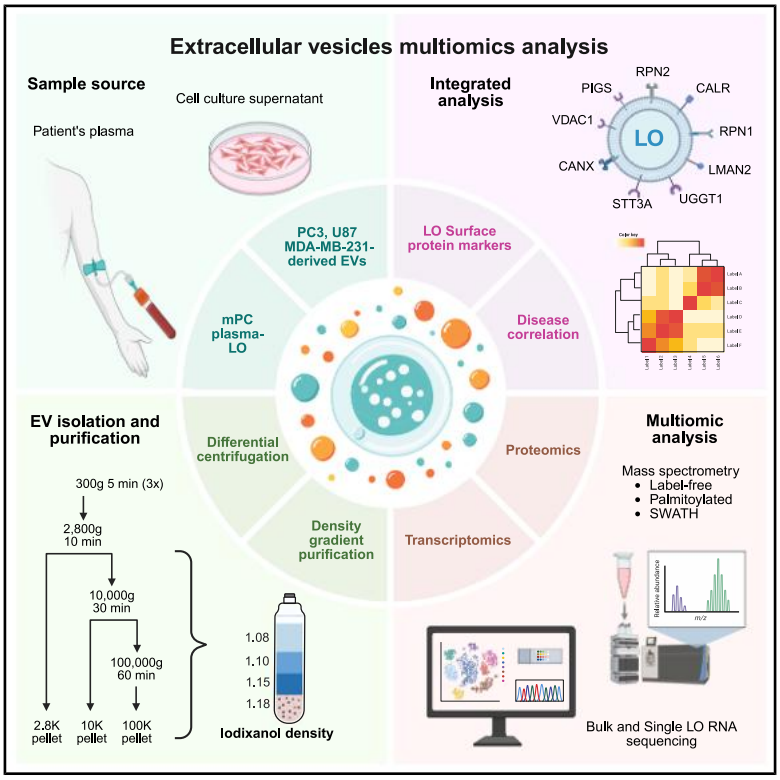
Cell Reports
Medicine

Article

Extracellular vesicle heterogeneity through the lens of multiomics

Graphical abstract



Highlights

- The LO cargo identified here is consistent across cancer cell lines and patient plasma
- Purified plasma LOs can be analyzed by data-independent acquisition mass spectrometry
- Single-LO RNA-seq confirms mitochondrial enrichment in LO-positive droplets
- Plasma LOs may aid in patient stratification and liquid biopsy development

Authors

Taylon F. Silva, Elizabeth Hutchins, Wenyan Zhao, ..., Paul C. Boutros, Kendall Van Keuren-Jensen, Dolores Di Vizio

Correspondence

pboutros@mednet.ucla.edu (P.C.B.), kjensen@tgen.org (K.V.K.-J.), dolores.divizio@cshs.org (D.D.V.)

In brief

Silva and colleagues refine an EV isolation method to optimize recovery of cancer-derived large oncosomes (LOs). This approach enables high-yield profiling by mass spectrometry and bulk/single-LO RNA-seq. A consistent set of credentialed proteins is enriched in LOs from various cancer sources, including patient plasma.

Silva et al., 2025, Cell Reports Medicine 6, 102161
July 15, 2025
https://doi.org/10.1016/j.xcrm.2025.102161

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Extracellular vesicle heterogeneity through the lens of multiomics

Taylon F. Silva,[1,21] Elizabeth Hutchins,[2,21] Wenyan Zhao,[3,21] Yari Ciani,[4] Minhyung Kim,[1] Emily Ko,[5] Javier Mariscal,[1] Zhuyu Qiu,[3,6] Fatima Bedier,[1] Agnes Kittel,[7] Bo Zhou,[1] Yang Wang,[1] Megan Hall,[2] Francesca Galasso,[2] Rebecca Reiman,[2] Michael R. Freeman,[1] Sarah Parker,[8] Jennifer Van Eyk,[8] Wei Yang,[1] Edwin Posadas,[9,10] Jlenia Guarnerio,[5,11,12] John Nolan,[13] Clotilde Théry,[14] Andries Zijlstra,[15] Shannon Stott,[16,17,18] Sungyong You,[1] Francesca Demichelis,[4] Paul C. Boutros,[3,6,19,20,*] Kendall Van Keuren-Jensen,[2,*] and Dolores Di Vizio[1,9,22,*]

[1]Department of Urology, Division of Cancer Biology and Therapeutics, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[2]Neurogenomics Division, Translational Genomics Research Institute (TGen), Phoenix, AZ, USA
[3]Jonsson Comprehensive Cancer Center, University of California: Los Angeles, Los Angeles, CA, USA
[4]Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Trento, Italy
[5]Department of Radiation Oncology, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[6]Department of Human Genetics, University of California: Los Angeles, Los Angeles, CA, USA
[7]HUN-REN Institute of Experimental Medicine, Budapest, Hungary
[8]Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[9]Samuel Oschin Comprehensive Cancer Institute, Division of Cancer Biology and Therapeutics, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[10]Department of Medicine, David Geffen School of Medicine, University of California: Los Angeles, Los Angeles, CA, USA
[11]Board of Governors Regenerative Medicine Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[12]Department of Hematology Oncology, David Geffen Medical School, University of California: Los Angeles, Los Angeles, CA, USA
[13]Scintillon Institute for Biomedical and Bioenergy Research, San Diego, CA, USA
[14]Institut Curie, PSL Research University, INSERMU932, Paris, France
[15]Department of Pathology, Vanderbilt University Medical Center, Nashville, TN, USA
[16]Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, MA, USA
[17]BioMEMS Resource Center, Center for Engineering in Medicine and Surgical Services, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[18]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[19]Department of Urology, University of California: Los Angeles, Los Angeles, CA, USA
[20]Institute for Precision Health, University of California: Los Angeles, Los Angeles, CA, USA
[21]These authors contributed equally
[22]Lead contact
*Correspondence: pboutros@mednet.ucla.edu (P.C.B.), kjensen@tgen.org (K.V.K.-J.), dolores.divizio@cshs.org (D.D.V.)
https://doi.org/10.1016/j.xcrm.2025.102161

SUMMARY

Extracellular vesicles (EVs) are heterogeneous in size, biogenesis, content, and function. Aggressive cancer cells release a distinct, poorly characterized, and particularly large EV subtype, namely large oncosomes (LOs). This study employs an optimized method to improve LO yields and integrates mass spectrometry and RNA sequencing (RNA-seq) to profile their molecular cargo. A consistent set of proteins enriched in LOs is identified across glioma, prostate, and breast cancer cell lines. These proteins are also present as mRNA in LOs from the prostate cancer model and are abundant in plasma LOs from 20 patients with metastasis. Single-LO RNA-seq confirms bulk LO cargo, demonstrating the utility of single-cell technologies for large vesicle analysis. Our patient study provides proof-of-principle evidence that we can use multiomics to delve into EV heterogeneity, biogenesis, and composition. It also suggests that plasma LOs help stratify patients, supporting their potential prognostic value for developing a multi-analyte approach for liquid biopsy.

INTRODUCTION

Extracellular vesicles (EVs) are released by virtually all cell types and play a role in intercellular communication. EVs are heterogeneous in biogenesis, size, molecular composition, and biological function.[1–4] EVs originating through exocytosis from the endocytic machinery, historically called exosomes, are generally smaller than 150 nm[5]. EVs originating by direct budding from the plasma membrane (ectosomes), on the other hand, range dramatically in size and can be as small as exosomes or as large as microvesicles[2–5], migrasomes,[6] and large oncosomes (LOs)[7,8]. LOs are atypically large EVs (L-EVs)

(>1 μm in diameter) that are released by highly migratory and metastatic cancer cells undergoing ameboid motility.[8,9] We previously reported that LOs isolated from prostate cancer (PC) cells are stronger promoters of growth and tumor angiogenesis than exosomes and achieve this by an MYC-induced reprogramming of cancer-associated fibroblasts.[10] Evidence from an independent team indicates that LOs from hepatocellular carcinoma play a crucial role in bone metastasis by promoting an osteoclastic pre-metastatic niche.[11] Another team has reported that LOs from glioma stem cells promote disease progression via reprogramming the tumor microenvironment.[12] Additionally, our group and others have identified cancer-derived LOs, which are not detectable in human plasma of healthy controls, in the circulation of patients with advanced cancer.[2,9,10,13] Notably, LOs have been shown to be particularly abundant in the circulation of glioblastoma patients with short survival time.[12]

Most studies characterizing EVs have focused on small EVs (S-EVs).[14–17] While a few studies have acknowledged EV heterogeneity and used large-scale profiling to compare EV cargo across various cell types, they have not specifically focused on cancer-derived EVs.[18–20] There are no studies that integrate both the proteome and transcriptome across different EV populations. Additionally, the few reports that have integrated proteomics and transcriptomics in S-EVs have primarily focused on non-coding RNA using small libraries[19–23] rather than on mRNA. Moreover, although a significant effort has recently been made to profile the proteomes of various S-EV populations from human and murine samples across different cancer types,[17] this has never been done for LOs.

We and others have demonstrated that particles in the size range of LOs contain more abundant proteins and RNA and DNA cargo when compared to S-EVs.[3,7,13,24] The association of LOs with disease progression and poor outcomes, along with their content of multiple intracellular analytes, underscores the need for more in-depth investigation in this EV population. In the current study, we optimized the isolation of the L-EV fraction, specifically targeting the largest EVs, including LOs, and conducted a comparative profiling against the S-EV fraction. We conducted a comprehensive quantitative proteome analysis of L-EVs and S-EVs from three different cancer cell lines to identify proteins specific to each EV population. Next, we performed a whole-transcriptome analysis of EVs derived from PC cells, followed by an integrated analysis of protein and transcript levels. This approach aimed to identify molecules present at both the RNA and protein levels, making them strong candidates for future multi-analyte studies. LO-specific transcripts were then analyzed using single-vesicle RNA sequencing (RNA-seq), by adapting a single-cell RNA-seq pipeline for application to single LOs. Finally, sequential window acquisition of all theoretical mass spectra (SWATH-MS) was applied as a data-independent acquisition (DIA) mass spectrometry (MS) to LOs isolated from the plasma of PC patients. LO proteins were then used to query the Prostate Cancer Transcriptome Atlas (PCTA), a database comprising gene expression data from 1,321 clinical specimens from 38 independent PC cohorts to assess their correlation with PC progression.

RESULTS

LO yield can be efficiently increased at low-speed ultracentrifugation

For this study, we followed the guidelines from the International Society for Extracellular Vesicles.[25] We adapted a protocol, previously applied to human primary dendritic cells, which are benign cells, to isolate L-EVs from S-EVs[15] from cancer cells. The protocol employs differential ultracentrifugation (dUC) to pellet EVs according to their size and density, followed by flotation on a discontinuous iodixanol gradient. This upward gradient (EVs are deposited at the bottom of the tube) is known to enable purification of EVs based on their buoyant densities and separation of the EVs from non-vesicular components[15] (Figure S1A). Using PC3 cell culture medium, we identified a similar number of large particles in the size range of LOs (>1.5 μm) in the low (2.8K)- and medium (10K)-speed fractions (Figure S1B, left panel), but particles >4 μm were significantly more abundant in the 2.8K fraction (Figure S1B, right panel), indicating, as expected, that the largest particles sediment at the lowest speed. However, the low-speed fraction also contained a few cells (Figure S1C). Adding three low-speed centrifugation spins (500 × *g*) prior to pellet down the L-EVs at 2.8K resulted in a significant decrease of particles >10 μm (Figure S1D), but also in the elimination of residual cells (Figures S1E and S1F). Spike-in of an increasing number of cells in the LO preparations demonstrated a gradual increased expression of the intracellular protein GM130, which was otherwise negligible (Figure S1G). This suggests that even a minor contamination with cell would be noticeable during our quality control (QC) analysis of the EV preparations. Cell viability prior to EV isolation was close to 100%, indicating that these EVs were not shed by dying cells (Figure S1H). Additionally, the number of S-EVs in the 2.8K fraction was significantly lower when compared to 10K and 100K fractions (Figure S1I). Importantly, adding these extra steps to reduce cell contamination in the 2.8K pellet did not alter the number (Figure S1J) or size distribution of EVs as measured with two different technologies (electrical sensing zone and tunable resistive pulse sensing (Figure S1K).

To identify proteins that are consistently enriched in LOs and S-EVs across more than one cancer cell types, we rigorously isolated and characterized three EV fractions from PC (PC3), glioma (U87), and breast cancer (MDA-MB-231) cell lines. Although S-EVs were three orders of magnitude more abundant than L-EVs (Figure S2A), the total protein content was comparable between the two (Figure S2B), supporting the idea that L-EVs, with their larger volume, can accommodate a greater molecular cargo. Both the mean and mode of the size distribution of L-EVs were >1 μm, demonstrating the presence of a population of LOs across the three different cell lines (Figure S2C). The enrichment of L-EVs in the low- and medium-speed fractions and the enrichment of S-EVs in the high-speed fraction were also confirmed by transmission electron microscopy (Figure S2C). In summary, we found that a population of L-EVs can be efficiently separated from S-EVs. This study is focused on cancer-derived EVs; therefore, we will use predominantly the term LOs to refer to L-EVs.

LOs sediment in the low-speed fraction

MS analysis of the three fractions from the three cancer cell lines revealed that over 50% of the proteins identified in EVs were present across all EV fractions (Data S1: Label-free proteomics dataset). This finding suggests that a significant portion of the total EV protein cargo of the cell is released into all types of EVs, which all reflect, at some level, the molecular composition and biological function of the originating cell. The number of unique proteins was higher in the fractions collected at low speed and containing the largest EVs (Figure S3A). The normalized label-free quantitation (LFQ) signal among technical replicates exhibited >95% correlation (Figure S3B).

The fraction containing the largest EVs and the highest number of proteins also exhibited the greatest total protein abundance (Figure S3C). Comparison of ranked normalized protein intensities across the three fractions revealed that EVs isolated at medium speed are more likely to exhibit middle-ranking protein intensities than those collected at low and high speeds. This might reflect the presence of a mixture of L-EVs and S-EVs in the medium-speed fraction. In contrast, both the low-speed and medium-speed EVs contained subsets of proteins with both high and low abundances (Figure S3D), exhibiting similar patterns across the different cancer types. Hierarchical clustering revealed a more similar protein profile in the low-speed fraction, which is enriched for LOs, across the three cancer cell lines (Figure 1A).

Immunoblotting demonstrated enrichment of previously identified LO proteins, including from different cancer cell types, such as heat shock protein family A5 (HSPA5) and keratin 18 (KRT18)[7,13,26] in the low- and medium-speed fractions versus high-speed fractions. This finding validates that LOs can be enriched not only at medium speed, as previously reported by us and others,[2,3,8–10,13] but even more efficiently at a lower speed, in a fraction that was previously discarded (Figure 1B). In line with rigorous guidelines from the International Society for Extracellular Vesicles, the lack of the intracellular protein cytochrome *c*1 in the low-speed fraction confirmed the lack of contamination with cells.[25] Enrichment of TSG101, CD81, and CD9 in the high-speed fraction is consistent with the enrichment of S-EVs, including exosomes, in this fraction (Figure S3E). These proteins were excluded from the L-EV fractions. CAV-1 was detected in all EV populations, corroborating our previous finding that this protein, which ubiquitously populates all cell membranes, is present in all EV types.[5,26]

Next, we determined the relative abundance of known LOs or exosome proteins in the EV fractions from all three cancer cell lines. Proteomics analysis confirmed an enrichment of previously identified LO proteins, such as HSPD1 and HSPA5,[7,10,13] in low-speed versus high-speed fractions across board, highlighting the low-speed fraction as the best source for LOs (Figure 1C). Conversely, S-EV-enriched proteins such as CD9 and ITGB1 were most abundant in high-speed fractions, in line with previous reports.[27–29]

Finally, gene set enrichment analysis (GSEA) for previously identified proteins enriched in LOs versus S-EVs[7] across each EV fraction confirmed the enrichment of the LO protein cargo in medium- vs. high-speed fractions. Moreover, it revealed an even stronger enrichment of the LO proteins in low-speed fractions, particularly pronounced in EVs derived from PC cells (Figures 1D and S3E). Altogether, our results demonstrate that a population of LOs can be efficiently recovered at low centrifugation speed.

The LO fraction from three cancer cell lines is enriched in mitochondrial proteins

Unsupervised clustering identified 104 unique proteins in L-EVs and 17 unique proteins in S-EVs across all cancer cell lines, with no unique proteins identified in the medium-speed fraction (Figure 2A). Gene Ontology (GO) analysis revealed a significant enrichment of mitochondrial components, along with notable enrichment of endoplasmic reticulum (ER), nuclear, and ribosomal components (Figure 2B) in the L-EV fraction. The proteins uniquely identified in the S-EV fraction were mostly associated with GO terms such as plasma membrane and extracellular vesicles (Figure S4A), confirming enrichment of S-EVs likely including a population of exosomes in this fraction.[14–16]

Consistent with the unique protein data, also the majority of differentially expressed proteins were found to be more abundant in L-EVs across the three cancer cells types (Figure 2C). GSEA on differentially abundant proteins showed high enrichment for the GO cell component term Mitochondrion (GO:0005739) in the fraction with the largest EVs, in comparison with the other two, further suggesting that mitochondrial proteins, which can be found in all EVs, are enriched in LOs (Figure 2D).

To identify general markers of cancer-derived LOs and S-EVs, we focused on the differentially abundant proteins (false discovery rate [FDR]-adjusted p value from t test <0.05) that were detected on the highest quartile fraction at the high absolute level rank in either the low- or the high-speed fraction. The top 25% proteins in the low- or high-speed fraction, common to all three cancer cell lines, were nominated *bona fide* markers of cancer-derived LOs and S-EVs.

The majority of the proteins enriched in LOs were of mitochondrial origin (ATP5A1, ATP5B, HSPD1, LRPPRC, and SLC25A5) with exception of RPN2, which is an ER protein (Figures 2E and S4B). Conversely, cell membrane proteins DIP2B, ERBB2IP, LIN7C, and PLXNA1 and the secreted protein CASK were enriched in S-EVs. These proteins have been previously identified in S-EVs at both the protein and mRNA level[30,31]; they are deposited in the Vesiclepedia database (data not shown) and can be considered *bone fide* markers of S-EVs. Using a published MS-derived cell surface protein atlas,[32] we then integrated the current proteomic dataset with this "EV surfaceome" and identified several surface proteins in LOs and S-EVs from all three cancer cell types. The number of surface proteins was higher in S-EVs than in LOs (Figures 2F and 2G). This is consistent with a recent study that, using a mathematical model, reported a higher membrane-to-lumen ratio for S-EVs in comparison to L-EVs.[33] In summary, L-EVs including LOs are the EV fraction with more abundant cytosolic and organelle-derived proteins, while S-EVs are enriched in membrane proteins.

