scientific reports

Check for updates

OPEN

Clinical and molecular correlates of tumor aneuploidy in metastatic non-small cell lung cancer

Liam F. Spurr[1,2] & Sean P. Pitroda[2,3]✉

Recent studies have linked elevated tumor aneuploidy to anti-tumor immune suppression and adverse survival following immunotherapy. Herein, we provide supportive evidence for tumor aneuploidy as a biomarker of response to immunotherapy in patients with non-small cell lung cancer (NSCLC). We identify a dose–response relationship between aneuploidy score and patient outcomes. In two independent NSCLC cohorts (n = 659 patients), we demonstrate a novel association between elevated aneuploidy and non-smoking-associated oncogenic driver mutations. Lastly, we report enrichment of *TERT* amplification and immune-suppressive phenotypes of highly aneuploid NSCLC. Taken together, our findings emphasize a potentially critical role for tumor aneuploidy in guiding immunotherapy treatment strategies.

Keywords Aneuploidy, Immunotherapy, Anti-tumor immunity, Lung cancer

Aneuploidy, an unbalanced chromosome or chromosome arm number, is the most common somatic alteration in human cancer[1]. While our understanding of the role of aneuploidy in the initiation and propagation of cancer cells is evolving, recent studies have indicated that elevated levels of aneuploidy may provide a fitness advantage for malignant cells[2], resulting in a more aggressive tumor phenotype and deleterious prognosis[3,4]. In addition, the available evidence suggests that aneuploidy may promote cancer growth in part through impairing innate and adaptive anti-tumor immunity[5–7]. For instance, aneuploidy has been shown to reduce CD8+ T cell and natural killer cell function and infiltration. This has raised questions about whether aneuploidy may impact the efficacy of immunotherapy, with recent studies showing an adverse prognosis following treatment with immune checkpoint inhibitors (ICIs)[8,9]. This work has indicated that aneuploidy provides synergistic prognostic value with established predictors of ICI response, such as tumor mutational burden (TMB)[8,10]. Validation of these findings in large, multi-institutional datasets will be critical in assessing the reliability and generalizability of aneuploidy as a biomarker of ICI response.

ICIs have been particularly central to modern therapeutic strategies for metastatic non-small cell lung cancer (NSCLC). However, patients exhibit strikingly heterogeneous responses to ICIs, and NSCLC remains the leading cause of cancer related death in the United States. In particular, patients whose tumors exhibit a lower TMB or tumor programmed-cell death ligand 1 (*PD-L1*) expression and patients without a history of smoking (whose tumors tend to be driven by oncogenic alterations such as *EGFR* mutations), frequently have a poor response to treatment with ICIs[11,12]. In addition, many patients with favorable biomarker profiles, such as high TMB, frequently do not respond to ICI. Based on the known associations with anti-tumor immunity and poor prognosis following immunotherapy, it is possible that the effects of aneuploidy on tumor biology may mediate these observations.

To alleviate the deleterious prognosis associated with highly aneuploid NSCLC, adjunctive treatment strategies are critical. Some potential strategies have been proposed, including targeting specific oncogenes[13] or taking advantage of vulnerability to cytotoxic therapies[7,14]. It is likely that there are additional analogous avenues through which highly aneuploid tumors can be targeted, and further understanding of the genomic and transcriptomic features associated with aneuploidy is important to identifying candidate therapeutic vulnerabilities.

Herein, we set out to validate previous work from our group and others[6,8,9] demonstrating that elevated tumor aneuploidy is associated with poor survival following treatment with ICIs in a multi-institutional cohort[15] of 309 patients (SU2C-MARK Foundation cohort) with NSCLC treated with immunotherapy. We extended these findings by comprehensively characterizing the clinicopathologic, immunologic, and molecular correlates of

[1]Pritzker School of Medicine, The University of Chicago, Chicago, IL, USA. [2]Department of Radiation and Cellular Oncology, The University of Chicago, Chicago, IL, USA. [3]Ludwig Center for Metastasis Research, The University of Chicago, 5758 S. Maryland Ave. MC 9006, Chicago, IL 60637, USA. ✉email: spitroda@uchicago.edu

elevated tumor aneuploidy in metastatic NSCLC. Importantly, we validated these novel findings in an independent cohort of 350 patients with non-small cell lung cancer treated with immunotherapy (MSK-IMPACT cohort).

Results

Validation of the prognostic value of aneuploidy score

We first confirmed that tumor mutational burden (TMB) and aneuploidy score (AS) independently predict survival in the SU2C-MARK cohort. We utilized previously established ASCETS[16] parameters for defining arm-level somatic copy-number alterations (aSCNAs)[7,8] and validated that both TMB and AS were significantly associated with overall survival (OS; Wald $P < 0.05$).

To optimize aSCNA delineation, we then used bootstrap resampling to tune the ASCETS parameters used to call aSCNAs for this cohort (Methods, Supplementary Fig. 1a,b, Supplementary Table 1–3). Using these optimized parameters, both TMB and AS were again associated with OS (HR [95% CI] for $\log_{10}$(TMB): 0.54 [0.36–0.80], Wald $P = 1.9e{-3}$; HR [95% CI] for AS: 3.08 [1.57–6.04], Wald $P = 1.1e{-3}$) in a multivariable Cox proportional hazards model. In addition, TMB (split at the top 20%ile[10]) and AS (split at the 50%ile[8], 0.49 in this cohort) were significantly associated with OS (Log-rank $P = 1.7e{-05}$; Fig. 1a). Specifically, among low-TMB (TMB-L) tumors, those with high AS (AS-H) exhibited inferior OS (median TMB-L/AS-H: 15.6 months vs median TMB-L/AS-L: 22.7 months, Log-rank $P = 0.012$). Among high TMB tumors, there was no difference in OS between AS-H vs. AS-L tumors (Log-rank $P = 0.38$).

We also assessed the relationship between AS and PFS in the SU2C-MARK cohort. In a multivariable Cox proportional hazards model both TMB and AS were associated with PFS (HR [95% CI] for $\log_{10}$(TMB): 0.38 [0.27–0.52], Wald $P = 5.6e{-9}$; HR [95% CI] for AS: 1.76 [1.02–3.04], Wald $P = 0.044$).When patients were binned into groups of high and low AS and TMB as above, there was a deleterious effect of aneuploidy on prognosis noted among both high and low TMB tumors (Supplementary Fig. 1a).

