FACULTY OF MEDICINE AND HEALTH SCIENCES

Optimizing and miniaturizing ultrafiltration-based extracellular vesicle isolation towards liquid biopsy cancer diagnostics

Freya De Muyer, Danaë De Pauw
Student number: 01903675, 01904324

Supervisor(s): Prof. Dr. Kris Gevaert (promotor) and Jarne Pauwels (mentor)

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master in Science of Biomedical Sciences

Academic year: 2023 - 2024

GHENT UNIVERSITY

Department of Biomolecular Medicine, Ghent University, Ghent, Belgium.
VIB-UGhent, Center for Medical Biotechnology, VIB, Ghent, Belgium.

1. Preface

Writing and conducting this thesis has been an instructive experience during our academic course. Therefore, we would like to express our gratitude to many people that have supported us in completing the research proposal and this thesis.

Firstly, we would like to thank our promotor, prof. Kris Gevaert, for providing us the opportunity to conduct this master's dissertation in his lab, for giving constructive feedback to ameliorate our writing, presenting and critical thinking skills and for letting us be part of his esteemed research group.

In line with this, we would like to express our appreciation to our mentor, Jarne Pauwels, for being a major supportive factor throughout these academic master years. Thank you for always being available for answering questions, for sharing your opinion while encouraging us to have our own and for all the time you spent contributing to our practical lab skills and revising our work or, as you would say it, thank you for the team work. You also taught us to efficiently conduct lab experiments by anticipating future handlings and to think critically about the steps throughout a protocol. In short, you supported our fundaments of being competent researchers.

Furthermore, the work environment would not have been as pleasant without the Gevaert lab. Theoretical, nor practical questions were left unanswered and we could always count on their expertise opinions and their help in mastering RStudio.

Since we completed a duo thesis, our work is not contributable to ourselves only. Speaking in my name, Freya, I also have to thank Danaë for being a dream of a thesis partner. In my opinion, we complemented each other's skills and I loved learning from your critical mindset, your logical thinking and your ability to explain complexities in a simple and straightforward manner. Our discussions raised the level of comprehensiveness and precision of the work we delivered and I can only look back with wonderful memories to our cooperation. I think it's safe to say that I made a friend for life. Speaking of friends, I should also thank them for motivating me during writing this thesis, as we often worked together in discord sessions, but also for taking my mind off it by organizing small but fun events and dragging me away from my desk as I tend to be a workaholic. Lastly, a special thanks to my mom for supporting me in every possible way to accomplish my academic achievements and for being proud on them.

As for me, Danaë, I would like to express my appreciation towards Freya for being the best thesis partner I could have imagined. I am beyond grateful to have shared this journey together with not only a dedicated lab partner, but also a friend. Our personalities found much common ground, and in areas where they diverged we complemented each other in a mutually enriching cooperation. I have learned a lot from your perspective on the matter, work ethic and scientific reasoning. I really admire your commitment and perseverance. Thank you for making this thesis process both productive and enjoyable. Last but not least, I want to profoundly thank my family and Tshwarelo for the endless love and emotional support throughout my personal development as a young researcher.

2. Table of contents

3. Scientific summary

Extracellular vesicles (EVs) are a promising source for circulating cancer protein biomarkers, yet their isolation from complex biological matrices, including plasma, for subsequent proteome analysis remains challenging. The low abundance of the EV-proteome relative to the non-EV plasma proteome, coupled with potential co-enrichment of contaminants, hinder EV protein identification by bottom-up mass spectrometry (MS). Here, we optimized an ultrafiltration-based EV isolation technique, termed filter-aided EV enrichment (FAEVEr), that combines washing steps supplemented with the mild detergent TWEEN-20 towards a high throughput screening setup for data-independent acquisition (DIA) MS-based biomarker discovery. We extensively compared FAEVEr in a low and high throughput format to ultracentrifugation, a widely used EV enrichment strategy, in terms of differences in isolation efficiency. We further applied FAEVEr for proteome analysis on EVs isolated from cell culture conditioned medium, and expanded the strategy to more complex biological matrices, including serum and plasma. Our results indicate that FAEVEr supplemented with TWEEN-20 is a valid approach for EV enrichment and purification, improving EV-proteome analysis by DIA liquid chromatography-tandem MS (LC-MS/MS), with the most prominent effect of TWEEN-20 in the high throughput setup. We found that FAEVEr from conditioned medium, as well as from serum and plasma, has potential to discover protein biomarkers for cancer and infectious disease, although further optimization is required. In conclusion, our findings highlight the promise of FAEVEr combined with TWEEN-20 for EV enrichment and subsequent proteome analysis with DIA LC-MS/MS, enhancing cancer protein biomarker discovery, even in a high throughput setting.

4. Layman summary with societal impact

Cancer is a leading cause of death worldwide, accounting for one in six deaths, according to the World Health Organization. Compared to more advanced stages of cancer, early diagnosis improves prognosis and survival rates while reducing therapy complexity and invasiveness, thereby limiting the personal burden, as morbidity decreases and quality of life improves. It also lowers healthcare costs, as treatment tends to be more expensive at advanced stages. Early diagnosis could be obtained by detecting early cancer-specific biomarkers in bodily fluids. Promising carriers of such biomarkers are extracellular vesicles (EVs), small particles that are secreted by virtually all cells in bodily fluids, thereby carrying representative information of the parental cell. This thesis focuses on optimizing an innovative setup to isolate EVs from complex biological matrices for protein analysis, termed filter-aided extracellular vesicle enrichment (FAEVEr). EV surface proteins are particularly interesting, as they are easily accessible to affinity reagents leading to targeted diagnostic tests that offer several advantages. More specifically, these tests could involve minimally invasive liquid biopsies, such as blood that can be collected by a general practitioner, thereby reducing reluctancy, and can be repeated to monitor treatment effectiveness. Moreover, targeted tests could provide clearer results, with fewer false positives and a lower need for confirmatory tests, thus reducing expenses. In conclusion, our goal is to advance EV-based diagnostic tests for early cancer detection by improving EV enrichment for protein biomarker discovery to improve patient outcomes and reduce the burden on both individuals and society.

5. Introduction

5.1 General introduction to extracellular vesicles and their biogenesis

Extracellular vesicles (EVs) are a heterogeneous population of nano-sized particles that are secreted in the extracellular environment by virtually all cells, cannot replicate and are delimited by a lipid bilayer, as defined by the MISEV2023 guidelines[1]. Therefore, they are present in all bodily fluids such as cerebrospinal fluid, urine and blood[2]. Furthermore, based on their biogenesis, EVs can be divided into three subpopulations: exosomes, microvesicles and apoptotic bodies (Figure 1). Exosomes are extracellular vesicles of endocytic origin. They are formed by invagination of the endosomal membrane with the formation of multivesicular bodies (MVBs) containing intraluminal vesicles (ILVs)[3]. MVBs then fuse with the plasma membrane, releasing the ILVs as exosomes into the extracellular environment. Microvesicles are formed by outward budding of the plasma membrane forming heterogeneous membrane vesicles that cover a large size range[3]. Apoptotic bodies are formed by blebbing of the plasma membrane[3]. In bodily fluids, the combination of these three biogenesis pathways results in a heterogeneous EV population in terms of size and potentially distinct in physiology. In practice, EVs are often divided into groups based on their size. Therefore, in this thesis, we will use the term EVs as an overarching term for small EVs (sEV) and large EVs (lEV) for particles with a size of < 200 nm, and > 200 nm, respectively.

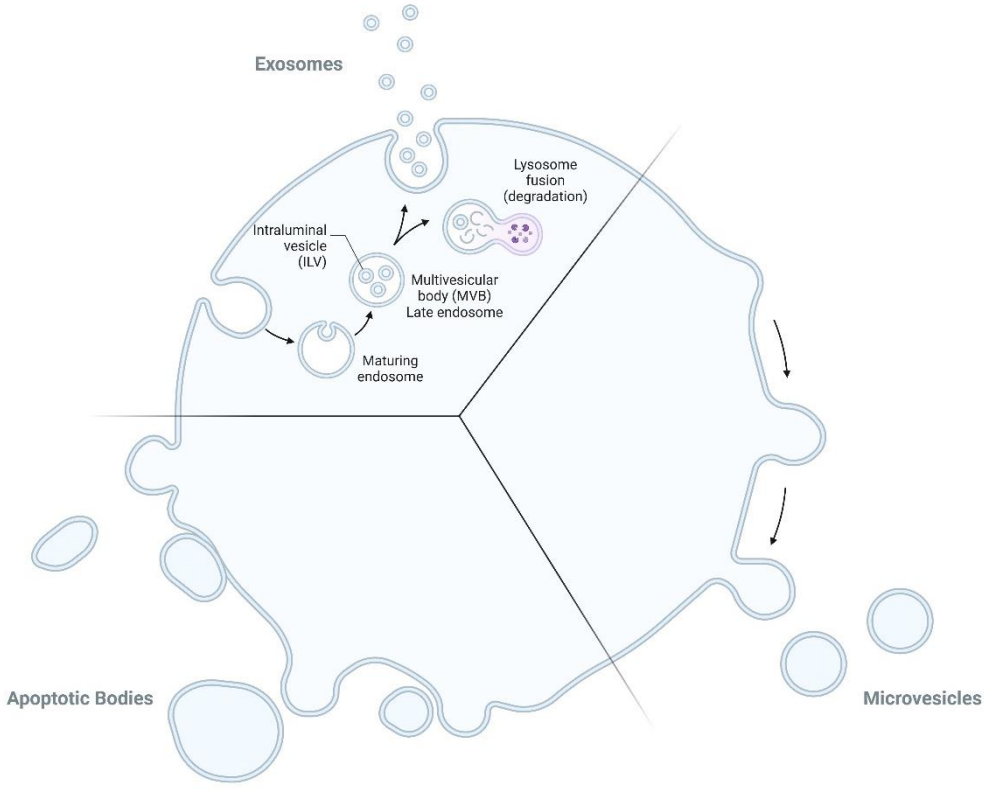


Figure 1: Overview of EV biogenesis of the three subpopulations: exosomes, microvesicles and apoptotic bodies.

Since EV biogenesis-related proteins of exosomes and microvesicles are taken into account during analyses to demonstrate their presence in the isolated fraction, the EV biogenesis pathways of exosomes and microvesicles are further elucidated. Orchestrated by the endosomal sorting complex required for transport (ESCRT) machinery, exosomes are

generated by the invagination of the endosomal membrane of a maturing endosome. This process results in the formation of multivesicular bodies (MVBs) which are late endosomes containing membrane-enclosed intraluminal vesicles (ILVs)[3–5]. On the one hand, due to the presence of specific surface proteins, such as the HSP70-HSP90 organizing protein complex, GTP-ase Ras-related protein RAB7A and protein members of the SNARE-complex, these MVBs fuse with lysosomes resulting in lysosomal degradation of the ILV cargo[5]. On the other hand, following the endosomal recycling pathway, MVBs release the ILVs as exosomes into the extracellular environment after fusion with the plasma membrane[3–5]. This exosome biogenesis pathway is regulated in an ESCRT-dependent or -independent manner[3–5]. The ESCRT-complex is composed of thirty proteins and can be divided into the ESCRT-0, -I, -II and -III subcomplexes[3,4] (Table S1). Alternatively, ESCRT-independent machineries sustain ILV formation and involve among others neutral sphingomyelinase, tetraspanins CD63 and CD9, ADP ribosylation factor 6 (ARF6) and phospholipase D2 (PLD2)[3–5]. Microvesicle (MV) biogenesis occurs as direct budding of vesicles from the plasma membrane following phospholipid rearrangement[3–5]. Therefore, their membrane and lipid content highly represent that of the plasma membrane itself, except for the plasma membrane asymmetry[6]. This MV biogenesis process remains less understood than the exosome biogenesis[3]. However, one of the known stimuli is the increase of intracellular calcium leading to disruption of plasma membrane anchorage to the cytoskeleton, thereby resulting in outward blebbing[3]. Scission of these blebs can be mediated by the ESCRT-complex protein TSG101, ESCRT-accessory protein PDCD6IP/ALIX, arrestin domain containing protein-1 (ARRDC1) and ADP-ribosylation factor 6 (ARF6)[3,5].

5.2 Functions of extracellular vesicles

The EV cargo, consisting of biomolecules such as oligonucleotides, lipids and proteins, is protected by a phospholipid membrane. In this way, EVs are present as stable entities in bodily fluids and their cargo remains functional and intact. By means of their cargo, EVs carry representative information of the parental cell and can contribute to intercellular communication[2]. More specifically, their lipid bilayer can incorporate (trans)membrane proteins that allow selective signal transduction and cargo delivery to receiving cells. In healthy conditions, EVs contribute to maintaining homeostasis of physiological functions. For instance, breast milk EVs influence neonatal immunity by increasing the number of a specific group of T-regulatory cells whereas the vasopressin-regulated water channel aquaporin-2 supervises the water permeability in the renal collecting duct cells aided by its transport by EVs[2]. However, EVs are also known to play a role in pathogenic processes such as carcinogenesis and tumor progression[7]. For example, breast tumor EVs induce the SMAD-mediated signaling pathway in mesenchymal stem cells by increasing tumor-promoting factors to transform them into tumor associated myofibroblasts, thereby promoting breast tumor progression[7]. Moreover, in brain pre-metastatic niches, tumor cells can be attracted by increased glucose availability due to suppression of the glucose uptake by stromal cells caused by tumor-derived microvesicles[8]. Lastly, hypoxic cancer cells were shown to promote growth and invasion by secreting a higher number of EVs[7]. Taken together, the EV stability, their representative information of parental cells and their suggested increased release by cancer cells make EVs a potential clinical asset for diagnostics.

5.3 Clinical purposes of extracellular vesicles

Liquid biopsies are non-solid biological samples, such as urine and blood, that can be analyzed for biomarkers to indicate the presence or severity of a disease state or to monitor the effect of treatment. Bodily fluids are easily accessible and their sampling is considered to be minimally to non-invasive, rapid, easy and cheap. Moreover, sampling can be performed repeatedly, which allows longitudinal follow-up of the patients to monitor disease progression or therapeutic response. Considering the stability of EVs in bodily fluids due to their protection by a phospholipid bilayer, EVs possess all the necessary characteristics to serve as a source

of biomarkers. Furthermore, they carry the specific fingerprint of the parental cell and may be secreted in higher quantities by cancer cells compared to healthy cells[9]. So, EVs could be analyzed for cancer diagnosis, evaluation of cancer progression and therapy response. Alternatively, EVs could be involved in therapeutic interventions as targets for chemoresistance in cancer cells[7] or as drug delivery vehicles to carry therapeutics to target cells.

5.4 Current research gap

5.4.1 The matrix complexity of plasma complicates extracellular vesicle isolation

Despite their biomarker-bearing potential, the implementation of EVs in clinical applications is limited by isolation challenges. Given the focus of this thesis on proteomics-based EV biomarker discovery towards plasma as the final, most challenging matrix, the most important pitfalls are further highlighted. Contamination of non-EV material is a major obstacle. On the one hand, EVs are very lowly abundant compared to plasma proteins like albumin which makes up 60% of the total plasma protein concentration of about 60-80 mg/mL[10]. Moreover, EVs, with an abundance of about 10^9 particles/mL plasma[11], are approximately 10^7 times less abundant in plasma compared to lipoprotein particles[11]. Consequently, EV cargo proteins are present in much lower numbers compared to plasma proteins and lipoproteins. On the other hand, contamination results from the overlap of EVs in size and density with other particles in plasma including protein aggregates, protein complexes and lipoprotein particles which are therefore potentially co-enriched[12,13]. Indeed, high density lipoproteins (HDLs) have a similar density to EVs while very low density lipoproteins (VLDLs), low density lipoproteins (LDLs) and chylomicrons (CMs) overlap in size[14] (Figure 2). By masking the EV proteins of interest, co-enriched highly abundant contaminating proteinspose a major consequence for bottom-up mass spectrometry (MS) analysis, which is still the method of choice for protein discovery. Indeed, the elution profile of peptides can be dominated by peptides from highly abundant proteins overshadowing less abundant peptides from biomarkers of interest. Furthermore, these high intensity signals can saturate the detector leading to dynamic range compression resulting in loss of lower abundant signals in noise. Also in DIA-mode, overabundant peptide material increases spectral complexity which complicates the identification of proteins of interest. In proteomics-based biomarker discovery, it is thus essential that EVs are enriched and purified prior to analysis. Over the years, multiple enrichment strategies have been developed, which are described in the next section.

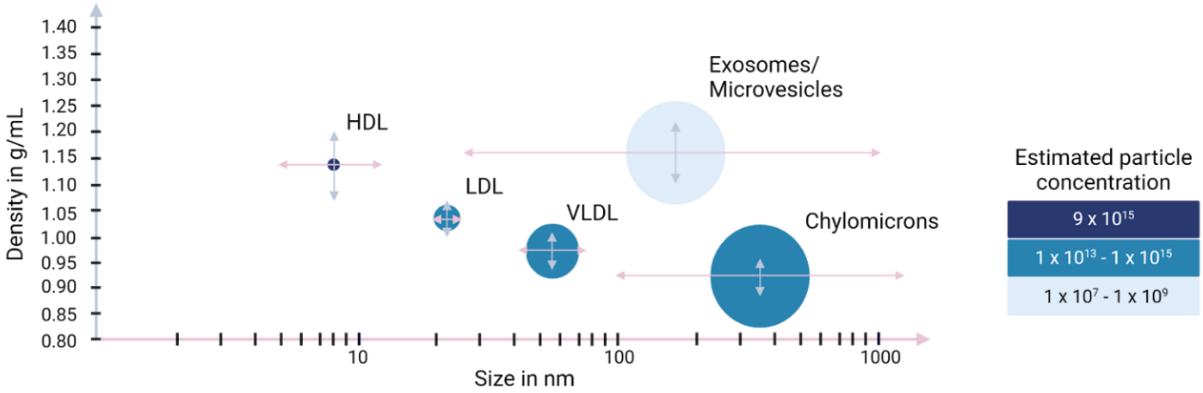


Figure 2: Overview of the overlap in size and density of lipoprotein particles (HDL, LDL, VLDL, chylomicrons) and extracellular vesicles (exosomes and microvesicles) in plasma. [Abbreviations: HDL = high density lipoproteins, LDL = low density lipoproteins, VLDL = very low density lipoproteins. (This figure is based on Simonsen, J. B et al.[11])

5.4.2 Different isolation techniques introduce technical variation

The majority of the existing EV isolation techniques exploit biophysical and biochemical EV characteristics[15]. Differential ultracentrifugation and density gradient centrifugation separate EVs based on density, whereas the aggregation of EVs upon addition of a hydrophilic polymer is employed in precipitation-based methods. Protein markers that are expressed on the EV membrane can also be exploited in immunoaffinity isolation methods. Lastly, the size of EVs is used for EV enrichment by filtration-based methods and size exclusion chromatography. As each of these approaches focuses on specific EV characteristics, each of them has inherent benefits and limitations. Therefore, orthogonal EV isolation techniques are often combined in practice[15].

a) Differential ultracentrifugation

Differential ultracentrifugation (dUC) involves sequential centrifugation rounds at increasing centrifugal forces or duration to separate particles based on sedimentation rate. More specifically, the low-medium speed steps remove cell debris and large particles, while the following high speed (100,000 - 200,000 x g) steps pellet the EVs. This technique is the most widespread method for EV isolation[13,16–19]. Although the equipment is expensive[13,17,19], the additional costs related to consumables are relatively low[16–19]. UC-based EV enrichments are highly reproducible when the same parameters are used[17,19]. The isolation efficiency and the characteristics of the isolated EVs are dependent on acceleration, rotor type, viscosity and centrifugation time[13,16,17,19]. Depending on the centrifugation tube length, a balance needs to be found between time and the loss of EVs that do not precipitate at the final high speed centrifugation[16–18]. Also, after each sequential step, the supernatant is manually removed, which potentially introduces inter-sample variability due to pellet disruption or incomplete removal of EV depleted supernatant. Additionally, the purity of the EV fraction is not optimal as non-EV particles with a similar sedimentation rate, such as lipoproteins and protein aggregates, tend to co-precipitate[13,16,17,19]. More centrifugation steps can be implemented to reduce contamination, though a trade-off has to be made between EV purity, EV yield[16,19] and time invested. Furthermore, because of high centrifugal forces and EV aggregate formation, it is described that dUC can damage EVs, thereby changing their morphological and functional properties[13,16,17,19]. Since the dUC procedure is labor-intensive, time-consuming, difficult to parallelize and not practical for handling small or large volumes of biomaterial , its clinical applicability remains limited[13,16–19].

b) Density gradient ultracentrifugation

The separation of particles by density gradient UC (dgUC) is performed in a single centrifugation step through distinct migration behavior in a viscosity or density gradient[16]. Two setups are possible. On the one hand, rate-zonal UC separates particles based on the speed of movement determined by particle size and mass, also known as sedimentation rate. This results in discrete zones each containing similarly sized particles. The centrifugation time should be optimized, because prolonged centrifugation can result in sedimentation of all particles at the bottom of the tube. On the other hand, isopycnic UC separates particles based on particle positioning in a layer with a specific density. This results in fractions containing particles with similar buoyant densities. dgUC leads to a higher purity of EVs compared to dUC[13,16–19]. A major advantage of isopycnic UC is the possibility to separate viral particles from EVs in a iodixanol gradient[17–19]. Another benefit is the greater preservation of vesicle morphology[17,18]. Still, with dgUC it is difficult to isolate EVs from similarly sized particles such as LDLs, IDLs, VLDLs or chylomicrons or particles with similar density such as HDLs[13,15,17,18,20]. Similar to dUC, dgUC uses expensive equipment. The preparation of a discontinuous gradient is a laborious and time-consuming procedure, and the fractionation adds to EV loss and inter-sample variability[17,18,21]. However, automation of these steps is possible, which significantly reduces these drawbacks[21]. Yet, a remaining disadvantage is the fact that the sample volume to be loaded on the gradient is very small. Therefore, often an extra EV concentration step

needs to be included. Moreover, the long centrifugation times and the use of an ultracentrifuge limit parallelization and render this technique less suitable for clinical applications[17].

c) Polyethylene glycol precipitation

Polyethylene glycol (PEG) precipitation relies on the aggregation of EVs in the presence of the highly hydrophilic polymer PEG, followed by precipitation using low-speed (1,500 x g) centrifugation. This isolation technique can be upscaled as it is cheap, simple and fast, and it allows parallelization of EV preparation using basic lab equipment[13,16–19]. Moreover, it greatly reduces sample volume[13,18] with a minimal loss of EVs that maintain their morphological and functional quality[16–19]. However, the purity of the EV preparations is very poor compared to the other discussed EV isolation techniques[16–19]. This is mainly due to co-precipitation of non-EV particles, poor solubility of the precipitated EV aggregates and the presence of PEG in the EV fraction, which additionally limits the compatibility with MS-analysis[13,16–19].

d) Immuno-affinity capture

To enrich the EVs from complex matrices, specific EV surface markers can be targeted by antibodies conjugated to a solid phase[13]. Commonly used antibodies include anti-CD9, anti-CD63 and anti-CD81. The washing capacity of immobilized EVs is related to the biological affinity of the antibody-antigen pair[13,16]. Immuno-affinity-based enrichment is specific and very sensitive in comparison with the other techniques. It is therefore capable of generating a particularly pure EV sample, even from very low input samples[13,16–19]. Thus, this isolation technique is an ideal candidate to investigate the specific role of disease-related EVs[18]. However, as the antibody-based selection introduces a bias towards marker-expressing EVs, these EV subtypes are overrepresented[16,19]. In addition, non-specific binding of contaminants to the antibodies on the solid phase can lead to lower EV purity[17,19]. Another issue is the recovery of EVs from the antibodies after isolation for other purposes[16,17]. The integrity of the EVs can be jeopardized, limiting downstream technical applications[19] and subsequent analyses. Despite the possibility of automation with a high reproducibility[17], limited availability, variability between vendors and high cost of antibodies restrict this method[13,16–19].

e) Size exclusion chromatography

Size exclusion chromatography (SEC) separates molecules that differ in hydrodynamic radius. Large molecules elute faster whereas small molecules are initially retained in the pores of a column resin and thereby have a higher retention time. The column length and width, together with the type of polymer bead determine the level of EV purification[16]. SEC is considered a loss-free method[17], but the inherent requirement to dilute the sample obligates further concentration of the obtained EV fraction, which typically results in lower yields[16,17]. Furthermore, SEC is a simple and rapid isolation method with a high reproducibility, provided that columns with the same properties are used[13,17–19]. Although sorbents are considered expensive[17], a column can be washed and reused[19]. Therefore, SEC can be deemed cost-effective. However, the main issue remains that the EV enriched fraction is often contaminated with similarly sized lipoproteins[13,16,17] and moreover, separation of different EV-types with similar sizes is impossible[16]. Together with a limited scalability, these disadvantages complicate the use of SEC for clinical applications.

f) Filtration-based methods

For the isolation of EVs, microfiltration (MF) is typically used for pre-clearing the sample, while ultrafiltration (UF) can be applied for the actual isolation of EVs. UF is often performed as a concentration step prior or post enrichment in other orthogonal EV isolation methods such as dgUC or SEC[15], indicating its substantial implementation in the field of EV research. MF and UF filters have pore sizes ranging from 10 µm to 0.1 µm and 0.1 µm to 0.01 µm, respectively. Membranes with different molecular weight cut-offs (MWCO) are also used[16,17]. Particles exceeding the pore size or MWCO are retained on the filter (the retentate), whereas the smaller particles are collected in the flow-through (the filtrate). There are some inherent challenges when using UF for EV isolation from complex protein mixtures. Firstly, accumulation and deposition of retained particles or biomolecules at the filter leads to membrane fouling and

might impair the EV isolation process or even result in irreversible filter clogging leading to physical obstruction of fluid passage. Both membrane fouling and clogging may reduce the purity of the EV fraction, because more contaminating proteins are retained[16–18]. Moreover, depending on the membrane type, EVs can be trapped by the filter, leading to inefficient recovery[13,16–18]. Despite the aforementioned complications, UF is a fast, straightforward and low-cost procedure for which no special equipment, nor special skills are required[16–18]. Therefore, UF could serve as a high throughput enrichment strategy for which we will establish a workflow in this thesis.

g) Standardization and downscaling of EV isolation remains a bottleneck

There are still some unmet needs for the development of EV-based clinical technologies[16]. Due to a heterogeneity in size, density, source, biogenesis, shape and composition of EVs, a standardized isolation technique for all EV subpopulations is not available[12,13,16–19]. Varying among labs, many efforts are made to create an optimal EV isolation method dependent on the desired application. However, inter-lab variability leads to inconsistent yield, purity and biophysical properties of the isolated EV fraction[12,16,18,19]. Therefore, subsequent analysis results in lowly reproducible and differing conclusions[18,19]. To address this inter-lab variability, minimal information for studies of extracellular vesicles (MISEV) guidelines have been created for EV research, to increase reproducibility and to ensure quality[1]. These guidelines provide an overview of the nomenclature, isolation strategy, concentration and characterization of EVs and give recommendations to contribute to standardized execution and reporting of EV experiments[1,22]. In addition, reproducibility and standardization is improved by EV-TRACK[23]. This knowledgebase was developed to exchange experimental data of EV research. It also facilitates communication between researchers and thereby advances EV research. Even though this renders EVs more useful to incorporate in diagnostic applications, the scale of EV isolation is not yet sufficiently adapted to clinical practice. Concretely, there is still room for improvement concerning parallelization, throughput, time consumption and costs[12], as discussed above and in Table S2.

5.5 FAEVEr as an ultrafiltration-based extracellular vesicle isolation setup with TWEEN-20

UF is a suitable EV isolation method to bridge the gap towards EV isolation on a higher scale. The method is time efficient and straightforward which allows for parallelization in contrast to UC where this is limited by the rotor size and loading capacity. Moreover, UF is considered cost-effective because it does not require expensive equipment. In this thesis, an EV isolation protocol, referred to as filter-aided EV enrichment or FAEVEr[24], is proposed (Figure 3). FAEVEr is based on UF using a 300 kDa MWCO polyether sulfone (PES) membrane which is hydrophilic with low protein adsorption characteristics and is resistant to a high pH range and a wide variety of detergents, solvents and buffers at different temperatures. The method implements the mild detergent TWEEN-20 to additionally reduce membrane fouling and to facilitate the UF flow-through by reducing non-specific interactions with the membrane. In this way, the higher purity of the EV fraction could enhance subsequent proteome analysis and proteomics-based EV marker discovery with data independent acquisition liquid chromatography-tandem MS analysis (DIA LC-MS/MS).

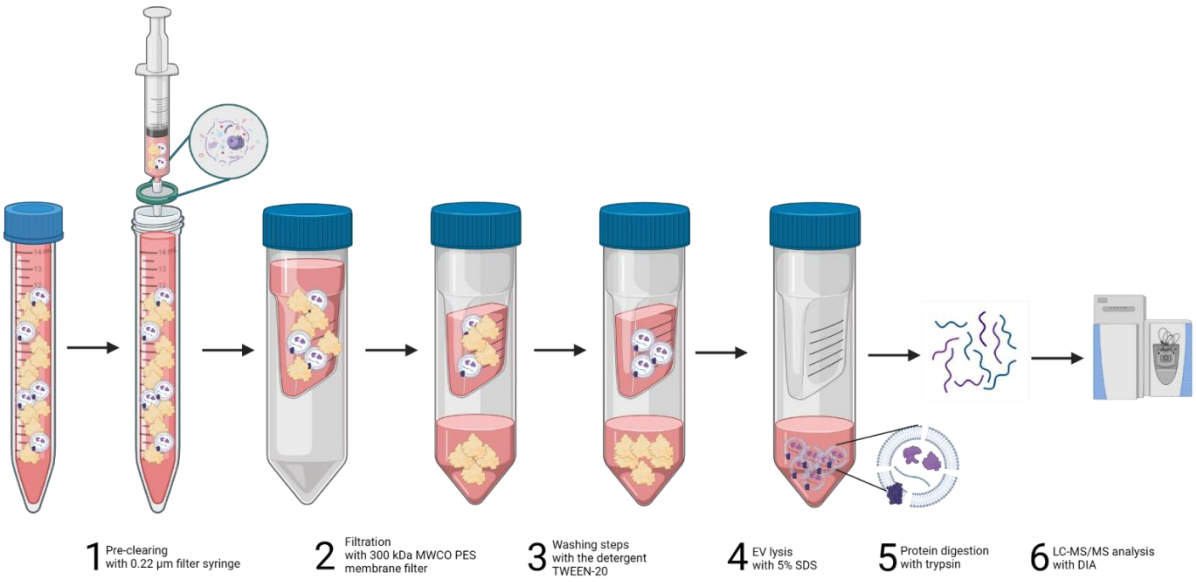


Figure 3: Overview of the EV isolation workflow of filter-aided EV enrichment or FAEVEr referring to ultrafiltration on 300 kDa molecular weight cut off (MWCO) PES membrane filters using the TWEEN-20 detergent for EV proteome analysis with data independent acquisition liquid chromatography-tandem MS analysis (DIA LC-MS/MS). 1) Input samples are pre-cleared with a syringe equipped with a 0.22 µm filter to remove cell debris. 2) Samples that remained undiluted or were diluted with TWEEN-20 are filtered using a 300 kDa MWCO PES membrane. 3) Subsequent washing steps with TWEEN-20 remove additional contaminating proteins. 4) The EVs in the EV-enriched fraction are lysed on the filter with 5% SDS. 5) After further sample preparation, the EV proteins are digested with trypsin. 6) Finally, DIA LC-MS/MS analysis is performed.