Proteins enriched in PC patient plasma LOs provide information about hormonal status

To identify the proteins that are consistently enriched in LOs from the plasma of PC patients with biologically aggressive disease, we applied a quantitative proteomic DIA method (SWATH-MS)

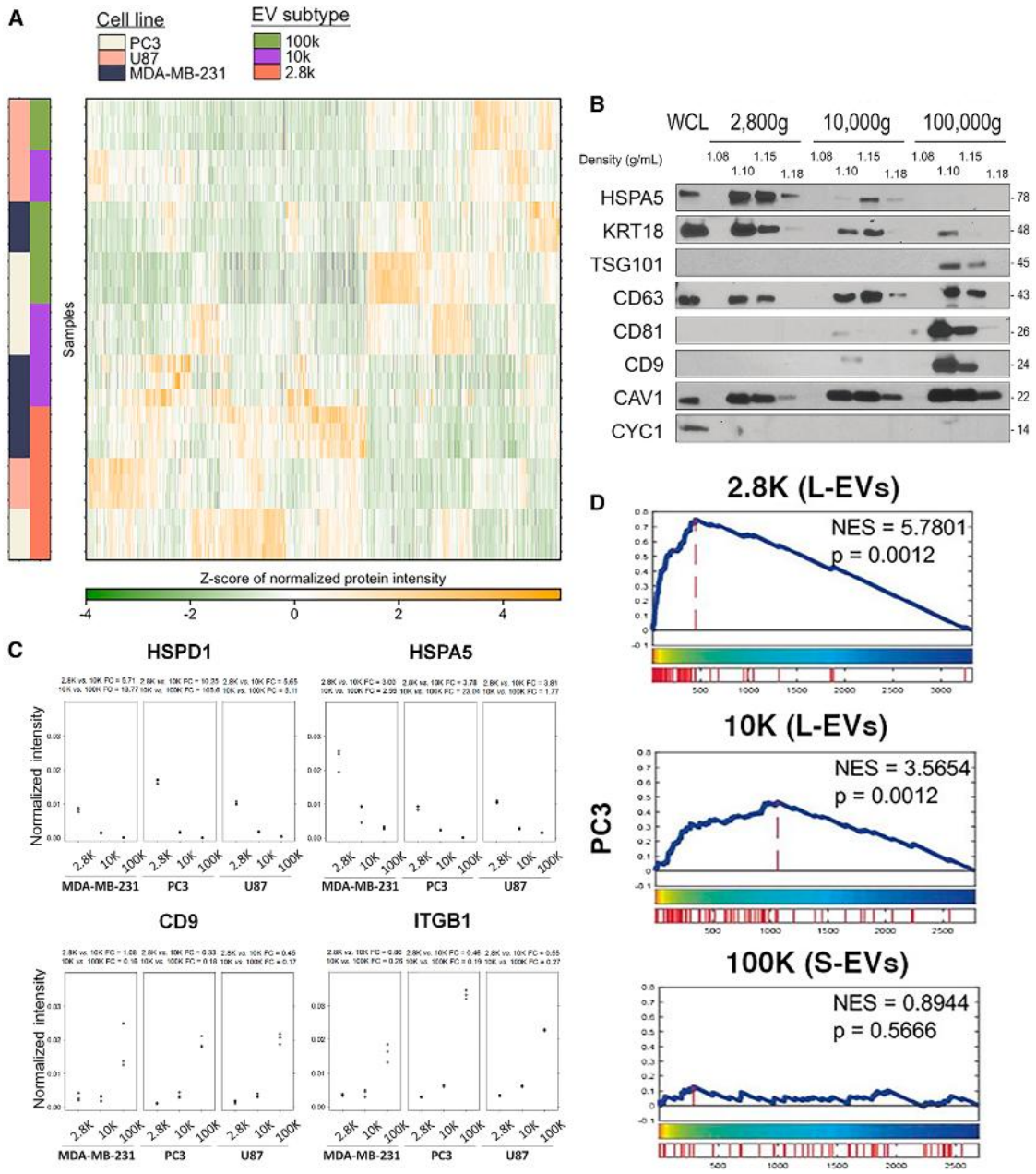


Figure 1. Large oncosome cargo is more evident in the low-speed than in the medium-speed fraction

(A) Divisive hierarchical clustering (DIANA) of the normalized relative abundance of the proteins identified across the indicated EV populations and cell lines. Pearson's correlation was used to calculate the correlation matrix.

(B) The three EV fractions and the whole-cell lysate (WCL) from PC3 cells were blotted with the indicated antibodies to characterize the presence of standard EV proteins in the different preps according to the minimal information for studies of extracellular vesicles (MISEV)[25] recommendations.

(C) Normalized relative abundance and fold change of the large EV (HSDP1 and HSPA5)- and small EV (CD9 and ITGB1)-enriched proteins across the indicated EV pellets and cell lines.

(D) The gene set enrichment analysis (GSEA) of the proteins differentially enriched in large oncosomes[7] (FC > 1, FDR < 0.05) across different EV populations from PC3 cell line. See also Figure S3.

to EVs from plasma samples of men with metastatic PC (n = 20), including castration-sensitive and castration-resistant patients. EVs were isolated using the same method described earlier for the cell line studies. With such a stringent approach, the S-EVs did not produce adequate amounts of protein for liquid chromatography-mass spectrometry (LC-MS) analysis by DIA MS. Therefore, we decided to apply SWATH-MS only on the LO fraction.[31]

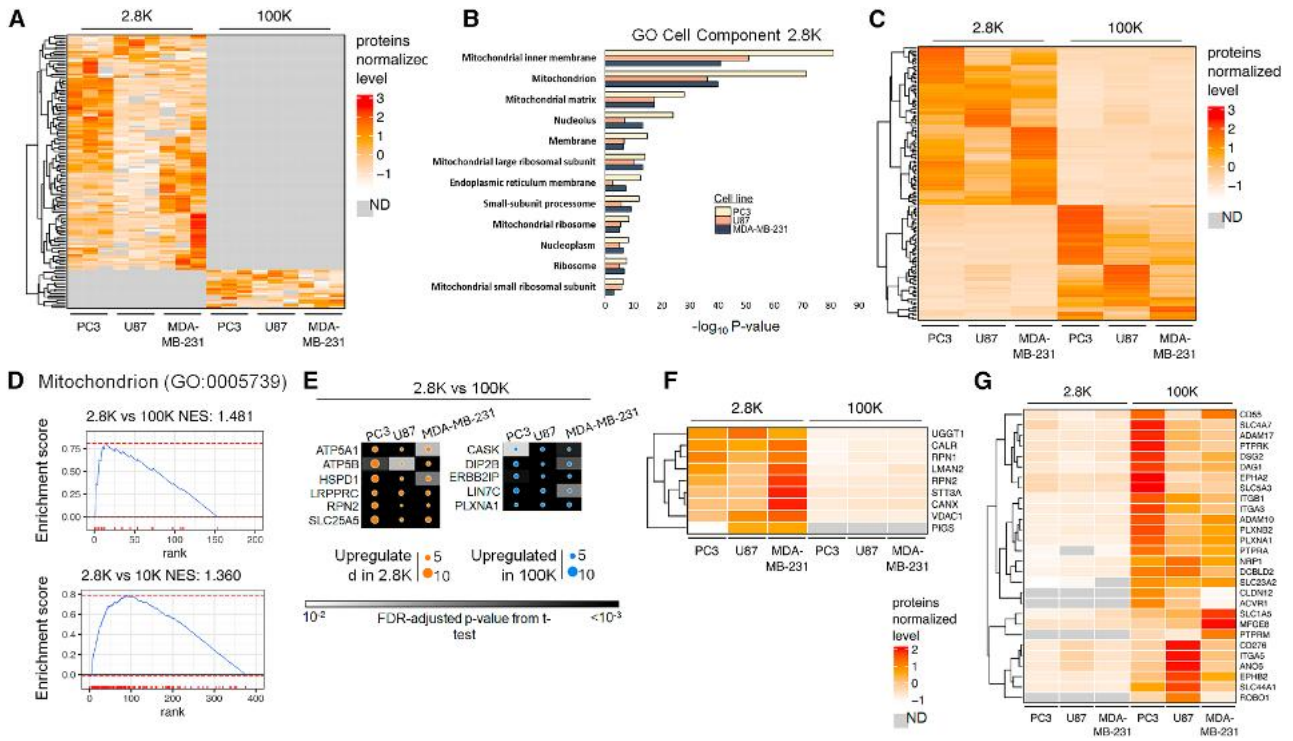


Figure 2. The low- and high-speed fractions contain the two most distinct L-EV and S-EV populations

(A) Heatmap showing normalized protein intensity of the proteins unique to the low- or high-speed EV fractions from the indicated cell lines. ND, not detected.

(B) Gene Ontology (GO) analysis of the proteins unique to the low-speed fraction of all cancer cell types demonstrates a robust representation of the cell component term mitochondria in L-EVs.

(C) Heatmap and dendrogram of differentially abundant proteins in L-EVs and S-EVs. ND, not detected.

(D) GSEA of the differentially abundant proteins using the term Mitochondrion (GO:0005739) showing that mitochondrial proteins are enriched in L-EVs.

(E) The most enriched (top 25%, fold change > upper quartile) and most highly expressed (top 25%, normalized expression > upper quartile) differentially abundant proteins in L-EVs and S-EVs common to all three cell lines.

(F and G). Surface proteins unique to or differentially abundant in either L-EVs (F) or S-EVs (G) and common to the PC3, U87, and MDA-MB-231 cell lines. ND, not detected. See also Figure S3. See also Figure S4.

We detected a total of 2,933 proteins, 56.6% of which were identified in all patients (Data S2: metastatic prostate cancer [mPC] plasma LO SWATH proteomic). The median number of proteins per sample was 2,630 ± 136. A subset of surface proteins that were significantly more abundant in LOs versus S-EVs from all the three cell lines (Figures 2F and 2G) was also abundant in LOs from the plasma patient specimens (Figure 3A). Additionally, an agglomerative hierarchical clustering analysis identified a set of proteins that were significantly upregulated in patients with hormone-resistant versus hormone-sensitive disease (t test $p \leq 0.05$) (Figure 3B). We further performed an analysis using an "in-house" prostate gene signature composed of 46 prostate-enriched genes identified using RNA-seq data from normal or normal adjacent to tumor prostate samples from the Genotype-Tissue Expression project and The Cancer Genome Atlas.[34] Among those, 6 proteins were enriched in L-EVs, suggesting L-EVs are a good surrogate of prostate tissue in the circulation of patients (Figure 3C).

In order to test if LO enrichment is essential for protein detection and improvement of signal-to-noise ratio in a complex biofluid such as the plasma, we performed a palmitoylated label-free MS analysis on three matched LOs and unfractionated plasma samples, randomly selected from the aforementioned cohort. We detected a total of 1,415 protein identities among the samples (Data S3: MS of LOs and unfractionated plasma), with 1,063 (75.1%) being detected in purified LOs. Out of all the proteins identified, 725 (51.2%) were detected only in purified LOs but not in unfractionated plasma (Figure 3D and S4C), suggesting that LO enrichment increases the identification of cancer cells-derived proteins that are not otherwise quantifiable in unfractionated plasma.

Using the same surface proteins atlas described earlier,[32] we assembled a list of candidate surface proteins for PC-derived LOs detected not only in LOs from PC3 cell line but also in LOs from patient plasma. In addition, we identified surface proteins detected in LOs from all PC patients, regardless of their response status, and expressed at relatively high levels. As these proteins were identified in all PC patients, they could be leveraged for specific isolation of PC-derived LOs from patient plasma by immune-affinity approaches. Among the total of 302 surface proteins in the dataset, we identified 39 candidates (Table S1) whose relative expression was above the median.

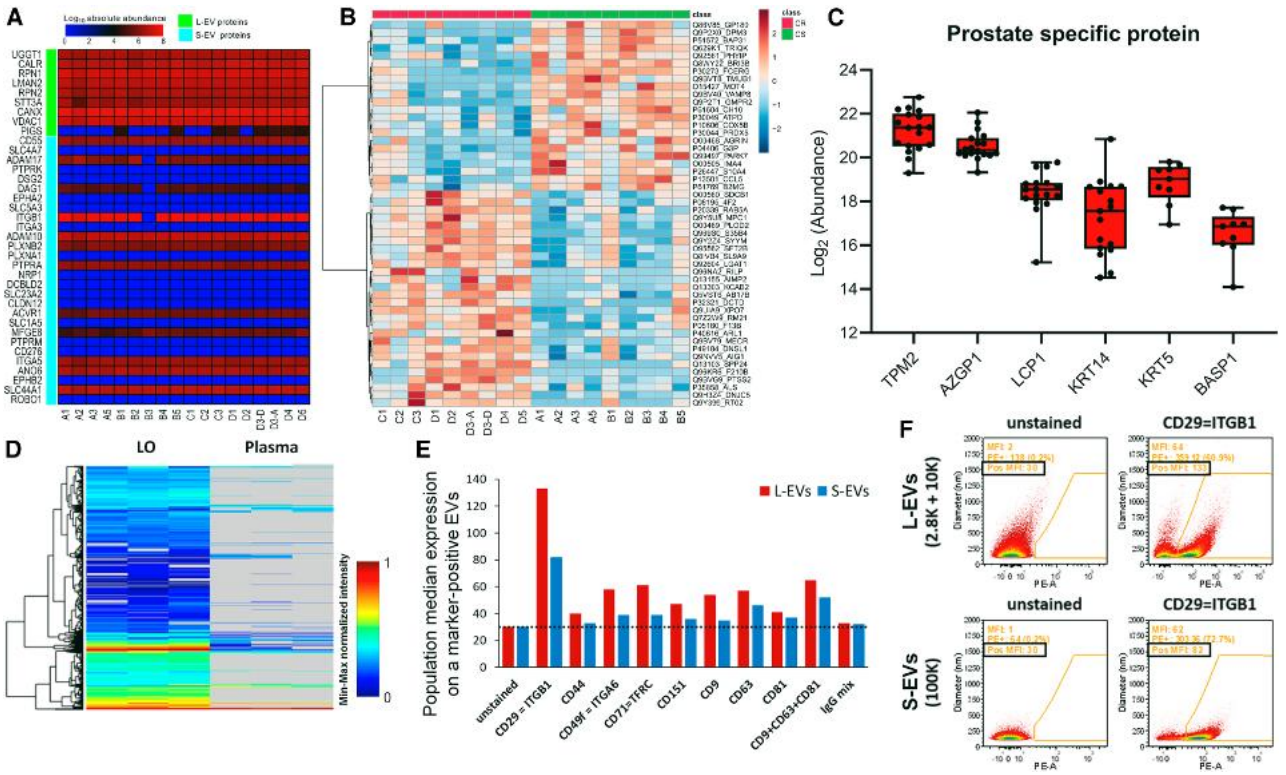


Figure 3. L-EV markers might be prognostic and predictive of treatment outcomes in metastatic prostate cancer patients

(A) Heatmap of the highly abundant surface proteins differentially expressed in either L-EVs or S-EVs identified in all three cell lines described in Figures 2F and 2G present in the LOs of metastatic PC patients.

(B) Agglomerative hierarchical clustering analysis using the Euclidean distance and Ward's linkage of the proteins differentially expressed (p value < 0.05) in plasma L-EVs from patients with castration-resistant or castration-sensitive PC.

(C) Comparison of $\log_2$ (abundance) of prostate-enriched gene proteins[34] across all samples from plasma patients-derived LOs. The plot displays all the proteins detected in our patient-derived LOs among these 46 prostate-enriched genes. Boxplots show $\log_2$ (abundance) median and interquartile range. Error bars represent min and max values.

(D) Heatmap of normalized LFQ intensity showing the proteins identified in LOs and in unfractionated plasma from the palmitoylated MS, showing that the majority of palm-proteins identified in purified LOs were not identified in unfractionated plasma.

(E and F) Single-EV flow cytometry analysis of the select proteins expressed on the EV surface. (E) Quantification of the expression of the indicated proteins in molecules per EV. (F) Representative dot plots showing the size of the EVs and the expression level of the target proteins.

CR, castration resistant; CS, castration sensitive.

Finally, we used single-vesicle flow cytometry to estimate the abundance of some of the surface proteins identified in both patient-derived EVs and PC3 cell-derived EVs. Using validated antibodies with a calibrated flow cytometer, we found that, in line with MS, cargo abundance was higher on LOs than S-EVs, even at the single-vesicle level. The median protein abundance was also higher in LOs than S-EVs owing to the higher total protein amount per vesicle in LOs (Figures 3E and 3F). Notably, protein expression was heterogeneous among EVs, with a fraction of EVs showing measurable antibody staining, while other EVs showed no detectable antigen. The measured single EV antigen expression was generally concordant with MS-based abundance rankings, with a few exceptions (e.g., CD44 and CD9), pointing to the importance of employing orthogonal approaches and understanding the quantitative limits of different methods. In summary, a subset of proteins enriched in plasma LOs from metastatic PC patients seems to provide information about castration resistance.

Distinct RNA composition of LOs versus S-EVs

We then performed a comprehensive, comparative analysis of the whole transcriptome of LOs and S-EVs using the PC3 cell line. Normalized counts among the technical replicates for each sample were highly correlated (Figure S5A and Data S4: PC3 EVS RNA-seq). We found that the length of the RNA transcripts (transcripts per million, TPM > 1) with >80% coverage belonging to different biotypes was nearly identical in the PC3 cells and in all three EV populations. However, the frequency of transcripts, which was lower in EVs compared to the cells, was slightly higher in the LO-enriched fraction (Figure S5B).

Most of the transcripts (10,074) were detected in all three EV populations and originating cells (Figure 4A). In agreement with the proteomic data, each EV fraction had a set of unique transcripts and a set of transcripts shared with other fractions, and unique transcripts were more abundant in the low- and high-speed fractions (Figure 4A). Principal-component analysis

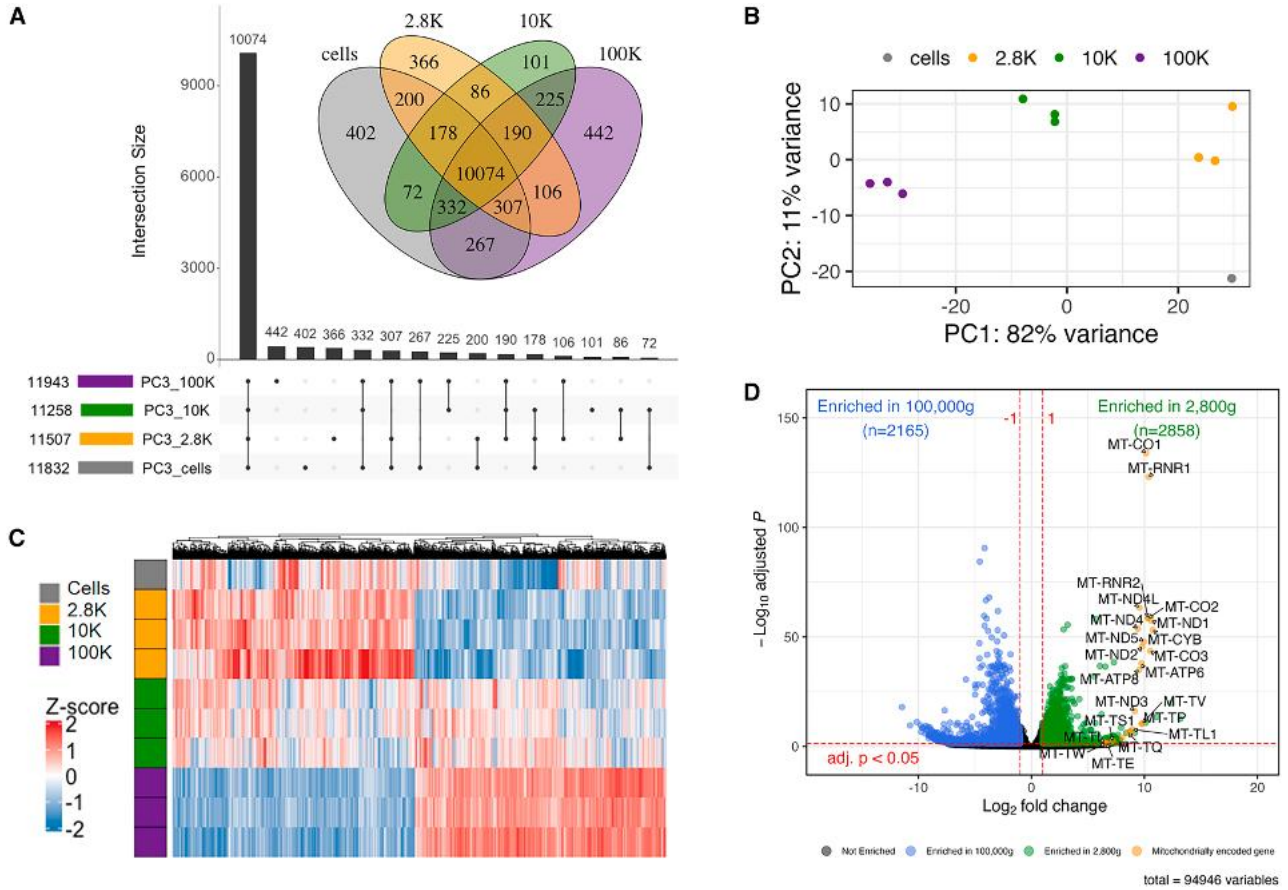


Figure 4. Whole transcriptome analysis identifies distinct RNA cargo in the 2.8K and 100K PC3 EVs

(A) Upset plot of the genes identified in the PC3 cell line (DESeq2 normalized counts > 10) between EV fractions and source cells.

