure, this study also emphasizes the need for a clinically measurable endpoint of toxicity for benchmarking the utility of an epigenetic biomarker. The strong discriminating performance of our epigenetic biomarker for predicting As skin lesion status in HEALS – slightly surpassing even the accuracy of measured urinary As levels – demonstrates its reliability in assessing toxicity risk and its potential to capture historical exposure. This finding, compounded with our observation that the epigenetic biomarker was associated with mortality in SHS, suggests this biomarker could help stratify and identify individuals at highest risk for As-related diseases. Consisting of 117 and 255 CpGs, our epigenetic biomarker of exposure could potentially complement traditional spot urine measurements of As as cost-effective screening approaches and enable large-scale, population-level screening.<sup>207</sup> Future studies to develop DNAm-based biomarkers of As exposure with increased sample sizes and diverse populations, will help create biomarkers that are increasingly precise and portable across populations.

While we identified many As-associated CpGs, the EPIC array only measures methylation at ~3% of CpG sites in the genome and is biased towards regions with gene annotations, potentially limiting resolution of specific CpGs in epigenomic regions most heavily impacted by As exposure.<sup>208</sup> Future studies using more comprehensive DNAm profiling approaches such as whole genome bisulfite sequencing are needed to identify As-associated regions not captured by EPIC array and resolve other As-associated CpGs around these signals, which may further improve performance of biomarkers developed with these approaches.

In conclusion, we demonstrate that As exposure and associated toxicities can be robustly captured in an epigenetic biomarker. This study not only provides guidelines for the development of robust epigenetic biomarkers of chemical exposures, but also accentuates our understanding of effects of As on the epigenome, as we identified >1,000 novel As-associated DNA methylation changes in whole blood from Bangladeshi adults chronically exposed to As-contaminated drinking water. Mendelian randomization and validation of epigenetic biomarkers across low- and high-exposure populations highlights the causal impact of As on the epigenome and the utility of these biomarkers for assessing As-related health risks. Further studies are needed to assess the impact of As on DNAm at CpGs not covered by the EPIC array, as well as the utility of our DNAm-based biomarkers in the prevention of As toxicity among high-risk individuals over time.

3.6 APPENDIX

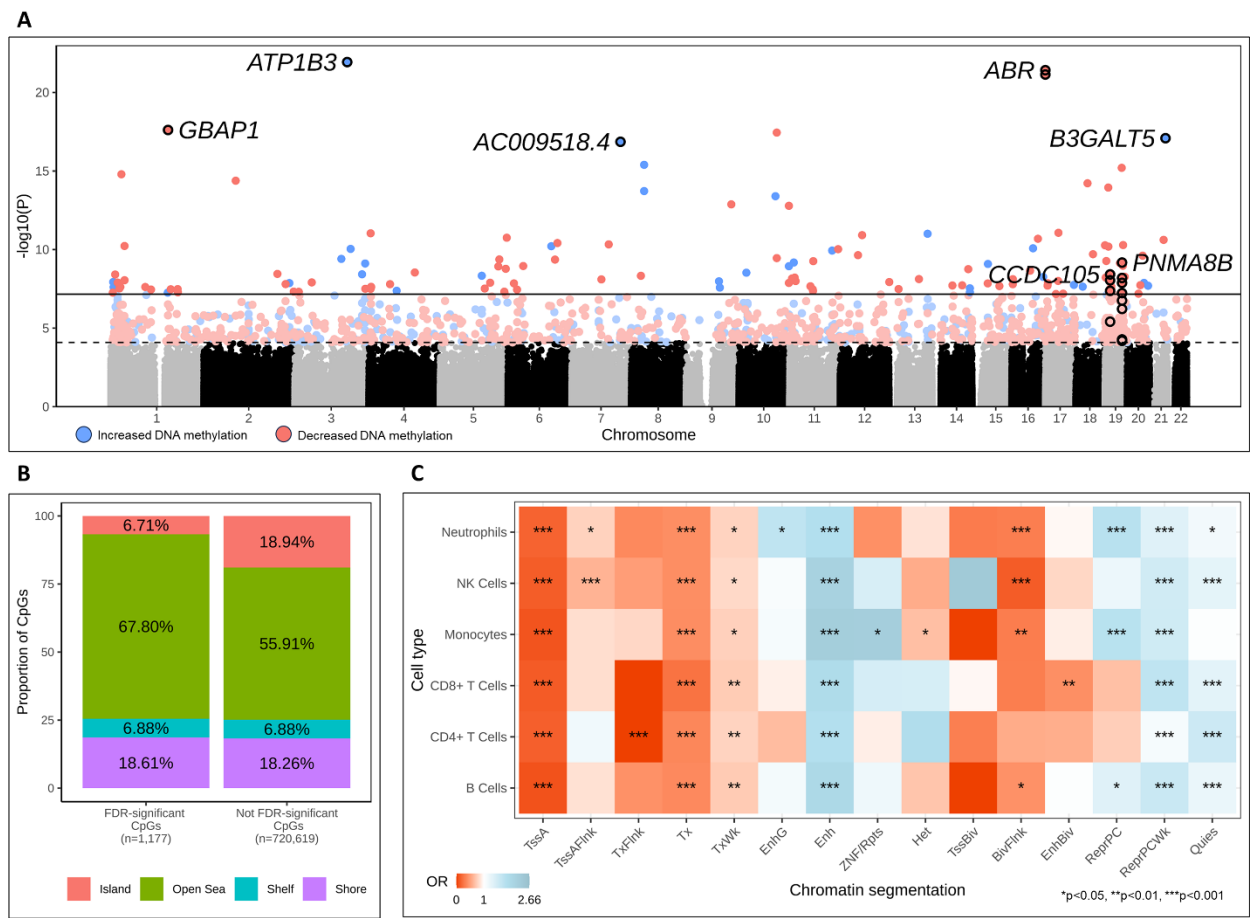


Figure 3.1 Epigenome-wide study of the association between urinary arsenic and whole blood DNAm in HEALS. **(A)**. Manhattan plot showing P-values for the association of urinary arsenic with DNAm in HEALS. Blue dots represent CpG sites where As exposure is associated with increased DNAm. Red dots represent CpGs sites where As exposure is associated with decreased DNAm. Dashed line represents FDR threshold ( $p=8.12E-05$ ). Solid line represents Bonferroni threshold ( $p=6.93E-08$ ). **(B)**. As-associated CpGs were enriched for in open seas and depleted in CpG islands. **(C)**. As-associated CpGs were enriched for in enhancer regions, based on chromatin segmentation categories, across several leukocyte cell types and depleted in regions involved in active transcription.



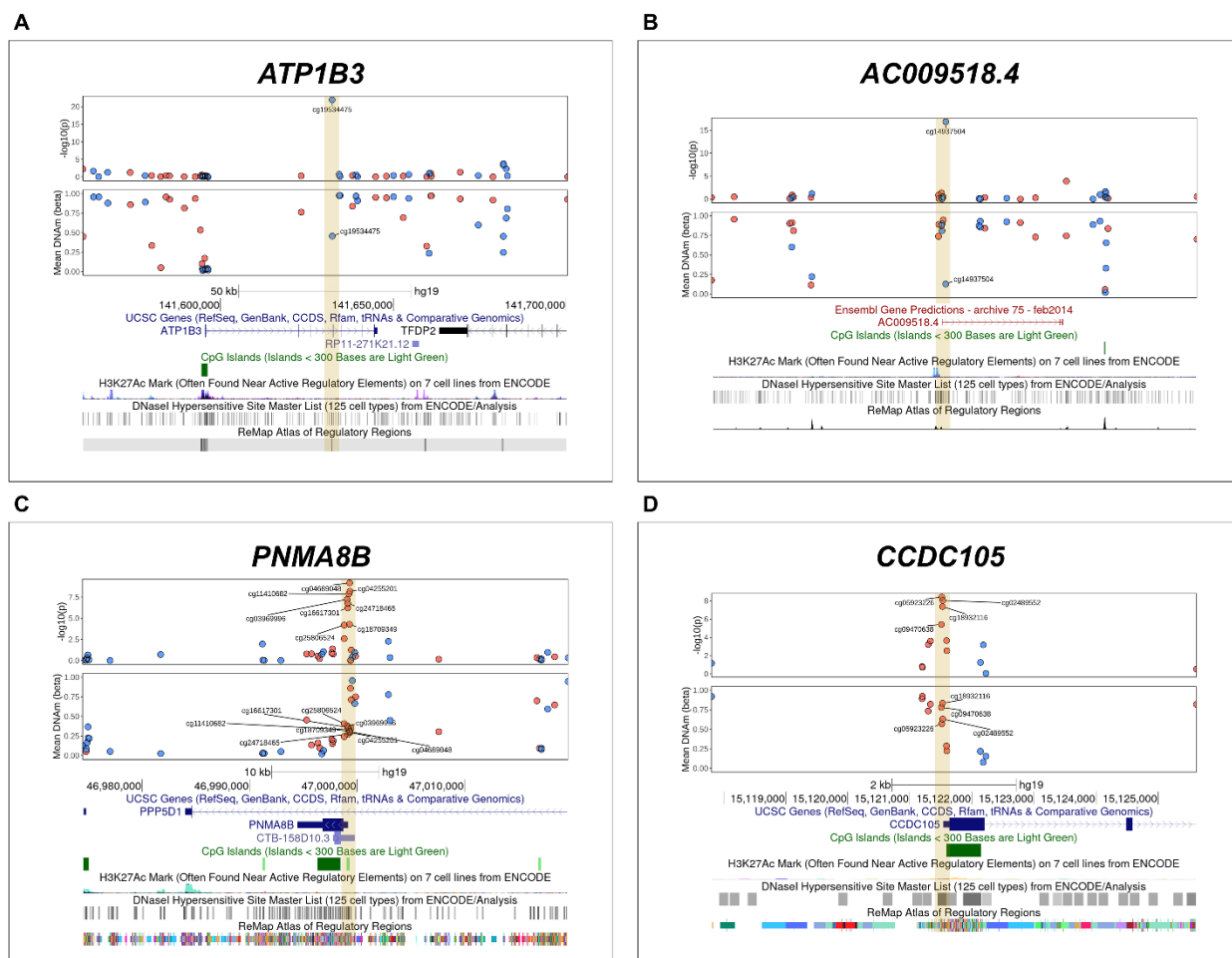


Figure 3.2 Regional plots of several regions containing CpGs showing association with urinary arsenic levels. In each panel, P-values (above) and mean DNAm levels (bottom) are shown above the gene map. FDR-significant CpGs are labeled. Genomic annotations were obtained from the UCSC genome browser. (A). *ATP1B3* region. (B). *AC009518.4* region. (C). *PNMA8B* region. (D). *CCDC105* region.

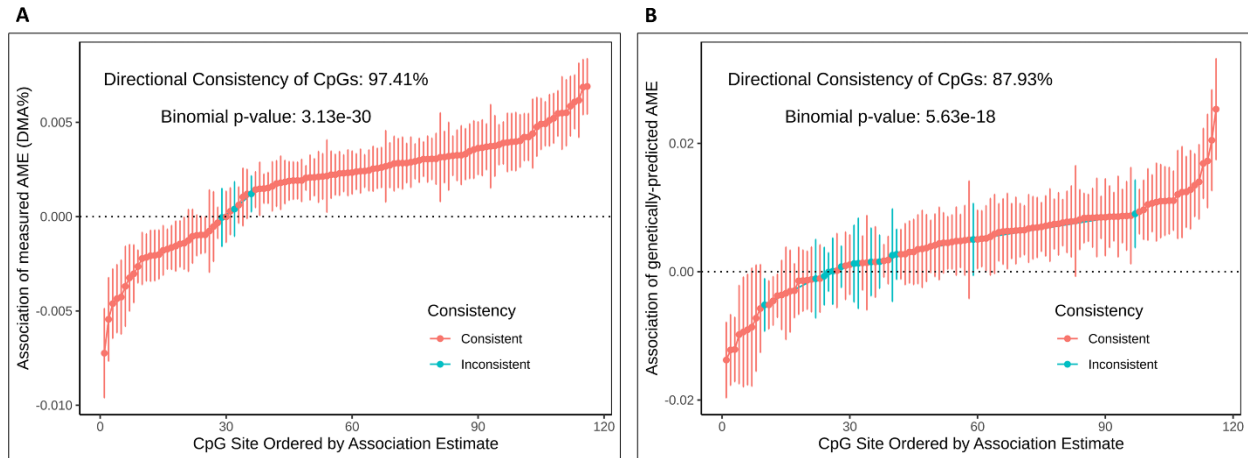


Figure 3.3 Association of arsenic metabolism efficiency (reflecting internal dose) with DNAm at arsenic-associated CpGs. The 116 CpGs plotted represent a high confidence set of As-associated CpGs that had a Bonferroni-adjusted  $p < 0.05$  in their association with urinary As and were also identified to associate with arsenic levels in well water ( $FDR < 0.05$ ). **(A)**. Beta values and confidence intervals for the association of measured arsenic metabolism efficiency (DMA%) with DNAm at As-associated CpGs. >97% of the associations observed are consistent with estimates from EWAS. **(B)**. Beta values and confidence intervals for the association between genetically-determined arsenic metabolism efficiency with DNAm at As-associated CpGs. >87% of CpGs show association that is directionally consistent with estimates from EWAS.

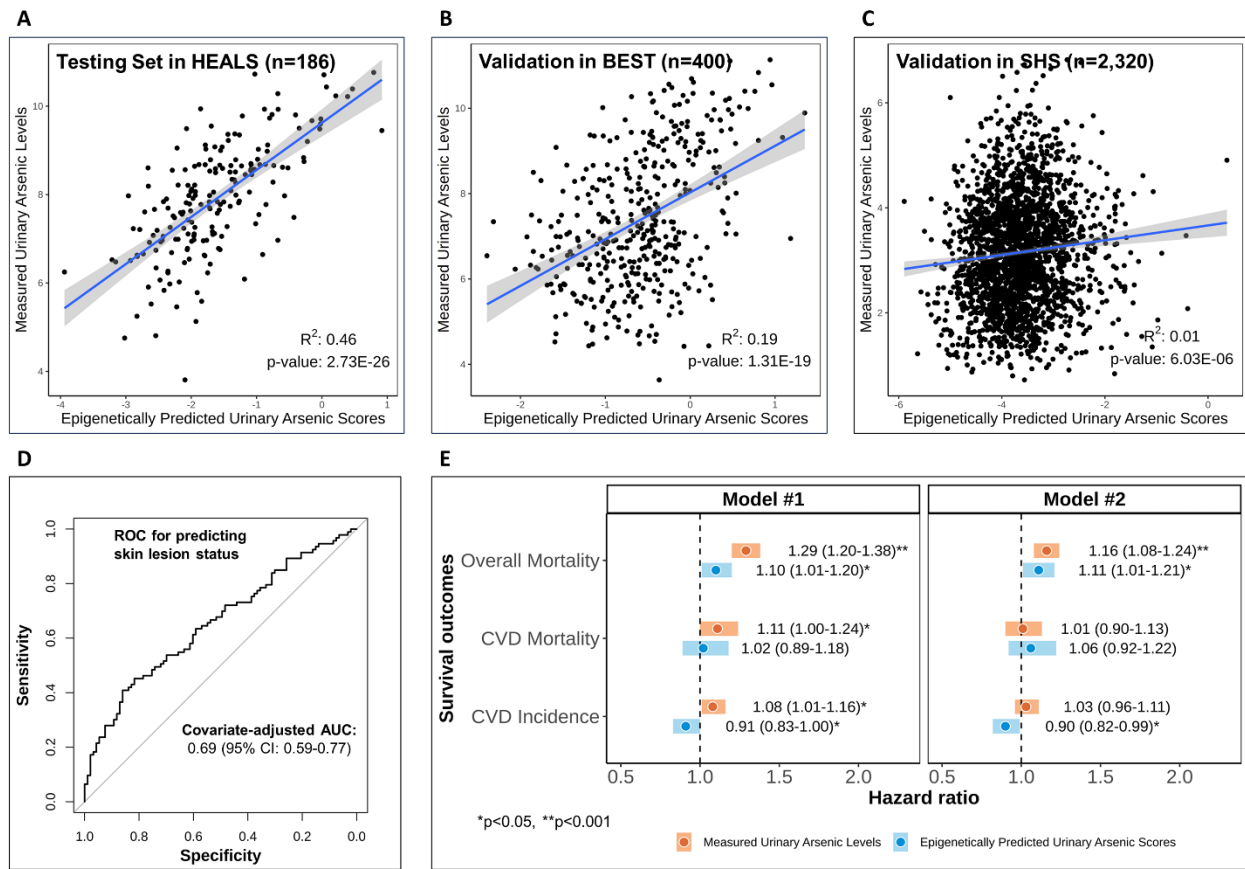


Figure 3.4 Validating DNAm-based biomarkers of As exposure. **(A)**. Association between measured (log<sub>2</sub>-transformed creatinine-adjusted) and predicted urinary As scores in an independent subset in the Health Effects of Arsenic Longitudinal Study (HEALS). **(B)**. Association between measured (log<sub>2</sub>-transformed creatinine-adjusted) and predicted urinary As scores among participants in the Bangladesh Vitamin E and Selenium Trial (BEST). **(C)**. Association between measured (log<sub>2</sub>-transformed creatinine-adjusted) and predicted urinary As scores among participants in the Strong Heart Study (SHS). **(D)**. Associations between urinary As scores predicted by the 255-CpG DNAm-based biomarker and arsenical skin lesion status in HEALS. **(E)**. Associations between urinary As scores predicted by the 255-CpG DNAm-based biomarker and survival outcomes in the SHS. Model #1 adjusts for covariates including age, study center, sex, smoking status, body mass index, estimated glomerular filtration rate. Model #2 controls for potential mediators in addition to Model #1 covariates including LDL cholesterol, HDL cholesterol, systolic blood pressure, diabetes status, hypertension history, albuminuria.

Table 3.1 Participant characteristics at baseline for a subcohort of HEALS participants [median (25th, 75th percentiles) or n (%)] compared to all HEALS participants.

|                                    | <b>HEALS subcohort with<br/>methylation data (n=1,186)</b> | <b>HEALS cohort (n=11,224)</b> |
|------------------------------------|------------------------------------------------------------|--------------------------------|
| <b>Age (years)</b>                 | 38.0 [17.0, 70.0]                                          | 36.0 [29.0,45.0]               |
| <b>Urinary arsenic (µg/g)</b>      | 205 [11.0, 8560]                                           | 199.0 [106.0, 352.0]           |
| <b>Water arsenic (µg/L)</b>        | 61.0 [0.100, 864]                                          | 60.0 (12.0,147.0)              |
| <b>Sex</b>                         |                                                            |                                |
| Male                               | 580 (48.9%)                                                | 4,855 (43.3%)                  |
| Female                             | 606 (51.1%)                                                | 6,369 (56.7%)                  |
| <b>Smoking</b>                     |                                                            |                                |
| Never                              | 671 (56.6%)                                                | 7,204 (64.2%)                  |
| Former                             | 88 (7.4%)                                                  | 743 (6.6%)                     |
| Current                            | 427 (36.0%)                                                | 3,271 (29.1%)                  |
| Unknown                            | 0 (0%)                                                     | 6 (0.1%)                       |
| <b>BMI (kg/m2)</b>                 |                                                            |                                |
| Normal (18.5–22.9)                 | 576 (48.6%)                                                | 5,067 (45.1%)                  |
| Underweight (<18.5)                | 486 (41.0%)                                                | 4,421 (39.4%)                  |
| Overweight (23.0–29.9)             | 105 (8.9%)                                                 | 1,575 (14.0%)                  |
| Obese (≥30)                        | 11 (0.9%)                                                  | 80 (0.7%)                      |
| Unknown                            | 8 (0.7%)                                                   | 81 (0.7%)                      |
| <b>Baseline Skin Lesion Status</b> |                                                            |                                |
| Yes                                | 93 (7.8%)                                                  | 714 (6.4%)                     |
| No                                 | 1081 (91.1%)                                               | 10,204 (90.9%)                 |
| Unknown                            | 12 (1.0%)                                                  | 306 (2.7%)                     |

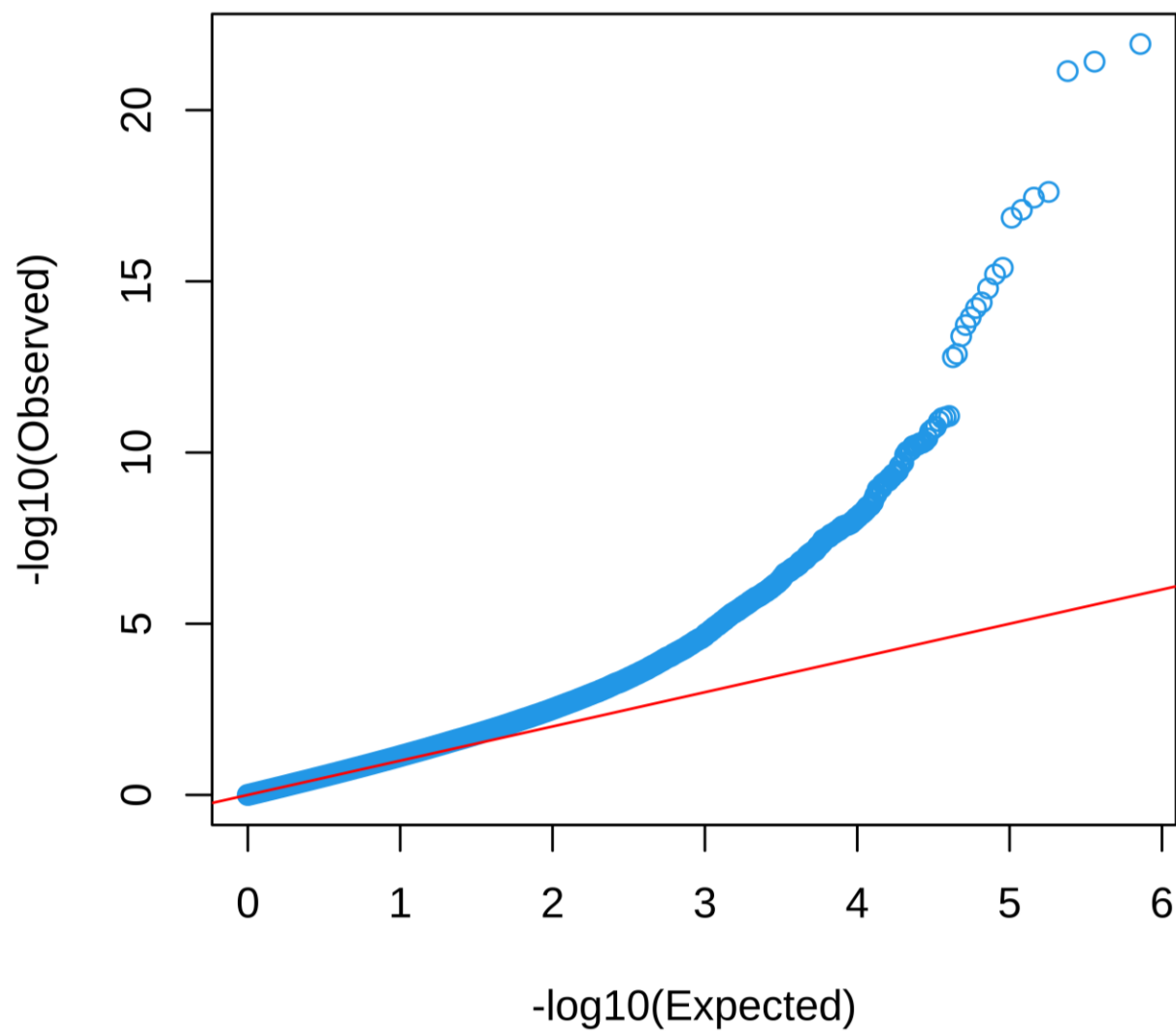


Figure S 3.1 QQ-plot for the epigenome-wide association study of log2-transformed creatinine-adjusted urinary arsenic.

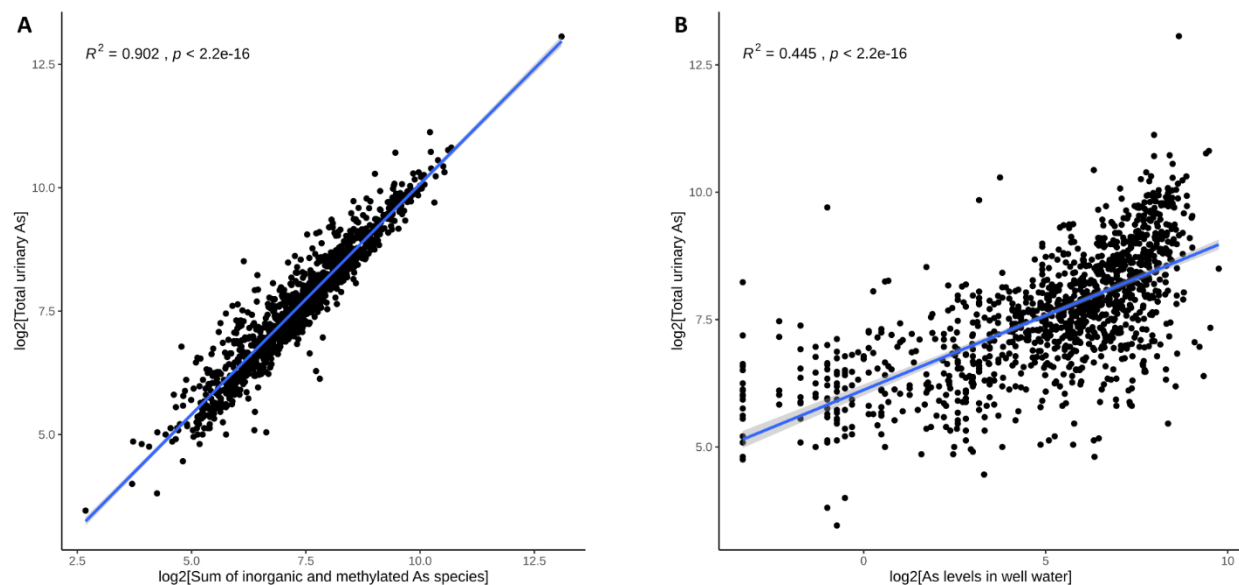


Figure S 3.2 Consistency in arsenic measurements between total urinary arsenic obtained from spot urine samples (utilized in the primary EWAS) and A) the sum of inorganic and methylated arsenic species and B) arsenic levels in well water.

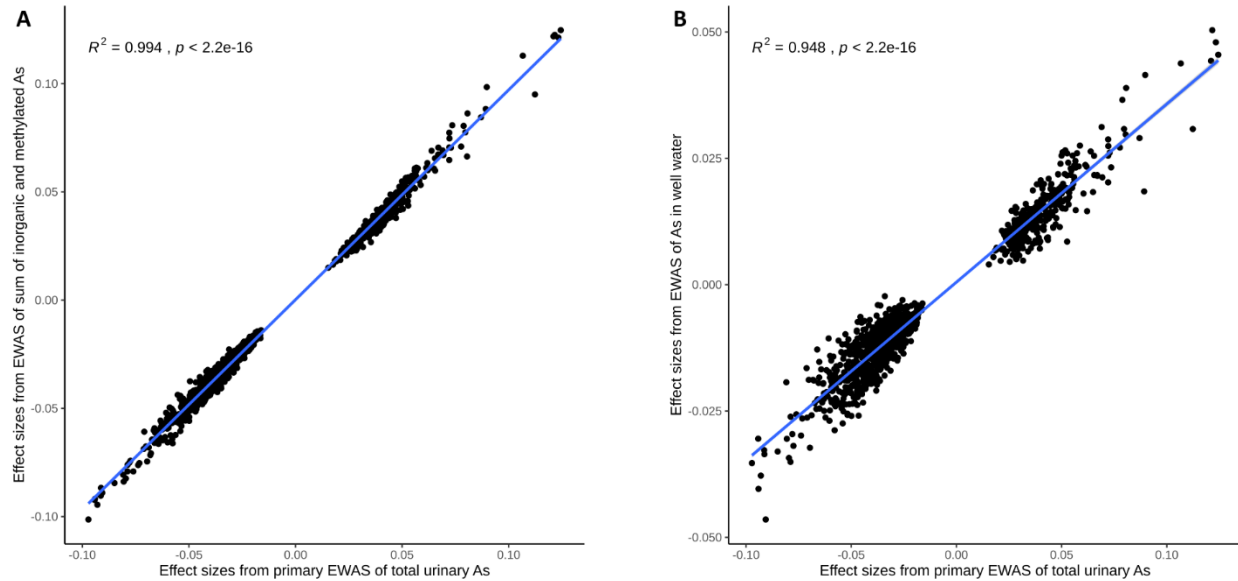


Figure S 3.3 Consistency in effect sizes of the 1,177 FDR-significant identified in the primary EWAS between primary EWAS and A) EWAS of the sum of inorganic and methylated As species and B) EWAS of water As.

Note: For A) 898 of the 1,177 FDR-significant As-associated CpGs in our primary EWAS were among the 1,086 CpGs also associated with the creatinine sum of inorganic and methylated As species (FDR<0.05) and for B) 370 of the 1,177 FDR-significant As-associated CpGs in our primary EWAS were among the 492 CpGs also associated with log2-transformed As in water (FDR<0.05).

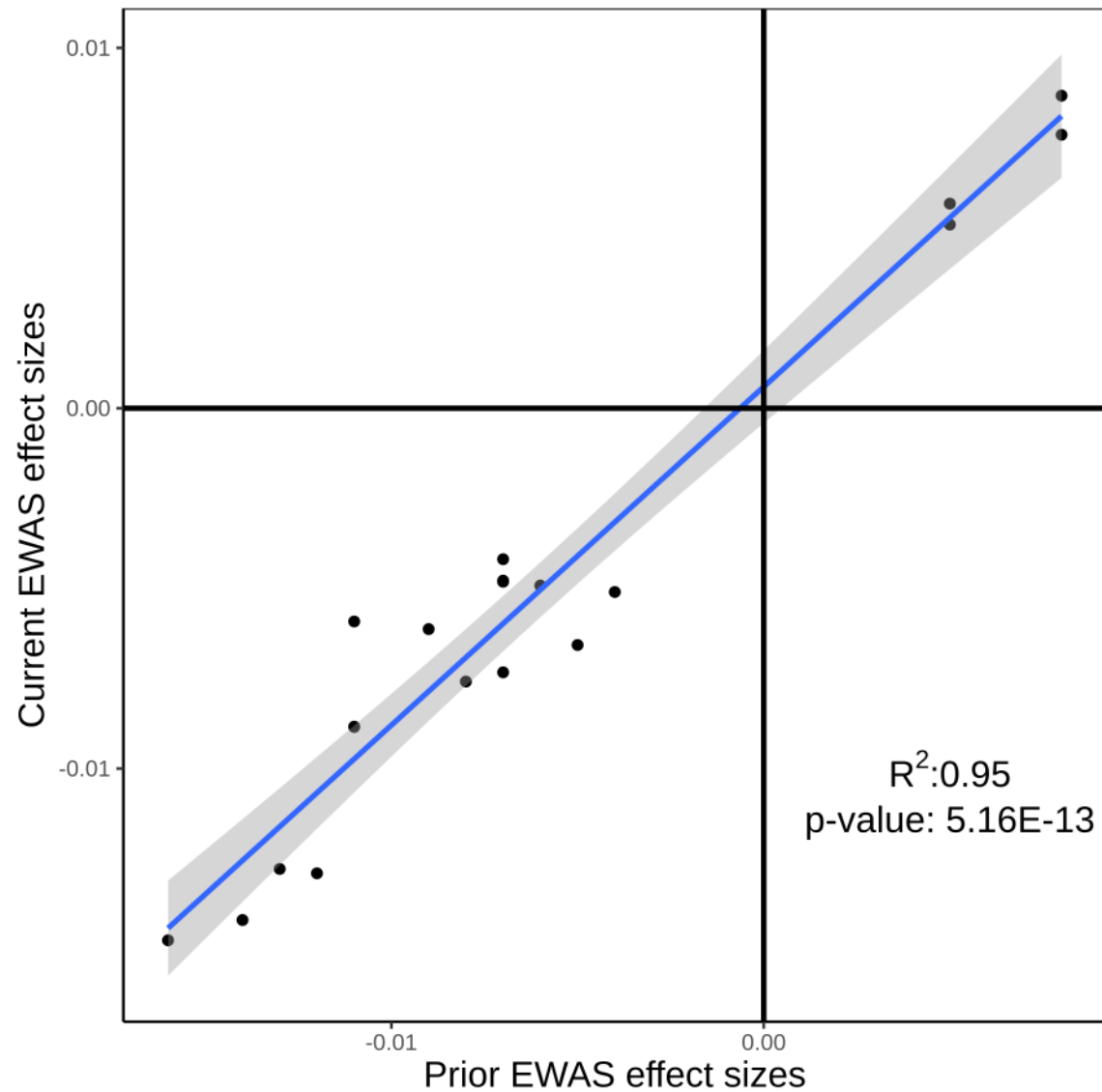


Figure S 3.4 Consistency of EWAS associations between urinary arsenic and DNA methylation (beta values) for the 20 CpGs identified in both the previous EWAS of a HEALS subcohort (n=396) and the current EWAS of a larger HEALS subcohort including the previous subcohort (n=1,186).



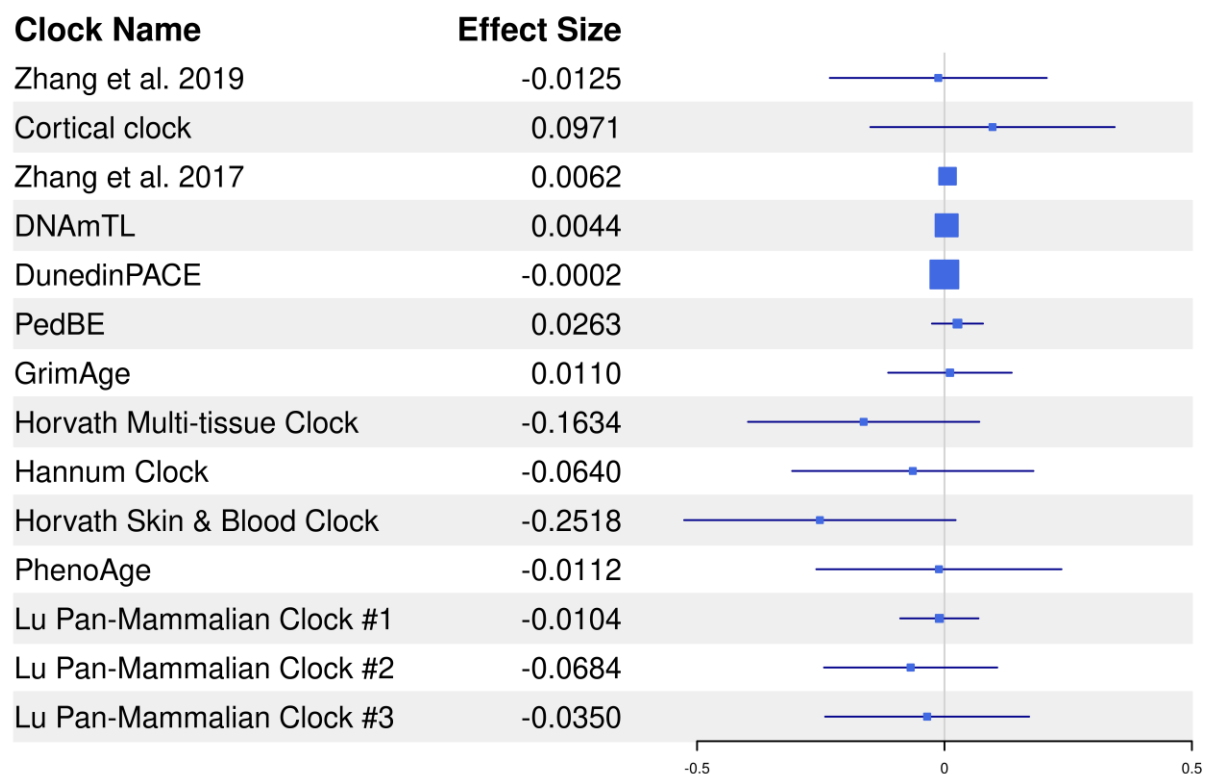


Figure S 3.5 Forest plot of the associations between epigenetic age acceleration predicted by various DNAm clocks and log2-transformed creatinine-adjusted urinary arsenic levels in HEALS.

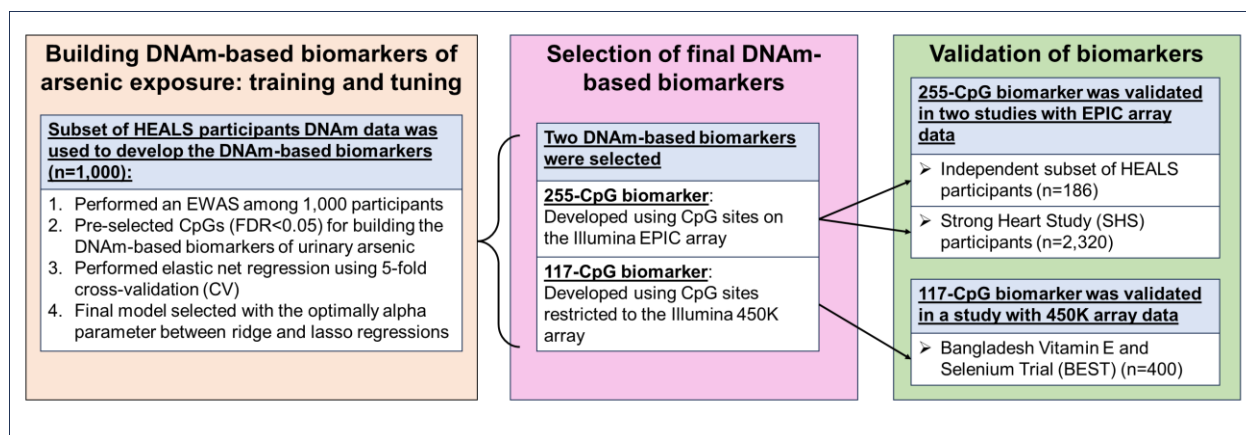


Figure S 3.6 Workflow for developing and validating DNAm-based biomarkers of arsenic exposure.

Table S 3.1 P-value matrices for the correlations between the first 20 surrogate variables (SVs) and inferred cell types (generated using EpiDISH) for the EWASs of urinary arsenic and arsenic in well water

| <i>EWAS of urinary arsenic</i>                                              |             |          |           |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
|-----------------------------------------------------------------------------|-------------|----------|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|                                                                             | SV1         | SV2      | SV3       | SV4      | SV5      | SV6      | SV7      | SV8      | SV9      | SV10     | SV11     | SV12     | SV13     | SV14     | SV15     | SV16     | SV17     | SV18     | SV19     | SV20     |
| <b>B</b>                                                                    | 6.35E-08    | 5.04E-03 | 1.96E-35  | 1.38E-05 | 1.95E-12 | 1.52E-28 | 2.54E-01 | 8.51E-03 | 3.47E-01 | 6.53E-11 | 3.28E-01 | 5.49E-13 | 1.13E-11 | 7.33E-15 | 1.22E-10 | 2.99E-01 | 8.89E-19 | 8.40E-05 | 8.64E-02 | 6.22E-04 |
| <b>NK</b>                                                                   | 9.49E-06    | 1.50E-13 | 2.63E-143 | 1.48E-08 | 6.02E-08 | 2.60E-02 | 5.95E-02 | 1.24E-09 | 6.04E-02 | 6.10E-01 | 3.87E-01 | 2.83E-08 | 9.60E-01 | 8.77E-04 | 1.25E-01 | 1.81E-01 | 4.78E-01 | 2.39E-02 | 1.23E-01 | 1.65E-01 |
| <b>CD4T</b>                                                                 | 3.11E-08    | 1.56E-02 | 2.76E-17  | 8.05E-06 | 1.45E-03 | 3.06E-95 | 1.30E-06 | 1.99E-14 | 5.40E-01 | 5.48E-10 | 3.46E-01 | 4.48E-04 | 3.18E-02 | 6.56E-02 | 2.81E-03 | 6.22E-03 | 5.43E-06 | 3.73E-01 | 1.07E-01 | 8.24E-01 |
| <b>CD8T</b>                                                                 | 1.71E-18    | 5.59E-31 | 3.06E-107 | 2.38E-04 | 2.73E-02 | 8.45E-06 | 2.51E-02 | 7.23E-02 | 5.24E-01 | 4.49E-01 | 6.53E-01 | 9.16E-01 | 1.33E-03 | 5.08E-01 | 1.27E-03 | 9.14E-02 | 5.38E-09 | 7.09E-01 | 1.60E-01 | 2.83E-01 |
| <b>Mono</b>                                                                 | 7.90E-01    | 5.43E-06 | 4.37E-20  | 2.30E-03 | 2.94E-02 | 3.63E-03 | 5.67E-04 | 3.77E-03 | 2.87E-01 | 1.43E-01 | 1.20E-04 | 7.56E-10 | 2.50E-03 | 6.01E-04 | 2.95E-01 | 9.39E-02 | 5.12E-04 | 1.51E-01 | 6.66E-01 | 1.16E-01 |
| <b>Neutro</b>                                                               | 5.36E-02    | 1.69E-14 | 6.38E-142 | 1.14E-10 | 7.81E-01 | 3.39E-60 | 4.42E-26 | 3.36E-16 | 8.23E-05 | 1.12E-07 | 3.79E-01 | 5.99E-02 | 1.11E-02 | 6.20E-02 | 3.13E-01 | 5.96E-01 | 1.55E-01 | 1.91E-01 | 7.73E-01 | 6.58E-01 |
| <i>EWAS of the sum of inorganic and methylated arsenic species in urine</i> |             |          |           |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| <b>B</b>                                                                    | 5.56E-08    | 0.004376 | 3E-35     | 1.71E-05 | 2.12E-12 | 1.91E-28 | 0.268109 | 0.007789 | 0.365183 | 8.35E-11 | 0.305742 | 2.99E-13 | 1.3E-11  | 3.12E-15 | 2.62E-10 | 0.297726 | 1.51E-18 | 0.000104 | 0.082735 | 0.000685 |
| <b>NK</b>                                                                   | 0.000011    | 8.78E-14 | 2.43E-143 | 1.34E-08 | 7.49E-08 | 0.027879 | 0.079728 | 9.02E-10 | 0.070101 | 0.549947 | 0.361783 | 3.66E-08 | 0.991066 | 0.000875 | 0.140095 | 0.181987 | 0.490021 | 0.025399 | 0.12334  | 0.182598 |
| <b>CD4T</b>                                                                 | 1.93E-08    | 0.014194 | 3.19E-17  | 1.18E-05 | 0.001528 | 1.4E-94  | 4.47E-06 | 4.92E-15 | 0.557402 | 5.73E-10 | 0.340383 | 0.00039  | 0.03516  | 0.05985  | 0.003284 | 0.006175 | 4.85E-06 | 0.382736 | 0.107326 | 0.813913 |
| <b>CD8T</b>                                                                 | 3.24E-18    | 7.31E-31 | 7.06E-107 | 0.000256 | 0.026988 | 9.21E-06 | 0.029157 | 0.064472 | 0.540872 | 0.433927 | 0.628102 | 0.958139 | 0.001326 | 0.461043 | 0.001365 | 0.091843 | 5.08E-09 | 0.723832 | 0.15579  | 0.277668 |
| <b>Mono</b>                                                                 | 0.822963072 | 5.2E-06  | 4.91E-20  | 0.002068 | 0.03312  | 0.004084 | 0.000905 | 0.003267 | 0.302538 | 0.127744 | 9.29E-05 | 6.76E-10 | 0.002944 | 0.000554 | 0.314948 | 0.094629 | 0.000613 | 0.152775 | 0.646724 | 0.119411 |
| <b>Neutro</b>                                                               | 0.049322585 | 7E-15    | 9.1E-142  | 1.99E-10 | 0.761138 | 1.28E-60 | 1.87E-26 | 2.41E-15 | 0.000247 | 1.21E-07 | 0.47063  | 0.051642 | 0.013701 | 0.060434 | 0.324406 | 0.602491 | 0.157678 | 0.194716 | 0.784288 | 0.666546 |
| <i>EWAS of arsenic in well water</i>                                        |             |          |           |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| <b>B</b>                                                                    | 5.50E-08    | 0.007055 | 7.67E-37  | 3.36E-05 | 1.22E-12 | 2.45E-27 | 0.094594 | 0.001493 | 0.214391 | 4.21E-11 | 0.164512 | 2.29E-14 | 1.26E-12 | 9.07E-20 | 5.99E-09 | 0.132343 | 3.69E-16 | 5.37E-03 | 0.052027 | 0.001858 |
| <b>NK</b>                                                                   | 8.89E-06    | 2.77E-14 | 1.62E-143 | 3.19E-07 | 2.75E-08 | 0.0414   | 0.961788 | 3.46E-11 | 0.147083 | 0.80324  | 0.312096 | 1.62E-08 | 0.990932 | 0.000327 | 0.266119 | 0.16933  | 0.548218 | 0.034033 | 0.104548 | 0.242134 |
| <b>CD4T</b>                                                                 | 2.77E-08    | 0.019246 | 2.79E-18  | 1.54E-05 | 0.000917 | 1.35E-92 | 2.98E-01 | 1.95E-21 | 0.874386 | 6.55E-10 | 0.228663 | 0.00309  | 0.035083 | 0.015219 | 0.004938 | 0.002711 | 2.71E-05 | 0.694982 | 0.085745 | 0.889493 |
| <b>CD8T</b>                                                                 | 2.02E-18    | 4.62E-32 | 5.55E-107 | 0.001166 | 0.025367 | 5.84E-06 | 0.399742 | 0.016547 | 0.534121 | 0.395514 | 0.601611 | 0.901657 | 0.000539 | 0.137671 | 0.000674 | 0.04537  | 3.16E-08 | 0.83616  | 0.116026 | 0.322713 |
| <b>Mono</b>                                                                 | 0.793426561 | 8.08E-06 | 4.92E-20  | 0.001183 | 0.026709 | 0.001615 | 0.033433 | 0.00028  | 0.354603 | 0.120499 | 0.000104 | 1.83E-10 | 0.001276 | 0.000214 | 0.37349  | 0.109672 | 0.000892 | 0.335305 | 0.630535 | 0.144843 |
| <b>Neutro</b>                                                               | 0.049442116 | 2.22E-15 | 3.46E-144 | 3.40E-09 | 0.722133 | 1.71E-85 | 1.12E-16 | 6.07E-08 | 2.94E-03 | 2.43E-06 | 0.922062 | 0.053009 | 0.017538 | 0.052667 | 0.449765 | 0.648119 | 0.197466 | 0.322025 | 0.883908 | 0.632368 |

Note: Red highlighted cells indicate p-values that are statistically significant (Bonferroni-adjusted  $p < 0.05$ ) after multiple testing correction across all SVs and cell types that were tested

Table S 3.2 CpGs associated with log2-transformed creatinine-adjusted urinary arsenic (FDR < 0.05)

| Name       | CpG Location   | Nearest Gene*       | Gene Feature* | log2[Fold-Change] | 95% CI Lower Bound | 95% CI Upper Bound | p-value  | FDR-adjusted p-value | Bonferroni-adjusted p-value | As-associated Region ID** | Number of CpGs in As-associated Region** |
|------------|----------------|---------------------|---------------|-------------------|--------------------|--------------------|----------|----------------------|-----------------------------|---------------------------|------------------------------------------|
| cg19534475 | Non-CpG Island | <i>ATP1B3</i>       | Body          | 0.051             | 4.07E-02           | 6.06E-02           | 1.17E-22 | 8.45E-17             | 8.45E-17                    | 1                         | 1                                        |
| cg10003262 | Shore          | <i>ABR</i>          | 5' UTR        | -0.094            | -1.13E-01          | -7.55E-02          | 3.84E-22 | 1.39E-16             | 2.77E-16                    | 2                         | 2                                        |
| cg01912040 | Shore          | <i>ABR</i>          | 5' UTR        | -0.091            | -1.10E-01          | -7.30E-02          | 7.25E-22 | 1.74E-16             | 5.23E-16                    | 2                         | 2                                        |
| cg06466147 | Non-CpG Island | <i>GBAP1</i>        | Body          | -0.091            | -1.12E-01          | -7.12E-02          | 2.45E-18 | 4.42E-13             | 1.77E-12                    | 3                         | 1                                        |
| cg05962511 | Shore          | <i>SEMA4G</i>       | 5' UTR        | -0.060            | -7.28E-02          | -4.63E-02          | 3.60E-18 | 5.19E-13             | 2.60E-12                    | 4                         | 2                                        |
| cg26390598 | Non-CpG Island | <i>B3GALT5</i>      | 5' UTR        | 0.122             | 9.43E-02           | 1.49E-01           | 8.18E-18 | 9.84E-13             | 5.91E-12                    | 5                         | 1                                        |
| cg14937504 | Non-CpG Island | <i>AC009518.4</i>   | 5' UTR        | 0.121             | 9.36E-02           | 1.48E-01           | 1.39E-17 | 1.43E-12             | 1.00E-11                    | 6                         | 1                                        |
| cg05889085 | Non-CpG Island | <i>RP11-473J6.1</i> | TSS1500       | 0.123             | 9.40E-02           | 1.53E-01           | 4.03E-16 | 3.64E-11             | 2.91E-10                    | 7                         | 2                                        |
| cg06381803 | Island         | <i>EML2</i>         | Body          | -0.078            | -9.66E-02          | -5.93E-02          | 6.26E-16 | 5.02E-11             | 4.52E-10                    | 8                         | 1                                        |
| cg19026587 | Non-CpG Island |                     | Intergenic    | -0.093            | -1.16E-01          | -7.04E-02          | 1.61E-15 | 1.16E-10             | 1.16E-09                    | 9                         | 1                                        |
| cg14931378 | Non-CpG Island | <i>RNF103</i>       | Body          | -0.056            | -6.92E-02          | -4.19E-02          | 4.13E-15 | 2.71E-10             | 2.98E-09                    | 10                        | 2                                        |
| cg17420142 | Non-CpG Island | <i>MAPRE2</i>       | Body          | -0.053            | -6.64E-02          | -4.00E-02          | 6.03E-15 | 3.63E-10             | 4.36E-09                    | 11                        | 1                                        |
| cg13480898 | Shore          | <i>C19orf66</i>     | TSS1500       | -0.047            | -5.87E-02          | -3.52E-02          | 1.13E-14 | 6.26E-10             | 8.14E-09                    | 12                        | 1                                        |
| cg24845774 | Non-CpG Island | <i>RP11-473J6.1</i> | TSS1500       | 0.124             | 9.30E-02           | 1.56E-01           | 1.89E-14 | 9.75E-10             | 1.36E-08                    | 7                         | 2                                        |
| cg15108641 | Shelf          | <i>UBTD1</i>        | Body          | 0.050             | 3.73E-02           | 6.30E-02           | 4.01E-14 | 1.93E-09             | 2.90E-08                    | 13                        | 1                                        |
| cg07650421 | Non-CpG Island |                     | Intergenic    | -0.061            | -7.72E-02          | -4.52E-02          | 1.32E-13 | 5.98E-09             | 9.56E-08                    | 14                        | 1                                        |
| cg00284153 | Island         | <i>C10orf125</i>    | TSS200        | -0.048            | -6.04E-02          | -3.52E-02          | 1.64E-13 | 6.98E-09             | 1.19E-07                    | 15                        | 4                                        |
| cg01529253 | Non-CpG Island |                     | Intergenic    | -0.055            | -7.03E-02          | -3.92E-02          | 8.57E-12 | 3.44E-07             | 6.19E-06                    | 16                        | 2                                        |
| cg21510186 | Non-CpG Island | <i>PPP2R2C</i>      | Body          | -0.063            | -8.08E-02          | -4.50E-02          | 9.22E-12 | 3.50E-07             | 6.66E-06                    | 17                        | 2                                        |
| cg02324238 | Non-CpG Island | <i>NALCN-AS1</i>    | Body          | 0.090             | 6.41E-02           | 1.15E-01           | 9.89E-12 | 3.57E-07             | 7.14E-06                    | 18                        | 1                                        |
| cg10185759 | Non-CpG Island |                     | Intergenic    | -0.044            | -5.72E-02          | -3.17E-02          | 1.21E-11 | 4.16E-07             | 8.74E-06                    | 19                        | 1                                        |
| cg01225779 | Shelf          | <i>SQSTM1</i>       | 5' UTR        | -0.066            | -8.52E-02          | -4.70E-02          | 1.76E-11 | 5.79E-07             | 1.27E-05                    | 20                        | 1                                        |
| cg12654412 | Island         | <i>RP11-510M2.1</i> | 3' UTR        | -0.050            | -6.40E-02          | -3.52E-02          | 2.05E-11 | 6.45E-07             | 1.48E-05                    | 21                        | 2                                        |
| cg10133876 | Shore          | <i>CLIC6</i>        | TSS1500       | -0.057            | -7.42E-02          | -4.07E-02          | 2.41E-11 | 7.26E-07             | 1.74E-05                    | 22                        | 2                                        |
| cg04334496 | Non-CpG Island |                     | Intergenic    | -0.059            | -7.58E-02          | -4.14E-02          | 3.84E-11 | 1.11E-06             | 2.77E-05                    | 23                        | 1                                        |
| cg16748130 | Non-CpG Island | <i>RP11-44M6.1</i>  | 5' UTR        | -0.050            | -6.42E-02          | -3.49E-02          | 4.68E-11 | 1.30E-06             | 3.38E-05                    | 24                        | 1                                        |
| cg07948401 | Shore          | <i>CD37</i>         | 3' UTR        | -0.063            | -8.19E-02          | -4.45E-02          | 5.18E-11 | 1.38E-06             | 3.74E-05                    | 25                        | 3                                        |
| cg16425342 | Shore          | <i>ADAMTSL5</i>     | Body          | -0.052            | -6.68E-02          | -3.63E-02          | 5.36E-11 | 1.38E-06             | 3.87E-05                    | 26                        | 1                                        |
| cg15649007 | Non-CpG Island |                     | Intergenic    | -0.057            | -7.43E-02          | -4.03E-02          | 5.93E-11 | 1.48E-06             | 4.28E-05                    | 27                        | 1                                        |
| cg07738234 | Shore          | <i>VGLL2</i>        | 3' UTR        | 0.079             | 5.54E-02           | 1.02E-01           | 6.13E-11 | 1.48E-06             | 4.43E-05                    | 28                        | 1                                        |
| cg03482769 | Island         | <i>RAVER1</i>       | TSS1500       | -0.064            | -8.28E-02          | -4.48E-02          | 6.47E-11 | 1.51E-06             | 4.67E-05                    | 29                        | 1                                        |
| cg05425326 | Non-CpG Island | <i>GLIS3</i>        | 3' UTR        | 0.046             | 3.26E-02           | 6.04E-02           | 8.31E-11 | 1.88E-06             | 6.00E-05                    | 30                        | 1                                        |
| cg01907945 | Non-CpG Island | <i>CLRN1-AS1</i>    | Body          | 0.048             | 3.38E-02           | 6.28E-02           | 9.09E-11 | 1.99E-06             | 6.56E-05                    | 31                        | 1                                        |
| cg07296705 | Non-CpG Island | <i>NTM</i>          | Body          | -0.054            | -7.06E-02          | -3.80E-02          | 9.45E-11 | 2.01E-06             | 6.82E-05                    | 32                        | 1                                        |
| cg13266323 | Non-CpG Island | <i>CADM1</i>        | Body          | 0.072             | 5.04E-02           | 9.40E-02           | 1.16E-10 | 2.40E-06             | 8.38E-05                    | 33                        | 1                                        |
| cg16245318 | Non-CpG Island | <i>MYO5B</i>        | Body          | -0.037            | -4.78E-02          | -2.54E-02          | 1.97E-10 | 3.96E-06             | 1.43E-04                    | 34                        | 1                                        |

## Continuation of Table S 3.2

|            |                |                      |            |        |           |           |          |          |          |    |   |
|------------|----------------|----------------------|------------|--------|-----------|-----------|----------|----------|----------|----|---|
| cg12964519 | Shore          | <i>WNT1</i>          | 3' UTR     | -0.054 | -7.02E-02 | -3.72E-02 | 2.28E-10 | 4.44E-06 | 1.64E-04 | 35 | 1 |
| cg25779645 | Non-CpG Island | <i>R3HDM1</i>        | TSS1500    | -0.048 | -6.24E-02 | -3.31E-02 | 2.45E-10 | 4.66E-06 | 1.77E-04 | 36 | 1 |
| cg05428706 | Shore          | <i>SEMA4G</i>        | 5' UTR     | -0.060 | -7.80E-02 | -4.11E-02 | 3.50E-10 | 6.47E-06 | 2.53E-04 | 4  | 2 |
| cg16525400 | Non-CpG Island | <i>RP11-71E19.1</i>  | 5' UTR     | 0.054  | 3.70E-02  | 7.03E-02  | 3.92E-10 | 7.08E-06 | 2.83E-04 | 37 | 1 |
| cg18394552 | Non-CpG Island |                      | Intergenic | -0.048 | -6.29E-02 | -3.30E-02 | 4.25E-10 | 7.48E-06 | 3.07E-04 | 38 | 1 |
| cg07617039 | Shore          | <i>RSPO3</i>         | TSS1500    | -0.097 | -1.27E-01 | -6.69E-02 | 4.35E-10 | 7.48E-06 | 3.14E-04 | 39 | 1 |
| cg24037821 | Non-CpG Island | <i>UHRF1</i>         | TSS1500    | -0.036 | -4.69E-02 | -2.45E-02 | 5.34E-10 | 8.96E-06 | 3.85E-04 | 40 | 1 |
| cg17060882 | Non-CpG Island | <i>NAALADL1</i>      | Body       | -0.032 | -4.16E-02 | -2.18E-02 | 5.47E-10 | 8.97E-06 | 3.95E-04 | 41 | 1 |
| cg20741479 | Non-CpG Island | <i>TEAD1</i>         | 5' UTR     | 0.033  | 2.25E-02  | 4.33E-02  | 6.67E-10 | 1.05E-05 | 4.82E-04 | 42 | 1 |
| cg04689048 | Island         | <i>PNMA8B</i>        | TSS200     | -0.040 | -5.24E-02 | -2.73E-02 | 6.71E-10 | 1.05E-05 | 4.84E-04 | 43 | 8 |
| cg25028129 | Non-CpG Island | <i>CYP2B6</i>        | Body       | -0.052 | -6.87E-02 | -3.57E-02 | 7.33E-10 | 1.13E-05 | 5.29E-04 | 44 | 1 |
| cg23683730 | Non-CpG Island |                      | Intergenic | 0.061  | 4.19E-02  | 8.07E-02  | 7.70E-10 | 1.16E-05 | 5.56E-04 | 45 | 1 |
| cg05438461 | Shore          | <i>BMF</i>           | TSS1500    | 0.046  | 3.14E-02  | 6.05E-02  | 8.26E-10 | 1.22E-05 | 5.96E-04 | 46 | 1 |
| cg17773655 | Shore          | <i>RASIP1</i>        | TSS200     | -0.045 | -5.88E-02 | -3.03E-02 | 1.07E-09 | 1.55E-05 | 7.75E-04 | 47 | 3 |
| cg06438159 | Non-CpG Island |                      | Intergenic | -0.061 | -7.99E-02 | -4.12E-02 | 1.12E-09 | 1.58E-05 | 8.06E-04 | 48 | 1 |
| cg05646745 | Shore          | <i>FUOM</i>          | TSS1500    | 0.041  | 2.76E-02  | 5.35E-02  | 1.15E-09 | 1.60E-05 | 8.30E-04 | 15 | 4 |
| cg19730422 | Non-CpG Island | <i>CYFIP2</i>        | Body       | -0.042 | -5.56E-02 | -2.86E-02 | 1.18E-09 | 1.61E-05 | 8.53E-04 | 49 | 1 |
| cg01995468 | Island         |                      | Intergenic | -0.062 | -8.15E-02 | -4.17E-02 | 1.67E-09 | 2.24E-05 | 1.21E-03 | 50 | 2 |
| cg04664595 | Non-CpG Island |                      | Intergenic | -0.047 | -6.22E-02 | -3.18E-02 | 1.76E-09 | 2.32E-05 | 1.27E-03 | 51 | 1 |
| cg02568341 | Non-CpG Island |                      | Intergenic | -0.035 | -4.63E-02 | -2.36E-02 | 2.23E-09 | 2.87E-05 | 1.61E-03 | 52 | 1 |
| cg04491800 | Non-CpG Island |                      | Intergenic | -0.036 | -4.75E-02 | -2.41E-02 | 2.88E-09 | 3.64E-05 | 2.08E-03 | 53 | 1 |
| cg26788044 | Non-CpG Island |                      | Intergenic | 0.057  | 3.81E-02  | 7.53E-02  | 2.97E-09 | 3.70E-05 | 2.15E-03 | 54 | 1 |
| cg16210758 | Non-CpG Island |                      | Intergenic | -0.041 | -5.48E-02 | -2.76E-02 | 3.54E-09 | 4.33E-05 | 2.56E-03 | 55 | 1 |
| cg15721584 | Non-CpG Island | <i>SOX2OT</i>        | TSS1500    | 0.054  | 3.64E-02  | 7.23E-02  | 3.74E-09 | 4.50E-05 | 2.70E-03 | 56 | 3 |
| cg05923226 | Shore          | <i>CCDC105</i>       | TSS200     | -0.039 | -5.18E-02 | -2.60E-02 | 3.82E-09 | 4.51E-05 | 2.76E-03 | 57 | 4 |
| cg01446044 | Non-CpG Island | <i>RN5S41</i>        | TSS1500    | -0.029 | -3.87E-02 | -1.95E-02 | 3.87E-09 | 4.51E-05 | 2.79E-03 | 58 | 1 |
| cg25542319 | Non-CpG Island | <i>MCC</i>           | Body       | 0.053  | 3.51E-02  | 7.00E-02  | 4.65E-09 | 5.27E-05 | 3.36E-03 | 59 | 1 |
| cg11688949 | Non-CpG Island | <i>STMN4</i>         | TSS200     | -0.050 | -6.61E-02 | -3.31E-02 | 4.67E-09 | 5.27E-05 | 3.37E-03 | 60 | 2 |
| cg14583825 | Island         | <i>PRKCG</i>         | Body       | -0.030 | -4.06E-02 | -2.03E-02 | 5.39E-09 | 5.98E-05 | 3.89E-03 | 61 | 1 |
| cg05864097 | Shelf          |                      | Intergenic | 0.048  | 3.22E-02  | 6.45E-02  | 5.51E-09 | 6.03E-05 | 3.98E-03 | 62 | 1 |
| cg24199203 | Shore          | <i>NFIC</i>          | Body       | -0.058 | -7.74E-02 | -3.86E-02 | 6.05E-09 | 6.48E-05 | 4.37E-03 | 63 | 1 |
| cg08946973 | Non-CpG Island | <i>INSC</i>          | Body       | -0.033 | -4.44E-02 | -2.21E-02 | 6.10E-09 | 6.48E-05 | 4.40E-03 | 64 | 1 |
| cg04255201 | Shore          | <i>PNMA8B</i>        | TSS200     | -0.040 | -5.31E-02 | -2.64E-02 | 6.43E-09 | 6.72E-05 | 4.64E-03 | 43 | 8 |
| cg08043020 | Shelf          | <i>RP11-236J17.3</i> | TSS200     | -0.041 | -5.54E-02 | -2.75E-02 | 6.98E-09 | 7.20E-05 | 5.04E-03 | 65 | 2 |
| cg26390944 | Non-CpG Island | <i>EDNRB-AS1</i>     | Body       | -0.048 | -6.42E-02 | -3.18E-02 | 7.55E-09 | 7.68E-05 | 5.45E-03 | 66 | 1 |
| cg27554551 | Shore          | <i>VASN</i>          | TSS1500    | -0.036 | -4.76E-02 | -2.36E-02 | 7.66E-09 | 7.68E-05 | 5.53E-03 | 67 | 1 |
| cg03834786 | Non-CpG Island |                      | Intergenic | -0.044 | -5.89E-02 | -2.92E-02 | 7.92E-09 | 7.83E-05 | 5.72E-03 | 68 | 1 |
| cg03535930 | Non-CpG Island |                      | Intergenic | -0.066 | -8.82E-02 | -4.36E-02 | 8.93E-09 | 8.70E-05 | 6.44E-03 | 69 | 1 |
| cg02489552 | Shore          | <i>CCDC105</i>       | TSS200     | -0.060 | -7.97E-02 | -3.94E-02 | 9.04E-09 | 8.70E-05 | 6.53E-03 | 57 | 4 |

## Continuation of Table S 3.2

|            |                |                      |            |        |           |           |          |          |          |     |   |
|------------|----------------|----------------------|------------|--------|-----------|-----------|----------|----------|----------|-----|---|
| cg03180777 | Non-CpG Island |                      | Intergenic | -0.070 | -9.33E-02 | -4.60E-02 | 9.94E-09 | 9.44E-05 | 7.17E-03 | 70  | 2 |
| cg09815261 | Non-CpG Island | <i>OTOG</i>          | Body       | -0.035 | -4.70E-02 | -2.31E-02 | 1.01E-08 | 9.49E-05 | 7.30E-03 | 71  | 1 |
| cg08897705 | Non-CpG Island |                      | Intergenic | 0.045  | 2.94E-02  | 5.97E-02  | 1.05E-08 | 9.70E-05 | 7.56E-03 | 72  | 1 |
| cg17442683 | Non-CpG Island | <i>RERE</i>          | Body       | 0.038  | 2.53E-02  | 5.15E-02  | 1.16E-08 | 1.05E-04 | 8.39E-03 | 73  | 1 |
| cg20697394 | Shelf          |                      | Intergenic | -0.037 | -5.01E-02 | -2.46E-02 | 1.17E-08 | 1.05E-04 | 8.43E-03 | 74  | 1 |
| cg19982061 | Non-CpG Island | <i>CCDC12</i>        | Body       | -0.029 | -3.84E-02 | -1.88E-02 | 1.23E-08 | 1.10E-04 | 8.89E-03 | 75  | 1 |
| cg04637372 | Non-CpG Island | <i>VWA5B1</i>        | 5' UTR     | -0.050 | -6.71E-02 | -3.29E-02 | 1.26E-08 | 1.11E-04 | 9.08E-03 | 76  | 2 |
| cg14455466 | Non-CpG Island | <i>CTC-321K16.1</i>  | 5' UTR     | -0.036 | -4.87E-02 | -2.38E-02 | 1.29E-08 | 1.12E-04 | 9.34E-03 | 77  | 1 |
| cg11410682 | Island         | <i>PNMA88</i>        | TSS200     | -0.052 | -6.95E-02 | -3.40E-02 | 1.32E-08 | 1.12E-04 | 9.54E-03 | 43  | 8 |
| cg09653610 | Non-CpG Island |                      | Intergenic | 0.072  | 4.74E-02  | 9.68E-02  | 1.33E-08 | 1.12E-04 | 9.58E-03 | 78  | 1 |
| cg21345949 | Island         | <i>C10orf125</i>     | TSS200     | -0.047 | -6.36E-02 | -3.11E-02 | 1.33E-08 | 1.12E-04 | 9.60E-03 | 15  | 4 |
| cg04211591 | Non-CpG Island |                      | Intergenic | 0.053  | 3.51E-02  | 7.17E-02  | 1.36E-08 | 1.13E-04 | 9.84E-03 | 79  | 1 |
| cg22761431 | Shore          | <i>EFNB3</i>         | Body       | -0.041 | -5.53E-02 | -2.71E-02 | 1.39E-08 | 1.14E-04 | 1.00E-02 | 80  | 1 |
| cg12662084 | Non-CpG Island | <i>KIF13A</i>        | Body       | -0.085 | -1.14E-01 | -5.58E-02 | 1.42E-08 | 1.14E-04 | 1.03E-02 | 81  | 1 |
| cg14768892 | Non-CpG Island | <i>FSIP1</i>         | Body       | -0.042 | -5.67E-02 | -2.77E-02 | 1.42E-08 | 1.14E-04 | 1.03E-02 | 82  | 1 |
| cg10887181 | Non-CpG Island | <i>TUBA4B</i>        | Body       | -0.036 | -4.80E-02 | -2.34E-02 | 1.58E-08 | 1.25E-04 | 1.14E-02 | 83  | 1 |
| cg17724147 | Non-CpG Island | <i>RP11-319E12.1</i> | TSS1500    | -0.035 | -4.67E-02 | -2.27E-02 | 1.60E-08 | 1.25E-04 | 1.15E-02 | 84  | 1 |
| cg20821903 | Shore          | <i>TFAP4</i>         | Body       | -0.043 | -5.78E-02 | -2.82E-02 | 1.61E-08 | 1.25E-04 | 1.16E-02 | 85  | 1 |
| cg26211349 | Non-CpG Island | <i>HRNBP3</i>        | 5' UTR     | 0.049  | 3.19E-02  | 6.56E-02  | 1.77E-08 | 1.36E-04 | 1.28E-02 | 86  | 1 |
| cg10663081 | Non-CpG Island |                      | Intergenic | -0.029 | -3.97E-02 | -1.93E-02 | 1.85E-08 | 1.40E-04 | 1.33E-02 | 87  | 1 |
| cg08119153 | Non-CpG Island | <i>NRXN3</i>         | 5' UTR     | -0.056 | -7.50E-02 | -3.64E-02 | 1.89E-08 | 1.42E-04 | 1.37E-02 | 88  | 1 |
| cg18369034 | Shore          | <i>PTGDR</i>         | TSS1500    | -0.073 | -9.87E-02 | -4.79E-02 | 1.93E-08 | 1.44E-04 | 1.40E-02 | 89  | 2 |
| cg00641210 | Island         | <i>BCL6B</i>         | Body       | -0.045 | -6.03E-02 | -2.92E-02 | 1.97E-08 | 1.45E-04 | 1.42E-02 | 90  | 1 |
| cg10546626 | Shelf          | <i>GNASAS</i>        | Body       | 0.032  | 2.12E-02  | 4.37E-02  | 2.01E-08 | 1.46E-04 | 1.45E-02 | 91  | 1 |
| cg07750834 | Non-CpG Island | <i>UACA</i>          | Body       | -0.043 | -5.85E-02 | -2.83E-02 | 2.11E-08 | 1.52E-04 | 1.52E-02 | 92  | 1 |
| cg25281677 | Shore          |                      | Intergenic | -0.030 | -4.01E-02 | -1.94E-02 | 2.16E-08 | 1.55E-04 | 1.56E-02 | 93  | 1 |
| cg05985069 | Non-CpG Island | <i>IL12RB1</i>       | Body       | -0.037 | -4.96E-02 | -2.40E-02 | 2.19E-08 | 1.55E-04 | 1.58E-02 | 94  | 1 |
| cg17929959 | Non-CpG Island |                      | Intergenic | 0.045  | 2.95E-02  | 6.11E-02  | 2.31E-08 | 1.62E-04 | 1.66E-02 | 95  | 1 |
| cg08268947 | Non-CpG Island | <i>SORCS2</i>        | Body       | -0.039 | -5.28E-02 | -2.55E-02 | 2.35E-08 | 1.63E-04 | 1.70E-02 | 96  | 2 |
| cg15166089 | Shore          | <i>CD164L2</i>       | TSS1500    | -0.040 | -5.45E-02 | -2.63E-02 | 2.40E-08 | 1.64E-04 | 1.73E-02 | 97  | 2 |
| cg00519063 | Shore          | <i>GCLM</i>          | Body       | -0.033 | -4.40E-02 | -2.12E-02 | 2.40E-08 | 1.64E-04 | 1.74E-02 | 98  | 1 |
| cg14176088 | Non-CpG Island | <i>PIK3CD</i>        | Body       | 0.041  | 2.69E-02  | 5.58E-02  | 2.48E-08 | 1.67E-04 | 1.79E-02 | 99  | 1 |
| cg21445325 | Non-CpG Island | <i>SORCS2</i>        | Body       | -0.041 | -5.47E-02 | -2.64E-02 | 2.51E-08 | 1.67E-04 | 1.81E-02 | 96  | 2 |
| cg06402858 | Non-CpG Island | <i>SHC3</i>          | Body       | 0.035  | 2.28E-02  | 4.73E-02  | 2.63E-08 | 1.74E-04 | 1.90E-02 | 100 | 1 |
| cg01228288 | Non-CpG Island |                      | Intergenic | 0.050  | 3.23E-02  | 6.72E-02  | 2.91E-08 | 1.91E-04 | 2.10E-02 | 101 | 2 |
| cg09956442 | Non-CpG Island |                      | Intergenic | -0.044 | -5.94E-02 | -2.85E-02 | 2.97E-08 | 1.92E-04 | 2.14E-02 | 102 | 2 |
| cg20692268 | Non-CpG Island |                      | Intergenic | -0.041 | -5.55E-02 | -2.66E-02 | 2.97E-08 | 1.92E-04 | 2.15E-02 | 103 | 1 |
| cg24976744 | Non-CpG Island |                      | Intergenic | -0.065 | -8.79E-02 | -4.22E-02 | 3.02E-08 | 1.93E-04 | 2.18E-02 | 104 | 1 |
| cg12110939 | Non-CpG Island |                      | Intergenic | -0.039 | -5.26E-02 | -2.52E-02 | 3.14E-08 | 1.98E-04 | 2.27E-02 | 105 | 1 |

