

# Comparative proteomic analysis of three major extracellular classes secreted from human adenocarcinoma and metastatic colorectal cancer cells: exosomes, microparticles and shed midbody remnants

Wittaya Suwakulsiri<sup>1</sup>, Rong Xu<sup>2</sup>, Alin Rai<sup>3</sup>, Maoshan Chen<sup>4</sup>, Adnan Shafiq<sup>5</sup>, David Greening<sup>3</sup>, and Richard Simpson<sup>5</sup>

<sup>1</sup>Monash University

<sup>2</sup>Monash University The Australian Centre for Blood Diseases

<sup>3</sup>Baker Heart and Diabetes Institute

<sup>4</sup>Army Medical University

<sup>5</sup>La Trobe University

May 3, 2023

## Abstract

Cell-derived extracellular vesicles (EVs) are evolutionary-conserved secretory organelles that, based on their molecular composition, are important intercellular signaling regulators. At least three classes of circulating EVs are known based on mechanism of biogenesis: exosomes (sEVs/Exos), microparticles (lEVs/MPs) and shed midbody remnants (sMB-Rs). sEVs/Exos are of endosomal pathway origin, microparticles (lEVs/MPs) from plasma membrane blebbing, and shed midbody remnants (sMB-Rs) arise from symmetric cytokinetic abscission. Here, we isolate sEVs/Exos, lEVs/MPs and sMB-Rs secreted from human isogenic primary (SW480) and metastatic (SW620) colorectal cancer (CRC) cell lines in milligram quantities for label-free MS/MS-based proteomic profiling. Purified EVs revealed selective composition packaging of exosomal protein markers in SW480/SW620-sEVs/Exos, metabolic enzymes in SW480/SW620-lEVs/MPs, while centralspindlin complex proteins, nucleoproteins, splicing factors, RNA granule proteins, translation-initiation factors, and mitochondrial proteins selectively traffic to SW480/SW620-sMB-Rs. Collectively, we identify 39 human cancer-associated genes in EVs; 17 associated with SW480-EVs, 22 with SW620-EVs. We highlight oncogenic receptors/transporters selectively enriched in sEVs/Exos (EGFR/ FAS in SW480-Exos and MET, TGFBR2, ABCB1 in SW620-sEVs/Exos). Interestingly, MDK, STAT1, and TGM2 are selectively enriched in SW480-sMB-Rs, and ADAM15 to SW620-sMB-Rs. Our study reveals sEVs/Exos, lEVs/MPs and sMB-Rs have distinct protein signatures that open potential diagnostic avenues of distinct types of EVs for clinical utility.

## Comparative proteomic analysis of three major extracellular classes secreted from human primary and metastatic colorectal cancer cells: exosomes, microparticles and shed midbody remnants

Wittaya Suwakulsiri<sup>1,2</sup>, Rong Xu<sup>3</sup>, Alin Rai<sup>4,5,6</sup>, Adnan Shafiq<sup>1</sup>, Maoshan Chen<sup>7</sup>, David W. Greening<sup>4,5,6</sup> and Richard Simpson<sup>1\*</sup>

<sup>1</sup>*Department of Biochemistry and Chemistry, La Trobe Institute for Molecular Science (LIMS), La Trobe University, Melbourne, Victoria 3086, Australia*

<sup>2</sup>*Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash Medical Centre, Monash University, Clayton, Victoria 3168, Australia*

<sup>3</sup>*Nanobiotechnology Laboratory, Australia Centre for Blood Diseases, Centre Clinical*

School, Monash University, Melbourne, Victoria, 3004, Australia

<sup>4</sup>Baker Heart and Diabetes Institute, Melbourne, Victoria 3004, Australia

<sup>5</sup>Baker Department of Cardiovascular Research, Translation and Implementation,  
La Trobe University, Melbourne, Victoria 3086, Australia

<sup>6</sup>Baker Department of Cardiometabolic Health, University of Melbourne, Melbourne, Victoria 3010, Australia

<sup>7</sup>Laboratory of Radiation Biology, Department of Blood Transfusion, Laboratory Medicine Centre, The Second  
Affiliated Hospital, Army Medical University, Chongqing 400037, China

\*To whom correspondence should be addressed:

Professor Richard J. Simpson

Department of Biochemistry and Chemistry |, La Trobe Institute for Molecular Science (LIMS),

LIMS 1, Room 412

La Trobe University, Melbourne, Victoria 3086, Australia

Tel: +61 03 9479 3099

Fax: +61 03 9479 1226

Email: [Richard.Simpson@latrobe.edu.au](mailto:Richard.Simpson@latrobe.edu.au)

## ABSTRACT

Cell-derived extracellular vesicles (EVs) are evolutionary-conserved secretory organelles that, based on their molecular composition, are important intercellular signaling regulators. At least three classes of circulating EVs are known based on mechanism of biogenesis: exosomes (sEVs/Exos), microparticles (IEVs/MPs) and shed midbody remnants (sMB-Rs). sEVs/Exos are of endosomal pathway origin, microparticles (IEVs/MPs) from plasma membrane blebbing, and shed midbody remnants (sMB-Rs) arise from symmetric cytokinetic abscission. Here, we isolate sEVs/Exos, IEVs/MPs and sMB-Rs secreted from human isogenic primary (SW480) and metastatic (SW620) colorectal cancer (CRC) cell lines in milligram quantities for label-free MS/MS-based proteomic profiling. Purified EVs revealed selective composition packaging of exosomal protein markers in SW480/SW620-sEVs/Exos, metabolic enzymes in SW480/SW620-IEVs/MPs, while centralspindlin complex proteins, nucleoproteins, splicing factors, RNA granule proteins, translation-initiation factors, and mitochondrial proteins selectively traffic to SW480/SW620-sMB-Rs. Collectively, we identify 39 human cancer-associated genes in EVs; 17 associated with SW480-EVs, 22 with SW620-EVs. We highlight oncogenic receptors/transporters selectively enriched in sEVs/Exos (EGFR/ FAS in SW480-Exos and MET, TGFBR2, ABCB1 in SW620-sEVs/Exos). Interestingly, MDK, STAT1, and TGM2 are selectively enriched in SW480-sMB-Rs, and ADAM15 to SW620-sMB-Rs. Our study reveals sEVs/Exos, IEVs/MPs and sMB-Rs have distinct protein signatures that open potential diagnostic avenues of distinct types of EVs for clinical utility. **Keywords:**

*Extracellular vesicles, proteomics, exosomes, microparticles, shed microvesicles, shed midbody remnants, colorectal cancer*

## 1 INTRODUCTION

Secretion and exchange of extracellular vesicles (EVs) by most cell types is a central paradigm for intercellular communication [1-3]. EVs play a critical role in normal physiology and pathophysiology such as cancer, neurodegenerative disorders, and infectious diseases. These secretory organelles are evolutionary conserved with the capacity to modulate recipient cell phenotype/function by horizontal transfer of intrinsic cargo constituents such as oncogenic proteins, cytokines, infectious proteins (amyloid- $\beta$  proteins, prions, malarial proteins), RNA species (miRNAs, mRNAs, lncRNA), lipids, and metabolites[4]. The utility of EV-based

derived biomarkers have gained significant attention in recent years and there are several notable examples of the translatability and diagnostic potential of these markers, demonstrating promise for clinical utility [1, 5-9].

As heterogeneously sized (~30–2000+ nm) particles, EVs include exosomes (Exos are a subset of small EVs, sEVs), microparticles (MPs) (a subset of large EVs (IEVs), also termed microvesicles and ectosomes), and shed midbody remnants (sMB-Rs - a subset of IEVs), migrasomes and apoptotic bodies [10-17], based predominantly on their biophysical properties (size and density), differing biochemical composition, and surface charge [18]. Of these distinct subtypes, sEVs/Exos, and microparticles (MPs, aka shed microvesicles, sMVs) have been well described [1, 19, 20]. Exosomes derive from intracellular endosomal compartments and are formed through inward invagination of late endosomes to form intraluminal vesicles (ILVs) within multivesicular bodies (MVBs), which subsequently fuse with the plasma membrane to release ILVs, now termed exosomes, into the extracellular space. Exosome biogenesis is mediated by ESCRT machinery (Endosomal Sorting Complexes Required for Transport) such as ALIX and TSG101. The precise nature of the ESCRT complexes involved in exosome biogenesis has been reviewed elsewhere [1, 21]. Morphologically, exosomes are spherical and relatively homogenous with a size distribution over the range 50 to 200 nm diameter and density 1.08-1.14 g/ml. ALIX and TSG101, along with CD81, and CD63, are widely used as stereotypic protein markers for sEVs/exosomes [22, 23].

Microparticles are formed by outward blebbing of the cell plasma membrane [1]. In contrast to exosomes, MPs are more ellipsoid in shape and heterogeneous with respect to size (50 to ~2000 nm diameter) [1], density 1.08-1.14 g/ml; to date, no stereotypic protein markers have been described for MPs.

While much is known about exosomes and MPs, little is known about recently reported shed midbody remnants (sMB-Rs) [14, 24]. During the final stages of cell division newly formed daughter cells remain connected by a thin intercellular bridge containing the midbody (MB), a microtubule-rich organelle responsible for cytokinetic abscission. For decades the prevailing view was that cytokinetic abscission occurred asymmetrically and that the MB remnant was inherited by one daughter cell, where it persists as a midbody remnant (MB-R) and is subsequently engulfed by autophagy whereupon it is degraded intracellularly [24, 25]. Accumulating evidence now shows that MB-Rs can also be released into the extracellular space (shed MB-Rs, sMB-Rs; diameter 200-600 nm, density: 1.22-1.30 g/ml) following symmetric cytokinetic abscission and potentially be taken up by non-sister cells [26]. An unanswered question in the EV field relates to our knowledge of what subset of cytoplasmic proteins selectively traffic to the different EV classes – are they distinct? – and, if so, how they might impact on EV functionality. As a first step towards addressing this question we have undertaken, a comprehensive comparative proteomic analysis of sEVs/Exos, IEVs/MPs and sMB-Rs secreted from human primary and metastatic colorectal cancer (CRC) cells using mass spectrometry. Using a combination of differential ultracentrifugation and isopycnic iodixanol density centrifugation [14, 27-29] sEVs/Exos, IEVs/MPs and sMB-Rs were purified from the culture media of isogenic CRC cells SW480 (surrogate of CRC adenocarcinoma) and SW620 cells (surrogate of lymph node-metastatic CRC cancer). We report here the identity of proteins selectively enriched in each of the three EV classes secreted from SW480 and SW620 cells, including oncoproteins.

## 2 EXPERIMENTAL PROCEDURES

### 2.1 Cell culture and large-scale purification of Exos, MPs and sMB-Rs

SW480 and SW620 cell lines were cultured in a CELLLine AD-1000 bioreactor device as described [30]. SW480 and SW620 culture media were sequentially centrifuged at 500 x *g* for 5 min (to remove floating cells), 2,000 x *g* for 10 min (to remove apoptotic body debris) and 10,000 x *g* for 30 min at 4 °C. The 10K pellet was subjected to buoyant density (isopycnic iodixanol (OptiPrep) gradient centrifugation to separate IEVs/MPs (low buoyant density, fraction #7, 1.10 g/mL) from sMB-Rs (high buoyant density, fraction #9, 1.14-1.15 g/mL) by centrifugation at 100,000 x *g* for 18 h at 4degC. IEVs/MPs and sMB-Rs were recovered from fractions #7 and #9, respectively, by centrifugation at 100,000 x *g* for 18 h at 4degC, and the pellets resuspended in PBS (500  $\mu$ L) and then re-centrifuged at 10,000 x *g*(30 min, 4°C). The pellets

were resuspended in PBS (150  $\mu$ L) for proteome analysis. The 10K supernatant was centrifuged at 100,000  $\times$   $g$  (1 h, 4°C) to harvest crude sEVs/Exos. The crude sEVs/Exos pellet was re-suspended in 500  $\mu$ L PBS and subjected to OptiPrep buoyant density gradient centrifugation as described above and purified sEVs/Exos harvested (fraction #7 at a buoyant density of 1.10 g/mL) for biophysical/biochemical characterization.

## 2.2 Protein quantification and western blotting

Protein quantification of EV samples, and western blot analyses were determined as described [30]. For Western blot analysis membranes were probed with primary antibodies (anti-mouse ALIX, 1:1,000, Cell Signaling, Cat. No. 2171), (anti-mouse TSG101, 1:1,000, BD Bioscience, Cat. No. 612697), (anti-rabbit GAPDH, 1:3,000, Cell Signaling, Cat. No. 2118), (anti-mouse KIF23, 1:1,000, Santa Cruz Biotechnology, Cat. No. sc-390113), (anti-mouse RACGAP1, 1:1,000, Santa Cruz Biotechnology, Cat. No. sc-271110) according to manufacturer's instructions. The secondary antibodies (IRDye800 goat anti-mouse IgG (Cat No. AP308P) or IRDye700 goat anti-rabbit IgG (Cat. No. SAB4600400)) were diluted (1:15,000) and the fluorescent signals were detected using the Odyssey Infrared Imaging System, v3.0 (Li-COR Biosciences, Nebraska USA).

## 2.3 Transmission electron microscopy (TEM)

TEM images of sEVs/Exos, IEVs/MPs and sMB-Rs were obtained as described [30]. Briefly, sEVs/Exos, IEVs/MPs and sMB-Rs from SW480 and SW620 cell lines (1  $\mu$ g in 10  $\mu$ L PBS) were applied to 400 mesh carbon-coated copper grids for 2 min. EV samples were negatively stained with 10  $\mu$ L of a 2% uranyl acetate solution for 10 min (ProSciTech, Queensland, Australia). The grids were dried and viewed using a JEOL JEM-2010 transmission electron microscope operated at 80 k.