5.1 Characterization of the extracellular vesicle fraction

To verify the quality of the EV isolation, SDS-PAGE, Western blot (WB), transmission and scanning electron microscopy (TEM and SEM), nanoparticle tracking analysis (NTA), and mass spectrometry (MS) are typically conducted. More precisely, SDS-PAGE is used to indicate the abundance of contaminating proteins at different stages of the protocol, while WB visualizes the EVs retained on the filter compared to the EVs lost during EV isolation. With NTA, the size distribution of nanoparticles is measured. Since NTA does not differentiate between EVs and other nanoparticles, SEM and TEM are used to complement this technique by visualizing the structure of these particles. Alternatively, DIA LC-MS/MS provides deeper insights on the EV proteome.

5.2 Recombinant extracellular vesicles for quality control

Recombinant extracellular vesicles (rEVs) are immature virus-like particles produced after the expression of a polyprotein of HIV-1 Gag linked to eGFP that hijacks the EV-releasing cell mechanism[9,25]. This is due to the accumulation of the Gag-eGFP polyprotein at the intracellular and endosomal membrane surface. In this way, rEVs are surrounded by a lipid bilayer and enriched for luminal Gag-eGFP molecules and EV-associated proteins. Therefore, they contain EV-like biophysical and biochemical characteristics and thus show a similar behavior to sample EVs. Moreover, they are trackable and can be used for monitoring EV integrity and for quantification of the EV recovery after isolation. Consequently, they form an ideal biological reference material for EV-based sample preparation and analysis, data normalization and method development.

5.6 Patient-derived xenograft mouse models

Patient-derived xenograft (PDX) mouse models are constructed by subcutaneous injection of suspended patient tumor tissue into immunodeficient mice. In this thesis, EVs will be isolated from plasma of a heterotopic breast cancer PDX mouse model. Similar to Barlin *et al.*, the goal is to differentiate between EVs from the murine host and human tumor to find biomarkers specifically related to breast cancer[26]. These PDX models preserve the heterogeneity of the primary patient tumor and contain representative phenotypic and molecular characteristics of the patient's tumor. Moreover, the differential species-related origin helps in the distinction between healthy (from the mouse) and tumor-related EVs. However, introduction of new mutations due to serial transplantation over mouse generations is possible. Therefore, identified human tumor markers could be directly linked to the primary tumor but could also result from random mutagenesis. Furthermore, some proteins are evolutionary conserved between both species, which may complicate proteome analysis. Other disadvantages inherent to a heterotopic subcutaneous model are the low occurrence of spontaneous metastases and the reduced clinical relevance of the tumor setting as relevant stroma interactions are lacking. Therefore, the role, interaction pattern and cargo of tumor EVs could be impacted.

5.7 Scope of the thesis

The overall aim of this work was to optimize the FAEVEr set-up using 300 kDa MWCO PES membrane filters in combination with TWEEN-20 for EV isolation and purification from matrices including cell medium and bodily fluids to advance cancer protein biomarker research. Additionally, we aimed to explore miniaturization of this setup to potentially facilitate transition to high throughput clinical applications. Both the low and high throughput setup will be applied to different biological matrices to test the EV isolation efficiency. Noteworthy, the scope generally excludes large EVs, like apoptotic bodies.

6. Materials and methods

6.1 Cell culture

a. Production of rEV material

A complete protocol for the transformation and production of recombinant extracellular vesicles (rEV) material is described by Geeurickx *et al.*[9] and Eyckerman *et al.*[27]. In short, approximately 4.0E+06 HEK293T cells of low passage number (<10) were seeded in T75 flasks, supplemented with Dulbecco's Modified Eagle Medium (DMEM, Gibco) containing 10% fetal bovine serum (FBS) and incubated in 5% CO_2 at 37 °C for 48 h. The HEK293T cells were transfected by refreshing the cell medium to 9.5 mL DMEM + 2% FBS, adding a mixture of 625 µL DMEM with 7 µg bait structure (pMET7-Gag-eGFP) and 625 µL DMEM with 37.5 µL polyethylene imine (PEI) per T75 plate and incubating the cells in 5% CO_2 at 37 °C for 6 h. Subsequently, the cell medium was discarded and replaced with 8 mL fresh DMEM containing 10% EV-depleted FBS (EDS, Thermo A2720801) followed by incubation of the cells in 5% CO_2 at 37 °C for 48 h. After successive evaluation of the transfection efficiency under UV light (excitation at 488 nm and emission at 507 nm), the conditioned medium (CM) was isolated and immediately pre-cleared.

b. Pancreatic ductal adenocarcinoma cell lines

Three different pancreatic ductal adenocarcinoma cell lines, PANC-1, CAPAN-1 and MIA-PaCa-2 were cultured using standard cell culture protocol. Both PANC-1 and CAPAN-1 were cultured in DMEM (Gibco, 11584516) supplemented with 1 mM sodium pyruvate (Gibco, 11360070), 5% Pen Strep (Gibco, 11548876) and 10% fetal bovine serum (FBS). MIA-PaCa-2 cells were cultured in DMEM supplemented with 1 mM sodium pyruvate, 5% Pen Strep, 10% FBS and 2.5% horse serum. When the cells reached 70% confluence, they were split (1:5) in T25 flasks. After enough cells were cultured (70% - 80% confluence), the medium was discarded and washed three times with PBS (room temperature) before adding 3 mL DMEM supplemented with 1 mM sodium pyruvate, 5% Pen Strep and 10% EDS. After 48 h, the CM was isolated and transferred to fresh tubes for pre-clearing. All cell cultures were checked by bright-field microscopy.

6.2 Biological samples

a. Mouse serum

Mice were sacrificed with an overdose of anesthetic (pentobarbital at dose of 300 mg/kg) after surviving influenza infection and terminal blood was collected from the orbital sinus by removing the eye and placed at 4 °C for approximately 24 h. After this, the samples were centrifuged at 6,800 x g for 5 min. The supernatant was transferred to a new tube, after which the centrifugation step was repeated. The supernatants were stored at -20 °C. The samples were not heat-inactivated as to prevent protein aggregation.

b. Patient-derived xenograft mouse models

Two surgically resected patient-derived tumors (TM00096 and TM00098) were heterotopically engrafted via subcutaneous injection into immunodeficient nod-scid-gamma (NSG) mice at the Jackson laboratory (JAX). These tumors, both grade three invasive breast carcinomas of no special type, originated from female patients with unspecified ages. TM00096 tumor cells were collected from the breast, while the TM00098 tumor cells were sourced from the lung. Serial transplantation of the tumor cells into other NSG mice was carried out over eight or nine passages following sacrifice of the preceding mice with isoflurane or CO_2 (Table S3). Blood was collected through cardiac puncture by the laboratory of Steven Goossens and Kaat Durinck. Subsequently, samples were centrifuged twice at 2,000 x g at 4 °C for 10 min in EDTA

tubes and stored frozen at -80 °C. In this thesis, biological replicates of generation F8 and F9 patient-derived xenograft (PDX) mice bearing these two tumor types were used to isolate EVs from mouse plasma using FAEVEr with TWEEN-20. The tumor size of the sacrificed F8 and F9 mice ranged from 1492.992 mm^3 to 2976.75 mm^3, with a median of 2078.85 mm^3 (Table S3). The volume of collected blood plasma ranged from 123 to 600 µL, with a median of 450 µL (Table S3).

6.3 Pre-clearing the sample

CM and sera were pre-cleared by a centrifugation step at 1,000 x g at room temperature for 5 min to remove loose cells and debris. Mouse plasma was centrifuged at 12,500 x g for 20 min at 4 °C. Subsequently, the supernatant was filtered through 0.22 µm syringe filters (Millex). Mouse plasma was filtered using Nanosep spin columns with a 0.2 µm wwPTFE membrane (*PALL Life Sciences,* ODPTFE02C34). This 0.22 µm filtration step was applied to all sample matrices including serum and plasma next to CM. The pre-cleared sample was divided over different aliquots if necessary and frozen in -80 °C until further use.

6.4 Ultracentrifugation

Pre-cleared CM was divided into *Beckman Coulter*® Polycarbonate thick-wall Centrifuge Tubes (REF: 343778), with each tube containing a final volume of 1 mL sample. The samples were centrifuged at 120,000 x g at 4 °C for 70 min using an OptimaTM TLX Ultracentrifuge with rotor SN381 TLA120.2. A fraction of the supernatant was collected and the remainder was discarded. The pellet was washed with different percentages of TWEEN-20 in phosphate buffered saline (PBS), followed by another round of centrifugation (120,000 x g at 4 °C for 70 min). Afterwards, a fraction of the supernatant was collected and the rest was discarded. The EV pellet was either recovered by adding 50 µL of 0.1% TWEEN-20 in PBS followed by vortexing, or the EVs were lysed by adding 50 µL of 5% SDS in 50 mM TEAB followed by 5 min of incubation at room temperature before collecting the lysate for proteomic applications.

6.5 FAEVEr on 300 kDa MWCO filter with TWEEN-20

FAEVEr was evaluated in a low and high throughput format. In the low throughput format (UF6), individual *Sartorius* Vivaspin® 300 kDa MWCO PES membrane filter tubes with a maximum volume of 6 mL were used (REF: VS0652), in contrast to the high throughput format (UF96), where an *Agilent Technologies* 96-well 300 kDa MWCO PES membrane filter plate with a maximum well volume of 400 µL (Material No: 201598 - 100) was used. After pre-clearing, for UF6 the filtrate was collected and diluted 1:1 with TWEEN-20 in PBS (final TWEEN-20 concentration within a range of 0% - 0.1% - 0.5% - 1.0% - 2.5%- 5.0%) to a final volume of 5 mL per diluted sample. For UF96, 500 or 600 µL of pre-cleared sample was collected and used undiluted. When using the UF6 filter tubes, the entire volume of the diluted samples was loaded onto the filter device and centrifuged at 4,000 x g at room temperature for 30 min. For the UF96 filter plate, samples were loaded twice and centrifuged at 1,000 x g at room temperature for 15 min. Subsequently, three washing steps were performed with 1 mL and 200 µL of the corresponding concentration of TWEEN-20 in PBS by centrifugation at 4,000 x g and 1,000 x g for 3 min or 7 min at room temperature for UF6 and UF96, respectively. A final washing step (same centrifugation settings) with similar volumes of PBS without TWEEN-20 was conducted. Finally, for UF6 and UF96, respectively, the EVs were either recovered by adding 125 µL or 50 µL of 0.1% TWEEN-20 in PBS, pipetting up and down along the membrane, and transferring the entire volume to Eppendorf tubes, or the EVs were lysed by adding 125 µL or 50 µL of 5% SDS in 50 mM TEAB, pipetting up and down, and centrifuging at 4,000 x g or 1,000 x g at room temperature for 10 min or 7 min. After each step, a fraction of the filtrate was transferred to a collection plate and the remainder was discarded. Various swing-out centrifuges were used for executing the protocol, including the Eppendorf 5804 R with rotor A-4-44 for the filter tubes and the Universal 320 with rotor 1460, the Eppendorf 5810R with rotor A-4-81 or rotor A-4-62, or the Eppendorf 5430 R with rotor A-2-MTP for the 96-well filter plate.

6.6 SDS-PAGE Coomassie and Western Blot

Approximately 25 µg of protein material was mixed with 4 x sample buffer and 20 x reducing agent (BIO RAD, 610791 and 1610792) and incubated for 5 min at 95 °C prior to SDS-PAGE analysis. Protein material was separated on 4-12% ExpressPlus™ PAGE Gels (M41215) or 4-20% *GenScript* SurePAGE™ (M00657) for 1h and 20 min at 120 V. Precision Plus Protein All Blue Standards (Cat. #1610373) was used as molecular weight marker. For Coomassie staining, gels were washed in dH_2O for 5 min, stained with InstantBlue® Coomassie Protein Stain (ISB1L, ab119211) for 30 min and washed again for 5 min with dH_2O to reduce background staining before imaging (*Odyssey*® LI-COR 9120 Imaging System or the GE Amersham Imager 680 RGB). For Western blotting, gel-separated proteins were transferred to an activated PVDF membrane at 100 V for 30 min and blocked for 30 min with 10 mL of blocking buffer (1:1, Tris buffered saline with 0.1% TWEEN-20 (TBS-T) and *LI-COR* Intercept® (TBS) Blocking Buffer (927-60001)). Primary antibodies were mouse anti-Gag HIV1 (AB63958, 1:1,000), rabbit anti-syntenin 1 (SDCBP) (AB19903, 1:1,000) and rabbit anti-calnexin (AB22595, 1:1,000). After overnight incubation at 4 °C, secondary antibodies, goat anti-mouse IRDye® (800CW, 1:5,000) and goat anti-rabbit IRDye® (680RD, 1:2,500) were added for 1 h at room temperature, washed three times with TBS-T and imaged using the *Odyssey*® LI-COR 9120 Imaging System or the *Odyssey*® Fc

6.7 Sample preparation for mass spectrometry

For the sample preparation preceding mass spectrometry (MS) analysis, the S-trap protocol (described in S4) was performed on the 5% SDS-mediated lysates, in which poorly soluble molecules are also dissolved. This protocol was executed using either S-trap™ mini columns (100-300 µg) or the 96-well S-trap™ plate (100-300 µg per well). Following peptide elution, the samples were transferred to MS vials, dried through vacuum evaporation in the *Savant* SpeedVac SC200 for approximately three hours and resuspended in 30 µL MS loading solvent containing 2% ACN and 0.1% trifluoroacetic acid (TFA) in dH_2O. Finally, the peptide concentration was measured at 230 - 450 nm using the UV/VIS spectrophotometry Lunatic device DropSense 16 (SN 100448) through a Lunatic chip (product code: 701-2009) with a pipetting volume of 2 µL and single path length of 0.5 mm per microwell.

6.8 Mass spectrometry analysis

The injected samples (15 µL) were analyzed using either the Orbitrap Fusion Lumos, the Q-Exactive HF or the Orbitrap Exploris 240 mass spectrometer through data independent acquisition (DIA). Proteomics data was processed using DIA-NN[28] based on project-specific spectral libraries generated from Swiss-Prot database FASTA files of the complete human proteome (20,598 sequences), Gag-eGFP polyprotein in cases of rEV isolation, bovine serum proteins in case fetal bovine serum (FBS) or EV depleted fetal bovine serum (EDS) was used as or was present in the sample matrix, and mouse proteome (21,709 sequences) in case mouse serum or plasma was used as matrix. Search parameters included two missed cleavages for trypsin/P, cysteine carbamidomethylation as fixed modification, both methionine oxidation and N-terminal acetylation as variable modifications, peptide length ranging from 7 to 30 amino acids, precursor charge ranging from 1 to 4, precursor m/z ranging from 375-900, fragment ion m/z ranging from 200-1800, mass accuracy of 20, MS1 accuracy of 10, protein inference based on genes, neural network classifier on double pass mode, cross-run normalization dependent on retention time (RT), isotopologues included, matching-between-runs (MBR) included and no shared spectra included.

6.9 Nanoparticle tracking analysis

To determine the size distribution and the mean number of particles present in the recovered samples, nanoparticle tracking analysis (NTA) was performed using the ZetaView device. The ZetaView protocol was followed (described in S5). In short, the instrument was calibrated by injecting 1 mL of 250,000x diluted ZetaView beads at 23 °C. The recovered samples were 25x

diluted with PBS to 1 mL and measured after injection into the instrument. For the measurements, the ZetaView software (version 8.05.16 SP2) was used. After video acquisition with 60 frames per second for 2 cycles (sensitivity 70, shutter 100) and laser wavelength at 520 nm, standardly all measurements were included in the analysis, even when they were indicated as outliers by the instrument. In between measurements, the instrument chamber was flushed with PBS. When the number of particles exceeded the threshold, the samples were diluted more than 25x.

6.10 Electron microscopy

b. Scanning electron microscopy

Nanosep centrifugal 300 kDa MWCO PES membrane filters containing EVs were placed in 2 mL Eppendorf tubes and fixed with 4% paraformaldehyde and 2% glutaraldehyde in 0.1 M sodium cacodylate buffer at room temperature. After 1 h, the fixative was replaced and filters were washed three times for 5 min with dH_2O. After dehydration in 30%, 50%, 70%, 95% and 2x 100% ethanol for 30 min each, the filters were critical point dried (EM CPD300, Leica) and imaged on a Crossbeam 540 SEM (Zeiss) at 1 kV.

c. Transmission electron microscopy

Aliquots (5 µL) of the solutions containing EVs isolated with Nanosep centrifugal 300 kDa MWCO PES membrane filters were blotted for 1 min on formvar- and carbon-coated Ni maze grids (EMS), which were glow discharged for 40 s at 15 mA. The grids were then washed five times in droplets of dH_2O and stained in a droplet of 1/4 Uranyl Acetate Replacement Stain (EMS)/dH_2O for 1 min. Excess stain was removed with filter paper and the grids were air dried for at least 4 h before viewing with the TEM. Imaging was done at 80 kV on a JEM1400plus (JEOL).

6.11 Statistical analysis and figures

To enhance the quality of protein identification and quantification, the label-free-quantification (LFQ) data generated with DIA-NN underwent further preprocessing using KNIME[29]. This involved the removal of non-proteotypic peptides and precursors with a q-value above 0.01. In addition, proteins that were identified with only one peptide across the different experimental conditions were removed. Subsequently, the protein LFQ values were log_2-transformed to obtain a normal distribution. For statistical analysis of the proteomic profiles of EV-enriched fractions, Perseus software was employed[30]. Two sample t-tests or ANOVA multiple sample tests were performed over all samples on either 100% valid values or imputed missing values using the normal distribution with width and downshift parameters set at 0.1 and 2.2, respectively. After calculating the Z-scores and selecting significantly ($p<0.05$) differential proteins, hierarchical clustering was performed (Pearson correlation). Clusters were manually defined and subjected to gene ontology enrichment analysis (Cellular Component) against human, bovine and mouse proteome databases using the online STRING database tool. If values were imputed, these imputed values were converted back into missing values and re-visualized with the defined clusters. Additionally, Perseus was employed to generate volcano plots with a false discovery rate (FDR) set at 0.05 and S0 at 0.1 (default). When creating these volcano plots, columns were selected according to the categorical row annotations that were tailored to the groups being compared. Other numerical graphics and figures were created with RStudio and BioRender, respectively.

7. Results

PART 7.1 Optimizing the FAEVEr setup with TWEEN-20 for proteome studies

Although ultrafiltration (UF) is well-implemented in other extracellular vesicle (EV)-enrichment strategies (e.g. dgUC, SEC) as a concentration step prior or post enrichment, it holds many challenges as a stand-alone EV enrichment and purification strategy. Here, we explored the possibilities for improving proteomics-based discovery of EV-biomarkers using filter-aided extracellular vesicle enrichment (FAEVEr) with 300 kDa MWCO polyethersulfone (PES) membranes. The EV enrichment and purification using FAEVEr heavily depends on membrane fouling caused by non-specific interactions of non-EV proteins with the filter membrane. To address this, we explored the use of TWEEN-20, a rather mild detergent known to reduce non-specific protein adsorption to PVDF membranes in Western blotting. We reasoned that TWEEN-20 addition could also reduce non-specific interactions of globular non-EV proteins and the 300 kDa MWCO PES membrane filter. Previous results already confirmed that the addition of 5% TWEEN-20 in combination with FAEVEr leads to improved EV purity[24]. The choice of this particular TWEEN-20 concentration was justified by results obtained with data dependent acquisition (DDA) mass spectrometry (MS) analysis.

We reassessed previous results for three main reasons. Firstly, data independent acquisition liquid chromatography-tandem mass spectrometry (DIA LC-MS/MS) analysis has recently made its entrance in the proteomics field. DIA LC-MS/MS is more capable of coping with large dynamic range differences of proteins and generally leads to less missing values. Secondly, we observed that 5% TWEEN-20 results in a smear on SDS-PAGE protein gels visualized with Coomassie staining, especially in the low molecular weight range, where multiple EV protein markers (CD9, CD81 and CD63) are found. This might affect the visualization of EV-specific markers in this molecular weight range. Thirdly, high concentrations of TWEEN-20 increase the sample viscosity and decrease the filtration performance. Therefore, using DIA-MS, we explored the use of lower TWEEN-20 concentrations for FAEVEr and their impact on EV proteome analysis. We initially performed a proof-of-principle experiment using 0.1% TWEEN-20 (described in S6) and continued with a more elaborate comparison between FAEVEr (on both 6 mL (UF6) and 96-well (UF96) format) and ultracentrifugation (UC) using different percentages of TWEEN-20.

7.1.1 Experimental setup for a large-scale comparison of EV enrichment efficiency between different enrichment strategies

Our results from a proof-of-principle experiment using 0.1% TWEEN-20 (S6) indicated that incorporation of lower TWEEN-20 concentrations hold promise for enhancing the FAEVEr set-up for proteomics-based EV biomarker discovery with DIA LC-MS/MS, as 0.1% TWEEN-20 decreases the complexity of the chromatographic profile compared to no TWEEN-20 addition, while obtaining a higher number of protein identifications compared to 5% TWEEN-20 (Figure S1**A**). Building further on this concept, we determined the optimal TWEEN-20 concentration towards minimizing contamination in the EV-enriched fraction and potentially increasing identifications of EV-specific proteins with DIA LC-MS/MS. Therefore, we conducted a comparative experiment employing different EV enrichment strategies in combination with increasing TWEEN-20 concentrations added to the wash buffer (Figure 4). Varying starting volumes of rEV-conditioned medium (CM) (1 mL, 2.5 mL and 600 µL) were used to compare the UC, UF6 and UF96 enrichment protocols, respectively. For each of the three isolation techniques, we tested a wide range of supplemented TWEEN-20 concentrations (0.0% - 0.1% - 0.5% - 1.0% - 5.0%) in quintuplicate, resulting in 25 samples for each EV enrichment strategy. The lysed rEV-enriched fractions were analyzed by SDS-PAGE, Western blotting (WB) and DIA LC-MS/MS analysis, while the recovered rEVs were subjected to nanoparticle tracking analysis (NTA) and scanning (SEM) or transmission (TEM) electron microscopy.

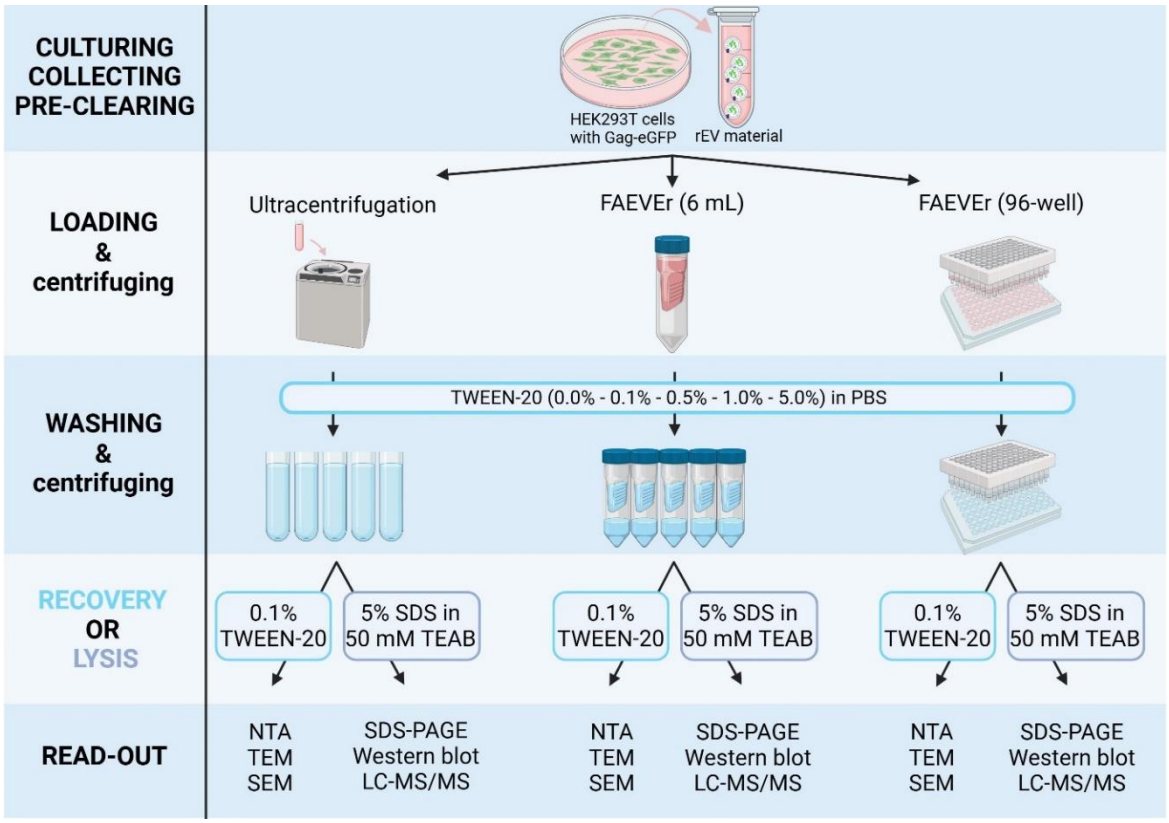


Figure 4: Overview of the comparative experimental setup for the optimization of FAEVEr. CM supplemented with 10% EDS containing rEVs secreted by cultured Gag-eGFP transfected HEK293T cells, is collected, pre-cleared and divided into 25 samples per EV-isolation protocol (Ultracentrifugation, FAEVEr on 6 mL or on 96-well format). The samples are loaded and centrifuged. Then, washing steps with one of the TWEEN-20 concentrations in PBS are alternated by centrifugation steps. Finally, the samples are either recovered with 0.1% TWEEN-20 in PBS or lysed with 5% SDS in 50 mM TEAB for further compatible read-out methods. [Abbreviations: CM = conditioned medium, EDS = EV-depleted FBS, rEV = recombinant extracellular vesicles, TEAB = triethylammonium bicarbonate, NTA = nanoparticle tracking analysis, SEM = scanning electron microscopy, TEM = transmission electron microscopy, LC-MS/MS = liquid chromatography-tandem mass spectrometry].

7.1.2 FAEVEr and UC show similar efficiency in removing non-EV protein material

To validate the removal of non-EV proteins in each of the experimental setups, the flow-through and supernatant (for FAEVEr and UC, respectively) was analyzed by SDS-PAGE (Figure 5). Prior to loading the fractions, the volumes were adjusted according to the initial sample dilution factor, the varying sample input volumes and the different washing volumes to load a similar amount of protein material per employed strategy, except for the UF6 wash fractions which contain 2.5 times more protein material compared to those of UC and UF96. As albumin is by far the most abundant bovine serum protein, we focused on the 65 kDa molecular weight (MW) bands, representing bovine albumin.

Figure 5 illustrates that similar amounts of protein material are removed during the first flow-through or ultracentrifugation step. For UC, the most protein material is present in the 0.1% TWEEN-20 wash condition with the least protein material present in the 0.0% TWEEN-20 condition (PBS-only). For UF6, the first wash fractions contain similar amounts of protein material, except for the 0.0% TWEEN-20, which has a slightly higher amount of protein material present. For UF96, the first wash fractions show more variability for the different TWEEN-20 concentrations with the least protein material present in 0.1% and 0.5% TWEEN-20. However, the performance of the UF96 protocol was suboptimal due to a vacuum suction force of the protective foil, resulting in more variability than expected. During the third washing steps, less protein material is washed away in both UF6 and UF96 with a slight increase in protein material in the higher TWEEN-20 concentrations (1% and 5% TWEEN-20) for both techniques.

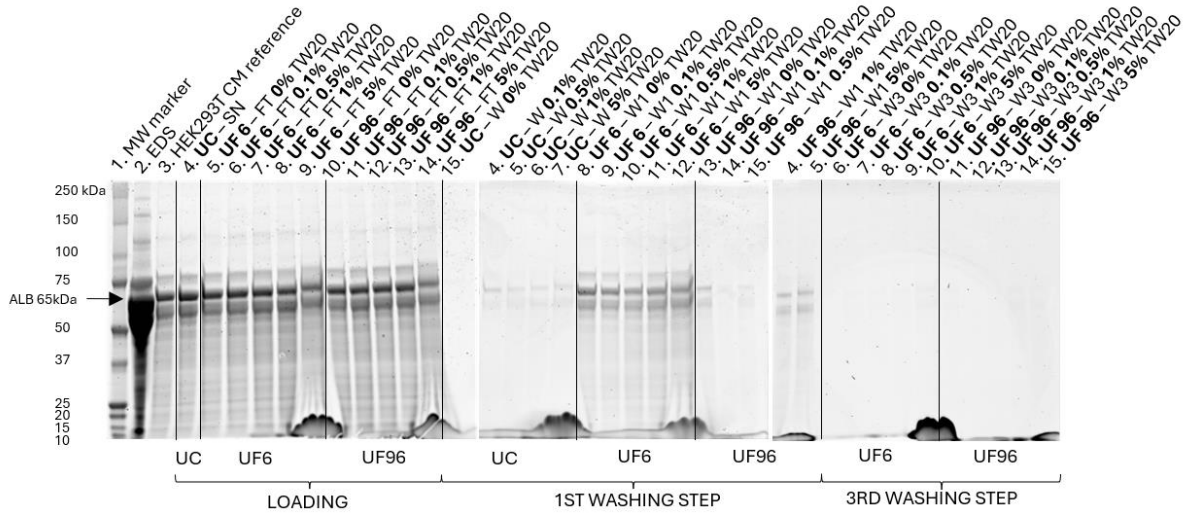


Figure 5: SDS-PAGE with Coomassie staining of the different fractions (supernatans (SN), flow-through (FT), (first and third) washing step (W(1,3))) obtained during ultracentrifugation (UC), FAEVEr on 6 mL 300 kDa MWCO filter tubes (UF6) and FAEVEr on a 300 kDa MWCO 96-well filter plate (UF96) with TWEEN-20 (TW20, 0.0% - 0.1% - 0.5% - 1.0% - 5.0%) in PBS. EV-depleted fetal bovine serum (EDS) and conditioned medium (CM) of HEK293T cells are used as negative and positive control, respectively. Albumin (ALB) is indicated with a black arrow at approximately 65 kDa.

