Proteomic dissection of large extracellular surfaceome unravels interactive surface platform

Alin Rai[1,2,3,*], Haoyun Fang[1], Bethany Claridge[1,4], Richard J Simpson[4], David W Greening[1,2,3,4*]

[1]Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia. [2]Central Clinical School, Monash University, Melbourne, Victoria, Australia. [3]Baker Department of Cardiometabolic Health, University of Melbourne, Melbourne, Victoria, Australia. [4]Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria, Australia

*To whom correspondence should be addressed:
Alin Rai, PhD
Molecular Proteomics, Baker Heart and Diabetes Institute
75 Commercial Road, Melbourne, 3004, Australia
Tel: +61-3-8532-1217
Email: alin.rai@baker.edu.au

David W. Greening, PhD
Molecular Proteomics
Baker Heart and Diabetes Institute
75 Commercial Road, Melbourne, 3004, Australia
Tel: +61-3-8532-1585
Email: david.greening@baker.edu.au

ORCID
Alin Rai - 0000-0001-7994-5151; Richard Simpson - 0000-0002-9834-0796, David W Greening - 0000-0001-7516-485X; Bethany Claridge 0000-0002-4589-1033

Keywords
extracellular vesicles, surfaceome, surface proteins, proteomics, vesicle heterogeneity

Figures: 6, **Tables**: 2

Abstract

The extracellular vesicle (EV) surface proteome (surfaceome) acts as a fundamental signalling gateway by bridging intra- and extracellular signalling networks, dictates EVs' capacity to communicate and interact with their environment, and is a source of potential disease biomarkers and therapeutic targets. However, our understanding of surface protein composition of large EVs (L-EVs, 100-800 nm, mean 310 nm, ATP5F1A, ATP5F1B, DHX9, GOT2, HSPA5, HSPD1, MDH2, STOML2), a major EV-subtype that are distinct from small EVs (S-EVs, 30-150 nm, mean 110 nm, CD44, CD63, CD81, CD82, CD9, PDCD6IP, SDCBP, TSG101) remains limited. Using a membrane impermeant derivative of biotin to capture surface proteins coupled to mass spectrometry analysis, we show that out of 4143 proteins identified in density-gradient purified L-EVs (1.07-1.11 g/mL, from multiple cancer cell lines), 961 proteins are surface accessible. The surface molecular diversity of L-EVs include i) bona fide plasma membrane anchored proteins (cluster of differentiation, transporters, receptors and GPI anchored proteins implicated in cell-cell and cell-ECM interactions), and ii) membrane surface-associated proteins (that are released by divalent ion chelator EDTA) implicated in actin cytoskeleton regulation, junction organization, glycolysis and platelet activation. Ligand-receptor analysis of L-EV surfaceome (e.g., ITGAV/ITGB1) uncovered interactome spanning 172 experimentally verified cognate binding partners (e.g., ANGPTL3, PLG and VTN) with highest tissue enrichment for liver. Assessment of biotin inaccessible L-EV proteome revealed enrichment for proteins belonging to COPI/II-coated ER/Golgi-derived vesicles and mitochondria. Additionally, despite common surface proteins identified in L-EVs and S-EVs, our data reveals surfaceome heterogeneity between the two EV-subtype. Collectively, our study provides critical insights into diverse proteins operating at the interactive platform of L-EVs and molecular leads for future studies seeking to decipher L-EV heterogeneity and function.

Introduction

Extracellular vesicles (EVs), membranous vesicles released by cells into extracellular space, are laden with proteins and nucleic acids which they can transfer between cells to elicit functional response and therefore serve as an effective means of intercellular signalling, both in physiological as well as pathological conditions[1]. Critical to EV function are the EV surface proteins which regulate their interaction with the extracellular environment (recipient cells, ECM)[2], dictate biodistribution[3], half-life in circulation[4] and pharmacokinetics[5]. In cancer, EV surface proteins not only mobilize pro-metastatic bone marrow cells (through transfer of active surface MET)[6] but also regulate EV homing to specific organs (to lungs via surface integrins ITGA6/B4 and ITGA6/B1, and to liver via integrins ITGAV/B5)[3] to establish pre-metastatic niches and enhance metastasis. EV surface proteins (e.g., PD-L1) also regulate immune escape to promote tumorigenesis in lung cancer[7]. Moreover, EV surface proteins are not only a source of potential disease biomarkers[8,9] but also enable capture of EVs as liquid biopsy[10] and systemic clearance of pathogenic EVs[11] in the clinic. EV surface proteins also have therapeutic potential in bone healing[12], provide cardio protection[13] and rescue neuronal function impairment[14]. Recently, functionalization of EV surface proteins is emerging as an effective means to design vehicles that can potentially deliver therapeutic drugs to target sites[15-18] or to develop potential cancer vaccines[19]. Because EVs are released in high number, their collective surface area represents a large interactive platform, engaging proteins, lipids and glycans; understanding players of this dynamic interactome is thus pivotal.

Based on their size, EVs can be broadly categorized into large EVs (L-EVs, 100-1000 nm) and small EVs (S-EVs, 30-150 nm)[20,21]. While several studies have catalogued role of S-EVs in many pathologies (e.g., cancer[1], cardiovascular disease[22] and neurodegenerative diseases[23]), L-EVs - although relatively understudied compared to S-EVs - are now also starting to gain prominence for their signalling role, either complementing or contrasting S-EV function[24]. We have previously shown that L-EVs can be further segregated based on their varying densities, namely L-EVs displaying buoyant density of 1.07-1.11 g/mL which are distinct from L-EVs of relatively higher buoyant density (1.22-1.30 g/mL)[25]. While heavy density L-EVs represent midbody remnants of cytokinetic origin, our understanding of low-density L-EVs remains limited.

Besides their size-based categorization, it is important to note that cells release heterogeneous sub-populations of EVs[25-28] which can be categorised based on their origin (e.g., plasma membrane budding that give rise to shed microvesicles[8], EVs of endosomal origin called exosomes[1,29], shed midbody remnants derived from cytokinetic bridges[25], migrasomes released by migrating cells[30]), floatation density (e.g., light vs heavy density EVs[21,25]) and/or biochemical compositions[21,29] (e.g., EPCAM/A33$^+$ EVs[28]). Depending on isolation strategies employed, heterogeneous vesicles co-purify (for example, S-EVs can arise from plasma membrane budding or are endosomally-derived) due to overlapping characteristics[24,26]. Thus, while these subtype categorizations are useful at an operational level, there is growing awareness of EV heterogeneity and the need to study it. Fundamental questions regarding L-EV form and function remain unanswered, these include our insight into their surface proteome landscape.

By definition, cellular surfaceome encompasses all plasma membrane proteins with at least one amino acid residue exposed to the extracellular space[31]. Several studies have characterized S-EVs surfaceome[11,32-38]; we previously characterized surfaceome of S-EVs by proteolytically "shaving" surface proteins using proteinase K[35]. However, we noted that proteinase K treatment compromises L-EVs integrity, calling for an alternative strategy to define surface landscape of L-EVs. The focus of this paper is directed at defining surface protein landscape of L-EVs (buoyant density of 1.07-1.11 g/mL) by using membrane impermeant biotin to label and enrich for surface proteins and identify them using mass spectrometry.

Methods

Cell culture

SW620 (CCL-227, ATCC) and LIM1863 cells[39] (Ludwig Institute for Cancer Research, Melbourne) cells were cultured in RPMI-1640 (Life Technologies). MDA MB 231 (HTB-26, ATCC) and U87 (HTB-14, ATCC) cells were cultured in DMEM (Life Technologies). Complete culture media included media supplemented with 5% (v/v) Fetal Bovine Serum (FBS, Life Technologies) and 1% (v/v) Penicillin Streptomycin (Pen/Strep, Life Technologies) at 37 °C with 10% CO_2. Cells were passaged with trypsin-EDTA (Gibco).

Generation of cell conditioned media

Cells (SW620, LIM1863, MDA MB 231 and U87) were cultured in CELLine AD-1000 Bioreactor classic flasks (Integra Biosciences) as previously described[25]. Cells (3×10^7) in 15 ml of complete culture media were added to the lower cell-cultivation chamber for 72 h at 37 °C with 10% CO_2. The upper nutrient supply chamber contained 500 ml RPMI or DMEM (5% FBS, 1% Pen/Strep) that was replaced every 4 days. The lower chamber was washed three times with serum-free media. For LIM1863 cells, which grow as floating organoids, cells were recovered and washed in serum free media (150 *g*, 5 mins) and re-introduced back into the lower chamber. Cells in the lower chamber were then cultured in 15 ml of media supplemented with 0.5% (v/v) insulin transferrin selenium (Invitrogen) and 1% Pen/Strep). Thereafter, conditioned medium (CM) in the cell cultivation chamber was collected every 2 days.

Isolation of large and small EVs

CM was centrifuged at 500 *g* (5 min, 4 °C) and 2,000 *g* (10 min, 4 °C) and the supernatant either processed for EV isolation or stored at -20°C until further use. The supernatant was centrifuged at 10,000 *g* (30 min, 4 °C, SW28 rotor; Optima XPN Ultracentrifuge, Beckman Coulter) to pellet crude L-EVs and then at 100,000 *g* (1 h, 4 °C, 41 Ti rotor; Optima XPN Ultracentrifuge) to pellet crude S-EVs. EV-pellets were resuspended in ~200 µl PBS and subjected to top-down isopycnic (iodixanol-density) ultracentrifugation[28,40], whereby EVs were then over (discontinuous gradient of OptiPrep™ (40% (3 ml), 20% (3 ml), 10% (3 ml) and 5% (2.5 ml) (diluent: 0.25 M sucrose / PBS solution)) and ultracentrifuged at 100,000 *g* for 18 h (4 °C, 41 Ti rotor; Optima XPN Ultracentrifuge). Twelve equal fractions were collected, diluted in PBS (2 ml) and centrifuged at either 10,000 *g* (30 min, 4 °C, Eppendorf 5430R) or 100,000 *g* (1 h, 4 °C, TLA-55 rotor; Optima MAX-MP Tabletop Ultracentrifuge) to collect L-EV and S-EV containing fractions, respectively. Density of each fraction was determined as previously described[25]. Pellets were further washed in PBS, reconstituted in PBS and stored at -80 °C until further use.