## 2.4 Nanoparticle tracking analysis (NTA)

sEV/Exo, IEV/MP and sMB-Rs particle diameters were obtained by nanoparticle tracking analysis (NTA) as described [30]. Briefly, EVs were diluted in 1x PBS ( $\sim 8 \times 10^8$  particles/ml) and loaded into a NanoSight NS300 (NanoSight Ltd, UK) using a Nanosight syringe pump. Three separate technical replicates (60 s/video) were recorded for each sample and analyzed by NanoSight NTA 2.3 software.

## 2.5 Label-free mass spectrometry and protein identification

GeLC-S/MS analysis was performed on EV samples (15  $\mu$ g) as described [30]. Briefly, sEVs/Exos, IEVs/MPs, sMB-Rs (15  $\mu$ g) were lysed in SDS sample buffer and proteins in each sample were separated by short-range SDS-PAGE (10  $\times$  6 mm). The samples were excised into equal fractions ( $n = 2$ ), reduced with 2 mM tri(2-carboxyethyl)phosphine hydrochloride (Sigma-Aldrich, C4706) at 22 °C for 4 h on gentle rotation, alkylated by treatment with 25 mM iodoacetamide (Sigma-Aldrich) for 30 min, and digested with 1  $\mu$ g bovine sequencing grade trypsin (Promega, V5111) at 37 °C for 18 h. Reverse-phase C18 StageTips (Sep-Park cartridges, Waters, MA) in 85% (v/v) acetonitrile (ACN) in 0.5% (v/v) formic acid (FA) were used to purified peptides. Subsequently, peptides were lyophilized and acidified with buffer containing 0.1% FA, 2% ACN.

Proteomic experiments were performed with The Minimal Information about a Proteomics Experiment (MIAPE) [31, 32]. A nanoflow UPLC instrument (Ultimate 3000 RSLCnano, Thermo Fisher Scientific) was coupled on-line to a Q-Exactive HF Orbitrap mass spectrometer (Thermo Fisher Scientific) with a nanoelectrospray ion source (Thermo Fisher Scientific). Peptides were loaded (Acclaim PepMap100, 5 mm  $\times$  300  $\mu$ m i.d.,  $\mu$ -Precolumn packed with 5  $\mu$ m C18 beads, (Thermo Fisher Scientific) and separated (BioSphere C18 1.9  $\mu$ m 120 Å, 360/75  $\mu$ m  $\times$  400 mm, NanoSeparations) with a 120-min gradient from 2 to 100% (v/v) phase B (0.1% (v/v) FA in 80% (v/v) ACN) (2–100% 0.1% FA in acetonitrile (2–40% from 0 to 100 mins, 40–80% from 100 to 110 mins at a flow rate of 250 nL/min operated at 55 °C).

The mass spectrometer was operated in data-dependent mode where the top 10 most abundant precursor ions in the survey scan (350–1500 Th) were selected for MS/MS fragmentation. Survey scans were acquired at a resolution of 60,000 with MS/MS resolution of 15,000.

Unassigned precursor ion charge states and singly charged species were rejected, and peptide match disabled. The isolation window was set to 1.4 Th and selected precursors fragmented by high-energy collision dissociation (HCD) with normalised collision energies of 25 with a maximum ion injection time of 110 msec. Ion target values were set to 3e6 and 1e5 for survey and MS/MS scans, respectively. Dynamic exclusion was activated for 30 s. Data was acquired using Xcalibur software v4.0 (Thermo Fisher Scientific)

Raw data were pre-processed as described [32] and processed using MaxQuant [33] (v1.6.0.1) with Andromeda (v1.5.6), using a Human-only (UniProt #133,798 entries) sequence database. Data were searched as described [34] with a parent tolerance of 10 ppm, fragment tolerance of 0.5 Da and minimum peptide length 6, with false discovery rate 1% at the peptide and protein levels, tryptic digestion with up to two missed cleavages, cysteine carbamidomethylation as fixed modification, and methionine oxidation and protein N-terminal acetylation as variable modifications, and data analyzed with label-free quantitation (LFQ). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD041529.

## 2.6 Differential protein enrichment analysis

LFQ intensities of peptide ions identified in sEVs/Exos, lEVs/MPs and sMB-R preparations were statistically analyzed using the edgeR software package [35]. Briefly, LFQ intensities of each protein were normalized based on 'effective library size'[35] for each sample using a trimmed mean of M-values (TMM) [36]. P-values were calculated using Benjamini-Hochberg method [37] and normalized LFQ intensities were presented as log2.

For EV class comparisons three specific comparisons were undertaken: -i) SW480-/SW620-sEVs/Exos vs SW480-/ SW620-lEVs/MPs, -ii) SW480-/ SW620-sEVs/Exos vs SW480-/ SW620-lEVs/sMB-Rs, and -iii) SW480-/ SW620-lEVs/MPs vs SW480-/SW620-sMB-Rs.

For CRC cancer progression-related protein identifications, detected proteins in this study were compared with the Cancer Gene Census from COSMIC database (GRCh38, COSMIC, version 97, <https://cancer.sanger.ac.uk/census#cl.search>)

For protein abundance analysis, SW480-combined EVs were compared with SW620-combined EVs. Highly-enriched proteins in each comparison were identified using the criteria - log2fold change < -1 or > 1 with pvalue < 0.05.

## 2.7 Gene ontology and KEGG pathway analyses

Gene ontology (cellular compartment, specific level 10) and KEGG pathways (organism: hsa, pvalue cutoff: 0.05) were analyzed based on highly-enriched proteins from the statistical analysis comparison (see Section 3.2.6) using clusterProfiler [38] (v.3.11, <https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html>) in R

## 2.8 Data visualization

Principle component analysis, dot, box, ridge and volcano plots were visualized using ggplot (v.3.3.2, <https://ggplot2.tidyverse.org/>) in R. Heatmaps were visualized using pheatmap (v.1.0.12, <https://www.rdocumentation.org/packages/pheatmap/versions/1.0.12>) in R. Venn diagram was generated using a web-based tool (<http://www.interactivenn.net/>) [39]. KEGG pathway analysis was visualized using pathview (v.3.1.2, <http://bioconductor.org/packages/release/bioc/html/pathview.html>) [40] in R.

## 3 RESULTS

### 3.1 Purification and characterization of sEVs/Exos, lEVs/MPs and sMB-Rs from SW480 and SW620 cell culture media

sEVs/Exos, lEVs/MPs and sMB-Rs were isolated in a scalable manner from SW480 and SW620 cells grown in continuous culture using the CELLLine AD-1000 bioreactor device. The three EV types were purified from 180 mL CM acquired over 18 days (30 mL harvested each day, 6 days for each biological replicate, n=3) using a combination of differential centrifugation and isopycnic density (iodixanol) fractionation -

see **Figure 1A** for purification pipeline. Using this approach crude IEVs/MPs and sMB-Rs were isolated from the 10K pellet and crude Exos from the 100K-10K pellet. IEVs/MPs were separated from sMB-Rs by buoyant density (isopycnic iodixanol) gradient centrifugation (**Figure 1B, C**). SDS-PAGE analysis of buoyant density fractions (SYPRO Ruby protein gel stain) showed IEVs/MPs distribute in fractions #6 and #7 (buoyant density  $\sim 1.10$  g/mL) and sMB-Rs in fraction #9 (buoyant density  $\sim 1.14$  g/mL) as evidenced by Western blot analysis using stereospecific marker antibodies for the centralspindlin complex of KIF23/ MKLP1 and RACGAP1, a key component of midbodies [41].

SW480- and SW620-Exos were purified from the 100K-10K pellet by buoyant density (isopycnic iodixanol) gradient centrifugation (fractions #6 and #7, buoyant density 1.08-1.10 g/mL) (**Figure 1D, E**).

Transmission electron microscopic analysis was used to investigate the morphology and size distribution of sEVs/Exos, IEVs/MPs, and sMB-Rs (**Figure 1F**). TEM revealed IEVs/MPs and sMB-Rs are ellipsoid in shape and heterogenous in size (100-500 nm diameter) compared to sEVs/Exos, which displayed round-like structures and a smaller size distribution range (50-200 nm diameter). Nanoparticle tracking analysis (**Figure 1G**) showed mean particle diameters of  $187.7 \pm 74.1$  nm /  $185.2 \pm 63.5$  nm,  $382 \pm 115.2$  nm /  $326.7 \pm 98.9$  nm and  $415.7 \pm 121.2$  nm /  $399.3 \pm 134.6$  nm for SW480/SW620-sEVs/Exos, IEVs/MPs, and sMB-Rs, respectively (biological replicates, n=3).

Western blot analysis revealed sEVs/Exos are ALIX<sup>+</sup>/ TSG101<sup>+</sup>/ CD63<sup>+</sup>/ CD9<sup>+</sup>/ CD81<sup>+</sup>, IEVs/MPs are CD9<sup>+</sup>, and sMB-Rs are KIF23<sup>+</sup>/ RACGAP1<sup>+</sup> (**Figure 1H**).

### 3.2 sEVs/Exos, IEVs/MPs and sMB-Rs secreted by SW480/ SW620 cells are molecularly distinct

**Figure 2** compares the protein profiles of sEVs/Exos, IEVs/MPs and sMB-Rs using a label-free MS approach. Overall, 1544, 1527, 1544, 1527, 1510, and 1760 proteins were identified from SW480-sEVs/Exos, SW620-sEVs/Exos, SW480-IEVs/MPs, SW620-IEVs/MPs, SW480-sMB-Rs and SW620-sMB-Rs, respectively (see Venn diagrams, **Figure 2A and B**). A complete list of global protein identifications in sEVs/Exos, IEVs/MPs and sMB-Rs is given in **Table S1** For the three vesicle types described here from SW480 cells, 101, 76, and 269 proteins were found to be uniquely present (based on presence/ absence of peptide ion intensity) in sEVs/Exos, IEVs/MPs and sMB-Rs, respectively, and 1162 proteins were found to be common to all three EV types (**Figure 2A**). For the three classes of EVs described here from SW620 cells, 186, 55, and 340 proteins are unique to sEVs/Exos, IEVs/MPs, and sMB-Rs, respectively, with 1224 proteins common to the three EV classes (**Figure 2B**). Principle component analysis (PCA) of these datasets shows that the proteomes of sMB-Rs, sEVs/Exos and IEVs/MPs are dissimilar, regardless of parental cell type (**Figure 2C**). These differences in proteomes are further highlighted in the clustering plot in **Figure 2D** which shows that the cluster patterns for sMB-R proteomes from SW480/SW620 cell types show higher degree of similarity than SW480-/SW620-sEVs/Exos and -IEVs/MPs cluster patterns that are more alike.

A full list of uniquely-identified proteins found in sEVs/Exos, IEVs/MPs, and sMB-Rs secreted from SW480 and SW620 cells is given in **Table S2, S3 and S4, respectively**.

Collectively, these data indicate selective trafficking of many SW480- and SW620-cellular proteins to their respective sEVs/Exos, IEVs/MPs and sMB-Rs.

### 3.3 Interrogation of highly-enriched proteins in SW480-/SW620-derived sEVs/Exos, IEVs/MPs and sMB-Rs

To gain insights into similarities and differences between the three EV classes, a deep interrogation of the protein profiles for SW480-/ SW620-derived sEVs/Exos, IEVs/MPs, and sMB-Rs was undertaken to gain insights into similarities and differences between the three EV subtypes. For this comparative analysis, datasets for exosomes from SW480 and SW620 cells were combined and, likewise, datasets for IEVs/MPs and sMB-Rs. Highly-enriched proteins for each combined vesicle subtype were identified using the criteria: log2fold change  $< -1.0$  or  $> 1.0$  with a pvalue  $< 0.05$ . A complete list of highly-enriched protein identifications in sEVs/Exos, IEVs/MPs and sMB-Rs is provided in **Tables S5, S6 and S7, respectively**.

The number of protein identifications highly-enriched in SW480-/SW620-sEVs/Exos (data combined) were compared to corresponding SW480-/SW620-derived IEVs/MPs (data combined) and the number of proteins enriched in SW480-/SW620-derived sEVs/Exos were compared to sMB-Rs. The criteria for proteins that preferentially traffic into one or another EV class was based on those proteins not being evident, or of very low abundance, in *one* EV class, but not the other *two* EV classes. In **Figure 3A**, it can be seen that there are 95 proteins highly-enriched in sEVs/Exos when compared to IEVs/MPs, and 409 proteins highly-enriched in sEVs/Exos when sEVs/Exos are compared to sMB-Rs. When these two datasets are compared (see Venn diagram) 80 proteins selectively traffic to sEVs/Exos. A list of the 80 selectively enriched sEVs/Exos protein identifications is given in **Table S5** .

In the next comparison, 108 proteins were found to be highly-enriched in IEVs/MPs compared to sEVs/Exos and 319 proteins were highly enriched in IEVs/MPs compared to sMB-Rs. When these two data sets were compared (Venn diagram, **Figure 3B** ) only 14 proteins were found to selectively traffic to IEVs/MPs. A list of the 14 proteins selectively enriched in MPs compared to sEVs/Exos and sMB-Rs is given in **Table S6** .

In a third comparative analysis, 604 proteins were found to be highly-enriched in sMB-Rs when sMB-Rs were compared to sEVs/Exos, and 533 when sMB-R proteins were compared to IEVs/MPs (**Figure 3C** ). Of these, 492 proteins were found to selectively traffic to sMB-Rs (a list of selectively-enriched proteins in sMB-Rs is given in **Table S7** ).

Selectively-enriched proteins found in sEVs/Exos (80), IEVs/MPs (14) and sMB-Rs (492) were subjected to GO and KEGG pathway analysis to gain insights into their possible biological roles.