# Continuation of Table S 3.2

|            |                |               |            |              |           |              |          |          |          |     |   |
|------------|----------------|---------------|------------|--------------|-----------|--------------|----------|----------|----------|-----|---|
| cg16198075 | Non-CpG Island | PPP2R2C       | Body       | -0.062       | -8.42E-02 | -4.03E-02    | 3.15E-08 | 1.98E-04 | 2.27E-02 | 17  | 2 |
| cg10613567 | Shore          | SACS          | TSS1500    | -0.048       | -6.44E-02 | -3.08E-02    | 3.26E-08 | 2.03E-04 | 2.36E-02 | 106 | 2 |
| cg16125894 | Non-CpG Island | KIAA1614      | 3' UTR     | -0.055       | -7.40E-02 | -3.54E-02    | 3.30E-08 | 2.03E-04 | 2.38E-02 | 107 | 1 |
| cg24807394 | Non-CpG Island |               | Intergenic | -0.038       | -5.10E-02 | -2.44E-02    | 3.37E-08 | 2.06E-04 | 2.44E-02 | 108 | 1 |
| cg23836403 | Shore          | RP11-697H9.3  | TSS1500    | -0.036       | -4.80E-02 | -2.30E-02    | 3.44E-08 | 2.09E-04 | 2.48E-02 | 109 | 1 |
| cg08655800 | Non-CpG Island | KIAA1324      | Body       | -0.048       | -6.48E-02 | -3.09E-02    | 3.52E-08 | 2.11E-04 | 2.54E-02 | 110 | 1 |
| cg05816193 | Shelf          | HIST1H1A      | TSS200     | -0.055       | -7.39E-02 | -3.53E-02    | 3.54E-08 | 2.11E-04 | 2.56E-02 | 111 | 1 |
| cg18152558 | Non-CpG Island | PARM1         | Body       | 0.057        | 3.67E-02  | 7.73E-02     | 4.21E-08 | 2.49E-04 | 3.04E-02 | 112 | 1 |
| cg18932116 | Shore          | CCDC105       | TSS200     | -0.063       | -8.53E-02 | -4.06E-02    | 4.26E-08 | 2.50E-04 | 3.08E-02 | 57  | 4 |
| cg00022866 | Shore          | CCDC88B       | Body       | -0.057       | -7.73E-02 | -3.67E-02    | 4.29E-08 | 2.50E-04 | 3.09E-02 | 113 | 2 |
| cg13716566 | Shore          |               | Intergenic | -0.030       | -4.02E-02 | -1.90E-02    | 4.75E-08 | 2.74E-04 | 3.43E-02 | 114 | 1 |
| cg13224253 | Non-CpG Island |               | Intergenic | -0.044       | -5.94E-02 | -2.82E-02    | 4.79E-08 | 2.74E-04 | 3.46E-02 | 115 | 1 |
| cg16340410 | Non-CpG Island | PPARG         | 5' UTR     | -0.042       | -5.70E-02 | -2.70E-02    | 4.88E-08 | 2.77E-04 | 3.52E-02 | 116 | 1 |
| cg16227392 | Non-CpG Island |               | Intergenic | 0.065        | 4.21E-02  | 8.89E-02     | 5.18E-08 | 2.92E-04 | 3.74E-02 | 101 | 2 |
| cg05473085 | Non-CpG Island | CACNA1E       | Body       | -0.027       | -3.72E-02 | -1.76E-02    | 5.26E-08 | 2.94E-04 | 3.80E-02 | 117 | 1 |
| cg01436424 | Non-CpG Island | CAMTA1        | Body       | -0.024       | -3.29E-02 | -1.55E-02    | 5.48E-08 | 3.03E-04 | 3.96E-02 | 118 | 1 |
| cg21889405 | Non-CpG Island | IL6R          | 3' UTR     | 0.033        | 2.11E-02  | 4.47E-02     | 5.51E-08 | 3.03E-04 | 3.97E-02 | 119 | 1 |
| cg03969996 | Island         | PNMA8B        | First Exon | -0.053       | -7.20E-02 | -3.39E-02    | 5.94E-08 | 3.25E-04 | 4.29E-02 | 43  | 8 |
| cg18318818 | Shore          | CACNA1G       | Body       | -0.027       | -3.71E-02 | -1.74E-02    | 6.42E-08 | 3.46E-04 | 4.63E-02 | 120 | 1 |
| cg11035100 | Non-CpG Island | C20orf114     | TSS200     | -0.029       | -3.88E-02 | -1.83E-02    | 6.43E-08 | 3.46E-04 | 4.64E-02 | 121 | 1 |
| cg15133043 | Non-CpG Island |               | Intergenic | -0.037       | -5.01E-02 | -2.35E-02    | 6.73E-08 | 3.60E-04 | 4.86E-02 | 122 | 1 |
| cg26358374 | Non-CpG Island | TRPV4         | 5' UTR     | -0.035314531 | -4.81E-02 | -0.022528154 | 7.29E-08 | 3.87E-04 | 0.05     | 123 | 1 |
| cg03736069 | Shelf          | MAPK4         | Body       | -0.049736985 | -6.78E-02 | -0.031715553 | 7.45E-08 | 3.92E-04 | 0.05     | 124 | 1 |
| cg09176309 | Shelf          | RFX2          | Body       | 0.034298123  | 2.19E-02  | 0.046730787  | 7.54E-08 | 3.95E-04 | 0.05     | 125 | 1 |
| cg11902728 | Island         | MAG           | Body       | -0.022790967 | -3.11E-02 | -0.014521757 | 7.75E-08 | 3.98E-04 | 0.06     | 126 | 1 |
| cg15512289 | Non-CpG Island |               | Intergenic | 0.046332447  | 2.95E-02  | 0.063144012  | 7.77E-08 | 3.98E-04 | 0.06     | 127 | 1 |
| cg00169725 | Island         |               | Intergenic | -0.062947473 | -8.58E-02 | -0.040106917 | 7.77E-08 | 3.98E-04 | 0.06     | 50  | 2 |
| cg07361315 | Non-CpG Island | CLMN          | Body       | -0.035443649 | -4.83E-02 | -0.022559398 | 8.20E-08 | 4.17E-04 | 0.06     | 128 | 1 |
| cg00492070 | Shore          | CD47          | TSS1500    | 0.072100414  | 4.59E-02  | 0.098333776  | 8.42E-08 | 4.25E-04 | 0.06     | 129 | 1 |
| cg04732968 | Shelf          |               | Intergenic | -0.035327206 | -4.82E-02 | -0.022469898 | 8.49E-08 | 4.25E-04 | 0.06     | 130 | 1 |
| cg01562540 | Non-CpG Island |               | Intergenic | -0.0436828   | -5.96E-02 | -0.027781934 | 8.53E-08 | 4.25E-04 | 0.06     | 131 | 1 |
| cg05685272 | Non-CpG Island | OPCML         | Body       | -0.038397993 | -5.24E-02 | -0.024416172 | 8.61E-08 | 4.26E-04 | 0.06     | 132 | 1 |
| cg05989795 | Shore          | LMF1          | TSS1500    | -0.052245025 | -7.13E-02 | -0.033201288 | 8.88E-08 | 4.36E-04 | 0.06     | 133 | 1 |
| cg04605617 | Non-CpG Island | PLA2G2C       | First Exon | 0.066928649  | 4.25E-02  | 0.091377289  | 9.45E-08 | 4.61E-04 | 0.07     | 134 | 3 |
| cg07684019 | Shore          | PSD2          | TSS1500    | -0.032429182 | -4.43E-02 | -0.020579804 | 9.53E-08 | 4.61E-04 | 0.07     | 135 | 1 |
| cg25744957 | Non-CpG Island | FLJ36777      | TSS1500    | -0.025186278 | -3.44E-02 | -0.015980022 | 9.63E-08 | 4.61E-04 | 0.07     | 136 | 1 |
| cg19420069 | Non-CpG Island | RP11-351A20.2 | TSS1500    | 0.055508916  | 3.52E-02  | 0.075799247  | 9.63E-08 | 4.61E-04 | 0.07     | 137 | 4 |
| cg27106513 | Island         | RGS20         | Body       | 0.027481006  | 1.74E-02  | 0.037545099  | 1.02E-07 | 4.83E-04 | 0.07     | 138 | 1 |
| cg22563815 | Shore          | CHRNA5        | TSS1500    | -0.044057296 | -6.02E-02 | -0.027916218 | 1.03E-07 | 4.85E-04 | 0.07     | 139 | 2 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg10766021 | Shore          | <i>ZNF783</i>        | Body       | -0.037089319 | -5.07E-02 | -0.023492219 | 1.05E-07 | 4.91E-04 | 0.08 | 140 | 1 |
| cg15662038 | Non-CpG Island |                      | Intergenic | -0.028330376 | -3.88E-02 | -0.017902506 | 1.18E-07 | 5.48E-04 | 0.08 | 141 | 1 |
| cg10965265 | Non-CpG Island |                      | Intergenic | 0.058102201  | 3.67E-02  | 0.079493749  | 1.19E-07 | 5.48E-04 | 0.09 | 142 | 1 |
| cg19696491 | Shore          | <i>CHRNA5</i>        | TSS1500    | -0.044369437 | -6.07E-02 | -0.027999826 | 1.26E-07 | 5.78E-04 | 0.09 | 139 | 2 |
| cg21284823 | Non-CpG Island | <i>LRFN2</i>         | 5' UTR     | -0.029134801 | -3.99E-02 | -0.018374964 | 1.29E-07 | 5.92E-04 | 0.09 | 143 | 1 |
| cg16764781 | Non-CpG Island | <i>TP63</i>          | TSS1500    | 0.049447603  | 3.12E-02  | 0.067716889  | 1.31E-07 | 5.93E-04 | 0.09 | 144 | 3 |
| cg12164925 | Non-CpG Island |                      | Intergenic | -0.027844939 | -3.81E-02 | -0.017555335 | 1.32E-07 | 5.93E-04 | 0.10 | 145 | 1 |
| cg10283165 | Non-CpG Island | <i>USHBP1</i>        | TSS200     | -0.031918056 | -4.37E-02 | -0.020121278 | 1.32E-07 | 5.93E-04 | 0.10 | 146 | 1 |
| cg01218012 | Non-CpG Island | <i>DLEU1</i>         | Body       | -0.03510008  | -4.81E-02 | -0.022115573 | 1.36E-07 | 6.05E-04 | 0.10 | 147 | 1 |
| cg08241514 | Island         | <i>IL17REL</i>       | Body       | -0.047175078 | -6.46E-02 | -0.029704924 | 1.40E-07 | 6.18E-04 | 0.10 | 148 | 2 |
| cg04189678 | Shore          | <i>ITGBL1</i>        | 5' UTR     | 0.054988805  | 3.46E-02  | 0.075355008  | 1.40E-07 | 6.18E-04 | 0.10 | 149 | 2 |
| cg27110790 | Non-CpG Island |                      | Intergenic | -0.044285881 | -6.07E-02 | -0.027862292 | 1.46E-07 | 6.32E-04 | 0.11 | 150 | 1 |
| cg27346773 | Non-CpG Island |                      | Intergenic | -0.038878397 | -5.33E-02 | -0.024460145 | 1.46E-07 | 6.32E-04 | 0.11 | 151 | 1 |
| cg17899417 | Non-CpG Island | <i>DNAH17</i>        | Body       | -0.041973869 | -5.75E-02 | -0.02640612  | 1.46E-07 | 6.32E-04 | 0.11 | 152 | 1 |
| cg24668753 | Non-CpG Island |                      | Intergenic | -0.036725183 | -5.04E-02 | -0.023095727 | 1.49E-07 | 6.39E-04 | 0.11 | 153 | 1 |
| cg07252792 | Shelf          | <i>CCDC116</i>       | TSS200     | -0.04270987  | -5.86E-02 | -0.02684772  | 1.52E-07 | 6.48E-04 | 0.11 | 154 | 2 |
| cg11670749 | Non-CpG Island |                      | Intergenic | -0.052632143 | -7.22E-02 | -0.033066877 | 1.56E-07 | 6.62E-04 | 0.11 | 155 | 1 |
| cg25073708 | Island         | <i>C9orf106</i>      | 5' UTR     | -0.038913589 | -5.34E-02 | -0.024440652 | 1.58E-07 | 6.67E-04 | 0.11 | 156 | 1 |
| cg21152304 | Non-CpG Island | <i>PLA2G2C</i>       | TSS200     | 0.051823608  | 3.25E-02  | 0.071116745  | 1.62E-07 | 6.82E-04 | 0.12 | 134 | 3 |
| cg24718465 | Island         | <i>PNMA88</i>        | First Exon | -0.040833436 | -5.61E-02 | -0.025591057 | 1.75E-07 | 7.30E-04 | 0.13 | 43  | 8 |
| cg17480291 | Shelf          | <i>TRIO</i>          | Body       | -0.055274157 | -7.60E-02 | -0.034598154 | 1.86E-07 | 7.70E-04 | 0.13 | 157 | 2 |
| cg26428830 | Non-CpG Island |                      | Intergenic | -0.054520654 | -7.49E-02 | -0.034109562 | 1.90E-07 | 7.83E-04 | 0.14 | 158 | 1 |
| cg25520960 | Shelf          |                      | Intergenic | -0.040259418 | -5.53E-02 | -0.025184349 | 1.91E-07 | 7.83E-04 | 0.14 | 159 | 1 |
| cg11321181 | Shelf          | <i>UNC13A</i>        | Body       | -0.065217507 | -8.97E-02 | -0.040778329 | 1.95E-07 | 7.95E-04 | 0.14 | 160 | 1 |
| cg04779737 | Non-CpG Island |                      | Intergenic | -0.04752872  | -6.54E-02 | -0.029702655 | 2.00E-07 | 8.10E-04 | 0.14 | 161 | 1 |
| cg07286682 | Non-CpG Island | <i>MIR548H4</i>      | Body       | -0.034695585 | -4.77E-02 | -0.021672353 | 2.04E-07 | 8.22E-04 | 0.15 | 162 | 2 |
| cg15501231 | Non-CpG Island | <i>C6orf27</i>       | Body       | -0.034450675 | -4.74E-02 | -0.021516471 | 2.05E-07 | 8.22E-04 | 0.15 | 163 | 1 |
| cg07621682 | Non-CpG Island | <i>CCDC97</i>        | Body       | -0.034821339 | -4.79E-02 | -0.021746051 | 2.06E-07 | 8.22E-04 | 0.15 | 164 | 1 |
| cg23608915 | Non-CpG Island | <i>MAP4K4</i>        | Body       | -0.048821909 | -6.72E-02 | -0.030457018 | 2.17E-07 | 8.55E-04 | 0.16 | 165 | 1 |
| cg13839793 | Non-CpG Island |                      | Intergenic | 0.045904304  | 2.86E-02  | 0.063175673  | 2.18E-07 | 8.55E-04 | 0.16 | 166 | 1 |
| cg04826368 | Non-CpG Island |                      | Intergenic | -0.038412074 | -5.29E-02 | -0.023958327 | 2.18E-07 | 8.55E-04 | 0.16 | 167 | 1 |
| cg27519981 | Shore          | <i>SACS</i>          | TSS1500    | -0.036194275 | -4.98E-02 | -0.02257371  | 2.19E-07 | 8.55E-04 | 0.16 | 106 | 2 |
| cg03843031 | Island         | <i>UBE2C</i>         | TSS1500    | -0.057146753 | -7.87E-02 | -0.035613685 | 2.27E-07 | 8.81E-04 | 0.16 | 168 | 1 |
| cg00884041 | Shore          | <i>SBK2</i>          | TSS200     | -0.032874706 | -4.53E-02 | -0.020481486 | 2.30E-07 | 8.87E-04 | 0.17 | 169 | 2 |
| cg10815470 | Non-CpG Island | <i>VTA1</i>          | Body       | -0.034580379 | -4.76E-02 | -0.021536798 | 2.33E-07 | 8.94E-04 | 0.17 | 170 | 1 |
| cg11921539 | Non-CpG Island |                      | Intergenic | -0.034907986 | -4.81E-02 | -0.02173868  | 2.35E-07 | 8.94E-04 | 0.17 | 171 | 1 |
| cg14320320 | Non-CpG Island | <i>RP11-194G10.1</i> | TSS200     | -0.055710556 | -7.67E-02 | -0.034687961 | 2.36E-07 | 8.94E-04 | 0.17 | 172 | 2 |
| cg09222732 | Non-CpG Island |                      | Intergenic | 0.042622901  | 2.65E-02  | 0.058707919  | 2.37E-07 | 8.94E-04 | 0.17 | 173 | 1 |
| cg22959742 | Non-CpG Island | <i>FRMD4A</i>        | Body       | 0.039728208  | 2.47E-02  | 0.054753788  | 2.51E-07 | 9.36E-04 | 0.18 | 174 | 1 |



## Continuation of Table S 3.2

|            |                |              |            |              |           |              |          |          |      |     |   |
|------------|----------------|--------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg26964636 | Non-CpG Island | KIAA1217     | Body       | 0.047885991  | 2.98E-02  | 0.066000287  | 2.52E-07 | 9.36E-04 | 0.18 | 175 | 1 |
| cg02417857 | Non-CpG Island | SOX2OT       | TSS1500    | 0.04900422   | 3.05E-02  | 0.067541862  | 2.53E-07 | 9.36E-04 | 0.18 | 56  | 3 |
| cg22588682 | Non-CpG Island |              | Intergenic | -0.034310965 | -4.73E-02 | -0.021330961 | 2.53E-07 | 9.36E-04 | 0.18 | 176 | 1 |
| cg22486366 | Non-CpG Island | RP3-340N1.2  | TSS1500    | 0.045558508  | 2.83E-02  | 0.062807163  | 2.58E-07 | 9.52E-04 | 0.19 | 177 | 1 |
| cg22078805 | Island         | FAM171A2     | Body       | -0.06015423  | -8.29E-02 | -0.037359612 | 2.65E-07 | 9.70E-04 | 0.19 | 178 | 1 |
| cg08419476 | Non-CpG Island | TEAD1        | 5' UTR     | 0.037033434  | 2.30E-02  | 0.051095035  | 2.79E-07 | 1.02E-03 | 0.20 | 179 | 1 |
| cg06394677 | Non-CpG Island |              | Intergenic | -0.051776202 | -7.15E-02 | -0.032097704 | 2.87E-07 | 1.03E-03 | 0.21 | 180 | 1 |
| cg20421058 | Non-CpG Island | SORCS2       | Body       | -0.047232383 | -6.52E-02 | -0.029280077 | 2.87E-07 | 1.03E-03 | 0.21 | 181 | 1 |
| cg23848052 | Non-CpG Island | MEGF11       | Body       | -0.036528021 | -5.04E-02 | -0.022644186 | 2.87E-07 | 1.03E-03 | 0.21 | 182 | 1 |
| cg05211864 | Shore          | P3H4         | TSS1500    | -0.033971722 | -4.69E-02 | -0.02105695  | 2.89E-07 | 1.03E-03 | 0.21 | 183 | 4 |
| cg23206630 | Non-CpG Island | TLN2         | Body       | 0.035298426  | 2.19E-02  | 0.048727505  | 2.95E-07 | 1.05E-03 | 0.21 | 184 | 1 |
| cg10314788 | Non-CpG Island | NRCAM        | 5' UTR     | -0.042148726 | -5.82E-02 | -0.026095013 | 3.04E-07 | 1.07E-03 | 0.22 | 185 | 1 |
| cg17951445 | Non-CpG Island | PCDH7        | Body       | 0.055921136  | 3.46E-02  | 0.077228467  | 3.07E-07 | 1.07E-03 | 0.22 | 186 | 1 |
| cg11909043 | Non-CpG Island | KIAA1217     | Body       | 0.030981941  | 1.92E-02  | 0.042787991  | 3.08E-07 | 1.07E-03 | 0.22 | 187 | 1 |
| cg16863678 | Shore          | PMVK         | TSS1500    | -0.038527786 | -5.32E-02 | -0.023845548 | 3.08E-07 | 1.07E-03 | 0.22 | 188 | 1 |
| cg19222775 | Non-CpG Island |              | Intergenic | 0.038732339  | 2.40E-02  | 0.05350023   | 3.13E-07 | 1.08E-03 | 0.23 | 189 | 1 |
| cg06383241 | Non-CpG Island | MAP1LC3B2    | TSS200     | -0.040236137 | -5.56E-02 | -0.024886851 | 3.17E-07 | 1.09E-03 | 0.23 | 190 | 1 |
| cg12382670 | Non-CpG Island | CSMD2        | Body       | -0.036892761 | -5.10E-02 | -0.022802458 | 3.27E-07 | 1.12E-03 | 0.24 | 191 | 1 |
| cg20874670 | Non-CpG Island |              | Intergenic | -0.042080954 | -5.82E-02 | -0.026008644 | 3.27E-07 | 1.12E-03 | 0.24 | 192 | 1 |
| cg01468567 | Shore          | TEAD2        | 3' UTR     | -0.029390327 | -4.06E-02 | -0.018159643 | 3.32E-07 | 1.13E-03 | 0.24 | 25  | 3 |
| cg01090161 | Non-CpG Island | GRM4         | Body       | -0.034148813 | -4.72E-02 | -0.021097455 | 3.33E-07 | 1.13E-03 | 0.24 | 193 | 1 |
| cg04814021 | Non-CpG Island |              | Intergenic | -0.023956268 | -3.31E-02 | -0.014794871 | 3.39E-07 | 1.14E-03 | 0.24 | 194 | 1 |
| cg16400794 | Shore          | CD164L2      | TSS200     | -0.040317771 | -5.57E-02 | -0.024897524 | 3.40E-07 | 1.14E-03 | 0.25 | 97  | 2 |
| cg09331496 | Non-CpG Island |              | Intergenic | 0.033855087  | 2.09E-02  | 0.046828656  | 3.58E-07 | 1.19E-03 | 0.26 | 195 | 1 |
| cg09221413 | Shore          | P3H4         | TSS1500    | -0.030093379 | -4.16E-02 | -0.018553696 | 3.64E-07 | 1.21E-03 | 0.26 | 183 | 4 |
| cg12818967 | Non-CpG Island | RP11-428F8.2 | 3' UTR     | -0.05276581  | -7.30E-02 | -0.032505158 | 3.77E-07 | 1.25E-03 | 0.27 | 196 | 1 |
| cg15379400 | Non-CpG Island | TNS4         | 5' UTR     | -0.026256212 | -3.63E-02 | -0.016170441 | 3.81E-07 | 1.26E-03 | 0.28 | 197 | 1 |
| cg05087792 | Non-CpG Island | DSE          | 5' UTR     | -0.043263238 | -5.99E-02 | -0.026624545 | 3.93E-07 | 1.29E-03 | 0.28 | 198 | 1 |
| cg26810816 | Island         |              | Intergenic | 0.039867993  | 2.45E-02  | 0.055220961  | 4.07E-07 | 1.33E-03 | 0.29 | 199 | 1 |
| cg17417600 | Shore          | GNPNAT1      | TSS200     | -0.053186939 | -7.37E-02 | -0.032666747 | 4.27E-07 | 1.39E-03 | 0.31 | 200 | 1 |
| cg01757312 | Island         | SOX1         | TSS1500    | 0.0293137    | 1.80E-02  | 0.04062818   | 4.32E-07 | 1.40E-03 | 0.31 | 201 | 1 |
| cg17897187 | Shore          | KIAA1026     | TSS1500    | 0.05007403   | 3.07E-02  | 0.069418197  | 4.42E-07 | 1.42E-03 | 0.32 | 202 | 1 |
| cg09082427 | Shore          | NSMF         | Body       | -0.035514838 | -4.92E-02 | -0.021794847 | 4.42E-07 | 1.42E-03 | 0.32 | 203 | 2 |
| cg13960680 | Non-CpG Island | SLC15A1      | Body       | 0.044708429  | 2.74E-02  | 0.062039322  | 4.84E-07 | 1.54E-03 | 0.35 | 204 | 1 |
| cg11990488 | Non-CpG Island |              | Intergenic | -0.047034713 | -6.53E-02 | -0.028798439 | 4.86E-07 | 1.55E-03 | 0.35 | 205 | 1 |
| cg25935178 | Island         | MIR4259      | TSS200     | -0.04731997  | -6.57E-02 | -0.028953289 | 5.00E-07 | 1.58E-03 | 0.36 | 206 | 3 |
| cg09707395 | Non-CpG Island | ARHGEF15     | TSS1500    | -0.033817636 | -4.69E-02 | -0.020685434 | 5.06E-07 | 1.60E-03 | 0.37 | 207 | 1 |
| cg07399628 | Non-CpG Island | ELMO1        | TSS1500    | -0.048987873 | -6.80E-02 | -0.029956595 | 5.12E-07 | 1.61E-03 | 0.37 | 208 | 1 |
| cg13938970 | Non-CpG Island | PRKCB        | Body       | 0.031066406  | 1.90E-02  | 0.043158182  | 5.37E-07 | 1.67E-03 | 0.39 | 209 | 1 |

## Continuation of Table S 3.2

|            |                |              |            |              |           |              |          |          |      |     |   |
|------------|----------------|--------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg22256784 | Non-CpG Island |              | Intergenic | 0.053116915  | 3.24E-02  | 0.073791737  | 5.38E-07 | 1.67E-03 | 0.39 | 210 | 1 |
| cg02879554 | Island         | MIR1225      | TSS1500    | -0.079590529 | -1.11E-01 | -0.048596722 | 5.44E-07 | 1.69E-03 | 0.39 | 211 | 2 |
| cg03180011 | Non-CpG Island | CROCCP2      | Body       | 0.029162828  | 1.78E-02  | 0.040521546  | 5.47E-07 | 1.69E-03 | 0.39 | 212 | 1 |
| cg01321673 | Island         | IL17REL      | Body       | -0.029044531 | -4.04E-02 | -0.017721378 | 5.60E-07 | 1.72E-03 | 0.40 | 148 | 2 |
| cg20831648 | Non-CpG Island | ODZ4         | Body       | -0.053996253 | -7.51E-02 | -0.032915752 | 5.81E-07 | 1.78E-03 | 0.42 | 213 | 1 |
| cg07036796 | Non-CpG Island |              | Intergenic | -0.040114898 | -5.58E-02 | -0.024445638 | 5.89E-07 | 1.79E-03 | 0.42 | 214 | 1 |
| cg16617301 | Island         | PNMA88       | First Exon | -0.040859406 | -5.68E-02 | -0.024883969 | 6.03E-07 | 1.83E-03 | 0.44 | 43  | 8 |
| cg22992022 | Non-CpG Island | C5AR2        | TSS1500    | -0.03650228  | -5.08E-02 | -0.022228004 | 6.06E-07 | 1.83E-03 | 0.44 | 215 | 1 |
| cg18849612 | Non-CpG Island | NUAK1        | Body       | -0.023503751 | -3.27E-02 | -0.014296104 | 6.34E-07 | 1.91E-03 | 0.46 | 216 | 1 |
| cg13409442 | Non-CpG Island | STAM         | Body       | -0.078802277 | -1.10E-01 | -0.047924953 | 6.38E-07 | 1.91E-03 | 0.46 | 217 | 1 |
| cg25735583 | Non-CpG Island | LOC101929710 | Body       | 0.043459069  | 2.64E-02  | 0.060494178  | 6.44E-07 | 1.92E-03 | 0.46 | 218 | 1 |
| cg01735357 | Shore          | HAR1B        | Body       | -0.04670142  | -6.50E-02 | -0.028374782 | 6.62E-07 | 1.97E-03 | 0.48 | 219 | 1 |
| cg19750024 | Non-CpG Island | BMP1         | Body       | -0.030420009 | -4.24E-02 | -0.018478734 | 6.68E-07 | 1.98E-03 | 0.48 | 220 | 1 |
| cg13761615 | Non-CpG Island | MCC          | Body       | 0.040072897  | 2.43E-02  | 0.055809234  | 6.74E-07 | 1.99E-03 | 0.49 | 221 | 1 |
| cg21042265 | Non-CpG Island | ZNF827       | Body       | -0.049158843 | -6.85E-02 | -0.029837974 | 6.89E-07 | 2.02E-03 | 0.50 | 222 | 1 |
| cg21599943 | Non-CpG Island | RP11-429O1.1 | TSS1500    | 0.056021663  | 3.40E-02  | 0.078044707  | 6.93E-07 | 2.02E-03 | 0.50 | 223 | 1 |
| cg12546556 | Non-CpG Island | STAB2        | Body       | -0.049883399 | -6.95E-02 | -0.03027338  | 6.93E-07 | 2.02E-03 | 0.50 | 224 | 1 |
| cg09319083 | Non-CpG Island |              | Intergenic | 0.036741313  | 2.23E-02  | 0.051188679  | 6.97E-07 | 2.02E-03 | 0.50 | 225 | 1 |
| cg09766547 | Non-CpG Island | NOD1         | 3' UTR     | -0.037531065 | -5.23E-02 | -0.02276892  | 7.02E-07 | 2.03E-03 | 0.51 | 226 | 1 |
| cg11940368 | Non-CpG Island | HACD3        | Body       | -0.032134416 | -4.48E-02 | -0.019485546 | 7.16E-07 | 2.06E-03 | 0.52 | 227 | 1 |
| cg16166958 | Non-CpG Island | TP63         | Body       | 0.072894043  | 4.42E-02  | 0.101607258  | 7.29E-07 | 2.09E-03 | 0.53 | 228 | 1 |
| cg16794139 | Non-CpG Island | KCNA2        | TSS200     | -0.041619431 | -5.80E-02 | -0.025221052 | 7.34E-07 | 2.09E-03 | 0.53 | 229 | 1 |
| cg25773935 | Non-CpG Island | ELN          | Body       | 0.028204254  | 1.71E-02  | 0.039324577  | 7.46E-07 | 2.12E-03 | 0.54 | 230 | 3 |
| cg04090409 | Non-CpG Island |              | Intergenic | -0.03322743  | -4.63E-02 | -0.020118064 | 7.59E-07 | 2.15E-03 | 0.55 | 231 | 1 |
| cg03189652 | Shore          | PTH2R        | TSS200     | -0.067794271 | -9.45E-02 | -0.041041497 | 7.63E-07 | 2.15E-03 | 0.55 | 232 | 4 |
| cg26750002 | Shore          | IZUMO1       | Body       | -0.044966853 | -6.28E-02 | -0.027178304 | 8.11E-07 | 2.27E-03 | 0.59 | 47  | 3 |
| cg00362690 | Non-CpG Island | MIR548H4     | Body       | -0.041404251 | -5.78E-02 | -0.025022563 | 8.14E-07 | 2.27E-03 | 0.59 | 162 | 2 |
| cg26205859 | Non-CpG Island | PION         | Body       | -0.038154669 | -5.33E-02 | -0.023058682 | 8.14E-07 | 2.27E-03 | 0.59 | 233 | 1 |
| cg15934389 | Non-CpG Island | PPARD        | 5' UTR     | 0.037411198  | 2.26E-02  | 0.052219258  | 8.23E-07 | 2.28E-03 | 0.59 | 234 | 1 |
| cg22869522 | Non-CpG Island | C10orf120    | TSS1500    | -0.046827841 | -6.54E-02 | -0.02828343  | 8.33E-07 | 2.30E-03 | 0.60 | 235 | 2 |
| cg06801305 | Non-CpG Island |              | Intergenic | -0.053779462 | -7.51E-02 | -0.032470707 | 8.44E-07 | 2.33E-03 | 0.61 | 236 | 1 |
| cg13679772 | Non-CpG Island | FOXN3        | 5' UTR     | -0.035716278 | -4.99E-02 | -0.021555556 | 8.58E-07 | 2.35E-03 | 0.62 | 237 | 1 |
| cg06014707 | Non-CpG Island | MIR2052HG    | Body       | 0.087021664  | 5.25E-02  | 0.121527598  | 8.60E-07 | 2.35E-03 | 0.62 | 238 | 1 |
| cg20729813 | Non-CpG Island | SORCS2       | Body       | 0.042535855  | 2.57E-02  | 0.059405228  | 8.64E-07 | 2.35E-03 | 0.62 | 239 | 1 |
| cg03031932 | Non-CpG Island | CMIP         | Body       | -0.027030349 | -3.78E-02 | -0.016302698 | 8.79E-07 | 2.39E-03 | 0.63 | 240 | 1 |
| cg12303649 | Non-CpG Island | DLEU1        | 5' UTR     | 0.047496133  | 2.86E-02  | 0.066376916  | 9.16E-07 | 2.48E-03 | 0.66 | 241 | 1 |
| cg17046263 | Non-CpG Island | PHLPP2       | 3' UTR     | 0.026815688  | 1.61E-02  | 0.037488197  | 9.43E-07 | 2.54E-03 | 0.68 | 242 | 1 |
| cg27656459 | Non-CpG Island |              | Intergenic | -0.049128094 | -6.87E-02 | -0.029571555 | 9.48E-07 | 2.54E-03 | 0.68 | 243 | 1 |
| cg03348540 | Non-CpG Island |              | Intergenic | 0.029622817  | 1.78E-02  | 0.041423642  | 9.65E-07 | 2.58E-03 | 0.70 | 244 | 1 |

## Continuation of Table S 3.2

|            |                |                     |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg19832381 | Shore          | <i>HIST1H2BL</i>    | TSS1500    | -0.049668938 | -6.95E-02 | -0.029880714 | 9.67E-07 | 2.58E-03 | 0.70 | 245 | 1 |
| cg23570968 | Non-CpG Island | <i>TEX15</i>        | 3' UTR     | -0.036162095 | -5.06E-02 | -0.021750517 | 9.75E-07 | 2.59E-03 | 0.70 | 246 | 1 |
| cg21626848 | Shore          | <i>SC65</i>         | TSS1500    | -0.029349425 | -4.11E-02 | -0.017648371 | 9.84E-07 | 2.60E-03 | 0.71 | 183 | 4 |
| cg05461268 | Non-CpG Island | <i>DNER</i>         | Body       | 0.039554077  | 2.38E-02  | 0.055324373  | 9.85E-07 | 2.60E-03 | 0.71 | 247 | 1 |
| cg26299044 | Shore          | <i>KRT12</i>        | Body       | 0.045712493  | 2.75E-02  | 0.063952947  | 1.01E-06 | 2.64E-03 | 0.73 | 248 | 1 |
| cg01297530 | Non-CpG Island | <i>AP000318.2</i>   | 5' UTR     | 0.033114439  | 1.99E-02  | 0.046333721  | 1.02E-06 | 2.65E-03 | 0.73 | 249 | 1 |
| cg10176717 | Non-CpG Island |                     | Intergenic | -0.044690693 | -6.25E-02 | -0.026849497 | 1.02E-06 | 2.65E-03 | 0.73 | 250 | 2 |
| cg04537173 | Non-CpG Island |                     | Intergenic | -0.045980148 | -6.44E-02 | -0.027601043 | 1.05E-06 | 2.72E-03 | 0.76 | 70  | 2 |
| cg16435459 | Non-CpG Island | <i>CRK</i>          | TSS1500    | -0.02050636  | -2.87E-02 | -0.012307791 | 1.05E-06 | 2.73E-03 | 0.76 | 251 | 1 |
| cg00647910 | Non-CpG Island |                     | Intergenic | -0.044383253 | -6.21E-02 | -0.026635313 | 1.06E-06 | 2.73E-03 | 0.76 | 252 | 2 |
| cg06520829 | Non-CpG Island |                     | Intergenic | -0.031604261 | -4.43E-02 | -0.018951923 | 1.09E-06 | 2.80E-03 | 0.79 | 253 | 1 |
| cg13672106 | Non-CpG Island | <i>TMEM99</i>       | 5' UTR     | 0.042489495  | 2.55E-02  | 0.059504043  | 1.10E-06 | 2.81E-03 | 0.79 | 254 | 1 |
| cg16881282 | Non-CpG Island | <i>NALCN-AS1</i>    | Body       | -0.076075293 | -1.07E-01 | -0.04559361  | 1.11E-06 | 2.84E-03 | 0.80 | 255 | 1 |
| cg00894715 | Non-CpG Island | <i>CAT</i>          | Body       | 0.043856733  | 2.63E-02  | 0.061439301  | 1.13E-06 | 2.87E-03 | 0.81 | 256 | 1 |
| cg15718287 | Shelf          | <i>DCLK1</i>        | 5' UTR     | -0.036665757 | -5.14E-02 | -0.021958119 | 1.14E-06 | 2.89E-03 | 0.83 | 257 | 1 |
| cg11801086 | Non-CpG Island | <i>KIF6</i>         | Body       | 0.07769848   | 4.65E-02  | 0.108867954  | 1.15E-06 | 2.89E-03 | 0.83 | 258 | 1 |
| cg20230708 | Non-CpG Island | <i>RP3-416H24.4</i> | 5' UTR     | -0.023137417 | -3.24E-02 | -0.013852505 | 1.15E-06 | 2.90E-03 | 0.83 | 259 | 1 |
| cg03210579 | Non-CpG Island |                     | Intergenic | -0.035409607 | -4.96E-02 | -0.021194696 | 1.17E-06 | 2.92E-03 | 0.84 | 260 | 1 |
| cg12101380 | Non-CpG Island | <i>C10orf71</i>     | 5' UTR     | -0.035926482 | -5.04E-02 | -0.021501617 | 1.17E-06 | 2.92E-03 | 0.84 | 261 | 1 |
| cg26282283 | Shelf          | <i>TTYH3</i>        | 3' UTR     | -0.040107203 | -5.62E-02 | -0.024002699 | 1.17E-06 | 2.92E-03 | 0.85 | 262 | 1 |
| cg21519787 | Island         | <i>SPHK2</i>        | 3' UTR     | -0.080839346 | -1.13E-01 | -0.048355245 | 1.19E-06 | 2.96E-03 | 0.86 | 263 | 2 |
| cg23020497 | Non-CpG Island | <i>AL034449.1</i>   | TSS200     | 0.043481283  | 2.60E-02  | 0.060962016  | 1.21E-06 | 2.98E-03 | 0.87 | 264 | 1 |
| cg04894218 | Non-CpG Island | <i>MCC</i>          | Body       | 0.054651623  | 3.27E-02  | 0.076627671  | 1.21E-06 | 2.99E-03 | 0.88 | 265 | 2 |
| cg03452365 | Non-CpG Island | <i>KANSL1</i>       | Body       | -0.033662128 | -4.72E-02 | -0.020124962 | 1.22E-06 | 2.99E-03 | 0.88 | 266 | 1 |
| cg14710896 | Non-CpG Island | <i>KIAA1217</i>     | Body       | 0.050466392  | 3.02E-02  | 0.070775886  | 1.24E-06 | 3.03E-03 | 0.89 | 267 | 1 |
| cg09452569 | Non-CpG Island |                     | Intergenic | 0.029783133  | 1.78E-02  | 0.041770456  | 1.24E-06 | 3.03E-03 | 0.90 | 268 | 1 |
| cg19020666 | Shelf          |                     | Intergenic | -0.042244009 | -5.93E-02 | -0.02523211  | 1.26E-06 | 3.06E-03 | 0.91 | 269 | 1 |
| cg16417374 | Shore          | <i>MAN1C1</i>       | Body       | -0.03830345  | -5.37E-02 | -0.022874878 | 1.26E-06 | 3.06E-03 | 0.91 | 270 | 1 |
| cg00997853 | Non-CpG Island | <i>ITPR2</i>        | Body       | -0.036312965 | -5.09E-02 | -0.021681506 | 1.27E-06 | 3.08E-03 | 0.92 | 271 | 1 |
| cg08877283 | Non-CpG Island | <i>GRIK4</i>        | Body       | -0.051726913 | -7.26E-02 | -0.030863742 | 1.31E-06 | 3.14E-03 | 0.94 | 272 | 1 |
| cg26442250 | Non-CpG Island |                     | Intergenic | 0.032662074  | 1.95E-02  | 0.045846255  | 1.33E-06 | 3.19E-03 | 0.96 | 273 | 1 |
| cg22005304 | Non-CpG Island | <i>ARHGEF10L</i>    | Body       | -0.037472231 | -5.26E-02 | -0.022341829 | 1.34E-06 | 3.20E-03 | 0.97 | 274 | 1 |
| cg00189117 | Island         | <i>HNRNP1L</i>      | TSS200     | 0.062117024  | 3.70E-02  | 0.087211546  | 1.36E-06 | 3.23E-03 | 0.98 | 275 | 1 |
| cg04249817 | Shore          | <i>FUOM</i>         | TSS1500    | 0.027653115  | 1.65E-02  | 0.03882656   | 1.36E-06 | 3.24E-03 | 0.98 | 15  | 4 |
| cg02637282 | Non-CpG Island |                     | Intergenic | -0.058025459 | -8.15E-02 | -0.034569059 | 1.38E-06 | 3.26E-03 | 0.99 | 276 | 1 |
| cg10223198 | Non-CpG Island | <i>AP000797.3</i>   | 5' UTR     | 0.030951855  | 1.84E-02  | 0.043471275  | 1.40E-06 | 3.30E-03 | 1.00 | 277 | 1 |
| cg09502182 | Non-CpG Island |                     | Intergenic | -0.03451332  | -4.85E-02 | -0.02055023  | 1.41E-06 | 3.30E-03 | 1.00 | 278 | 1 |
| cg22338587 | Non-CpG Island | <i>LOC101927070</i> | Body       | -0.037496927 | -5.27E-02 | -0.022324482 | 1.41E-06 | 3.31E-03 | 1.00 | 279 | 2 |
| cg13626246 | Non-CpG Island | <i>NTRK2</i>        | Body       | 0.04871091   | 2.90E-02  | 0.068431375  | 1.43E-06 | 3.34E-03 | 1.00 | 280 | 1 |

## Continuation of Table S 3.2

|            |                |                      |               |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|---------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg20114054 | Shelf          |                      | Intergenic    | -0.057941061 | -8.14E-02 | -0.034479323 | 1.43E-06 | 3.34E-03 | 1.00 | 281 | 1 |
| cg11848931 | Non-CpG Island | <i>PRTFDC1</i>       | Body          | -0.044264374 | -6.22E-02 | -0.02633645  | 1.44E-06 | 3.35E-03 | 1.00 | 282 | 1 |
| cg02803228 | Non-CpG Island |                      | Intergenic    | -0.058890115 | -8.28E-02 | -0.035025478 | 1.46E-06 | 3.38E-03 | 1.00 | 283 | 1 |
| cg07236324 | Non-CpG Island | <i>CUEDC1</i>        | 5' UTR        | -0.019717541 | -2.77E-02 | -0.011720064 | 1.49E-06 | 3.44E-03 | 1.00 | 284 | 1 |
| cg01230907 | Non-CpG Island | <i>TPO</i>           | Body          | -0.026709054 | -3.75E-02 | -0.015868578 | 1.52E-06 | 3.48E-03 | 1.00 | 285 | 1 |
| cg02348119 | Non-CpG Island | <i>TBC1D16</i>       | 5' UTR        | -0.033614234 | -4.73E-02 | -0.019971001 | 1.52E-06 | 3.48E-03 | 1.00 | 286 | 1 |
| cg08326132 | Non-CpG Island |                      | Intergenic    | -0.0537646   | -7.56E-02 | -0.031939183 | 1.52E-06 | 3.48E-03 | 1.00 | 287 | 1 |
| cg21122437 | Shore          | <i>TEAD2</i>         | 5' UTR        | -0.03760785  | -5.29E-02 | -0.022337919 | 1.53E-06 | 3.49E-03 | 1.00 | 288 | 2 |
| cg07207669 | Shore          | <i>EFNA1</i>         | Body          | -0.027970198 | -3.93E-02 | -0.016611607 | 1.54E-06 | 3.49E-03 | 1.00 | 289 | 1 |
| cg13631484 | Non-CpG Island |                      | Intergenic    | -0.05125319  | -7.21E-02 | -0.030418922 | 1.57E-06 | 3.56E-03 | 1.00 | 290 | 1 |
| cg02267536 | Island         | <i>BANP</i>          | Body          | 0.030174177  | 1.79E-02  | 0.042441381  | 1.58E-06 | 3.56E-03 | 1.00 | 291 | 1 |
| cg09831562 | Non-CpG Island | <i>SOX2OT</i>        | TSS1500       | 0.063822522  | 3.79E-02  | 0.08977671   | 1.59E-06 | 3.57E-03 | 1.00 | 56  | 3 |
| cg26210250 | Non-CpG Island | <i>CIB4</i>          | Body          | -0.029863746 | -4.20E-02 | -0.01771522  | 1.60E-06 | 3.58E-03 | 1.00 | 292 | 1 |
| cg15330438 | Non-CpG Island | <i>RNF103</i>        | Body          | -0.030106546 | -4.24E-02 | -0.017858794 | 1.60E-06 | 3.58E-03 | 1.00 | 10  | 2 |
| cg10255761 | Shore          | <i>KLHDC8B</i>       | 5' UTR        | -0.050039224 | -7.04E-02 | -0.029665062 | 1.64E-06 | 3.64E-03 | 1.00 | 293 | 1 |
| cg23923308 | Non-CpG Island | <i>ELMO1</i>         | Body          | -0.050683283 | -7.13E-02 | -0.030044631 | 1.64E-06 | 3.64E-03 | 1.00 | 294 | 1 |
| cg20936142 | Shelf          | <i>AGAP1</i>         | Body          | 0.040307759  | 2.39E-02  | 0.056723086  | 1.64E-06 | 3.64E-03 | 1.00 | 295 | 1 |
| cg02856716 | Non-CpG Island | <i>PAX7</i>          | Body          | -0.034615634 | -4.87E-02 | -0.020510943 | 1.67E-06 | 3.68E-03 | 1.00 | 296 | 1 |
| cg08476354 | Non-CpG Island |                      | Intergenic    | 0.04086148   | 2.42E-02  | 0.057517035  | 1.68E-06 | 3.70E-03 | 1.00 | 297 | 1 |
| cg06891056 | Non-CpG Island | <i>RP5-884C9.2</i>   | 5' UTR        | -0.036226426 | -5.10E-02 | -0.021455751 | 1.69E-06 | 3.70E-03 | 1.00 | 298 | 1 |
| cg09558509 | Island         | <i>ANKRD13B</i>      | Exon Boundary | -0.060118283 | -8.46E-02 | -0.035605699 | 1.69E-06 | 3.70E-03 | 1.00 | 299 | 2 |
| cg19066876 | Non-CpG Island |                      | Intergenic    | 0.03360988   | 1.99E-02  | 0.047315803  | 1.70E-06 | 3.70E-03 | 1.00 | 300 | 1 |
| cg17873222 | Non-CpG Island | <i>ODZ3</i>          | 5' UTR        | 0.051889796  | 3.07E-02  | 0.073054889  | 1.71E-06 | 3.70E-03 | 1.00 | 301 | 1 |
| cg14660461 | Non-CpG Island | <i>SLC10A7</i>       | Body          | -0.036409177 | -5.13E-02 | -0.021558208 | 1.71E-06 | 3.70E-03 | 1.00 | 302 | 1 |
| cg13210763 | Non-CpG Island | <i>RP11-319E12.1</i> | 5' UTR        | -0.053444328 | -7.52E-02 | -0.031639291 | 1.72E-06 | 3.70E-03 | 1.00 | 303 | 1 |
| cg25762706 | Non-CpG Island | <i>STMN4</i>         | First Exon    | -0.043790298 | -6.17E-02 | -0.025923803 | 1.72E-06 | 3.70E-03 | 1.00 | 60  | 2 |
| cg02017417 | Shore          | <i>IBA57</i>         | TSS200        | -0.094068241 | -1.32E-01 | -0.055671019 | 1.74E-06 | 3.73E-03 | 1.00 | 304 | 1 |
| cg15483907 | Non-CpG Island | <i>ANGPT1</i>        | TSS1500       | 0.042517747  | 2.51E-02  | 0.059893777  | 1.79E-06 | 3.83E-03 | 1.00 | 305 | 1 |
| cg20533553 | Shore          | <i>KCNQ1</i>         | TSS1500       | -0.037746743 | -5.32E-02 | -0.022315287 | 1.80E-06 | 3.84E-03 | 1.00 | 306 | 1 |
| cg01973394 | Island         | <i>MEX3B</i>         | Body          | 0.028195797  | 1.67E-02  | 0.03972473   | 1.81E-06 | 3.85E-03 | 1.00 | 307 | 1 |
| cg17043444 | Non-CpG Island |                      | Intergenic    | 0.072114453  | 4.26E-02  | 0.101622611  | 1.84E-06 | 3.90E-03 | 1.00 | 134 | 3 |
| cg25293430 | Non-CpG Island | <i>ANK2</i>          | 5' UTR        | 0.029777303  | 1.76E-02  | 0.04196548   | 1.85E-06 | 3.92E-03 | 1.00 | 308 | 1 |
| cg25784136 | Shelf          | <i>CCDC116</i>       | First Exon    | -0.080597972 | -1.14E-01 | -0.047599481 | 1.86E-06 | 3.93E-03 | 1.00 | 154 | 2 |
| cg15253752 | Non-CpG Island |                      | Intergenic    | -0.032780101 | -4.62E-02 | -0.019351413 | 1.89E-06 | 3.98E-03 | 1.00 | 309 | 1 |
| cg15035117 | Non-CpG Island |                      | Intergenic    | -0.037359784 | -5.27E-02 | -0.022051634 | 1.90E-06 | 3.98E-03 | 1.00 | 310 | 1 |
| cg07346426 | Non-CpG Island | <i>ARL6IP6</i>       | Body          | -0.032828118 | -4.63E-02 | -0.019373386 | 1.91E-06 | 4.00E-03 | 1.00 | 311 | 1 |
| cg16411103 | Non-CpG Island | <i>STARD13</i>       | 5' UTR        | -0.09063017  | -1.28E-01 | -0.05347374  | 1.92E-06 | 4.00E-03 | 1.00 | 312 | 1 |
| cg15322932 | Non-CpG Island | <i>ALDH3B1</i>       | TSS200        | -0.037519414 | -5.29E-02 | -0.022136079 | 1.93E-06 | 4.00E-03 | 1.00 | 313 | 1 |
| cg19246944 | Non-CpG Island |                      | Intergenic    | -0.039826924 | -5.62E-02 | -0.023496699 | 1.93E-06 | 4.00E-03 | 1.00 | 314 | 1 |

## Continuation of Table S 3.2

|            |                |               |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg00404400 | Shore          | DGKG          | 5' UTR     | -0.047269318 | -6.67E-02 | -0.027868642 | 1.97E-06 | 4.07E-03 | 1.00 | 315 | 2 |
| cg26723174 | Non-CpG Island |               | Intergenic | 0.031615517  | 1.86E-02  | 0.044591982  | 1.98E-06 | 4.07E-03 | 1.00 | 316 | 1 |
| cg26572392 | Non-CpG Island | KIAA1217      | 5' UTR     | 0.03824017   | 2.25E-02  | 0.053936252  | 1.98E-06 | 4.07E-03 | 1.00 | 317 | 1 |
| cg22290528 | Shore          | HEXA          | TSS1500    | 0.025438407  | 1.50E-02  | 0.035887556  | 2.01E-06 | 4.13E-03 | 1.00 | 318 | 1 |
| cg01933515 | Non-CpG Island | C14orf132     | Body       | 0.04999156   | 2.94E-02  | 0.070534526  | 2.03E-06 | 4.15E-03 | 1.00 | 319 | 1 |
| cg27162196 | Shore          | ZFP1          | 5' UTR     | 0.052872187  | 3.11E-02  | 0.074602     | 2.04E-06 | 4.15E-03 | 1.00 | 320 | 1 |
| cg05324516 | Non-CpG Island |               | Intergenic | -0.043238467 | -6.10E-02 | -0.025445917 | 2.10E-06 | 4.26E-03 | 1.00 | 321 | 1 |
| cg23465295 | Non-CpG Island | RP11-774D14.1 | TSS200     | -0.031739632 | -4.48E-02 | -0.01867713  | 2.10E-06 | 4.26E-03 | 1.00 | 322 | 1 |
| cg08792629 | Non-CpG Island | LOC100134368  | Body       | 0.035543462  | 2.09E-02  | 0.050172669  | 2.11E-06 | 4.26E-03 | 1.00 | 323 | 1 |
| cg00939525 | Non-CpG Island | GPR176        | Body       | 0.046418258  | 2.73E-02  | 0.065526083  | 2.11E-06 | 4.26E-03 | 1.00 | 324 | 1 |
| cg10680196 | Non-CpG Island | ACTN4         | Body       | -0.032758338 | -4.63E-02 | -0.019261609 | 2.16E-06 | 4.34E-03 | 1.00 | 325 | 1 |
| cg04422903 | Shore          | CCDC88B       | Body       | -0.039290902 | -5.55E-02 | -0.023097532 | 2.17E-06 | 4.36E-03 | 1.00 | 113 | 2 |
| cg06287505 | Shore          |               | Intergenic | -0.046902832 | -6.62E-02 | -0.027567027 | 2.19E-06 | 4.37E-03 | 1.00 | 326 | 1 |
| cg19004883 | Shelf          |               | Intergenic | -0.02702757  | -3.82E-02 | -0.015881429 | 2.21E-06 | 4.39E-03 | 1.00 | 327 | 1 |
| cg21771140 | Non-CpG Island |               | Intergenic | 0.080462464  | 4.73E-02  | 0.113645124  | 2.21E-06 | 4.39E-03 | 1.00 | 328 | 1 |
| cg20077275 | Non-CpG Island | CHST4         | TSS1500    | -0.042723503 | -6.03E-02 | -0.025097912 | 2.22E-06 | 4.41E-03 | 1.00 | 329 | 1 |
| cg16668176 | Non-CpG Island | PTPRS         | 5' UTR     | -0.028315837 | -4.00E-02 | -0.016630918 | 2.24E-06 | 4.42E-03 | 1.00 | 330 | 1 |
| cg20798375 | Non-CpG Island |               | Intergenic | 0.034220362  | 2.01E-02  | 0.048342862  | 2.24E-06 | 4.42E-03 | 1.00 | 331 | 1 |
| cg14799766 | Non-CpG Island | KCNB1         | Body       | 0.050553555  | 2.97E-02  | 0.071446731  | 2.32E-06 | 4.56E-03 | 1.00 | 332 | 1 |
| cg26153532 | Non-CpG Island |               | Intergenic | -0.037020623 | -5.23E-02 | -0.02170036  | 2.39E-06 | 4.67E-03 | 1.00 | 333 | 1 |
| cg01772439 | Non-CpG Island | FLJ41603      | 3' UTR     | -0.024747164 | -3.50E-02 | -0.014505179 | 2.39E-06 | 4.67E-03 | 1.00 | 334 | 1 |
| cg01695994 | Shelf          |               | Intergenic | -0.031624209 | -4.47E-02 | -0.018535686 | 2.40E-06 | 4.67E-03 | 1.00 | 335 | 1 |
| cg13165009 | Shore          | C17orf102     | Body       | -0.042336632 | -5.99E-02 | -0.024807887 | 2.42E-06 | 4.70E-03 | 1.00 | 336 | 1 |
| cg19911880 | Non-CpG Island | ASAP1         | Body       | -0.040749401 | -5.76E-02 | -0.023868093 | 2.45E-06 | 4.75E-03 | 1.00 | 337 | 1 |
| cg15705813 | Non-CpG Island | PCBP1-AS1     | 3' UTR     | 0.025338529  | 1.48E-02  | 0.035839241  | 2.47E-06 | 4.77E-03 | 1.00 | 338 | 1 |
| cg22953687 | Shelf          | ZFYVE28       | Body       | -0.050691781 | -7.17E-02 | -0.029683572 | 2.47E-06 | 4.77E-03 | 1.00 | 339 | 1 |
| cg02525267 | Non-CpG Island |               | Intergenic | -0.04325313  | -6.12E-02 | -0.025325608 | 2.48E-06 | 4.77E-03 | 1.00 | 340 | 1 |
| cg08170023 | Shore          | RASGEF1A      | Body       | -0.066536544 | -9.41E-02 | -0.038945056 | 2.50E-06 | 4.80E-03 | 1.00 | 341 | 2 |
| cg11463292 | Non-CpG Island | SULF1         | 5' UTR     | 0.046969304  | 2.75E-02  | 0.06644789   | 2.51E-06 | 4.80E-03 | 1.00 | 342 | 1 |
| cg03348845 | Non-CpG Island |               | Intergenic | -0.019617524 | -2.78E-02 | -0.011477738 | 2.54E-06 | 4.85E-03 | 1.00 | 343 | 1 |
| cg15706604 | Non-CpG Island | CCDC33        | Body       | -0.033639157 | -4.76E-02 | -0.019666116 | 2.60E-06 | 4.96E-03 | 1.00 | 344 | 1 |
| cg00724936 | Non-CpG Island | MARK4         | Body       | -0.06071474  | -8.59E-02 | -0.035481502 | 2.63E-06 | 5.00E-03 | 1.00 | 345 | 1 |
| cg00155731 | Non-CpG Island |               | Intergenic | -0.043573691 | -6.17E-02 | -0.025460415 | 2.65E-06 | 5.02E-03 | 1.00 | 346 | 1 |
| cg17205574 | Non-CpG Island | MEGF11        | 5' UTR     | -0.031383119 | -4.44E-02 | -0.018334134 | 2.66E-06 | 5.03E-03 | 1.00 | 347 | 1 |
| cg18422750 | Non-CpG Island | PDGFC         | Body       | -0.03951438  | -5.59E-02 | -0.023079391 | 2.68E-06 | 5.05E-03 | 1.00 | 348 | 1 |
| cg17603132 | Shore          | RIC3          | Body       | 0.039649308  | 2.31E-02  | 0.056153253  | 2.73E-06 | 5.12E-03 | 1.00 | 349 | 1 |
| cg09298313 | Non-CpG Island |               | Intergenic | -0.023373644 | -3.31E-02 | -0.01364392  | 2.73E-06 | 5.12E-03 | 1.00 | 350 | 2 |
| cg16298481 | Shore          | LRRC8B        | TSS1500    | -0.027090414 | -3.84E-02 | -0.015809924 | 2.75E-06 | 5.14E-03 | 1.00 | 351 | 1 |
| cg19433225 | Shore          | GIPC2         | Body       | -0.059306373 | -8.40E-02 | -0.03460703  | 2.76E-06 | 5.14E-03 | 1.00 | 352 | 1 |

## Continuation of Table S 3.2

|            |                |                     |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg12351310 | Non-CpG Island | <i>NDUFS4</i>       | Body       | -0.033745383 | -4.78E-02 | -0.019688459 | 2.78E-06 | 5.14E-03 | 1.00 | 353 | 1 |
| cg08064860 | Non-CpG Island | <i>STRN3</i>        | Body       | 0.043923063  | 2.56E-02  | 0.062220507  | 2.78E-06 | 5.14E-03 | 1.00 | 354 | 1 |
| cg01121022 | Non-CpG Island | <i>CTD-2134A5.4</i> | 3' UTR     | -0.026278571 | -3.72E-02 | -0.015330629 | 2.78E-06 | 5.14E-03 | 1.00 | 355 | 1 |
| cg03864443 | Non-CpG Island | <i>KCND2</i>        | Body       | 0.056691742  | 3.31E-02  | 0.080310938  | 2.79E-06 | 5.14E-03 | 1.00 | 356 | 1 |
| cg03566527 | Shelf          | <i>GLI2</i>         | 3' UTR     | 0.035501392  | 2.07E-02  | 0.050295005  | 2.80E-06 | 5.15E-03 | 1.00 | 357 | 1 |
| cg03172796 | Non-CpG Island |                     | Intergenic | -0.03817741  | -5.41E-02 | -0.022260756 | 2.83E-06 | 5.19E-03 | 1.00 | 358 | 1 |
| cg12043561 | Non-CpG Island | <i>CREB3L3</i>      | TSS1500    | -0.042930311 | -6.08E-02 | -0.025030533 | 2.84E-06 | 5.19E-03 | 1.00 | 359 | 1 |
| cg22517730 | Non-CpG Island | <i>LOC101927123</i> | Body       | -0.053230032 | -7.54E-02 | -0.031014328 | 2.90E-06 | 5.30E-03 | 1.00 | 360 | 1 |
| cg27565185 | Non-CpG Island | <i>CP</i>           | Body       | -0.020610338 | -2.92E-02 | -0.012004597 | 2.93E-06 | 5.34E-03 | 1.00 | 361 | 1 |
| cg05629849 | Shore          | <i>STBD1</i>        | TSS200     | 0.039433837  | 2.30E-02  | 0.055911081  | 2.98E-06 | 5.41E-03 | 1.00 | 362 | 1 |
| cg08428972 | Non-CpG Island |                     | Intergenic | -0.037719415 | -5.35E-02 | -0.021953448 | 3.00E-06 | 5.44E-03 | 1.00 | 102 | 2 |
| cg21810778 | Non-CpG Island |                     | Intergenic | -0.018075105 | -2.56E-02 | -0.010515423 | 3.04E-06 | 5.49E-03 | 1.00 | 363 | 1 |
| cg09525260 | Non-CpG Island | <i>ADAMTSL3</i>     | Body       | 0.03223273   | 1.88E-02  | 0.045714024  | 3.04E-06 | 5.49E-03 | 1.00 | 364 | 1 |
| cg25509871 | Shore          | <i>PLD3</i>         | 5' UTR     | -0.049409167 | -7.01E-02 | -0.02873893  | 3.06E-06 | 5.50E-03 | 1.00 | 365 | 2 |
| cg11551236 | Island         | <i>PAX5</i>         | First Exon | 0.031665571  | 1.84E-02  | 0.044918357  | 3.09E-06 | 5.54E-03 | 1.00 | 366 | 1 |
| cg02510044 | Shore          | <i>KDELR1</i>       | TSS1500    | -0.025106308 | -3.56E-02 | -0.014598079 | 3.09E-06 | 5.54E-03 | 1.00 | 367 | 2 |
| cg03221025 | Shore          | <i>CBFA2T3</i>      | Body       | -0.029414513 | -4.17E-02 | -0.017101265 | 3.10E-06 | 5.54E-03 | 1.00 | 368 | 1 |
| cg13978777 | Shelf          |                     | Intergenic | -0.033628886 | -4.77E-02 | -0.019548206 | 3.12E-06 | 5.55E-03 | 1.00 | 369 | 2 |
| cg15827779 | Shore          | <i>CDC42BPB</i>     | TSS1500    | -0.035330627 | -5.01E-02 | -0.02053665  | 3.12E-06 | 5.55E-03 | 1.00 | 370 | 1 |
| cg12307174 | Non-CpG Island | <i>RAD51B</i>       | Body       | -0.056308159 | -7.99E-02 | -0.03272348  | 3.14E-06 | 5.56E-03 | 1.00 | 371 | 1 |
| cg19451566 | Non-CpG Island | <i>ATP2B2</i>       | Body       | -0.040919295 | -5.81E-02 | -0.023780137 | 3.14E-06 | 5.56E-03 | 1.00 | 372 | 1 |
| cg03018753 | Non-CpG Island |                     | Intergenic | 0.049584136  | 2.88E-02  | 0.070363018  | 3.18E-06 | 5.61E-03 | 1.00 | 373 | 1 |
| cg06228645 | Non-CpG Island | <i>KIAA1217</i>     | Body       | 0.056539879  | 3.28E-02  | 0.080263538  | 3.27E-06 | 5.74E-03 | 1.00 | 374 | 1 |
| cg06804344 | Non-CpG Island | <i>GP9</i>          | TSS1500    | -0.034271779 | -4.87E-02 | -0.019891593 | 3.27E-06 | 5.74E-03 | 1.00 | 375 | 1 |
| cg25839606 | Shelf          |                     | Intergenic | -0.039916322 | -5.67E-02 | -0.023149555 | 3.35E-06 | 5.86E-03 | 1.00 | 376 | 1 |
| cg15241106 | Non-CpG Island | <i>EWSAT1</i>       | Body       | -0.048550553 | -6.89E-02 | -0.028155424 | 3.35E-06 | 5.86E-03 | 1.00 | 377 | 1 |
| cg13905211 | Non-CpG Island |                     | Intergenic | -0.022525677 | -3.20E-02 | -0.013062334 | 3.36E-06 | 5.86E-03 | 1.00 | 378 | 1 |
| cg05483509 | Island         | <i>ATP6V0C</i>      | TSS1500    | -0.035838266 | -5.09E-02 | -0.020780477 | 3.37E-06 | 5.86E-03 | 1.00 | 379 | 1 |
| cg17001328 | Shore          |                     | Intergenic | -0.031625044 | -4.49E-02 | -0.018323272 | 3.45E-06 | 5.99E-03 | 1.00 | 369 | 2 |
| cg06901200 | Non-CpG Island | <i>WWC2</i>         | Body       | -0.022973313 | -3.26E-02 | -0.013306387 | 3.48E-06 | 6.03E-03 | 1.00 | 380 | 1 |
| cg16231873 | Non-CpG Island | <i>PSAPL1</i>       | 3' UTR     | -0.046695674 | -6.63E-02 | -0.027044192 | 3.49E-06 | 6.03E-03 | 1.00 | 381 | 1 |
| cg08620095 | Non-CpG Island | <i>KLK11</i>        | 5' UTR     | -0.031995424 | -4.55E-02 | -0.018528435 | 3.50E-06 | 6.03E-03 | 1.00 | 382 | 1 |
| cg19005955 | Non-CpG Island | <i>FSTL4</i>        | Body       | -0.046207939 | -6.57E-02 | -0.026757935 | 3.51E-06 | 6.03E-03 | 1.00 | 383 | 1 |
| cg03550951 | Non-CpG Island | <i>RGS7</i>         | Body       | 0.052877186  | 3.06E-02  | 0.075138535  | 3.52E-06 | 6.04E-03 | 1.00 | 384 | 1 |
| cg00753112 | Non-CpG Island | <i>KHDRBS3</i>      | Body       | -0.025624715 | -3.64E-02 | -0.014832354 | 3.55E-06 | 6.08E-03 | 1.00 | 385 | 1 |
| cg21798630 | Non-CpG Island | <i>IL1RAP</i>       | Body       | 0.038983446  | 2.26E-02  | 0.055405403  | 3.57E-06 | 6.09E-03 | 1.00 | 386 | 1 |
| cg13850368 | Non-CpG Island | <i>ELN</i>          | Body       | 0.0297843    | 1.72E-02  | 0.042337766  | 3.61E-06 | 6.15E-03 | 1.00 | 230 | 3 |
| cg07667666 | Non-CpG Island | <i>PLD1</i>         | Body       | 0.036530733  | 2.11E-02  | 0.051931353  | 3.63E-06 | 6.17E-03 | 1.00 | 387 | 1 |
| cg00275479 | Shelf          |                     | Intergenic | -0.057529372 | -8.18E-02 | -0.033255084 | 3.70E-06 | 6.27E-03 | 1.00 | 388 | 2 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg15254759 | Non-CpG Island |                      | Intergenic | 0.036421313  | 2.10E-02  | 0.051800547  | 3.76E-06 | 6.36E-03 | 1.00 | 389 | 1 |
| cg27001818 | Non-CpG Island | <i>LARGE</i>         | Body       | 0.033729865  | 1.95E-02  | 0.047975907  | 3.78E-06 | 6.38E-03 | 1.00 | 390 | 1 |
| cg07023764 | Non-CpG Island | <i>CCDC26</i>        | Body       | -0.033486344 | -4.76E-02 | -0.019340031 | 3.80E-06 | 6.38E-03 | 1.00 | 391 | 1 |
| cg18548879 | Shore          | <i>MGA</i>           | 5' UTR     | -0.04161618  | -5.92E-02 | -0.024035386 | 3.80E-06 | 6.38E-03 | 1.00 | 392 | 1 |
| cg22260782 | Non-CpG Island | <i>DTHD1</i>         | TSS1500    | -0.066171342 | -9.41E-02 | -0.038206906 | 3.83E-06 | 6.41E-03 | 1.00 | 393 | 1 |
| cg27335600 | Island         | <i>CACNA1D</i>       | TSS200     | 0.027689603  | 1.60E-02  | 0.039391961  | 3.84E-06 | 6.41E-03 | 1.00 | 394 | 1 |
| cg01233571 | Non-CpG Island | <i>DEPDC1B</i>       | Body       | -0.043283375 | -6.16E-02 | -0.024986646 | 3.85E-06 | 6.43E-03 | 1.00 | 395 | 1 |
| cg19997281 | Non-CpG Island | <i>MKNK1</i>         | 5' UTR     | 0.039402259  | 2.27E-02  | 0.056062259  | 3.87E-06 | 6.44E-03 | 1.00 | 396 | 1 |
| cg09470638 | Shore          | <i>CCDC105</i>       | TSS200     | -0.036319424 | -5.17E-02 | -0.02096038  | 3.89E-06 | 6.45E-03 | 1.00 | 57  | 4 |
| cg26247064 | Non-CpG Island | <i>RNF169</i>        | Body       | -0.033849421 | -4.82E-02 | -0.019532414 | 3.90E-06 | 6.46E-03 | 1.00 | 397 | 1 |
| cg06130787 | Shore          | <i>KLK10</i>         | TSS1500    | -0.034181673 | -4.87E-02 | -0.019710865 | 3.98E-06 | 6.58E-03 | 1.00 | 398 | 1 |
| cg23958942 | Shore          | <i>RAP1GAP</i>       | Body       | -0.039859783 | -5.68E-02 | -0.022960715 | 4.11E-06 | 6.75E-03 | 1.00 | 399 | 1 |
| cg04246167 | Non-CpG Island | <i>C3orf67</i>       | 5' UTR     | -0.023889063 | -3.40E-02 | -0.013759156 | 4.13E-06 | 6.75E-03 | 1.00 | 400 | 1 |
| cg18228693 | Shore          |                      | Intergenic | -0.039977412 | -5.69E-02 | -0.023025391 | 4.13E-06 | 6.75E-03 | 1.00 | 401 | 1 |
| cg11990561 | Non-CpG Island | <i>EEA1</i>          | Body       | -0.037655622 | -5.36E-02 | -0.021687701 | 4.13E-06 | 6.75E-03 | 1.00 | 402 | 1 |
| cg07561162 | Non-CpG Island | <i>GRIN2A</i>        | Body       | -0.035029841 | -4.99E-02 | -0.020174613 | 4.14E-06 | 6.75E-03 | 1.00 | 403 | 1 |
| cg24661860 | Non-CpG Island | <i>CASZ1</i>         | 5' UTR     | -0.045625484 | -6.50E-02 | -0.02626993  | 4.17E-06 | 6.79E-03 | 1.00 | 404 | 1 |
| cg13569146 | Non-CpG Island |                      | Intergenic | -0.052575489 | -7.49E-02 | -0.030268003 | 4.18E-06 | 6.80E-03 | 1.00 | 405 | 2 |
| cg11947712 | Shore          | <i>C1orf59</i>       | 5' UTR     | -0.054235347 | -7.72E-02 | -0.0312213   | 4.19E-06 | 6.80E-03 | 1.00 | 406 | 1 |
| cg17501210 | Non-CpG Island | <i>RPS6KA2</i>       | Body       | -0.046753099 | -6.66E-02 | -0.026911248 | 4.21E-06 | 6.81E-03 | 1.00 | 407 | 1 |
| cg26893502 | Non-CpG Island | <i>RAB3C</i>         | Body       | 0.045670558  | 2.63E-02  | 0.065071435  | 4.29E-06 | 6.93E-03 | 1.00 | 408 | 1 |
| cg09829382 | Island         | <i>BRUNOL5</i>       | Body       | -0.032562635 | -4.64E-02 | -0.018728498 | 4.30E-06 | 6.93E-03 | 1.00 | 409 | 1 |
| cg03383434 | Island         | <i>EXD3</i>          | TSS1500    | -0.035674581 | -5.08E-02 | -0.020516918 | 4.31E-06 | 6.93E-03 | 1.00 | 410 | 1 |
| cg07476379 | Non-CpG Island | <i>SOX2-OT</i>       | Body       | 0.044795773  | 2.58E-02  | 0.063830866  | 4.32E-06 | 6.93E-03 | 1.00 | 411 | 1 |
| cg16313343 | Shore          | <i>BRF1</i>          | TSS1500    | -0.046988288 | -6.70E-02 | -0.027017479 | 4.34E-06 | 6.94E-03 | 1.00 | 412 | 1 |
| cg10760040 | Non-CpG Island |                      | Intergenic | -0.036947953 | -5.27E-02 | -0.021243697 | 4.35E-06 | 6.94E-03 | 1.00 | 413 | 1 |
| cg15950727 | Non-CpG Island | <i>RP11-114G22.1</i> | TSS200     | -0.05552744  | -7.91E-02 | -0.031912831 | 4.40E-06 | 7.00E-03 | 1.00 | 414 | 1 |
| cg23070738 | Non-CpG Island | <i>RP11-73O6.4</i>   | 5' UTR     | -0.032988901 | -4.70E-02 | -0.018958488 | 4.41E-06 | 7.00E-03 | 1.00 | 415 | 1 |
| cg03221779 | Non-CpG Island | <i>SETD3</i>         | Body       | 0.03599959   | 2.07E-02  | 0.051311597  | 4.41E-06 | 7.00E-03 | 1.00 | 416 | 1 |
| cg20681708 | Shore          | <i>EML2-AS1</i>      | TSS1500    | -0.035921333 | -5.12E-02 | -0.020641547 | 4.42E-06 | 7.00E-03 | 1.00 | 417 | 2 |
| cg03320078 | Non-CpG Island |                      | Intergenic | -0.039021113 | -5.56E-02 | -0.022415172 | 4.46E-06 | 7.05E-03 | 1.00 | 418 | 1 |
| cg00382138 | Non-CpG Island | <i>CFI</i>           | First Exon | -0.032205802 | -4.59E-02 | -0.018497805 | 4.48E-06 | 7.06E-03 | 1.00 | 419 | 1 |
| cg26994484 | Shore          |                      | Intergenic | -0.047601776 | -6.79E-02 | -0.027326207 | 4.55E-06 | 7.16E-03 | 1.00 | 420 | 1 |
| cg17885144 | Non-CpG Island |                      | Intergenic | -0.046594099 | -6.64E-02 | -0.02673966  | 4.59E-06 | 7.21E-03 | 1.00 | 421 | 1 |
| cg00600584 | Non-CpG Island | <i>TP63</i>          | TSS1500    | 0.054333331  | 3.12E-02  | 0.077492828  | 4.62E-06 | 7.24E-03 | 1.00 | 144 | 3 |
| cg17191060 | Non-CpG Island |                      | Intergenic | -0.037631701 | -5.37E-02 | -0.021573798 | 4.73E-06 | 7.37E-03 | 1.00 | 422 | 1 |
| cg11731280 | Shore          | <i>TMEM210</i>       | TSS1500    | -0.037636384 | -5.37E-02 | -0.021576106 | 4.74E-06 | 7.37E-03 | 1.00 | 423 | 1 |
| cg20677824 | Non-CpG Island | <i>EFNA5</i>         | Body       | -0.051589505 | -7.36E-02 | -0.029574724 | 4.74E-06 | 7.37E-03 | 1.00 | 424 | 1 |
| cg11755558 | Non-CpG Island | <i>TNIP3</i>         | Body       | -0.057274566 | -8.17E-02 | -0.032825463 | 4.77E-06 | 7.38E-03 | 1.00 | 425 | 1 |