Immunoblotting

Protein quantification (microBCA™ Protein Assay Kit (23235, Thermo Fisher Scientific)) and Western blotting (iBlot 2 Dry Blotting System, Thermo Fisher Scientific) were performed as per manufacturer's instructions. Dot blot analysis was performed using 96-well Bio-Dot (Bio-Rad Laboratories) as per manufacturer's instructions with proteins were lysed in 50 mM HEPES (1% SDS). Rabbit antibodies raised against, MET (Santa Cruz Biotechnology), CD63 (Santa Cruz), ANXA1 (Abcam), GAPDH (Cell Signalling), and GFP (Abcam) were used.

Mouse antibodies ALIX (BD Biosciences), TSG101 (BD Biosciences) were used. Secondary antibodies used were IRDye 800 goat anti-mouse IgG or IRDye 700 goat anti-rabbit IgG (1:15000, LI-COR Biosciences).

Biophysical characterization of EVs

Cryo-electron microscopy (Tecnai G2 F30) on EVs (2 µg) was performed as described[20]. Vesicle particle size was determined using a NanoSight NS300, Nanoparticle tracking analysis (NTA) (Malvern) system fitted with a NS300 flow-cell top plate with a 405 nm laser as described[25]. Samples (1 µg μl^{-1}) were diluted in 500 µl PBS (1:10,000) and injected using 1 ml syringes (BD Biosciences) (detection threshold = 10, flowrate = 50, temperature = 25 ºC). Each analysis consisted of 60 s video captures. Data was analysed using NTA software 3.0 (Malvern).

Global proteomic sample preparation of EVs

Global mass spectrometry-based proteomics of EVs (10 µg in 50 µL) was performed as previously described[41] using single-pot solid-phase-enhanced sample preparation (SP3) method[42]. Briefly, samples were solubilised in 1% (v/v) sodium dodecyl sulphate (SDS), 50 mM HEPES pH 8.0, incubated at 95 °C for 5 mins and cooled. Samples were reduced with 10 mM dithiothreitol (DTT) for 45 min at 25 °C followed by alkylation with 20 mM iodoacetamide for 30 min at 25 °C in the dark. The reaction was quenched to a final concentration of 20 mM DTT. Magnetic beads were prepared by mixing SpeedBeads™ magnetic carboxylate modified particles (65152105050250, 45152105050250, Cytiva) at 1:1 (v:v) ratio and washing twice with 200 µL MS-water. Magnetic beads were reconstituted to a final concentration of 100 µg/µl. Magnetic beads were added to the samples at 10:1 beads-to-protein ratio and 100% ethanol (EA043, ChemSupply) added for a final concentration of 50% ethanol (v/v). Protein-bound magnetic beads were washed three times with 200 µl of 80% ethanol and reconstituted in 50 mM TEAB and digested with trypsin (Promega, V5111) at a 1:50 enzyme-to-substrate ratio for 16 h at 37 °C with constant shaking (1000 rpm). The peptide mixture was acidified to a final concentration of 2% formic acid (pH ~1-2) and centrifuged at 20,000*g* for 1 min. The peptide digests were frozen at -80°C and dried by vacuum centrifugation (Savant SPD121P, Thermo Fisher Scientific), reconstituted in 0.07% trifluoroacetic acid, and quantified by Fluorometric Peptide Assay (23290, Thermo Fisher Scientific) as per manufacturer's instruction.

Surface biotin-labelling of EVs and proteomic sample preparation

EV surface proteins were biotinylated using Pierce™ Cell Surface Biotinylation and Isolation Kit (A44390, Thermo Fisher Scientific) as per manufacturer's recommendation. Briefly, 100 µg EVs (L-EVs or S-EVs) were labeled with 0.25 mg/mL EZ-Link Sulfo-NHS-SS-Biotin for 10 mins at room temperature. EVs were then washed twice with ice-cold TBS (10,000 g for 30 mins for L-EVs or 100,000 g for 1 h for S-EVs). Labelled EVs were then lysed (30 mins on ice) as per manufacturer's instructions. To assess bound and released EV surface proteins, biotinylated EVs in PBS were treated with 5 mM EDTA for 15 min at room temperature[43]. Samples were centrifuged at either 10,000 *g* (30 min, 4 °C, Eppendorf 5430R) or 100,000 *g* (1 h, 4 °C, TLA-55 rotor; Optima MAX-MP Tabletop Ultracentrifuge) to collect L-EV and S-EV with their membrane bound proteins (pellet fraction), respectively. Supernatants (containing the released proteins) were processed for SP3-based tryptic protein digestion (as above). The pellet fractions were resuspended in 1 ml PBS, re-centrifuged and pellet fractions reconstituted in 50 µl of PBS.

Biotin-labelled proteins were then captured onto NeutrAvidin Agarose slurry (30 mins at room temperature with end-over-end mixing on a rotor). Samples were loaded onto epTIPS (Eppendorf, 200 µl) fitted with 20 µm nylon net (NY2004700, Merck Millipore), washed 3 times as per manufacturers instruction, and reduced in 10 mM DTT in 100 mM triethylamonium bicarbonate (TEAB) for 45 min at 25°C. Eluted proteins were then alkylated with 20 mM iodoacetamide (IAA) for 30 min at 25°C in the dark. The reaction was quenched to a final concentration of 20 mM DTT and digested with trypsin (Promega, V5111) at a 1:50 enzyme-to-substrate ratio for 16 h at 37 °C. The resultant peptides were acidified to a final concentration of 2% formic acid (FA), peptides desalted using SDB-RPS Stage-Tips[44] followed by elution with 30-80% acetonitrile (ACN), 0.1% trifluoroacetic acid, and dried by vacuum centrifugation. Peptides were reconstituted in 0.07% trifluoroacetic acid and quantified by Fluorometric Peptide Assay (23290, Thermo Fisher Scientific). SP3-based tryptic protein digestion for global proteome was performed as previously described[45].

Proteomic liquid chromatography–tandem mass spectrometry

Peptides were analysed on a Dionex UltiMate NCS-3500RS nanoUHPLC coupled to a Q-Exactive HF-X hybrid quadrupole-Orbitrap mass spectrometer equipped with nanospray ion

source in positive mode as described.[46,47] Peptides were loaded (Acclaim PepMap100 C18 3 μm beads with 100 Å pore-size, Thermo Fisher Scientific) and separated (1.9-μm particle size C18, 0.075 × 250 mm, Nikkyo Technos Co. Ltd) with a gradient of 2–28% acetonitrile containing 0.1% formic acid over 95 mins at 300 nl min^{-1} followed by 28-80% from 95-98 mins at 300 nL min-1 at 55°C (butterfly portfolio heater, Phoenix S&T). An MS1 scan was acquired from 350–1,650 *m*/*z* (60,000 resolution, 3×10^6 automatic gain control (AGC), 128 msec injection time) followed by MS/MS data-dependent acquisition (top 25) with collision-induced dissociation and detection in the ion trap (30,000 resolution, 1×10^5 AGC, 60 msec injection time, 28% normalized collision energy, 1.3 *m*/*z* quadrupole isolation width). Unassigned, 1, 6-8 precursor ions charge states were rejected and peptide match disabled. Selected sequenced ions were dynamically excluded for 30 sec. Data was acquired using Xcalibur software v4.0 (Thermo Fisher Scientific). The MS-based proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository and are available via ProteomeXchange with identifier (PXD026658).

Data processing

MaxQuant (v1.6.6.0) with its built-in search engine Andromeda[48] was used to perform peptide identification and quantification as described[46]. Human-only (UniProt #74,823 entries) sequence database (Jan 2020) with a contaminants database was employed. Cysteine carbamidomethylation was set as a fixed modification and N-terminal acetylation and methionine oxidations as variable modifications; for biotin surface proteome analysis additional Thioacyl (DSP) (C(3)H(4)OS)) was employed. False discovery rate (FDR) was 0.01 for protein and peptide levels. Enzyme specificity was set as C-terminal to arginine and lysine using trypsin protease, and a maximum of two missed cleavages allowed. Peptides were identified with an initial precursor mass deviation of up to 7 ppm and a fragment mass deviation of 20 ppm. Protein identification required at least one unique or razor peptide per protein group. Contaminants, and reverse identification were excluded from further data analysis. 'Match between run algorithm' in MaxQuant[49] and label-free protein quantitation (maxLFQ) was performed. All proteins and peptides matching to the reversed database were filtered out.

For both L-EVs and S-EVs, we investigated 3 biological replicates for global proteome (3 cell lines: LIM1863, SW620, U87), 3 biological replicates for biotin capture proteome (3 cell lines: SW620, MDA MB 231, U87), 2 biological replicates for EDTA bound biotin capture (2 cell

lines: MDA MB 231, SW620), 2 biological replicate for EDTA released biotin capture proteome (2 cell lines: MDA MB 231, SW620).

For global surface proteome (**Supplementary Table 1**), for each cell line, protein was identified at least once in one biological replicate and detected at least once in EVs from at least two cell lines. For biotin-captured surfaceome, EVs from 3 cell lines were analysed, with 3 biological replicates for each cell line. Proteins selected for downstream bioinformatics analysis include those that were identified in at least 2 biological replicates, irrespective of the donor cells. For biotin-captured surfaceome that are removed by EDTA, EVs from 2 cell lines were analysed, with 2 biological replicates for each cell line. Proteins selected for downstream bioinformatics analysis include those that were identified in at least 2 biological replicates, irrespective of the donor cells.

Experimental parameters are submitted to EV-TRACK knowledgebase (EV-TRACK ID: EV210261)[50].