### 3.3.1 Proteins selectively identified in sEVs/Exos

GO terms such as vacuolar membrane and nuclear envelope were identified in all EV subtypes (**Figure 3D** , red stars). GO terms related to endoplasmic reticulum membrane, endosome membrane, coated vesicle membrane, nuclear membrane and secretory granules were co-identified in sEVs/Exos and sMB-Rs (**Figure 3D** , yellow stars). sEVs/Exos showed enriched GO terms such as endocytic membrane, transport vesicle membrane and synaptic vesicles (**Figure 3D** ) and proteins related to exosome biogenesis and trafficking – for example, CD63, CD81, CD82, CHMP4B, VAMP3, VPS25, CD9, SDCBP(Syntenin), ARRDC1, TSG101, CHMP1A, VPS28, VPS37B, VAMP5, TSPAN1, TSPAN6, TSPAN14 (**Figure 3E** , for a list of proteins selectively enriched in sEVs/Exos, IEVs/MPs, and sMB-Rs, see **Table 1** ).

KEGG pathway analysis of proteins highly-enriched in sEVs/Exos compared to IEVs/MPs (95 proteins) and sEVs/Exos compared to sMB-Rs (409 proteins) (**Figure 3A** ) showed mutual KEGG pathways such as *category-i*) Endocytosis (hsa04144), *category-ii*) SNARE interactions in vesicular transport (hsa04130) and, *category-iii*) Cell adhesion molecules (hsa04514) enriched in sEVs/Exos compared to both IEVs/MPs and sMB-Rs (**Figure S2A-F** and **Table S8** ).

**KEGG category-i) Endocytosis** (hsa04144) showed enriched ESCRT proteins in sEVs/Exos (compared to IEVs/MPs and sMB-Rs) (CHMP1 and VPS37, CHMP1, CHMP2, CHMP3, CHMP4, CHMP5, and Clathrin-dependent/-independent endocytosis (PLD, MHCI, E3 ligase, and PLD, MHCI, E3 ligase, SRC) (**Figure S2A , B** ) .

**KEGG category-ii) SNARE interactions in vesicular transport** (hsa04130) displayed enriched VAMP3, VAMP5, VAMP8 and STX7 proteins (**Figure S2C, D** ) .

**KEGG category-iii) Cell adhesion molecules** (hsa04514) showed highly-enriched proteins such as MHCI, PVRL1, PVRL2, CD99 in sEVs/Exos compared to IEVs/MPs and MHCI, PVR, PVRL1, PVRL2, CD99, OCLN, L1CAM in sEVs/Exos compared to sMB-Rs (**Figure S2E, F** ) .

### 3.3.2 Proteins that selectively traffic to IEVs/MPs

GO terms such as vacuolar membrane and nuclear envelope were identified IEVs/MPs (**Figure 3D** , red stars). Several proteins such as DOCK1, DTYMK, FGFR4, FHL1 (**Figure 3E** ), SLC29A2, IMPA1, MRI1

(**Table 1**) were enriched in IEVs/MPs. However, KEGG pathway analysis on 108 and 319 proteins enriched in IEVs/MPs compared to sEVs/Exos and sMB-Rs, respectively did not show mutual KEGG pathways.

### 3.3.3 Proteins that selectively traffic to sMB-Rs

Highly-enriched proteins identified in SW480-/SW620- sMB-Rs exhibited enriched GO terms such as ‘Golgi-associated vesicle membrane’, ‘Intrinsic component of endoplasmic reticulum membrane’ (**Figure 3D**). These identified proteins include AP2A1, CANX, MGST1, PDIA6 (**Figure 3E**), HSPA5 and ERP44 (**Table 1**) mitochondrial membrane, inner mitochondrial membrane protein complex such as SLC25A11, SLC25A13 (**Figure 3E**), TOMM22, VDAC1, VDAC2 (**Table 1**) and spliceosomal complex, ribonucleoproteins such as SNRNP40, U2AF2, HNRNPAB, HNRNPH3, HNRNPL (**Figure 3E**), SF3B1, SF3B2, SF3B3, U2AF2, U2AF1L5, HNRNPA1, HNRNPK, HNRNPD and HNRNPU (**Table 1**). Interestingly, midbody related proteins (AURKB, CEP55, KIF23, PLK1, RACGAP1), histones such as HIST1H4A), HIST1H1C, HIST2H3PS2, HIST2H3A, HIST1H2AC (**Figure 3E, Table 1**) are preeminent. RNA granule proteins such as IGF2BP1, IGF2BP2, FUS and TARDBP (**Table 1**) and several translation initiation factors (EIF families) (**Figure S3**) were also highly-enriched in sMB-Rs.

KEGG pathway analysis on 604 and 533 highly-enriched proteins in sMB-Rs (**Figure 3C**) compared to sEVs/Exos and IEVs/MPs, respectively, showed mutual KEGG pathways such as *category-iv*) RNA transport (hsa03013), *category-v*) Spliceosome (hsa03040), *category-vi*) Protein processing in endoplasmic reticulum (hsa04141), and *KEGG category-vii*) Citrate cycle (TCA cycle) (hsa00020) compared with both sEVs/Exos and IEVs/MPs (**Table S8**).

**KEGG *category-iv*) RNA transport** (hsa03013) revealed highly-enriched proteins in sMB-Rs such as RAE1, SEC13, UBC9, RANGAP (nuclear pore complex), EIF2, EIF2B, EIF3, EIF5 (translation initiation factor), Y14, MAGOH, EIF4A3, PININ, RNPS1 (exon-junction complex) and FUS, TDP43 (Pre-mRNA processing complex) enriched in sMB-Rs (**Figure S2G, H**).

**KEGG *category-v*) Spliceosome** (hsa03040) showed enriched spliceosome components in sMB-Rs such as SM, U1-70K, U1A, FUS (U1 complex), U2A, SF3B, U2AF, PRP43 (U2 complex), SMU13, SAD1 (U4/U6 complex), SNULL4, PRP8BP (U5 complex) and HNRNP families, SR (common components) (**Figure S2I, J**).

**KEGG *category-vi*) Protein processing in endoplasmic reticulum** (hsa04141) demonstrated enriched proteins in sMB-Rs such as OST, CLIMP63, GLCI, GLCII, UGGT, HSP40, PERK, WFS1, SAR1, SEC13/31, NEF, TRAP and BAP31 (**Figure S2K, L**).

**KEGG *category-vii*) TCA cycle** (hsa00020) revealed highly-enriched proteins and mitochondrial enzymes such as PCK1 (4.1.1.32), PDC1 (1.2.4.1), LPD1 (1.8.1.4), IDH1 (1.1.1.42), FUM1 (4.2.1.2), DLST (2.3.1.61), and OGDH (1.2.4.2) (**Figure S2M, N**).

### 3.4 Highly-enriched cancer associated cargo proteins in SW480-/SW620-EV classes that modulate CRC progression

It is well recognized that sEVs/Exos secreted from human metastatic colorectal cancer cells harbor metastatic factors and signaling pathway components that engage in crosstalk between tumor and stromal cells in the tumor microenvironment [42]. To gain insights into how the proteome of individual EV classes might impact on CRC progression, a differential protein enrichment analysis between SW480-EVs (combined Exos, MPs, and sMB-Rs) and SW620-EVs (combined) was performed.

This analysis revealed 230 cancer-associated proteins that are highly enriched in combined SW480-EVs (such as CD44, STAT1, MDK, TGM2, EGFR, FAS, CLDN7) and 264 cancer-associated proteins in SW620-EVs such as RICTOR, MACC1, PRKACA, TGFB2, MET (**Figure 4A**). A complete list of highly-enriched protein identifications in this analysis is provided in **Table S9**.

Using human protein atlas cancer database resource of genes linked from COSMIC cancer database, we showed 14 commonly identified proteins in our proteomic data and COSMIC cancer database (**Table S1**).

Of these 14 commonly identified proteins include KRAS, beta-catenin (CTNNB1), Proto-oncogene SRC, and DNA mismatch repair proteins MSH2 and MSH6 (**Table 2**). Interestingly, the MSH proteins were detected only in sMB-Rs (**Table 2**), while KRAS, CTNNB1 and SRC protein abundance is higher in sEVs/Exos and lEVs/MPs compared to sMB-Rs (**Table 2**).

Next, KEGG pathway analysis was performed on these cancer-associated cargo proteins to gain insights into possible functional roles. KEGG pathway analysis on SW480-EV highly-enriched cancer-associated proteins (230 proteins) revealed involvement of 7 pathways including Ribosome (hsa03010), Cell adhesion molecule (hsa04514), Proteasome (hsa03050) and Phagosome (hsa04145) (**Figure 4B**). In contrast, KEGG pathway analysis on proteins enriched in SW620-EVs (264 proteins) showed 16 pathways that include MAPK signaling pathway (hsa04010), Proteoglycans in cancer (hsa05205), Insulin signaling pathway (hsa04931) and ErbB signaling pathway (hsa04012) (**Figure 4B**).

We next questioned whether the EV classes exhibit different functionalities in cancer progression. Proteomic analysis showed that many cancer-progression-associated proteins are highly or uniquely sorted in the separate EV classes. For instance, protein-related to genetic stability such as PARP1 was highly-enriched in SW480-sMB-Rs and histone deacetylase (HDAC1) and its substrate MSH6 (DNA mismatch repair protein) were uniquely detected in SW480-sMB-Rs (**Table 2**). Several receptors and transporters are highly enriched in sEVs/Exos and lEVs/MPs (CD44, EGFR, FAS in SW480-sEVs/Exos and -lEVs/MPs and MET, FGFR4 in SW620-sEVs/Exos and -lEVs/MPs) (**Table 2**). Interestingly, chemokine receptor (CXCR4) was uniquely detected in SW480- sEVs/Exos and TGFBR2, AXL and ABCB1 were uniquely detected in SW620-sEVs/Exos (**Table 2**). Signal transduction proteins such as CLDN7 was highly enriched in sEVs/Exos and lEVs/MPs derived from SW480 cells (**Table 2**). Phospholipase D1 (PLD1) was only detected in SW480-sEVs/Exos and STAT1, GYS1 and CTNBL1 were uniquely detected in SW480-sMB-Rs (**Table 2**). Signal transduction proteins such as PLD2, SMAD5 and TP53RK were uniquely detected in SW620-sEVs/Exos and LAMTOR3 is highly-enriched in SW620-sEVs/Exos (**Table 2**). PRKACA, MACC1 and RICTOR were highly enriched in sEVs/Exos, lEVs/MPs and sMB-Rs from SW620 compared to sEVs/Exos, lEVs/MPs and sMB-Rs from SW480. Midkine (MDK) growth factors was highly-enriched in SW480-sMB-Rs and growth/differentiation factor-15 (GDF15) was uniquely detected in SW620-sMB-Rs (**Table 2**). Interestingly, ECM remodeling proteins such as tissue transglutaminase-2 (TGM2) and matrix metalloproteinase MMP14 are enriched in SW480-EVs with selective enrichment in sMB-Rs and lEVs/MPs, respectively (**Table 2**). ADAM15 and tenascin (TNC) were uniquely identified in SW620-sEVs/Exos and SW620-sMB-Rs, respectively.

In summary, proteomic analysis showed selective enrichment of cancer -associated proteins in different EV classes from primary CRC (SW480) and metastatic CRC (SW620) cell lines.

#### 4 DISCUSSION

In this study we performed a detailed comparative proteome analysis of three EV classes (sEVs/Exos, lEVs/MPs and sMB-Rs) secreted from isogenic human primary and metastatic colorectal cancer cells - SW480 cells (surrogate of adenocarcinoma) and SW620 (surrogate of metastatic colon cancer) [43].

sEVs/Exos, lEVs/MPs and sMB-Rs were enriched and purified in high yield from the same preparation of SW480/ SW620 CM using a combination of differential centrifugation and isopycnic buoyant density (iodoxinol/ OptiPrep) centrifugation (**Figure 1A**). The yields of SW480-/ SW620-derived sEVs/Exos (from resuspended 100K-10K pellet at buoyant density = 1.10 g/mL), lEVs/MPs (from resuspended 10K pellet at buoyant density = 1.10 g/mL), and sMB-Rs (from resuspended 10K pellet at buoyant density = 1.14 g/mL) from 180 mL CM were 868-987  $\mu$ g, 749-827  $\mu$ g, and 463-660  $\mu$ g, respectively. TEM revealed lEVs/MPs and sMB-Rs are more ellipsoid in shape and heterogenous in size (100-500 nm) compared to sEVs/Exos, which displayed a smaller size distribution range (50-200 nm) (**Figure 1F**). Nanoparticle tracking analysis showed mean particle diameters of  $187.7 \pm 74.1$  nm/ $185.2 \pm 63.5$ ,  $382 \pm 115.2$  nm/  $326.7 \pm 98.9$  nm and  $415.7 \pm 121.2$  nm/ $399.3 \pm 134.6$  nm for SW480/SW620-sEVs/Exos, lEVs/MPs, and sMB-Rs, respectively (**Figure 1G**). The diameters of sEVs/Exos, and lEVs/MPs are similar to those reported in previous studies[1] Western blot

analysis revealed ALIX, TSG101, CD63, CD9 and CD81 were more enriched in sEVs/Exos than in IEVs/MPs (**Figure 1H**), a finding consistent with previous studies [27, 30, 44]. The midbody centraspindlin complex components KIF23/MKLP1 and RACGAP1 [24, 41] are selectively enriched in sMB-Rs (**Figure 1H**). While there is no gold standard protocol to acquire sEVs/Exos, IEVs/MPs and sMB-Rs at a high level of purity, the EV characterization results from this study demonstrated that a combination of differential centrifugation and isopycnic buoyant density (iodoxinol/ OptiPrep) centrifugation can be used to enrich the abundance of sEVs/Exos, IEVs/MPs and sMB-Rs.