(B) PCA plot of replicates showing grouping by triplicate.

(C) Heatmap (unsupervised clustering) of *Z*-scored gene expression of differentially expressed genes between the 2.8K and 100K fractions.

(D) Volcano plot of differentially expressed genes (adjusted $p < 0.05$, log2 fold change > 1, base mean > 10) the 2,800*g* (green) and 100,000*g* fractions (blue). See also Figure S5.

(PCA) demonstrated that the three EV populations clustered separately from each other and from their producing cells (Figure 4B), suggesting that the RNA shed in EVs is not a mere representation of the cellular RNA. Unsupervised clustering analysis of the transcripts differentially enriched in the three EV populations (relative to each other) revealed the most striking differences in the low- and high-speed EV fractions, while the medium-speed EV fraction was less distinct and appeared transitional between LO and S-EV profiles (Figure 4C). This was in line with the proteomic data (Figure 1A). Interestingly, both PCA and clustering analysis showed that the mRNAs of the originating cells were more similar to those enriched in the low-speed fractions compared to the other two fractions (Figures 4B and 4C). This is in line with the concept that LOs contain several structural cytosolic and organelle-derived molecules that reflect the cell of origin better than S-EVs. In line with proteomics, mitochondrial transcripts were mostly enriched (Figure 4D) in the L-EV fraction. In summary, whole-transcriptome analysis revealed that the low- and high-speed fractions contain the most distinct EV populations also at the RNA level and that LOs are enriched in mitochondrial transcripts, in line with the proteomic analysis.

Single-EV RNA-seq analysis of LOs confirms the abundance of mitochondrial transcripts

To further investigate differential enrichment of mitochondrial transcripts in LOs vs. S-EVs, we performed GSEA using the GO cell component category for Mitochondrion (GO:0005739). The analysis showed that the low- and medium-speed fractions, which both contain L-EVs, were enriched in mitochondrial transcripts in comparison to the high-speed fraction (Figure 5A, top and middle). However, a higher enrichment of mitochondrial transcripts was found in the medium-speed than in the low-speed fraction (Figure 5A, bottom). To determine if these mitochondrial transcripts are measurable in single LOs obtained from PC3 cells, we analyzed this fraction using a Chromium 10× technology. To enable encapsulation of single LOs into individual droplets, the EV sample was diluted to 1:10,000, as typically done for cells (Data S5: single PC3 LO RNA-seq).

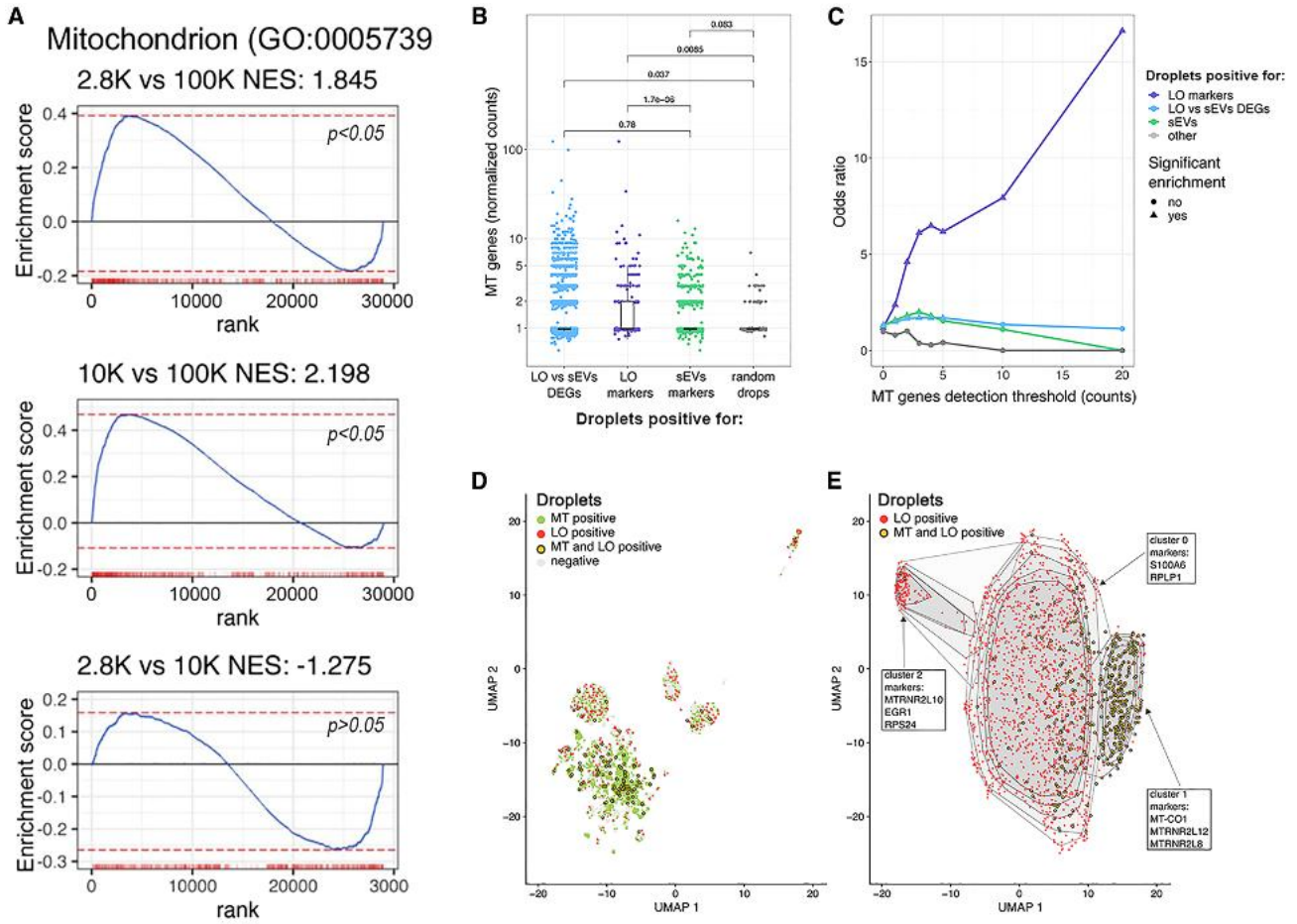


Figure 5. Single-EV RNA-seq analysis of L-EVs confirms the abundance of mitochondrial transcripts

(A) GSEA using the GO cell component category for Mitochondrion (GO:0005739) showing an enrichment of mitochondrial genes in both 2.8K and 10K when compared to the 100K (top and middle) and a higher enrichment of mitochondrial genes in the 10K fraction when compared to the 2.8K. Genes are ranked based on Wald's test statistic of differential expression analysis.

(B) Expression level of mitochondrial (MT) genes in droplets positive for LOs vs. S-EV differentially expressed genes (DEGs), LO markers, S-EV markers, and a set of randomly selected droplets. Statistical significance is obtained with Mann-Whitney two-tailed test.

(C) Co-occurrence analysis of MT reads and reads of genes classified as (1) LOs vs. S-EV DEGs, (2) LO markers, and (3) S-EVs markers. *x* axis represents the number of reads requested to classify a drop as "MT positive." *y* axis represents the odds ratio of each Fisher's test, testing the independence of the positivity for MT markers ("MT positive") and the positivity for other markers ("LO markers," "LO vs. sEVs DEGs," "sEVs," and "other"). A positive odds ratio indicates enrichment for EVs positive for MT markers. Significance is shown by the shape of each point, with a threshold of 0.05.

(D) Dimensionality reduction of single-EV RNA sequencing data (only drops containing high number [≥50] of reads are shown). Drops are colored based on the presence of MT and LO transcripts. MT-LO double-positive vesicles (yellow) are observed.

(E) Dimensionality reduction of single-EV RNA sequencing data limited to vesicles positive for LO markers (only vesicles containing high number [≥50] of reads and positive for LO markers are shown). Drops are colored based on the presence of MT and LO transcripts. Clustering analysis identifies 3 clusters. Convex hulls indicate area containing 95%, 90%, 80%, 70%, and 60% of points of each cluster. Significant markers of each cluster are show in text boxes.

Because there are no methods available for the analysis of single-EV sequencing data, we did not use standard data filters but rather applied custom thresholds. Common quality control techniques would, in fact, filter out the signal from single EVs since they contain fewer transcripts than a single cell (Table S2A). From the bulk RNA-seq data, we first identified genes only identified at the RNA level in LOs (LO markers), only in S-EVs (S-EV markers), or differentially abundant between the two groups ($p < 0.05$; Benjamini and Hochberg adjusted test; LOs vs. S-EV DEGs). We then classified single EVs based on the presence of reads of LO markers, S-EV markers, and LOs vs. S-EV DEGs (Figures 5B–5D) and finally investigated the presence of reads associated to mitochondrial transcripts. Globally, 16.3% of droplets contained at least one mitochondrial read (Table S2B). We observed that the number of normalized counts associated with mitochondrial transcripts was higher in LOs than in S-EVs (Mann-Whitney two-tailed test, $p = 1.7 \times 10^{-6}$), thus confirming the observed enrichment of mitochondrial signal in LOs (Figure 2D). To confirm this hypothesis at the single-EV level, we performed an enrichment analysis to observe the co-occurrence

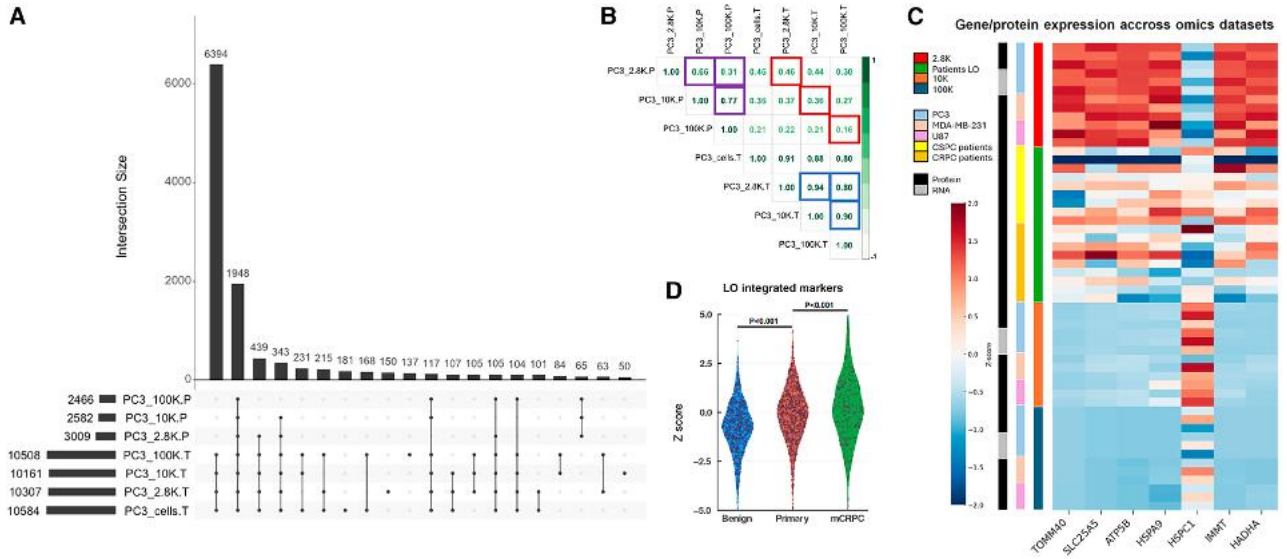


Figure 6. Comparison of protein and mRNA cargo in EV populations in the PC3 cell line

(A) Upset plot of protein-coding genes identified in the proteome (P) and transcriptome (T; DESeq2 normalized counts > 10).

(B) Spearman's rank correlation of genes expressed in both the proteome and the transcriptome. Blue boxes: correlation of different sized EVs within the transcriptome or proteome. Red boxes: correlation of corresponding fractions between the proteome and transcriptome.

(C) Heatmap of the LO-enriched analytes derived from the intersection of protein profiles of LOs obtained from different cancer cell lines with the mRNA profile of LOs obtained from the PC cell lines and with the protein profile of LOs from PC patient plasma.

(D) PCTA analysis shows that the proteins and mRNAs enriched in the PC-derived 2.8K are significantly upregulated in patients with primary cancer vs. cancer-free individuals and are further upregulated in patients with mCRPC vs. primary cancer.

between LO reads and mitochondrial reads. We observed that, independently of the threshold used for the detection of mitochondrial signal (based on the number of reads associated with mitochondrial genes in each EV), droplets positive for LO signal showed a significant enrichment for mitochondrial transcripts (Figure 5C). Uniform Manifold Approximation and Projection (UMAP) analysis allowed the detection of different clusters (Figure S6A and Table S2C) of single EVs and the widespread presence of EVs positive for both LOs and MT markers (Figure 5D). When analyzing only the droplets positive for LO markers, we identified a cluster enriched for MT-LO double-positive droplets (Figures 5E and S6B). Interestingly, the gene MT-CO1 (mitochondrially encoded cytochrome *c* oxidase I) emerged as a marker of the MT-LO-enriched cluster (Table S2D). The list of genes used for clustering is described in Table S2E. In summary, LOs are suitable for single-EV RNA-seq, which confirmed enrichment of mitochondrial transcripts in LOs.

LO protein/mRNA levels correlate with cancer progression

Next, to integrate proteomic and transcriptomic profiles of L-EVs and S-EVs, we correlated the RNA transcripts and their correspondent encoded proteins within each EV population as well as between different EV populations. To achieve this, we identified the protein-coding RNAs that also had corresponding proteins in our datasets (Figure 6A). Most mRNAs (8,070 out of 11,241) were identified only in the transcriptome and not in the proteome (Figure S5C), with only 1,948 genes identified both as mRNA and as protein in all three fractions. To determine the EV fraction distribution of transcripts and proteins, we focused on these 1,948 protein-coding mRNAs found at the intersection of the transcriptome and proteome data (Figure 6A). The Spearman's rank correlation between proteins and transcripts in different EV fractions was much higher for transcripts than proteins (Figure 6B, blue vs. purple squares). It was also higher for L-EVs, especially in the 2.8K fraction, than in S-EVs (Figure 6B, red frames), in line with the idea that EVs with a larger volume might accommodate both protein and mRNA for the same molecules. Additionally, the mRNAs in the low-speed fraction were slightly more similar to the mRNAs in the medium speed ($r = 0.94$) compared to the high-speed fraction ($r = 0.80$) (Figure 6B, blue frames). Conversely, the proteins in the medium- and high-speed fractions were more similar to each other ($r = 0.77$) than the proteins in the low- and medium-speed fractions (0.66) (Figure 6B, purple frames).

Lastly, we generated an integrated list of LOs *bona fide* candidate markers by selecting proteins enriched in LOs both from cultured cells and from the plasma of PC patients (Data S1, S2, and S4). Then, we narrowed down this subset of proteins to the ones that had also high levels of the correspondent mRNA, in LOs. Surprisingly, these are all mitochondrial molecules: TOMM40, SLC25A5, ATP5B, HSPA9, HSPC1, IMMT, and HADHA (Figure 6C). To explore the potential association of LO-enriched cargo with clinical outcomes, we used the aforementioned list of targets to interrogate the PCTA (http://www.thepcta.org/),[35,36] a large transcriptome database comprising 1,321 clinical specimens from 38 PC cohorts. This query returned significantly upregulated signal in primary PC versus

Table 1. Patient characteristics: castration-sensitive and castration-resistant prostate cancer cohort. PSA, serum PSA concentration; RT, radiation therapy; ADT, androgen deprivation therapy; LN, lymph nodes; RP, radical prostatectomy

ID	Treatment	TTP (mo)	Metastatic burden	RP (Y/N)	Radiation	Previous systemic treatment	PSA (ng/mL)
CSPC – treatment responsive							
A1	ADT	78.2	none (prostatic bed recurrence by PSMA)	Y	salvage RT	none	0.1
A2	ADT + docetaxel	101	high (heavy bone + liver)	Y	salvage RT + metastasis directed RT to bone	none	0.03
A3	ADT + docetaxel	75	high (heavy bone + lung)	N	none prior to blood samples	none	0.4
A4	ADT + abiraterone	50	high (bone)	N	palliative for cord compression	none	0.04
A5	ADT + abiraterone	60	high (bone)	N	none	none	0.14
CSPC – treatment resistant							
B1	ADT	8	low (LN)	N	none	none	0.6
B2	ADT	12	high (LN + bone)	N	none	none	1.3
B3	ADT	8.3	high (bone)	N	primary RT + metastasis-directed RT	none	4.61
B4	ADT + docetaxel	8.7	high (bone + LN)	N	palliative RT to femur	none	0.45
B5	ADT + apalutamide	2.8	high (LN + bone)	N	primary RT + ADT	none	25.98
CRPC – treatment responsive							
C1	enzalutamide	86.5	none (biochemical disease only)	Y	adjuvant RT	none	0.06
C2	enzalutamide	58	low (LN only)	Y	adjuvant RT	none	0.03
C3	abiraterone	49	low (bone)	Y	salvage RT +ADT + MDT	none (progressed during ADT + RT)	0.03
C4	bicalutamide	49	low (bone)	N	primary RT + salvage HD RT (no ADT) + MDT	sipuleucel-T, bicalutamide, experimental stem cell therapy (China)	1.18
CRPC – treatment resistant							
D1	abiraterone	0.9	high (lung + bone)	Y	salvage RT (no ADT) MDT	docetaxel, enzalutamide, MDT (bone), carotuximab + enzalutamide, abiraterone	16.83
D2	enzalutamide	0.9	high (bone)	Y	salvage RT	bicalutamide, apalutamide + abiraterone, MDT (bone), enzalutamide	16.0
D3-D	docetaxel	0.9	high (bone)	N	primary RT + ADT	bicalutamide, apalutamide, docetaxel	71.91
D3-A	apalutamide	7.6	high (Bone)	N	primary RT + ADT	bicalutamide, apalutamide, docetaxel	0.50
D4	docetaxel	5.2	high (bone + LN)	N	none	enzalutamide, docetaxel	4.91
D5	docetaxel	1.8	high (LN + bone)	N	primary RT + ADT	none	774.17

benign prostate tissues and in metastatic castration-resistant PC (mCRPC) versus organ-confined disease (Figure 6D), suggesting that the circulating LOs carry information with potential prognostic value.

DISCUSSION

This is a first-in-field study in which DIA MS was applied to EVs, in this case LOs, rigorously isolated from patient plasma. Indeed, few studies have utilized MS on patient plasma. The subset of proteins identified here was consistently found highly abundant in LOs from different cancer cell lines; they were present also as mRNA in LOs from the PC model and were also abundant in plasma LOs from all 20 patients with metastatic PC.

The current literature on LO molecular profiling is scarce. L-EVs, in general, have been neglected till recently. However, L-EVs are as abundant and as important as S-EVs,[37] and we now know that different EV populations might play distinct functions and report different types of information in clinical samples. This has contributed to a strong interest in LO profiling to identify circulating cancer markers. One of the most commonly used methods for EV isolation is dUC,[25] which sediments L-EVs and S-EVs at medium- and high-speed centrifugation, respectively. Kowal et al.[15] were the first to report the presence of L-EVs at such a low-speed centrifugation. To fully remove the cells, they added extra centrifugation steps, a process that we have implemented in this study. Even though these EVs were <400 nm and therefore not comparable to LOs in size, and even though they were isolated from benign cells (dendritic cells), we reasoned that, perhaps, also LOs would sediment at low speed. Our data indicate that even larger EVs, in the size range of LOs (1–10 μm), indeed mostly sediment at low-speed centrifugation.