Bioinformatics and statistics

Cellular surfaceome data used include proteins previously experimentally verified as a cell-surface (CSPA)[51] or predicted as surfaceome proteins based on SURFY[31]. Voronoi tree maps (i.e., *in silico* surfaceome tree map) were generated using http://wlab.ethz.ch/surfaceome. Venn diagrams were created using www.interactivenn.net. Gene Ontologies, KEGG and Reactome pathways were obtained using g:Profiler[52] or DAVID analysis[53]. Cytoscape[54] was used to generate EnrichmentMap[55] (plugin v3.7.1) and GeneMania-based radial interaction map[56] (plugin v3.5.1). Sankey diagram of ligand-receptor interactions was generated using http://www.rna-society.org/cellinker/[57]. Volcano plot, principal component analysis plot, Pearson correlation matrix and hierarchical clustering was performed using Perseus[58]. Data were analyzed and bar plots/violin plots generated using GraphPad Prism (v8.0.1) or Microsoft Excel. One-way ANOVA (multiple comparisons) test was performed using Perseus and statistical significance defined at $p < 0.05$.

Fluorescence microscopy

Immunofluorescence was performed, as previously described[25]. Generation of SW620 cells stably-expressing plasma membrane-targeting Growth Associated protein 43 or GAP43 (1-20 a.a.)[59]fused to GFP (SW620-GAP-GFP cells) is described previously[25]. Briefly, cells (cultured

on Nunc® Lab-Tek® Chamber Slide™ (Sigma-Aldrich) to 60-80% confluency) fixed (4% formaldehyde for 5 min), permeabilized (0.2% (v/v) Triton X-100 in PBS, 5 min) and blocked (3% (w/v) bovine serum albumin (BSA, Sigma) in TTBS (0.2% (v/v) Triton X-100) (blocking solution) for 30 min at room temperature. Cells were then incubated with primary antibodies (1:100) MET (Santa Cruz Biotechnology), ALIX (BD Biosciences), TSG101 (BD Biosciences) or EEA1 (Cell Signalling) antibodies in blocking solution for 1 h at room temperature. Cells were washed and incubated with secondary antibodies (1:200) (Alexa Fluor 488-conjugated goat anti-mouse IgG or Alexa Fluor 568-conjugated goat anti-rabbit IgG (Invitrogen) in blocking solution for 20 min at room temperature (in the dark). Cells were washed three times in TTBS. Where indicated, nuclei were stained with Hoechst stain (10 μg ml^{-1}) for 1 min. Cells were imaged using a Zeiss AxioObserver Z1 microscope (Zeiss) or Zeiss Confocal LSM 780 PicoQuant FLIM (Zeiss) and images were analysed using Zen 2011 (Blue edition, Zeiss).

Results

Isolation and characterization of Large EVs

We isolated large EVs (L-EVs) and small EVs (S-EVs) released by 4 different cell lines using differential centrifugation coupled to density gradient separation **(Fig. 1A)** as previously described[28,60,61] and characterized them (for their buoyant densities, size, morphology, presence of specific proteins) to meet the experimental requirements as set out by the International Society for Extracellular Vesicles (ISEV) guidelines[26]. While both L-EVs and S-EVs displayed buoyant densities of 1.07-1.11 g/mL and were positive for EV marker ANXA1 **(Supplementary Fig. 1)**, S-EVs were enriched in CD63 consistent with previous reports[25,60,61]. Cryo-EM revealed that both EV subtypes were spherical in shape and morphologically intact with L-EVs (mean 310 nm) significantly ($p<0.0001$) larger than S-EVs (mean 110 nm) **(Fig. 1B-C**, Interestingly, there was no striking difference in mode size (i.e., size of the majority of EVs in the samples, e.g. 185.5 nm (L-EVs) vs 168.7 nm (S-EVs) for MDA-MB-231, **Supplementary Fig. 2-3**); this could be due to bias of NTA towards certain particle size ranges (especially 50–150 nm[26,62]. However, NTA analysis of EVs from three cell lines suggests that percentage of EVs >200 nm was higher in L-EV vs S-EVs, similar to cryoEM analysis of EVs from SW620 cells (**Supplementary Fig. 4)**. Although the two EV subtypes from different cell

lines clustered based on donor cells **(Supplementary Fig. 5)**, mass-spectrometry revealed that S-EVs are enriched in classical markers of small EVs[20,21,25,63], (CD44, CD63, CD81, CD82, CD9, PDCD6IP, SDCBP, TSG101) (**Fig. 1D-F, Supplementary Fig. 6, Supplementary Table 1-2**) whereas L-EVs were enriched in previously reported large EV proteins[20,21,25,64] (ATP5F1A, ATP5F1B, DHX9, GOT2, HSPA5, HSPD1, MDH2, STOML2), ribosomal (RPS5, RPL9/35), endoplasmic reticulum (CANX, CALR, TMEM33), mitochondrial (TIM44, ALDH2, PUS1), spliceosomal (DDX5, DDX39B, SNRPD1/2/3) and ribonucleoproteins (HNRNPU, HNRNPK) **(Fig 1D-F, Supplementary Fig. 6, Supplementary Table 1-2)**, as previously described[25,61].

Workflow for defining EV surfaceome

The workflow used in this study to define EV surfaceome using quantitative mass spectrometry-based proteomics is outlined in **Fig. 2**. We mapped experimentally verified cell-surface proteins (Cell Surface Protein Atlas (CSPA))[51] and surfaceome proteins based on cell surfaceome predictor SURFY[31] onto global EV protein profiles. Next, we verified their EV surface localization by enriching surface proteins using amine reactive, membrane-impermeant and thiol-cleavable Sulfo-NHS-SS-Biotin and subsequently identifying using mass spectrometry. Because EDTA can successfully chelate metal ions such as calcium to release surface bound proteins (such as ANXA1[43]) without compromising EV integrity[43,65], we also incubated EVs with EDTA and define the released proteins as the "EV surface-associated proteome". This led to the identification of 291 and 456 surface, and 122 and 67 surface-associated proteins in L-EVs and S-EVs, respectively.

Large EVs contain a sub-set of cellular surfaceome

We first interrogated distribution of known cellular surface proteins in global proteome of both EV-subtypes isolated from three different cell lines (SW620, LIM1863, U87) **(Fig. 3, Supplementary Fig. 5)**. In L-EVs, 2539/4143 proteins were detected in at least 2/3 datasets, of which 284 are cellular surfaceome (CSPA proteins[51]) and 190 predicted as surfaceome (SURFY proteins[31]) **(Fig. 3A)**. In S-EVs, 2449/3086 proteins were detected in at least 2/3 datasets, of which 295 were CSPA and 236 SURFY proteins **(Fig. 3B)**. Mapping L-EVs and S-EVs data onto the functionally annotated *in silico* surfaceome tree map revealed striking diversity in protein classes identified, which include transporters, adhesion molecules

(cadherins), integrins, ephrins, receptors (including RTKs) and proteases **(Fig. 3C-D)**. These proteins include bona fide cell-surface proteins such as cluster of differentiation (CD) proteins (54 in L-EVs and 60 in S-EVs), GPI-anchored proteins (10 in L-EVs and 11 in S-EVs) that do not have transmembrane domains, integrins and membrane transporters **(Fig. 3C-D, Supplementary Table 1)**. Of 638 human surfaceome identified between both EV-subtypes, 61 and 57 were significantly enriched ($q<0.05$) in L-EVs and S-EVs, respectively **(Fig. 3E, Supplementary Table 1)**, with their gene ontology (biological process) enrichment analysis presented in **Supplementary Fig. 7** Interestingly, we also found that 170/190 SURFY proteins in L-EVs displayed differential abundance (FDR < 0.05) in L-EVs from 3 different cell lines (SW620, LIM1863, U87, **Supplementary Fig. 8A, Supplementary Table 3**). Moreover, several of these proteins (with higher abundance in L-EVs from U87 vs SW620 and LIM1863) were also detected in greater abundance (e.g., APLP2, ERMP1, HM13, LMAN2, RPN1) in U87 L-EVs as compared to U87 S-EVs (**Supplementary Fig. 8A**). Amongst proteins with higher abundance in L-EVs from CRC cell lines (SW620 or LIM1863) vs U87 L-EVs, none displayed higher abundance when compared with CRC S-EVs, however several of these proteins (MUC13, CDH17, ACE, ACE2, TSPAN8) are reported as intestine-enriched proteins in Human Protein Atlas[66] (**Supplementary Fig. 8B**).

Biotin capture of large EVs surface proteome

Next, we employed Sulfo-NHS-SS-Biotin-based capture coupled with proteomic profiling to ascertain surface localization and identify additional EV surface proteins (not annotated as cellular surfaceome by CSPA/SURFY) **(Fig. 4, Supplementary Table 4)**. In L-EVs from different cell lines (SW620, MDA MB 231, U87), 291 proteins were detected in at least 2 data sets, whereas 456/1365 proteins were identified for S-EVs **(Fig. 4A, Supplementary Fig. 9A)**.

An important feature of membrane anchored proteins is the presence of membrane spanning transmembrane domain (TM). Upon cursory inspection, within the biotin-captured L-EV proteome, 36/291 proteins were annotated by UniProt to contain at least single TM region and 37/291 as surface PM protein by SURFY (**Fig. 4A, Supplementary Table 4**); proteins ranged from 'containing single' to multiple TM domains (e.g., 1 TM for EPHA, 4 for ABCC1, 13 for NPC1L1) (**Supplementary Table 4**). In contrast, 238/291 L-EVs surface biotin proteins not predicted to contain TM domain (by UniProt/SURFY) are cellular surface proteins demonstrated by CSPA **(Fig. 4A)**.