Interestingly, the proteome of sMB-Rs is dissimilar to sEVs/Exos and IEVs/MPs as indicated in PCA and clustering plots (**Figure 2C and D**). Differential protein enrichment analysis revealed 80, 14 and 492 proteins are selectively enriched (using a cutoff,  $>1$  or  $<-1$  log<sub>2</sub> fold change and pvalue  $< 0.05$ ) in sEVs/Exos, IEVs/MPs and sMB-Rs, respectively (**Figure 3A, B and C**).

When compared to combined SW480/SW620 IEVs/MPs and sMB-Rs, 80 proteins were found to be selectively-enriched in sEVs/Exos. These include exosomal protein markers [19, 27, 45] such as CD63, CD81, CHMP4B, TSG101 (**Figure 3E**). KEGG pathway analysis of these selectively-enriched proteins showed they are involved in pathways such as Endocytosis (hsa04144) and SNARE interactions in vesicular transport (hsa04130) and Cell adhesion molecules (hsa04514) (**Figure S2A-F**). This finding further supports the notion that sEVs/Exos biogenesis involves endosome dynamics and trafficking, as previously reported [1, 46].

When SW480/SW620 combined datasets for IEVs/MPs were compared to sEVs/Exos and sMB-Rs 14 proteins were found to be selectively-enriched (**Figure 3B**, **Table S6**); GO analysis revealed that these 14 proteins are involved in vacuolar membranes and nuclear envelopes (**Figure 3D**). IEVs/MPs did not share any KEGG pathways identified sEVs/Exos and sMB-Rs. IEVs/MPs contain enriched proteins such as DOCK1, DTYMK, FGFR4, FHL2, IMPA2, MRI1 and SLC29A2 not observed in sEVs/Exos and sMB-Rs (**Figure 3E**). Interestingly, proteins purported to be involved in IEVs/MP biogenesis such as ARRDC1 [47], ARF1 [48], ARF6 [49], RHOA [50] and ANXA5 [11] are not only enriched in IEVs/MPs but also in sEVs/Exos derived from SW480 and SW620 in this current study. This paradox raises the specter of EV purity – for example, possible cross contamination with other EV types is highly probable, especially in the case of IEVs/MPs and sEVs/Exos which overlap in size distribution and have similar biophysical properties such as buoyant density.

Concerning the 492 selectively-enriched proteins found in sMB-Rs (compared to sEVs/Exos and IEVs/MPs, **Figure 3C**), GO analysis revealed terms such as small nuclear ribonucleoprotein (RNP) complex, spliceosomal complex, and organelles such as mitochondrial membrane part, Golgi-associated vesicle membrane, endoplasmic reticulum membrane and nuclear membrane (**Figure 3D and E**). KIF23/ MKLP1, CEP55, and RACGAP1, which are core structural proteins of midbody/midbody remnants were highly enriched in sMB-Rs, as well as key enzymes involved in midbody formation (AURKB, PLK1) [41] (**Figure 3E**). KEGG pathway analysis of the 492 proteins that selectively traffic to sMB-Rs showed pathways such as ‘RNA transport’ (hsa03013), ‘Spliceosome’ (hsa03040), ‘Protein processing in endoplasmic reticulum’ (hsa04141) and ‘Citrate cycle (TCA cycle) and mitochondrial enzymes’ (hsa00020) (**Figure S2G-N**).

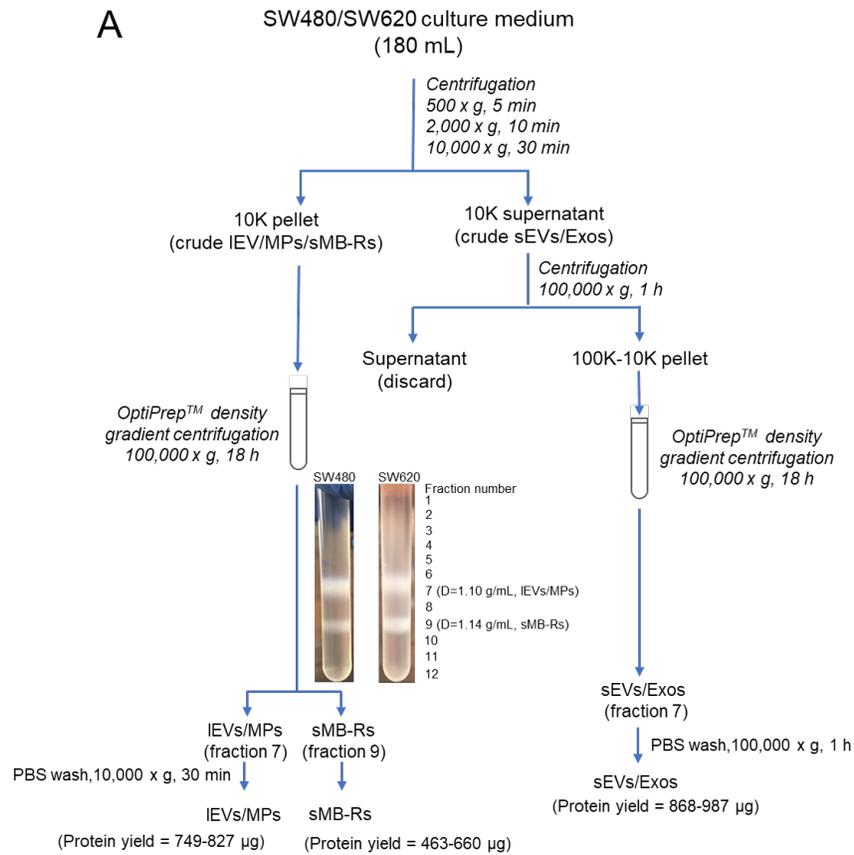
The observed enrichment of RNA granule and mitochondrial proteins in sMB-Rs is intriguing. While typical RNA granule and mitochondria isolation methods [51, 52] are similar to the EV subtype isolation methods used in this study, TEM images of sMB-Rs derived from both SW480 and SW620 cells (**Figure 1F**) did not show any evidence of intact mitochondria (see additional TEM images of sMB-Rs in **Figure S1**). However, it is evident that mitochondrial bodies can be released from cells as mitovesicles and transferred to recipient cells [53, 54]. As these mitochondrial bodies are ~100-200 nm (small than IEVs/MPs and sMB-Rs) [54] and size ranges of EVs and mitovesicles overlap [55], it is possible that mitochondrial bodies/mitovesicles are co-isolated or incorporated in IEVs/MPs and sMB-Rs. Biological experiments are needed to prove this concept. Interestingly, Skop and colleagues reported the identification of mitochondrial (26%), nuclear (16%) and ribosomal (13%) proteins in the midbody proteome [2] and RNA localization, nuclear transport, RNA splicing, mitochondrion organization rate as top fold enrichment GO annotation profiles of midbody proteome and

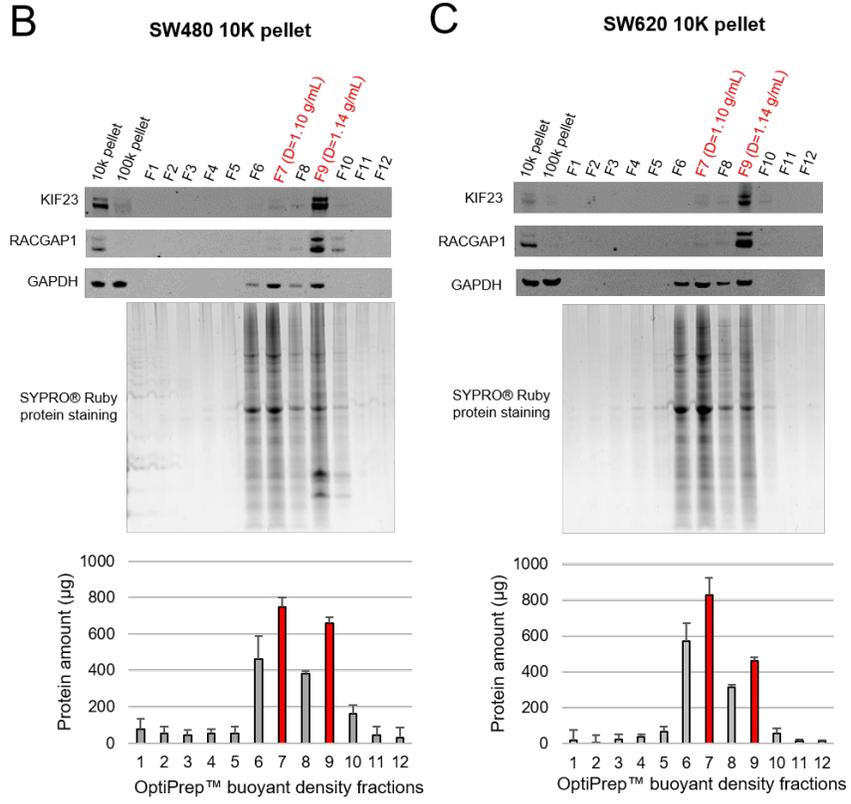
interactome using PANTHER [56, 57]. Furthermore, 25/492 selectively-enriched sMB-R proteins identified (RACGAP1, KIF4A, KIF23, CEP55, PLK1, for example) co-identified with proteins listed in the MiCroKITS v.4.0 database [58] (a database of proteins temporally and spatially localized in distinct subcellular positions including midbody, centrosome, kinetochore, telomere, and mitotic spindle during cytokinesis (cell division/mitosis) (<http://microkit.biocuckoo.org>, **Table S7** ). In another midbody remnant study enriched vesicular traffic transport and protein-translation related proteins were reported in the ‘Flemmingsome’ (referred as ‘post-abscission midbody’ or ‘midbody remnants’) using STRING functional association network [59]. In this present study also compared our sMB-R proteome with the ‘Flemmingsome’ proteome from Addi and colleagues. Interestingly, 271/492 enriched proteins in our purified sMB-Rs are found in the ‘Flemmingsome’ proteome. These identified proteins include RNA granule proteins (FUS, IGF2BP1, TARDBP), ribonucleoproteins and splicing factors (HNRNPs, SFPQ, SF3B3) endoplasmic reticulum proteins (CALR, CANX), mitochondrial proteins (VDAC1, VDAC2, SLC25A3, SLC25A5, SLC25A6), histones (HIST1H1C, HIST2H3A) and midbody proteins (KIF2A, KIF4A, KIF23, RACGAP1, PLK1, CEP55) (**Table S7** ). While there is accumulating evidence of the presence mitochondrial proteins [60], RNA binding proteins (RBPs) [61] in EVs such as sEVs/Exos, the comparative proteomic analysis in this study reveals that mitochondrial and RBPs are also enriched in sMB-Rs. Further studies are required to discern whether RNA granules, mitochondrial proteins and other cellular organelle proteins are physiologically relevant components of sMB-Rs.

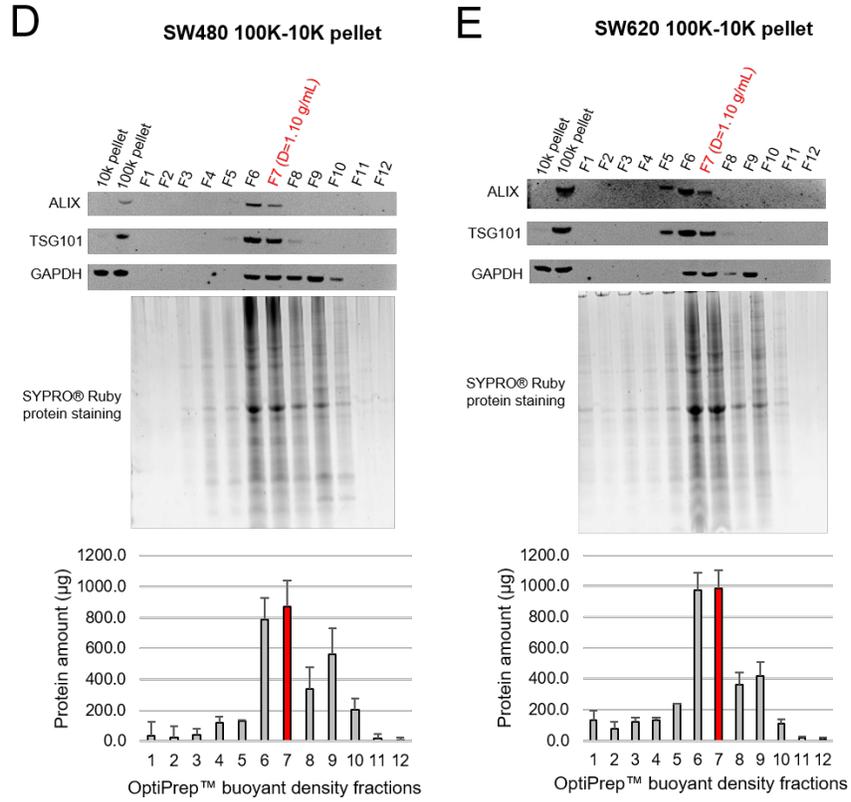
Cancer progression-associated proteins have been previously shown to selectively traffic to EVs [30, 42, 61]. In our present study receptors such as CXCR4 and TGFBR2/AXL were uniquely detected in SW480-sEVs/Exos and SW620-sEVs/Exos, respectively. A possible reason for receptor enrichment in sEVs/Exos is that receptors such as EGFR, HER2, ERBB3 and ERBB4 bind to their cognate ligands and are then internalized into the intercellular early endosome which further develops to late endosome and multivesicular body (MVB), respectively [62]. MVBs that contain receptor-intact intraluminal vesicles (exosomes) can fuse either with lysosomes (leading to proteolytic degradation) or traffic to the plasma membrane whereupon sEVs/Exos are released into the extracellular milieu [63].