## Continuation of Table S 3.2

|            |                |               |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg19835136 | Island         | FAM129C       | Body       | -0.040835883 | -5.83E-02 | -0.023402794 | 4.78E-06 | 7.38E-03 | 1.00 | 426 | 1 |
| cg10753966 | Non-CpG Island | ERC2          | 3' UTR     | 0.058705922  | 3.36E-02  | 0.083767867  | 4.78E-06 | 7.38E-03 | 1.00 | 427 | 1 |
| cg20734611 | Non-CpG Island | PHF21B        | Body       | -0.028493999 | -4.07E-02 | -0.016329367 | 4.78E-06 | 7.38E-03 | 1.00 | 428 | 1 |
| cg15840151 | Non-CpG Island | C10orf120     | TSS1500    | -0.038657403 | -5.52E-02 | -0.022149648 | 4.81E-06 | 7.40E-03 | 1.00 | 235 | 2 |
| cg01323312 | Non-CpG Island | ELN           | Body       | 0.028208512  | 1.62E-02  | 0.040261783  | 4.88E-06 | 7.46E-03 | 1.00 | 230 | 3 |
| cg04954111 | Non-CpG Island | ABTB2         | Body       | -0.038729619 | -5.53E-02 | -0.022180545 | 4.88E-06 | 7.46E-03 | 1.00 | 429 | 1 |
| cg15320277 | Non-CpG Island | NRXN3         | Body       | -0.047972023 | -6.85E-02 | -0.027472851 | 4.88E-06 | 7.46E-03 | 1.00 | 430 | 1 |
| cg05729480 | Shore          | 9-Sep         | TSS1500    | -0.02276092  | -3.25E-02 | -0.01303263  | 4.90E-06 | 7.48E-03 | 1.00 | 431 | 1 |
| cg06673536 | Shelf          | CEACAM16      | 3' UTR     | -0.028953976 | -4.13E-02 | -0.016572831 | 4.96E-06 | 7.55E-03 | 1.00 | 432 | 1 |
| cg17064161 | Non-CpG Island | RALB          | Body       | -0.039694544 | -5.67E-02 | -0.022717006 | 4.98E-06 | 7.55E-03 | 1.00 | 433 | 1 |
| cg02639218 | Non-CpG Island | GBE1          | Body       | -0.058491992 | -8.35E-02 | -0.033473877 | 4.98E-06 | 7.55E-03 | 1.00 | 434 | 1 |
| cg01135863 | Non-CpG Island |               | Intergenic | -0.031769025 | -4.54E-02 | -0.018176915 | 5.01E-06 | 7.58E-03 | 1.00 | 435 | 1 |
| cg25305703 | Non-CpG Island | RP11-382A18.2 | 5' UTR     | -0.041033691 | -5.86E-02 | -0.023470954 | 5.05E-06 | 7.63E-03 | 1.00 | 436 | 1 |
| cg16998839 | Non-CpG Island | RAPGEF1       | Body       | 0.039400684  | 2.25E-02  | 0.056265749  | 5.06E-06 | 7.63E-03 | 1.00 | 437 | 1 |
| cg16698186 | Non-CpG Island | MSC-AS1       | Body       | -0.035849494 | -5.12E-02 | -0.020499631 | 5.10E-06 | 7.67E-03 | 1.00 | 438 | 1 |
| cg07328569 | Non-CpG Island | NFIA          | 5' UTR     | -0.032940616 | -4.70E-02 | -0.018834106 | 5.11E-06 | 7.67E-03 | 1.00 | 439 | 1 |
| cg27383544 | Non-CpG Island | C5orf67       | Body       | -0.041600191 | -5.94E-02 | -0.023777182 | 5.16E-06 | 7.73E-03 | 1.00 | 440 | 1 |
| cg24262469 | Shore          | TIPARP        | TSS1500    | 0.027290515  | 1.56E-02  | 0.038983333  | 5.17E-06 | 7.73E-03 | 1.00 | 441 | 2 |
| cg00281776 | Shore          | PTH2R         | TSS1500    | -0.06813673  | -9.74E-02 | -0.038914693 | 5.28E-06 | 7.87E-03 | 1.00 | 232 | 4 |
| cg02489630 | Non-CpG Island | LINC00446     | Body       | -0.042375291 | -6.05E-02 | -0.024200621 | 5.29E-06 | 7.87E-03 | 1.00 | 442 | 1 |
| cg13526488 | Shelf          | RPTOR         | Body       | -0.036490484 | -5.21E-02 | -0.020834138 | 5.33E-06 | 7.91E-03 | 1.00 | 443 | 1 |
| cg07657727 | Non-CpG Island | ST6GALNAC3    | Body       | 0.040304176  | 2.30E-02  | 0.057600267  | 5.35E-06 | 7.93E-03 | 1.00 | 444 | 1 |
| cg06478823 | Non-CpG Island | ACSM3         | TSS1500    | -0.05338998  | -7.63E-02 | -0.030458662 | 5.45E-06 | 8.06E-03 | 1.00 | 445 | 1 |
| cg05340191 | Non-CpG Island |               | Intergenic | -0.033336616 | -4.77E-02 | -0.019009302 | 5.52E-06 | 8.13E-03 | 1.00 | 446 | 1 |
| cg22327175 | Shore          | SORL1         | Body       | 0.034380488  | 1.96E-02  | 0.049156916  | 5.53E-06 | 8.13E-03 | 1.00 | 447 | 1 |
| cg12903825 | Non-CpG Island | PHYHD1        | TSS1500    | -0.031135728 | -4.45E-02 | -0.017753223 | 5.53E-06 | 8.13E-03 | 1.00 | 448 | 1 |
| cg27048989 | Non-CpG Island |               | Intergenic | -0.037099557 | -5.31E-02 | -0.021140109 | 5.63E-06 | 8.27E-03 | 1.00 | 449 | 1 |
| cg04069951 | Island         | CD81          | TSS1500    | -0.028051366 | -4.01E-02 | -0.01597919  | 5.68E-06 | 8.32E-03 | 1.00 | 450 | 1 |
| cg17192520 | Shore          | GALC          | TSS1500    | 0.026201969  | 1.49E-02  | 0.037482379  | 5.73E-06 | 8.37E-03 | 1.00 | 451 | 1 |
| cg09944942 | Non-CpG Island |               | Intergenic | -0.038759593 | -5.55E-02 | -0.022064576 | 5.79E-06 | 8.44E-03 | 1.00 | 452 | 1 |
| cg18158739 | Non-CpG Island | RP11-696D21.2 | TSS1500    | 0.067005507  | 3.81E-02  | 0.095875559  | 5.83E-06 | 8.48E-03 | 1.00 | 453 | 1 |
| cg07663850 | Shelf          | FBXL18        | 3' UTR     | 0.022387346  | 1.27E-02  | 0.032036806  | 5.87E-06 | 8.53E-03 | 1.00 | 454 | 1 |
| cg22877937 | Non-CpG Island | PKD1L2        | TSS200     | -0.031860676 | -4.56E-02 | -0.01812583  | 5.89E-06 | 8.54E-03 | 1.00 | 455 | 1 |
| cg16493238 | Non-CpG Island | RP11-494M8.4  | TSS1500    | -0.05264873  | -7.54E-02 | -0.02994361  | 5.94E-06 | 8.59E-03 | 1.00 | 456 | 1 |
| cg15706574 | Non-CpG Island | RCAN2         | Body       | -0.060337927 | -8.64E-02 | -0.034314056 | 5.96E-06 | 8.59E-03 | 1.00 | 457 | 1 |
| cg01791524 | Non-CpG Island | AGL           | 3' UTR     | -0.06426562  | -9.20E-02 | -0.036545455 | 5.97E-06 | 8.59E-03 | 1.00 | 458 | 1 |
| cg16057826 | Island         | CLIC6         | TSS1500    | -0.026193711 | -3.75E-02 | -0.014893564 | 5.99E-06 | 8.61E-03 | 1.00 | 22  | 2 |
| cg01763410 | Shore          | CCDC178       | TSS200     | 0.041530779  | 2.36E-02  | 0.059449033  | 6.00E-06 | 8.61E-03 | 1.00 | 459 | 1 |
| cg10587449 | Shore          |               | Intergenic | -0.046474802 | -6.65E-02 | -0.02641899  | 6.03E-06 | 8.63E-03 | 1.00 | 460 | 1 |



## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg10529613 | Non-CpG Island | <i>STXBP6</i>        | Body       | -0.030580441 | -4.38E-02 | -0.017366321 | 6.20E-06 | 8.86E-03 | 1.00 | 461 | 1 |
| cg18206291 | Shelf          | <i>LINC01315</i>     | Body       | -0.041888679 | -6.00E-02 | -0.023785667 | 6.21E-06 | 8.86E-03 | 1.00 | 462 | 1 |
| cg00124095 | Non-CpG Island |                      | Intergenic | -0.023899063 | -3.42E-02 | -0.013563809 | 6.30E-06 | 8.97E-03 | 1.00 | 463 | 1 |
| cg00090792 | Non-CpG Island |                      | Intergenic | -0.035603558 | -5.10E-02 | -0.020197845 | 6.38E-06 | 9.05E-03 | 1.00 | 464 | 1 |
| cg05673287 | Non-CpG Island | <i>SGK269</i>        | Body       | 0.069374595  | 3.94E-02  | 0.0993933    | 6.38E-06 | 9.05E-03 | 1.00 | 465 | 1 |
| cg25520324 | Shelf          | <i>REM2</i>          | TSS200     | -0.026893144 | -3.85E-02 | -0.015252551 | 6.42E-06 | 9.08E-03 | 1.00 | 466 | 1 |
| cg03823863 | Non-CpG Island |                      | Intergenic | -0.061733612 | -8.85E-02 | -0.03500746  | 6.45E-06 | 9.08E-03 | 1.00 | 467 | 1 |
| cg11126410 | Shore          | <i>RAB43</i>         | TSS1500    | -0.053495777 | -7.67E-02 | -0.030335753 | 6.45E-06 | 9.08E-03 | 1.00 | 468 | 1 |
| cg13431464 | Shelf          | <i>PTPN2</i>         | Body       | 0.030226213  | 1.71E-02  | 0.043312267  | 6.45E-06 | 9.08E-03 | 1.00 | 469 | 1 |
| cg18424286 | Non-CpG Island | <i>TRPC4</i>         | Body       | 0.052441531  | 2.97E-02  | 0.075148815  | 6.47E-06 | 9.09E-03 | 1.00 | 470 | 1 |
| cg06924548 | Non-CpG Island | <i>SLC44A5</i>       | Body       | 0.050660601  | 2.87E-02  | 0.072617694  | 6.60E-06 | 9.25E-03 | 1.00 | 471 | 1 |
| cg25393494 | Island         | <i>PLD6</i>          | TSS1500    | -0.070960051 | -1.02E-01 | -0.04020157  | 6.62E-06 | 9.25E-03 | 1.00 | 472 | 1 |
| cg04246882 | Non-CpG Island |                      | Intergenic | -0.02458767  | -3.52E-02 | -0.013929258 | 6.63E-06 | 9.25E-03 | 1.00 | 473 | 1 |
| cg21064451 | Shore          | <i>LGALS1</i>        | TSS200     | -0.025912501 | -3.71E-02 | -0.014678327 | 6.64E-06 | 9.26E-03 | 1.00 | 474 | 1 |
| cg10241937 | Non-CpG Island | <i>FHIT</i>          | Body       | -0.036723901 | -5.27E-02 | -0.020796474 | 6.70E-06 | 9.31E-03 | 1.00 | 475 | 1 |
| cg20040976 | Non-CpG Island |                      | Intergenic | -0.023080963 | -3.31E-02 | -0.013066886 | 6.75E-06 | 9.37E-03 | 1.00 | 476 | 1 |
| cg26101468 | Shore          | <i>ANP32B</i>        | TSS1500    | -0.026240848 | -3.76E-02 | -0.014851228 | 6.81E-06 | 9.43E-03 | 1.00 | 477 | 1 |
| cg17913504 | Non-CpG Island |                      | Intergenic | -0.02443195  | -3.50E-02 | -0.013820819 | 6.90E-06 | 9.53E-03 | 1.00 | 478 | 1 |
| cg03505501 | Shore          | <i>CADM1</i>         | Body       | 0.047137803  | 2.67E-02  | 0.067611454  | 6.90E-06 | 9.53E-03 | 1.00 | 479 | 1 |
| cg22305150 | Non-CpG Island | <i>FNDC3B</i>        | 5' UTR     | -0.033729527 | -4.84E-02 | -0.019074347 | 6.96E-06 | 9.58E-03 | 1.00 | 480 | 1 |
| cg11308227 | Non-CpG Island | <i>ENTHD2</i>        | 3' UTR     | -0.034974344 | -5.02E-02 | -0.01977647  | 6.97E-06 | 9.59E-03 | 1.00 | 481 | 2 |
| cg09225336 | Non-CpG Island | <i>RP11-351A20.2</i> | TSS200     | 0.043056167  | 2.43E-02  | 0.061775387  | 7.05E-06 | 9.67E-03 | 1.00 | 137 | 4 |
| cg02582500 | Non-CpG Island | <i>RP11-65J21.3</i>  | TSS1500    | -0.025542927 | -3.67E-02 | -0.01443572  | 7.08E-06 | 9.69E-03 | 1.00 | 482 | 1 |
| cg14566801 | Island         | <i>PTH2R</i>         | 5' UTR     | -0.035204652 | -5.05E-02 | -0.019885016 | 7.18E-06 | 9.82E-03 | 1.00 | 232 | 4 |
| cg11935615 | Non-CpG Island | <i>KIAA0430</i>      | Body       | 0.04514581   | 2.55E-02  | 0.064808946  | 7.32E-06 | 9.99E-03 | 1.00 | 483 | 1 |
| cg15350922 | Non-CpG Island |                      | Intergenic | -0.02718498  | -3.90E-02 | -0.015340981 | 7.37E-06 | 1.00E-02 | 1.00 | 484 | 1 |
| cg18186343 | Non-CpG Island | <i>MIR770</i>        | TSS1500    | -0.034099212 | -4.90E-02 | -0.019239667 | 7.40E-06 | 1.01E-02 | 1.00 | 485 | 1 |
| cg07382588 | Non-CpG Island | <i>CTDSPL2</i>       | 5' UTR     | -0.031055131 | -4.46E-02 | -0.01751145  | 7.52E-06 | 1.02E-02 | 1.00 | 486 | 1 |
| cg10637292 | Non-CpG Island | <i>KRT74</i>         | 3' UTR     | -0.022643131 | -3.25E-02 | -0.012767523 | 7.53E-06 | 1.02E-02 | 1.00 | 487 | 1 |
| cg15551092 | Non-CpG Island |                      | Intergenic | 0.027230872  | 1.53E-02  | 0.039117954  | 7.67E-06 | 1.04E-02 | 1.00 | 488 | 1 |
| cg12485221 | Non-CpG Island | <i>SH3TC2</i>        | Body       | -0.035630278 | -5.12E-02 | -0.020069481 | 7.74E-06 | 1.04E-02 | 1.00 | 489 | 1 |
| cg05371159 | Non-CpG Island | <i>ROR2</i>          | 3' UTR     | 0.052173689  | 2.94E-02  | 0.074965325  | 7.78E-06 | 1.04E-02 | 1.00 | 490 | 1 |
| cg27455444 | Non-CpG Island |                      | Intergenic | -0.036789067 | -5.29E-02 | -0.02071779  | 7.79E-06 | 1.04E-02 | 1.00 | 491 | 1 |
| cg14525458 | Shore          | <i>KDELR1</i>        | TSS1500    | -0.027045114 | -3.89E-02 | -0.015229535 | 7.80E-06 | 1.04E-02 | 1.00 | 367 | 2 |
| cg10723259 | Non-CpG Island | <i>APBA2</i>         | 5' UTR     | -0.040023981 | -5.75E-02 | -0.022538037 | 7.80E-06 | 1.04E-02 | 1.00 | 492 | 1 |
| cg14800005 | Non-CpG Island | <i>FYN</i>           | 5' UTR     | -0.032358233 | -4.65E-02 | -0.018220843 | 7.81E-06 | 1.04E-02 | 1.00 | 493 | 1 |
| cg09148948 | Non-CpG Island | <i>7SK</i>           | TSS1500    | -0.043710422 | -6.28E-02 | -0.024604041 | 7.88E-06 | 1.05E-02 | 1.00 | 494 | 1 |
| cg07806820 | Non-CpG Island | <i>ANO1</i>          | Body       | 0.026115773  | 1.47E-02  | 0.037538396  | 7.99E-06 | 1.06E-02 | 1.00 | 495 | 2 |
| cg26103696 | Non-CpG Island |                      | Intergenic | -0.038356109 | -5.51E-02 | -0.021572737 | 8.06E-06 | 1.07E-02 | 1.00 | 496 | 1 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg03023089 | Island         |                      | Intergenic | -0.035290452 | -5.07E-02 | -0.019842363 | 8.12E-06 | 1.08E-02 | 1.00 | 497 | 1 |
| cg16983110 | Non-CpG Island | <i>PTPRK</i>         | Body       | 0.040147984  | 2.26E-02  | 0.057725302  | 8.15E-06 | 1.08E-02 | 1.00 | 498 | 1 |
| cg13949829 | Island         | <i>MIR1225</i>       | TSS1500    | -0.073831574 | -1.06E-01 | -0.041489267 | 8.24E-06 | 1.09E-02 | 1.00 | 211 | 2 |
| cg19536401 | Shelf          |                      | Intergenic | -0.062791421 | -9.03E-02 | -0.035285146 | 8.24E-06 | 1.09E-02 | 1.00 | 499 | 1 |
| cg10828175 | Non-CpG Island | <i>RP11-13P5.1</i>   | TSS1500    | -0.029935796 | -4.31E-02 | -0.01681732  | 8.31E-06 | 1.09E-02 | 1.00 | 500 | 1 |
| cg03230469 | Island         | <i>GDNF</i>          | 5' UTR     | 0.045000444  | 2.53E-02  | 0.06472735   | 8.37E-06 | 1.10E-02 | 1.00 | 501 | 1 |
| cg16423910 | Island         | <i>CD37</i>          | 3' UTR     | -0.045177802 | -6.50E-02 | -0.025372296 | 8.37E-06 | 1.10E-02 | 1.00 | 25  | 3 |
| cg23372422 | Non-CpG Island | <i>MFN2</i>          | Body       | 0.031725613  | 1.78E-02  | 0.045635653  | 8.40E-06 | 1.10E-02 | 1.00 | 502 | 1 |
| cg27491025 | Shore          | <i>TMED1</i>         | Body       | -0.028007627 | -4.03E-02 | -0.015723732 | 8.45E-06 | 1.10E-02 | 1.00 | 503 | 1 |
| cg08373187 | Shelf          | <i>CXCL14</i>        | Body       | -0.029358477 | -4.22E-02 | -0.016481883 | 8.46E-06 | 1.10E-02 | 1.00 | 504 | 1 |
| cg15029812 | Non-CpG Island |                      | Intergenic | -0.020783881 | -2.99E-02 | -0.011667634 | 8.47E-06 | 1.10E-02 | 1.00 | 505 | 1 |
| cg19343186 | Non-CpG Island | <i>RNF220</i>        | Body       | -0.033739334 | -4.85E-02 | -0.018930746 | 8.58E-06 | 1.12E-02 | 1.00 | 506 | 1 |
| cg01953698 | Non-CpG Island |                      | Intergenic | 0.036484035  | 2.05E-02  | 0.052500753  | 8.62E-06 | 1.12E-02 | 1.00 | 507 | 1 |
| cg14165660 | Non-CpG Island |                      | Intergenic | -0.038762688 | -5.58E-02 | -0.021744055 | 8.64E-06 | 1.12E-02 | 1.00 | 250 | 2 |
| cg22271479 | Non-CpG Island | <i>TJP3</i>          | First Exon | -0.045086414 | -6.49E-02 | -0.025287311 | 8.67E-06 | 1.12E-02 | 1.00 | 508 | 1 |
| cg23360332 | Non-CpG Island |                      | Intergenic | -0.03925908  | -5.65E-02 | -0.022013803 | 8.73E-06 | 1.13E-02 | 1.00 | 509 | 1 |
| cg04970434 | Non-CpG Island | <i>GP1BA</i>         | 3' UTR     | -0.029089277 | -4.19E-02 | -0.01630604  | 8.80E-06 | 1.13E-02 | 1.00 | 510 | 1 |
| cg03303857 | Non-CpG Island | <i>GRIK4</i>         | Body       | -0.031851054 | -4.58E-02 | -0.017852515 | 8.82E-06 | 1.13E-02 | 1.00 | 511 | 1 |
| cg01370241 | Non-CpG Island | <i>VWCE</i>          | Body       | -0.037256308 | -5.36E-02 | -0.020862988 | 9.03E-06 | 1.16E-02 | 1.00 | 512 | 2 |
| cg18866560 | Non-CpG Island | <i>AC026904.1</i>    | 5' UTR     | -0.03668751  | -5.28E-02 | -0.020534523 | 9.15E-06 | 1.17E-02 | 1.00 | 513 | 1 |
| cg05294243 | Shore          | <i>KLK13</i>         | TSS1500    | -0.033008753 | -4.75E-02 | -0.018470372 | 9.21E-06 | 1.18E-02 | 1.00 | 514 | 1 |
| cg26613723 | Island         | <i>SREBF2</i>        | TSS1500    | -0.032115418 | -4.63E-02 | -0.017970297 | 9.22E-06 | 1.18E-02 | 1.00 | 515 | 1 |
| cg24313472 | Non-CpG Island | <i>RHOBTB1</i>       | Body       | -0.028542482 | -4.11E-02 | -0.015968206 | 9.26E-06 | 1.18E-02 | 1.00 | 516 | 1 |
| cg04650889 | Non-CpG Island | <i>APTR</i>          | Body       | -0.037440428 | -5.39E-02 | -0.020935201 | 9.39E-06 | 1.19E-02 | 1.00 | 517 | 1 |
| cg03466342 | Non-CpG Island | <i>KRT85</i>         | 3' UTR     | 0.027209533  | 1.52E-02  | 0.039208582  | 9.45E-06 | 1.20E-02 | 1.00 | 518 | 1 |
| cg12170419 | Non-CpG Island | <i>NAPB</i>          | Body       | -0.029731511 | -4.28E-02 | -0.016616577 | 9.51E-06 | 1.21E-02 | 1.00 | 519 | 1 |
| cg26657240 | Non-CpG Island | <i>IL16</i>          | First Exon | 0.10662329   | 5.96E-02  | 0.153672908  | 9.57E-06 | 1.21E-02 | 1.00 | 520 | 2 |
| cg09090090 | Non-CpG Island | <i>RP11-246K15.1</i> | 5' UTR     | -0.025798752 | -3.72E-02 | -0.014408592 | 9.68E-06 | 1.22E-02 | 1.00 | 521 | 1 |
| cg06918740 | Non-CpG Island |                      | Intergenic | -0.031917228 | -4.60E-02 | -0.017824016 | 9.70E-06 | 1.22E-02 | 1.00 | 522 | 1 |
| cg07096868 | Non-CpG Island | <i>ACVR2A</i>        | 5' UTR     | -0.024476742 | -3.53E-02 | -0.01366868  | 9.71E-06 | 1.22E-02 | 1.00 | 523 | 1 |
| cg14279641 | Non-CpG Island | <i>RP11-367F23.2</i> | 5' UTR     | -0.025371944 | -3.66E-02 | -0.014164619 | 9.78E-06 | 1.23E-02 | 1.00 | 524 | 1 |
| cg21505925 | Non-CpG Island | <i>DYRK1A</i>        | Body       | -0.042065176 | -6.06E-02 | -0.023483765 | 9.78E-06 | 1.23E-02 | 1.00 | 525 | 1 |
| cg18062205 | Shelf          | <i>KCNG1</i>         | 5' UTR     | -0.029276097 | -4.22E-02 | -0.016343835 | 9.78E-06 | 1.23E-02 | 1.00 | 526 | 1 |
| cg02402946 | Non-CpG Island | <i>CRYL1</i>         | Body       | -0.033892819 | -4.89E-02 | -0.018913607 | 9.88E-06 | 1.24E-02 | 1.00 | 527 | 1 |
| cg05938624 | Shelf          |                      | Intergenic | 0.023938449  | 1.34E-02  | 0.034519898  | 9.91E-06 | 1.24E-02 | 1.00 | 528 | 1 |
| cg11387709 | Non-CpG Island | <i>RP11-351A20.2</i> | TSS1500    | 0.055982081  | 3.12E-02  | 0.080737987  | 1.00E-05 | 1.24E-02 | 1.00 | 137 | 4 |
| cg07655853 | Non-CpG Island | <i>GRID1</i>         | Body       | -0.036452562 | -5.26E-02 | -0.02033261  | 1.00E-05 | 1.24E-02 | 1.00 | 529 | 1 |
| cg22570405 | Non-CpG Island | <i>PRKCE</i>         | Body       | -0.028292726 | -4.08E-02 | -0.015778913 | 1.00E-05 | 1.25E-02 | 1.00 | 530 | 1 |
| cg03365937 | Shore          | <i>CHST1</i>         | TSS1500    | -0.040465161 | -5.84E-02 | -0.022554334 | 1.02E-05 | 1.26E-02 | 1.00 | 531 | 1 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg07312099 | Shore          | <i>ASAH1</i>         | TSS200     | 0.112354334  | 6.26E-02  | 0.162089428  | 1.02E-05 | 1.26E-02 | 1.00 | 532 | 1 |
| cg25453686 | Non-CpG Island |                      | Intergenic | -0.02809626  | -4.05E-02 | -0.015654487 | 1.03E-05 | 1.27E-02 | 1.00 | 533 | 1 |
| cg10243322 | Shelf          | <i>KDM4B</i>         | Body       | -0.06335537  | -9.14E-02 | -0.035296625 | 1.03E-05 | 1.27E-02 | 1.00 | 534 | 1 |
| cg14054075 | Non-CpG Island | <i>HRAT56</i>        | Body       | 0.029140022  | 1.62E-02  | 0.042048709  | 1.04E-05 | 1.27E-02 | 1.00 | 535 | 1 |
| cg16947791 | Shore          | <i>ARHGEF7</i>       | Body       | 0.037370609  | 2.08E-02  | 0.053925708  | 1.04E-05 | 1.27E-02 | 1.00 | 536 | 1 |
| cg08227353 | Island         | <i>KAZALD1</i>       | First Exon | -0.028157681 | -4.07E-02 | -0.015664148 | 1.07E-05 | 1.31E-02 | 1.00 | 537 | 1 |
| cg04484639 | Non-CpG Island | <i>GBE1</i>          | Body       | -0.035916898 | -5.19E-02 | -0.019971772 | 1.08E-05 | 1.33E-02 | 1.00 | 538 | 1 |
| cg26384544 | Non-CpG Island |                      | Intergenic | -0.037930739 | -5.48E-02 | -0.021088328 | 1.09E-05 | 1.33E-02 | 1.00 | 539 | 1 |
| cg04624603 | Non-CpG Island | <i>ARPP21</i>        | 5' UTR     | 0.057335813  | 3.19E-02  | 0.082806605  | 1.10E-05 | 1.33E-02 | 1.00 | 540 | 1 |
| cg24470403 | Shore          | <i>PTGDR</i>         | TSS1500    | -0.0323652   | -4.67E-02 | -0.017987185 | 1.10E-05 | 1.33E-02 | 1.00 | 89  | 2 |
| cg14963371 | Shore          | <i>GPR3</i>          | 5' UTR     | -0.032835658 | -4.74E-02 | -0.018248428 | 1.10E-05 | 1.33E-02 | 1.00 | 541 | 1 |
| cg19643211 | Island         | <i>PPP1R13B</i>      | TSS1500    | -0.030901007 | -4.46E-02 | -0.017173191 | 1.10E-05 | 1.33E-02 | 1.00 | 542 | 1 |
| cg19839825 | Shore          | <i>PNPLA6</i>        | Body       | -0.077435648 | -1.12E-01 | -0.043009681 | 1.11E-05 | 1.35E-02 | 1.00 | 543 | 1 |
| cg00099427 | Island         | <i>CCDC19</i>        | TSS1500    | -0.037489829 | -5.42E-02 | -0.020821544 | 1.11E-05 | 1.35E-02 | 1.00 | 206 | 3 |
| cg09205359 | Non-CpG Island |                      | Intergenic | 0.032962588  | 1.83E-02  | 0.047618243  | 1.12E-05 | 1.35E-02 | 1.00 | 544 | 1 |
| cg02202721 | Shore          | <i>PEAR1</i>         | 5' UTR     | -0.021396139 | -3.09E-02 | -0.011880488 | 1.12E-05 | 1.35E-02 | 1.00 | 545 | 1 |
| cg17615115 | Non-CpG Island |                      | Intergenic | -0.024136242 | -3.49E-02 | -0.013395987 | 1.13E-05 | 1.37E-02 | 1.00 | 546 | 1 |
| cg15086607 | Non-CpG Island |                      | Intergenic | -0.036898641 | -5.33E-02 | -0.020477615 | 1.14E-05 | 1.37E-02 | 1.00 | 547 | 1 |
| cg17178468 | Shore          |                      | Intergenic | -0.026788863 | -3.87E-02 | -0.014866632 | 1.14E-05 | 1.37E-02 | 1.00 | 548 | 1 |
| cg08525461 | Non-CpG Island |                      | Intergenic | -0.029235186 | -4.22E-02 | -0.016221371 | 1.14E-05 | 1.37E-02 | 1.00 | 549 | 1 |
| cg21121708 | Shore          | <i>MPPED1</i>        | Body       | -0.04781284  | -6.91E-02 | -0.026527188 | 1.14E-05 | 1.37E-02 | 1.00 | 550 | 1 |
| cg18265326 | Non-CpG Island | <i>RP11-351A20.2</i> | TSS200     | 0.041803378  | 2.32E-02  | 0.060416085  | 1.15E-05 | 1.37E-02 | 1.00 | 137 | 4 |
| cg22477023 | Non-CpG Island | <i>ERBB4</i>         | Body       | -0.039238589 | -5.67E-02 | -0.021762605 | 1.15E-05 | 1.37E-02 | 1.00 | 551 | 1 |
| cg24117197 | Shelf          | <i>7SK</i>           | TSS1500    | -0.033037116 | -4.78E-02 | -0.018323005 | 1.15E-05 | 1.37E-02 | 1.00 | 552 | 1 |
| cg05411199 | Shore          | <i>C6orf145</i>      | Body       | -0.039349361 | -5.69E-02 | -0.021815292 | 1.17E-05 | 1.39E-02 | 1.00 | 553 | 1 |
| cg15344192 | Shore          |                      | Intergenic | 0.034381949  | 1.91E-02  | 0.049708538  | 1.17E-05 | 1.39E-02 | 1.00 | 554 | 1 |
| cg09848638 | Non-CpG Island | <i>PRKAR1B</i>       | Body       | 0.042428427  | 2.35E-02  | 0.061342407  | 1.18E-05 | 1.39E-02 | 1.00 | 555 | 1 |
| cg18253885 | Non-CpG Island | <i>PPARGC1A</i>      | TSS1500    | 0.027145546  | 1.50E-02  | 0.039250308  | 1.18E-05 | 1.40E-02 | 1.00 | 556 | 1 |
| cg16294731 | Non-CpG Island | <i>LOC102724691</i>  | Body       | -0.030041116 | -4.34E-02 | -0.016642882 | 1.19E-05 | 1.40E-02 | 1.00 | 557 | 1 |
| cg20015040 | Non-CpG Island |                      | Intergenic | -0.043111665 | -6.23E-02 | -0.023874596 | 1.20E-05 | 1.41E-02 | 1.00 | 558 | 1 |
| cg05235171 | Non-CpG Island | <i>TMPRSS4</i>       | Body       | -0.049498727 | -7.16E-02 | -0.027411202 | 1.20E-05 | 1.41E-02 | 1.00 | 559 | 1 |
| cg19403323 | Non-CpG Island | <i>C1orf107</i>      | Body       | 0.035263014  | 1.95E-02  | 0.051000442  | 1.20E-05 | 1.41E-02 | 1.00 | 560 | 1 |
| cg08282375 | Non-CpG Island | <i>LINC01272</i>     | TSS200     | -0.020931938 | -3.03E-02 | -0.01158953  | 1.20E-05 | 1.41E-02 | 1.00 | 561 | 1 |
| cg07088882 | Shelf          | <i>RP11-236J17.3</i> | TSS200     | -0.034426262 | -4.98E-02 | -0.019051625 | 1.22E-05 | 1.43E-02 | 1.00 | 65  | 2 |
| cg16824480 | Non-CpG Island | <i>RNF157</i>        | Body       | 0.034037785  | 1.88E-02  | 0.049241806  | 1.22E-05 | 1.43E-02 | 1.00 | 562 | 1 |
| cg12277336 | Shelf          | <i>MCOLN2</i>        | Body       | -0.044439416 | -6.43E-02 | -0.024582417 | 1.23E-05 | 1.44E-02 | 1.00 | 563 | 1 |
| cg21732535 | Shore          | <i>PTH2R</i>         | TSS1500    | -0.047852371 | -6.92E-02 | -0.026468348 | 1.23E-05 | 1.44E-02 | 1.00 | 232 | 4 |
| cg13620631 | Non-CpG Island | <i>NTRK2</i>         | 3' UTR     | 0.037016639  | 2.05E-02  | 0.053559846  | 1.24E-05 | 1.44E-02 | 1.00 | 564 | 1 |
| cg04277753 | Non-CpG Island | <i>CMIP</i>          | Body       | 0.031420037  | 1.74E-02  | 0.045474277  | 1.26E-05 | 1.46E-02 | 1.00 | 565 | 1 |

## Continuation of Table S 3.2

|            |                |               |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg23054584 | Shore          | AC092296.1    | 3' UTR     | 0.044845594  | 2.48E-02  | 0.064907734  | 1.26E-05 | 1.46E-02 | 1.00 | 566 | 1 |
| cg00003287 | Non-CpG Island | TNNT2         | 5' UTR     | -0.028259489 | -4.09E-02 | -0.01561593  | 1.26E-05 | 1.46E-02 | 1.00 | 567 | 1 |
| cg24181174 | Non-CpG Island | TACC2         | Body       | 0.030231345  | 1.67E-02  | 0.043757194  | 1.26E-05 | 1.46E-02 | 1.00 | 568 | 1 |
| cg14986541 | Shelf          |               | Intergenic | -0.035307559 | -5.11E-02 | -0.019509868 | 1.26E-05 | 1.46E-02 | 1.00 | 569 | 1 |
| cg07100050 | Non-CpG Island | NEBL          | Body       | 0.033401553  | 1.84E-02  | 0.048354826  | 1.28E-05 | 1.47E-02 | 1.00 | 570 | 1 |
| cg11359094 | Non-CpG Island | OPRD1         | Body       | -0.037012203 | -5.36E-02 | -0.020440705 | 1.28E-05 | 1.48E-02 | 1.00 | 571 | 1 |
| cg09808211 | Non-CpG Island | XKR6          | Body       | 0.038306506  | 2.11E-02  | 0.055472405  | 1.30E-05 | 1.50E-02 | 1.00 | 572 | 1 |
| cg21723486 | Non-CpG Island | TP63          | First Exon | 0.055241616  | 3.05E-02  | 0.080008178  | 1.32E-05 | 1.51E-02 | 1.00 | 144 | 3 |
| cg00776917 | Non-CpG Island | CTD-2251F13.1 | 3' UTR     | -0.032771227 | -4.75E-02 | -0.018073771 | 1.33E-05 | 1.52E-02 | 1.00 | 573 | 1 |
| cg19402463 | Non-CpG Island | GPR107        | Body       | 0.038101411  | 2.10E-02  | 0.055193068  | 1.33E-05 | 1.52E-02 | 1.00 | 574 | 1 |
| cg11993828 | Non-CpG Island | LINC01258     | TSS1500    | -0.03069348  | -4.45E-02 | -0.016915952 | 1.35E-05 | 1.54E-02 | 1.00 | 575 | 1 |
| cg00213082 | Non-CpG Island |               | Intergenic | 0.035733014  | 1.97E-02  | 0.05178819   | 1.37E-05 | 1.57E-02 | 1.00 | 576 | 1 |
| cg17656881 | Non-CpG Island | ENTHD2        | 3' UTR     | -0.030919626 | -4.48E-02 | -0.017022714 | 1.38E-05 | 1.57E-02 | 1.00 | 481 | 2 |
| cg09958060 | Non-CpG Island | DTX1          | Body       | -0.039162163 | -5.68E-02 | -0.021558011 | 1.39E-05 | 1.58E-02 | 1.00 | 577 | 1 |
| cg22828783 | Non-CpG Island | DCUN1D5       | 3' UTR     | 0.019073854  | 1.05E-02  | 0.027653989  | 1.41E-05 | 1.60E-02 | 1.00 | 578 | 1 |
| cg09377704 | Shore          | AC010336.1    | TSS1500    | -0.031429856 | -4.56E-02 | -0.017288093 | 1.41E-05 | 1.60E-02 | 1.00 | 579 | 1 |
| cg00889363 | Non-CpG Island | PARD3B        | Body       | -0.019760731 | -2.87E-02 | -0.010869226 | 1.41E-05 | 1.60E-02 | 1.00 | 580 | 1 |
| cg24119717 | Shore          | CTB-131B5.4   | TSS1500    | 0.028073655  | 1.54E-02  | 0.040706823  | 1.42E-05 | 1.60E-02 | 1.00 | 581 | 3 |
| cg26920605 | Non-CpG Island | TRPM2         | TSS1500    | -0.02801489  | -4.06E-02 | -0.015399243 | 1.44E-05 | 1.62E-02 | 1.00 | 582 | 1 |
| cg17725163 | Non-CpG Island | ALPL          | Body       | -0.031278204 | -4.54E-02 | -0.017190977 | 1.44E-05 | 1.62E-02 | 1.00 | 583 | 1 |
| cg21110915 | Non-CpG Island |               | Intergenic | -0.031078821 | -4.51E-02 | -0.017081241 | 1.44E-05 | 1.62E-02 | 1.00 | 584 | 1 |
| cg22356061 | Non-CpG Island | SNAP47        | Body       | -0.028939111 | -4.20E-02 | -0.015904309 | 1.44E-05 | 1.62E-02 | 1.00 | 585 | 2 |
| cg15345567 | Non-CpG Island | AC018712.3    | 3' UTR     | 0.031860072  | 1.75E-02  | 0.046216007  | 1.45E-05 | 1.63E-02 | 1.00 | 586 | 1 |
| cg09043768 | Non-CpG Island |               | Intergenic | -0.031660986 | -4.59E-02 | -0.017391764 | 1.46E-05 | 1.63E-02 | 1.00 | 587 | 1 |
| cg16786154 | Non-CpG Island | HCAR1         | 3' UTR     | 0.037516957  | 2.06E-02  | 0.054437954  | 1.48E-05 | 1.65E-02 | 1.00 | 588 | 1 |
| cg17447841 | Non-CpG Island |               | Intergenic | -0.035584683 | -5.16E-02 | -0.019534721 | 1.48E-05 | 1.65E-02 | 1.00 | 589 | 1 |
| cg18942298 | Non-CpG Island | JPH2          | Body       | -0.028005798 | -4.06E-02 | -0.015371876 | 1.49E-05 | 1.65E-02 | 1.00 | 590 | 1 |
| cg12261095 | Non-CpG Island | FCER2         | Body       | -0.031758775 | -4.61E-02 | -0.017431376 | 1.49E-05 | 1.65E-02 | 1.00 | 591 | 1 |
| cg10799510 | Non-CpG Island |               | Intergenic | 0.027698118  | 1.52E-02  | 0.040194849  | 1.49E-05 | 1.65E-02 | 1.00 | 592 | 1 |
| cg17865555 | Island         | SPHK2         | 3' UTR     | -0.07015974  | -1.02E-01 | -0.038495801 | 1.50E-05 | 1.66E-02 | 1.00 | 263 | 2 |
| cg17274885 | Non-CpG Island | SLC15A1       | Body       | 0.046624715  | 2.56E-02  | 0.067669382  | 1.50E-05 | 1.66E-02 | 1.00 | 593 | 1 |
| cg16990252 | Non-CpG Island | AC100802.3    | 5' UTR     | -0.03538829  | -5.14E-02 | -0.019413179 | 1.51E-05 | 1.66E-02 | 1.00 | 594 | 1 |
| cg14484792 | Non-CpG Island | CTD-2058B24.2 | 3' UTR     | 0.060390409  | 3.31E-02  | 0.087652647  | 1.51E-05 | 1.66E-02 | 1.00 | 595 | 1 |
| cg13108341 | Non-CpG Island | DNAH9         | Body       | 0.079671113  | 4.37E-02  | 0.115639962  | 1.51E-05 | 1.66E-02 | 1.00 | 596 | 1 |
| cg25006851 | Non-CpG Island | MYST4         | Body       | -0.03341719  | -4.85E-02 | -0.018330099 | 1.51E-05 | 1.66E-02 | 1.00 | 597 | 1 |
| cg27027151 | Non-CpG Island | IL21R         | 3' UTR     | -0.030660626 | -4.45E-02 | -0.016810919 | 1.52E-05 | 1.68E-02 | 1.00 | 598 | 1 |
| cg12124090 | Non-CpG Island |               | Intergenic | 0.04417292   | 2.42E-02  | 0.064134347  | 1.54E-05 | 1.69E-02 | 1.00 | 599 | 2 |
| cg14467966 | Non-CpG Island | PDGFC         | Body       | -0.041606554 | -6.04E-02 | -0.02279059  | 1.56E-05 | 1.71E-02 | 1.00 | 600 | 1 |
| cg03348792 | Non-CpG Island | KRT1          | TSS1500    | 0.026544635  | 1.45E-02  | 0.038549408  | 1.56E-05 | 1.71E-02 | 1.00 | 601 | 1 |

## Continuation of Table S 3.2

|            |                |                     |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg15600670 | Non-CpG Island | <i>RP11-167B3.1</i> | TSS1500    | -0.027404513 | -3.98E-02 | -0.015007862 | 1.57E-05 | 1.71E-02 | 1.00 | 602 | 1 |
| cg00572387 | Non-CpG Island | <i>SIRPB2</i>       | TSS1500    | -0.071213325 | -1.03E-01 | -0.038985309 | 1.58E-05 | 1.72E-02 | 1.00 | 603 | 1 |
| cg20696080 | Shore          | <i>TEAD2</i>        | 5' UTR     | -0.033876601 | -4.92E-02 | -0.018545174 | 1.58E-05 | 1.72E-02 | 1.00 | 288 | 2 |
| cg18474095 | Shore          | <i>FAM181B</i>      | TSS1500    | -0.027308082 | -3.97E-02 | -0.014937127 | 1.61E-05 | 1.75E-02 | 1.00 | 604 | 1 |
| cg06676511 | Shore          | <i>ID2-AS1</i>      | Body       | -0.02841401  | -4.13E-02 | -0.015537546 | 1.62E-05 | 1.76E-02 | 1.00 | 605 | 1 |
| cg22279720 | Non-CpG Island | <i>NRP1</i>         | Body       | -0.032508386 | -4.72E-02 | -0.017773894 | 1.63E-05 | 1.77E-02 | 1.00 | 606 | 1 |
| cg12019614 | Island         | <i>DOCK6</i>        | Body       | -0.053116057 | -7.72E-02 | -0.029007607 | 1.67E-05 | 1.81E-02 | 1.00 | 607 | 1 |
| cg24086717 | Non-CpG Island | <i>GLIS1</i>        | 5' UTR     | -0.027455239 | -3.99E-02 | -0.014989876 | 1.68E-05 | 1.82E-02 | 1.00 | 608 | 1 |
| cg19046904 | Shelf          |                     | Intergenic | -0.024403809 | -3.55E-02 | -0.013319609 | 1.70E-05 | 1.83E-02 | 1.00 | 609 | 1 |
| cg19180827 | Non-CpG Island | <i>MSH2</i>         | Body       | -0.024792069 | -3.61E-02 | -0.013531231 | 1.70E-05 | 1.83E-02 | 1.00 | 610 | 1 |
| cg19026961 | Non-CpG Island | <i>ANO1</i>         | Body       | 0.021274594  | 1.16E-02  | 0.030938723  | 1.70E-05 | 1.83E-02 | 1.00 | 495 | 2 |
| cg15243205 | Shore          | <i>HIST1H3J</i>     | TSS1500    | -0.045819779 | -6.66E-02 | -0.025000082 | 1.71E-05 | 1.83E-02 | 1.00 | 611 | 1 |
| cg05992638 | Non-CpG Island | <i>LINC01162</i>    | Body       | -0.051755019 | -7.53E-02 | -0.028232484 | 1.72E-05 | 1.84E-02 | 1.00 | 612 | 1 |
| cg16284880 | Non-CpG Island |                     | Intergenic | -0.040688805 | -5.92E-02 | -0.022195025 | 1.72E-05 | 1.84E-02 | 1.00 | 252 | 2 |
| cg15991632 | Non-CpG Island | <i>PHKB</i>         | Body       | -0.031589451 | -4.59E-02 | -0.017229019 | 1.72E-05 | 1.84E-02 | 1.00 | 613 | 1 |
| cg06749277 | Non-CpG Island |                     | Intergenic | -0.039465079 | -5.74E-02 | -0.021520412 | 1.73E-05 | 1.85E-02 | 1.00 | 614 | 1 |
| cg03147691 | Shore          | <i>HSD11B1L</i>     | Body       | -0.032656105 | -4.75E-02 | -0.017799931 | 1.75E-05 | 1.86E-02 | 1.00 | 615 | 1 |
| cg11241750 | Non-CpG Island | <i>GRAMD3</i>       | TSS1500    | 0.028065058  | 1.53E-02  | 0.040836386  | 1.76E-05 | 1.87E-02 | 1.00 | 616 | 1 |
| cg25698975 | Non-CpG Island |                     | Intergenic | -0.053666197 | -7.81E-02 | -0.029239944 | 1.77E-05 | 1.88E-02 | 1.00 | 617 | 1 |
| cg21501831 | Non-CpG Island |                     | Intergenic | -0.027784405 | -4.04E-02 | -0.015137173 | 1.77E-05 | 1.88E-02 | 1.00 | 618 | 1 |
| cg09661126 | Non-CpG Island | <i>CTIF</i>         | Body       | 0.034593772  | 1.88E-02  | 0.050342873  | 1.77E-05 | 1.88E-02 | 1.00 | 619 | 1 |
| cg02711042 | Non-CpG Island |                     | Intergenic | -0.025711994 | -3.74E-02 | -0.014005556 | 1.78E-05 | 1.88E-02 | 1.00 | 620 | 1 |
| cg22303565 | Shelf          | <i>NCOR2</i>        | Body       | -0.02295167  | -3.34E-02 | -0.01250195  | 1.78E-05 | 1.88E-02 | 1.00 | 621 | 1 |
| cg08993368 | Shore          |                     | Intergenic | 0.037022996  | 2.02E-02  | 0.053880301  | 1.78E-05 | 1.88E-02 | 1.00 | 622 | 1 |
| cg22355342 | Non-CpG Island | <i>EXOC4</i>        | Body       | -0.033218077 | -4.83E-02 | -0.018090067 | 1.79E-05 | 1.88E-02 | 1.00 | 623 | 1 |
| cg23171506 | Non-CpG Island | <i>LOC388282</i>    | TSS1500    | -0.019138786 | -2.79E-02 | -0.010420045 | 1.80E-05 | 1.89E-02 | 1.00 | 624 | 1 |
| cg05224015 | Non-CpG Island | <i>PTPRN2</i>       | Body       | -0.029626889 | -4.31E-02 | -0.016123364 | 1.81E-05 | 1.90E-02 | 1.00 | 625 | 1 |
| cg19789753 | Island         |                     | Intergenic | -0.059539182 | -8.67E-02 | -0.032401136 | 1.81E-05 | 1.90E-02 | 1.00 | 626 | 1 |
| cg26682566 | Non-CpG Island |                     | Intergenic | 0.065400902  | 3.56E-02  | 0.09521343   | 1.82E-05 | 1.90E-02 | 1.00 | 627 | 1 |
| cg24783018 | Non-CpG Island | <i>ROR1</i>         | Body       | 0.039147413  | 2.13E-02  | 0.057001911  | 1.84E-05 | 1.92E-02 | 1.00 | 628 | 1 |
| cg22667835 | Shore          | <i>JAKMIP1</i>      | Body       | -0.033905617 | -4.94E-02 | -0.018437406 | 1.85E-05 | 1.93E-02 | 1.00 | 629 | 2 |
| cg10592521 | Shelf          | <i>MEX3D</i>        | Body       | -0.036289139 | -5.29E-02 | -0.019725599 | 1.86E-05 | 1.94E-02 | 1.00 | 630 | 1 |
| cg02343335 | Non-CpG Island | <i>IL1F8</i>        | Body       | -0.035209795 | -5.13E-02 | -0.019137684 | 1.87E-05 | 1.94E-02 | 1.00 | 631 | 1 |
| cg15656455 | Non-CpG Island |                     | Intergenic | -0.027794282 | -4.05E-02 | -0.015106196 | 1.87E-05 | 1.94E-02 | 1.00 | 632 | 1 |
| cg03554283 | Island         | <i>HNF4G</i>        | TSS1500    | 0.02474535   | 1.34E-02  | 0.03604199   | 1.87E-05 | 1.94E-02 | 1.00 | 633 | 1 |
| cg05068452 | Shore          | <i>UGT8</i>         | 5' UTR     | 0.019128446  | 1.04E-02  | 0.027862783  | 1.88E-05 | 1.95E-02 | 1.00 | 634 | 1 |
| cg12950271 | Non-CpG Island | <i>ATP2B2</i>       | 5' UTR     | 0.038167092  | 2.07E-02  | 0.055599474  | 1.89E-05 | 1.95E-02 | 1.00 | 635 | 1 |
| cg24680570 | Non-CpG Island |                     | Intergenic | -0.027872146 | -4.06E-02 | -0.015141704 | 1.89E-05 | 1.95E-02 | 1.00 | 636 | 1 |
| cg13735166 | Non-CpG Island |                     | Intergenic | 0.022648497  | 1.23E-02  | 0.032995796  | 1.90E-05 | 1.96E-02 | 1.00 | 637 | 1 |

## Continuation of Table S 3.2

|            |                |                       |            |              |           |              |          |          |      |     |   |
|------------|----------------|-----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg12054453 | Non-CpG Island | <i>TMEM49</i>         | Body       | -0.065181175 | -9.50E-02 | -0.03539611  | 1.90E-05 | 1.96E-02 | 1.00 | 638 | 1 |
| cg26485048 | Non-CpG Island | <i>AC006465.3</i>     | 5' UTR     | -0.03329312  | -4.85E-02 | -0.018075323 | 1.91E-05 | 1.97E-02 | 1.00 | 639 | 1 |
| cg17241935 | Shore          | <i>GPR4</i>           | 5' UTR     | -0.019355184 | -2.82E-02 | -0.01050805  | 1.92E-05 | 1.97E-02 | 1.00 | 640 | 1 |
| cg25316901 | Non-CpG Island | <i>C1orf127</i>       | Body       | -0.037101344 | -5.41E-02 | -0.020138975 | 1.92E-05 | 1.97E-02 | 1.00 | 641 | 1 |
| cg02634272 | Island         | <i>RP11-510M2.1</i>   | 3' UTR     | -0.037919916 | -5.53E-02 | -0.020579988 | 1.93E-05 | 1.98E-02 | 1.00 | 21  | 2 |
| cg08263205 | Non-CpG Island |                       | Intergenic | -0.018850378 | -2.75E-02 | -0.010227327 | 1.94E-05 | 1.99E-02 | 1.00 | 642 | 1 |
| cg24022624 | Non-CpG Island |                       | Intergenic | -0.04107755  | -5.99E-02 | -0.022286296 | 1.94E-05 | 1.99E-02 | 1.00 | 643 | 1 |
| cg03793277 | Non-CpG Island | <i>APOC1</i>          | TSS1500    | -0.026203739 | -3.82E-02 | -0.014214688 | 1.95E-05 | 1.99E-02 | 1.00 | 644 | 1 |
| cg07920076 | Non-CpG Island | <i>TGIF2-C20orf24</i> | Body       | -0.040702478 | -5.93E-02 | -0.0220751   | 1.96E-05 | 2.00E-02 | 1.00 | 645 | 1 |
| cg13397738 | Non-CpG Island | <i>DDX10</i>          | Body       | -0.04778801  | -6.97E-02 | -0.0259081   | 1.98E-05 | 2.01E-02 | 1.00 | 646 | 1 |
| cg25514301 | Shore          | <i>GRIN3A</i>         | TSS1500    | 0.050202036  | 2.72E-02  | 0.073194443  | 1.99E-05 | 2.02E-02 | 1.00 | 647 | 1 |
| cg11754344 | Non-CpG Island |                       | Intergenic | -0.045207762 | -6.59E-02 | -0.024502689 | 1.99E-05 | 2.02E-02 | 1.00 | 648 | 1 |
| cg19146711 | Non-CpG Island |                       | Intergenic | -0.04223055  | -6.16E-02 | -0.022887497 | 1.99E-05 | 2.02E-02 | 1.00 | 649 | 1 |
| cg09399176 | Shore          | <i>INSC</i>           | TSS1500    | 0.029998263  | 1.63E-02  | 0.043740201  | 2.00E-05 | 2.02E-02 | 1.00 | 650 | 1 |
| cg04233270 | Non-CpG Island | <i>SLIT1</i>          | Body       | -0.030816076 | -4.49E-02 | -0.016687297 | 2.03E-05 | 2.05E-02 | 1.00 | 651 | 1 |
| cg02746002 | Non-CpG Island |                       | Intergenic | -0.036318817 | -5.30E-02 | -0.019655922 | 2.06E-05 | 2.07E-02 | 1.00 | 652 | 1 |
| cg27042983 | Non-CpG Island |                       | Intergenic | 0.028523     | 1.54E-02  | 0.041611339  | 2.06E-05 | 2.08E-02 | 1.00 | 653 | 1 |
| cg04810224 | Non-CpG Island | <i>PRKG1</i>          | Body       | -0.048219716 | -7.04E-02 | -0.026042795 | 2.15E-05 | 2.16E-02 | 1.00 | 654 | 1 |
| cg10044623 | Non-CpG Island | <i>LOC100133286</i>   | Body       | -0.042942636 | -6.27E-02 | -0.023191953 | 2.15E-05 | 2.16E-02 | 1.00 | 655 | 1 |
| cg12608784 | Shore          | <i>NSMF</i>           | Body       | -0.027450653 | -4.01E-02 | -0.014824186 | 2.16E-05 | 2.17E-02 | 1.00 | 203 | 2 |
| cg17930194 | Island         | <i>LHX6</i>           | Body       | -0.062377051 | -9.11E-02 | -0.033678194 | 2.17E-05 | 2.17E-02 | 1.00 | 656 | 1 |
| cg27165960 | Non-CpG Island | <i>FYB</i>            | Body       | -0.028168139 | -4.11E-02 | -0.015205924 | 2.17E-05 | 2.18E-02 | 1.00 | 657 | 1 |
| cg14907703 | Non-CpG Island | <i>LINC01561</i>      | Body       | -0.028648347 | -4.18E-02 | -0.015462872 | 2.18E-05 | 2.18E-02 | 1.00 | 658 | 1 |
| cg00234726 | Non-CpG Island |                       | Intergenic | -0.032024054 | -4.68E-02 | -0.017280163 | 2.19E-05 | 2.19E-02 | 1.00 | 659 | 1 |
| cg03761216 | Shore          | <i>EPS15</i>          | TSS1500    | -0.038220198 | -5.58E-02 | -0.020612751 | 2.22E-05 | 2.21E-02 | 1.00 | 660 | 1 |
| cg09380122 | Non-CpG Island | <i>DICER1</i>         | Body       | 0.030680638  | 1.65E-02  | 0.044829775  | 2.27E-05 | 2.26E-02 | 1.00 | 661 | 1 |
| cg19753867 | Non-CpG Island | <i>DSN1</i>           | Body       | 0.023175616  | 1.25E-02  | 0.033865743  | 2.27E-05 | 2.26E-02 | 1.00 | 662 | 1 |
| cg09139810 | Non-CpG Island |                       | Intergenic | -0.031865264 | -4.66E-02 | -0.017164163 | 2.28E-05 | 2.26E-02 | 1.00 | 663 | 1 |
| cg08758963 | Non-CpG Island | <i>MRVI1</i>          | 5' UTR     | -0.023068372 | -3.37E-02 | -0.012425512 | 2.28E-05 | 2.26E-02 | 1.00 | 664 | 1 |
| cg06061791 | Shore          | <i>UQCRRF51</i>       | Body       | -0.0421621   | -6.16E-02 | -0.022710004 | 2.28E-05 | 2.26E-02 | 1.00 | 665 | 1 |
| cg01894038 | Non-CpG Island | <i>KCTD14</i>         | 5' UTR     | -0.024324071 | -3.55E-02 | -0.013099748 | 2.29E-05 | 2.26E-02 | 1.00 | 666 | 1 |
| cg11538149 | Shore          | <i>GLRA3</i>          | TSS1500    | 0.043496133  | 2.34E-02  | 0.063568892  | 2.29E-05 | 2.26E-02 | 1.00 | 667 | 1 |
| cg10442251 | Shelf          | <i>IGSF21</i>         | Body       | -0.031510421 | -4.61E-02 | -0.01696225  | 2.31E-05 | 2.28E-02 | 1.00 | 668 | 1 |
| cg17252620 | Non-CpG Island |                       | Intergenic | -0.033566553 | -4.91E-02 | -0.018067192 | 2.32E-05 | 2.28E-02 | 1.00 | 16  | 2 |
| cg09440862 | Island         | <i>NTRK2</i>          | 5' UTR     | 0.044986886  | 2.42E-02  | 0.065762338  | 2.32E-05 | 2.29E-02 | 1.00 | 669 | 1 |
| cg23756527 | Non-CpG Island |                       | Intergenic | -0.021670366 | -3.17E-02 | -0.011655873 | 2.35E-05 | 2.31E-02 | 1.00 | 670 | 1 |
| cg12750791 | Non-CpG Island |                       | Intergenic | 0.04974682   | 2.68E-02  | 0.072738342  | 2.36E-05 | 2.31E-02 | 1.00 | 671 | 1 |
| cg00407468 | Non-CpG Island | <i>DENND1A</i>        | TSS1500    | -0.03680771  | -5.38E-02 | -0.01979092  | 2.37E-05 | 2.32E-02 | 1.00 | 672 | 1 |
| cg13054321 | Non-CpG Island | <i>JAKMIP1</i>        | 5' UTR     | -0.031298035 | -4.58E-02 | -0.016828448 | 2.37E-05 | 2.32E-02 | 1.00 | 673 | 1 |

## Continuation of Table S 3.2

|            |                |                      |               |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|---------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg19108718 | Island         | <i>RFESD</i>         | TSS200        | -0.044702158 | -6.54E-02 | -0.024031665 | 2.38E-05 | 2.33E-02 | 1.00 | 674 | 1 |
| cg08500818 | Shore          | <i>SLC12A4</i>       | TSS1500       | -0.033971369 | -4.97E-02 | -0.018261218 | 2.39E-05 | 2.33E-02 | 1.00 | 675 | 1 |
| cg03473532 | Shelf          | <i>MKLN1</i>         | Body          | 0.053651744  | 2.88E-02  | 0.07846621   | 2.39E-05 | 2.33E-02 | 1.00 | 676 | 1 |
| cg11492723 | Non-CpG Island | <i>RP11-194G10.1</i> | TSS200        | -0.034451687 | -5.04E-02 | -0.018516805 | 2.39E-05 | 2.33E-02 | 1.00 | 172 | 2 |
| cg11891925 | Non-CpG Island | <i>CCDC3</i>         | Exon Boundary | -0.024095802 | -3.52E-02 | -0.012942045 | 2.43E-05 | 2.36E-02 | 1.00 | 677 | 1 |
| cg26103104 | Shore          | <i>PURA</i>          | 5' UTR        | -0.039611802 | -5.79E-02 | -0.02127551  | 2.43E-05 | 2.36E-02 | 1.00 | 581 | 3 |
| cg08559711 | Non-CpG Island | <i>CDKAL1</i>        | Body          | -0.031558097 | -4.62E-02 | -0.016946924 | 2.44E-05 | 2.36E-02 | 1.00 | 678 | 1 |
| cg16931200 | Shore          | <i>RASGEF1C</i>      | Body          | -0.034319974 | -5.02E-02 | -0.018428379 | 2.44E-05 | 2.36E-02 | 1.00 | 679 | 1 |
| cg08547137 | Non-CpG Island | <i>PREX1</i>         | Body          | 0.022417693  | 1.20E-02  | 0.032800645  | 2.45E-05 | 2.37E-02 | 1.00 | 680 | 1 |
| cg17125270 | Shore          | <i>PLD3</i>          | 5' UTR        | -0.051740159 | -7.57E-02 | -0.027769867 | 2.47E-05 | 2.38E-02 | 1.00 | 365 | 2 |
| cg13951597 | Shore          |                      | Intergenic    | -0.037517807 | -5.49E-02 | -0.020135937 | 2.47E-05 | 2.38E-02 | 1.00 | 681 | 1 |
| cg11453712 | Island         | <i>PHPT1</i>         | TSS1500       | -0.039065984 | -5.72E-02 | -0.020965066 | 2.47E-05 | 2.38E-02 | 1.00 | 682 | 1 |
| cg21773245 | Non-CpG Island | <i>NDNF</i>          | 5' UTR        | 0.038744604  | 2.08E-02  | 0.056707292  | 2.50E-05 | 2.40E-02 | 1.00 | 683 | 1 |
| cg14492992 | Shelf          |                      | Intergenic    | -0.037517932 | -5.49E-02 | -0.020115906 | 2.52E-05 | 2.42E-02 | 1.00 | 684 | 1 |
| cg16993186 | Non-CpG Island | <i>CEL2F2</i>        | Body          | -0.026656403 | -3.90E-02 | -0.014291739 | 2.52E-05 | 2.42E-02 | 1.00 | 685 | 1 |
| cg11854372 | Non-CpG Island | <i>NCS1</i>          | Body          | -0.025649449 | -3.75E-02 | -0.01375073  | 2.53E-05 | 2.42E-02 | 1.00 | 686 | 1 |
| cg19115017 | Non-CpG Island | <i>INSC</i>          | Body          | -0.025216968 | -3.69E-02 | -0.013517983 | 2.53E-05 | 2.42E-02 | 1.00 | 687 | 1 |
| cg08513698 | Non-CpG Island | <i>ATP11A</i>        | Body          | -0.041740468 | -6.11E-02 | -0.022366275 | 2.55E-05 | 2.44E-02 | 1.00 | 688 | 1 |
| cg12230203 | Non-CpG Island |                      | Intergenic    | -0.039308183 | -5.76E-02 | -0.021059049 | 2.56E-05 | 2.44E-02 | 1.00 | 689 | 1 |
| cg06064155 | Non-CpG Island |                      | Intergenic    | -0.016528621 | -2.42E-02 | -0.008854106 | 2.57E-05 | 2.45E-02 | 1.00 | 690 | 1 |
| cg18993113 | Non-CpG Island | <i>SBF2</i>          | Body          | -0.031977596 | -4.68E-02 | -0.017126099 | 2.58E-05 | 2.46E-02 | 1.00 | 691 | 1 |
| cg26288449 | Shore          | <i>SBK2</i>          | TSS1500       | -0.031621798 | -4.63E-02 | -0.016933266 | 2.59E-05 | 2.46E-02 | 1.00 | 169 | 2 |
| cg00741432 | Shore          | <i>PARM1</i>         | Body          | -0.04325598  | -6.33E-02 | -0.023162037 | 2.59E-05 | 2.46E-02 | 1.00 | 692 | 1 |
| cg14864041 | Non-CpG Island |                      | Intergenic    | 0.030918256  | 1.65E-02  | 0.045293071  | 2.63E-05 | 2.49E-02 | 1.00 | 693 | 1 |
| cg04114034 | Non-CpG Island |                      | Intergenic    | -0.047110302 | -6.90E-02 | -0.025203062 | 2.64E-05 | 2.50E-02 | 1.00 | 694 | 1 |
| cg04413584 | Shelf          |                      | Intergenic    | -0.0313796   | -4.60E-02 | -0.016784362 | 2.65E-05 | 2.51E-02 | 1.00 | 695 | 1 |
| cg03231199 | Non-CpG Island |                      | Intergenic    | 0.064922233  | 3.47E-02  | 0.095121407  | 2.66E-05 | 2.51E-02 | 1.00 | 696 | 1 |
| cg00984392 | Non-CpG Island | <i>PLXNA4</i>        | Body          | -0.033326273 | -4.88E-02 | -0.017822075 | 2.66E-05 | 2.51E-02 | 1.00 | 697 | 1 |
| cg26621780 | Shelf          | <i>RASAL1</i>        | Body          | -0.033048597 | -4.84E-02 | -0.017672787 | 2.67E-05 | 2.51E-02 | 1.00 | 698 | 1 |
| cg11652343 | Non-CpG Island |                      | Intergenic    | -0.030880033 | -4.52E-02 | -0.016511419 | 2.67E-05 | 2.51E-02 | 1.00 | 699 | 1 |
| cg13012015 | Non-CpG Island | <i>AC104820.2</i>    | 5' UTR        | -0.028093711 | -4.12E-02 | -0.015020128 | 2.68E-05 | 2.51E-02 | 1.00 | 700 | 1 |
| cg22345623 | Non-CpG Island | <i>MRRF</i>          | Body          | -0.039414999 | -5.78E-02 | -0.021070714 | 2.68E-05 | 2.51E-02 | 1.00 | 701 | 1 |
| cg23120547 | Non-CpG Island | <i>ARHGAP22</i>      | Body          | -0.022540016 | -3.30E-02 | -0.012049095 | 2.69E-05 | 2.51E-02 | 1.00 | 702 | 1 |
| cg12062266 | Non-CpG Island |                      | Intergenic    | 0.04205931   | 2.25E-02  | 0.061635552  | 2.69E-05 | 2.51E-02 | 1.00 | 599 | 2 |
| cg00617180 | Non-CpG Island |                      | Intergenic    | -0.029949896 | -4.39E-02 | -0.016007053 | 2.70E-05 | 2.52E-02 | 1.00 | 703 | 1 |
| cg05733575 | Shore          |                      | Intergenic    | -0.042651876 | -6.25E-02 | -0.022784558 | 2.73E-05 | 2.54E-02 | 1.00 | 704 | 1 |
| cg15088991 | Shelf          | <i>PIGV</i>          | Body          | 0.024399283  | 1.30E-02  | 0.035768023  | 2.74E-05 | 2.55E-02 | 1.00 | 705 | 1 |
| cg01052993 | Non-CpG Island | <i>MGAM</i>          | 5' UTR        | -0.029798253 | -4.37E-02 | -0.015912907 | 2.74E-05 | 2.55E-02 | 1.00 | 706 | 1 |
| cg08390622 | Non-CpG Island | <i>MIR4716</i>       | TSS1500       | -0.032010995 | -4.69E-02 | -0.017093088 | 2.75E-05 | 2.55E-02 | 1.00 | 707 | 1 |