In addition, we examined the effects of AS and TMB on objective response rate (ORR, the fraction of patients who exhibited a partial or complete response by RECIST criteria) and disease control rate (DCR, the fraction of patients who exhibited a partial or complete response or stable disease as their best response). High TMB tumors exhibited higher DCR and ORR rates than low TMB tumors regardless of AS. Within TMB groups, high AS tumors exhibited lower rates of response by both measures. For example, among patients with high TMB tumors, those with a low AS had an ORR of 76.2% compared to 63.4% for high AS and a DCR of 95.2% versus 80.5% (Fig. 1b). Taken together, these findings suggested that highly aneuploid tumors are less likely to respond to immunotherapy regardless of TMB, but that subsequent salvage therapies are more efficacious in high TMB tumors leading to similar overall survival between AS-L/TMB-H and AS-H/TMB-H tumors. By contrast, salvage therapies are less effective in low TMB tumors given the significant difference in overall survival observed between AS-L/TMB-L and AS-H/TMB-L tumors.

Multivariable analysis

To assess the relative importance of aneuploidy score as a predictor of survival while controlling for important clinicopathologic variables, we performed least absolute shrinkage and selection operator (LASSO) Cox regression to control for potential confounding variables, which confirmed AS is an independent predictor of OS, in addition to TMB, smoking history (pack-years), number of lines of prior therapy, immunotherapy agent type, PD-L1 tumor proportion score (TPS), prior use of platinum agents, as well as tumor histology (Supplementary Table 1).

Given the heterogeneous nature of the immunotherapy treatment paradigms among these patients and the differential effects on survival noted above, we repeated the survival analysis separately among patients treated with single agent PD-(L)1 directed checkpoint inhibitors versus combination anti-PD-(L)1 and CTLA-4 therapy (with the exception of 1% of patients treated with combination chemotherapy and anti-PD-(L)1 agents). Interestingly, among low TMB, but not high TMB, tumors treated with single-agent immune checkpoint inhibitors, patients with highly aneuploid tumors had worse survival than those with low levels of aneuploidy. On the other hand, we found that there was no difference in survival by aneuploidy grouping among either high or low TMB tumors in patients treated with combination therapy, suggesting that perhaps the use of combination therapy could at least partially attenuate the negative effects of elevated aneuploidy on survival following immunotherapy (Supplementary Fig. 1b).

We also examined whether specific aSCNAs drive the AS associations with survival. Only *2q* deletion among adenocarcinomas, *11q* deletion across all tumors, and *13q* amplification across all tumors were associated with survival (Wald $Q < 0.05$, Supplementary Table 5). We repeated the analysis controlling for overall AS, as in general, any specific aSCNA is more likely to be present by chance in a tumor with high aneuploidy. We found that no specific aSCNAs were significantly associated with survival in a multivariable Cox proportional hazards model (Wald $Q > 0.05$, Supplementary Table 6), thus highlighting the importance of global aneuploidy on clinical outcomes.

Dose–response relationship between aneuploidy score and clinical outcomes

While binarizing continuous values is clinically useful, it also presents challenges in terms of statistical rigor. Thus, we investigated whether there was a gradient of response probabilities with increasing AS. At each unique AS, we quantified (1) the fraction of patients exhibiting complete or partial responses by RECIST criteria and (2) two-year OS in patients with an AS greater than or equal to the candidate score. We identified a dose–response relationship between AS and both RECIST response and 2-year OS probability. The probability of 2-year OS increased from 10–20% to 40–50% as AS decreased from 1.0 to approximately 0.46 (Fig. 1c, Supplementary Table 7). Notably, an AS of 0.46 is similar to both the cohort median of 0.49 as used above and our previously derived optimized threshold of 0.42[7]. Similarly, the probability of response increased from ~ 10–20% to ~ 35%