HDAC2 and MSH6 [64] were uniquely detected in SW480-sMB-Rs (**Table 2** ). Deacetylation of DNA mismatch repair proteins (MSHs) by HDAC1 leads to destabilization of MSHs, resulting in genetic instability [65]. Furthermore, cancer progression-proteins such as STAT1 [66], TGM2 [67], MDK [68] were found to be highly enriched in sMB-Rs from SW480 cells and PRKACA [69] SW620 cell-derived sMB-Rs. Compared with Cancer Gene Census from COSMIC database, we found 14 proteins commonly identified with this study (**Table S1** ). Of these 14 proteins, colorectal oncoproteins such as KRAS[70] and beta-catenin[71] are abundant in sEVs/Exos and IEVs/MPs compared to sMB-Rs. Oncogenic KRAS has been shown to be delivered by sEVs/Exos[72]. Our study demonstrated that KRAS protein abundance in IEVs/MPs is as high as in sEVs/Exos (**Table 2** ), suggesting an implication of IEVs/MPs in colorectal cancer progression. A salient finding was the detection of the DNA mismatch repair protein MSH2 and MSH6 only in sMB-Rs (**Table 2** ). MSH2 and MSH6 form a dimeric complex [73] which is implicated in hereditary nonpolyposis colorectal cancer (HNPCC) [74].

In summary, the proteome of SW480/ SW620 sMB-Rs is distinct from sEVs/Exos and IEVs/MPs proteomes. The sMB-R proteome is high enriched with mitochondrial proteins (membrane proteins and enzymes), RNA granule proteins, splicing factors, ribonucleoproteins, histone subunits, translation initiation factors and integral components of midbodies. SW480/SW620 cell-derived sEVs/Exos are highly-enriched in tetraspanins/glycoproteins (TSPAN1, TSPAN14, CD63, CD81, CD82) and ESCRT components (TSG101, CHMP1A, CHMP4B). IEVs/MPs are highly-enriched in enzymes (DTYMK, IMPA1 and MRI) and membrane-associated proteins (SLC29A2, FGFR4). This study provides, for the first time, an in-depth comparative proteomic analysis of three EV classes (sEVs/Exos, IEVs/MPs and sMB-Rs) which were purified simultaneously from two CRC cell types (SW480 and SW620 cells). This comparative proteome study paves the way to advancing the characterization of EV classes and in doing so may impact on our understanding of intercellular communication.

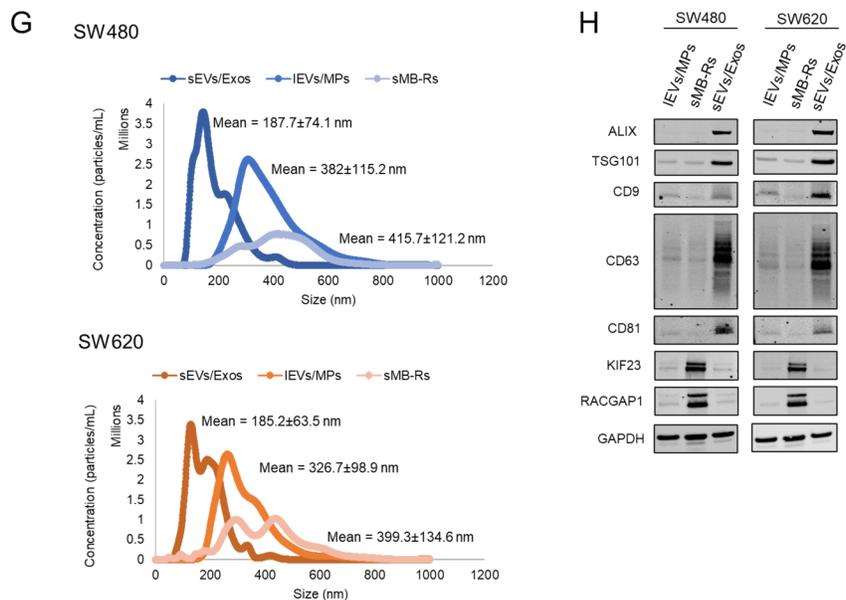






### Hosted file

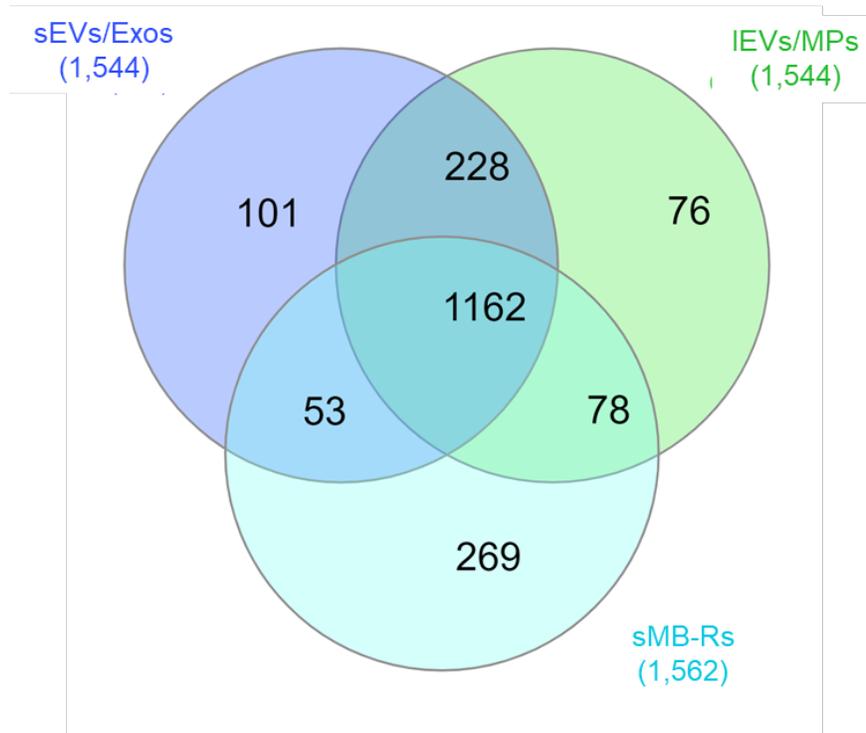
image4.png available at <https://authorea.com/users/614198/articles/641154-comparative-proteomic-analysis-of-three-major-extracellular-classes-secreted-from-human-adenocarcinoma-and-metastatic-colorectal-cancer-cells-exosomes-microparticles-and-shed-midbody-remnants>



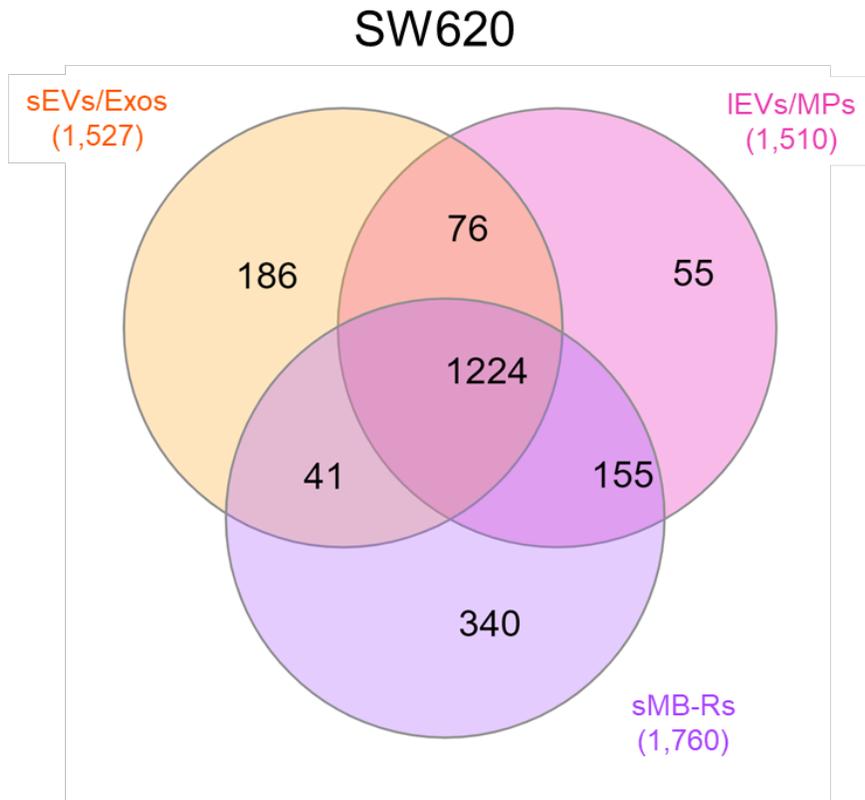
**Figure 1. Isolation and characterization of SW480 and SW620 cell-derived sEVs/Exos, IEVs/MPs, and sMB-Rs.** (A) Experimental workflow used for isolation of sEVs/Exos, IEVs/MPs and sMB-Rs from culture media of SW480 and SW620 using differential centrifugation in combination with OptiPrep density gradient centrifugation. (B, C) Western blot analysis of purified SW480-/ SW620-IEVs/MPs (fraction #7), -sMB-Rs (fraction #9), crude 10K pellet, crude sEVs/Exos (100K pellet) using anti-KIF23, anti-RACGAP1 and anti-GAPDH antibodies, protein quantification of OptiPrep fractions was performed using SDS-PAGE and SYPRO quantitative protein staining, n=3. (D, E) Western blot analysis of purified SW480-/ SW620-sEVs/Exos (fraction #7) using anti-ALIX, anti-TSG101 and anti-GAPDH antibodies, protein quantification of OptiPrep fractions was performed using SDS-PAGE and SYPRO quantitative protein staining, n=3. (F) Transmission electron microscopic analysis of SW480-/ SW620-sEVs/Exos, -IEVs/MPs and -sMB-Rs. (G) Nanoparticle tracking analysis (NTA) of SW480-/ SW620-sEVs/Exos, -IEVs/MPs, and -sMB-Rs (mean ± SD). (H) Western blot analysis of purified SW480-/ SW620-sEVs/Exos, -IEVs/MPs, and sMB-Rs using anti-ALIX, anti-TSG101, anti-CD9, anti-CD63, anti-CD81, anti-KIF23, anti-RACGAP1 and anti-GAPDH antibodies (Protein load: 20 µg protein per lane, n=3).