Overall, 86/291 proteins are CSPA/SURFY proteins (i.e. plasma membrane proteins) *versus* 5/291 annotated as non-surface membrane proteins by SURFY i.e. intracellular origin; this suggests that L-EVs mainly arise from plasma membrane (**Fig. 4A, Supplementary Table 4**). We refer to CSPA/SURFY proteins (86/291) as "classical surface proteins" and the rest 205/291 as "non-classical surface proteins". Cursory inspection revealed that classical proteins were mainly membrane anchored proteins (e.g., integrins, solute carrier (SLC) transporters, cluster of differentiation (CDs)), whereas non-classical were proteins that themselves lack TM domain but help in organization of junction proteins (e.g., DSG1, RAC1, RHOA, CTNNA1) and secreted proteins such as annexins, chemokine (S100A8), 14-3-3 proteins and enzymes (PGK1) (**Supplementary Table 4**).

Differential abundance of biotin surface proteins in S-EVs and L-EVs is also provided in **Supplementary Fig. 5, Supplementary Table 4**. We note that compared to L-EVs, several receptors were enriched in S-EVs **(Supplementary Fig. 9B-E)**. Moreover, fluorescence microscopy revealed, that consistent with their enrichment in S-EVs, MET (a receptor tyrosine kinase) co-localized with endosomal/exosomal markers (EEA1, ALIX and TSG101) **(Supplementary Fig. 9F)**, which supports their active sorting into S-EVs.

Next, we compared L-EV surface proteome with global EV proteome; biotin captured proteomes were significantly enriched in CD proteins, receptors and transporters, which further demonstrates successful enrichment of membrane surface proteins **(Fig. 4B)**. Importantly, 75 proteins that were annotated as CSPA/SURFY surface proteins in the global L-EV proteome, we verify as surface accessible proteins **(Fig. 4C)**. An additional 168 proteins in global L-EV proteome not annotated as cell surface proteins were also biotin accessible **(Fig. 4C)**. Biotin capture also enabled identification of an additional 48 surface proteins which were not detected in the global EV proteome (**Supplementary Table 5**).

Bioinformatic assessment of biotin accessible/inaccessible proteomes in L-EVs

We next constructed EnrichmentMap of Cellular Ontologies in 291 biotin captured proteins vs 2052 biotin inaccessible proteins **(Fig. 4D)**. Biotin captured proteins were enriched for "plasma membrane" proteins, "membrane raft" proteins, "cell projection", "leading edge" and "trailing edge" **(Fig. 4D, Table 1**), which again suggests that these L-EVs likely arise from budding of

plasma membrane. KEGG/Reactome terms enriched include “cell adhesion”, “tight junction”, “glycolysis”, “MAPK signalling” and “cell-ECM interactions” **(Fig. 4E, Table 1, Supplementary Table 6)**.

We argue that biotin inaccessible proteins represent luminal proteins. In this regard, biotin inaccessible proteins were enriched for “COPI/II-coated vesicles”, “proteasome”, “ribosomes”, “EIF complex”, “spliceosome” and “ER/Golgi” **(Fig 4D, Table 1)**. These proteins were implicated in “translation”, “protein processing in ER” and “MAPK signalling cascade” **(Fig. 4E)**. While how these proteins/structures are sorted into EVs is not understood, enrichment of “membrane trafficking”, and “vesicle-mediated transport” in biotin inaccessible proteome supports for an active sorting mechanism.

Surface associated proteins of L-EVs

Previous studies have shown that a pool of proteins can associate with EV surface and are functional[43,67-69]. Presence of CSPA/SURFY proteins (86/291 as classical surface proteins) *versus* 205/291 as non-CSPA/SURFY proteins strongly suggests that L-EVs also contain a subset of proteins that are associated on the surface. Indeed, we found that EDTA-treatment effectively released 122 proteins (now referred to as EV surface-associated) from the surface of L-EVs **(Fig. 5A, Table 2, Supplementary Table 4)**, including annexins and 14-3-3 proteins (**Fig. 5B**), but not membrane proteins including CDs, receptors and transporters **(Fig. 5C)**.

EV surface-associated proteome included both classical and non-classical surface proteins **(Fig. 5D)**; we noted that a subset of both classical and non-classical proteins could be partly released by EDTA treatment. Bound classical L-EV surface proteins that are prone to removal by EDTA (24, **set 1**) are predominantly non-membrane anchored and include actinins (ACTN1/4) and enzymes (GAPDH, PKM, ENO1, PGK), whereas bound non-classical proteins (non-CSPA/SURFY proteins) that are removed by EDTA (83, **set 2**) include annexins, enzymes (CALM3/5, CAT, LDHA/B, TGM1/3), junction assembly proteins (DSG1), cytoskeleton proteins (EZR, ACTC1/G1, MSN), chemokine (S100A9), 14-3-3 proteins; these proteins are implicated in “glycolysis”, “regulation of actin cytoskeleton” “junction organization” and “platelet activation” **(Fig. 5D-E, Supplementary Table 7)**. In contrast, bound classical cellular surface proteins on L-EVs resistant to removal by EDTA (43, **Fig. 5D set 3**) were membrane proteins (CDs, receptors such as integrins, LAPM1, SLC transporters,

GPI anchored proteins), which are implicated in "adhesion" and "ECM organization" **(Fig. 5E, Supplementary Table 7)**. While we do not understand how these surface associated proteins bind to L-EV surface, they have been previously reported to physically interact with several proteins of classical surface proteome **(Fig. 5F, Table 2)**. Furthermore, we also noted that 93/291 biotin-capture L-EV proteome displayed differential abundance in L-EVs from different cell lines (SW620, U87, MDA MB 231, **Supplementary Table 8**) which also includes several proteins (CTNNB1, EZR, SLCA5, SLC9A3R1, VIL1) that were readily removed by EDTA (**Supplementary Fig. 10**).

Bioinformatics assessment of Large EV surface proteome interactome

Interaction of surface proteins in cancer S-EVs to their cognate binding partners at distal sites have been shown to dictate EV-homing to specific organs[3] to develop pre-metastatic niches and enhance metastasis[3,6,67,70-72]. To gain insight into potential interactome of the surfaceome of cancer L-EVs, we next employed manually curated resource Cellinker that catalogues literature-supported ligand-receptor interactions[57] to retrieve potential surface engagements for 52 classical surfaceome (detected in >3 biotin capture experiments and annotated as SURFY/CSPA proteins). This resulted in 172 experimentally verified binding partners, encompassing 108 cell-adhesion interactions, 65 ECM-receptor interactions, 39 cytokine-cytokine receptor interactions and 36 secreted protein-receptor interactions **(Fig. 6A-C, Supplementary Table 9)**. Protein receptor interactions for L-EV surface-associated proteins are listed in **Supplementary Table 10**. DAVID-based analysis of the cognate binding partners revealed tissue enrichment for liver **(Fig. 6D)**. These interactions involved L-EVs ITGAV/ITGB1 with ANGPTL3, PLG and VTN, which display tissue enrichment in the liver **(Fig 6E-F)**, a primary metastatic site for colorectal and breast cancers[73-75]. Because exosomal integrin ITGAV/B5 was shown to be linked to liver metastasis[3], whether both EV-subtypes may share similar organ-specific homing and enhance metastasis warrants future investigation.

Discussion

In this study, we found that the protein diversity on surface of L-EVs (buoyant density of 1.07-1.11 g/mL) includes bona fide plasma membrane-anchored proteins (CDs, transporters, receptors) implicated in cell-cell and cell-ECM interactions, as well as membrane surface-

associated proteins (actin cytoskeleton components, enzymes, chemokines) implicated in actin filament organization, platelet activation and glycolysis. A pre-requisite for successful identification of surface proteins by membrane impermeant Sulfo-NHS-SS-biotin, a well-established agent which has been extensively used to identify cell surface proteins[76] as well as small EVs without compromising their integrity[34,77], is purified EV preparations that are morphologically intact as demonstrated in **Fig. 1**. Significant enrichment of bona fide membrane proteins in affinity-purification using avidin-beads, as compared to global L-EV proteome, indicates high surface labeling specificity (**Fig. 3-4**). Based on these results, we are confident that our experimental data reliably provides a comprehensive map of L-EV surfaceome.

Expectedly, L-EVs surface proteome comprised proteins of TM domain containing surface receptors, CDs, transporters and GPI anchored proteins. Several of these surface proteins are drivers of cancer progression from primary tumour growth, remodelling of the tumour microenvironment, invasive outgrowth (MET[6], EGFR[78]), immune regulation and metastatic spread (integrins[79]). The diversity of transporters encompassing ABC transporters, metal transporters (CNNM3/4), lactate transporters (SLC16A3), carboxylic acids transporters (SLC16A1/MTC) and SLC transporters (amino acid). S-EV-mediated horizontal transfer of membrane-embedded drug efflux pumps to sensitive cancer cells is well-known, leading to acquired drug resistance *in vitro* and *in vivo*[80-84]. Besides their role in chemoresistance, membrane transporters also dictate organ tropism of circulating EVs[85].