Our study suggests that LOs sedimented in low-speed fractions from different cancer cell lines are more similar to each other than smaller EVs isolated from the same cancer cells. In addition, those L-EVs are enriched with cytosolic proteins including organelle-derived proteins and metabolic enzymes, with an even higher enrichment for mitochondrial components, at both protein and mRNA levels. Similar to our data, both Cocozza et al.[38] and Jimenez et al.[20] demonstrated a similar enrichment trend for mitochondrial proteins identified in medium- vs. high-speed fractions in murine and human cancer cells lines, confirming that mitochondrial proteins are associated with L-EVs. Recent studies reported the presence of functional mitochondrial material in S-EVs and that this can be transferred from cell to cell in various models.[39–43] Our results indicate that mitochondrial proteins and mRNAs are significantly more abundant in L-EVs than in S-EVs. We were able to confirm this result by single-EV RNA-seq that we performed to leverage the large volume of LOs. The observed significant enrichment of mitochondrial reads only in the droplets containing LO-specific reads suggests that single-EV RNA-seq is feasible, at least for L-EVs. These results might simply reflect the larger volume of LOs, which could randomly incorporate organelles during plasma membrane blebbing. In line with that, a recent study showed functional mitochondria in blebbisomes, another extremely L-EV type with cell-autonomous functions.[44,45] Whether mitochondria are functional in LOs is currently unknown but is the object of investigation.

The finding that the transcriptome profile of LOs is more similar to the originating cells than that of S-EVs suggests that LOs reflect the cellular transcriptome better than S-EVs. This is corroborated by the higher abundance of RNA transcripts with over 80% coverage in L-EVs compared to S-EVs. This is in line with a study reporting that, while EV transcripts are overall shorter than intracellular transcripts, the largest transcripts are more abundant in L-EVs than S-EVs.[46] Additionally, the correlation between mRNA and protein, which is relatively poor even in cells and tissues,[47] was higher in LOs than in S-EVs. On the other hand, the result that more than half of the proteins identified in EVs are present in all EV fractions suggests that most EV protein cargo is shed into all EV types indiscriminately. However, this could also be an indication of the limitations of dUC in separating different classes of EVs.[48,49] Therefore, even though LOs emerge as an ideal source of circulating biomarkers to be probed with multi-analyte platforms, new methodologies for EV isolation that can be easily applied to the clinic are needed. The surface proteins/mRNAs identified in LOs obtained not only from cancer cell lines but also from cancer patient plasma are solid and suitable candidates for development of affinity purification methods. Additionally, the use of highly specific surface proteins for immune capture of LOs from patient plasma effectively addresses a major challenge in the blood EV field—platelet contamination—which is notoriously difficult to eliminate. Immune capture can be seamlessly implemented in clinical laboratories to isolate LOs for prognostic and diagnostic purposes without requiring additional purification steps. In that regard, this study represents a key step in the liquid biopsy field as a proof-of-concept experiment. We received National Institutes of Health funds to interrogate a large cohort of PC patients and are in the process to request more funds to extend our studies to different cancer types.

Our approach to intersect cell line and patient plasma data enabled identification of a robust subset of proteins that are enriched in LOs from different sources. Our findings in the plasma of cancer patients show an association of LOs proteins with castration resistance. They also reveal a significant correlation between expression levels of the mRNA codifying for these proteins and metastasis. For these experiments, we used a well-curated collection of gene expression data from over 1,000 clinical specimens from multiple independent PC cohorts (PCTA),[35,36] highlight the potential value of LOs in liquid biopsy. Importantly, our identified set of markers can be used to interrogate clinical specimens by microfluidics, flow cytometry, or other antibody-based technologies that aim to select a population of EVs with a distinct role in cancer.

In summary, this study is a profile of LOs by parallel multiomics and represents a key step in the liquid biopsy field. This study provides a quantitative platform of LO and S-EV cargo from different cancer cell types and of LOs obtained from the plasma of patients with metastatic PC. It also provides a promising avenue for future research into developing better diagnostic and therapeutic tools for cancer.

Limitations of the study

One of the main limitations of this study is the limited statistical power of the patient cohort investigated here. This reflects the

inherent challenge of working with human plasma and Clinical Laboratory Improvement Amendments (CLIA)-laboratory incompatible, technically demanding ultra-purification methods that pose a barrier to clinical translation. However, the highly reliable LO proteins identified here and derived from the intersection between independent sets of cell lines and cancer patient data, as well as protein and RNA profiles, will be validated in larger, multicenter cohorts to facilitate clinical application. For this purpose, we were recently awarded a 5-year R01 (R01CA218526) to validate the LO proteins in large and independent cohorts using CLIA-compliant methodologies.

An additional limitation is the likely presence of platelets in the plasma preps. Platelets are the major source of contamination for any type of plasma-derived EVs. Despite efforts to minimize platelet interference in EV studies, complete removal remains challenging. However, the subset of LO proteins identified here will be used to develop immune-capture-based technologies, with or without negative selection with platelet-specific markers such as CD31, CD41, CD61, and GP1BB.

Despite strong evidence, both at protein and mRNA levels, that mitochondrial content is higher in LOs than S-EVs, we have to keep in mind the technical limitations inherent to the EV fractionation. S-EVs were isolated after a low-speed depletion step that might contain free mitochondria besides LOs and other L-EVs.

One limitation of the single-LO RNA-seq study is that LOs were not captured as single EVs from single cells but rather diluted to 1:10,000 as for single-cell RNA-seq. Therefore, it is not possible to rule out the presence of apoptotic bodies or the contamination from S-EVs in droplets containing LOs.

Despite these limitations, our findings provide valuable insights into LO composition and lay the groundwork for future advancements in EV-based biomarker discovery. Other teams of investigators will be able to explore the molecules identified here in patients with PC, while also possibly expanding this to other cancers.

RESOURCE AVAILABILITY

Lead contact

Requests for further information, resources, and reagents should be directed to and will be fulfilled by the lead contact, Dolores Di Vizio (Dolores.DiVizio@cshs.org).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The label-free proteomics (Pride: PXD038518; analyzed data available as Data S1) and the SWATH-MS datasets (Pride: PXD038011; analyzed data available as Data S2) were deposited on Proteomics Identifications Database (PRIDE) and are available through https://www.ebi.ac.uk/pride/. Palmitoylated proteomics of LOs and unfractionated human plasma dataset are available in supplementary data as Data S3. Both bulk RNA-seq (GEO: GSE214804; analyzed data available as Data S4) and single-EV RNA-seq datasets (GEO: GSE231846; analyzed data available as Data S5) were deposited on Gene Expression Omnibus (GEO) database and are available through https://www.ncbi.nlm.nih.gov/geo/.
- This study did not involve the reporting of novel or original code.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the Center Proteomics and Metabolomics Shared Resource of Cedars-Sinai Cancer Center.

This work was supported by grants from the National Institutes of Health (R01CA218526 to D.D.V., R01CA234557 to D.D.V., U2CCA271894 to P.C.B., U24CA248265 to P.C.B., and R01CA270108 to P.C.B.) and by Cancer Research UK and Fondazione AIRC per la ricerca sul cancro ETS: Accelerator Award 20218 (A22792 to F.D.).

AUTHOR CONTRIBUTIONS

Conceptualization, T.F.S. and D.D.V.; methodology and validation, T.F.S., E.H., W.Z., Y.C., E.K., Z.Q., M.K., S.Y., J.M., F.B., F.D., P.C.B., and K.V.K.-J.; formal analysis, E.H., W.Z., Y.C., E.K., Z.Q., M.K., and S.Y.; investigation, T.F.S., J.M., A.K., J.N., S.P., M.H., F.G., R.R., B.Z., Y.W., and W.Y.; resources, E.P.; writing – original draft, T.F.S., E.H., W.Z., and D.D.V.; writing – review and editing, M.R.F., J.V.E., J.G., C.T., A.Z., S.S., F.D., P.C.B., and K.V.K.-J.; visualization, T.F.S. and D.D.V.; supervision, P.C.B., K.V.K.-J., and D.D.V.; project administration, D.D.V.; funding acquisition, D.D.V.

DECLARATION OF INTERESTS

P.C.B. sits on the Scientific Advisory Board of Intersect Diagnostics Inc. and previously sat on those of BioSymetrics Inc. and Sage Bionetworks.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Cell culture
  - Human subjects
- METHOD DETAILS
  - Enrichment of EVs from cell culture media
  - Enrichment of large EVs (L-EVs) from patient plasma
  - Whole cell protein lysates
  - Immunoblotting
  - Tunable Resistive Pulse Sensing (TRPS) measurements
  - Electrical Sensing Zone (ESZ) measurements
  - Transmission electron microscopy (TEM)
  - Vesicle flow cytometry
  - Label-free proteomic profiling of cell line-derived EVs
  - SWATH-MS instrumentation of patient-derived L-EVs
  - LB-ABE enrichment of palmitoyl-proteins
  - Bulk RNA-seq instrumentation
  - Single-EV RNA-seq instrumentation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Label-free proteomic analysis
  - Normalization
  - Statistical analysis
  - SWATH-MS analysis of patient-derived L-EVs
  - LB-ABE LC-MS/MS analysis and data processing
  - Bulk RNA-seq analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2025.102161.

Received: March 11, 2024
Revised: November 19, 2024
Accepted: May 8, 2025
Published: June 3, 2025

REFERENCES

1. Couch, Y., Buzàs, E.I., Vizio, D.D., Gho, Y.S., Harrison, P., Hill, A.F., Lötvall, J., Raposo, G., Stahl, P.D., Théry, C., et al. (2021). A brief history of nearly EV-erything - The rise and rise of extracellular vesicles. J. Extracell. Vesicles *10*, e12144. https://doi.org/10.1002/JEV2.12144.
2. Gerdtsson, A.S., Setayesh, S.M., Malihi, P.D., Ruiz, C., Carlsson, A., Nevarez, R., Matsumoto, N., Gerdtsson, E., Zurita, A., Logothetis, C., et al. (2021). Large Extracellular Vesicle Characterization and Association with Circulating Tumor Cells in Metastatic Castrate Resistant Prostate Cancer. Cancers (Basel) *13*, 1056. https://doi.org/10.3390/CANCERS13051056.
3. Yekula, A., Minciacchi, V.R., Morello, M., Shao, H., Park, Y., Zhang, X., Muralidharan, K., Freeman, M.R., Weissleder, R., Lee, H., et al. (2020). Large and small extracellular vesicles released by glioma cells in vitro and in vivo. J. Extracell. Vesicles *9*, 1689784. https://doi.org/10.1080/20013078.2019.1689784.
4. Setayesh, S.M., Hart, O., Naghdloo, A., Higa, N., Nieva, J., Lu, J., Hwang, S., Wilkinson, K., Kidd, M., Anderson, A., et al. (2022). Multianalyte liquid biopsy to aid the diagnostic workup of breast cancer. NPJ Breast Cancer *8*, 112. https://doi.org/10.1038/S41523-022-00480-4.
5. Zijlstra, A., and Di Vizio, D. (2018). Size matters in nanoscale communication. Nat. Cell Biol. *20*, 228–230. https://doi.org/10.1038/S41556-018-0049-8.
6. Jiao, H., and Yu, L. (2024). Migrasomes: Biogenesis, physiological roles, and therapeutic potentials. J. Cell Biol. *223*, e202403051. https://doi.org/10.1083/JCB.202403051.
7. Minciacchi, V.R., You, S., Spinelli, C., Morley, S., Zandian, M., Aspuria, P. J., Cavallini, L., Ciardiello, C., Reis Sobreiro, M., Morello, M., et al. (2015). Large oncosomes contain distinct protein cargo and represent a separate functional class of tumor-derived extracellular vesicles. Oncotarget *6*, 11327–11341. https://doi.org/10.18632/ONCOTARGET.3598.
8. Ciardiello, C., Leone, A., Lanuti, P., Roca, M.S., Moccia, T., Minciacchi, V. R., Minopoli, M., Gigantino, V., De Cecio, R., Rippa, M., et al. (2019). Large oncosomes overexpressing integrin alpha-V promote prostate cancer adhesion and invasion via AKT activation. J. Exp. Clin. Cancer Res. *38*, 317. https://doi.org/10.1186/S13046-019-1317-6.
9. Di Vizio, D., Morello, M., Dudley, A.C., Schow, P.W., Adam, R.M., Morley, S., Mulholland, D., Rotinen, M., Hager, M.H., Insabato, L., et al. (2012). Large oncosomes in human prostate cancer tissues and in the circulation of mice with metastatic disease. Am. J. Pathol. *181*, 1573–1584. https://doi.org/10.1016/J.AJPATH.2012.07.030.
10. Minciacchi, V.R., Spinelli, C., Reis-Sobreiro, M., Cavallini, L., You, S., Zandian, M., Li, X., Mishra, R., Chiarugi, P., Adam, R.M., et al. (2017). MYC Mediates Large Oncosome-Induced Fibroblast Reprogramming in Prostate Cancer. Cancer Res. *77*, 2306–2317. https://doi.org/10.1158/0008-5472.CAN-16-2942.
11. Zhang, S., Liao, X., Chen, S., Qian, W., Li, M., Xu, Y., Yang, M., Li, X., Mo, S., Tang, M., et al. (2022). Large Oncosome-Loaded VAPA Promotes Bone-Tropic Metastasis of Hepatocellular Carcinoma Via Formation of Osteoclastic Pre-Metastatic Niche. Adv. Sci. *9*, e2201974. https://doi.org/10.1002/ADVS.202201974.
12. Bertolini, I., Terrasi, A., Martelli, C., Gaudioso, G., Di Cristofori, A., Storaci, A.M., Formica, M., Braidotti, P., Todoerti, K., Ferrero, S., et al. (2019). A GBM-like V-ATPase signature directs cell-cell tumor signaling and reprogramming via large oncosomes. EBioMedicine *41*, 225–235. https://doi.org/10.1016/J.EBIOM.2019.01.051.
13. Vagner, T., Spinelli, C., Minciacchi, V.R., Balaj, L., Zandian, M., Conley, A., Zijlstra, A., Freeman, M.R., Demichelis, F., De, S., et al. (2018). Large extracellular vesicles carry most of the tumour DNA circulating in prostate cancer patient plasma. J. Extracell. Vesicles *7*, 1505403. https://doi.org/10.1080/20013078.2018.1505403.
14. Kugeratski, F.G., Hodge, K., Lilla, S., McAndrews, K.M., Zhou, X., Hwang, R.F., Zanivan, S., and Kalluri, R. (2021). Quantitative proteomics identifies the core proteome of exosomes with syntenin-1 as the highest abundant protein and a putative universal biomarker. Nat. Cell Biol. *23*, 631–641. https://doi.org/10.1038/S41556-021-00693-Y.
15. Kowal, J., Arras, G., Colombo, M., Jouve, M., Morath, J.P., Primdal-Bengtson, B., Dingli, F., Loew, D., Tkach, M., and Théry, C. (2016). Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. Proc. Natl. Acad. Sci. USA *113*, E968–E977. https://doi.org/10.1073/PNAS.1521230113.
16. Tkach, M., Thalmensi, J., Timperi, E., Gueguen, P., Névo, N., Grisard, E., Sirven, P., Cocozza, F., Gouronnec, A., Martin-Jaular, L., et al. (2022). Extracellular vesicles from triple negative breast cancer promote pro-inflammatory macrophages associated with better clinical outcome. Proc. Natl. Acad. Sci. USA *119*, e2107394119. https://doi.org/10.1073/PNAS.2107394119.
17. Hoshino, A., Kim, H.S., Bojmar, L., Gyan, K.E., Cioffi, M., Hernandez, J., Zambirinis, C.P., Rodrigues, G., Molina, H., Heissel, S., et al. (2020). Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. Cell *182*, 1044–1061.e18. https://doi.org/10.1016/J.CELL.2020.07.009.
18. Jeppesen, D.K., Fenix, A.M., Franklin, J.L., Higginbotham, J.N., Zhang, Q., Zimmerman, L.J., Liebler, D.C., Ping, J., Liu, Q., Evans, R., et al. (2019). Reassessment of Exosome Composition. Cell *177*, 428–445. https://doi.org/10.1016/J.CELL.2019.02.029.
19. Crescitelli, R., Lässer, C., and Lötvall, J. (2021). Isolation and characterization of extracellular vesicle subpopulations from tissues. Nat. Protoc. *16*, 1548–1580. https://doi.org/10.1038/S41596-020-00466-1.
20. Jimenez, L., Yu, H., McKenzie, A.J., Franklin, J.L., Patton, J.G., Liu, Q., and Weaver, A.M. (2019). Quantitative Proteomic Analysis of Small and Large Extracellular Vesicles (EVs) Reveals Enrichment of Adhesion Proteins in Small EVs. J. Proteome Res. *18*, 947–959. https://doi.org/10.1021/ACS.JPROTEOME.8B00647.
21. Sork, H., Corso, G., Krjutskov, K., Johansson, H.J., Nordin, J.Z., Wiklander, O.P.B., Lee, Y.X.F., Westholm, J.O., Lehtiö, J., Wood, M.J.A., et al. (2018). Heterogeneity and interplay of the extracellular vesicle small RNA transcriptome and proteome. Sci. Rep. *8*, 10813. https://doi.org/10.1038/S41598-018-28485-9.
22. Jiang, P., Ma, X., Han, S., Ma, L., Ai, J., Wu, L., Zhang, Y., Xiao, H., Tian, M., Tao, W.A., et al. (2022). Characterization of the microRNA transcriptomes and proteomics of cochlear tissue-derived small extracellular vesicles from mice of different ages after birth. Cell. Mol. Life Sci. *79*, 154. https://doi.org/10.1007/S00018-022-04164-X.
23. Cifola, I., Fratini, F., Cardinali, B., Palmieri, V., Gatti, G., Selmi, T., Donzelli, S., Sacconi, A., Cesarini, V., Marei, H.E., et al. (2022). miRNome and Proteome Profiling of Small Extracellular Vesicles Secreted by Human Glioblastoma Cell Lines and Primary Cancer Stem Cells. Biomedicines *10*, 1886. https://doi.org/10.3390/BIOMEDICINES10081886.
24. Conley, A., Minciacchi, V.R., Lee, D.H., Knudsen, B.S., Karlan, B.Y., Citrigno, L., Viglietto, G., Tewari, M., Freeman, M.R., Demichelis, F., and Di Vizio, D. (2017). High-throughput sequencing of two populations of extracellular vesicles provides an mRNA signature that can be detected in the circulation of breast cancer patients. RNA Biol. *14*, 305–316. https://doi.org/10.1080/15476286.2016.1259061.
25. Welsh, J.A., Goberdhan, D.C.I., O'Driscoll, L., Buzas, E.I., Blenkiron, C., Bussolati, B., Cai, H., Di Vizio, D., Driedonks, T.A.P., Erdbrügger, U., et al. (2024). Minimal information for studies of extracellular vesicles (MISEV2023): From basic to advanced approaches. J. Extracell. Vesicles *13*, e12404. https://doi.org/10.1002/JEV2.12404.
26. Mariscal, J., Vagner, T., Kim, M., Zhou, B., Chin, A., Zandian, M., Freeman, M.R., You, S., Zijlstra, A., Yang, W., and Di Vizio, D. (2020). Comprehensive palmitoyl-proteomic analysis identifies distinct protein signatures for large and small cancer-derived extracellular vesicles. J. Extracell. Vesicles *9*, 1764192. https://doi.org/10.1080/20013078.2020.1764192.
27. Gao, Q., An, K., Gao, Z., Wang, Y., Ding, C., Niu, P., and Lei, F. (2023). Rectal cancer-derived exosomes activate the nuclear factor kappa

B pathway and lung fibroblasts by delivering integrin beta-1. KOREAN J. PHYSIOL. PHARMACOL. *27*, 375–381. https://doi.org/10.4196/KJPP.2023.27.4.375.

