

RESEARCH ARTICLE

Variability in protein cargo detection in technical and biological replicates of exosome-enriched extracellular vesicles

Suma Tiruvayipati¹, Don Wolfgeher², Ming Yue¹, FangFang Duan¹, Jorge Andrade³, Hui Jiang⁴, Lucia Schuger^{1*}

1 Biological Sciences Division, Department of Pathology, The University of Chicago, Chicago, Illinois, United States of America, **2** Proteomics Core Laboratory, Cummings Life Science Center, The University of Chicago, Chicago, Illinois, United States of America, **3** The Center for Research Informatics, The University of Chicago, Chicago, Illinois, United States of America, **4** Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, United States of America

* lschuger@bsd.uchicago.edu



OPEN ACCESS

Citation: Tiruvayipati S, Wolfgeher D, Yue M, Duan F, Andrade J, Jiang H, et al. (2020) Variability in protein cargo detection in technical and biological replicates of exosome-enriched extracellular vesicles. *PLoS ONE* 15(3): e0228871. <https://doi.org/10.1371/journal.pone.0228871>

Editor: Aleksandra Nita-Lazar, NIH, UNITED STATES

Received: May 7, 2019

Accepted: January 24, 2020

Published: March 2, 2020

Copyright: © 2020 Tiruvayipati et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All mass spectrometry proteomics data files are available from the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012798 and [10.6019/PXD012798](https://doi.org/10.6019/PXD012798)

Funding: This study was supported by the National Institutes of Health (<https://www.nih.gov/>) under the Grant R01 HL132870 (L.S.) and Grant R01 HL128228 (L.S.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Exosomes are extracellular vesicles (EVs) of ~20–200 nm diameter that shuttle DNAs, RNAs, proteins and other biomolecules between cells. The large number of biomolecules present in exosomes demands the frequent use of high-throughput analysis. This, in turn, requires technical replicates (TRs), and biological replicates (BRs) to produce accurate results. As the number and abundance of identified biomolecules varies between replicates (Rs), establishing the replicate variability predicted for the event under study is essential in determining the number of Rs required. Although there have been few reports of replicate variability in high throughput biological data, none of them focused on exosomes. Herein, we determined the replicate variability in protein profiles found in exosomes released from 3 lung adenocarcinoma cell lines, H1993, A549 and H1975. Since exosome isolates are invariably contaminated by a small percentage of ~200–300 nm microvesicles, we refer to our samples as exosome-enriched EVs (EE-EVs). We generated BRs of EE-EVs from each cell line, and divided each group into 3 TRs. All Rs were analyzed by liquid chromatography/mass spectrometry (LC/MS/MS) and customized bioinformatics and biostatistical workflows (raw data available via ProteomeXchange: PXD012798). We found that the variability among TRs as well as BRs, was largely qualitative (protein present or absent) and higher among BRs. By contrast, the quantitative (protein abundance) variability was low, save for the H1975 cell line where the quantitative variability was significant. Importantly, our replicate strategy identified 90% of the most abundant proteins, thereby establishing the utility of our approach.

Introduction

Extracellular vesicles (EVs) include microvesicles and exosomes[1–3]. Exosomes are bilayered membrane-bound nanovesicles [4] of endocytic origin, ranging from ~20 to 200 nm[3, 5, 6],

Competing interests: The authors have declared that no competing interests exist.

emerging during the formation of multivesicular bodies and secreted into the extracellular space as a result of the fusion of multivesicular bodies with the plasma membrane [7, 8]. Exosomes have specific surface proteins such as flotillin-1 and tetraspanin family proteins CD9, CD81, and CD63 [7, 9]. In addition, membrane proteins such as annexins and components of the ESCRT complex, and integrins were also characterized in these EVs [10]. Exosomes are constantly released by all types of cells, normal or diseased, and are present in all body fluids [7]. They contain DNA, RNAs, proteins, and lipids [11, 12] and provide local and distal biological signals to tissues via endocytic transfer of their contents [11].

Exosomes participate in multiple normal biological processes [11] and play a significant role in a myriad of pathological conditions such as cancer progression, autoimmune and infectious diseases, obesity, and neurodegenerative diseases [11]. Since exosomes reflect the phenotype of their donor cells [2, 13–15], they became an important part of the newest repertoire of what is referred as the “tumor circulome” in liquid biopsies with a promising potential in cancer management [16]. Moreover, exosomes are being studied as agents for gene therapy, vaccines, and drug delivery [2, 4].

The Exocarta (<http://www.exocarta.org/>) database, inception in 2009, compiles the RNAs and proteins from a wide range of exosomal data [17]. Following which, EVpedia (<http://evpedia.info>) an integrated proteome, transcriptome, and lipidome database has led to considerable improvement in EVs research [18].

It remains worthy to point to the existence of technical caveats in exosome research, as currently all the conventional methods of exosome isolation pervasively retain a small percentage of ~200–300 nm microvesicles [6, 19, 20]. Therefore, research is now being conducted to obtain pure types of EVs [3]. Due to this well-established factor, we refer to the exosome isolates obtained by current techniques as exosome enriched EVs (EE-EVs).

Due to the large biomolecular cargo carried by these EVs, exosomal research often relies on the generation of high-throughput data. A proper interpretation of data generated by high-throughput analysis requires the use of replicate samples (Rs). These include technical replicates (TRs) and biological replicates (BRs). TRs help to understand the reproducibility of an assay, whereas BRs inform about the reproducibility of the phenomenon [21]. Therefore, both have to be included in the design of any experiment in order to reach accurate conclusions.

As variability in the number and abundance of identified biomolecules (here respectively referred to as qualitative and quantitative variability) is always encountered among Rs, it is essential to know the variability in sampling expected for the specific phenomenon under study in order to determine the number of Rs adequate to generate accurate results.

Although there have been few reports of Rs variability in high throughput biological data [22–24], none of them focused on exosomes. Therefore, here we used custom bioinformatics and biostatistics workflows of LC/MS/MS data to determine the qualitative and quantitative variability in proteins in TRs and BRs of EE-EVs from lung adenocarcinoma cell lines.

Materials and methods

Cell culture

The cells used in the current study were human lung adenocarcinoma cell lines: H1993 (ATCC CRL-5909), H1975 (ATCC CRL-5908) and A549 (ATCC CCL-185). H1975 and H1993 cells were cultured in RPMI media (Gibco, 11875119) while, A549 cells were cultured in Ham's F-12K media (Corning, 10-025-CV) with 100 Units/ml penicillin– 100 µg/ml streptomycin (Gibco, 15140122) and 0.25 µg/ml amphotericin (Gibco, 15290026). Media was supplemented with 10% fetal bovine serum (FBS) (Atlanta biologicals, S11150). Cells were grown to 85–95% confluency in 8 150 mm cell culture dishes (NuncTM, 157150), trypsinized (using

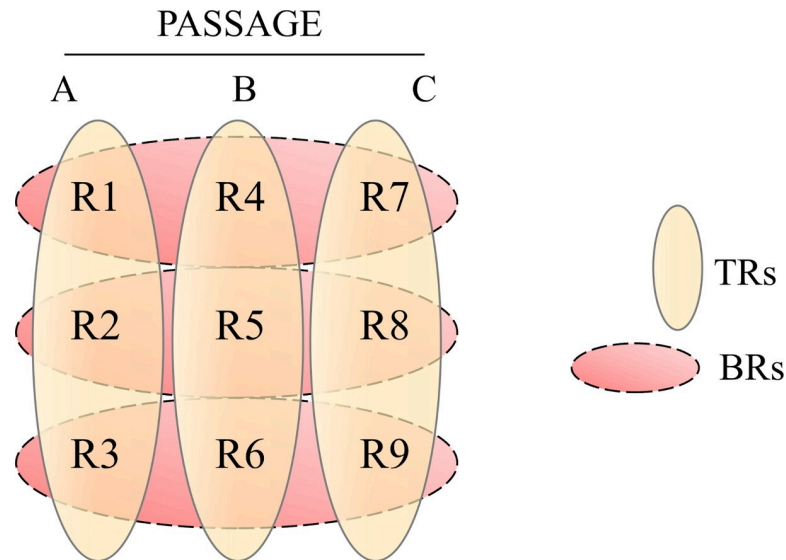


Fig 1. Schematic representation of exosomal samples collection. Exosomes were collected from three different cell lines at three different passages (A, B, C) to provide three biological replicates (BRs). The exosome lysate from each of them was further divided into three technical replicates (TRs). The nine samples were grouped so that the TRs were TR1: R1, R2, R3, TR2: R4, R5, R6, TR3: R7, R8, and R9. The BRs were BR1: R1, R4, R7, B2: R2, R5, R8, BR3: R3, R6, and R9.