A

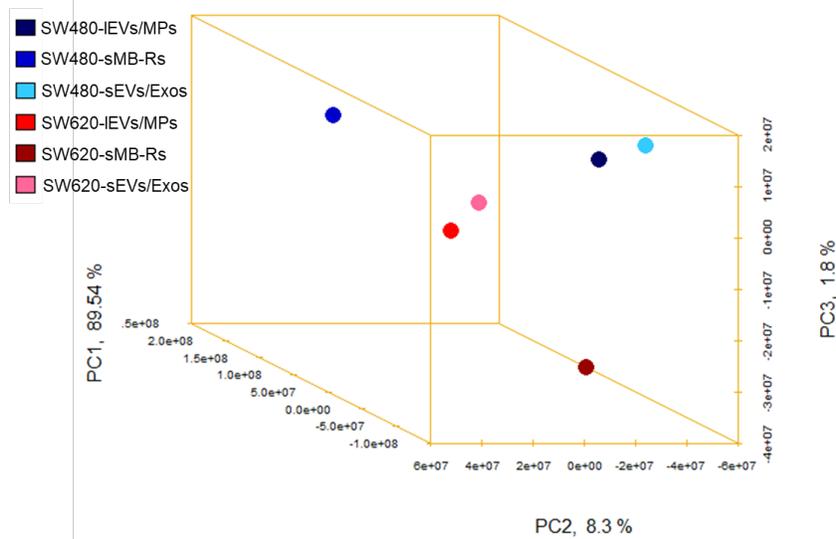
SW480

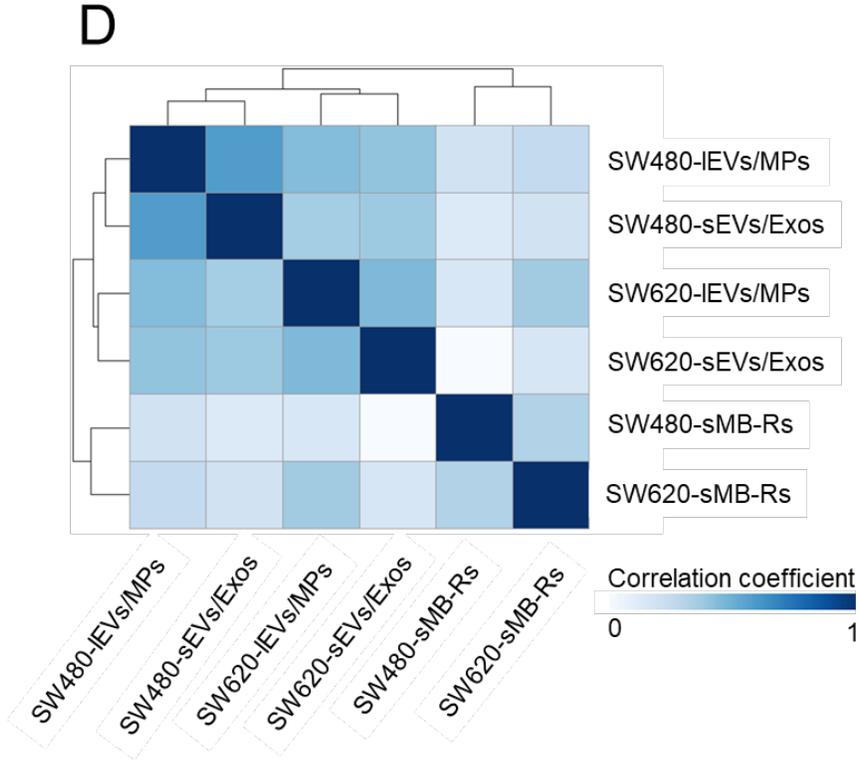


**B**

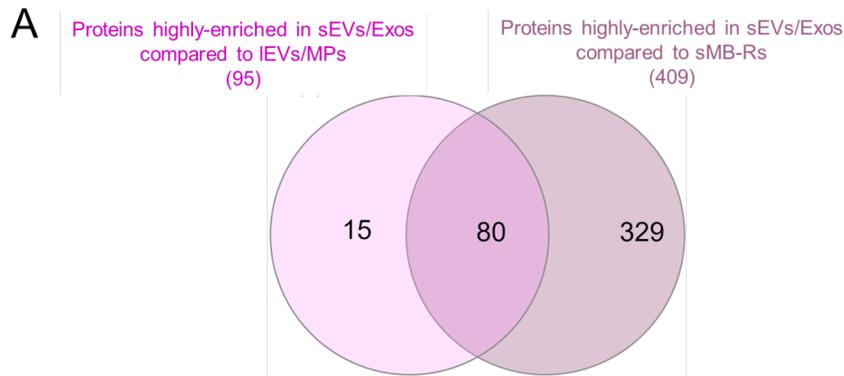


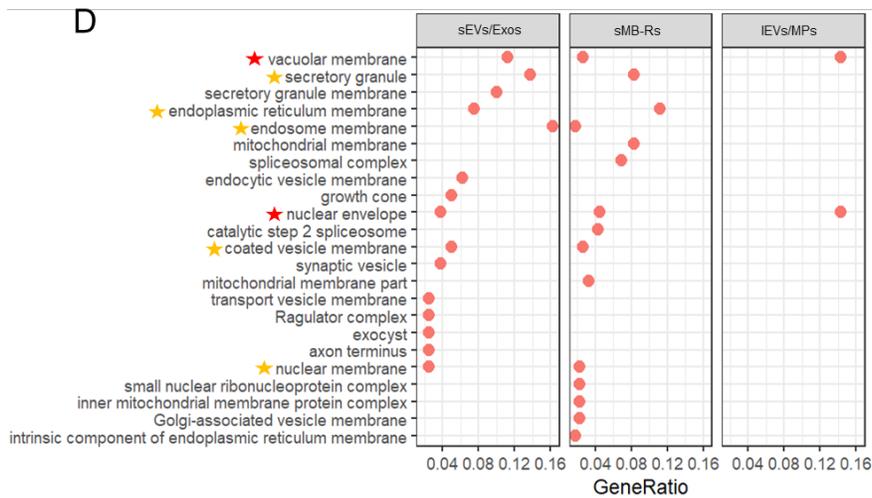
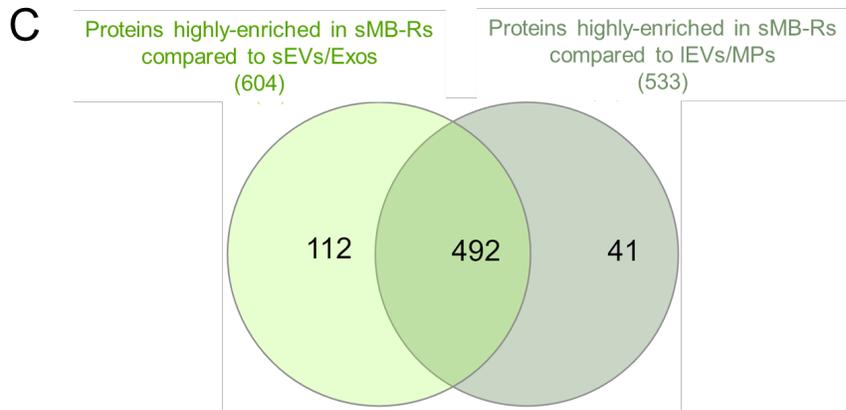
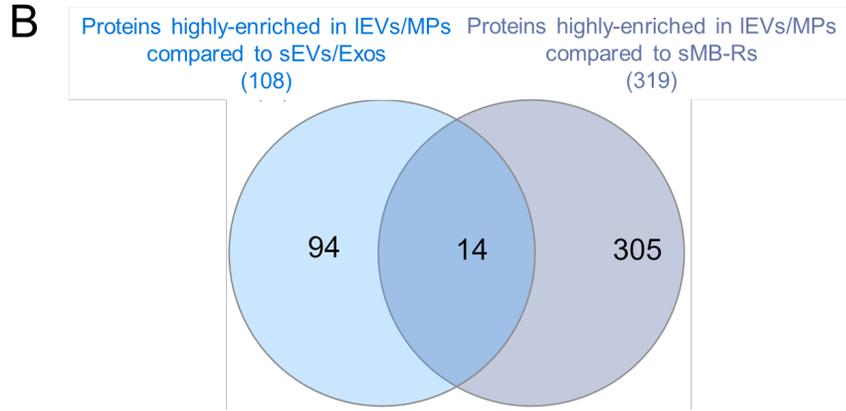
**C**

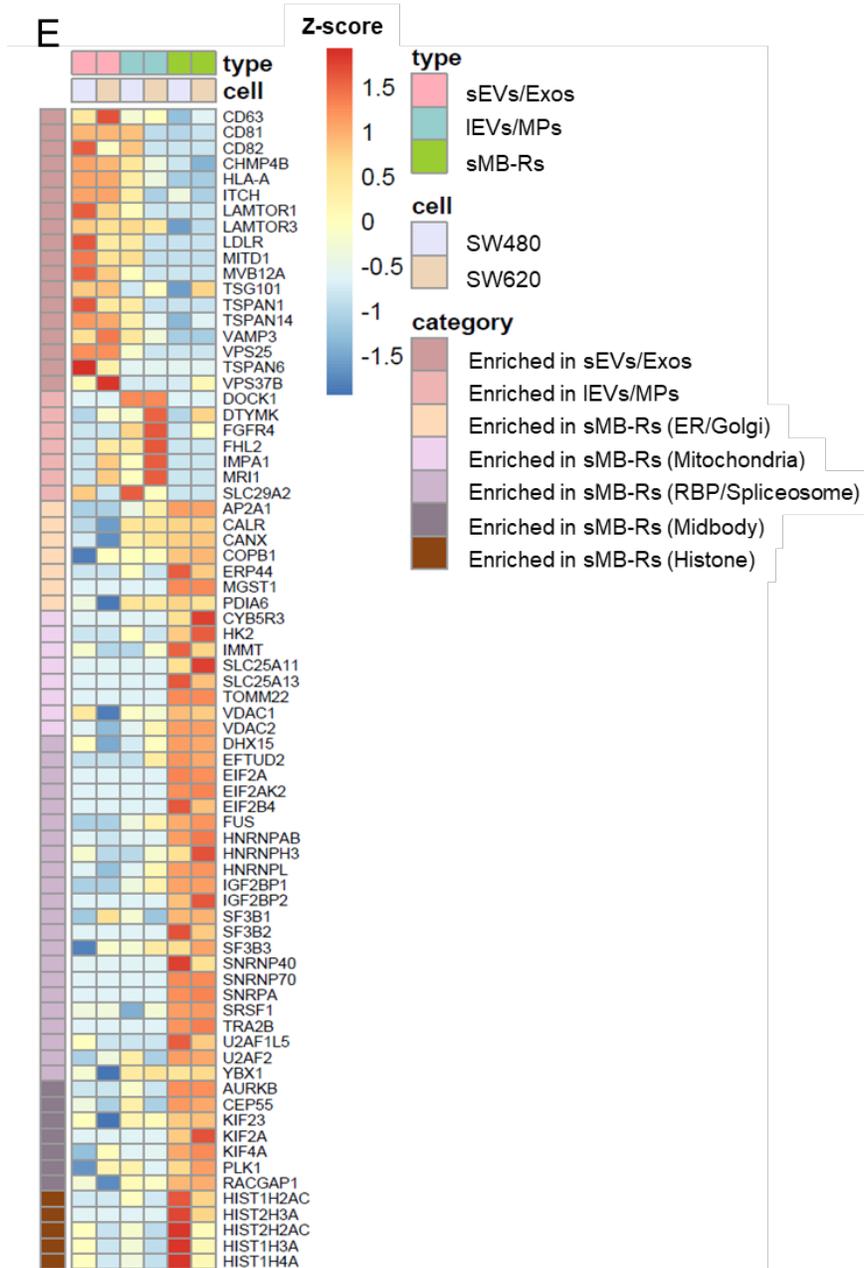




**Figure 2. Proteomic profiling of purified sEVs/Exos, IEVs/MPs and sMB-Rs derived from SW480 and SW620 cell lines.** (A) A three-way Venn diagram of proteins identified in SW480-sEVs/Exos, -IEVs/MPs and -sMB-Rs reveals 1162 proteins were commonly identified, while 101, 76, and 269 proteins were uniquely identified in SW480-sEVs/Exos, IEVs/MPs, and sMB-Rs, respectively. (B) A three-way Venn diagram of proteins identified in SW620-sEVs/Exos, -IEVs/MPs reveals 186, 55, and 340 proteins were uniquely identified in SW620-sEVs/Exos, IEVs/MPs and sMB-Rs, respectively. (C) Principal component analysis (PCA) and (D) clustering analysis of total proteomes of sEVs/Exos, IEVs/MPs and sMB-Rs derived from SW480 and SW620 cell lines.

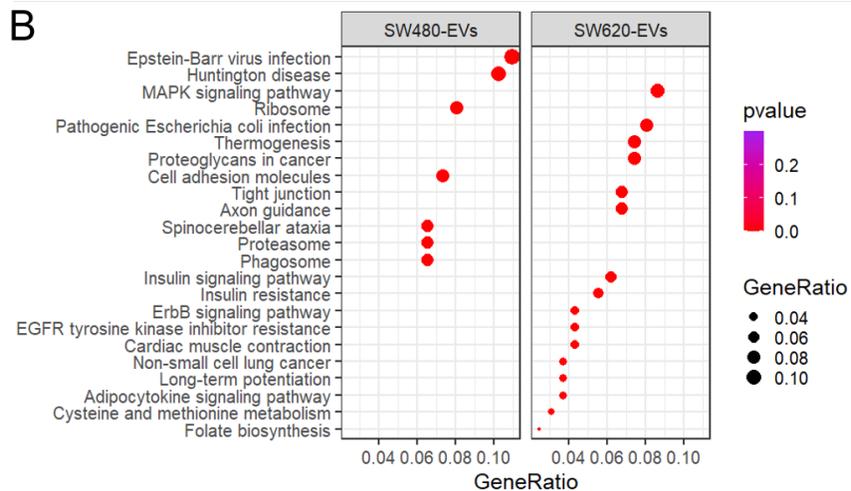


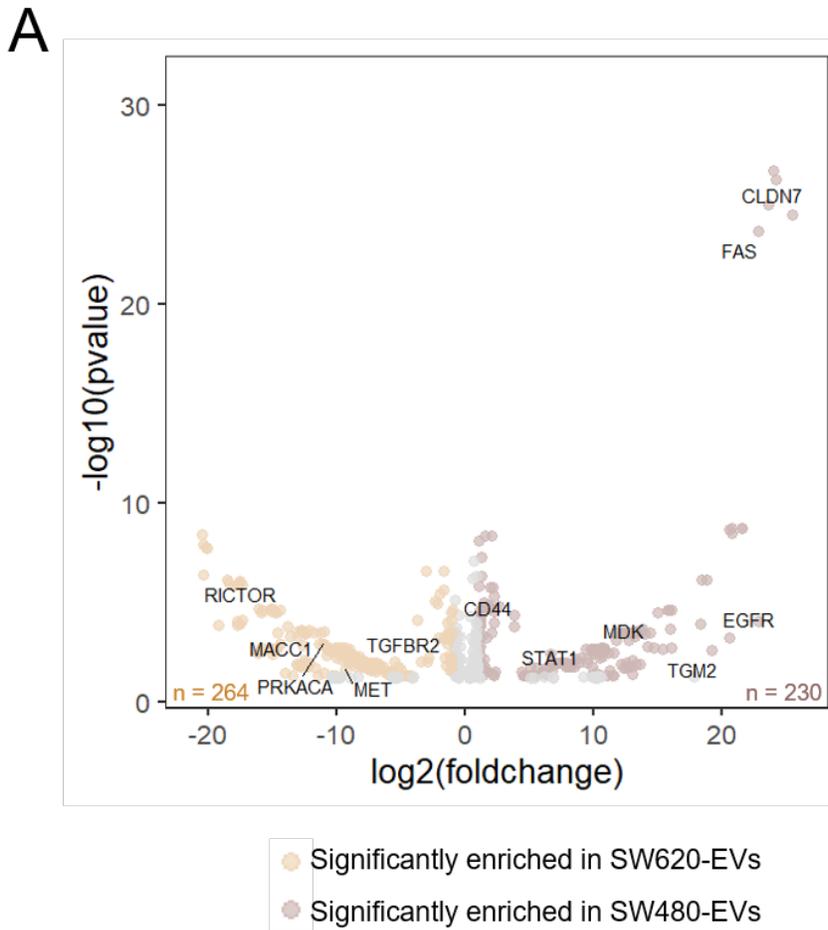




**Figure 3. Comparative proteomic analysis of sEVs/Exos, IEVs/MPs and sMB-Rs derived from SW480 and SW620 cell lines. For this analysis EV datasets from SW480/SW620 were combined. (A)** A two-way Venn diagram of selectively-enriched proteins ( $\log_2$  fold change  $>1$  or  $<-1$  and  $p$ value  $< 0.05$ ) in sEVs/Exos compared to IEVs/MPs and sMB-Rs reveals 80 proteins selectively enriched in sEVs/Exos. **(B)** Two-way Venn diagram of selectively-enriched proteins in IEVs/MPs compared to sEVs/Exos and sMB-Rs reveals 14 proteins selectively-enriched in IEVs/MPs. **(C)** Two-way Venn diagram of selectively-enriched proteins in sMB-Rs compared to sEVs/Exos and IEVs/MPs reveals 492 proteins selectively-enriched in sMB-Rs. **(D)** Identification of enriched gene ontology (GO) term (ranked by protein counts) in sEVs/Exos, IEVs/MPs and sMB-Rs based on selectively-enriched proteins in sEVs/Exos (80 proteins), IEVs/MPs (14 proteins) and sMB-Rs (492 proteins). Red stars indicate commonly-identified

GO terms in sEVs/Exos, IEVs/MPs and sMB-Rs. Yellow stars indicate commonly-identified GO terms in 2 EV classes. **(E)** Heat map illustration of selectively-enriched proteins in SW480-/SW620-derived sEVs/Exos, IEVs/MPs and sMB-Rs (scale shown is average normalized LFQ subtracted by mean and divided by standard deviation).