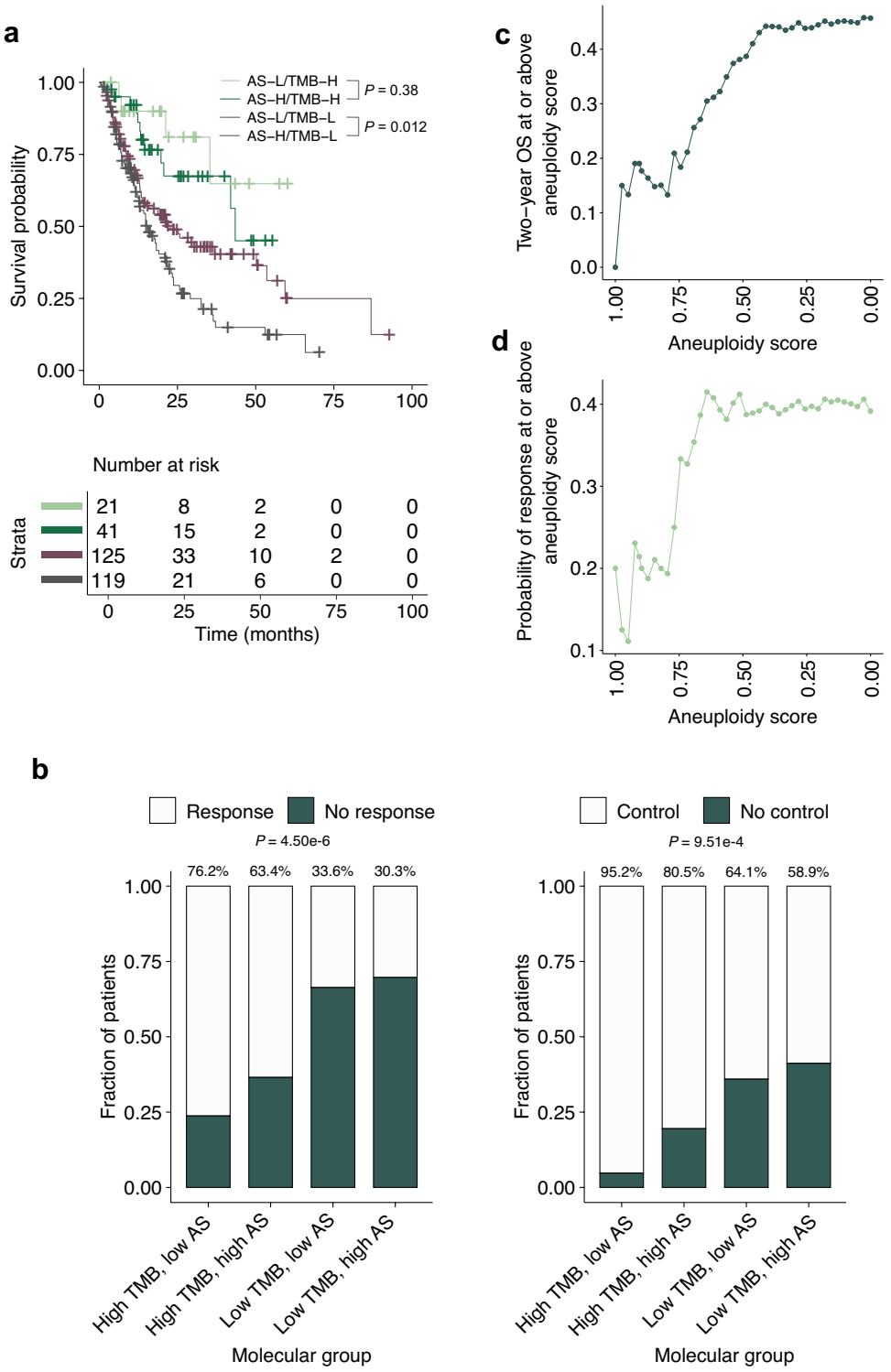
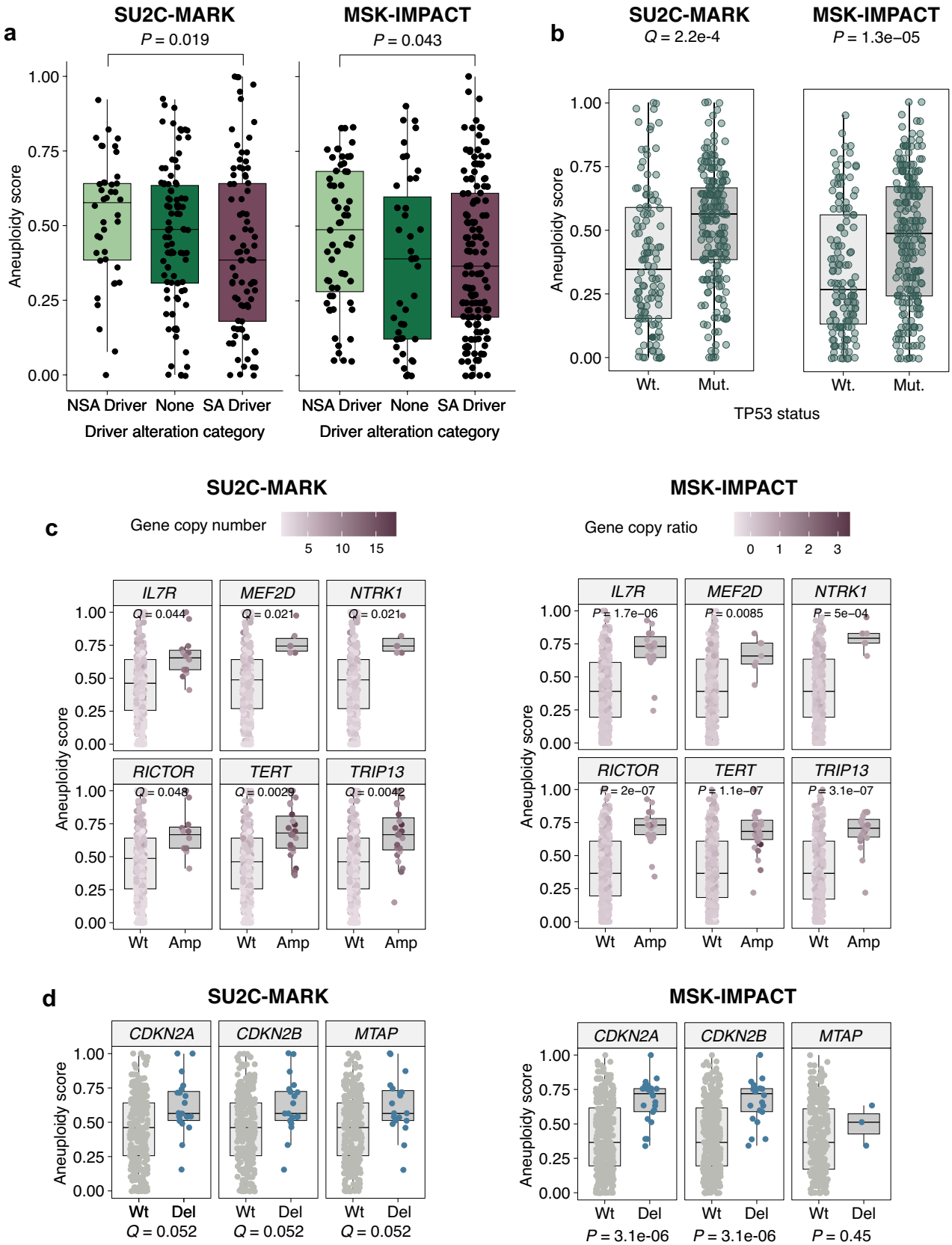


Figure 1. Association between aneuploidy score and survival. (**a**) Kaplan–Meier curves showing association of aneuploidy score (AS, binned at the median) and tumor mutational burden (TMB, binned at the top 20th percentile) with overall survival (OS) in the SU2C-MARK cohort. Multivariable Wald *P* values for pairwise comparisons are shown. (H: high, L: low) (**b**) Bar plots of objective response rate (ORR) and disease control rate (DCR) among the prognostic groups (TMB: tumor mutational burden, AS: aneuploidy score). Global Fisher's exact test *P* values and response rates for each group are displayed above each bar (n = 309 samples). (**c**) Relationship between AS and probability of response (n = 309 samples) and (**d**) two-year overall survival (OS) among tumors with aneuploidy scores at least as great as the x-axis value (n = 306 samples).


a
SU2C-MARK
P = 0.019
MSK-IMPACT
P = 0.043
Aneuploidy score
NSA Driver None SA Driver
Driver alteration category
b
SU2C-MARK
Q = 2.2e-4
MSK-IMPACT
P = 1.3e–05
Wt. Mut.
TP53 status
c
SU2C-MARK
Gene copy number
MSK-IMPACT
Gene copy ratio
IL7R
MEF2D
NTRK1
RICTOR
TERT
TRIP13
Wt Amp
d
SU2C-MARK
MSK-IMPACT
CDKN2A
CDKN2B
MTAP
Wt Del
Q = 0.052
P = 3.1e-06
P = 0.45