<https://doi.org/10.1371/journal.pone.0228871.g001>

0.25% trypsin EDTA (Gibco, 25200114)), counted with a hemocytometer and seeded in 3 corning 224 mm cell culture dishes with 10% exosome depleted FBS (SBI, EXO-FBS-250A-1). Cells were incubated at 37°C and 5% CO₂ and after 24 hours, the medium was collected for exosome purification. Triplicate exosome enriched extracellular vesicles (EE-EVs) samples from each of the 3 lung adenocarcinoma cell lines H1993, A549 and H1975 generated at different passages were used to generate 3 technical replicates (TRs) and 3 biological replicates (BRs) (Fig 1). In total, 9 replicate samples (Rs) per cell line that is, 27 Rs in total were analyzed for this study. The 9 samples were grouped so that the TRs were TR1: R1, R2, R3, TR2: R4, R5, R6, TR3: R7, R8, and R9. The BRs were BR1: R1, R4, R7, BR2: R2, R5, R8, BR3: R3, R6, and R9.

Exosome enriched extracellular vesicles (EE-EVs) purification

Cell debris from the media was removed by centrifugation at 300 x g for 5 minutes followed by 3,000 x g for 45 minutes. Care was taken not to touch the pellets. As an additional quality control measure, the portion of the media touching the pellet was discarded to avoid contamination. The supernatant was transferred to 50 ml centrifuge tubes and mixed with exosome precipitation solution exoquick-TC (SBI, EXOTC50A-1) [25–27] at a 5:1 ratio and incubated for 16 hours at 4°C to precipitate exosomes followed by centrifugation at 1,500 x g for 60 minutes to collect the exosome pellets. The supernatant was discarded. The exosomes were then lysed using RIPA buffer (Thermo Fisher Scientific, PI89900) with protease and phosphatase inhibitors (Pierce, A32961) and the protein concentration was measured by Pierce BCA protein assay kit (Thermo Fisher Scientific, PI23227). Western blots were performed with cells and EE-EVs lysed in RIPA lysis buffer (Sigma-Aldrich) containing proteinase inhibitors (Roche). Lysates were cleared by centrifugation for 10 minutes at 12,000 rpm and supernatant fluids were collected. Immunoblots were performed as previously described [28]. The following antibodies (Abs) were used for immunoblotting: anti-CD 81, anti-CD 63, anti-CD 9 (all from System Biosciences), anti-Flotillin 1, anti-TSG 101, and anti-Calnexin (all from Abcam).

All Abs were used at the dilutions recommended by the manufacturers. Further, the diameter of the isolated vesicles was determined using the qNano-Tunable Resistive Pulse Sensing (TRPS) at Izon Science, USA. EE-EVs were analyzed by nanopore NP150, which has a pore size of 150 nm at 5 mbar pressure. Data acquisition and analysis were performed using the Izon Control Suite software version 3.3.2.2001. The reason for electing this method is because TRPS has the advantage of detecting EVs at a higher precision than other methodologies [29–31].

Proteomic analysis

A 100 µg of total protein of exosomes was heated with 4x SDS loading buffer (Thermo Fisher Scientific, NP0007) for 10 minutes at 70°C and loaded on 4–12% Bis-tris protein gels (Thermo Fisher Scientific, NP0335). Gels were then prefixed in 1:2:1 methanol, acetic acid, and water overnight followed by staining with Brilliant blue G solution (Sigma, B8522) for 2 hours and further destained in 10% acetic acid for 4 hours. The protein bands were then excised, placed in 1.5 ml individual centrifuge tubes with 100 µl 5% acetic acid and sent for liquid chromatography-mass spectrometry (LC/MS/MS) performed at the Michigan State University (MSU) Proteomics Core Facility. The experimental protocol was as follows: Gel bands were digested in-gel according to a previously reported study [32]. Briefly, the gel bands were dehydrated using 100% acetonitrile (ACN) and incubated with 10 mM dithiothreitol in 100 mM ammonium bicarbonate, pH~8, at 56°C for 45 minutes. The gel bands were dehydrated again in 100% ACN to force out all aqueous buffers and allow the addition of 50 mM iodoacetamide in 100 mM ammonium bicarbonate to equilibrate all the protein and incubated in the dark for 20 minutes. The gel bands were then washed with ammonium bicarbonate and dehydrated again in 100% ACN followed by an overnight incubation at 37°C with sequencing grade modified trypsin (Promega, V5111) prepared in 50 mM ammonium bicarbonate and added at ~1:50 ratio. Peptides were then extracted from the gel by water bath sonication in a solution of 60% ACN and 1% trifluoroacetic acid (TFA) and vacuum dried to ~2 µL. Dried peptides were then re-suspended in 2% ACN/0.1% TFA to 25 µL. From this, 5 µL was automatically injected by a Thermo EASYnLC 1000 liquid chromatography system onto a Thermo Acclaim PepMap 100 C18 trapping column (0.1 mm x 20 mm, 5 µm, 100A) and washed with buffer A (99.9% water/0.1% formic acid) for ~5 minutes. Bound peptides were then eluted onto a Thermo Acclaim PepMap RSLC C18 resolving column (0.075 mm x 500 mm, 3 µm, 100A) for over 125 minutes with a gradient of 5% buffer B to 28% buffer B (99.9% ACN/0.1% formic acid) for 114 minutes, ramping to 90% buffer B at 115 minutes and held at 90% buffer B for the duration of the run at a constant flow rate of 300 nl/min.

Eluted peptides were sprayed into a Thermo Fisher Q-Exactive mass spectrometer using a FlexSpray nano-spray ion source. Survey scans were taken by the ion trap, a second mass analyzer of the mass spectrometer i.e. the Orbitrap (70,000 resolutions, determined at m/z 200). In each survey scan, the top ten most intense peptide ions were automatically selected and subjected to higher energy collision induced dissociation with fragment spectra acquired at 17,500 resolutions. Conversion of MS/MS spectra to peak lists was done using Mascot Distiller version 2.6.1 (www.matrixscience.com). Peptide-to-spectrum matching was done using the Mascot search algorithm version 2.6, against a database containing all human protein sequences available from UniProt (www.uniprot.org, downloaded on 11-13-2017) and appended with common laboratory contaminants (www.thegpm.org). The search output was then analyzed using Scaffold Q+S version 4.8.4 (www.proteomesoftware.com) to probabilistically validate protein identification and quantification. Assignments validated using the default confidence filter of 1% False Discovery Rate (FDR) at the protein level in order to allow maximum discovery at reasonable stringency were considered true.

Mascot parameters for all databases were as follows: allow up to 2 missed tryptic sites, fixed modification of carbamidomethyl cysteine, variable modification of oxidation of methionine, peptide tolerance of +/- 10 ppm, MS/MS tolerance of 0.3 Da, peptide charge state limited to +2/+3.

The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE [33] partner repository with the dataset identifier PXD012798 and 10.6019/PXD012798.

Bioinformatics analysis

Protein identification. Label-free quantitative (LFQ) intensity values were generated with the tool MaxQuant (version 1.6.0.0) [34] (www.biochem.mpg.de/5111795/maxquant) using “.raw” files provided by the MSU proteomics facility and searching against a Uniprot human database (downloaded on 2/1/2018). The parameters in MaxQuant were set as follows: oxidation of methionine and protein N-terminal acetylation were allowed as variable modifications, and cysteine carbamidomethyl was set as a fixed modification. The option for proteases was chosen as trypsin/P (proline) [35] which marks the cleavage at the carboxyl side of the lysine and arginine amino acids with 2 missed cleavages allowed. The parameter for label-free modification to check for protein presence was selected as LFQ values. The FDR with a *p*-value less than 0.01 was determined as significant. For protein quantification firstly, label minimum ratio count was set to 2, secondly, both unique and razor peptides were selected, and thirdly, the modifications were once again set to oxidation of methionine and protein N-terminal acetylation, along with their unmodified peptides. The obtained proteingroups.xlsx output file was sorted by descending LFQ values and used as an LFQ value reference list for further bioinformatics analysis in addition to which the LFQ values were processed with custom shell scripts for further biostatistical analysis of qualitative data.

Protein quantification. Absolute protein expression (APEX) values were generated using “protXML” files, a required file format by the APEX tool [36, 37]. For this, “.sf3” files provided by the MSU proteomics facility were processed to generate “protXML” files using the Scaffold 4.8.4 [38] (www.proteomesoftware.com) software, set with a cut off corresponding to a peptide and protein FDR corrected *p*-value of less than 0.01. The peptide FDR was calculated by the tool as a percentage of the sum of the exclusive spectral counts of decoy proteins divided by the sum of exclusive spectral counts of target proteins. The protein FDR was calculated as a percentage of the number of decoy proteins divided by the number of target proteins. The “protXML” files and the Uniprot human database (downloaded on 2/1/2018) was used to calculate the APEX values by following the apex protocol [36, 37] using the tool APEX_1_1_0 (<https://sourceforge.net/projects/apexqpt/>). The top 50 proteins from the LFQ intensity list from MaxQuant were considered to build a reference list for use with the APEX tool to generate an “.apex” file. The protein’s abundance was usually presented as *relative* to all protein within the sample, here the multiplicative normalization factor *C*, which multiplies the protein’s abundance by *C*, places the abundance values into *absolute* terms where *C* corresponds to “1.0E8”. The “.apex” file was further processed with custom shell scripts to proceed with biostatistical analysis. The absolute counts obtained in an “.apex” file are directly proportional to the protein levels, and used for biostatistical analysis of quantitative data.