Resonating S-EVs capturing tissue-specific signatures of organs they originate from[86], we also noted that L-EVs from different cell lines clustered based on donor cells **(Supplementary Fig. 5)**. Moreover, we identified 170/190 SURFY proteins in L-EVs displaying differential abundance (FDR < 0.05) in L-EVs based on donor cells (SW620, LIM1863, U87). Several of these proteins, for example, ICAM1[87], SLC9A1[88] and TPBG[89] enriched in U87 L-EVs (vs SW620 and LIM186 L-EVs) have been implicated in glioblastoma disease progression. Interestingly, similar abundance was observed in U87-derived S-EVs and L-EVs for these proteins (ICAM1, SLC9A1, TPBG), highlighting potential of both EV-subtypes to recapitulate parental cell proteome. On the other hand, several proteins (APLP2, HM13, LMAN2 and RPN1) displayed higher abundance in U87 L-EVs compared to U87 S-EVs. HM13 was recently shown to enhance tumour progression by regulating secretion profile in glioblastoma[90]. APLP2 (Amyloid precursor-like protein 2, a type I transmembrane protein and

a member of the amyloid precursor protein family) have been shown to exert oncogenic function in diverse types of cancer[91]. Although APLP2 expression is similar between glioblastoma and normal brain, however, a high level of APLP2 is associated with poor prognosis in the glioblastoma patient group[92]. On the other hand, amongst proteins with higher abundance in L-EVs from CRC cell lines (SW620 or LIM1863) vs U-87 L-EVs, none displayed greater abundance when compared with CRC cell line S-EVs, however several of these proteins (MUC13, CDH17, ACE, ACE2, TSPAN8) are reported as intestine-enriched proteins in Human Protein Atlas[66]. Moreover, in our study, upon manual inspection, we noted that colon cancer cell-derived EVs (compared to EVs derived from U87 glioma or MDA MB 231 breast cancer cells) carried 17 intestinal specific proteins, including ACE/CD143, CDHR2 and MUC13. CDHR2 is a cell adhesion molecule expressed in intestinal epithelium in a tissue-specific way (https://www.proteinatlas.org/ENSG00000074276-CDHR2/tissue)[66] and has the potential to diagnose gastrointestinal adenocarcinomas[93,94]. Thus, L-EV surface proteins such as CDHR2 could potentially enable capture/detection of intestinal EV sub-population (amongst the bulk circulating EVs). Combined with our previous report of L-EVs carrying oncogenic KRAS (G12V mutant variant)[25], L-EVs surfaceome could potentially provide immunocapture-based avenue for cancer biomarker diagnostics.

Besides transmembrane anchored proteins, we also identified proteins that were associated with the surface of L-EVs. Several surface-associated proteins have been documented to dictate EV function; whereas EV surface ANXA1 promotes microcalcification[43], FN1 facilitates cellular migration through tissues[68], TGFB1 triggers fibroblast differentiation[69] and MIF establishes pre-metastatic niche in the liver[67]. We show that major components of glycolytic pathways were also present on L-EV surface (ALDH7A1, ALDOA, ENO1, GAPDH, LDHA, LDHB, PGAM1, PGK1, PKM, TPI1). Secretion of glycolytic enzymes (GAPDH, PKM, TPI1) in sera of colon cancer patients has been associated with 5-fluorouracil resistance[95]. The mechanism of glycolytic protein secretion is only beginning to emerge, which includes export via EVs (exosomes via tetraspanins[96] and shed microvesicles via caveolin-1[97]) or via SNARE-driven unconventional secretion[98]. These proteins then localize to surface of cells and/or potentially EVs[99]. Besides their well-known role in glycolysis, these proteins (GAPDH, TPI, PKM) also carry out multiple moonlighting functions in cancer such as attachment[100], proliferation[101] and migration via surface localization[99]. Currently, how glycolytic proteins associate with cell or EV outer surface remains unknown, but are potentially mediated by direct

interactions with surface proteins (**Fig. 6**); indeed, extracellular PKM is a ligand for cell surface EGFR[101].

We also found that non-classical surfaceome of L-EVs include proteins that assist in the assembly of junction complex (DLG1/3, CTNNA1/B1/D1, RHOA, NECTIN2, VCL) and components of actomyosin bundle (ACTN1, ACTN4, MYH6, MYH9) **(Table 1)**. While these are classically annotated in the cytoplasmic side of plasma membrane, they were also removed from EV surface following EDTA treatment **(Table 1)**; likely due to a subpopulation of EVs displaying inside-out topology[34]. Along this line, L-EV surface associated proteins also include actin and actin binding proteins that are typically intracellular proteins. While cytosolic proteins likely originate from damaged cells, attach to the surface of EVs and are hence co-purified, it is not uncommon for cytosolic proteins such as functional histones to be released extracellularly and bind to extracellular structure such as neutrophil extracellular traps or protein granules[102,103]. Alternatively, actins are also found extracellularly, co-localizes with and is a ligand for TREM-1 (a potent amplifier of pro-inflammatory innate immune response) on the surface of activated mouse macrophage[104]. Our current data also concurs our previous finding that actins are peripheral surface exposed proteins on S-EVs removed by proteinase K treatment[35], a finding that is also corroborated by other groups employing biotin labelling of surface proteins[34,36]. In contrast, classical $CD63^+$, $CD81^+$, and $CD9^+$ S-EVs were shown to lack cytoskeletal constituents that make up actin filaments, microtubules, or intermediate filaments[29], however we have previously shown that cytoskeleton constituents including actins are found in $EPCAM^+/A33^+$ S-EVs[28].

EVs can arise from different parts of cells (for e.g., exosomes are endosomally derived, microvesicles are plasma membrane derived, midbody remnants are derived from cytokinetic bridges), with several extracellular EV-subtypes being discovered whose route of biogenesis is yet to be unravelled. Based on our data, we propose that the majority of L-EVs arise from plasma membrane as we see striking enrichment of plasma membrane proteins in L-EV surfaceome. Their involvement in "cell projection", "leading edge" and "trailing edge" **(Fig. 4D, Table 1)** also suggests that a subset of L-EVs could also represent EVs released by migrating cells also known as migrasomes[105]. It should be noted that although midbody remnants are also L-EVs by definition (200-600 nm)[25], midbody remnants display higher buoyant density (1.22-1.30 g/mL) compared to L-EVs purified in this study (1.07-1.11 g/mL). Whether these L-EV subtypes have distinct surfaceome warrants future investigation.

Phenotypic diversity in EVs is gaining vested interest with S-EVs carrying multiple smaller EVs within its lumen.: such morphological diversity is observed in EVs released by single cell-type[26] as well isolated from body fluids (plasma[106], serum[107], cerebrospinal fluid[107,108] and semen[109]). Some studies suggest that such EV phenotypic diversity could be an artefact due to high speed centrifugation[110], however, cells have been shown to release EVs encapsulating smaller EVs[30,111], with such EV diversity also evident in unprocessed EVs from ejaculates[112]. , In our data, assessment of biotin inaccessible proteins (enriched for "COPI/II-coated vesicles", "ribosomes" that are associated with "ER/Golgi") strongly suggest that sub-populations of L-EVs could potentially contain smaller EVs within its lumen. Indeed, mitochondria can be actively sorted into plasma membrane blebs that are subsequently released in EVs[113,114]. Alternatively, L-EVs not only encapsulate mitochondria but also co-purify with free mitochondria themselves[115]. While enrichment of mitochondrial proteins in L-EVs is in accord with our previous findings[20], we did not find discernible cristae-containing mitochondria from our cryo EM imaging. In our data, we do note that VDAC1/2 were also labelled with biotin, thus L-EV subtype, in the absence of cristae phenotypes, could potentially include mitochondria-derived vesicles (VDAC enriched, ~200 nm, lack cristae)[116] that package mitochondrial matrix proteins (hence explaining their biotin inaccessibility). Mito-vesicles are known to selectively package mitochondrial proteins (exclude complex proteins and TOM20 which were also not detected in L-EV biotin surface proteome)[116] and are released extracellularly[117]. Although, we do not know the mechanisms underlying their biogenesis, intracellular mitovesicles are targeted either to lysosomes or MVB[118,119] which could facilitate their release to extracellular space.

In summary, our study provides a first comprehensive map of L-EV surface proteins. While density gradient-based purifications of EVs are regarded as a gold-standard by ISEV[26] and can separate EVs from soluble secreted proteins and heavy density EVs/protein aggregates, it is now apparent that other cellular organelles could potentially share biophysical properties which results in co-enrichment and preclude L-EV assessment for their form and function. Moreover, while size-based (i.e., large *versus* small) broad categorization of EVs is useful at an operational level, however future dissection calls for immunocapture of L-EV sub-populations if we are to resolve their origins, cargo and function which has been successfully used to classify S-EV subpopulations.

Acknowledgements

We thank Bio21 Molecular Science and Biotechnology Institute for assisting cryo-electron microscopy (University of Melbourne).

Author Contribution

A.R. and D.W.G conceptualised the idea. A.R. designed the experiments and wrote the manuscript. A.R., H.F., B.C. and D.W.G. performed experiments. A.R. performed bioinformatics analysis A.R., H.F., B.C., R.J.S. and D.W.G. authors reviewed manuscript for submission.

ORCID

Alin Rai 000-0001-7994-5151; Bethany Claridge 0000-0002-4589-1033; David W Greening - 0000-0001-7516-485X; Richard Simpson 0000-0002-9834-0796

Competing Interests

The authors have no conflicts to declare.

Data and Software Availability: The MS-based proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository[120] and are available via ProteomeXchange with identifier PXD024733. Experimental parameters submitted to EV-TRACK knowledgebase (EV-TRACK ID: EV210261)[50].

Funding

This work was supported by National Health and Medical Research Council Project grants 1057741 and 1139489 (to D.W.G), and Future Fund grant 1201805 (to D.W.G). We note further support by Helen Amelia Hains Fellowship (to D.W.G), and by the Victorian State Government Operational Infrastructure funding to the Baker Institute. BC is supported by an Australian Government Training Program (RTP) scholarship and Baker Institute Bright Sparks Scholarship Top Up.

Figure legends

Figure 1. Isolation and characterization of large and small EVs. A. Workflow for isolation of L-EVs and S-EVs. **B**. Cryo electron microscopic images of L-EVs and S-EVs isolated from SW620 cells. **C**. Histogram represents diameter of L-EVs and S-EVs based on cryo-EM images. Data presented as mean ± s.e.m (standard error of mean). **D**. Volcano plot of protein abundance between L-EVs and S-EVs using mass spectrometry-based proteomics; comparisons of log2 fold changes versus p-values (Student's t-test). **E**. Heat map of selected proteins enriched in L-EVs or S-EVs from three indicated cell lines ($p\text{-value} < 0.05$). **F**. EnrichmentMap of Gene Ontology (cellular components) processes overrepresented in L-EVs or S-EVs. Node size represents gene number.