◂**Figure 2.** Relationships between aneuploidy and genomic features in the SU2C-MARK and MSK-IMPACT cohorts. (**a**) Boxplot comparing aneuploidy scores (AS) by tumors with smoking-associated driver mutations (SA Drivers), driver mutations not associated with smoking (NSA Drivers) and tumors with no known driver mutations (None) in patients with lung adenocarcinoma; Kruskal–Wallis test (SU2C-MARK: n = 220 patients, MSK-IMPACT: n = 252 patients). (**b**) Boxplots comparing AS in tumors with oncogenic *TP53* mutations (Mut) to wild-type (Wt) tumors. Benjamini–Hochberg adjusted *P* value (*Q*) for Wilcoxon rank sum text. The top and bottom edges represent the first and third quartiles, respectively. The center line represents the median. Whiskers extend to the furthest data points before any outliers (within 1.5 × the interquartile range). (**c**) Boxplots comparing AS in tumors with and without amplifications (amp, SU2C-MARK: ≥ 5 copies above average sample ploidy, MSK-IMPACT: $\log_2$ copy ratio > 0.75). Benjamini–Hochberg adjusted *P* value (*Q*) for Wilcoxon rank sum text. (**d**) Boxplots comparing AS in tumors with and without deletions (del, SU2C-MARK: 0 copies, MSK-IMPACT: $\log_2$ copy ratio < -−1.2). Benjamini–Hochberg adjusted *P*-value (*Q*) for Wilcoxon rank sum text (SU2C-MARK: n = 309 samples, MSK-IMPACT: n = 350 patients).

when AS decreased from 1.0 to 0.5; thereafter, the probability of response remained constant (Fig. 1d, Supplementary Fig. 2c, Supplementary Table 8).

Association of aneuploidy with clinical and genomic variables

Next, we investigated the association between AS and clinicopathologic and genomic variables. AS and TMB were weakly correlated (Spearman rho = 0.23, *P* = 4.4e−5, Supplementary Fig. 3a). Large-cell neuroendocrine (LC-NE) tumors showed the highest median AS (Median AS: LC-NE 0.72 vs. squamous: 0.56 vs. adenocarcinoma: 0.46 vs. other: 0.44; Kruskal–Wallis *P* = 0.016, Supplementary Fig. 3b). AS was not associated with PD-L1 TPS (Supplementary Fig. 3c).

As smoking-associated and non-smoking associated tumors have distinct biological underpinnings, wherein patients with NSCLCs driven by mutations not associated with smoking often exhibit worse response to immunotherapy, we investigated whether aneuploidy may play a role in this phenomenon. Indeed, among patients with lung adenocarcinoma (the histology most frequently driven by oncogenes), tumors with classically smoking-associated drivers had a significantly lower aneuploidy score than those in patients with non-smoking associated drivers (median AS in tumors with smoking-associated driver alterations: 0.39 vs non-smoking associated: 0.58, Wilcoxon *P* = 0.019, Fig. 2a, Supplementary Table 9). Tumors with no identified oncogene driver had an intermediate median AS (0.49). When controlling for *PD-L1* status and TMB in a multivariable logistic regression model, AS was negatively associated with tumors driven by smoking-associated alterations vs non-smoking alterations (Beta = – 2.24, *P* = 0.0064). Similar results were obtained in the MSK-IMPACT cohort (median AS in tumors with smoking-associated driver alterations: 0.37 vs non-smoking associated: 0.49, Wilcoxon *P* = 0.043). In the multivariable model there was also a significant relationship, albeit PD-L1 status was not available for inclusion (Beta = – 1.76, *P* = 0.0082).

We further assessed whether AS was associated with specific oncogenic mutations and focal copy-number alterations. Only *TP53* mutations were associated with elevated aneuploidy score (median AS mutant: 0.56 vs. median wild-type: 0.36, Wilcoxon *Q* = 2.2e−4, Fig. 2b, Supplementary Table 10), consistent with existing literature[1]. This finding was recapitulated in the MSK-IMPACT cohort[10] (Fig. 2b). Focal amplifications (at least 5 gene copies greater than the average tumor ploidy) in *IL7R*, *MEF2D*, *NTRK1*, *RICTOR*, *TERT*, and *TRIP13* were also associated with higher median AS (Wilcoxon *Q* < 0.05, Fig. 2c, Supplementary Table 11); no associations between AS and focal deletions (defined as zero gene copies) were significant at Q < 0.05, but deletions in the co-located *CDKN2A, CDKN2B,* and *MTAP* genes were associated with increased AS (Wilcoxon *Q* = 0.052, Fig. 2d, Supplementary Table 11*)*. The relationships between these focal copy-number alterations and aneuploidy score were also statistically significant in the MSK-IMPACT cohort, apart from *MTAP* deletion (Fig. 2c,d).

Based on previous work[5–7] demonstrating a negative relationship between tumor aneuploidy and immune cell content, we then examined whether AS was associated with immune signaling in the SU2C-MARK cohort. AS negatively correlated with ESTIMATE immune score[17], indicating that highly aneuploid tumors may have reduced total immune cell content (Spearman rho = – 0.43, *P* = 2.3e−4). We also found that the expression of 1254 genes were significantly associated with AS (Spearman *Q* < 0.05), 555 of which were positively correlated and 699 were negatively correlated (Fig. 3a, Supplementary Table 12). Of these 699 genes, 103 were linked to immune signaling[18], including cytokines/chemokines, antimicrobials (such as cell-surface receptor proteins), and genes representing B/T/NK cell function (Fig. 3a). In addition, genes expressed in low AS, but not high AS, NSCLC were enriched for immune-related pathways (odds ratio [95% CI] 10.47 [5.24–23.78], Fisher *P* < 2.2e−16). Both innate and adaptive immune gene expression programs, including IFN gamma, interleukin, TNF alpha, and inflammatory signaling, were significantly downregulated in highly aneuploid tumors (Spearman *Q* < 0.05). In contrast, elevated G2M checkpoint, E2F targets, mitotic spindle, MYC targets, and MTORC1 signaling correlated with higher AS (Spearman *Q* < 0.05; Fig. 3b, Supplementary Table 13). We specifically examined curated immune cell signatures from the SU2C-MARK dataset[15] and found cytotoxic cell, memory T cell, dendritic cell, B cell, and plasma cell signatures were decreased in high AS NSCLC (Spearman *Q* < 0.05; Fig. 3c, Supplementary Table 14).