The LC/MS/MS spectra database matching identifies peptides, and not proteins [39]. Hence, the protein list reported by LFQ values is only tentative, as several peptides can be assigned to more than 1 protein [39]. For an absolute quantitative count, the APEX proteomics tool [37] was used, which calculates abundances of protein expression based upon machine learning correction factors, LC/MS/MS spectral counts, and correct identification of protein

probability. Hence, the protein list reported by APEX values is more reliable in view of identification of a complete protein sequence.

Biostatistical analysis

Statistical analysis was performed by using the LFQ (protein identification) and APEX (protein abundance) values of LC/MS/MS data. The LFQ values were used to perform statistical analysis to show the qualitative variability while, the APEX values were used to perform statistical analysis to show the quantitative variability. The replicates were grouped into technical replicates (TRs) and biological replicates (BRs) (Fig 1) to perform statistical tests. For graphical representation and analysis, Microsoft excel and R-studio with R-version 3.4.3 were used. Venn diagrams and heatmaps for qualitative data were plotted with the LFQ values (S1–S3 Tables) comparing all 9 Rs. The heatmaps for quantitative data were plotted with the APEX values to compare TRs and BRs. To generate data for heatmaps a reference list was made by pooling the protein abundance values from all the 9 Rs in decreasing order of abundance. On the other hand, to filter the topmost abundant proteins, an arbitrary cut-off of 2.0E6 was considered for all the 3 cell lines. The APEX abundance values of the 9 Rs per cell line were averaged, and abundance of 2.0E6 and above was considered the most abundant. The arbitrary cut-off was set based on user-defined (1.0E8 in our data) normalization factor C, which is an estimate of the total protein abundance in 1 sample.

The variance of TRs and BRs was calculated in terms of relative standard deviation (RSD) [22, 40]. RSD was calculated using the following equation:

$$RSD = \frac{SD}{\bar{x}} \times 100$$

Where SD is the standard deviation of a dataset, and \bar{x} is the mean value of a dataset. SD was calculated using the following equation:

$$SD = \sqrt{\frac{\sum_{i=0}^n (x_i - \bar{x})^2}{n - 1}}$$

Where x_i is the observed value of 1 sample item, \bar{x} is the mean value of the observations, and n is the total number of observations.

The RSD of LFQ values in TRs was calculated as the percentage of the SD of the number of proteins identified in R1, R2, R3 / R4, R5, R6 / R7, R8, R9 divided by the average number of proteins identified in R1, R2, R3 / R4, R5, R6 / R7, R8, R9 and used as a numeric representation of the technical variance. RSD of LFQ values in TRs was calculated as follows:

$$RSD_{LFQ_TRs} = \frac{\sqrt{\frac{\sum_{i=0}^n ((TR_i - \overline{TRs}))^2}{n-1}}}{\overline{TRs}} \times 100$$

Where TR_i is the observed value of the sum of total proteins identified in each TR (i.e., $\Sigma TR_1, \Sigma TR_2,$ and ΣTR_3), \overline{TRs} is the mean value of the proteins identified in 3 TRs, and n is the total number of observations.

Whereas the RSD in BRs was calculated as the percentage of the SD of number of proteins identified in R1, R4, R7 / R2, R5, R8 / R3, R6, R9 divided by the average number of proteins identified in R1, R4, R7 / R2, R5, R8 / R3, R6, R9 and used as a numeric representation of the

biological variance. RSD of LFQ values in BRs was calculated as follows:

$$RSD_{LFQ_BRs} = \frac{\sqrt{\frac{\sum_{i=0}^n ((BR_i - \overline{BRs}))^2}{n-1}}}{\overline{BRs}} \times 100$$

Where BR_i is the observed value of the sum of total proteins identified in each BR (i.e, $\Sigma BR_1, \Sigma BR_2,$ and ΣBR_3), \overline{BRs} is the mean value of the proteins identified in 3 BRs, and n is the total number of observations.

The RSD of APEX values in TRs was calculated as the percentage of the SD of the protein abundance in R1, R2, R3 / R4, R5, R6 / R7, R8, R9 divided by the average of protein abundance in R1, R2, R3 / R4, R5, R6 / R7, R8, R9 and used as a numeric representation of the technical variance. RSD of APEX values in TRs was calculated as follows:

$$RSD_{APEX_TRs} = \frac{\sqrt{\frac{((\sum TR_1 - \overline{TRs}))^2 + (\sum TR_2 - \overline{TRs}))^2 + (\sum TR_3 - \overline{TRs}))^2}{n-1}}}{\overline{TRs}} \times 100$$

Where ΣTR_1 is the observed value of the sum of the protein abundances of $\Sigma R_1, \Sigma R_2,$ and $\Sigma R_3, \Sigma TR_2$ is the observed value of the sum of the protein abundances of $\Sigma R_4, \Sigma R_5,$ and $\Sigma R_6, \Sigma TR_3$ is the observed value of the sum of the protein abundances of $\Sigma R_7, \Sigma R_8,$ and $\Sigma R_9, \overline{TRs}$ is the mean value of the proteins identified in 3 TRs, and n is the total number of observations.

Whereas the RSD in BRs was calculated as the percentage of the SD of the protein abundance in R1, R4, R7 / R2, R5, R8 / R3, R6, R9 divided by the average of protein abundance in R1, R4, R7 / R2, R5, R8 / R3, R6, R9 and used as a numeric representation of the biological variance. RSD of APEX values in BRs was calculated as follows:

$$RSD_{APEX_BRs} = \frac{\sqrt{\frac{((\sum BR_1 - \overline{BRs}))^2 + (\sum BR_2 - \overline{BRs}))^2 + (\sum BR_3 - \overline{BRs}))^2}{n-1}}}{\overline{BRs}} \times 100$$

Where ΣBR_1 is the observed value of the sum of the protein abundances of $\Sigma R_1, \Sigma R_4,$ and $\Sigma R_7, \Sigma BR_2$ is the observed value of the sum of the protein abundances of $\Sigma R_2, \Sigma R_5,$ and $\Sigma R_8, \Sigma BR_3$ is the observed value of the sum of the protein abundances of $\Sigma R_3, \Sigma R_6,$ and $\Sigma R_9, \overline{BRs}$ is the mean value of the proteins identified in 3 BRs, and n is the total number of observations.

One-way repeated measures analysis of variance (ANOVA) was conducted to check for statistically significant differences in the means of proteins levels of TRs, BRs, and the 9 Rs of all the cell lines. A p -value of less than 0.016 was held as the threshold for identifying significant changes among TRs and BRs by applying the standard Bonferroni [41, 42] correction ($\alpha/3 = 0.05/3$) considering 3 groups. A p -value of less than 0.0055 was held as the threshold for identifying significant changes between 9 Rs by applying the standard Bonferroni correction ($\alpha/9 = 0.05/9$) considering a total of 9 groups.

Power analysis was performed on the total proteins identified in the 9 Rs per cell line using the power ANOVA test in the R-statistical package. This was performed at a significance level of 0.016, and a power of 0.8 to identify how many Rs will be needed to confidently identify all possible EE-EV proteins in the 3 cell lines. Further, the power analysis was applied to determine how many more folds of proteins would be obtained relative to the effect size when using 9 Rs.

Results

Replicates generation

Triplicate EE-EV samples from lung adenocarcinoma cell lines H1993, A549 and H1975 generated at different passages, here referred as A, B and C, were used to generate technical replicates (TRs) and biological replicates (BRs) as shown in Fig 1. Altogether, 27 replicate samples (Rs), representing 9 Rs per cell line, were analyzed for this study.

Identification of exosomal markers

CD9 and CD81 were identified by LC/MS/MS in EE-EV Rs from all 3 cell lines, while TSG101 was detected in Rs from cell line H1993. The presence of these and other exosomal markers such as CD63, Flotillin 1, and Calnexin was confirmed by western blot analysis (S1 Fig).

Particle size distribution of EE-EVs

The minimum and maximum diameters of vesicles isolated from H1993 were 82 nm and 656 nm, respectively, with a mean diameter of 157 ± 73.3 nm and a d90 value of 249 nm (90% of the vesicles showed a diameter below 249 nm) (S2 Fig). The minimum and maximum diameters of vesicles isolated from A549 were 65 nm and 564 nm respectively, with a mean diameter of 145 ± 82.3 nm, and a d90 value of 249 nm (90% of the vesicles showed a diameter below 249 nm) (S2 Fig). The minimum and maximum diameters of vesicles isolated from H1975 were 63 nm and 560 nm, respectively, with a mean diameter of 146 ± 76.8 nm, and a d90 value of 231 nm (90% of the vesicles showed a diameter below 231 nm) (S2 Fig). These diameters are consistent with that of exosomes with a small microvesicle contamination, as known to happen in exosome isolations [6, 43].