Figure 2. Workflow for capturing surface proteins of large and small EVs. The workflow used in this study to define large and small EV surfaceome is outlined in **Fig. 3**. Experimentally verified cell-surface proteins (CSPA)[51] and cellular surfaceome predicted by SURFY[31] were mapped onto global EV protein profiles. Surface proteins were captured using membrane-impermeant Sulfo-NHS-SS-Biotin and subsequently identified by mass spectrometry. We also define a pool of EV-surface associated proteome by treating EVs with EDTA and assessing released proteins. CD; cluster of differentiation, GPI; Glycosylphosphatidylinositol.

Figure 3. Large EVs contain classical cellular surfaceome. Venn diagram of total proteins identified in **A.** L-EVs or **B**. S-EVs *versus* surface proteins reported in CSPA[51], SURFY[31] or UniProt. **C.** Voronoi tree maps for L-EV and S-EV proteomes generated on wlab.ethz.ch/surfaceome. Light color indicates low expression; dark color indicates strong expression. White genes are not expressed. **D**. Different classes of SURFY[31] predicted cellular surface proteins identified in each EV-subtype. **E**. Volcano plot of differentially abundant surfaceome proteins in L-EVs and S-EVs.

Figure 4. Proteome profiling of biotin accessible and inaccessible large EV proteomes. A. Venn diagram of biotin-captured surface proteins in L-EVs *versus* surface proteins reported in CSPA[51], SURFY[31] or UniProt. **B**. Bar plot of relative abundance of number of indicated classes of proteins identified *versus* total number of proteins identified in global or biotin-captured surface proteome. **C**. Venn diagram of biotin captured L-EV surface proteins and global L-EV proteome. Inner circle in L-EV proteome represents a subset of SURFY/CSPA proteins

detected in L-EV global proteome. **D.** Enrichment Map of Gene Ontology (cellular components) terms overrepresented in L-EV surface proteome *versus* biotin inaccessible L-EV proteome. **E.** KEGG and Reactome pathways overrepresented in L-EV surface proteome *versus* biotin inaccessible L-EV proteome.

Figure 5. Surface-associated proteins in large EVs. A. Venn diagram of proteins identified in L-EV surface proteome, and proteins that remain either bound or are released from L-EV surface following EDTA treatment. Bar plot of protein numbers of indicated classes that are either released **(B)** or remain bound **(C)** with EVs following EDTA treatment. **D.** L-EV biotin surface proteome was divided into CSPA/SURFY proteins (i.e. classical cellular surfaceome) or non-CSPA/SURFY proteins (i.e. non-classical surface proteins); Venn diagram reveals pool of these proteins that remain either bound or are released from L-EV surface following EDTA treatment. **E.** EnrichmentMap of KEGG and Reactome pathways overrepresented in L-EV proteins that remain either bound or are released from L-EV surface following EDTA treatment. **F.** GeneMania-based radial interaction map of membrane-bound and membrane-associated (i.e., released) proteins in L-EVs. Nodes represent the proteins, and the edges represent evidence-based direct physical interactions. Pathway involvement for indicated proteins are asterisk-color coded.

Figure 6. Receptor-ligand interactome of large EV surface proteome. A. Venn diagram of L-EV biotin surface proteins (classical) and their experimentally-verified cognate binding partners based on Cellinker that catalogues literature-supported ligand-receptor interactions[57]. **B.** Distribution of interaction-type identified in L-EVs. **C.** Sankey diagram of ligand-receptor interactions identified in L-EVs; complete list provided in Supplementary Table 7. **D.** Bar plot represents tissue enrichment (based on DAVID analysis) of 172 cognate interacting partners of L-EV surface proteins. **E.** Cellinker-based interacting ligands identified for L-EV surface ITGB1/ITGAV proteins. **F.** RNA expression and tissue specificity for ANGPTL3, PLG and VTN in different human tissues was obtained from The Human Protein Atlas (http://www.proteinatlas.org, image credit: Human Protein Atlas[66]).

Supplementary Figure Legends

Supplementary Figure 1. A. Immunoblot detection of CD63 and ANXA1 in large and small EVs purified using density-gradient separation from conditioned media of SW620 cells. B. Immunoblot detection of CD63 and ANXA1 in MDA MB 231 whole cell lysate and derived large and small EVs.

Supplementary Figure 2. Nanoparticle tracking analysis of purified large and small EVs from indicated cell lines.

Supplementary Figure 3. Nanoparticle tracking analysis data of purified large and small EVs from indicated cell lines represented as log10 transformed data (concentration (particle/mL) and diameter (size, nm).

Supplementary Figure 4. Bar plot depicting percentage of EVs of indicated size (>200 nm, >300 nm, >400 nm) within each EV-subtype from indicated cell lines based on A. cryo-EM analysis or B. single particle tracking (NTA) data.

Supplementary Figure 5. A. Venn diagram of proteins identified in L-EVs and S-EVs. B. Violin plot of LFQ intensities of proteins identified in L-EVs and S-EVs. C. Pearson correlation matrix of L-EV and S-EV proteomes for indicated cell lines.

Supplementary Figure 6. Bar plot depicting normalised LFQ intensities of indicated proteins in L-EVs or S-EVs from three cell lines MDA MB 231, SW620 and U87. p-value < 0.05 (*), 0.005 (**), 0.001 (***), 0.0001 (****), ns representing non-significant.

Supplementary Figure 7. Bar plot of Gene ontology-based biological processes overrepresented in surface proteins enriched in L-EVs or S-EVs. Middle panel represents biological processes enriched in surface proteins commonly identified in both L-EVs and S-EVs.

Supplementary Figure 8. A. Hierarchical clustering analysis of differentially abundant SURFY proteins (FDR <0.05) in global proteome of L-EVs from U87, SW620 and LIM1863

cell lines. Corresponding boxes list proteins that are found in associated clusters. *Proteins that are found in higher abundance in U87 L-EVs vs U87 S-EVs. B. Venn diagram of Human Protein Atlas-based "brain enriched" or "intestine enriched" proteins and Cluster 2 (U87 enriched) and CRC cell lines enriched (Cluster 1/3/4) proteins. Bar plots depicts LFQ intensities of indicated proteins in L-EVs and S-EVs from LIM1863.

Supplementary Figure 9. A. Venn diagram of biotin-captured surface proteins in S-EVs *versus* surface proteins reported in CSPA[51], SURFY[31] or UniProt. B. Relative abundance of transporters, CDs and receptors in L-EVs and S-EVs biotin-captured proteome. C. Volcano plot of differentially abundant biotin-captured receptors in L-EVs and S-EVs. D. Relative abundance of LFQ intensities of indicated proteins in S-EVs *versus* L-EVs. E. Western blot analysis of L-EVs and S-EVs released by SW620 cells using indicated antibodies. F. Fluorescence microscopic analysis of SW620 cells expressing plasma membrane targeting GAP-GFP proteins using indicated antibodies.

Supplementary Figure 10. Heat map of differentially abundant biotin captured proteins (FDR <0.05) in L-EVs from U87, SW620 and MDA MB 231 cell lines. Corresponding boxes list proteins that are found in associated clusters. *Proteins that are released from MDA MB 231 L-EVs following EDTA treatment.

Supplementary Table Legends

Supplementary Table 1. Global proteomic analysis of L-EVs and S-EVs

Supplementary Table 2. Gene ontology (cellular component) enriched in L-EVs and S-EVs

Supplementary Table 3. Differentially abundant SURFY proteins in L-EVs

Supplementary Table 4. Proteomic analysis of EV-surface proteins captured using membrane impermeant biotin

Supplementary Table 5. List of surface proteins identified in L-EVs

Supplementary Table 6. Pathways enriched in L-EVs and S-EVs

Supplementary Table 7. List of proteins from Figure 5D

Supplementary Table 8. Differentially abundant biotin captured proteins in L-EVs

Supplementary Table 9. Receptor-ligand interactome of L-EV surface proteins

Supplementary Table 10. Receptor-ligand interactome of L-EV surface-associated proteins

References

1. Xu, R. et al. Extracellular vesicles in cancer - implications for future improvements in cancer care. *Nat Rev Clin Oncol* **15**, 617-638 (2018).
2. Hu, Q. et al. Clinical applications of exosome membrane proteins. *Precis Clin Med* **3**, 54-66 (2020).
3. Hoshino, A. et al. Tumour exosome integrins determine organotropic metastasis. *Nature* **527**, 329-335 (2015).
4. Buzas, E.I., Toth, E.A., Sodar, B.W. & Szabo-Taylor, K.E. Molecular interactions at the surface of extracellular vesicles. *Semin Immunopathol* **40**, 453-464 (2018).
5. Charoenviriyakul, C., Takahashi, Y., Morishita, M., Nishikawa, M. & Takakura, Y. Role of Extracellular Vesicle Surface Proteins in the Pharmacokinetics of Extracellular Vesicles. *Mol Pharm* **15**, 1073-1080 (2018).
6. Peinado, H. et al. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med* **18**, 883-891 (2012).
7. Kim, D.H. et al. Exosomal PD-L1 promotes tumor growth through immune escape in non-small cell lung cancer. *Exp Mol Med* **51**, 1-13 (2019).
8. Al-Nedawi, K. et al. Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat Cell Biol* **10**, 619-624 (2008).
9. Melo, S.A. et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* **523**, 177-182 (2015).
10. Yoshioka, Y. et al. Ultra-sensitive liquid biopsy of circulating extracellular vesicles using ExoScreen. *Nat Commun* **5**, 3591 (2014).
11. Zaborowski, M.P. et al. Methods for Systematic Identification of Membrane Proteins for Specific Capture of Cancer-Derived Extracellular Vesicles. *Cell Rep* **27**, 255-268 e256 (2019).
12. Furuta, T. et al. Mesenchymal Stem Cell-Derived Exosomes Promote Fracture Healing in a Mouse Model. *Stem Cells Transl Med* **5**, 1620-1630 (2016).
13. Vicencio, J.M. et al. Plasma exosomes protect the myocardium from ischemia-reperfusion injury. *J Am Coll Cardiol* **65**, 1525-1536 (2015).
14. Yang, Y., Qin, M., Bao, P., Xu, W. & Xu, J. Secretory carrier membrane protein 5 is an autophagy inhibitor that promotes the secretion of alpha-synuclein via exosome. *PLoS One* **12**, e0180892 (2017).
15. Tian, T. et al. Surface functionalized exosomes as targeted drug delivery vehicles for cerebral ischemia therapy. *Biomaterials* **150**, 137-149 (2018).
16. Gao, X. et al. Anchor peptide captures, targets, and loads exosomes of diverse origins for diagnostics and therapy. *Sci Transl Med* **10** (2018).
17. Ohno, S. et al. Systemically injected exosomes targeted to EGFR deliver antitumor microRNA to breast cancer cells. *Mol Ther* **21**, 185-191 (2013).
18. Mentkowski, K.I. & Lang, J.K. Exosomes Engineered to Express a Cardiomyocyte Binding Peptide Demonstrate Improved Cardiac Retention in Vivo. *Sci Rep* **9**, 10041 (2019).
19. Yaddanapudi, K. et al. Exosomes from GM-CSF expressing embryonic stem cells are an effective prophylactic vaccine for cancer prevention. *Oncoimmunology* **8**, 1561119 (2019).
20. Xu, R., Greening, D.W., Rai, A., Ji, H. & Simpson, R.J. Highly-purified exosomes and shed microvesicles isolated from the human colon cancer cell line LIM1863 by sequential centrifugal ultrafiltration are biochemically and functionally distinct. *Methods* **87**, 11-25 (2015).
21. Kowal, J. et al. Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc Natl Acad Sci U S A* **113**, E968-977 (2016).
22. Bellin, G. et al. Exosome in Cardiovascular Diseases: A Complex World Full of Hope. *Cells* **8** (2019).