**Figure 4. Identification of cancer progression-related proteins and KEGG pathways in EVs derived from SW480 and SW620 cell lines.** (A) Differential protein enrichment analysis of highly-enriched ( $\log_2$  fold change  $> 1$ ,  $p$ value  $< 0.05$ ) cancer-associated cargo proteins in SW480-EVs (230 proteins) and SW620-EVs (264 proteins). (B) KEGG pathway analysis (ranked by  $p$ value) of highly-enriched cancer-associated proteins found in SW480-EVs and SW620-EVs.

**Table 1 Selectively-enriched proteins in sEVs/Exos, lEVs/MPs and sMB-Rs secreted from SW480 and SW620 cells**

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup> SW480-sEVs/Exos	Normalized LFQ intensity <sup>a</sup> SW480-lEVs/MPs	Normalized LFQ intensity <sup>a</sup> SW480-sMB-Rs	Normalized LFQ intensity <sup>a</sup> SW620-sEVs/Exos	Normalized LFQ intensity <sup>a</sup> SW620-lEVs/MPs
sEVs/Exos	Exosomal biogenesis	CD9	Tetraspanin	A6NNI4	25,519,078	15,249,895	4,228,553	8,346,331	3,882,327
		SDCBP	Syntenin-1	O00560	5,704,720	2,535,526	267,311	3,124,564	64,040

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
		SDCBP2*	Syntenin-2	Q9H190	29,811	-	-	137,412	-
		CD81	Tetraspanin	E9PJK1	2,906,182	2,072,424	721,091	2,243,535	968,283
		CD63	CD63 antigen	F8VV56	962,989	376,641	-	525,147	36,030
		ARRDC1	Arrestin domain-containing protein 1	Q8N5I2	646,143	360,359	9,284	178,366	47,319
		CD82	CD82 antigen	P27701	476,529	279,614	-	120,106	35,493
		CHMP4B	Charged multi-vesicular body protein 4b	Q9H444	135,214	115,238	19,315	127,888	30,932
		TSG101	Tumor susceptibility gene 101 protein	Q99816	132,686	113,310	50,004	256,676	100,816
		CHMP1A	Charged multi-vesicular body protein 1a	F8VUA2	32,297	12,682	4,571	39,378	-
	Vesicle-associated proteins	VPS28	Vacuolar protein sorting-associated protein 28	Q9UK41	250,149	108,190	36,386	336,014	71,659
		VPS37B	homolog Vacuolar protein sorting-associated protein 37B	Q9H9H4	96,613	36,452	-	106,115	16,535

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
		VPS25	Vacuolar protein-sorting-associated protein 25	Q9BRG1	63,915	29,509	-	99,477	39,172
		VAMP3	Vesicle-associated membrane protein 3	Q15836	61,516	36,033	-	28,717	-
		VAMP5	Vesicle-associated membrane protein 5	O95183	55,443	26,760	-	45,797	-
	Membrane associated proteins	TSPAN1	Tetraspanin-O606351		128,571	73,976	-	28,886	-
		TSPAN14	Tetraspanin-Q8NG1114		67,265	19,594	-	79,319	-
		TSPAN6	Tetraspanin-O436576		32,107	14,280	-	20,362	-
		CD99*	CD99 antigen	P14209	3,780	-	-	16,497	-
		NECTIN1*	Nectin cell adhesion molecule 1 (CD antigen CD111)	Q15223	2,277	-	-	7,040	-
IEVs/MPs	Membrane associated proteins	SLC29A2	Equilibrative nucleoside transporter 2	Q14542	12,057	19,309	-	-	4,976
		FGFR4	Fibroblast growth factor receptor 4	P22455	-	2,420	-	6,308	10,476

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
sMB-Rs	Enzyme	DTYMK*	Thymidylate kinase	P23919	-	18,102	-	-	15,154
		IMPA1	Inositol monophosphatase 1	P29218	-	9,695	-	13,144	47,096
		MRI1	Methylthioribose 1-phosphate isomerase	Q9BV20	-	7,804	-	-	34,210
	Histone	HIST1H1C	Histone H1.2	P16403	146,951	103,434	17,457,590	-	-
		HIST2H3PS2	Histone H3	Q5TEC6	172,367	99,191	7,126,765	30,482	19,164
		HIST1H2AC	Histone H2A type 1-C	Q93077	-	138,759	2,051,667	-	-
		HIST2H3A*	Histone H3.2	Q71DI3	-	-	1,149,510	-	-
	RNA granule protein	IGF2BP1	Insulin-like growth factor 2 mRNA-binding protein 1	Q9NZI8	-	7,184	50,988	-	9,292
		FUS	RNA-binding protein FUS	P35637	-	4,148	17,101	-	6,866
		TARDBP	TAR DNA-binding protein 43	Q13148	-	-	7,698	-	8,158
		IGF2BP2*	Insulin-like growth factor 2 mRNA-binding protein 2	F8W930	-	-	6,518	-	-

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
	Ribonucleoprotein	HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	F8W6I7	189,305	280,594	667,383	204,208	577,883
		HNRNPK	Heterogeneous nuclear ribonucleoprotein K	R61978	102,706	143,624	623,067	63,661	164,451
		SRSF7*	Serine/arginine-rich-splicing factor 7	A0A0B4J1Z1	-	-	329,624	-	-
		HNRNPD	Heterogeneous nuclear ribonucleoprotein D0	Q14103	26,432	48,492	211,402	38,889	103,100
		SRSF1	Serine/arginine-rich-splicing factor 1	I3KTL2	22,406	5,156	176,279	17,424	21,063
		HNRNPU	Heterogeneous nuclear ribonucleoprotein U	Q00839	6,567	6,574	162,877	-	3,206
		U2AF1L5	Splicing factor U2AF 35 kDa subunit-like protein	P0DN76	11,168	-	89,469	-	-
		U2AF2	Splicing factor U2AF 65 kDa subunit	K7ENG2	-	16,353	54,726	4,994	-

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
		LARP1*	La ribonucleo-protein domain family member 1	Q6PKG0	-	-	5,240	-	-
		SFPQ	Splicing factor, proline- and glutamine-rich	P23246	-	698	44,513	-	-
		SRSF7*	Serine/arginine-rich-splicing factor 7	A0A0B4J1Z1	-	-	299,849	-	-
		SF3B1	Splicing factor 3B subunit 1	O75533	3,472	7,735	33,276	5,462	-
		SRSF10*	Serine/arginine-rich-splicing factor 10	Q5JRI1	-	-	20,123	-	-
		SF3B3	Splicing factor 3B subunit 3	Q15393	870	1,516	18,003	1,698	10,518
		SF3B2*	Splicing factor 3B subunit 2	E9PPJ0	-	-	6,861	-	-
	Organelle-associated protein	HSPA5	Endoplasmic reticulum luminal Ca(2+)-binding protein grp78	P11021	17,406	190,663	1,089,414	32,950	305,576

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
		CYB5B	Cytochrome b5 type B	J3KNF8	341,891	406,353	976,467	-	156,024
		VDAC1	Voltage-dependent anion-selective channel protein 1	P21796	113,357	204,643	943,804	-	81,780
		CANX	Calnexin	P27824	16,262	57,326	343,148	5,218	108,205
		VDAC2	Voltage-dependent anion-selective channel protein 2	P45880	7,296	11,766	189,210	-	29,286
		ERP44	Endoplasmic reticulum resident protein 44	Q9BS26	-	17,941	77,026	-	-
		TOMM22*	Mitochondrial import receptor subunit TOM22 homolog	Q9NS69	-	-	53,531	-	-
	Midbody	RACGAP1	Rac GTPase-activating protein 1	Q9H0H5	18,088	85,367	1,689,821	2,708	38,445
		KIF23	Kinesin-like protein KIF23	Q02241	11,504	38,477	1,062,271	91	14,435
		CEP55	Centrosomal protein of 55 kDa	Q53EZ4	9,834	12,795	97,651	-	-

EV class	Category	Gene name	Protein description	Protein accession (UniProt)	Normalized LFQ intensity <sup>a</sup>				
		AURKB	Aurora kinase B (Fragment)	J3KT86	-	4,807	52,033	-	-
		KIF4A	Chromosome associated kinesin KIF4A	Q95239	-	793	40,044	2,015	2,394
		PLK1	Polo-like kinase 1	P53350	429	3,604	21,875	2,589	2,893
		KIF2A*	Kinesin-like protein KIF2A	O00139	-	-	11,375	-	-

*a* = LFQ (label free precursor intensity) normalized with protein length

\* = uniquely identified proteins

- = undetected in samples

**Table 2 Cancer progression-associated proteins in sEVs/Exos, lEVs/MPs and sMB-Rs secreted from SW480 and SW620 cells**

Category	Genename	Protein description
Genetic instability	PARP1	Poly [ADP-ribose] polymerase 1
	HDAC2*	Histone deacetylase 2
	MSH2#	DNA mismatch repair protein
	MSH6#*	DNA mismatch repair protein Msh6
Receptor/transportor	CD44	CD44 antigen
	EGFR	Epidermal growth factor receptor
	FAS	Tumor necrosis factor receptor superfamily member 6
	MET	Hepatocyte growth factor receptor
	CXCR4*	C-X-C chemokine receptor type 4
	FGFR4	Fibroblast growth factor receptor 4
	TGFBR2#*	TGF-beta receptor type-2
	AXL*	Tyrosine-protein kinase receptor
	ABCB1*	Multidrug resistance protein 1
Signal transduction	CLDN7	Claudin (Fragment)
	LAMTOR3	Ragulator complex protein LAMTOR3
	PRKACA	cAMP-dependent protein kinase catalytic subunit alpha
	MACC1	Metastasis-associated in colon cancer protein 1
	PLD1*	Phospholipase D1
	RICTOR	Rapamycin-insensitive companion of mTOR
CTNBL1*	Beta-catenin-like protein 1	

Category	Genename	Protein description
Growth factor	STAT1*	Signal transducer and activator of transcription 1-alpha/beta
	GYS1*	Glycogen synthase
	PLD2*	Phospholipase D2 (Fragment)
	SMAD5*	Mothers against decapentaplegic homolog 5
	KRAS#	GTPase Kras
	CTNNB1#	Beta-catenin
	SRC#	Proto-oncogene tyrosine-protein kinase Src
	PTPRK#	Receptor-type tyrosine-protein phosphatase kappa
	EIF3E#	Eukaryotic translation initiation factor 3 subunit E
	USP9X#	Ubiquitin-specific protease 9
	AKT1#	RAC-alpha serine/threonine-protein kinase (EC 2.7.11.1) (Protein kinase B)
	ARHGAP5#	Rho GTPase-activating protein 5 (Rho-type GTPase-activating protein GEF)
	B2M#	Beta-2-microglobulin [Cleaved into: Beta-2-microglobulin form pI 12 and 14]
	TP53RK*	TP53-regulating kinase
	MDK	Midkine (Fragment)
ECM remodeling enzyme/ECM protein	GDF15*	Growth/differentiation factor 15
	TGM2	Transglutaminase-2
	MMP14	Matrix metalloproteinase-14
	ADAM15*	Disintegrin and metalloproteinase domain-containing protein 15
	TNC*	Tenascin

$a$  = LFQ (label free precursor intensity) normalized with protein length

\* = uniquely identified proteins

# = commonly identified colorectal cancer-related proteins with COSMIC cancer database

- = undetected in samples

#### Author Contributions

W.S. and R.J.S. conceived and designed the study. W.S., R.X. and D.W.G. conducted the experiments. W.S., R.X. and R.J.S. analyzed and interpreted the data. R.J.S. and D.W.G. supervised the project. W.S. and R.J.S. wrote the first draft of the manuscript.

#### Notes

The authors declare no competing financial interest.

#### REFERENCES

- [1] Xu, R., Rai, A., Chen, M., Suwakulsiri, W., *et al.* , Extracellular vesicles in cancer - implications for future improvements in cancer care. *Nat Rev Clin Oncol* 2018, *15* , 617-638.
- [2] Skop, A. R., Liu, H., Yates, J., Meyer, B. J., Heald, R., Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. *Science* 2004, *305* , 61-66.
- [3] Colombo, M., Raposo, G., Thery, C., Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. *Annu Rev Cell Dev Biol* 2014, *30* , 255-289.
- [4] van Niel, G., Carter, D. R. F., Clayton, A., Lambert, D. W., *et al.* , Challenges and directions in studying cell-cell communication by extracellular vesicles. *Nat Rev Mol Cell Biol* 2022, *23* , 369-382.
- [5] Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., *et al.* , Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* 2018, *362* .

- [6] Zhou, B., Xu, K., Zheng, X., Chen, T., *et al.* , Application of exosomes as liquid biopsy in clinical diagnosis. *Signal Transduct Target Ther* 2020, *5* , 144.
- [7] Chen, S. L., Huang, Q. S., Huang, Y. H., Yang, X., *et al.* , GYS1 induces glycogen accumulation and promotes tumor progression via the NF-kappaB pathway in Clear Cell Renal Carcinoma. *Theranostics* 2020, *10* , 9186-9199.
- [8] Yu, W., Hurley, J., Roberts, D., Chakraborty, S. K., *et al.* , Exosome-based liquid biopsies in cancer: opportunities and challenges. *Ann Oncol* 2021, *32* , 466-477.
- [9] Armand, E. J., Li, J., Xie, F., Luo, C., Mukamel, E. A., Single-Cell Sequencing of Brain Cell Transcriptomes and Epigenomes. *Neuron* 2021, *109* , 11-26.
- [10] Xu, R., Greening, D. W., Zhu, H. J., Takahashi, N., Simpson, R. J., Extracellular vesicle isolation and characterization: toward clinical application. *J Clin Invest* 2016, *126* , 1152-1162.
- [11] Jeppesen, D. K., Fenix, A. M., Franklin, J. L., Higginbotham, J. N., *et al.* , Reassessment of Exosome Composition. *Cell* 2019, *177* , 428-445 e418.
- [12] Martinez-Greene, J. A., Hernandez-Ortega, K., Quiroz-Baez, R., Resendis-Antonio, O., *et al.* , Quantitative proteomic analysis of extracellular vesicle subgroups isolated by an optimized method combining polymer-based precipitation and size exclusion chromatography. *J Extracell Vesicles* 2021, *10* , e12087.
- [13] Al-Nedawi, K., Meehan, B., Micallef, J., Lhotak, V., *et al.* , Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat Cell Biol* 2008, *10* , 619-624.
- [14] Rai, A., Greening, D. W., Xu, R., Chen, M., *et al.* , Secreted midbody remnants are a class of extracellular vesicles molecularly distinct from exosomes and microparticles. *Commun Biol* 2021, *4* , 400.
- [15] van Niel, G., D'Angelo, G., Raposo, G., Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol* 2018, *19* , 213-228.
- [16] Willms, E., Johansson, H. J., Mager, I., Lee, Y., *et al.* , Cells release subpopulations of exosomes with distinct molecular and biological properties. *Sci Rep* 2016, *6* , 22519.
- [17] Jiao, H., Jiang, D., Hu, X., Du, W., *et al.* , Mitocytosis, a migrasome-mediated mitochondrial quality-control process. *Cell* 2021, *184* , 2896-2910 e2813.
- [18] Thery, C., Witwer, K. W., Aikawa, E., Alcaraz, M. J., *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* 2018, *7* , 1535750.
- [19] Willms, E., Cabanas, C., Mager, I., Wood, M. J. A., Vader, P., Extracellular Vesicle Heterogeneity: Subpopulations, Isolation Techniques, and Diverse Functions in Cancer Progression. *Front Immunol* 2018, *9* , 738.
- [20] Cocucci, E., Meldolesi, J., Ectosomes and exosomes: shedding the confusion between extracellular vesicles. *Trends in cell biology* 2015, *25* , 364-372.
- [21] Hurley, J. H., ESCRTs are everywhere. *EMBO J* 2015, *34* , 2398-2407.
- [22] Greening, D. W., Simpson, R. J., Understanding extracellular vesicle diversity - current status. *Expert Rev Proteomics* 2018, *15* , 887-910.
- [23] Zhao, Y., Xu, J., Synovial fluid-derived exosomal lncRNA PCGEM1 as biomarker for the different stages of osteoarthritis. *Int Orthop* 2018, *42* , 2865-2872.
- [24] Peterman, E., Gibieza, P., Schafer, J., Skeberdis, V. A., *et al.* , The post-abscission midbody is an intracellular signaling organelle that regulates cell proliferation. *Nat Commun* 2019, *10* , 3181.

- [25] Dionne, L. K., Wang, X. J., Prekeris, R., Midbody: from cellular junk to regulator of cell polarity and cell fate. *Curr Opin Cell Biol* 2015, *35* , 51-58.
- [26] Peterman, E., Prekeris, R., Understanding post-mitotic roles of the midbody during cell differentiation and polarization. *Methods Cell Biol* 2017, *137* , 173-186.
- [27] 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* 2015, *87* , 11-25.
- [28] Tauro, B. J., Greening, D. W., Mathias, R. A., Ji, H., *et al.* , Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. *Methods* 2012, *56* , 293-304.
- [29] Tauro, B. J., Greening, D. W., Mathias, R. A., Mathivanan, S., *et al.* , Two distinct populations of exosomes are released from LIM1863 colon carcinoma cell-derived organoids. *Mol Cell Proteomics* 2013, *12* , 587-598.
- [30] Suwakulsiri, W., Rai, A., Xu, R., Chen, M., *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* 2019, *1867* , 140171.
- [31] Martinez-Bartolome, S., Binz, P. A., Albar, J. P., The Minimal Information about a Proteomics Experiment (MIAPE) from the Proteomics Standards Initiative. *Methods Mol Biol* 2014, *1072* , 765-780.
- [32] Gorshkov, V., Verano-Braga, T., Kjeldsen, F., SuperQuant: A Data Processing Approach to Increase Quantitative Proteome Coverage. *Anal Chem* 2015, *87* , 6319-6327.
- [33] Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008, *26* , 1367-1372.
- [34] Gopal, S. K., Greening, D. W., Mathias, R. A., Ji, H., *et al.* , YBX1/YB-1 induces partial EMT and tumourigenicity through secretion of angiogenic factors into the extracellular microenvironment. *Oncotarget* 2015, *6* , 13718-13730.
- [35] Robinson, M. D., McCarthy, D. J., Smyth, G. K., edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, *26* , 139-140.
- [36] Robinson, M. D., Oshlack, A., A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010, *11* , R25.
- [37] Benjamini, Y., Hochberg, Y., Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 1995, *57* , 289-300.
- [38] Yu, G., Wang, L. G., Han, Y., He, Q. Y., clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012, *16* , 284-287.
- [39] Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P., Minghim, R., InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 2015, *16* , 169.
- [40] Luo, W., Brouwer, C., Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 2013, *29* , 1830-1831.
- [41] Glotzer, M., The 3Ms of central spindle assembly: microtubules, motors and MAPs. *Nat Rev Mol Cell Biol* 2009, *10* , 9-20.
- [42] Ji, H., Greening, D. W., Barnes, T. W., Lim, J. W., *et al.* , Proteome profiling of exosomes derived from human primary and metastatic colorectal cancer cells reveal differential expression of key metastatic factors and signal transduction components. *Proteomics* 2013, *13* , 1672-1686.

- [43] Hewitt, R. E., McMarlin, A., Kleiner, D., Wersto, R., *et al.* , Validation of a model of colon cancer progression. *J Pathol*2000, *192* , 446-454.
- [44] Kowal, J., Arras, G., Colombo, M., Jouve, M., *et al.* , Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc Natl Acad Sci U S A* 2016, *113* , E968-977.
- [45] Doyle, L. M., Wang, M. Z., Overview of Extracellular Vesicles, Their Origin, Composition, Purpose, and Methods for Exosome Isolation and Analysis. *Cells* 2019, *8* .
- [46] Hessvik, N. P., Llorente, A., Current knowledge on exosome biogenesis and release. *Cell Mol Life Sci* 2018, *75* , 193-208.
- [47] Nabhan, J. F., Hu, R., Oh, R. S., Cohen, S. N., Lu, Q., Formation and release of arrestin domain-containing protein 1-mediated microvesicles (ARMMs) at plasma membrane by recruitment of TSG101 protein. *Proc Natl Acad Sci U S A* 2012, *109* , 4146-4151.
- [48] Schlienger, S., Campbell, S., Claing, A., ARF1 regulates the Rho/MLC pathway to control EGF-dependent breast cancer cell invasion. *Mol Biol Cell* 2014, *25* , 17-29.
- [49] Muralidharan-Chari, V., Clancy, J., Plou, C., Romao, M., *et al.* , ARF6-regulated shedding of tumor cell-derived plasma membrane microvesicles. *Curr Biol* 2009, *19* , 1875-1885.
- [50] Li, B., Antonyak, M. A., Zhang, J., Cerione, R. A., RhoA triggers a specific signaling pathway that generates transforming microvesicles in cancer cells. *Oncogene* 2012, *31* , 4740-4749.
- [51] Wheeler, J. R., Jain, S., Khong, A., Parker, R., Isolation of yeast and mammalian stress granule cores. *Methods* 2017, *126* , 12-17.
- [52] Djafarzadeh, S., Jakob, S. M., Isolation of Intact Mitochondria from Skeletal Muscle by Differential Centrifugation for High-resolution Respirometry Measurements. *J Vis Exp* 2017.
- [53] D'Souza, A., Burch, A., Dave, K. M., Sreeram, A., *et al.* , Microvesicles transfer mitochondria and increase mitochondrial function in brain endothelial cells. *J Control Release* 2021, *338* , 505-526.
- [54] Melki, I., Allaey, I., Tessandier, N., Levesque, T., *et al.* , Platelets release mitochondrial antigens in systemic lupus erythematosus. *Sci Transl Med* 2021, *13* .
- [55] D'Acunzo, P., Perez-Gonzalez, R., Kim, Y., Hargash, T., *et al.* , Mitovesicles are a novel population of extracellular vesicles of mitochondrial origin altered in Down syndrome. *Sci Adv* 2021, *7* .
- [56] Capalbo, L., Bassi, Z. I., Geymonat, M., Todesca, S., *et al.* , The midbody interactome reveals unexpected roles for PP1 phosphatases in cytokinesis. *Nature communications* 2019, *10* , 1-17.
- [57] Capalbo, L., Bassi, Z. I., Geymonat, M., Todesca, S., *et al.* , The midbody interactome reveals unexpected roles for PP1 phosphatases in cytokinesis. *Nat Commun* 2019, *10* , 4513.
- [58] Huang, Z., Ma, L., Wang, Y., Pan, Z., *et al.* , MiCroKiTS 4.0: a database of midbody, centrosome, kinetochore, telomere and spindle. *Nucleic Acids Res* 2015, *43* , D328-334.
- [59] Addi, C., Presle, A., Frémont, S., Cuvelier, F., *et al.* , The Flemmingsome reveals an ESCRT-to-membrane coupling via ALIX/syntenin/syndecan-4 required for completion of cytokinesis. *Nature Communications* 2020, *11* , 1-15.
- [60] Hough, K. P., Trevor, J. L., Strenkowski, J. G., Wang, Y., *et al.* , Exosomal transfer of mitochondria from airway myeloid-derived regulatory cells to T cells. *Redox Biol* 2018, *18* , 54-64.
- [61] Statello, L., Maugeri, M., Garre, E., Nawaz, M., *et al.* , Identification of RNA-binding proteins in exosomes capable of interacting with different types of RNA: RBP-facilitated transport of RNAs into exosomes. *PLoS One* 2018, *13* , e0195969.

- [62] Ceresa, B. P., Schmid, S. L., Regulation of signal transduction by endocytosis. *Curr Opin Cell Biol* 2000, *12* , 204-210.
- [63] Keller, S., Sanderson, M. P., Stoeck, A., Altevogt, P., Exosomes: from biogenesis and secretion to biological function. *Immunol Lett* 2006, *107* , 102-108.
- [64] Nalawansha, D. A., Zhang, Y., Herath, K., Pflum, M. K. H., HDAC1 Substrate Profiling Using Proteomics-Based Substrate Trapping. *ACS Chem Biol* 2018, *13* , 3315-3324.
- [65] Li, S., Shi, B., Liu, X., An, H. X., Acetylation and Deacetylation of DNA Repair Proteins in Cancers. *Front Oncol* 2020, *10* , 573502.
- [66] Wang, S., Darini, C., Desaubry, L., Koromilas, A. E., STAT1 Promotes KRAS Colon Tumor Growth and Susceptibility to Pharmacological Inhibition of Translation Initiation Factor eIF4A. *Mol Cancer Ther* 2016, *15* , 3055-3063.
- [67] Antonyak, M. A., Li, B., Boroughs, L. K., Johnson, J. L., *et al.* , Cancer cell-derived microvesicles induce transformation by transferring tissue transglutaminase and fibronectin to recipient cells. *Proc Natl Acad Sci U S A* 2011, *108* , 4852-4857.
- [68] Filippou, P. S., Karagiannis, G. S., Constantinidou, A., Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene* 2020, *39* , 2040-2054.
- [69] Berthon, A. S., Szarek, E., Stratakis, C. A., PRKACA: the catalytic subunit of protein kinase A and adrenocortical tumors. *Front Cell Dev Biol* 2015, *3* , 26.
- [70] Zhu, G., Pei, L., Xia, H., Tang, Q., Bi, F., Role of oncogenic KRAS in the prognosis, diagnosis and treatment of colorectal cancer. *Mol Cancer* 2021, *20* , 143.
- [71] Disoma, C., Zhou, Y., Li, S., Peng, J., Xia, Z., Wnt/beta-catenin signaling in colorectal cancer: Is therapeutic targeting even possible? *Biochimie* 2022, *195* , 39-53.
- [72] Demory Beckler, M., Higginbotham, J. N., Franklin, J. L., Ham, A. J., *et al.* , Proteomic analysis of exosomes from mutant KRAS colon cancer cells identifies intercellular transfer of mutant KRAS. *Mol Cell Proteomics* 2013, *12* , 343-355.
- [73] Edelbrock, M. A., Kaliyaperumal, S., Williams, K. J., Structural, molecular and cellular functions of MSH2 and MSH6 during DNA mismatch repair, damage signaling and other noncanonical activities. *Mutat Res* 2013, *743-744* , 53-66.
- [74] Bonis, P. A., Trikalinos, T. A., Chung, M., Chew, P., *et al.* , Hereditary nonpolyposis colorectal cancer: diagnostic strategies and their implications. *Evid Rep Technol Assess (Full Rep)* 2007, 1-180.

#### Hosted file

Singh\_Proteomics\_Figures\_10\_04\_23.pptx available at <https://authorea.com/users/614198/articles/641154-comparative-proteomic-analysis-of-three-major-extracellular-classes-secreted-from-human-adenocarcinoma-and-metastatic-colorectal-cancer-cells-exosomes-microparticles-and-shed-midbody-remnants>