In conclusion, all three EV-isolation techniques succeed in removing contaminating protein material and TWEEN-20 slightly improves the washing efficiency during UC. In UF6 and UF96, even though PBS seems to remove slightly more non-EV protein material during the first washing step, higher TWEEN-20 concentrations (1%-5%) increase the removal of protein material during subsequent washing steps. Lastly, UF96 shows the highest variability in its washing efficiency and requires further optimization. For this reason, we repeated the UF96 protocol including puncturing of the protective foil covering the 96-well plate to prevent a vacuum suction force, leading to a better filtration performance of the UF96 protocol. Unfortunately, the washing efficiency was not visualized with a Coomassie staining.

7.1.3 Minimal loss and efficient enrichment of EVs

We validated if TWEEN-20 leads to an undesired loss of rEV material by WB using antibodies against Gag-eGFP. Since Gag-eGFP multimerizes at the luminal side of the vesicles, it should not be present in the filtrate or supernatant of the loading steps, nor in any of the washing steps, as long as the EV membrane remains intact. On UF6, a volume of 2.5 mL rEV input material was diluted 1:1 with either 0%, 0.1% or 5% TWEEN-20 in PBS and enriched. As a reference, we included 1 mL undiluted rEV input material enriched by UC. The individual fractions were analyzed by WB. We observed a clear lack of signal, demonstrating the absence of Gag-eGFP (84 kDa) in the flow-through or supernatant and in the wash fractions (Figure 6). We concluded that for both UF6 and UC no rEV material is lost during EV enrichment. In addition, it shows that rEVs remain intact when TWEEN-20 is added in concentrations up to 5%. Conversely, in the lysis fractions of both UC and UF6 samples, an abundant presence of green fluorescent signal at 84 kDa depicts the successful lysis with the release of luminal Gag-eGFP. Remarkably, even though no rEVs seem to be lost in both techniques and, for UC, 2.5 times less protein material is loaded on the gel, the lysates of UC show a higher fluorescent signal intensity compared to the UF6 lysates. Although only the fractions of 0%, 0.1% and 5% TWEEN-20 in PBS are visualized, we expect similar results for intermediate TWEEN-20 concentrations.

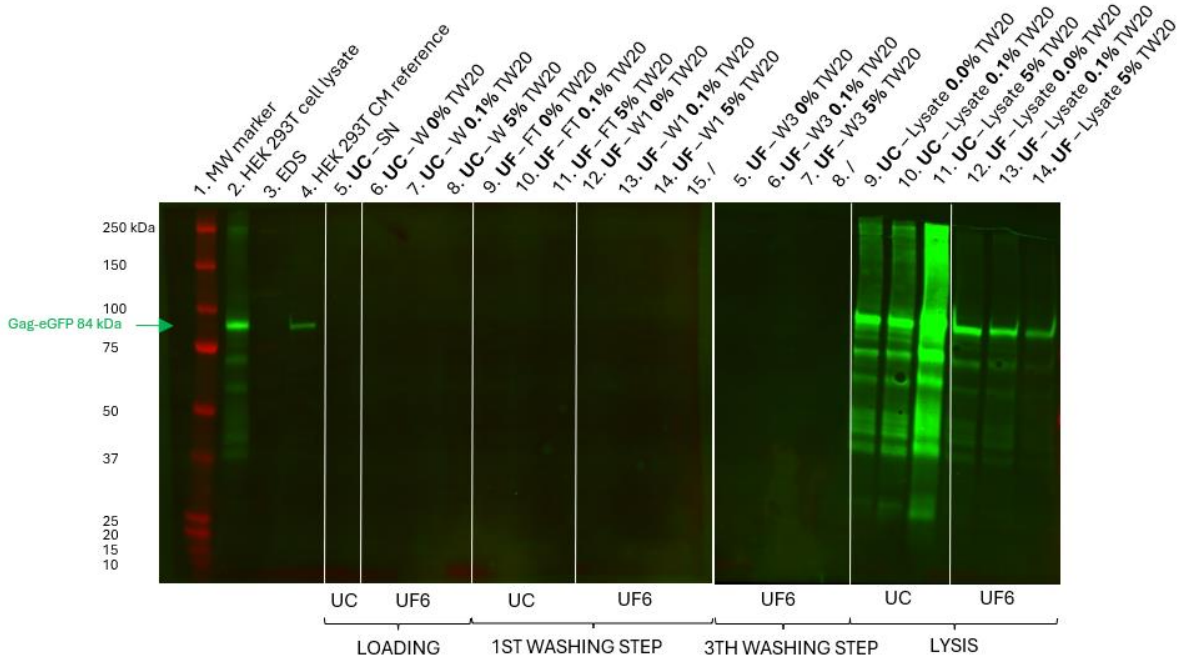


Figure 6: Western blot of the different fractions (supernatans (SN), flow-through (FT), (first and third) washing step (W(1,3)), Lysate) obtained during ultracentrifugation (UC) and FAEVEr on 6 mL 300 kDa MWCO PES membrane filter tubes (UF) with TWEEN-20 (TW20, 0.0%-0.1%-5.0%) in PBS. EV-depleted fetal bovine serum (EDS) and conditioned medium (CM) of HEK293T cells are used as negative and positive control, respectively. Gag-eGFP is indicated with a green arrow at approximately 84 kDa.

FAEVEr on a 96-well filter plate in combination with 0.5% TWEEN-20 also showed to enrich rEVs efficiently, as illustrated with WB in Figure 7. Of note, this considers the UF96 experiment in which the covering foil was punctured resulting in an optimal flow-through. Similar to UC and UF6, no fluorescent signal is detected in the supernatant and flow-through fractions. However, contrary to UC and UF6, the fluorescent signal at 800 nm (green channel) at 84 kDa is observed in the lysate fractions, as well as in the wash fractions, albeit with observable lower intensity. This suggests that some rEV material may be lost during the isolation. However, the luminal Gag-eGFP is notably enriched in the lysate fractions. In addition, fluorescent signal at 680 nm (red channel) is visualized at 75 kDa and 32 kDa, corresponding to calnexin and the EV-specific marker syntenin-1 (SDCBP), respectively. Compared to the green fluorescent signal associated with Gag-eGFP, the red signal related to SDCBP is less intense in both the wash and lysate fractions. Thus, the intensities of these red and green fluorescent signals do not correlate, even though both should represent the rEV material. This disparity in intensity between red and green fluorescent signals may be attributed to a higher abundance of Gag-eGFP in rEVs compared to syntenin-1. In contrast to Gag-eGFP, SDCBP is slightly more intensely visualized in the first wash fractions compared to the lysate fractions. This could be due to SDCBP's presence on the surface of the rEVs rather than in the lumen. Therefore, it might be that its epitope is sheared from the rEV surface leading to lower intensity of its signal in the lysate fractions compared to the wash fractions.

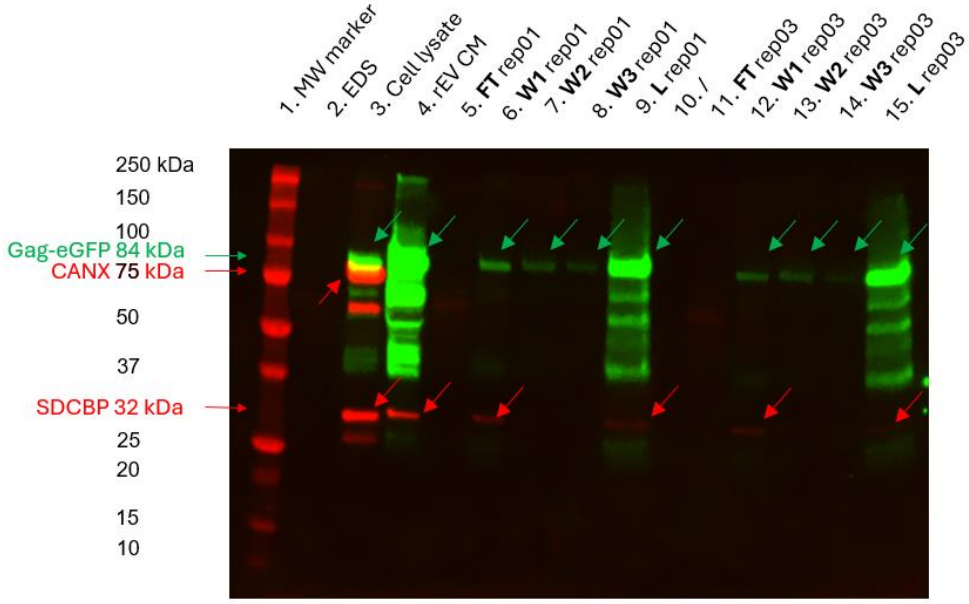


Figure 7: Western blot of the different fractions (flow-through (FT), first, second and third washing step (W1,2,3) and lysate (L)) obtained during FAEVEr on a 96-well 300 kDa MWCO PES membrane filter plate (UF96) with 0.5% TWEEN-20 in PBS. EV-depleted fetal bovine serum (EDS) and conditioned medium (CM) of HEK293T cells are used as negative and positive control, respectively. Gag-eGFP is indicated with green arrows at approximately 84 kDa. Calnexin (CANX) and syntenin-1 (SDCBP) are indicated with red arrows at approximately 75 kDa and 32 kDa, respectively.

7.1.4 Recombinant EVs are retained on the filter and can be recovered without jeopardizing their integrity

SEM was performed to evaluate retention of EVs on the Nanosep centrifugal 300 kDa MWCO PES membrane filter (500 µL sample volume) and the potential impact of FAEVEr on their morphology[24]. The SEM images show an empty filter as negative control, successful retention of rEVs using FAEVEr with 5% TWEEN-20 or with PBS, and successful retention of 100 nm sized nylon beads that are also used for calibration of the Zetaview device for NTA (Figure 8). The morphology of the rEVs remained intact.

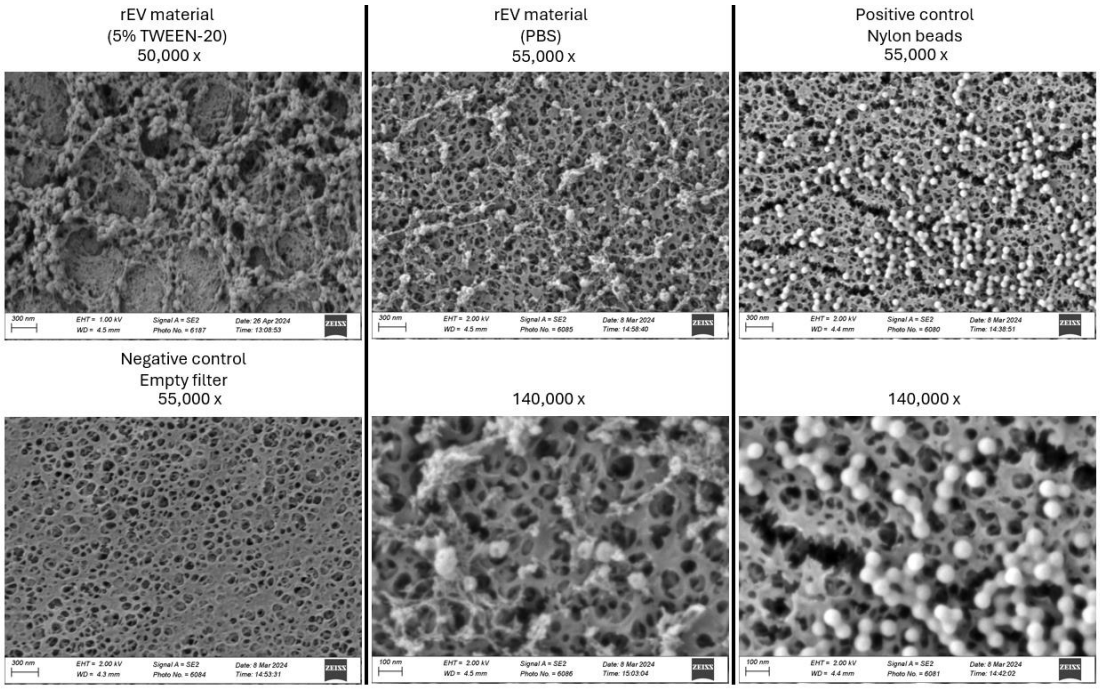


Figure 8: Scanning electron microscopy images at a magnitude of 50,000 times, 55,000 times and 140,000 times are shown. An empty nanosep filter (bottom left), retained recombinant extracellular vesicles (rEVs) enriched using FAEVEr with either 5% TWEEN-20 or PBS (upper left and middle) and retained 100 nm sized nylon beads (right) are illustrated. The microscopy settings such as accelerating voltage of the electron beam (EHT), working distance (WD) and the signal of secondary electrons and the scale of each image are indicated on the bottom in a white balk.

TEM images further validated the ability to recover EVs from 300 kDa MWCO PES membrane filters, after washing with both PBS and 5% TWEEN-20 (Figure 9). In FAEVEr with PBS wash steps, nanoparticles of different sizes are illustrated with smaller particles of about 20 nm in size that appear to be white, presumably due to a high-lipid composition, potentially indicating the presence of lipid-rich HDLs. In 5% TWEEN-20, a highly contrasted vesicle of approximately 200 nm in size is visualized. Even though Nanosep centrifugal 300 kDa MWCO PES membrane filters with 500 µL sample volume were used, these concepts can be extrapolated to the 6 mL filter tubes and 96-well filter plate that also contain 300 kDa MWCO PES membrane filters.

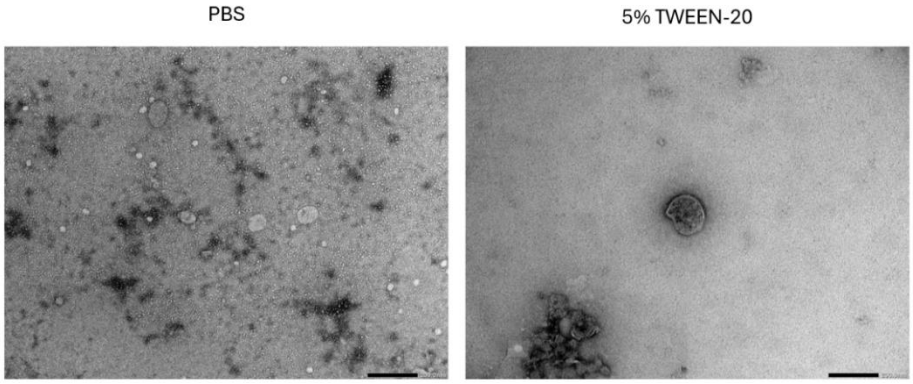


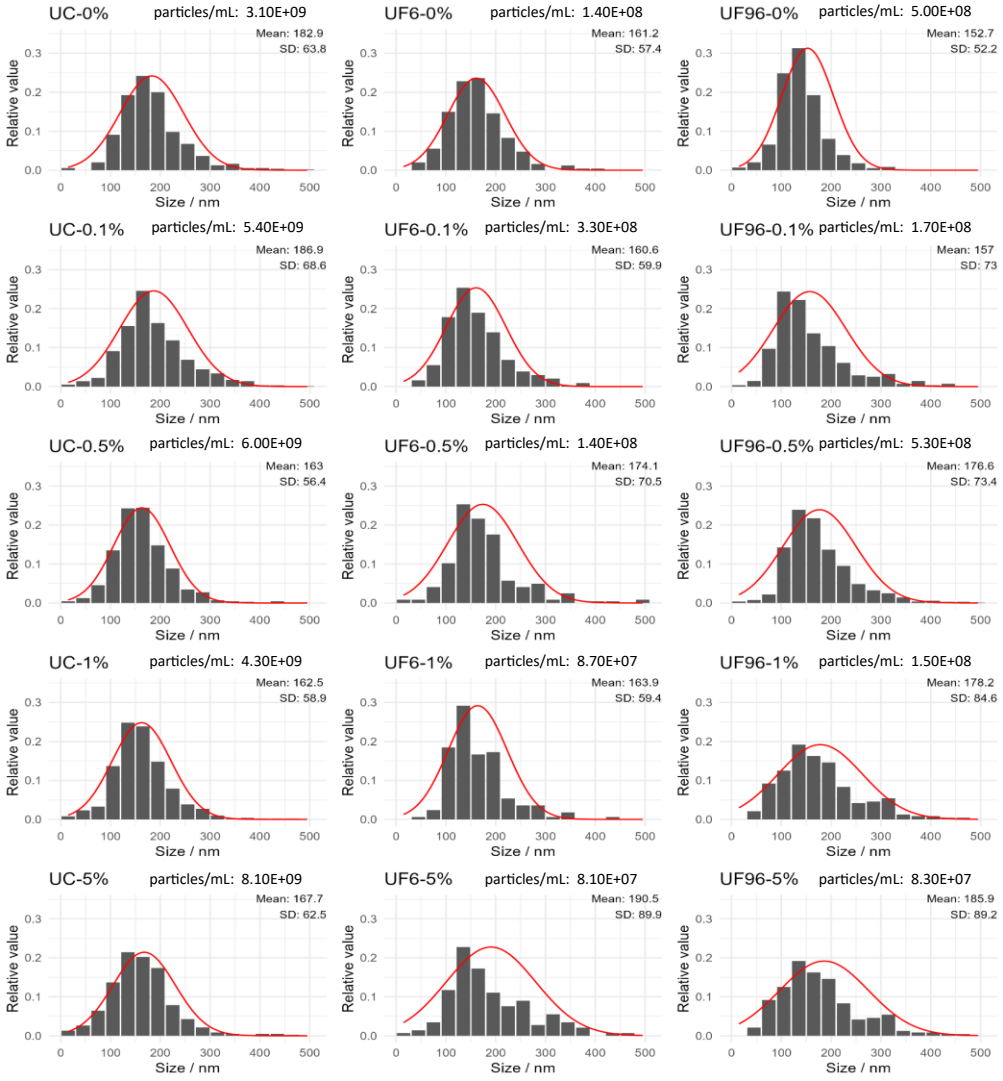
Figure 9: Transmission electron microscopy images at 80 kV, magnified up to 30,000 times (scale indicates 200 nm), of recovered recombinant extracellular vesicles from HEK293T cell medium supplemented with 10% EV-depleted FBS and subjected to FAEVEr with either PBS or 5% TWEEN-20.

In conclusion, we gathered visual evidence that rEV material is successfully retained on the 300 kDa MWCO filter and can be recovered while the vesicle integrity is preserved after multiple washing steps including 5% TWEEN-20.

7.1.5 TWEEN-20 does not consistently affect the concentration and size-based characterization of isolated nanoparticle populations for FAEVEr and UC

For FAEVEr, the EV cargo can be collected after on-filter lysis or the retained intact EVs can be recovered from the filter. To get an indication of the concentration and the size of the isolated particle subpopulation, an NTA was performed on one replicate per isolation technique with one concentration of TWEEN-20. Of note, the results obtained with UF96 were generated with the suboptimal flow-through. The particle count and size distribution are shown in Table S4 and Figure 10. These results suggest that in all three isolation methods, both the recovered particle concentration and the particle size distribution do not follow a linear trend for increasing TWEEN-20 concentrations. Of note, prior findings suggest that the recovery of nanoparticles from the membrane filter is not complete[24], which could potentially influence the NTA results for FAEVEr-processed samples.

Figure 10: Particle concentration in particles/mL and size distribution of isolated nanoparticles from one



recovered sample replicate per experimental condition are displayed. The ultracentrifugation (UC), FAEVEr on 6 mL capacity (UF6) and on 96-well capacity (UF96) samples are represented on the left, middle and right respectively, following increasing TWEEN-20 concentrations (0% - 0.1% - 0.5% - 1.0% - 5.0%) from top to bottom. The relative number of particles is presented in a bar plot in function of the particle size in nm. The corresponding Gaussian distributions for the given values of the mean and standard deviation are indicated in red.

7.1.6 Proteome comparison

To enhance the identification of EV-specific protein biomarkers using bottom-up DIA LC-MS/MS, we aimed to optimize the FAEVEr TWEEN-20 setup for EV isolation. Besides the assessment of general isolation aspects discussed above, we compared the proteomes of rEV-enriched fractions from UC (1 mL), UF6 (2.5 mL), and UF96 (600 µL). Again, different percentages of TWEEN-20 supplemented to the wash buffer were tested (0.0% - 0.1% - 0.5% - 1.0% - 5.0%). Here, we utilized the results of the repeated UF96 protocol on 600 µL samples for proteome comparison. Since this protocol was executed separately, it additionally included a 2.5% TWEEN-20 concentration in the range of TWEEN-20 concentrations.

a) Differential impact of isolation techniques and TWEEN-20 on recombinant EV proteomic profiles

The proteomic profiles of the lysed rEV-enriched fractions exhibit variations across the different isolation methods employed (UC, UF6, UF96), as illustrated in a principal component analysis (PCA, Figure 11). The samples tend to cluster according to the isolation technique employed, with UC and UF6 displaying closer association compared to UF96. Interestingly, TWEEN-20 demonstrates substantial effects on the rEV proteomes in UF6 and UF96 but negligible effect in UC. More specifically, all UC enriched samples cluster together, whereas in UF6 and UF96, samples without TWEEN-20 are segregated from the samples with TWEEN-20. More particularly, this variability introduced by TWEEN-20 is more pronounced for UF96 compared to UF6. Remarkably, in UF96, samples with 0.1% TWEEN-20 cluster between higher TWEEN-20 concentrations and those without TWEEN-20. To investigate the cause of the variability introduced by TWEEN-20 in more depth, we explored the origin of the identified proteins.

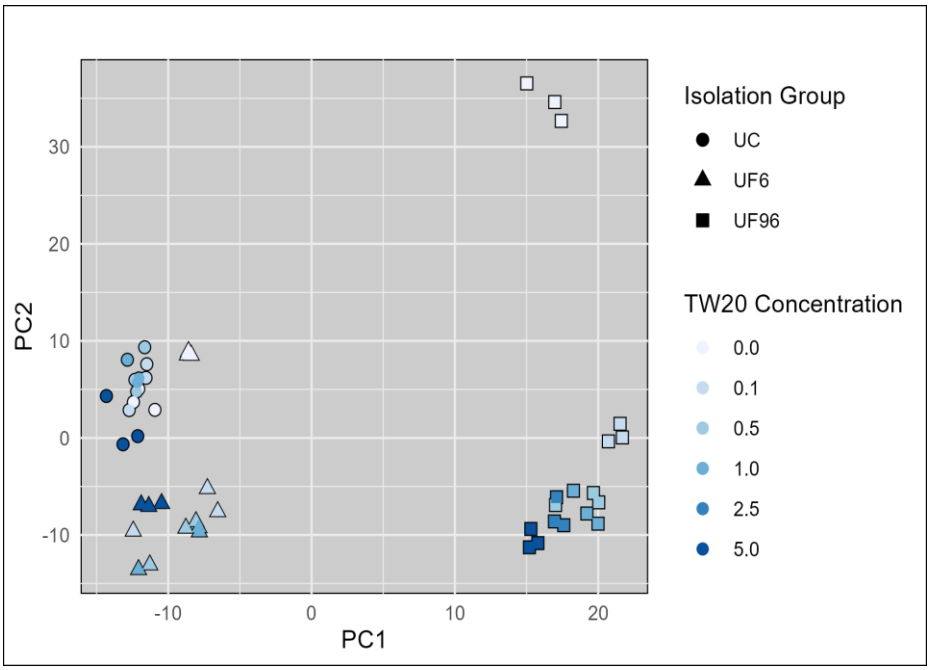


Figure 11: Principle component analysis (PCA) of the log2 transformed LFQ values of all triplicates per experimental condition. The isolation groups are indicated by their filled shapes: a circle, triangle and square for ultracentrifugation (UC), FAEVEr on 6 mL filter tubes (UF6) and on 96-well filter plate (UF96) respectively, whereas the TWEEN-20 concentration groups (0% - 0.1% - 0.5% - 1.0% - 5.0%) are indicated by a color gradient following increasing TWEEN-20 concentration.

b) Isolation-based impact: UC yields more protein identifications while UF96 yields more reproducible protein identifications

Generally, UC-based EV isolation yields the highest absolute number of human and bovine protein identifications over all the different TWEEN-20 experimental conditions (Figure 12**A**). However, UF6-based EV isolation with PBS yields the maximum number of human and bovine identifications observed in this experiment (averaging over 4,000 identifications (Figure 12**A**)). When comparing PBS with or without addition of TWEEN-20, we found that the addition of 0.5% to 5.0% TWEEN-20 during UC delivered a slightly lower average number of human protein identifications compared to 0.0% and 0.1% TWEEN-20. For UF6 and UF96, the absolute number of human and bovine protein identifications follows a decreasing parabolic trend with increasing TWEEN-20 concentrations, reaching a minimum at 1.0% TWEEN-20.

Remarkably, UF6 samples exhibit the most variability in the number of human protein identifications, whereas the UF96 samples show the least variability, as illustrated by the coefficient of variation plotted in Figure 12**B**. These differences in variation could be attributed to the discrepancy in the filter surface area, which is smaller for the filter-containing wells compared to the filter-containing tubes. These variations, ranging from large to small on average for UF6, UC and UF96, respectively, are also observed in the number of unique peptides, as depicted in Figure 12**C**. Since the coefficient of variation is smallest for UF96, both the absolute number of identifications as well as the number of identified human unique peptides are most reproducible for this technique, thereby suggesting more accurate protein quantification. Noteworthy, in both UF6 and UF96, the coefficient of variation is the smallest for 5% TWEEN-20 compared to the other TWEEN-20 concentrations.

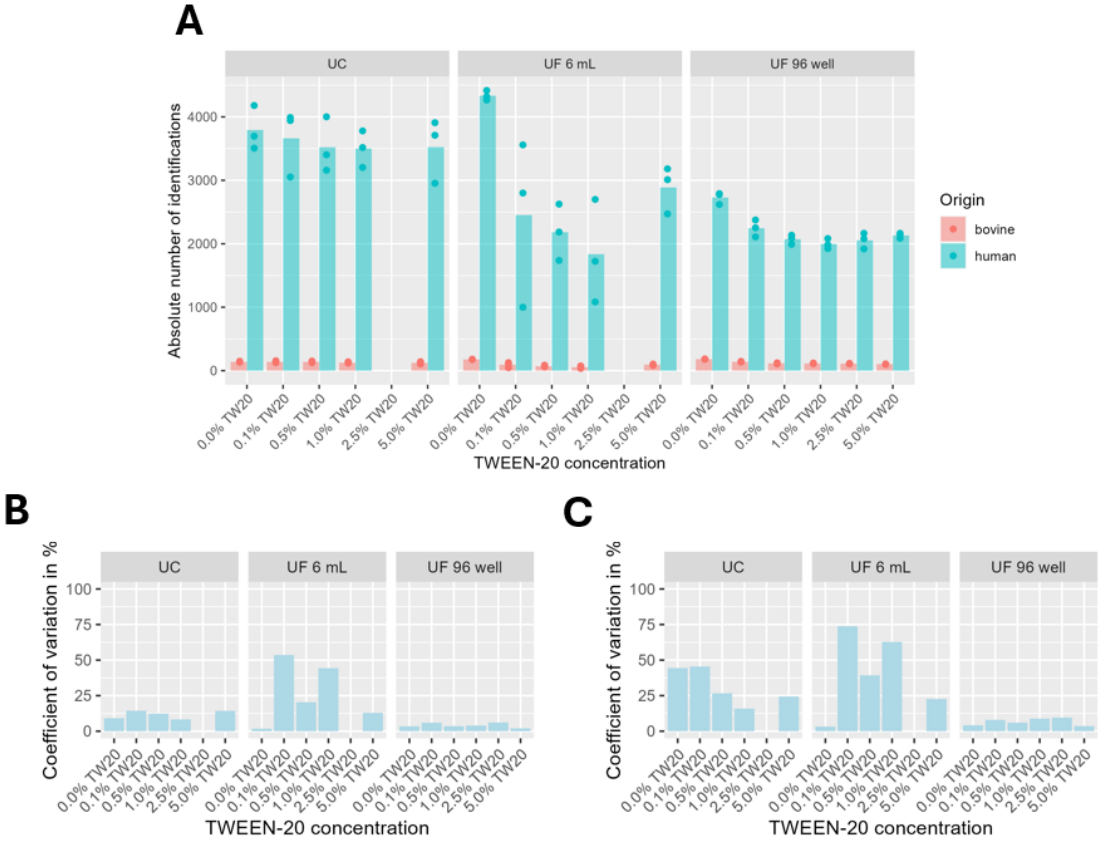


*Figure 12: **A)** Representation of the absolute number of identified proteins (human in blue and bovine in red) with data independent acquired (DIA) liquid chromatography tandem (LC-MS/MS) analysis per TWEEN-20 concentration (0% - 0.1% - 0.5% - 1.0% - 2.5% - 5.0%) per EV enrichment method: ultracentrifugation (UC), FAEVEr on 6 mL (UF 6 mL) or 96-well (UF 96 well) capacity. **B)** Coefficient of variation (in %) of the average number of identified human proteins per TWEEN-20 concentration perenrichment strategy. **C)** Coefficient of variation (in %) of the average number of detected unique peptides from human proteins and Gag-eGFP per TWEEN-20 concentration per enrichment strategy.*

c) TWEEN-20 increases the proportion of human protein identifications and the percentage of transmembrane proteins

The differential impact of TWEEN-20 across the three isolation methods is highlighted by the relative number of human and bovine identifications in Figure S2. Its addition shows no significant effect for UC, while UF6 and UF96, exhibit a linear decrease and increase in the percentage of bovine and human protein identifications, respectively (Figure 13). This linear trend is more pronounced for UF96. Moreover, examining the gene ontology of the identified human proteins, categorized in four subcellular location groups (transmembrane proteins, intracellular proteins, secreted proteins and other), reveals a non-linear increase in the percentage of transmembrane protein identifications upon addition of TWEEN-20 (Figure 16). This increase is also more pronounced for UF96 with a subtle in-between step at 0.1% TWEEN-20. Also, this increase in transmembrane protein identifications is out-balanced by a decrease in secreted and intracellular protein identifications. So, despite fewer proteins being identified through UF with TWEEN-20 compared to UC, a higher proportion of these identified proteins are of interest. Firstly, human proteins, originating from HEK293T cell-derived rEVs are prioritized over bovine proteins, originating from EDS. Secondly, among these human EV proteins, transmembrane proteins hold significant biomarker potential due to their accessibility to affinity reagents. Conversely, in UF6 and UF96, a lower proportion of the identified proteins are human secreted proteins that are also considered to contribute to the contaminating potential of the EV-embedding matrix.