23. Howitt, J. & Hill, A.F. Exosomes in the Pathology of Neurodegenerative Diseases. *J Biol Chem* **291**, 26589-26597 (2016).
24. Tkach, M., Kowal, J. & Thery, C. Why the need and how to approach the functional diversity of extracellular vesicles. *Philos Trans R Soc Lond B Biol Sci* **373** (2018).
25. Rai, A. et al. Secreted midbody remnants are a class of extracellular vesicles molecularly distinct from exosomes and microparticles. *Commun Biol* **4**, 400 (2021).
26. Thery, C. et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles* **7**, 1535750 (2018).
27. Ji, H. et al. Deep sequencing of RNA from three different extracellular vesicle (EV) subtypes released from the human LIM1863 colon cancer cell line uncovers distinct miRNA-enrichment signatures. *PLoS One* **9**, e110314 (2014).
28. Tauro, B.J. et al. Two distinct populations of exosomes are released from LIM1863 colon carcinoma cell-derived organoids. *Mol Cell Proteomics* **12**, 587-598 (2013).
29. Jeppesen, D.K. et al. Reassessment of Exosome Composition. *Cell* **177**, 428-445 e418 (2019).
30. Wiklander, O.P. et al. Extracellular vesicle in vivo biodistribution is determined by cell source, route of administration and targeting. *J Extracell Vesicles* **4**, 26316 (2015).
31. Bausch-Fluck, D. et al. The in silico human surfaceome. *Proc Natl Acad Sci U S A* **115**, E10988-E10997 (2018).
32. Wu, D. et al. Profiling surface proteins on individual exosomes using a proximity barcoding assay. *Nat Commun* **10**, 3854 (2019).
33. Jung, A.L. et al. Surface Proteome of Plasma Extracellular Vesicles as Biomarkers for Pneumonia and Acute Exacerbation of Chronic Obstructive Pulmonary Disease. *J Infect Dis* **221**, 325-335 (2020).
34. Cvjetkovic, A. et al. Detailed Analysis of Protein Topology of Extracellular Vesicles-Evidence of Unconventional Membrane Protein Orientation. *Sci Rep* **6**, 36338 (2016).
35. Xu, R. et al. Surfaceome of Exosomes Secreted from the Colorectal Cancer Cell Line SW480: Peripheral and Integral Membrane Proteins Analyzed by Proteolysis and TX114. *Proteomics* **19**, e1700453 (2019).
36. Castillo, J. et al. Surfaceome profiling enables isolation of cancer-specific exosomal cargo in liquid biopsies from pancreatic cancer patients. *Ann Oncol* **29**, 223-229 (2018).
37. Jeppesen, D.K. et al. Quantitative proteomics of fractionated membrane and lumen exosome proteins from isogenic metastatic and nonmetastatic bladder cancer cells reveal differential expression of EMT factors. *Proteomics* **14**, 699-712 (2014).
38. Hu, S. et al. Purification and Identification of Membrane Proteins from Urinary Extracellular Vesicles using Triton X-114 Phase Partitioning. *J Proteome Res* **17**, 86-96 (2018).
39. Whitehead, R.H., Jones, J.K., Gabriel, A. & Lukies, R.E. A new colon carcinoma cell line (LIM1863) that grows as organoids with spontaneous differentiation into crypt-like structures in vitro. *Cancer Res* **47**, 2683-2689 (1987).
40. Ford, T., Graham, J. & Rickwood, D. Iodixanol: a nonionic iso-osmotic centrifugation medium for the formation of self-generated gradients. *Anal Biochem* **220**, 360-366 (1994).
41. Poh, Q.H., Rai, A., Carmichael, II, Salamonsen, L.A. & Greening, D.W. Proteome reprogramming of endometrial epithelial cells by human trophectodermal small extracellular vesicles reveals key insights into embryo implantation. *Proteomics* **21**, e2000210 (2021).
42. Hughes, C.S. et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68-85 (2019).
43. Rogers, M.A. et al. Annexin A1-dependent tethering promotes extracellular vesicle aggregation revealed with single-extracellular vesicle analysis. *Sci Adv* **6** (2020).
44. Rai, A. et al. Proteomic profiling of human uterine extracellular vesicles reveal dynamic regulation of key players of embryo implantation and fertility during menstrual cycle. *Proteomics* **21**, e2000211 (2021).
45. Claridge, B. et al. Proteome characterisation of extracellular vesicles isolated from heart. *Proteomics* **21**, e2100026 (2021).

46. Kompa, A.R. et al. Sustained subcutaneous delivery of secretome of human cardiac stem cells promotes cardiac repair following myocardial infarction. *Cardiovasc Res* **117**, 918-929 (2021).
47. Greening, D.W. et al. Chronic methamphetamine interacts with BDNF Val66Met to remodel psychosis pathways in the mesocorticolimbic proteome. *Mol Psychiatry*, In Press (2019).
48. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).
49. Nagaraj, N. et al. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Molecular & Cellular Proteomics* **11**, M111. 013722 (2012).
50. Consortium, E.-T. et al. EV-TRACK: transparent reporting and centralizing knowledge in extracellular vesicle research. *Nat Methods* **14**, 228-232 (2017).
51. Bausch-Fluck, D. et al. A mass spectrometric-derived cell surface protein atlas. *PLoS One* **10**, e0121314 (2015).
52. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191-W198 (2019).
53. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).
54. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
55. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).
56. Warde-Farley, D. et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* **38**, W214-220 (2010).
57. Zhang, Y. et al. Cellinker: a platform of ligand-receptor interactions for intercellular communication analysis. *Bioinformatics* (2021).
58. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016).
59. Zuber, M.X., Strittmatter, S.M. & Fishman, M.C. A membrane-targeting signal in the amino terminus of the neuronal protein GAP-43. *Nature* **341**, 345-348 (1989).
60. Rai, A. et al. Exosomes Derived from Human Primary and Metastatic Colorectal Cancer Cells Contribute to Functional Heterogeneity of Activated Fibroblasts by Reprogramming Their Proteome. *Proteomics* **19**, e1800148 (2019).
61. Suwakulsiri, W. et al. Proteomic profiling reveals key cancer progression modulators in shed microvesicles released from isogenic human primary and metastatic colorectal cancer cell lines. *Biochim Biophys Acta Proteins Proteom* **1867**, 140171 (2019).
62. van der Pol, E. et al. Particle size distribution of exosomes and microvesicles determined by transmission electron microscopy, flow cytometry, nanoparticle tracking analysis, and resistive pulse sensing. *J Thromb Haemost* **12**, 1182-1192 (2014).
63. Kugeratski, F.G. et al. 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 (2021).
64. Minciacchi, V.R. et al. Large oncosomes contain distinct protein cargo and represent a separate functional class of tumor-derived extracellular vesicles. *Oncotarget* **6**, 11327-11341 (2015).
65. Santucci, L. et al. Biological surface properties in extracellular vesicles and their effect on cargo proteins. *Sci Rep* **9**, 13048 (2019).
66. Uhlen, M. et al. Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
67. Costa-Silva, B. et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat Cell Biol* **17**, 816-826 (2015).
68. Sung, B.H., Ketova, T., Hoshino, D., Zijlstra, A. & Weaver, A.M. Directional cell movement through tissues is controlled by exosome secretion. *Nat Commun* **6**, 7164 (2015).