28. Rai, A., Fang, H., Claridge, B., Simpson, R.J., and Greening, D.W. (2021). Proteomic dissection of large extracellular vesicle surfaceome unravels interactive surface platform. J. Extracell. Vesicles *10*, e12164. https://doi.org/10.1002/JEV2.12164.
29. Soloveva, N., Novikova, S., Farafonova, T., Tikhonova, O., and Zgoda, V. (2023). Proteomic Signature of Extracellular Vesicles Associated with Colorectal Cancer. Molecules *28*, 4227. https://doi.org/10.3390/MOLECULES28104227.
30. Chitti, S.V., Gummadi, S., Kang, T., Shahi, S., Marzan, A.L., Nedeva, C., Sanwlani, R., Bramich, K., Stewart, S., Petrovska, M., et al. (2024). Vesiclepedia 2024: an extracellular vesicles and extracellular particles repository. Nucleic Acids Res. *52*, D1694–D1698. https://doi.org/10.1093/NAR/GKAD1007.
31. Pathan, M., Fonseka, P., Chitti, S.V., Kang, T., Sanwlani, R., Van Deun, J., Hendrix, A., and Mathivanan, S. (2019). Vesiclepedia 2019: a compendium of RNA, proteins, lipids and metabolites in extracellular vesicles. Nucleic Acids Res. *47*, D516–D519. https://doi.org/10.1093/NAR/GKY1029.
32. Bausch-Fluck, D., Hofmann, A., Bock, T., Frei, A.P., Cerciello, F., Jacobs, A., Moest, H., Omasits, U., Gundry, R.L., Yoon, C., et al. (2015). A mass spectrometric-derived cell surface protein atlas. PLoS One *10*, e0121314. https://doi.org/10.1371/JOURNAL.PONE.0121314.
33. Zendrini, A., Guerra, G., Sagini, K., Vagner, T., Di Vizio, D., and Bergese, P. (2022). On the surface-to-bulk partition of proteins in extracellular vesicles. Colloids Surf. B Biointerfaces *218*, 112728. https://doi.org/10.1016/J.COLSURFB.2022.112728.
34. Khoo, A., Govindarajan, M., Qiu, Z., Liu, L.Y., Ignatchenko, V., Waas, M., Macklin, A., Keszei, A., Neu, S., Main, B.P., et al. (2024). Prostate cancer reshapes the secreted and extracellular vesicle urinary proteomes. Nat. Commun. *15*, 5069. https://doi.org/10.1038/S41467-024-49424-5.
35. You, S., Knudsen, B.S., Erho, N., Alshalalfa, M., Takhar, M., Al-Deen Ashab, H., Davicioni, E., Karnes, R.J., Klein, E.A., Den, R.B., et al. (2016). Integrated Classification of Prostate Cancer Reveals a Novel Luminal Subtype with Poor Outcome. Cancer Res. *76*, 4948–4958. https://doi.org/10.1158/0008-5472.CAN-16-0902.
36. Yoon, J., Kim, M., Posadas, E.M., Freedland, S.J., Liu, Y., Davicioni, E., Den, R.B., Trock, B.J., Karnes, R.J., Klein, E.A., et al. (2021). A comparative study of PCS and PAM50 prostate cancer classification schemes. Prostate Cancer Prostatic Dis. *24*, 733–742. https://doi.org/10.1038/S41391-021-00325-4.
37. Di Vizio, D., Schoppet, M., Weeraratna, A., and Witwer, K.W. (2023). Blebs and former blebs: From surface protrusions to extracellular vesicles in cancer signalling, anoikis resistance and beyond. J. Extracell. Biol. *2*, e112. https://doi.org/10.1002/JEX2.112.
38. Cocozza, F., Martin-Jaular, L., Lippens, L., Di Cicco, A., Arribas, Y.A., Ansart, N., Dingli, F., Richard, M., Merle, L., Jouve San Roman, M., et al. (2023). Extracellular vesicles and co-isolated endogenous retroviruses from murine cancer cells differentially affect dendritic cells. EMBO J. *42*, e113590. https://doi.org/10.15252/EMBJ.2023113590.
39. Todkar, K., Chikhi, L., Desjardins, V., El-Mortada, F., Pépin, G., and Germain, M. (2021). Selective packaging of mitochondrial proteins into extracellular vesicles prevents the release of mitochondrial DAMPs. Nat. Commun. *12*, 1971. https://doi.org/10.1038/S41467-021-21984-W.
40. Peruzzotti-Jametti, L., Bernstock, J.D., Willis, C.M., Manferrari, G., Rogall, R., Fernandez-Vizarra, E., Williamson, J.C., Braga, A., van den Bosch, A., Leonardi, T., et al. (2021). Neural stem cells traffic functional mitochondria via extracellular vesicles. PLoS Biol. *19*, e3001166. https://doi.org/10.1371/JOURNAL.PBIO.3001166.
41. Nicolás-Ávila, J.A., Lechuga-Vieco, A.V., Esteban-Martínez, L., Sánchez-Díaz, M., Díaz-García, E., Santiago, D.J., Rubio-Ponce, A., Li, J.L., Balachander, A., Quintana, J.A., et al. (2020). A Network of Macrophages Supports Mitochondrial Homeostasis in the Heart. Cell *183*, 94–109. https://doi.org/10.1016/J.CELL.2020.08.031.
42. Phinney, D.G., Di Giuseppe, M., Njah, J., Sala, E., Shiva, S., St Croix, C.M., Stolz, D.B., Watkins, S.C., Di, Y.P., Leikauf, G.D., et al. (2015). Mesenchymal stem cells use extracellular vesicles to outsource mitophagy and shuttle microRNAs. Nat. Commun. *6*, 8472. https://doi.org/10.1038/NCOMMS9472.
43. Jang, S.C., Crescitelli, R., Cvjetkovic, A., Belgrano, V., Olofsson Bagge, R., Sundfeldt, K., Ochiya, T., Kalluri, R., and Lötvall, J. (2019). Mitochondrial protein enriched extracellular vesicles discovered in human melanoma tissues can be detected in patient plasma. J. Extracell. Vesicles *8*, 1635420. https://doi.org/10.1080/20013078.2019.1635420.
44. D'Souza-Schorey, C., and Di Vizio, D. (2025). A class of large cell-like extracellular vesicles. Nat. Cell Biol. *27*, 372–374. https://doi.org/10.1038/s41556-025-01611-2.
45. Jeppesen, D.K., Sanchez, Z.C., Kelley, N.M., Hayes, J.B., Ambroise, J., Koory, E.N., Krystofiak, E., Taneja, N., Zhang, Q., Dungan, M.M., et al. (2025). Blebbisomes are large, organelle-rich extracellular vesicles with cell-like properties. Nat. Cell Biol. *27*, 438–448. https://doi.org/10.1038/s41556-025-01621-0.
46. Wei, Z., Batagov, A.O., Schinelli, S., Wang, J., Wang, Y., El Fatimy, R., Rabinovsky, R., Balaj, L., Chen, C.C., Hochberg, F., et al. (2017). Coding and noncoding landscape of extracellular RNA released by human glioma stem cells. Nat. Commun. *8*, 1145. https://doi.org/10.1038/S41467-017-01196-X.
47. Sinha, A., Huang, V., Livingstone, J., Wang, J., Fox, N.S., Kurganovs, N., Ignatchenko, V., Fritsch, K., Donmez, N., Heisler, L.E., et al. (2019). The Proteogenomic Landscape of Curable Prostate Cancer. Cancer Cell *35*, 414–427. https://doi.org/10.1016/J.CCELL.2019.02.005.
48. Livshts, M.A., Khomyakova, E., Evtushenko, E.G., Lazarev, V.N., Kulemin, N.A., Semina, S.E., Generozov, E.V., and Govorun, V.M. (2015). Isolation of exosomes by differential centrifugation: Theoretical analysis of a commonly used protocol. Sci. Rep. *5*, 1–14. https://doi.org/10.1038/srep17319.
49. Konoshenko, M.Y., Lekchnov, E.A., Vlassov, A.V., and Laktionov, P.P. (2018). Isolation of Extracellular Vesicles: General Methodologies and Latest Trends. BioMed Res. Int. *2018*, 8545347. https://doi.org/10.1155/2018/8545347.
50. Crescitelli, R., Lässer, C., Jang, S.C., Cvjetkovic, A., Malmhäll, C., Karimi, N., Höög, J.L., Johansson, I., Fuchs, J., Thorsell, A., et al. (2020). Subpopulations of extracellular vesicles from human metastatic melanoma tissue identified by quantitative proteomics after optimized isolation. J. Extracell. Vesicles *9*, 1722433. https://doi.org/10.1080/20013078.2020.1722433.
51. Zhou, B., Yan, Y., Wang, Y., You, S., Freeman, M.R., and Yang, W. (2019). Quantitative proteomic analysis of prostate tissue specimens identifies deregulated protein complexes in primary prostate cancer. Clin. Proteomics *16*, 15. https://doi.org/10.1186/S12014-019-9236-2.
52. Di Vizio, D., Adam, R.M., Kim, J., Kim, R., Sotgia, F., Williams, T., Demichelis, F., Solomon, K.R., Loda, M., Rubin, M.A., et al. (2008). Caveolin-1 interacts with a lipid raft-associated population of fatty acid synthase. Cell Cycle *7*, 2257–2267. https://doi.org/10.4161/CC.7.14.6475.
53. Németh, A., Orgovan, N., Sódar, B.W., Osteikoetxea, X., Pálóczi, K., Szabó-Taylor, K., Vukman, K.V., Kittel, Á., Turiák, L., Wiener, Z., et al. (2017). Antibiotic-induced release of small extracellular vesicles (exosomes) with surface-associated DNA. Sci. Rep. *7*, 8202. https://doi.org/10.1038/S41598-017-08392-1.
54. Sandau, U.S., Duggan, E., Shi, X., Smith, S.J., Huckans, M., Schutzer, W. E., Loftis, J.M., Janowsky, A., Nolan, J.P., and Saugstad, J.A. (2020). Methamphetamine use alters human plasma extracellular vesicles and their microRNA cargo: An exploratory study. J. Extracell. Vesicles *10*, e12028. https://doi.org/10.1002/JEV2.12028.
55. Shpigelman, J., Lao, F.S., Yao, S., Li, C., Saito, T., Sato-Kaneko, F., Nolan, J.P., Shukla, N.M., Pu, M., Messer, K., et al. (2021). Generation and Application of a Reporter Cell Line for the Quantitative Screen of Extracellular

Vesicle Release. Front. Pharmacol. *12*, 668609. https://doi.org/10.3389/FPHAR.2021.668609.

56. Yan, Y., Zhou, B., Qian, C., Vasquez, A., Kamra, M., Chatterjee, A., Lee, Y. J., Yuan, X., Ellis, L., Di Vizio, D., et al. (2022). Receptor-interacting protein kinase 2 (RIPK2) stabilizes c-Myc and is a therapeutic target in prostate cancer metastasis. Nat. Commun. *13*, 669. https://doi.org/10.1038/S41467-022-28340-6.
57. Wiśniewski, J.R., Zougman, A., and Mann, M. (2009). Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. J. Proteome Res. *8*, 5674–5678. https://doi.org/10.1021/PR900748N.
58. Zhou, B., Wang, Y., Yan, Y., Mariscal, J., Di Vizio, D., Freeman, M.R., and Yang, W. (2019). Low-Background Acyl-Biotinyl Exchange Largely Eliminates the Coisolation of Non- S-Acylated Proteins and Enables Deep S-Acylproteomic Analysis. Anal. Chem. *91*, 9858–9866. https://doi.org/10.1021/ACS.ANALCHEM.9B01520.
59. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. https://doi.org/10.1038/NCOMMS14049.
60. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587. https://doi.org/10.1016/J.CELL.2021.04.048.
61. Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. *4*, 1184–1191. https://doi.org/10.1038/NPROT.2009.97.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
HSPA5 antibody	Cell Signaling	Cat# 3177; RRID: AB_2119845
Cytochrome *c* antibody	Cell Signaling	Cat# 4280; RRID: AB_1039672
TSG101 antibody	Santa Cruz	Cat# sc-7964; RRID: AB_671473
CD9 antibody	Santa Cruz	Cat# sc-13118; RRID: AB_627245
Cav-1 antibody	Santa Cruz	Cat# sc-894; RRID: AB_631407
KRT18 antibody	Abcam	Cat# ab93741; RRID: AB_10563566
GM-130 antibody	Santa Cruz	Cat# sc-55591; RRID: AB_630303
HSP60 antibody	Cell Signaling	Cat# 4870S; RRID: AB_2085240
CD81 antibody	Abcam	Cat# ab79559; RRID: AB_1604021
CD63 Antibody	Cell Signaling	Cat# 52090; RRID: AB_2799392
Biological samples		
Plasma of metastatic prostate cancer patients	Cedars-Sinai Center for Uro-Oncology Research Excellence clinics	N/A
Chemicals, peptides, and recombinant proteins		
Dulbecco's Modified Eagle Medium (DMEM)	Invitrogen	11965118
Minimum Essential Medium (MEM)	Invitrogen	11095080
Fetal bovine serum (FBS)	Invitrogen	A5670801
L-glutamine	Invitrogen	A2916801
Penicillin-Streptomycin	Invitrogen	15140122
Trypan Blue	Sigma	T8154
PBS	Invitrogen	10010023
Optiprep (Iodixanol)	Sigma	D1556
Tris base	Sigma	9210-OP
Sodium dodecyl sulfate	Sigma	71725
Hydrochloric acid	Sigma	258148
ISOTON II Diluent	Beckman Coulter	C96980
DAPI	Thermo Fisher	62248
Paraformaldehyde	Electron Microscopy Sciences	15712
Osmium tetroxide (OsO4)	Sigma	O5500
Taab 812	Taab	T024
Urea	Sigma	51456
DTT	Sigma	3860-OP
Triethylammonium bicarbonate (TEAB)	Sigma	T7408
Iodoacetamide	Sigma	I6125
Tris(2-carboxyethyl)phosphine (TCEP)	Sigma	75259
N-ethylmaleimide (NEM)	Sigma	4259
2,2′Dithiodipyridine (DTDP)	Sigma	43791
Biotin-HPDP	Sigma	SML3797
Streptavidin	Thermo Fisher	21122
MS-grade trypsin	Promega	V5111
Trizol	Thermo Fisher	15596026
TURBO DNase	ThermoFisher	AM2238
Minimum Essential Medium (MEM)	Invitrogen	11095080

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fetal bovine serum (FBS)	Invitrogen	A5670801
L-glutamine	Invitrogen	A2916801
Penicillin-Streptomycin	Invitrogen	15140122
Trypan Blue	Sigma	T8154
PBS	Invitrogen	10010023
Optiprep (Iodixanol)	Sigma	D1556
Tris base	Sigma	9210-OP
Sodium dodecyl sulfate	Sigma	71725
Hydrochloric acid	Sigma	258148
ISOTON II Diluent	Beckman Coulter	C96980
DAPI	Thermo Fisher	62248
Paraformaldehyde	Electron Microscopy Sciences	15712
Osmium tetroxide (OsO4)	Sigma	O5500
Taab 812	Taab	T024
Urea	Sigma	51456
DTT	Sigma	3860-OP
Triethylammonium bicarbonate (TEAB)	Sigma	T7408
Iodoacetamide	Sigma	I6125
Tris(2-carboxyethyl)phosphine (TCEP)	Sigma	75259
N-ethylmaleimide (NEM)	Sigma	4259
2,2′Dithiodipyridine (DTDP)	Sigma	43791
Biotin-HPDP	Sigma	SML3797
Streptavidin	Thermo Fisher	21122
MS-grade trypsin	Promega	V5111
Trizol	Thermo Fisher	15596026
TURBO DNase	ThermoFisher	AM2238
Critical commercial assays		
MycoAlert PLUS Mycoplasma Detection Kit	Lonza	LT07-318
Pierce 660nm Protein Assay	Thermo Fisher	22660
vFC Assay kit	Cellarcus Biosciences	CBS4-1
Protifi S-TRAP 96-well plates	Protifi	C02-96well-1
Pierce BCA protein assay	Thermo Fisher	23227
Zymo RNA Clean and Concentrator	Zymo Research	R1014
Quant-it™ RiboGreen Reagent and RNA Assay Kit	ThermoFisher	R11490
SMARTer Pico V2 (Takara 634412)	Takara	634412
Kapa Library Quantification Kits (KK4824)	Roche	KK4824
Pierce 660nm Protein Assay	Thermo Fisher	22660
vFC Assay kit	Cellarcus Biosciences	CBS4-1
Protifi S-TRAP 96-well plates	Protifi	C02-96well-1
Pierce BCA protein assay	Thermo Fisher	23227
Zymo RNA Clean and Concentrator	Zymo Research	R1014
Quant-it™ RiboGreen Reagent and RNA Assay Kit	ThermoFisher	R11490
SMARTer Pico V2 (Takara 634412)	Takara	634412
Kapa Library Quantification Kits (KK4824)	Roche	KK4824
Deposited data		
A mass spectrometric-derived cell surface protein atlas	Bausch-Fluck et al.[29]	DOI: https://doi.org/10.1371/journal.pone.0121314
Prostate cancer reshapes the secreted and extracellular vesicle urinary proteomes	Khoo et al.[31]	DOI: https://doi.org/10.1038/s41467-024-49424-5

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Label-free proteomics of different EV populations from three different cell lines	This paper	PXD038518
SWATH mass spectrometry of LOs from metastatic prostate cancer patients	This paper	PXD038011
Palmitoylated proteomics of LOs and unfractionated human plasma	This paper	Supplementary file 3
Bulk RNA sequencing of different EV populations from PC3 cell line	This paper	GSE214804
Single EV RNA sequencing of LO isolated from PC3 cell line	This paper	GSE231846
Prostate cancer reshapes the secreted and extracellular vesicle urinary proteomes	Khoo et al.[31]	DOI: https://doi.org/10.1038/s41467-024-49424-5
Label-free proteomics of different EV populations from three different cell lines	This paper	PXD038518
SWATH mass spectrometry of LOs from metastatic prostate cancer patients	This paper	PXD038011
Palmitoylated proteomics of LOs and unfractionated human plasma	This paper	Supplementary file 3
Bulk RNA sequencing of different EV populations from PC3 cell line	This paper	GSE214804
Single EV RNA sequencing of LO isolated from PC3 cell line	This paper	GSE231846
Experimental models: Cell lines		
PC3 cells	American Type Culture Collection	CRL-11435
MDA-MB-231 cells	American Type Culture Collection	HTB-26
U87 cells	American Type Culture Collection	HTB-14
MDA-MB-231 cells	American Type Culture Collection	HTB-26
U87 cells	American Type Culture Collection	HTB-14
Software and algorithms		
FCS Express	*De Novo* Software	A48515
MaxQuant (v1.5.5.1)	Max Planck Institute of Biochemistry	https://www.maxquant.org/
Andromeda	The European Unionseventh Framework Program	https://bioinformaticshome.com/tools/proteomics/descriptions/Andromeda.html#gsc.tab=0
R (v4.0.3, v3.6.1, v4.2.1)	The R Foundation for Statistical Computing	https://www.r-project.org/
DESeq2 (v1.26.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
ComplexHeatmap (v2.12.1)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html
UpSetR (v1.4.0)	CRAN	https://cran.r-project.org/web/packages/UpSetR/index.html
ClusterProfiler (v3.10.1)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
limma package (v3.42.2)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/limma.html
ConsensusClusterPlus (v1.58.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html
survival package (v3.5.0)	CRAN	https://cran.r-project.org/web/packages/survival/index.html
CellRanger (v6.0.1)	10x Genomics	https://support.10xgenomics.com/software/cell-ranger/latest/overview
Seurat	Satija Lab	https://satijalab.org/seurat/
ggplot2	CRAN	https://ggplot2.tidyverse.org/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ggforce	CRAN	https://cran.r-project.org/web/packages/ggforce/index.html
UMAP	McInnes, Healy, & Melville	https://github.com/lmcinnes/umap
umap-learn	Python UMAP implementation	https://umap-learn.readthedocs.io/en/latest/
bcl2fastq (v2.19.1.403)	Illumina	https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html
cutadapt (v2.7)	Marcel Martin	https://cutadapt.readthedocs.io/en/stable/
STAR (v2.6.1day)	Alexander Dobin et al.	https://github.com/alexdobin/STAR
featureCounts (v1.6.3)	Wei Shi et al.	http://subread.sourceforge.net/
salmon quant (v0.11.3)	Rob Patro, Devin Petersohn, and Carl Kingsford	https://salmon.readthedocs.io/en/latest/
bedtools coverage (v2.29.0)	Quinlan and Hall	https://bedtools.readthedocs.io/en/latest/
MaxQuant (v1.5.5.1)	Max Planck Institute of Biochemistry	https://www.maxquant.org/
Andromeda	The European Unionseventh Framework Program	https://bioinformaticshome.com/tools/proteomics/descriptions/Andromeda.html#gsc.tab=0
R (v4.0.3, v3.6.1, v4.2.1)	The R Foundation for Statistical Computing	https://www.r-project.org/
DESeq2 (v1.26.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
ComplexHeatmap (v2.12.1)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html
UpSetR (v1.4.0)	CRAN	https://cran.r-project.org/web/packages/UpSetR/index.html
ClusterProfiler (v3.10.1)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
limma package (v3.42.2)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/limma.html
ConsensusClusterPlus (v1.58.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html
survival package (v3.5.0)	CRAN	https://cran.r-project.org/web/packages/survival/index.html
CellRanger (v6.0.1)	10x Genomics	https://support.10xgenomics.com/software/cell-ranger/latest/overview
Seurat	Satija Lab	https://satijalab.org/seurat/
ggplot2	CRAN	https://ggplot2.tidyverse.org/
ggforce	CRAN	https://cran.r-project.org/web/packages/ggforce/index.html
UMAP	McInnes, Healy, & Melville	https://github.com/lmcinnes/umap
umap-learn	Python UMAP implementation	https://umap-learn.readthedocs.io/en/latest/
bcl2fastq (v2.19.1.403)	Illumina	https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html
cutadapt (v2.7)	Marcel Martin	https://cutadapt.readthedocs.io/en/stable/
STAR (v2.6.1day)	Alexander Dobin et al.	https://github.com/alexdobin/STAR
featureCounts (v1.6.3)	Wei Shi et al.	http://subread.sourceforge.net/
salmon quant (v0.11.3)	Rob Patro, Devin Petersohn, and Carl Kingsford	https://salmon.readthedocs.io/en/latest/
bedtools coverage (v2.29.0)	Quinlan and Hall	https://bedtools.readthedocs.io/en/latest/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
NP250 nanopore	iZON Science	NP250
NP 2000 nanopore	iZON Science	NP2000
Izon calibration particles 200 nm	iZON Science	CPC200
Izon calibration particles 2 μm	iZON Science	CPC2000
EASY-Spray analytical column	Thermo Scientific	ES903
C18 column	Phenomenex	00A-4462-AC
Ultracentrifuge tubes	Beckman Coulter	344058
NP 2000 nanopore	iZON Science	NP2000
Izon calibration particles 200 nm	iZON Science	CPC200
Izon calibration particles 2 μm	iZON Science	CPC2000
EASY-Spray analytical column	Thermo Scientific	ES903
C18 column	Phenomenex	00A-4462-AC
Ultracentrifuge tubes	Beckman Coulter	344058