Discussion

Taken together, across more than 300 metastatic NSCLCs in the SU2C-MARK cohort, we validate that aneuploidy score is an independent, complementary prognostic factor for patients treated with immune checkpoint inhibitors, consistent with our previous findings in the 350 patient MSK-IMPACT cohort[8]. Importantly, aneuploidy score exhibited a dose–response relationship with objective response rate and overall survival. Patients with

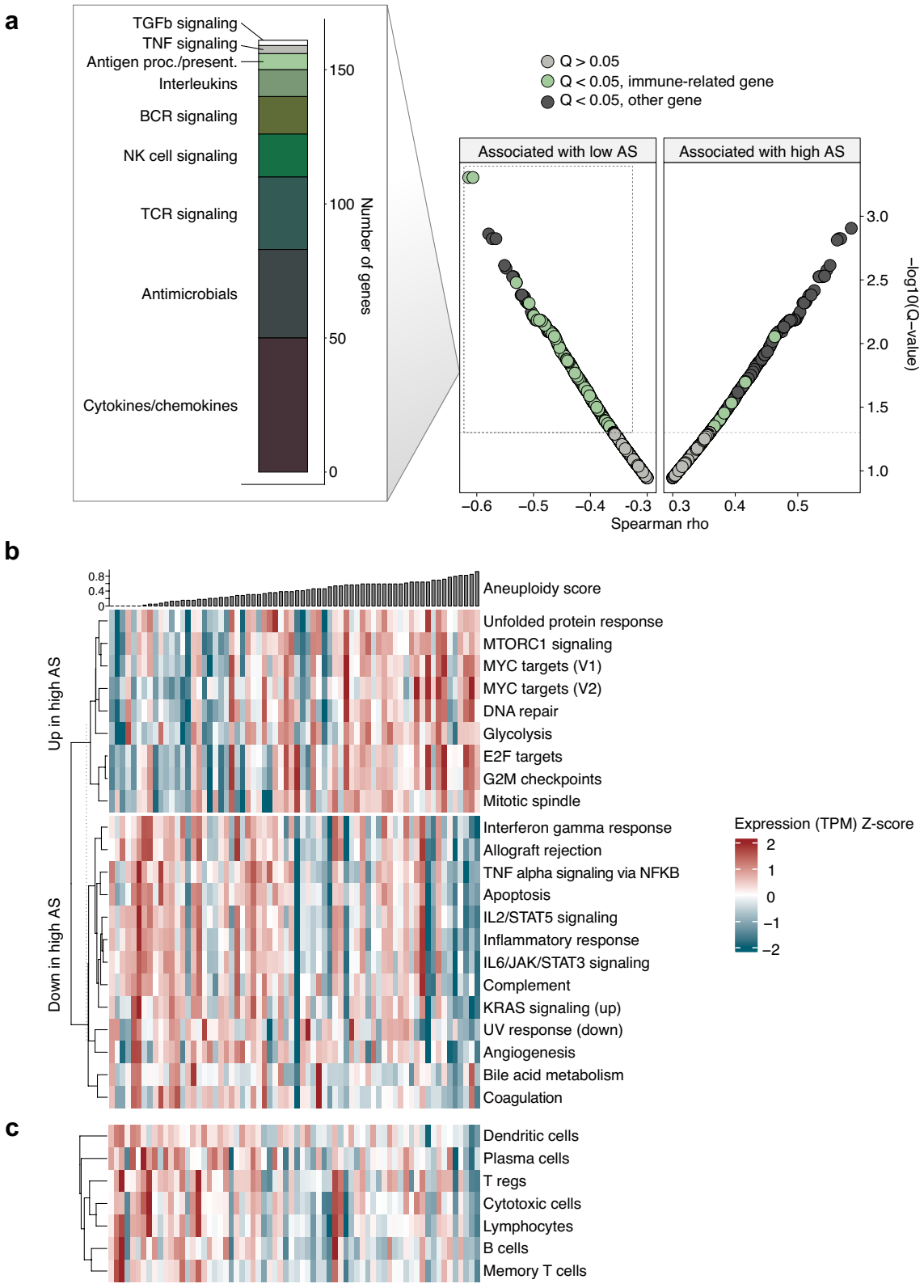


Figure 3. Immune and pathway correlates of aneuploidy in the SU2C-MARK cohort. (**a**) Spearman correlations between expression of individual genes and aneuploidy score (AS). Genes with a correlation coefficient of at least 0.3 in magnitude are plotted on the scatter plot (n = 68 patients). Benjamini–Hochberg adjusted *P*-values (*Q*) for Spearman correlation are displayed. The stacked bar plot illustrates categories of immune-associated genes (as defined by Immport) significantly associated with AS. Genes may be represented in multiple categories (n = 103 unique genes). (Proc: processing, present: presentation). (**b**) Heatmap of ssGSEA pathways significantly correlated with AS at Spearman $Q < 0.05$. Columns represent individual patients (n = 68); (TPM: transcripts per million). (**c**) Heatmap of curated immune cell signatures significantly associated with AS at Spearman $Q < 0.05$ (n = 68 samples).

aneuploidy affecting more than half of the genome were especially unlikely to exhibit an objective response to immunotherapy. However, the deleterious impact of aneuploidy on survival appeared to be attenuated in patients treated with combination anti-PD-(L)1 and anti-CTLA-4 immune checkpoint blockade (and potentially combination chemotherapy and anti-PD-(L)1 agents) as opposed to single agent PD-1/PD-L1 directed therapy.

Whereas there was no significant impact of aneuploidy score on overall survival among high TMB patients, interestingly, we found that there was a numerical difference in progression-free survival between high and low AS among both high and low TMB tumors. One possible explanation is that highly aneuploid tumors are less likely to respond to immunotherapy regardless of their TMB, but that poor responses to immunotherapy of tumors with both high TMB and high aneuploidy tumors can be salvaged with subsequent lines of therapy, such as chemotherapy and/or radiotherapy[7,14], whereas this approach is potentially less successful in highly aneuploid, low TMB tumors.