## Continuation of Table S 3.2

|            |                |                     |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg03173169 | Island         | <i>HK2</i>          | 5' UTR     | -0.027697315 | -4.06E-02 | -0.014789455 | 2.75E-05 | 2.55E-02 | 1.00 | 708 | 1 |
| cg08106907 | Non-CpG Island |                     | Intergenic | -0.022179911 | -3.25E-02 | -0.011842556 | 2.75E-05 | 2.55E-02 | 1.00 | 709 | 1 |
| cg16617774 | Non-CpG Island | <i>PPP2R2B</i>      | 5' UTR     | -0.027705463 | -4.06E-02 | -0.014791442 | 2.76E-05 | 2.55E-02 | 1.00 | 710 | 1 |
| cg08712483 | Non-CpG Island |                     | Intergenic | 0.069004118  | 3.68E-02  | 0.101172879  | 2.77E-05 | 2.56E-02 | 1.00 | 711 | 1 |
| cg04286826 | Non-CpG Island | <i>LOC101927070</i> | Body       | -0.038865485 | -5.70E-02 | -0.02074526  | 2.77E-05 | 2.56E-02 | 1.00 | 279 | 2 |
| cg04962528 | Shelf          |                     | Intergenic | -0.044846379 | -6.58E-02 | -0.023937363 | 2.77E-05 | 2.56E-02 | 1.00 | 712 | 1 |
| cg27246364 | Non-CpG Island | <i>ZNF19</i>        | 5' UTR     | 0.02360575   | 1.26E-02  | 0.034612185  | 2.78E-05 | 2.56E-02 | 1.00 | 713 | 1 |
| cg05702597 | Non-CpG Island | <i>MME</i>          | Body       | -0.039844863 | -5.84E-02 | -0.021265246 | 2.78E-05 | 2.56E-02 | 1.00 | 714 | 1 |
| cg11312836 | Non-CpG Island | <i>FMO6P</i>        | Body       | -0.018977768 | -2.78E-02 | -0.010127536 | 2.79E-05 | 2.56E-02 | 1.00 | 715 | 1 |
| cg18400281 | Non-CpG Island | <i>KIAA1217</i>     | Body       | 0.037878833  | 2.02E-02  | 0.055544081  | 2.79E-05 | 2.56E-02 | 1.00 | 716 | 1 |
| cg01041061 | Non-CpG Island |                     | Intergenic | -0.044116924 | -6.47E-02 | -0.023541344 | 2.79E-05 | 2.56E-02 | 1.00 | 717 | 1 |
| cg05788267 | Non-CpG Island | <i>SEMA3C</i>       | Body       | -0.03733477  | -5.47E-02 | -0.019919717 | 2.80E-05 | 2.56E-02 | 1.00 | 718 | 1 |
| cg00365642 | Non-CpG Island | <i>LINC00299</i>    | Body       | -0.056015064 | -8.22E-02 | -0.02987806  | 2.81E-05 | 2.57E-02 | 1.00 | 719 | 1 |
| cg25444821 | Shore          | <i>KCNK3</i>        | 3' UTR     | -0.037403592 | -5.49E-02 | -0.019950349 | 2.81E-05 | 2.57E-02 | 1.00 | 720 | 1 |
| cg07250758 | Non-CpG Island | <i>VWCE</i>         | Body       | -0.048835632 | -7.16E-02 | -0.026042586 | 2.83E-05 | 2.58E-02 | 1.00 | 512 | 2 |
| cg27541551 | Non-CpG Island | <i>AC093171.1</i>   | TSS1500    | -0.026871294 | -3.94E-02 | -0.014328198 | 2.83E-05 | 2.58E-02 | 1.00 | 721 | 1 |
| cg02483931 | Non-CpG Island | <i>CCL5</i>         | Body       | -0.02816882  | -4.13E-02 | -0.015019292 | 2.84E-05 | 2.58E-02 | 1.00 | 722 | 1 |
| cg19560945 | Non-CpG Island | <i>LINC00939</i>    | Body       | 0.039870042  | 2.13E-02  | 0.058483358  | 2.84E-05 | 2.58E-02 | 1.00 | 723 | 1 |
| cg27549887 | Island         | <i>DUT</i>          | TSS1500    | 0.03636272   | 1.94E-02  | 0.053342062  | 2.85E-05 | 2.59E-02 | 1.00 | 724 | 1 |
| cg12995604 | Shore          | <i>ZBTB7A</i>       | Body       | -0.025595926 | -3.76E-02 | -0.013634394 | 2.89E-05 | 2.62E-02 | 1.00 | 725 | 1 |
| cg11364889 | Non-CpG Island |                     | Intergenic | -0.034305068 | -5.03E-02 | -0.018271017 | 2.90E-05 | 2.62E-02 | 1.00 | 726 | 1 |
| cg03990746 | Non-CpG Island |                     | Intergenic | -0.024985997 | -3.67E-02 | -0.013307358 | 2.90E-05 | 2.62E-02 | 1.00 | 727 | 1 |
| cg21212234 | Non-CpG Island |                     | Intergenic | 0.042517706  | 2.26E-02  | 0.062391671  | 2.91E-05 | 2.62E-02 | 1.00 | 728 | 1 |
| cg03485217 | Non-CpG Island | <i>CLNK</i>         | First Exon | -0.032818938 | -4.82E-02 | -0.01747328  | 2.92E-05 | 2.63E-02 | 1.00 | 729 | 1 |
| cg22300252 | Non-CpG Island | <i>SEMA6D</i>       | 5' UTR     | 0.044096582  | 2.35E-02  | 0.06472145   | 2.94E-05 | 2.64E-02 | 1.00 | 730 | 1 |
| cg26624021 | Non-CpG Island | <i>CETP</i>         | TSS200     | -0.050916061 | -7.47E-02 | -0.027095197 | 2.95E-05 | 2.65E-02 | 1.00 | 731 | 1 |
| cg03281572 | Non-CpG Island | <i>VTI1A</i>        | Body       | -0.046796468 | -6.87E-02 | -0.024898052 | 2.97E-05 | 2.66E-02 | 1.00 | 732 | 1 |
| cg15000547 | Non-CpG Island |                     | Intergenic | -0.04401874  | -6.46E-02 | -0.023408088 | 3.00E-05 | 2.68E-02 | 1.00 | 733 | 1 |
| cg05061804 | Shore          | <i>TTF1</i>         | TSS200     | 0.026777916  | 1.42E-02  | 0.039316164  | 3.00E-05 | 2.68E-02 | 1.00 | 734 | 1 |
| cg08603845 | Shelf          | <i>GRHL3</i>        | TSS1500    | -0.028469969 | -4.18E-02 | -0.015138566 | 3.00E-05 | 2.68E-02 | 1.00 | 735 | 1 |
| cg05064853 | Non-CpG Island | <i>C5orf64</i>      | 3' UTR     | -0.022200893 | -3.26E-02 | -0.011799066 | 3.03E-05 | 2.70E-02 | 1.00 | 736 | 1 |
| cg11818303 | Non-CpG Island | <i>EPB41L1</i>      | Body       | -0.024104677 | -3.54E-02 | -0.012810467 | 3.03E-05 | 2.70E-02 | 1.00 | 737 | 1 |
| cg21291672 | Island         | <i>TMEM200B</i>     | TSS1500    | -0.025877137 | -3.80E-02 | -0.013751919 | 3.04E-05 | 2.70E-02 | 1.00 | 738 | 1 |
| cg11667451 | Island         | <i>TOX2</i>         | TSS1500    | 0.028091935  | 1.49E-02  | 0.041254943  | 3.04E-05 | 2.70E-02 | 1.00 | 739 | 1 |
| cg18848606 | Non-CpG Island |                     | Intergenic | -0.028510367 | -4.19E-02 | -0.015145786 | 3.06E-05 | 2.72E-02 | 1.00 | 740 | 1 |
| cg08790320 | Shore          | <i>ITGBL1</i>       | 5' UTR     | 0.061860035  | 3.29E-02  | 0.090865867  | 3.08E-05 | 2.73E-02 | 1.00 | 149 | 2 |
| cg01287555 | Shore          | <i>KBTBD6</i>       | First Exon | 0.025598747  | 1.36E-02  | 0.037602761  | 3.08E-05 | 2.73E-02 | 1.00 | 741 | 1 |
| cg01023548 | Non-CpG Island |                     | Intergenic | -0.034631989 | -5.09E-02 | -0.018389793 | 3.09E-05 | 2.73E-02 | 1.00 | 742 | 1 |
| cg18212920 | Non-CpG Island |                     | Intergenic | 0.028374189  | 1.51E-02  | 0.041686633  | 3.11E-05 | 2.75E-02 | 1.00 | 743 | 1 |



## Continuation of Table S 3.2

|            |                |                     |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg19703023 | Non-CpG Island |                     | Intergenic | -0.030099679 | -4.42E-02 | -0.015976183 | 3.12E-05 | 2.75E-02 | 1.00 | 744 | 1 |
| cg23204810 | Non-CpG Island | <i>TCF7L1</i>       | Body       | -0.033373383 | -4.90E-02 | -0.017713035 | 3.12E-05 | 2.75E-02 | 1.00 | 745 | 1 |
| cg24028809 | Non-CpG Island | <i>RGS17</i>        | 5' UTR     | 0.040889702  | 2.17E-02  | 0.06007816   | 3.12E-05 | 2.75E-02 | 1.00 | 746 | 1 |
| cg12245084 | Shore          | <i>ODZ4</i>         | Body       | -0.025202218 | -3.70E-02 | -0.013374226 | 3.13E-05 | 2.75E-02 | 1.00 | 747 | 1 |
| cg24931191 | Non-CpG Island |                     | Intergenic | -0.041363147 | -6.08E-02 | -0.021948394 | 3.13E-05 | 2.75E-02 | 1.00 | 748 | 1 |
| cg09765245 | Shore          | <i>RNF149</i>       | TSS1500    | -0.029239932 | -4.30E-02 | -0.015515373 | 3.13E-05 | 2.75E-02 | 1.00 | 749 | 1 |
| cg16421283 | Non-CpG Island | <i>LINC01221</i>    | Body       | -0.029374077 | -4.32E-02 | -0.015584963 | 3.14E-05 | 2.75E-02 | 1.00 | 750 | 1 |
| cg19025470 | Non-CpG Island | <i>ZNF787</i>       | 5' UTR     | -0.037498658 | -5.51E-02 | -0.019894921 | 3.14E-05 | 2.75E-02 | 1.00 | 751 | 1 |
| cg08499559 | Non-CpG Island | <i>SMAD1</i>        | Body       | -0.026725501 | -3.93E-02 | -0.014177673 | 3.15E-05 | 2.75E-02 | 1.00 | 752 | 1 |
| cg19110943 | Non-CpG Island |                     | Intergenic | -0.023840362 | -3.50E-02 | -0.012646644 | 3.15E-05 | 2.75E-02 | 1.00 | 350 | 2 |
| cg18632152 | Non-CpG Island | <i>GSDMA</i>        | Body       | -0.027249856 | -4.00E-02 | -0.014454809 | 3.15E-05 | 2.75E-02 | 1.00 | 753 | 1 |
| cg18940157 | Non-CpG Island | <i>PACS2</i>        | Body       | -0.018693049 | -2.75E-02 | -0.009915302 | 3.16E-05 | 2.75E-02 | 1.00 | 754 | 1 |
| cg23493133 | Shelf          |                     | Intergenic | -0.033521576 | -4.93E-02 | -0.017778296 | 3.17E-05 | 2.76E-02 | 1.00 | 755 | 1 |
| cg07757007 | Non-CpG Island | <i>HNMT</i>         | Body       | -0.04373963  | -6.43E-02 | -0.023192246 | 3.18E-05 | 2.76E-02 | 1.00 | 756 | 1 |
| cg17942388 | Non-CpG Island | <i>CLDN25</i>       | TSS1500    | -0.034009298 | -5.00E-02 | -0.01803266  | 3.18E-05 | 2.76E-02 | 1.00 | 757 | 1 |
| cg09308842 | Island         | <i>TRIM67</i>       | TSS1500    | 0.03054411   | 1.62E-02  | 0.044900015  | 3.21E-05 | 2.79E-02 | 1.00 | 758 | 1 |
| cg03833948 | Shelf          | <i>C12orf39</i>     | TSS1500    | 0.027741152  | 1.47E-02  | 0.040781378  | 3.22E-05 | 2.79E-02 | 1.00 | 759 | 1 |
| cg00724111 | Shore          | <i>PACSIN3</i>      | TSS1500    | -0.041157355 | -6.05E-02 | -0.021808854 | 3.22E-05 | 2.79E-02 | 1.00 | 760 | 1 |
| cg08409615 | Non-CpG Island | <i>STK32B</i>       | Body       | 0.027513828  | 1.46E-02  | 0.040450512  | 3.23E-05 | 2.79E-02 | 1.00 | 761 | 1 |
| cg04320150 | Non-CpG Island |                     | Intergenic | -0.047372538 | -6.96E-02 | -0.025095576 | 3.24E-05 | 2.80E-02 | 1.00 | 762 | 1 |
| cg21466789 | Non-CpG Island | <i>C8orf75</i>      | 3' UTR     | 0.049110152  | 2.60E-02  | 0.072211707  | 3.26E-05 | 2.81E-02 | 1.00 | 763 | 1 |
| cg07475825 | Non-CpG Island | <i>CYBR4</i>        | Body       | -0.051153659 | -7.52E-02 | -0.027072585 | 3.30E-05 | 2.85E-02 | 1.00 | 764 | 1 |
| cg09767076 | Shelf          | <i>ATP6V0A2</i>     | Body       | -0.05679071  | -8.35E-02 | -0.03005308  | 3.31E-05 | 2.85E-02 | 1.00 | 765 | 1 |
| cg19998151 | Non-CpG Island |                     | Intergenic | -0.068826074 | -1.01E-01 | -0.036420367 | 3.31E-05 | 2.85E-02 | 1.00 | 766 | 1 |
| cg25343965 | Non-CpG Island | <i>NCAM1</i>        | Body       | -0.023439973 | -3.45E-02 | -0.012402714 | 3.32E-05 | 2.85E-02 | 1.00 | 767 | 1 |
| cg24055030 | Non-CpG Island | <i>KIAA1671</i>     | Body       | -0.052668137 | -7.75E-02 | -0.027862872 | 3.33E-05 | 2.86E-02 | 1.00 | 768 | 1 |
| cg22022957 | Shore          | <i>CDKN1C</i>       | TSS1500    | -0.043776076 | -6.44E-02 | -0.023156681 | 3.34E-05 | 2.86E-02 | 1.00 | 769 | 1 |
| cg08692903 | Non-CpG Island |                     | Intergenic | -0.018269837 | -2.69E-02 | -0.009663269 | 3.35E-05 | 2.86E-02 | 1.00 | 770 | 1 |
| cg09081665 | Non-CpG Island | <i>COQ2</i>         | Body       | 0.028917966  | 1.53E-02  | 0.042542628  | 3.35E-05 | 2.87E-02 | 1.00 | 771 | 1 |
| cg03871754 | Island         |                     | Intergenic | -0.026166426 | -3.85E-02 | -0.013835722 | 3.37E-05 | 2.87E-02 | 1.00 | 772 | 1 |
| cg21018834 | Non-CpG Island |                     | Intergenic | -0.025781763 | -3.79E-02 | -0.013628238 | 3.39E-05 | 2.89E-02 | 1.00 | 773 | 1 |
| cg13187651 | Shore          | <i>COASY</i>        | TSS1500    | -0.050620756 | -7.45E-02 | -0.026755285 | 3.39E-05 | 2.89E-02 | 1.00 | 774 | 1 |
| cg20212624 | Shelf          | <i>CNP</i>          | Body       | 0.021603833  | 1.14E-02  | 0.031790596  | 3.40E-05 | 2.89E-02 | 1.00 | 775 | 1 |
| cg04328144 | Non-CpG Island | <i>LMBRD1</i>       | Body       | -0.02635576  | -3.88E-02 | -0.01392636  | 3.41E-05 | 2.90E-02 | 1.00 | 776 | 1 |
| cg01427126 | Non-CpG Island | <i>TRAP1</i>        | Body       | -0.029534026 | -4.35E-02 | -0.015603405 | 3.42E-05 | 2.90E-02 | 1.00 | 777 | 1 |
| cg19987923 | Non-CpG Island | <i>RP5-881L22.5</i> | 3' UTR     | -0.034631986 | -5.10E-02 | -0.01828962  | 3.45E-05 | 2.92E-02 | 1.00 | 778 | 1 |
| cg01442064 | Island         | <i>EVC</i>          | Body       | 0.0261658    | 1.38E-02  | 0.038516025  | 3.46E-05 | 2.93E-02 | 1.00 | 779 | 2 |
| cg27598407 | Non-CpG Island | <i>ATF7IP</i>       | 5' UTR     | -0.043011896 | -6.33E-02 | -0.022704964 | 3.48E-05 | 2.94E-02 | 1.00 | 780 | 1 |
| cg15695956 | Non-CpG Island |                     | Intergenic | 0.021033157  | 1.11E-02  | 0.030964999  | 3.49E-05 | 2.95E-02 | 1.00 | 781 | 1 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg04891961 | Island         | <i>ANKRD13B</i>      | Body       | -0.036066493 | -5.31E-02 | -0.019035195 | 3.49E-05 | 2.95E-02 | 1.00 | 299 | 2 |
| cg12295374 | Non-CpG Island |                      | Intergenic | -0.036128829 | -5.32E-02 | -0.019062416 | 3.51E-05 | 2.96E-02 | 1.00 | 782 | 1 |
| cg05874698 | Non-CpG Island |                      | Intergenic | 0.031693307  | 1.67E-02  | 0.046665799  | 3.52E-05 | 2.96E-02 | 1.00 | 783 | 1 |
| cg11711040 | Non-CpG Island |                      | Intergenic | 0.029847951  | 1.57E-02  | 0.043955404  | 3.55E-05 | 2.98E-02 | 1.00 | 784 | 1 |
| cg11530006 | Non-CpG Island |                      | Intergenic | -0.033490034 | -4.93E-02 | -0.017655909 | 3.57E-05 | 3.00E-02 | 1.00 | 785 | 1 |
| cg05663480 | Non-CpG Island | <i>NFASC</i>         | 5' UTR     | -0.036188255 | -5.33E-02 | -0.019074807 | 3.58E-05 | 3.01E-02 | 1.00 | 786 | 1 |
| cg08586855 | Non-CpG Island | <i>SPSB1</i>         | 5' UTR     | -0.04982982  | -7.34E-02 | -0.026260224 | 3.60E-05 | 3.01E-02 | 1.00 | 787 | 1 |
| cg25887355 | Non-CpG Island | <i>SYN3</i>          | Body       | -0.034274092 | -5.05E-02 | -0.018059808 | 3.61E-05 | 3.02E-02 | 1.00 | 788 | 1 |
| cg07964115 | Shelf          | <i>NINJ2</i>         | Body       | -0.034197012 | -5.04E-02 | -0.018011703 | 3.64E-05 | 3.04E-02 | 1.00 | 789 | 1 |
| cg01290904 | Shore          | <i>EVC2</i>          | 5' UTR     | 0.030505712  | 1.61E-02  | 0.044947898  | 3.66E-05 | 3.05E-02 | 1.00 | 779 | 2 |
| cg03815267 | Non-CpG Island | <i>SEMA5A</i>        | Body       | -0.033605117 | -4.95E-02 | -0.01769412  | 3.66E-05 | 3.05E-02 | 1.00 | 790 | 1 |
| cg05930965 | Non-CpG Island |                      | Intergenic | -0.032625486 | -4.81E-02 | -0.017177664 | 3.67E-05 | 3.05E-02 | 1.00 | 791 | 1 |
| cg27638835 | Non-CpG Island | <i>AXIN1</i>         | Body       | -0.028414664 | -4.19E-02 | -0.014960164 | 3.67E-05 | 3.05E-02 | 1.00 | 792 | 1 |
| cg09235426 | Shelf          | <i>LRP1</i>          | Body       | -0.025003331 | -3.68E-02 | -0.013162776 | 3.67E-05 | 3.05E-02 | 1.00 | 793 | 1 |
| cg19839325 | Shore          | <i>KLHL26</i>        | Body       | -0.026593313 | -3.92E-02 | -0.013994244 | 3.70E-05 | 3.07E-02 | 1.00 | 794 | 1 |
| cg05218995 | Non-CpG Island | <i>TMPRSS2</i>       | Body       | -0.028603889 | -4.22E-02 | -0.015051221 | 3.71E-05 | 3.07E-02 | 1.00 | 795 | 1 |
| cg19339407 | Shore          | <i>KRT85</i>         | Body       | -0.022764702 | -3.36E-02 | -0.011976718 | 3.72E-05 | 3.08E-02 | 1.00 | 796 | 1 |
| cg26750073 | Non-CpG Island | <i>RP11-305P14.1</i> | 5' UTR     | 0.044599148  | 2.35E-02  | 0.065740356  | 3.74E-05 | 3.09E-02 | 1.00 | 797 | 1 |
| cg07692435 | Non-CpG Island | <i>7SK</i>           | TSS1500    | -0.037959232 | -5.60E-02 | -0.019965354 | 3.74E-05 | 3.09E-02 | 1.00 | 798 | 1 |
| cg04582369 | Non-CpG Island | <i>ITPR1</i>         | Body       | -0.036380645 | -5.36E-02 | -0.019131122 | 3.76E-05 | 3.10E-02 | 1.00 | 799 | 1 |
| cg11955249 | Non-CpG Island |                      | Intergenic | 0.02728747   | 1.43E-02  | 0.040226427  | 3.76E-05 | 3.10E-02 | 1.00 | 800 | 1 |
| cg02409479 | Shore          |                      | Intergenic | -0.054040141 | -7.97E-02 | -0.02841437  | 3.76E-05 | 3.10E-02 | 1.00 | 801 | 1 |
| cg08263398 | Shore          | <i>B3GNT3</i>        | Body       | -0.032329889 | -4.77E-02 | -0.016991503 | 3.80E-05 | 3.12E-02 | 1.00 | 802 | 1 |
| cg05178561 | Non-CpG Island | <i>DOCK8</i>         | Body       | -0.033349632 | -4.92E-02 | -0.017522887 | 3.82E-05 | 3.13E-02 | 1.00 | 803 | 1 |
| cg21748207 | Shore          | <i>FAM59B</i>        | TSS1500    | -0.036558964 | -5.39E-02 | -0.019203514 | 3.84E-05 | 3.15E-02 | 1.00 | 804 | 1 |
| cg14575656 | Non-CpG Island |                      | Intergenic | -0.041440401 | -6.11E-02 | -0.021763315 | 3.85E-05 | 3.16E-02 | 1.00 | 405 | 2 |
| cg23398508 | Shore          | <i>AC007639.1</i>    | 3' UTR     | -0.028686191 | -4.23E-02 | -0.015063635 | 3.86E-05 | 3.16E-02 | 1.00 | 805 | 1 |
| cg01564343 | Non-CpG Island | <i>TREML1</i>        | Body       | -0.032937333 | -4.86E-02 | -0.017293128 | 3.87E-05 | 3.17E-02 | 1.00 | 806 | 1 |
| cg16505573 | Non-CpG Island |                      | Intergenic | -0.030621739 | -4.52E-02 | -0.016076512 | 3.88E-05 | 3.17E-02 | 1.00 | 807 | 1 |
| cg10759796 | Non-CpG Island | <i>MTA3</i>          | 5' UTR     | -0.024406542 | -3.60E-02 | -0.012812507 | 3.88E-05 | 3.17E-02 | 1.00 | 808 | 1 |
| cg15183692 | Non-CpG Island | <i>RHOH</i>          | First Exon | 0.037878153  | 1.99E-02  | 0.055872712  | 3.89E-05 | 3.17E-02 | 1.00 | 809 | 1 |
| cg20552296 | Shore          | <i>HPSE2</i>         | First Exon | 0.031368207  | 1.65E-02  | 0.046270476  | 3.89E-05 | 3.17E-02 | 1.00 | 810 | 1 |
| cg03986427 | Non-CpG Island | <i>TULP2</i>         | Body       | -0.043796758 | -6.46E-02 | -0.022987605 | 3.90E-05 | 3.17E-02 | 1.00 | 811 | 1 |
| cg09866569 | Non-CpG Island | <i>KCTD5</i>         | Body       | 0.037691654  | 1.98E-02  | 0.055604755  | 3.92E-05 | 3.18E-02 | 1.00 | 812 | 1 |
| cg11854860 | Shelf          | <i>PPP2R2C</i>       | Body       | -0.04136062  | -6.10E-02 | -0.021700712 | 3.93E-05 | 3.18E-02 | 1.00 | 813 | 1 |
| cg07884459 | Non-CpG Island | <i>SMYD3</i>         | Body       | -0.027212039 | -4.01E-02 | -0.014276497 | 3.93E-05 | 3.18E-02 | 1.00 | 814 | 1 |
| cg13409439 | Shore          | <i>NTNG2</i>         | 5' UTR     | -0.040637877 | -6.00E-02 | -0.021312743 | 3.96E-05 | 3.20E-02 | 1.00 | 815 | 1 |
| cg21995800 | Non-CpG Island | <i>WDR77</i>         | Body       | 0.027354784  | 1.43E-02  | 0.040372218  | 4.01E-05 | 3.24E-02 | 1.00 | 816 | 1 |
| cg23680518 | Shelf          | <i>SBSN</i>          | TSS1500    | -0.039971717 | -5.90E-02 | -0.020949868 | 4.01E-05 | 3.24E-02 | 1.00 | 817 | 2 |

# Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg14145614 | Non-CpG Island | <i>LPP</i>           | 5' UTR     | 0.025276765  | 1.32E-02  | 0.037312846  | 4.05E-05 | 3.27E-02 | 1.00 | 818 | 1 |
| cg14580608 | Non-CpG Island | <i>SMCR2</i>         | Body       | -0.019858953 | -2.93E-02 | -0.010400728 | 4.07E-05 | 3.28E-02 | 1.00 | 819 | 1 |
| cg16239257 | Island         | <i>LTBP2</i>         | Body       | 0.024045256  | 1.26E-02  | 0.035498462  | 4.07E-05 | 3.28E-02 | 1.00 | 820 | 1 |
| cg20228731 | Non-CpG Island | <i>FLJ43663</i>      | Body       | -0.067332861 | -9.94E-02 | -0.035256983 | 4.08E-05 | 3.28E-02 | 1.00 | 821 | 1 |
| cg09476130 | Island         | <i>CCDC19</i>        | TSS200     | -0.035571625 | -5.25E-02 | -0.018626048 | 4.08E-05 | 3.28E-02 | 1.00 | 206 | 3 |
| cg09080114 | Shore          | <i>SPSB2</i>         | TSS1500    | -0.036485637 | -5.39E-02 | -0.019102809 | 4.09E-05 | 3.28E-02 | 1.00 | 822 | 1 |
| cg18994250 | Shore          | <i>JAKMIP1</i>       | Body       | -0.048394093 | -7.15E-02 | -0.025332352 | 4.11E-05 | 3.29E-02 | 1.00 | 629 | 2 |
| cg20327845 | Non-CpG Island | <i>PFKP</i>          | Body       | -0.031479505 | -4.65E-02 | -0.016474816 | 4.12E-05 | 3.30E-02 | 1.00 | 823 | 1 |
| cg11839163 | Non-CpG Island | <i>CALD1</i>         | Body       | -0.046960946 | -6.93E-02 | -0.024575278 | 4.13E-05 | 3.30E-02 | 1.00 | 824 | 1 |
| cg22320795 | Non-CpG Island | <i>CLEC19A</i>       | Body       | 0.031583895  | 1.65E-02  | 0.046644454  | 4.15E-05 | 3.31E-02 | 1.00 | 825 | 1 |
| cg25107667 | Non-CpG Island | <i>RP4-723E3.1</i>   | 5' UTR     | 0.052616952  | 2.75E-02  | 0.077709099  | 4.16E-05 | 3.31E-02 | 1.00 | 826 | 1 |
| cg11545768 | Non-CpG Island | <i>PRKCA</i>         | Body       | 0.021705661  | 1.14E-02  | 0.032056993  | 4.16E-05 | 3.31E-02 | 1.00 | 827 | 1 |
| cg24990672 | Shore          | <i>DGKG</i>          | 5' UTR     | -0.020771374 | -3.07E-02 | -0.010865416 | 4.16E-05 | 3.31E-02 | 1.00 | 315 | 2 |
| cg12580228 | Non-CpG Island | <i>NDUFAF7</i>       | Body       | 0.02584158   | 1.35E-02  | 0.038166808  | 4.17E-05 | 3.31E-02 | 1.00 | 828 | 1 |
| cg26931208 | Shore          | <i>IZUMO1</i>        | Body       | -0.035990413 | -5.32E-02 | -0.018823802 | 4.17E-05 | 3.31E-02 | 1.00 | 47  | 3 |
| cg08012287 | Shore          | <i>ACTN3</i>         | Body       | -0.023628804 | -3.49E-02 | -0.012357076 | 4.18E-05 | 3.31E-02 | 1.00 | 829 | 1 |
| cg00096898 | Non-CpG Island | <i>XYLB</i>          | Body       | -0.036267689 | -5.36E-02 | -0.018966639 | 4.18E-05 | 3.31E-02 | 1.00 | 830 | 1 |
| cg24302752 | Non-CpG Island | <i>BCKDHB</i>        | Body       | -0.045003319 | -6.65E-02 | -0.023535035 | 4.18E-05 | 3.31E-02 | 1.00 | 831 | 1 |
| cg27325787 | Non-CpG Island | <i>PRKCB</i>         | Body       | -0.033488636 | -4.95E-02 | -0.017510908 | 4.19E-05 | 3.32E-02 | 1.00 | 832 | 1 |
| cg18733931 | Non-CpG Island |                      | Intergenic | -0.036234295 | -5.35E-02 | -0.018938911 | 4.23E-05 | 3.34E-02 | 1.00 | 833 | 1 |
| cg12065777 | Shore          |                      | Intergenic | -0.021004258 | -3.10E-02 | -0.010974451 | 4.26E-05 | 3.36E-02 | 1.00 | 834 | 1 |
| cg19142740 | Non-CpG Island | <i>CADM4</i>         | Body       | -0.029047381 | -4.29E-02 | -0.015176274 | 4.26E-05 | 3.36E-02 | 1.00 | 835 | 1 |
| cg03460109 | Non-CpG Island | <i>SORCS2</i>        | Body       | -0.024334111 | -3.60E-02 | -0.012712437 | 4.27E-05 | 3.36E-02 | 1.00 | 836 | 1 |
| cg14590560 | Non-CpG Island | <i>GNA15</i>         | Body       | -0.021079598 | -3.11E-02 | -0.011010626 | 4.28E-05 | 3.36E-02 | 1.00 | 837 | 1 |
| cg05202858 | Non-CpG Island |                      | Intergenic | 0.073534824  | 3.84E-02  | 0.108662024  | 4.29E-05 | 3.36E-02 | 1.00 | 838 | 1 |
| cg20220458 | Non-CpG Island | <i>CDC42BPB</i>      | Body       | -0.025436606 | -3.76E-02 | -0.013285403 | 4.29E-05 | 3.36E-02 | 1.00 | 839 | 1 |
| cg19024056 | Non-CpG Island | <i>MMRN1</i>         | Body       | -0.031089932 | -4.59E-02 | -0.016234592 | 4.30E-05 | 3.37E-02 | 1.00 | 840 | 1 |
| cg23513000 | Non-CpG Island | <i>TMEM51</i>        | 5' UTR     | 0.032184326  | 1.68E-02  | 0.047565295  | 4.32E-05 | 3.37E-02 | 1.00 | 841 | 1 |
| cg04250661 | Non-CpG Island | <i>RAPGEF2</i>       | Body       | -0.044203145 | -6.53E-02 | -0.023077385 | 4.32E-05 | 3.37E-02 | 1.00 | 842 | 1 |
| cg01021271 | Non-CpG Island | <i>RP11-179A16.1</i> | 3' UTR     | 0.080671373  | 4.21E-02  | 0.119226969  | 4.32E-05 | 3.37E-02 | 1.00 | 843 | 1 |
| cg04136369 | Non-CpG Island | <i>HNF1B</i>         | Body       | -0.029049941 | -4.29E-02 | -0.01516577  | 4.32E-05 | 3.37E-02 | 1.00 | 844 | 1 |
| cg18610541 | Non-CpG Island | <i>KIAA1279</i>      | Body       | 0.025736328  | 1.34E-02  | 0.038040775  | 4.35E-05 | 3.39E-02 | 1.00 | 845 | 1 |
| cg18255353 | Island         |                      | Intergenic | 0.041496047  | 2.17E-02  | 0.06133806   | 4.36E-05 | 3.39E-02 | 1.00 | 846 | 1 |
| cg13468041 | Shore          | <i>CXCL3</i>         | 3' UTR     | -0.028651965 | -4.24E-02 | -0.014944424 | 4.40E-05 | 3.42E-02 | 1.00 | 847 | 1 |
| cg05404728 | Non-CpG Island | <i>CCDC33</i>        | Body       | -0.027665903 | -4.09E-02 | -0.014425301 | 4.43E-05 | 3.44E-02 | 1.00 | 848 | 1 |
| cg14029057 | Non-CpG Island | <i>RN5S208</i>       | TSS200     | -0.04895712  | -7.24E-02 | -0.02552512  | 4.43E-05 | 3.44E-02 | 1.00 | 849 | 1 |
| cg09134819 | Shelf          |                      | Intergenic | -0.035163444 | -5.20E-02 | -0.018331504 | 4.44E-05 | 3.44E-02 | 1.00 | 850 | 1 |
| cg11962116 | Non-CpG Island |                      | Intergenic | -0.024344118 | -3.60E-02 | -0.012689211 | 4.45E-05 | 3.45E-02 | 1.00 | 851 | 1 |
| cg10682119 | Non-CpG Island | <i>FAM174B</i>       | Body       | -0.07884399  | -1.17E-01 | -0.041094324 | 4.46E-05 | 3.45E-02 | 1.00 | 852 | 1 |

## Continuation of Table S 3.2

|            |                |                     |            |              |           |              |          |          |      |     |   |
|------------|----------------|---------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg16760981 | Shelf          |                     | Intergenic | -0.030107943 | -4.45E-02 | -0.015687641 | 4.49E-05 | 3.47E-02 | 1.00 | 853 | 1 |
| cg24184634 | Non-CpG Island | <i>IL12A-AS1</i>    | Body       | 0.029173886  | 1.52E-02  | 0.043151257  | 4.51E-05 | 3.48E-02 | 1.00 | 854 | 1 |
| cg12263409 | Non-CpG Island |                     | Intergenic | -0.03854178  | -5.70E-02 | -0.020075339 | 4.52E-05 | 3.48E-02 | 1.00 | 855 | 1 |
| cg15363332 | Non-CpG Island |                     | Intergenic | 0.047082117  | 2.45E-02  | 0.069641811  | 4.52E-05 | 3.48E-02 | 1.00 | 856 | 1 |
| cg05469118 | Non-CpG Island | <i>RP11-820L6.1</i> | 3' UTR     | -0.05193882  | -7.68E-02 | -0.027043461 | 4.55E-05 | 3.50E-02 | 1.00 | 857 | 1 |
| cg12384562 | Non-CpG Island | <i>PARD3B</i>       | Body       | -0.039982325 | -5.91E-02 | -0.020816134 | 4.55E-05 | 3.50E-02 | 1.00 | 858 | 1 |
| cg11115502 | Island         | <i>SF3B14</i>       | TSS200     | 0.043060554  | 2.24E-02  | 0.063725043  | 4.64E-05 | 3.56E-02 | 1.00 | 859 | 1 |
| cg12613020 | Shore          |                     | Intergenic | -0.038650543 | -5.72E-02 | -0.020099532 | 4.65E-05 | 3.57E-02 | 1.00 | 388 | 2 |
| cg06984107 | Non-CpG Island | <i>PSME4</i>        | Body       | -0.047329752 | -7.00E-02 | -0.024612584 | 4.66E-05 | 3.57E-02 | 1.00 | 860 | 1 |
| cg11628564 | Non-CpG Island |                     | Intergenic | 0.024513003  | 1.27E-02  | 0.036281259  | 4.67E-05 | 3.58E-02 | 1.00 | 861 | 1 |
| cg05131696 | Non-CpG Island |                     | Intergenic | -0.036187686 | -5.36E-02 | -0.018806345 | 4.71E-05 | 3.60E-02 | 1.00 | 862 | 1 |
| cg02964085 | Non-CpG Island | <i>C5orf56</i>      | Body       | 0.040299714  | 2.09E-02  | 0.059656509  | 4.71E-05 | 3.60E-02 | 1.00 | 863 | 1 |
| cg01133720 | Non-CpG Island |                     | Intergenic | -0.025600519 | -3.79E-02 | -0.013303764 | 4.72E-05 | 3.60E-02 | 1.00 | 864 | 1 |
| cg01658566 | Shore          | <i>SGSM1</i>        | Body       | 0.03163871   | 1.64E-02  | 0.046836549  | 4.72E-05 | 3.60E-02 | 1.00 | 865 | 2 |
| cg09099627 | Non-CpG Island |                     | Intergenic | -0.036598766 | -5.42E-02 | -0.019017307 | 4.73E-05 | 3.60E-02 | 1.00 | 866 | 1 |
| cg19910679 | Non-CpG Island | <i>PDE6A</i>        | Body       | -0.040728095 | -6.03E-02 | -0.021158796 | 4.74E-05 | 3.60E-02 | 1.00 | 867 | 1 |
| cg21526749 | Shore          | <i>SST</i>          | TSS1500    | 0.046023253  | 2.39E-02  | 0.068137066  | 4.74E-05 | 3.60E-02 | 1.00 | 868 | 1 |
| cg15067124 | Shore          | <i>ATP2A3</i>       | Body       | 0.031896444  | 1.66E-02  | 0.047224304  | 4.75E-05 | 3.61E-02 | 1.00 | 869 | 1 |
| cg19748455 | Non-CpG Island | <i>LOC100996291</i> | TSS1500    | -0.042014028 | -6.22E-02 | -0.021819982 | 4.77E-05 | 3.62E-02 | 1.00 | 870 | 1 |
| cg12804278 | Non-CpG Island | <i>C10orf67</i>     | Body       | 0.04665173   | 2.42E-02  | 0.069079298  | 4.79E-05 | 3.63E-02 | 1.00 | 871 | 1 |
| cg06927323 | Shore          | <i>TBX21</i>        | TSS1500    | -0.03047006  | -4.51E-02 | -0.015820076 | 4.80E-05 | 3.63E-02 | 1.00 | 872 | 1 |
| cg14618634 | Non-CpG Island |                     | Intergenic | -0.034207972 | -5.07E-02 | -0.017760446 | 4.80E-05 | 3.63E-02 | 1.00 | 873 | 1 |
| cg00166372 | Shore          | <i>GNAT1</i>        | 3' UTR     | -0.040666285 | -6.02E-02 | -0.021112316 | 4.80E-05 | 3.63E-02 | 1.00 | 874 | 1 |
| cg10024876 | Island         | <i>FRMD4B</i>       | TSS1500    | 0.034476166  | 1.79E-02  | 0.051055879  | 4.81E-05 | 3.63E-02 | 1.00 | 875 | 1 |
| cg13309020 | Shore          | <i>SMAD1</i>        | TSS1500    | -0.028268115 | -4.19E-02 | -0.014673366 | 4.82E-05 | 3.63E-02 | 1.00 | 876 | 1 |
| cg23838005 | Shore          | <i>MGAT5B</i>       | Body       | -0.029355778 | -4.35E-02 | -0.015235602 | 4.83E-05 | 3.64E-02 | 1.00 | 877 | 1 |
| cg24365525 | Non-CpG Island |                     | Intergenic | 0.027772029  | 1.44E-02  | 0.041134118  | 4.85E-05 | 3.65E-02 | 1.00 | 878 | 1 |
| cg01718920 | Shelf          | <i>PFKFB4</i>       | TSS1500    | -0.03005253  | -4.45E-02 | -0.015591258 | 4.87E-05 | 3.65E-02 | 1.00 | 879 | 1 |
| cg21923568 | Non-CpG Island | <i>SMURF1</i>       | Body       | -0.023660611 | -3.50E-02 | -0.012274577 | 4.87E-05 | 3.65E-02 | 1.00 | 880 | 1 |
| cg01891124 | Non-CpG Island | <i>LOC158435</i>    | Body       | -0.022841499 | -3.38E-02 | -0.011848574 | 4.88E-05 | 3.66E-02 | 1.00 | 881 | 1 |
| cg15773890 | Non-CpG Island |                     | Intergenic | 0.029150274  | 1.51E-02  | 0.043190693  | 4.95E-05 | 3.70E-02 | 1.00 | 882 | 1 |
| cg27158867 | Island         | <i>LRRC56</i>       | 5' UTR     | -0.048925008 | -7.25E-02 | -0.025355665 | 4.96E-05 | 3.71E-02 | 1.00 | 883 | 1 |
| cg17288560 | Non-CpG Island | <i>CDK12</i>        | 5' UTR     | 0.030058632  | 1.56E-02  | 0.044540136  | 4.97E-05 | 3.71E-02 | 1.00 | 884 | 1 |
| cg03182466 | Non-CpG Island | <i>LINC01088</i>    | Body       | -0.046334768 | -6.87E-02 | -0.024005558 | 4.99E-05 | 3.73E-02 | 1.00 | 885 | 2 |
| cg04514392 | Non-CpG Island |                     | Intergenic | -0.041516221 | -6.15E-02 | -0.021507727 | 5.00E-05 | 3.73E-02 | 1.00 | 886 | 1 |
| cg24310434 | Shore          | <i>RP11-689K5.3</i> | 3' UTR     | -0.040842911 | -6.05E-02 | -0.021158143 | 5.00E-05 | 3.73E-02 | 1.00 | 887 | 1 |
| cg09684262 | Shore          | <i>TIPARP</i>       | TSS1500    | 0.023801034  | 1.23E-02  | 0.035273269  | 5.01E-05 | 3.73E-02 | 1.00 | 441 | 2 |
| cg02244028 | Non-CpG Island | <i>SCN11A</i>       | TSS200     | -0.040505568 | -6.01E-02 | -0.020958581 | 5.11E-05 | 3.80E-02 | 1.00 | 888 | 1 |
| cg24810429 | Non-CpG Island |                     | Intergenic | -0.035454027 | -5.26E-02 | -0.018342624 | 5.12E-05 | 3.80E-02 | 1.00 | 889 | 1 |

## Continuation of Table S 3.2

|            |                |                   |            |              |           |              |          |          |      |     |   |
|------------|----------------|-------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg09503767 | Non-CpG Island | <i>NCS1</i>       | Body       | -0.027676616 | -4.10E-02 | -0.014318685 | 5.12E-05 | 3.80E-02 | 1.00 | 890 | 1 |
| cg01336268 | Non-CpG Island | <i>NRXN3</i>      | Body       | -0.047583976 | -7.06E-02 | -0.024617076 | 5.13E-05 | 3.80E-02 | 1.00 | 891 | 1 |
| cg14488626 | Non-CpG Island | <i>SLIT3</i>      | Body       | -0.031015448 | -4.60E-02 | -0.016044241 | 5.13E-05 | 3.80E-02 | 1.00 | 892 | 1 |
| cg18774154 | Non-CpG Island | <i>AC097499.1</i> | 5' UTR     | -0.057585509 | -8.54E-02 | -0.029777456 | 5.17E-05 | 3.82E-02 | 1.00 | 893 | 1 |
| cg23889047 | Non-CpG Island | <i>CAMTA1</i>     | Body       | -0.02445468  | -3.63E-02 | -0.012644375 | 5.18E-05 | 3.82E-02 | 1.00 | 894 | 2 |
| cg12227872 | Non-CpG Island | <i>ITGAM</i>      | Body       | -0.029984122 | -4.45E-02 | -0.015503003 | 5.18E-05 | 3.82E-02 | 1.00 | 895 | 1 |
| cg11543399 | Non-CpG Island | <i>F12</i>        | TSS1500    | -0.03258725  | -4.83E-02 | -0.016844296 | 5.21E-05 | 3.83E-02 | 1.00 | 896 | 1 |
| cg10928075 | Shore          | <i>ZFP36L1</i>    | TSS200     | -0.031183873 | -4.62E-02 | -0.016118596 | 5.21E-05 | 3.83E-02 | 1.00 | 897 | 1 |
| cg04417677 | Shore          | <i>RARRES2</i>    | TSS1500    | -0.057080737 | -8.47E-02 | -0.029503972 | 5.21E-05 | 3.83E-02 | 1.00 | 898 | 1 |
| cg17646797 | Non-CpG Island |                   | Intergenic | -0.041857343 | -6.21E-02 | -0.021634185 | 5.22E-05 | 3.83E-02 | 1.00 | 899 | 1 |
| cg07095995 | Island         | <i>COL25A1</i>    | TSS200     | 0.031867344  | 1.65E-02  | 0.047270417  | 5.25E-05 | 3.85E-02 | 1.00 | 900 | 1 |
| cg18709349 | Island         | <i>PNMA8B</i>     | TSS200     | -0.038352306 | -5.69E-02 | -0.019814603 | 5.25E-05 | 3.85E-02 | 1.00 | 43  | 8 |
| cg25253857 | Non-CpG Island |                   | Intergenic | -0.035817524 | -5.31E-02 | -0.018504426 | 5.26E-05 | 3.85E-02 | 1.00 | 901 | 1 |
| cg18077086 | Non-CpG Island | <i>ATP10D</i>     | 5' UTR     | -0.037276901 | -5.53E-02 | -0.019254747 | 5.28E-05 | 3.86E-02 | 1.00 | 902 | 1 |
| cg24920294 | Non-CpG Island | <i>IL20RB</i>     | 5' UTR     | -0.03484743  | -5.17E-02 | -0.017999543 | 5.28E-05 | 3.86E-02 | 1.00 | 903 | 1 |
| cg18219562 | Non-CpG Island |                   | Intergenic | -0.022428241 | -3.33E-02 | -0.011583684 | 5.29E-05 | 3.86E-02 | 1.00 | 904 | 1 |
| cg16875568 | Non-CpG Island | <i>ITGBL1</i>     | 5' UTR     | 0.034381181  | 1.78E-02  | 0.051011771  | 5.32E-05 | 3.88E-02 | 1.00 | 905 | 1 |
| cg02544453 | Non-CpG Island | <i>MYLK3</i>      | TSS1500    | -0.017884407 | -2.65E-02 | -0.00923203  | 5.34E-05 | 3.89E-02 | 1.00 | 906 | 1 |
| cg07754492 | Non-CpG Island | <i>CASP10</i>     | 5' UTR     | 0.050379623  | 2.60E-02  | 0.074756627  | 5.35E-05 | 3.90E-02 | 1.00 | 907 | 1 |
| cg00843264 | Shore          | <i>TRPM2</i>      | Body       | 0.030268179  | 1.56E-02  | 0.04491528   | 5.36E-05 | 3.90E-02 | 1.00 | 908 | 1 |
| cg00483304 | Shore          | <i>SERTAD2</i>    | 3' UTR     | 0.04117321   | 2.12E-02  | 0.061098185  | 5.36E-05 | 3.90E-02 | 1.00 | 909 | 1 |
| cg21860785 | Non-CpG Island | <i>ESRRG</i>      | Body       | -0.033376585 | -4.95E-02 | -0.017220214 | 5.39E-05 | 3.91E-02 | 1.00 | 910 | 1 |
| cg12258811 | Island         | <i>PER3</i>       | TSS200     | -0.031631998 | -4.69E-02 | -0.016317089 | 5.41E-05 | 3.92E-02 | 1.00 | 911 | 1 |
| cg18039296 | Shore          | <i>LPFR2</i>      | TSS1500    | -0.016134958 | -2.39E-02 | -0.008322865 | 5.41E-05 | 3.92E-02 | 1.00 | 912 | 1 |
| cg06462065 | Shelf          |                   | Intergenic | -0.019583885 | -2.91E-02 | -0.010099404 | 5.43E-05 | 3.93E-02 | 1.00 | 913 | 1 |
| cg02987216 | Non-CpG Island | <i>TRAF3IP2</i>   | TSS1500    | -0.024317381 | -3.61E-02 | -0.012539737 | 5.44E-05 | 3.93E-02 | 1.00 | 914 | 1 |
| cg19820543 | Shore          |                   | Intergenic | -0.038935815 | -5.78E-02 | -0.020075977 | 5.45E-05 | 3.94E-02 | 1.00 | 915 | 1 |
| cg07831886 | Non-CpG Island |                   | Intergenic | -0.052698956 | -7.82E-02 | -0.027167738 | 5.47E-05 | 3.95E-02 | 1.00 | 916 | 1 |
| cg27039257 | Non-CpG Island | <i>IGF2BP2</i>    | Body       | -0.024309246 | -3.61E-02 | -0.012530002 | 5.48E-05 | 3.95E-02 | 1.00 | 917 | 1 |
| cg26701785 | Non-CpG Island | <i>SYNJ2</i>      | Body       | -0.048602425 | -7.22E-02 | -0.025044131 | 5.51E-05 | 3.97E-02 | 1.00 | 918 | 2 |
| cg26080737 | Shore          | <i>SGSM1</i>      | Body       | -0.033267624 | -4.94E-02 | -0.017141989 | 5.52E-05 | 3.97E-02 | 1.00 | 865 | 2 |
| cg03331229 | Non-CpG Island | <i>MMP13</i>      | Body       | 0.028455325  | 1.47E-02  | 0.042249428  | 5.52E-05 | 3.97E-02 | 1.00 | 919 | 1 |
| cg24288310 | Island         | <i>ST8SIA2</i>    | TSS1500    | 0.028096534  | 1.45E-02  | 0.04171792   | 5.53E-05 | 3.97E-02 | 1.00 | 920 | 1 |
| cg04332373 | Shore          | <i>CD38</i>       | TSS1500    | -0.030733406 | -4.56E-02 | -0.015832624 | 5.54E-05 | 3.97E-02 | 1.00 | 921 | 1 |
| cg11848788 | Shelf          |                   | Intergenic | -0.026896125 | -3.99E-02 | -0.013854079 | 5.55E-05 | 3.97E-02 | 1.00 | 922 | 1 |
| cg03856144 | Non-CpG Island | <i>CAMTA1</i>     | Body       | -0.031120955 | -4.62E-02 | -0.016029749 | 5.55E-05 | 3.97E-02 | 1.00 | 894 | 2 |
| cg13223043 | Shore          |                   | Intergenic | -0.019875533 | -2.95E-02 | -0.010237432 | 5.55E-05 | 3.97E-02 | 1.00 | 923 | 1 |
| cg01001286 | Non-CpG Island | <i>IL16</i>       | 5' UTR     | 0.051713902  | 2.66E-02  | 0.07679506   | 5.57E-05 | 3.98E-02 | 1.00 | 520 | 2 |
| cg07581623 | Non-CpG Island | <i>DPYSL2</i>     | Body       | 0.036386341  | 1.87E-02  | 0.05403453   | 5.57E-05 | 3.98E-02 | 1.00 | 924 | 1 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg19039578 | Non-CpG Island | <i>CTD-3088G3.3</i>  | 3' UTR     | -0.032036344 | -4.76E-02 | -0.016496426 | 5.58E-05 | 3.98E-02 | 1.00 | 925 | 1 |
| cg10254445 | Non-CpG Island |                      | Intergenic | 0.025265696  | 1.30E-02  | 0.037521443  | 5.58E-05 | 3.98E-02 | 1.00 | 926 | 1 |
| cg25773223 | Non-CpG Island |                      | Intergenic | -0.035478081 | -5.27E-02 | -0.018268421 | 5.58E-05 | 3.98E-02 | 1.00 | 927 | 1 |
| cg14625801 | Non-CpG Island | <i>ADGRB2</i>        | Body       | -0.030322748 | -4.50E-02 | -0.015612385 | 5.59E-05 | 3.98E-02 | 1.00 | 928 | 1 |
| cg05343811 | Shelf          | <i>NEK9</i>          | Body       | 0.026859123  | 1.38E-02  | 0.039890233  | 5.60E-05 | 3.98E-02 | 1.00 | 929 | 1 |
| cg16024613 | Non-CpG Island | <i>HAUS8</i>         | Body       | 0.023495844  | 1.21E-02  | 0.034896953  | 5.62E-05 | 3.99E-02 | 1.00 | 930 | 1 |
| cg25132244 | Non-CpG Island | <i>TNK2</i>          | 5' UTR     | 0.030544841  | 1.57E-02  | 0.045370637  | 5.64E-05 | 4.00E-02 | 1.00 | 931 | 1 |
| cg07398622 | Non-CpG Island |                      | Intergenic | -0.038012002 | -5.65E-02 | -0.019561261 | 5.65E-05 | 4.00E-02 | 1.00 | 932 | 1 |
| cg14254452 | Shore          | <i>BBC3</i>          | TSS1500    | -0.027040543 | -4.02E-02 | -0.013912259 | 5.67E-05 | 4.01E-02 | 1.00 | 933 | 1 |
| cg03230255 | Non-CpG Island | <i>INSC</i>          | 5' UTR     | 0.028009051  | 1.44E-02  | 0.041610314  | 5.69E-05 | 4.02E-02 | 1.00 | 934 | 1 |
| cg05486252 | Shelf          | <i>RP11-203M5.6</i>  | 5' UTR     | -0.031125148 | -4.62E-02 | -0.016004911 | 5.73E-05 | 4.04E-02 | 1.00 | 935 | 1 |
| cg01153053 | Non-CpG Island |                      | Intergenic | -0.041871427 | -6.22E-02 | -0.0215226   | 5.76E-05 | 4.07E-02 | 1.00 | 936 | 1 |
| cg07332023 | Shelf          | <i>SLC39A13</i>      | Body       | 0.025539201  | 1.31E-02  | 0.037951438  | 5.77E-05 | 4.07E-02 | 1.00 | 937 | 1 |
| cg22994495 | Non-CpG Island | <i>RPTOR</i>         | Body       | -0.027185588 | -4.04E-02 | -0.013971037 | 5.79E-05 | 4.07E-02 | 1.00 | 938 | 1 |
| cg18135341 | Non-CpG Island | <i>IL20RA</i>        | Body       | 0.042545842  | 2.19E-02  | 0.063227257  | 5.79E-05 | 4.07E-02 | 1.00 | 939 | 1 |
| cg19800916 | Non-CpG Island | <i>GRK5</i>          | Body       | -0.046219182 | -6.87E-02 | -0.023752068 | 5.79E-05 | 4.07E-02 | 1.00 | 940 | 1 |
| cg23045307 | Non-CpG Island |                      | Intergenic | -0.018342904 | -2.73E-02 | -0.009418721 | 5.87E-05 | 4.12E-02 | 1.00 | 941 | 1 |
| cg23055986 | Shore          | <i>CEP41</i>         | Body       | 0.028673852  | 1.47E-02  | 0.042624404  | 5.87E-05 | 4.12E-02 | 1.00 | 942 | 1 |
| cg06225967 | Non-CpG Island | <i>CEP95</i>         | Body       | 0.018408558  | 9.45E-03  | 0.027365102  | 5.88E-05 | 4.12E-02 | 1.00 | 943 | 1 |
| cg14721232 | Non-CpG Island | <i>PPP2R2C</i>       | Body       | -0.029353091 | -4.36E-02 | -0.015067512 | 5.91E-05 | 4.13E-02 | 1.00 | 944 | 1 |
| cg00070899 | Shore          | <i>GRM4</i>          | Body       | -0.038109403 | -5.67E-02 | -0.019562094 | 5.91E-05 | 4.13E-02 | 1.00 | 945 | 1 |
| cg08490859 | Non-CpG Island | <i>VWA5B1</i>        | TSS200     | -0.02214375  | -3.29E-02 | -0.011366426 | 5.91E-05 | 4.13E-02 | 1.00 | 76  | 2 |
| cg26607113 | Non-CpG Island |                      | Intergenic | 0.035340462  | 1.81E-02  | 0.052541406  | 5.91E-05 | 4.13E-02 | 1.00 | 946 | 1 |
| cg14438453 | Non-CpG Island | <i>DAAM2</i>         | 5' UTR     | 0.028666477  | 1.47E-02  | 0.04261931   | 5.92E-05 | 4.13E-02 | 1.00 | 947 | 1 |
| cg10817916 | Shore          | <i>CYB561</i>        | 3' UTR     | -0.017599803 | -2.62E-02 | -0.009031752 | 5.94E-05 | 4.14E-02 | 1.00 | 948 | 1 |
| cg15866363 | Shelf          | <i>LOC100506801</i>  | Body       | -0.031508981 | -4.68E-02 | -0.016167966 | 5.95E-05 | 4.14E-02 | 1.00 | 949 | 1 |
| cg07677157 | Non-CpG Island | <i>RP11-221N13.4</i> | 5' UTR     | -0.059479191 | -8.84E-02 | -0.030519602 | 5.95E-05 | 4.14E-02 | 1.00 | 950 | 1 |
| cg27086356 | Shore          | <i>EIF5A2</i>        | TSS1500    | -0.033892847 | -5.04E-02 | -0.017389286 | 5.96E-05 | 4.14E-02 | 1.00 | 951 | 1 |
| cg01658523 | Shore          | <i>POLL</i>          | TSS1500    | 0.028479571  | 1.46E-02  | 0.042353909  | 6.01E-05 | 4.17E-02 | 1.00 | 952 | 1 |
| cg01892389 | Non-CpG Island | <i>SYNJ2</i>         | 5' UTR     | -0.026300638 | -3.91E-02 | -0.013486934 | 6.01E-05 | 4.17E-02 | 1.00 | 918 | 2 |
| cg19208876 | Non-CpG Island | <i>KCNMA1</i>        | Body       | 0.041611155  | 2.13E-02  | 0.061884843  | 6.02E-05 | 4.17E-02 | 1.00 | 953 | 1 |
| cg18586479 | Non-CpG Island | <i>MYLK3</i>         | TSS200     | -0.025169128 | -3.74E-02 | -0.012904535 | 6.03E-05 | 4.17E-02 | 1.00 | 954 | 1 |
| cg23120101 | Non-CpG Island | <i>MCC</i>           | Body       | -0.035145107 | -5.23E-02 | -0.018018006 | 6.04E-05 | 4.17E-02 | 1.00 | 265 | 2 |
| cg15866367 | Shore          | <i>MLNR</i>          | Body       | -0.051320337 | -7.63E-02 | -0.026309965 | 6.04E-05 | 4.17E-02 | 1.00 | 955 | 1 |
| cg23846712 | Shore          | <i>TBXA2R</i>        | TSS1500    | -0.029577647 | -4.40E-02 | -0.015162819 | 6.05E-05 | 4.17E-02 | 1.00 | 956 | 1 |
| cg22816857 | Island         | <i>ELAVL3</i>        | Body       | -0.026527292 | -3.95E-02 | -0.013597414 | 6.06E-05 | 4.18E-02 | 1.00 | 957 | 1 |
| cg15458577 | Shelf          | <i>LINC00505</i>     | 5' UTR     | -0.023677021 | -3.52E-02 | -0.012135077 | 6.07E-05 | 4.18E-02 | 1.00 | 958 | 1 |
| cg06373716 | Non-CpG Island |                      | Intergenic | -0.036040766 | -5.36E-02 | -0.018470214 | 6.08E-05 | 4.18E-02 | 1.00 | 959 | 1 |
| cg22215119 | Non-CpG Island |                      | Intergenic | 0.025613167  | 1.31E-02  | 0.038100656  | 6.09E-05 | 4.18E-02 | 1.00 | 960 | 1 |

## Continuation of Table S 3.2

|            |                |                      |               |              |           |              |          |          |      |     |   |
|------------|----------------|----------------------|---------------|--------------|-----------|--------------|----------|----------|------|-----|---|
| cg07959890 | Non-CpG Island |                      | Intergenic    | -0.057863615 | -8.61E-02 | -0.029646347 | 6.11E-05 | 4.19E-02 | 1.00 | 961 | 1 |
| cg21753297 | Non-CpG Island | <i>DOCK6</i>         | Body          | -0.029524234 | -4.39E-02 | -0.015124835 | 6.12E-05 | 4.20E-02 | 1.00 | 962 | 1 |
| cg07222015 | Non-CpG Island | <i>LINC00944</i>     | Body          | 0.051522026  | 2.64E-02  | 0.076655675  | 6.14E-05 | 4.21E-02 | 1.00 | 963 | 1 |
| cg24292594 | Non-CpG Island | <i>C19orf81</i>      | Body          | -0.028199276 | -4.20E-02 | -0.014442774 | 6.15E-05 | 4.21E-02 | 1.00 | 964 | 1 |
| cg04862404 | Non-CpG Island | <i>NAALADL2</i>      | Body          | 0.026002581  | 1.33E-02  | 0.03868753   | 6.15E-05 | 4.21E-02 | 1.00 | 965 | 1 |
| cg13569867 | Shelf          |                      | Intergenic    | -0.05157446  | -7.67E-02 | -0.026411432 | 6.16E-05 | 4.21E-02 | 1.00 | 966 | 1 |
| cg25093216 | Non-CpG Island | <i>MCC</i>           | Body          | 0.026215933  | 1.34E-02  | 0.039008372  | 6.17E-05 | 4.22E-02 | 1.00 | 967 | 1 |
| cg00232402 | Shore          |                      | Intergenic    | -0.03158772  | -4.70E-02 | -0.016164821 | 6.24E-05 | 4.26E-02 | 1.00 | 968 | 1 |
| cg02555215 | Shelf          | <i>DTX1</i>          | Body          | -0.03716448  | -5.53E-02 | -0.019015168 | 6.26E-05 | 4.26E-02 | 1.00 | 969 | 1 |
| cg14100916 | Non-CpG Island |                      | Intergenic    | -0.01643658  | -2.45E-02 | -0.008409489 | 6.26E-05 | 4.26E-02 | 1.00 | 970 | 1 |
| cg14608778 | Non-CpG Island | <i>FSIP1</i>         | Body          | -0.046185921 | -6.87E-02 | -0.023629897 | 6.26E-05 | 4.26E-02 | 1.00 | 971 | 1 |
| cg06425263 | Non-CpG Island | <i>GRIN2A</i>        | Body          | -0.039250983 | -5.84E-02 | -0.020080604 | 6.27E-05 | 4.26E-02 | 1.00 | 972 | 1 |
| cg14179786 | Non-CpG Island | <i>RHBDD1</i>        | Body          | 0.037932581  | 1.94E-02  | 0.056468589  | 6.32E-05 | 4.29E-02 | 1.00 | 973 | 1 |
| cg07201717 | Non-CpG Island |                      | Intergenic    | -0.023397884 | -3.48E-02 | -0.01196403  | 6.33E-05 | 4.29E-02 | 1.00 | 974 | 1 |
| cg24832862 | Non-CpG Island | <i>MS4A6E</i>        | 3' UTR        | -0.034877153 | -5.19E-02 | -0.017833055 | 6.33E-05 | 4.29E-02 | 1.00 | 975 | 1 |
| cg04451864 | Non-CpG Island | <i>TIGIT</i>         | TSS200        | -0.034194149 | -5.09E-02 | -0.01748092  | 6.35E-05 | 4.30E-02 | 1.00 | 976 | 1 |
| cg20910352 | Non-CpG Island | <i>AMHR2</i>         | TSS200        | -0.030920928 | -4.60E-02 | -0.015807187 | 6.35E-05 | 4.30E-02 | 1.00 | 977 | 1 |
| cg10853566 | Shelf          | <i>SLC7A7</i>        | First Exon    | -0.036232236 | -5.39E-02 | -0.018518195 | 6.38E-05 | 4.31E-02 | 1.00 | 978 | 1 |
| cg17425590 | Non-CpG Island |                      | Intergenic    | -0.021506049 | -3.20E-02 | -0.01099147  | 6.38E-05 | 4.31E-02 | 1.00 | 979 | 1 |
| cg25806524 | Shore          | <i>PNMA88</i>        | First Exon    | -0.027513137 | -4.10E-02 | -0.014059693 | 6.39E-05 | 4.31E-02 | 1.00 | 43  | 8 |
| cg19779810 | Shore          | <i>SC65</i>          | TSS1500       | -0.021613788 | -3.22E-02 | -0.011044535 | 6.40E-05 | 4.31E-02 | 1.00 | 183 | 4 |
| cg26906737 | Non-CpG Island | <i>TMEM206</i>       | Body          | 0.017592182  | 8.99E-03  | 0.026195219  | 6.40E-05 | 4.31E-02 | 1.00 | 980 | 1 |
| cg20465027 | Non-CpG Island |                      | Intergenic    | -0.030365668 | -4.52E-02 | -0.015512927 | 6.43E-05 | 4.32E-02 | 1.00 | 981 | 1 |
| cg05942199 | Non-CpG Island | <i>FNDC3A</i>        | Body          | -0.037293442 | -5.55E-02 | -0.019048792 | 6.45E-05 | 4.33E-02 | 1.00 | 982 | 1 |
| cg04718705 | Non-CpG Island | <i>TNKS</i>          | Body          | -0.040783062 | -6.07E-02 | -0.020824323 | 6.48E-05 | 4.35E-02 | 1.00 | 983 | 1 |
| cg03828221 | Island         | <i>ZNF497</i>        | Body          | -0.059020756 | -8.79E-02 | -0.030134622 | 6.49E-05 | 4.35E-02 | 1.00 | 984 | 1 |
| cg18043014 | Non-CpG Island | <i>DEAF1</i>         | Body          | -0.034971877 | -5.21E-02 | -0.017851388 | 6.52E-05 | 4.37E-02 | 1.00 | 985 | 1 |
| cg22623303 | Island         | <i>KHSRP</i>         | Body          | -0.036203789 | -5.39E-02 | -0.018479634 | 6.52E-05 | 4.37E-02 | 1.00 | 986 | 1 |
| cg07173688 | Non-CpG Island |                      | Intergenic    | -0.030668321 | -4.57E-02 | -0.015653583 | 6.53E-05 | 4.37E-02 | 1.00 | 987 | 1 |
| cg27490646 | Non-CpG Island |                      | Intergenic    | -0.028425107 | -4.23E-02 | -0.014507636 | 6.54E-05 | 4.37E-02 | 1.00 | 988 | 1 |
| cg09982291 | Shore          | <i>MIR203A</i>       | TSS1500       | -0.039674047 | -5.91E-02 | -0.02024791  | 6.54E-05 | 4.37E-02 | 1.00 | 989 | 1 |
| cg10733692 | Non-CpG Island |                      | Intergenic    | -0.02698354  | -4.02E-02 | -0.013769507 | 6.56E-05 | 4.37E-02 | 1.00 | 990 | 1 |
| cg05587665 | Non-CpG Island |                      | Intergenic    | -0.030441471 | -4.53E-02 | -0.01553331  | 6.56E-05 | 4.37E-02 | 1.00 | 991 | 1 |
| cg11608390 | Non-CpG Island | <i>GMDS</i>          | Body          | -0.031178967 | -4.64E-02 | -0.015908977 | 6.57E-05 | 4.37E-02 | 1.00 | 992 | 1 |
| cg16857134 | Non-CpG Island | <i>DKFZp434J0226</i> | TSS1500       | -0.026440154 | -3.94E-02 | -0.013488733 | 6.58E-05 | 4.38E-02 | 1.00 | 993 | 1 |
| cg05728295 | Non-CpG Island |                      | Intergenic    | -0.023263179 | -3.47E-02 | -0.011867146 | 6.59E-05 | 4.38E-02 | 1.00 | 994 | 1 |
| cg26643325 | Shore          | <i>RASGEF1A</i>      | Body          | -0.040507735 | -6.04E-02 | -0.020662538 | 6.60E-05 | 4.38E-02 | 1.00 | 341 | 2 |
| cg13793503 | Non-CpG Island | <i>SLC4A10</i>       | Exon Boundary | -0.031001844 | -4.62E-02 | -0.015813357 | 6.60E-05 | 4.38E-02 | 1.00 | 995 | 1 |
| cg18616702 | Shore          | <i>ADAMTSL4</i>      | 5' UTR        | -0.024576082 | -3.66E-02 | -0.012532567 | 6.63E-05 | 4.40E-02 | 1.00 | 996 | 1 |

## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |      |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|------|---|
| cg00982484 | Non-CpG Island |                      | Intergenic | 0.034256649  | 1.75E-02  | 0.051045879  | 6.64E-05 | 4.40E-02 | 1.00 | 997  | 1 |
| cg15170994 | Non-CpG Island | <i>DGKG</i>          | Body       | 0.042382982  | 2.16E-02  | 0.063157876  | 6.66E-05 | 4.41E-02 | 1.00 | 998  | 1 |
| cg06534670 | Non-CpG Island | <i>DLGAP4</i>        | 5' UTR     | -0.025818516 | -3.85E-02 | -0.013162464 | 6.66E-05 | 4.41E-02 | 1.00 | 999  | 1 |
| cg11757214 | Shelf          |                      | Intergenic | -0.037970614 | -5.66E-02 | -0.019354223 | 6.69E-05 | 4.41E-02 | 1.00 | 1000 | 1 |
| cg08569022 | Non-CpG Island | <i>ACSS2</i>         | Body       | -0.039386478 | -5.87E-02 | -0.020071906 | 6.71E-05 | 4.43E-02 | 1.00 | 1001 | 1 |
| cg17674725 | Shore          | <i>LYPD1</i>         | 5' UTR     | 0.025704851  | 1.31E-02  | 0.038312158  | 6.73E-05 | 4.43E-02 | 1.00 | 1002 | 1 |
| cg06775471 | Non-CpG Island | <i>OSTN</i>          | Body       | 0.033293318  | 1.70E-02  | 0.04962557   | 6.75E-05 | 4.44E-02 | 1.00 | 1003 | 1 |
| cg09954323 | Shore          | <i>BRINP2</i>        | TSS200     | 0.028504991  | 1.45E-02  | 0.042488918  | 6.75E-05 | 4.44E-02 | 1.00 | 1004 | 1 |
| cg20519597 | Non-CpG Island | <i>CSTF3</i>         | Body       | -0.037065202 | -5.53E-02 | -0.018879415 | 6.77E-05 | 4.45E-02 | 1.00 | 1005 | 1 |
| cg08575141 | Shelf          |                      | Intergenic | -0.037704007 | -5.62E-02 | -0.019199224 | 6.80E-05 | 4.47E-02 | 1.00 | 1006 | 1 |
| cg24999870 | Shore          | <i>EML2</i>          | TSS1500    | -0.025327935 | -3.78E-02 | -0.01289576  | 6.81E-05 | 4.47E-02 | 1.00 | 417  | 2 |
| cg06269299 | Non-CpG Island | <i>G6PC2</i>         | TSS1500    | 0.036324769  | 1.85E-02  | 0.054156278  | 6.82E-05 | 4.47E-02 | 1.00 | 1007 | 1 |
| cg21573101 | Non-CpG Island |                      | Intergenic | 0.04021526   | 2.05E-02  | 0.059956815  | 6.83E-05 | 4.47E-02 | 1.00 | 1008 | 1 |
| cg09685073 | Non-CpG Island | <i>DKFZp686K1684</i> | Body       | -0.032800412 | -4.89E-02 | -0.016698419 | 6.83E-05 | 4.47E-02 | 1.00 | 1009 | 1 |
| cg14491582 | Non-CpG Island | <i>STARD13</i>       | 5' UTR     | -0.049864474 | -7.43E-02 | -0.025383298 | 6.84E-05 | 4.47E-02 | 1.00 | 1010 | 1 |
| cg08036284 | Non-CpG Island |                      | Intergenic | -0.065514648 | -9.77E-02 | -0.033348502 | 6.84E-05 | 4.47E-02 | 1.00 | 1011 | 1 |
| cg06100542 | Non-CpG Island | <i>RXRG</i>          | Body       | -0.0356296   | -5.31E-02 | -0.018128208 | 6.90E-05 | 4.50E-02 | 1.00 | 1012 | 1 |
| cg14339051 | Non-CpG Island | <i>PDE4D</i>         | Body       | 0.032905606  | 1.67E-02  | 0.049069651  | 6.90E-05 | 4.50E-02 | 1.00 | 1013 | 1 |
| cg00404703 | Non-CpG Island |                      | Intergenic | -0.020173296 | -3.01E-02 | -0.010260875 | 6.93E-05 | 4.52E-02 | 1.00 | 1014 | 1 |
| cg05136561 | Non-CpG Island | <i>OTOG</i>          | Body       | -0.028736457 | -4.29E-02 | -0.014615674 | 6.94E-05 | 4.52E-02 | 1.00 | 1015 | 1 |
| cg04748834 | Island         | <i>CHL1</i>          | 5' UTR     | 0.028218857  | 1.43E-02  | 0.042091168  | 6.99E-05 | 4.54E-02 | 1.00 | 1016 | 1 |
| cg19567415 | Non-CpG Island | <i>TANC2</i>         | Body       | -0.036422981 | -5.43E-02 | -0.018516783 | 6.99E-05 | 4.54E-02 | 1.00 | 1017 | 1 |
| cg19359526 | Non-CpG Island | <i>ITGAD</i>         | Body       | -0.035851523 | -5.35E-02 | -0.018221174 | 7.03E-05 | 4.56E-02 | 1.00 | 1018 | 1 |
| cg09325749 | Non-CpG Island |                      | Intergenic | 0.03040158   | 1.55E-02  | 0.045352577  | 7.03E-05 | 4.56E-02 | 1.00 | 1019 | 1 |
| cg23630423 | Shore          | <i>RTN4RL1</i>       | Body       | -0.033134536 | -4.94E-02 | -0.016838532 | 7.04E-05 | 4.56E-02 | 1.00 | 1020 | 1 |
| cg25847702 | Non-CpG Island | <i>LTBP1</i>         | Body       | -0.029150421 | -4.35E-02 | -0.014813774 | 7.04E-05 | 4.56E-02 | 1.00 | 1021 | 1 |
| cg11270364 | Non-CpG Island |                      | Intergenic | 0.023120971  | 1.17E-02  | 0.034492383  | 7.04E-05 | 4.56E-02 | 1.00 | 1022 | 1 |
| cg07303548 | Shore          | <i>TMEM134</i>       | 3' UTR     | -0.023998104 | -3.58E-02 | -0.012190277 | 7.09E-05 | 4.58E-02 | 1.00 | 1023 | 1 |
| cg00738869 | Shore          | <i>PIK3CB</i>        | 5' UTR     | 0.037880115  | 1.92E-02  | 0.056521396  | 7.11E-05 | 4.59E-02 | 1.00 | 1024 | 1 |
| cg10421346 | Non-CpG Island |                      | Intergenic | -0.036954477 | -5.51E-02 | -0.018766936 | 7.12E-05 | 4.60E-02 | 1.00 | 1025 | 1 |
| cg18041642 | Non-CpG Island |                      | Intergenic | -0.025654325 | -3.83E-02 | -0.013025896 | 7.15E-05 | 4.60E-02 | 1.00 | 1026 | 1 |
| cg17093914 | Non-CpG Island | <i>SORCS2</i>        | Body       | -0.030448457 | -4.54E-02 | -0.015459922 | 7.15E-05 | 4.60E-02 | 1.00 | 1027 | 1 |
| cg07345969 | Non-CpG Island |                      | Intergenic | 0.0483014    | 2.45E-02  | 0.07207881   | 7.15E-05 | 4.60E-02 | 1.00 | 1028 | 1 |
| cg02759482 | Non-CpG Island |                      | Intergenic | 0.028840187  | 1.46E-02  | 0.043037922  | 7.15E-05 | 4.60E-02 | 1.00 | 1029 | 1 |
| cg21510141 | Non-CpG Island | <i>EIF2C1</i>        | 5' UTR     | -0.037752521 | -5.63E-02 | -0.019161538 | 7.19E-05 | 4.62E-02 | 1.00 | 1030 | 1 |
| cg19717347 | Shelf          | <i>SLC15A4</i>       | Body       | 0.015402588  | 7.82E-03  | 0.02298767   | 7.19E-05 | 4.62E-02 | 1.00 | 1031 | 1 |
| cg20559634 | Non-CpG Island | <i>LIMD1</i>         | 5' UTR     | 0.047046578  | 2.39E-02  | 0.070217438  | 7.21E-05 | 4.62E-02 | 1.00 | 1032 | 1 |
| cg00242263 | Non-CpG Island | <i>LINC01088</i>     | Body       | -0.023261687 | -3.47E-02 | -0.011804467 | 7.21E-05 | 4.62E-02 | 1.00 | 885  | 2 |
| cg01520578 | Non-CpG Island |                      | Intergenic | -0.050207951 | -7.49E-02 | -0.025477911 | 7.22E-05 | 4.62E-02 | 1.00 | 1033 | 1 |



## Continuation of Table S 3.2

|            |                |                      |            |              |           |              |          |          |      |      |   |
|------------|----------------|----------------------|------------|--------------|-----------|--------------|----------|----------|------|------|---|
| cg02205348 | Shelf          | <i>TRIO</i>          | Body       | -0.031409265 | -4.69E-02 | -0.015938125 | 7.22E-05 | 4.62E-02 | 1.00 | 157  | 2 |
| cg15072822 | Shore          |                      | Intergenic | 0.027107628  | 1.38E-02  | 0.040464011  | 7.26E-05 | 4.64E-02 | 1.00 | 1034 | 1 |
| cg08694513 | Non-CpG Island |                      | Intergenic | -0.022801358 | -3.40E-02 | -0.011565444 | 7.27E-05 | 4.64E-02 | 1.00 | 1035 | 1 |
| cg23089913 | Non-CpG Island | <i>NANOS1</i>        | TSS1500    | -0.044877785 | -6.70E-02 | -0.022762344 | 7.28E-05 | 4.64E-02 | 1.00 | 1036 | 1 |
| cg14913407 | Non-CpG Island | <i>SNAP47</i>        | Body       | -0.029285537 | -4.37E-02 | -0.01485048  | 7.30E-05 | 4.65E-02 | 1.00 | 585  | 2 |
| cg11228392 | Non-CpG Island | <i>RP11-343/18.2</i> | TSS1500    | -0.028523206 | -4.26E-02 | -0.014460904 | 7.33E-05 | 4.66E-02 | 1.00 | 1037 | 1 |
| cg14515812 | Shore          | <i>RAP1GAP</i>       | 5' UTR     | -0.035921999 | -5.36E-02 | -0.01821137  | 7.33E-05 | 4.66E-02 | 1.00 | 1038 | 1 |
| cg01014262 | Shelf          |                      | Intergenic | -0.025518601 | -3.81E-02 | -0.012936968 | 7.34E-05 | 4.66E-02 | 1.00 | 1039 | 1 |
| cg09697696 | Shore          | <i>CTB-131B5.4</i>   | TSS1500    | 0.028367687  | 1.44E-02  | 0.042358326  | 7.37E-05 | 4.68E-02 | 1.00 | 581  | 3 |
| cg03746015 | Non-CpG Island | <i>CLEC16A</i>       | Body       | -0.044244282 | -6.61E-02 | -0.022421953 | 7.38E-05 | 4.68E-02 | 1.00 | 1040 | 1 |
| cg05394800 | Shore          | <i>DLEU1</i>         | 5' UTR     | 0.025790306  | 1.31E-02  | 0.038512488  | 7.40E-05 | 4.69E-02 | 1.00 | 1041 | 1 |
| cg25012097 | Shore          | <i>FREM2</i>         | First Exon | 0.027576637  | 1.40E-02  | 0.041180967  | 7.41E-05 | 4.69E-02 | 1.00 | 1042 | 1 |
| cg25991374 | Non-CpG Island |                      | Intergenic | -0.036287264 | -5.42E-02 | -0.018384472 | 7.42E-05 | 4.69E-02 | 1.00 | 1043 | 1 |
| cg06237102 | Non-CpG Island |                      | Intergenic | -0.045010426 | -6.72E-02 | -0.022800623 | 7.44E-05 | 4.70E-02 | 1.00 | 1044 | 1 |
| cg27307975 | Non-CpG Island | <i>LIPC</i>          | Body       | -0.035354078 | -5.28E-02 | -0.017907442 | 7.45E-05 | 4.70E-02 | 1.00 | 1045 | 1 |
| cg02268561 | Shore          | <i>NMT2</i>          | TSS1500    | -0.04272766  | -6.38E-02 | -0.021633274 | 7.50E-05 | 4.73E-02 | 1.00 | 1046 | 1 |
| cg04247018 | Non-CpG Island | <i>IQCJ-SCHIP1</i>   | Body       | -0.031146168 | -4.65E-02 | -0.015762741 | 7.55E-05 | 4.76E-02 | 1.00 | 1047 | 1 |
| cg14507310 | Non-CpG Island |                      | Intergenic | -0.043054346 | -6.43E-02 | -0.021789046 | 7.56E-05 | 4.76E-02 | 1.00 | 1048 | 1 |
| cg15883678 | Non-CpG Island | <i>LASP1</i>         | Body       | -0.027897039 | -4.17E-02 | -0.014116493 | 7.57E-05 | 4.76E-02 | 1.00 | 1049 | 1 |
| cg14153041 | Non-CpG Island | <i>HSPG2</i>         | Body       | -0.035104931 | -5.24E-02 | -0.017763371 | 7.57E-05 | 4.76E-02 | 1.00 | 1050 | 1 |
| cg10674392 | Non-CpG Island | <i>IL12RB2</i>       | Body       | 0.035386477  | 1.79E-02  | 0.052871264  | 7.60E-05 | 4.77E-02 | 1.00 | 1051 | 1 |
| cg03601752 | Shore          | <i>GNG7</i>          | 5' UTR     | -0.044121507 | -6.59E-02 | -0.022320564 | 7.60E-05 | 4.77E-02 | 1.00 | 1052 | 1 |
| cg05848656 | Shore          | <i>CREB3L1</i>       | TSS1500    | -0.020347096 | -3.04E-02 | -0.010292763 | 7.61E-05 | 4.77E-02 | 1.00 | 1053 | 1 |
| cg25186292 | Non-CpG Island | <i>HSPBAP1</i>       | Body       | -0.031560353 | -4.72E-02 | -0.015965081 | 7.61E-05 | 4.77E-02 | 1.00 | 1054 | 1 |
| cg20080548 | Shelf          | <i>SBSN</i>          | TSS1500    | -0.060041782 | -8.97E-02 | -0.030370886 | 7.62E-05 | 4.77E-02 | 1.00 | 817  | 2 |
| cg19442101 | Non-CpG Island | <i>CTB-330/18.2</i>  | TSS200     | -0.030569134 | -4.57E-02 | -0.015462343 | 7.62E-05 | 4.77E-02 | 1.00 | 1055 | 1 |
| cg21515958 | Non-CpG Island | <i>TMEM207</i>       | Body       | 0.027913177  | 1.41E-02  | 0.041707825  | 7.63E-05 | 4.77E-02 | 1.00 | 1056 | 1 |
| cg12899747 | Non-CpG Island |                      | Intergenic | 0.05086543   | 2.57E-02  | 0.076008157  | 7.65E-05 | 4.77E-02 | 1.00 | 1057 | 1 |
| cg16717480 | Shore          | <i>KIAA1429</i>      | TSS1500    | 0.089227006  | 4.51E-02  | 0.13333331   | 7.66E-05 | 4.77E-02 | 1.00 | 1058 | 1 |
| cg15421037 | Non-CpG Island | <i>PRMT8</i>         | Body       | -0.03752447  | -5.61E-02 | -0.018974884 | 7.66E-05 | 4.77E-02 | 1.00 | 1059 | 1 |
| cg03357219 | Non-CpG Island | <i>BANK1</i>         | TSS200     | -0.060038407 | -8.97E-02 | -0.03035138  | 7.69E-05 | 4.79E-02 | 1.00 | 1060 | 1 |
| cg13695646 | Non-CpG Island | <i>BHMT</i>          | TSS200     | -0.031811081 | -4.75E-02 | -0.016078878 | 7.72E-05 | 4.80E-02 | 1.00 | 1061 | 1 |
| cg00200515 | Non-CpG Island |                      | Intergenic | -0.033762899 | -5.05E-02 | -0.017064669 | 7.72E-05 | 4.80E-02 | 1.00 | 1062 | 1 |
| cg07006163 | Shore          | <i>CRAMP1L</i>       | Body       | -0.032604286 | -4.87E-02 | -0.016474582 | 7.76E-05 | 4.82E-02 | 1.00 | 1063 | 1 |
| cg01005523 | Non-CpG Island | <i>PDGFC</i>         | Body       | -0.027750881 | -4.15E-02 | -0.014020429 | 7.77E-05 | 4.82E-02 | 1.00 | 1064 | 1 |
| cg15294485 | Non-CpG Island | <i>CNNM1</i>         | Body       | -0.057656898 | -8.62E-02 | -0.029129641 | 7.77E-05 | 4.82E-02 | 1.00 | 1065 | 1 |
| cg00338002 | Shore          | <i>C1orf38</i>       | TSS1500    | -0.03131018  | -4.68E-02 | -0.015815845 | 7.80E-05 | 4.83E-02 | 1.00 | 1066 | 1 |
| cg22439390 | Non-CpG Island |                      | Intergenic | -0.029481141 | -4.41E-02 | -0.014891289 | 7.80E-05 | 4.83E-02 | 1.00 | 1067 | 1 |
| cg25799433 | Non-CpG Island | <i>SIRPB1</i>        | 5' UTR     | 0.034072991  | 1.72E-02  | 0.050947051  | 7.89E-05 | 4.88E-02 | 1.00 | 1068 | 1 |

Continuation of Table S 3.2

|            |                |                |            |              |           |              |          |          |      |      |   |
|------------|----------------|----------------|------------|--------------|-----------|--------------|----------|----------|------|------|---|
| cg19082401 | Non-CpG Island |                | Intergenic | -0.029650737 | -4.43E-02 | -0.014960381 | 7.95E-05 | 4.91E-02 | 1.00 | 1069 | 1 |
| cg12073370 | Shore          | <i>LRCOL1</i>  | Body       | -0.042537055 | -6.36E-02 | -0.021458239 | 7.97E-05 | 4.92E-02 | 1.00 | 1070 | 1 |
| cg10001287 | Non-CpG Island | <i>RELN</i>    | Body       | 0.043397935  | 2.19E-02  | 0.064912212  | 8.03E-05 | 4.95E-02 | 1.00 | 1071 | 1 |
| cg01057126 | Non-CpG Island |                | Intergenic | -0.019034421 | -2.85E-02 | -0.00959758  | 8.04E-05 | 4.95E-02 | 1.00 | 1072 | 1 |
| cg11134246 | Shore          | <i>CCDC24</i>  | Body       | -0.025138424 | -3.76E-02 | -0.012672941 | 8.06E-05 | 4.96E-02 | 1.00 | 1073 | 1 |
| cg13908635 | Non-CpG Island | <i>CYP27A1</i> | 3' UTR     | 0.028517633  | 1.44E-02  | 0.042660769  | 8.08E-05 | 4.97E-02 | 1.00 | 1074 | 1 |
| cg04238199 | Non-CpG Island | <i>CSorf49</i> | 3' UTR     | -0.024255019 | -3.63E-02 | -0.01222493  | 8.09E-05 | 4.97E-02 | 1.00 | 1075 | 1 |
| cg07418863 | Shore          | <i>RNF182</i>  | TSS1500    | -0.037118754 | -5.55E-02 | -0.018704413 | 8.12E-05 | 4.98E-02 | 1.00 | 1076 | 1 |
| cg24121503 | Shelf          | <i>AKAP12</i>  | Body       | -0.029221594 | -4.37E-02 | -0.014724678 | 8.12E-05 | 4.98E-02 | 1.00 | 1077 | 1 |
| cg25697823 | Shore          | <i>HIST4H4</i> | 3' UTR     | 0.052631812  | 2.65E-02  | 0.078742616  | 8.12E-05 | 4.98E-02 | 1.00 | 1078 | 1 |

Footnote:

\* Nearest gene and gene feature were annotated using UCSC annotations in the Illumina EPIC array manifest file. For CpGs without these UCSC annotations, GENCODEv12 annotations were used instead.

\*\* As-associated regions were defined as regions where As-associated CpGs were located within 5kb of each other

Table S 3.3 List of CpGs previously identified in whole blood to associate with urinary arsenic levels in adults after accounting for multiple testing

| Name       | Study               | PMID     |
|------------|---------------------|----------|
| cg04605617 | Argos et al. 2015   | 25325195 |
| cg06121226 | Argos et al. 2015   | 25325195 |
| cg01225779 | Argos et al. 2015   | 25325195 |
| cg13651690 | Argos et al. 2015   | 25325195 |
| cg27413543 | Kaushal et al. 2017 | 28242323 |
| cg23671167 | Kaushal et al. 2017 | 28242323 |
| cg11516038 | Kaushal et al. 2017 | 28242323 |
| cg02347881 | Kaushal et al. 2017 | 28242323 |
| cg22443249 | Kaushal et al. 2017 | 28242323 |
| cg06582411 | Kaushal et al. 2017 | 28242323 |
| cg01983216 | Kaushal et al. 2017 | 28242323 |
| cg24221507 | Kaushal et al. 2017 | 28242323 |
| cg05903720 | Kaushal et al. 2017 | 28242323 |
| cg14981312 | Kaushal et al. 2017 | 28242323 |
| cg12106731 | Bozack et al. 2020  | 32603190 |
| cg07317306 | Bozack et al. 2020  | 32603190 |
| cg18616702 | Bozack et al. 2020  | 32603190 |
| cg04940901 | Bozack et al. 2020  | 32603190 |
| cg06690548 | Bozack et al. 2020  | 32603190 |
| cg01538969 | Bozack et al. 2020  | 32603190 |
| cg14827056 | Bozack et al. 2020  | 32603190 |
| cg09280971 | Bozack et al. 2020  | 32603190 |
| cg14595618 | Bozack et al. 2020  | 32603190 |
| cg27178850 | Bozack et al. 2020  | 32603190 |

Continuation of Table S 3.3

|            |                       |          |
|------------|-----------------------|----------|
| cg03036214 | Bozack et al. 2020    | 32603190 |
| cg00500428 | Bozack et al. 2020    | 32603190 |
| cg03497652 | Bozack et al. 2020    | 32603190 |
| cg07021906 | Bozack et al. 2020    | 32603190 |
| cg12116137 | Bozack et al. 2020    | 32603190 |
| cg21369801 | Bozack et al. 2020    | 32603190 |
| cg20493718 | Bozack et al. 2020    | 32603190 |
| cg08059112 | Bozack et al. 2020    | 32603190 |
| cg22294740 | Bozack et al. 2020    | 32603190 |
| cg20509831 | Bozack et al. 2020    | 32603190 |
| cg01912040 | Demanelis et al. 2019 | 31135185 |
| cg05962511 | Demanelis et al. 2019 | 31135185 |
| cg10003262 | Demanelis et al. 2019 | 31135185 |
| cg17420142 | Demanelis et al. 2019 | 31135185 |
| cg06466147 | Demanelis et al. 2019 | 31135185 |
| cg09082427 | Demanelis et al. 2019 | 31135185 |
| cg04193083 | Demanelis et al. 2019 | 31135185 |
| cg19534475 | Demanelis et al. 2019 | 31135185 |
| cg12608784 | Demanelis et al. 2019 | 31135185 |
| cg14891900 | Demanelis et al. 2019 | 31135185 |
| cg11308227 | Demanelis et al. 2019 | 31135185 |
| cg13832772 | Demanelis et al. 2019 | 31135185 |
| cg15108641 | Demanelis et al. 2019 | 31135185 |
| cg09658504 | Demanelis et al. 2019 | 31135185 |
| cg05438461 | Demanelis et al. 2019 | 31135185 |
| cg10663081 | Demanelis et al. 2019 | 31135185 |

Continuation of Table S 3.3

|            |                       |          |
|------------|-----------------------|----------|
| cg08759026 | Demanelis et al. 2019 | 31135185 |
| cg11644394 | Demanelis et al. 2019 | 31135185 |
| cg05428706 | Demanelis et al. 2019 | 31135185 |
| cg12865207 | Demanelis et al. 2019 | 31135185 |
| cg09183146 | Demanelis et al. 2019 | 31135185 |
| cg04622454 | Demanelis et al. 2019 | 31135185 |
| cg06381803 | Demanelis et al. 2019 | 31135185 |
| cg02772605 | Demanelis et al. 2019 | 31135185 |
| cg10283165 | Demanelis et al. 2019 | 31135185 |
| cg04891961 | Demanelis et al. 2019 | 31135185 |
| cg12746706 | Demanelis et al. 2019 | 31135185 |
| cg02330195 | Demanelis et al. 2019 | 31135185 |
| cg08077890 | Demanelis et al. 2019 | 31135185 |
| cg20433952 | Demanelis et al. 2019 | 31135185 |
| cg10185759 | Demanelis et al. 2019 | 31135185 |
| cg05646745 | Demanelis et al. 2019 | 31135185 |
| cg22345623 | Demanelis et al. 2019 | 31135185 |
| cg13480898 | Demanelis et al. 2019 | 31135185 |

Table S 3.4 Inflation factors for the EWAS approaches for assessing the association between DNAm and log2-transformed creatinine-adjusted urinary arsenic levels

| <b>EWAS Approach</b>                                           | <b>Inflation factors</b> |
|----------------------------------------------------------------|--------------------------|
| Cell Type Composition without SVs                              | 1.644                    |
| Only SVs without other covariates                              | 1.263                    |
| Covariates + default # of SVs obtained from<br>the sva package | 1.210                    |
| Specified # of SVs                                             | 1.141                    |

Table S 3.5 Enrichment of As-associated CpGs in DNaseI hypersensitivity sites and EWASs of human health and disease

| <b><i>Enrichment in DNaseI hypersensitivity sites</i></b> |                                                  |                                                      |                                                                    |                                    |
|-----------------------------------------------------------|--------------------------------------------------|------------------------------------------------------|--------------------------------------------------------------------|------------------------------------|
|                                                           | <b>Located in a DNaseI hypersensitivity site</b> | <b>Not located in a DNaseI hypersensitivity site</b> | <b>Proportion of CpGs located in DNaseI hypersensitivity sites</b> | <b>Fisher's Exact Test p-value</b> |
| <b>FDR-significant As-associated CpGs</b>                 | 908                                              | 269                                                  | 0.77                                                               | 2.74E-36                           |
| <b>Not FDR-significant CpGs</b>                           | 430,786                                          | 289,833                                              | 0.60                                                               |                                    |

| <b><i>Enrichment in EWASs of human health and disease in the MRC-IEU EWAS catalog</i></b> |                                                           |                                                               |                                                                           |                                    |
|-------------------------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------|---------------------------------------------------------------------------|------------------------------------|
|                                                                                           | <b>Identified in another EWAS of human health/disease</b> | <b>Not identified in another EWAS of human health/disease</b> | <b>Proportion of CpGs identified in EWASs of human health and disease</b> | <b>Fisher's Exact Test p-value</b> |
| <b>FDR-significant As-associated CpGs</b>                                                 | 880                                                       | 297                                                           | 0.75                                                                      | 3.51E-27                           |
| <b>Not FDR-significant CpGs</b>                                                           | 431,283                                                   | 290,513                                                       | 0.60                                                                      |                                    |

Table S 3.6 Hits for As-associated CpGs in EWASs of other health conditions and diseases in the MRC-IEU EWAS Catalog

| Minimum p-values of CpGs that were also identified across studies in the EWAS catalog |                     |                 |             |       |       |                                                |
|---------------------------------------------------------------------------------------|---------------------|-----------------|-------------|-------|-------|------------------------------------------------|
| Name                                                                                  | Chromosome:Position | Gene            | Type        | Beta  | SE    | Minimum p-value across studies in EWAS catalog |
| cg00003287                                                                            | chr1:201346149      | <i>TNNT2</i>    | Open sea    | 0.002 | 1E-04 | 0                                              |
| cg00099427                                                                            | chr1:159870134      | <i>CCDC19</i>   | Island      | 0.011 | 4E-04 | 0                                              |
| cg00338002                                                                            | chr1:28198813       | <i>C1orf38</i>  | North shore | -1.81 | 0.116 | 5.80E-27                                       |
| cg00519063                                                                            | chr1:94373434       | <i>GCLM</i>     | North shore | -0.01 | 2E-04 | 0                                              |
| cg01436424                                                                            | chr1:7669070        | <i>CAMTA1</i>   | Open sea    | 0.004 | 1E-04 | 0                                              |
| cg02856716                                                                            | chr1:18993307       | <i>PAX7</i>     | Open sea    | NA    | NA    | 4.61E-36                                       |
| cg03761216                                                                            | chr1:51986130       | <i>EPS15</i>    | South shore | -2.91 | 0.082 | 2.64E-53                                       |
| cg04605617                                                                            | chr1:20501558       | <i>PLA2G2C</i>  | Open sea    | -0    | 2E-04 | 0                                              |
| cg04637372                                                                            | chr1:20617456       | <i>VWA5B1</i>   | Open sea    | 0.005 | 2E-04 | 0                                              |
| cg07201717                                                                            | chr1:25031947       | -               | Open sea    | 0.003 | 2E-04 | 0                                              |
| cg07207669                                                                            | chr1:155102388      | <i>EFNA1</i>    | South shore | -0.01 | 3E-04 | 0                                              |
| cg08263205                                                                            | chr1:175854934      | -               | Open sea    | -0    | 2E-04 | 0                                              |
| cg08655800                                                                            | chr1:109735314      | <i>KIAA1324</i> | Open sea    | -0    | 2E-04 | 0                                              |
| cg09308842                                                                            | chr1:231297968      | <i>TRIM67</i>   | Island      | 0.001 | ##### | 0                                              |
| cg09476130                                                                            | chr1:159870086      | <i>CCDC19</i>   | Island      | 0.016 | 4E-04 | 0                                              |
| cg10442251                                                                            | chr1:18700625       | <i>IGSF21</i>   | North shelf | 0.007 | 2E-04 | 0                                              |
| cg11134246                                                                            | chr1:44458001       | <i>CCDC24</i>   | South shore | 0.018 | 0.001 | 0                                              |
| cg11848788                                                                            | chr1:39041399       | -               | North shelf | -0    | 2E-04 | 0                                              |
| cg11947712                                                                            | chr1:109203268      | <i>C1orf59</i>  | North shore | 0.002 | 2E-04 | 0                                              |
| cg12258811                                                                            | chr1:7844735        | <i>PER3</i>     | Island      | -0.02 | 0.002 | 0                                              |
| cg12277336                                                                            | chr1:85458682       | <i>MCOLN2</i>   | North shelf | 0.016 | 0.002 | 1.50E-14                                       |
| cg13223043                                                                            | chr1:26492308       | -               | South shore | -2.58 | 0.094 | 9.41E-45                                       |
| cg14176088                                                                            | chr1:9785289        | <i>PIK3CD</i>   | Open sea    | -0    | 3E-04 | 0                                              |
| cg14515812                                                                            | chr1:21978863       | <i>RAP1GAP</i>  | South shore | 0.005 | 2E-04 | 0                                              |
| cg15166089                                                                            | chr1:27710030       | <i>CD164L2</i>  | South shore | -0.01 | 0.001 | 0                                              |
| cg15458577                                                                            | chr1:46947998       | -               | North shelf | 0.006 | 2E-04 | 0                                              |
| cg15512289                                                                            | chr1:97026071       | -               | Open sea    | 0.007 | 3E-04 | 0                                              |



Continuation of Table S 3.6

|            |                |                 |             |       |       |          |
|------------|----------------|-----------------|-------------|-------|-------|----------|
| cg16298481 | chr1:89989717  | <i>LRRC8B</i>   | North shore | 0.003 | 2E-04 | 0        |
| cg16400794 | chr1:27709850  | <i>CD164L2</i>  | South shore | 0.001 | ##### | 0        |
| cg16417374 | chr1:26098310  | <i>MAN1C1</i>   | South shore | -0    | 2E-04 | 0        |
| cg17425590 | chr1:229830303 | -               | Open sea    | -0    | 1E-04 | 1.63E-15 |
| cg17897187 | chr1:14923736  | <i>KIAA1026</i> | North shore | -0.02 | 0.002 | 0        |
| cg19020666 | chr1:178452552 | -               | North shelf | -0    | 2E-04 | 0        |
| cg19403323 | chr1:210024060 | <i>C1orf107</i> | Open sea    | -0.37 | 0.049 | 3.05E-14 |
| cg20692268 | chr1:25358981  | -               | Open sea    | -0.01 | 3E-04 | 0        |
| cg21291672 | chr1:29451713  | <i>TMEM200B</i> | Island      | 0.007 | 2E-04 | 0        |
| cg21995800 | chr1:111984376 | <i>WDR77</i>    | Open sea    | NA    | NA    | 2.90E-25 |
| cg22356061 | chr1:227954102 | <i>SNAP47</i>   | Open sea    | -0    | ##### | 0        |
| cg23889047 | chr1:7705756   | <i>CAMTA1</i>   | Open sea    | -0.01 | 3E-04 | 0        |
| cg24661860 | chr1:10784363  | <i>CASZ1</i>    | Open sea    | 0.005 | 3E-04 | 0        |
| cg24680570 | chr1:232726838 | -               | Open sea    | -0    | 2E-04 | 0        |
| cg26906737 | chr1:212541551 | <i>TMEM206</i>  | Open sea    | -0    | 2E-04 | 0        |
| cg00281776 | chr2:209224225 | -               | North shore | -0    | ##### | 0        |
| cg00889363 | chr2:206255841 | <i>PARD3B</i>   | Open sea    | 0.004 | 2E-04 | 0        |
| cg02343335 | chr2:113785852 | <i>IL1F8</i>    | Open sea    | 0.003 | 2E-04 | 0        |
| cg03566527 | chr2:121749280 | <i>GLI2</i>     | South shelf | 0.002 | 3E-04 | 0        |
| cg07173688 | chr2:119564329 | -               | Open sea    | 0.009 | 2E-04 | 0        |
| cg07754492 | chr2:202047754 | <i>CASP10</i>   | Open sea    | 0.02  | 0.003 | 3.58E-14 |
| cg13716566 | chr2:241772244 | -               | South shore | 0.005 | 2E-04 | 0        |
| cg15344192 | chr2:121347210 | -               | South shore | 0.03  | 0.003 | 0        |
| cg15705813 | chr2:70297499  | -               | Open sea    | -0.02 | 4E-04 | 0        |
| cg17674725 | chr2:133426418 | <i>LYPD1</i>    | North shore | 0.002 | 1E-04 | 0        |
| cg18774154 | chr2:126504079 | -               | Open sea    | NA    | NA    | 1.09E-30 |
| cg19146711 | chr2:44473071  | -               | Open sea    | 7E-04 | NA    | 3.90E-06 |
| cg19180827 | chr2:47684718  | <i>MSH2</i>     | Open sea    | 0.002 | 2E-04 | 4.40E-16 |
| cg19789753 | chr2:242805962 | -               | Island      | -0.84 | 0.073 | 5.55E-31 |
| cg25520960 | chr2:219863984 | -               | North shelf | 0.013 | ##### | 0        |
| cg26210250 | chr2:26863503  | <i>CIB4</i>     | Open sea    | 0.003 | 2E-04 | 0        |
| cg00124095 | chr3:72305082  | -               | Open sea    | -0.01 | 3E-04 | 0        |

Continuation of Table S 3.6

|            |                |                            |             |       |       |          |
|------------|----------------|----------------------------|-------------|-------|-------|----------|
| cg00166372 | chr3:50232459  | <i>GNAT1</i>               | South shore | 1.331 | 0.098 | 3.38E-23 |
| cg00404400 | chr3:186078568 | <i>DGKG</i>                | North shore | -0    | 1E-04 | 0        |
| cg00492070 | chr3:107810716 | <i>CD47</i>                | South shore | -0    | 2E-04 | 0        |
| cg02417857 | chr3:181326706 | <i>SOX2OT</i>              | Open sea    | 0.006 | 4E-04 | 0        |
| cg04246167 | chr3:58984465  | <i>C3orf67</i>             | Open sea    | -0.01 | 3E-04 | 0        |
| cg04514392 | chr3:107086058 | -                          | Open sea    | -0.01 | 3E-04 | 0        |
| cg04732968 | chr3:13933460  | -                          | North shelf | 2.4   | 0.307 | 5.70E-15 |
| cg04748834 | chr3:239229    | <i>CHL1</i>                | Island      | 0.002 | ##### | 0        |
| cg06804344 | chr3:128778699 | <i>GP9</i>                 | Open sea    | -2.06 | 0.095 | 1.14E-36 |
| cg09831562 | chr3:181327125 | <i>SOX2OT</i>              | Open sea    | 0.004 | 3E-04 | 0        |
| cg10024876 | chr3:69592049  | -                          | Island      | 7E-04 | ##### | 0        |
| cg10255761 | chr3:49210029  | <i>KLHDC8B</i>             | South shore | -0.01 | 2E-04 | 0        |
| cg10753966 | chr3:55556197  | <i>ERC2</i>                | Open sea    | 0.01  | 4E-04 | 0        |
| cg11492723 | chr3:16577697  | -                          | Open sea    | -0.01 | 3E-04 | 0        |
| cg12899747 | chr3:25391527  | -                          | Open sea    | -0    | 3E-04 | 0        |
| cg15721584 | chr3:181326755 | <i>SOX2OT</i>              | Open sea    | 0.004 | 2E-04 | 0        |
| cg16764781 | chr3:189347820 | <i>TP63</i>                | Open sea    | 0.006 | 3E-04 | 0        |
| cg20559634 | chr3:45609400  | -                          | Open sea    | NA    | NA    | 2.84E-55 |
| cg21723486 | chr3:189349323 | <i>TP63</i>                | Open sea    | -0    | 4E-04 | 0        |
| cg21798630 | chr3:190363742 | <i>IL1RAP</i>              | Open sea    | -0.01 | 3E-04 | 0        |
| cg24262469 | chr3:156391694 | <i>TIPARP;LOC100287227</i> | North shore | 0.006 | 0.001 | 1.30E-07 |
| cg25698975 | chr3:139440415 | -                          | Open sea    | 0.007 | 2E-04 | 0        |
| cg00382138 | chr4:110723299 | <i>CFI</i>                 | Open sea    | -0    | 3E-04 | 0        |
| cg01290904 | chr4:5708474   | <i>EVC2</i>                | North shore | -0    | 2E-04 | 0        |
| cg01442064 | chr4:5713450   | <i>EVC</i>                 | Island      | 0.001 | 1E-04 | 0        |
| cg03460109 | chr4:7427313   | <i>SORCS2</i>              | Open sea    | 0.004 | ##### | 0        |
| cg03485217 | chr4:10686300  | <i>CLNK</i>                | Open sea    | 0.011 | 3E-04 | 0        |
| cg04332373 | chr4:15779642  | <i>CD38</i>                | North shore | 0.002 | 1E-04 | 0        |
| cg05068452 | chr4:115521241 | <i>UGT8</i>                | South shore | 0.001 | NA    | 8.10E-16 |
| cg07095995 | chr4:110223980 | <i>COL25A1</i>             | Island      | 8E-04 | ##### | 0        |
| cg13054321 | chr4:6123352   | <i>JAKMIP1</i>             | Open sea    | NA    | NA    | 1.79E-12 |
| cg13210763 | chr4:58309519  | -                          | Open sea    | 0.005 | 3E-04 | 0        |

Continuation of Table S 3.6

|            |                |                 |             |       |       |          |
|------------|----------------|-----------------|-------------|-------|-------|----------|
| cg13468041 | chr4:74902951  | <i>CXCL3</i>    | North shore | -3    | 0.09  | 2.95E-51 |
| cg17724147 | chr4:58290751  | -               | Open sea    | -0    | 3E-04 | 0        |
| cg20421058 | chr4:7337225   | <i>SORCS2</i>   | Open sea    | 0.006 | 2E-04 | 0        |
| cg20729813 | chr4:7502429   | <i>SORCS2</i>   | Open sea    | -0.01 | 3E-04 | 0        |
| cg22667835 | chr4:6107791   | <i>JAKMIP1</i>  | South shore | 0.2   | 0.026 | 1.00E-14 |
| cg22953687 | chr4:2363130   | <i>ZFYVE28</i>  | North shelf | -0    | 2E-04 | 1.30E-10 |
| cg25744957 | chr4:7106195   | <i>FLJ36777</i> | Open sea    | -0    | 2E-04 | 0        |
| cg00169725 | chr5:176543901 | -               | Island      | 0.002 | 1E-04 | 0        |
| cg00275479 | chr5:134882430 | -               | South shelf | 0.013 | 3E-04 | 0        |
| cg01225779 | chr5:179238472 | <i>SQSTM1</i>   | South shelf | 0.002 | 2E-04 | 2.20E-16 |
| cg01772439 | chr5:149012102 | <i>FLJ41603</i> | Open sea    | -0.01 | 3E-04 | 0        |
| cg01995468 | chr5:176544135 | -               | Island      | 7E-04 | ##### | 0        |
| cg04238199 | chr5:7831537   | <i>C5orf49</i>  | Open sea    | NA    | NA    | 5.86E-44 |
| cg07684019 | chr5:139174529 | <i>PSD2</i>     | North shore | 0.009 | 2E-04 | 0        |
| cg09697696 | chr5:139488623 | -               | South shore | 0.004 | 2E-04 | 0        |
| cg11241750 | chr5:125694354 | <i>GRAMD3</i>   | Open sea    | 0.003 | 2E-04 | 0        |
| cg12351310 | chr5:52937992  | <i>NDUFS4</i>   | Open sea    | -0.01 | 3E-04 | 0        |
| cg12613020 | chr5:134881662 | -               | South shore | 0.011 | 2E-04 | 0        |
| cg13695646 | chr5:78407509  | <i>BHMT</i>     | Open sea    | 0.005 | 2E-04 | 0        |
| cg18394552 | chr5:159428643 | -               | Open sea    | -0.01 | 3E-04 | 0        |
| cg19005955 | chr5:132938694 | <i>FSTL4</i>    | Open sea    | 0.007 | 2E-04 | 0        |
| cg19910679 | chr5:149240728 | <i>PDE6A</i>    | Open sea    | -0.01 | 4E-04 | 0        |
| cg24119717 | chr5:139488731 | -               | South shore | 0.007 | ##### | 0        |
| cg25542319 | chr5:112540429 | <i>MCC</i>      | Open sea    | 1.467 | 0.099 | 1.69E-25 |
| cg26103104 | chr5:139490895 | -               | North shore | 0.021 | 4E-04 | 0        |
| cg00070899 | chr6:34024479  | <i>GRM4</i>     | South shore | 0.006 | 3E-04 | 0        |
| cg01090161 | chr6:34032747  | <i>GRM4</i>     | Open sea    | 0.003 | 2E-04 | 0        |
| cg01564343 | chr6:41121920  | <i>TREML1</i>   | Open sea    | -0    | 2E-04 | 0        |
| cg04334496 | chr6:133953605 | -               | Open sea    | 0.007 | 2E-04 | 0        |
| cg04826368 | chr6:27130208  | -               | Open sea    | -0    | 2E-04 | 0        |
| cg05411199 | chr6:3749715   | <i>C6orf145</i> | North shore | 0.004 | 2E-04 | 0        |
| cg09222732 | chr6:466893    | -               | Open sea    | -0    | 1E-04 | 0        |

Continuation of Table S 3.6

|            |                |                   |             |       |       |          |
|------------|----------------|-------------------|-------------|-------|-------|----------|
| cg11608390 | chr6:2021942   | <i>GMDS</i>       | Open sea    | -0    | 2E-04 | 0        |
| cg12662084 | chr6:17809126  | <i>KIF13A</i>     | Open sea    | -0    | 2E-04 | 0        |
| cg13569146 | chr6:26330589  | -                 | Open sea    | 0.031 | 0.004 | 2.04E-12 |
| cg14438453 | chr6:39786566  | <i>DAAM2</i>      | Open sea    | 3.7   | 0.349 | 3.60E-26 |
| cg14575656 | chr6:26330584  | -                 | Open sea    | 0.001 | 1E-04 | 0        |
| cg15501231 | chr6:31737630  | <i>C6orf27</i>    | Open sea    | -0    | 1E-04 | 0        |
| cg15773890 | chr6:17259549  | -                 | Open sea    | -0    | 2E-04 | 0        |
| cg16983110 | chr6:128813319 | <i>PTPRK</i>      | Open sea    | 2.055 | 0.165 | 4.42E-21 |
| cg17501210 | chr6:166970252 | <i>RPS6KA2</i>    | Open sea    | -0    | 2E-04 | 0        |
| cg24028809 | chr6:153414101 | <i>RGS17</i>      | Open sea    | -0    | 3E-04 | 0        |
| cg24121503 | chr6:151650834 | <i>AKAP12</i>     | South shelf | -0    | 2E-04 | 0        |
| cg26701785 | chr6:158448455 | <i>SYNJ2</i>      | Open sea    | -0.01 | 2E-04 | 0        |
| cg27042983 | chr6:44063147  | -                 | Open sea    | -0    | 2E-04 | 0        |
| cg03473532 | chr7:131008743 | <i>MKLN1</i>      | North shelf | -0    | ##### | 0        |
| cg09766547 | chr7:30465150  | <i>NOD1</i>       | Open sea    | -0.04 | 0.004 | 9.80E-16 |
| cg09848638 | chr7:733065    | <i>PRKAR1B</i>    | Open sea    | NA    | NA    | 6.77E-19 |
| cg20228731 | chr7:130646051 | <i>FLJ43663</i>   | Open sea    | NA    | NA    | 2.20E-28 |
| cg26205859 | chr7:76967860  | <i>PION</i>       | Open sea    | 0.002 | 2E-04 | 0        |
| cg26282283 | chr7:2704368   | <i>TTYH3</i>      | South shelf | 0.021 | 0.002 | 0        |
| cg03554283 | chr8:76319451  | -                 | Island      | 0.001 | ##### | 0        |
| cg05938624 | chr8:1314612   | -                 | South shelf | NA    | NA    | 1.06E-31 |
| cg07312099 | chr8:17942583  | <i>ASAH1</i>      | South shore | 0.028 | 0.004 | 4.95E-11 |
| cg08525461 | chr8:40958085  | -                 | Open sea    | -0    | 2E-04 | 0        |
| cg11688949 | chr8:27115956  | <i>STMN4</i>      | Open sea    | 0.014 | 3E-04 | 0        |
| cg16717480 | chr8:95566380  | <i>KIAA1429</i>   | South shore | -0    | 1E-04 | 2.20E-15 |
| cg19911880 | chr8:131353969 | <i>ASAP1</i>      | Open sea    | NA    | NA    | 9.47E-30 |
| cg25305703 | chr8:128378218 | -                 | Open sea    | 0.003 | 3E-04 | 0        |
| cg25762706 | chr8:27115867  | <i>STMN4</i>      | Open sea    | 0.01  | 3E-04 | 0        |
| cg27106513 | chr8:54790372  | <i>RGS20</i>      | Island      | 0.001 | 1E-04 | 0        |
| cg00407468 | chr9:126692955 | <i>DENND1A</i>    | Open sea    | 0.113 | 0.021 | 3.28E-08 |
| cg03383434 | chr9:140318662 | <i>EXD3;NOXA1</i> | Island      | -0    | ##### | 0        |
| cg05061804 | chr9:135282417 | <i>TTF1</i>       | South shore | -0.01 | 3E-04 | 0        |

Continuation of Table S 3.6

|            |                 |                  |             |       |       |          |
|------------|-----------------|------------------|-------------|-------|-------|----------|
| cg13409439 | chr9:135040296  | <i>NTNG2</i>     | South shore | -0.17 | 0.026 | 1.76E-11 |
| cg14507310 | chr9:121285184  | -                | Open sea    | 0.005 | 2E-04 | 0        |
| cg14618634 | chr9:89193057   | -                | Open sea    | 0.002 | 2E-04 | 0        |
| cg17930194 | chr9:124982834  | <i>LHX6</i>      | Island      | 0.012 | 3E-04 | 0        |
| cg25073708 | chr9:132083538  | <i>C9orf106</i>  | Island      | NA    | NA    | 2.76E-38 |
| cg00284153 | chr10:135171713 | <i>C10orf125</i> | Island      | 0.008 | 3E-04 | 0        |
| cg02268561 | chr10:15212066  | <i>NMT2</i>      | South shore | -1.39 | 0.107 | 3.56E-22 |
| cg03281572 | chr10:114502318 | <i>VTI1A</i>     | Open sea    | -0    | 1E-04 | 0        |
| cg05324516 | chr10:29421357  | -                | Open sea    | -0.01 | 1E-04 | 0        |
| cg05428706 | chr10:102730130 | -                | South shore | -0.02 | 0.002 | 2.90E-14 |
| cg05962511 | chr10:102730022 | -                | South shore | 0.003 | 2E-04 | 0        |
| cg08227353 | chr10:102821670 | <i>KAZALD1</i>   | Island      | -2.47 | 0.086 | 3.52E-46 |
| cg11891925 | chr10:13091949  | -                | Open sea    | -0.01 | ##### | 0        |
| cg14165660 | chr10:112889349 | -                | Open sea    | -0    | 2E-04 | 0        |
| cg15108641 | chr10:99263320  | <i>UBTD1</i>     | South shelf | -0    | 3E-04 | 0        |
| cg15840151 | chr10:124459681 | <i>C10orf120</i> | Open sea    | 0.011 | 3E-04 | 0        |
| cg18400281 | chr10:24633303  | <i>KIAA1217</i>  | Open sea    | 0.002 | 1E-04 | 0        |
| cg20327845 | chr10:3138418   | <i>PFKP</i>      | Open sea    | -0.02 | 5E-04 | 0        |
| cg20552296 | chr10:100995305 | <i>HPSE2</i>     | South shore | 6E-04 | NA    | 3.40E-13 |
| cg21345949 | chr10:135171718 | <i>C10orf125</i> | Island      | 0.01  | 3E-04 | 0        |
| cg22959742 | chr10:13913931  | <i>FRMD4A</i>    | Open sea    | 0.019 | 0.002 | 0        |
| cg23089913 | chr10:120788366 | <i>NANOS1</i>    | Open sea    | -1.31 | 0.119 | 3.55E-18 |
| cg24181174 | chr10:123900861 | <i>TACC2</i>     | Open sea    | -0    | 2E-04 | 0        |
| cg26572392 | chr10:24496943  | <i>KIAA1217</i>  | Open sea    | 0.01  | 3E-04 | 0        |
| cg26964636 | chr10:24639300  | <i>KIAA1217</i>  | Open sea    | 0.002 | 2E-04 | 0        |
| cg00022866 | chr11:64108440  | <i>CCDC88B</i>   | North shore | 0.008 | ##### | 0        |
| cg01894038 | chr11:77741873  | -                | Open sea    | -0.02 | ##### | 0        |
| cg03303857 | chr11:120619307 | <i>GRIK4</i>     | Open sea    | 0.005 | 2E-04 | 0        |
| cg03331229 | chr11:102821111 | <i>MMP13</i>     | Open sea    | -0.02 | 0.002 | 1.90E-14 |
| cg04422903 | chr11:64108550  | <i>CCDC88B</i>   | North shore | 0.005 | 2E-04 | 0        |
| cg04954111 | chr11:34363786  | <i>ABTB2</i>     | Open sea    | -1.64 | 0.065 | 5.75E-42 |
| cg05469118 | chr11:122215103 | -                | Open sea    | -0    | 1E-04 | 0        |

Continuation of Table S 3.6

|            |                 |                  |             |       |       |          |
|------------|-----------------|------------------|-------------|-------|-------|----------|
| cg07088882 | chr11:8038043   | -                | North shelf | 0.008 | 2E-04 | 0        |
| cg07250758 | chr11:61049251  | <i>VWCE</i>      | Open sea    | -2.25 | 0.088 | 3.30E-42 |
| cg08012287 | chr11:66315298  | <i>ACTN3</i>     | South shore | 0.01  | 3E-04 | 0        |
| cg08043020 | chr11:8038083   | -                | North shelf | 0.01  | 3E-04 | 0        |
| cg11921539 | chr11:116328665 | -                | Open sea    | 0.007 | 2E-04 | 0        |
| cg12245084 | chr11:78672632  | <i>ODZ4</i>      | North shore | 0.004 | 2E-04 | 0        |
| cg15322932 | chr11:67777682  | <i>ALDH3B1</i>   | Open sea    | 0.002 | 1E-04 | 0        |
| cg16760981 | chr11:63688531  | -                | South shelf | -0    | 2E-04 | 1.80E-15 |
| cg17603132 | chr11:8190152   | <i>RIC3</i>      | North shore | NA    | NA    | 7.45E-41 |
| cg18993113 | chr11:9944707   | <i>SBF2</i>      | Open sea    | -0    | ##### | 0        |
| cg19115017 | chr11:15229502  | <i>INSC</i>      | Open sea    | -0    | 1E-04 | 0        |
| cg19998151 | chr11:267613    | -                | Open sea    | 9E-04 | 1E-04 | 1.20E-14 |
| cg20831648 | chr11:78761473  | <i>ODZ4</i>      | Open sea    | NA    | NA    | 1.66E-49 |
| cg22022957 | chr11:2908102   | <i>CDKN1C</i>    | South shore | 0.004 | 3E-04 | 0        |
| cg22327175 | chr11:121323934 | <i>SORL1</i>     | South shore | -0.01 | 0.001 | 2.10E-15 |
| cg25281677 | chr11:57248038  | -                | North shore | NA    | NA    | 7.43E-12 |
| cg26247064 | chr11:74466198  | <i>RNF169</i>    | Open sea    | -0    | 3E-04 | 0        |
| cg26428830 | chr11:45728183  | -                | Open sea    | 0.009 | 2E-04 | 0        |
| cg27158867 | chr11:537638    | <i>LRRC56</i>    | Island      | NA    | NA    | 8.02E-14 |
| cg00997853 | chr12:26844579  | <i>ITPR2</i>     | Open sea    | -0    | ##### | 0        |
| cg02555215 | chr12:113497560 | <i>DTX1</i>      | South shelf | -0    | ##### | 2.10E-14 |
| cg02637282 | chr12:98768240  | -                | Open sea    | -0.02 | 6E-04 | 0        |
| cg03348792 | chr12:53075482  | <i>KRT1</i>      | Open sea    | -0.01 | 3E-04 | 0        |
| cg03466342 | chr12:52753899  | <i>KRT85</i>     | Open sea    | -0.01 | 3E-04 | 0        |
| cg05131696 | chr12:31499320  | -                | Open sea    | -0    | 2E-04 | 0        |
| cg06383241 | chr12:116997022 | <i>MAP1LC3B2</i> | Open sea    | NA    | NA    | 6.15E-39 |
| cg07677157 | chr12:66050928  | -                | Open sea    | -0.01 | 3E-04 | 0        |
| cg09080114 | chr12:6983111   | <i>SPSB2</i>     | South shore | NA    | NA    | 7.87E-17 |
| cg09767076 | chr12:124199914 | <i>ATP6V0A2</i>  | South shelf | -0    | 2E-04 | 0        |
| cg10637292 | chr12:52960529  | <i>KRT74</i>     | Open sea    | 0.003 | 2E-04 | 0        |
| cg19717347 | chr12:129304956 | <i>SLC15A4</i>   | North shelf | NA    | NA    | 3.36E-13 |
| cg20697394 | chr12:132918305 | -                | North shelf | NA    | NA    | 7.88E-36 |

Continuation of Table S 3.6

|            |                 |                    |             |       |       |          |
|------------|-----------------|--------------------|-------------|-------|-------|----------|
| cg26358374 | chr12:110252616 | <i>TRPV4</i>       | Open sea    | 0.002 | ##### | 0        |
| cg26621780 | chr12:113547430 | <i>RASAL1</i>      | North shelf | -0.13 | 0.017 | 5.40E-15 |
| cg27598407 | chr12:14570669  | <i>ATF7IP</i>      | Open sea    | NA    | NA    | 4.91E-10 |
| cg27656459 | chr12:116796562 | -                  | Open sea    | -0    | 2E-04 | 0        |
| cg01757312 | chr13:112720565 | <i>SOX1</i>        | Island      | 0.001 | 1E-04 | 0        |
| cg02402946 | chr13:21035079  | <i>CRYL1</i>       | Open sea    | 0.023 | 0.003 | 1.10E-15 |
| cg05394800 | chr13:50707050  | -                  | North shore | 0.002 | ##### | 0        |
| cg08993368 | chr13:28364872  | -                  | North shore | 6E-04 | ##### | 0        |
| cg15718287 | chr13:36702658  | <i>DCLK1</i>       | North shelf | 0.002 | 3E-04 | 1.10E-15 |
| cg15866367 | chr13:49796387  | <i>MLNR</i>        | South shore | 0.002 | 2E-04 | 0        |
| cg00776917 | chr14:30691759  | -                  | Open sea    | 0.007 | 2E-04 | 0        |
| cg01121022 | chr14:104338788 | -                  | Open sea    | 0.005 | 2E-04 | 0        |
| cg04962528 | chr14:21098715  | -                  | North shelf | 0.002 | 1E-04 | 0        |
| cg05343811 | chr14:75590896  | <i>NEK9</i>        | North shelf | NA    | NA    | 5.20E-81 |
| cg09298313 | chr14:55569959  | -                  | Open sea    | -0    | 2E-04 | 0        |
| cg10529613 | chr14:25287305  | <i>STXBP6</i>      | Open sea    | NA    | NA    | 3.98E-43 |
| cg10853566 | chr14:23288993  | <i>SLC7A7</i>      | North shelf | 0.002 | 1E-04 | 0        |
| cg13679772 | chr14:89933549  | <i>FOXN3</i>       | Open sea    | -0.02 | 0.002 | 0        |
| cg15827779 | chr14:103524899 | <i>CDC42BPB</i>    | South shore | 0.004 | 2E-04 | 0        |
| cg16239257 | chr14:75052674  | <i>LTBP2</i>       | Island      | NA    | NA    | 1.66E-25 |
| cg16313343 | chr14:105768109 | <i>BRF1</i>        | South shore | 0.002 | 1E-04 | 0        |
| cg18186343 | chr14:101317620 | <i>MIR770;MEG3</i> | Open sea    | -0.01 | 3E-04 | 0        |
| cg18369034 | chr14:52733216  | <i>PTGDR</i>       | North shore | -0    | 2E-04 | 0        |
| cg19643211 | chr14:104314557 | <i>PPP1R13B</i>    | Island      | 0.001 | ##### | 0        |
| cg00939525 | chr15:40129363  | <i>GPR176</i>      | Open sea    | -0    | 1E-04 | 4.10E-14 |
| cg01001286 | chr15:81475638  | -                  | Open sea    | 0.002 | NA    | 5.40E-15 |
| cg01973394 | chr15:82335947  | <i>MEX3B</i>       | Island      | 0.007 | 0.001 | 8.60E-06 |
| cg07286682 | chr15:69447132  | <i>MIR548H4</i>    | Open sea    | 0.002 | ##### | 0        |
| cg14492992 | chr15:53101548  | -                  | South shelf | 0.01  | 3E-04 | 0        |
| cg19696491 | chr15:78857124  | <i>CHRNA5</i>      | North shore | 0.032 | 0.002 | 0        |
| cg22563815 | chr15:78856949  | <i>CHRNA5</i>      | North shore | 0.022 | 0.002 | 0        |
| cg23206630 | chr15:62942319  | <i>TLN2</i>        | Open sea    | -0    | 3E-04 | 0        |

Continuation of Table S 3.6

|            |                |                   |                   |       |       |          |
|------------|----------------|-------------------|-------------------|-------|-------|----------|
| cg23848052 | chr15:66318701 | <i>MEGF11</i>     | Open sea          | 0.012 | 3E-04 | 0        |
| cg24288310 | chr15:92936678 | <i>ST8SIA2</i>    | Island            | 6E-04 | NA    | 4.00E-30 |
| cg26657240 | chr15:81475061 | -                 | Open sea          | -0.01 | 3E-04 | 0        |
| cg27307975 | chr15:58797839 | <i>LIPC</i>       | Open sea          | -5    | 0.323 | 1.60E-54 |
| cg01014262 | chr16:48533930 | -                 | South shelf       | 2.915 | 0.082 | 1.42E-53 |
| cg02267536 | chr16:88039689 | <i>BANP</i>       | Island            | -0    | 2E-04 | 0        |
| cg02544453 | chr16:46782760 | <i>MYLK3</i>      | Open sea          | -0    | ##### | 1.55E-17 |
| cg03031932 | chr16:81547138 | <i>CMIP</i>       | Open sea          | -0.01 | 0.002 | 2.82E-11 |
| cg03746015 | chr16:11141000 | <i>CLEC16A</i>    | Open sea          | 0.003 | 3E-04 | 0        |
| cg05425326 | chr16:58439361 | <i>GINS3</i>      | Open sea          | -0.01 | 0.001 | 0        |
| cg05989795 | chr16:1021386  | <i>LMF1</i>       | South shore       | 0.003 | 1E-04 | 0        |
| cg06478823 | chr16:20774873 | <i>ACSM3</i>      | Open sea          | -0.02 | 5E-04 | 0        |
| cg07561162 | chr16:10174417 | <i>GRIN2A</i>     | Open sea          | 7E-04 | NA    | 1.10E-23 |
| cg08692903 | chr16:85424553 | -                 | Open sea          | -0.01 | 3E-04 | 0        |
| cg09866569 | chr16:2737340  | <i>KCTD5</i>      | Open sea          | 7E-04 | ##### | 0        |
| cg12654412 | chr16:71459826 | -                 | Island            | 0.012 | 3E-04 | 0        |
| cg26624021 | chr16:56995739 | <i>CETP</i>       | Open sea          | -0.01 | 3E-04 | 0        |
| cg27027151 | chr16:27461638 | <i>IL21R</i>      | Open sea          | -0    | 1E-04 | 1.61E-13 |
| cg27162196 | chr16:75183231 | <i>ZFP1</i>       | South shore       | -0    | 3E-04 | 0        |
| cg27554551 | chr16:4421486  | <i>VASN;CORO7</i> | North shore       | NA    | NA    | 1.23E-13 |
| cg01912040 | chr17:1106553  | -                 | North shore       | 0.005 | 3E-04 | 0        |
| cg02348119 | chr17:78003528 | <i>TBC1D16</i>    | Open sea          | -0.01 | 3E-04 | 0        |
| cg02483931 | chr17:34202461 | <i>CCL5</i>       | Open sea          | 0.007 | 0.001 | 1.90E-12 |
| cg03871754 | chr17:79320652 | -                 | Island            | 0.005 | 2E-04 | 0        |
| cg04136369 | chr17:36070600 | <i>HNF1B</i>      | Open sea          | 0.004 | 3E-04 | 0        |
| cg04891961 | chr17:27939900 | <i>ANKRD13B</i>   | Island            | 0.006 | 2E-04 | 0        |
| cg04970434 | chr17:4838096  | <i>GP1BA</i>      | Open sea          | -0.01 | 2E-04 | 0        |
| cg05729480 | chr17:75276428 |                   | 9-Sep North shore | -1.9  | 0.083 | 2.44E-38 |
| cg06927323 | chr17:45809727 | <i>TBX21</i>      | North shore       | 0.002 | 1E-04 | 0        |
| cg10587449 | chr17:2654874  | -                 | South shore       | 0.01  | 3E-04 | 0        |
| cg10817916 | chr17:61511069 | <i>CYB561</i>     | North shore       | 0.002 | 2E-04 | 0        |
| cg12054453 | chr17:57915717 | <i>TMEM49</i>     | Open sea          | 0.004 | 3E-04 | 0        |



Continuation of Table S 3.6

|            |                |                    |             |       |       |          |
|------------|----------------|--------------------|-------------|-------|-------|----------|
| cg12065777 | chr17:16311838 | -                  | South shore | -0.01 | 2E-04 | 0        |
| cg13165009 | chr17:32905661 | <i>C17orf102</i>   | North shore | 0.005 | 1E-04 | 0        |
| cg13526488 | chr17:78793807 | <i>RPTOR</i>       | North shelf | -0    | 2E-04 | 0        |
| cg13672106 | chr17:38984421 | <i>TMEM99</i>      | Open sea    | -2.61 | 0.079 | 4.33E-51 |
| cg15379400 | chr17:38655492 | <i>TNS4</i>        | Open sea    | 0.003 | 2E-04 | 0        |
| cg17178468 | chr17:43049442 | -                  | South shore | 0.009 | 2E-04 | 0        |
| cg17288560 | chr17:37720009 | -                  | Open sea    | 0.004 | 3E-04 | 0        |
| cg17615115 | chr17:60905352 | -                  | Open sea    | -0.01 | 3E-04 | 0        |
| cg18041642 | chr17:62062565 | -                  | Open sea    | -0    | 3E-04 | 0        |
| cg18219562 | chr17:41773643 | -                  | Open sea    | 0.002 | 2E-04 | 0        |
| cg18318818 | chr17:48640257 | <i>CACNA1G</i>     | South shore | 0.009 | 2E-04 | 0        |
| cg19779810 | chr17:39969264 | <i>SC65;FKBP10</i> | South shore | 0.005 | 2E-04 | 0        |
| cg20212624 | chr17:40123227 | <i>CNP</i>         | South shelf | -0.01 | 2E-04 | 0        |
| cg21626848 | chr17:39969267 | <i>SC65;FKBP10</i> | South shore | 0.006 | 2E-04 | 0        |
| cg22078805 | chr17:42432046 | <i>FAM171A2</i>    | Island      | 0.004 | 1E-04 | 0        |
| cg22761431 | chr17:7609416  | <i>EFNB3</i>       | South shore | 0.002 | ##### | 0        |
| cg23398508 | chr17:70536199 | -                  | North shore | 0.004 | 2E-04 | 0        |
| cg23630423 | chr17:1903022  | <i>RTN4RL1</i>     | South shore | 0.006 | 3E-04 | 0        |
| cg23838005 | chr17:74868604 | <i>MGAT5B</i>      | South shore | -0.01 | 3E-04 | 0        |
| cg25393494 | chr17:17109936 | <i>PLD6</i>        | Island      | -0.01 | 3E-04 | 0        |
| cg26211349 | chr17:77404555 | <i>HRNBP3</i>      | Open sea    | -0.01 | 2E-04 | 0        |
| cg13431464 | chr18:12880767 | <i>PTPN2</i>       | North shelf | 0.003 | 1E-04 | 0        |
| cg01468567 | chr19:49843922 | <i>TEAD2</i>       | South shore | -0    | 3E-04 | 4.70E-06 |
| cg02489552 | chr19:15121531 | <i>CCDC105</i>     | North shore | 0.016 | 3E-04 | 0        |
| cg03969996 | chr19:46999055 | <i>PNMAL2</i>      | Island      | 0.008 | 2E-04 | 0        |
| cg04255201 | chr19:46999307 | <i>PNMAL2</i>      | South shore | 0.006 | 1E-04 | 0        |
| cg04689048 | chr19:46999289 | <i>PNMAL2</i>      | Island      | 0.008 | 2E-04 | 0        |
| cg05294243 | chr19:51569106 | <i>KLK13</i>       | South shore | 0.003 | 2E-04 | 0        |
| cg05923226 | chr19:15121516 | <i>CCDC105</i>     | North shore | 0.01  | ##### | 0        |
| cg06130787 | chr19:51523550 | <i>KLK10</i>       | South shore | NA    | NA    | 4.67E-36 |
| cg06381803 | chr19:46119475 | <i>EML2</i>        | Island      | 0.012 | 4E-04 | 0        |
| cg06462065 | chr19:55885168 | -                  | North shelf | 0.003 | 2E-04 | 0        |

Continuation of Table S 3.6

|            |                |                      |             |       |       |          |
|------------|----------------|----------------------|-------------|-------|-------|----------|
| cg06673536 | chr19:45213941 | <i>CEACAM16</i>      | South shelf | 0.003 | 2E-04 | 0        |
| cg08263398 | chr19:17919173 | <i>B3GNT3</i>        | South shore | -0    | 1E-04 | 6.10E-11 |
| cg08620095 | chr19:51530970 | <i>KLK11</i>         | Open sea    | -0.08 | 0.015 | 1.70E-08 |
| cg09377704 | chr19:7936683  | -                    | North shore | 0.002 | 2E-04 | 0        |
| cg09470638 | chr19:15121509 | <i>CCDC105</i>       | North shore | 0.013 | 3E-04 | 0        |
| cg09829382 | chr19:3286070  | <i>BRUNOL5</i>       | Island      | 0.003 | 1E-04 | 0        |
| cg10243322 | chr19:5131279  | <i>KDM4B</i>         | North shelf | 0.006 | 2E-04 | 0        |
| cg10592521 | chr19:1560658  | <i>MEX3D</i>         | South shelf | -0    | 2E-04 | 0        |
| cg11410682 | chr19:46999224 | <i>PNMAL2</i>        | Island      | 0.006 | 2E-04 | 0        |
| cg11902728 | chr19:35786580 | <i>MAG</i>           | Island      | 0.008 | 2E-04 | 0        |
| cg12019614 | chr19:11353996 | <i>DOCK6</i>         | Island      | NA    | NA    | 6.74E-26 |
| cg13480898 | chr19:10195914 | <i>C19orf66</i>      | North shore | 2.5   | 0.29  | 1.10E-17 |
| cg13978777 | chr19:2491825  | -                    | North shelf | -0.01 | ##### | 0        |
| cg16617301 | chr19:46999109 | <i>PNMAL2</i>        | Island      | 0.009 | 2E-04 | 0        |
| cg16668176 | chr19:5308893  | <i>PTPRS</i>         | Open sea    | -0    | 2E-04 | 0        |
| cg17865555 | chr19:49133647 | <i>SPHK2</i>         | Island      | 0.013 | 4E-04 | 0        |
| cg18039296 | chr19:11465809 | <i>LPPR2</i>         | North shore | -0.01 | 1E-03 | 3.20E-10 |
| cg18709349 | chr19:46999274 | <i>PNMAL2</i>        | Island      | 0.003 | ##### | 0        |
| cg19004883 | chr19:14448422 | -                    | South shelf | 0.002 | 2E-04 | 0        |
| cg19536401 | chr19:6504797  | -                    | South shelf | -0.01 | ##### | 0        |
| cg19835136 | chr19:17650995 | <i>FAM129C</i>       | Island      | 0.006 | 2E-04 | 0        |
| cg19839325 | chr19:18761132 | <i>KLHL26</i>        | North shore | 0.005 | 2E-04 | 0        |
| cg19839825 | chr19:7616011  | <i>PNPLA6</i>        | South shore | 0.004 | 2E-04 | 0        |
| cg21519787 | chr19:49133421 | <i>SPHK2</i>         | Island      | 0.015 | 5E-04 | 0        |
| cg22816857 | chr19:11565391 | <i>ELAVL3</i>        | Island      | -0.02 | 0.002 | 4.70E-15 |
| cg23680518 | chr19:36020232 | <i>SBSN</i>          | North shelf | -0    | 2E-04 | 0        |
| cg23846712 | chr19:3607608  | <i>TBXA2R</i>        | South shore | NA    | NA    | 1.08E-16 |
| cg24199203 | chr19:3387811  | <i>NFIC</i>          | North shore | -0.01 | 3E-04 | 0        |
| cg24718465 | chr19:46999118 | <i>PNMAL2</i>        | Island      | 0.006 | 2E-04 | 0        |
| cg25509871 | chr19:40871557 | <i>PLD3</i>          | North shore | 0.003 | 1E-04 | 0        |
| cg26288449 | chr19:56048453 | <i>SBK2</i>          | South shore | 0.008 | 3E-04 | 0        |
| cg26750002 | chr19:49244674 | <i>IZUMO1;RASIP1</i> | South shore | 0.004 | 3E-04 | 0        |