69. Ringuette Goulet, C. et al. Exosomes Induce Fibroblast Differentiation into Cancer-Associated Fibroblasts through TGFbeta Signaling. *Mol Cancer Res* **16**, 1196-1204 (2018).
70. Jung, T. et al. CD44v6 dependence of premetastatic niche preparation by exosomes. *Neoplasia* **11**, 1093-1105 (2009).
71. Hood, J.L., San, R.S. & Wickline, S.A. Exosomes released by melanoma cells prepare sentinel lymph nodes for tumor metastasis. *Cancer Res* **71**, 3792-3801 (2011).
72. Grange, C. et al. Microvesicles released from human renal cancer stem cells stimulate angiogenesis and formation of lung premetastatic niche. *Cancer Res* **71**, 5346-5356 (2011).
73. Liu, J. et al. Identification of liver metastasis-associated genes in human colon carcinoma by mRNA profiling. *Chin J Cancer Res* **30**, 633-646 (2018).
74. Ku, X. et al. In-Depth Characterization of Mass Spectrometry-Based Proteomic Profiles Revealed Novel Signature Proteins Associated with Liver Metastatic Colorectal Cancers. *Anal Cell Pathol (Amst)* **2019**, 7653230 (2019).
75. Chen, M.T. et al. Comparison of patterns and prognosis among distant metastatic breast cancer patients by age groups: a SEER population-based analysis. *Sci Rep* **7**, 9254 (2017).
76. Li, Y. et al. Sensitive profiling of cell surface proteome by using an optimized biotinylation method. *J Proteomics* **196**, 33-41 (2019).
77. Diaz, G., Wolfe, L.M., Kruh-Garcia, N.A. & Dobos, K.M. Changes in the Membrane-Associated Proteins of Exosomes Released from Human Macrophages after Mycobacterium tuberculosis Infection. *Sci Rep* **6**, 37975 (2016).
78. Liu, B. et al. Metformin induces unique biological and molecular responses in triple negative breast cancer cells. *Cell Cycle* **8**, 2031-2040 (2009).
79. Sokeland, G. & Schumacher, U. The functional role of integrins during intra- and extravasation within the metastatic cascade. *Mol Cancer* **18**, 12 (2019).
80. Bebawy, M. et al. Membrane microparticles mediate transfer of P-glycoprotein to drug sensitive cancer cells. *Leukemia* **23**, 1643-1649 (2009).
81. Levchenko, A. et al. Intercellular transfer of P-glycoprotein mediates acquired multidrug resistance in tumor cells. *Proc Natl Acad Sci U S A* **102**, 1933-1938 (2005).
82. Sousa, D., Lima, R.T. & Vasconcelos, M.H. Intercellular Transfer of Cancer Drug Resistance Traits by Extracellular Vesicles. *Trends Mol Med* **21**, 595-608 (2015).
83. Lv, M.M. et al. Exosomes mediate drug resistance transfer in MCF-7 breast cancer cells and a probable mechanism is delivery of P-glycoprotein. *Tumour Biol* **35**, 10773-10779 (2014).
84. Zhao, Y., Chen, Y., Wang, J. & Liu, L. Effects of ATP-binding cassette transporter G2 in extracellular vesicles on drug resistance of laryngeal cancer cells in in vivo and in vitro. *Oncol Lett* **21**, 364 (2021).
85. Wu, A.Y. et al. Multiresolution Imaging Using Bioluminescence Resonance Energy Transfer Identifies Distinct Biodistribution Profiles of Extracellular Vesicles and Exomeres with Redirected Tropism. *Adv Sci (Weinh)* **7**, 2001467 (2020).
86. Hoshino, A. et al. Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. *Cell* **182**, 1044-1061 e1018 (2020).
87. Piao, Y. et al. Targeting intercellular adhesion molecule-1 prolongs survival in mice bearing bevacizumab-resistant glioblastoma. *Oncotarget* **8**, 96970-96983 (2017).
88. Guan, X. et al. Elevated Na/H exchanger 1 (SLC9A1) emerges as a marker for tumorigenesis and prognosis in gliomas. *J Exp Clin Cancer Res* **37**, 255 (2018).
89. Zhang, C. et al. A Novel Glucose Metabolism-Related Gene Signature for Overall Survival Prediction in Patients with Glioblastoma. *Biomed Res Int* **2021**, 8872977 (2021).
90. Wei, J.W. et al. Signal Peptide Peptidase, Encoded by HM13, Contributes to Tumor Progression by Affecting EGFRvIII Secretion Profiles in Glioblastoma. *CNS Neurosci Ther* **23**, 257-265 (2017).
91. Pandey, P. et al. Amyloid precursor protein and amyloid precursor-like protein 2 in cancer. *Oncotarget* **7**, 19430-19444 (2016).
92. Chen, Y. et al. Expression of amyloid precursor-like protein 2 (APLP2) in glioblastoma is associated with patient prognosis. *Folia Neuropathol* **56**, 30-38 (2018).
93. Su, M.C., Yuan, R.H., Lin, C.Y. & Jeng, Y.M. Cadherin-17 is a useful diagnostic marker for adenocarcinomas of the digestive system. *Mod Pathol* **21**, 1379-1386 (2008).

94. Panarelli, N.C., Yantiss, R.K., Yeh, M.M., Liu, Y. & Chen, Y.T. Tissue-specific cadherin CDH17 is a useful marker of gastrointestinal adenocarcinomas with higher sensitivity than CDX2. *Am J Clin Pathol* **138**, 211-222 (2012).
95. Shin, Y.K. et al. Upregulation of glycolytic enzymes in proteins secreted from human colon cancer cells with 5-fluorouracil resistance. *Electrophoresis* **30**, 2182-2192 (2009).
96. Perez-Hernandez, D. et al. The intracellular interactome of tetraspanin-enriched microdomains reveals their function as sorting machineries toward exosomes. *J Biol Chem* **288**, 11649-11661 (2013).
97. Raikar, L.S., Vallejo, J., Lloyd, P.G. & Hardin, C.D. Overexpression of caveolin-1 results in increased plasma membrane targeting of glycolytic enzymes: the structural basis for a membrane associated metabolic compartment. *J Cell Biochem* **98**, 861-871 (2006).
98. Miura, N. et al. Tracing putative trafficking of the glycolytic enzyme enolase via SNARE-driven unconventional secretion. *Eukaryot Cell* **11**, 1075-1082 (2012).
99. Gomez-Arreaza, A. et al. Extracellular functions of glycolytic enzymes of parasites: unpredicted use of ancient proteins. *Mol Biochem Parasitol* **193**, 75-81 (2014).
100. Grundel, A., Jacobs, E. & Dumke, R. Interactions of surface-displayed glycolytic enzymes of Mycoplasma pneumoniae with components of the human extracellular matrix. *Int J Med Microbiol* **306**, 675-685 (2016).
101. Hsu, M.C. et al. Extracellular PKM2 induces cancer proliferation by activating the EGFR signaling pathway. *Am J Cancer Res* **6**, 628-638 (2016).
102. Yousefi, S. et al. In vivo evidence for extracellular DNA trap formation. *Cell Death Dis* **11**, 300 (2020).
103. Nakazawa, D. et al. Histones and Neutrophil Extracellular Traps Enhance Tubular Necrosis and Remote Organ Injury in Ischemic AKI. *J Am Soc Nephrol* **28**, 1753-1768 (2017).
104. Fu, L. et al. Identification of Extracellular Actin As a Ligand for Triggering Receptor Expressed on Myeloid Cells-1 Signaling. *Front Immunol* **8**, 917 (2017).
105. Jiang, D. et al. Migrasomes provide regional cues for organ morphogenesis during zebrafish gastrulation. *Nat Cell Biol* **21**, 966-977 (2019).
106. Gallart-Palau, X. et al. Extracellular vesicles are rapidly purified from human plasma by PRotein Organic Solvent PRecipitation (PROSPR). *Sci Rep* **5**, 14664 (2015).
107. Chen, W.W. et al. BEAMing and Droplet Digital PCR Analysis of Mutant IDH1 mRNA in Glioma Patient Serum and Cerebrospinal Fluid Extracellular Vesicles. *Mol Ther Nucleic Acids* **2**, e109 (2013).
108. Emelyanov, A. et al. Cryo-electron microscopy of extracellular vesicles from cerebrospinal fluid. *PLoS One* **15**, e0227949 (2020).
109. Poliakov, A., Spilman, M., Dokland, T., Amling, C.L. & Mobley, J.A. Structural heterogeneity and protein composition of exosome-like vesicles (prostasomes) in human semen. *Prostate* **69**, 159-167 (2009).
110. Issman, L., Brenner, B., Talmon, Y. & Aharon, A. Cryogenic transmission electron microscopy nanostructural study of shed microparticles. *PLoS One* **8**, e83680 (2013).
111. Zabeo, D. et al. Exosomes purified from a single cell type have diverse morphology. *J Extracell Vesicles* **6**, 1329476 (2017).
112. Hoog, J.L. & Lotvall, J. Diversity of extracellular vesicles in human ejaculates revealed by cryo-electron microscopy. *J Extracell Vesicles* **4**, 28680 (2015).
113. Islam, M.N. et al. Mitochondrial transfer from bone-marrow-derived stromal cells to pulmonary alveoli protects against acute lung injury. *Nat Med* **18**, 759-765 (2012).
114. Nicolas-Avila, J.A. et al. A Network of Macrophages Supports Mitochondrial Homeostasis in the Heart. *Cell* **183**, 94-109 e123 (2020).
115. Puhm, F. et al. Mitochondria Are a Subset of Extracellular Vesicles Released by Activated Monocytes and Induce Type I IFN and TNF Responses in Endothelial Cells. *Circ Res* **125**, 43-52 (2019).
116. Soubannier, V., Rippstein, P., Kaufman, B.A., Shoubridge, E.A. & McBride, H.M. Reconstitution of mitochondria derived vesicle formation demonstrates selective enrichment of oxidized cargo. *PLoS One* **7**, e52830 (2012).

117. D'Acunzo, P. et al. Mitovesicles are a novel population of extracellular vesicles of mitochondrial origin altered in Down syndrome. *Sci Adv* **7** (2021).
118. Sugiura, A., McLelland, G.L., Fon, E.A. & McBride, H.M. A new pathway for mitochondrial quality control: mitochondrial-derived vesicles. *EMBO J* **33**, 2142-2156 (2014).
119. Matheoud, D. et al. Parkinson's Disease-Related Proteins PINK1 and Parkin Repress Mitochondrial Antigen Presentation. *Cell* **166**, 314-327 (2016).
120. Jarnuczak, A.F. & Vizcaino, J.A. Using the PRIDE Database and ProteomeXchange for Submitting and Accessing Public Proteomics Datasets. *Curr Protoc Bioinformatics* **59**, 13 31 11-13 31 12 (2017).