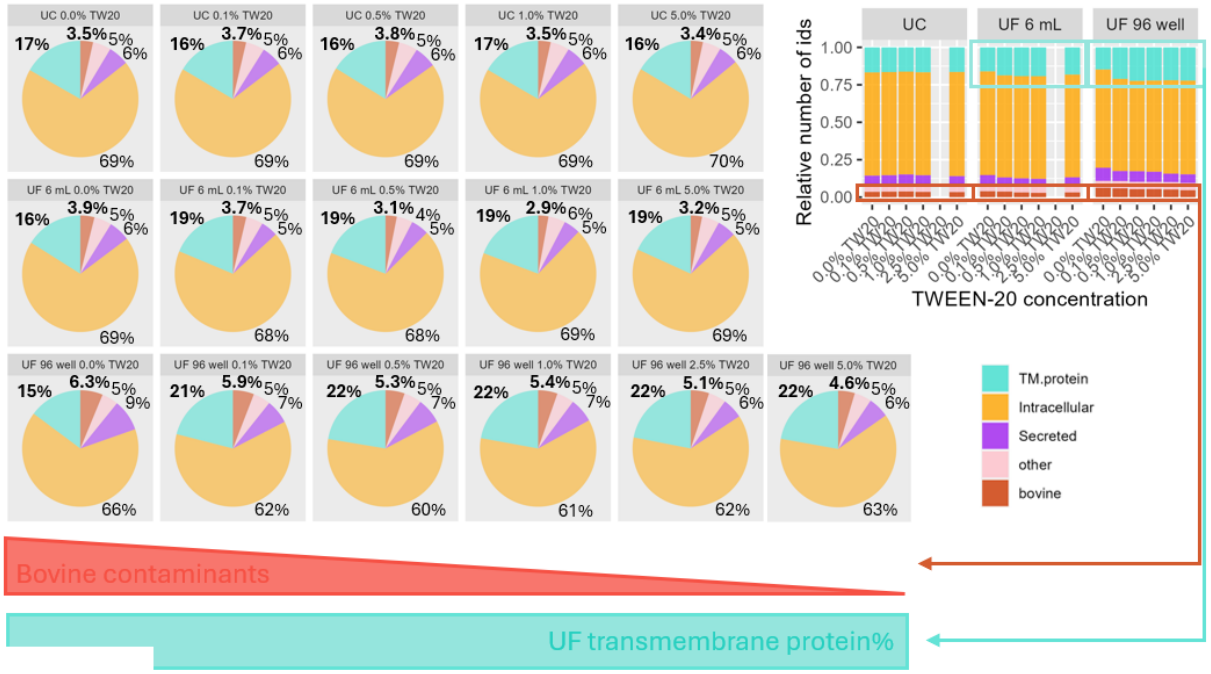


Figure 13: Visualization of the relative number of identified proteins in the EV-enriched fractions from ultracentrifugation (UC), FAEVEr on 6 mL capacity (UF6) or on 96-well capacity (UF96) in bar plots per TWEEN-20 concentration (0% - 0.1% - 0.5% - 1.0% - 2.5% - 5.0%). The fraction of identified human proteins is divided into four subcellular location groups: transmembrane (TM protein), intracellular, secreted and other proteins. The fraction of identified bovine proteins is colored in red and is represented by two decimals.

d) TWEEN-20 improves the spectral space occupancy for proteins of interest

Although the number of human and bovine proteins offers valuable insight into sample purity, analyzing the spectral space, which considers both the number and intensity of peptide-spectral matches, provides a more qualitative assessment. By plotting the protein precursor quantities relatively to the sum of the precursor quantities per sample, the relative spectral space can be examined. Figure 14 depicts the proportion of the spectral space occupied by bovine and non-bovine (human and Gag-eGFP) protein precursors. This spectral space

occupancy is one of the determinants of protein discovery by bottom-up mass spectrometry. Highly abundant contaminating proteins such as bovine albumin from EDS can generate more detectable peptides thereby increasing spectral space occupancy, yet resulting in only a few protein identifications.

In the case of UC, the addition of lower TWEEN-20 concentrations (0%-0.1%-0.5%) appears to have minimal effects on the relative spectral space occupancy. However, higher TWEEN-20 concentrations (1%-5%) slightly elevate the fraction of human precursors occupying the spectral space. In UF6, the addition of any TWEEN-20 concentration increases the proportion of the spectral space occupied by human precursors by approximately 15-20%, reaching a plateau at 5% TWEEN-20. In UF96, the relative occupancy by human precursors shows a more significant and linear increase with the addition of higher TWEEN-20 concentrations. Notably, higher concentrations of TWEEN-20 result in lower fractions of the spectral space occupied by bovine precursors, thus raising the fraction of the spectral space occupied by relevant human precursors. So, the average gain in spectral space proportion occupied by non-bovine proteins (human and Gag-eGFP) through the addition of any concentration of TWEEN-20 compared to PBS only, is the largest for UF96, followed by UF6 and is the smallest for UC.

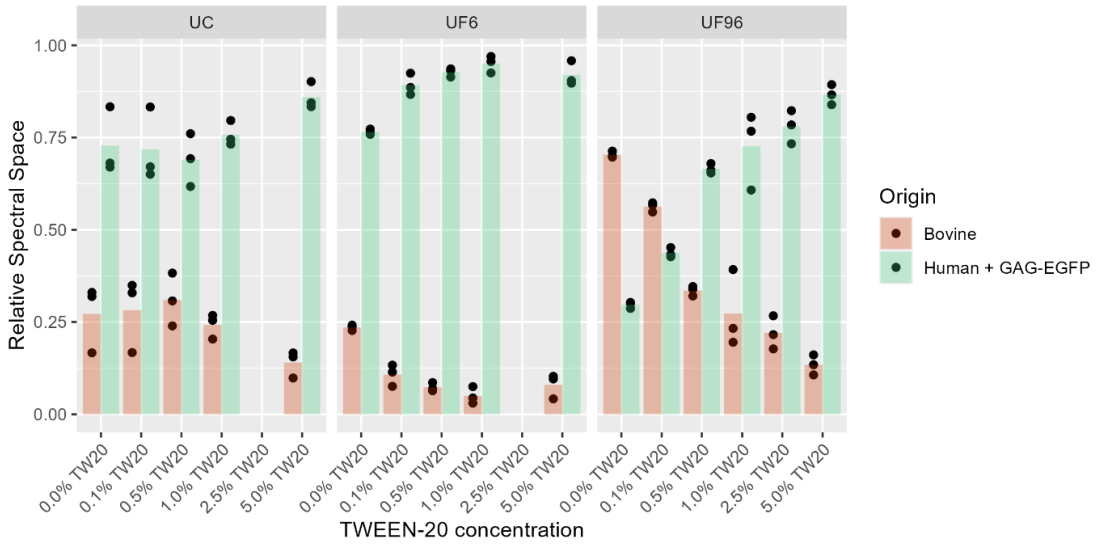


Figure 14: The proportion of the spectral space occupied by bovine (red) or non-bovine (human and Gag-eGFP, green) derived precursors is demonstrated per TWEEN-20 concentration (0% - 0.1% - 0.5% - 1.0% - 2.5% - 5.0%) per EV enrichment strategy: ultracentrifugation (UC), FAEVEr on 6 mL capacity (UF6) or on 96-well capacity (UF96).

Moreover, human protein precursors were categorized into four subcellular location groups each occupying a different fraction of the spectral space, as demonstrated in Figure 15. This figure illustrates that, for UC, primarily the proportion of intracellular protein occupancy in the spectral space increases. Conversely, for UF6, the proportion of transmembrane protein occupancy increases, accompanied by a decrease in the proportion of secreted protein occupancy in the spectral space. For UF96, all subcellular location protein fractions in the spectral space occupancy show an increase. This can be attributed to the larger gain in the proportion of the spectral space occupied by human precursors compared to UF6.

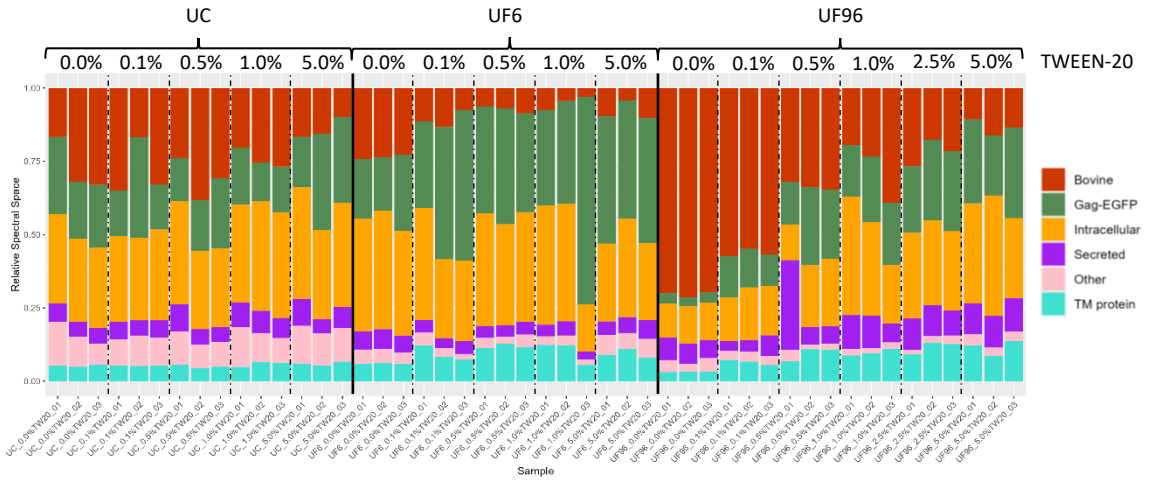


Figure 15: The proportion of the spectral space occupied by bovine (red), Gag-eGFP (dark green) or human derived precursors is demonstrated per TWEEN-20 concentration (0% - 0.1% - 0.5% - 1.0% - 2.5% - 5.0%) per EV enrichment strategy: ultracentrifugation (UC), FAEVEr on 6 mL capacity (UF6) or on 96-well capacity (UF96). The human precursors are further divided related to their subcellular location being transmembrane (TM protein), intracellular, secreted or anything other.

This differential effect of TWEEN-20 on the spectral space occupancy for the three isolation methods is also elucidated by the volcano plots in Figure 16. The higher the impact of TWEEN-20, the greater the number of human proteins that are significantly differentially present in the TWEEN-20 conditions compared to samples with PBS only (Figure 16**A**). In contrast, significant bovine proteins are differentially more abundant towards the PBS only condition, including apolipoprotein A1 (APOA1). Moreover, using UF6 and UF96 in combination with TWEEN-20, significantly differentially present human proteins are predominantly enriched in transmembrane proteins towards TWEEN-20 (Figure 16**B**), in line with the relative spectral space. For UF96, a few secreted and intracellular proteins are also more significantly differentially abundant towards TWEEN-20, reflecting the effect size of TWEEN-20 on the proportional spectral space occupancy.

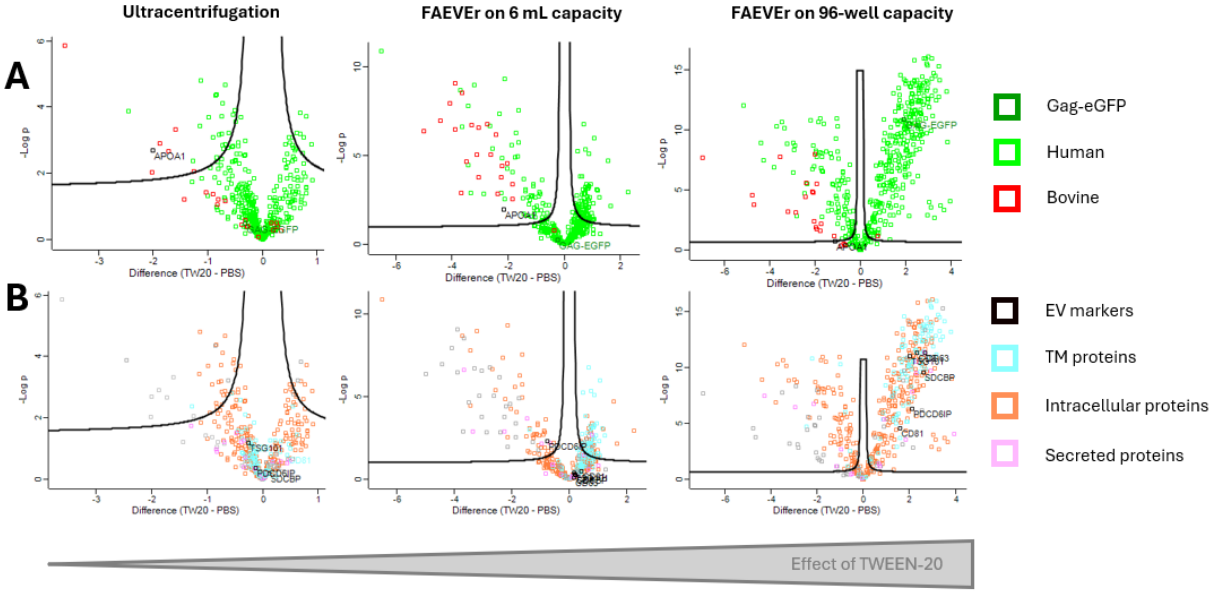


*Figure 16: Volcano plots (FDR = 0.05, S0 = 0.1) of all valid values of the PBS-only condition (left) compared to grouped TWEEN-20 (TW20) conditions (right) per EV enrichment strategy. In the upper panels (**A**), proteins are indicated as bovine (red, with APOA1 in black), human (light green) or Gag-eGFP (dark green). In the lower panels (**B**), proteins are indicated as EV markers (labelled in black), transmembrane proteins (TM proteins in blue), intracellular proteins (orange) or secreted proteins (purple).*

e) TWEEN-20 facilitates enrichment of extracellular vesicle-specific markers in UF96

Notably, these volcano plots (Figure 16) demonstrate that, in UF96 ,TWEEN-20 aids in the enrichment of EV-specific markers (CD9, CD63, CD81, TSG101, SDCBP, PDCD6IP) and rEV-related luminal Gag-eGFP compared to PBS. Gag-eGFP and the specific EV-markers show strong significantly differential presence towards TWEEN-20 conditions compared to PBS, as additionally illustrated by the EV-marker profile plot (Figure 17). Alternatively, in UC and UF6, all EV markers and Gag-eGFP are also detected, with similar relative abundance regardless of the presence or absence of TWEEN-20 (Figures 16 and 17). Thus, FAEVEr emerges as a valid approach for EV enrichment since it offers comparable outcomes to UC. However, in UF96, the inclusion of TWEEN-20 becomes imperative to achieve similar levels of rEV-specific protein markers as observed in UC and UF6.

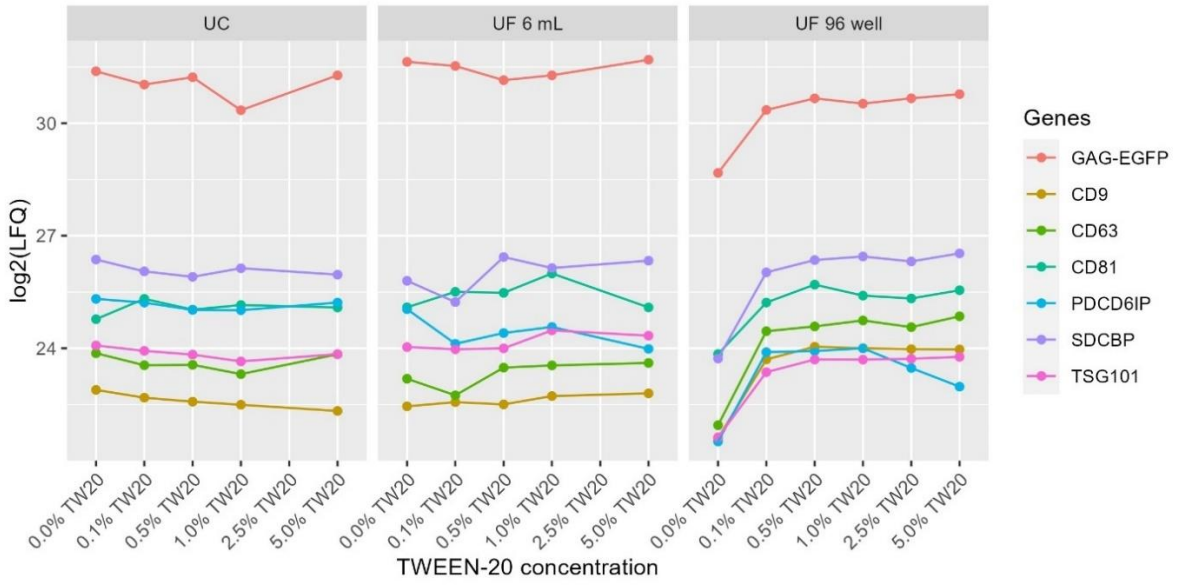


Figure 17: Profile plot of the log2 transformed LFQ-values for the EV markers CD9, CD63, CD81, PDCD6IP, SDCBP and TSG101 and for Gag-eGFP averaged per TWEEN-20 concentration for each enrichment strategy: ultracentrifugation (UC), FAEVEr on 6 mL filter tubes (UF6) and on 96-well filter plate (UF96).

f) Differential gene ontology enrichment of protein profiles from EV enriched sample fractions for UC and FAEVEr with TWEEN-20

Hierarchical clustering of the proteins detected in all samples results in four distinct sample clusters: UF96 with TWEEN-20, UF6 with TWEEN-20, UF96 and UF6 with PBS only and UC with both TWEEN-20 and PBS (Figure 18). This clustering reaffirms the clear separation between PBS and TWEEN-20 conditions for UF6 and UF96, while indicating the negligible impact of TWEEN-20 on sample clustering for UC. Moreover, together with gene ontology analysis in STRING, it underscores the impact of the different isolation techniques and their combination with or without TWEEN-20 on the proteomic profiles of the isolated EV fractions. The hierarchical protein clusters were manually divided in four sub-clusters (Figure 18). The proteins in cluster 1 (gold) are notably enriched in UF96 with PBS only, but less in UF6 with 5% TWEEN-20 and in UC. This cluster is enriched in extracellular and secreted proteins such as lipoprotein particles and secretory granule-associated proteins. Cluster 2 (pink) appears to be enriched in proteins associated with the extracellular space like blood microparticles, secretory granule-related proteins and lipoprotein particles and these proteins are slightly enriched in UC and in FAEVEr with PBS only, as well as in UF96 with 0.1% TWEEN-20. In addition, cluster 2 also includes terms such as extrinsic components of the cell membrane and cell surface-related membrane proteins and also cytosolic proteins associated with the proteasome and ribosomes. Cluster 3 (blue) contains mainly cytosolic proteins associated with

the ribonucleoprotein complexes, ribosomes, proteasome and the nucleus, along with membrane proteins related to cell and anchoring junctions and these proteins are relatively more abundant in UF6 and UC compared to UF96. In cluster 4 (light pink), for UF96 with PBS only, secretory granule-related proteins, plasma membrane proteins related to cell junction and cell projection on the one hand and intrinsic plasma membrane components on the other hand, as well as endomembrane system-related proteins including ESCRT- and multivesicular body (MVB)- related proteins are visibly less abundant. This observation is consistent with the findings of the EV-marker profile plots and volcano plots, where many transmembrane proteins and the EV-specific markers showed significant differentially abundance towards UF96 samples with TWEEN-20. Interestingly, the ESCRT-related, MVB-related and intrinsic plasma membrane proteins are part of cluster 4 (below the dashed line in Figure 18). So, these particular proteins of interest are slightly more present in UF6 combined with TWEEN-20, even more so in UF96 combined with TWEEN-20 and are relatively less abundant in FAEVEr with PBS only and in UC.

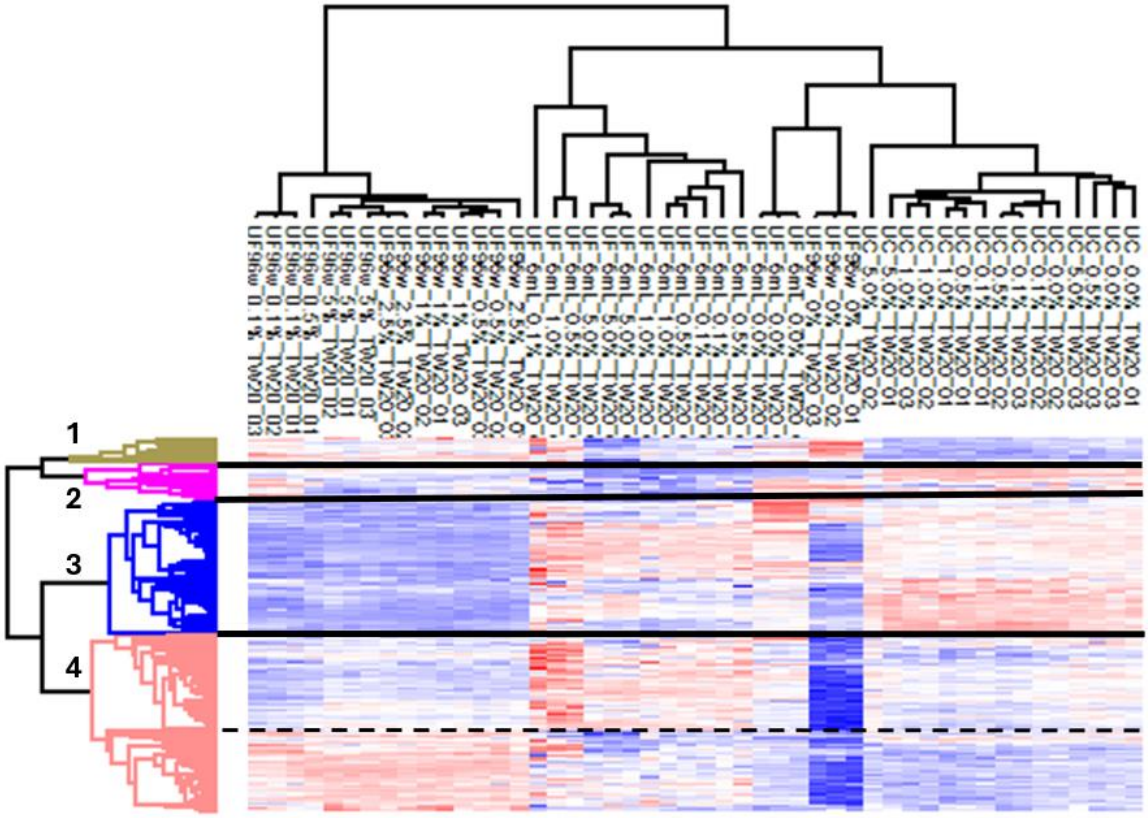


Figure 18: Hierarchical clustering of all valid values shows four sample clusters, from left to right: FAEVEr 96-well capacity with TWEEN-20, FAEVEr 6 mL capacity with TWEEN-20, FAEVEr with PBS only and UC. Moreover, four protein clusters, colored in gold (cluster 1), pink (cluster 2), blue (cluster 3) and light pink (cluster 4) from top to bottom are defined by black lines. The light pink cluster is further divided into two subclusters by a dashed black line. The highest z-scores are colored in red and lowest in blue.

To elucidate the effect of TWEEN-20 on FAEVEr further, samples subjected to UF6 or UF96 were hierarchically clustered (Figure 19). In both techniques, clustering analysis reveals distinguishable clusters between samples with and without TWEEN-20. However, the impact of high TWEEN-20 concentrations is more pronounced in UF6, with 5% TWEEN-20 forming an additional sample cluster. Conversely, in UF96, the effect of 5% TWEEN-20 is less obvious, since it clusters more closely with lower TWEEN-20 concentrations. Furthermore, it is noteworthy that, in UF96, small amounts of TWEEN-20 demonstrate a more pronounced and consistent effect on the protein clusters of the samples, contrasting with UF6 where the response to TWEEN-20 concentration variations appears less coherent. This differential

sensitivity to TWEEN-20 between UF6 and UF96 techniques suggests that UF6 exhibits a clearer impact on the proteome of the isolated EV-enriched fraction in response to extremities of TWEEN-20 concentration, with a greater response variation for different TWEEN-20 concentrations. In contrast, UF96 shows more consistent effects, already at small TWEEN-20 concentrations. Moreover, in UF96, 0.1% TWEEN-20 seems to fit the protein level patterns of both sample clusters with and without TWEEN-20, suggesting it could be considered a transitional concentration step. These findings are consistent with the PCA cluster, volcano plots and variation observed in the numbers of identified proteins and unique peptides.

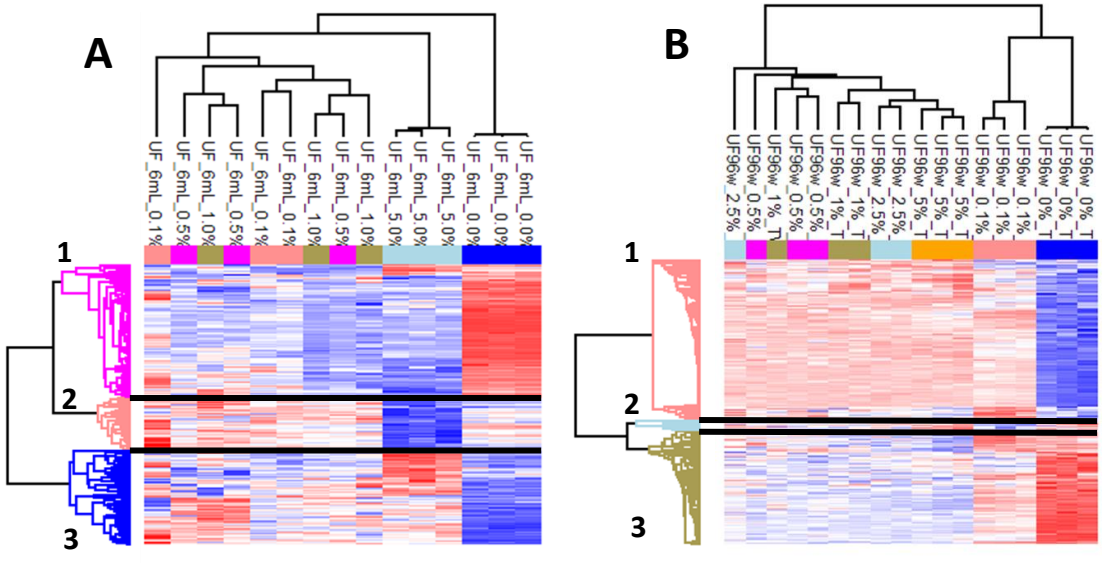


*Figure 19: Hierarchical clustering of all valid values shows three main sample clusters for each FAEVEr format. On the left figure (**A**), from left to right: FAEVEr on 6 mL capacity with mediate TWEEN-20 concentrations (0.1%-0.5%-1%), FAEVEr on 6 mL capacity with 5% TWEEN-20 and FAEVEr on 6 mL capacity with PBS only. On the right figure (**B**), from left to right: FAEVEr on 96-well capacity with higher TWEEN-20 concentrations (0.5%-1%-2.5%-5%), FAEVEr on 96-well capacity with 0.1% TWEEN-20 and FAEVEr on 96-well capacity with PBS only. Moreover, three protein clusters are colored from top to bottom in pink (cluster 1), light pink (cluster 2) and blue (cluster 3) for UF6 and in light pink (cluster 1), light blue (cluster 2) and gold (cluster 3) for UF96. These protein clusters are additionally defined by black lines. The highest z-scores are colored in red and lowest in blue.*

Furthermore, gene ontology analysis of the defined protein clusters in UF6 and UF96 reveals distinct enrichment patterns related to TWEEN-20. In PBS without TWEEN-20 addition, the proteome of the EV-enriched fraction seems to contain higher levels of contaminating matrix proteins. In both UF6 and UF96 with PBS only (pink cluster 1 Figure 19**A**, gold cluster 3 Figure 19**B**), lipoprotein particle-associated proteins are enriched. In UF96 (gold cluster 3 Figure 19**B**), blood microparticle and collagen fiber-associated proteins are enriched without TWEEN-20. In contrast, FAEVEr with TWEEN-20 (blue cluster 3 Figure 19**A**, light pink cluster 1 Figure 19**B**) aids in the enrichment of our proteins of interest, such as endosomal system and ESCRT-related proteins (CHMP3, CHMP2B, CHMP4B in UF6) and intrinsic plasma membrane proteins compared to PBS only. Also, membrane proteins related to cell projection are found in higher levels for TWEEN-20 addition (light pink cluster 2 Figure 19**A**, light pink cluster 1 Figure 19**B**). For UF6, these particular proteins are even more enriched with 5% TWEEN-20 (blue cluster 3 Figure 19**A**). However, in this more extreme TWEEN-20 condition (light pink cluster 2 Figure 19**A**), some membrane proteins related to cell projection and cell junction, and a few endomembrane system-associated proteins, are less abundant.

Strikingly, in the absence of TWEEN-20 in UF6 and UF96 (hot pink cluster 1 Figure 19**A**, gold cluster 3 Figure 19**B**), cytosolic ribosome-, spliceosome-, proteasome- and nucleus-related

proteins are highly enriched. Moreover, 91 and 32 proteins were only identified in UF6 or UF96 without TWEEN-20 addition to PBS, and these were also mainly related to the nucleus, ribosomes or spliceosome. Also, a few ribosomal subunit and nucleus-associated proteins are slightly enriched in the TWEEN-20 conditions (blue cluster 3 Figure 19**A**). However, with addition of 5% TWEEN-20 to UF6, these proteins are again relatively more abundant compared to the other TWEEN-20 concentrations and PBS only (blue cluster 3 Figure 19**A**). Additionally, in UF96, a separate protein cluster arises, less abundant in 0.1% TWEEN-20 and more in 5% TWEEN-20 and PBS (light blue cluster 2 Figure 19**B**). This protein cluster also mainly contains ribosome, spliceosome and nucleus-related proteins.

g) The optimal TWEEN-20 concentration for subsequent applications of FAEVEr

Since TWEEN-20 is found to have a beneficial impact on FAEVEr-mediated EV proteomic profiles that form the fundament for EV protein biomarker discovery, MS-results were used to select an optimal TWEEN-20 concentration for further applications of the FAEVEr setup on more complex biological matrices. Although 0.5% TWEEN-20 does not yield the highest number of identifications, it yields over 2000 protein identifications on average with minimal variation in both the number of human identifications and the number of unique human peptides identified (Figure 12), thereby enhancing reproducibility and accuracy in protein quantification. This complements the mediocre effect of 0.5% TWEEN-20 on the increase in the relative number of human identified proteins (Figure 13) and the proportion of spectral space occupied by human peptide precursor ions (Figure 14). Moreover, in the case of UF96, 0.5% TWEEN-20 is the initial concentration that surpasses the transitional threshold set by 0.1% TWEEN-20. Therefore, its impact on the augmentation of the relative number of transmembrane proteins is maximal, as a further increase in TWEEN-20 concentration does not significantly enhance this proportion anymore (Figure 13). Also for that reason, 0.5% TWEEN-20 exerts a more pronounced impact on the proteomic profile obtained with FAEVEr combined with TWEEN-20 compared to PBS only and 0.1% TWEEN-20. Therefore, it leads to relatively higher levels of endosomal system-, ESCRT-related proteins and intrinsic plasma membrane proteins and to relatively lower levels of potentially contaminating matrix proteins. In conclusion, we propose 0.5% TWEEN-20 as the optimal concentration for further applications of the FAEVEr with TWEEN-20 setup, for both low and high throughput format.