Continuation of Table S 3.6

|            |                 |                      |             |       |       |          |
|------------|-----------------|----------------------|-------------|-------|-------|----------|
| cg26931208 | chr19:49244592  | <i>IZUMO1;RASIP1</i> | South shore | 0.004 | 3E-04 | 0        |
| cg26994484 | chr19:1873591   | -                    | South shore | -0.01 | 2E-04 | 0        |
| cg00572387 | chr20:1472914   | <i>SIRPB2</i>        | Open sea    | NA    | NA    | 8.54E-32 |
| cg01735357 | chr20:61732467  | <i>HAR1B;HAR1A</i>   | North shore | 0.011 | ##### | 0        |
| cg03843031 | chr20:44440947  | <i>UBE2C</i>         | Island      | -0    | ##### | 1.61E-07 |
| cg10546626 | chr20:57424521  | <i>GNASAS;GNAS</i>   | North shelf | -0    | 2E-04 | 0        |
| cg11035100 | chr20:31870935  | <i>C20orf114</i>     | Open sea    | -0.03 | 0.002 | 0        |
| cg11667451 | chr20:42543232  | <i>TOX2</i>          | Island      | 5E-04 | NA    | 1.60E-17 |
| cg18062205 | chr20:49630819  | <i>KCNG1</i>         | South shelf | 0.006 | 3E-04 | 0        |
| cg18942298 | chr20:42809578  | <i>JPH2</i>          | Open sea    | 0.01  | 3E-04 | 0        |
| cg19753867 | chr20:35383166  | <i>DSN1</i>          | Open sea    | -0.01 | 2E-04 | 0        |
| cg25779645 | chr20:42965025  | <i>R3HDML</i>        | Open sea    | 0.01  | 2E-04 | 0        |
| cg25799433 | chr20:1600623   | <i>SIRPB1</i>        | Open sea    | 1E-03 | NA    | 3.80E-70 |
| cg16057826 | chr21:36041334  | <i>CLIC6</i>         | Island      | 0.011 | 3E-04 | 0        |
| cg21505925 | chr21:38807423  | <i>DYRK1A</i>        | Open sea    | NA    | NA    | 2.60E-37 |
| cg26390598 | chr21:41032396  | <i>B3GALT5</i>       | Open sea    | -0.02 | 4E-04 | 0        |
| cg01321673 | chr22:50438363  | <i>IL17REL</i>       | Island      | 0.007 | 2E-04 | 0        |
| cg03023089 | chr22:36806580  | -                    | Island      | 0.001 | 2E-04 | 1.90E-10 |
| cg07252792 | chr22:21987048  | <i>CCDC116</i>       | South shelf | 0.002 | 2E-04 | 0        |
| cg08241514 | chr22:50438769  | <i>IL17REL</i>       | Island      | 0.002 | 2E-04 | 0        |
| cg21064451 | chr22:38071534  | <i>LGALS1</i>        | North shore | -0    | 2E-04 | 0        |
| cg25784136 | chr22:21987106  | <i>CCDC116</i>       | South shelf | 0.003 | ##### | 0        |
| cg07757007 | chr2:138734074  | <i>HNMT</i>          | Open sea    | 0.018 | 0.005 | 9.50E-05 |
| cg11652343 | chr9:96919970   | -                    | Open sea    | NA    | NA    | 1.70E-06 |
| cg08547137 | chr20:47367586  | <i>PREX1</i>         | Open sea    | -2.3  | 0.382 | 1.30E-09 |
| cg01021271 | chr11:82268518  | -                    | Open sea    | 0.68  | 0.163 | 3.60E-05 |
| cg01695994 | chr17:80246403  | -                    | North shelf | -0    | 2E-04 | 0        |
| cg14963371 | chr1:27719463   | <i>GPR3</i>          | South shore | 0.003 | 2E-04 | 0        |
| cg05235171 | chr11:117958104 | <i>TMPPSS4</i>       | Open sea    | -0    | 2E-04 | 4.00E-13 |
| cg16423910 | chr19:49843627  | <i>CD37</i>          | Island      | 0.002 | NA    | 5.20E-12 |
| cg00724111 | chr11:47209415  | <i>PACSN3</i>        | South shore | -0.01 | 0.001 | 3.40E-14 |
| cg26299044 | chr17:39021588  | <i>KRT12</i>         | South shore | -0    | 2E-04 | 0        |

Continuation of Table S 3.6

|            |                |                     |             |       |       |           |
|------------|----------------|---------------------|-------------|-------|-------|-----------|
| cg11359094 | chr1:29172578  | <i>OPRD1</i>        | Open sea    | 0.033 | 0.007 | 4.90E-06  |
| cg27455444 | chr19:8268929  | -                   | Open sea    | 5E-04 | ##### | 1.50E-05  |
| cg19142740 | chr19:44139323 | <i>CADM4</i>        | Open sea    | 1.3   | 0.308 | 3.60E-05  |
| cg03834786 | chr7:80571772  | -                   | Open sea    | -1.6  | 0.186 | 5.10E-17  |
| cg09082427 | chr9:140349184 | <i>NSMF</i>         | South shore | -0.12 | 0.028 | 1.30E-05  |
| cg12608784 | chr9:140349197 | <i>NSMF</i>         | South shore | -0.06 | 0.015 | 4.70E-05  |
| cg24470403 | chr14:52733119 | <i>PTGDR</i>        | North shore | -0.04 | 0.006 | 6.60E-08  |
| cg19748455 | chr17:76274856 | <i>LOC100996291</i> | Open sea    | -6    | 0.231 | 1.50E-144 |
| cg21748207 | chr2:26395359  | <i>FAM59B</i>       | North shore | 4E-04 | ##### | 2.30E-07  |
| cg25006851 | chr10:76622627 | <i>MYST4</i>        | Open sea    | -0    | 2E-04 | 0         |
| cg14320320 | chr3:16577787  | -                   | Open sea    | -0    | ##### | 0         |
| cg27383544 | chr5:55813120  | <i>C5orf67</i>      | Open sea    | -5.6  | 0.573 | 8.60E-23  |
| cg00362690 | chr15:69447139 | <i>MIR548H4</i>     | Open sea    | 1.8   | 0.169 | 7.60E-26  |
| cg10254445 | chr1:37197260  | -                   | Open sea    | NA    | NA    | 8.00E-08  |
| cg06402858 | chr9:91692642  | <i>SHC3</i>         | Open sea    | 1.9   | 0.207 | 2.30E-19  |
| cg13735166 | chr3:9680117   | -                   | Open sea    | 2.1   | 0.253 | 2.60E-16  |
| cg19208876 | chr10:79093530 | <i>KCNMA1</i>       | Open sea    | 2.5   | 0.188 | 3.20E-39  |
| cg20040976 | chr15:70766226 | -                   | Open sea    | 2.6   | 0.404 | 4.80E-09  |
| cg20677824 | chr5:106724406 | <i>EFNA5</i>        | Open sea    | -3.2  | 0.262 | 1.10E-34  |
| cg07328569 | chr1:61524438  | -                   | Open sea    | -3.6  | 0.266 | 5.80E-41  |
| cg18152558 | chr4:75956216  | <i>PARM1</i>        | Open sea    | 1.8   | 0.161 | 4.20E-29  |
| cg19750024 | chr8:22068189  | <i>BMP1</i>         | Open sea    | -3.7  | 0.25  | 2.00E-49  |
| cg16125894 | chr1:180915357 | -                   | Open sea    | 2.1   | 0.28  | 2.50E-14  |
| cg25444821 | chr2:26952869  | <i>KCNK3</i>        | South shore | -0.01 | 1E-03 | 0         |
| cg08268947 | chr4:7710890   | <i>SORCS2</i>       | Open sea    | NA    | NA    | 8.90E-05  |
| cg21732535 | chr2:209224280 | <i>PTH2R</i>        | North shore | 0.025 | 0.005 | 7.60E-08  |
| cg12303649 | chr13:51106368 | -                   | Open sea    | NA    | NA    | 4.90E-05  |
| cg13793503 | chr2:162762397 | <i>SLC4A10</i>      | Open sea    | NA    | NA    | 6.50E-05  |
| cg24668753 | chr4:142385640 | -                   | Open sea    | 0.99  | 0.202 | 9.10E-07  |
| cg03550951 | chr1:241111248 | <i>RGS7</i>         | Open sea    | NA    | NA    | 1.70E-08  |
| cg03189652 | chr2:209224422 | <i>PTH2R</i>        | North shore | 0.029 | 0.004 | 1.30E-10  |
| cg20015040 | chr6:138226252 | -                   | Open sea    | -3.6  | 0.318 | 2.40E-29  |

Continuation of Table S 3.6

|            |                 |                  |             |       |       |           |
|------------|-----------------|------------------|-------------|-------|-------|-----------|
| cg09661126 | chr18:46238526  | <i>CTIF</i>      | Open sea    | 0.022 | 0.003 | 1.90E-13  |
| cg12546556 | chr12:104053179 | <i>STAB2</i>     | Open sea    | NA    | NA    | 3.00E-06  |
| cg14580608 | chr17:17579573  | <i>SMCR2</i>     | Open sea    | NA    | NA    | 3.70E-06  |
| cg04484639 | chr3:81773021   | <i>GBE1</i>      | Open sea    | 0.024 | 0.005 | 6.50E-07  |
| cg03864443 | chr7:120190397  | <i>KCND2</i>     | Open sea    | -0.03 | 0.003 | 7.50E-15  |
| cg06237102 | chr12:96472140  | -                | Open sea    | NA    | NA    | 1.50E-08  |
| cg11757214 | chr12:132291338 | -                | North shelf | NA    | NA    | 1.50E-10  |
| cg16881282 | chr13:101535847 | <i>NALCN-AS1</i> | Open sea    | NA    | NA    | 3.40E-05  |
| cg12230203 | chr20:48242474  | -                | Open sea    | -0.97 | 0.169 | 8.80E-09  |
| cg15600670 | chr15:98824655  | -                | Open sea    | NA    | NA    | 1.90E-05  |
| cg23372422 | chr1:12061772   | <i>MFN2</i>      | Open sea    | -0.03 | 0.006 | 7.10E-09  |
| cg14525458 | chr19:48896100  | <i>KDELR1</i>    | North shore | 0.002 | 2E-04 | 0         |
| cg16947791 | chr13:111905195 | <i>ARHGEF7</i>   | North shore | -0    | 1E-04 | 2.60E-13  |
| cg25773223 | chr10:3680061   | -                | Open sea    | -0    | ##### | 0         |
| cg00738869 | chr3:138552648  | -                | North shore | -0    | 7E-04 | 6.50E-09  |
| cg13108341 | chr17:11608711  | <i>DNAH9</i>     | Open sea    | 0.169 | 0.03  | 1.40E-08  |
| cg10254445 | NA              | NA               | NA          | NA    | NA    | 5.70E-08  |
| cg00232402 | chr18:77385290  | -                | North shore | 0.001 | 1E-04 | 0         |
| cg00483304 | chr2:64976962   | -                | North shore | 0.011 | 0.002 | 3.10E-09  |
| cg00641210 | chr17:6927329   | <i>BCL6B</i>     | Island      | 0.006 | 2E-04 | 0         |
| cg01791524 | chr1:100389264  | <i>AGL</i>       | Open sea    | 0.002 | NA    | 4.10E-07  |
| cg02244028 | chr3:38992200   | <i>SCN11A</i>    | Open sea    | 0.003 | 1E-04 | 0         |
| cg02634272 | chr16:71459870  | -                | Island      | ##### | ##### | 0         |
| cg03147691 | chr19:5686568   | <i>HSD11B1L</i>  | North shore | 0.001 | NA    | 3.70E-105 |
| cg03221025 | chr16:88955224  | <i>CBFA2T3</i>   | North shore | -0    | 2E-04 | 0         |
| cg03365937 | chr11:45688659  | <i>CHST1</i>     | South shore | 0.004 | ##### | 0         |
| cg03823863 | chr5:1672612    | -                | Open sea    | 0.002 | ##### | 0         |
| cg03833948 | chr12:21678044  | <i>C12orf39</i>  | North shelf | 0.002 | NA    | 3.70E-15  |
| cg04069951 | chr11:2398336   | <i>CD81</i>      | Island      | 0.004 | 2E-04 | 0         |
| cg05224015 | chr7:158235423  | <i>PTPRN2</i>    | Open sea    | 0.003 | 2E-04 | 0         |
| cg05483509 | chr16:2563435   | <i>ATP6V0C</i>   | Island      | 0.002 | ##### | 0         |
| cg05663480 | chr1:204803813  | <i>NFASC</i>     | Open sea    | 0.004 | 3E-04 | 0         |

Continuation of Table S 3.6

|            |                |                      |             |       |       |          |
|------------|----------------|----------------------|-------------|-------|-------|----------|
| cg05673287 | chr15:77411982 | <i>SGK269</i>        | Open sea    | -0    | ##### | 2.30E-13 |
| cg05816193 | chr6:26018127  | <i>HIST1H1A</i>      | North shelf | 0.002 | 2E-04 | 0        |
| cg07303548 | chr11:67231878 | <i>TMEM134</i>       | North shore | -0    | 2E-04 | 4.70E-08 |
| cg07831886 | chr16:9278007  | -                    | Open sea    | -0    | ##### | 0        |
| cg08373187 | chr5:134912048 | <i>CXCL14</i>        | North shelf | 0.005 | 2E-04 | 0        |
| cg09134819 | chr5:5497464   | -                    | South shelf | 0.003 | 2E-04 | 0        |
| cg10003262 | chr17:1106589  | -                    | North shore | 0.004 | 3E-04 | 0        |
| cg10766021 | chr7:148983373 | <i>ZNF783</i>        | South shore | -0    | 1E-04 | 0        |
| cg11115502 | chr2:24299379  | <i>SF3B14</i>        | Island      | -0    | ##### | 0        |
| cg11453712 | chr9:139741803 | <i>PHPT1</i>         | Island      | 0.002 | 1E-04 | 0        |
| cg11848931 | chr10:25210092 | <i>PRTFDC1</i>       | Open sea    | -0.01 | 3E-04 | 0        |
| cg11854860 | chr4:6384165   | <i>PPP2R2C</i>       | South shelf | -0    | 2E-04 | 3.20E-11 |
| cg13620631 | chr9:87489528  | <i>NTRK2</i>         | Open sea    | -0.01 | 0.001 | 4.60E-14 |
| cg13908635 | chr2:219679764 | <i>CYP27A1</i>       | Open sea    | -0    | 1E-04 | 1.50E-06 |
| cg14484792 | chr14:38511543 | -                    | Open sea    | -0    | 2E-04 | 0        |
| cg14566801 | chr2:209225123 | -                    | Island      | -0    | 2E-04 | 4.00E-09 |
| cg14583825 | chr19:54393040 | <i>PRKCG</i>         | Island      | 0.001 | 1E-04 | 0        |
| cg14913407 | chr1:227953546 | <i>SNAP47</i>        | Open sea    | -0    | 2E-04 | 0        |
| cg15695956 | chr10:11826122 | -                    | Open sea    | -0    | 1E-04 | 0        |
| cg16411103 | chr13:33752544 | <i>STARD13</i>       | Open sea    | -0    | 2E-04 | 0        |
| cg16493238 | chr11:7928236  | -                    | Open sea    | 0.002 | 3E-04 | 4.20E-09 |
| cg16617774 | chr5:146435656 | <i>PPP2R2B</i>       | Open sea    | 0.004 | 2E-04 | 0        |
| cg16857134 | chr19:46712292 | <i>DKFZp434J0226</i> | Open sea    | -0    | 2E-04 | 2.30E-11 |
| cg17417600 | chr14:53258566 | <i>GNPNAT1</i>       | South shore | 0.001 | 2E-04 | 0        |
| cg17442683 | chr1:8664311   | <i>RERE</i>          | Open sea    | -0.01 | 3E-04 | 0        |
| cg18255353 | chr4:154712422 | -                    | Island      | 5E-04 | ##### | 0        |
| cg18994250 | chr4:6107021   | <i>JAKMIP1</i>       | North shore | NA    | NA    | 4.80E-19 |
| cg19108718 | chr5:94982330  | <i>RFESD</i>         | Island      | 0.001 | 1E-04 | 0        |
| cg19820543 | chr6:168120556 | -                    | North shore | -0    | ##### | 3.90E-06 |
| cg20080548 | chr19:36020200 | <i>SBSN</i>          | North shelf | -0    | 1E-04 | 0        |
| cg20533553 | chr11:2464822  | <i>KCNQ1</i>         | North shore | -0.01 | 0.001 | 1.30E-15 |
| cg20910352 | chr12:53817598 | <i>AMHR2</i>         | Open sea    | ##### | ##### | 2.00E-06 |

Continuation of Table S 3.6

|            |                 |                           |             |       |       |          |
|------------|-----------------|---------------------------|-------------|-------|-------|----------|
| cg21526749 | chr3:187389513  | <i>SST</i>                | South shore | -0    | 2E-04 | 0        |
| cg22260782 | chr4:36285339   | <i>DTHD1</i>              | Open sea    | -0    | 3E-04 | 8.40E-13 |
| cg22623303 | chr19:6415017   | <i>KHSRP</i>              | Island      | 3E-04 | NA    | 4.80E-10 |
| cg23054584 | chr19:36822311  | -                         | North shore | 7E-04 | ##### | 0        |
| cg23465295 | chr5:20611854   | -                         | Open sea    | -0.03 | 0.003 | 6.00E-16 |
| cg24302752 | chr6:80853625   | <i>BCKDHB</i>             | Open sea    | -0.01 | 3E-04 | 0        |
| cg27335600 | chr3:53528857   | <i>CACNA1D</i>            | Island      | 1E-03 | ##### | 0        |
| cg01233571 | chr5:59947488   | <i>DEPDC1B</i>            | Open sea    | 0.91  | 0.214 | 2.00E-05 |
| cg01370241 | chr11:61049520  | <i>VWCE</i>               | Open sea    | 0.002 | NA    | 1.20E-23 |
| cg03505501 | chr11:115371789 | <i>CADM1</i>              | North shore | 0.002 | NA    | 7.30E-18 |
| cg12804278 | chr10:23633326  | <i>C10orf67</i>           | Open sea    | ##### | NA    | 1.50E-98 |
| cg15320277 | chr14:79316418  | <i>NRXN3</i>              | Open sea    | 0.002 | NA    | 1.50E-05 |
| cg12261095 | chr19:7764345   | <i>FCER2</i>              | Open sea    | NA    | NA    | 4.10E-13 |
| cg22290528 | chr15:72669149  | <i>HEXA;C15orf34</i>      | South shore | NA    | NA    | 1.00E-08 |
| cg18548879 | chr15:41914454  | -                         | South shore | 0.003 | NA    | 5.70E-05 |
| cg26810816 | chr20:708630    | -                         | Island      | 0.002 | NA    | 1.60E-05 |
| cg27491025 | chr19:10946246  | <i>TMED1</i>              | North shore | 0.006 | 9E-04 | 2.90E-09 |
| cg03793277 | chr19:45416910  | <i>APOC1</i>              | Open sea    | 1.7   | 0.245 | 3.90E-12 |
| cg24931191 | chr3:36248859   | -                         | Open sea    | 0.024 | 0.005 | 1.60E-07 |
| cg18228693 | chr19:39810872  | -                         | North shore | 0.007 | 0.001 | 7.70E-07 |
| cg19046904 | chr22:46426201  | -                         | South shelf | 0.052 | 0.01  | 3.20E-07 |
| cg11731280 | chr9:140066934  | <i>TMEM210</i>            | North shore | 0.123 | 0.02  | 2.30E-09 |
| cg27165960 | chr5:39177381   | <i>FYB</i>                | Open sea    | -0.07 | 0.013 | 5.40E-09 |
| cg15350922 | chr12:46054221  | -                         | Open sea    | 0.093 | 0.016 | 8.90E-09 |
| cg19832381 | chr6:27776524   | <i>HIST1H2BL;HIST1H3H</i> | South shore | -0.05 | 0.004 | 2.20E-34 |
| cg07663850 | chr7:5517419    | <i>FBXL18</i>             | North shelf | -0.08 | 0.008 | 3.20E-24 |
| cg14768892 | chr15:39915931  | <i>FSIP1</i>              | Open sea    | -0.04 | 0.005 | 7.50E-16 |
| cg04664595 | chr14:95147289  | -                         | Open sea    | -0.02 | 0.003 | 1.40E-11 |
| cg25991374 | chr1:178476197  | -                         | Open sea    | 0.033 | 0.005 | 2.20E-11 |
| cg14455466 | chr5:134895387  | -                         | Open sea    | -0.03 | 0.004 | 7.70E-12 |
| cg15133043 | chr17:29405668  | -                         | Open sea    | 0.063 | 0.006 | 6.00E-27 |
| cg22486366 | chr1:20479732   | -                         | Open sea    | -0.05 | 0.006 | 9.40E-18 |

Continuation of Table S 3.6

|            |                |                     |             |       |       |          |
|------------|----------------|---------------------|-------------|-------|-------|----------|
| cg09956442 | chr19:18133156 | -                   | Open sea    | -3.2  | 0.26  | 3.30E-35 |
| cg27086356 | chr3:170627562 | <i>EIF5A2</i>       | South shore | -1.8  | 0.194 | 4.30E-20 |
| cg05371159 | chr9:94342441  | -                   | Open sea    | NA    | NA    | 6.60E-08 |
| cg08428972 | chr19:18133184 | -                   | Open sea    | -4.2  | 0.349 | 2.30E-33 |
| cg26485048 | chr7:7951246   | -                   | Open sea    | 2     | 0.237 | 1.60E-16 |
| cg03221779 | chr14:99897083 | <i>SETD3</i>        | Open sea    | 2.6   | 0.37  | 1.00E-12 |
| cg01892389 | chr6:158448301 | <i>SYNJ2</i>        | Open sea    | 0.039 | 0.007 | 8.50E-09 |
| cg06520829 | chr1:21680109  | -                   | Open sea    | 2.4   | 0.284 | 1.30E-17 |
| cg01297530 | chr21:35724208 | -                   | Open sea    | 2.6   | 0.364 | 8.20E-13 |
| cg19703023 | chr17:8323910  | -                   | Open sea    | -0.03 | 0.005 | 7.30E-08 |
| cg06014707 | chr8:75667618  | <i>MIR2052HG</i>    | Open sea    | -0.02 | 0.003 | 2.50E-13 |
| cg06918740 | chr19:18509973 | -                   | Open sea    | -2    | 0.19  | 3.70E-25 |
| cg17885144 | chr2:165245254 | -                   | Open sea    | NA    | NA    | 3.60E-06 |
| cg16993186 | chr10:11224102 | <i>CELF2</i>        | Open sea    | 3.2   | 0.24  | 2.10E-40 |
| cg09081665 | chr4:84187982  | <i>COQ2</i>         | Open sea    | -0.05 | 0.006 | 2.50E-17 |
| cg15706574 | chr6:46231809  | <i>RCAN2</i>        | Open sea    | 0.017 | 0.003 | 1.30E-08 |
| cg05202858 | chr3:191773926 | -                   | Open sea    | -0.02 | 0.002 | 4.60E-10 |
| cg12384562 | chr2:205707683 | <i>PARD3B</i>       | Open sea    | -0.03 | 0.004 | 9.30E-12 |
| cg16698186 | chr8:72836376  | <i>MSC-AS1</i>      | Open sea    | -2.3  | 0.295 | 1.90E-14 |
| cg23608915 | chr2:102333795 | <i>MAP4K4</i>       | Open sea    | 3.9   | 0.27  | 3.40E-47 |
| cg27325787 | chr16:23960325 | <i>PRKCB</i>        | Open sea    | 0.027 | 0.005 | 2.70E-08 |
| cg01135863 | chr4:17788258  | -                   | Open sea    | 2.9   | 0.387 | 4.20E-14 |
| cg24832862 | chr11:60138281 | -                   | Open sea    | -0.03 | 0.005 | 5.30E-13 |
| cg09653610 | chr20:46631362 | -                   | Open sea    | 2     | 0.297 | 1.20E-11 |
| cg16198075 | chr4:6418840   | <i>PPP2R2C</i>      | Open sea    | 0.036 | 0.005 | 2.90E-15 |
| cg08119153 | chr14:78913301 | <i>NRXN3</i>        | Open sea    | 0.154 | 0.018 | 1.40E-12 |
| cg00753112 | chr8:136593301 | <i>KHDRBS3</i>      | Open sea    | 0.03  | 0.004 | 3.70E-12 |
| cg03231199 | chr5:8634333   | -                   | Open sea    | 0.135 | 0.023 | 6.90E-08 |
| cg01446044 | chr1:13951012  | -                   | Open sea    | -0.06 | 0.007 | 2.80E-10 |
| cg04286826 | chr2:99386505  | <i>LOC101927070</i> | Open sea    | -0.03 | 0.007 | 9.50E-06 |
| cg16421283 | chr1:199009523 | <i>LINC01221</i>    | Open sea    | 0.85  | 0.197 | 1.80E-05 |
| cg24810429 | chr15:80911444 | -                   | Open sea    | 0.303 | 0.067 | 6.70E-06 |



Continuation of Table S 3.6

|            |                |                            |             |       |       |          |
|------------|----------------|----------------------------|-------------|-------|-------|----------|
| cg08419476 | chr11:12739670 | <i>TEAD1</i>               | Open sea    | 0.99  | 0.175 | 1.40E-08 |
| cg20741479 | chr11:12768311 | <i>TEAD1</i>               | Open sea    | 2.7   | 0.248 | 6.40E-27 |
| cg08499559 | chr4:146452508 | <i>SMAD1</i>               | Open sea    | -1.1  | 0.223 | 6.20E-07 |
| cg07345969 | chr4:64583157  | -                          | Open sea    | -0.02 | 0.006 | 9.70E-05 |
| cg13309020 | chr4:146402262 | <i>SMAD1</i>               | North shore | 0.095 | 0.021 | 6.10E-06 |
| cg07096868 | chr2:148634326 | <i>ACVR2A</i>              | Open sea    | 0.1   | 0.019 | 7.20E-08 |
| cg21889405 | chr1:154438777 | <i>IL6R</i>                | Open sea    | -0.21 | 0.046 | 5.70E-06 |
| cg25132244 | chr3:195627583 | <i>TNK2</i>                | Open sea    | -0.15 | 0.036 | 2.40E-05 |
| cg08946973 | chr11:15177163 | <i>INSC</i>                | Open sea    | -1.4  | 0.245 | 5.40E-09 |
| cg02582500 | chr16:14416517 | -                          | Open sea    | 2     | 0.31  | 2.20E-10 |
| cg19730422 | chr5:156718560 | <i>CYFIP2</i>              | Open sea    | 3.4   | 0.28  | 4.30E-34 |
| cg07692435 | chr16:29743682 | -                          | Open sea    | 6.2   | 0.683 | 1.30E-14 |
| cg14931378 | chr2:86831346  | <i>RNF103;RNF103-CHMP3</i> | Open sea    | -0.06 | 0.006 | 6.30E-18 |
| cg17064161 | chr2:121036791 | <i>RALB</i>                | Open sea    | 2.7   | 0.347 | 7.90E-12 |
| cg06287505 | chr5:40757936  | -                          | South shore | 1.9   | 0.259 | 7.10E-11 |
| cg05788267 | chr7:80459066  | <i>SEMA3C</i>              | Open sea    | 2.5   | 0.343 | 8.10E-11 |
| cg15330438 | chr2:86831407  | <i>RNF103;RNF103-CHMP3</i> | Open sea    | -0.06 | 0.006 | 2.00E-16 |
| cg12170419 | chr20:23367377 | <i>NAPB</i>                | Open sea    | -1.9  | 0.213 | 4.20E-18 |
| cg09765245 | chr2:101926154 | <i>RNF149</i>              | South shore | -2.7  | 0.269 | 3.00E-23 |
| cg03482769 | chr19:10445593 | <i>RAVER1;ICAM3</i>        | Island      | 0.049 | 0.008 | 2.90E-10 |
| cg07236324 | chr17:55986035 | <i>CUEDC1</i>              | Open sea    | -0.05 | 0.006 | 9.50E-14 |
| cg10733692 | chr20:52328616 | -                          | Open sea    | 1.8   | 0.358 | 2.50E-07 |
| cg06394677 | chr16:57636195 | -                          | Open sea    | 3     | 0.597 | 2.30E-06 |
| cg21018834 | chr15:70108281 | -                          | Open sea    | 1.4   | 0.3   | 9.80E-06 |
| cg06225967 | chr17:62511734 | <i>CEP95</i>               | Open sea    | 1.1   | 0.245 | 2.00E-05 |
| cg22439390 | chr19:14480797 | -                          | Open sea    | -1.1  | 0.185 | 2.50E-09 |
| cg26607113 | chr11:44702952 | -                          | Open sea    | 0.065 | 0.006 | 2.60E-17 |
| cg05178561 | chr9:365007    | <i>DOCK8</i>               | Open sea    | -0.06 | 0.007 | 3.60E-14 |
| cg24976744 | chr5:120518655 | -                          | Open sea    | -0.06 | 0.007 | 1.30E-14 |
| cg12227872 | chr16:31298223 | <i>ITGAM</i>               | Open sea    | -0.06 | 0.006 | 8.30E-16 |
| cg23923308 | chr7:37006604  | <i>ELMO1</i>               | Open sea    | -0.05 | 0.007 | 5.80E-11 |
| cg24313472 | chr10:62738287 | <i>RHOBTB1</i>             | Open sea    | -0.05 | 0.006 | 4.80E-13 |

Continuation of Table S 3.6

|            |                 |                           |             |       |       |          |
|------------|-----------------|---------------------------|-------------|-------|-------|----------|
| cg10759796 | chr2:42802828   | <i>MTA3</i>               | Open sea    | 0.053 | 0.008 | 2.50E-09 |
| cg25186292 | chr3:122492971  | <i>HSPBAP1</i>            | Open sea    | -0.05 | 0.005 | 3.00E-16 |
| cg27638835 | chr16:386373    | <i>AXIN1</i>              | Open sea    | 2.8   | 0.205 | 9.10E-41 |
| cg21860785 | chr1:216779666  | <i>ESRRG</i>              | Open sea    | 0.027 | 0.004 | 4.60E-10 |
| cg04211591 | chr2:230159830  | -                         | Open sea    | 2.2   | 0.18  | 1.40E-34 |
| cg07750834 | chr15:71040854  | <i>UACA</i>               | Open sea    | 2     | 0.22  | 7.10E-20 |
| cg26723174 | chr2:16669690   | -                         | Open sea    | -0.05 | 0.005 | 1.50E-19 |
| cg04328144 | chr6:70468286   | <i>LMBRD1</i>             | Open sea    | -0.04 | 0.004 | 5.90E-14 |
| cg27519981 | chr13:23950139  | <i>SACS</i>               | South shore | 2.3   | 0.238 | 3.10E-21 |
| cg22345623 | chr9:125050297  | <i>MRRF</i>               | Open sea    | 2.5   | 0.201 | 6.00E-35 |
| cg22305150 | chr3:171789943  | <i>FNDC3B</i>             | Open sea    | -2.1  | 0.265 | 1.60E-15 |
| cg01230907 | chr2:1522652    | <i>TPO</i>                | Open sea    | -0.05 | 0.006 | 3.70E-14 |
| cg14279641 | chr9:93763689   | -                         | Open sea    | 0.035 | 0.006 | 4.20E-08 |
| cg07964115 | chr12:754928    | <i>NINJ2;LOC100049716</i> | South shelf | -0.04 | 0.005 | 1.00E-10 |
| cg17191060 | chr12:846450    | -                         | Open sea    | 1.1   | 0.179 | 2.00E-09 |
| cg20465027 | chr1:19372076   | -                         | Open sea    | -2    | 0.353 | 8.60E-09 |
| cg25316901 | chr1:11039738   | <i>C1orf127</i>           | Open sea    | 0.033 | 0.005 | 2.10E-08 |
| cg01658566 | chr22:25202611  | <i>SGSM1</i>              | South shore | 0.032 | 0.005 | 5.20E-08 |
| cg27541551 | chr2:230989210  | -                         | Open sea    | -0.03 | 0.003 | 4.40E-17 |
| cg14590560 | chr19:3146674   | <i>GNA15;LOC100996351</i> | Open sea    | -0.03 | 0.003 | 1.10E-13 |
| cg23493133 | chr19:50056114  | -                         | North shelf | 1.9   | 0.217 | 5.20E-19 |
| cg22869522 | chr10:124459711 | <i>C10orf120</i>          | Open sea    | -0.03 | 0.005 | 4.20E-08 |
| cg10799510 | chr1:201236058  | -                         | Open sea    | 0.027 | 0.004 | 1.40E-08 |
| cg22303565 | chr12:124817985 | <i>NCOR2</i>              | North shelf | 2.4   | 0.313 | 2.90E-14 |
| cg15866363 | chr1:21620071   | <i>LOC100506801;ECE1</i>  | South shelf | 0.027 | 0.004 | 3.70E-11 |
| cg09944942 | chr8:130284642  | -                         | Open sea    | -0.03 | 0.003 | 4.80E-17 |
| cg10314788 | chr7:107933976  | <i>NRCAM</i>              | Open sea    | -0.03 | 0.003 | 1.20E-13 |
| cg18610541 | chr10:70767279  | <i>KIAA1279</i>           | Open sea    | -0.03 | 0.003 | 5.10E-13 |
| cg04246882 | chr14:99592323  | -                         | Open sea    | -2.5  | 0.28  | 1.50E-19 |
| cg09176309 | chr19:6022308   | <i>RFX2</i>               | North shelf | -0.02 | 0.004 | 5.60E-08 |
| cg15345567 | chr2:173562256  | -                         | Open sea    | -0.02 | 0.003 | 2.10E-13 |
| cg21515958 | chr3:190161974  | <i>TMEM207</i>            | Open sea    | -0.02 | 0.003 | 1.30E-14 |

Continuation of Table S 3.6

|            |                 |                         |             |       |       |          |
|------------|-----------------|-------------------------|-------------|-------|-------|----------|
| cg15183692 | chr4:40192660   | <i>RHOH</i>             | Open sea    | -0.03 | 0.005 | 4.60E-08 |
| cg12964519 | chr12:49376120  | <i>WNT1</i>             | South shore | -0.03 | 0.004 | 4.10E-11 |
| cg04250661 | chr4:160199879  | <i>RAPGEF2</i>          | Open sea    | 0.023 | 0.003 | 8.10E-14 |
| cg05486252 | chr14:20906926  | -                       | South shelf | -0.02 | 0.003 | 6.00E-11 |
| cg05461268 | chr2:230266391  | <i>DNER</i>             | Open sea    | 0.021 | 0.002 | 2.00E-15 |
| cg22300252 | chr15:47705256  | <i>SEMA6D</i>           | Open sea    | -0.02 | 0.002 | 1.50E-14 |
| cg00155731 | chr15:66981279  | -                       | Open sea    | -0.02 | 0.003 | 4.70E-08 |
| cg05646745 | chr10:135172466 | <i>FUOM</i>             | South shore | -0.02 | 0.003 | 1.90E-08 |
| cg04277753 | chr16:81585553  | <i>CMIP</i>             | Open sea    | 3.1   | 0.303 | 3.90E-24 |
| cg19534475 | chr3:141632139  | <i>ATP1B3</i>           | Open sea    | -0.05 | 0.008 | 2.10E-11 |
| cg06228645 | chr10:24652097  | <i>KIAA1217</i>         | Open sea    | 0.014 | 0.002 | 3.60E-15 |
| cg17646797 | chr1:241970344  | -                       | Open sea    | -0.06 | 0.006 | 9.50E-16 |
| cg22279720 | chr10:33603872  | <i>NRP1</i>             | Open sea    | -0.05 | 0.007 | 2.00E-12 |
| cg06676511 | chr2:8820292    | <i>ID2-AS1</i>          | North shore | 0.053 | 0.006 | 8.30E-13 |
| cg23756527 | chr19:8422410   | -                       | Open sea    | -0.03 | 0.004 | 4.50E-13 |
| cg22994495 | chr17:78662780  | <i>RPTOR</i>            | Open sea    | -0.03 | 0.005 | 1.50E-08 |
| cg16425342 | chr19:1509775   | <i>ADAMTSL5</i>         | North shore | 1.4   | 0.256 | 4.40E-08 |
| cg16505573 | chr19:34573738  | -                       | Open sea    | 0.028 | 0.003 | 2.90E-12 |
| cg17656881 | chr17:79202227  | <i>ENTHD2</i>           | Open sea    | -0.03 | 0.004 | 3.20E-08 |
| cg22215119 | chr3:14822735   | -                       | Open sea    | -0.03 | 0.003 | 1.30E-11 |
| cg11940368 | chr15:65843703  | <i>HACD3</i>            | Open sea    | 2.6   | 0.311 | 1.80E-17 |
| cg18043014 | chr11:681173    | <i>DEAF1</i>            | Open sea    | 0.02  | 0.003 | 4.10E-08 |
| cg05438461 | chr15:40401720  | <i>BMF</i>              | South shore | -0.01 | 0.002 | 4.10E-08 |
| cg01427126 | chr16:3741213   | <i>TRAP1</i>            | Open sea    | 2.1   | 0.23  | 2.80E-19 |
| cg19560945 | chr12:126467559 | <i>LINC00939</i>        | Open sea    | 0.064 | 0.016 | 7.10E-05 |
| cg18616702 | chr1:150523808  | <i>ADAMTSL4;MIR4257</i> | South shore | -37.5 | 9.1   | 4.10E-05 |
| cg08282375 | chr20:48883904  | <i>LINC01272</i>        | Open sea    | -2.1  | 0.227 | 1.10E-19 |
| cg11962116 | chr3:152889183  | -                       | Open sea    | -1.1  | 0.239 | 2.50E-06 |
| cg07346426 | chr2:153587510  | <i>ARL6IP6</i>          | Open sea    | -2.87 | 0.709 | 6.00E-05 |
| cg13012015 | chr2:181983380  | -                       | Open sea    | -3.75 | 0.915 | 4.80E-05 |
| cg08559711 | chr6:20978738   | <i>CDKAL1</i>           | Open sea    | -2.1  | 0.203 | 1.60E-24 |
| cg00242263 | chr4:80132081   | <i>LINC01088</i>        | Open sea    | 1.3   | 0.248 | 2.20E-07 |

Continuation of Table S 3.6

|            |                 |                        |             |       |       |          |
|------------|-----------------|------------------------|-------------|-------|-------|----------|
| cg10965265 | chr7:128306480  | -                      | Open sea    | 1.4   | 0.138 | 3.40E-25 |
| cg25520324 | chr14:23352253  | <i>REM2</i>            | North shelf | 0.067 | 0.013 | 5.30E-07 |
| cg06064155 | chr20:47463403  | -                      | Open sea    | -0.06 | 0.01  | 3.00E-10 |
| cg10176717 | chr10:112888974 | -                      | Open sea    | 0.032 | 0.005 | 8.40E-11 |
| cg23513000 | chr1:15530445   | <i>TMEM51</i>          | Open sea    | -0.05 | 0.007 | 4.30E-14 |
| cg16024613 | chr19:17163492  | <i>HAUS8</i>           | Open sea    | -0.06 | 0.009 | 3.80E-11 |
| cg10682119 | chr15:93182494  | <i>FAM174B</i>         | Open sea    | -0.02 | 0.003 | 1.10E-14 |
| cg09139810 | chr12:108998904 | -                      | Open sea    | -0.01 | 0.003 | 3.40E-05 |
| cg04650889 | chr7:77313418   | <i>APTR</i>            | Open sea    | 2.2   | 0.307 | 2.30E-12 |
| cg16875568 | chr13:102173110 | <i>ITGBL1</i>          | Open sea    | -0.01 | 0.002 | 1.10E-05 |
| cg10815470 | chr6:142508602  | <i>VTA1</i>            | Open sea    | 0.084 | 0.018 | 1.60E-06 |
| cg11670749 | chr7:73303364   | -                      | Open sea    | 0.06  | 0.006 | 5.00E-23 |
| cg15243205 | chr6:27859478   | <i>HIST1H3J</i>        | North shore | -3.7  | 0.349 | 1.00E-25 |
| cg15241106 | chr15:69385944  | <i>EWSAT1;MIR548H4</i> | Open sea    | 2.3   | 0.251 | 1.60E-20 |
| cg04814021 | chr17:75693889  | -                      | Open sea    | 0.08  | 0.011 | 6.20E-13 |
| cg16990252 | chr8:19963288   | -                      | Open sea    | -2.2  | 0.206 | 3.90E-26 |
| cg16863678 | chr1:154910368  | <i>PMVK</i>            | South shore | -0.04 | 0.006 | 2.80E-10 |
| cg02205348 | chr5:14465012   | <i>TRIO</i>            | South shelf | 1.7   | 0.21  | 5.00E-16 |
| cg03990746 | chr17:47533409  | -                      | Open sea    | 0.059 | 0.012 | 3.60E-07 |
| cg14608778 | chr15:39933963  | <i>FSIP1</i>           | Open sea    | -0.05 | 0.008 | 7.10E-09 |
| cg04249817 | chr10:135172491 | <i>FUOM</i>            | South shore | 0.073 | 0.015 | 9.50E-07 |
| cg19359526 | chr16:31414218  | <i>ITGAD</i>           | Open sea    | -1.9  | 0.215 | 6.90E-19 |
| cg17480291 | chr5:14464838   | <i>TRIO</i>            | South shelf | 2.2   | 0.274 | 4.30E-15 |
| cg07581623 | chr8:26498830   | <i>DPYSL2</i>          | Open sea    | 1.9   | 0.355 | 1.50E-07 |
| cg23570968 | chr8:30726234   | -                      | Open sea    | -1.5  | 0.286 | 8.10E-08 |
| cg03601752 | chr19:2621851   | <i>GNG7</i>            | North shore | -0.03 | 0.005 | 6.70E-10 |
| cg14864041 | chr17:65796703  | -                      | Open sea    | -1.9  | 0.234 | 2.40E-15 |
| cg20798375 | chr4:23947535   | -                      | Open sea    | 2.7   | 0.357 | 1.60E-14 |
| cg10613567 | chr13:23950128  | <i>SACS</i>            | South shore | 0.038 | 0.007 | 7.20E-09 |
| cg24990672 | chr3:186076897  | <i>DGKG</i>            | North shore | -2.4  | 0.308 | 1.10E-14 |
| cg23836403 | chr11:63383471  | -                      | South shore | -1.4  | 0.277 | 7.90E-07 |
| cg14054075 | chr5:117882587  | <i>HRAT56</i>          | Open sea    | -1.1  | 0.19  | 3.60E-09 |

Continuation of Table S 3.6

|            |                 |                      |             |       |       |          |
|------------|-----------------|----------------------|-------------|-------|-------|----------|
| cg08513698 | chr13:113524002 | <i>ATP11A</i>        | Open sea    | 0.028 | 0.007 | 7.60E-05 |
| cg08694513 | chr10:21497747  | -                    | Open sea    | -1.3  | 0.225 | 1.30E-08 |
| cg04862404 | chr3:174577336  | <i>NAALADL2</i>      | Open sea    | -3.2  | 0.337 | 8.50E-21 |
| cg05218995 | chr21:42851888  | <i>TMPRSS2</i>       | Open sea    | -0.03 | 0.007 | 9.40E-05 |
| cg20821903 | chr16:4313878   | <i>TFAP4</i>         | South shore | 0.037 | 0.007 | 3.00E-08 |
| cg26442250 | chr5:27639425   | -                    | Open sea    | -2.5  | 0.321 | 2.60E-15 |
| cg01907945 | chr3:150726740  | <i>CLRN1-AS1</i>     | Open sea    | 2.3   | 0.225 | 2.90E-24 |
| cg27001818 | chr22:33724826  | <i>LARGE</i>         | Open sea    | -1.3  | 0.244 | 4.10E-08 |
| cg24117197 | chr1:229547029  | -                    | South shelf | -0.02 | 0.006 | 2.90E-05 |
| cg14721232 | chr4:6426056    | <i>PPP2R2C</i>       | Open sea    | 0.064 | 0.007 | 1.50E-18 |
| cg18135341 | chr6:137341992  | <i>IL20RA</i>        | Open sea    | -0.03 | 0.007 | 5.00E-07 |
| cg09043768 | chr22:37943067  | -                    | Open sea    | 0.025 | 0.006 | 2.80E-05 |
| cg21510186 | chr4:6417164    | <i>PPP2R2C</i>       | Open sea    | 1.2   | 0.168 | 6.60E-12 |
| cg21110915 | chr17:8888894   | -                    | Open sea    | -0.03 | 0.006 | 8.90E-07 |
| cg25806524 | chr19:46998786  | <i>PNMAL2;PPP5D1</i> | South shore | 0.94  | 0.21  | 7.10E-06 |
| cg17046263 | chr16:71679605  | <i>PHLPP2</i>        | Open sea    | 1.5   | 0.318 | 2.00E-06 |
| cg24184634 | chr3:159721504  | <i>IL12A-AS1</i>     | Open sea    | -3.4  | 0.444 | 4.90E-14 |
| cg12124090 | chr2:121359919  | -                    | Open sea    | 1.9   | 0.254 | 2.50E-13 |
| cg15254759 | chr1:38763833   | -                    | Open sea    | -1.8  | 0.202 | 8.00E-19 |
| cg04090409 | chr4:153109679  | -                    | Open sea    | 3     | 0.288 | 9.80E-26 |
| cg16227392 | chr14:99501573  | -                    | Open sea    | 2.3   | 0.207 | 5.80E-29 |
| cg07332023 | chr11:47433972  | <i>SLC39A13</i>      | South shelf | -2.2  | 0.361 | 6.30E-10 |
| cg02746002 | chr2:219718531  | -                    | Open sea    | 1.2   | 0.19  | 2.60E-10 |
| cg18212920 | chr11:134430123 | -                    | Open sea    | 2.2   | 0.271 | 1.50E-16 |
| cg16166958 | chr3:189408443  | <i>TP63</i>          | Open sea    | 1     | 0.118 | 1.30E-18 |
| cg09331496 | chr2:121255962  | -                    | Open sea    | 3.3   | 0.303 | 1.80E-27 |
| cg10887181 | chr2:220134324  | <i>TUBA4B</i>        | Open sea    | 0.038 | 0.009 | 5.50E-05 |
| cg18866560 | chr8:49612348   | -                    | Open sea    | -2.1  | 0.425 | 7.30E-07 |
| cg00096898 | chr3:38396823   | <i>XYLB</i>          | Open sea    | 3.1   | 0.317 | 4.70E-23 |
| cg13905211 | chr11:68070255  | -                    | Open sea    | 1.4   | 0.281 | 3.90E-07 |
| cg11312836 | chr1:171122542  | <i>FMO6P</i>         | Open sea    | -1.8  | 0.236 | 1.20E-13 |
| cg19039578 | chr16:11483144  | -                    | Open sea    | -2.2  | 0.216 | 8.00E-24 |

Continuation of Table S 3.6

|            |                 |                         |             |       |       |          |
|------------|-----------------|-------------------------|-------------|-------|-------|----------|
| cg18848606 | chr17:32993105  | -                       | Open sea    | 1.4   | 0.309 | 1.10E-05 |
| cg12101380 | chr10:50517170  | <i>C10orf71</i>         | Open sea    | 2.4   | 0.375 | 1.00E-10 |
| cg15662038 | chr19:13085771  | -                       | Open sea    | 2.2   | 0.354 | 2.50E-10 |
| cg07650421 | chr9:122324712  | -                       | Open sea    | -0.03 | 0.005 | 3.80E-07 |
| cg04413584 | chr9:82183423   | -                       | North shelf | 2.9   | 0.307 | 2.50E-21 |
| cg12062266 | chr2:121355307  | -                       | Open sea    | 2.2   | 0.18  | 1.50E-35 |
| cg25343965 | chr11:113126621 | <i>NCAM1</i>            | Open sea    | -1.4  | 0.256 | 4.20E-08 |
| cg09558509 | chr17:27939837  | <i>ANKRD13B</i>         | Island      | 0.021 | 0.004 | 2.70E-06 |
| cg19402463 | chr9:132872679  | <i>GPR107</i>           | Open sea    | 2.8   | 0.34  | 4.30E-16 |
| cg11126410 | chr3:128841191  | <i>RAB43;ISY1-RAB43</i> | South shore | 0.013 | 0.003 | 3.00E-05 |
| cg16231873 | chr4:7432628    | <i>PSAPL1;SORCS2</i>    | Open sea    | 0.019 | 0.005 | 6.30E-05 |
| cg09148948 | chr2:218115006  | -                       | Open sea    | 0.032 | 0.007 | 1.70E-06 |
| cg11935615 | chr16:15699179  | <i>KIAA0430</i>         | Open sea    | -0.02 | 0.005 | 3.60E-05 |
| cg09503767 | chr9:132947677  | <i>NCS1</i>             | Open sea    | 0.028 | 0.007 | 3.80E-05 |
| cg24365525 | chr2:121227771  | -                       | Open sea    | -0.03 | 0.006 | 7.40E-06 |
| cg01562540 | chr14:56274296  | -                       | Open sea    | 0.022 | 0.005 | 9.50E-06 |
| cg13960680 | chr13:99394346  | <i>SLC15A1</i>          | Open sea    | 1.9   | 0.24  | 5.30E-15 |
| cg24086717 | chr1:54154485   | <i>GLIS1</i>            | Open sea    | -0.03 | 0.006 | 2.00E-06 |
| cg23120547 | chr10:49833689  | <i>ARHGAP22</i>         | Open sea    | -0.04 | 0.008 | 1.40E-07 |
| cg08586855 | chr1:9375862    | <i>SPSB1</i>            | Open sea    | -1.6  | 0.268 | 2.90E-09 |
| cg15934389 | chr6:35362046   | <i>PPARD</i>            | Open sea    | -0.03 | 0.005 | 2.40E-07 |
| cg12485221 | chr5:148421151  | <i>SH3TC2</i>           | Open sea    | 2.1   | 0.42  | 4.50E-07 |
| cg05930965 | chr2:12938268   | -                       | Open sea    | 1.2   | 0.179 | 2.90E-11 |
| cg00090792 | chr1:38753536   | -                       | Open sea    | -0.02 | 0.005 | 3.50E-06 |
| cg19567415 | chr17:61327103  | <i>TANC2</i>            | Open sea    | -0.03 | 0.004 | 1.90E-11 |
| cg23070738 | chr6:128260883  | -                       | Open sea    | -0.03 | 0.006 | 8.50E-08 |
| cg23045307 | chr4:141079760  | -                       | Open sea    | -0.04 | 0.007 | 2.40E-08 |
| cg20734611 | chr22:45297139  | <i>PHF21B</i>           | Open sea    | -0.04 | 0.006 | 2.90E-10 |
| cg16786154 | chr12:123205797 | -                       | Open sea    | -0.04 | 0.007 | 7.70E-11 |
| cg17420142 | chr18:32702783  | <i>MAPRE2</i>           | Open sea    | -1.2  | 0.228 | 2.30E-07 |
| cg13266323 | chr11:115204072 | <i>CADM1</i>            | Open sea    | -0.02 | 0.004 | 6.40E-05 |
| cg17192520 | chr14:88461468  | <i>GALC</i>             | South shore | -0.05 | 0.006 | 3.00E-13 |

Continuation of Table S 3.6

|            |                 |                           |             |       |       |          |
|------------|-----------------|---------------------------|-------------|-------|-------|----------|
| cg08897705 | chr9:89382059   | -                         | Open sea    | 2.1   | 0.212 | 1.60E-22 |
| cg22355342 | chr7:133717055  | <i>EXOC4</i>              | Open sea    | 1.1   | 0.209 | 3.80E-07 |
| cg11308227 | chr17:79202435  | <i>ENTHD2</i>             | Open sea    | -0.03 | 0.006 | 3.20E-09 |
| cg17725163 | chr1:21900811   | <i>ALPL</i>               | Open sea    | 2.3   | 0.454 | 5.20E-07 |
| cg07006163 | chr16:1665950   | <i>CRAMP1L</i>            | South shore | -0.02 | 0.005 | 5.10E-05 |
| cg06924548 | chr1:75812997   | <i>SLC44A5</i>            | Open sea    | -0.03 | 0.004 | 1.40E-14 |
| cg24807394 | chr1:162458159  | -                         | Open sea    | 2.9   | 0.324 | 2.00E-19 |
| cg22828783 | chr11:102926593 | <i>DCUN1D5</i>            | Open sea    | -1.9  | 0.329 | 5.10E-09 |
| cg12580228 | chr2:37471322   | <i>NDUFAF7</i>            | Open sea    | -0.03 | 0.006 | 7.60E-07 |
| cg26103696 | chr20:32469992  | -                         | Open sea    | 0.04  | 0.006 | 8.20E-10 |
| cg10241937 | chr3:59825287   | <i>FHIT</i>               | Open sea    | 0.041 | 0.008 | 4.20E-07 |
| cg18849612 | chr12:106496921 | <i>NUAK1</i>              | Open sea    | -2.8  | 0.328 | 4.60E-17 |
| cg08758963 | chr11:10681378  | <i>MRVI1</i>              | Open sea    | -0.03 | 0.007 | 2.10E-06 |
| cg13938970 | chr16:24048204  | <i>PRKCB</i>              | Open sea    | -0.05 | 0.006 | 2.10E-18 |
| cg01323312 | chr7:73476657   | <i>ELN</i>                | Open sea    | -0.04 | 0.006 | 1.30E-09 |
| cg05889085 | chr8:35718561   | -                         | Open sea    | 0.77  | 0.116 | 2.90E-11 |
| cg06984107 | chr2:54172785   | <i>PSME4</i>              | Open sea    | 1     | 0.185 | 8.00E-08 |
| cg14153041 | chr1:22231285   | <i>HSPG2</i>              | Open sea    | 0.058 | 0.005 | 8.20E-27 |
| cg03815267 | chr5:9343843    | <i>SEMA5A</i>             | Open sea    | -0.03 | 0.004 | 1.10E-09 |
| cg26384544 | chr8:120660846  | -                         | Open sea    | -0.02 | 0.005 | 1.70E-05 |
| cg01052993 | chr7:141615491  | -                         | Open sea    | 4.9   | 0.702 | 3.40E-12 |
| cg15649007 | chr1:38979009   | -                         | Open sea    | 0.19  | 0.042 | 6.80E-06 |
| cg15551092 | chr11:124405865 | -                         | Open sea    | -0.05 | 0.006 | 3.70E-13 |
| cg09684262 | chr3:156391685  | <i>TIPARP;TIPARP-AS1</i>  | North shore | 0.029 | 0.006 | 3.40E-07 |
| cg27110790 | chr19:35701964  | -                         | Open sea    | 0.033 | 0.005 | 4.80E-10 |
| cg07657727 | chr1:76775210   | <i>ST6GALNAC3</i>         | Open sea    | -0.04 | 0.008 | 2.00E-07 |
| cg08792629 | chr16:441986    | <i>LOC100134368</i>       | Open sea    | -0.05 | 0.007 | 1.10E-12 |
| cg03452365 | chr17:44178504  | <i>KANSL1</i>             | Open sea    | -0.02 | 0.004 | 6.10E-06 |
| cg04247018 | chr3:159352535  | <i>IQCJ-SCHIP1;SCHIP1</i> | Open sea    | -0.02 | 0.005 | 6.70E-06 |
| cg04894218 | chr5:112614911  | <i>MCC</i>                | Open sea    | -0.02 | 0.003 | 3.90E-07 |
| cg15088991 | chr1:27118222   | <i>PIGV</i>               | South shelf | -0.06 | 0.007 | 2.00E-14 |
| cg05587665 | chr6:170252085  | -                         | Open sea    | -1    | 0.241 | 3.40E-05 |

Continuation of Table S 3.6

|            |                 |                       |             |       |       |          |
|------------|-----------------|-----------------------|-------------|-------|-------|----------|
| cg09502182 | chr20:16195955  | -                     | Open sea    | -0.03 | 0.006 | 3.90E-09 |
| cg01287555 | chr13:41704639  | <i>KBTBD6</i>         | North shore | 1.7   | 0.22  | 2.00E-15 |
| cg14937504 | chr7:131583452  | -                     | Open sea    | 1.8   | 0.273 | 1.90E-11 |
| cg23683730 | chr3:189632845  | -                     | Open sea    | -0.03 | 0.007 | 5.40E-07 |
| cg09958060 | chr12:113524330 | <i>DTX1</i>           | Open sea    | 2.1   | 0.27  | 2.00E-15 |
| cg04582369 | chr3:4856946    | <i>ITPR1</i>          | Open sea    | 0.033 | 0.008 | 2.50E-05 |
| cg16824480 | chr17:74189931  | <i>RNF157</i>         | Open sea    | 0.031 | 0.007 | 2.30E-05 |
| cg17043444 | chr1:20505852   | -                     | Open sea    | -0.02 | 0.003 | 1.30E-05 |
| cg05685272 | chr11:132511149 | <i>OPCML</i>          | Open sea    | -0.03 | 0.006 | 1.30E-05 |
| cg21510141 | chr1:36342676   | -                     | Open sea    | -0.02 | 0.005 | 4.80E-05 |
| cg07920076 | chr20:35224478  | <i>TGIF2-C20orf24</i> | Open sea    | 0.028 | 0.005 | 5.40E-08 |
| cg08476354 | chr2:13059185   | -                     | Open sea    | -0.04 | 0.007 | 3.40E-11 |
| cg23171506 | chr16:57843964  | <i>LOC388282</i>      | Open sea    | 0.046 | 0.01  | 4.40E-06 |
| cg24845774 | chr8:35718698   | -                     | Open sea    | 0.65  | 0.111 | 4.40E-09 |
| cg04779737 | chr17:6082551   | -                     | Open sea    | -0.02 | 0.005 | 7.60E-06 |
| cg14660461 | chr4:147379854  | <i>SLC10A7</i>        | Open sea    | -0.02 | 0.005 | 6.50E-05 |
| cg19343186 | chr1:44984971   | <i>RNF220</i>         | Open sea    | -2.4  | 0.386 | 6.60E-10 |
| cg01520578 | chr19:41536777  | -                     | Open sea    | -0.05 | 0.012 | 7.20E-05 |
| cg00365642 | chr2:8456058    | <i>LINC00299</i>      | Open sea    | -0.07 | 0.016 | 6.50E-06 |
| cg07023764 | chr8:130567184  | <i>CCDC26</i>         | Open sea    | 0.07  | 0.013 | 2.20E-08 |
| cg01953698 | chr2:230148091  | -                     | Open sea    | 1.8   | 0.196 | 5.80E-21 |
| cg03173169 | chr2:75060921   | <i>HK2</i>            | Island      | 0.077 | 0.013 | 4.20E-09 |
| cg12307174 | chr14:68555650  | <i>RAD51B</i>         | Open sea    | -1.1  | 0.149 | 1.70E-14 |
| cg12818967 | chr6:111550921  | -                     | Open sea    | -1.7  | 0.339 | 4.50E-07 |
| cg25839606 | chr9:114364161  | -                     | South shelf | 0.05  | 0.012 | 3.40E-05 |
| cg11990561 | chr12:93293779  | <i>EEA1</i>           | Open sea    | 1.9   | 0.276 | 3.50E-12 |
| cg01218012 | chr13:50772687  | <i>DLEU1</i>          | Open sea    | -4.7  | 0.338 | 1.80E-44 |
| cg02759482 | chr8:40978162   | -                     | Open sea    | -2.1  | 0.388 | 3.80E-08 |
| cg05473085 | chr1:181483325  | <i>CACNA1E</i>        | Open sea    | -0.04 | 0.008 | 3.50E-05 |
| cg07884459 | chr1:246110773  | <i>SMYD3</i>          | Open sea    | -0.04 | 0.008 | 2.20E-05 |
| cg14488626 | chr5:168569433  | <i>SLIT3</i>          | Open sea    | -0.03 | 0.008 | 5.90E-05 |
| cg21810778 | chr1:31251307   | -                     | Open sea    | -2    | 0.413 | 8.50E-07 |



Continuation of Table S 3.6

|            |                |                       |             |       |       |          |
|------------|----------------|-----------------------|-------------|-------|-------|----------|
| cg25847702 | chr2:33591533  | <i>LTBP1</i>          | Open sea    | -0.95 | 0.211 | 7.30E-06 |
| cg19420069 | chr16:65635834 | -                     | Open sea    | 0.067 | 0.015 | 8.10E-06 |
| cg21923568 | chr7:98721688  | <i>SMURF1</i>         | Open sea    | 0.28  | 0.045 | 1.10E-09 |
| cg00200515 | chr3:101592635 | -                     | Open sea    | 0.15  | 0.028 | 3.10E-07 |
| cg02489630 | chr13:78620363 | <i>LINC00446</i>      | Open sea    | 1.5   | 0.184 | 2.50E-15 |
| cg16998839 | chr9:134531983 | <i>RAPGEF1</i>        | Open sea    | 0.15  | 0.032 | 2.40E-06 |
| cg04624603 | chr3:35696119  | <i>ARPP21</i>         | Open sea    | 1.9   | 0.198 | 1.20E-21 |
| cg19246944 | chr2:57926355  | -                     | Open sea    | 2.4   | 0.377 | 9.20E-11 |
| cg03180011 | chr1:16955678  | <i>CROCCP2</i>        | Open sea    | -1.2  | 0.285 | 1.50E-05 |
| cg25107667 | chr20:51042037 | -                     | Open sea    | 0.99  | 0.167 | 3.00E-09 |
| cg19982061 | chr3:46970947  | <i>CCDC12</i>         | Open sea    | 0.09  | 0.018 | 9.20E-07 |
| cg24037821 | chr19:4901791  | <i>UHRF1;ARRDC5</i>   | Open sea    | 1.3   | 0.29  | 1.10E-05 |
| cg09225336 | chr16:65635750 | -                     | Open sea    | 0.065 | 0.015 | 9.10E-06 |
| cg18265326 | chr16:65635738 | -                     | Open sea    | 0.073 | 0.016 | 4.70E-06 |
| cg07475825 | chr6:84577581  | <i>CYB5R4</i>         | Open sea    | -1.8  | 0.276 | 3.60E-11 |
| cg22271479 | chr19:3708509  | <i>TJP3</i>           | Open sea    | 3.3   | 0.374 | 4.20E-19 |
| cg03856144 | chr1:7705567   | <i>CAMTA1</i>         | Open sea    | -3.6  | 0.303 | 8.80E-33 |
| cg11801086 | chr6:39656197  | <i>KIF6</i>           | Open sea    | 2.1   | 0.16  | 1.30E-39 |
| cg07621682 | chr19:41827759 | <i>CCDC97</i>         | Open sea    | -1.7  | 0.226 | 1.60E-14 |
| cg01005523 | chr4:157867829 | <i>PDGFC</i>          | Open sea    | -1.3  | 0.2   | 1.00E-10 |
| cg00234726 | chr2:216756432 | -                     | Open sea    | 2.1   | 0.243 | 7.30E-18 |
| cg09399176 | chr11:15135329 | <i>INSC</i>           | North shore | 1.9   | 0.331 | 1.40E-08 |
| cg05992638 | chr7:21003764  | <i>LINC01162</i>      | Open sea    | 1.6   | 0.143 | 2.40E-28 |
| cg17252620 | chr17:36125170 | -                     | Open sea    | -1.4  | 0.256 | 6.90E-08 |
| cg17929959 | chr18:19873627 | -                     | Open sea    | 1.1   | 0.166 | 6.60E-11 |
| cg01228288 | chr14:99501901 | -                     | Open sea    | 2.6   | 0.273 | 2.50E-21 |
| cg00724936 | chr19:45776576 | <i>MARK4</i>          | Open sea    | 2.3   | 0.32  | 8.10E-13 |
| cg25293430 | chr4:113810918 | <i>ANK2</i>           | Open sea    | 1.9   | 0.23  | 3.20E-16 |
| cg05211864 | chr17:39969239 | <i>P3H4;FKBP10</i>    | South shore | 1.6   | 0.245 | 5.70E-11 |
| cg25935178 | chr1:159869902 | <i>MIR4259;CFAP45</i> | Island      | 0.65  | 0.146 | 8.50E-06 |
| cg09221413 | chr17:39969252 | <i>P3H4;FKBP10</i>    | South shore | 1.4   | 0.218 | 8.00E-11 |
| cg25012097 | chr13:39263863 | <i>FREM2</i>          | South shore | -1.9  | 0.438 | 1.20E-05 |

Continuation of Table S 3.6

|            |                 |                          |             |       |       |          |
|------------|-----------------|--------------------------|-------------|-------|-------|----------|
| cg11955249 | chr6:122403090  | -                        | Open sea    | -1.9  | 0.461 | 3.30E-05 |
| cg23120101 | chr5:112617123  | <i>MCC</i>               | Open sea    | -0.98 | 0.244 | 5.50E-05 |
| cg08712483 | chr3:59411984   | -                        | Open sea    | 1.3   | 0.241 | 8.30E-08 |
| cg14907703 | chr10:122357656 | <i>LINC01561;MIR5694</i> | Open sea    | -1.7  | 0.289 | 3.10E-09 |
| cg11990488 | chr12:2159166   | -                        | Open sea    | 1.4   | 0.295 | 9.40E-07 |
| cg26750073 | chr5:65854680   | -                        | Open sea    | 1.5   | 0.222 | 2.60E-11 |
| cg01933515 | chr14:96532698  | <i>C14orf132</i>         | Open sea    | 1.8   | 0.408 | 7.80E-06 |
| cg21212234 | chr16:65706984  | -                        | Open sea    | 1.6   | 0.249 | 2.80E-10 |
| cg08603845 | chr1:24644596   | <i>GRHL3</i>             | North shelf | 1.9   | 0.417 | 3.30E-06 |
| cg09954323 | chr1:177140494  | <i>BRINP2</i>            | South shore | 2.2   | 0.439 | 4.70E-07 |
| cg02568341 | chr16:50459768  | -                        | Open sea    | 1.2   | 0.243 | 3.70E-07 |
| cg09707395 | chr17:8213604   | <i>ARHGEF15</i>          | Open sea    | 1.1   | 0.263 | 6.40E-05 |
| cg24999870 | chr19:46147326  | <i>EML2</i>              | South shore | 0.96  | 0.246 | 9.70E-05 |
| cg26390944 | chr13:78433828  | <i>EDNRB-AS1</i>         | Open sea    | -1.6  | 0.314 | 2.30E-07 |
| cg12903825 | chr9:131684260  | <i>PHYHD1</i>            | Open sea    | -1.8  | 0.322 | 2.90E-08 |
| cg25093216 | chr5:112430577  | <i>MCC</i>               | Open sea    | 1.4   | 0.32  | 6.20E-06 |
| cg10760040 | chr17:44983068  | -                        | Open sea    | 1.2   | 0.195 | 1.00E-09 |
| cg01891124 | chr9:98836086   | <i>LOC158435</i>         | Open sea    | 1.7   | 0.356 | 1.60E-06 |
| cg13626246 | chr9:87623942   | <i>NTRK2</i>             | Open sea    | 1.7   | 0.375 | 7.90E-06 |
| cg17125270 | chr19:40871421  | <i>PLD3</i>              | North shore | 1.1   | 0.244 | 3.00E-06 |
| cg11270364 | chr8:80672475   | -                        | Open sea    | 1.8   | 0.367 | 1.30E-06 |
| cg16340410 | chr3:12345716   | <i>PPARG</i>             | Open sea    | 1.2   | 0.206 | 4.00E-09 |
| cg09325749 | chr18:57549610  | -                        | Open sea    | 3.8   | 0.699 | 4.20E-08 |
| cg03357219 | chr4:102734940  | <i>BANK1</i>             | Open sea    | 0.53  | 0.106 | 6.20E-07 |
| cg16284880 | chr20:49738043  | -                        | Open sea    | 0.64  | 0.153 | 2.80E-05 |
| cg00600584 | chr3:189347814  | <i>TP63</i>              | Open sea    | 0.61  | 0.145 | 2.80E-05 |
| cg17773655 | chr19:49244121  | <i>RASIP1</i>            | South shore | 1     | 0.205 | 7.70E-07 |
| cg00982484 | chr10:22548542  | -                        | Open sea    | -1.5  | 0.3   | 1.10E-06 |
| cg05340191 | chr13:76080287  | -                        | Open sea    | 1.2   | 0.19  | 1.90E-10 |
| cg21501831 | chr12:108383714 | -                        | Open sea    | 1.1   | 0.276 | 4.10E-05 |
| cg17274885 | chr13:99399450  | <i>SLC15A1</i>           | Open sea    | 1.3   | 0.271 | 1.80E-06 |
| cg00894715 | chr11:34489880  | <i>CAT</i>               | Open sea    | 1.4   | 0.359 | 7.50E-05 |