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell culture

All cell lines were obtained from American Type Culture Collection. Cell line identity was validated using STR analysis. Cell lines were maintained at 37°C in 5% CO2. PC3 and MDA-MB-231 cells were grown in Dulbecco's Modified Eagle Medium (Invitrogen), and U87 cells were grown in Minimum Essential Medium (Invitrogen). The media for all cell lines were supplemented with 10% fetal bovine serum (Invitrogen), 2 mM L-glutamine (Invitrogen) and 1% PenStrep (Invitrogen). Cell viability of the EV-producing cells was tested with the 0.4% Trypan Blue (Sigma) exclusion method. All cell lines were routinely tested for mycoplasma contamination by using the MycoAlert PLUS Mycoplasma Detection Kit (Lonza).

Human subjects

All human subjects provided written informed consent and research was performed under Cedars Sinai Medical Center Institutional Review Board-approved protocols Pro00033050 and Pro00042197. Subjects provided serial specimens at standard clinical visits with timing set by the treating clinicians. The samples utilized in this study were collected between January 2016 - December 2020. We included two groups of patients with metastatic PC: hormone-sensitive (n = 10) and hormone-resistant (n = 10) with a pathologically proven diagnosis of metastatic PC with evidence of cancer metastasis either on next-generation imaging (PSMA PET/CT, ^{68}Ga or ^{18}F-based) or conventional radiographs (CT/MRI, 99Tc-singlephoton emission computed tomography [SPECT] bone scan). For patients classified by PSMA PET/CT, a metastatic lesion was defined by the presence of a PET-detectable lesion with an uptake greater than background and an SUV minimum of 10. This was consistent with previous studies utilizing PSMA PET as a screening modality for detection of metastatic disease.

Samples were selected based upon the availability of plasma before and during therapy to the point of progression mandating a change in treatment. Selected patients were identified as hormone-sensitive if they were treatment naive and had serum testosterone concentrations >250 ng/dL prior to therapy then exhibited a response (biochemical and/or clinical) while on hormone therapy. Patients were identified as hormone-resistant if they exhibited biochemical progression with or without radiographic progression while undergoing effective hormone therapy (i.e., serum testosterone <50 ng/dL). Details of the subjects are presented in Table 1. For each individual, plasma aliquots utilized for L-EVs isolation were obtained from the blood draws at a timepoint within 3 months of starting the assigned therapy (noted on Table 1). Time to progression for progressive disease was calculated as the difference in months between treatment start date and treatment failure. Patients without disease recurrency were considered stable and time to progression was calculated considering the most updated clinical information available as of October/2024 at least four years post treatment. Details of timing are presented in Table 1.

METHOD DETAILS

Enrichment of EVs from cell culture media

Cells were plated into 150 mm tissue culture dishes (n = 18) and allowed to proliferate in serum-supplemented media for at least 48 h. After cells reached ~90% confluency, they were washed twice in PBS, and placed in serum-free media 24 h before EV harvest. The enrichment and density gradient purification of EVs was performed as previously reported[7,13,15,26] with minor modifications described below. The conditioned media was centrifuged three times at 300 g, for 5 min each to pellet down floating cells. This was followed by centrifugation at 2,800 g for 10 min to collect the first fraction. The supernatant was spun in an ultracentrifuge at

10,000 *g* for 30 min (k-factor 2547.2) to collect the second fraction, and the supernatant was then spun at 100,000 *g* for 60 min (k-factor 254.7) to collect the third fraction. All differential centrifugation steps were performed at 4°C. All pellets (2.8K, 10K, and 100K) were then subjected to iodixanol (Optiprep, Sigma) density gradient purification. Freshly pelleted EVs were resuspended in 0.2 μm-filtered PBS and deposited at the bottom of an ultracentrifuge tube (Cat#344058, Beckman Coulter). Next, 30% (4.3 mL, 1.20 g/mL), 25% (3 mL, 1.15 g/mL), 15% (2.5 mL, 1.10 g/mL), and 5% (6 mL, 1.08 g/mL) iodixanol solutions were sequentially layered at decreasing density to form a discontinuous gradient. Separation was performed by ultracentrifugation at 100,000 *g* for 230 min (4°C, k-factor 254.7) and L-EVs derived from the 2.8K and 10K fractions separated into the 1.10–1.15 g/mL density fractions while S-EVs derived from the 100K fraction separated into the 1.10 g/mL density fraction. Purified EVs were then washed in PBS (100,000 *g*, 60 min, 4°C) and resuspended in 0.2 μm-filtered PBS, or in 25 μL of DTT-free 4% SDS/Tris-HCl lysis buffer[50] for proteomic analysis, or in 1 μL TRIzol for transcriptomic analysis. All ultracentrifugation spins were performed in an SW28 swinging rotor (Beckman Coulter).

Enrichment of large EVs (L-EVs) from patient plasma

Blood samples were collected in BD Vacutainer ACD tubes and processed within 2 h of phlebotomy. Platelet-poor plasma was prepared from blood by centrifugation at 2000 x *g*, 4°C, for 5 min, followed by another spin at 2000 x *g*, 4°C, for 5 min. Circulating L-EVs were isolated from platelet-poor plasma by differential centrifugation. Three milliliters of plasma were diluted in the 0.2 μm-filtered PBS to the total volume of 16 mL and centrifuged at 10,000 *g* for 30 min (k-factor 2547.2). Resulting crude LO pellets were resuspended in 200 μL of 0.2 μm-filtered PBS and purified from contaminating non-EV proteins by iodixanol density gradient, as described above.

Whole cell protein lysates

Whole cell lysate (WCL) was obtained after growing cells in serum-free medium for 24 h and collection of conditioned cell media. Cell monolayers were scraped and washed in chilled PBS ×3 times. Cells were lysed in DTT-free 4% SDS/Tris-HCl lysis buffer.[51]

Immunoblotting

Immunoblotting analysis was performed as previously described.[52] Protein concentration was measured using Pierce 660nm Protein Assay (#22660, ThermoFisher Scientific). The following primary antibodies were used: HSPA5 (#3177, 1:1,000 dilution) and Cytochrome *c* (#4280, 1:1000 dilution) from Cell Signaling; TSG101 (sc-7964, 1:1,000 dilution), CD9 (sc-13118, 1:1,000 dilution), and Cav-1 (sc-894, 1:10,000 dilution) from Santa Cruz; KRT18 (ab93741, 1:10,000 dilution) and CD81 (ab79559, 1:10,000 dilution) from Abcam.

Tunable Resistive Pulse Sensing (TRPS) measurements

EV concentrations and size distributions were measured using the qNanoGold instrument (iZON Science, New Zealand). Freshly purified EVs were diluted 1:40 in 0.2-μm filtered PBS. L-EVs (2.8K and 10K) were analyzed using an NP 2000 nanopore (resolution window 0.9–5.7 μm) and S-EVs (100K) were analyzed using an NP250 nanopore (resolution window 110–630 nm). Membranes were stretched at 47 mm and voltage set either at 0.04 V for L-EVs or 0.5 V for S-EVs in order to achieve a stable current baseline of about 120 nA. Particle size and concentrations were calibrated using Izon calibration particles (1:1,000 diluted CPC2000 for L-EVs and 1:100 diluted CPC200 for S-EVs) and a minimum of 500 events were recorded for each sample with a positive pressure of 5 mbar.

Electrical Sensing Zone (ESZ) measurements

EV concentration and size distribution was also orthogonally measured using the Beckman Coulter z series instrument. Briefly, EV preparations were diluted 1:1000 in ISOTON II Diluent (Beckman Coulter). The measurements were performed using the following parameters: lower detection limit – 1 μm, upper detection limit – 40 μm, gain – 512, current – 2.828 A, preamp gain – 179.2.

Transmission electron microscopy (TEM)

EV pellets were analyzed by TEM as described previously.[53] Briefly, EVs fixed with 4% paraformaldehyde were postfixed in 0.5% osmium tetroxide (OsO_4, Sigma). The pellets were dehydrated in graded ethanol including block staining with 1% uranyl-acetate in 50% ethanol for 30 min and embedded in Taab 812 (Taab). An overnight polymerization of samples at 60°C was followed by sectioning, and the ultrathin sections were analyzed using a Hitachi 7100 electron microscope (Hitachi, Chiyoda City, Tokyo, Japan) equipped by Veleta, a 2000 × 2000 MegaPixel side-mounted TEM CCD camera (Olympus, Shinjuku City, Tokyo, Japan).

Vesicle flow cytometry

EV concentration, size, and surface marker abundance were measured by single vesicle flow cytometry[54,55] using a commercial kit (vFC Assay kit, Cellarcus Biosciences, La Jolla, CA) and flow cytometer (CytoFLEX, Beckman Coulter) Briefly, samples were stained with the fluorogenic membrane stain vFRed and one or more fluorescent antibodies for 1 h at RT and analyzed using vFRed fluorescence to trigger detection. Controls included buffer only, reagent only, and positive and negative controls for antigen abundance. Spectral compensation was performed using antibody-stained antibody capture beads and validated using single stained controls. Data were analyzed using FCS Express 7 (*De Novo* Software) and included calibration using a vesicle size and fluorescence intensity

standards. The analysis included a pre-stain dilution series to determine the optimal initial sample dilution and multiple positive and negative controls, per guidelines of the International Society for Extracellular Vesicles (ISEV).[25] A detailed description of vFC methods and controls, as well as the MIFlowCyt and MIFlowCyt-EV checklists, as requested by the guidelines are provided in Supporting Information.

Label-free proteomic profiling of cell line-derived EVs

Label-free proteomic comparison of EV samples in biological triplicates was conducted essentially as described.[56] For each replicate, 4 μg EV proteins were digested into tryptic peptides by filter-aided sample preparation.[57] About 1 μg tryptic peptides in 10 μL solution was loaded onto a 2 cm trap column (Thermo Scientific) and then separated on a 50 cm EASY-Spray analytical column (Thermo Scientific) heated to 55°C, using a 3 h gradient at the flow rate of 250 nL/min. The resolved tryptic peptides were ionized by an EASY-Spray ion source (Thermo Scientific), and mass spectra (MS1 and MS2) were acquired in a data-dependent manner in an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). EASY-IC was used for internal mass calibration. MS1 scans were acquired in 240,000 resolution at m/z of 400 Th, with an ion packet setting of 4×10^5 for automatic gain control and a maximum injection time of 50 ms. Most intense peptide ions with charge state of 2–7 were automatically selected for MS2 fragmentation by higher energy collisional dissociation, using 32% normalized collision energy. MS2 spectra were acquired in the ion trap, using rapid ion trap scan at 1×10^4 automatic gain control and 35 ms maximum injection time. To minimize redundant MS2 acquisition, dynamic exclusion was enabled (repeat count = 1, exclusion during = 60 s, repeat during = 60 s).

The acquired MS data (27 RAW files) were searched against the Uniprot_Human database using MaxQuant (v1.5.5.1). The searching parameters were set as follows: trypsin/P as the protease; carbamidomethyl (C) as variable modification; oxidation (M), acetyl (protein N-term), and deamidation (NQ) as variable modifications; minimal peptide length as 7; up to two missed cleavages; mass tolerance for MS1 was 4.5 ppm and for MS2 was 0.5 Da; identification of second peptides enabled; label free quantification enabled, with match-between-runs within 0.7 min. A standard false discovery rate of 0.01 was used to filter peptide-spectrum matches, peptide identifications, and protein identifications.

SWATH-MS instrumentation of patient-derived L-EVs

Patient-derived L-EV samples were dehydrated in a speed vacuum and resuspended in 25 μL lysis buffer containing 8 M Urea, 5% SDS and 100 mM DTT in triethylammonium bicarbonate (TEAB). Following brief sonication, samples were alkylated with 10 mM iodoacetamide and processed on Protifi S-TRAP 96-well plates for tryptic digestion according to the manufacturer's protocol. Eluted peptides were assayed for concentration using the BCA protein assay, dried, and resuspended in 0.1% formic acid in water for injection on LC MS/MS. LC MS/MS analysis was performed on an Orbitrap Exploris 480 (Thermo Scientific, BRE725539) mass spectrometer interfaced with a Nanospray Flex ionization source (Thermo Scientific, P/N ES072) coupled to an Ultimate 3000 ultra-high-pressure chromatography system with 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. Peptides were loaded at 2 μg, based on estimated recoveries from S-Trap processing protocol, and separated on a gradient of 1% B organic phase for 2 min, 1–5% B for 0.5 min, 5–9% B for 3.5 min, 9–27% B for 39 min, 27–44% B for 15 min, on a C18 column (15 cm length, 300 μm ID, 100 Å pore size, Phenomenex) over the course of 60 min with a constant flow rate of 9.5 μL/min. Source parameters were set to a voltage of 3000 kV, capillary temp of 300°C, and a sweep gas of 2 Arb. MS1 resolution was set to 60,000 and AGC was set to "standard" with ion transmission of 100 ms and RF Lens 40%. Mass range of 400–1000 and AGC target value for fragment spectra of 300% was used. Peptide ions were fragmented at a normalized collision energy of 30%. Fragmented ions were detected across 50 non-overlapping data independent acquisition precursor windows of size 12 Da MS2 Resolution was set to 15,000 with an ion transmission time of 22 ms. All data was acquired in profile mode using positive polarity.

LB-ABE enrichment of palmitoyl-proteins

Plasma-purified LO and unfractionated plasma from matched mCRPC patients were subjected to LB-ABE coupled to label-free mass spectrometry as described.[26,58] Briefly, 40 μg proteins were reduced with 50 mM tris(2-carboxyethyl)phosphine (TCEP), sequentially alkylated with 50 mM N-ethylmaleimide (NEM) and 25 mM 2,2'-dithiodipyridine (DTDP), and biotinylated with 1 mM biotin-HPDP in the presence or absence of 2 M neutral hydroxylamine (Hyd). Palmitoyl-proteins were enriched by streptavidin affinity purification, eluted by 50 mM TCEP, and precipitated by methanol/chloroform. Enriched palmitoyl-proteins were digested with MS-grade trypsin (Promega) by filter-aided sample preparation (FASP). Tryptic peptides were then recovered, processed and loaded onto a 2-cm trap column and separated on a 50-cm EASY-Spray analytical column heated to 55°C, using a 2 h-gradient consisting of 2–40% B in 150 min, 40–100% B in 20 min, and 100% B in 10 min at the flow rate of 150 nL/min. Separated peptides were ionized with an EASYSpray ion source. Mass spectra were acquired in a data dependent manner, with automatic switching between MS and MS/MS scans. In MS scans, the lock mass at m/z 445.120025 was applied to provide real-time internal mass calibration.

Bulk RNA-seq instrumentation

RNA was isolated from EV samples using Trizol (Thermo Fisher, Waltham, MA, USA) for RNA sequencing. RNA samples were treated with TURBO DNase (ThermoFisher AM2238) followed by Zymo RNA Clean and Concentrator (R1014). RNA concentration was quantified by Ribogreen reagent (ThermoFisher R11490) and by Agilent TapeStation (5067–5579). Between 2.5 and 10 ng of RNA were used in sample preparation. cDNA libraries for sequencing were made were made using SMARTer Pico V2 (Takara 634412). Libraries

were quantified for paired-end sequencing (101bp x 101bp) on the Illumina NovaSeq 6000 using Kapa Library Quantification Kits (Roche KK4824).

Single-EV RNA-seq instrumentation

10x Genomics Chromium library construction and sequencing

L-EVs were isolated from the PC3 cell conditioned medium, purified by iodixanol density gradient, and the number of particles quantified as described above. To enable encapsulation of single L-EVs into individual droplets, the EV sample was diluted to contain 10,000 particles total. A 10x Genomics Chromium machine was used for single-EV capture. cDNA preparations were obtained according to the Single Cell 3′ Protocol recommended by the manufacturer. Silane magnetic beads and solid-phase reversible immobilization (SPRI) beads were used to clean up the GEM reaction mixture, and the barcoded cDNA was then amplified in a PCR step. The P7 and R2 primers were added during the GEM incubation and the P5, and R1 during library construction via end repair, A-tailing, adaptor ligation, and PCR. The final libraries contain the P5 and P7 primers used in Illumina bridge amplification. Sequencing was carried out on an Illumina Novaseq 6000 with paired-end with read lengths of 28 and 91 bp for read 1 and read 2, respectively.

QUANTIFICATION AND STATISTICAL ANALYSIS

Label-free proteomic analysis

In order to standardize the proteomic data, normalization was performed on a per-sample basis.

Normalization

Normalization of protein was performed on a per-sample basis, calculated by Equation 1.

$$(Normalized\ intensity)_{i,j} = \frac{(Protein\ intensity)_{i,j}}{\left(\sum_{q=0}^{n}(Protein\ intensity_q)\right)_j}, i = 1,\ldots n_{proteins}; j = 1,\ldots n_{samples} \quad \text{(Equation 1)}$$

Fraction of proteins in each rank across the nine samples was calculated by Equation 2.

$$(Fraction\ of\ Rank\ x)_{i,j} = \frac{(Number\ of\ proteins\ with\ rank\ x)_{i,j}}{Number\ of\ proteins\ detected\ in\ three\ subtypes},$$

$$x = 1,\ldots 9; i = 1,\ldots n_{proteins}; j = 1,\ldots n_{samples} \quad \text{(Equation 2)}$$

Proteins only detected in one of the triplicates were excluded from statistical analysis. A linear regression model was fit using the R limma package (version 3.42.2). Test statistics from F-test and t-test were moderated by the eBayes function, which applies empirical Bayes method to shrink the protein-wise sample variances toward a common value, thereby augmenting the degrees of freedom for the individual variances. For proteins detected in all three EV subtypes, a multiple linear regression model was fit:

$(Normalized\ intensity)_{i,j} = \beta_{0,i}+\beta_{1,i}\times x_{1,j}+\beta_{2,i}\times x_{2,j}$, $i = 1, \ldots n_{proteins}; j = 1, \ldots n_{samples}$[26]

If sample$_j$ was in the 2.8K EV subtype, then $x_{1,j} = 0$, $x_{2,j} = 0$

if sample$_j$ was in the 10K EV subtype, then $x_{1,j} = 1$, $x_{2,j} = 0$

if sample$_j$ was in the 100K EV subtype, then $x_{1,j} = 0$, $x_{2,j} = 1$.