Qualitative variability analysis

The EE-EV proteins identified in triplicates of TRs and BRs from each of the 3 cell lines (S3 and S4 Figs) were subjected to qualitative variability analysis (Fig 2) The average qualitative

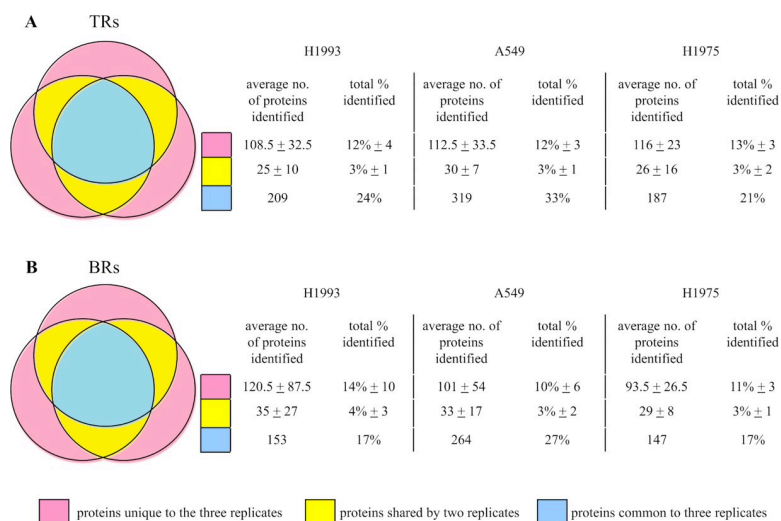


Fig 2. Venn diagrams and tables showing the average number of proteins (in total and in percentage of the total) identified across three TRs and three BRs out of the total proteins detected in the three cell lines H1993 (886), A549 (976) and H1975 (879), respectively for (A) TRs (B) BRs. The proteins unique to the three replicates is shown in pink, the proteins shared by 2 replicates is shown in yellow, and the proteins common to three replicates is shown in blue.

<https://doi.org/10.1371/journal.pone.0228871.g002>

variability from each of the 3 TRs and BRs, the average number of proteins unique to each R, shared by 2 Rs, and common to 3 Rs was calculated. Venn diagrams and tables were used to show the average number of proteins (in total and in percentage of the total) identified across 3 TRs and 3 BRs out of the total proteins detected in the 3 cell lines H1993 (886), A549 (976) and H1975 (879), respectively for TRs (Fig 2A) and for BRs (Fig 2B). These results indicated an average 6% higher qualitative variability in BRs than in TRs for the 3 cell lines studied.

Variance analysis showed qualitative variability in 3 cell line EE-EVs TRs and BRs (Fig 3A and 3B, respectively). The RSD values in BRs of H1993 (30.2%), A549 (13.8%) and H1975 (15.6%) EE-EVs can be observed to be higher than in their respective TRs (H1993–15.2%, A549–9.4% and H1975–12.7%). Therefore, RSD analysis of the qualitative data also indicated that compared to TRs (Fig 3A), BRs (Fig 3B) showed a higher variance for all the 3 cell lines studied (7%).

In addition, qualitative heatmaps of all the proteins identified in the 9 Rs (+ve in 9 Rs) obtained from each cell line indicated that a high number of proteins identified in a single R (+ve in 1 R) were absent (-ve) in the other 2, whereas the number of proteins in common among the 9 Rs was relatively small (S1–S3 Tables and Figs 4–6). Heatmap for a total of 886 proteins identified in 9 Rs of H1993 EE-EVs (Fig 4) showed 117 proteins common to all 9 Rs (+ve in 9 Rs), 312 proteins present in >1 R and <9 Rs (+ve in <9 Rs and >1 R) and 457 proteins present only in 1 R (+ve in 1 R). The number of proteins identified only in 1 R (457) was higher than those identified in all 9 Rs (117) together. Heatmap for a total of 976 proteins identified in 9 Rs of A549 EE-EVs (Fig 5) showed 223 proteins common to all 9 Rs (+ve in 9 Rs), 359 proteins present in >1 R and <9 Rs (+ve in <9 Rs and >1 R) and 394 proteins present only in 1 R (+ve in 1 R). Therefore, also in this second cell line the number of proteins identified only in 1 R (394) was higher than those identified in all 9 Rs (223) together. Heatmap for a total of 879 proteins identified in 9 Rs of H1975 EE-EVs (Fig 6) showed 108 proteins common to all 9 Rs (+ve in 9 Rs), 305 proteins present in >1 R and <9 Rs (+ve in <9 Rs and >1 R) and 466 proteins present only in 1 R (+ve in 1 R). Hence, in the third cell line, the number of proteins identified only in 1 R (466) was higher than those identified in all 9 Rs (108) together. Therefore, all the 9 Rs, obtained per cell line contributed to the completeness of the EE-EV profile in each of the 3 cell lines studied.

Quantitative variability analysis

Heatmaps were used to show the quantitative variability among exosomal proteins identified in triplicates of TRs and BRs from the 3 cell lines studied (Figs 7–9). EE-EVs proteins identified in H1993 (Fig 7), A549 (Fig 8) and H1975 (Fig 9) are shown in the order of decreasing levels in TRs (R1-R2-R3, R4-R5-R6, R7-R8-R9) as well as in BRs (R1-R4-R7, R2-R5-R6, R3-R6-R9). From these heatmaps, it can be concluded that the abundance of each protein was

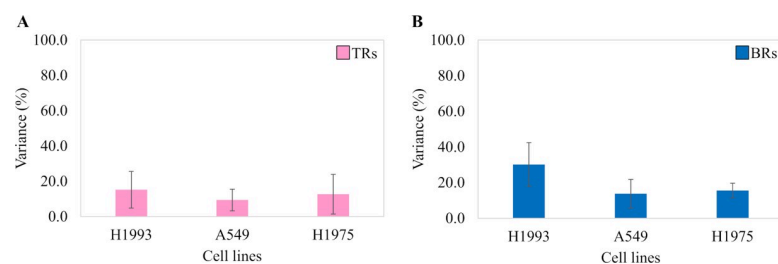


Fig 3. Variance analysis showing qualitative variability in TRs and BRs of three H1993, A549, and H1975 cell line exosomes.

<https://doi.org/10.1371/journal.pone.0228871.g003>

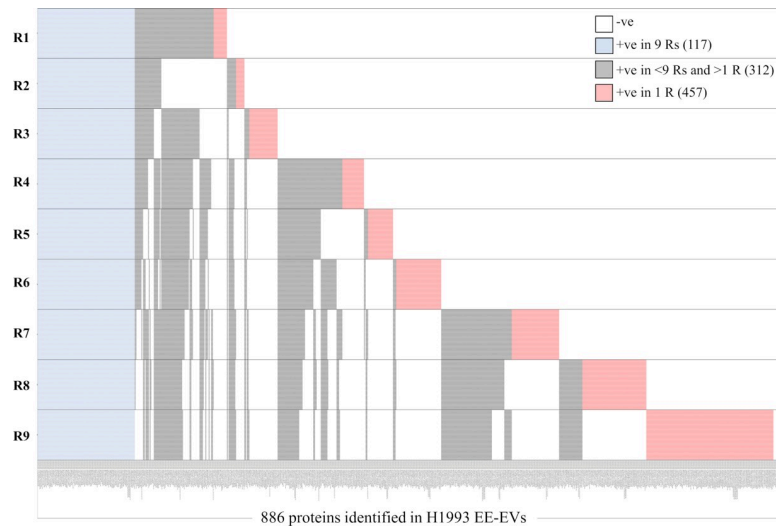


Fig 4. Heatmap showing number of samples in which each of the total 886 proteins was identified in H1993 exosomes. Proteins common to all nine replicates are shown in blue (117). Proteins present in >1 replicate and <9 replicates are shown in grey (312) and proteins present only in 1 replicate are shown in red (457).

<https://doi.org/10.1371/journal.pone.0228871.g004>

similar across all the Rs. However, the BRs showed more quantitative variability compared to TRs. The APEX values for each of the Rs are presented in [S4 Table](#).

To evaluate the variance between Rs, we assessed the quantitative variability by RSD analysis of the quantitative data from all the 3 cell line EE-EVs TRs and BRs ([Fig 10A and 10B](#), respectively). This figure ([Fig 10](#)) shows that the RSD values in BRs of H1993 (68.2%), A549 (75.8%) and H1975 (59.8%) EE-EVs were higher than in their respective TRs (H1993–20.8%, A549–21.4% and H1975–22.0%). In conclusion, RSD analysis indicated that compared to TRs ([Fig 10A](#)), BRs ([Fig 10B](#)) showed 47% higher variance in BRs than in TRs for all the 3 cell lines studied.