Continuation of Table S 3.6

|            |                 |                 |             |       |       |          |
|------------|-----------------|-----------------|-------------|-------|-------|----------|
| cg13187651 | chr17:40712736  | <i>COASY</i>    | North shore | 0.66  | 0.13  | 3.20E-07 |
| cg07738234 | chr6:117593953  | <i>VGLL2</i>    | South shore | 0.93  | 0.2   | 3.40E-06 |
| cg11551236 | chr9:37034177   | <i>PAX5</i>     | Island      | 3.6   | 0.495 | 6.30E-13 |
| cg11755558 | chr4:122103800  | <i>TNIP3</i>    | Open sea    | -2.4  | 0.262 | 2.40E-20 |
| cg17205574 | chr15:66427768  | <i>MEGF11</i>   | Open sea    | -2.2  | 0.359 | 3.70E-10 |
| cg15072822 | chr6:168614849  | -               | South shore | 1.3   | 0.273 | 1.30E-06 |
| cg18940157 | chr14:105807576 | <i>PACS2</i>    | Open sea    | 2.1   | 0.363 | 9.50E-09 |
| cg11909043 | chr10:24607515  | <i>KIAA1217</i> | Open sea    | 2.8   | 0.378 | 3.80E-13 |
| cg19442101 | chr5:173001933  | -               | Open sea    | 0.95  | 0.241 | 7.70E-05 |
| cg27039257 | chr3:185533592  | <i>IGF2BP2</i>  | Open sea    | -1.2  | 0.205 | 6.70E-09 |
| cg20519597 | chr11:33156606  | <i>CSTF3</i>    | Open sea    | 1.2   | 0.295 | 7.00E-05 |
| cg15170994 | chr3:186025138  | <i>DGKG</i>     | Open sea    | 0.8   | 0.144 | 2.80E-08 |
| cg01153053 | chr4:54669219   | -               | Open sea    | 2.1   | 0.344 | 4.90E-10 |
| cg01763410 | chr18:31020209  | <i>CCDC178</i>  | North shore | 1.8   | 0.315 | 1.90E-08 |
| cg19024056 | chr4:90851085   | <i>MMRN1</i>    | Open sea    | -0.73 | 0.187 | 9.50E-05 |
| cg07100050 | chr10:21390114  | <i>NEBL</i>     | Open sea    | 0.79  | 0.187 | 2.20E-05 |
| cg14710896 | chr10:24671718  | <i>KIAA1217</i> | Open sea    | -0.55 | 0.135 | 4.70E-05 |
| cg06891056 | chr1:38560790   | -               | Open sea    | 1.8   | 0.346 | 2.60E-07 |
| cg11463292 | chr8:70397688   | <i>SULF1</i>    | Open sea    | 2.1   | 0.522 | 8.10E-05 |
| cg09525260 | chr15:84523205  | <i>ADAMTSL3</i> | Open sea    | 19    | 3.5   | 6.50E-08 |
| cg00617180 | chr6:44642051   | -               | Open sea    | 15    | 3.7   | 5.00E-05 |
| cg02510044 | chr19:48896080  | <i>KDELR1</i>   | South shore | 28    | 7.1   | 7.20E-05 |

## Continuation of Table S 3.6

### All EWAS catalog study IDs containing significant hits

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>26500701_sex_kora_f4, 27717397_gestational_age_ultrasoundestimated_gestational_age, 25282492_fetal_vs_adult_liver, 27717397_gestational_age_last_menstrual_periodestimated_gestational_age, 27651444_smoking_current_vs_never_smoking, 25650246_gestational_age, 28056824_age_4_vs_age_0, 25325195_urinary_arsenic_concentration_discovery, 23334450_rheumatoid_arthritis, 23526956_clear_cell_renal_carcinoma, 27843151_alcohol_consumption_per_day_european_ancestry, 25325195_blood_arsenic_concentration_discovery, 28002404_body_mass_index_european_ancestry_discovery, 28002404_body_mass_index_all_ancestry_discovery, 28002404_body_mass_index_discovery_and_replication, 26857698_primary_sjogrens_syndrome_whole_blood, 28225026_tobacco_smoking, 27105112_hiv_infection_discovery, 25404168_age_t_cells, 25404168_age_monocytes, 25855720_maternal_underweight_adolescence, 29086770_educational_achievement_basic_model, 26500701_sex_metaanalysis, 25888029_ageing, 26500701_sex_kora_f3, 25249537_sex_discovery, 25249537_sex_replication, 25129077_alzheimers_disease_braak_stage_discovery, 28811542_age, 25855720_maternal_underweight_birth, 23691101_smoking_current_vs_never_smoking_discovery, 27843151_alcohol_consumption_per_day_african_ancestry, 28603561_cleft_palate_vs_cleft_lip, 24500968_pancreatic_ductal_adenocarcinoma, 28002404_body_mass_index_south_asian_ancestry_discovery, 28002404_body_mass_index_all_ancestry_replication, 26864933_smoking_current_vs_never_smoking, 27040690_maternal_smoking_in_pregnancy_sustained_maternal_smoking_in_pregnancy_effect_on_newborns_adjusted_for_cell_counts, 28002404_body_mass_index_european_ancestry_replication, 27826092_body_mass_index, 25935004_waist_circumference_discovery, 28686328_smoking, 26864933_smoking_former_vs_never_smoking, 27955697_creative_protein_discovery, 27955697_creative_protein_replication, 27651444_smoking_former_vs_never_smoking, 25556184_smoking_current_vs_never_smoking, 26110892_body_mass_index_discovery_goldn, 26110892_body_mass_index_metaanalysis, 26110892_waist_circumference_discovery_goldn, 26110892_waist_circumference_aric, 26110892_waist_circumference_metaanalysis, 27843151_alcohol_consumption_per_day_cd14_monocytes, 26756918_smoking, 27040690_maternal_smoking_in_pregnancy_any_maternal_smoking_in_pregnancy_effect_on_newborns, 28095459_body_mass_index_framingham_heart_study_batch_2, 28095459_body_mass_index_lothian_birth_cohort_1936, 28095459_body_mass_index_metaanalysis, 28213390_serum_highdensity_lipoprotein_cholesterol_discovery, 28213390_serum_triglycerides_discovery, 28213390_serum_highdensity_lipoprotein_cholesterol_discovery_and_replication, 28213390_serum_triglycerides_discovery_and_replication, 27824951_creative_protein_kora_f4_discovery, 27886173_inflammatory_bowel_disease, 26119815_body_mass_index_replication, 25855716_prenatal_exposure_to_aflatoxin_b1, 26667048_lung_cancer, 23691101_smoking_current_vs_never_smoking_replication, 27040690_maternal_smoking_in_pregnancy_sustained_maternal_smoking_in_pregnancy_effect_on_newborns, 25650246_sex, 28194238_highdensity_lipoprotein_cholesterol_discovery_and_replication, 27572077_schizophrenia_phase_1, 27572077_schizophrenia_phase_2, 27015559_oldage_mortality_255_years, 27564456_chronic_obstructive_pulmonary_disease, 26661806_growth_differentiation_factor_15_nsphs_study, 26661806_growth_differentiation_factor_15_pivus_study, 26661806_growth_differentiation_factor_15_metaanalysis, 25935004_body_mass_index_discovery, 28603561_cleft_palate_vs_cleft_lip_palate, Rijlaarsdam-J.trajectories_of_social_communication_deficits_discovery_in_cord_blood, Battram-T.cholesterol_esters_in_idl, Battram-T.cholesterol_esters_in_large_idl, Battram-T.cholesterol_esters_in_medium_idl, Battram-T.cholesterol_esters_in_small_idl, Battram-T.cholesterol_esters_in_very_small_vldl, Battram-T.cholesterol_esters_to_total_lipids_ratio_in_medium_idl, Battram-T.cholesterol_fasting, Battram-T.concentration_of_idl_particles, Battram-T.concentration_of_very_large_hdl_particles, Battram-T.concentration_of_very_small_vldl_particles, Battram-T.lg_length, Battram-T.diastolic_blood_pressure, Battram-T.albumin, Battram-T.free_cholesterol_in_very_small_vldl, Battram-T.free_cholesterol_to_total_lipids_ratio_in_medium_hdl, Battram-T.hip_axis_length, Battram-T.hip_cortical_width_calcar, Battram-T.android_fat_mass, Battram-T.hip_total_area, Battram-T_idl_cholesterol_fasting, Battram-T.phospholipids_in_idl, Battram-T.phospholipids_in_large_idl, Battram-T.phospholipids_in_very_small_vldl, Battram-T.phospholipids_to_total_lipids_ratio_in_small_hdl, Battram-T_ratio_of_apolipoprotein_b_to_apolipoprotein_a_i, Battram-T_total_cholesterol_in_idl, Battram-T_total_cholesterol_in_large_idl, Battram-T_total_cholesterol_in_idl, Battram-T_total_cholesterol_in_medium_idl, Battram-T_total_cholesterol_in_very_small_vldl, Battram-T_total_cholesterol_in_very_small_vldl, Battram-T_total_lipids_in_idl, Battram-T_total_lipids_in_large_idl, Battram-T_total_lipids_in_medium_idl, Battram-T_total_lipids_in_very_small_vldl, Battram-T_body_mass_index, Battram-T_free_cholesterol_to_total_lipids_ratio_in_small_hdl, Battram-T_glutamine, Battram-T_glycoprotein_acetyls, Battram-T_mean_diameter_for_hdl_particles, 31691802_Albaio-D_type_ii_diabetes_mellitus_linear_regression, 32066674_Mooney-M_attention_deficit_hyperactivity_disorder_adhd_prs-by-sex_interactions, 31426852_Tajuddin-SM_age_in_african_american_individuals, 31426852_Tajuddin-SM_age_in_white_individuals, 29097680_Chu-A_chronic_kidney_disease_meta-analysis_of_estimated_glomerular_filtration_rate_ewas, 29097680_Chu-A_chronic_kidney_disease_meta-analysis_of_prevalent_ckd_ewas, 31536415_Sikdar-S_prenatal_smoke_exposure_inverse_variance_weighted_fixed-effects_meta-analysis, 32484729_Eze-Inkenna_nitrogen_dioxide_air_pollution, 31536415_Sikdar-S_smoking_random_effects_meta-analysis, 30376510_Van-Roekel-E_time_spent_sitting, 31230546_Kazmi-N_hypertensive_disorders_of_pregnancy_meta_analysis_adjusted_model, 31230546_Kazmi-N_hypertensive_disorders_of_pregnancy_meta_analysis_unadjusted_model, 31230546_Kazmi-N_preeclampsia_disorders_of_pregnancy_meta_analysis_unadjusted_model, 31282290_Huan-T_mir-296-5p, 31282290_Huan-T_mir-125b-5p, 31282290_Huan-T_mir-100-5p, 31594917_Ju-C_response_to_antidepressants_responders_vs_non-responders, 31910897_Cronjé-HT_alcohol_consumption, 31910897_Cronjé-HT_age, 32190749_Chen-S_pubertal_timing_males_only_pt_measured_by_voice_change, 32190749_Chen-S_pubertal_timing_males_only_pt_measured_by_skin_change, 32190749_Chen-S_pubertal_timing_females_only_pt_measured_by_age_at_menarche, 32484362_Domingo-Relloso-A_smoking_current_vs_never, 32484362_Domingo-Relloso-A_smoking_former_vs_never, 29740313_Mendelson-MM_soluble_tumor_necrosis_factor_receptor_2, 24014485_Petersen-A-K_m32388_dodecanedioate_no_snps_model, 24014485_Petersen-A-K_m15749_3-phenylpropionate_hydrocinamate_no_snps_model, 24014485_Petersen-A-K_m33957_1-heptadecanoylglycerophosphocholine_no_snps_model, 24014485_Petersen-A-K_m01444_pipecolate_no_snps_model, 24014485_Petersen-A-K_m32761_x-11444_no_snps_model, 24014485_Petersen-A-K_m32388_dodecanedioate_snps_as_covariates_model, 24014485_Petersen-A-K_m33957_1-heptadecanoylglycerophosphocholine_snps_as_covariates_model, 24014485_Petersen-A-K_m33960_1-oleoylglycerophosphocholine_snps_as_covariates_model, 24014485_Petersen-A-K_m32761_x-11444_snps_as_covariates_model, 31985870_Raffield-LM_coagulation_factor_viii, 32144264_Valleaga-CL_parkinson_s_disease_meta_analysis_using_moa_model_from_osca, 30343628_Carmona-JJ_lung_function_decline_discovery_fvc_lung_function_decline_in_this_paper_is_defined_as_a_decrease_in_fvc_between_visits_31197173_Liu-J_fasting_glucose_base_31197173_Liu-J_fasting_insulin_base_31197173_Liu-J_fasting_insulin_base_plus_bmi_30936141_Ma-J_hepatic_fat_meta-analysis_model_31789449_Dugue-P-A_alcohol_consumption_lifetime_consumption_discovery_31789449_Dugue-P-A_alcohol_consumption_previous_week_consumption_discovery_29198723_Richard-MA_systolic_blood_pressure_meta-analysis_29198723_Richard-MA_diastolic_blood_pressure_meta-analysis_31510868_Sun-D_body_mass_index_discovery_in_africans_31510868_Sun-D_body_mass_index_meta-analysis_in_europeans_31510868_Sun-D_body_mass_index_discovery_follow_up_in_europeans_31789449_Dugue-P-A_alcohol_consumption_previous_decade_consumption_discovery_31789449_Dugue-P-A_alcohol_consumption_replication_31510868_Sun-D_body_mass_index_discovery_in_europeans_31510868_Sun-D_body_mass_index_replication_in_europeans_31015461_Kupers-L_birthweight_meta_analysis_all_ancestries_31015461_Kupers-L_birthweight_meta_analysis_european_ancestry_only_30779925_Somineni-H_crohn_s_disease_30779925_Somineni-H_crohn_s_disease_inflammatory_crohn_s_disease_30264654_Wang-Y_aging_31324826_Shiau-S_perinatally_acquired_hiv_32453742_Natril-H_papuan_ancestry_29311653_Marioni-R_cognitive_abilities_digit_test_basic_adjusted_model_29419728_Svane-AM_mortality_discovery_cox_regression_29016858_Sharp-GC_maternal_body_mass_index_meta-analysis_29016858_Sharp-GC_maternal_overweight_obesity_meta-analysis_29410382_Lichtenfels-AJ_nitrogen_dioxide_exposure_30736859_Lund-JB_mortality_cox_regression_30935889_Birmingham-M_fev1_30935889_Birmingham-M_fvc_29278407_Geurts-YM_bmi_discovery_30966880_Hannon-E_gestational_age_31552803_Dugue-PA_smoking_current_vs_never_smoking_31552803_Dugue-PA_smoking_former_vs_never_smoking_31386212_Xu-K_alcohol_consumption_not_adjusted_for_bmi_31386212_Xu-K_alcohol_consumption_adjusted_for_bmi_32034291_Liang-X_alcohol_consumption_meta_analysis_with_phosphatidylethanol_used_as_measure_for_ac_33450751_Mulder-RH_age_model_1_with_age_as_a_fixed_effect_33450751_Mulder-RH_age_model_1_with_age_as_a_random_effect_33450751_Mulder-RH_age_model_2_at_9y_with_age_as_a_fixed_effect_33450751_Mulder-RH_age_model_2_at_9y_with_age_as_a_random_effect_33450751_Mulder-RH_sex_discovery_33450751_Mulder-RH_age_sex_discovery_33397400_Walker-R_apoe_e2_vs_e4_meta-analysis_33371772_Kandaswamy_sever_adolescent_victimization_replication_epic_33420481_van-Dongen-J_aggressive_behaviour_meta-analysis_in_whole_blood_33564054_Kashima-K_gestational_age_33564054_Kashima-K_birth_weight_33547282_Muskens-I_downs_syndrome_33547282_Muskens-I_downs_syndrome_latino_individuals_only_33547282_Muskens-I_downs_syndrome_individuals_with_gata1_mutation_removed_33547282_Muskens-I_downs_syndrome_basic_model_plus_birthweight_and_gestational_age_adjustment_33593402_Christiansen-C_smoking_current_vs_never_smoking_meta-analysis_in_european_samples_33593402_Christiansen-C_smoking_current_vs_never_smoking_meta-analysis_in_trans-ethnic_samples_33517419_Sharp-G_maternal_body_mass_index_adjusted_for_paternal_bmi_using_cord_blood_in_males_only_33517419_Sharp-G_maternal_body_mass_index_adjusted_for_paternal_bmi_using_cord_blood_in_females_only_33517419_Sharp-G_maternal_body_mass_index_adjusted_for_paternal_bmi_using_cord_blood_in_childhood_in_females_only_33752734_Yao-C_gene_expression_of_affymetrix_id_3944882_gene_symbol_lgals11_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3455674_gene_symbol_krt73_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3455651_gene_symbol_krt72_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_2426676_gene_symbol_henmt11_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3842141_gene_symbol_rpl28_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_2906333_gene_symbol_daam21_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3894637_gene_symbol_nslf1c1c_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3894727_gene_symbol_sirpb1_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3848492_gene_symbol_fcer21c1c_meta-analysis_33752734_Yao-C_gene_expression_of_affymetrix_id_3303652_gene_symbol_mrlp431_meta-analysis_33646943_Hannon-E_psychosis_random_effects_meta-analysis_33646943_Hannon-E_psychosis_fixed_effects_meta-analysis_33646943_Hannon-E_schizophrenia_random_effects_meta-analysis_33646943_Hannon-E_schizophrenia_fixed_effects_meta-analysis_33319642_Lin-Y_iga_nephropathy_33823916_Song-N_cancer_treatment_alkylating_agents_classic_33823916_Song-N_cancer_treatment_anti-metabolites_33823916_Song-N_cancer_treatment_asparaginase_enzymes_33823916_Song-N_cancer_treatment_chest-rt_33823916_Song-N_cancer_treatment_pelvis-rt_33823916_Song-N_cancer_treatment_antimetabolites_and_asparaginase_enzymes_33823916_Song-N_cancer_treatment_asparaginase_enzymes_and_epidodaphyllotoxins_33257653_Zhang-L_alzheimers_disease_braak_stage_meta-analysis_33279932_Li-M_schizophrenia_main_33279932_Li-M_schizophrenia_adjusted_for_cell_composition_33151971_Antoun-E_gestational_diabetes_milletus_33151971_Antoun-E_gestational_diabetes_milletus_adjusted_for_bmi_33151971_Antoun-E_1-hour_glucose_33151971_Antoun-E_gestational_diabetes_milletus_adjusted_for_gestational_age_at_gtt_33151971_Antoun-E_1-hour_glucose_adjusted_for_gestational_age_at_gtt_33151971_Antoun-E_1-hour_glucose_adjusted_for_intervention_33151971_Antoun-E_gestational_diabetes_milletus_adjusted_for_intervention_31506065_Jansen-R_age_males_only_methylspectrum_to_derive_cell_composition_33239103_Vehmeijer-FOL_body_mass_index_childhood_bmi_and_methylation_33239103_Vehmeijer-FOL_body_mass_index_adolescent_bmi_and_methylation_32916427_Lu-X_mehp_33079621_Madden-RA_birthweight_childhood_bmi_and_methylation_33079621_Madden-RA_birthweight_childhood_bmi_and_methylation_33079621_Madden-RA_birthweight_sensitivity_meta-analysis_of_generation_scotland_dna_methylation_set_1_and_set_2_adjusting_for_bmi_33079621_Madden-RA_birthweight_sensitivity_meta-analysis_of_generation_scotland_dna_methylation_set_1_and_set_2_adjusting_for_education_years_33079621_Madden-RA_birthweight_sensitivity_meta-analysis_of_generation_scotland_dna_methylation_set_1_and_set_2_adjusting_for_deprivation_simd_33079621_Madden-RA_birthweight_sensitivity_meta-analysis_of_generation_scotland_dna_methylation_set_1_and_set_2_adjusting_for_bmi_education_years_and_deprivation_simd_33079621_Madden-</p> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

RA\_birthweight\_birthweight\_set\_1\_preterm\_individuals\_removed,33079621\_Madden-RA\_birthweight\_meta-analysis\_preterm\_individuals\_removed,33079621\_Madden-RA\_birthweight\_birthweight\_set\_1,33902726\_Zhang-L\_braak\_stage\_females,34116986\_Canoui-M\_gestational\_diabetes\_mellitus\_ewas\_in\_mothers,34044884\_Inkster-A\_sex\_discovery,34008478\_Zhao-N\_lung\_cancer\_smoking\_score\_unadjusted,34008478\_Zhao-N\_lung\_cancer,32080920\_Witt-S\_alcohol\_withdrawal\_recovery\_discovery,33990564\_Karabegovic-I\_coffee\_and\_tea\_consumption\_meta-analysis,33990564\_Karabegovic-I\_coffee\_and\_tea\_consumption\_discovery,32697766\_Colicino-E\_ail\_cause\_mortality\_fully\_adjusted\_model,32697766\_Colicino-E\_ail\_cause\_mortality\_basic\_model,32795589\_Kyung-WK\_eosinophilia\_discovery,32795589\_Kyung-WK\_atopy\_discovery,33890633\_Penova-Veselinovic\_in\_vitro\_fertilisation\_vs\_intracytoplasmic\_sperm\_injection\_not\_adjusted\_for\_cell\_counts,33883000\_Fernandez-Sanles-A\_acute\_myocardial\_infarction\_meta-analysis\_model3,33883000\_Fernandez-Sanles-A\_acute\_myocardial\_infarction\_meta-analysis\_model2,33883000\_Fernandez-Sanles-A\_acute\_myocardial\_infarction\_meta-analysis\_model1,33931109\_Breeze-C\_estimated\_glomerular\_filtration\_rate\_egfr\_summary\_results\_from\_lmms\_were\_meta-analyzed\_within\_and\_across\_ethnicities,34427971\_Tan-Q\_age\_discovery,34569420\_Bozack\_infant\_sex\_umbilical\_artery,34569420\_Bozack\_infant\_sex\_placenta,34569420\_Bozack\_infant\_sex\_cord\_blood,34516295\_Hoang\_dicamba\_discovery,34584077\_Jenny-van-Dongen\_monozygotic\_twinning,34583751\_Rebecca-Richmond\_smoking\_smoking\_vs\_non-smoking,34523531\_Paulina-Jedynak\_pregnancy\_exposure\_to\_triclosan\_adjusted\_for\_cell\_composition,34523531\_Paulina-Jedynak\_pregnancy\_exposure\_to\_triclosan\_unadjusted\_for\_cell\_counts,34670603\_Chen-Y\_waist\_circumference\_multi-ethnic\_asian\_meta-analysis,34670603\_Chen-Y\_body\_mass\_index\_multi-ethnic\_asian\_meta-analysis,34689037\_Jeong-A\_ndvi\_of\_500\_m\_circular\_buffer\_surrounding\_residential\_address,34689037\_Jeong-A\_ndvi\_of\_30x30m\_cell\_surrounding\_residential\_address,30833390\_Knijnenburg\_idiopathic\_preterm\_birth\_discovery,30833390\_Knijnenburg\_very\_early\_preterm\_birth\_\_28\_weeks\_discovery,30833390\_Knijnenburg\_early\_preterm\_birth\_\_28\_weeks\_discovery,30796027\_Jiantao-Ma\_hepatic\_fat\_replication,30796027\_Jiantao-Ma\_hepatic\_fat\_discovery,31300640\_Cardenas-A\_bronchodilator\_response\_bdr,31300640\_Cardenas-A\_allergic\_asthma,31300640\_Cardenas-A\_fractional\_exhaled\_nitric\_oxide\_feno,31300640\_Cardenas-A\_current\_asthma,Fernandez-Jimenez-N\_maternal\_pre-pregnancy\_body\_mass\_index\_inverse\_variance-weighted\_fixed\_effects\_meta-analysis,Robert-F-Hillary\_Incident\_Stroke\_Basic\_model,Robert-F-Hillary\_Incident\_Osteoarthritis\_Basic\_model,Robert-F-Hillary\_Incident\_Lung\_Cancer\_Basic\_model,Robert-F-Hillary\_Incident\_Liver\_Cirrhosis\_Fully-adjusted\_model,Robert-F-Hillary\_Incident\_Liver\_Cirrhosis\_Basic\_model,Robert-F-Hillary\_Incident\_Ischemic\_Heart\_Disease\_Fully-adjusted\_model,Robert-F-Hillary\_Incident\_Ischemic\_Heart\_Disease\_Basic\_model,Robert-F-Hillary\_Incident\_Type\_2\_Diabetes\_Fully-adjusted\_model,Robert-F-Hillary\_Incident\_Type\_2\_Diabetes\_Basic\_model,Robert-F-Hillary\_Incident\_COPD\_Fully-adjusted\_model,Robert-F-Hillary\_Incident\_COPD\_Basic\_model,Robert-F-Hillary\_Incident\_Chronic\_Kidney\_Disease\_Basic\_model,Robert-F-Hillary\_Incident\_Prostate\_Cancer\_Fully-adjusted\_model,Robert-F-Hillary\_Incident\_Prostate\_Cancer\_Basic\_model,Robert-F-Hillary\_Incident\_Breast\_Cancer\_Basic\_model,Robert-F-Hillary\_Prevalent\_Prostate\_Cancer\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_Prostate\_Cancer\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Stroke\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Rheumatoid\_Arthritis\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Ischemic\_Heart\_Disease\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_Ischemic\_Heart\_Disease\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Type\_2\_Diabetes\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_Type\_2\_Diabetes\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Colorectal\_Cancer\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_Colorectal\_Cancer\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_COPD\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_COPD\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Chronic\_Pain\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_Chronic\_Pain\_(Self-report)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Chronic\_Kidney\_Disease\_(Estimated)\_Fully-adjusted\_Model,Robert-F-Hillary\_Prevalent\_Chronic\_Kidney\_Disease\_(Estimated)\_Basic\_model,Robert-F-Hillary\_Prevalent\_Breast\_Cancer\_(Self-report)\_Fully-adjusted\_model,Robert-F-Hillary\_Prevalent\_Breast\_Cancer\_(Self-report)\_Basic\_model,Robert-F-Hillary\_c-reactive\_protein\_crp\_levels\_fully-adjusted\_model,Robert-F-Hillary\_c-reactive\_protein\_crp\_levels\_basic\_model,Reho-P\_Lewy\_body\_dementia,38052982\_Priyanka-Choudhary\_maternal\_education\_at\_the\_time\_of\_pregnancy\_model1\_adolescence,38052982\_Priyanka-Choudhary\_maternal\_education\_at\_the\_time\_of\_pregnancy\_model1\_cord\_european,38052982\_Priyanka-Choudhary\_maternal\_education\_at\_the\_time\_of\_pregnancy\_model1\_cord\_all,38184627\_Anna-Ulrich\_idiopathic\_heritable\_pulmonary\_arterial\_hypertension\_discovery

Table S 3.7 Gene set and pathway enrichment analysis

| Description                                         | Number of<br>genes in gene<br>set | Number of<br>differentially<br>methylated genes | Enrichment<br>p-value | FDR-adjusted<br>p-value |
|-----------------------------------------------------|-----------------------------------|-------------------------------------------------|-----------------------|-------------------------|
| <b><i>Hallmark Gene Sets</i></b>                    |                                   |                                                 |                       |                         |
| Downregulation of genes in response to UV radiation | 142                               | 14                                              | 0.04890885            | 0.972088899             |
| <b><i>KEGG Pathways</i></b>                         |                                   |                                                 |                       |                         |
| Pancreatic secretion                                | 94                                | 12                                              | 0.00051781            | 0.183650909             |
| Aldosterone synthesis and secretion                 | 94                                | 13                                              | 0.001537105           | 0.183650909             |
| Calcium signaling pathway                           | 240                               | 23                                              | 0.001761525           | 0.183650909             |
| Thyroid hormone synthesis                           | 73                                | 10                                              | 0.002069306           | 0.183650909             |
| Salivary secretion                                  | 85                                | 10                                              | 0.003312069           | 0.235156902             |
| Hematopoietic cell lineage                          | 89                                | 8                                               | 0.005576708           | 0.262612719             |
| GnRH secretion                                      | 62                                | 9                                               | 0.005740589           | 0.262612719             |
| Vascular smooth muscle contraction                  | 132                               | 13                                              | 0.005918033           | 0.262612719             |
| Carbohydrate digestion and absorption               | 41                                | 6                                               | 0.00841711            | 0.332008244             |
| Glutamatergic synapse                               | 114                               | 13                                              | 0.010479369           | 0.358441855             |
| Sphingolipid signaling pathway                      | 120                               | 12                                              | 0.011106649           | 0.358441855             |
| Amphetamine addiction                               | 66                                | 8                                               | 0.014129794           | 0.408090947             |
| Gastric acid secretion                              | 75                                | 9                                               | 0.014944176           | 0.408090947             |
| Cholinergic synapse                                 | 111                               | 12                                              | 0.017959461           | 0.455400621             |
| Starch and sucrose metabolism                       | 30                                | 4                                               | 0.021654347           | 0.496582206             |
| cGMP-PKG signaling pathway                          | 161                               | 14                                              | 0.02238117            | 0.496582206             |
| Parathyroid hormone synthesis, secretion and action | 105                               | 11                                              | 0.026498371           | 0.546615862             |
| Amino sugar and nucleotide sugar metabolism         | 48                                | 5                                               | 0.027857342           | 0.546615862             |
| Aldosterone-regulated sodium reabsorption           | 36                                | 5                                               | 0.029562406           | 0.546615862             |
| Dopaminergic synapse                                | 128                               | 12                                              | 0.031370606           | 0.546615862             |
| Cholesterol metabolism                              | 51                                | 5                                               | 0.03377076            | 0.546615862             |
| Type II diabetes mellitus                           | 45                                | 6                                               | 0.035422432           | 0.546615862             |
| Circadian entrainment                               | 96                                | 10                                              | 0.038444844           | 0.546615862             |
| Growth hormone synthesis, secretion and action      | 115                               | 11                                              | 0.03915044            | 0.546615862             |
| Cytokine-cytokine receptor interaction              | 268                               | 12                                              | 0.041900284           | 0.546615862             |

Continuation of Table S 3.7

|                                                                  |      |     |             |             |
|------------------------------------------------------------------|------|-----|-------------|-------------|
| Viral protein interaction with cytokine and cytokine receptor    | 91   | 5   | 0.042669591 | 0.546615862 |
| Axon guidance                                                    | 175  | 16  | 0.042785188 | 0.546615862 |
| Glycolysis / Gluconeogenesis                                     | 64   | 5   | 0.046143311 | 0.546615862 |
| Long-term potentiation                                           | 64   | 7   | 0.049246325 | 0.546615862 |
| HIF-1 signaling pathway                                          | 105  | 9   | 0.04975856  | 0.546615862 |
| Sphingolipid metabolism                                          | 53   | 5   | 0.049908991 | 0.546615862 |
| <b>Gene ontology (GO) Terms</b>                                  |      |     |             |             |
| CC - plasma membrane                                             | 5161 | 248 | 2.19E-05    | 0.261058606 |
| CC - intrinsic component of plasma membrane                      | 1650 | 96  | 2.29E-05    | 0.261058606 |
| CC - cell periphery                                              | 5619 | 265 | 5.18E-05    | 0.392933523 |
| CC - integral component of plasma membrane                       | 1573 | 88  | 0.000164428 | 0.936253639 |
| BP - blood coagulation, intrinsic pathway                        | 7    | 3   | 0.00022057  | 1           |
| CC - integral component of postsynaptic density membrane         | 51   | 10  | 0.000297486 | 1           |
| CC - intrinsic component of postsynaptic density membrane        | 54   | 10  | 0.00044747  | 1           |
| BP - forebrain neuron development                                | 22   | 6   | 0.000706432 | 1           |
| CC - integral component of postsynaptic specialization membrane  | 72   | 11  | 0.000731007 | 1           |
| BP - negative regulation of dendritic spine development          | 12   | 5   | 0.000797686 | 1           |
| CC - intrinsic component of postsynaptic specialization membrane | 75   | 11  | 0.001002163 | 1           |
| MF - ion channel activity                                        | 422  | 33  | 0.001004206 | 1           |
| MF - microfibril binding                                         | 4    | 3   | 0.001108704 | 1           |
| BP - neuron projection extension                                 | 162  | 19  | 0.001185338 | 1           |
| MF - calcium ion transmembrane transporter activity              | 131  | 16  | 0.001483042 | 1           |
| MF - heparan sulfate proteoglycan binding                        | 18   | 5   | 0.001592932 | 1           |
| BP - signaling                                                   | 6027 | 269 | 0.00163469  | 1           |
| CC - glycoprotein Ib-IX-V complex                                | 4    | 2   | 0.001986383 | 1           |
| BP - cell communication                                          | 6085 | 270 | 0.002070111 | 1           |
| MF - gated channel activity                                      | 319  | 27  | 0.00212896  | 1           |
| MF - channel activity                                            | 463  | 33  | 0.002327818 | 1           |

Continuation of Table S 3.7

|                                                                                |      |     |             |   |
|--------------------------------------------------------------------------------|------|-----|-------------|---|
| MF - cation channel activity                                                   | 331  | 27  | 0.002359752 | 1 |
| MF - passive transmembrane transporter activity                                | 464  | 33  | 0.002382763 | 1 |
| MF - calcium channel activity                                                  | 116  | 14  | 0.002483591 | 1 |
| BP - semaphorin-plexin signaling pathway                                       | 41   | 8   | 0.00260204  | 1 |
| CC - transporter complex                                                       | 375  | 28  | 0.00269793  | 1 |
| CC - platelet dense tubular network membrane                                   | 9    | 4   | 0.002803527 | 1 |
| CC - anchored component of membrane                                            | 160  | 14  | 0.002875159 | 1 |
| BP - positive regulation of macrophage inflammatory protein 1 alpha production | 2    | 2   | 0.002938308 | 1 |
| CC - anchored component of plasma membrane                                     | 59   | 8   | 0.003063205 | 1 |
| BP - central nervous system neuron development                                 | 81   | 11  | 0.00316685  | 1 |
| BP - sensory perception of pain                                                | 101  | 10  | 0.00316854  | 1 |
| CC - transmembrane transporter complex                                         | 351  | 26  | 0.00366071  | 1 |
| BP - developmental growth involved in morphogenesis                            | 219  | 22  | 0.003831354 | 1 |
| MF - prostanoid receptor activity                                              | 11   | 3   | 0.004250383 | 1 |
| CC - plasma membrane protein complex                                           | 552  | 38  | 0.004264584 | 1 |
| BP - trans-synaptic signaling                                                  | 702  | 48  | 0.004274317 | 1 |
| BP - regulation of extent of cell growth                                       | 101  | 13  | 0.004295989 | 1 |
| CC - platelet dense tubular network                                            | 11   | 4   | 0.004299702 | 1 |
| MF - calcium channel inhibitor activity                                        | 11   | 3   | 0.004344872 | 1 |
| BP - synaptic signaling                                                        | 723  | 49  | 0.004501066 | 1 |
| BP - developmental cell growth                                                 | 216  | 21  | 0.004581422 | 1 |
| MF - outward rectifier potassium channel activity                              | 24   | 5   | 0.004582983 | 1 |
| BP - membrane depolarization during SA node cell action potential              | 5    | 3   | 0.004617238 | 1 |
| CC - postsynaptic density membrane                                             | 86   | 11  | 0.004672225 | 1 |
| CC - ion channel complex                                                       | 281  | 23  | 0.005119029 | 1 |
| CC - integral component of postsynaptic membrane                               | 114  | 13  | 0.005147071 | 1 |
| CC - intrinsic component of membrane                                           | 5067 | 203 | 0.005224695 | 1 |
| BP - myelin assembly                                                           | 22   | 5   | 0.00528181  | 1 |
| BP - chemical synaptic transmission                                            | 696  | 47  | 0.005478761 | 1 |



Continuation of Table S 3.7

|                                                                                    |      |     |             |   |
|------------------------------------------------------------------------------------|------|-----|-------------|---|
| BP - anterograde trans-synaptic signaling                                          | 696  | 47  | 0.005478761 | 1 |
| BP - regulation of biological quality                                              | 3561 | 170 | 0.00549142  | 1 |
| BP - brain-derived neurotrophic factor receptor signaling pathway                  | 6    | 3   | 0.005627461 | 1 |
| BP - lipoprotein transport                                                         | 20   | 4   | 0.005760545 | 1 |
| BP - axon extension involved in axon guidance                                      | 36   | 7   | 0.005789232 | 1 |
| BP - neuron projection extension involved in neuron projection guidance            | 36   | 7   | 0.005789232 | 1 |
| BP - blood coagulation, fibrin clot formation                                      | 15   | 3   | 0.005882902 | 1 |
| BP - axon guidance                                                                 | 229  | 23  | 0.005995913 | 1 |
| MF - calcium-dependent protein serine/threonine kinase activity                    | 21   | 5   | 0.006054347 | 1 |
| BP - calcium ion transport                                                         | 419  | 29  | 0.006160108 | 1 |
| BP - positive regulation of ARF protein signal transduction                        | 2    | 2   | 0.006164199 | 1 |
| BP - neuron projection guidance                                                    | 230  | 23  | 0.006317484 | 1 |
| BP - cell-cell signaling                                                           | 1586 | 89  | 0.006770046 | 1 |
| CC - vesicle                                                                       | 3858 | 169 | 0.006794857 | 1 |
| MF - calcium-dependent protein kinase activity                                     | 22   | 5   | 0.0068087   | 1 |
| MF - icosanoid receptor activity                                                   | 14   | 3   | 0.006825052 | 1 |
| BP - postganglionic parasympathetic fiber development                              | 2    | 2   | 0.006830441 | 1 |
| BP - regulation of isotype switching to IgG isotypes                               | 13   | 3   | 0.007031356 | 1 |
| BP - lipoprotein localization                                                      | 21   | 4   | 0.007064238 | 1 |
| BP - negative regulation of axonogenesis                                           | 60   | 9   | 0.007066449 | 1 |
| MF - signaling receptor binding                                                    | 1390 | 70  | 0.007274789 | 1 |
| MF - hydrolase activity, acting on acid phosphorus-nitrogen bonds                  | 3    | 2   | 0.007297798 | 1 |
| CC - intrinsic component of postsynaptic membrane                                  | 119  | 13  | 0.00744548  | 1 |
| MF - voltage-gated calcium channel activity involved SA node cell action potential | 2    | 2   | 0.007483541 | 1 |
| BP - positive regulation of viral entry into host cell                             | 14   | 3   | 0.007647478 | 1 |
| BP - positive regulation by symbiont of entry into host                            | 14   | 3   | 0.007647478 | 1 |
| CC - postsynaptic specialization membrane                                          | 110  | 12  | 0.007729461 | 1 |

Continuation of Table S 3.7

|                                                                     |      |     |             |   |
|---------------------------------------------------------------------|------|-----|-------------|---|
| MF - histone kinase activity (H3-T6 specific)                       | 2    | 2   | 0.007990656 | 1 |
| BP - isotype switching to IgG isotypes                              | 14   | 3   | 0.008033994 | 1 |
| BP - chemical homeostasis                                           | 1011 | 54  | 0.008034706 | 1 |
| BP - regulation of cell size                                        | 172  | 17  | 0.008057784 | 1 |
| BP - nervous system development                                     | 2372 | 132 | 0.008523105 | 1 |
| BP - axon extension                                                 | 113  | 13  | 0.008628156 | 1 |
| BP - neuron development                                             | 1061 | 70  | 0.008796597 | 1 |
| BP - peptidyl-arginine omega-N-methylation                          | 4    | 2   | 0.008925831 | 1 |
| BP - synaptic vesicle priming                                       | 19   | 4   | 0.00904733  | 1 |
| MF - calcium channel regulator activity                             | 45   | 6   | 0.009069296 | 1 |
| BP - uronic acid metabolic process                                  | 23   | 3   | 0.009322276 | 1 |
| BP - glucuronate metabolic process                                  | 23   | 3   | 0.009322276 | 1 |
| BP - chylomicron remnant clearance                                  | 6    | 2   | 0.009401882 | 1 |
| BP - triglyceride-rich lipoprotein particle clearance               | 6    | 2   | 0.009401882 | 1 |
| BP - regulation of neuron death                                     | 310  | 22  | 0.009530956 | 1 |
| BP - substance P catabolic process                                  | 3    | 2   | 0.009710774 | 1 |
| BP - negative regulation of phosphatidylcholine catabolic process   | 1    | 1   | 0.009750776 | 1 |
| BP - negative regulation of axon extension                          | 43   | 7   | 0.009837755 | 1 |
| BP - positive regulation of phosphatidylinositol 3-kinase signaling | 78   | 9   | 0.009919723 | 1 |
| BP - regulation of axon extension involved in axon guidance         | 32   | 6   | 0.01010916  | 1 |
| MF - peptide:proton symporter activity                              | 4    | 2   | 0.010152093 | 1 |
| MF - dipeptide transmembrane transporter activity                   | 4    | 2   | 0.010152093 | 1 |
| BP - positive regulation of activation of Janus kinase activity     | 6    | 2   | 0.010218789 | 1 |
| MF - protein binding involved in heterotypic cell-cell adhesion     | 12   | 3   | 0.010292719 | 1 |
| MF - signaling receptor activity                                    | 1332 | 55  | 0.010318737 | 1 |
| MF - molecular transducer activity                                  | 1332 | 55  | 0.010318737 | 1 |
| BP - regulation of axon extension                                   | 87   | 11  | 0.010341713 | 1 |
| BP - calcium ion transmembrane transport                            | 328  | 24  | 0.010436166 | 1 |

Continuation of Table S 3.7

|                                                            |      |     |             |   |
|------------------------------------------------------------|------|-----|-------------|---|
| MF - N-acetylglucosamine 6-O-sulfotransferase activity     | 6    | 2   | 0.010691859 | 1 |
| BP - cellular metal ion homeostasis                        | 396  | 25  | 0.010978697 | 1 |
| MF - calcium-dependent protein kinase C activity           | 15   | 4   | 0.011142763 | 1 |
| BP - dendrite extension                                    | 34   | 6   | 0.011703011 | 1 |
| CC - TEAD-YAP complex                                      | 3    | 2   | 0.011777572 | 1 |
| MF - protein kinase C activity                             | 16   | 4   | 0.011972498 | 1 |
| MF - clathrin heavy chain binding                          | 10   | 3   | 0.012041557 | 1 |
| MF - testosterone 16-alpha-hydroxylase activity            | 1    | 1   | 0.012068612 | 1 |
| BP - calcium ion transmembrane import into cytosol         | 174  | 15  | 0.012111694 | 1 |
| BP - regulation of T-helper 17 cell lineage commitment     | 7    | 2   | 0.012122895 | 1 |
| MF - anti-Mullerian hormone receptor activity              | 1    | 1   | 0.012179847 | 1 |
| BP - T-helper 17 cell lineage commitment                   | 14   | 3   | 0.012193636 | 1 |
| MF - propionate-CoA ligase activity                        | 4    | 2   | 0.012323071 | 1 |
| BP - signal transduction                                   | 5548 | 239 | 0.012360205 | 1 |
| BP - cell projection morphogenesis                         | 616  | 47  | 0.012362232 | 1 |
| BP - peptidyl-arginine N-methylation                       | 5    | 2   | 0.012366092 | 1 |
| BP - regulation of neuron apoptotic process                | 208  | 16  | 0.012374051 | 1 |
| BP - cellular component morphogenesis                      | 736  | 53  | 0.012399921 | 1 |
| BP - glucuronate catabolic process                         | 5    | 2   | 0.012451293 | 1 |
| BP - glucuronate catabolic process to xylulose 5-phosphate | 5    | 2   | 0.012451293 | 1 |
| BP - xylulose 5-phosphate metabolic process                | 5    | 2   | 0.012451293 | 1 |
| BP - xylulose 5-phosphate biosynthetic process             | 5    | 2   | 0.012451293 | 1 |
| BP - regulation of semaphorin-plexin signaling pathway     | 3    | 2   | 0.012577385 | 1 |
| BP - natural killer cell chemotaxis                        | 8    | 2   | 0.012600778 | 1 |
| BP - central nervous system morphogenesis                  | 3    | 2   | 0.012697062 | 1 |
| MF - transmembrane transporter binding                     | 121  | 12  | 0.012727015 | 1 |
| BP - metal ion homeostasis                                 | 493  | 30  | 0.012907747 | 1 |
| BP - regulation of dendrite extension                      | 25   | 5   | 0.012963629 | 1 |
| BP - high-density lipoprotein particle remodeling          | 16   | 3   | 0.013175999 | 1 |

Continuation of Table S 3.7

|                                                                                |      |    |             |   |
|--------------------------------------------------------------------------------|------|----|-------------|---|
| BP - cranial nerve structural organization                                     | 9    | 3  | 0.013276261 | 1 |
| CC - multimerin complex                                                        | 1    | 1  | 0.013386457 | 1 |
| BP - adenylate cyclase-modulating G protein-coupled receptor signaling pathway | 241  | 17 | 0.013757832 | 1 |
| BP - neuropeptide catabolic process                                            | 4    | 2  | 0.014045625 | 1 |
| BP - protein activation cascade                                                | 18   | 3  | 0.014190634 | 1 |
| BP - N-acetylglucosamine metabolic process                                     | 16   | 3  | 0.014306396 | 1 |
| CC - bounding membrane of organelle                                            | 2013 | 98 | 0.014335022 | 1 |
| CC - secretory granule                                                         | 835  | 40 | 0.014360152 | 1 |
| BP - regulation of GTPase activity                                             | 339  | 27 | 0.014535775 | 1 |
| CC - intrinsic component of synaptic membrane                                  | 160  | 15 | 0.014645394 | 1 |
| BP - multicellular organismal signaling                                        | 164  | 14 | 0.014680286 | 1 |
| BP - cell adhesion                                                             | 1466 | 81 | 0.01486748  | 1 |
| BP - positive regulation of cysteine-type endopeptidase activity               | 139  | 11 | 0.015024643 | 1 |
| BP - positive regulation of hydrolase activity                                 | 554  | 35 | 0.015105311 | 1 |
| BP - oligopeptide import across plasma membrane                                | 5    | 2  | 0.015180656 | 1 |
| BP - dipeptide import across plasma membrane                                   | 5    | 2  | 0.015180656 | 1 |
| BP - macrophage inflammatory protein-1 alpha production                        | 6    | 2  | 0.01519338  | 1 |
| BP - regulation of macrophage inflammatory protein 1 alpha production          | 6    | 2  | 0.01519338  | 1 |
| BP - transmission of nerve impulse                                             | 75   | 8  | 0.015400406 | 1 |
| BP - positive regulation of chemokine (C-C motif) ligand 5 production          | 5    | 2  | 0.015545901 | 1 |
| BP - inorganic ion transmembrane transport                                     | 870  | 49 | 0.015658738 | 1 |
| BP - positive regulation of striated muscle contraction                        | 14   | 3  | 0.015795828 | 1 |
| BP - neuron projection morphogenesis                                           | 595  | 45 | 0.015828092 | 1 |
| MF - chemokine receptor antagonist activity                                    | 1    | 1  | 0.015940124 | 1 |
| BP - negative regulation of axon extension involved in axon guidance           | 26   | 5  | 0.016237724 | 1 |
| MF - serine-type endopeptidase activity                                        | 164  | 10 | 0.016278376 | 1 |
| BP - positive regulation of platelet activation                                | 9    | 2  | 0.016478513 | 1 |

Continuation of Table S 3.7

|                                                                                            |      |    |             |   |
|--------------------------------------------------------------------------------------------|------|----|-------------|---|
| BP - cellular chemical homeostasis                                                         | 637  | 36 | 0.01654462  | 1 |
| CC - integral component of synaptic membrane                                               | 149  | 14 | 0.016577969 | 1 |
| BP - plasma membrane bounded cell projection morphogenesis                                 | 611  | 46 | 0.016641851 | 1 |
| BP - histamine catabolic process                                                           | 1    | 1  | 0.016706654 | 1 |
| MF - histamine N-methyltransferase activity                                                | 1    | 1  | 0.016706654 | 1 |
| BP - imidazole-containing compound catabolic process                                       | 1    | 1  | 0.016706654 | 1 |
| BP - cellular calcium ion homeostasis                                                      | 270  | 19 | 0.017020394 | 1 |
| BP - cell part morphogenesis                                                               | 634  | 47 | 0.017078091 | 1 |
| MF - phospholipase D activity                                                              | 5    | 2  | 0.017110358 | 1 |
| BP - trigeminal nerve morphogenesis                                                        | 3    | 2  | 0.017136378 | 1 |
| BP - trigeminal nerve structural organization                                              | 3    | 2  | 0.017136378 | 1 |
| MF - transmembrane signaling receptor activity                                             | 1124 | 44 | 0.017436826 | 1 |
| BP - cellular homeostasis                                                                  | 731  | 40 | 0.01758808  | 1 |
| BP - negative regulation of platelet-derived growth factor receptor-beta signaling pathway | 4    | 2  | 0.017668062 | 1 |
| BP - negative chemotaxis                                                                   | 46   | 7  | 0.017768393 | 1 |
| BP - forebrain neuron differentiation                                                      | 41   | 6  | 0.018251964 | 1 |
| BP - secretion                                                                             | 903  | 49 | 0.018283723 | 1 |
| BP - polyamine homeostasis                                                                 | 1    | 1  | 0.018389507 | 1 |
| MF - inorganic molecular entity transmembrane transporter activity                         | 698  | 41 | 0.018507938 | 1 |
| BP - regulation of cellular respiration                                                    | 46   | 5  | 0.018564159 | 1 |
| MF - serine-type peptidase activity                                                        | 182  | 11 | 0.018648214 | 1 |
| MF - mRNA 3'-UTR AU-rich region binding                                                    | 26   | 4  | 0.018718841 | 1 |
| BP - regulation of activation of Janus kinase activity                                     | 8    | 2  | 0.018767442 | 1 |
| BP - negative regulation of cell projection organization                                   | 182  | 16 | 0.018877797 | 1 |
| MF - betaine-homocysteine S-methyltransferase activity                                     | 1    | 1  | 0.018913513 | 1 |
| BP - regulation of homocysteine metabolic process                                          | 1    | 1  | 0.018913513 | 1 |
| BP - 'de novo' L-methionine biosynthetic process                                           | 1    | 1  | 0.018913513 | 1 |

Continuation of Table S 3.7

|                                                                              |      |     |             |   |
|------------------------------------------------------------------------------|------|-----|-------------|---|
| BP - glycosylceramide metabolic process                                      | 17   | 3   | 0.018925304 | 1 |
| CC - integral component of membrane                                          | 4909 | 191 | 0.018975059 | 1 |
| CC - photoreceptor outer segment membrane                                    | 16   | 3   | 0.018976475 | 1 |
| BP - sympathetic neuron axon guidance                                        | 4    | 2   | 0.019080792 | 1 |
| BP - sleep                                                                   | 29   | 4   | 0.019117948 | 1 |
| MF - brain-derived neurotrophic factor binding                               | 4    | 2   | 0.019223776 | 1 |
| BP - oocyte construction                                                     | 9    | 2   | 0.019345438 | 1 |
| BP - oocyte axis specification                                               | 9    | 2   | 0.019345438 | 1 |
| BP - intracellular signal transduction                                       | 2525 | 125 | 0.019544334 | 1 |
| BP - peptidyl-arginine methylation, to symmetrical-dimethyl arginine         | 1    | 1   | 0.020095879 | 1 |
| BP - negative regulation of neuron apoptotic process                         | 149  | 12  | 0.020224342 | 1 |
| MF - calcium-release channel activity                                        | 17   | 4   | 0.020245656 | 1 |
| MF - inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity | 3    | 2   | 0.020352506 | 1 |
| BP - peripheral nervous system development                                   | 80   | 9   | 0.020480287 | 1 |
| BP - small GTPase mediated signal transduction                               | 473  | 34  | 0.020502846 | 1 |
| BP - translational frameshifting                                             | 1    | 1   | 0.020503324 | 1 |
| BP - viral translational frameshifting                                       | 1    | 1   | 0.020503324 | 1 |
| BP - regulation of translational frameshifting                               | 1    | 1   | 0.020503324 | 1 |
| BP - negative regulation of translational frameshifting                      | 1    | 1   | 0.020503324 | 1 |
| MF - propionate CoA-transferase activity                                     | 1    | 1   | 0.020523621 | 1 |
| MF - methylbutanoate-CoA ligase activity                                     | 1    | 1   | 0.020523621 | 1 |
| MF - voltage-gated ion channel activity                                      | 189  | 16  | 0.020531183 | 1 |
| MF - voltage-gated channel activity                                          | 189  | 16  | 0.020531183 | 1 |
| BP - neuron death                                                            | 349  | 23  | 0.020565662 | 1 |
| MF - L-fucose mutarotase activity                                            | 1    | 1   | 0.020861865 | 1 |
| BP - divalent inorganic cation homeostasis                                   | 342  | 22  | 0.020883329 | 1 |
| BP - calcium ion homeostasis                                                 | 295  | 20  | 0.02094137  | 1 |
| MF - inorganic cation transmembrane transporter activity                     | 572  | 35  | 0.020974341 | 1 |
| MF - NuRD complex binding                                                    | 4    | 2   | 0.021115521 | 1 |

# Continuation of Table S 3.7

|                                                                         |      |    |             |   |
|-------------------------------------------------------------------------|------|----|-------------|---|
| MF - ligand-gated calcium channel activity                              | 26   | 5  | 0.021389228 | 1 |
| BP - positive regulation of immune response                             | 499  | 26 | 0.021568388 | 1 |
| CC - intrinsic component of endosome membrane                           | 4    | 2  | 0.021569297 | 1 |
| BP - positive regulation of catalytic activity                          | 1113 | 61 | 0.021577274 | 1 |
| BP - glutamate receptor signaling pathway                               | 47   | 7  | 0.021710161 | 1 |
| BP - positive regulation of endopeptidase activity                      | 169  | 12 | 0.021829683 | 1 |
| CC - cellularization cleavage furrow                                    | 1    | 1  | 0.021842046 | 1 |
| BP - cellular divalent inorganic cation homeostasis                     | 309  | 20 | 0.021860222 | 1 |
| BP - negative regulation of neurogenesis                                | 140  | 13 | 0.021886011 | 1 |
| BP - negative regulation of response to interferon-gamma                | 7    | 2  | 0.021958208 | 1 |
| BP - negative regulation of interferon-gamma-mediated signaling pathway | 7    | 2  | 0.021958208 | 1 |
| BP - inorganic cation transmembrane transport                           | 782  | 44 | 0.022094302 | 1 |
| CC - synapse                                                            | 1268 | 76 | 0.022256316 | 1 |
| BP - secretion by cell                                                  | 774  | 43 | 0.022698877 | 1 |
| BP - positive regulation of molecular function                          | 1487 | 78 | 0.022700829 | 1 |
| BP - cellular ion homeostasis                                           | 485  | 28 | 0.022744687 | 1 |
| MF - phosphomevalonate kinase activity                                  | 1    | 1  | 0.022894352 | 1 |
| BP - ubiquitin-dependent SMAD protein catabolic process                 | 6    | 2  | 0.023124386 | 1 |
| MF - glutamate-cysteine ligase catalytic subunit binding                | 1    | 1  | 0.023186445 | 1 |
| BP - regulation of glutamate-cysteine ligase activity                   | 1    | 1  | 0.023186445 | 1 |
| BP - positive regulation of glutamate-cysteine ligase activity          | 1    | 1  | 0.023186445 | 1 |
| BP - dipeptide transmembrane transport                                  | 7    | 2  | 0.023272367 | 1 |
| BP - dipeptide transport                                                | 7    | 2  | 0.023272367 | 1 |
| BP - peptidyl-arginine methylation                                      | 18   | 3  | 0.023369926 | 1 |
| MF - serine hydrolase activity                                          | 186  | 11 | 0.023533489 | 1 |
| BP - cellular cation homeostasis                                        | 468  | 27 | 0.023617223 | 1 |
| BP - peptidyl-histidine dephosphorylation                               | 1    | 1  | 0.023944205 | 1 |

Continuation of Table S 3.7

|                                                                                    |     |    |             |   |
|------------------------------------------------------------------------------------|-----|----|-------------|---|
| BP - negative regulation of ATP citrate synthase activity                          | 1   | 1  | 0.023944205 | 1 |
| CC - secretory granule membrane                                                    | 299 | 18 | 0.024116161 | 1 |
| BP - negative regulation of neuron projection development                          | 135 | 13 | 0.024119221 | 1 |
| CC - C-fiber                                                                       | 1   | 1  | 0.024186129 | 1 |
| BP - fibrinolysis                                                                  | 27  | 3  | 0.024214961 | 1 |
| BP - positive regulation of peptidase activity                                     | 189 | 13 | 0.02459058  | 1 |
| BP - cation homeostasis                                                            | 566 | 32 | 0.024795135 | 1 |
| BP - phosphatidylinositol 3-kinase signaling                                       | 140 | 12 | 0.024889882 | 1 |
| MF - cytokine receptor activity                                                    | 88  | 7  | 0.0251414   | 1 |
| CC - anchored component of external side of plasma membrane                        | 19  | 3  | 0.02542406  | 1 |
| MF - immune receptor activity                                                      | 126 | 8  | 0.025536719 | 1 |
| BP - SA node cell action potential                                                 | 12  | 3  | 0.025674595 | 1 |
| BP - SA node cell to atrial cardiac muscle cell signaling                          | 12  | 3  | 0.025674595 | 1 |
| MF - cholesterol 26-hydroxylase activity                                           | 1   | 1  | 0.025749104 | 1 |
| MF - 3-alpha,7-alpha,12-alpha-trihydroxycholestan-26-al 26-oxidoreductase activity | 1   | 1  | 0.025749104 | 1 |
| MF - cholestanetetraol 26-dehydrogenase activity                                   | 1   | 1  | 0.025749104 | 1 |
| MF - cholestanetriol 26-monooxygenase activity                                     | 1   | 1  | 0.025749104 | 1 |
| BP - cell-cell adhesion                                                            | 911 | 51 | 0.025786479 | 1 |
| CC - high-density lipoprotein particle                                             | 26  | 3  | 0.025787794 | 1 |
| BP - neuron apoptotic process                                                      | 241 | 17 | 0.025953805 | 1 |
| BP - neuron migration                                                              | 167 | 15 | 0.026051643 | 1 |
| CC - tertiary granule membrane                                                     | 68  | 6  | 0.026093408 | 1 |
| BP - sympathetic nervous system development                                        | 22  | 4  | 0.026652632 | 1 |
| CC - postsynaptic membrane                                                         | 251 | 20 | 0.026710847 | 1 |
| MF - chemokine activity                                                            | 44  | 3  | 0.026723036 | 1 |
| MF - anandamide 8,9 epoxidase activity                                             | 3   | 1  | 0.026773182 | 1 |
| MF - anandamide 11,12 epoxidase activity                                           | 3   | 1  | 0.026773182 | 1 |
| BP - nucleosome assembly                                                           | 89  | 6  | 0.026888523 | 1 |



Continuation of Table S 3.7

|                                                            |     |    |             |   |
|------------------------------------------------------------|-----|----|-------------|---|
| BP - galactosylceramide metabolic process                  | 8   | 2  | 0.026945258 | 1 |
| MF - cardiolipin hydrolase activity                        | 1   | 1  | 0.027135945 | 1 |
| MF - tropomyosin binding                                   | 16  | 3  | 0.027137161 | 1 |
| MF - glutamate receptor activity                           | 26  | 5  | 0.027170969 | 1 |
| BP - negative regulation of nervous system development     | 145 | 13 | 0.02736182  | 1 |
| BP - regulation of translational termination               | 10  | 2  | 0.02741705  | 1 |
| BP - positive regulation of macrophage chemotaxis          | 19  | 3  | 0.027484896 | 1 |
| BP - regulation of dendritic spine development             | 51  | 7  | 0.027627262 | 1 |
| MF - glucosamine 6-phosphate N-acetyltransferase activity  | 1   | 1  | 0.027701909 | 1 |
| CC - myelin sheath adaxonal region                         | 6   | 2  | 0.027786142 | 1 |
| BP - branchiomotor neuron axon guidance                    | 4   | 2  | 0.027789951 | 1 |
| MF - glycine-gated chloride ion channel activity           | 1   | 1  | 0.027812332 | 1 |
| BP - commissural neuron axon guidance                      | 11  | 3  | 0.028038682 | 1 |
| MF - pantetheine-phosphate adenyltransferase activity      | 1   | 1  | 0.028260133 | 1 |
| BP - regulation of mesonephros development                 | 5   | 2  | 0.028376372 | 1 |
| BP - positive regulation of natural killer cell chemotaxis | 3   | 1  | 0.028579813 | 1 |
| BP - synaptic membrane adhesion                            | 22  | 4  | 0.028648477 | 1 |
| CC - secretory vesicle                                     | 993 | 47 | 0.028660337 | 1 |
| BP - peptidyl-lysine hydroxylation                         | 9   | 2  | 0.028689601 | 1 |
| BP - negative regulation of chemotaxis                     | 66  | 7  | 0.02873213  | 1 |
| BP - signal release                                        | 449 | 28 | 0.028917729 | 1 |
| BP - ensheathment of neurons                               | 141 | 12 | 0.028986616 | 1 |
| BP - axon ensheathment                                     | 141 | 12 | 0.028986616 | 1 |
| BP - glucosamine-containing compound metabolic process     | 23  | 3  | 0.029053131 | 1 |
| MF - scavenger receptor activity                           | 47  | 5  | 0.029094878 | 1 |
| MF - metal ion transmembrane transporter activity          | 418 | 28 | 0.029098615 | 1 |
| BP - hindbrain formation                                   | 1   | 1  | 0.0292257   | 1 |
| BP - cerebellum formation                                  | 1   | 1  | 0.0292257   | 1 |

Continuation of Table S 3.7

|                                                                                      |      |     |             |   |
|--------------------------------------------------------------------------------------|------|-----|-------------|---|
| BP - midbrain-hindbrain boundary maturation                                          | 1    | 1   | 0.0292257   | 1 |
| BP - midbrain-hindbrain boundary maturation during brain development                 | 1    | 1   | 0.0292257   | 1 |
| CC - serine-pyruvate aminotransferase complex                                        | 1    | 1   | 0.029326862 | 1 |
| MF - xylulokinase activity                                                           | 1    | 1   | 0.029465838 | 1 |
| BP - xylulose catabolic process                                                      | 1    | 1   | 0.029465838 | 1 |
| MF - voltage-gated cation channel activity                                           | 146  | 13  | 0.02950565  | 1 |
| MF - glycosaminoglycan binding                                                       | 231  | 16  | 0.029536123 | 1 |
| BP - negative regulation of cell development                                         | 181  | 15  | 0.029660828 | 1 |
| MF - ionotropic glutamate receptor activity                                          | 18   | 4   | 0.029797357 | 1 |
| BP - cell death                                                                      | 2028 | 94  | 0.029923034 | 1 |
| MF - 2',3'-cyclic-nucleotide 3'-phosphodiesterase activity                           | 1    | 1   | 0.02993129  | 1 |
| MF - cyclic nucleotide binding                                                       | 36   | 5   | 0.029973041 | 1 |
| BP - release of sequestered calcium ion into cytosol                                 | 118  | 10  | 0.029999598 | 1 |
| BP - positive regulation of isotype switching to IgG isotypes                        | 10   | 2   | 0.030119413 | 1 |
| BP - positive regulation of vascular associated smooth muscle cell apoptotic process | 9    | 2   | 0.030120046 | 1 |
| BP - chemotaxis                                                                      | 586  | 36  | 0.030140597 | 1 |
| BP - positive regulation of muscle contraction                                       | 48   | 5   | 0.03014653  | 1 |
| MF - actinin binding                                                                 | 15   | 3   | 0.030326008 | 1 |
| MF - interleukin-12 receptor activity                                                | 1    | 1   | 0.030365217 | 1 |
| MF - structural constituent of skin epidermis                                        | 36   | 3   | 0.030578708 | 1 |
| CC - actinin receptor complex                                                        | 7    | 2   | 0.030663689 | 1 |
| BP - taxis                                                                           | 588  | 36  | 0.030687793 | 1 |
| BP - negative regulation of microtubule polymerization                               | 14   | 3   | 0.030768767 | 1 |
| BP - galactolipid metabolic process                                                  | 9    | 2   | 0.030867995 | 1 |
| BP - axonogenesis                                                                    | 417  | 32  | 0.030884556 | 1 |
| CC - membrane                                                                        | 8637 | 342 | 0.030892133 | 1 |
| MF - proteoglycan binding                                                            | 36   | 5   | 0.031122324 | 1 |
| BP - cellular response to interleukin-12                                             | 6    | 2   | 0.031185919 | 1 |

Continuation of Table S 3.7

|                                                                                     |      |     |             |   |
|-------------------------------------------------------------------------------------|------|-----|-------------|---|
| BP - interneuron migration                                                          | 14   | 3   | 0.031287264 | 1 |
| BP - positive regulation of GTPase activity                                         | 254  | 20  | 0.031300882 | 1 |
| BP - metanephric distal tubule morphogenesis                                        | 1    | 1   | 0.031355811 | 1 |
| CC - extracellular region                                                           | 4136 | 156 | 0.03141071  | 1 |
| BP - dichotomous subdivision of terminal units involved in salivary gland branching | 4    | 2   | 0.031564    | 1 |
| BP - regulation of cAMP-mediated signaling                                          | 29   | 4   | 0.03172992  | 1 |
| BP - triglyceride-rich lipoprotein particle remodeling                              | 12   | 2   | 0.031780825 | 1 |
| BP - very-low-density lipoprotein particle remodeling                               | 12   | 2   | 0.031780825 | 1 |
| MF - benzaldehyde dehydrogenase (NADP+) activity                                    | 1    | 1   | 0.031794565 | 1 |
| CC - synaptic membrane                                                              | 355  | 26  | 0.031803493 | 1 |
| MF - 4-hydroxybenzoate decaprenyltransferase activity                               | 1    | 1   | 0.03191095  | 1 |
| MF - 4-hydroxybenzoate octaprenyltransferase activity                               | 1    | 1   | 0.03191095  | 1 |
| MF - 4-hydroxybenzoate nonaprenyltransferase activity                               | 1    | 1   | 0.03191095  | 1 |
| BP - regulation of peptidyl-tyrosine phosphorylation                                | 253  | 17  | 0.032001542 | 1 |
| BP - exocytic process                                                               | 81   | 8   | 0.032145033 | 1 |
| BP - positive regulation of viral life cycle                                        | 23   | 3   | 0.032173778 | 1 |
| BP - negative regulation of sequestering of calcium ion                             | 119  | 10  | 0.032289659 | 1 |
| MF - pH-gated chloride channel activity                                             | 1    | 1   | 0.032318595 | 1 |
| BP - sphinganine-1-phosphate biosynthetic process                                   | 1    | 1   | 0.032379167 | 1 |
| BP - histone H2B-K12 acetylation                                                    | 1    | 1   | 0.032379167 | 1 |
| MF - AMP binding                                                                    | 14   | 3   | 0.032470306 | 1 |
| BP - regulation of atrial cardiac muscle cell membrane repolarization               | 6    | 2   | 0.032484741 | 1 |
| CC - vesicle membrane                                                               | 1142 | 59  | 0.032504116 | 1 |
| MF - 2-hydroxyacylsphingosine 1-beta-galactosyltransferase activity                 | 1    | 1   | 0.032597837 | 1 |

Continuation of Table S 3.7

|                                                                               |      |    |             |   |
|-------------------------------------------------------------------------------|------|----|-------------|---|
| MF - N-acylsphingosine galactosyltransferase activity                         | 1    | 1  | 0.032597837 | 1 |
| BP - positive regulation of potassium ion transmembrane transport             | 42   | 5  | 0.032631742 | 1 |
| BP - SA node cell to atrial cardiac muscle cell communication                 | 14   | 3  | 0.032637049 | 1 |
| CC - cytoplasmic vesicle membrane                                             | 1125 | 58 | 0.032657715 | 1 |
| BP - response to mitochondrial depolarisation                                 | 19   | 3  | 0.032662034 | 1 |
| BP - programmed cell death                                                    | 1871 | 87 | 0.032917849 | 1 |
| BP - regulation of vascular permeability                                      | 49   | 6  | 0.033025818 | 1 |
| BP - negative regulation of growth rate                                       | 1    | 1  | 0.033142375 | 1 |
| BP - amyloid-beta metabolic process                                           | 64   | 6  | 0.033356913 | 1 |
| BP - regulation of sensory perception of pain                                 | 35   | 4  | 0.03339724  | 1 |
| BP - regulation of sensory perception                                         | 35   | 4  | 0.03339724  | 1 |
| BP - neuron differentiation                                                   | 1322 | 78 | 0.033917392 | 1 |
| BP - regulation of sequestering of calcium ion                                | 121  | 10 | 0.033922575 | 1 |
| MF - CCR4 chemokine receptor binding                                          | 2    | 1  | 0.034058941 | 1 |
| MF - cation transmembrane transporter activity                                | 619  | 36 | 0.034062969 | 1 |
| BP - neuron projection development                                            | 920  | 59 | 0.03419088  | 1 |
| MF - growth factor receptor binding                                           | 136  | 10 | 0.034203989 | 1 |
| MF - inositol 1,4,5 trisphosphate binding                                     | 12   | 3  | 0.034246254 | 1 |
| BP - inorganic ion homeostasis                                                | 581  | 32 | 0.034263597 | 1 |
| BP - maintenance of location                                                  | 326  | 20 | 0.034293519 | 1 |
| MF - racemase and epimerase activity, acting on carbohydrates and derivatives | 10   | 2  | 0.03444002  | 1 |
| MF - thromboxane receptor activity                                            | 1    | 1  | 0.034455983 | 1 |
| MF - thromboxane A2 receptor activity                                         | 1    | 1  | 0.034455983 | 1 |
| BP - thromboxane A2 signaling pathway                                         | 1    | 1  | 0.034455983 | 1 |
| BP - microvillus assembly                                                     | 15   | 3  | 0.034571494 | 1 |
| BP - gonadotrophin-releasing hormone neuronal migration to the hypothalamus   | 5    | 2  | 0.034833274 | 1 |
| BP - hypothalamic tangential migration using cell-axon interactions           | 5    | 2  | 0.034833274 | 1 |
| BP - T-helper cell lineage commitment                                         | 18   | 3  | 0.034913478 | 1 |