For proteins detected in only two EV subtypes (e.g., 2.8K pellet and 10K pellet), a linear regression model was fit:

$(Normalized\ intensity)_{i,j} = \beta_{0,i} + \beta_{1,i} \times x_{1,j}$, $i = 1, \ldots n_{proteins}; j = 1, \ldots n_{samples}$ (4)

If sample$_j$ was in the 2.8K EV subtype, then $x_{1,j} = 0$

if sample$_j$ was in the 10K EV subtype, then $x_{1,j} = 1$.

False discovery rate (FDR) was used to adjust *p*-values in multiple hypothesis testing. For proteins detected in three EV subtypes, *p*-values of t-test were adjusted together for all three pairwise comparisons (2.8K vs. 10K, 10K vs. 100K and 2.8K vs. 100K).

All statistical analyses were performed in the programming language R (version 3.6.1). Data was visualized using the BPG library (v5.3.4).[27]

Statistical analysis

Statistical analysis was performed based on normalized intensities. Proteins detected in exactly one of the triplicates were excluded from statistical analysis. For proteins detected in all three EV subtypes, a multiple linear regression model was fit to evaluate differential normalized intensity. An F-test was used to test if the mean of normalized intensity was equal across three EV subtypes. A T-test with Welch's adjustment for heteroscedasticity was then applied to test if the mean of the normalized intensities were equal between every pair of EV subtypes. Threshold of FDR-adjusted *p*-value <0.05 from F-tests, FDR-adjusted *p*-value <0.01 from t-tests and $|\log_2(\text{Fold change})| > 2$ were used in identifying differentially abundant proteins.

For proteins detected only in two EV subtypes, a linear regression model was fit and a t-test with Welch's adjustment for heteroscedasticity was performed to assess the difference in normalized intensity between the two EV subtypes. Threshold of

FDR-adjusted p-value <0.01 from t-tests and $|\log_2(\text{Fold change})|$ >2 were used to identify differentially abundant proteins. The MS data was deposited to the ProteomeXchange Consortium via the PRIDE partner repository with dataset identifier PXD038518.

SWATH-MS analysis of patient-derived L-EVs

All analyses were carried out in the programming language R (v4.2.1). Consensus clustering (max_k = 10; Spearman's ρ as the similarity metric; pItem = 0.8; pFeature = 1; seed = 2023; reps = 1000; ConsensusClusterPlus v1.58.0) was performed using hierarchical clustering on 30 samples and the z-scored 75% most variable protein groups that were detected in all samples. The samples from patient A2 are outliers and were therefore, excluded from the analysis. Survival analysis was performed using the R package survival (v3.5.0) to identify the time to progression. Cox proportional models were applied to the before-treatment abundance of the 75% most variable protein groups, adjusting for treatment_ADT and treatment_Docetaxel. C index was used to evaluate the goodness of fit for Cox models. p values were adjusted for multiple comparisons using FDR. Visualizations were created using the R package BPG (v7.0.5). MS data was deposited to the ProteomeXchange Consortium via the PRIDE[28] partner repository with the dataset identifier PXD038011.

LB-ABE LC-MS/MS analysis and data processing

The acquired MS data were searched against the Uniprot_Human database using the Andromeda [30] algorithm in the MaxQuant (v1.5.5.1) environment. The searching parameters were set as follows: trypsin/P as the protease; oxidation (M), acetyl (protein N-term), NEM (C) and carbamidomethyl (C) as variable modifications; up to two missed cleavages; minimal peptide length as 7; mass tolerance for MS1 was 4.5 ppm for main search and for MS2 was 0.5 Da; identification of second peptides enabled; label free quantification (LFQ) enabled, and match-between runs within 2 min were enabled. A stringent false discovery rate (FDR) < 0.01 was used to filter PSM, peptide, and protein identifications.

To retain high-confidence proteins from the raw LFQ data, we applied criteria that proteins should be detected with at least 2 peptides and in at least 2 replicates, and low abundant proteins less than 5% of absolute protein abundance for each identification were discarded for subsequent analyses. Absolute protein abundance for each identification was defined as the median value of replicates. To compare the protein abundance between groups, a max-min normalization method was used to scale the protein expression values between 0 and 1. Protein expression value 1 was given to the most abundant protein within each sample and protein value 0 was given to the least abundant protein identified. To determine the cut-off value of the protein expression difference, we computed the protein expression differences of randomly permuted samples, fitted a Gaussian distribution to the random protein expression difference, and then calculated the 99th percentile corresponding to $\alpha = 0.01$ (protein expression difference cut-off = 0.1). MS data is available as a supplementary file (Data S1).

Bulk RNA-seq analysis

FASTQ generation and gene and transcript quantification

FASTQ files were generated using bcl2fastq v2.19.1.403 using default parameters. Reads were trimmed with cutadapt v2.7 according to kit recommendations: -U 3 for SMARTerPicov2. Trimmed FASTQ files were then aligned to GRCh38 primary assembly genome from GENCODE with STAR v2.6.1day with the following parameters: –runMode alignReads –outSAMtype BAM Unsorted –outSAMmode Full –outSAMstrandField intronMotif –outFilterType BySJout –outSAMunmapped Within –outSAMmapqUnique 255 –outFilterMultimapNmax 20 –outFilterMismatchNmax 999 –outFilterMismatchNoverLmax 0.1 –alignMatesGapMax 1000000 –seedSearchStartLmax 50 –alignIntronMin 20 –alignIntronMax 1000000 –alignSJoverhangMin 18 –alignSJDBoverhangMin 18 –chimSegmentMin 18 –chimJunctionOverhangMin 18 –outSJfilterOverhangMin 18 18 18 18 –alignTranscriptsPerReadNmax 50000. Following genome alignment, reads were counted with featureCounts v1.6.3 (part of the subread package) using a non-redundant genome annotation combined from GENCODE 29 and LNCipedia (v5.2) and the following parameters: -p -t exon -g gene_id -s 2. Trimmed fastqs were also quasimapped to the same combined GENCODE 29 and LNCipedia5.2 annotation using salmon quant v0.11.3 to estimate transcripts per million (TPMs) with the following parameters: –libType A –numBootstraps 100 –seqBias –gcBias -dumpEq. Transcript coverage was estimated using bedtools coverage v2.29.0 using default parameters.

Analysis

Count data was loaded into R v4.0.3 for analysis. Normalized counts, variance stabilizing transformation, and differential abundance analysis were performed using DESeq2 v1.26.0. Differentially abundant genes with a Benjamini and Hochberg adjusted $p < 0.05$ and mean-normalized count >10 were kept for downstream analysis. Heatmaps were created using the ComplexHeatmap (v2.12.1) package in R with Z-scored variance stabilized values as input. UpSet plots were generated using UpSetR (v1.4.0). Pathway analysis was performed on differentially expressed genes using ClusterProfiler (v3.10.1) with Gene Ontology Biological Processes. GSEA was performed using the rank of the Wald test statistic for all genes tested for differential abundance analysis. Data were deposited and approved by GEO (ID GSE214804).

Single-EV RNA-seq analysis

Data was analyzed using CellRanger (version 6.0.1)[59] using GRCh38-2020-A provided by 10x Genomics as reference genome. Since default filters are optimized for the detection of cells, we applied custom filters on barcodes and features in the downstream analyses instead, as explained below. Seurat[60] was used to filter features and barcodes (min.cells = 1, min.features = 3). Mitochondrial markers are defined by annotation obtained through Biomart[61] based on human genome hg38. Markers of large oncosomes (LO),

small EVs (S-EVs) and LO vs. S-EV differentially abundant genes (DEGs) were obtained from bulk RNA sequencing data using the following thresholds.

(1) LO markers are defined as genes with mean normalized counts = 0 in the 100K fraction and mean normalized counts >5 in the 10K fraction,
(2) S-EV markers are defined as genes with mean normalized counts = 0 in the 10K fraction and mean normalized counts >5 in the 100K fraction,
(3) LO vs. S-EV DEGs were calculated as explained in Bulk RNA sequencing paragraph: briefly, differential abundance analysis was performed using DESeq2 v1.26.0. Differentially abundant genes with a Benjamini and Hochberg adjusted $p < 0.05$ and mean normalized counts >10 are retained. All the mitochondrial markers are excluded.

Co-occurrence of LO and S-EV marker reads with mitochondrial reads was performed using Fisher's Test on raw counts, testing different thresholds of detection for mitochondrial reads (from 1 to 20 reads). Data were deposited and approved by GEO (ID GSE231846).

For dimensionality reduction and clustering analysis we proceeded as follows: we filtered out all barcodes with less than 50 counts (n = 9343, we introduce this filter in order to remove signal putatively originating from small nanoparticles or ambient RNA and keep signal from LO: this threshold is conservative but does not take into account the presence of doublets since, to our knowledge, there are no methods for excluding aggregates in single EVs sequencing data), then data was normalized and scaled using the NormalizeData (normalization.method = "LogNormalize") and Scale functions in Seurat. FindVariableFeatures was used to select most informative features (selection.method = "vst", nfeatures = 1000), then RunPCA was run using the selected features. Following FindNeighbors (reduction = "pca", dims = 1:10) and FindClusters (resolution = 0.002, algorithm = 2 and resolution = 0.2, algorithm = 2 for analysis of all drops with reads $\geq$50 and drops positive for LO markers, respectively) were run to find clusters. Finally, UMAP was generated using the RunUMAP function (dims = 1:10, method = "umap-learn", n.neighbors = 30, and min.dist = 0.5 or min.dist = 1 for analysis of all drops with reads $\geq$50 and drops positive for LO markers, respectively). Markers of each cluster were found using FindAllMarkers (only.pos = TRUE, min.pct = 0.25, thresh.use = 0.25). Visualization was obtained using custom scripts and the ggplot2 and ggforce packages. Statistical analysis was performed using R v4.0.2.

Cell Reports Medicine, Volume *6*

Supplemental information

Extracellular vesicle heterogeneity through the lens of multiomics

Taylon F. Silva, Elizabeth Hutchins, Wenyan Zhao, Yari Ciani, Minhyung Kim, Emily Ko, Javier Mariscal, Zhuyu Qiu, Fatima Bedier, Agnes Kittel, Bo Zhou, Yang Wang, Megan Hall, Francesca Galasso, Rebecca Reiman, Michael R. Freeman, Sarah Parker, Jennifer Van Eyk, Wei Yang, Edwin Posadas, Jlenia Guarnerio, John Nolan, Clotilde Théry, Andries Zijlstra, Shannon Stott, Sungyong You, Francesca Demichelis, Paul C. Boutros, Kendall Van Keuren-Jensen, and Dolores Di Vizio

Supplemental figures and figure legends

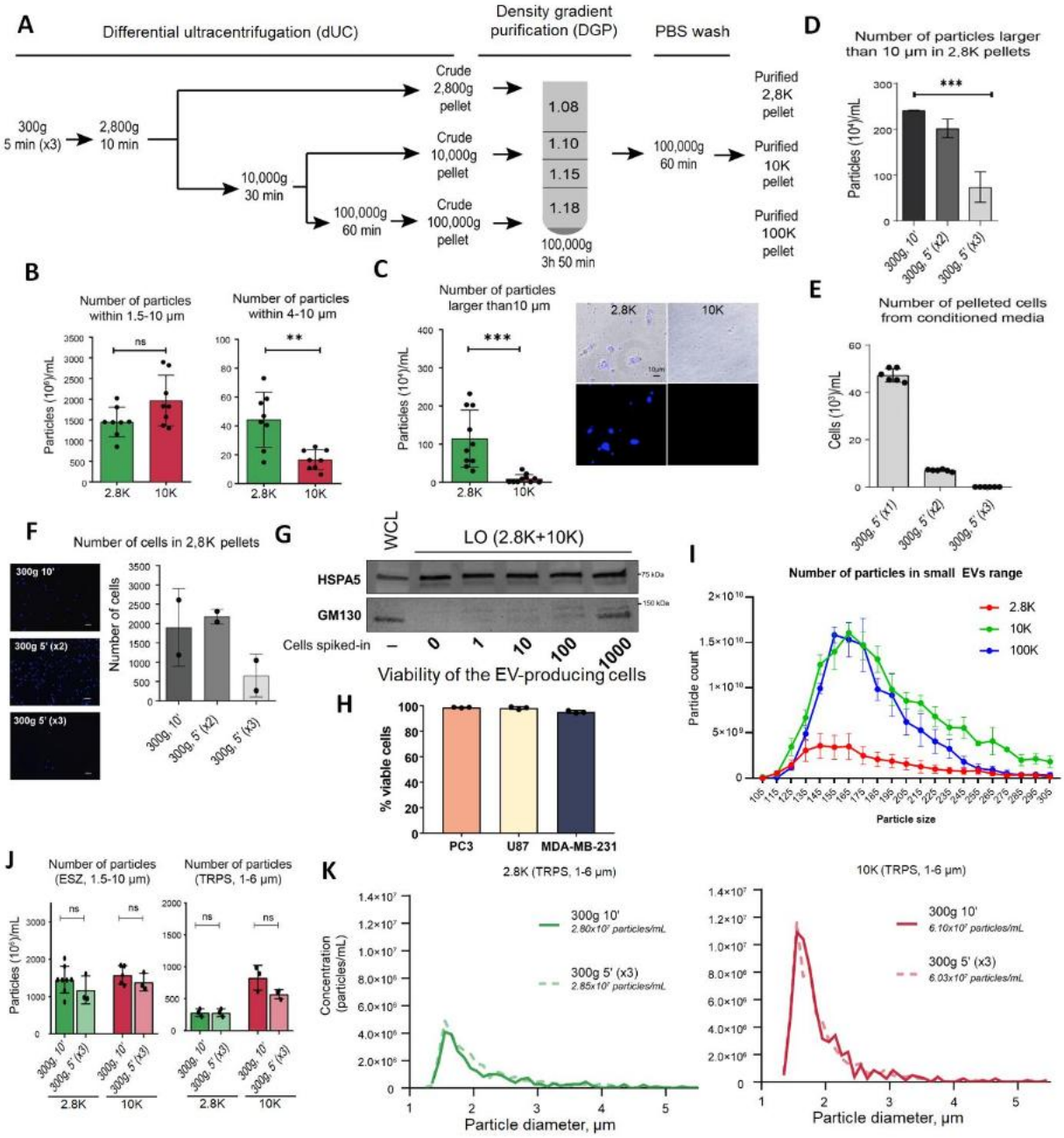


Figure S1. Differential centrifugation protocol optimization to exclude cells from the 2.8K fraction. Related to Figure 1. (**A**) Workflow for purification of three EV populations by differential ultracentrifugation (dUC) at different speeds (2,800g/2.8K, 10,000g/10K, 100,000g/100K) followed by density gradient purification (DGP). (**B**) The electrical sensing zone (ESZ) quantitation of particles within 1.5-10 µm (left) and 4-10 µm[18] in the 2.8K and 10K pellets obtained from PC3 cell conditioned media after depleting cells at 300g, 10 min. (**C**) (Left) ESZ quantitation of the particles >10 µm in the 2.8K and 10K pellets. Hoechst staining of the 2.8K and 10K pellets cultured for 24 hours showing the presence of some cells in the 2.8K pellet. Scale bar: 10 µm. (**D**) ESZ quantification of the particles >10 µm in 2.8K pellets following different combinations of low-speed centrifugation spins to deplete cells. (**E**) Bar plot showing the number of cells pelleted at sequential centrifugations spins at 300g, 5min, in 50 mL of PC3 conditioned

media. (**F**) (Top) Hoechst staining and (bottom) quantification of the cell numbers in the 2.8K pellets following different combinations of low-speed centrifugation spins to deplete cells. Scale bar: 50 µm. (**G**) Western blot analysis of LO preps spiked-in with increasing number of cells. HSPA5 (LO marker) and GM130 (cellular marker) were used to confirm the lack of cell contaminants in the LO preps and to determine de sensitivity of the contaminant detection. (**H**) Average cell numbers and viability in a 150 cm diameter culture dish after 24 hours of serum starvation, prior to collection of the conditioned media for EV isolation. (**I**) Total particle count of the three population of EVs from PC3 cells in the range of small EVs (100-300 nm) using TRPS. (**J**) Cell depletion at 3 x (300g, 5 min) does not affect the recovery of large particles in 2.8K and 10K pellets quantified by ESZ (left) and tunable resistive pulse sensing (TRPS, right). (**K**) TRPS quantification and particle size distribution of the 2.8K and 10K PC3 EVs before and after implementing an optimized protocol for cell depletion demonstrates that the optimized protocol does not alter the recovery of large EVs. Data represents mean ± standard deviation.

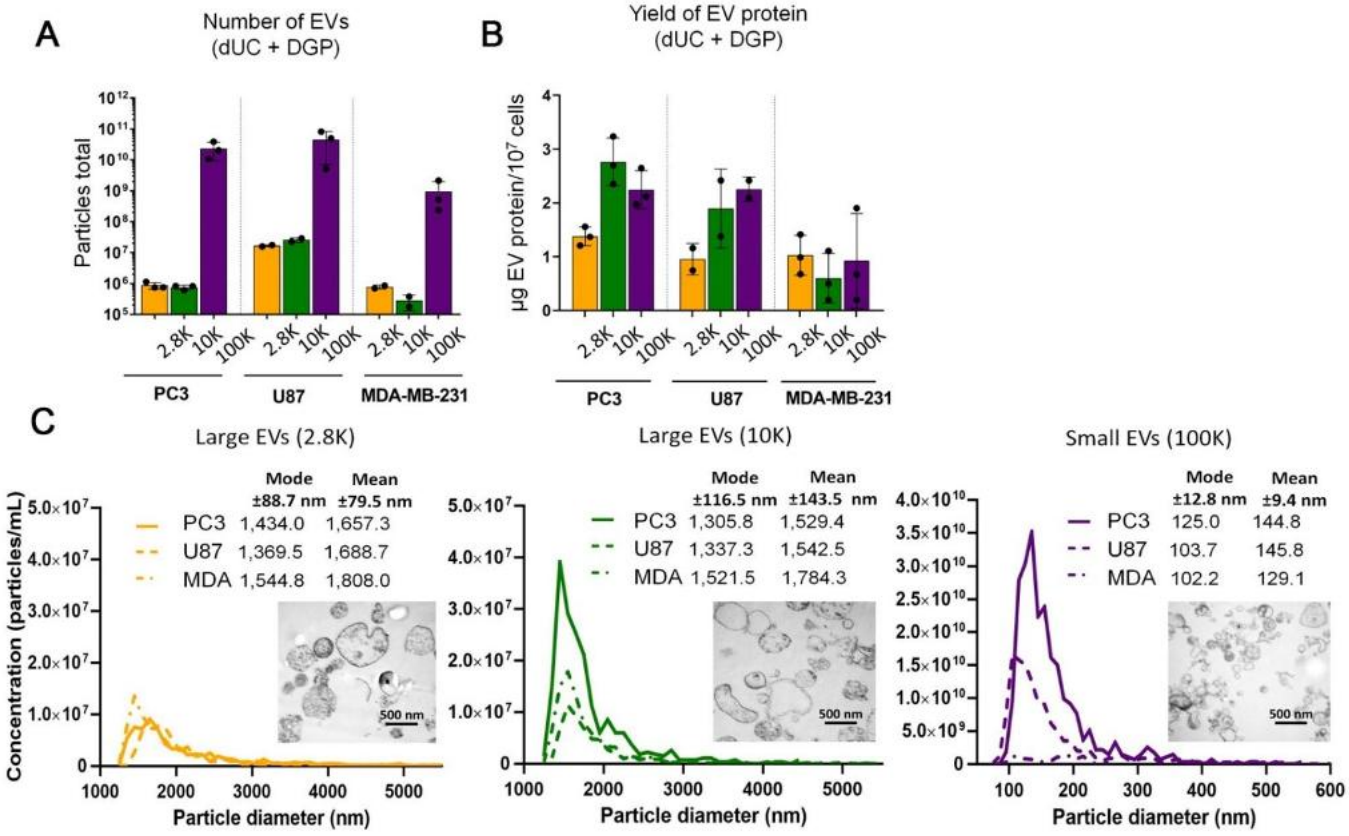


Figure S2. Low-speed centrifugation enriches for a population of L-EVs containing LO. **Related to Figure 1.** (**A**) Quantitation of the particles by Tunable Resistive Pulse Sensing (TRPS) and EV protein amount (**B**) in the indicated fractions from the indicated cell lines (see methods for details). (**C**) TRPS size distributions showing variable amounts of L-EVs and S-EVs in the three EV fractions from the three indicated cancer cell lines. In the insets, transmission electron microscopy (TEM) show enrichment of L-EVs in the 2.8K and 10K fractions (left and middle panel) and of S-EVs in the 100K fraction (right panel). Data represents mean ± standard deviation.

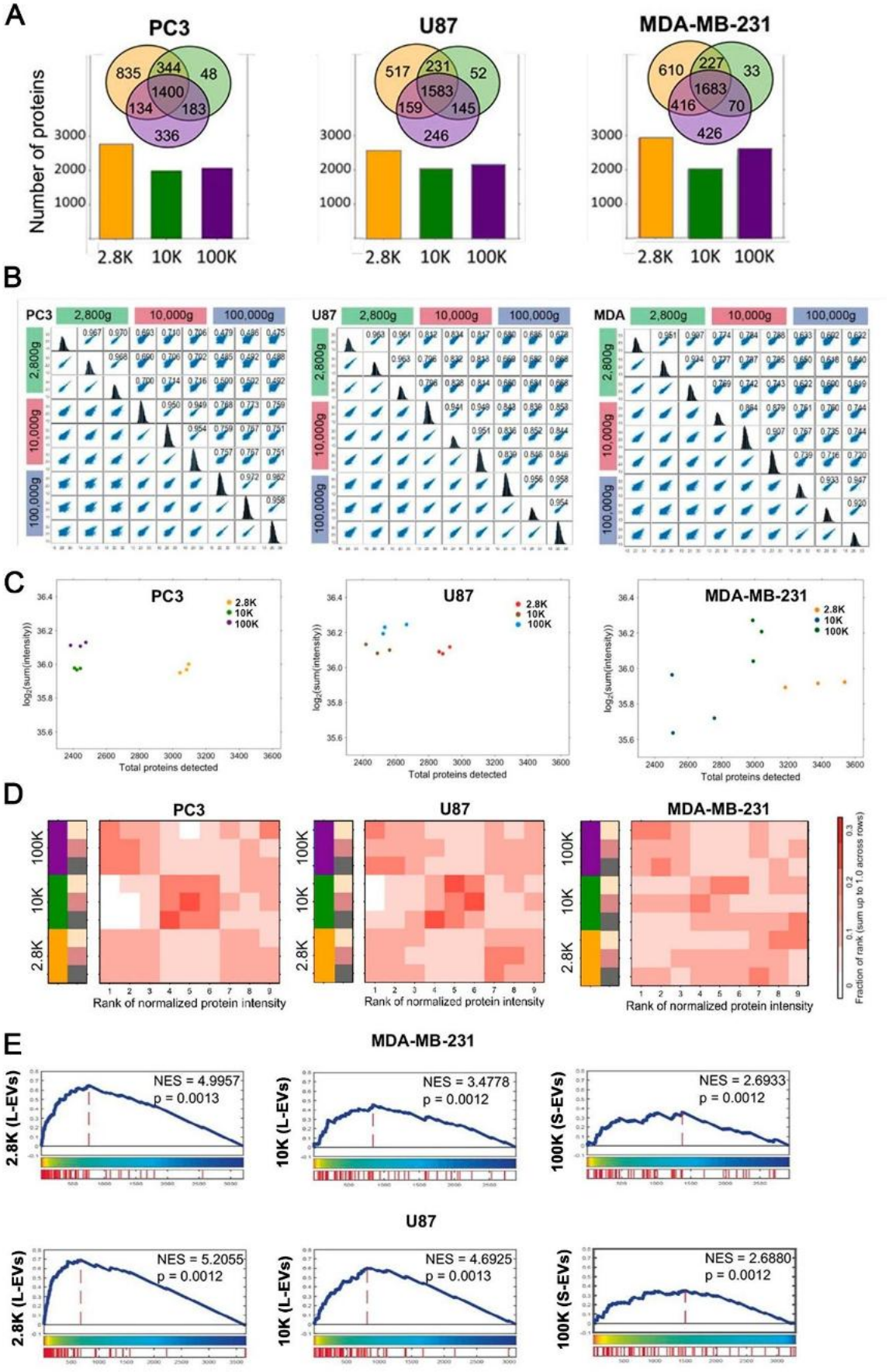

A
PC3
U87
MDA-MB-231
Number of proteins
835
344
48
1400
134
183
336
517
231
52
1583
159
145
246
610
227
33
1683
416
70
426
3000
2000
1000
2.8K
10K
100K
B
PC3
U87
MDA
2,800g
10,000g
100,000g
C
PC3
U87
MDA-MB-231
2.8K
10K
100K
log2(sum(intensity))
Total proteins detected
D
PC3
U87
MDA-MB-231
Rank of normalized protein intensity
Fraction of rank (sum up to 1.0 across rows)
E
MDA-MB-231
U87
2.8K (L-EVs)
10K (L-EVs)
100K (S-EVs)
NES = 4.9957
p = 0.0013
NES = 3.4778
p = 0.0012
NES = 2.6933
p = 0.0012
NES = 5.2055
p = 0.0012
NES = 4.6925
p = 0.0013
NES = 2.6880
p = 0.0012

Figure S3. Related to Figure 1. 10K EVs have a distinct protein intensity pattern compared to the 2.8K and 100K EVs. Related to Figure 2. (A) Venn diagrams showing the extent of overlapping proteins in each EV fraction (top) from the indicated cells lines. The bar plots summarize the total number of proteins in each fraction (**B**) Spearman's correlation plots and coefficients of the normalized LFQ signal among technical replicates in the indicated EV populations in PC3, U87, and MDA-MB-231 cells lines (p-value<0.001), and frequency histograms of the normalized LFQ signal for each technical replicate. (**C**) Association between the sum of protein intensity and number of detected proteins in the PC3, U87, and MDA-MB-231 cells lines. (**D**) Rank plots of normalized protein intensity showing a fraction of proteins in each rank across 9 samples for proteins detected in all three EV subtypes in PC3, U87, and MDA-MB-231 cells lines. Each row represents a sample, and each column is the fraction of proteins in each rank. (**E**) The gene set enrichment analysis (GSEA) of the proteins differentially enriched in large oncosomes[7] (FC>1 FDR<0.05) across different EV populations from breast cancer and glioma cell lines.

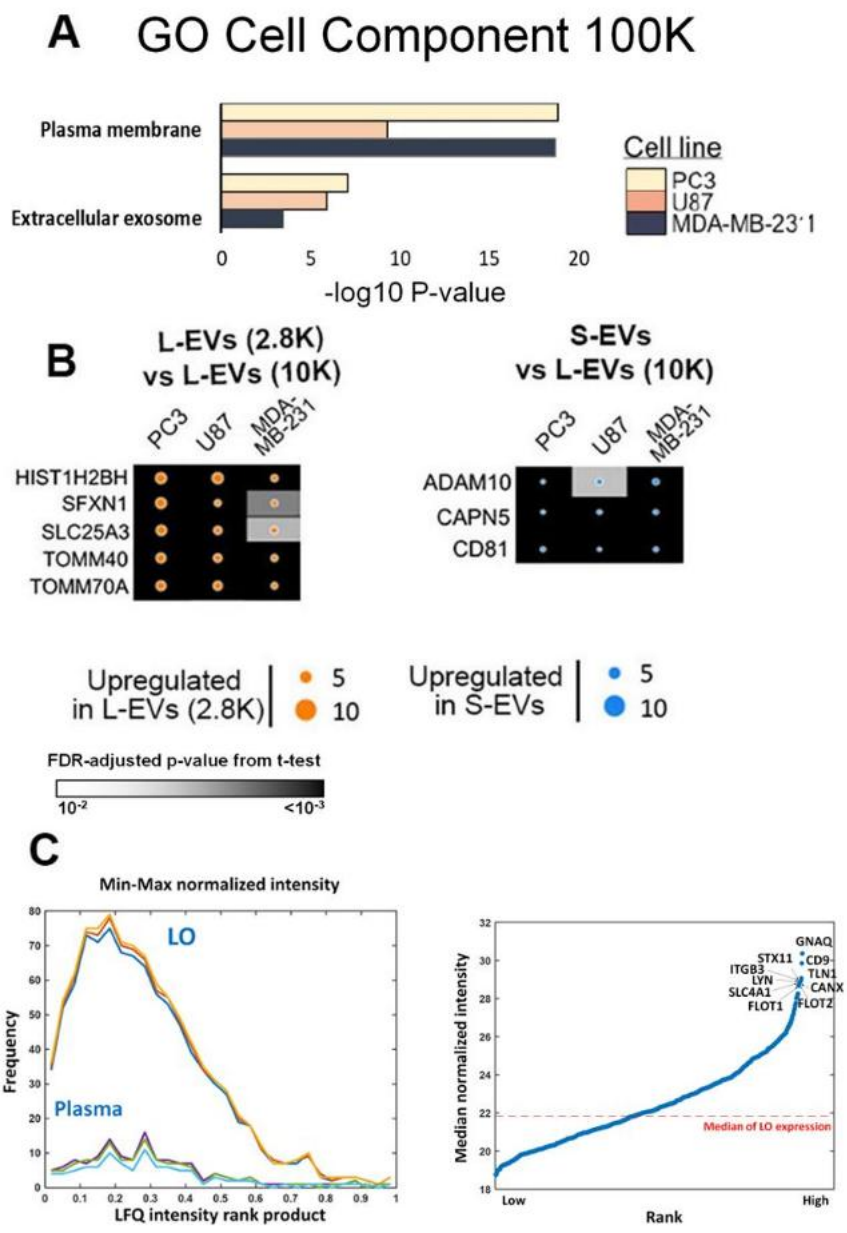
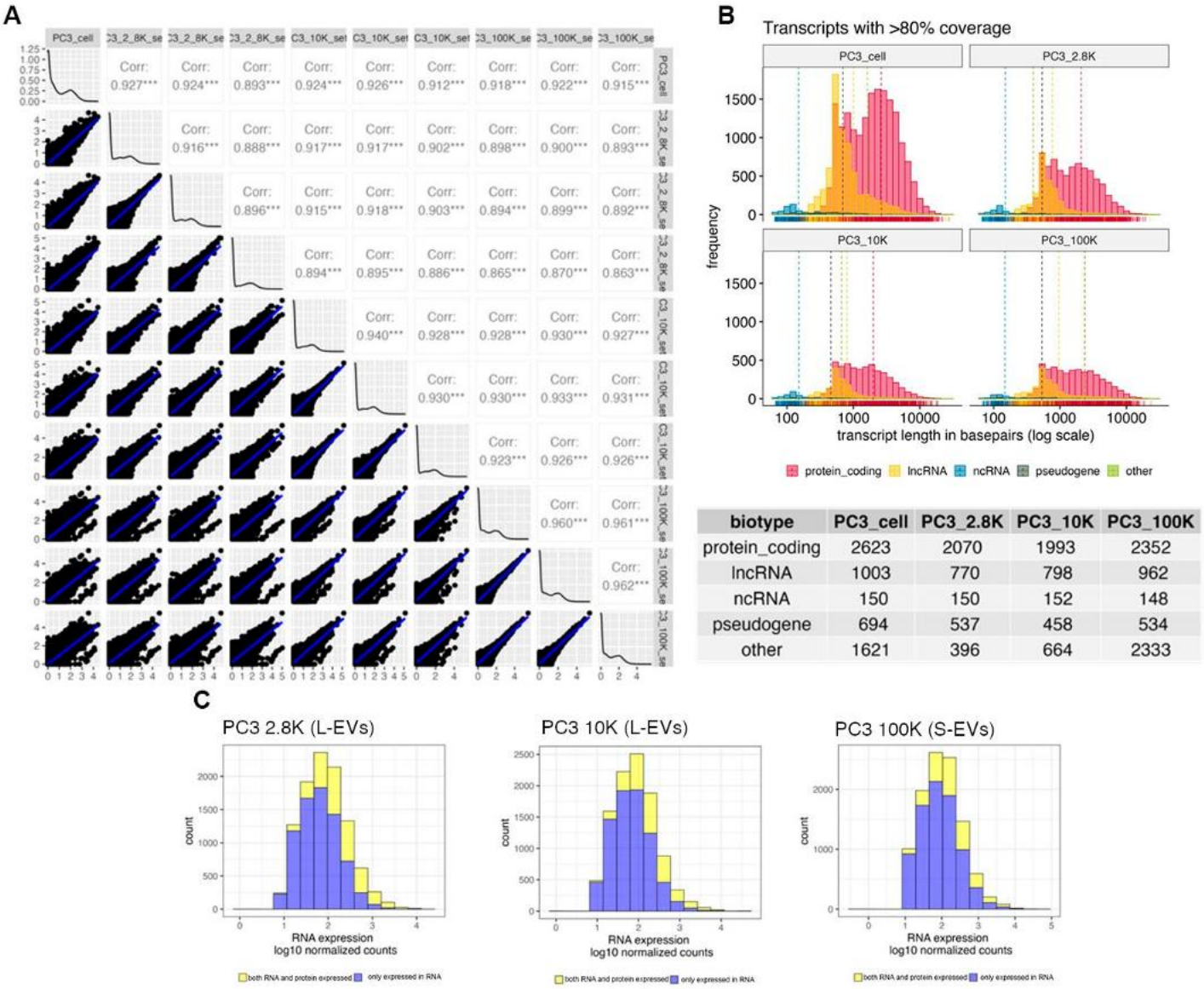


Figure S4. Related to Figure 2. The 2.8K and 100K contain the two most distinct EV populations on the proteome level. Related to Figure 2 and 3. (**A**) Gene ontology (GO) analysis for cell component on the proteins unique to 100K demonstrating the enrichment of the plasma membrane and exosome terms across the three cell lines (PC3, U87, and MDA-MB-231). (**B**) The most enriched (top 25%, fold change > upper quartile) and most highly expressed (top 25%, normalized expression > upper quartile) differentially abundant proteins in 2.8K *vs.* 10K and 100K *vs.* 10K common to the PC3, U87, and MDA-MB-231 cell lines. (**C**) Histogram of LFQ Minimum and maximum normalized intensity of all tree samples from each group (left). Normalized intensity of the 725 proteins detected in all the triplicates of LO with expression larger than low 5% and not detected in unfractionated plasma.



biotype	PC3_cell	PC3_2.8K	PC3_10K	PC3_100K
protein_coding	2623	2070	1993	2352
lncRNA	1003	770	798	962
ncRNA	150	150	152	148
pseudogene	694	537	458	534
other	1621	396	664	2333

Figure S5. Transcriptomic analysis reveals a similar RNA composition across the EV fractions. Related to Figures 4. (**A**) Spearman's rank correlation matrix of transcriptome replicates for the PC3 cell line. (**B**) Histogram of the transcript length of transcripts (TPM > 1) with >80% coverage in different sized EVs from the PC3 cell line (top). The mean transcript length (bp) is shown in the table (bottom). (**C**) Histogram of genes expressed only in the transcriptome (blue) and both the transcriptome and proteome (yellow) with RNA expression on the X-axis.

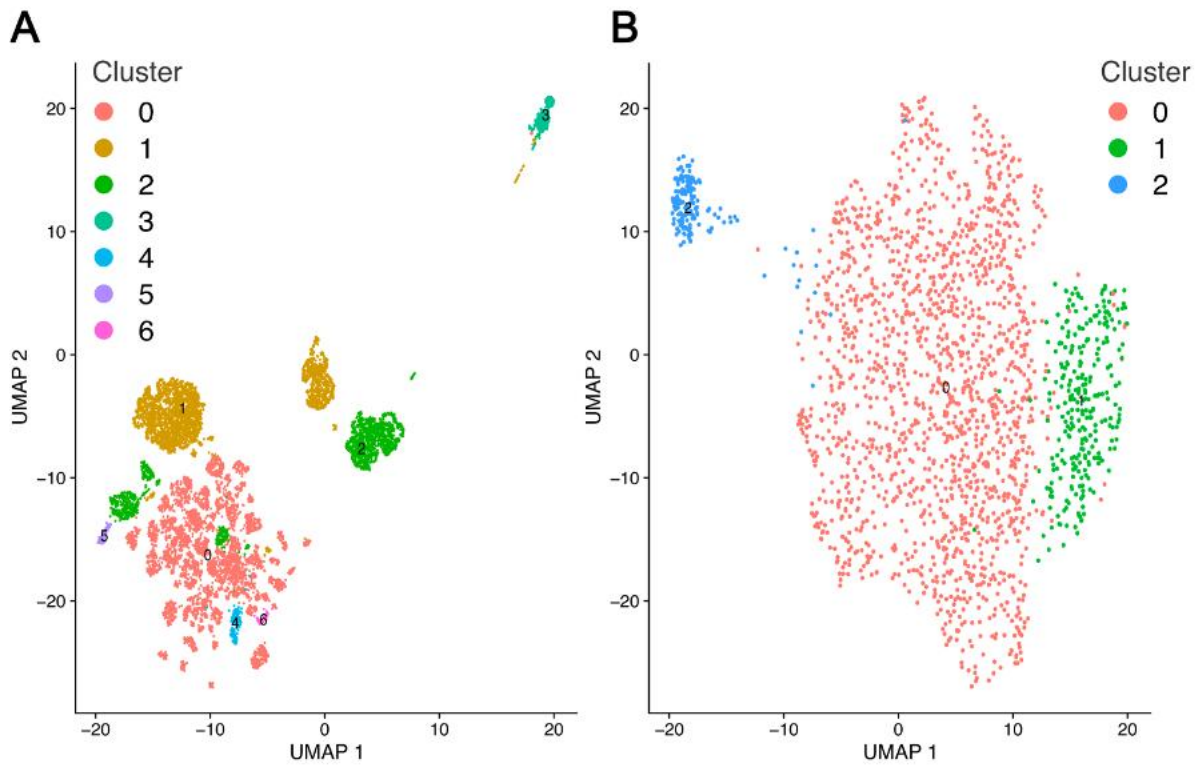


Figure S6: Single-EV RNA-seq analysis of L-EVs from PC3 cells. Related to Figure 5. (A) Dimensionality reduction of S-EV RNA sequencing data (only drops containing high number (≥50) of reads are shown). Drops are colored by cluster. **(B)** Dimensionality reduction of S-EV RNA sequencing data limited to vesicles positive for LO markers (only vesicles containing high number (≥50) of reads and positive for LO markers are shown). Drops are colored by cluster.

Supplemental tables

Abundant & not detected in non-fractionated human plasma by MS			
Detected only in patient-derived L-EVs	ADCY6	EPDR1	TMEM63
	ATP6V0A1	F2RL3	ATMX4
	CD68	P2RX1	TRPC6
	CD69	PLXNA4	TSPAN14
	CLEC1B	TBXA2R	TSPAN33
	ENPP4	TGFB1	
Detected in patient- and PC3-derived L-EVs	ABCC1	ITPR1	SPPL2A
	ABCC4	ITPR3	SYPL1
	ATL1	NCLN	TMEM30A
	CD151	PANX1	TOR1AIP2
	CD40	PNPLA6	TSPAN15
	CD63	PTPRA	TSPAN9
	CRTAP	RPN2	
	ERMP1	SLC44A2	

Supplemental table 1. Related to Figure 4. Candidate surface proteins in PC-derived L-EVs