

# Combination of size-exclusion chromatography and ultracentrifugation improves the proteomic profiling of plasma derived small extracellular vesicles

**Rui Wei**

Capital Medical University Affiliated Beijing Friendship Hospital

**Libo Zhao**

Echo Biotech Co., Ltd

**Guanyi Kong**

Echo Biotech Co., Ltd

**Xiang Liu**

Echo Biotech Co., Ltd

**Shengtao Zhu**

Capital Medical University Affiliated Beijing Friendship Hospital

**Shutian Zhang**

Capital Medical University Affiliated Beijing Friendship Hospital

**Li Min** (✉ [minli@ccmu.edu.cn](mailto:minli@ccmu.edu.cn))

Capital Medical University Affiliated Beijing Friendship Hospital

---

## Research

**Keywords:** Extracellular vesicles; Size-exclusion chromatography; Ultracentrifugation; Proteomics

**Posted Date:** April 9th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-21362/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on June 23rd, 2020. See the published version at <https://doi.org/10.1186/s12575-020-00125-5>.

# Abstract

**Background:** Circulating small extracellular vesicles (sEVs) and its associated proteins are of great interest in the early detection of many diseases. However, there is no gold standard for plasma sEVs isolation, especially for proteomic profiling which could be largely affected by contamination such as lipoproteins. Previous studies suggested combinations of different sEVs isolation methods could improve the purity of the isolated fractions. Nevertheless, there is no systematic evaluation of size-exclusion chromatography (SEC), ultracentrifugation (UC) and their combination in a proteomic perspective.

**Results:** Here we exhibited that SEC+UC showed comparable recovery of sEVs with higher purity in contrast to single-step UC or SEC isolation. In our proteomic analysis, there are 992 protein species identified in the sEVs fractions isolated by SEC+UC, much more than the sEVs fractions isolated by UC (453) or SEC (682) alone. As compared to Vesiclepedia and Exocarta databases, SEC+UC kept 584 previously identified sEV-associated proteins and 360 other proteins. Furthermore, detailed analysis suggested that sEV-associated proteins, such as CD9, CD81 and ITGB1, showed the better protein rank in SEC+UC group than UC group and SEC group. Lipoproteins, the most common contamination in sEVs proteomic analysis, along with other free-floating proteins in the plasma, were largely removed in SEC+UC.

**Conclusions:** We suggested that combining SEC with UC could significantly improve the performance of mass-spectrum (MS)-based proteomic profiling in analyzing plasma-derived sEVs.

## Background

Liquid biopsy is one of the most widely used approaches for non-invasive clinical cancer diagnosis via fast acquiring a minimal body fluid <sup>1</sup>. The human blood sample contains various specific biomarkers, like bioactive lipids, cell-free DNA, messenger RNA and non-coding RNA, soluble proteins, small extracellular vesicles (sEVs) <sup>2</sup>. Among all liquid biopsy targets, sEVs attract many attentions during the past decades, which are characterized as nano-sized lipid bilayer vesicles (40–200 nm in diameter) stuffed with RNAs, proteins, and lipids, regardless of their origination <sup>3,4</sup>.

Most biomarker-oriented studies focused on the sEV-associated nucleic acids, which could be amplified *in vitro* and easy to detect <sup>5–7</sup>. However, sEV-associated protein biomarkers are also irreplaceable due to their potential integrability with isolation procedure such as nano-flow cytometry <sup>8,9</sup>. Nevertheless, the discovery of sEV-associated protein biomarkers is largely restricted by recent sEVs isolation procedures, especially for plasma samples, which contain a dominating pool of impurities such as lipoprotein particles <sup>10,11</sup>. High levels of plasma lipoproteins ( $\approx 10^{16}$  lipoproteins/mL plasma) are secreted from the liver and intestine, then released and matured in circulation. They are classified into HDL, LDL, IDL, VLDL, and chylomicrons (CM), according to their mass and density <sup>12</sup>.

Ultracentrifugation (UC) could separate different particles based on a compounding effect of size and density, which is considered as a classic method in sEVs separation<sup>13</sup>. UC is widely adopted in sEV-associated RNA biomarker discovery studies, but it alone is unable to remove abundant lipoprotein particles, such as HDL, which could affect the downstream mass-spectrum (MS) analysis of proteins<sup>14</sup>.

Size-exclusion chromatography (SEC), also known as “gel filtration”, as an old size-based separation tool, is more and more widely used in the field of EV studies after its potential in separation of sEVs from HDL and proteins being fully characterized in 2014<sup>15</sup>. However, SEC itself also has a limitation. It becomes clear that SEC-based isolation of sEVs can't remove CM and VLDL, which overlap in size with sEVs.

Theoretically, combining SEC purification with UC enrichment could achieve a considerable isolation efficiency of sEVs with higher purity than a single-step isolation process by removing a large number of HDL, VLDL, and CM. As compared to other combined sEVs isolation pipelines, in our SEC + UC strategy, the biofluid viscosity of samples shows a significant decrease after loaded on the SEC column, so there is no need to dilute samples with 7-fold volume PBS in the next UC process. Thus, SEC + UC helps to simplify the procedures and operators could deal with a larger volume of plasma at one time. To fully characterize the combined strategy in sEVs isolation, here we compared SEC + UC with single-step UC/SEC isolation procedures in a proteomic perspective.

## Results

### SEC + UC successfully isolated sEVs from plasma

The sEVs fractions were isolated directly by UC, SEC and SEC + UC methods (Fig. 1), and characterized according to MISEV2018 guideline<sup>16</sup>. Transmission electron microscope (TEM) images showed that the sEVs isolated from human plasma by three different isolation pipelines had intact membrane structures and similar morphology (Fig. 2A-C). We applied nanoparticle tracking analysis (NTA) to measure the mean diameters of isolated fractions, and most particles were at 60–100 nm (Fig. 2D-F). Coomassie blue staining showed the total protein level of SEC + UC, UC and SEC fractions visually (Fig. 2G), and verified SEC + UC had the minimum protein content. Western blotting confirmed several sEVs positive markers, including CD9 (60232-1, 1:1000, Proteintech), CD81 (66866-1, 1:1000, Proteintech), and HSP90 (60318-1, 1:1000, Proteintech) (Fig. 2H). Under the premise of the same protein quantity, UC preserved the highest level of CD81, as compared to SEC and SEC + UC. Meanwhile, the CD9 and HSP90 levels of the SEC + UC sEVs fraction were similar to that of UC and much higher than that of SEC. Generally, our SEC + UC isolation pipeline kept all three sEV-associated biomarkers, revealing the successful sEVs recovery of SEC + UC isolation assay.