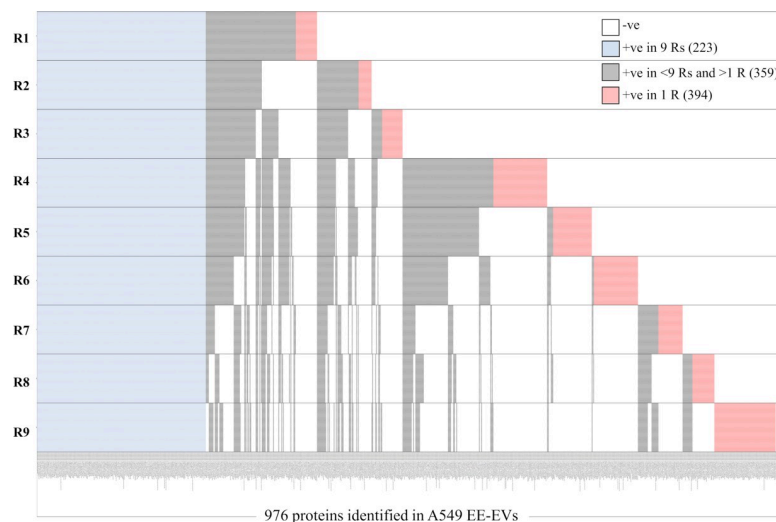


Fig 5. Heatmap showing number of samples in which each of the total 976 proteins was identified in A549 exosomes. Proteins common to all nine replicates are shown in blue (223). Proteins present in >1 replicate and <9 replicates are shown in grey (359) and proteins present only in 1 replicate are shown in red (394).

<https://doi.org/10.1371/journal.pone.0228871.g005>

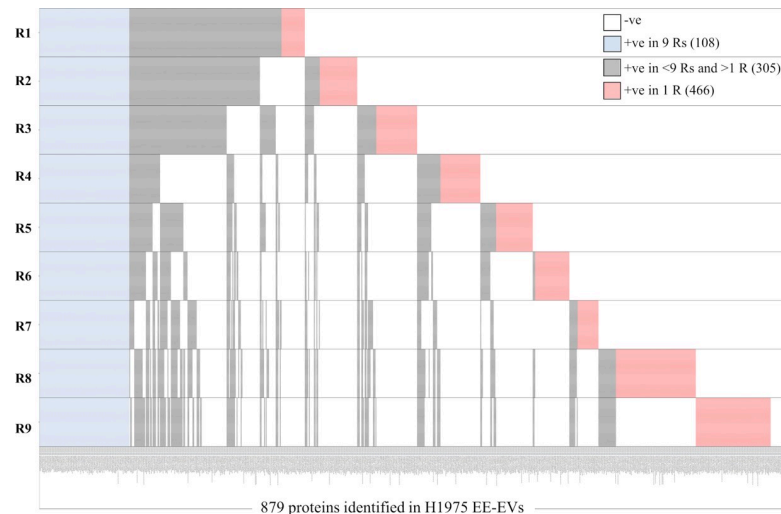


Fig 6. Heatmap showing number of samples in which each of the total 879 proteins was identified in H1975 exosomes. Proteins common to all nine replicates are shown in blue (108). Proteins present in >1 replicate and <9 replicates are shown in grey (305) and proteins present only in 1 replicate are shown in red (466).

<https://doi.org/10.1371/journal.pone.0228871.g006>

However, there was no statistically significant quantitative variability observed among the triplicates of TRs and BRs, except for the H1975 cell line (Table 1), which showed statistically significant quantitative variability in TRs, BRs as well as all the 9 Rs. All these data show a cumulatively higher degree of variability in Rs of H1975 cell line EE-EVs when compared to the H1993 and A549 cell line EE-EVs.

Power analysis

Finally, we assessed how much the number of EE-EVs proteins would increase if we analyzed all the Rs indicated by the power analysis, which was 18, 27 and 252 Rs for cell lines H1993, A549 and H1975, respectively. We found that the analysis of all these Rs would only result in 0.2 fold increase in the number of proteins identified in EE-EVs from cell line H1993, 0.3 fold increase in the number of proteins from cell line A549 and 0.4 fold increase in the number of proteins from the cell line H1975. Therefore, increasing the number of Rs would not increase significantly the number of total proteins identified in their EE-EVs.

Discussion

The study of exosomal cargo often requires high-throughput analysis of replicate samples (Rs). As the number and abundance of identified biomolecules varies between Rs, establishing the replicate variability predicted for the event under study is essential in determining the number of Rs required for reaching accurate conclusions. Since, to the best of our knowledge, the variability between Rs of any of the various types of exosomal cargo has not been previously reported; in this study, we used LC/MS/MS analysis of exosome enriched EVs (EE-EVs) technical replicates (TRs) and biological replicates (BRs) from 3 different lung adenocarcinoma cell lines to determine the qualitative and quantitative variability in the detected proteins. To maximize protein identification, we analyzed 100 µg of total exosomal protein per replicate [44].

Our workflow started by establishing the qualitative variability among Rs. Venn diagrams and RSD analysis showed considerable variability in the proteins identified in TRs; an

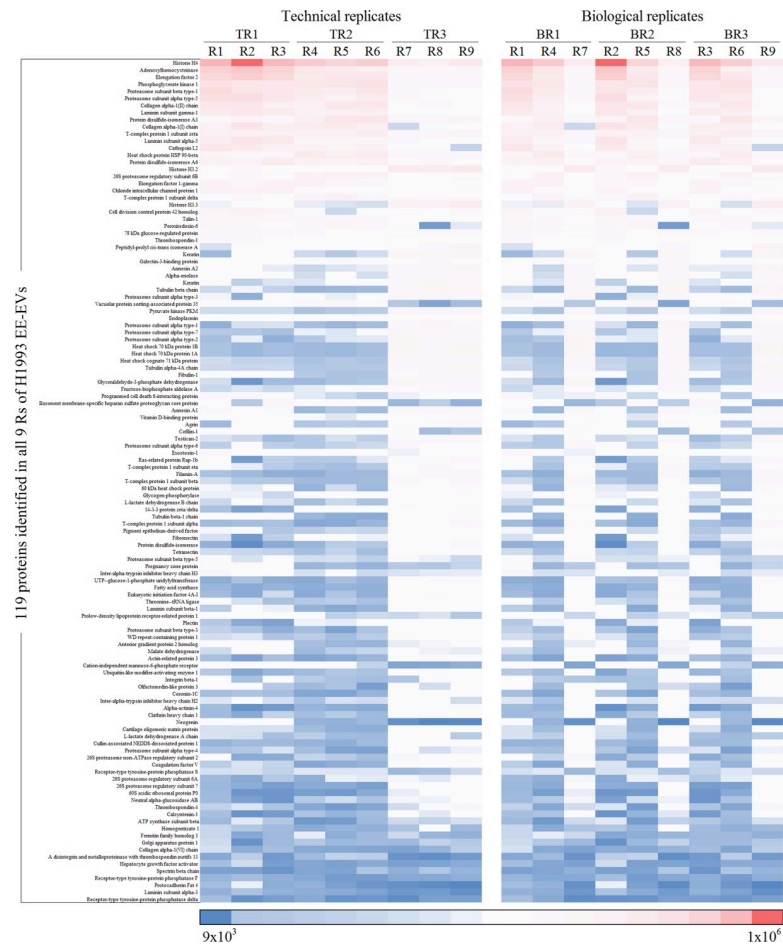


Fig 7. Heatmap representing exosomal proteins identified in H1993 in order of decreasing levels in TRs (R1-R2-R3, R4-R5-R6, R7-R8-R9) and in BRs (R1-R4-R7, R2-R5-R6, R3-R6-R9). Boxes highlighted in dark red represent the most abundant, while boxes in dark blue represent the least abundant proteins.

<https://doi.org/10.1371/journal.pone.0228871.g007>

unexpected finding considering that each set of TRs originates from a single EE-EV sample. Therefore, technical factors have an impact on protein identification by LC/MS/MS analysis, which cannot be overlooked. Among the technical sources of TRs variability, including extraction, digestion, instrumental variance and instrumental stability, a study conducted by Piehowski et al concluded that the main source of variability is instrumental variance, and mainly involves ion-suppression and chromatographic disturbances [45].

The qualitative variability among BRs was, in average, 6% higher than that of TRs for the 3 cell lines studied, pointing to the existence of a small set of passage-dependent proteins. In vivo, changes in exosomal protein cargo occur due to a variety of causes, including viral infections [46], internal diseases [47–49], radiation [50], and ageing [51]. Interestingly, when we generated a heatmap of the proteins present in each of the 9 Rs obtained per cell line, we found that the number of proteins unique to each replicate was higher than the number of proteins common to all 9. Therefore, each of the Rs, whether TRs or BRs, contributed to generate a more complete EE-EVs protein profile. This finding had an impact on the power analysis, which is discussed in a later paragraph.