PART 7.2 Applying FAEVEr to more complex biological matrices

To move up the translational ladder, it was evaluated whether the FAEVEr setup with TWEEN-20 can enrich EVs from biological matrices by focusing on physical obstruction of the filter, contaminant removal and the proteome of the EV-enriched fraction. As a pilot experiment, a comparison was made between FBS and EDS to confirm EV enrichment from serum. Later on, FAEVEr was performed on higher concentrations of FBS to assess the limits of UF6 and UF96 and to find the optimal balance between filter clogging and the identification of relevant EV-proteins. To investigate the potential of FAEVEr with TWEEN-20 to yield biologically relevant protein identifications with DIA LC-MS/MS, the CM of three pancreatic ductal adenocarcinoma (PDAC) cell lines, serum of influenza infected mice and plasma from breast cancer PDX mouse models were used as input material and analyzed for the identification of protein biomarkers.

7.2.1 Confirming EV enrichment with FAEVEr by comparing the enriched fractions from FBS and EDS

The FAEVEr setup on 6 mL capacity (UF6) was performed on 1 mL FBS and EDS diluted 1:4 (resulting in 20% FBS and 20% EDS) with TWEEN-20 in PBS to a final concentration of 0.1% TWEEN-20, which was selected before the large-scale comparison, based on the results of the proof of principle experiment (S6). Moreover, the effect of pre-clearing using a 0.22 µm filter was assessed by analyzing both pre-cleared (PC) and non-pre-cleared samples (non-PC) in triplicate. As FBS contains a higher protein concentration compared to EDS (40 µg/µL versus 5 µg/µL), which was part of the EV-embedding matrix in the large scale comparison

experiment, the EV isolation efficiency of FAEVEr was again evaluated by loading the wash fractions on SDS-PAGE and by comparing the enriched proteome of both input materials.

a) Wash fractions from FBS contain more contaminating protein material compared to EDS

The flow-through, wash and lysate fractions were visualized on Coomassie Blue stained gels Figure S3. In the flow-through, except for one replicate of each input material, both PC and non-PC show the same amount of contaminants (Figure S3**B**). Similarly, in the wash fractions and lysates, no different profile is observed for PC or non-PC input material (Figure S3**B-D**). However, comparing the input matrices, less material is removed in the flow-through for EDS compared to FBS (Figure S3**A**), likely due to the lower protein concentration of EDS. In the first washing step, a large amount of protein material is washed away for FBS compared to EDS (Figure S3**A**). In contrast, by the fifth wash fraction almost no contaminants are visible for both input materials (Figure S3**C**). Aside from a small band in the FBS wash fractions, no additional bovine serum albumin (BSA, 65 kDa) is filtered through, indicating that the majority of the contaminating proteins are removed over successive washing steps. Remarkably, a small band at 25 kDa appears in the lysates (Figure S3**D**), which is slightly present in the flow-through and first wash fraction, but not in the fifth wash fraction. This suggests partial retention of this material on the filter, despite excessive TWEEN-20 washing steps. Via mass spectrometry, the 25 kDa protein was identified as apolipoprotein A1 (APOA1), primarily associated with HDLs[11] that are smaller than EVs and can be washed away during the flow-through and the first washing steps. However, HDLs may non-specifically interact with the filter and be retained, or APOA1 might be present on similarly-sized chylomicrons leading to APOA1 contamination in the EV protein fraction. In conclusion, the wash test shows similar behavior between PC and non-PC samples. For FBS observably more protein material is washed away in the flow-though and initial washing step compared to EDS, consistent with their respective protein concentrations. The lysate fractions reveal that FAEVEr does not effectively remove lipoprotein particles, resulting in APOA1 contamination. We further investigated the lysate fractions by analyzing the full proteome.

b) Proteome analysis confirms that FAEVEr is capable of enriching EVs

Considering the pre-clearing step, no effect is observed on the sample clustering of both serum types in a PCA plot (PCA, Figure 20**A**), aligning with SDS-PAGE results that show no difference in washing efficiency. Remarkably, a volcano plot comparing PC versus non-PC FBS reveals more enriched transmembrane proteins in the PC condition (Figure 20**D**) and EV markers are slightly shifted towards the PC condition, suggesting that pre-clearing aids in enriching EVs. Comparing the input materials, FBS samples cluster together, while the EDS samples are more spread but clearly separate from the FBS samples (PCA, Figure 20**A**). Together with the presence and absence of EV markers in FBS and EDS samples, respectively (Figure 20**B**), this clustering pattern confirms that FAEVEr enriches EVs and it indicates that the EDS matrix is EV depleted. Also, the total number of protein identifications is about three times higher in (PC-)FBS samples compared to (PC-)EDS samples, pointing to an EV protein enrichment from FBS (Figure 20**B**). Still, not all additional proteins in FBS originate from EVs, as the depletion process also removes other proteins that are thus absent in EDS. The subcellular location analysis shows relatively fewer transmembrane proteins and more secreted proteins in the isolated fraction from EDS (Figure 20**B**), which is also visualized in a volcano plot comparing FBS and EDS (Figure 20**C**). Indeed, transmembrane proteins are highly enriched in FBS, further indicating EV enrichment from (PC) FBS. However, a significant background signal is expected due to unintended enrichment of non-EV proteins with FAEVEr from (PC) EDS samples.

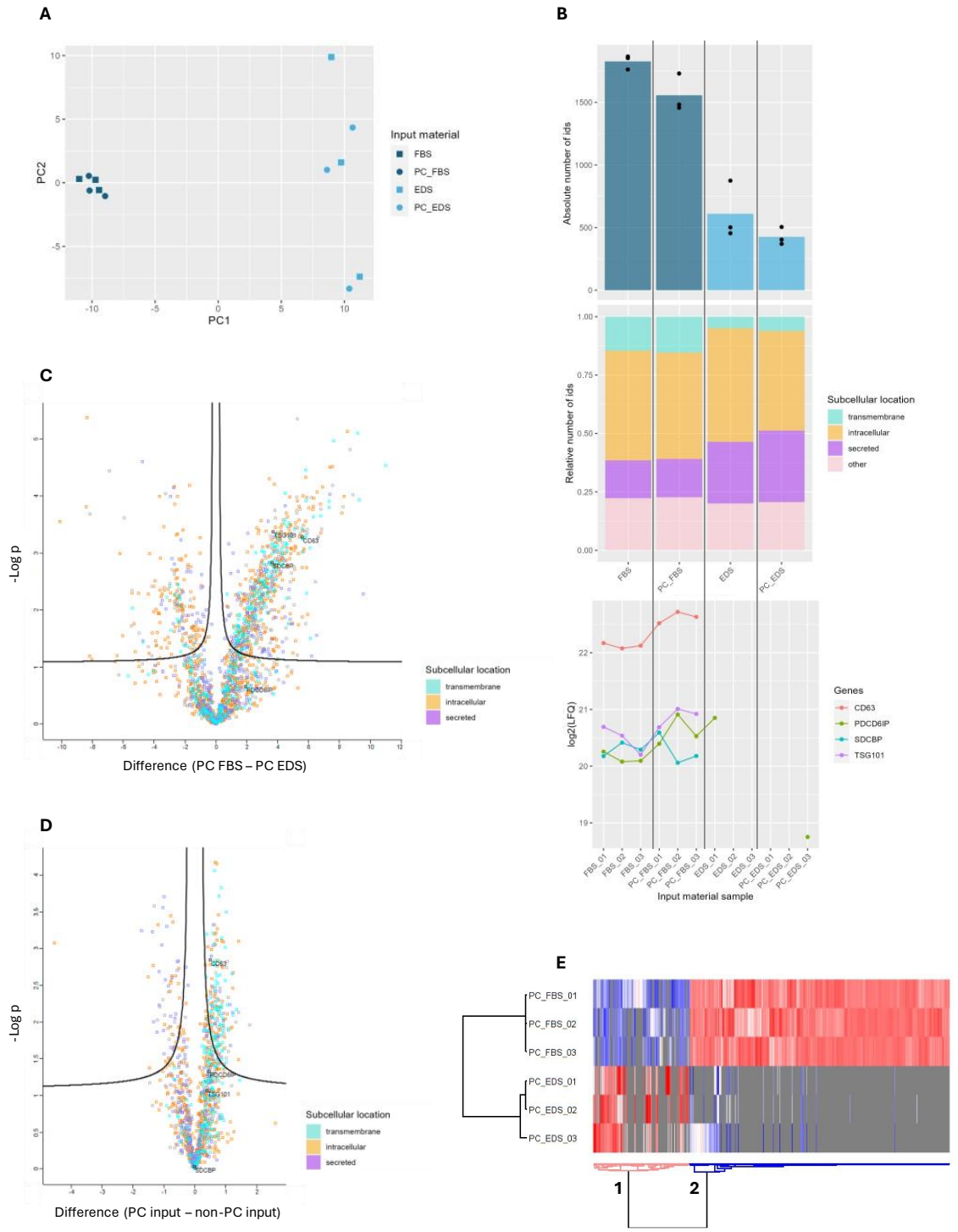


Figure 20: Proteome analysis of FAEVEr on 6 mL filter tubes from 20% (PC) FBS and 20% (PC) EDS samples with 0.1% TWEEN-20. ***A)*** *PCA plot of log2 transformed LFQ values of the (PC) FBS and (PC) EDS samples.* ***B)*** *The first panel shows the number of identifications per (PC) input material, the second panel depicts the relative subcellular location of these identified proteins, the third panel is a profile plot of the EV markers that were detected over all samples.* ***C)*** *Volcano plot (FDR = 0.05, S0 = 0.1) of imputed values comparing the EV proteome from PC FBS and PC EDS.* ***D)*** *Volcano plot (FDR = 0.05, S0 = 0.1) of all valid values comparing the EV proteome from PC and non-PC input material.* ***E)*** *Hierarchical clustering on the imputed values from PC input material: replicates cluster per input material. Moreover, two protein clusters colored in light pink (cluster 1) and blue (cluster 2) can be distinguished. The highest z-scores are colored in red, lowest in blue and missing values in grey. [Abbreviations: (PC) FBS = (pre-cleared) fetal bovine serum, (PC) EDS = (pre-cleared) EV depleted fetal bovine serum, PCA: principal component analysis.]*

Following imputation, the significantly differing protein levels between PC FBS and PC EDS were hierarchically clustered. Both PC FBS and PC EDS samples cluster together, forming two protein clusters (Figure 20**E**). Gene ontology analysis in STRING reveals that cluster 1 (pink), enriched in PC EDS samples, contains proteins from the extracellular region, cell periphery and cytoplasm, resulting in no specific enrichment. Although some EV-related cellular component terms including '(Cytoplasmic) vesicle' and 'Clathrin coat of coated pit' show up, they are non-specifically enriched in the majority of the data and overlap with other non-EV-related components. Therefore, they do not specifically point to EV-biogenesis. Conversely, cluster 2 (blue) is strongly enriched for transmembrane and intracellular proteins, and particularly ESCRT complex-related proteins. These results confirm that EDS is truly EV-depleted and that the FAEVEr setup on 6 mL capacity is capable of enriching EVs from complex matrices like FBS.

7.2.2 Determining the maximum filtration potential of FAEVEr by increasing the FBS concentration of the starting material

Our final goal is to enrich EVs from human plasma on a high throughput scale through parallelization of sample processing from low starting volumes, whilst maximizing the number of isolated EVs. This comes down to minimizing the sample dilution to concentrate the highest number of EVs in the smallest volume possible. Still, this theoretical principle is constrained by the filtration potential of the UF6 and UF96 filtration devices. Therefore, by executing the protocol on different concentrations of FBS, we determined the minimum sample dilution for UF6. This dilution correlates with the maximum protein concentration that allows filtration without clogging and will serve as a measure for EV enrichment from more biologically relevant samples. In short, the goal is to determine the optimal protein concentration, as high as possible to increase the number of relevant protein identifications, but still low enough to avoid filter blocking. Since previous results show that UF6 can handle 20% FBS, we prepared a dilution series of FBS with final concentrations ranging from 20% expanding to 30%-40%-50% FBS by using different starting volumes diluted to 5 mL per FBS condition (Figure S4). These samples were diluted in 0.1% TWEEN-20 in PBS, which was also used as the wash buffer during the protocol.

a) Concentrations up to 30% FBS elude filter clogging while efficiently removing non-EV proteins

During the protocol, the filtration efficiency was profoundly monitored. After loading the different concentrations of FBS, in all cases except for 20% FBS, visual fouling of the filter occurred to some extent (Figure S5**A**). Still, during the washing steps this membrane fouling was removed for the 30% FBS samples, resulting in filtration behavior similar to the 20% FBS samples, which were completely filtered through in every step. However, for 40-50% FBS, particle accumulation and protein aggregation at the filter membrane led to filter clogging during the protocol (Figure S5**B**). Note that the 40% FBS replicates showed high variability in the extent of filter clogging as shown in Figure S5**B**, suggesting that 40% FBS exceeds the limit of the filtration capacity. Further, it is a transitional concentration between 20%-30%, where no problems are observed, and 50% FBS, where consistent filter blocking occurs.

The consequences of these observations are further elucidated by a Coomassie Blue staining of the flow-through, wash and lysate fractions after SDS-PAGE in Figure S5**C-F**. In the flow-through fraction, less protein material is visualized for the 50% FBS replicates which can be attributed to filter clogging that blocks protein passage (Figure S5**C**). Moreover, during the washing steps, increasing amounts of protein material seem to be washed away for increasing FBS concentrations (Figure S5**D-E**). In the third wash step, no additional protein material is washed away for the 20% FBS replicates (Figure S5**E**). Finally, in the EV lysate fraction of the 20% and 30% FBS replicates, albumin contamination is noticeably lower compared to the 40% and 50% FBS replicates (Figure S5**F**). However, in the EV fraction of the 20% and 30% FBS replicates, APOA1 is still clearly present.

b) Proteome analysis of the enriched fraction from 30% FBS reveals most EV-related proteins of interest

In line with the observations during the execution of the FAEVEr protocol, the replicates of 20%, 30% and 50% FBS are tightly clustered together by concentration, whereas the 40% FBS replicates are the most widespread (PCA, Figure 21**A**). Except for one replicate of 40% FBS that lies in-between the lower (20% - 30%) and higher (40% - 50%) FBS concentrations, the proteome of the 20% and 30% FBS conditions is highly similar and separately clustered from the higher percentages of FBS. Indeed, the absolute number of identifications (Figure 21**B**) and uniquely identified peptides (Figure S6**A**) show a distinction between the lower and the higher percentages of FBS. Remarkably, the 30% FBS condition outperforms the other conditions for these parameters likely due to its higher starting volume compared to 20% FBS. Sketching a rough idea of the nature of these identified proteins, the subcellular location shows that more intracellular and transmembrane proteins are present in the 20% and 30% FBS conditions compared to the 40% and 50% FBS conditions as visualized in Figure 21**B**. Also, relatively more secreted proteins are identified in these latter conditions. A profile plot demonstrates that 20% and 30% FBS samples contain relatively higher EV marker levels compared to the 40% and 50% FBS conditions (Figure 21**B**). Furthermore, higher CHMP levels and less contaminants including albumin, globin C1, serotransferrin and alpha-2-HS-glycoprotein are present in the 20% and 30% FBS conditions. This depicts how contaminating proteins can mask the less abundant EV-related proteins in the 40% and 50% FBS conditions and highlights the identification of lowly abundant EV-biogenesis-associated proteins in the 30% FBS condition. Hierarchical clustering and gene ontology analysis confirm the above trends (Figure 21**C**). Indeed, two FBS percentage clusters are formed containing 20% and 30% FBS on the one hand and 40% and 50% FBS on the other hand. Additionally, two protein clusters, differentially enriched between those groups are distinguished. Similarly to the PCA plot, one deviating replicate of 40% FBS clusters separately from other 40% and 50% FBS samples. STRING gene ontology analysis indicates that both protein clusters contain extracellular and secreted proteins including extracellular matrix proteins, secretory granules and particularly in cluster 1 (blue) lipoprotein particles. However, cluster 2 (pink), representing proteins enriched in the lower FBS concentrations, includes cytoplasmic and transmembrane (cell projection and membrane associated proteins) cellular components. Furthermore, proteins uniquely identified in the 20%-30% FBS samples show clear endosome- and vesicle-related enrichment (Figure S6**B**). Since bovine extracellular vesicles are poorly annotated, these endosome- and vesicle-related annotations could be an indication of EV proteins.

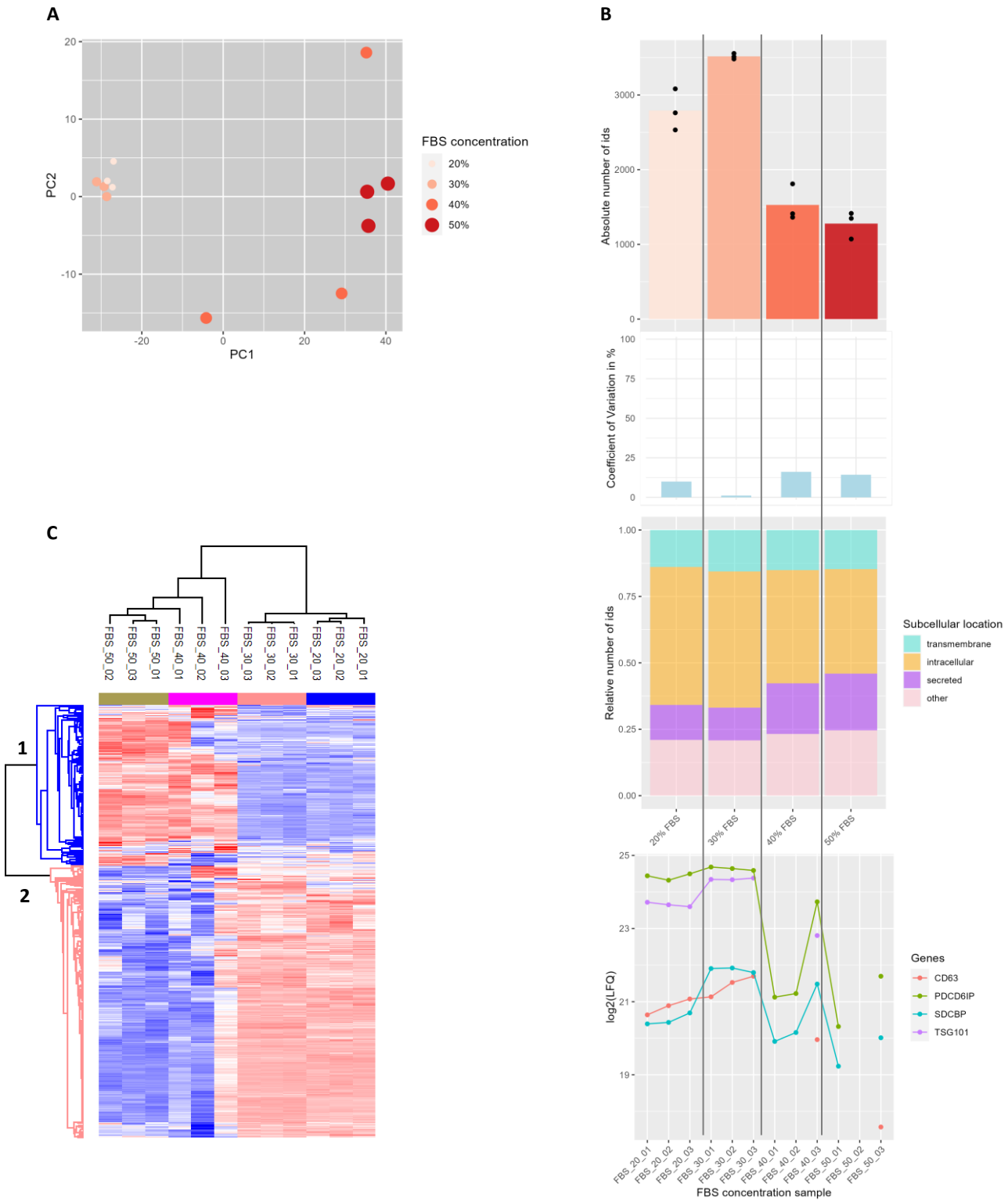


Figure 21: Proteome analysis of FAEVEr on 6 mL filter tubes from 20%-30%-40%-50% FBS samples with 0.1% TWEEN-20. A) PCA plot of log2 transformed LFQ values of all FBS concentration samples. B) The first panel shows the number of identifications per FBS concentration, the second panel depicts the relative subcellular location of these identified proteins, the third panel is a profile plot of the EV markers that were detected over all samples. C) Hierarchical clustering of all FBS concentration samples: replicates cluster per FBS concentration. Moreover, two protein clusters colored in blue (cluster 1) and light pink (cluster 2) can be distinguished. The highest z-scores are colored in red and lowest in blue. [Abbreviations: FBS = fetal bovine serum, PCA = principal component analysis]

In conclusion, 30% FBS, or roughly 10 mg/mL protein material, is the optimal protein concentration that avoids filter clogging during FAEVEr with 0.1% TWEEN-20, while allowing the highest number of relevant EV protein identifications from serum samples containing a total protein amount of about 50 mg. This protein concentration will be used as a measure for filtering more complex biological matrices to optimally map the EV proteome.

c) Downscaling FAEVEr from UF6 to UF96 with 50% FBS showing the highest filtration efficiency

Given the successful EV enrichment from 20% and 30% FBS with UF6, we further downscaled this protocol on a 96-well filter plate and assessed whether the optimal protein concentration of 10 mg/mL for FAEVEr can be extrapolated from UF6 to UF96 format, using the same range of FBS concentrations. However, compared to the UF6 experiment, we made some substantial changes that were supported by our previous results. Firstly, we employed 0.5% TWEEN-20 instead of 0.1% TWEEN-20, as the large-scale comparison of EV enrichment methods concluded that 0.5% TWEEN-20 in PBS was the optimal concentration to use for UF96. Secondly, the dilution approach differs from the previous UF6 experiment, in which the total loading volumes were kept constant and the FBS input volumes were varied to reach the final percentages of FBS, whereas in UF96, a constant FBS input volume of 500 µL was used and diluted to increasing loading volumes (Figure S4). Ideally, 1 mL diluted sample volume is loaded in three times 15 minutes. However, since the wells of a 96-well filtration plate can maximally contain 400 µL, highly diluted FBS concentrations require an increasing number of loading steps. Concretely, seven, five, four and three loading steps were performed for the 20%, 30%, 40% and 50% FBS replicates, respectively.

To assess contaminant removal, the flow-through and wash fractions were analyzed by SDS-PAGE (Figure S7**A-B**). Over successive washing steps, less protein material is washed away, except for 40% FBS replicate 02 (Figure S7**A**). In contrast to UF6, during the protocol, we observed variability in filter clogging across replicates for 20%, 30% and 40% FBS in particular. These trends align with the large difference in protein material between replicates visualized in the lysate fractions of the stained gels (Figure S7**B**). Both 20% and 30% FBS showed efficient flow-through until the last loading step and all 40% FBS replicates started clogging during the second loading step. Unexpectedly, for 50% FBS, only the third replicate demonstrated filter blocking. Of note, since all replicates of the four different experimental conditions are loaded onto the same 96-well plate, the higher FBS concentration replicates were centrifuged more thoroughly without adding additional sample volume, which helped to reduce retained material on the filter for some of the 30% and 40% FBS replicates. However, in the end, the first replicate of the 20%, the second and third replicate of the 30% and all three replicates of the 40% FBS condition still retained material on the filter. These are the replicates with the highest amount of protein material during the third wash step and in the lysate fractions. It can thus be stated that filter clogging should be avoided at all times to ensure a minimally contaminated lysate fraction.

In conclusion, the FAEVEr setup with 0.5% TWEEN-20 on the 96-well plate format shows high variability in filter clogging. Since the 40% FBS replicates performed the worst in this wash test and during the protocol, it could be argued that this is the protein concentration threshold for these filters. However, two of the 50% FBS replicates outperformed the 40% FBS condition. This behavior is not similar to UF6 and emphasizes the need for a separate UF96 optimization. Therefore, prior to progressing to more complex matrices or higher concentrations of protein material, further experiments should be conducted to reduce the variability in filter clogging and to determine the optimal protein concentration for EV enrichment using the 96-well filter plate.

7.2.3 Investigating biological differences between EVs enriched from conditioned medium from pancreatic ductal adenocarcinoma cell lines

EVs were enriched from conditioned medium (CM) from three PDAC cell lines, MIA PaCA-2, PANC-1 and CAPAN-1 representing pancreatic cancer cells in a poorly differentiated, moderately poorly differentiated and well-differentiated state, respectively[31,32]. As the culture medium contains only 10% EDS, this matrix is relatively simple compared to lowly diluted serum or plasma samples. This was deemed an ideal opportunity to explore the potential of the UF96 protocol for biomarker discovery. The pre-cleared CM samples (500 µL) were not diluted before being loaded on the filter plate in two rounds. During the washing steps the EVs were washed on the filter with 0.5% TWEEN-20 in PBS. To ensure a smooth flow-through, the plate was centrifuged at 1,500 x g (the upper limit of the used centrifuge). Indeed, despite a small volume remaining on the filter during intermediate centrifugation steps for PANC-1 and CAPAN-1 samples, in the end, all samples showed a complete flow-through before addition of the 5% SDS lysis buffer.

a) UF96 shows optimal washing efficiency but potential loss of EVs when applied to CM of pancreatic ductal adenocarcinoma cell lines

On Coomassie stained gels, optimal washing efficiency of UF96 is demonstrated by the regular pattern of gradually fewer removal of protein material over successive washing steps and the absence of observable albumin or APOA1 in the lysate fractions (Figure S8**A-B**). Note that the detection limit of a Coomassie staining ranges from 0.1 to 0.5 µg protein[33], which means that contaminating proteins can still be present in low quantities. Moreover, Western blotting against SDCBP revealed a similar pattern as for the UF96 isolation of rEVs (Figure S8**B**). Here, however, the largest share of the SDCBP signal is present in the first wash fraction for the MIA PaCA-2 cell line, and present in the flow-through and first wash fraction for the CAPAN-1 and PANC-1 cell line. In contrast to rEVs, where Gag-eGFP levels are evaluated, these EV lysate fractions were not assessed for preservation of luminal cargo. Therefore, this does not necessarily indicate that the EV integrity is harmed. Indeed, contradictory to these Western blot results, DIA LC-MS/MS analysis revealed that SDCBP is present in the lysate of all samples and, even more, it turns out to be the most abundant EV marker (Figure 22**B**). It is possible that the amount of SDCBP in the lysate was not enough to transcend the Western Blot detection limit compared to the detection limit in mass spectrometry. Consistent presence of other EV-markers like PDCD6IP and TSG101 also contradict the loss of EVs during the washing steps. Still, these results stress the need for further investigation of the effect of the UF96 protocol on SDCBP and other surface proteins and thus ask for cautiousness when interpreting the EV surfaceome.

b) High variability between the number of protein identifications of EV-enriched fractions from different pancreatic ductal adenocarcinoma cell lines

Proteome analysis revealed a widely ranging number of protein identifications over the selected samples going from 1,074 to 2,822 (Figure 22**B**). Note that due to the availability of a higher CM volume, for the PANC-1 and CAPAN-1 cell line, more than four replicates were initially subjected to FAEVEr and subsequent MS-analysis of which the four replicates with the highest identifications were selected for further proteome analysis as illustrated in Figure S8**C**. Alternatively, the MIA-PaCa-2 condition included only four replicates because of a lower CM volume at our disposal, possibly penalizing this cell line with a lower mean in number of identifications and a higher coefficient of variation (Figures 22**B**). Indeed, for the selected samples, PANC-1 showed the highest number of identifications followed by CAPAN-1 and MIA PaCa-2 (Figure 22**B**). Furthermore, the coefficient of variation indicates the highest variability in number of identifications for the MIA PaCa-2 cell line (Figure 22**B**). Looking at the subcellular location of the proteins, it can be observed that for CAPAN-1 and PANC-1 a higher percentage of transmembrane proteins was identified compared to MIA-PaCa-2, for which more secreted proteins show up (Figure 22**B**). This could be attributed to the biological differences between the cell lines, however, as the MIA PaCa-2 replicates exhibit a lower absolute number of identifications, we suspect that the secreted proteins might mask the identification of lowly abundant EV-related transmembrane proteins.

When a hierarchical clustering is performed on the protein identifications, the replicates of each cell line cluster together, distinguishing three sample clusters corresponding to the three PDAC cell lines (Figure 22**D**). The enrichment pattern of CAPAN-1 and MIA PaCA-2 unexpectedly seems more similar whereas PANC-1 has a small overlap with MIA PaCA-2. Concretely, four protein clusters are formed. Remarkably, cluster 1 (gold) shows that the ESCRT complex proteins are overrepresented in the CAPAN-1 cell line (Figure 22**D**). This is in line with the volcano plots showing an enrichment for EV markers (with TSG101 and PDCD6IP associated with ESCRT) towards the CAPAN-1 cell line (Figure 22**C**). Also, the CHMPs and the Rabs representing the EV-biogenesis proteins were overrepresented in CAPAN-1 indicating an upregulated EV metabolism in this cell line. Biologically, this could be linked to the role of EVs in metastasis as CAPAN-1 is collected from a liver metastasis and we hypothesize that its secreted EVs might have prepared secondary metastatic niches, which can be linked to a more aggressive character of the cancer cell line. Clusters 2 and 3 (blue, Figure 22**D**) are not exclusively enriched in one cell line, but indicate that there is an overlap between the MIA PaCA-2 and the CAPAN-1 cell line on the one hand and an overlap between the PANC-1 and the CAPAN-1 cell line. The lack of overlap between PANC-1 and Mia PaCa-2 is less expected, as CAPAN-1 is the most differentiated epithelial cell line, while Mia PaCa-2 and PANC-1 are rather considered to be an epithelial-mesenchymal cell type.

Looking further into some highly differentially expressed key players (Figure 22**C**), we found insights in FAEVEr's ability to sustain EV protein biomarker research. Firstly, comparing the significantly differentially abundant transmembrane proteins for Mia PaCa-2 and PANC-1, we observed a typical epithelial-related cell-adhesion protein desmoglein 1 (DSG1) to be significantly differentially abundant towards Mia PaCa-2 whereas DAG1, RTN4 and APP, proteins involved in regulation of cell migration, are significantly differentially abundant towards PANC-1. Secondly, we observed that PLOD1 and COL18A1 are significantly differentially abundant towards Mia PaCa-2 and PANC-1, respectively, compared to CAPAN-1. Both play a role in positive regulation of cell migration as PLOD1 is required for cross-linking of collagen fibrils through catalyzation of the hydroxylation of lysine residues and COL18A1 can influence the structural integrity of the extracellular matrix thereby potentially affecting cell adhesion and migration.