We found that individual aSCNAs were not significantly associated with overall survival. While this result could be cohort-specific, as a recent study showed a relationship between *9p* loss and survival in the setting of ICI treatment[6], our data suggests that aneuploidy elicits immune suppression by downregulating several critical immune signaling pathways[5,14]. We posit that at a certain level of aneuploidy, the probability of disrupting key anti-tumor immune pathways becomes sufficient such that tumors are highly likely to acquire sufficient gene dosage imbalances[3] and exhibit a resistant phenotype. This phenomenon may be particularly deleterious to patient survival in cases where few neoantigens are present to stimulate anti-tumor immunity, such as those with low TMB.

Expression signatures of innate immune pathways, such as dendritic cell activation and various cell surface receptors were inversely correlated with aneuploidy score. Our analysis corroborates prior findings that aneuploidy is linked to decreased cytotoxic (CD8 + and natural killer) cell infiltration[5,7]. The results additionally suggest that aneuploidy is associated with impaired dendritic cell recruitment, which has been previously described to result in decreased T cell priming[19] and thereby act as an ICI resistance mechanism in melanoma[20]. In addition, the strong correlation of aneuploidy with cell cycle and proliferation pathways is in line with previous studies demonstrating that highly aneuploid tumors exhibit an aggressive phenotype[3,4].

In both the SU2C-MARK and MSK-IMPACT cohorts, adenocarcinomas driven by oncogenes that predominantly occur in never-smokers demonstrated an elevated degree of aneuploidy. This patient population is known to often exhibit a poor response to immunotherapy, which has partially been attributed to a lower TMB[11,21]. Given the observed synergistic predictive value of TMB and aneuploidy, it is possible that aneuploidy also plays a role in mediating this immunotherapy resistant phenotype.

In two independent cohorts, we found that tumors with *TP53* mutations exhibited a higher aneuploidy score than their wild-type counterparts, indicating mutant p53 may promote tolerance of aneuploidy[1]. We also confirmed that a high aneuploidy score was associated with focal deletions in *CDKN2A/B,* which have previously been linked to poor response to immunotherapy[22]. Amplifications of *TERT*, often of 10–20 copies, were also highly associated with increased aneuploidy score, and *TERT* amplifications have been previously linked to poor prognosis in NSCLC[23,24]. Specifically, *TERT* amplifications only occurred in tumors above the 33rd percentile of AS and 19/22 of the amplifications in our cohort were in highly aneuploid tumors in the SU2C-MARK cohort (greater than or equal to the median aneuploidy score). While increasing levels of aneuploidy often confer a survival advantage for tumor cells, they also induce replication stress and premature senescence. *TERT* expression has previously been shown to bypass this phenomenon and allow continued cell division[24]. Perhaps the mechanism of adverse prognosis associated with *TERT* amplification could be mediated partially by telomerase-permitted tumor aneuploidy. If validated in future studies, dependence on *TERT* overexpression could represent a common therapeutic vulnerability among highly aneuploid tumors, which could be exploited with adjunctive treatment with emerging small molecule inhibitors[25] to improve survival following immunotherapy.

Collectively, our findings provide important supporting evidence for the role of high tumor aneuploidy in predicting reduced response to, and overall survival following, treatment with immune checkpoint inhibitors. These data confirm previous studies from our group and others[5–8,12,26–28] demonstrating the deleterious effects of aneuploidy (and related measures of copy-number alteration burden, such as the fraction of genome altered) on patient outcomes following immunotherapy. Importantly, we build upon this foundation to elucidate the genomic and immunologic correlates of tumor aneuploidy, including key associations with *TERT* amplification and non-smoking associated tumors. We propose these findings will contribute to emerging translational evidence indicating the potential of utilizing tumor aneuploidy to guide treatment selection for patients with NSCLC.

Methods

Data sources

All clinicopathologic, genomic, and transcriptomic data for the SU2C-MARK cohort[15] were obtained from the repository associated with the original manuscript[15]: https://zenodo.org/records/7849582, including clinical outcomes, patient-level clinical information, tumor mutational burden (TMB), mutation calls, normalized gene expression data, segmented copy-number data, and ABSOLUTE purity and ploidy estimations. Further details on the generation of this data can be found in the original manuscript.

Clinical outcomes and mutation data for the MSK-IMPACT[10] cohort were downloaded from the associated cBioPortal repository, and segmented copy-number data were downloaded from AACR Project GENIE v15.0 (see *Data Availability* statement below).

All methods were carried out in accordance with relevant guidelines and regulations. No institutional approval was required as all patient data in this publication was publicly available.

Gene level mutation and copy-number analysis

For the SU2C-MARK cohort, mutation calls were obtained from the manuscript's Zenodo repository as above. Only mutations categorized as "Oncogenic" or "Likely Oncogenic" per the OncoKB annotation were considered for analysis. For the MSK-IMPACT cohort, the OncoKB-annotated list of driver mutations as presented in the study's cBioPortal repository were included. Oncogenic drivers were classified as smoking- or non-smoking-associated based on previous work[21,29–33].

Gene-level copy-number calls were generated from the segmented copy-number data from the Zenodo repository for the SU2C-MARK cohort and from AACR Project GENIE v14.0 for the MSK-IMPACT cohort. Segments overlapping the coordinates of the gene of interest (obtained from ensemble Biomart) were identified and the genes were assigned the copy-number ("corrected_total_cn" for the SU2C-MARK cohort and the $\log_2$ copy ratio for the MSK-IMPACT cohort as integer copy number calls were unavailable) of their overlapping segment. In cases where multiple segments overlapped a gene, the length-weighted average of the segment copy-number was taken. Genes were considered amplified if their copy-number was at least five copies above the average sample ploidy rounded to the nearest whole number or their $\log_2$ copy ratio was > 0.75[34]; they were considered deleted if their copy-number was zero or their $\log_2$ copy ratio was < -1.2. Only genes with oncogenic mutations present in at least 5 out of the 309 patients with exome data were assessed for correlations with aneuploidy score. An analogous approach was taken for focal amplifications and deletions: only genes identified as known tumor suppressor (deletions) or oncogenes (amplifications) per OncoKB annotation which were altered in at least five patients were further analyzed. Only genes that were identified as having a statistically significant relationship with aneuploidy in the SU2C-MARK cohort were assessed in the MSK-IMPACT cohort.

Arm-level copy number and aneuploidy score quantification

Arm-level somatic copy number alterations (aSCNAs) were called using ASCETS[16] on segmented copy-number data (log2 copy ratio threshold = ± 0.1; arm alteration fraction threshold = 0.5; min breadth of coverage [BOC] = 0.5) as per optimization below. Aneuploidy scores were calculated for each sample by computing the fraction of evaluable arms (ASCETS call ≠ LOWCOV) affected by aSCNAs (ASCETS call = AMP or DEL).

ASCETS parameter optimization

Bootstrap resampling was used to optimize the $\log_2$ copy ratio amplitude and alteration fraction thresholds. From the set of 306 samples in the SU2C-MARK cohort with exome and overall survival data, 10,000 sets of 306 samples were randomly selected with replacement. On each bootstrap sample, Cox proportional hazards models with aneuploidy score (AS) and tumor mutational burden ($\log_{10}$ scale) as independent variables were constructed at multiple amplitude thresholds (0.1–0.5 in increments of 0.1) and alteration fraction thresholds (0.5–0.9 in increments of 0.1). Multivariable Wald *P* values and hazard ratios for overall survival were aggregated across the 10,000 bootstrap samples (Supplementary Fig. 1a-b). The pairing of amplitude and alteration thresholds which resulted in the smallest mean *P* value across the bootstrap samples was selected as the optimal threshold.

Based on the optimization results, regions with $\log_2$ copy ratios of magnitude ≥ 0.1 are considered altered (the same as above). Orthogonally, copy ratio threshold of 0.1 appears appropriate as it estimates the full-width at half maximum (FWHM) of the central peak of aggregated copy-number segment means across the cohort, as is recommended per the ASCETS manuscript[16] (Supplementary Fig. 1c).

To constitute an aSCNA, the sum of all alterations must cover at least 50% (alteration fraction threshold: 0.5) of the genomic territory on a chromosome arm. Although 50% of an arm may appear to be a lenient threshold to constitute an arm-level event, it is likely that focal copy-number gains superimposed on arm-level losses and vice versa, as well as subclonal aSCNAs, confound more stringent aSCNA calling thresholds, and thus a more lenient threshold is appropriate in this setting. Thresholds requiring up to 90% of the arm to be altered were significantly associated with survival, however 50% showed the best performance with the greatest mean hazard ratio and lowest mean *P*-value across bootstraps (Supplementary Tables 1–2). In addition, aSCNA calls using these parameters identified chromosome arm alteration frequencies consistent with previous studies[1], further supporting the notion that these parameters are appropriate for calling aSCNAs (Supplementary Table 3).

Hallmark pathways

Single-sample GSEA was performed TPM RNA-seq data using the *gsva* package in R on the set of Hallmark gene sets (h.all.v7.4.symbols.gmt).

Immune analysis

The Estimation of Stromal and Immune cells in Malignant Tumours using Expression data (ESTIMATE) algorithm[35] was implemented using the R *tidyestimate* package. Curated cell type signatures were obtained from the Supplementary Tables of the original manuscript[15]. Genes were identified as "immune genes" and were categorized using the *Immport* database[18].

Statistical analysis

All analyses were performed using R version 4.3.1. Comparisons between a discrete independent variable and continuous independent variable were analyzed using Kruskal–Wallis and Wilcoxon tests. Correlations between continuous variables were assessed using a Spearman correlation. *P* values were corrected for multiple comparisons using the Benjamini–Hochberg method for false discovery rate; these corrected *P* values are denoted as *Q*-values. All tests were two-tailed and *P* or $Q < 0.05$ was considered significant.

LASSO Cox proportional hazards regression was performed using the *glmnet* R package. The following variables were included in the model as potential predictors: TMB, AS, pack-years of smoking, stage at diagnosis, patient sex, patient race, patient age, line of therapy, immunotherapy treatment paradigm, PD-L1 tumor proportion score, prior treatment with platinum agents, prior treatment with tyrosine kinase inhibitors, and tumor histology. Missing value imputation was performed as implemented in the *glmnet* package; 25 patients were missing race annotations, 20 were missing staging annotations, and 116 were missing PD-L1 annotations. Leave-one-out cross validation was performed using the cv.glmnet function, and the minimal lambda value was selected as the optimal value for the final model.

Per the original manuscript[15], overall survival was defined as the interval between date of treatment start with a PD(L)1 agent and death (event) or last follow-up (censor); response data were assessed using RECIST v1.1.

Kaplan–Meier curves were generated using the *survminer* package in R; a log-rank test was used to compare survival between groups. Hazard ratios and confidence intervals were calculated using a Cox proportional hazards regression model.

Data availability

All original data for the SU2C-MARK Foundation cohort[15] are available from the original manuscript's Zenodo repository: https://zenodo.org/records/7849582. All data for the MSK-IMPACT cohort[10] is deposited at the following cBioPortal repository: https://www.cbioportal.org/study/summary?id=tmb_mskcc_2018 and AACR Project GENIE: https://www.synapse.org/#!Synapse:syn7222066/wiki/405659. All modified and/or newly generated data are available at this manuscript's associated Zenodo repository: https://zenodo.org/records/11671610[36].

Code availability

All proprietary code necessary to replicate the results and figures in this manuscript are available at this manuscript's associated Zenodo repository: https://zenodo.org/records/11671610[36].

Received: 7 April 2024; Accepted: 26 June 2024
Published online: 21 August 2024

References

1. Taylor, A. M. *et al.* Genomic and functional approaches to understanding cancer aneuploidy. *Cancer Cell* **33**, 676-689.e3 (2018).
2. Shih, J. *et al.* Cancer aneuploidies are shaped primarily by effects on tumour fitness. *Nature* **619**, 793–800 (2023).
3. Vasudevan, A. *et al.* Aneuploidy as a promoter and suppressor of malignant growth. *Nat. Rev. Cancer* **21**, 89–103 (2021).
4. Ben-David, U. & Amon, A. Context is everything: aneuploidy in cancer. *Nat. Rev. Genet.* **21**, 44–62 (2020).
5. Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017).
6. Alessi, J. V. *et al.* Impact of aneuploidy and chromosome 9p loss on tumor immune microenvironment and immune checkpoint inhibitor efficacy in NSCLC. *J. Thorac. Oncol.* https://doi.org/10.1016/j.jtho.2023.05.019 (2023).
7. Spurr, L. F. *et al.* Highly aneuploid non-small cell lung cancer shows enhanced responsiveness to concurrent radiation and immune checkpoint blockade. *Nat. Cancer* **3**, 1498–1512 (2022).
8. Spurr, L. F., Weichselbaum, R. R. & Pitroda, S. P. Tumor aneuploidy predicts survival following immunotherapy across multiple cancers. *Nat. Genet.* **54**, 1782–1785 (2022).
9. Liu, L. *et al.* Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin. Cancer Res.* **25**, 7413–7423 (2019).
10. Samstein, R. M. *et al.* Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* **51**, 202–206 (2019).
11. Popat, S. *et al.* Association between smoking history and overall survival in patients receiving pembrolizumab for first-line treatment of advanced non-small cell lung cancer. *JAMA Netw. Open* **5**, e2214046–e2214046 (2022).
12. Chowell, D. *et al.* Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat. Biotechnol.* **40**, 499–506 (2022).
13. Girish, V. *et al.* Oncogene-like addiction to aneuploidy in human cancers. *Science* **381**, eadg4521 (2023).
14. Spurr, L. F. & Pitroda, S. P. Exploiting tumor aneuploidy as a biomarker and therapeutic target in patients treated with immune checkpoint blockade. *Npj Precis. Oncol.* **8**, 1 (2024).
15. Ravi, A. *et al.* Genomic and transcriptomic analysis of checkpoint blockade response in advanced non-small cell lung cancer. *Nat. Genet.* **55**, 807–819 (2023).
16. Spurr, L. F. *et al.* Quantification of aneuploidy in targeted sequencing data using ASCETS. *Bioinformatics* https://doi.org/10.1093/bioinformatics/btaa980 (2020).
17. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612–2612 (2013).
18. Bhattacharya, S. *et al.* ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci. Data* **5**, 180015 (2018).
19. Fares, C. M., Van Allen, E. M., Drake, C. G., Allison, J. P. & Hu-Lieskovan, S. Mechanisms of resistance to immune checkpoint blockade: why does checkpoint inhibitor immunotherapy not work for all patients?. *Am. Soc. Clin. Oncol. Educ. Book* **39**, 147–164. https://doi.org/10.1200/EDBK_240837 (2019).
20. Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235 (2015).
21. Wang, X. *et al.* Association between smoking history and tumor mutation burden in advanced non-small cell lung cancer. *Cancer Res.* **81**, 2566–2573 (2021).
22. Gutiontov, S. I. *et al.* CDKN2A loss-of-function predicts immunotherapy resistance in non-small cell lung cancer. *Sci. Rep.* **11**, 20059 (2021).
23. Zhu, C.-Q. *et al.* Amplification of telomerase (hTERT) gene is a poor prognostic marker in non-small-cell lung cancer. *Br. J. Cancer* **94**, 1452–1459 (2006).
24. Meena, J. K. *et al.* Telomerase abrogates aneuploidy-induced telomere replication stress, senescence and cell depletion. *EMBO J.* **34**, 1371–1384 (2015).
25. Liu, M. *et al.* The regulations of telomerase reverse transcriptase (TERT) in cancer. *Cell Death Dis.* **15**, 90 (2024).

26. Chang, T.-G. *et al.* Optimizing cancer immunotherapy response prediction by tumor aneuploidy score and fraction of copy number alterations. *Npj Precis. Oncol.* **7**, 54 (2023).
27. Jia, Q. *et al.* Mutational burden and chromosomal aneuploidy synergistically predict survival from radiotherapy in non-small cell lung cancer. *Commun. Biol.* **4**, 131 (2021).
28. Lu, Z. *et al.* Tumor copy-number alterations predict response to immune-checkpoint-blockade in gastrointestinal cancer. *J. Immunother. Cancer* **8**, e000374 (2020).
29. Zhu, Q., Zhan, P., Zhang, X., Lv, T. & Song, Y. Clinicopathologic characteristics of patients with ROS1 fusion gene in non-small cell lung cancer: a meta-analysis. *Transl. Lung Cancer Res.* **4**, 300–309 (2015).
30. Wong, S. K. *et al.* MET exon 14 skipping mutation positive non-small cell lung cancer: response to systemic therapy. *Lung Cancer* **154**, 142–145 (2021).
31. Shaw, A. T. & Engelman, J. A. ALK in lung cancer: past, present, and future. *J. Clin. Oncol.* **31**, 1105–1111 (2013).
32. Zhou, C., Li, W., Shao, J., Zhao, J. & Chen, C. Analysis of the clinicopathologic characteristics of lung adenocarcinoma With CTNNB1 mutation. *Front. Genet.* **10**, 1367 (2020).
33. Servetto, A. *et al.* RET rearrangements in non-small cell lung cancer: evolving treatment landscape and future challenges. *Biochim. Biophys. Acta. BBA - Rev Cancer* **1877**, 188810 (2022).
34. Mina, M., Iyer, A., Tavernari, D., Raynaud, F. & Ciriello, G. Discovering functional evolutionary dependencies in human cancers. *Nat. Genet.* **52**, 1198–1207 (2020).
35. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218–218 (2016).
36. Spurr, L. & Pitroda, S. Clinical and molecular correlates of tumor aneuploidy in metastatic non-small cell lung cancer. Zenodo https://doi.org/10.5281/zenodo.11671610 (2024).

Acknowledgements

This work was supported by the Ludwig Cancer Research Foundation (S.P.P.), a Career Development Award from the LUNGevity Foundation (S.P.P.), an Ullman Scholarship in Translational Cancer Immunology from the University of Chicago Comprehensive Cancer Center (UCCCC) (S.P.P.), a Cancer Spotlight Grant (S.P.P.) from the UCCCC, a Lung Cancer Discovery Award from the American Lung Association (S.P.P), a Catalyst Award from the Falk Medical Research Trust (S.P.P.), and the John D. Arnold, MD Scientific Research Prize from the University of Chicago Pritzker School of Medicine (L.F.S.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

Both authors contributed to the conceptualization, investigation, visualization, writing and editing of this work.

Competing interests

L.F.S. and S.P.P. have filed patents related to the use of tumor aneuploidy as a predictor of response to radiotherapy and immunotherapy. S.P.P. holds other patents outside the submitted work.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-66062-5.

Correspondence and requests for materials should be addressed to S.P.P.