Continuation of Table S 3.7

|                                                                      |      |     |             |   |
|----------------------------------------------------------------------|------|-----|-------------|---|
| BP - positive regulation of immune system process                    | 842  | 40  | 0.034959297 | 1 |
| BP - positive regulation of long-term synaptic potentiation          | 23   | 4   | 0.035117817 | 1 |
| CC - calyx of Held                                                   | 18   | 3   | 0.035249451 | 1 |
| BP - positive regulation of dendrite extension                       | 22   | 4   | 0.035834747 | 1 |
| MF - GTPase regulator activity                                       | 445  | 32  | 0.035874589 | 1 |
| MF - nucleoside-triphosphatase regulator activity                    | 445  | 32  | 0.035874589 | 1 |
| BP - response to interleukin-12                                      | 7    | 2   | 0.035922219 | 1 |
| BP - Factor XII activation                                           | 2    | 1   | 0.036538246 | 1 |
| MF - I-SMAD binding                                                  | 14   | 3   | 0.0366232   | 1 |
| BP - regulation of small GTPase mediated signal transduction         | 285  | 23  | 0.036659451 | 1 |
| MF - acyl binding                                                    | 2    | 1   | 0.036729155 | 1 |
| MF - amyloid-beta binding                                            | 78   | 8   | 0.037045314 | 1 |
| MF - prostaglandin receptor activity                                 | 10   | 2   | 0.037064791 | 1 |
| BP - response to amino acid                                          | 108  | 9   | 0.037498953 | 1 |
| BP - positive regulation of blood-brain barrier permeability         | 5    | 2   | 0.037833281 | 1 |
| BP - regulation of interleukin-2-mediated signaling pathway          | 1    | 1   | 0.038020611 | 1 |
| BP - negative regulation of interleukin-2-mediated signaling pathway | 1    | 1   | 0.038020611 | 1 |
| BP - negative regulation of interleukin-4-mediated signaling pathway | 1    | 1   | 0.038020611 | 1 |
| BP - negative regulation of positive thymic T cell selection         | 1    | 1   | 0.038020611 | 1 |
| BP - response to external stimulus                                   | 2595 | 112 | 0.038123023 | 1 |
| BP - granulocyte migration                                           | 146  | 9   | 0.038425611 | 1 |
| BP - detoxification of hydrogen peroxide                             | 1    | 1   | 0.03842783  | 1 |
| BP - cellular detoxification of hydrogen peroxide                    | 1    | 1   | 0.03842783  | 1 |
| CC - catalase complex                                                | 1    | 1   | 0.03842783  | 1 |
| CC - nuclear microtubule                                             | 2    | 1   | 0.038592581 | 1 |
| BP - positive regulation of protein catabolic process in the vacuole | 5    | 2   | 0.038627881 | 1 |
| BP - central nervous system myelin formation                         | 6    | 2   | 0.038686635 | 1 |

Continuation of Table S 3.7

|                                                              |     |    |             |   |
|--------------------------------------------------------------|-----|----|-------------|---|
| BP - potassium ion export across plasma membrane             | 18  | 3  | 0.038908903 | 1 |
| CC - galectin complex                                        | 2   | 1  | 0.038925849 | 1 |
| BP - paranodal junction assembly                             | 6   | 2  | 0.03895768  | 1 |
| CC - neuronal cell body                                      | 474 | 31 | 0.039067182 | 1 |
| CC - cell cortex                                             | 302 | 23 | 0.039180788 | 1 |
| BP - retinal ganglion cell axon guidance                     | 22  | 4  | 0.039203779 | 1 |
| BP - ventricular system development                          | 30  | 4  | 0.039315037 | 1 |
| BP - cerebral cortex tangential migration                    | 5   | 2  | 0.039368193 | 1 |
| BP - regulated exocytosis                                    | 226 | 16 | 0.039718262 | 1 |
| BP - bradykinin catabolic process                            | 7   | 2  | 0.039867604 | 1 |
| MF - G protein-coupled enkephalin receptor activity          | 1   | 1  | 0.03991091  | 1 |
| MF - interleukin-21 receptor activity                        | 1   | 1  | 0.039925508 | 1 |
| BP - interleukin-21-mediated signaling pathway               | 1   | 1  | 0.039925508 | 1 |
| BP - response to interleukin-21                              | 1   | 1  | 0.039925508 | 1 |
| BP - cellular response to interleukin-21                     | 1   | 1  | 0.039925508 | 1 |
| BP - hypothalamus cell migration                             | 6   | 2  | 0.040283847 | 1 |
| BP - developmental growth                                    | 622 | 39 | 0.040380616 | 1 |
| MF - transmitter-gated ion channel activity                  | 59  | 6  | 0.040502216 | 1 |
| MF - transmitter-gated channel activity                      | 59  | 6  | 0.040502216 | 1 |
| BP - chemokine (C-C motif) ligand 5 production               | 10  | 2  | 0.04051183  | 1 |
| BP - regulation of chemokine (C-C motif) ligand 5 production | 10  | 2  | 0.04051183  | 1 |
| BP - sequestering of calcium ion                             | 125 | 10 | 0.04058766  | 1 |
| MF - polyubiquitin modification-dependent protein binding    | 49  | 5  | 0.040637115 | 1 |
| BP - exocytosis                                              | 344 | 22 | 0.040929824 | 1 |
| BP - inflammatory response                                   | 805 | 35 | 0.040977327 | 1 |
| BP - forebrain generation of neurons                         | 50  | 6  | 0.040995941 | 1 |
| MF - dUTP diphosphatase activity                             | 2   | 1  | 0.041181541 | 1 |
| BP - dUTP metabolic process                                  | 2   | 1  | 0.041181541 | 1 |
| BP - dUTP catabolic process                                  | 2   | 1  | 0.041181541 | 1 |

Continuation of Table S 3.7

|                                                                                     |      |    |             |   |
|-------------------------------------------------------------------------------------|------|----|-------------|---|
| BP - Golgi lumen acidification                                                      | 9    | 2  | 0.041391818 | 1 |
| BP - positive regulation of leukocyte tethering or rolling                          | 12   | 2  | 0.041429993 | 1 |
| BP - regulation of microvillus assembly                                             | 6    | 2  | 0.041464169 | 1 |
| BP - Ras protein signal transduction                                                | 329  | 23 | 0.041564265 | 1 |
| CC - specific granule membrane                                                      | 87   | 7  | 0.04183083  | 1 |
| MF - CCR1 chemokine receptor binding                                                | 4    | 1  | 0.041980841 | 1 |
| CC - membrane protein complex                                                       | 1123 | 55 | 0.042114398 | 1 |
| BP - negative regulation of neuron death                                            | 210  | 14 | 0.042225315 | 1 |
| MF - UDP-galactose:beta-N-acetylglucosamine beta-1,3-galactosyltransferase activity | 12   | 2  | 0.042620556 | 1 |
| CC - intrinsic component of external side of plasma membrane                        | 23   | 3  | 0.042854859 | 1 |
| BP - mesonephric duct development                                                   | 8    | 2  | 0.042984326 | 1 |
| BP - apoptotic process                                                              | 1814 | 84 | 0.043031833 | 1 |
| CC - chloride channel complex                                                       | 47   | 5  | 0.043110329 | 1 |
| BP - ion homeostasis                                                                | 590  | 32 | 0.043136903 | 1 |
| CC - plasma membrane signaling receptor complex                                     | 176  | 13 | 0.043249732 | 1 |
| BP - regulation of Ras protein signal transduction                                  | 181  | 15 | 0.043376689 | 1 |
| BP - calcium ion import                                                             | 46   | 5  | 0.043549082 | 1 |
| BP - autophagosome assembly                                                         | 106  | 8  | 0.043570552 | 1 |
| BP - keratinocyte migration                                                         | 20   | 3  | 0.043606224 | 1 |
| BP - cell recognition                                                               | 155  | 11 | 0.043830328 | 1 |
| MF - RNA polymerase III cis-regulatory region sequence-specific DNA binding         | 7    | 2  | 0.043846796 | 1 |
| MF - RNA polymerase III type 3 promoter sequence-specific DNA binding               | 7    | 2  | 0.043846796 | 1 |
| MF - cargo receptor activity                                                        | 81   | 7  | 0.044285958 | 1 |
| BP - negative regulation of relaxation of muscle                                    | 7    | 2  | 0.044311996 | 1 |
| MF - small RNA 2'-O-methyltransferase activity                                      | 2    | 1  | 0.044314869 | 1 |
| BP - central nervous system development                                             | 1001 | 58 | 0.044476652 | 1 |
| MF - low-density lipoprotein particle receptor activity                             | 15   | 3  | 0.044701847 | 1 |
| MF - histone threonine kinase activity                                              | 7    | 2  | 0.045065295 | 1 |

Continuation of Table S 3.7

|                                                                               |     |    |             |   |
|-------------------------------------------------------------------------------|-----|----|-------------|---|
| BP - hypothalamus gonadotrophin-releasing hormone neuron differentiation      | 7   | 2  | 0.0451046   | 1 |
| BP - hypothalamus gonadotrophin-releasing hormone neuron development          | 7   | 2  | 0.0451046   | 1 |
| MF - heparin binding                                                          | 168 | 12 | 0.045438884 | 1 |
| BP - export from cell                                                         | 843 | 44 | 0.046485389 | 1 |
| MF - hydrolase activity, acting on glycosyl bonds                             | 121 | 8  | 0.046513562 | 1 |
| BP - positive regulation of glucose catabolic process to lactate via pyruvate | 2   | 1  | 0.046656189 | 1 |
| MF - testosterone 16-beta-hydroxylase activity                                | 2   | 1  | 0.046836718 | 1 |
| BP - peptidyl-arginine modification                                           | 25  | 3  | 0.047016123 | 1 |
| BP - apoptotic cell clearance                                                 | 50  | 5  | 0.047132556 | 1 |
| BP - regulation of axonogenesis                                               | 145 | 14 | 0.047325286 | 1 |
| BP - siRNA processing                                                         | 9   | 2  | 0.047483435 | 1 |
| MF - complement component C5a receptor activity                               | 2   | 1  | 0.047526112 | 1 |
| BP - complement component C5a signaling pathway                               | 2   | 1  | 0.047526112 | 1 |
| MF - ion transmembrane transporter activity                                   | 804 | 43 | 0.047592749 | 1 |
| BP - exocyst assembly                                                         | 1   | 1  | 0.047615844 | 1 |
| BP - regulation of exocyst assembly                                           | 1   | 1  | 0.047615844 | 1 |
| BP - regulation of exocyst localization                                       | 1   | 1  | 0.047615844 | 1 |
| BP - innervation                                                              | 26  | 4  | 0.047679812 | 1 |
| MF - prostaglandin J receptor activity                                        | 2   | 1  | 0.047705346 | 1 |
| MF - prostaglandin D receptor activity                                        | 2   | 1  | 0.047705346 | 1 |
| BP - positive regulation of macrophage migration                              | 25  | 3  | 0.047874133 | 1 |
| BP - amino sugar metabolic process                                            | 39  | 4  | 0.047902076 | 1 |
| BP - generation of precursor metabolites and energy                           | 464 | 22 | 0.047955932 | 1 |
| CC - pseudopodium                                                             | 16  | 3  | 0.047989883 | 1 |
| BP - xylulose metabolic process                                               | 2   | 1  | 0.048233129 | 1 |
| BP - regulation of rhodopsin mediated signaling pathway                       | 9   | 2  | 0.048488185 | 1 |
| BP - relaxation of vascular associated smooth muscle                          | 8   | 2  | 0.048555562 | 1 |



Continuation of Table S 3.7

|                                                                                                                    |     |    |             |   |
|--------------------------------------------------------------------------------------------------------------------|-----|----|-------------|---|
| MF - delayed rectifier potassium channel activity                                                                  | 31  | 4  | 0.04857873  | 1 |
| MF - RNA polymerase III transcription regulatory region sequence-specific DNA binding                              | 8   | 2  | 0.048689017 | 1 |
| BP - regulation of focal adhesion disassembly                                                                      | 6   | 2  | 0.048702272 | 1 |
| BP - positive regulation of focal adhesion disassembly                                                             | 6   | 2  | 0.048702272 | 1 |
| BP - protein-lipid complex remodeling                                                                              | 30  | 3  | 0.048799665 | 1 |
| BP - plasma lipoprotein particle remodeling                                                                        | 30  | 3  | 0.048799665 | 1 |
| CC - cation channel complex                                                                                        | 213 | 16 | 0.048805165 | 1 |
| BP - ventral spinal cord interneuron specification                                                                 | 7   | 2  | 0.049001268 | 1 |
| BP - cell fate specification involved in pattern specification                                                     | 7   | 2  | 0.049001268 | 1 |
| BP - rhythmic behavior                                                                                             | 47  | 5  | 0.049011597 | 1 |
| BP - regulation of developmental growth                                                                            | 314 | 22 | 0.049036039 | 1 |
| MF - 1,4-alpha-glucan branching enzyme activity                                                                    | 1   | 1  | 0.049250533 | 1 |
| MF - 1,4-alpha-glucan branching enzyme activity (using a glucosylated glycogenin as primer for glycogen synthesis) | 1   | 1  | 0.049250533 | 1 |
| BP - neurotransmitter secretion                                                                                    | 144 | 12 | 0.049308234 | 1 |
| BP - signal release from synapse                                                                                   | 144 | 12 | 0.049308234 | 1 |
| BP - negative regulation of T-helper 2 cell cytokine production                                                    | 4   | 1  | 0.049412845 | 1 |
| MF - open rectifier potassium channel activity                                                                     | 1   | 1  | 0.049451017 | 1 |
| BP - CD4-positive, alpha-beta T cell lineage commitment                                                            | 20  | 3  | 0.049496096 | 1 |
| CC - mesaxon                                                                                                       | 1   | 1  | 0.049503857 | 1 |
| CC - ensheathing process                                                                                           | 1   | 1  | 0.049503857 | 1 |
| BP - creatinine metabolic process                                                                                  | 1   | 1  | 0.049514429 | 1 |
| BP - neuropeptide processing                                                                                       | 1   | 1  | 0.049514429 | 1 |
| BP - semaphorin-plexin signaling pathway involved in axon guidance                                                 | 12  | 3  | 0.049531187 | 1 |
| MF - D4 dopamine receptor binding                                                                                  | 1   | 1  | 0.049673165 | 1 |
| BP - iodide transport                                                                                              | 6   | 2  | 0.049855058 | 1 |
| BP - positive regulation of hippocampal neuron apoptotic process                                                   | 2   | 1  | 0.049921368 | 1 |
| BP - vesicle docking involved in exocytosis                                                                        | 45  | 5  | 0.04995249  | 1 |

Table S 3.8 Enrichment of As-associated CpGs in regions in relation to UCSC CpG islands for As-associated CpGs located in regions with multiple vs. one As-associated CpG

| <i>Enrichment in relation to CpG islands</i>           |             |        |         |          |                                                  |                          |
|--------------------------------------------------------|-------------|--------|---------|----------|--------------------------------------------------|--------------------------|
|                                                        | CpG Islands | Shores | Shelves | Open Sea | Proportion of CpGs located in CpG islands/shores | Chi-squared Test p-value |
| As-associated CpG located in region with multiple hits | 27          | 65     | 10      | 72       | 0.53                                             | 5.26E-18                 |
| As-associated CpG located in region with one hit       | 52          | 154    | 71      | 726      | 0.21                                             |                          |

Table S 3.9 CpGs involved in arsenic and cell type interactions (AxCT)

| Name       | AxCT<br>Cell Type | Chromosome | Position  | AxCT<br>Effect<br>Size<br>Estimate | AxCT p-<br>value | CpG Location   | Nearest<br>Gene | Gene<br>Feature |
|------------|-------------------|------------|-----------|------------------------------------|------------------|----------------|-----------------|-----------------|
| cg01092293 | B cell            | 2          | 160761427 | 0.161                              | 7.64E-07         | Island         | LY75            | Promoter        |
| cg01203426 | B cell            | 2          | 62081861  | -0.076                             | 8.06E-07         | Shore          | FAM161A         | Promoter        |
| cg01427769 | B cell            | 6          | 31830600  | -0.192                             | 4.32E-07         | Island         | NEU1            | Body            |
| cg02379540 | B cell            | 10         | 1010456   | 0.106                              | 1.05E-06         | Non-CpG Island | GTPBP4          | Upstream        |
| cg02535507 | B cell            | 11         | 110166799 | -0.163                             | 1.45E-06         | Island         | RDX             | Body            |
| cg03902817 | B cell            | 12         | 10902749  | -0.070                             | 4.67E-07         | Non-CpG Island | YBX3            | Upstream        |
| cg05716702 | B cell            | 2          | 29438753  | 0.142                              | 7.48E-07         | Non-CpG Island | ALK             | Body            |
| cg06204745 | B cell            | 17         | 79448740  | -0.131                             | 1.11E-06         | Island         | BAHCC1          | Downstream      |
| cg07152259 | B cell            | 15         | 55700689  | -0.132                             | 9.67E-07         | Island         | CCPG1           | Body            |
| cg09259525 | B cell            | 1          | 202199085 | 0.112                              | 3.27E-07         | Non-CpG Island | LGR6            | Body            |
| cg09784579 | B cell            | 8          | 145911240 | -0.097                             | 3.59E-08         | Island         | ARHGAP39        | Promoter        |
| cg10832007 | B cell            | 1          | 179198356 | -0.081                             | 1.18E-09         | Island         | ABL2            | Body            |
| cg11448127 | B cell            | 14         | 104603563 | -0.225                             | 1.36E-07         | Island         | KIF26A          | Promoter        |
| cg12801917 | B cell            | 17         | 15417503  | 0.184                              | 1.36E-06         | Non-CpG Island | TVP23C          | Body            |
| cg15076218 | B cell            | 1          | 46599002  | -0.052                             | 4.05E-07         | Island         | PIK3R3          | Body            |
| cg15470344 | B cell            | 1          | 203056504 | 0.150                              | 5.10E-07         | Non-CpG Island | MYOG            | Promoter        |
| cg18104104 | B cell            | 2          | 10499465  | 0.086                              | 3.04E-07         | Non-CpG Island | HPCAL1          | Body            |
| cg18702827 | B cell            | 8          | 75227933  | 0.140                              | 1.20E-06         | Island         | JPH1            | Body            |
| cg19339179 | B cell            | 4          | 113324678 | -0.138                             | 1.06E-08         | Non-CpG Island | ALPK1           | Body            |
| cg22504696 | B cell            | 10         | 118956514 | 0.135                              | 9.64E-07         | Non-CpG Island | KCNK18          | Promoter        |
| cg25211435 | B cell            | 16         | 84095350  | 0.172                              | 1.05E-06         | Non-CpG Island | MBTPS1          | Body            |
| cg02039370 | NK cell           | 15         | 32158966  | 0.065                              | 3.75E-07         | Non-CpG Island | OTUD7A          | Body            |
| cg03305634 | NK cell           | 17         | 77975090  | 0.076                              | 5.35E-07         | Non-CpG Island | TBC1D16         | Body            |
| cg03612522 | NK cell           | 12         | 108733261 | 0.089                              | 5.10E-07         | Non-CpG Island | CMKLR1          | Promoter        |
| cg03977056 | NK cell           | 2          | 12595966  | 0.162                              | 2.51E-09         | Non-CpG Island | TRIB2           | Upstream        |
| cg09353563 | NK cell           | 15         | 70994435  | 0.059                              | 2.90E-12         | Non-CpG Island | UACA            | Body            |
| cg10769057 | NK cell           | 19         | 34950976  | 0.074                              | 1.80E-08         | Non-CpG Island | UBA2            | Body            |
| cg12091956 | NK cell           | 16         | 88803110  | -0.091                             | 4.86E-07         | Shore          | PIEZO1          | Body            |
| cg17753476 | NK cell           | 10         | 7621949   | -0.070                             | 1.31E-07         | Island         | ITIH5           | Body            |

Continuation of Table S 3.9

|            |           |    |           |        |          |                |           |            |
|------------|-----------|----|-----------|--------|----------|----------------|-----------|------------|
| cg18693395 | NK cell   | 14 | 52735421  | 0.112  | 1.99E-07 | Island         | PTGDR     | Body       |
| cg18724158 | NK cell   | 22 | 49905790  | -0.102 | 2.57E-10 | Non-CpG Island | MIR3667HG | Downstream |
| cg01817885 | CD8T cell | 5  | 59165136  | 0.101  | 3.35E-10 | Non-CpG Island | PDE4D     | Body       |
| cg00590639 | Monocyte  | 19 | 10464771  | 0.069  | 3.31E-08 | Island         | TYK2      | Body       |
| cg03889617 | Monocyte  | 19 | 10463828  | 0.102  | 6.10E-07 | Island         | TYK2      | Body       |
| cg06568565 | Monocyte  | 12 | 109304203 | -0.115 | 6.62E-09 | Non-CpG Island | SVOP      | Downstream |
| cg13971552 | Monocyte  | 19 | 10463352  | 0.124  | 2.05E-08 | Shore          | TYK2      | Body       |
| cg14499928 | Monocyte  | 12 | 131455619 | -0.140 | 9.22E-07 | Non-CpG Island | ADGRD1    | Body       |
| cg16851046 | Monocyte  | 13 | 21291173  | 0.084  | 1.80E-07 | Non-CpG Island | IL17D     | Body       |
| cg18038269 | Monocyte  | 19 | 10464320  | 0.078  | 1.09E-08 | Island         | TYK2      | Body       |
| cg18242219 | Monocyte  | 18 | 23666144  | -0.111 | 5.16E-07 | Shelf          | SS18      | Body       |
| cg19844101 | Monocyte  | 11 | 12028069  | -0.125 | 6.24E-07 | Shore          | DKK3      | Body       |
| cg19925475 | Monocyte  | 19 | 10464767  | 0.060  | 2.23E-07 | Island         | TYK2      | Body       |
| cg19977501 | Monocyte  | 14 | 23233723  | -0.118 | 2.59E-08 | Shelf          | OXA1L     | Promoter   |
| cg22770862 | Monocyte  | 1  | 244753038 | -0.169 | 6.49E-07 | Non-CpG Island | CATSPERE  | Body       |
| cg25259479 | Monocyte  | 7  | 157408632 | -0.171 | 8.69E-07 | Shore          | PTPRN2    | Body       |
| cg27309038 | Monocyte  | 2  | 233244424 | -0.147 | 1.82E-07 | Island         | ALPP      | Body       |

Table S 3.10 List of CpGs and effect sizes for the 255-CpG and 117-CpG biomarkers of urinary arsenic levels developed using the Illumina EPIC and 450k array, respectively.

| 255-CpG biomarker of As for the Illumina EPIC array |                                            | 117-CpG biomarker of As for the Illumina 450k array |                                            |
|-----------------------------------------------------|--------------------------------------------|-----------------------------------------------------|--------------------------------------------|
| CpG Name                                            | Beta Coefficient (using M-values as input) | CpG Name                                            | Beta Coefficient (using M-values as input) |
| cg04893611                                          | 0.06547207                                 | cg14176088                                          | 0.289857416                                |
| cg02641839                                          | 0.000532531                                | cg17897187                                          | 0.060641975                                |
| cg13057875                                          | 0.056594133                                | cg17493193                                          | -0.020978885                               |
| cg03945895                                          | 0.005440264                                | cg06216065                                          | -0.136214605                               |
| cg23513000                                          | 0.001983111                                | cg16651877                                          | 0.071579896                                |
| cg03180011                                          | 0.030800694                                | cg03761216                                          | -0.117419967                               |
| cg21152304                                          | 0.011314758                                | cg04550950                                          | 0.204815241                                |
| cg06216065                                          | -0.082422837                               | cg23598089                                          | 0.045745689                                |
| cg10711145                                          | 0.029255709                                | cg15487498                                          | 0.059102474                                |
| cg10254445                                          | 0.091017334                                | cg11115502                                          | 0.09666852                                 |
| cg03761216                                          | -0.041606364                               | cg04461334                                          | 0.006035556                                |
| cg07480639                                          | 0.006355035                                | cg06224721                                          | -0.018677881                               |
| cg04550950                                          | 0.031929144                                | cg14709239                                          | -0.003004494                               |
| cg15512289                                          | -0.078517238                               | cg10633051                                          | -0.070471919                               |
| cg01791524                                          | -0.004514313                               | cg24758280                                          | -0.086642585                               |
| cg11947712                                          | -0.015633703                               | cg14320320                                          | -0.067544249                               |
| cg09036468                                          | 0.004811784                                | cg13531735                                          | 0.005583504                                |
| cg24430920                                          | 0.051880125                                | cg04385523                                          | 0.012957227                                |
| cg06466147                                          | -0.216917919                               | cg10255761                                          | -0.185984929                               |
| cg13968233                                          | 0.015676844                                | cg23153546                                          | 0.012190647                                |

Continuation of Table S 3.10

|            |              |            |              |
|------------|--------------|------------|--------------|
| cg09421232 | 0.13636866   | cg27402634 | 0.043690788  |
| cg05663480 | -0.145019681 | cg15721584 | 0.013391215  |
| cg17894719 | 0.035320122  | cg09831562 | 0.117422684  |
| cg26224923 | 0.083432613  | cg17470397 | 0.10002376   |
| cg26093964 | 0.153677448  | cg22953687 | -0.104698516 |
| cg11115502 | 0.034170589  | cg06690548 | -0.072724699 |
| cg22717723 | 0.03040898   | cg23360388 | -0.008099784 |
| cg00189117 | 0.048787409  | cg08942527 | -0.044473373 |
| cg14651327 | -0.098943993 | cg18394552 | -0.09756639  |
| cg03289816 | 0.009210659  | cg09222732 | 0.136983588  |
| cg14931378 | -0.4176492   | cg12213680 | 0.073770141  |
| cg04260584 | 0.02414848   | cg20715953 | -0.021095955 |
| cg12124090 | -0.002193287 | cg20405697 | 0.045137528  |
| cg07346426 | -0.131188065 | cg09261794 | 0.108303773  |
| cg21571995 | -0.043597802 | cg14438453 | 0.03197262   |
| cg14108285 | 0.002368777  | cg04334496 | -0.167944161 |
| cg19066876 | 0.091958044  | cg00638845 | 0.031193081  |
| cg06801305 | -0.003093199 | cg26701785 | -0.048559401 |
| cg11907276 | 0.048854009  | cg06864151 | -0.026151194 |
| cg09319083 | 0.030190957  | cg10152684 | 0.04368518   |
| cg20936142 | 0.072534724  | cg11079619 | -0.074277809 |
| cg01095763 | 0.053979453  | cg13459606 | 0.002966981  |
| cg04582369 | -0.050742534 | cg11573518 | 0.061257183  |
| cg00231361 | 0.026123127  | cg12445565 | -0.0347303   |

Continuation of Table S 3.10

|            |              |            |              |
|------------|--------------|------------|--------------|
| cg14320320 | -0.028696771 | cg24491618 | -0.154361704 |
| cg05930631 | 0.022202925  | cg06819296 | 0.186835949  |
| cg08436654 | -0.016376733 | cg14269813 | -0.148361199 |
| cg25019890 | 0.065247046  | cg14405625 | 0.172838607  |
| cg11256374 | 0.077246929  | cg03651219 | 0.069551235  |
| cg10255761 | -0.070281339 | cg22959742 | 0.186633986  |
| cg18308510 | -0.026730379 | cg03717584 | -0.120427167 |
| cg10753966 | 0.036218606  | cg00366037 | 0.045848913  |
| cg10241937 | -0.063888848 | cg15108641 | 0.450710313  |
| cg02639218 | -0.022648219 | cg05962511 | -0.282832476 |
| cg22517730 | 0.03152608   | cg14165660 | -0.096444352 |
| cg24920294 | -0.07331879  | cg24181174 | 0.109582859  |
| cg24525065 | -0.009525115 | cg15289899 | -0.032017879 |
| cg19534475 | 0.372922672  | cg25281677 | -0.24043751  |
| cg01907945 | 0.10321882   | cg08540622 | 0.181067604  |
| cg27402634 | 0.010672728  | cg24009118 | -0.009450682 |
| cg18711583 | 0.005439825  | cg03695871 | -0.08882111  |
| cg16166958 | 0.061133796  | cg24368962 | 0.000425731  |
| cg23683730 | 0.01221375   | cg05469118 | -0.062947862 |
| cg03733470 | -0.019424445 | cg04574480 | 0.105959409  |
| cg22953687 | -0.02167406  | cg16190510 | -0.015026899 |
| cg21510186 | -0.097417131 | cg06996254 | 0.001594865  |
| cg05440526 | -0.121199555 | cg24028828 | 0.050503199  |
| cg01164417 | 0.047637093  | cg18774692 | -0.135248804 |

Continuation of Table S 3.10

|            |              |            |              |
|------------|--------------|------------|--------------|
| cg01153053 | -0.012525553 | cg17413460 | 0.079989673  |
| cg17894854 | 0.009623738  | cg02637282 | -0.093070738 |
| cg04151727 | 0.14473424   | cg26358374 | -0.223470289 |
| cg22256784 | 0.042024734  | cg10101109 | 0.004548796  |
| cg10102128 | 0.000118827  | cg27656459 | -0.113861843 |
| cg25689838 | 0.004824155  | cg21575634 | 0.037592217  |
| cg15848159 | 0.001821332  | cg16411103 | -0.063111217 |
| cg07261946 | 0.043254946  | cg18067103 | 0.059131569  |
| cg21773245 | 0.071224307  | cg27148665 | 0.084006268  |
| cg04491800 | -0.088348646 | cg24125468 | -0.037561998 |
| cg06690548 | -0.103084553 | cg23206630 | 0.120455648  |
| cg24668753 | -0.11966115  | cg01973394 | 0.157125762  |
| cg11538149 | 0.064098633  | cg11183227 | -0.055894396 |
| cg05887366 | 0.031588535  | cg05989795 | -0.172110523 |
| cg26225478 | 0.037909959  | cg02529919 | 0.110012464  |
| cg17842821 | -0.068202528 | cg04045544 | 0.07748422   |
| cg03231199 | 0.014481345  | cg03099790 | 0.132668341  |
| cg27165960 | -0.189697066 | cg05425326 | 0.096519642  |
| cg07872489 | 0.000253081  | cg27162196 | 0.134544422  |
| cg26160660 | 0.012504218  | cg10462803 | 0.030903588  |
| cg13761615 | -0.008310552 | cg01912040 | -0.192747079 |
| cg07033790 | 0.003753794  | cg10003262 | -0.295123697 |
| cg03746860 | 0.01873468   | cg18963239 | -0.026649066 |
| cg27610206 | 0.073488686  | cg06040872 | -0.116619698 |



Continuation of Table S 3.10

|            |              |            |              |
|------------|--------------|------------|--------------|
| cg11441280 | 0.025006095  | cg26299044 | 0.060396106  |
| cg19730422 | -0.132348437 | cg07050712 | 0.00085959   |
| cg13667243 | 0.207961513  | cg06407137 | -0.103309381 |
| cg06938464 | 0.000324842  | cg26211349 | 0.076314561  |
| cg10137337 | -0.066273754 | cg02348119 | -0.035118479 |
| cg16931200 | -0.205746812 | cg05613273 | 0.000511999  |
| cg13583602 | 0.006515583  | cg07352016 | -0.066186055 |
| cg09041851 | -0.018479312 | cg12272104 | -0.087971387 |
| cg18891604 | 0.001836896  | cg13480898 | -0.140499385 |
| cg18704138 | 0.00827114   | cg24143196 | 0.024117439  |
| cg06438159 | -0.050068292 | cg16597993 | 0.042641356  |
| cg15706574 | -0.000396619 | cg09777776 | 0.071678752  |
| cg16756949 | 0.052235918  | cg13252583 | 0.038233429  |
| cg06367607 | 0.044649247  | cg06381803 | -0.048695068 |
| cg05087792 | -0.035651578 | cg24718465 | 0.108353545  |
| cg07617039 | -0.032575182 | cg01865484 | -0.037214634 |
| cg04334496 | -0.031124509 | cg26288449 | -0.068354139 |
| cg02123704 | -0.022205645 | cg00572387 | -0.116406279 |
| cg26701785 | -0.00628479  | cg11086127 | 0.009873887  |
| cg09848638 | 0.155481076  | cg09462575 | 0.133522838  |
| cg11079619 | -0.025347659 | cg10546626 | 0.287385284  |
| cg12856392 | 0.037290472  | cg18563860 | -0.119419098 |
| cg11670749 | -0.05556028  | cg26390598 | 0.111881501  |
| cg25773935 | 0.066916867  | cg13915314 | -0.148473599 |

Continuation of Table S 3.10

|            |              |            |             |
|------------|--------------|------------|-------------|
| cg05788267 | -0.076177638 | cg03318904 | 0.097998478 |
| cg03834786 | -0.185718479 |            |             |
| cg11543304 | 0.034105433  |            |             |
| cg16748130 | -0.006099383 |            |             |
| cg13732043 | 0.029860096  |            |             |
| cg10965265 | 0.118507281  |            |             |
| cg23055986 | 0.176877881  |            |             |
| cg14937504 | 0.006741273  |            |             |
| cg24491618 | -0.040886378 |            |             |
| cg06819296 | 0.078265475  |            |             |
| cg07312099 | 0.011648383  |            |             |
| cg00078857 | 0.093157506  |            |             |
| cg16225800 | 0.04839794   |            |             |
| cg04903019 | 0.017753155  |            |             |
| cg14269813 | -0.078875658 |            |             |
| cg04825341 | 0.021088474  |            |             |
| cg09440862 | 0.070769784  |            |             |
| cg06402858 | 0.092070847  |            |             |
| cg11088536 | -0.010710958 |            |             |
| cg22345623 | -0.060745235 |            |             |
| cg14405625 | 0.145643595  |            |             |
| cg18892873 | 0.002530641  |            |             |
| cg10454881 | 0.005917694  |            |             |
| cg10955407 | 0.040374576  |            |             |

Continuation of Table S 3.10

|            |              |
|------------|--------------|
| cg13409442 | -0.058696746 |
| cg06228645 | 0.043560173  |
| cg20626045 | 0.02006761   |
| cg03717584 | -0.091900886 |
| cg26510238 | 0.042472596  |
| cg01740552 | 0.008732796  |
| cg15108641 | 0.329326434  |
| cg23969297 | 0.005362552  |
| cg05962511 | -0.129408182 |
| cg14165660 | -0.044322015 |
| cg09066266 | -0.071912242 |
| cg05646745 | 0.020361695  |
| cg04249817 | 0.10822615   |
| cg15289899 | -0.020677883 |
| cg21239418 | 0.024749829  |
| cg16493238 | -0.021641085 |
| cg17603132 | 0.043107953  |
| cg22169659 | 0.016704448  |
| cg03700417 | 0.013730491  |
| cg11559886 | -0.03056576  |
| cg25281677 | -0.147803434 |
| cg08540622 | 0.066222965  |
| cg07602847 | -0.044643935 |
| cg17191060 | -0.031512077 |

Continuation of Table S 3.10

|            |              |
|------------|--------------|
| cg08559782 | 0.017750173  |
| cg22898516 | -0.108192494 |
| cg04574480 | 0.003537715  |
| cg12635154 | -0.008309559 |
| cg00997853 | -0.15673624  |
| cg20389635 | 0.056824871  |
| cg12750791 | 0.008233309  |
| cg00213082 | 0.046361129  |
| cg19222775 | 0.135429601  |
| cg22827806 | 0.145263198  |
| cg03348792 | 0.064877312  |
| cg10185759 | -0.127342715 |
| cg18774692 | -0.05558972  |
| cg02637282 | -0.00745782  |
| cg26358374 | -0.19738375  |
| cg27656459 | -0.07548471  |
| cg27519981 | -0.016913518 |
| cg07891540 | -0.001080232 |
| cg08302053 | 0.127940196  |
| cg02489630 | -0.074242616 |
| cg13960680 | 0.126334272  |
| cg21049840 | 0.072099973  |
| cg08790320 | -0.008487021 |
| cg16875568 | -0.035829574 |

Continuation of Table S 3.10

|            |              |
|------------|--------------|
| cg26696488 | -0.042851179 |
| cg01242348 | -0.048848046 |
| cg19743728 | 0.017437936  |
| cg02525267 | -0.093883902 |
| cg25504557 | 0.005214342  |
| cg18268625 | 0.00203935   |
| cg03221779 | 0.079463022  |
| cg19515081 | -0.042602879 |
| cg00884574 | -0.030630608 |
| cg05438461 | 0.007901489  |
| cg17514855 | 0.061279237  |
| cg13077072 | 0.021490252  |
| cg27238311 | -0.039175597 |
| cg26657240 | 0.000988472  |
| cg05989795 | -0.154187497 |
| cg13938970 | 0.025804813  |
| cg14918689 | -0.028209585 |
| cg27246364 | 0.024197416  |
| cg27162196 | 0.092598199  |
| cg05864097 | 0.004495688  |
| cg01912040 | -0.07114724  |
| cg10003262 | -0.148230436 |
| cg26705394 | 0.031162869  |
| cg21110915 | -0.024376129 |

Continuation of Table S 3.10

|            |              |
|------------|--------------|
| cg06040872 | -0.001027095 |
| cg26299044 | 0.01888151   |
| cg22078805 | -0.013843935 |
| cg03452365 | -0.095791005 |
| cg00512837 | -0.029040687 |
| cg22011053 | 0.027924045  |
| cg09662588 | 0.1437508    |
| cg07559942 | 0.053960752  |
| cg18632189 | -0.026947797 |
| cg20671304 | 0.028524085  |
| cg02348119 | -0.025547001 |
| cg01695994 | -0.091282611 |
| cg07142466 | -0.059725277 |
| cg17420142 | -0.233688844 |
| cg16616857 | 0.110090214  |
| cg09661126 | 0.11495487   |
| cg02382818 | -0.101038899 |
| cg07352016 | -0.001865059 |
| cg12272104 | -0.016501795 |
| cg09176309 | 0.319999236  |
| cg09777776 | 0.042008851  |
| cg00244260 | -0.036479436 |
| cg01520578 | -0.045587941 |
| cg00724936 | -0.032365775 |

Continuation of Table S 3.10

|            |              |
|------------|--------------|
| cg06381803 | -0.059415909 |
| cg22992022 | -0.074501786 |
| cg26288449 | -0.200455271 |
| cg00572387 | -0.117493726 |
| cg10516161 | 0.015757187  |
| cg15031924 | 0.041168915  |
| cg24118151 | 0.030420983  |
| cg25779645 | -0.093418182 |
| cg09462575 | 0.06881343   |
| cg09653610 | 0.004597933  |
| cg26322579 | 0.174693576  |
| cg10546626 | 0.106175586  |
| cg18563860 | -0.051684221 |
| cg14813872 | -0.011568011 |
| cg26390598 | 0.051642581  |
| cg13915314 | -0.072963758 |
| cg21980274 | 0.079192719  |
| cg09691433 | 0.009479671  |
| cg07513672 | -0.095702131 |

Table S 3.11 Participant characteristics of the two study populations used for the external validation of the DNAm-based biomarkers of Arsenic exposure [median (25th, 75th percentiles) or n (%)]

|                               | <b>Bangladesh Vitamin E and Selenium Trial Subcohort<br/>(n=400)</b> | <b>Strong Heart Study Subcohort (n=2,320)</b> |
|-------------------------------|----------------------------------------------------------------------|-----------------------------------------------|
| <b>Age (years)</b>            | 44.0 [35.0, 50.3]                                                    | 55.0 [49.2, 62.0]                             |
| <b>Urinary arsenic (µg/g)</b> | 137.3 [76.2, 394.4]                                                  | 11.6 [7.0, 19.1]                              |
| <b>Sex</b>                    |                                                                      |                                               |
| Male                          | 212 (53.0%)                                                          | 962 (41.5%)                                   |
| Female                        | 188 (47.0%)                                                          | 1358 (58.5%)                                  |
| <b>BMI (kg/m<sup>2</sup>)</b> | 19.8 [17.6, 22.2]                                                    | 29.6 [26.2, 33.7]                             |



## **4 SUMMARY AND FUTURE DIRECTIONS**

In summary, this dissertation advances our understanding of cancer etiology through two distinct research foci: the genetic epidemiology of breast cancer and the molecular epidemiology of As toxicity. Through studies that evaluated the effects of inherited variation, the continued improvement of cancer risk prediction models, and the biological effects of an environmental carcinogen, we demonstrate the importance of leveraging information across ancestries and methodologies.

In Chapter 1, we present to our knowledge, the first and largest cross-ancestry GWAS and TWAS of breast cancer intrinsic subtypes, comprising 123,669 cases and 110,277 controls of European and African ancestry from the BCAC and AABCG consortia. We identified 116 GWAS loci associated with intrinsic subtypes, including 13 loci not previously linked to these subtypes. Our TWAS further identified 90 susceptibility genes at 56 loci associated with these subtypes, including 48 novel associations. By integrating findings from this study with prior TWAS results, we observed that 73% of GWAS loci are located within 2 Mb of a TWAS-identified gene, highlighting considerable overlap between genetic risk loci and mapped susceptibility genes. These findings collectively improve our understanding of the genetic basis underlying these subtypes.

In Chapter 2, we developed PRS models of overall BC, estrogen receptor-positive, estrogen receptor-negative, and triple-negative breast cancer for African-ancestry women displaying predictive performance comparable to models developed for European-ancestry women. We demonstrate the ability of these models to stratify individuals by absolute risks of developing

each BC phenotype and validated these models in multiple external populations. Our findings underscore the potential of PRS models to reduce breast cancer disparities by improving risk stratification for guiding targeted screenings and ultimately, earlier interventions. Further validation of these PRS models in additional African-ancestry populations is needed.

In Chapter 3, we advance our understanding of the impacts of As toxicity on the epigenome by identifying >1,000 novel As-associated DNAm changes in whole blood in a high exposure population. By integrating Mendelian Randomization analyses and developing a biomarker of As exposure, we demonstrate the effects of As on the epigenome are causal and portable across populations. Lastly, we provide guidelines for the future development of effective biomarkers including 1) training on a large sample size, 2) including participants with a broad range of exposure levels, 3) ensuring high-quality exposure assessments, including both environmental and biospecimen-derived exposure measures if possible, 4) utilizing individuals with chronic exposure, and 5) replicating biomarkers across multiple external populations. Additional studies are necessary to evaluate the effects of As on DNAm at CpG sites not included on the EPIC array, as well as to determine the long-term utility of our DNAm-based biomarkers for preventing As toxicity in high-risk populations.

## 5 REFERENCES

1. Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA. Cancer J. Clin.* **74**, 12–49 (2024).
2. Bray, F. *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **74**, 229–263 (2024).
3. Cooper, G. M. The Development and Causes of Cancer. in *The Cell: A Molecular Approach. 2nd edition* (Sinauer Associates, 2000).
4. Anand, P. *et al.* Cancer is a Preventable Disease that Requires Major Lifestyle Changes. *Pharm. Res.* **25**, 2097–2116 (2008).
5. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).
6. DeSantis, C. E. *et al.* Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA. Cancer J. Clin.* **66**, 31–42 (2016).
7. DeSantis, C. E. *et al.* Breast cancer statistics, 2019. *CA. Cancer J. Clin.* **69**, 438–451 (2019).
8. Chen, L. & Li, C. I. Racial disparities in breast cancer diagnosis and treatment by hormone receptor and HER2 status. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **24**, 1666–1672 (2015).
9. Dent, R. *et al.* Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clin. Cancer Res.* **13**, 4429–4434 (2007).
10. Anyigba, C. A., Awandare, G. A. & Paemka, L. Breast cancer in sub-Saharan Africa: The current state and uncertain future. *Exp. Biol. Med.* **246**, 1377–1387 (2021).
11. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **68**, 394–424 (2018).
12. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci.* **98**, 10869–10874 (2001).
13. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
14. Zhang, H. *et al.* Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* **52**, 572–581 (2020).
15. Sun, X. *et al.* Case-Case Genome-Wide Analyses Identify Subtype-Informative Variants That Confer Risk for Breast Cancer. *Cancer Res.* **84**, 2533–2548 (2024).
16. The Broad Scope of Health Effects from Chronic Arsenic Exposure: Update on a Worldwide Public Health Problem | Environmental Health Perspectives | Vol. 121, No. 3. <https://ehp.niehs.nih.gov/doi/full/10.1289/ehp.1205875>.
17. Flanagan, S. V., Johnston, R. B. & Zheng, Y. Arsenic in tube well water in Bangladesh: health and economic impacts and implications for arsenic mitigation. *Bull. World Health Organ.* **90**, 839–846 (2012).
18. Gamboa-Loira, B., Cebrián, M. E., Franco-Marina, F. & López-Carrillo, L. Arsenic metabolism and cancer risk: A meta-analysis. *Environ. Res.* **156**, 551–558 (2017).

19. Lamm, S. H., Ferdosi, H., Dissen, E. K., Li, J. & Ahn, J. A Systematic Review and Meta-Regression Analysis of Lung Cancer Risk and Inorganic Arsenic in Drinking Water. *Int. J. Environ. Res. Public. Health* **12**, 15498–15515 (2015).
20. Shajil, C., Chen, R. J. & Mahabal, G. D. Arsenical Keratosis. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
21. Ruiz de Luzuriaga, A. M., Ahsan, H. & Shea, C. R. Arsenical keratoses in Bangladesh--update and prevention strategies. *Dermatol. Clin.* **29**, 45–51 (2011).
22. dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease | International Journal of Epidemiology | Oxford Academic.  
<https://academic.oup.com/ije/article/46/6/1924/4210158>.
23. Argos, M. *et al.* Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study. *The Lancet* **376**, 252–258 (2010).
24. Cortessis, V. K. *et al.* Environmental epigenetics: prospects for studying epigenetic mediation of exposure–response relationships. *Hum. Genet.* **131**, 1565–1589 (2012).
25. Martin, E. M. & Fry, R. C. Environmental Influences on the Epigenome: Exposure-Associated DNA Methylation in Human Populations. *Annu. Rev. Public Health* **39**, 309–333 (2018).
26. Demanelis, K. *et al.* Association of Arsenic Exposure with Whole Blood DNA Methylation: An Epigenome-Wide Study of Bangladeshi Adults. *Environ. Health Perspect.* **127**, 57011 (2019).
27. Bick, A. G. *et al.* Genomic data in the All of Us Research Program. *Nature* **627**, 340–346 (2024).
28. Lee, E. T. *et al.* The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. *Am. J. Epidemiol.* **132**, 1141–1155 (1990).
29. Guo, L. *et al.* Breast cancer heterogeneity and its implication in personalized precision therapy. *Exp. Hematol. Oncol.* **12**, 3 (2023).
30. Orrantia-Borunda, E., Anchondo-Nuñez, P., Acuña-Aguilar, L. E., Gómez-Valles, F. O. & Ramírez-Valdespino, C. A. Subtypes of Breast Cancer. in *Breast Cancer* (ed. Mayrovitz, H. N.) (Exon Publications, Brisbane (AU), 2022).
31. Iqbal, J., Ginsburg, O., Rochon, P. A., Sun, P. & Narod, S. A. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. *JAMA* **313**, 165–173 (2015).
32. Howard, F. M. & Olopade, O. I. Epidemiology of Triple-Negative Breast Cancer: A Review. *Cancer J. Sudbury Mass* **27**, 8–16 (2021).
33. Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).
34. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
35. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).
36. Li, J. L., McClellan, J. C., Zhang, H., Gao, G. & Huo, D. Multi-tissue transcriptome-wide association studies identified 235 genes for intrinsic subtypes of breast cancer. *JNCI J. Natl. Cancer Inst.* **116**, 1105–1115 (2024).

37. Head, S. T. *et al.* Cis- and trans-eQTL TWASs of breast and ovarian cancer identify more than 100 susceptibility genes in the BCAC and OCAC consortia. *Am. J. Hum. Genet.* **111**, 1084–1099 (2024).
38. Ping, J. *et al.* Using genome and transcriptome data from African-ancestry female participants to identify putative breast cancer susceptibility genes. *Nat. Commun.* **15**, 3718 (2024).
39. Jia, G. *et al.* Genome-wide association analyses of breast cancer in women of African ancestry identify new susceptibility loci and improve risk prediction. *Nat. Genet.* **56**, 819–826 (2024).
40. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
41. Bhat-Nakshatri, P. *et al.* Acquisition, processing, and single-cell analysis of normal human breast tissues from a biobank. *STAR Protoc.* **3**, 101047 (2021).
42. Zhang, H. *et al.* A mixed-model approach for powerful testing of genetic associations with cancer risk incorporating tumor characteristics. *Biostat. Oxf. Engl.* **22**, 772–788 (2021).
43. Pérez-Benavente, B. *et al.* New roles for AP-1/JUNB in cell cycle control and tumorigenic cell invasion via regulation of cyclin E1 and TGF- $\beta$ 2. *Genome Biol.* **23**, 252 (2022).
44. Lekan Deprez, R. H. *et al.* Cloning and characterization of MN1, a gene from chromosome 22q11, which is disrupted by a balanced translocation in a meningioma. *Oncogene* **10**, 1521–1528 (1995).
45. Zhang, Z. *et al.* XRCC5 cooperates with p300 to promote cyclooxygenase-2 expression and tumor growth in colon cancers. *PloS One* **12**, e0186900 (2017).
46. Liu, T. *et al.* The NRF2-dependent transcriptional axis, XRCC5/hTERT drives tumor progression and 5-Fu insensitivity in hepatocellular carcinoma. *Mol. Ther. Oncolytics* **24**, 249–261 (2021).
47. Chen, Q., Zhou, X.-W., Zhang, A.-J. & He, K. ACTN1 supports tumor growth by inhibiting Hippo signaling in hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* **40**, 23 (2021).
48. Foster, H. *et al.* ATMIN is a tumor suppressor gene in lung adenocarcinoma. *Cancer Res.* **79**, 5159–5166 (2019).
49. Ullah, F., Khurshid, N., Fatimi, Q., Loidl, P. & Saeed, M. Mutations in the acetylation hotspots of Rbl2 are associated with increased risk of breast cancer. *PLoS ONE* **17**, e0266196 (2022).
50. Gonzalez-Gomez, P. *et al.* Retinoblastoma-related gene RB2/p130 exons 19-22 are rarely mutated in glioblastomas. *Oncol. Rep.* **9**, 951–954 (2002).
51. Barbie, T. U., Barbie, D. A., MacLaughlin, D. T., Maheswaran, S. & Donahoe, P. K. Mullerian Inhibiting Substance inhibits cervical cancer cell growth via a pathway involving p130 and p107. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15601–15606 (2003).
52. Zhao, T.-T. *et al.* TRIM32 promotes proliferation and confers chemoresistance to breast cancer cells through activation of the NF- $\kappa$ B pathway. *J. Cancer* **9**, 1349–1356 (2018).
53. ShahidSales, S. *et al.* A genetic variant in CDKN2A/B gene is associated with the increased risk of breast cancer. *J. Clin. Lab. Anal.* **32**, e22190 (2017).
54. Nielsen, N. H., Roos, G., Emdin, S. O. & Landberg, G. Methylation of the p16(Ink4a) tumor suppressor gene 5'-CpG island in breast cancer. *Cancer Lett.* **163**, 59–69 (2001).
55. Yang, J. *et al.* High expression of small GTPase Rab3D promotes cancer progression and metastasis. *Oncotarget* **6**, 11125–11138 (2015).

56. Wu, Y. & Chen, G. Y. The Novel Role of GDI2: A Mini-Review. *Ann. Med. Med. Res.* **7**, 1083 (2024).
57. Gopal Krishnan, P. D., Golden, E., Woodward, E. A., Pavlos, N. J. & Blancafort, P. Rab GTPases: Emerging Oncogenes and Tumor Suppressive Regulators for the Editing of Survival Pathways in Cancer. *Cancers* **12**, 259 (2020).
58. Yao, T. *et al.* Identifying therapeutic targets for breast cancer: insights from systematic Mendelian randomization analysis. *Front. Oncol.* **14**, 1407795 (2024).
59. Peng, Y. *et al.* LINC01376 promotes nasopharyngeal carcinoma tumorigenesis by competitively binding to the SP1/miR-4757/IGF1 axis. *IUBMB Life* **75**, 702–716 (2023).
60. Mello, A. C., Freitas, M., Coutinho, L., Falcon, T. & Matte, U. Machine Learning Supports Long Noncoding RNAs as Expression Markers for Endometrial Carcinoma. *BioMed Res. Int.* **2020**, 3968279 (2020).
61. Qin, J. *et al.* The Role of the Linc01376/Mir-15b-3p\_R-1/Fgf2 Axis in A549 and H1299 Cells Emt Induced by Lmw-Pahs. SSRN Scholarly Paper at <https://doi.org/10.2139/ssrn.4767257> (2024).
62. Garcia-Closas, M. *et al.* Genome-wide association studies identify four ER negative–specific breast cancer risk loci. *Nat. Genet.* **45**, 392–398 (2013).
63. Huo, D. *et al.* Comparison of Breast Cancer Molecular Features and Survival by African and European Ancestry in The Cancer Genome Atlas. *JAMA Oncol.* **3**, 1654–1662 (2017).
64. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
65. Zhou, H., Alexander, D. & Lange, K. A quasi-Newton acceleration for high-dimensional optimization algorithms. *Stat. Comput.* **21**, 261–273 (2011).
66. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
67. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, s13742-015-0047–8 (2015).
68. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
69. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375, S1-3 (2012).
70. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
71. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
72. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
73. Krueger, F. FelixKrueger/TrimGalore. (2025).
74. BioInfoTools/BBMap. BioInfoTools (2025).
75. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
76. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).
77. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
78. broadinstitute/picard. Broad Institute (2025).

79. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
80. Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
81. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
82. Zhou, H. J., Li, L., Li, Y., Li, W. & Li, J. J. PCA outperforms popular hidden variable inference methods for molecular QTL mapping. *Genome Biol.* **23**, 210 (2022).
83. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
84. Aran, D., Hu, Z. & Butte, A. J. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **18**, 220 (2017).
85. Song, X. *et al.* MiXcan: a framework for cell-type-aware transcriptome-wide association studies with an application to breast cancer. *Nat. Commun.* **14**, 377 (2023).
86. Privé, F., Aschard, H., Ziyatdinov, A. & Blum, M. G. B. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. *Bioinforma. Oxf. Engl.* **34**, 2781–2787 (2018).
87. Barbeira, A. N. *et al.* Fine-mapping and QTL tissue-sharing information improves the reliability of causal gene identification. *Genet. Epidemiol.* **44**, 854–867 (2020).
88. Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA. Cancer J. Clin.* **74**, 12–49 (2024).
89. Mavaddat, N. *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am. J. Hum. Genet.* **104**, 21–34 (2019).
90. Du, Z. *et al.* Evaluating Polygenic Risk Scores for Breast Cancer in Women of African Ancestry. *J. Natl. Cancer Inst.* **113**, 1168–1176 (2021).
91. Shifman, S., Kuypers, J., Kokoris, M., Yakir, B. & Darvasi, A. Linkage disequilibrium patterns of the human genome across populations. *Hum. Mol. Genet.* **12**, 771–776 (2003).
92. Huo, D. *et al.* Genome-wide association studies in women of African ancestry identified 3q26.21 as a novel susceptibility locus for oestrogen receptor negative breast cancer. *Hum. Mol. Genet.* **25**, 4835–4846 (2016).
93. Stacey, S. N. *et al.* Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.* **39**, 865–869 (2007).
94. Zheng, W. *et al.* Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat. Genet.* **41**, 324–328 (2009).
95. Chen, F. *et al.* Fine-mapping of breast cancer susceptibility loci characterizes genetic risk in African Americans. *Hum. Mol. Genet.* **20**, 4491–4503 (2011).
96. Zheng, W. *et al.* Evaluation of 11 breast cancer susceptibility loci in African-American women. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **18**, 2761–2764 (2009).
97. Gao, G. *et al.* Polygenic risk scores for prediction of breast cancer risk in women of African ancestry: a cross-ancestry approach. *Hum. Mol. Genet.* **31**, 3133–3143 (2022).
98. Zirpoli, G. R. *et al.* Addition of polygenic risk score to a risk calculator for prediction of breast cancer in US Black women. *Breast Cancer Res. BCR* **26**, 2 (2024).

99. Jia, G. *et al.* Genome-wide association analyses of breast cancer in women of African ancestry identify new susceptibility loci and improve risk prediction. *Nat. Genet.* **56**, 819–826 (2024).
100. Zhang, H. *et al.* A new method for multi-ancestry polygenic prediction improves performance across diverse populations. *Nat. Genet.* **55**, 1757–1768 (2023).
101. Ruan, Y. *et al.* Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).
102. Zhang, H. *et al.* Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* **52**, 572–581 (2020).
103. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
104. Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience* **8**, giz082 (2019).
105. Privé, F., Arbel, J., Aschard, H. & Vilhjálmsson, B. J. Identifying and correcting for misspecifications in GWAS summary statistics and polygenic scores. *Hum. Genet. Genomics Adv.* **3**, 100136 (2022).
106. Privé, F., Albiñana, C., Arbel, J., Pasaniuc, B. & Vilhjálmsson, B. J. Inferring disease architecture and predictive ability with LDpred2-auto. *Am. J. Hum. Genet.* **110**, 2042–2055 (2023).
107. Tomlinson-Hansen, S. E., Budh, D. P. & Sapra, A. Breast Cancer Screening in the Average-Risk Patient. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
108. Choudhury, P. P. *et al.* iCARE: An R package to build, validate and apply absolute risk models. *PLOS ONE* **15**, e0228198 (2020).
109. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
110. Mooney, J. A., Agranat-Tamir, L., Pritchard, J. K. & Rosenberg, N. A. On the number of genealogical ancestors tracing to the source groups of an admixed population. *Genetics* **224**, iyad079 (2023).
111. Oeffinger, K. C. *et al.* Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society. *JAMA* **314**, 1599–1614 (2015).
112. Hu, C. *et al.* A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N. Engl. J. Med.* **384**, 440–451 (2021).
113. Breast Cancer Association Consortium *et al.* Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N. Engl. J. Med.* **384**, 428–439 (2021).
114. Guindalini, R. S. C. *et al.* Intensive Surveillance with Biannual Dynamic Contrast-Enhanced Magnetic Resonance Imaging Downstages Breast Cancer in BRCA1 Mutation Carriers. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **25**, 1786–1794 (2019).
115. null, null. The “All of Us” Research Program. *N. Engl. J. Med.* **381**, 668–676 (2019).
116. Roden, D. M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.* **84**, 362–369 (2008).
117. Kasimatis, K. R. *et al.* Evaluating human autosomal loci for sexually antagonistic viability selection in two large biobanks. *Genetics* **217**, iyaa015 (2021).
118. Pal, T. *et al.* A high frequency of BRCA mutations in young black women with breast cancer residing in Florida. *Cancer* **121**, 4173–4180 (2015).



119. Trowbridge, J. *et al.* Exposure to perfluoroalkyl substances in a cohort of women firefighters and office workers in San Francisco. *Environ. Sci. Technol.* **54**, 3363–3374 (2020).
120. Gao, G. *et al.* Polygenic risk scores for prediction of breast cancer risk in women of African ancestry: a cross-ancestry approach. *Hum. Mol. Genet.* **31**, 3133–3143 (2022).
121. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
122. Du, Z. *et al.* Evaluating Polygenic Risk Scores for Breast Cancer in Women of African Ancestry. *JNCI J. Natl. Cancer Inst.* **113**, 1168 (2021).
123. Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer risk loci. *Nature* **551**, 92–94 (2017).
124. Milne, R. L. *et al.* Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat. Genet.* **49**, 1767–1778 (2017).
125. Ruan, Y. *et al.* Improving Polygenic Prediction in Ancestrally Diverse Populations. *Nat. Genet.* **54**, 573–580 (2022).
126. Enhancing polygenic risk prediction in diverse populations: opportunities and challenges. *Nat. Genet.* **55**, 1621–1622 (2023).
127. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
128. Rodríguez-Álvarez, M. X. & Inácio, V. ROCnReg: An R Package for Receiver Operating Characteristic Curve Inference With and Without Covariates. *R J.* **13**, 525–555 (2021).
129. Mavaddat, N. *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J. Natl. Cancer Inst.* **107**, djv036 (2015).
130. Mavaddat, N. *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am. J. Hum. Genet.* **104**, 21 (2019).
131. Hopper, J. L. *et al.* Variance of age-specific log incidence decomposition (VALID): a unifying model of measured and unmeasured genetic and non-genetic risks. *Int. J. Epidemiol.* **52**, 1557–1568 (2023).
132. Ryu, H.-W. *et al.* Influence of Toxicologically Relevant Metals on Human Epigenetic Regulation. *Toxicol. Res.* **31**, 1 (2015).
133. Kim, N., Filipovic, D., Bhattacharya, S. & Cuddapah, S. Epigenetic toxicity of heavy metals – implications for embryonic stem cells. *Environ. Int.* **193**, 109084 (2024).
134. Elkin, E. R., Higgins, C., Aung, M. T. & Bakulski, K. M. Metals exposures and DNA methylation: Current evidence and future directions. *Curr. Environ. Health Rep.* **9**, 673 (2022).
135. Ruiz-Hernandez, A. *et al.* Environmental chemicals and DNA methylation in adults: a systematic review of the epidemiologic evidence. *Clin. Epigenetics* **7**, 55 (2015).
136. Colicino, E. *et al.* Blood DNA Methylation Biomarkers of Cumulative Lead Exposure in Adults. *J. Expo. Sci. Environ. Epidemiol.* **31**, 108 (2019).
137. Moon, K. A. *et al.* A dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease. *Int. J. Epidemiol.* **47**, 1013 (2018).
138. Ahsan, H. *et al.* Health Effects of Arsenic Longitudinal Study (HEALS): description of a multidisciplinary epidemiologic investigation. *J. Expo. Sci. Environ. Epidemiol.* **16**, 191–205 (2006).

139. Pierce, B. L. *et al.* Genome-Wide Association Study Identifies Chromosome 10q24.32 Variants Associated with Arsenic Metabolism and Toxicity Phenotypes in Bangladesh. *PLoS Genet.* **8**, e1002522 (2012).
140. Chen, Y. *et al.* Reduction in urinary arsenic levels in response to arsenic mitigation efforts in Arai hazar, Bangladesh. *Environ. Health Perspect.* **115**, 917–923 (2007).
141. Nixon, D. E., Mussmann, G. V., Eckdahl, S. J. & Moyer, T. P. Total arsenic in urine: palladium-persulfate vs nickel as a matrix modifier for graphite furnace atomic absorption spectrophotometry. *Clin. Chem.* **37**, 1575–1579 (1991).
142. Gamble, M. V. *et al.* Folic acid supplementation lowers blood arsenic. *Am. J. Clin. Nutr.* **86**, 1202–1209 (2007).
143. Pidsley, R. *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* **17**, 208 (2016).
144. Fortin, J.-P., Triche, T. J. & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558–560 (2017).
145. Teschendorff, A. E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinforma. Oxf. Engl.* **29**, 189–196 (2013).
146. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
147. Teschendorff, A. E., Breeze, C. E., Zheng, S. C. & Beck, S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinformatics* **18**, 105 (2017).
148. Reinius, L. E. *et al.* Differential DNA Methylation in Purified Human Blood Cells: Implications for Cell Lineage and Studies on Disease Susceptibility. *PLOS ONE* **7**, e41361 (2012).
149. Teschendorff, A. E., Breeze, C. E., Zheng, S. C. & Beck, S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinformatics* **18**, 105 (2017).
150. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
151. Buja, A. & Eyuboglu, N. Remarks on Parallel Analysis. *Multivar. Behav. Res.* **27**, 509–540 (1992).
152. McGregor, K. *et al.* An evaluation of methods correcting for cell-type heterogeneity in DNA methylation studies. *Genome Biol.* **17**, 84 (2016).
153. Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
154. Leek, J. T. & Storey, J. D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* **3**, 1724–1735 (2007).
155. Nassar, L. R. *et al.* The UCSC Genome Browser database: 2023 update. *Nucleic Acids Res.* **51**, D1188–D1195 (2023).
156. Phipson, B., Maksimovic, J. & Oshlack, A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* **32**, 286–288 (2016).
157. Battram, T. *et al.* The EWAS Catalog: a database of epigenome-wide association studies. *Wellcome Open Res.* **7**, 41 (2022).

158. Zheng, S. C., Breeze, C. E., Beck, S. & Teschendorff, A. E. Identification of differentially methylated cell-types in Epigenome-Wide Association Studies. *Nat. Methods* **15**, 1059–1066 (2018).
159. Wang, Y., Grant, O. A., Zhai, X., McDonald-Maier, K. D. & Schalkwyk, L. C. Insights into ageing rates comparison across tissues from recalibrating cerebellum DNA methylation clock. *GeroScience* **46**, 39–56 (2024).
160. Hannum, G. *et al.* Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Mol. Cell* **49**, 359–367 (2013).
161. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, 3156 (2013).
162. Yang, Z. *et al.* Correlation of an epigenetic mitotic clock with cancer risk. *Genome Biol.* **17**, 205 (2016).
163. Zhang, Y. *et al.* DNA methylation signatures in peripheral blood strongly predict all-cause mortality. *Nat. Commun.* **8**, 14617 (2017).
164. Horvath, S. *et al.* Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging* **10**, 1758–1775 (2018).
165. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
166. McEwen, L. M. *et al.* The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proc. Natl. Acad. Sci.* **117**, 23329–23335 (2020).
167. Zhang, Q. *et al.* Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Med.* **11**, 54 (2019).
168. Lu, A. T. *et al.* DNA methylation-based estimator of telomere length. *Aging* **11**, 5895–5923 (2019).
169. Shireby, G. L. *et al.* Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. *Brain* **143**, 3763–3775 (2020).
170. Belsky, D. W. *et al.* DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife* **11**, e73420 (2022).
171. Higgins-Chen, A. T. *et al.* A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nat. Aging* **2**, 644–661 (2022).
172. Bernabeu, E. *et al.* Refining epigenetic prediction of chronological and biological age. *Genome Med.* **15**, 12 (2023).
173. Lu, A. T. *et al.* Universal DNA methylation age across mammalian tissues. *Nat. Aging* **3**, 1144–1166 (2023).
174. Tseng, C.-H. A review on environmental factors regulating arsenic methylation in humans. *Toxicol. Appl. Pharmacol.* **235**, 338–350 (2009).
175. Pierce, B. L. *et al.* A missense variant in FTCD is associated with arsenic metabolism and toxicity phenotypes in Bangladesh. *PLoS Genet.* **15**, e1007984 (2019).
176. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
177. Janes, H. & Pepe, M. S. Adjusting for covariate effects on classification accuracy using the covariate-adjusted receiver operating characteristic curve. *Biometrika* **96**, 371–382 (2009).
178. Argos, M. *et al.* Baseline comorbidities in a skin cancer prevention trial in Bangladesh. *Eur. J. Clin. Invest.* **43**, 579–588 (2013).

179. Bozack, A. K. *et al.* Locus-Specific Differential DNA Methylation and Urinary Arsenic: An Epigenome-Wide Association Study in Blood among Adults with Low-to-Moderate Arsenic Exposure. *Environ. Health Perspect.* **128**, 67015 (2020).
180. Lee, E. T. *et al.* The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. *Am. J. Epidemiol.* **132**, 1141–1155 (1990).
181. Kuo, C.-C. *et al.* The association of arsenic exposure and arsenic metabolism with all-cause, cardiovascular and cancer mortality in the Strong Heart Study. *Environ. Int.* **159**, 107029 (2022).
182. Domingo-Relloso, A. *et al.* Arsenic Exposure, Blood DNA Methylation, and Cardiovascular Disease. *Circ. Res.* **131**, e51–e69 (2022).
183. Moon, K. A. *et al.* A dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease. *Int. J. Epidemiol.* **46**, 1924–1939 (2017).
184. Navas-Acien, A. *et al.* Urine arsenic concentrations and species excretion patterns in American Indian communities over a 10-year period: the Strong Heart Study. *Environ. Health Perspect.* **117**, 1428–1433 (2009).
185. Nigra, A. E. *et al.* The effect of the Environmental Protection Agency maximum contaminant level on arsenic exposure in the USA from 2003 to 2014: an analysis of the National Health and Nutrition Examination Survey (NHANES). *Lancet Public Health* **2**, e513–e521 (2017).
186. Nigra, A. E. *et al.* Inequalities in Public Water Arsenic Concentrations in Counties and Community Water Systems across the United States, 2006-2011. *Environ. Health Perspect.* **128**, 127001 (2020).
187. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
188. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
189. Hammal, F., de Langen, P., Bergon, A., Lopez, F. & Ballester, B. ReMap 2022: a database of Human, Mouse, Drosophila and Arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.* **50**, D316–D325 (2022).
190. Yu, Q. *et al.* Long non-coding RNA LHX1-DT regulates cardiomyocyte differentiation through H2A.Z-mediated LHX1 transcriptional activation. *iScience* **26**, 108051 (2023).
191. Hillary, R. F. *et al.* Blood-based epigenome-wide analyses of chronic low-grade inflammation across diverse population cohorts. *Cell Genomics* **4**, 100544 (2024).
192. Hillary, R. F. *et al.* Blood-based epigenome-wide analyses on the prevalence and incidence of nineteen common disease states. 2023.01.10.23284387 Preprint at <https://doi.org/10.1101/2023.01.10.23284387> (2023).
193. Xu, C.-J. *et al.* The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics* **18**, 25 (2017).
194. Wozniak, M. B. *et al.* Integrative genome-wide gene expression profiling of clear cell renal cell carcinoma in Czech Republic and in the United States. *PloS One* **8**, e57886 (2013).
195. Bonder, M. J. *et al.* Genetic and epigenetic regulation of gene expression in fetal and adult human livers. *BMC Genomics* **15**, 860 (2014).
196. Ortiz-Fernández, L. *et al.* The role of a functional variant of TYK2 in vasculitides and infections. *Clin. Exp. Rheumatol.* **38**, 949–955 (2020).

197. Argos, M. *et al.* Gene-Specific Differential DNA Methylation and Chronic Arsenic Exposure in an Epigenome-Wide Association Study of Adults in Bangladesh. *Environ. Health Perspect.* **123**, 64–71 (2015).
198. Kaushal, A. *et al.* Genome-wide DNA methylation at birth in relation to in utero arsenic exposure and the associated health in later life. *Environ. Health Glob. Access Sci. Source* **16**, 50 (2017).
199. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).
200. Parvez, F. *et al.* Arsenic Exposure and Impaired Lung Function. Findings from a Large Population-based Prospective Cohort Study. *Am. J. Respir. Crit. Care Med.* **188**, 813–819 (2013).
201. Navas-Acien, A. *et al.* Arsenic Exposure and Type 2 Diabetes: A Systematic Review of the Experimental and Epidemiologic Evidence. *Environ. Health Perspect.* **114**, 641–648 (2006).
202. Zhang, Z. *et al.* Association between arsenic exposure and inflammatory cytokines and C-reaction protein: A systematic review and meta-analysis. *Medicine (Baltimore)* **101**, e32352 (2022).
203. Palma-Lara, I. *et al.* Arsenic exposure: A public health problem leading to several cancers. *Regul. Toxicol. Pharmacol. RTP* **110**, 104539 (2020).
204. Smith, G. D. & Ebrahim, S. ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
205. Mohammed Abdul, K. S., Jayasinghe, S. S., Chandana, E. P. S., Jayasumana, C. & De Silva, P. M. C. S. Arsenic and human health effects: A review. *Environ. Toxicol. Pharmacol.* **40**, 828–846 (2015).
206. Ladd-Acosta, C. Epigenetic Signatures as Biomarkers of Exposure. *Curr. Environ. Health Rep.* **2**, 117–125 (2015).
207. Baccarelli, A., Dolinoy, D. C. & Walker, C. L. A precision environmental health approach to prevention of human disease. *Nat. Commun.* **14**, 2449 (2023).
208. Stevens, M. *et al.* Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. *Genome Res.* **23**, 1541–1553 (2013).