Next, we focused on the quantitative analysis of the data. Heatmaps showed that the abundance of each protein was similar across all the Rs. Although, the BRs showed more variability

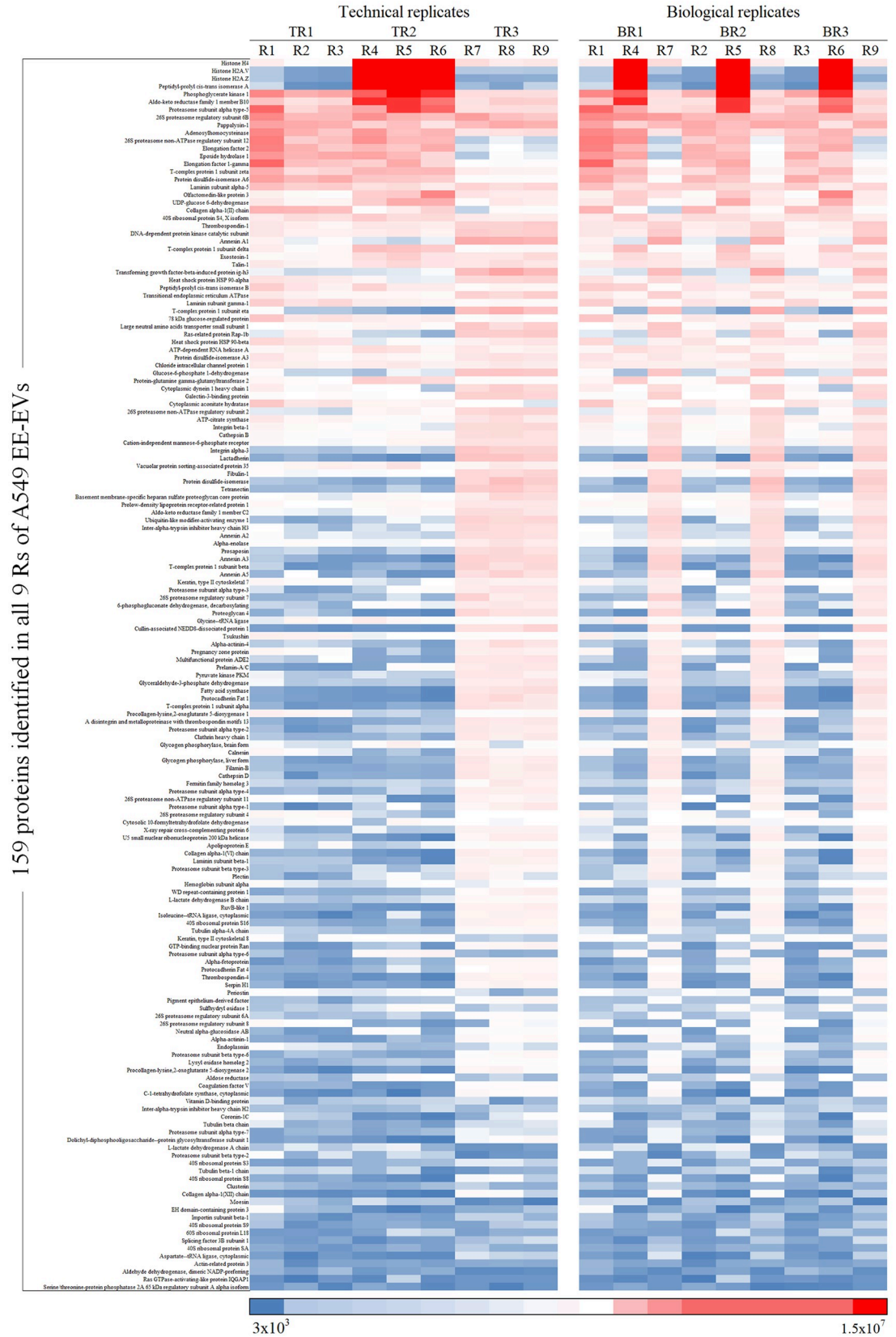


Fig 8. Heatmap representing exosomal proteins identified in A549 in order of decreasing levels in TRs (R1-R2-R3, R4-R5-R6, R7-R8-R9) and in BRs (R1-R4-R7, R2-R5-R6, R3-R6-R9). Boxes highlighted in dark red represent the most abundant, while boxes in dark blue represent the least abundant proteins.

<https://doi.org/10.1371/journal.pone.0228871.g008>

than the TRs, a finding previously reported for quantitative LC/MS/MS [52]. Our results, however, stood in contrast with a cell metabolome study in which the variability in BRs was found to be lower than that of the TRs [22]. Such a discrepancy between studies suggests that variability among BRs may hinge on the biological system under study, with some more stable than others. With regard to the additional biostatistical analyses performed here, the RSD analysis supported what we observed in the quantitative heatmaps, although it showed a 47% higher variance in BRs than in TRs for all the 3 cell lines studied, which is also likely to be a passage-dependent effect. The ANOVA analysis, however, showed no statistically significant differences in protein abundance among Rs, except for the H1975 cell line.

Importantly, for all 3 cell lines studied, the abundance of the 90% top proteins was similar in BRs and TRs, an observation consistent with the general concept in mass spectrometry studies that the top 75% most abundant proteins in Rs from a complex sample are very reproducibly detected, but the bottom 25% are quite variable [53].



Fig 9. Heatmap representing exosomal proteins identified in H1975 in order of decreasing levels in TRs (R1-R2-R3, R4-R5-R6, R7-R8-R9) and in BRs (R1-R4-R7, R2-R5-R6, R3-R6-R9). Boxes highlighted in dark red represent the most abundant, while boxes in dark blue represent the least abundant proteins.

<https://doi.org/10.1371/journal.pone.0228871.g009>

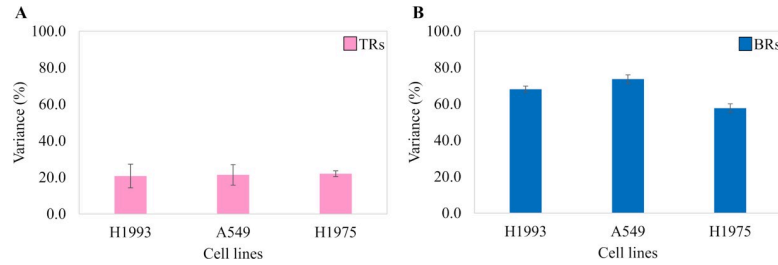


Fig 10. Variance analysis showing quantitative variability in TRs and BRs of three H1993, A549, and H1975 cell line exosomes.

<https://doi.org/10.1371/journal.pone.0228871.g010>

We finally determined the power analysis for each of the 3 cell lines studied based on the 9 Rs collected from each one. Such analysis indicated that 23 Rs were required to identify the maximum number of EE-EV proteins in H1993 and A549 cell lines, and approximately a 10-fold higher number of Rs was required for the H1975 cell line, which showed the highest qualitative and quantitative variability between Rs. These numbers were expected based on the qualitative heatmaps previously discussed, and for the first 2 cell lines it was slightly below of $n = thirty$, which statisticians consider appropriate to get a feeling for the mean and its distribution [21]. Nevertheless, the generation of as many Rs as indicated by our power analysis, is unrealistic, both for practical and financial reasons. Therefore, it is important to stress that performing all these Rs will only produce a 0.3-folds increase in EE-EV protein detection for all the 3 cell lines studied.

In conclusion, we found that the variability among TRs as well as BRs was largely qualitative and higher among BRs. By contrast, the quantitative variability was low, except for a single cell line where the quantitative variability was significant. Importantly, our replicate strategy of

Table 1. Table shows one-way repeated measures ANOVA F-values and corresponding p-values for the levels of proteins identified in EE-EVs in 3 TRs and 3 BRs for H1993, A549 and H1975 cell lines.

A.		
Sample	One-way repeated ANOVA	
	F—value	p—value
H1993 TRs	2.2	0.113
A549 TRs	2.35	0.097
H1975 TRs	11.38	1.98E-005**

B.		
Sample	One-way repeated ANOVA	
	F—value	p—value
H1993 BRs	1.891	0.153
A549 BRs	2.923	0.0552
H1975 BRs	6.413	0.00916*

C.		
Sample	One-way repeated ANOVA	
	F—value	p—value
H1993 9 Rs	2.373	0.0156
A549 9 Rs	2.385	0.015
H1975 9 Rs	10.99	7.37E-15***

* represents $p < 0.016$, which is considered statistically significant. Table 1A represents the analysis of TRs. Table 1B represents the analysis of BRs, and Table 1C represents the analysis of 9 Rs.

<https://doi.org/10.1371/journal.pone.0228871.t001>

analyzing 3 BRs, each divided into 3 TRs, identified 90% of the most abundant proteins, thereby establishing the utility of our approach.

Supporting information

S1 Fig. Western blot showing the presence of exosomal markers in EE-EVs lysate (EL) and total cell lysate (TCL) obtained from H1993, A549 and H1975 cell lines.

(TIF)

S2 Fig. TRPS analysis of EE-EVs from H1993, A549 and H1975 cell lines.

(TIF)

S3 Fig. Venn diagrams showing the number of proteins identified in 3 EE-EV TRs from the total proteins detected in the 3 cell lines H1993 (886), A549 (976) and H1975 (879), respectively.

(TIF)

S4 Fig. Venn diagrams showing the number of proteins identified in 3 EE-EV BRs from the total proteins detected in the 3 cell lines H1993 (886), A549 (976) and H1975 (879), respectively.

(TIF)

S1 Table. LFQ intensity values for all the 9 Rs in H1993 EE-EVs for checking presence/absence of proteins.

(XLS)

S2 Table. LFQ intensity values for all the 9 Rs in A549 EE-EVs for checking presence/absence of proteins.