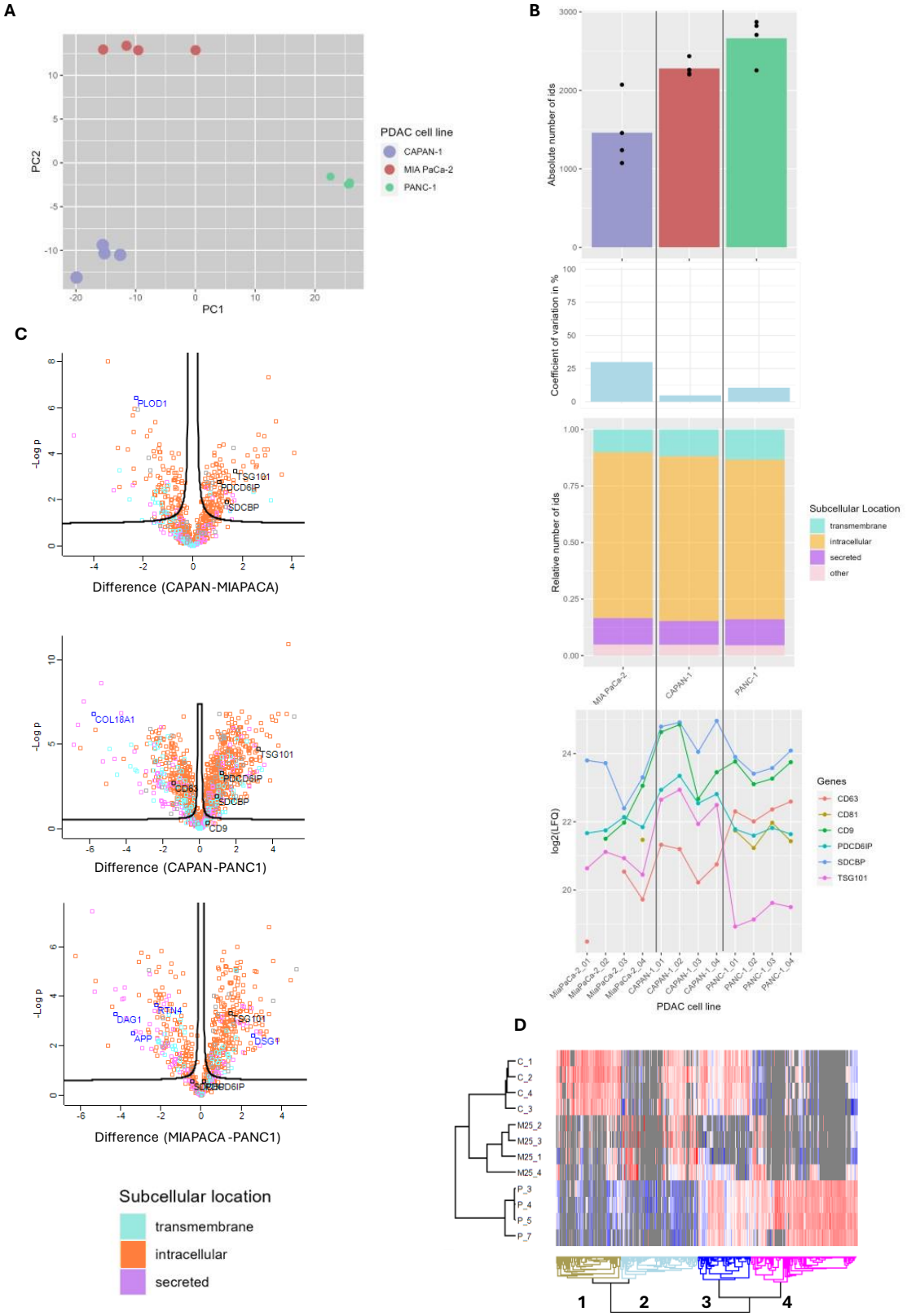


*Figure 22: Proteome analysis of FAEVEr on a 96-well filter plate from CM supplemented with 10% EDS from three PDAC cell lines: MiaPaCa-2, CAPAN-1, PANC-1 with 0.5% TWEEN-20. **A)** PCA plot of the four selected replicates per PDAC type. **B)** Overview of (from top to bottom) the absolute number of identified proteins for four selected replicates per PDAC type showing a large difference in number of identified proteins between the PDAC types, the coefficient of variation of this number of identifications for the four selected replicates per cell line, the relative number of protein identifications per subcellular*

location per PDAC type and the log2 transformed LFQ values of the identified EV markers. ***C)*** *Volcano plots of all valid values (FDR = 0.05, S0 = 0.1) for (from top to bottom) CAPAN versus MIAPACA, CAPAN versus PANC1 and MIAPACA versus PANC1.* ***D)*** *Hierarchical clustering performed on all identified proteins that were significantly differentially expressed between the PDAC cell lines. The highest z-scores are colored in red, the lowest in blue and missing values in grey. [Abbreviations: CM = conditioned medium, EDS = EV-depleted FBS, PDAC = pancreatic ductal adenocarcinoma, PCA = principle component analysis]*

7.2.4 EV enrichment from serum of influenza infected mice

To further expand the complexity of the chosen biologically relevant matrix, we enriched EVs from serum of mice that were infected with influenza compared to non-infected mice. EVs were enriched with UF6, as UF96 is not yet optimized for such complexity, as discussed earlier. The samples (500 µL) were diluted 10 times prior to enrichment to a final protein concentration of 4 mg/mL and were purified using 0.5% TWEEN-20. We observed a smooth flow-through during all centrifugation steps, resulting in an observably high EV purity in the lysate fractions on a Coomassie Blue-stained gel (Figure S9) with remaining minimal contamination of APOA1 and a 75 kDa protein band, considered to be immunoglobulin heavy mu chain based on the MS-analysis results (Table S5).

Approximately 2,000 proteins were identified with DIA LC-MS/MS analysis for both mouse types with low variation between the technical duplicates indicating good reproducibility, with high abundance of the different EV markers (Figure 23**A**). This figure also shows that for both biological replicates approximately 10%, 20% and 50% of these proteins are secreted, transmembrane and intracellular proteins, respectively, which is in line with the results obtained in EV enrichment from bovine serum. This high percentage of transmembrane proteins could be an indication of a successful EV enrichment. Remarkably, when comparing the levels of the identified proteins in a volcano plot, the pulmonary surfactant-associated protein D is highly enriched in the non-infected mouse compared to the influenza exposed mouse (Figure 23**B**). This might be a biological lead to further investigate, indicating that UF6 from mouse serum could function for detection of biologically relevant target proteins. In conclusion, FAEVEr is capable of enriching EVs from complex serum, whilst removing the bulk of high abundant serum proteins in a reproducible way, leading to the identification of potentially relevant proteins from a low starting volume (500 µL). Lastly, it is relevant for extrapolation to a clinical context knowing that 0.5 mL serum suffices to identify EV proteins. On the one hand, less blood should be taken for this analysis and on the other hand, more samples could be processed in parallel.

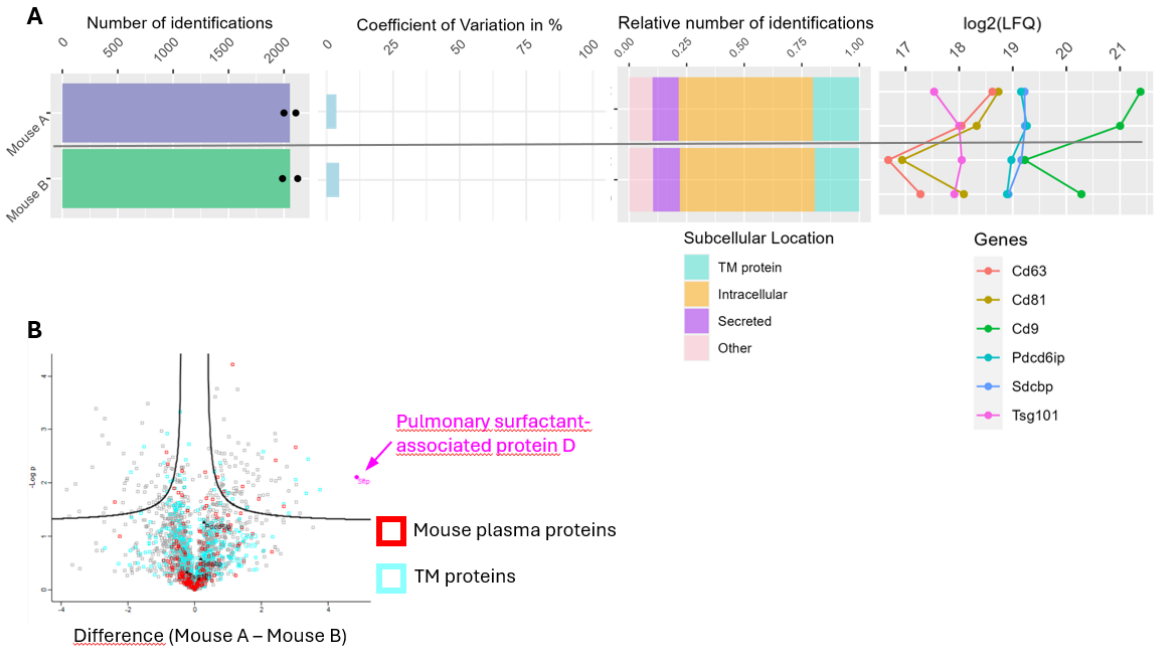


Figure 23: Proteome analysis of FAEVEr on 6 mL filter tubes from 500 µL mouse serum (non-infected and influenza infected) diluted in 5 mL with 0.5% TWEEN-20 washing steps. ***A)*** *Overview of (from left to right) the number of identified proteins per mouse type, the coefficient of variation (in %) on this number of identifications for both mouse types, the relative distribution per subcellular location of the identified proteins and the log2 transformed LFQ values of the identified EV markers.* ***B)*** *Volcano plot (FDR = 0.5, S0 = 0.1) of differentially expressed mouse plasma proteins (red) and transmembrane proteins (turquoise) with pulmonary surfactant-associated protein D indicated in purple. [Abbreviations: Mouse A = non-infected mouse, Mouse B = influenza-infected mouse, TM proteins = transmembrane proteins.]*

7.2.5 Performance of UF6 on plasma from breast cancer patient-derived xenograft mouse models

Plasma is considered one of the most challenging matrices for EV isolation, given its complexity and the abundance of contaminating proteins. Mouse plasma from two consecutive generations of serially transplanted patient-derived breast cancer tumor cells (F8 and F9) was utilized. Various volumes of mouse plasma (ranging from 123 µL to 600 µL, with a median of 450 µL) were collected through cardiac puncture, diluted to 5 mL in 0.5% TWEEN-20 in PBS and subjected to FAEVEr on 6 mL filter-containing tubes. Coomassie staining after SDS-PAGE revealed reproducible gradual reduction in protein material washed away across washing steps for both mouse generations (Figure S10**A**). Even though this indicates efficient removal of contaminants, the lysate fractions still demonstrated visible albumin and APOA1 contamination, in line with serum-derived EV enrichment. Similar to the stained fractions of the PDAC cell lines, a Western Blot targeting the EV-marker SDCBP showed a red fluorescent signal at 32 kDa mainly in the flow-through and first wash fractions but not in the lysate fractions (Figure S10**B**), suggesting potential loss of EVs during protocol execution. Still, in the lysate fraction a low amount of SDCBP might be present below the detection limit of Western Blot.

Indeed, at least partial preservation of the EVs is indicated by the detection of EV-specific markers. SDCBP is detected through DIA LC-MS/MS analysis in three out of four samples from the F8 generation and in five out of six samples from the F9 generation, as illustrated in Figure 24**B**. Also, whereas the EV-marker CD9 is detected in every sample, CD63 is detected in none of the samples. Moreover, CD81, PDCD6IP and TSG101 are sporadically detected in

few samples. Remarkably, all EV-specific markers were identified as mouse proteins, suggesting that no human EVs were enriched. This apparent absence of human EVs could be the consequence of highly abundant mouse proteins masking the lowly abundant human EV proteome. Furthermore, UF6 with 0.5% TWEEN-20 shows potential for isolating human EVs for detecting breast cancer-specific biomarkers, with approximately 100 human proteins identified on average (Figure 24**A**). However, these proteins account for only 16% of the total number of identified proteins, as demonstrated in Figure 24**C**. Moreover, the coefficient of variation for the number of identified proteins is higher for human proteins compared to mouse proteins (Figure 24**A**), suggesting that identification of the latter is more reproducible and such proteins can thus be more accurately quantified. Regarding subcellular localization of the identified proteins, the majority of mouse proteins are secreted. This is not surprising since the plasma proteome consists mainly of secreted proteins such as albumin, apolipoproteins and fibrinogen, which is also represented in the top 10 most abundant proteins identified over all samples (Table S6**A**). These 10 most abundant proteins are all secreted plasma proteins. Moreover, similar to previous experiments with UF6, approximately 20% of the identified proteins in the EV-enriched isolation fraction are of transmembrane origin. These could be associated with the mouse extracellular vesicles, indicating the ability to isolate EVs and contradicting the Western blot results. The identified human proteins are predominantly intracellular, with some tumor secreted proteins and few transmembrane proteins. The top 10 most abundant human proteins mirror these ratios, with the majority being intracellular and serum paraoxonase/lactonase 3 as a secreted plasma protein (Table S6**B**). Of note, we notice that the technical keratin contamination is prominently present, which should be avoided by precautionary sample handling measures. In conclusion, picking up the EVs coming from a tumor still remains challenging, as we used low sample volumes of mouse plasma and the EVs make up a small fraction from the total number of secreted EVs in blood. Therefore, this expresses the need for higher sample volumes, a purer EV fraction and highly sensitive mass spectrometers.

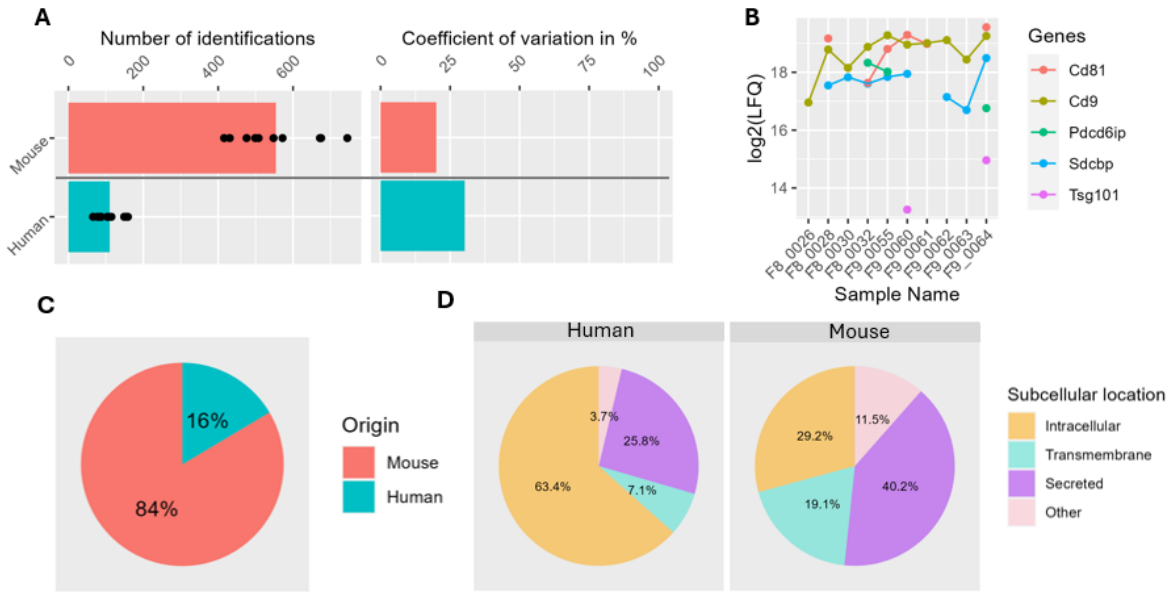


*Figure 24: Proteome analysis of FAEVEr on 6 mL filter tubes from varying volumes (123 µL – 600 µL) of mouse plasma diluted to 5 mL with 0.5% TWEEN-20 washing steps. **A**) Overview of (on the left) the number of identified proteins per origin (mouse in red and human in blue) and (on the right) the coefficient of variation (in %) of this number of identifications per origin. **B**) Profile plot displaying the log2 transformed LFQ values of the identified EV markers of mouse origin. **C**) The relative number of all protein identifications per origin displayed in a pie plot. **D**) Pie plots demonstrate the relative distribution of the identified proteins per subcellular location per origin.*

8. Discussion

Extracellular vesicles (EVs) hold clinical potential as carriers of circulating cancer protein biomarkers, but their isolation from blood plasma and subsequent proteome analysis is complicated by the matrix complexity, as plasma proteins exceed the EV proteome in several orders of magnitude. As a result, efficient EV enrichment and purification strategies are of paramount importance. Still, technical challenges inherent to the employed isolation method can impact EV purity. Indeed, ultrafiltration with 300 kDa MWCO polyethersulfone (PES) membrane filters can co-enrich contaminants up to 300 kDa in molecular weight. These filter membranes can foul or clog due to accumulation of aggregates at the membrane and even non-specific interactions to the membrane can occur. Therefore, we described an optimized strategy for isolating EVs from complex medium for proteome analysis, termed filter-aided EV enrichment (FAEVEr). In FAEVEr, 300 kDa MWCO PES membrane filters are used that retain particles on the filter membrane whilst allowing globular non-EV proteins to pass through. In addition, by extensive comparison, we showed that TWEEN-20 in different percentages benefits the overall purity of the samples. Although previous efforts already indicated that supplementing 5% TWEEN-20 to the wash buffer during the FAEVEr protocol on 6 mL filter tubes (UF6) was beneficial to the overall EV purity[24], these results were obtained in data-dependent acquisition (DDA) and therefore, the setup was revisited towards data independent acquisition liquid chromatography-tandem mass spectrometry (DIA LC-MS/MS) analysis. Furthermore, we explored the potential of FAEVEr for miniaturization to a 300 kDa MWCO PES membrane 96-well filter plate (UF96), since high throughput screening is a feature that is currently unavailable in the EV research field. Yet, increased parallelization reduces time, costs and inter-sample variability. These practical advantages related to parallelization were also experienced during our large-scale comparison experiment (Table 4).

Table 1: Overview of practicalities for ultracentrifugation (UC), FAEVEr on 6 mL 300 kDa MWCO PES membrane filter tubes (UF6) and FAEVEr on a 96-well 300 kDa MWCO PES membrane filter plate (UF96).

Parameter	*UC*	*UF6*	*UF96*
Duration (min.)	140	60	60
Maximum number of concurrent samples	10 tubes	16 tubes	96 wells
Volume processed per centrifugation cycle	Up to 1mL	Up to 6 mL	Up to 500 µL
Contaminant removal	Manual pipetting steps	Filtrate discarded	Filtrate discarded
Equipment costs (EUR)	€10.000 – €25.000 (ultracentrifuge)	€6.000 – €10.000 (centrifuge)	€6.000 – €10.000 (centrifuge)
Additional Material costs (EUR)	€190 euro per 100 tubes	€ 622 per 100 tubes	€ 930 per 25 plates
Cost per sample (EUR)	€1.90 per tube	€6.22 per tube	€0.39 per well

We compared both FAEVEr setups (UF6 and UF96) to ultracentrifugation (UC) as a control, since this is a widespread isolation technique used in EV research. In this comparative experiment, we also implemented a broad range of TWEEN-20 concentrations (0.0% - 0.1% - 0.5% - 1.0% - 5.0%) to enhance FAEVEr-yielded EV purity and to evaluate the effect of TWEEN-20 as an additional control in UC. For UC, the experimental protocol resulted in over seven hours of centrifugation time, as only ten samples can be processed simultaneously and require two centrifugation rounds of at least 70 min. In contrast, the UF6 isolation protocol was completed in two hours as the centrifuge rotor accommodates sixteen filter tubes. However, maximum parallelization was obtained with UF96 resulting in a duration of one hour, as it benefits from the substantial capacity of a 96-well filter plate, and offers the shortest relative processing time per sample. Next to these substantial time differences, we hypothesize that the inter-sample variability could be reduced for both FAEVEr setups by simply discarding the complete filtrate after each step instead of manually pipetting and discarding the supernatant per sample, as this serial disposal might be inconsistent or potentially disturbs the pellet. For UF96, improved reproducibility was indeed observed which could be attributed to both the absence of separately processed sample batches and the straightforward way of removing contaminating material. Lastly, by increasing parallelization, the costs per sample are also reduced.

Next to the beneficial practicalities of FAEVEr, some general aspects of the isolation performance starting from rEV conditioned medium (CM) supplemented with 10% EV-depleted fetal bovine serum (EDS) support that FAEVEr with TWEEN-20 is a valid approach for EV isolation. Both UF6 and UF96 succeed in removing the bulk of contaminating protein material like albumin that makes up 60% of the serum protein concentration[10]. During the first wash step on the filter, TWEEN-20 does not seem to improve washing efficiency compared to PBS. However, in later washing steps, addition of higher TWEEN-20 percentages to PBS washes away more contaminating protein material, thereby efficiently increasing EV purity. Of note, using 5% TWEEN-20, no particles with high lipid content were observed using transmission electron microscopy (TEM), in contrast to PBS, suggesting the detergent's ability to increase EV purity by lowering HDL levels, as apolipoprotein A1 (APOA1) was also found to be significantly differentially more abundant towards PBS compared to TWEEN-20 conditions. Moreover, when adding up to 5% TWEEN-20, rEVs can be retained on the 300 kDa MWCO filter while maintaining their integrity, as illustrated by both Western blotting and scanning electron microscopy. Even though minimal rEV loss is observed for UF96 with 0.5% TWEEN-20, both UF6 and UF96 succeed in efficiently enriching rEVs in the lysed fractions. In addition, TEM illustrates the ability to also recover rEVs washed with PBS or 5% TWEEN-20 from the filter without jeopardizing their integrity, besides the ability to collect EV cargo by on-filter-lysis. However, a lower number of recovered particles is measured for UF6 and UF96 compared to UC. This could be attributed to potential trapping of the rEVs by the filter. Therefore, on-filter lysis is preferred to ensure full recovery of the EV proteome for subsequent DIA LC-MS/MS analysis.

We have confirmed that FAEVEr with TWEEN-20 is a viable method for EV enrichment since the general aspects of rEV isolation exhibit similarities across UC, UF6 and UF96 with no detrimental effects of TWEEN-20 on the rEV integrity. However, substantial variations do emerge in the rEV proteomic profiles, which is relevant for further EV protein biomarker discovery. On the one hand, these differences are attributable to the isolation method used. With regard to the protein identifications and quantification, UC yields most protein identifications, whereas UF96 yields more accurate protein quantification due to its greater reproducibility. On the other hand, for UF6 and for UF96 to an even greater extent, the proteomes are also clearly impacted by the addition of TWEEN-20. Firstly, the relative number of human protein identifications, including the subset of transmembrane proteins, increases in a linear and non-linear trend, respectively, by adding TWEEN-20 during FAEVEr. Similarly, the proportion of the spectral space occupied by human precursors exhibits a linear association with TWEEN-20 concentration added. However, for UF6, this enlargement reaches a plateau. Moreover, FAEVEr with TWEEN-20 may be elevating levels of endosomal system, ESCRT-

related proteins and intrinsic plasma membrane proteins, while reducing levels of contaminating matrix proteins compared to FAEVEr with only PBS. Additionally, in UF96, TWEEN-20 proves crucial for enriching EV-specific markers to levels comparable to UC and UF6. Conversely, FAEVEr with only PBS tends to enrich ribosome-, spliceosome-, proteasome- and nucleus-related proteins compared to the addition of TWEEN-20 in both 6 mL and 96-well formats. Therefore, we propose a new hypothesis suggesting possible adherence of these specific proteins to the EV surface. Their interactions could potentially be interrupted using TWEEN-20, thereby reducing the detected protein levels of these specific proteins in the lysed EV fractions compared to PBS without TWEEN-20. To test this hypothesis, an additional pilot MS-based FAEVEr experiment could be conducted using PBS with and without addition of TWEEN-20, comparing the sheared EV surface fraction prior to lysis with the fractions containing subsequently lysed EVs. Furthermore, based on its moderate effects and superior reproducibility, alongside its distinction as the minimum concentration to exceed the transitional threshold observed at 0.1% TWEEN-20 in UF96, 0.5% TWEEN-20 emerges as the optimal concentration for further applications of FAEVEr, whether in low or high throughput format.

Even though FAEVEr with TWEEN-20 has proven to facilitate enrichment of EVs from CM supplemented with 10% EDS with a beneficial impact on the EV proteomic profile, plasma remains the desired biofluid for clinically relevant applications. FAEVEr with TWEEN-20 was therefore tested on matrices with increasing complexities. We demonstrated the capability of reproducibly enriching EVs from different complex biological matrices using UF6. We established an optimal starting material protein concentration for EV enrichment of 30% FBS, corresponding to approximately 10 mg/mL, which demonstrated a continuous flow-through and yielded optimal DIA LC-MS/MS results with an enrichment of EV markers and ESCRT complex-related proteins. In addition, a proteome comparison between the lysed FAEVEr retentate of FBS and EDS revealed different relative distributions of the detected proteins in subcellular location categories. We observed an increase in transmembrane proteins (with a final proportion of approximately 15-20% transmembrane proteins) together with a decrease in secreted proteins, indicating efficient EV enrichment. However, it should be noted that FAEVEr on EDS still yielded quite high levels of transmembrane and intracellular proteins and, therefore, considerable background could be present when EDS is part of the EV-embedding matrix used for biological applications. UF6 yielded similar results on diluted mouse serum. Once again, all EV markers were detected reproducibly in all mouse replicates and approximately 15-20% of the identified proteins were categorized as transmembrane proteins. In contrast, only 10% were secreted proteins, illustrating efficient reduction in contaminating proteins. Even when applied on mouse plasma, UF6 resulted in detection of all EV-specific markers of mouse origin, though not in every sample. Also similar to previous matrices, a proportion of 20% transmembrane proteins for the enriched mouse proteome was observed. Of note, FAEVEr with TWEEN-20 does not exclude contamination of the EV enriched fraction with apolipoproteins as APOA1 has been found to be abundantly present in the lysed fractions of FBS, mouse serum and mouse plasma using both SDS-PAGE with Coomassie staining and DIA LC-MS/MS analysis. This is potentially due to non-specific interactions of HDLs with the filter or because of retention of similarly-sized chylomicrons containing APOA1.

Lastly, FAEVEr with TWEEN-20 has potential to detect biologically relevant protein biomarkers carried by EVs. As a first example, this enrichment strategy led to the discovery of pulmonary surfactant-associated protein D as a potential lead for further research since it was observed to be enriched in mouse serum from non-infected mice compared to influenza-infected mice. Furthermore, when applied to CM from three different pancreatic ductal adenocarcinoma cell lines, it allowed analysis of discrepancies in the corresponding EV proteomic profiles such as the differences in the fraction of transmembrane proteins and secreted proteins. However, these variations could also be attributed to the variability in the number of identifications for the three cell lines observed in this experiment. Moreover, EV markers and EV biogenesis related proteins were overrepresented in the CAPAN-1-derived lysed fractions which could be linked

to the role of EVs in metastasis. Comparing significantly differentially abundant transmembrane proteins of Mia PaCa-2 and PANC-1, we observed a typical epithelial-associated desmoglein (DSG1), involved in formation of desmosomes that are present in the basolateral membrane of epithelial cells, to be more abundant in Mia PaCa-2 samples. Alternatively, proteins linked to regulation of cell migration (DAG1, RTN4, APP) were found to be significantly differentially abundant towards PANC-1. This could be linked to the superior EMT potential of PANC-1 compared to Mia Paca-2[32]. In the Mia PaCa-2 samples compared to CAPAN-1 samples, we also detected significantly differential presence of PLOD1 that previously has been linked to poor prognosis[34] as increased collagen deposition can enhance tumor progression through cell migration and invasion. Another potential link to this collagen-dependent cell migration is the significantly higher differential abundance of COL18A1 in PANC-1 compared to CAPAN-1. However, we lack a method to measure the background, which gives uncertainty whether the nature of the identified protein is secreted or EV-linked. Ideally, the EV-depleted fraction, obtained during the protocol is analyzed on MS and used as a background, though the MS analysis would still be dominated by the most abundant secreted proteins leading to masking of a large share of the contaminating proteome. As final sustaining evidence, FAEVEr employed on mouse plasma resulted in the detection of approximately 100 human proteins specifically related to the tumor, yet the relatively more abundant mouse proteome considerably limits more in-depth biomarker discovery[26]. For future applications towards the clinic, it could be beneficial to consider an additional step that allows better discrimination between cancer-related and host-related EVs after FAEVEr from PDX mouse plasma to limit masking of relevant tumor-specific proteins by the mouse EV proteome, although this could be complicated by high conservation between humans and mice.

In conclusion, FAEVEr with TWEEN-20, optimally 0.5% TWEEN-20, holds promise to enhance proteomics-based EV cancer biomarker discovery. On 6 mL format, it succeeds in efficiently and reproducibly enriching EVs from serum up until a protein concentration of approximately 10 mg/mL. We demonstrated that this protein concentration functions as a balance between avoidance of filter clogging and maximal discrepancy in the protein dynamic range of the EV-enriched fraction, which is preferred for an optimal EV proteome analysis. Addition of TWEEN-20 to FAEVEr aids in the removal of contaminating proteins and results in a yield of approximately 15-20% of the enriched proteome as transmembrane protein identifications in all tested matrices. Of note, transmembrane proteins hold significant biomarker potential as they are easily accessible for affinity reagents used in targeted diagnostic tests. Contrary to serum, when starting from plasma, the enrichment method shows potential, but still exhibits more variability in detecting EV-specific markers. Of note, FAEVEr on a 96-well plate benefits most from TWEEN-20 addition. Even more, TWEEN-20 addition is required to obtain detected relative EV marker levels similar to UC and FAEVEr on 6 mL filter tubes. Although this setup requires further optimization for its application on more complex biological matrices indicated by high variability in filtration efficiency at the level of serum, this FAEVEr high throughput setup has potential to further bridge the gap between research and clinic by advancing proteomics-based EV cancer biomarker discovery from relatively complex matrices.