### SEC + UC showed comparable recovery of sEVs with higher purity in contrast to a single-step UC or SEC isolation

According to the NTA quantification results (Fig. 3A), the SEC + UC procedure isolated  $3.03 \times 10^{10}$  particles from 1 mL plasma, slightly lower than UC ( $5.87 \times 10^{10}$ /mL plasma), while SEC assay isolated the largest number of particles at  $5.97 \times 10^{11}$ /mL plasma. BCA protein quantification was also performed, the results showed that the sEVs fraction isolated by SEC + UC (6.15  $\mu$ g/mL plasma) had a significantly lower protein amount, as compared to UC (209.8  $\mu$ g/mL plasma) and SEC (65.6  $\mu$ g/mL plasma, Fig. 3B), which suggested a successful removal of lipoproteins and other contaminant particles. In addition, protein abundance per particle was also calculated in Fig. 3C. Considering lipoprotein particles have a higher protein proportion than sEVs, our results showed that SEC (0.11 fg protein/particles) and SEC + UC (0.20 fg protein/particles) isolated particles had better purity of sEVs than UC (3.58 fg protein/particles).

Common putative contaminants, such as Albumin (16475-1, 1:5000, Proteintech) and apolipoprotein A1 (Apo-A1) (66206-1, 1:1000, Proteintech), and one sEVs negative biomarker, Calnexin (10427-2, 1:500, Proteintech) were all detected in the sEVs fractions isolated by UC, SEC, and SEC + UC, along with the plasma input. Albumin is the most abundant protein in human blood plasma, and considered as the most commonly polluted indicator. Apo-A1 is a major component of HDL and CM. Western blotting showed lipoprotein particles were co-isolated with sEVs (Fig. 3D), and one-step SEC and SEC + UC pipelines had fewer contaminants of lipoproteins, as compared to UC. In addition, all three pipelines showed the negative level of Calnexin. In summary, SEC + UC combination assay could acquire a higher yield of sEVs with fewer contaminants, as compared to a single-step UC or SEC.

## **SEC + UC kept more protein species detectable by MS than a single-step UC or SEC isolation**

We continued by analyzing proteins with MS. In proteomic analysis, there are 992 protein species identified in the sEVs fractions isolated by SEC + UC, much higher than the sEVs fractions isolated by UC (453) or SEC (682) alone (Fig. 4A). The heatmap shows all protein species of three isolation groups had a variation expression level (Fig. 4B). All protein species were displayed according to SEC + UC protein level from high to low. Generally, 345 protein species were identified in SEC + UC but not in SEC, and 591 protein species were identified in SEC + UC but not in UC. Those results suggested that SEC + UC kept more protein species detectable by MS than single-step UC or SEC isolation, especially for the proteins of low-abundance.

Then, we investigated what roles these newly identified low-abundance proteins mainly played. Gene Ontology (GO) / Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment of the 345 identified sEVs proteins in SEC + UC but not in SEC was performed to explore their functions (Fig. 4C). The enriched GO terms were analysed into three categories: biological process (BP), cellular component (CC), molecular function (MF). For BP, GO terms include mRNA splicing, via spliceosome (GO:0000398), RNA export from nucleus (GO:0006405), osteoblast differentiation (GO:0001649). For CC, GO terms include extracellular exosome (GO:0070062), membrane (GO:0016020), nucleoplasm (GO:0005654). For MF, GO

terms include poly(A) RNA binding (GO:0044822), protein binding (GO:0005515), RNA binding (GO:0003723). The top 10 enriched GO terms were shown in Fig. 4C, left. Additionally, the KEGG pathway analysis indicated 21 significantly enriched pathways, including spliceosome (hsa03040), dopaminergic synapse (hsa04728), platelet activation (hsa04611) (Fig. 4C, right).

We also performed the GO/KEGG enrichment of 591 identified proteins in SEC + UC but not in UC as well. The enriched GO terms were slightly different. For BP, GO terms include cell-cell adhesion (GO:0098609), viral transcription (GO:0019083), nuclear-transcribed mRNA catabolic process, nonsense-mediated decay (GO:0000184). For CC, GO terms include extracellular exosome (GO:0070062), membrane (GO:0016020), focal adhesion (GO:0005925). For MF, GO terms include poly(A) RNA binding (GO:0044822), protein binding (GO:0005515), GTPase activity (GO:0003924). The top 10 enriched GO terms were shown in Fig. 4D, left. Additionally, the KEGG pathway analysis indicated 46 significantly enriched pathways, including platelet activation (hsa04611), regulation of actin cytoskeleton (hsa04810), spliceosome (hsa03040) (Fig. 4D, right).

We also compared our results with the current public EV datasets, Vesiclepedia and Exocarta (Fig. 4E-G). There were 584 species of previously identified sEV-associated proteins (the intersecting area in Venn diagrams) identified by SEC + UC, much more than those of UC (180) and SEC (366) alone. Additionally, there were also more proteins which not reported before in the SEC + UC group (360). Thus, we demonstrated that the SEC + UC group maintained more protein species to be detectable in MS, including the most identified proteins and some unidentified low-abundant proteins, as compared to UC and SEC.

## **SEC + UC kept EV-associated proteins more detectable by MS and discarded more contaminants than a single-step UC or SEC isolation**

Then we ranked all detected proteins by their abundance and marked some common EV biomarkers in all three groups. Detailed analysis of MS results suggested that sEV-associated proteins showed higher rank in SEC + UC group than both UC group and SEC group evidently (Fig. 5A, C). Generally, CD9 and CD81 ranked highest in all three groups, suggested them as robust sEVs biomarkers. Flotillin-1 (Flot1) is considered as a highly specific exosome biomarker, which does not appear in other EVs. Here our results showed that Flot1 was only detected by MS in SEC + UC group, suggested a better detection ability for exosome components of our SEC + UC pipeline.

Lipoproteins (Apo-E, Apo-L1, Apo-A1, Apo-A2) and IgG antibody fragments (IgG L chain and IgG H chain) are the most common contaminations in plasma sEVs proteomic analysis. MS results analyzed that all contaminants in UC showed a higher abundance rank than SEC and SEC + UC, except Apo-E, which suggests UC can remove Apo-E efficiently. The most contaminations in SEC were Apo-E and Apo-L1, while SEC showed a lower abundance rank in Apo-A1, Apo-A2 and IgG H chain. In terms of SEC + UC, they showed a similar protein rank among Apo-E, Apo-L1, and Apo-A1. As compared to SEC, the combination

method showed a better removal effect on Apo-E and two IgG fractions (Fig. 5B, D). Additionally, isolated sEVs from all three pipelines did not contain Argonaute1, Argonaute2, GM130, PMP70, or Tamm-Horsfall protein, suggested that they were not sEVs associated, as previous reported<sup>17</sup>. Moreover, MS analysis also revealed that SEC + UC kept more histones, such as HMGN2, HIST1H1E, H1FX, H1F0, RBBP4, which might be derived from EVs released from cell death (eg. apoptosis, necrosis, NETosis).

## Discussion

Plasma sEVs have attracted much more attention as novel diagnostic biomarkers for cancer, inflammation, and other diseases, since the contained cargoes truly reflect the status of diseased individuals<sup>14,18</sup>. Currently, more studies are eager to investigate the potential function of proteins contained in plasma sEVs, especially some low abundant proteins<sup>19</sup>. However, plasma is a complex system and contains a large pool of plasma proteins, such as Albumin, Apo family, which brings great difficulties to the analysis of plasma sEVs derived proteins.