(XLS)

S3 Table. LFQ intensity values for all the 9 Rs in H1975 EE-EVs for checking presence/absence of proteins.

(XLS)

S4 Table. APEX values of the list of proteins identified in common across nine replicates in H1993, A549 and H1975 exosomes. Protein names represented in Figs 7–9 are given here.

(XLS)

Author Contributions

Conceptualization: Lucia Schuger.

Data curation: Suma Tiruvayipati.

Formal analysis: Suma Tiruvayipati.

Methodology: Suma Tiruvayipati.

Supervision: Don Wolfgeher, Jorge Andrade, Hui Jiang, Lucia Schuger.

Validation: Suma Tiruvayipati, Don Wolfgeher, Ming Yue, FangFang Duan.

Visualization: Suma Tiruvayipati.

Writing – original draft: Suma Tiruvayipati, Lucia Schuger.

Writing – review & editing: Ming Yue, Lucia Schuger.

References

1. Lucchetti D, Fattorossi A, Sgambato A. Extracellular Vesicles in Oncology: Progress and Pitfalls in the Methods of Isolation and Analysis. *Biotechnol J*. 2019; 14(1):e1700716. Epub 2018/06/08. <https://doi.org/10.1002/biot.201700716> PMID: 29878510.
2. van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol*. 2018; 19(4):213–28. Epub 2018/01/18. <https://doi.org/10.1038/nrm.2017.125> [pii]. PMID: 29339798.
3. Xu R, Greening DW, Zhu HJ, Takahashi N, Simpson RJ. Extracellular vesicle isolation and characterization: toward clinical application. *J Clin Invest*. 2016; 126(4):1152–62. <https://doi.org/10.1172/JCI81129> PMID: 27035807; PubMed Central PMCID: PMC4811150.
4. Zhang HG, Grizzle WE. Exosomes: a novel pathway of local and distant intercellular communication that facilitates the growth and metastasis of neoplastic lesions. *The American journal of pathology*. 2014; 184(1):28–41. Epub 2013/11/26. <https://doi.org/10.1016/j.ajpath.2013.09.027> [pii]. PMID: 24269592; PubMed Central PMCID: PMC3873490.
5. de la Torre Gomez C, Goreham RV, Bech Serra JJ, Nann T, Kussmann M. "Exosomics"-A Review of Biophysics, Biology and Biochemistry of Exosomes With a Focus on Human Breast Milk. *Frontiers in genetics*. 2018; 9:92. <https://doi.org/10.3389/fgene.2018.00092> PMID: 29636770; PubMed Central PMCID: PMC5881086.
6. Yan Z, Dutta S, Liu Z, Yu X, Mesgarzadeh N, Ji F, et al. A Label-Free Platform for Identification of Exosomes from Different Sources. *ACS sensors*. 2019; 4(2):488–97. <https://doi.org/10.1021/acssensors.8b01564> PMID: 30644736.
7. Colombo M, Raposo G, Thery C. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. *Annu Rev Cell Dev Biol*. 2014; 30:255–89. Epub 2014/10/08. <https://doi.org/10.1146/annurev-cellbio-101512-122326> PMID: 25288114.
8. Kalamvoki M, Du T, Roizman B. Cells infected with herpes simplex virus 1 export to uninfected cells exosomes containing STING, viral mRNAs, and microRNAs. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(46):E4991–6. <https://doi.org/10.1073/pnas.1419338111> PMID: 25368198; PubMed Central PMCID: PMC4246290.
9. Wu AY, Ueda K, Lai CP. Proteomic Analysis of Extracellular Vesicles for Cancer Diagnostics. *Proteomics*. 2019; 19(1–2):e1800162. <https://doi.org/10.1002/pmic.201800162> PMID: 30334355.
10. Kowal J, Arras G, Colombo M, Jouve M, Morath JP, Primdal-Bengtson B, et al. Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proceedings of the National Academy of Sciences of the United States of America*. 2016; 113(8):E968–77. <https://doi.org/10.1073/pnas.1521230113> PMID: 26858453; PubMed Central PMCID: PMC4776515.
11. Boriachek K, Islam MN, Moller A, Salomon C, Nguyen NT, Hossain MSA, et al. Biological Functions and Current Advances in Isolation and Detection Strategies for Exosome Nanovesicles. *Small*. 2018; 14(6). Epub 2017/12/29. <https://doi.org/10.1002/sml.201702153> PMID: 29282861.
12. Vagner T, Spinelli C, Minciacchi VR, Balaj L, Zandian M, Conley A, et al. Large extracellular vesicles carry most of the tumour DNA circulating in prostate cancer patient plasma. *J Extracell Vesicles*. 2018; 7(1):1505403. Epub 2018/08/16. <https://doi.org/10.1080/20013078.2018.1505403> [pii]. PMID: 30108686; PubMed Central PMCID: PMC6084494.
13. Lobb RJ, Hastie ML, Norris EL, van Amerongen R, Gorman JJ, Moller A. Oncogenic transformation of lung cells results in distinct exosome protein profile similar to the cell of origin. *Proteomics*. 2017; 17(23–24). Epub 2017/07/20. <https://doi.org/10.1002/pmic.201600432> PMID: 28722786.
14. Valadi H, Ekstrom K, Bossios A, Sjostrand M, Lee JJ, Lotvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol*. 2007; 9(6):654–9. <https://doi.org/10.1038/ncb1596> PMID: 17486113.
15. Wen SW, Lima LG, Lobb RJ, Norris EL, Hastie ML, Krumeich S, et al. Breast cancer-derived exosomes reflect the cell-of-origin phenotype. *Proteomics*. 2019:e1800180. Epub 2019/01/24. <https://doi.org/10.1002/pmic.201800180> PMID: 30672117.
16. De Rubis G, Rajeev Krishnan S, Bebawy M. Liquid Biopsies in Cancer Diagnosis, Monitoring, and Prognosis. *Trends Pharmacol Sci*. 2019. Epub 2019/02/10. S0165-6147(19)30017-3 [pii] <https://doi.org/10.1016/j.tips.2019.01.006> PMID: 30736982.
17. Mathivanan S, Simpson RJ. ExoCarta: A compendium of exosomal proteins and RNA. *Proteomics*. 2009; 9(21):4997–5000. Epub 2009/10/08. <https://doi.org/10.1002/pmic.200900351> PMID: 19810033.
18. Gho YS, Lee C. Emergent properties of extracellular vesicles: a holistic approach to decode the complexity of intercellular communication networks. *Mol Biosyst*. 2017; 13(7):1291–6. Epub 2017/05/11. <https://doi.org/10.1039/c7mb00146k> PMID: 28488707.

19. Konoshenko MY, Lekchnov EA, Vlassov AV, Laktionov PP. Isolation of Extracellular Vesicles: General Methodologies and Latest Trends. *BioMed research international*. 2018; 2018:8545347. <https://doi.org/10.1155/2018/8545347> PMID: 29662902; PubMed Central PMCID: PMC5831698.
20. Li P, Kaslan M, Lee SH, Yao J, Gao Z. Progress in Exosome Isolation Techniques. *Theranostics*. 2017; 7(3):789–804. <https://doi.org/10.7150/thno.18133> PMID: 28255367; PubMed Central PMCID: PMC5327650.
21. Naegle K, Gough NR, Yaffe MB. Criteria for biological reproducibility: what does "n" mean? *Science signaling*. 2015; 8(371):fs7. Epub 2015/04/09. <https://doi.org/10.1126/scisignal.aab1125> [pii]. PMID: 25852186.
22. Le H, Jerums M, Goudar CT. Characterization of intrinsic variability in time-series metabolomic data of cultured mammalian cells. *Biotechnol Bioeng*. 2015; 112(11):2276–83. Epub 2015/05/16. <https://doi.org/10.1002/bit.25646> PMID: 25976859.
23. McIntyre LM, Lopiano KK, Morse AM, Amin V, Oberg AL, Young LJ, et al. RNA-seq: technical variability and sampling. *BMC Genomics*. 2011; 12:293. Epub 2011/06/08. <https://doi.org/10.1186/1471-2164-12-293> [pii]. PMID: 21645359; PubMed Central PMCID: PMC3141664.
24. Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. *J Proteome Res*. 2011; 10(2):637–45. Epub 2010/12/04. <https://doi.org/10.1021/pr100835s> PMID: 21126025.
25. Richards KE, Zeleniak AE, Fishel ML, Wu J, Littlepage LE, Hill R. Cancer-associated fibroblast exosomes regulate survival and proliferation of pancreatic cancer cells. *Oncogene*. 2017; 36(13):1770–8. Epub 2016/09/27. <https://doi.org/10.1038/onc.2016.353> [pii]. PMID: 27669441; PubMed Central PMCID: PMC5366272.
26. Zhou CF, Ma J, Huang L, Yi HY, Zhang YM, Wu XG, et al. Cervical squamous cell carcinoma-secreted exosomal miR-221-3p promotes lymphangiogenesis and lymphatic metastasis by targeting VASH1. *Oncogene*. 2019; 38(8):1256–68. Epub 2018/09/27. <https://doi.org/10.1038/s41388-018-0511-x> [pii]. PMID: 30254211; PubMed Central PMCID: PMC6363643.
27. Yukawa H, Suzuki K, Aoki K, Arimoto T, Yasui T, Kaji N, et al. Imaging of angiogenesis of human umbilical vein endothelial cells by uptake of exosomes secreted from hepatocellular carcinoma cells. *Sci Rep*. 2018; 8(1):6765. Epub 2018/05/02. <https://doi.org/10.1038/s41598-018-24563-0> [pii]. PMID: 29713019; PubMed Central PMCID: PMC5928189.
28. Cheng T, Yue M, Aslam MN, Wang X, Shekhawat G, Varani J, et al. Neuronal Protein 3.1 Deficiency Leads to Reduced Cutaneous Scar Collagen Deposition and Tensile Strength due to Impaired Transforming Growth Factor-beta1 to -beta3 Translation. *The American journal of pathology*. 2017; 187(2):292–303. <https://doi.org/10.1016/j.ajpath.2016.10.004> PMID: 27939132; PubMed Central PMCID: PMC5389364.
29. Akers JC, Ramakrishnan V, Nolan JP, Duggan E, Fu CC, Hochberg FH, et al. Comparative Analysis of Technologies for Quantifying Extracellular Vesicles (EVs) in Clinical Cerebrospinal Fluids (CSF). *PLoS One*. 2016; 11(2):e0149866. <https://doi.org/10.1371/journal.pone.0149866> PMID: 26901428; PubMed Central PMCID: PMC4763994.
30. Lane RE, Korbie D, Anderson W, Vaidyanathan R, Trau M. Analysis of exosome purification methods using a model liposome system and tunable-resistive pulse sensing. *Sci Rep*. 2015; 5:7639. <https://doi.org/10.1038/srep07639> PMID: 25559219; PubMed Central PMCID: PMC4648344.
31. Vogel R, Pal AK, Jambhrunkar S, Patel P, Thakur SS, Reategui E, et al. High-Resolution Single Particle Zeta Potential Characterisation of Biological Nanoparticles using Tunable Resistive Pulse Sensing. *Sci Rep*. 2017; 7(1):17479. <https://doi.org/10.1038/s41598-017-14981-x> PMID: 29234015; PubMed Central PMCID: PMC5727177.
32. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*. 1996; 68(5):850–8. Epub 1996/03/01. <https://doi.org/10.1021/ac950914h> PMID: 8779443.
33. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 2019; 47(D1):D442–D50. Epub 2018/11/06. <https://doi.org/10.1093/nar/gky1106> [pii]. PMID: 30395289; PubMed Central PMCID: PMC6323896.
34. Tyanova S, Temu T, Carlson A, Sinitcyn P, Mann M, Cox J. Visualization of LC-MS/MS proteomics data in MaxQuant. *Proteomics*. 2015; 15(8):1453–6. Epub 2015/02/04. <https://doi.org/10.1002/pmic.201400449> PMID: 25644178; PubMed Central PMCID: PMC5024039.
35. Chiva C, Ortega M, Sabido E. Influence of the digestion technique, protease, and missed cleavage peptides in protein quantitation. *J Proteome Res*. 2014; 13(9):3979–86. Epub 2014/07/06. <https://doi.org/10.1021/pr500294d> PMID: 24986539.

36. Braisted JC, Kuntumalla S, Vogel C, Marcotte EM, Rodrigues AR, Wang R, et al. The APEX Quantitative Proteomics Tool: generating protein quantitation estimates from LC-MS/MS proteomics results. *BMC Bioinformatics*. 2008; 9:529. Epub 2008/12/11. <https://doi.org/10.1186/1471-2105-9-529> [pii]. PMID: 19068132; PubMed Central PMCID: PMC2639435.
37. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol*. 2007; 25(1):117–24. Epub 2006/12/26. nbt1270 [pii] <https://doi.org/10.1038/nbt1270> PMID: 17187058.
38. Searle BC. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics*. 2010; 10(6):1265–9. Epub 2010/01/16. <https://doi.org/10.1002/pmic.200900437> PMID: 20077414.
39. Cottrell JS. Protein identification using MS/MS data. *J Proteomics*. 2011; 74(10):1842–51. Epub 2011/06/04. <https://doi.org/10.1016/j.jprot.2011.05.014> [pii]. PMID: 21635977.
40. Chavez-Servin JL, Castellote AI, Lopez-Sabater MC. Analysis of mono- and disaccharides in milk-based formulae by high-performance liquid chromatography with refractive index detection. *Journal of chromatography A*. 2004; 1043(2):211–5. <https://doi.org/10.1016/j.chroma.2004.06.002> PMID: 15330094.
41. Perneger TV. What's wrong with Bonferroni adjustments. *BMJ*. 1998; 316(7139):1236–8. Epub 1998/05/16. <https://doi.org/10.1136/bmj.316.7139.1236> PMID: 9553006; PubMed Central PMCID: PMC1112991.
42. Robert JC, Randall JM. To Bonferroni or Not to Bonferroni: When and How are the Questions. *Bulletin of the Ecological Society of America*. 2000; 81(3):246–8.
43. Kim H, Lee KH, Han SI, Lee D, Chung S, Lee D, et al. Origami-paper-based device for microvesicle/exosome preconcentration and isolation. *Lab on a chip*. 2019; 19(23):3917–21. <https://doi.org/10.1039/c9lc00796b> PMID: 31650155.
44. Humphrey SJ, Karayel O, James DE, Mann M. High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nat Protoc*. 2018; 13(9):1897–916. Epub 2018/09/08. <https://doi.org/10.1038/s41596-018-0014-9> [pii]. PMID: 30190555.
45. Piehowski PD, Petyuk VA, Orton DJ, Xie F, Moore RJ, Ramirez-Restrepo M, et al. Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis. *J Proteome Res*. 2013; 12(5):2128–37. <https://doi.org/10.1021/pr301146m> PMID: 23495885; PubMed Central PMCID: PMC3695475.
46. Meckes DG Jr., Gunawardena HP, Dekroon RM, Heaton PR, Edwards RH, Ozgur S, et al. Modulation of B-cell exosome proteins by gamma herpesvirus infection. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110(31):E2925–33. Epub 2013/07/03. <https://doi.org/10.1073/pnas.1303906110> [pii]. PMID: 23818640; PubMed Central PMCID: PMC3732930.
47. He Z, Guan X, Liu Y, Tao Z, Liu Q, Wu J, et al. Alteration of exosomes secreted from renal tubular epithelial cells exposed to high-concentration oxalate. *Oncotarget*. 2017; 8(54):92635–42. Epub 2017/12/02. <https://doi.org/10.18632/oncotarget.21517> [pii]. PMID: 29190944; PubMed Central PMCID: PMC5696210.
48. Lakhter AJ, Pratt RE, Moore RE, Doucette KK, Maier BF, DiMeglio LA, et al. Beta cell extracellular vesicle miR-21-5p cargo is increased in response to inflammatory cytokines and serves as a biomarker of type 1 diabetes. *Diabetologia*. 2018; 61(5):1124–34. Epub 2018/02/16. <https://doi.org/10.1007/s00125-018-4559-5> [pii]. PMID: 29445851; PubMed Central PMCID: PMC5878132.
49. Levanen B, Bhakta NR, Torregrosa Paredes P, Barbeau R, Hiltbrunner S, Pollack JL, et al. Altered microRNA profiles in bronchoalveolar lavage fluid exosomes in asthmatic patients. *J Allergy Clin Immunol*. 2013; 131(3):894–903. Epub 2013/01/22. <https://doi.org/10.1016/j.jaci.2012.11.039> [pii]. PMID: 23333113; PubMed Central PMCID: PMC4013392.
50. Jelonek K, Wojakowska A, Marczak L, Muer A, Tinhofer-Keilholz I, Lysek-Gladysinska M, et al. Ionizing radiation affects protein composition of exosomes secreted in vitro from head and neck squamous cell carcinoma. *Acta Biochim Pol*. 2015; 62(2):265–72. Epub 2015/06/23. https://doi.org/10.18388/abp.2015_970 [pii]. PMID: 26098714.
51. Tietje A, Maron KN, Wei Y, Feliciano DM. Cerebrospinal fluid extracellular vesicles undergo age dependent declines and contain known and novel non-coding RNAs. *PLoS One*. 2014; 9(11):e113116. Epub 2014/11/25. <https://doi.org/10.1371/journal.pone.0113116> [pii]. PMID: 25420022; PubMed Central PMCID: PMC4242609.
52. Bullard JH, Purdom E, Hansen KD, Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics*. 2010; 11:94. Epub 2010/02/20. <https://doi.org/10.1186/1471-2105-11-94> [pii]. PMID: 20167110; PubMed Central PMCID: PMC2838869.
53. Tabb DL, Vega-Montoto L, Rudnick PA, Variyath AM, Ham AJ, Bunk DM, et al. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res*. 2010; 9(2):761–76. Epub 2009/11/20. <https://doi.org/10.1021/pr9006365> PMID: 19921851; PubMed Central PMCID: PMC2818771.