9. References

1. Welsh, J. A. *et al.* Minimal information for studies of extracellular vesicles (MISEV2023): From basic to advanced approaches. *J. Extracell. Vesicles* **13**, e12404 (2024).
2. Yáñez-Mó, M. *et al.* Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. Vesicles* **4**, 27066 (2015).
3. Popa, S. J. & Stewart, S. E. Socially Distanced Intercellular Communication: Mechanisms for Extracellular Vesicle Cargo Delivery. in *New Frontiers: Extracellular Vesicles* (eds. Mathivanan, S., Fonseka, P., Nedeva, C. & Atukorala, I.) 179–209 (Springer International Publishing, Cham, 2021). doi:10.1007/978-3-030-67171-6_8.
4. Mir, B. & Goettsch, C. Extracellular Vesicles as Delivery Vehicles of Specific Cellular Cargo. *Cells* **9**, 1601 (2020).
5. Xu, R. *et al.* Extracellular vesicles in cancer — implications for future improvements in cancer care. *Nat. Rev. Clin. Oncol.* **15**, 617–638 (2018).
6. Record, M., Silvente-Poirot, S., Poirot, M. & Wakelam, M. O. Extracellular vesicles: lipids as key components of their biogenesis and functions. *J. Lipid Res.* **59**, 1316–1324 (2018).
7. Zha, Q. B., Yao, Y. F., Ren, Z. J., Li, X. J. & Tang, J. H. Extracellular vesicles: An overview of biogenesis, function, and role in breast cancer. *Tumor Biol.* **39**, 1010428317691182 (2017).
8. Peinado, H. *et al.* Pre-metastatic niches: organ-specific homes for metastases. *Nat. Rev. Cancer* **17**, 302–317 (2017).
9. Geeurickx, E. *et al.* The generation and use of recombinant extracellular vesicles as biological reference material. *Nat. Commun.* **10**, 3288 (2019).
10. Leeman, M., Choi, J., Hansson, S., Storm, M. U. & Nilsson, L. Proteins and antibodies in serum, plasma, and whole blood—size characterization using asymmetrical flow field-flow fractionation (AF4). *Anal. Bioanal. Chem.* **410**, 4867–4873 (2018).
11. Simonsen, J. B. What Are We Looking At? Extracellular Vesicles, Lipoproteins, or Both? *Circ. Res.* **121**, 920–922 (2017).
12. Brennan, K. *et al.* A comparison of methods for the isolation and separation of extracellular vesicles from protein and lipid particles in human serum. *Sci. Rep.* **10**, 1039 (2020).
13. Zhang, Y. *et al.* Exosome: A Review of Its Classification, Isolation Techniques, Storage, Diagnostic and Targeted Therapy Applications. *Int. J. Nanomedicine* **15**, 6917–6934 (2020).
14. Wei, R. *et al.* Combination of Size-Exclusion Chromatography and Ultracentrifugation Improves the Proteomic Profiling of Plasma-Derived Small Extracellular Vesicles. *Biol. Proced. Online* **22**, 12 (2020).
15. Tulkens, J., De Wever, O. & Hendrix, A. Analyzing bacterial extracellular vesicles in human body fluids by orthogonal biophysical separation and biochemical characterization. *Nat. Protoc.* **15**, 40–67 (2020).
16. Yakubovich, E. I., Polischouk, A. G. & Evtushenko, V. I. Principles and Problems of Exosome Isolation from Biological Fluids. *Biochem. Mosc. Suppl. Ser. Membr. Cell Biol.* **16**, 115–126 (2022).
17. Konoshenko, M. Y., Lekchnov, E. A., Vlassov, A. V. & Laktionov, P. P. Isolation of Extracellular Vesicles: General Methodologies and Latest Trends. *BioMed Res. Int.* **2018**, e8545347 (2018).
18. Ma, X., Chen, Z., Chen, W., Chen, Z. & Meng, X. Exosome subpopulations: The isolation and the functions in diseases. *Gene* **893**, 147905 (2024).
19. Sidhom, K., Obi, P. O. & Saleem, A. A Review of Exosomal Isolation Methods: Is Size Exclusion Chromatography the Best Option? *Int. J. Mol. Sci.* **21**, 6466 (2020).
20. Geeurickx, E. & Hendrix, A. Targets, pitfalls and reference materials for liquid biopsy tests in cancer diagnostics. *Mol. Aspects Med.* **72**, 100828 (2020).
21. Van Dorpe, S. *et al.* Integrating automated liquid handling in the separation workflow of extracellular vesicles enhances specificity and reproducibility. *J. Nanobiotechnology* **21**, 157 (2023).
22. 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).
23. Van Deun, J. *et al.* EV-TRACK: transparent reporting and centralizing knowledge in extracellular vesicle research. *Nat. Methods* **14**, 228–232 (2017).
24. Pauwels, J. *et al.* Filter-aided extracellular vesicle enrichment (FAEVEr) for proteomics. 2023.07.06.547926 Preprint at https://doi.org/10.1101/2023.07.06.547926 (2024).
25. Geeurickx, E. *et al.* Recombinant extracellular vesicles as biological reference material for method development, data normalization and assessment of (pre-)analytical variables. *Nat. Protoc.* **16**, 603–633 (2021).

26. Barlin, M. *et al.* Proteins in Tumor-Derived Plasma Extracellular Vesicles Indicate Tumor Origin. *Mol. Cell. Proteomics* **22**, 100476 (2023).
27. Eyckerman, S. *et al.* Trapping mammalian protein complexes in viral particles. *Nat. Commun.* **7**, 11416 (2016).
28. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
29. Berthold, M. R. *et al.* KNIME - the Konstanz information miner: version 2.0 and beyond. *ACM SIGKDD Explor. Newsl.* **11**, 26–31 (2009).
30. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
31. Deer, E. L. *et al.* Phenotype and Genotype of Pancreatic Cancer Cell Lines. *Pancreas* **39**, 425–435 (2010).
32. Gradiz, R., Silva, H. C., Carvalho, L., Botelho, M. F. & Mota-Pinto, A. MIA PaCa-2 and PANC-1 – pancreas ductal adenocarcinoma cell lines with neuroendocrine differentiation and somatostatin receptors. *Sci. Rep.* **6**, 21648 (2016).
33. Brunelle, J. L. & Green, R. Chapter Thirteen - Coomassie Blue Staining. in *Methods in Enzymology* (ed. Lorsch, J.) vol. 541 161–167 (Academic Press, 2014).
34. Zhang, J., Tian, Y., Mo, S. & Fu, X. Overexpressing PLOD Family Genes Predict Poor Prognosis in Pancreatic Cancer. *Int. J. Gen. Med.* **15**, 3077 (2022).
35. Zhai, Q., Landesman, M. B., Robinson, H., Sundquist, W. I. & Hill, C. P. Structure of the Bro1 Domain Protein BROX and Functional Analyses of the ALIX Bro1 Domain in HIV-1 Budding. *PLoS ONE* **6**, e27466 (2011).

10. Poster

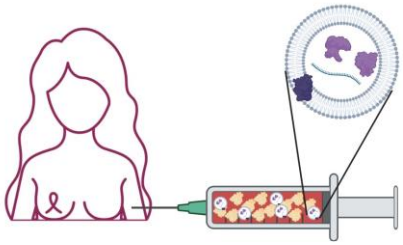
Improving Isolation of Extracellular Vesicles with Ultrafiltration towards Liquid Biopsy Diagnostics for Breast Cancer

Freya De Muyer[1,2], Danaë De Pauw[1,2], Jarne Pauwels[1,2], Kris Gevaert[1,2]

1 Department of Biomolecular Medicine, Ghent University, Ghent, Belgium.
2 VIB-UGhent, Center for Medical Biotechnology, VIB, Ghent, Belgium.

Societal impact of breast cancer diagnostics

With an incidence of more than 30%, breast cancer (BC) is the most common diagnosed cancer among females between 20 and 59 years old. Today, mammography screening is used for diagnosing BC. However, this method only detects more advanced BC cases. Therefore, alternative methods such as early biomarkers detection in liquid biopsies are needed for early diagnosis and thus a better prognosis. For this, extracellular vesicles (EVs) could be used.



What are extracellular vesicles?

Extracellular vesicles (EVs) are a heterogeneous group of nanovesicles with a lipid bilayer that are naturally released by all cell types. Depending on their biogenesis, they are divided into three subcategories: exosomes, microvesicles and apoptotic bodies. EVs also carry information linked to the parental cell type because they are loaded with molecules including:

- nucleic acids
- luminal proteins
- surface proteins

Function of extracellular vesicles?

- contribute to homeostasis because of role in intercellular communication
- contribute to disease initiation and progression

⟶ Relevance for early detection of breast cancer

Problems of EV isolation

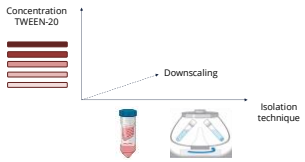
1. Low abundance of EVs compared to contaminants

2. Overlap in size with contaminants

Ultrafiltration workflow for the isolation and purification of extracellular vesicles[1]

1 Filtration with 300 kDa MWCO filter
2 Washing steps with detergent TWEEN-20
3 EV lysis with 5% SDS
4 Protein digestion with trypsin
5 LC-MS/MS analysis - DIA

Project 1
Optimization of EV isolation with ultrafiltration

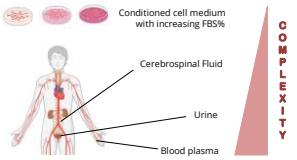


Dimension 1: Comparison of EV isolation techniques
The presented ultrafiltration (UF) workflow will be compared to ultracentrifugation (UC). UC is considered the most widespread technique for EV isolation.

Dimension 2: Comparison increasing TWEEN-20 concentrations
TWEEN-20 detergent is used to reduce non-EV proteins in the retained EV fraction. It also facilitates the ultrafiltration flow-through. A range of concentrations is used to find the best suited TWEEN-20 concentration for optimal EV isolation.

Dimension 3: Upscaling of the ultrafiltration workflow
Finally, this workflow will be tested on high throughput level with 96-well plate filters. In this way, the isolation method would be more applicable in a clinical setting.

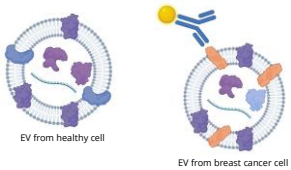
Project 2
More complex biological matrices



Cell culture
The optimized UF protocol for EV isolation will be performed on different matrix complexities in cell culture. The complexity is increased by adding more fetal bovine serum (FBS) to the cell medium. In this way, the optimal protein concentration is determined to prevent filter blocking.

Biological samples
To further augment the clinical relevance, the workflow can be tested on different biological matrices. If the matrix is more complex, the filter is more prone to blocking and the EV fraction is subject to more contamination. This means that a lower sample volume needs to be diluted and less EVs will be captured. So, a balance needs to be found between filter blocking and EV yield.

Project 3
Focusing on the surface proteins for biomarker discovery



EVs as carriers of protein biomarkers
As a result of their biogenesis, EVs contain information of the cell of origin. Therefore, EV cargo differs between cell types. By analyzing these differences, a distinction can be made between healthy and cancer cells present in body fluids. Both luminal and surface protein biomarkers can be discovered with mass spectrometry.

From mass spectrometry to clinical application
Mass spectrometry is less feasible in a clinical high throughput setting. So, an antibody based diagnostic test would be more pragmatic. Therefore, in this research, surface proteins are an important point of focus to accomplish this practicality.

In summary, these three projects contribute to the improvement of EV isolation with ultrafiltration. Since high throughput scale and different body fluids are also taken into account, it is possible to move up on the translational ladder. Finally, the discovered EV biomarkers could be used for early diagnosis of breast cancer to improve prognosis and survival chance of these patients.

[1] Pauwels, J., Van de Steene, T., Van de Velde, J.,Eyckerman, S. & Gevaert, K. Filter-aided extracellular vesicle enrichment (FAEVEr). bioRxiv, 2023.2007.2006.547926, doi:10.1101/2023.07.06.547926 (2023).

VIB-UGENT CENTER FOR MEDICAL BIOTECHNOLOGY
GHENT UNIVERSITY

SCIENCE MEETS LIFE

11. Supplementary

S1. Table ESCRT complex

(cross-reference: pg.2)

Table S1: Overview of the proteins that are part of the ESCRT complex which is involved in the biogenesis of exosomes and microvesicles.

ESCRT subcomplex	Human proteins	Extra information
ESCRT-0	HRS/HGS, STAM1,2	
ESCRT-I	TSG101, VPS28, VPS37A-D, MVB12A/B, UBAP1, UEVLD, UBAP1L	
ESCRT-II	EAP20/VPS25, EAP30/SNF8, EAP45/VPS36	
ESCRT-III	CHMP2A/B, CHMP3, CHMP4A-C, CHMP6	
Accessory ESCRT-III	CHMP7, IST1, CHMP1A/B	ESCRT-III-like, ESCRT-III-like and ESCRT-III-associated protein
ESCRT-III Regulatory	VPS4A/B	ESCRT-III disassembly, late endosome
Bro1 domain-containing proteins	ALIX/PDCD6IP, HD-PTP/PTN23, BROX	Function in association with ESCRT pathway to help mediate intraluminal vesicle formation at MVBs. Human Bro1 domain proteins share the ability to bind the CHMP4 subset of ESCRT-III proteins[35].

S2. Table EV isolation methods

(cross-reference: pg.7)

Table S2: Overview of the advantages and disadvantages of the most well-known and implemented techniques for EV isolation based on specific EV characteristics.

EV isolation technique	Principles Based on:	Advantages	Disadvantages
dUC	Sedimentation rate	✓ Costs related to consumables are relatively low ✓ High reproducibility when same parameters used	× Expensive equipment × Loss of EVs that do not precipitate × Inter-sample variability induced by manual removal of the supernatant × Co-precipitation of non-EV particles × Morphological and functional properties possibly altered due to high centrifugal forces × Labor-intensive, time-consuming × Not suitable for handling small volumes × Limited parallelization × Overlap in sedimentation rate with contaminants
dgUC	Distinct migration behavior in a gradient	✓ Higher EV purity compared to dUC ✓ Separation of viral particles is possible ✓ Compared to dUC: better preservation of EV morphology ✓ automation of fractionation steps is possible	× Expensive equipment × Laborious / time-consuming × EV loss due to manual fractionation × High inter-sample variability × Extra concentration step to reduce sample volume × Limited parallelization × Overlap in size or density with contaminants
PEG precipitation	Aggregation in the presence of a hydrophilic polymer	✓ Downscaling / parallelization possible ✓ Cheap and straightforward sample preparation ✓ Reduces sample volume with minimal loss of EVs ✓ EVs maintain their morphological and functional quality	× Low purity of the EV fraction due to PEG contamination and co-precipitation of non-EV particles × Poor solubility of the precipitated EV aggregates
Immunoaffinity capture	Antibody-based enrichment	✓ Very specific, sensitive leading to a high purity of the EV (subpopulation) fraction	× Bias towards marker expressing EVs × Separation of EVs from the antibodies × Expensive antibodies
SEC	Particle radius	✓ Straightforward EV isolation method	× Limited scalability × Variable purity of the EV fraction depending on the column properties and type of polymer bead × EV loss due to obligatory final concentration step because of sample dilution × Expensive sorbents (but reusable) × Overlap in size with contaminants
Filtration-based methods	Molecular weight & Size/Radius	✓ Straightforward EV isolation method ✓ Downscaling / parallelization possible	× Fouling and clogging of the membrane × Non-specific binding to the membrane × Potential deformation by pressure and contact with the membrane × Overlap in size with contaminants

S3. Table PDX mouse model

(cross-reference pg.10)

Table S3: Overview of the characteristics of the samples from PDX mouse models used during FAEVEr

Sample Number	Tumour type	Tumor size (mm^3)	Plasma (EDTA) Volume (µL)	Method of sacrifice
PDXG.BC.0026	TM00096 F8	1960	250	Isoflurane
PDXG.BC.0028	TM00096 F8	1533.072	350	Isoflurane
PDXG.BC.0030	TM00096 F8	2099.2335	450	Isoflurane
PDXG.BC.0032	TM00098 F8	1605.289	450	Isoflurane
PDXG.BC.0055	TM00096 F9	1492.992	150	CO2
PDXG.BC.0060	TM00098 F9	2323.0575	564	Isoflurane
PDXG.BC.0061	TM00098 F9	2121.808	123	Isoflurane
PDXG.BC.0062	TM00098 F9	2976.75	450	Isoflurane
PDXG.BC.0063	TM00098 F9	2122.2635	360	CO2
PDXG.BC.0064	TM00098 F9	2058.462	600	Isoflurane

S4. S-Trap protocol for mass spectrometry sample preparation

(cross-reference pg.12)

For the sample preparation preceding mass spectrometry (MS) analysis, the S-trap protocol was performed on the 5% SDS-mediated lysates, in which poorly soluble molecules are also dissolved. This protocol was executed using either S-trap™ mini columns (100-300 µg) or the 96-well S-trap™ plate (100-300 µg per well). Firstly, the proteins were reduced, alkylated and acidified by adding tris(2-carboxyethyl)phosphine (TCEP), iodoacetamide (IAA) and phosphoric acid to final concentrations of 5 mM, 20 mM and 1.2%, respectively. The samples were then diluted with a binding/wash buffer containing 90% methanol (MeOH) in 100 mM Tetraethylammonium bromide (TEAB) at pH 7.4 in a 1:7 ratio and loaded per 400 µL onto the filter of the separate columns or the wells of the 96-well plate. The filters of the columns and 96-well plate underwent three washes with 200 µL of binding/wash buffer. These loading and wash steps were alternated with centrifugation steps at 4,000 x g and 1,500 x g at room temperature for 1-2 min and 5 min for the columns and plate, respectively. Following the final wash step, an additional centrifugation step was performed to completely dry the filters. Subsequently, 125 µL of 0.5 µg trypsin in 50 mM TEAB was added to digest the proteins bound to the S-trap filter and the samples were incubated overnight at 37°C. For peptide elution, the S-trap filter was transferred to a clean tube or a clean collection plate. Three elution buffers, namely 50 mM TEAB, 0.2% formic acid (FA) in dH_2O and 50% acetonitrile (ACN) with 0.1% or 0.2% FA in dH_2O, were added in 80 µL to the filters, alternated by centrifugation at 4,000 x g and 1,500 x g at room temperature for 1-2 min and 5 min for columns and plate, respectively. The eluted samples were then transferred to MS vials and further prepared for peptide concentration measurement and DIA LC-MS/MS analysis.

S5. Nanoparticle Tracking Analysis Zetaview protocol

(cross-reference pg.12)

Start-up routine

1. Switch instrument on (button at the back)
2. Start computer and enter password
3. Open Zetaview software (not the Zetaview analysis software!)
4. The instrument will ask you to load a water sample
   a. 5 ml syringe
   a. Push 5-10 ml MilliQ through
   b. click ok
5. Calibrate the instrument if you are the first user of the day
   Important: your calibration standard should be at the same temperature as the samples you will measure. If you want to measure samples at room temperature, use MilliQ stored at room temperature. If your samples are stored on ice, use MilliQ coming from the fridge.
   - Calibration beads: 100 nm beads - have to be diluted 250.000x. Beads are stored at room temperature, on the shelf above the Zetaview.
     Vortex the beads first!
   - step 1: make a 1000x dilution: 10 ml MilliQ + 10 µl beads.
     The dilution can be stored at 4°C for 1 week
   - step 2: make the 250.000x dilution: 5 ml MilliQ + 20 µl 1000x diluted beads.
     Can only be used for 30 minutes!
   - step 3: load the 250.000x diluted beads onto the instrument
     (5ml syringe - push 2.5 ml in)
   - click ok - the instrument will autofocus
   - The standard contains $1{,}81 \times 10^{13}$ particles. Check if the instrument measures this amount (sensitivity and shutter: 70; dilution factor: 250.000x)

   Remark: if the temperature in the room changes during the day or when you see the particles you are measuring becoming blurry: load beads again and autofocus by pressing the “optimized focus” button.

Adjust the exit part of the chamber depending on your need:

- samples that you want to recover (for example CSF samples): use the short tube with a waste beaker underneath
- samples that you don’t want to recover: use the waste bottle connected to the chamber

Loading a sample

1. Flush the chamber with 15-20 ml PBS - check if the chamber is clean (particle free)
   This step is highly important to avoid horizontal drift. Make sure that the chamber is completely saturated with the buffer you are diluting in your sample in (in our case, this is PBS). Do not flush the chamber with MilliQ or another buffer in between, otherwise you will create a horizontal drift.
2. Flush the chamber with air. This step is only needed if you want to recover your sample. If this is not the case, you can immediately load your sample after the PBS.
3. Dilute your sample (total volume: 900 µl) - turn eppendorf a few times upside down (do not vortex to avoid nanobubbles!)
   sample preparation:
   a. add 960 µL PBS to 40 µL sample (total volume of 1 mL: 25 dilution)
   b. pipet up & down to get all the particles homogeneously in suspension
   c. push up & down with the syringe (for the same reason)

d. load the syringe (impossible to avoid an air bubble)
e. invert the syringe and tick against it to make the air bubble move towards the exit of the entrance
f. carefully push out the air bubble

4. Push sample in completely (900µl) (remark: if you have air bubbles, push sample out again with air and restart) and evaluate the calibration results.
   For CSF samples, we are able to dilute the sample in a total volume of 600µl. After your sample, push a little bit of air in (around 300µl) to lower the drift (reason: 600 µl is too little to fill the chamber completely - push in some extra air to make sure your sample is in the middle of the chamber).

Tab cell check

1. Position: one of the 11 positions the camera is at. You can check some positions to make sure there is no air bubble.
2. Scattering intensity: should be green. If not: dilute sample more.
3. Number of detected particles should ideally be between 100-150 (green between 50-200). Red signal: dilute sample more.
4. Dilute
5. Check particle drift: click check -> green = ok. Not ok: check if the valves are still connected well. Click button again to stop! You don't have to do this before every measurement, but check regularly. Sometimes you have to wait a little bit until the sample stabilizes.

Measurement

Tab pump & temp

Set the temperature to 23°C - switch on
Wait until this temperature is reached before starting your measurement (you can check the current temperature of the instrument in the left bottom corner)
Remark: The assembly we use makes cell recovery possible, but we can't use a pump or measure the zeta potential

Tab measurement

It is very important to use exactly the same settings if you want to compare your samples/experiments! SOPs will be made for all our sample types in basal conditions (CSF, medium, plasma) - use the right SOP with the right settings!

Run video acquisition

1. Custom entry: name you want to give your entry
2. Path where you want to store your data. We store our data on the z-drive
3. SOP: load the SOP for your sample type. Important: always check if the right SOP is selected!
4. Reload SOP: button becomes green

Analysis

During the analysis, you can already do the following steps:

1. Push your sample out with a 1ml syringe of air - collect sample. If you don't need to recover your sample, you can immediately flush with PBS.
2. Flush again with PBS (1 ml is sufficient)
3. Load your next sample

Important: check from time to time if the chamber is still clean. If not, flush with more PBS.

Analysis

- The software itself shows which positions it thinks are outliers - check for yourself! if you want to include an outlier, press “use” and an “X” will appear. Standardly, we decided to always use all measurements.
- The size average should be more or less the same in every position, otherwise the chamber is not clean. If this is the case: clean the instrument and start again.
- The amount of traces (“traces found”) has to be between 500 and 1000, otherwise your sample is too diluted.
  Improving the amount of traces:
    higher the amount of cycles
    higher the length of the video
- If you are not happy with your analysis, you can change some parameters (remark: to compare this analysis with other data, also the other data will have to be analyzed again!) and re-analyse the data.

Shutting down

Cleaning solution: stored on the shelf above the Zetaview

1. Flush the chamber with MilliQ x2
2. Flush the chamber with air x2
3. Add 2 drops of the blue bottle (cleaning kit) to 50 ml of MilliQ. This dilution is stored on the shelf above the Zetaview, mastermix is stored in the cleaning kit.
4. Make a swab (cleaning kit) wet by dipping it in the cleaning solution
5. Remove the valves on the entrance and exit of the chamber by turning and pulling them off
6. Put swab in chamber - move around 5-10 times - take swab out
7. Put valves back on
8. Push MilliQ through the chamber (3x 10ml)
9. Push air (3x10ml) through the chamber and leave it like this
10. Shut down the instrument and the computer

What if the chamber is dirty?

- Little bit dirty: swab dipped in MilliQ or cleaning solution (MilliQ + 2 drops blue bottle) from cleaning kit (see shutting down procedure)
  Person from the company recommended to always use the blue bottle and not the green bottle.
- Very very dirty:
    - remove chamber by pulling the button underneath the chamber down - switch it to red position
    - take chamber off carefully!
    - remove dust or open up the chamber - clean (20-50% acetone on a cotton swap)
    - put chamber back on - turn the button underneath again to the white position
    - put your fingers underneath the chamber and carefully push it back until you hear a click

S6. Proof-of-principle using 0.1% TWEEN-20

(cross-reference Part 7.1 pg.14, Part 7.2 pg.29)

The effect of 0.1% TWEEN-20 on the EV purity employing FAEVEr was explored using 2.5 mL of recombinant extracellular vesicle (rEV) conditioned medium diluted 1:1 with varying concentrations of TWEEN-20 (final concentrations of 0%, 0.1% and 5%) in PBS. Our results indicated that 5% TWEEN-20 results in the most drastic reduction in chromatographic profile complexity compared to 0% TWEEN-20, and that this effect is already observed at the lower concentration of 0.1% TWEEN-20, as illustrated in Figure S1**A**. Furthermore, 0.1% TWEEN-20 leads to a higher number of protein identifications compared to 5% TWEEN-20, with a greater fraction of identified human proteins, as shown in Figure S1**B**. In conclusion, lower concentrations of TWEEN-20 starting from 0.1% hold promise for enhancing the FAEVEr setup for proteomics-based EV protein biomarker discovery with DIA LC-MS/MS.

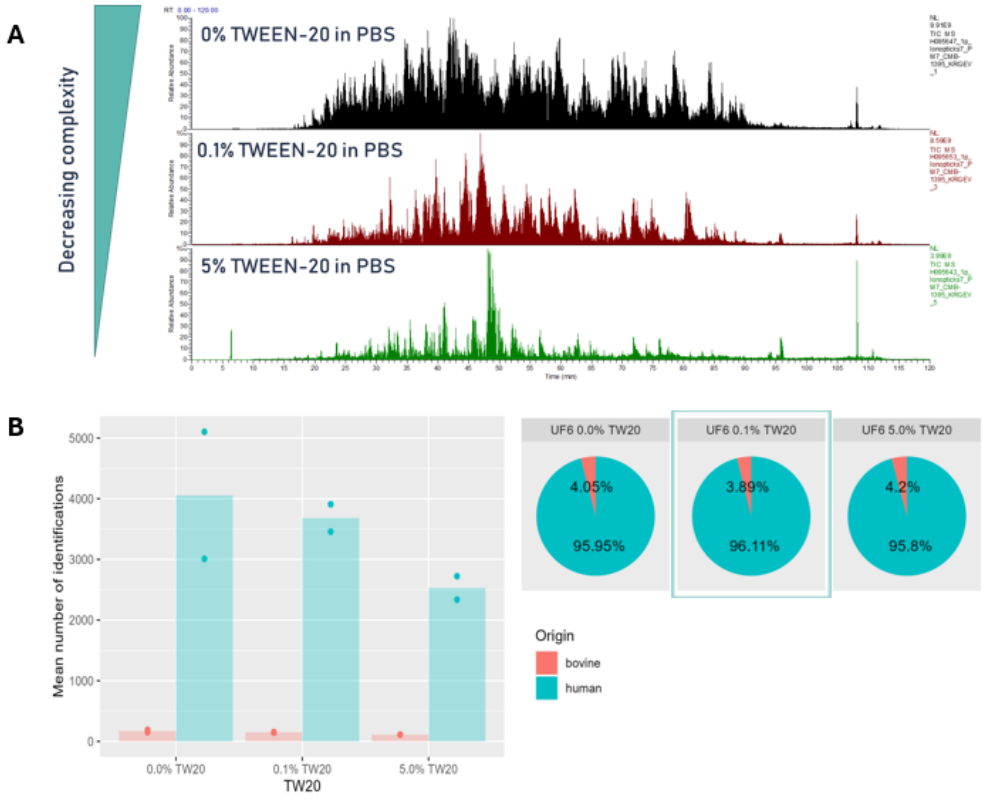


*Figure S1: **A)** Visual representation of the decreasing complexity in the chromatographic profile over increasing concentrations of TWEEN-20 (0% - 0.1% - 5%) in PBS using FAEVEr on 6 mL filter tubes. **B)** The average absolute and relative number of identifications per TWEEN-20 concentration (0% - 0.1% - 5%) in PBS using FAEVEr represented in bar plots and pie plots, respectively, with bovine proteins in red and human proteins in blue.*

S7. Large-scale comparison UF6-UF96-UC

Table Nanoparticle concentrations

(cross-reference pg.20)

Table S4: Original concentration of nanoparticles calculated with the ZetaView Software for each TWEEN-20 concentration (0% - 0.1% - 0.5% - 1.0% - 5.0%) per EV enrichment strategy: ultracentrifugation, FAEVEr (filter-aided EV enrichment) on 6 mL filter tubes or on a 96-well filter plate.

Ultracentrifugation	**original concentration (particles / mL)**	***FAEVEr on 6 mL filter tubes***	**original concentration (particles / mL)**	***FAEVEr on a 96-well filter plate***	**original concentration (particles / mL)**
0% TWEEN-20	3.10E+09	*0% TWEEN-20*	1.40E+08	*0% TWEEN-20*	5.00E+08
0.1% TWEEN-20	5.40E+09	*0.1% TWEEN-20*	3.30E+08	*0.1% TWEEN-20*	1.70E+08
0.5% TWEEN-20	6.00E+09	*0.5% TWEEN-20*	1.40E+08	*0.5% TWEEN-20*	5.30E+08
1% TWEEN-20	4.30E+09	*1% TWEEN-20*	8.70E+07	*1% TWEEN-20*	1.50E+08
5% TWEEN-20	8.10E+09	*5% TWEEN-20*	8.10E+07	*5% TWEEN-20*	8.30E+07

Figure relative number of protein identifications human-bovine

(cross-reference pg.23)

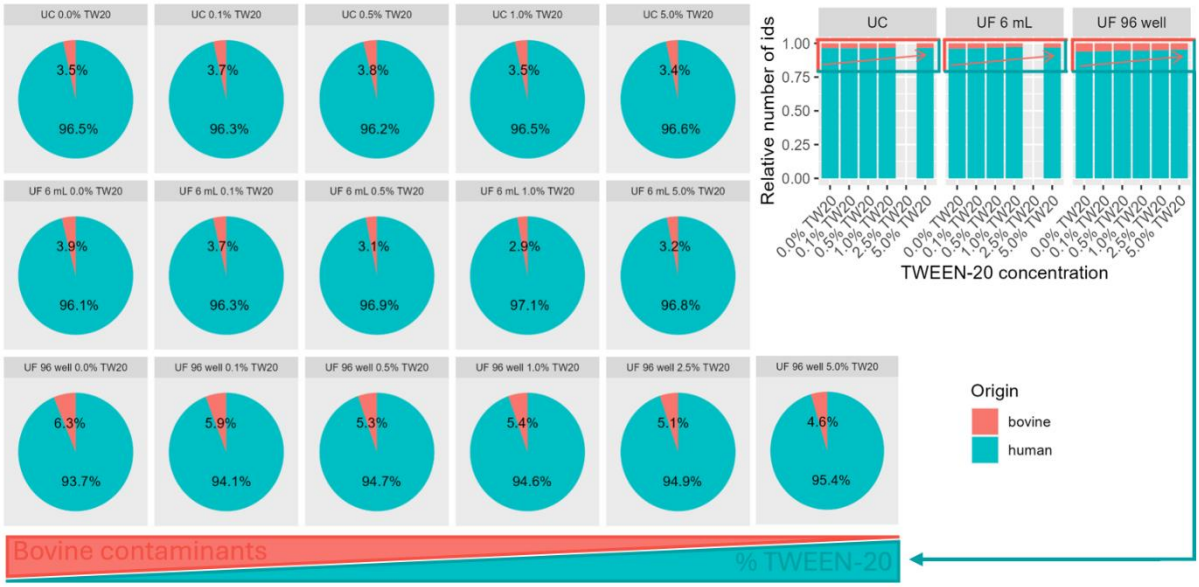


Figure S2: Visualization of the relative number of protein identifications in the EV-enriched fractions from ultracentrifugation (UC), FAEVEr on 6 mL capacity (UF6) or on 96-well capacity (UF96) in pie plots per TWEEN-20 concentration (TW20, 0% - 0.1% - 0.5% - 1.0% - 2.5% - 5.0%). The fraction of identified bovine proteins is colored in red and the fraction of identified human proteins in blue.