Different isolation methods could influence the purity and yield of plasma sEVs, thus there is still no gold method for isolating sEVs with no contaminants. Traditional UC assay was once the gold standard, but the limitations such as time-consuming, easily influenced by the operator's experience, and large abundance of lipoproteins-HDL, restrict it to be widely utilized in clinical analyzation<sup>20</sup>. The density of HDL is reported to be 1.06–1.21 g/ml, pretty closed to the density of sEVs (1.13–1.21 g/ml)<sup>21</sup>. Theoretically, the different particle sizes resulted a 100-fold faster sedimentation rate of sEVs (100 nm) as compared to HDL (10 nm). However, sEVs are still commonly co-isolated with a significant amount of HDL by UC isolation, for a far more abundant concentration of HDL ( $10^{16}$ /ml) as compared to sEVs ( $10^{10}$ /ml) in human plasma<sup>12</sup>. Unlike UC, SEC method separates sEVs and other particles purely by their diameter and size. The advantage of SEC is that it can separate particles and soluble small molecules with a diameter smaller or larger than sEVs (40–400 nm) well, such as HDL (5–12 nm), LDL (18–25 nm), and IDL (25–35 nm). Recent studies proved SEC had a promising potential because it allowed a purer sEVs preparation by removing nearly all HDL particles<sup>22</sup>. However, the molecular sieve effect of SEC filler sepharose CL-2B is incompetent to differentiate particles in the range of 50–100 nm, which showed a poor removal effect for CM and VLDL (30–80 nm). CM is almost the same size as sEVs, but the density of CM is much lower than water. The floating effect of CM, superimposed with the sedimentation effect of sEVs, could well separate CM from sEVs by UC. The density of VLDL is similar to water and could not subsided by UC. Thus, the UC procedure could also separate most VLDL from sEVs.

In this work, we compared the yield and purity of three sEVs isolated assays: one-step UC, one-step SEC and a combination of SEC + UC. By comparing the effects of different methods, we found SEC + UC was an optimized method. After most of the lipoproteins were filtered out by the SEC column, sEVs were then further isolated and enriched by UC. Our results showed that SEC + UC obtained a higher purity of sEVs. In addition, MS-based proteomics analysis revealed that sEVs isolated by the combination assay obtain large amounts of sEV-associated proteins, especially for the low-abundant proteins.

In the traditional UC assay, pellets acquired at the final step were washed by resuspension and re-collected by centrifugation, and fewer washing times could reduce the wasting, while the amount of contaminants also increases<sup>23,24</sup>. Therefore, to achieve a balance between sEVs yield and purity is essential. It is noteworthy that the total particles and total protein concentration of sEVs we acquired by SEC + UC were relatively less, while with strong expression of sEV-associated markers, such as CD9, CD81. These results demonstrated we isolated a higher sEVs yield with pretty purity. Moreover, MS results proved that SEC + UC isolated almost 1000 protein species, higher than traditional UC or one-step SEC, suggesting an improved performance in maintaining low abundant proteins. Previous studies used different isolation methods only identified 66–330 protein species associated with sEVs from human plasma by MS<sup>25,26</sup>. Thus, here we exhibited a considerable advantage of SEC + UC approach over previous reports in the sample preparing of MS analysis.

We further sorted out proteins only identified in SEC + UC, but not in SEC. Those genes were enriched in GO/KEGG terms of extracellular vesicular exosome localized proteins, spliceosomal proteins, splicing process and transcriptional mechanisms. Currently, several studies investigated the differential expression of splicing proteins in serum derived-sEVs and ascites, which could be potential biomarkers of early cancer<sup>19,27</sup>, and attractive targets for anticancer therapy<sup>28</sup>. We also sorted out proteins only identified in SEC + UC, but not in UC, and found that those proteins participate in more biological processes, including cell-cell adhesion, splicing, transcription, platelet activation, regulation of actin cytoskeleton and spliceosome. The sEVs of healthy individuals in the blood is mainly from erythrocyte and platelets, they selectively release proteins into vesicles to achieve cell-cell communication<sup>2</sup>. Thus, the present results of protein functions indicated sEVs isolated by SEC + UC acquired more kinds of low abundant proteins, and reflected transcription status of derived cells and blood microenvironment. In general, seeking low-abundant protein is mainly limited by the sensitivity of MS technology and protein degradability<sup>29</sup>. Our combined approach in plasma sEVs isolation not only provided a simple operating step for increasing sEVs yield, but also identified a few of low-abundant proteins, which need to be further studied.

Then we focused on common sEVs biomarkers, such as tetraspanins (CD9, CD63, and CD81), integrins (ITGB1), heat shock proteins (HSP90-AA1), Rabs (Rab27B), Alix and Flotillin-1 (FLOT1). SEC + UC acquired a higher abundance of all sEV-associated proteins, including CD9, CD81, etc. While, UC only identified two higher expression proteins, CD9 and CD81. FLOT1 was rare and could not detected in SEC group, which was considered as a specific sEVs marker in the previous study. Meantime, we analyzed the removal effect of lipoproteins (Apo-E, Apo-L1, Apo-A1 and Apo-A2) and IgG fractions (IgG L chain, IgG H chain). According to the MS results, UC kept nearly all above contaminants, while with less Apo-E. SEC was with high pollution of Apo-E, Apo-L1, and IgG L chain. Combined method of SEC + UC showed a better removal effect on Apo-E, Apo-A2, IgG fractions, as compared to SEC.

Previous MS studies suggested density-gradient UC may be a crucial step to further remove more contaminants, but a horizontal rotor with a speed more than 120,000 g and a solvent for suspending

sEVs pellets is needed, which makes the operation tedious. Cecilia isolated sEVs by UC + IDC (iodixanol density cushion) + SEC, a more complex procedure than ours, and only successfully identified about 1100 proteins, along with contaminants of lipoproteins. Here we used a two-step isolation procedure, combining SEC followed by UC, and identified almost 1000 proteins. The first step-SEC of our procedure largely decreased the bioliquid viscosity of plasma, which enable the collected fractions to be UC centrifugated directly without any dilution. Therefore, the SEC + UC pipeline can increase the total amount of plasma processed by a same centrifuge by 7 times, which make it more time-saving and labor-saving than single-step UC. Generally, the SEC + UC pipeline is probably the most cost-efficient sEVs separation method for MS proteomics profiling of plasma samples.

## Conclusions

Here we found that SEC + UC method could increase the total sEVs output and purity to make some low-abundant proteins detectable. MS analysis indicated sEVs isolated by SEC + UC contained not only previously identified proteins, but also some proteins associated with gene transcription, mRNA splicing, and platelet activation, which had not been detected by previous MS analysis of sEVs fractions. These newly identified sEV-associated proteins could provide a new pool of diagnostic biomarkers for many diseases. Therefore, our study demonstrated the combination of SEC and UC method could largely improve the proteomics profiling of plasma sEVs.

## Materials And Methods

The principal aim of this study was to improve the proteomics profiling of sEVs from human plasma, under sEVs acquired with higher yield and purity, helping detect and identify novel biomarkers. To compare the isolation efficacy of three methods (UC, SEC, and SEC+UC), NTA and MS were employed.

## Blood collection and sample handling

Blood was donated by three health individuals voluntarily. Written informed consent was obtained and this study was approved by the ethics committee of Beijing Friendship Hospital. Peripheral blood samples were collected in EDTA tubes following a regular venipuncture procedure. After centrifugation at 3,000 ×g for 15 min at 4 °C, the supernatant was transferred to new tubes and centrifuged at 3,000×g for 15 min at 4 °C again to minimise the platelets contamination.

## sEVs isolation by UC

The UC procedure was optimized according to the method previously described. Plasma samples were centrifugated at 3,000 ×g for 15 min. Then, the supernatant was diluted by 7-fold volume of phosphate-buffered saline (PBS), centrifuged at 13,000 ×g for 30 min, and processed through a 0.22 μm filter. The supernatant was ultracentrifuged using a P50A72-986 rotor (CP100NX; Hitachi, Brea, CA, USA) at 150,000

×g, 4 °C for 4 h to pellet the sEVs. The pellet was resuspended in PBS and centrifuged again at 150,000 ×g 4 °C for 2h. After PBS washing, the sEVs pellet was re-suspended in 100 µl PBS.

## sEVs isolation by SEC

1 mL fresh plasma samples were filtered through a 0.22 µm filter, then the supernatant was loaded into the column (Echo9103A-10ml; ECHO BIOTECH, China) which was prewashed with more than 20 mL sterile PBS in advance. After all samples were into the column and no fluid flowed out from the column bottom, PBS were used to eluate sEVs and other fractions. Each 500 µl of effluent represents one fraction. Then, 4~7 fractions were collected and added into a 100KD ultrafiltration tube. After centrifuged at 4,000 ×g for 15 min, the enriched sEVs were collected into a tube for experiments.

## sEVs isolation by SEC+UC

Plasma samples were filtered through a 0.22 µm filter and added to the SEC column, as we depicted before. Then 2mL collected fractions were centrifuged at 150,000 ×g, 4 °C for 4 h to further pellet the sEVs. The pellet was resuspended in PBS and centrifuged again 150,000 ×g, 4 °C for 2h. Finally, the supernatant was removed and resuspended in 100 µl PBS.

## Characterization of sEVs

The sEVs suspensions were characterized according to the MISEV2018 guideline<sup>16</sup>, the detailed procedures of NTA, TEM, Western blot analysis were performed according to our previous publication<sup>30</sup>.

## Protein isolation and quantification

The protein concentrations of UC, SEC and SEC+UC fractions were measured with Pierce™ BCA Protein Assay Kit (Product No. 23,225, Thermo Scientific, USA). 10 µl standard samples and sEVs enriched fractions of three methods were pipetted into 96-well plates, then added 200 µl BCA kit to each well. Then the plate was covered and incubated at 37 °C for 30min. The absorbance was set at 562 nm on the plate reader, and the standard curve was used to measure the protein concentrations of each isolated samples.

## Trypsin treatment and MS Analysis

Each protein sample was added 3 µL of 1 µg/µL trypsin and 500 µL of 100 mM TEAB buffer and then digested at 37 °C overnight. Equal volume of 1% formic acid was mixed with digested sample and centrifuged at 12000 g for 5 min at room temperature. The supernatant was slowly loaded to the C18 desalting column, washed with 1 mL of washing solution (0.1% formic acid, 4% acetonitrile) for three

times, and then eluted twice by 0.4 mL of elution buffer (0.1% formic acid, 75% acetonitrile). The eluents were combined and lyophilized. The lyophilized protein powder was dissolved in 10  $\mu$ L 0.1% formic acid in water (solvent A), and then injected into a home-made C18 Nano-Trap column (2 cm $\times$ 75  $\mu$ m, 3  $\mu$ m). Peptides were separated in a home-made analytical column (15 cm $\times$ 150  $\mu$ m, 1.9  $\mu$ m) with a mobile phase of 0.1% formic in 80% acetonitrile (solvent B). Sample elution was performed at a flow rate of 600 nL/min by increasing the solvent B concentration from 6 to 100% over 60 min. The separated peptides were analyzed by Q Exactive HF-X mass spectrometer (Thermo Fisher), with ion source of Nanospray Flex™ (ESI), spray voltage of 2.3 kV.

The raw data of MS detection were searched against UniProt database (<http://www.uniprot.org>). Carbamidomethyl was specified as fixed modifications. Oxidation of methionine (M) and acetylation of the N-terminus were specified as variable modifications. The identified protein contains at least 1 unique peptide with FDR no more than 0.01. GO analysis was conducted, and the databases of COG (Clusters of Orthologous Groups) and KEGG were used to annotate the protein family and pathway.

## Statistical analysis

Statistical tests were performed using R 3.5.1 ([www.r-project.org](http://www.r-project.org)). All tests were two-tailed and False Discovery Rate (FDR) was controlled for multiple comparisons.  $P < 0.05$  was considered significant. Packages plyr and reshape2 were used for data sorting and restructuring. VennDiagram, pheatmap, and ggplot2 were used for visualization of results.

## Abbreviations

**sEVs:** small extracellular vesicles

**SEC:** size-exclusion chromatography

**UC:** ultracentrifugation

**MS:** mass-spectrum

**TEM:** transmission electron microscope

**NTA:** nanoparticle tracking analysis

**Apo-A1:** apolipoprotein A1

**CM:** chylomicrons

**GO:** Gene Ontology

**KEGG:** Kyoto Encyclopedia of Genes and Genomes

**BP:** biological process

**CC:** cellular component

**MF:** molecular function

**Flot1:** flotillim-1

**IDC:** iodixanol density cushion

**PBS:** phosphate-buffered saline

**COG:** Clusters of Orthologous Groups

**FDR:** False Discovery Rate

## **Declarations**

### **Ethical Approval and Consent to participate**

This study was approved by the ethics committee of Beijing Friendship Hospital. The written informed consent was obtained from each participant.

### **Consent for Publication**

All authors have read and approved the final manuscript.

### **Availability of Data and Materials**

All data generated or analysed during this study are included in this published article.

### **Competing Interests**

The authors declare that they have no competing interests.

### **Funding**

This work was supported by grants from the Beijing Science and Technology Nova Program (Z191100001119128); Beijing Municipal Science and Technology Project (Z191100006619081); National Natural Science Foundation of China (81702314); Beijing Municipal Administration of Hospitals' Youth Programme (QML20180108); The Digestive Medical Coordinated Development Center of Beijing Municipal Administration of Hospitals (XXZ0201); Funding Program for Excellent Talents of Beijing (2017000021469G212).

### **Authors' contributions**

LM conceived and designed the study. RW, LZ, GK and XL performed all experiments. RW and LM draft the manuscript. S. Zhu and S. Zhang helped to revise the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

Not applicable.

## Author's information

*Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Disease, Beijing Digestive Disease Center, Beijing Key Laboratory for Precancerous Lesion of Digestive Disease, Beijing, 100050, P. R. China.*

## Rui Wei, Shengtao Zhu, Shutian Zhang, Li Min

*Echo Biotech Co., Ltd, Beijing, 100010, P. R. China.*

## Libo Zhao, Guanyi Kong, Xiang Liu

## ORCID

Li Min <https://orcid.org/0000-0001-9595-5536>

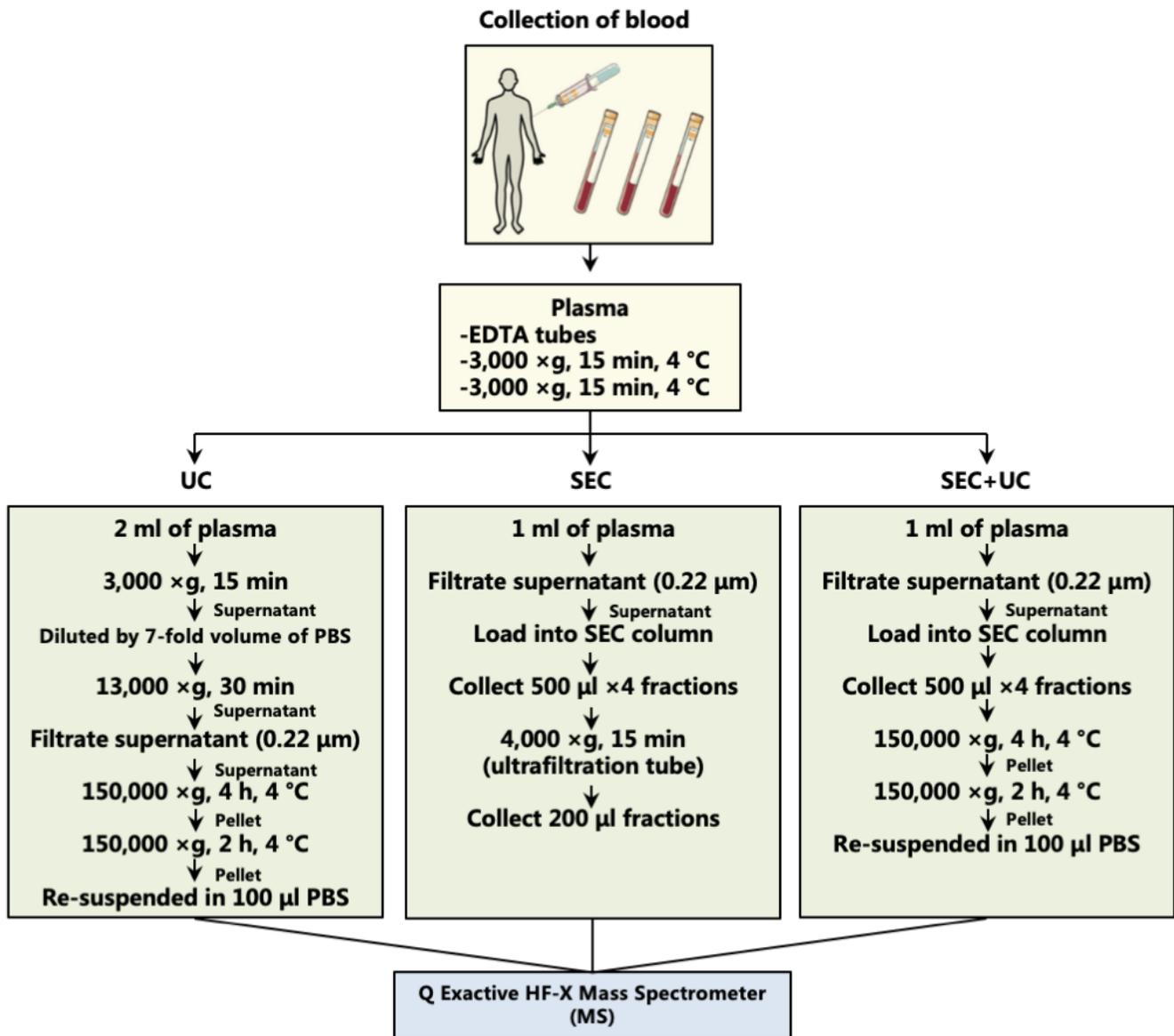
## References

1. Perakis, S. & Speicher, M. R. Emerging concepts in liquid biopsies. *Bmc Med.* **15**, 75 (2017).
2. Karimi, N. *et al.* Detailed analysis of the plasma extracellular vesicle proteome after separation from lipoproteins. *Cell. Mol. Life Sci. CMLS* **75**, 2873–2886 (2018).
3. Katsuda, T., Kosaka, N. & Ochiya, T. The roles of extracellular vesicles in cancer biology: toward the development of novel cancer biomarkers. *Proteomics* **14**, 412–425 (2014).
4. Biological properties of extracellular vesicles and their physiological functions. *J Extracell Vesicles* **4**, 27066 (2015).
5. Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).
6. Takahashi, R.-U., Prieto-Vila, M., Hironaka, A. & Ochiya, T. The role of extracellular vesicle microRNAs in cancer biology. *Clin. Chem. Lab. Med.* **55**, 648–656 (2017).
7. Salehi, M. & Sharifi, M. Exosomal miRNAs as novel cancer biomarkers: Challenges and opportunities. *J. Cell. Physiol.* **233**, 6370–6380 (2018).
8. Choi, D., Spinelli, C., Montermini, L. & Rak, J. Oncogenic Regulation of Extracellular Vesicle Proteome and Heterogeneity. *Proteomics* **19**, e1800169 (2019).

9. Tian, Y. *et al.* Quality and efficiency assessment of six extracellular vesicle isolation methods by nano-flow cytometry. *J. Extracell. Vesicles* **9**, 1697028 (2020).
10. Kalra, H. *et al.* Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. *Proteomics* **13**, 3354–3364 (2013).
11. Welton, J. L. *et al.* Proteomics analysis of vesicles isolated from plasma and urine of prostate cancer patients using a multiplex, aptamer-based protein array. *J. Extracell. Vesicles* **5**, 31209 (2016).
12. Simonsen Jens B. What Are We Looking At? Extracellular Vesicles, Lipoproteins, or Both? *Circ. Res.* **121**, 920–922 (2017).
13. Witwer, K. W. *et al.* Standardization of sample collection, isolation and analysis methods in extracellular vesicle research. *J. Extracell. Vesicles* **2**, (2013).
14. Takov, K., Yellon, D. M. & Davidson, S. M. Comparison of small extracellular vesicles isolated from plasma by ultracentrifugation or size-exclusion chromatography: yield, purity and functional potential. *J. Extracell. Vesicles* **8**, (2019).
15. Böing, A. N. *et al.* Single-step isolation of extracellular vesicles by size-exclusion chromatography. *J. Extracell. Vesicles* **3**, (2014).
16. Théry, 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).
17. Jeppesen, D. K. *et al.* Reassessment of Exosome Composition. *Cell* **177**, 428-445.e18 (2019).
18. Niel, G. V., D'Angelo, G. & Raposo, G. Shedding light on the cell biology of extracellular vesicles. *Nat. Rev. Mol. Cell Biol.* doi:10.1038/nrm.2017.125.
19. Khan, S. *et al.* Early diagnostic value of survivin and its alternative splice variants in breast cancer. *BMC Cancer* **14**, 176 (2014).
20. Monguió-Tortajada, M., Gálvez-Montón, C., Bayes-Genis, A., Roura, S. & Borràs, F. E. Extracellular vesicle isolation methods: rising impact of size-exclusion chromatography. *Cell. Mol. Life Sci. CMLS* **76**, 2369–2382 (2019).
21. Kreimer, S. *et al.* Mass-Spectrometry-Based Molecular Characterization of Extracellular Vesicles: Lipidomics and Proteomics. *J. Proteome Res.* **14**, 2367–2384 (2015).
22. Welton, J. L., Webber, J. P., Botos, L.-A., Jones, M. & Clayton, A. Ready-made chromatography columns for extracellular vesicle isolation from plasma. *J. Extracell. Vesicles* **4**, 27269 (2015).
23. Li, J., He, X., Deng, Y. & Yang, C. An Update on Isolation Methods for Proteomic Studies of Extracellular Vesicles in Biofluids. *Mol. Basel Switz.* **24**, (2019).
24. Langevin, S. M. *et al.* Balancing yield, purity and practicality: a modified differential ultracentrifugation protocol for efficient isolation of small extracellular vesicles from human serum. *RNA Biol.* **16**, 5–12 (2019).

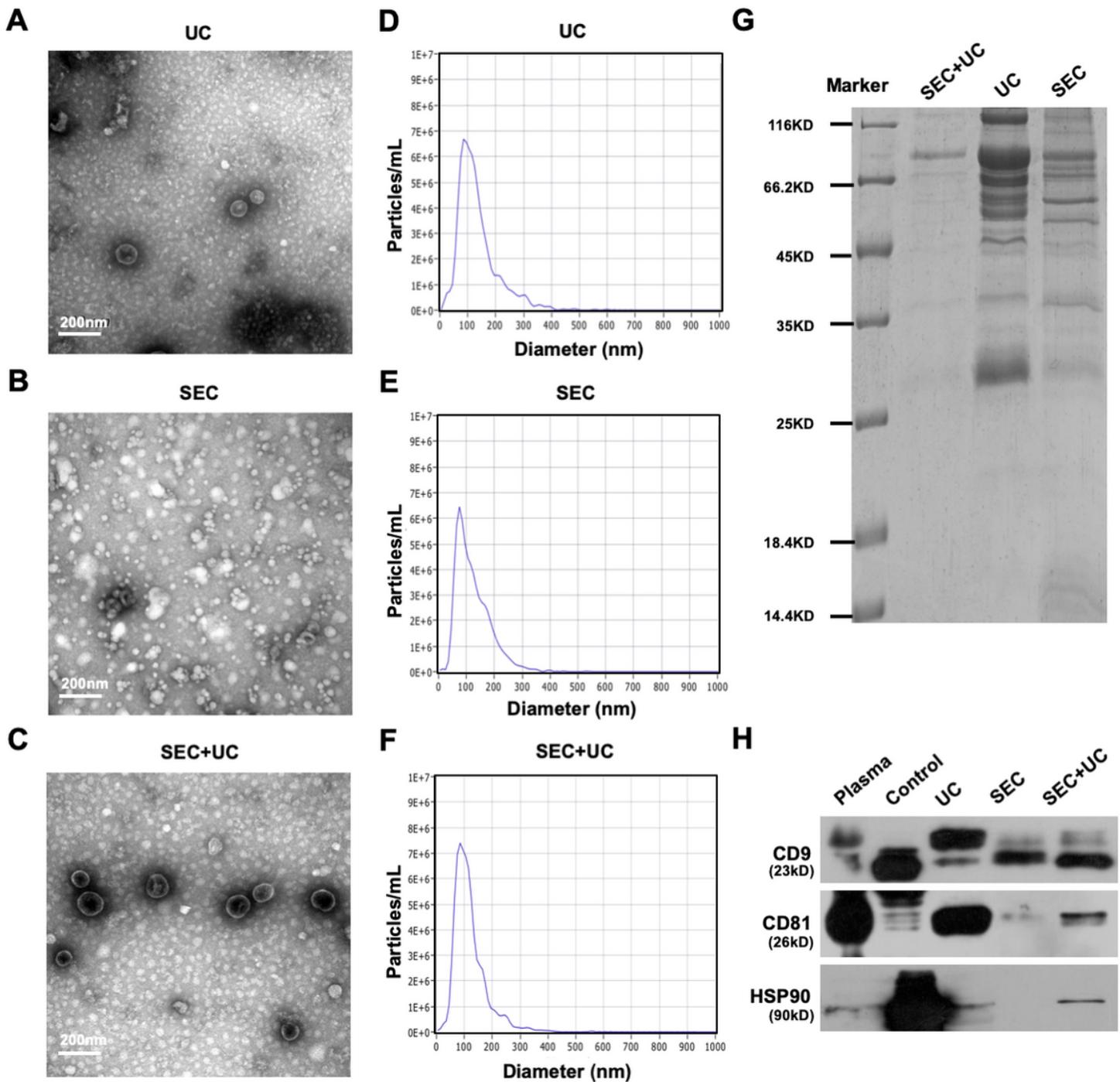
25. Looze, C. *et al.* Proteomic profiling of human plasma exosomes identifies PPARgamma as an exosome-associated protein. *Biochem. Biophys. Res. Commun.* **378**, 433–438 (2009).
26. de Menezes-Neto, A. *et al.* Size-exclusion chromatography as a stand-alone methodology identifies novel markers in mass spectrometry analyses of plasma-derived vesicles from healthy individuals. *J. Extracell. Vesicles* **4**, 27378 (2015).
27. Yu, D. *et al.* Exosomes in development, metastasis and drug resistance of breast cancer. *Cancer Sci.* **106**, 959–964 (2015).
28. Shender, V. O. *et al.* Proteome-metabolome profiling of ovarian cancer ascites reveals novel components involved in intercellular communication. *Mol. Cell. Proteomics MCP* **13**, 3558–3571 (2014).
29. Liotta, L. A. & Petricoin, E. F. Mass spectrometry-based protein biomarker discovery: solving the remaining challenges to reach the promise of clinical benefit. *Clin. Chem.* **56**, 1641–1642 (2010).
30. Min, L. *et al.* Evaluation of circulating small extracellular vesicles derived miRNAs as biomarkers of early colon cancer: a comparison with plasma total miRNAs. *J. Extracell. Vesicles* **8**, 1643670 (2019).

## Figures



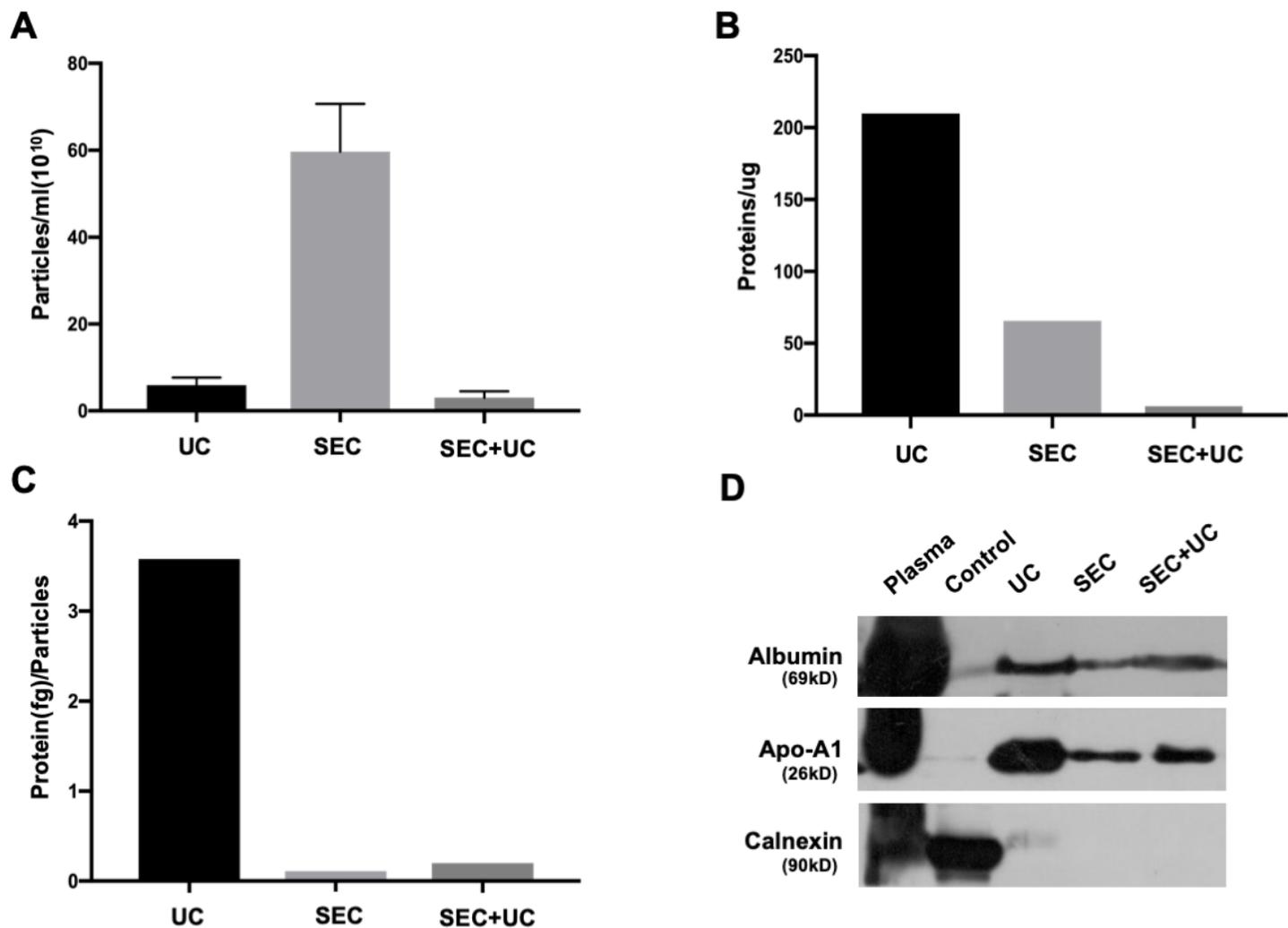
**Figure 1**

Schematic overview of the experiment pipeline. Blood was collected from three healthy individuals in EDTA tubes, and plasma were isolated. Several isolation approaches were used to separate sEVs from human plasma and to separate them from lipoproteins, plasma proteins. Mass spectrometer helped obtain the proteomic profiling. UC ultracentrifugation, SEC size-exclusion chromatography, MS mass spectrometer.



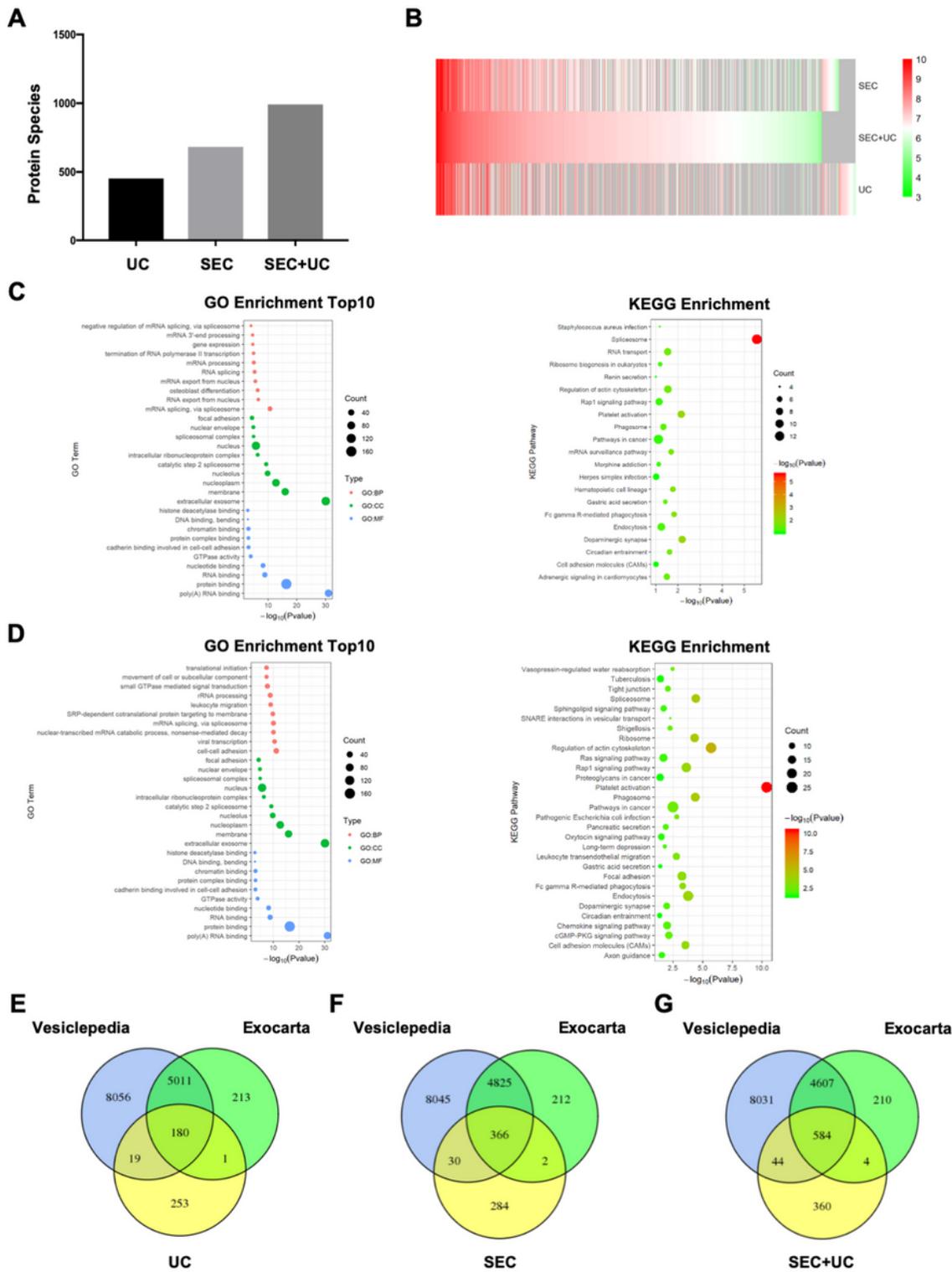
**Figure 2**

Characterization of plasma sEVs isolated by UC, SEC and SEC+UC. (A-C) sEVs isolated by three methods were observed by TEM. All sEVs fractions contained vesicles of variable sizes in the range of 60-100 nm. (D-F) Size distribution of sEVs was measured using NTA. The images shown were representative of three experiments. (G) Coomassie blue staining exhibited the total protein level of SEC+UC, UC and SEC fractions. (H) Western blotting showed different expression levels of three positive markers (CD9, CD81 and HSP90) in sEVs fractions isolated by UC, SEC and SEC+UC.



**Figure 3**

The protein purity of plasma sEVs isolated by UC, SEC and SEC+UC. (A) The statistic results of particle numbers in three groups by NTA (each group n=3). (B) The total protein concentration of isolated particles by BCA assay. (C) The ratio of BCA protein concentration/particle number. (D) Comparison of common lipoprotein contaminations (Albumin and Apo-A1) in sEVs fractions, and one negative marker of sEVs (Calnexin) by western blotting.



**Figure 4**

Proteomic analysis of sEVs isolated by UC, SEC and SEC+UC. (A) MS-analysis confirmed the total protein species of plasma sEVs. 453 in UC, 682 in SEC, 992 in SEC+UC. (B) A heatmap of all proteins' abundance in plasma sEVs isolated by SEC, SEC+UC and UC. (C) GO/KEGG enrichment of proteins identified by SEC+UC but not by SEC (left: a bubble plot of GOs; right: a bubble plot of 21 KEGG pathways enriched). (D) GO/KEGG enrichment of proteins identified by SEC+UC but not by UC (left: a bubble plot of GOs; right:

a bubble plot of 46 KEGG pathways enriched). (E-G) Bioinformatic analysis exhibited the sEV-associated proteins species and numbers. In total, at the intersecting area of Venn diagram, 180 EV-associated proteins of UC, 366 EV-associated proteins of SEC, and 584 EV-associated proteins of SEC+UC were identified.

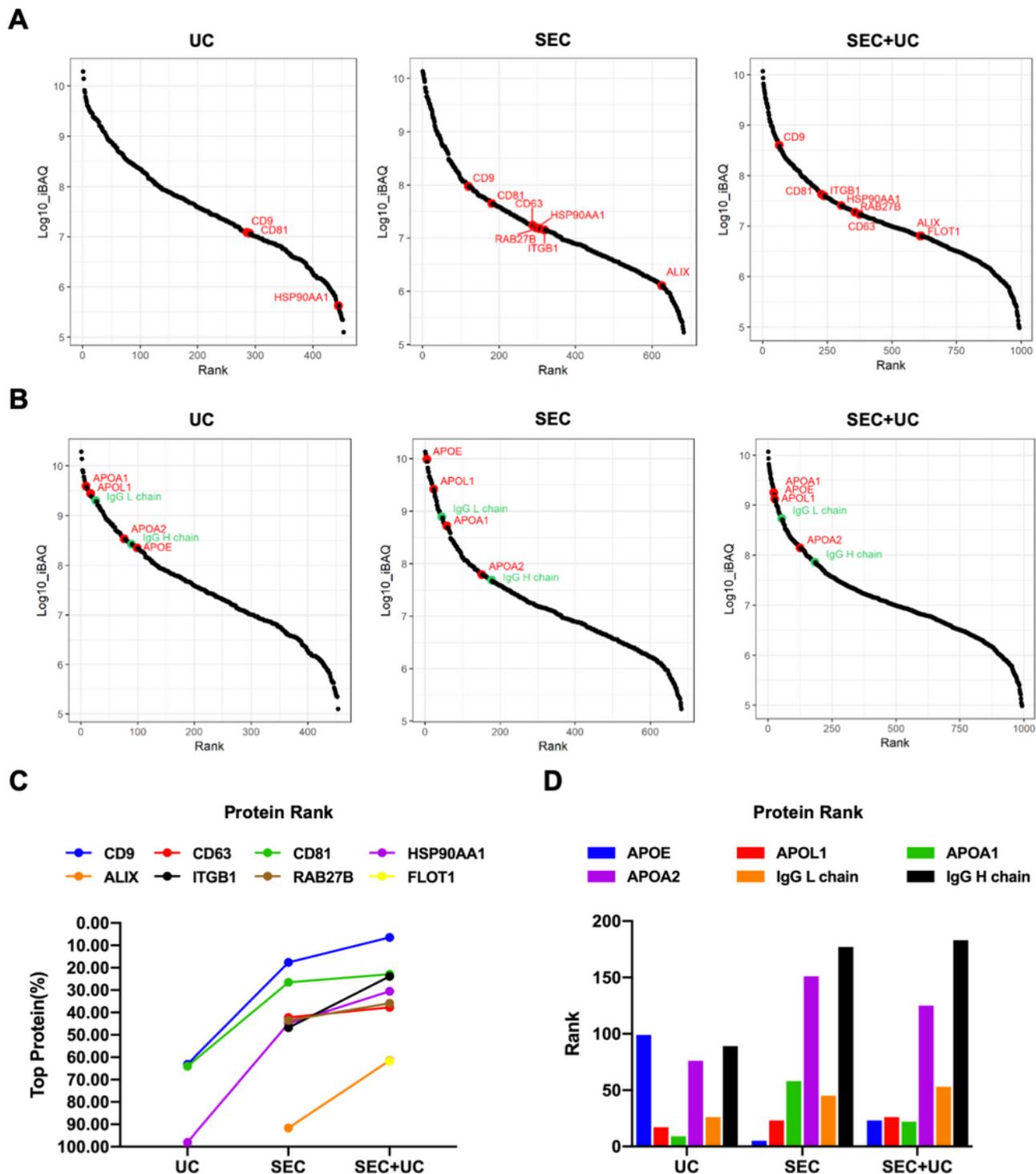


Figure 5

The species of sEV-associated proteins and common contaminants identified by MS. (A) Diverse sEV-associated proteins rank in three isolated assays. The dot plot highlighted common sEV-associated proteins rank, including CD9, CD81, ITGB1, HSP90AA1, RAB27B, CD63, ALIX and FLOT1. (B) Diverse sEVs contaminants of plasma lipoproteins and other free-floating proteins rank in the three isolated assays. The dot plot highlighted proteins rank, including APOE, APOL1, APOA1, APOA2, IgG L chain and IgG H chain. (C) The curve showed 8 sEV-associated proteins rank in total detected proteins of UC, SEC and SEC+UC groups. Top 0% represents the protein expression is the highest in MS-analysis, and Top 100% means the lowest. Data not shown of undetected proteins. (D) The curve showed 4 lipoproteins and 2 IgG fractions rank in total detected proteins of UC, SEC and SEC+UC groups. Top 0% represents the protein expression is the highest in MS-analysis, and Top 100% means the lowest