S8. Applications

S8.1 EV enrichment from 20% FBS and 20% EDS with FAEVEr on 6 mL filter tubes using 0.1% TWEEN-20

(cross-reference pg.30)

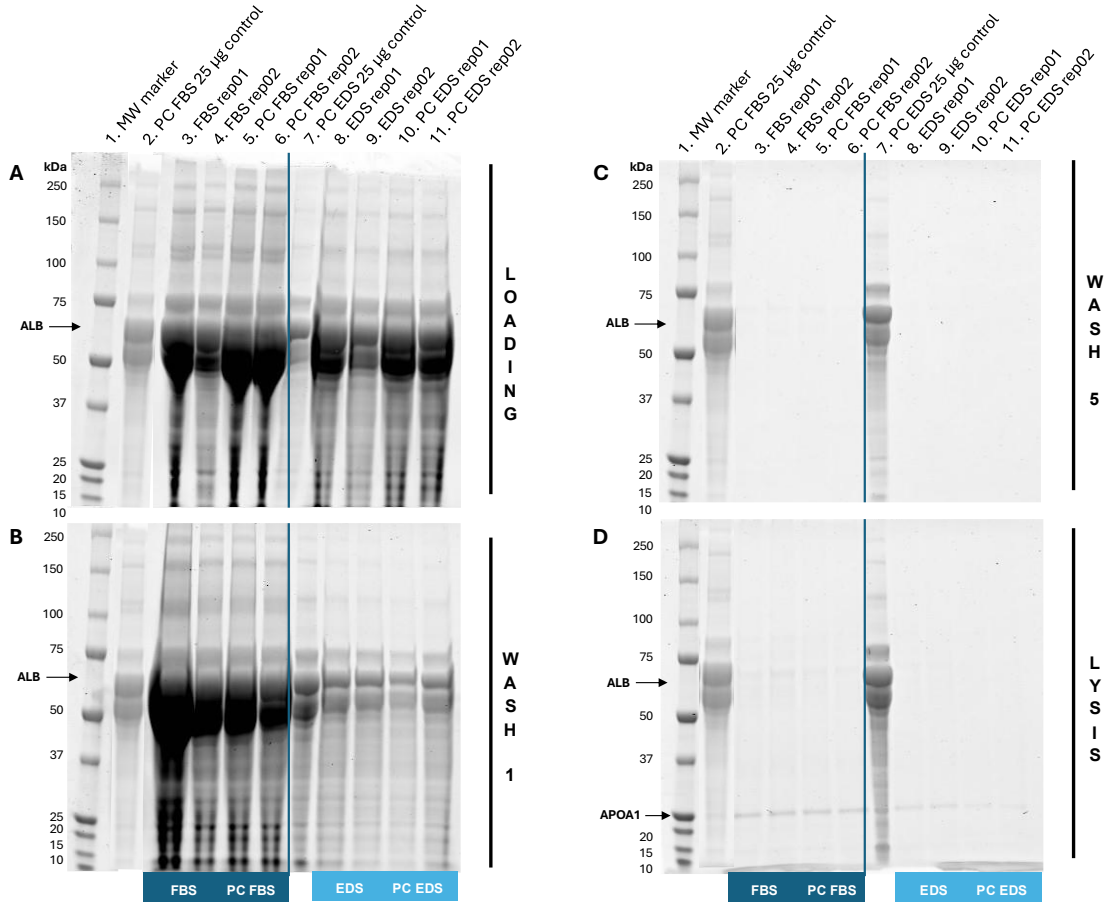


*Figure S3: SDS-PAGE with Coomassie staining of the different fractions (**A.** FT, **B.** W1, **C.** W5, **D.** Lysate) obtained during FAEVEr on 6 mL filter tubes (UF6) from 20% (PC) FBS and 20% (PC) EDS with 0.1% TWEEN-20 in PBS. PC FBS/PC EDS is used as positive control. ALB and APOA1 are indicated with black arrows at approximately 65 kDa and 25kDa. [Abbreviations: FT = flow-through, W = washing step, (PC) FBS = (pre-cleared) fetal bovine serum, (PC) EDS = (pre-cleared) EV depleted fetal bovine serum, ALB = albumin, APOA1 = apolipoprotein A1]*

S8.2 EV enrichment from different FBS concentrations

Differential experimental setup for low and high throughput format

(cross-reference 7.2.2 UF6 (a, b) pg.32 and 7.2.2 UF96 (c) pg.35)

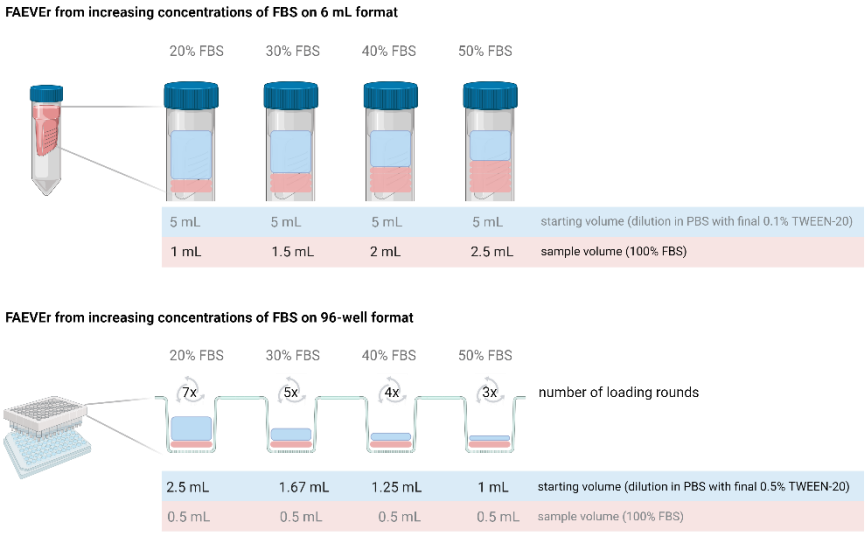


Figure S4: Experimental setup of filter-aided EV enrichment (FAEVEr) on UF6 and UF96 format from FBS concentrations. The different dilution approaches are illustrated.

FAEVEr on 6 mL filter tubes using 0.1% TWEEN-20
(cross-reference pg.32)

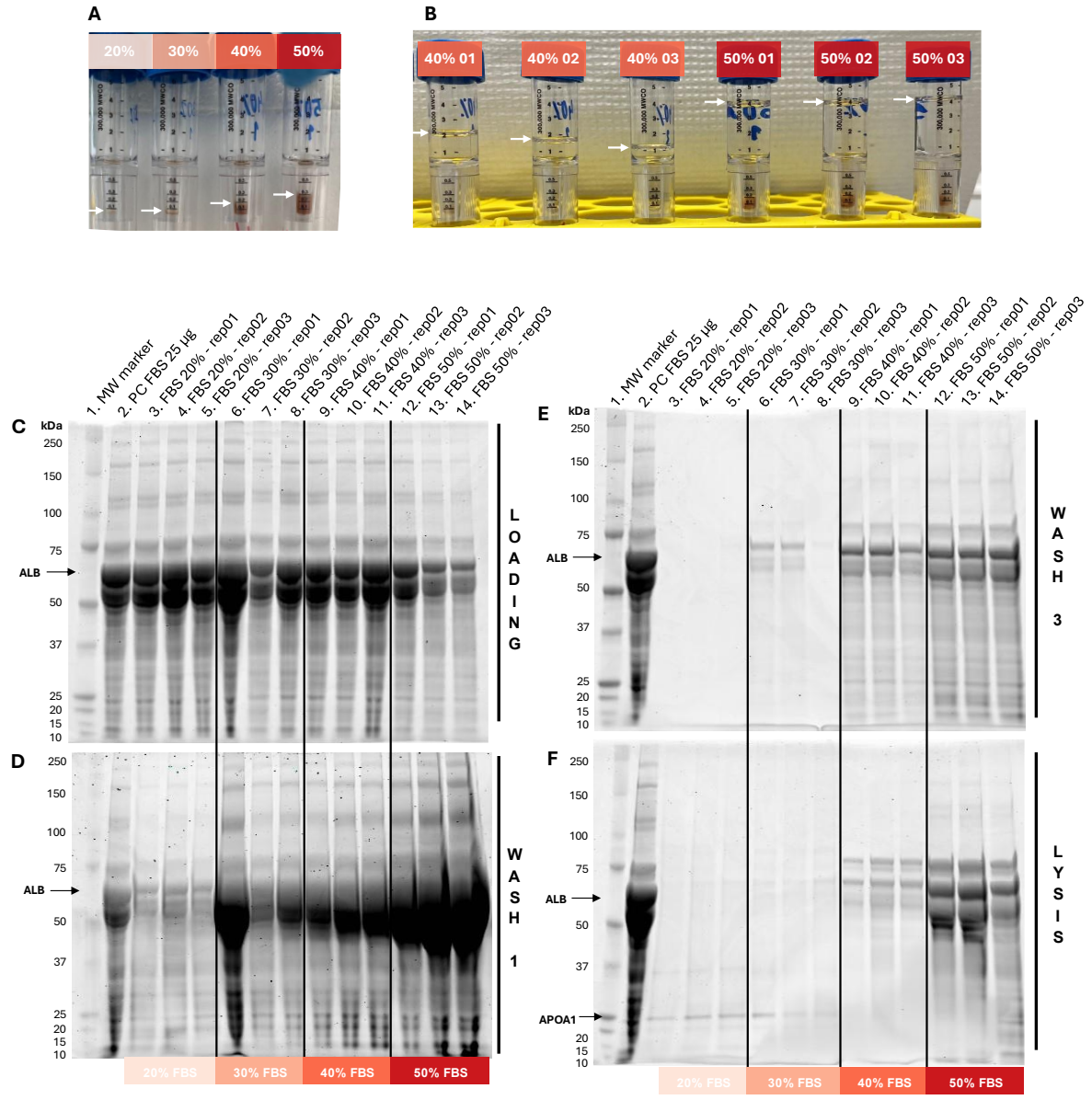


Figure S5: Evaluation of the filtration performance of FAEVEr on 6 mL filter tubes (UF6) from 20%-30%-40%-50% FBS with 0.1% TWEEN-20 in PBS. ***A)*** *All FBS concentrations equal to or higher than 30% show fouling of the filter with protein material to some extent.* ***B)*** *The 40% and 50% FBS replicates demonstrate filter clogging after the second wash step with great replicate variability for the 40% FBS condition. SDS-PAGE with Coomassie staining of the different fractions (****C.*** *FT,* ***D.*** *W1,* ***E.*** *W3,* ***F.*** *Lysate) obtained during the protocol. PC FBS is used as positive control. ALB and APOA1 are indicated with black arrows at approximately 65 kDa and 25 kDa, respectively. [Abbreviations: FT = flow-through, W = washing step, (PC) FBS = (pre-cleared) fetal bovine serum, (PC) EDS = (pre-cleared) EV depleted fetal bovine serum, ALB = albumin, APOA1 = Apolipoprotein A-I]*

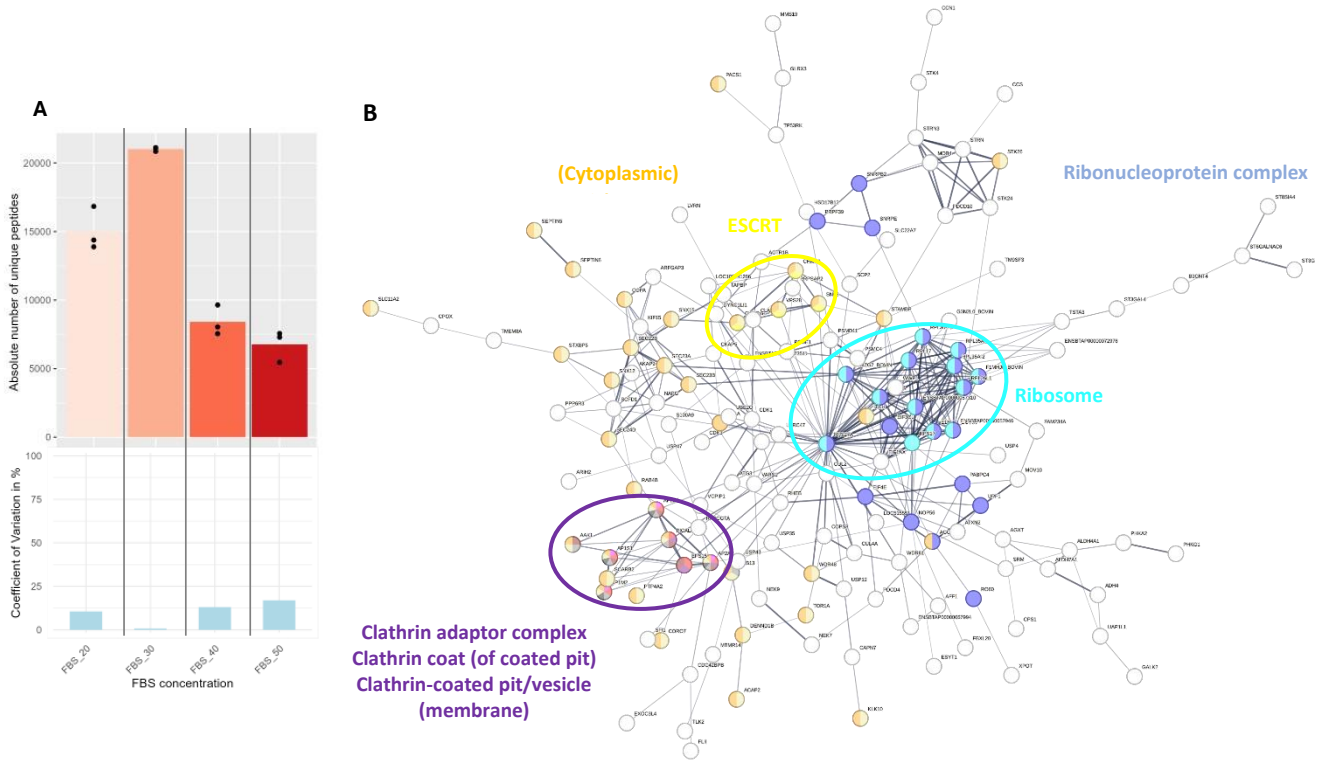


Figure S6: Additional parameters indicating more optimal EV enrichment for proteome analysis from 20% and 30% FBS using FAEVEr on 6 mL filter tubes with 0.1% TWEEN-20 washing steps. ***A)*** *Representation of the number of unique peptides per FBS concentration on average in bars and per replicate in dots in the upper panel. The corresponding coefficient of variation (in %) is indicated in the lower panel.* ***B)*** *Protein-protein interaction network in STRING of all proteins that are uniquely identified in the 20% and 30% FBS conditions indicates enrichment of ESCRT-related proteins (in yellow).*

FAEVEr on 96-well filter plate using 0.5% TWEEN-20
(cross-reference pg.35)

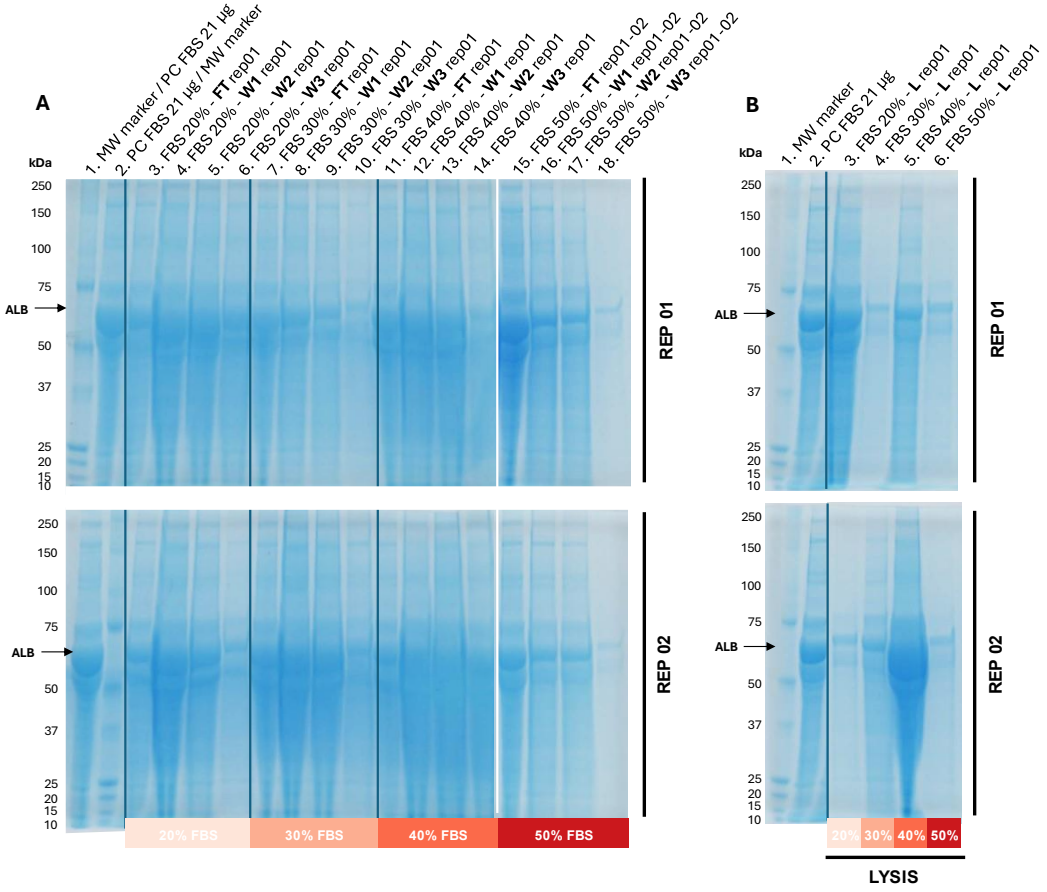


*Figure S7: SDS-PAGE with Coomassie staining of the different fractions (**A.** FT-W1-W2-W3, **B.** Lysate) obtained during FAEVEr on a 96-well filter plate (UF96) from 20%-30%-40%-50% FBS with 0.5% TWEEN-20 in PBS. PC FBS is used as positive control. ALB is indicated with a black arrow at approximately 65 kDa. [Abbreviations: FT = flow-through, W = washing step, (PC) FBS = (pre-cleared) fetal bovine serum, ALB = albumin]*

S8.3 EV enrichment from pancreatic ductal adenocarcinoma cell lines conditioned medium with FAEVEr on a 96-well filter plate using 0.5% TWEEN-20

(cross-reference pg.36)

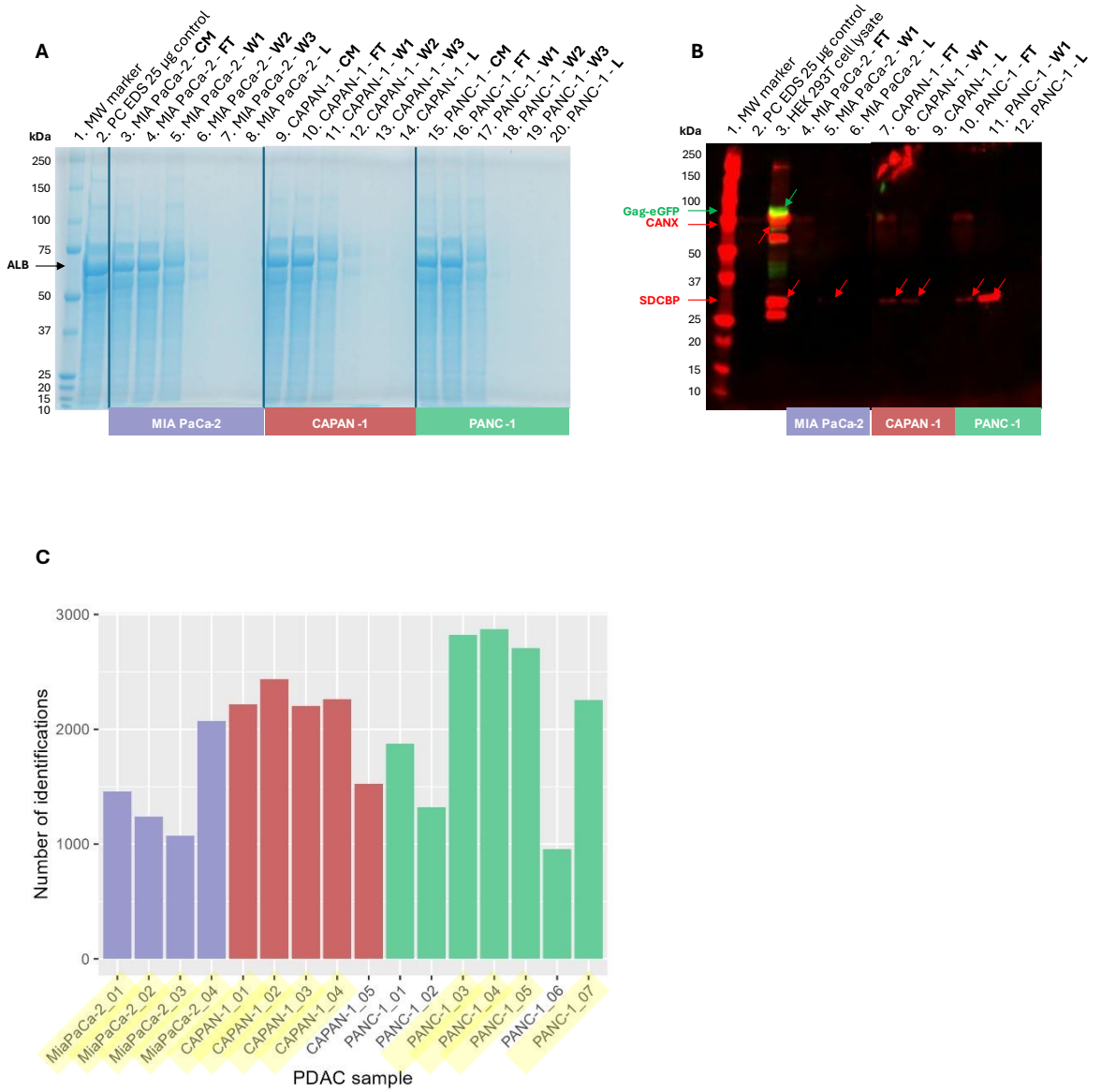


Figure S8: Evaluation of the performance of FAEVEr on a 96-well filter plate (UF96) from 500 µL CM of MIA PaCa-2, CAPAN-1 and PANC-1 PDAC cell lines with 0.5% TWEEN-20 in PBS. ***A)*** *SDS-PAGE with Coomassie staining of the different fractions (FT, W1-W3, L) obtained during FAEVEr. ALB is indicated with a black arrow at approximately 65 kDa.* ***B)*** *Western blot of these fractions for SDCBP, CANX and Gag-eGFP that are indicated with arrows at 32 kDa, 75 kDa and 84 kDa respectively* ***C)*** *Absolute number of identifications for all initially analysed PDAC samples. Those selected for further data analysis (four for every condition) are highlighted. [Abbreviations: CM = conditioned medium, PDAC = pancreatic ductal adenocarcinoma, FT = flow-through, W = washing step, L = lysate, EDS = EV-depleted fetal bovine serum, ALB = albumin, SDCBP = syntenin-1, CANX = calnexin]*

S8.4 EV enrichment from non-infected and influenza-infected mice sera with FAEVEr on 6 mL filter tubes using 0.5% TWEEN-20

(cross-reference pg.39)

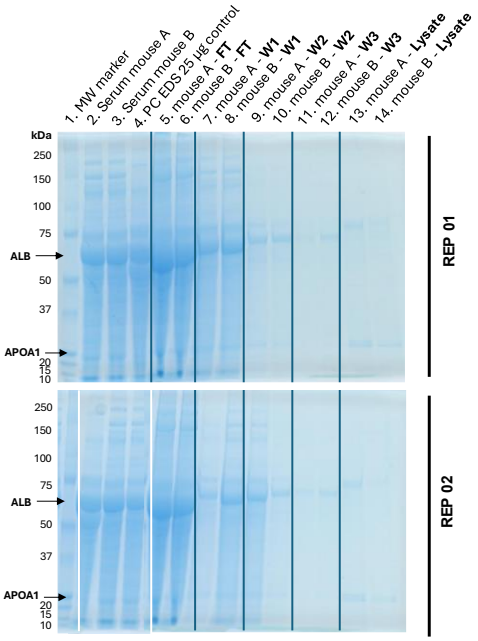


Figure S9: SDS-PAGE with Coomassie staining of the different fractions (FT, W1-W3, Lysate) obtained during FAEVEr on 6 mL filter tubes (UF6) from 500 µL mouse serum of WT – (mouse A) and influenza infected mice (mouse B) with 0.5% TWEEN-20 in PBS per duplicate. ALB and APOA1 are indicated with black arrows at approximately 65 kDa and 25 kDa. [Abbreviations: WT = wild-type, EDS = EV-depleted fetal bovine serum, ALB = albumin, APOA1 = apolipoprotein A1]

Table S5: Overview of the top 10 abundant proteins from all identified proteins with LC-MS/MS DIA after EV enrichment employing FAEVEr with 0.5% TWEEN-20 from 500 µL of mouse serum from influenza infected mice.

Protein	Gene
Apolipoprotein A-I	Apoa1
Immunoglobulin heavy constant mu	Ighm
Hemoglobin subunit alpha	Hba
Immunoglobulin kappa constant	Igkc
Immunoglobulin kappa chain variable 6-17	Igkv6-17
Hemoglobin subunit beta-1	Hbb-b1
Apolipoprotein E	Apoe
Albumin	Alb
Apolipoprotein A-IV	Apoa4
Hemoglobin subunit beta-2	Hbb-b2

S8.5 EV enrichment from mouse plasma of breast cancer PDX mouse models with FAEVEr on 6 mL filter tubes using 0.5% TWEEN-20

(cross-reference pg.40)

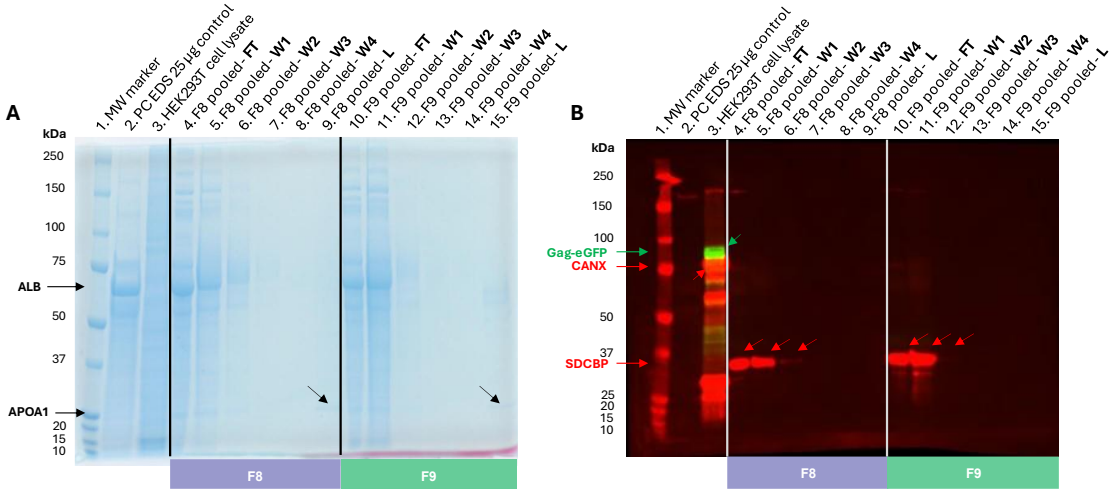


Figure S10: Evaluation of the performance of FAEVEr on filter tubes (UF6), with 0.5% TWEEN-20 in PBS, from 123 µL – 600 µL mouse plasma of breast cancer PDX models. ***A)*** *SDS-PAGE with Coomassie staining of the different fractions (FT, W1-W3, L) obtained during FAEVEr but pooled per mouse generation (F8 or F9). ALB and APOA1 are indicated with black arrows at approximately 65 kDa and 25 kDa, respectively.* ***B)*** *Western blot of these fractions for SDCBP, CANX and Gag-eGFP that are indicated with arrows at 32 kDa, 75 kDa and 84 kDa respectively. [Abbreviations: EDS = EV-depleted fetal bovine serum, FT = flow-through, W = washing step, L = lysate, ALB = albumin, APOA1 = apolipoprotein A1, SDCBP = syntenin-1, CANX = Calnexin]*

Table S6: Overview of the top 10 abundant proteins from all identified proteins ***(A)*** *and from all identified human proteins* ***(B)*** *with LC-MS/MS DIA after EV enrichment employing FAEVEr with 0.5% TWEEN-20 from varying volumes of mouse plasma (123 µL – 600 µL) from breast cancer PDX models.*

A

Protein	Gene
Apolipoprotein A-I	Apoa1
Albumin	Alb
Kininogen-1	Kng1
Apolipoprotein E	Apoe
Clusterin	Clu
Fibrinogen gamma chain	Fgg
Fibrinogen alpha chain	Fga
Histidine-rich glycoprotein	Hrg
Apolipoprotein A-IV	Apoa4
Fibrinogen beta chain	Fgb

B

Protein	Gene
Keratin, type II cytoskeletal 1	KRT1
Keratin, type I cytoskeletal 10	KRT10
Keratin, type I cytoskeletal 9	KRT9
Calpain-1 catalytic subunit	CAPN1
Elongation factor Tu, mitochondrial	TUFM
Cilium assembly protein DZIP1	DZIP1
Serum paraoxonase/lactonase 3	PON3
Intermembrane lipid transfer protein VPS13A	VPS13A
Keratin, type II cytoskeletal 2 epidermal	KRT2
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH