

## Article

# Characterization of extracellular vesicles secreted in lentiviral producing HEK293SF cell cultures

Aline Do Minh <sup>1</sup>, Alexandra T. Star <sup>2</sup>, Jacek Stupak <sup>2</sup>, Jean-François Gélinas <sup>1,a</sup>, Jianjun Li <sup>2</sup>, Susan M. Twine <sup>2</sup> and Amine A. Kamen <sup>2,\*</sup>

<sup>1</sup> Department of Bioengineering, McGill University, Montreal, QC, Canada

<sup>2</sup> Human Health Therapeutics Research Centre, National Research Council Canada, Ottawa, ON, Canada

\* Correspondance: amine.kamen@mcgill.ca

<sup>a</sup> Current address: Département de Médecine, Université de Montréal, Montréal, QC, Canada

**Abstract:** Lentiviral vectors (LVs) are a powerful tool for gene and cell therapy and human embryonic kidney cells (HEK293) have been extensively used as a platform for production of these vectors. Like most cells and cellular tissues, HEK293 cells release extracellular vesicles (EVs). EVs released by cells share similar size, biophysical characteristics and even a biogenesis pathway with cell-produced enveloped viruses, making it a challenge to efficiently separate EVs from LVs. Thus, EVs co-purify with LVs during downstream processing, becoming “impurities” in the context of cell therapy. To characterize EVs from an inducible lentivirus producing cell line, two conditions were studied: non-induced and induced. EVs’ identity was confirmed by transmission electron microscopy and western blot. Seven proteins were identified by mass spectrometry as potential EV markers. Lipid composition of EVs and LVs showed similar enrichment in phosphatidylserine. RNA cargos in EVs showed enrichment in genes involved in viral processes and binding functions. Flow virometry, GTA and ddPCR results also confirmed the heterogenic nature of EVs and LVs populations. These findings provide insights on the product profile of lentiviral preparation and could help develop separation strategies of co-produced EVs.

**Keywords:** extracellular vesicles, enveloped viruses, lentiviral vectors, exosome, proteomic, lipidomic, transcriptomic

## 1. Introduction

In the past decade, gene and cell therapies have become increasingly popular tools to treat diseases such as genetic disorders, cancer, cardiovascular disease, as well as a wide spectrum of orphan diseases [1]. Recently, the cell therapy field reported significant clinical achievements, as it is the case with Chimeric Antigen Receptor T cell (CAR-T) therapy, where the patient’s own immune cells are modified to express a surface receptor to stimulate an immune response against cancer cells [2]. Different viruses have been engineered to be used in gene and cell therapies as delivery vectors. Adenovirus, adeno-associated virus (AAV) and lentiviral vectors (LV) have become dominant in the field [3].

LV have several advantages over other viral vectors [4]. Their ability to mediate long-term therapeutic transgene expression [5] makes them the ideal candidate for cell therapy. However, challenges such as achieving sufficiently high yield and suitable purity for *in vivo* and *ex vivo* clinical applications need to be addressed. This is particularly crucial for large scale productions to meet the needs of large population treatments other than orphan diseases [6]. In this context, the discovery of and renewed interest in extracellular vesicles (EVs) raise unprecedented questions as some EVs, such as exosomes and small shedding microvesicles, not only share size distribution with lentiviral particles, but also biochemical and biophysical properties [7].

EVs are cell membrane-derived vesicles that are found in most cells and body fluids. The field of EVs has gained considerable attention in the past few years and their potential as drug delivery



vehicles and biomarkers for diseases is actively investigated [8]. EVs are known to transport lipids, proteins and nucleic acids. The cargo composition of EVs depends on many features, such as cell type and the cell environment or medium for in vitro cultures. However, the mechanism behind cargo sorting is not well understood [9].

Databases have been created to compile researchers' findings on EV characterization [10,11]. Definitive markers are currently not established, only enriched proteins can be noted, and they depend on the EV subtype. For instance, Endosomal Sorting Complex Required for Transport (ESCRT) machinery proteins (ALIX, TSG101, CD63, CD81 and CD9) are highly enriched in exosomes, while MMP2 and CK18 are mostly found in shedding microvesicles. EVs also have the ability to transport RNA. Coding RNA, as well as non-coding RNA, were reported in next-generation sequencing studies, revealing the presence of miRNAs in EVs' cargo which are involved in transcription regulation, post-transcription regulation and sometimes viral defence [12]. The lipid content of EVs is also important as EVs are enclosed within a single phospholipid bilayer with the lipid composition resembling that of the cell plasma membrane. In addition, exosomes are highly enriched in glycosphingolipids, sphingomyelin, cholesterol and phosphatidylserine. EV membranes also contain lipid-raft micro-domains, which are notably involved in virus morphogenesis and budding [13].

EVs and retroviruses share a biogenesis pathway using the ESCRT machinery, they incorporate similar host cell components as well as viral components [14], and also share biophysical and biochemical properties, making their separation challenging. Typical purification methods, such as chromatography based on charge or size will be ineffective at discriminating EVs and LVs. This problem needs to be addressed since EVs are released concomitantly by the cells and, thus, will be found in lentiviral preparations. As lentiviral-mediated gene therapies are intended for human use, they are strictly regulated by health authorities and any impurities in the viral preparation have to be documented as per regulatory requirements [15]. Indeed, impurities such as host cell proteins and host cell DNA are accepted on defined level. EVs, which contain both, should also be extensively characterized as to be able to set product specifications.

Many studies have been conducted to characterize EVs isolated from different biological fluids, tissues and even cultured cells. However, only few studies focus on cell lines used to produce viruses for vaccination or gene and cell therapy [16,17]. Moreover, these studies centered their attention on virus-like-particles versus EVs, which influenced their choice of separation technique. Methods such as step ultracentrifugation (UC), sucrose cushion used by Venereo et al. [16], or processes involving the qEV SEC column with a sample loading volume of  $< 500 \mu\text{L}$  have very low throughputs due to the volume limitation of the techniques. Additionally, these methods are labor intensive, not controlled and, therefore, would induce high variability in the yield of isolated EVs. These processes are also not scalable to produce large volume of samples when extensive analyse is required. Here, we want to emphasize the use of the human embryonic kidney (HEK293) cell line to produce viral vectors for gene and cell therapies [18]. Thus, they were selected as the cell culture platform in this study. Like most cells, HEK293 cells continuously generate EVs, which will be difficult to separate from LV concomitantly produced in these cell cultures. Therefore, gaining an understanding of the characteristics of EVs generated during LV production will provide an accurate product profile for LV-mediated gene therapies, and eventually, insights to improving the LV purification process. LV production in HEK293 cells can be achieved by different methods [19]: by transient transfection using 3 to 4 plasmids, using packaging cell lines where necessary genetic elements for the assembly and functioning of the vectors have been stably integrated, or using producer cell lines where the remaining transgene plasmid has been integrated. In this study, we developed a scalable process to isolate EVs from cultures of an inducible HEK293 lentivirus (Clone 92) producing cell under no-inducing conditions on one hand to extensively characterize isolated EVs for proteomic, lipidomic, and transcriptomic content. Furthermore, EVs from Clone 92 cells under no-inducing conditions were compared to the resulting products after induction of lentiviral production. This will shed light on markers that can be exploited for separation approaches.

## 2. Materials and Methods

### 2.1. Cell culture of HEK293SF cells in suspension

As a platform for lentiviral vector (LV) production, HEK293SF cell line (abbreviated hereafter as 293SF) and a stable producer cell line developed by the National Research Council Canada (NRC), HEK293SF-LVP-CMVGFPq-92 (abbreviated hereafter as Clone 92) were used in this study [20,21]. Production of the LVR2-GFP (rHIV.VSV-g CMV GFP) lentiviral vector is induced in the Clone 92 cell line by the addition 1  $\mu$ g/mL (w/v) doxycycline hydulate (Millipore Sigma, Etobicoke, Canada) (from a 1 mg/mL stock in nuclease-free water) and 10  $\mu$ g/mL (w/v) 4-isopropylbenzoic acid (cumate) (Millipore Sigma) (from a 10 mg/mL stock in ethanol absolute) to produce a third-generation SIN HIV-based lentiviral vector which expresses the green fluorescence protein (GFP). 293SF and Clone 92 cells were cultured in shake flasks (from 20 to 300 mL working volumes) in HyCell TransFx-H medium (GE Healthcare, Chicago, Illinois) supplemented with 4-6 mM L-Glutamine or GlutaMAX™ (ThermoFisher Scientific, Waltham, Massachusetts) and 0.1% Kolliphor (Millipore Sigma) without serum or antibiotics, or in HEK GM medium (Xell AG, Bielefeld, Germany) supplemented with 4-6 mM L-Glutamine or GlutaMAX™ (ThermoFisher Scientific). Cell growth was monitored by determining live cell density based on the principle of trypan blue dye exclusion on a Vi-Cell XR cell counter (Beckman Coulter, Brea, CA, United States of America (USA)). Cells were passaged twice a week by diluting to  $2.0 \times 10^5$  live cells per mL in fresh medium.

HEK293A cells (American Type Culture Collection, Manassas, VA, USA) were used for the gene transfer assay (GTA) [22]. They were maintained in a humidified incubator at 5% CO<sub>2</sub> and 37 °C in Dulbecco's Modified Eagle's Medium (DMEM) (Wisent, St-Bruno, Canada), supplemented with 2 mM L-Glutamine and 5% Fetal Bovine Serum (FBS) (Corning Inc., Corning, New York, NY, USA) without antibiotics. Cells were passaged twice a week.

### 2.2. Production of conditioned medium containing EVs

293SF and Clone 92 (under non-induced conditions) cell lines were cultivated and the cell density was measured every day. When the cell density reached  $1 \times 10^6$  cells/mL, the cells were kept in culture for 2 additional days before harvest.

### 2.3. EV isolation

#### 2.3.1. Ultrafiltration (UF) and size exclusion chromatography (SEC)

EVs in non-LV producing conditions from Clone 92 cell cultures were isolated using a combination of ultrafiltration followed by size exclusion chromatography as it was reported that this technique could yield more intact and pure particles [23,24]. The cells were first removed by centrifugation. The cell pellet was kept at -20°C for further analysis and the supernatant was filtered through a 0.45  $\mu$ m vacuum polyvinylidene fluoride (PVDF) filter (VWR, Ville Mont-Royal, Canada) to remove large particles. The filtrate was then subjected to ultrafiltration and diafiltration (DF) using a Vivaflow™ 50R membrane (Sartorius) with a 100 kDa MWCO pre-flushed with MilliQ water and phosphate-buffered saline (PBS) buffer (Wisent) containing 0.005% Kolliphor. The pressure and volume were monitored throughout the process. This membrane also allowed for large scale processing with volumes up to 1.5 L and reusability. The diafiltered concentrate was then loaded onto a HiScreen™ Capto™ Core 700 SEC column (GE Healthcare) which resin exhibits both size exclusion and binding properties. The Capto Core 700 column was operated in flowthrough mode on an ÄKTA avant (GE Healthcare), providing further control and allowing large volumes to be processed. The flowthrough was collected and stored at -80°C until further analysis. In some cases, the fractions of interest were pooled and subjected to an additional concentration step using a MicroKros 10 kDa MWCO hollow fiber (Repligen, Rancho Dominguez, California) or an Amicon Ultra-4 centrifugal filter unit (Millipore Sigma).

#### 2.3.2. Ultracentrifugation

The induction of Clone 92 cell cultures with cimate and doxycycline generates LV particles which are classified as biosafety level 2 (BSL2) material. As the isolation process described earlier was specifically designed for EVs, involving open handling and use of equipment not suitable for BSL 2 material, ultracentrifugation was used in order to compare EVs in non-LV producing conditions with EVs upon induction of LV production. The supernatant of Clone 92 cell culture, with and without induction, obtained after centrifugation at 1200xg for 5 min, was filtered through a 0.45 $\mu$ m filter and then subjected to a 100,000xg centrifugation for 70 min at 4°C. The pellet was then washed with PBS and centrifuged again at 100,000xg for 70 min at 4°C. The pellet was resuspended in 1 mL of PBS and stored at either 2-8°C or -80 °C until further analysis.

#### 2.4. Nomenclature

Table 1 presents the nomenclature that will be used hereafter for the purpose of clarification. As the result of Clone 92 induction with cimate and doxycycline is a mixed population of EVs and LVs, the nomenclature was chosen to highlight that fact. When designating Clone 92 EVs in general without a specific isolation method, the abbreviation  $C^92EVs$  will be used.

**Table 1.** Nomenclature for EV and LV samples in Clone 92 cell line using different isolation methods in two condition: without induction of LV production or after induction of LV production using cimate and doxycycline.

	No induction	Induction
<b>No isolation</b>	$C^92EV^{sup}$	$C^92EV/LV^{sup}$
<b>Isolation by UF/SEC</b>	$C^92EV^{SEC}$	N/A <sup>1</sup>
<b>Isolation by UC</b>	$C^92EV^{UC}$	$C^92EV/LV^{UC}$

<sup>1</sup> N/A: not applicable

#### 2.5. Quantification of functional viral titer by gene transfer assay (GTA)

A flow cytometry-based GTA was used to determine functional viral titer [21]. Each well of a 24-well plate was seeded with  $1 \times 10^5$  cells of HEK 293A. After leaving the cells adhere to the plate for 5h, the medium was removed. EV and LV samples were serially diluted in DMEM (Wisent) supplemented with 8  $\mu$ g/mL of polybrene (Millipore Sigma) and incubated at 37°C for 30 min. 200  $\mu$ L of diluted sample were then added to the cells for transduction and the plates were incubated overnight at 37°C before addition of 800  $\mu$ L of fresh culture medium in each well the next day. Three days post-transduction (therefore, 48 h after medium addition), cells were harvested and run on the Accuri flow cytometer (Becton Dickinson, Franklin Lakes, New Jersey) to quantify GFP expressing cells. Accepted values ranged between 2-20% fluorescent cells out of total cell count to avoid signal due to super infection.

#### 2.6. Quantification of total particles by digital drop polymerase chain reaction (ddPCR)

RNA was first extracted from LV samples using the High Pure Viral Nucleic Acid Kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. The extracted RNA was then reverse transcribed into cDNA using the iScript™ Select cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, California) according to the manufacturer's instructions and using gene-specific primers targeted towards the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) amplifying a 589-base pair fragment. Primer sequences were: forward primer (5'-GTCCTTCCATGGCTGCTC-3'), reverse primer (5'-CCGAAGGGACGTAGCAGA-3') (Integrated DNA Technologies, Inc., Coralville, IA, USA). Serial dilutions of cDNA were prepared in nuclease-free water. ddPCR reactions were prepared with the QX200™ ddPCR™ EvaGreen Supermix (Bio-Rad) and the WPRE primer set. PCR mixtures (22  $\mu$ L) were prepared for the QX200™ Droplet Generator (Bio-Rad), with final primer concentration of 0.8  $\mu$ M. After droplet generation, the following PCR program was run: one cycle of 95°C for 10 min; 40 cycles of 95°C for 30s and 60°C for

30s; followed by a final extension at 72°C for 10 min and a 4°C hold. PCR results were analyzed with the Droplet reader and QuantaSoft (Bio-Rad).

### 2.7. Quantification of total particles by flow virometry

Few studies used FM4-64FX and reported that the unbound fractions of the dye do not interfere with the flow cytometry measurements [25,26]. Moreover, FM4-64FX was shown to efficiently label EVs as well as the retrovirus under study [26]. CTV is a similar dye to Carboxyfluoresceinsuccinimidyl ester (CFSE), which has been used in many flow cytometry studies on EVs [25,27]. CTV was reported as more efficient and it has a different fluorescence spectrum than GFP, which is helpful in avoiding crosstalk, since the samples bear GFP.

A double staining experiment was performed by labeling Clone 92 EV samples with a generic lipophilic dye, FM4-64FX (Thermofisher Scientific), and a protein-binding dye, Cell Trace Violet (CTV) (Thermofisher Scientific) according to the manufacturer's instructions. A three-laser BD LSRFortessa™ X-20 was used for acquisition and results were analyzed by FlowJo V10.2 (FlowJo LLC, Ashland, Oregon, USA). 405 nm filter with 450/50 fluorescent channel, and 488 nm filter with 530/30 and 780/60 fluorescent channels were used.

For small particle detection, a Cytoflex flow cytometer (Beckman Coulter, Indianapolis, IN, USA) with a photomultiplier tube (PMT) for forward scatter detection was used. Specifications for laser wavelengths and power were as follows: 488 nm–300 mW, 525/40 fluorescent channel. Acquisition was done with CytExpert (Beckman Coulter). Samples, unless otherwise indicated, were acquired at the lowest flow rate 10  $\mu$ l/min. The instrument cleaning procedure prior to acquisition was as follows: 20 min with Cleaning solution (Beckman Coulter) or 20 min with 0.1% bleach followed by 20 min with distilled water.

### 2.8. Imaging of EVs by transmission electron microscopy (TEM)

EV samples were prepared for negative staining TEM imaging according to Théry *et al.* [28]. Imaging was done on a CM 100 Transmission Electron Microscope (Philips, Eindhoven, The Netherlands) operating at 80kV. Briefly, 10 $\mu$ L samples in 2% PFA were fixed on Formvar-carbon coated EM grids in 1% glutaraldehyde. Samples were then stained first in a solution of uranyl oxalate then embedded in a mixture of 4% uranyl acetate and 2% methyl cellulose for 10 min on ice. The stain was then removed by touching gently the edge of the grids on a filter paper. The grids were air dried prior to the TEM observation.

### 2.9. Immunoblot analysis

Proteins were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), transferred to nitrocellulose membranes, blocked with 5% non-fat powdered milk in PBS-tween (PBS-T) and probed for western blot (WB) with antibodies against EV-enriched proteins (anti-CD9 (rabbit), anti-CD81 (mouse) and anti-TSG101 (rabbit) (Abcam, Cambridge, United Kingdom)) and against non-EV enriched proteins (anti-Calnexin (rabbit) (Cell Signaling, Danvers, MA, USA)).

### 2.10. Protein and nucleic acid quantification

Protein concentration was determined using the RC/DC™ Protein Assay (Bio-Rad, Hercules, CA, USA) according to the manufacturer's instructions.

For DNA quantification, the nucleic acids of EVs were extracted using the High Pure Viral Nucleic Acid Kit (Roche, Mannheim, Germany). Then, the DNA content was quantified with the Quant-iT™ PicoGreen™ dsDNA Assay Kit (ThermoFisher Scientific) following the manufacturer's instructions.

RNA extraction using the High Pure Nucleic Acid kit has been done previously [16]. This technique was however deemed not suitable for that purpose since poly(A) is used in this kit in a non-negligible concentration to precipitate the RNA. This would compromise RNA quantification since the Ribogreen kit used for total RNA quantification has a high affinity for poly(A) fractions.

RNA was extracted using the Exosomal RNA Isolation Kit (Norgen, Thorold, Canada). The extracted RNA was quantified with the Quant-iT<sup>TM</sup> RiboGreen<sup>TM</sup> RNA Assay Kit (ThermoFisher Scientific) or the Qubit<sup>TM</sup> RNA assay (ThermoFisher).

### 2.11. EV identification

Protein markers from the Minimal Information for Studies of Extracellular Vesicles 2018 (MISEV 2018) guidelines were used to confirm enrichment of EVs from their parent cells [29]. For general EV characterization, MISEV 2018 recommends showing three positive protein markers of EVs to demonstrate EV enrichment with ideally one transmembrane/lipid bound protein and one cytosolic protein. In addition to demonstrating protein enrichment, MISEV 2018 also recommends the depletion of cellular proteins using at least one negative protein marker for EVs.

### 2.12. Proteomic analysis

#### 2.12.1. Filter-Aided Sample Preparation (FASP)

All analyses were done on three biological replicates. The samples were thawed on ice then boiled to ensure deactivation of virus and subsequently aliquoted for proteomic and phospholipid analyses. The samples used in proteomics studies were treated with 4X lysis buffer (14% SDS, 400 mM Tris-HCl (pH 8.5), 400 mM dithiothreitol (DTT), and Protease Inhibitor Cocktail (Millipore Sigma)). The samples were then diluted with water to reduce the lysis buffer concentration to 1X. The samples were sonicated on ice with Sonic Dismembrator (Fisher Scientific, Ottawa, Canada) and subsequently boiled at 95°C for 10 minutes. The samples were alkylated with 20 mM iodoacetamide and then digested using a modified filter-aided sample preparation (FASP) method [30]. Briefly, the samples were first buffer exchanged with 8 M urea using a 10 kDa MWCO filter in order to remove all detergent and alkylating reagents. A buffer exchange into 50 mM ammonium bicarbonate was then performed several times. Protein concentrations were determined by the Bradford protein assay (Bio-Rad). The protein suspensions were then digested with sequencing grade modified trypsin (Promega, Madison, WI, USA) at 37°C overnight, and the resulting peptides that were collected following centrifugation were acidified with formic acid (final concentration of 0.25%).

#### 2.12.2. Liquid chromatography-tandem mass spectrometry (LC MS/MS) analysis

The acidified peptides were analyzed on a reversed-phase nanoAcuity UPLC (Waters, Milford, MA, USA) coupled to LTQ-Orbitrap-XL ETD mass spectrometer (ThermoFisher Scientific) with a nano-electrospray interface operated in positive ion mode. The analysis involved injection and loading of approximately 0.1 µg of the peptide sample onto an inline Pepmap100 300µm x 5 mm C8 Acclaim 5µm 100Å precolumn (ThermoFisher Scientific), and Nano-Acuity Symmetry C18, 5µm, 180 um x 2cm Trap (Waters) followed by separation using a 100 µm I.D. x 10 cm 1.7 µm BEH130C18 nanoLC column (Waters). The mobile phase consisted of 0.1% (v/v) formic acid in HPLC grade water as solvent A and 0.1% (v/v) formic acid in acetonitrile as solvent B. The peptides were separated using a gradient ramping from 0.2% to 40% solvent B over 45 min, 40% to 95% solvent B over 4 min, and then re-equilibrating from 95% to 0.2% solvent B over 11 min at a flow rate of 500 nL/min. A blank with a 30-min gradient was run between samples to minimize carryover. Data was acquired on ions with mass/charge (m/z) ratio between 400 and 2,000 Da in profile mode at a resolution of 60,000 followed by data-dependent analysis (DDA) MS/MS scans of the top three ions using collision-induced dissociation (CID) for fragmentation with the following settings: isolation width of 3.0, normalized collision energy of 35.0, activation Q of 0.250, and activation time of 30.000 ms.

#### 2.12.3. Mascot Database Search

The raw files generated by MS analysis were converted to mascot generic files (mgf) and mzXML files using ProteoWizard [31] (version 3.0.18250, ProteoWizard Software Foundation, Palo Alto, CA, USA). Files were submitted to Mascot search engine [32] (version 2.6.2, Matrix Science, London,

United Kingdom) to search against protein sequence databases consisting of target and decoy sequences. The target sequences consisted of human Uniprot database [33] (release 2019) combined with HIV genome translated genome sequence and GFP sequences. The decoy database was constructed with reverse sequences from the target database. Searches were restricted to trypsin cleavage with one missed cleavage accepted. The peptide tolerance was set to  $\pm 5$  ppm with a fragment mass tolerance of  $\pm 0.8$  Da. Carbamidomethylation on cysteine residues was set as a fixed modification while oxidation of methionine residues was set as a variable modification. False discovery rate (FDR) in Mascot searching was calculated as follows:

$$FDR = \frac{N_{decoy}}{N_{target}}, \quad (1)$$

where  $N_{decoy}$  is the number of decoy hits identified and  $N_{target}$  is the number of target hits identified. To maximize the number of true positive peptides and minimize false-positives, FDR was chosen to be  $\leq 1\%$  per sample, which corresponded to an average Mascot ion scores  $\geq 40$ .

#### 2.12.4. Proteomics Data Processing

Proteomics data analysis involved measurement and assignment of MS intensity signal to each identified protein and was performed using MatchRX software (MatchRX, Royal Oak, MI, USA) as described previously [34]. Briefly, peak intensities of all the ions in each MS run were extracted from the mzXML files and assigned to Mascot-identified proteins using the MatchRX software using their  $m/z$ , retention times and neighbouring peak coordinates. Each MS intensity was adjusted using total median normalization as described previously [34]. For each sample, total MS intensity signal was also calculated by summing intensities of all the MS intensity signals in the run and was used to estimate fraction of MS intensity (FMSI) of each protein as follows:

$$FMSI \text{ of a protein} = \frac{\text{sum of all intensities specific to the protein in the sample}}{\text{sum of all intensities in the sample}}. \quad (2)$$

Venn diagram was generated using the BioVenn website [35].

#### 2.13. Liquid chromatography-mass spectrometry (LC-MS) of phospholipids

LC-MS was carried out using a Synapt G2-Si mass spectrometer (Waters) coupled to a Dionex3000 HPLC (ThermoFisher Scientific) using Waters' ESI source. Separations were performed on a  $50 \times 1$  mm internal diameter  $3.5 \mu\text{m}$  Zorbax XDB-C8 column (Agilent, Santa Clara, CA, USA). Solvent A was 5:1:4 IPA:MeOH:H<sub>2</sub>O (0.2% Formic Acid/0.028 NH<sub>4</sub>OH); while solvent B was IPA (0.2% Formic Acid/0.028 NH<sub>4</sub>OH). The following gradient program was used: 0% solvent B over 3 min, 0-95% solvent B over 12 min, 95% solvent B over 5 min, and re-equilibration at 0% solvent B for 10 min. Phospholipids were analyzed in negative-ion mode. A rolling collision energy between 45 and 160 eV was used for automated DDA MS/MS. Data interpretation was done manually using LIPID MAPS® Online Tools [36]. Data was normalized by first applying correction factors based on ionization efficiencies and response factors for each type of phospholipid, then percent compositions for each fraction were calculated.

#### 2.14. Transcriptomics and Bioinformatics Analysis

The quality of the RNA was assessed with the Qubit RNA assay. The sequencing library was prepared using the SMARTer smRNA-Seq kit for Illumina (Takara Bio USA, Mountain View, CA, USA), following the manufacturer's instructions, for miRNA samples, and the SMART-Seq v4 Ultra Low Input RNA kit (Takara Bio USA) for mRNA samples. The quality of the libraries was assessed using Qubit DNA assay (ThermoFisher Scientific), Bioanalyzer 2100 (Agilent), and qPCR. Sequencing was performed on the NextSeq 500 system (Illumina, San Diego, CA, USA), using a 1x75bp SE sequencing strategy.

The gene expression levels in each mRNA sample were evaluated by aligning reads to the human GRCh38 reference genome and following published methods.[37] The gene expression level was normalized by the number of fragments per kilobase per million mapped reads (FPKMs). Enrichment analyses were performed using the GO Enrichment Analysis tool and Metascape Express

Analysis [38-41]. Protein hits were classified by protein class using the Protein Analysis Through Evolutionary Relationships (PANTHER) tool [42].

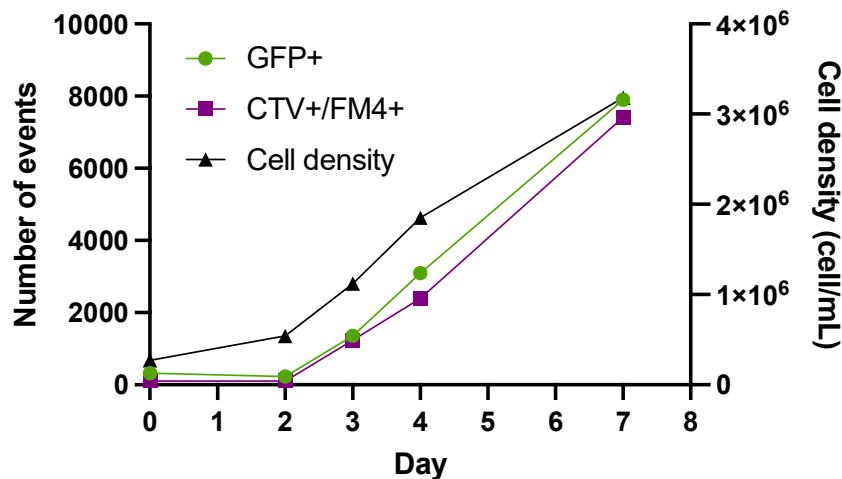
### 3. Results

#### 3.1. Characterization of Clone 92 EVs in the absence of lentiviral particles

Although efforts have been dedicated to segregate EVs from retrovirus particles [14,43], it is currently not possible to strictly separate EVs from LVs, especially in large scale processes. It is therefore important to understand what basal EVs, meaning under no-inducing conditions, are composed of, as they will constitute a subpopulation that will be found in LV preparation. Thus, the first part of the study focuses on the characterization of EVs generated by Clone 92 in the absence of lentiviral particles. As described in the materials and methods section, Clone 92 cells are cultured in suspension and serum-free medium, to avoid contamination by EVs associated with serum supplementation [28]. It is also important to note that the viability of the cell cultures was maintained and monitored above 95% at all time to avoid the presence of apoptotic bodies.

##### 3.1.1. Quantification of EVs using GFP signal by flow virometry

Clone 92 cells express GFP constitutively, allowing the detection of particles released by the cells as Clone 92 EVs will emit a fluorescence signal. The flow virometry quantification method was first validated using a double-labeling strategy. Samples of a non- induced Clone 92 culture supernatant referred to as  $^{C92}EV_{sup}$  were taken on day 0, 2, 3, 4 and 7 and labeled with FM4-64FX and CTV. Samples were then analyzed by flow cytometry without purification. Results are presented in Figure 1.



**Figure 1.** Analysis of Clone 92 supernatant by flow virometry: Quantification of GFP+ events and CTV+/FM4+ events over cell culture days as measured on the BD Fortessa flow cytometer.

Gating is shown in supplementary Figure S1. In Figure S1(a), the gate represents GFP positive events. In Figure S1(b), gating was done such as FM4-64FX positive and CTV positive events are found in quadrant Q2. In Figure S1(a) and S1(b), HyCell medium serves as a negative control and shows no GFP+ signal nor FM4-64FX+/CTV+ signal before and after staining. The analysis was done on the samples mentioned above and the GFP+ events, CTV+/FM4-64FX+ events and the cell density were plotted over time on Figure 1. Figure 1 shows that GFP positive events correlated to FM4-64FX / CTV double positive events and are increasing as the cell density increases over time. Thus, this preliminary experiment showed the feasibility of detecting  $^{C92}EVs$  using GFP fluorescence signal to enumerate the number of total particles. Subsequent flow virometry measurements were then done using only GFP signal.

Flow virometry was then used in order to estimate the number of particles bearing GFP, since GFP, which is constitutively expressed in that cell line, is being randomly incorporated into  $\text{C}92\text{EVs}$ . The gating used is presented in supplementary Figure S2, PBS being used as a negative control. Samples were diluted with PBS to keep a low abort rate (ideally below 2%) and the concentrations were corrected for the dilution.

### 3.1.2. Development of a scalable EV isolation process using size exclusion chromatography (SEC)

For consistency and reproducibility, it was desirable that all analyses be performed on a single batch, thus requiring a large volume with high yield of isolated EVs to proceed with extensive characterization.

The isolation process involving UF/DF and SEC described in the materials and methods section, with or without the final concentration step, yielded EVs with an adequate volume and concentration according to the protein content and was considered as an appropriate process to produce EVs for further characterization.

This isolation process was performed 3 times and yielded 3 batches of  $\text{C}92\text{EV}_{\text{SEC}}$ .

Table 2 presents the mass balance of the EV isolation process for Clone 92 culture showing recoveries at different steps in the process. Quantification was done by flow virometry in order to estimate the amount of in-process and  $\text{C}92\text{EV}_{\text{SEC}}$ .

**Table 2.** In-process quantification of GFP+ particles by flow virometry and total protein by RC/DC during Clone 92 EVs isolation process.

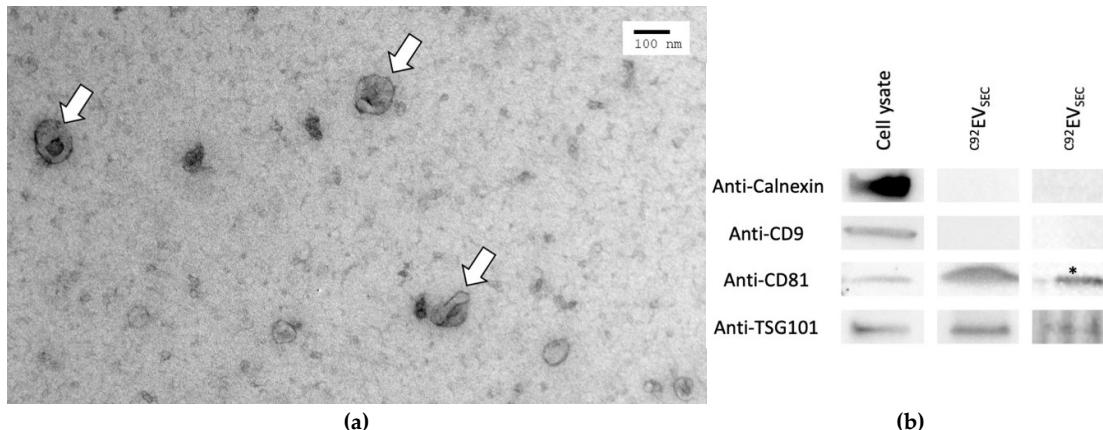
In-process Sample	Volume (mL)	GFP+ particles (part/mL)	GFP+ particles Step recovery (%)	Total protein (µg/mL)
<b>Supernatant</b>	1478	$1.37 \times 10^9$	N/A <sup>1</sup>	85
<b>Supernatant after 0.45µm filtration</b>	1473	$7.73 \times 10^8$	56	73
<b>UF/DF<sup>2</sup> product</b>	115	$4.01 \times 10^9$	68	181
<b>UF/DF<sup>2</sup> product after 0.45µm filtration</b>	108	$2.90 \times 10^9$	84	188
<b>UF/DF<sup>2</sup> permeate</b>	1759	$2.81 \times 10^8$	-	32
<b>SEC<sup>3</sup> EV peak</b>	101	$3.15 \times 10^9$	102	75
<b>SEC<sup>3</sup> post-EV peak</b>	8	$2.99 \times 10^8$	103	44
<b>Final concentrated EVs</b>	10	$2.38 \times 10^{10}$	74	795

<sup>1</sup> N/A: not applicable, <sup>2</sup> UF/DF: ultrafiltration/diafiltration, <sup>3</sup> SEC: size exclusion chromatography

Total protein quantification by RC/DC showed a reduction of 63% in the  $\text{C}92\text{EV}_{\text{SEC}}$  peak as compared to the starting material. Gene transfer assay (GTA) was performed on undiluted  $\text{C}92\text{EV}_{\text{SEC}}$  samples and did not show any functional titer confirming the absence of lentiviral activity.

### 3.1.3. Preliminary characterization confirms EV identity

$\text{C}92\text{EV}_{\text{SEC}}$  were imaged by transmission electron microscopy (TEM). In Figure 2(a), EVs are visible as cup-shaped indicated by white arrows. Their sizes range from about 50 to 100 nm.



**Figure 2.** Preliminary characterization of C92EV<sub>SEC</sub>. **(a)** Electron microscopy images of C92EV<sub>SEC</sub>. Scale bar: 100 nm. **(b)** Enriched proteins in C92EV<sub>SEC</sub> identified by Western Blot. \*Sample was concentrated before loading on the gel.

CD81 was detected in cell lysate and C92EV<sub>SEC</sub> samples, with an expected enrichment in C92EV<sub>SEC</sub> samples (Figure 2(b)). TSG101, also commonly found in EVs, was present in C92EV<sub>SEC</sub>. Calnexin is a protein embedded in the endoplasmic reticulum membrane and serves here as a negative marker to assess EVs purity. It was only found in cell lysate samples. The WB did not exhibit CD9 in C92EV<sub>SEC</sub> samples, the concentration of this common marker in the samples was either too low for detection or C92EV<sub>SEC</sub> might not be enriched in CD9.

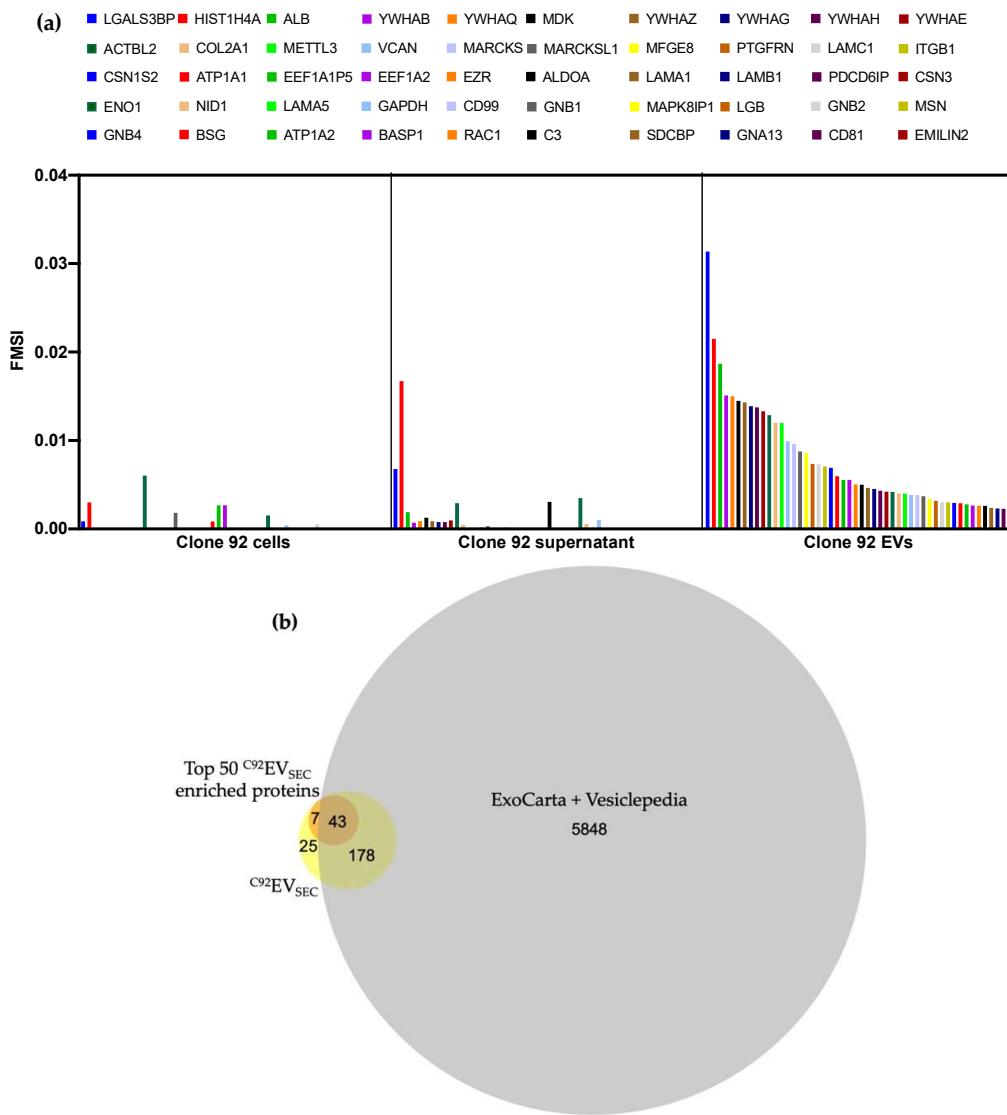
### 3.1.4. Proteomic cargo of C92EV<sub>SEC</sub>

Mass spectrometry (MS) was used to estimate enrichment of extracellular vesicles (EVs) by looking at the fraction of individual MS intensity contributed by each protein to the total MS intensity in each sample. The transmembrane proteins cluster of differentiation 81 (CD81), basigin (BSG), and the cytosolic protein, programmed cell death 6 interacting protein (PDCD6IP), were examined in order to confirm enrichment of EVs. The fraction of MS intensity (FMSI) of both CD81 and BSG was found to be higher in C92EV<sub>SEC</sub> when compared to the Clone 92 cells, as well as the conditioned media prior to EV isolation called “supernatant” (supplementary Figure S3). Additionally, protein PDCD6IP had a higher FMSI in the C92EV<sub>SEC</sub> than in the parental cells and associated supernatants (supplementary Figure S3). Endoplasmic reticulum chaperone BiP (HSPA5) and endoplasmic reticulum chaperone HSP90B1 proteins are both good candidates as negative protein markers as they are both found in the endoplasmic reticulum of cells and are not associated with the plasma membrane or endosomes. In this set of data both HSPA5 and HSP90B1 appear depleted in C92EV<sub>SEC</sub> compared to the parental cells (supplementary Figure S3). Taken together, these data indicate that the samples have been enriched for EVs.

MS was also used to detect the presence of GFP in the samples and confirm its presence in the C92EV<sub>SEC</sub>. GFP was detected in Clone 92 cells, C92EV<sub>sup</sub>, and C92EV<sub>SEC</sub> and not from the 293SF original cell line, as expected (supplementary Figure S4). No HIV proteins were identified in any of the samples.

The FMSI of all identified proteins in C92EV<sub>SEC</sub> was plotted to identify enrichment in the EV samples compared to the parental cells. Keratins were excluded from the data analysis (32 proteins in this study). Out of the 253 proteins identified, 192 showed enrichment in EVs based on their FMSI, with the top 50 of enriched proteins in C92EV<sub>SEC</sub> shown in Figure 3(a) based on their FMSI.

The total number of identified proteins in C92EV<sub>SEC</sub> and the top 50 enriched proteins were compared to the combined Vesiclepedia database [44] and ExoCarta database [45] in Figure 3(b).



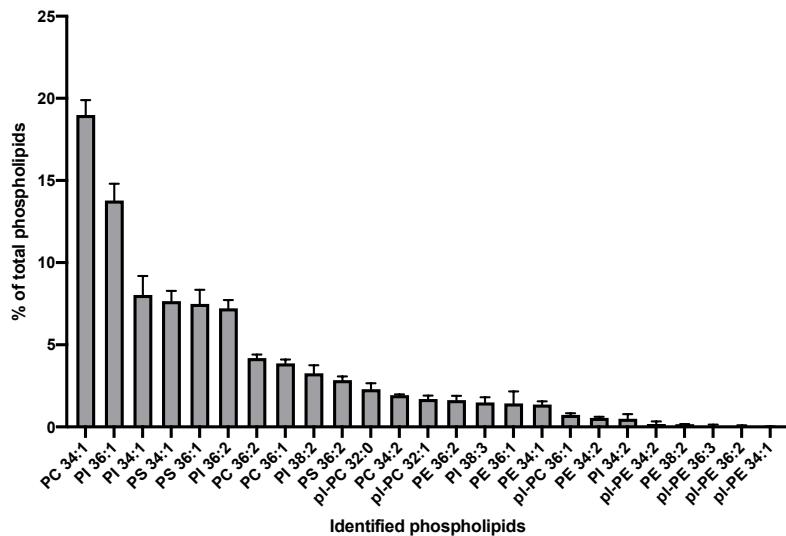
**Figure 3.** Proteomic analysis of  $C^{92}EV_{SEC}$ . (a) Top 50 enriched proteins identified in  $C^{92}EV_{SEC}$  based on their FMSI. (b) Area-proportional Venn diagram for the total number of identified proteins in  $C^{92}EV_{SEC}$  and the top 50 enriched proteins in  $C^{92}EV_{SEC}$  within the combined Vesiclepedia and ExoCarta database.

Among the total identified proteins in  $C^{92}EV_{SEC}$ , 32 were not found in the combined database, and 7 were in the top 50 enriched proteins in  $C^{92}EV_{SEC}$ : MDK, METTL3, CSN1S2, CD99, MAPK8IP1, LGB and EMILIN2. These proteins might be potential markers for  $C^{92}EVs$ .

### 3.1.5. Lipidomic composition of $C^{92}EV_{SEC}$

EVs are formed by a lipid bilayer membrane. Given the size of EVs, lipids are a significant component of EVs and may play important biological roles. The field is still young, however any data on lipids structuring EVs may give critical information related to their biogenesis.

The phospholipid species were quantified by liquid chromatography-mass spectrometry (LC-MS) in three samples of  $C^{92}EV_{SEC}$  (Figure 4). LIPID MAPS consortium guidelines were followed for lipid nomenclature and the annotation of lipid species was as follow: lipid class followed by total number of carbons and degree unsaturation of respective acyl chains (e.g., PS 34:1) [46].



**Figure 4:** Phospholipids identified in  $\text{C}92\text{EV}_{\text{SEC}}$ . PC: phosphatidylcholine, PI: phosphatidylinositol, PS: phosphatidylserine, pl-PC: plasmalogen-phosphatidylcholine, PE: phosphatidylethanolamine, pl-PE: plasmalogen-phosphatidylethanolamine. Error bars indicate standard error of the mean (SEM).

The most abundant phospholipids identified in  $\text{C}92\text{EV}_{\text{SEC}}$  were phosphatidylcholine (PC) 34:1 and phosphatidylinositol (PI) 36:1. Hexose-ceramide (sphingolipids (SL)) were also abundantly detected at levels comparable to PL, however they could not be quantified reliably.

### 3.1.6. Nucleic acid content and gene ontology

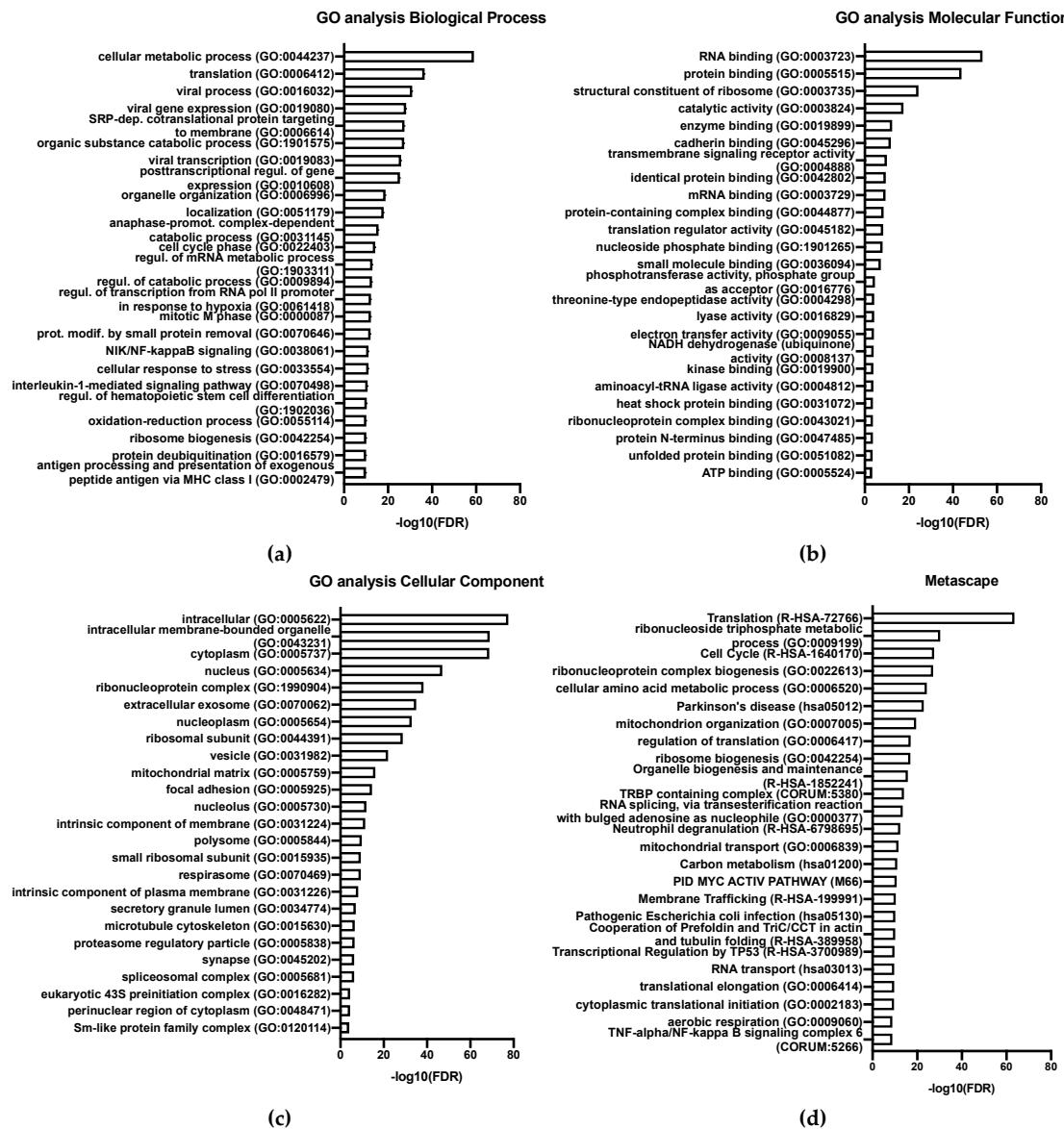
DNA and RNA extracted from  $\text{C}92\text{EV}_{\text{SEC}}$  were quantified using the Picogreen and Qubit assay respectively (Table 2). The analyses were performed on two different batches of  $\text{C}92\text{EV}_{\text{SEC}}$  in duplicate.

**Table 2.** Nucleic acid quantification in  $\text{C}92\text{EV}_{\text{SEC}}$ .

$\text{C}92\text{EV}_{\text{SEC}}^1$	
dsDNA ( $\mu\text{g/mL}$ )	$0.4 \pm 0.1$
Total RNA ( $\mu\text{g/mL}$ )	$9.7 \pm 1.7$

<sup>1</sup> Mean  $\pm$  SD.

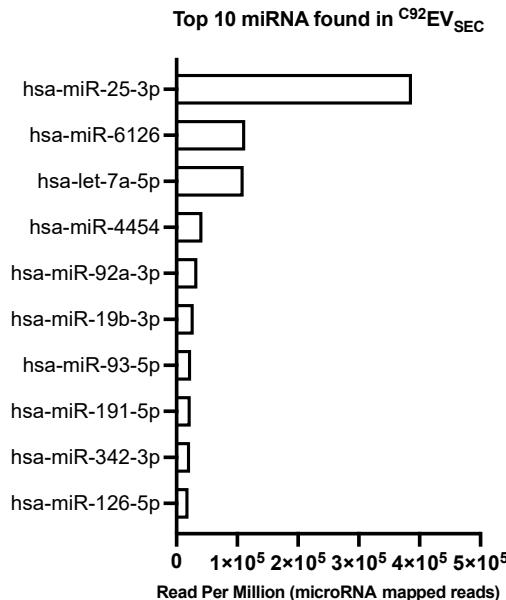
The 3000 most expressed genes present in replicate samples of  $\text{C}92\text{EV}_{\text{SEC}}$  and ranked by FPKM were analyzed for enrichment. The GO enrichment analysis tool and Metascape were both used to provide a broader search in available databases. The top 25 ontology terms are shown on Figure 5(a), 5(b), 5(c) and 5(d).



**Figure 5:** Gene ontology (GO) enrichment analysis in  $C92EV_{SEC}$ , top 25 ontology terms. **(a)** GO, biological process, **(b)** GO, molecular function, **(c)** GO, cellular component, **(d)** Metascape. Only terms with an FDR < 0.01 were selected.

$C92EV_{SEC}$  are enriched in genes involved in viral process, viral gene expression and viral transcription as seen in Figure 5(a). The GO enrichment analysis for molecular function in Figure 5(b) reveals that many genes represented in EVs have a binding function such as RNA binding, protein binding and enzyme binding. Many intracellular components are abundantly found in  $C92EV_{SEC}$  including intracellular membrane-bounded organelle and cytoplasm components, as well as genes associated with extracellular exosome (Figure 5(c)). Genes involved in DNA- and RNA-related functions are highly represented: RNA transport, viral transcription, regulation of mRNA metabolic process, transcription regulation activity, regulation of translation etc. Other enriched genes are involved in immune system process and cellular response such as NIK/NF-kappaB signaling, anaphase-promoting complex-dependent catabolic process.

miRNAs are highly conserved, non-coding, small single-stranded RNA molecules and have the ability to regulate gene expression. They were also characterized in  $C92EV_{SEC}$ . The 10 most abundant miRNAs found in  $C92EV_{SEC}$  are shown in Figure 6.



**Figure 6:** Top 10 miRNA found in  $C^{92}$ EV<sub>SEC</sub>.

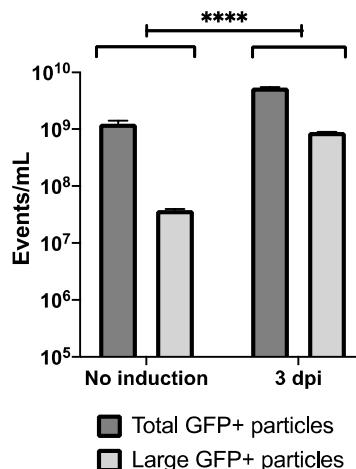
The most abundant miRNA found in  $C^{92}$ EV<sub>SEC</sub> was hsa-miR-25-3p, with over 3 times more read per million than the next most abundant miRNA species hsa-miR-6126 and hsa-let-7a-5p.

### 3.2. Characterization of Clone 92 EVs during lentiviral particles production

As previously indicated, it is not yet feasible to effectively separate EVs from LVs in a production process. In the second part of this study, we compared EVs from Clone 92 in absence of LV induction ( $C^{92}$ EV<sub>UC</sub>), and Clone 92 co-produced EVs following induction of LV production ( $C^{92}$ EV/LV<sub>UC</sub>). For consistency in the sample preparations, ultracentrifugation was used as described in section 2.3.2.

#### 3.2.1. Heterogeneity of EV and LV populations

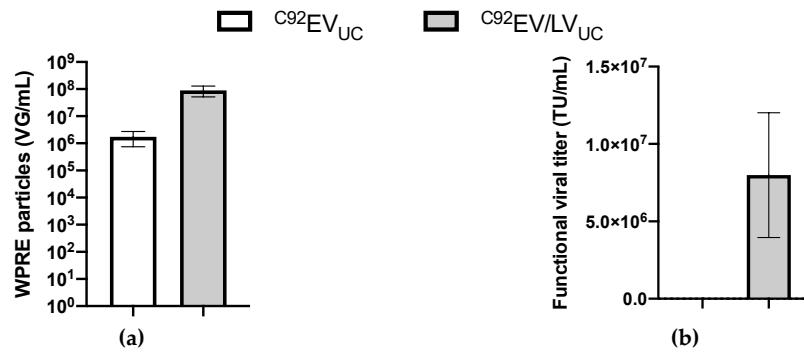
As flow virometry is based on GFP+ events, analysis can be performed directly on supernatant material.  $C^{92}$ EV<sub>sup</sub> were therefore compared to  $C^{92}$ EV/LV<sub>sup</sub> 3 days post-induction (3dpi). Results are shown in Figure 7.



**Figure 7:** Comparison between Clone 92 supernatants with no induction and 3 days post-induction (3dpi) by flow virometry: Quantification of flow virometry subpopulations of large particles and total GFP+ particles in each studied condition. Error bars indicate SEM. Significance is indicated by \*\*\*\* and is calculated via two-way ANOVA.

Using the same gating as in the first part of the study, total GFP+ events were higher in  $\text{C}^92\text{EV/LV}_{\text{UC}}$  3dpi. Another population was additionally observed after induction (supplementary Figure S5, still fluorescent but larger in size. A third population which was not gated in Figure S5 would include non-fluorescent even larger particles. This population was also observed in some in-process samples without induction from Table 2, suggesting large particles with no GFP but their proportion could not be estimated due to their overlap with the noise.

$\text{C}^92\text{EV}_{\text{UC}}$  and  $\text{C}^92\text{EV/LV}_{\text{UC}}$  samples were analyzed by digital drop polymerase chain reaction (ddPCR) and gene transfer assay (GTA). Results are shown in Figure 8.



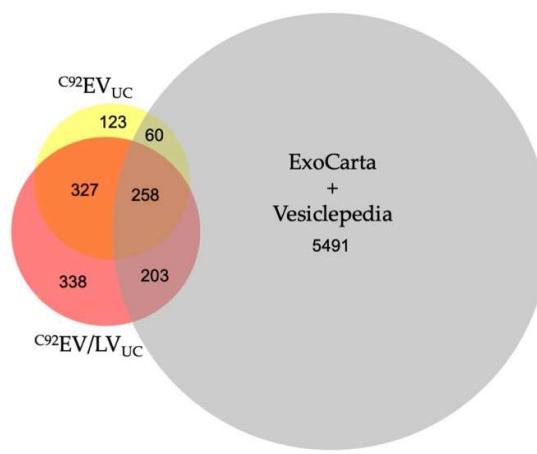
**Figure 8:** Comparison between  $\text{C}^92\text{EV}_{\text{UC}}$  and  $\text{C}^92\text{EV/LV}_{\text{UC}}$ . (a) Quantification of WPRE particles by ddPCR. (b) Quantification of functional viral titer by gene transfer assay (GTA). Error bars indicate SEM.

ddPCR allowed the quantification of particles containing the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE). As seen on Figure 7(a), both  $\text{C}^92\text{EV}_{\text{UC}}$  and  $\text{C}^92\text{EV/LV}_{\text{UC}}$  show a titer by ddPCR.  $\text{C}^92\text{EV/LV}_{\text{UC}}$ 's titer is greater than  $\text{C}^92\text{EV}_{\text{UC}}$ 's titer by two orders of magnitude.

GTA measures transgene expression (here GFP by flow cytometry) in transduced target cells to report functional viral vector particles. As in the first part of the study,  $\text{C}^92\text{EV}_{\text{UC}}$  samples did not show any functional titer, confirming the absence of functional LVs particles when there is no induction.  $\text{C}^92\text{EV/LV}_{\text{UC}}$  on the other hand confirmed the functionality of the produced LVs particles.

### 3.2.2. Protein cargos of EVs and LVs have common features

$\text{C}^92\text{EV}_{\text{UC}}$  and  $\text{C}^92\text{EV/LV}_{\text{UC}}$  were also compared using MS. Enrichment of EVs was estimated by looking at FMSI contributed by each protein to the total MS intensity. Proteins identified with a Mascot score over 30 and with more than 1 peptide were analyzed, excluding keratins. Both samples contained protein markers from the MISEV 2018 guidelines: CD81 and PDCD6IP were found to be enriched in  $\text{C}^92\text{EV}_{\text{UC}}$  and  $\text{C}^92\text{EV/LV}_{\text{UC}}$ . Additionally, prostaglandin F2 receptor inhibitor (PTGFRN), a protein from the ExoCarta database was also found in both. CD9 was not identified in the samples, indicating that this tetraspanin is not enriched in  $\text{C}^92\text{EVs}$ . In addition, Calnexin and HSP90B1, common EVs "negative markers", were not identified in any of the samples. This suggests that either EVs are indeed recovered in  $\text{C}^92\text{EV/LV}_{\text{UC}}$  samples or that LVs package the same proteins as EVs. A total of 768 proteins were identified in  $\text{C}^92\text{EV}_{\text{UC}}$  and 1126 proteins were identified in  $\text{C}^92\text{EV/LV}_{\text{UC}}$ , with an overlap of 44.7% as shown on Figure 9.



**Figure 9:** Area-proportional Venn diagram for the total number of identified proteins in  $C^{92}EV_{UC}$  and  $C^{92}EV/LV_{UC}$  with a Mascot score over 30 and with more than 1 peptide within the combined Vesiclepedia and ExoCarta database.

Among all identified proteins in  $C^{92}EV_{UC}$  and  $C^{92}EV/LV_{UC}$ , 183 were uniquely identified in  $C^{92}EV_{UC}$  and 541 were uniquely identified in  $C^{92}EV/LV_{UC}$  (Supplementary Table S1).

All identified proteins in  $C^{92}EV_{UC}$  and  $C^{92}EV/LV_{UC}$  were classified into 23 PANTHER protein classes (Table 3).

**Table 3.** Summary of protein classes identified in both  $C^{92}EV_{UC}$  and  $C^{92}EV/LV_{UC}$ , or only in either  $C^{92}EV_{UC}$  or  $C^{92}EV/LV_{UC}$ .

Category name (Accession)	Protein hits		
	Only in $C^{92}EV_{UC}$	Only in $C^{92}EV/LV_{UC}$	In both $C^{92}EV_{UC}$ and $C^{92}EV/LV_{UC}$
extracellular matrix protein (PC00102)	4	11	9
cytoskeletal protein (PC00085)	6	27	49
transporter (PC00227)	7	17	25
scaffold/adaptor protein (PC00226)	8	18	21
cell adhesion molecule (PC00069)	3	4	6
nucleic acid metabolism protein (PC00171)	9	49	33
intercellular signal molecule (PC00207)	0	2	2
protein-binding activity modulator (PC00095)	9	20	18
viral or transposable element protein (PC00237)	0	3	2
calcium-binding protein (PC00060)	2	3	3
gene-specific transcriptional regulator (PC00264)	6	20	18
defense/immunity protein (PC00090)	1	3	5
translational protein (PC00263)	2	50	28
metabolite interconversion enzyme (PC00262)	15	45	51
protein modifying enzyme (PC00260)	18	33	43
chromatin/chromatin-binding, or -regulatory protein (PC00077)	4	7	13
transfer/carrier protein (PC00219)	1	3	6
membrane traffic protein (PC00150)	5	11	21
chaperone (PC00072)	1	12	11
cell junction protein (PC00070)	1	1	1
structural protein (PC00211)	1	0	2
storage protein (PC00210)	0	0	1
transmembrane signal receptor (PC00197)	4	14	8

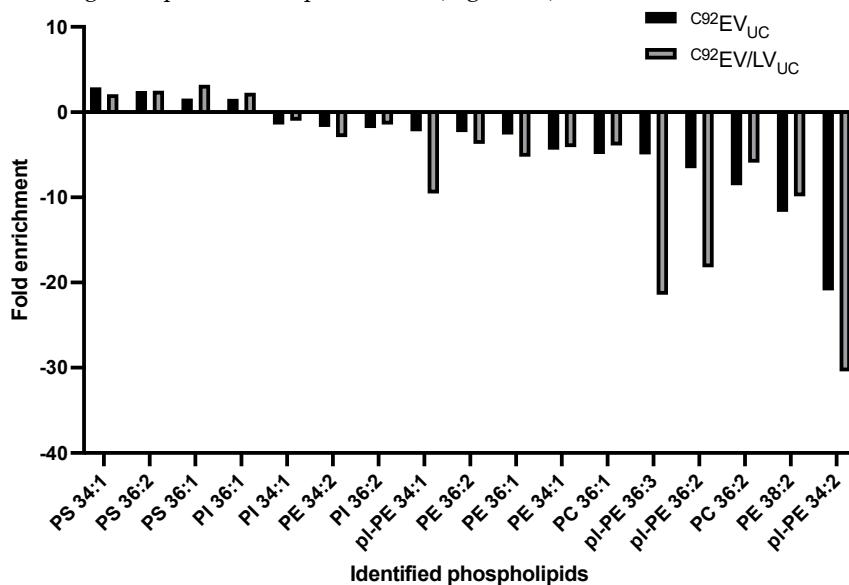
Metabolite interconversion enzymes and protein modifying enzymes were highly represented in all three categories (Table 3). Although also abundant in only  $C92EV_{UC}$  and in both  $C92EV_{UC}$  and  $C92EV/LV_{UC}$ , nucleic acid metabolism proteins were even more enriched in  $C92EV/LV_{UC}$ . Translational proteins were abundantly found in  $C92EV/LV_{UC}$  only and cytoskeletal proteins were dominant in the overlap population.

Additionally, GAG-POL was used to identify enrichment for LV particles. GAG-POL was found enriched only in samples after LV induction in  $C92EV/LV_{UC}$ .

GFP was identified in both  $C92EV_{UC}$  and  $C92EV/LV_{UC}$ . Lower level of GFP enrichment was seen in samples before induction.

### 3.2.2. Phospholipid content in EVs and LVs

The phospholipid species were quantified by liquid chromatography-mass spectrometry (LC-MS) and compared between Clone 92 cells (cell pellet),  $C92EV_{UC}$  and  $C92EV/LV_{UC}$ . The identified phospholipids in  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  were ranked by highest positive fold change to most negative fold change compared to the parent cells (Figure 10).



**Figure 10:** Phospholipids identified in Clone 92 EVs and LVs compared to Clone 92 parent cells. PS: phosphatidylserine, PI: phosphatidylinositol, PE: phosphatidylethanolamine, pl-PE: plasmalogen-phosphatidylethanolamine PC: phosphatidylcholine. Error bars indicate SEM.

Differences were not statistically significant but could have biological meaning.  $C92EV_{UC}$  membranes and  $C92EV/LV_{UC}$  membranes are enriched in the same PL compared to the cell membrane: phosphatidylserine (PS) 34:1, PS 36:2, PS 36:1 and phosphatidylinositol (PI) PI 36:1. Interestingly,  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  are enriched and depleted in the same PL compared to their parent cell. Plasmalogen-PE (pl-PE) are 1.5 to almost 5 times more depleted in  $C92EV/LV_{UC}$  than in  $C92EV_{UC}$ .

## 4. Discussion

EVs have gained a lot of attention in the past few years, as potential biomarkers and as drug delivery vehicles. Many studies have been carried out on EVs isolated from biofluids or even cultured cells. Yet, investigations do not report on EVs as secondary products in viral vaccines or viral vectors production. Most cell lines, especially mammalian cell lines, are known to release EVs and cell lines used as platform for biological products are no exception. The experiments completed in this study aim for a comprehensive characterization of EVs produced in HEK293SF cell lines that are widely used in viral vectors and viral vaccines production. Enveloped viruses-based products including lentiviral vectors are especially targeted here for their biophysical similarities to EVs as the preparations most certainly contain both EVs and viruses. To this end a large set of experiments has

been done to characterize EVs associated with an inducible HEK293SF lentivirus producing cell line (Clone 92) cultured under non-induced conditions.

The characterization of EVs is greatly impacted by the isolation method [47]. Here we decided to develop a process that would allow all selected analyses to be performed on one single batch of EVs for results consistency. The isolation method combining SEC and UF was selected for its scalability. Moreover, an additional advantage of developing a scalable process applicable to isolation of EVs associated with HEK293SF human cell line is the generalization of this process to multiple therapeutic products derived from the HEK293SF manufacturing platform. Indeed, EVs produced in HEK293SF cell cultures might be loaded with therapeutic cargos and used as drug delivery vehicles [48]. EVs associated with the two cell lines HEK293SF and HEK293-derived lentivirus producing cell, Clone 92 cultures were investigated. Since no significant differences were found between EVs isolated from the two cell lines and because of the intrinsic GFP labeling property of Clone 92 allowing for flow virometry measurements, these studies focused on Clone 92.

EVs reported in the literature have different cellular origins and therefore no definite markers of populations have been identified. Enriched proteins are, however, observed. In this study, although we did not discriminate between exosomes and microvesicles, only enriched proteins associated with exosomes were considered for identification. Additionally, the study focused on EVs co-produced with enveloped virus products, more specifically lentiviral vectors, consequently the size of the particles observed ranged from 80 to 100 nm, which mainly corresponds to the size of exosomes and only small microvesicles.

Other orthogonal methods are available for EV and LV quantification. However, significant discrepancies in absolute values with other techniques should be expected. For example, nanoparticle tracking analysis (NTA) is based on the Brownian motion of particles in suspension and is used to determine the size distribution of purified EVs [49] and for quantification [50]. This method lacks specificity and often leads to overestimation of the total particles measured. A method for in-process LV quantification was recently published [51] involving High-Performance Liquid Chromatography (HPLC). Although the authors optimized the method for minimizing the impact of EVs, they did acknowledge the presence of EVs in the quantification of LV particles and their proportion could not be estimated since the measure of a sample with no LV particles falls outside of the claimed linear range of the method.

The different methods used in this study highlight different features of EVs. Flow virometry results reflect the presence of GFP in <sup>C92</sup>EVs. As reported, the GFP+ analysis would be a better estimate of the total particles. However, it is likely that intermediate populations that do not carry GFP or have slightly different size or granularity are excluded. Moreover, this quantification method is applicable to <sup>C92</sup>EVs because of the fluorescence detection and is not applicable to EVs that do not carry GFP due to the challenges associated with signal detection which does not allow differentiating EVs from the signal background in the flow cytometer analyses. ddPCR analysis targeted WPRE as a probe. Indeed, as mentioned before, the GFP transgene and therefore the WPRE element which ensures high level transgene expression, are expressed constitutively. The quantification of WPRE therefore indicates the presence of the transgene, usually referred as "viral genome" when dealing with LVs particles. ddPCR results revealed that the "viral genome" is being incorporated in a fraction of EVs, although no viral protein or viral activity is present in EVs based on the proteomic and GTA analysis. This observation might be of interest for the design and development of therapeutic EVs for delivery of specific nucleic acid cargos. The results by flow virometry differ from the ddPCR data by at least 3 orders of magnitude in <sup>C92</sup>EV<sub>SEC</sub> suggesting that all EVs do not incorporate the "viral genome" sequences. The GTA and ddPCR data in LVs also reveals a difference. Indeed, the functional viral titer is lower than the VG titer as previously documented in Transfiguracion et al. [51]. This underlines the difficulty in assessing absolute quantification of EVs and LVs, but it also underlines the heterogeneous nature of EVs and LVs. In that respect, EVs and LVs are not unique populations but rather a broad distribution of populations that incorporate different cellular components. Here, the results suggest that Clone 92 LV preparations are at least composed of EVs which have incorporated the "viral genome", EVs which don't have the "viral genome", LVs with the viral genome but are not functional, and fully functional LV particles.

Proteomic results of  $C92EV_{SEC}$  showed that GFP was indeed detected in these EVs; however, no HIV proteins were found. Although Gag-Pol is under a constitutive promoter, Rev, which is tightly regulated by the cumate switch in the design of Clone 92 [21,52], induction is required for Gag efficient expression. Thus, HIV proteins are not expected to be found in Clone 92 EVs when there is no induction by cumate and doxycycline. Results confirm here the tight regulation from the switches. Proteomic analyses of Clone 92 EVs not only confirmed EVs identity, thus validating the isolation process, but they also revealed the presence of proteins commonly found in EV databases. More proteins enriched in EVs compared to the conditioned medium and parental cells were also identified (Figure 3) and could be potential new markers for  $C92EVs$ , such as Midkine (MDK in Figure 3(a)), a secreted protein that regulates multiple biological processes including cell proliferation, cell adhesion, cell growth, cell survival, and cell migration [53]. Another potential marker could be N6-adenosine-methyltransferase catalytic subunit (METTL3 in Figure 3(a)) which is involved in many biological processes including RNA processes [54]. Discrepancies between proteins identified in  $C92EV_{SEC}$  and  $C92EV_{UC}$  were observed. Only 78 proteins overlapped between  $C92EV_{SEC}$  and  $C92EV_{UC}$ . The lack of overlap is likely due to the difference in the EVs isolation methods underlining again the importance of this step. Among the 78 proteins, 7 were not found in the combined ExoCarta and Vesiclepedia database. MDK, CSN1S2 and LGB were again in the shortlist. The high percentage of protein overlap in  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  reinforces the observation that EVs and LVs have a lot of common features. Interestingly, a number of proteins identified in  $C92EV/LV_{UC}$  were previously reported to be associated with HIV-1 virus, including EEF1A1, a translational protein [55], H2AFV, a chromatin/chromatin-binding, or -regulatory protein [56], NONO, a nucleic acid metabolism protein [57], GAPDH, a metabolite interconversion enzyme [56], PPIA, a protein involved in host-virus interaction [58]. NONO, GAPDH and PPIA were also found in  $C92EV_{UC}$  thus indicating once again the similarities between EVs and LVs. The high abundance of cytoskeletal proteins in both  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  was expected as cytoskeletal proteins have been implicated in virus transport and release [59], indicating that the budding mechanism of both LV and EV rely on cytoskeletal proteins for the translocation process.

Among the most abundant proteins uniquely found in  $C92EV_{UC}$  were SYDE2 (involved in cell migration), RGPD3 (scaffold/adaptor protein), FHOD3 and SNTB1 (proteins with actin-binding function), HEATR5B (involved in endocytosis), TRMT12 (nucleic acid metabolism protein), CASC3 (RNA-related function), AVEN (cell death regulator), and CHRD (cytokine-related protein). ASB2 and KLHL18 (involved in protein ubiquitination), SMARCA2 (nucleic acid metabolism protein), DOCK6, ZNF814 (gene-specific transcriptional regulator), TNRC6B (RNA-related function), PDE4B (metabolite interconversion enzyme), ATN1, HDAC4, DEPDC1 and CDYL (involved in transcription), C16ORF62 (involved in protein transport), HIVEP1 which binds specific DNA sequences in the promoters and enhancer regions of several genes and viruses, including human immunodeficiency virus (HIV) [60], UBE3B and ERAP2 (protein modifying enzymes), were among the most abundant proteins uniquely found in  $C92EV/LV_{UC}$ . These proteins could be investigated as markers for developing further purification processes.

Lipid composition of EVs has mainly been described in biological fluids but not in EVs associated with HEK293SF cell cultures [61].  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  share a similar lipid composition, with an enrichment in phosphatidylserine as compared to the parental cells, consistent with the findings of other studies [62].  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  also contained less phosphatidylcholine and phosphatidylethanolamine than their parental cells. It has been reported that the change in distribution of these lipids was involved in the budding of microvesicles [63]. Sphingolipid and cholesterol analysis in LVs/EVs samples would be a good addition to this lipidomic characterization to confirm enrichment in ceramide and cholesterol in EVs and LVs as reported in these studies on lipids involved in the budding process [62,64]. The higher depletion of plasmalogen-PE in  $C92EV/LV_{UC}$  compared to  $C92EV_{UC}$  might be interesting to further study as pl-PE could play an important role in membrane dynamics and intracellular signaling [65]. Discrepancies in the lipidomic profiles observed between  $C92EV_{SEC}$  and  $C92EV_{UC}$  is again likely due to the difference in the EV isolation methods. Techniques for studying lipids should also be further improved to

quantify more accurately lipid species, which could conduct to identifying lipid markers for Clone 92 EVs or LVs.

DNA quantification is of importance especially when it comes to biologics and viral vectors and vaccines particularly because of the stringent regulation. In the field of EVs, DNA identification is often investigated with the perspective of using them as biomarkers. Additional DNA sequencing can be expected in the future.  $C92EV_{SEC}$  cargoes also revealed different types of RNA, including miRNA. The gene ontology analyses of  $C92EV_{SEC}$  confirmed the main components and functions attributed to EVs. For instance, the abundance of genes with binding functions can explain a mechanism of cargo sorting by which RNAs will interact with specific proteins to be packaged into EVs for cell-to-cell transport. The enrichment in genes involved in viral process, viral gene expression and viral transcription can be linked to the fact that EVs and some viruses including retroviruses share the same biogenesis pathways, including the ESCRT-dependant pathway. miRNAs are highly conserved, non-coding, small single-stranded RNA molecules and have the ability to regulate gene expression. They are also involved in diseases mechanisms and have been previously identified in EVs [66]. It was therefore critical to characterize them in  $C92EV_{SEC}$ . Most miRNA found in  $C92EV_{SEC}$  were also found in biofluids [67]. The most abundant miRNAs identified in  $C92EV_{SEC}$  (Figure 6) play a role in all sort of diseases: miR-25-3p and miR-93-5p in gastric cancer [68,69], miR-19b-3p and let-7a-5p in colon cancer [70,71]. Multiple cancers showed abnormal expression of miR-92a-3p while ovarian cancer cells are suggested to release exosomes containing miR-6126 abundantly [72,73]. Some miRNAs found in  $C92EV_{SEC}$  may have a positive regulating role, such as miR-93-5p in glioma or myocardial damage [74,75], miR-191-5p in lung cancer [76], or miR-342-3p in liver cancer [77]. Although it has been suggested that miRNAs are packaged into EVs as a way to dispose of excessive miRNAs, the TRBP containing complex, a member of the RNA-induced silencing complex (RISC) involved in RNA silencing [78] is also enriched in  $C92EV_{SEC}$ . So not only do  $C92EVs$  contain miRNA but they could also provide recipient cells with the miRNA processing machinery which is needed to process those miRNAs [79]. More studies on miRNA uptake from EVs should be conducted. Until then, the effect of miRNA on recipient cells cannot be excluded given the role of miRNAs in a number of diseases.

The fact that EVs share biogenesis pathways and biophysical properties with viral products produced in cell culture platforms such as lentiviral vectors produced in HEK293SF cells and derived cell lines, supports the need to characterize host cell EVs. As discussed above, the production of viral products will induce changes to EVs. In the context of cell and gene therapy, for future *in vivo* gene delivery of LVs, it will be critical to further investigate EV changes and the subsequent intermediate populations upon virus production to determine accurately the product profile and specifications. The effect of co-purified EVs in LV preparations on recipient cells also needs to be evaluated. Indeed, if EVs are proven to be safe, as an associated component to enveloped viral vectors and viral vaccines, they might also have a possible adjuvanting role in the vaccine formulation.

**Supplementary Materials:** The following are available online at [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Figure S1: Analysis of Clone 92 supernatant by flow virometry, Figure S2: Flow virometry density plots showing size (violet side scatter) and green fluorescence (GFP) in the negative control PBS and an in-process Clone 92 sample (here supernatant at a dilution factor of 100), Figure S3: Enrichment of EVs shown by the increase in signal for CD81, BSG, and PDCD6IP and the depletion of cellular protein in EVs shown by a decrease in signal for HSPA5 and HSP90B1 compared to parent cells, Figure S4: Presence of GFP in cells, supernatant and EVs from Clone 92, Figure S5: Flow virometry density plots showing size (violet side scatter) and green fluorescence (GFP) in the negative control PBS and in Clone 92 supernatants, Table S1: List of all proteins identified in  $C92EV_{UC}$  and  $C92EV/LV_{UC}$  with a Mascot score > 30 and more than 1 peptide.

**Author Contributions:** Conceptualization, A.D.M. and A.A.K.; Methodology, A.D.M., S.M.T. and A.A.K.; Investigation, A.D.M., A.T.S. and J.S., Writing – Original Draft, A.D.M. and A.T.S.; Writing – Review & Editing, J.F.G., J.L., S.M.T. and A.A.K.; Funding Acquisition, A.D.M. and A.A.K.; Resources, S.M.T., J.L. and A.K.; and Supervision A.A.K.

**Funding:** A.D.M. is funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) doctoral scholarship. A.A.K. is a recipient of a Canada Research Chair (CRC/240394) of government of Canada.

**Acknowledgments:** We would like to thank Xavier Elisseeff and Alice Bernier for their help with some EV isolations. We also thank Marie-Angélique Sené and David Sharon for their input in processing transcriptomics data. We would like to acknowledge Sam Williamson, and Luc Tessier for their help with mass spectrometry set-up at NRC, and Simon Foote for the data-file conversion pipeline for submission to Mascot.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Naldini, L. Gene therapy returns to centre stage. *Nature* **2015**, *526*, 351-360, doi:10.1038/nature15818.
2. Wang, Z.; Guo, Y.; Han, W. Current status and perspectives of chimeric antigen receptor modified T cells for cancer treatment. *Protein Cell* **2017**, *8*, 896-925, doi:10.1007/s13238-017-0400-z.
3. Sharon, D.; Kamen, A. Advancements in the design and scalable production of viral gene transfer vectors. *Biotechnology and bioengineering* **2017**, 10.1002/bit.26461, doi:10.1002/bit.26461.
4. Ansorge, S.; Lanthier, S.; Transfiguracion, J.; Durocher, Y.; Henry, O.; Kamen, A. Development of a scalable process for high-yield lentiviral vector production by transient transfection of HEK293 suspension cultures. *J. Gene Med.* **2009**, *11*, 868-876, doi:10.1002/jgm.1370.
5. Escors, D.; Breckpot, K. Lentiviral vectors in gene therapy: their current status and future potential. *Archivum immunologiae et therapiae experimentalis* **2010**, *58*, 107-119, doi:10.1007/s00005-010-0063-4.
6. Merten, O.W.; Hebben, M.; Bovolenta, C. Production of lentiviral vectors. *Mol Ther Methods Clin Dev* **2016**, *3*, 16017, doi:10.1038/mtm.2016.17.
7. Nolte-t Hoen, E.; Cremer, T.; Gallo, R.C.; Margolis, L.B. Extracellular vesicles and viruses: Are they close relatives? *Proc Natl Acad Sci U S A* **2016**, *113*, 9155-9161, doi:10.1073/pnas.1605146113.
8. Raposo, G.; Stoorvogel, W. Extracellular vesicles: Exosomes, microvesicles, and friends. *The Journal of Cell Biology* **2013**, *200*, 373-383, doi:10.1083/jcb.201211138.
9. Hessvik, N.P.; Llorente, A. Current knowledge on exosome biogenesis and release. *Cell Mol Life Sci* **2017**, 10.1007/s00018-017-2595-9, doi:10.1007/s00018-017-2595-9.
10. Kalra, H.; Simpson, R.J.; Ji, H.; Aikawa, E.; Altevogt, P.; Askenase, P.; Bond, V.C.; Borràs, F.E.; Breakefield, X.; Budnik, V., et al. Vesiclepedia: A Compendium for Extracellular Vesicles with Continuous Community Annotation. *PLoS Biol.* **2012**, *10*, e1001450, doi:10.1371/journal.pbio.1001450.
11. Mathivanan, S.; Simpson, R.J. ExoCarta: A compendium of exosomal proteins and RNA. *Proteomics* **2009**, *9*, 4997-5000, doi:10.1002/pmic.200900351.
12. Lasser, C.; Thery, C.; Buzas, E.I.; Mathivanan, S.; Zhao, W.; Gho, Y.S.; Lotvall, J. The International Society for Extracellular Vesicles launches the first massive open online course on extracellular vesicles. *Journal of extracellular vesicles* **2016**, *5*, 34299, doi:10.3402/jev.v5.34299.
13. Izquierdo-Useros, N.; Puertas, M.C.; Borràs, F.E.; Blanco, J.; Martínez-Picado, J. Exosomes and retroviruses: The chicken or the egg? *Cell. Microbiol.* **2011**, *13*, 10-17, doi:10.1111/j.1462-5822.2010.01542.x.
14. Cantin, R.; Diou, J.; Belanger, D.; Tremblay, A.M.; Gilbert, C. Discrimination between exosomes and HIV-1: Purification of both vesicles from cell-free supernatants. *J. Immunol. Methods* **2008**, *338*, 21-30, doi:10.1016/j.jim.2008.07.007.
15. White, M.; Whittaker, R.; Gandara, C.; Stoll, E.A. A Guide to Approaching Regulatory Considerations for Lentiviral-Mediated Gene Therapies. *Human gene therapy methods* **2017**, *28*, 163-176, doi:10.1089/hgtb.2017.096.

16. Venereo-Sánchez, A.; Fulton, K.; Koczka, K.; Twine, S.; Chahal, P.; Ansorge, S.; Gilbert, R.; Henry, O.; Kamen, A. Characterization of influenza H1N1 Gag virus-like particles and extracellular vesicles co-produced in HEK-293SF. *Vaccine* **2019**, <https://doi.org/10.1016/j.vaccine.2019.07.057>, doi:<https://doi.org/10.1016/j.vaccine.2019.07.057>.
17. Lavado-García, J.; González-Domínguez, I.; Cervera, L.; Jorge, I.; Vázquez, J.; Gòdia, F. Molecular Characterization of the Coproduced Extracellular Vesicles in HEK293 during Virus-Like Particle Production. *J Proteome Res* **2020**, 10.1021/acs.jproteome.0c00581, doi:10.1021/acs.jproteome.0c00581.
18. Schweizer, M.; Merten, O.W. Large-scale production means for the manufacturing of lentiviral vectors. *Curr. Gene Ther.* **2010**, *10*, 474-486.
19. Do Minh, A.; Tran, M.Y.; Kamen, A.A. Lentiviral Vector Production in Suspension Culture Using Serum-Free Medium for the Transduction of CAR-T Cells. *Methods Mol Biol* **2020**, *2086*, 77-83, doi:10.1007/978-1-0716-0146-4\_6.
20. Cote, J.; Garnier, A.; Massie, B.; Kamen, A. Serum-free production of recombinant proteins and adenoviral vectors by 293SF-3F6 cells. *Biotechnol. Bioeng.* **1998**, *59*, 567-575.
21. Manceur, A.P.; Kim, H.; Misic, V.; Andreev, N.; Dorion-Thibaudeau, J.; Lanthier, S.; Bernier, A.; Tremblay, S.; Gelinas, A.M.; Broussau, S., et al. Scalable lentiviral vector production using stable HEK293SF producer cell lines. *Human gene therapy methods* **2017**, 10.1089/hgtb.2017.086, doi:10.1089/hgtb.2017.086.
22. Graham, F.L.; Smiley, J.; Russell, W.C.; Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol* **1977**, *36*, 59-74, doi:10.1099/0022-1317-36-1-59.
23. Lobb, R.J.; Becker, M.; Wen Wen, S.; Wong, C.S.F.; Wiegmans, A.P.; Leimgruber, A.; Möller, A. Optimized exosome isolation protocol for cell culture supernatant and human plasma. *Journal of extracellular vesicles* **2015**, *4*, 27031, doi:10.3402/jev.v4.27031.
24. Nordin, J.Z.; Lee, Y.; Vader, P.; Mäger, I.; Johansson, H.J.; Heusermann, W.; Wiklander, O.P.B.; Hällbrink, M.; Seow, Y.; Bultema, J.J., et al. Ultrafiltration with size-exclusion liquid chromatography for high yield isolation of extracellular vesicles preserving intact biophysical and functional properties. *Nanomed. Nanotechnol. Biol. Med.* **2015**, *11*, 879-883, doi:<https://doi.org/10.1016/j.nano.2015.01.003>.
25. Pospichalova, V.; Svoboda, J.; Dave, Z.; Kotrbova, A.; Kaiser, K.; Klemova, D.; Ilkovics, L.; Hampl, A.; Crha, I.; Jandakova, E., et al. Simplified protocol for flow cytometry analysis of fluorescently labeled exosomes and microvesicles using dedicated flow cytometer. *Journal of extracellular vesicles* **2015**, *4*, 25530, doi:10.3402/jev.v4.25530.
26. Tang, V.A.; Renner, T.M.; Fritzsche, A.K.; Burger, D.; Langlois, M.A. Single-Particle Discrimination of Retroviruses from Extracellular Vesicles by Nanoscale Flow Cytometry. *Sci Rep* **2017**, *7*, 17769, doi:10.1038/s41598-017-18227-8.
27. Morales-Kastresana, A.; Telford, B.; Musich, T.A.; McKinnon, K.; Clayborne, C.; Braig, Z.; Rosner, A.; Demberg, T.; Watson, D.C.; Karpova, T.S., et al. Labeling Extracellular Vesicles for Nanoscale Flow Cytometry. *Scientific Reports* **2017**, *7*, 1878, doi:10.1038/s41598-017-01731-2.
28. Thery, C.; Amigorena, S.; Raposo, G.; Clayton, A. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. *Curr. Protoc. Cell Biol.* **2006**, Chapter 3, Unit 3.22, doi:10.1002/0471143030.cb0322s30.
29. Thery, C.; Witwer, K.W.; Aikawa, E.; Alcaraz, M.J.; Anderson, J.D.; Andriantsitohaina, R.; Antoniou, A.; Arab, T.; Archer, F.; Atkin-Smith, G.K., 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. *Journal of extracellular vesicles* **2018**, *7*, 1535750, doi:10.1080/20013078.2018.1535750.

30. Wisniewski, J.R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359-362, doi:10.1038/nmeth.1322.

31. Chambers, M.C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D.L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J., et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **2012**, *30*, 918-920, doi:10.1038/nbt.2377.

32. Perkins, D.N.; Pappin, D.J.C.; Creasy, D.M.; Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551-3567, doi:10.1002/(sici)1522-2683(19991201)20:18<3551::Aid-elps3551>3.0.co;2-2.

33. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2018**, *47*, D506-D515, doi:10.1093/nar/gky1049.

34. Haqqani, A.S.; Kelly, J.F.; Stanimirovic, D.B. Quantitative protein profiling by mass spectrometry using label-free proteomics. *Methods Mol Biol* **2008**, *439*, 241-256, doi:10.1007/978-1-59745-188-8\_17.

35. Hulsen, T.; de Vlieg, J.; Alkema, W. BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **2008**, *9*, 488, doi:10.1186/1471-2164-9-488.

36. Sud, M.; Fahy, E.; Subramaniam, S. Template-based combinatorial enumeration of virtual compound libraries for lipids. *Journal of Cheminformatics* **2012**, *4*, 23, doi:10.1186/1758-2946-4-23.

37. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650-1667, doi:10.1038/nprot.2016.095.

38. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T., et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25-29, doi:10.1038/75556.

39. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic acids research* **2019**, *47*, D330-d338, doi:10.1093/nar/gky1055.

40. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic acids research* **2019**, *47*, D419-d426, doi:10.1093/nar/gky1038.

41. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **2019**, *10*, 1523, doi:10.1038/s41467-019-09234-6.

42. Mi, H.; Ebert, D.; Muruganujan, A.; Mills, C.; Albou, L.-P.; Mushayamaha, T.; Thomas, P.D. PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **2020**, *49*, D394-D403, doi:10.1093/nar/gkaa1106.

43. Konadu, K.A.; Huang, M.B.; Roth, W.; Armstrong, W.; Powell, M.; Villinger, F.; Bond, V. Isolation of Exosomes from the Plasma of HIV-1 Positive Individuals. *J. Vis. Exp.* **2016**, 10.3791/53495, doi:10.3791/53495.

44. Pathan, M.; Fonseka, P.; Chitti, S.V.; Kang, T.; Sanwlani, R.; Van Deun, J.; Hendrix, A.; Mathivanan, S. Vesiclepedia 2019: a compendium of RNA, proteins, lipids and metabolites in extracellular vesicles. *Nucleic acids research* **2018**, 10.1093/nar/gky1029, doi:10.1093/nar/gky1029.

45. Keerthikumar, S.; Chisanga, D.; Ariyaratne, D.; Al Saffar, H.; Anand, S.; Zhao, K.; Samuel, M.; Pathan, M.; Jois, M.; Chilamkurti, N., et al. ExoCarta: A Web-Based Compendium of Exosomal Cargo. *J. Mol. Biol.* **2016**, *428*, 688-692, doi:<https://doi.org/10.1016/j.jmb.2015.09.019>.

46. Fahy, E.; Subramaniam, S.; Murphy, R.C.; Nishijima, M.; Raetz, C.R.H.; Shimizu, T.; Spener, F.; van Meer, G.; Wakelam, M.J.O.; Dennis, E.A. Update of the LIPID MAPS comprehensive classification system for lipids. *J. Lipid Res.* **2009**, *50 Suppl*, S9-S14, doi:10.1194/jlr.R800095-JLR200.

47. Van Deun, J.; Mestdagh, P.; Sormunen, R.; Cocquyt, V.; Vermaelen, K.; Vandesompele, J.; Bracke, M.; De Wever, O.; Hendrix, A. The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. *Journal of extracellular vesicles* **2014**, *3*, 10.3402/jev.v3403.24858, doi:10.3402/jev.v3.24858.

48. Luan, X.; Sansanaphongpricha, K.; Myers, I.; Chen, H.; Yuan, H.; Sun, D. Engineering exosomes as refined biological nanoplatforms for drug delivery. *Acta Pharmacologica Sinica* **2017**, *38*, 754-763, doi:10.1038/aps.2017.12.

49. Soo, C.Y.; Song, Y.; Zheng, Y.; Campbell, E.C.; Riches, A.C.; Gunn-Moore, F.; Powis, S.J. Nanoparticle tracking analysis monitors microvesicle and exosome secretion from immune cells. *Immunology* **2012**, *136*, 192-197, doi:10.1111/j.1365-2567.2012.03569.x.

50. van der Pol, E.; Hoekstra, A.G.; Sturk, A.; Otto, C.; van Leeuwen, T.G.; Nieuwland, R. Optical and non-optical methods for detection and characterization of microparticles and exosomes. *Journal of thrombosis and haemostasis : JTH* **2010**, *8*, 2596-2607, doi:10.1111/j.1538-7836.2010.04074.x.

51. Transfiguracion, J.; Tran, M.Y.; Lanthier, S.; Tremblay, S.; Coulombe, N.; Acchione, M.; Kamen, A.A. Rapid In-Process Monitoring of Lentiviral Vector Particles by High-Performance Liquid Chromatography. *Molecular Therapy - Methods & Clinical Development* **2020**, *18*, 803-810, doi:10.1016/j.omtm.2020.08.005.

52. Broussau, S.; Jabbour, N.; Lachapelle, G.; Durocher, Y.; Tom, R.; Transfiguracion, J.; Gilbert, R.; Massie, B. Inducible packaging cells for large-scale production of lentiviral vectors in serum-free suspension culture. *Molecular therapy : the journal of the American Society of Gene Therapy* **2008**, *16*, 500-507, doi:10.1038/sj.mt.6300383.

53. Sakaguchi, N.; Muramatsu, H.; Ichihara-Tanaka, K.; Maeda, N.; Noda, M.; Yamamoto, T.; Michikawa, M.; Ikematsu, S.; Sakuma, S.; Muramatsu, T. Receptor-type protein tyrosine phosphatase zeta as a component of the signaling receptor complex for midkine-dependent survival of embryonic neurons. *Neurosci Res* **2003**, *45*, 219-224, doi:10.1016/s0168-0102(02)00226-2.

54. Chien, C.S.; Li, J.Y.; Chien, Y.; Wang, M.L.; Yarmishyn, A.A.; Tsai, P.H.; Juan, C.C.; Nguyen, P.; Cheng, H.M.; Huo, T.I., et al. METTL3-dependent N(6)-methyladenosine RNA modification mediates the atherogenic inflammatory cascades in vascular endothelium. *Proc Natl Acad Sci U S A* **2021**, *118*, doi:10.1073/pnas.2025070118.

55. Zhao, Y.; Azam, S.; Thorpe, R. Comparative studies on cellular gene regulation by HIV-1 based vectors: implications for quality control of vector production. *Gene Ther.* **2005**, *12*, 311-319, doi:10.1038/sj.gt.3302414.

56. Saphire, A.C.; Gallay, P.A.; Bark, S.J. Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *J Proteome Res* **2006**, *5*, 530-538, doi:10.1021/pr050276b.

57. Zolotukhin, A.S.; Michalowski, D.; Bear, J.; Smulevitch, S.V.; Traish, A.M.; Peng, R.; Patton, J.; Shatsky, I.N.; Felber, B.K. PSF acts through the human immunodeficiency virus type 1 mRNA instability

elements to regulate virus expression. *Mol. Cell. Biol.* **2003**, *23*, 6618-6630, doi:10.1128/mcb.23.18.6618-6630.2003.

58. Thali, M.; Bukovsky, A.; Kondo, E.; Rosenwirth, B.; Walsh, C.T.; Sodroski, J.; Göttlinger, H.G. Functional association of cyclophilin A with HIV-1 virions. *Nature* **1994**, *372*, 363-365, doi:10.1038/372363a0.

59. Sasaki, H.; Nakamura, M.; Ohno, T.; Matsuda, Y.; Yuda, Y.; Nonomura, Y. Myosin-actin interaction plays an important role in human immunodeficiency virus type 1 release from host cells. *Proceedings of the National Academy of Sciences* **1995**, *92*, 2026-2030, doi:10.1073/pnas.92.6.2026.

60. Allen, C.E.; Wu, L.-C. ZAS Zinc Finger Proteins: The Other κB-Binding Protein Family. In *Zinc Finger Proteins: From Atomic Contact to Cellular Function*, Iuchi, S., Kuldell, N., Eds. Springer US: Boston, MA, 2005; 10.1007/0-387-27421-9\_29pp. 213-220.

61. Skotland, T.; Sandvig, K.; Llorente, A. Lipids in exosomes: Current knowledge and the way forward. *Progress in Lipid Research* **2017**, *66*, 30-41, doi:<https://doi.org/10.1016/j.plipres.2017.03.001>.

62. Kreimer, S.; Belov, A.M.; Ghiran, I.; Murthy, S.K.; Frank, D.A.; Ivanov, A.R. Mass-spectrometry-based molecular characterization of extracellular vesicles: lipidomics and proteomics. *J Proteome Res* **2015**, *14*, 2367-2384, doi:10.1021/pr501279t.

63. Hugel, B.; Martínez, M.C.; Kunzelmann, C.; Freyssinet, J.M. Membrane microparticles: two sides of the coin. *Physiology (Bethesda)* **2005**, *20*, 22-27, doi:10.1152/physiol.00029.2004.

64. Bianco, F.; Perrotta, C.; Novellino, L.; Francolini, M.; Riganti, L.; Menna, E.; Saglietti, L.; Schuchman, E.H.; Furlan, R.; Clementi, E., et al. Acid sphingomyelinase activity triggers microparticle release from glial cells. *EMBO J.* **2009**, *28*, 1043-1054, doi:10.1038/embj.2009.45.

65. Nagan, N.; Zoeller, R.A. Plasmalogens: biosynthesis and functions. *Progress in Lipid Research* **2001**, *40*, 199-229, doi:[https://doi.org/10.1016/S0163-7827\(01\)00003-0](https://doi.org/10.1016/S0163-7827(01)00003-0).

66. Valadi, H.; Ekström, K.; Bossios, A.; Sjöstrand, M.; Lee, J.J.; Lötvall, J.O. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **2007**, *9*, 654-659, doi:10.1038/ncb1596.

67. Momen-Heravi, F.; Getting, S.J.; Moschos, S.A. Extracellular vesicles and their nucleic acids for biomarker discovery. *Pharmacology & Therapeutics* **2018**, doi:<https://doi.org/10.1016/j.pharmthera.2018.08.002>, doi:<https://doi.org/10.1016/j.pharmthera.2018.08.002>.

68. Ning, L.; Zhang, M.; Zhu, Q.; Hao, F.; Shen, W.; Chen, D. miR-25-3p inhibition impairs tumorigenesis and invasion in gastric cancer cells in vitro and in vivo. *Bioengineered* **2020**, *11*, 81-90, doi:10.1080/21655979.2019.1710924.

69. Li, L.; Zhao, J.; Huang, S.; Wang, Y.; Zhu, L.; Cao, Y.; Xiong, J.; Deng, J. MiR-93-5p promotes gastric cancer-cell progression via inactivation of the Hippo signaling pathway. *Gene* **2018**, *641*, 240-247, doi:10.1016/j.gene.2017.09.071.

70. Jiang, T.; Ye, L.; Han, Z.; Liu, Y.; Yang, Y.; Peng, Z.; Fan, J. miR-19b-3p promotes colon cancer proliferation and oxaliplatin-based chemoresistance by targeting SMAD4: validation by bioinformatics and experimental analyses. *J Exp Clin Cancer Res* **2017**, *36*, 131, doi:10.1186/s13046-017-0602-5.

71. Ghanbari, R.; Mosakhani, N.; Sarhadi, V.K.; Armengol, G.; Nouraei, N.; Mohammadkhani, A.; Khorrami, S.; Arefian, E.; Paryan, M.; Malekzadeh, R., et al. Simultaneous Underexpression of let-7a-5p and let-7f-5p microRNAs in Plasma and Stool Samples from Early Stage Colorectal Carcinoma. *Biomark Cancer* **2016**, *7*, 39-48, doi:10.4137/BIC.S25252.

72. Cun, J.; Yang, Q. Bioinformatics-based interaction analysis of miR-92a-3p and key genes in tamoxifen-resistant breast cancer cells. *Biomed Pharmacother* **2018**, *107*, 117-128, doi:10.1016/j.biopharm.2018.07.158.

73. Kanlikilicer, P.; Rashed, M.H.; Bayraktar, R.; Mitra, R.; Ivan, C.; Aslan, B.; Zhang, X.; Filant, J.; Silva, A.M.; Rodriguez-Aguayo, C., et al. Ubiquitous Release of Exosomal Tumor Suppressor miR-6126 from Ovarian Cancer Cells. *Cancer Res* **2016**, *76*, 7194-7207, doi:10.1158/0008-5472.CAN-16-0714.

74. Wu, H.; Liu, L.; Zhu, J.M. MiR-93-5p inhibited proliferation and metastasis of glioma cells by targeting MMP2. *Eur Rev Med Pharmacol Sci* **2019**, *23*, 9517-9524, doi:10.26355/eurrev\_201911\_19446.

75. Liu, J.; Jiang, M.; Deng, S.; Lu, J.; Huang, H.; Zhang, Y.; Gong, P.; Shen, X.; Ruan, H.; Jin, M., et al. miR-93-5p-Containing Exosomes Treatment Attenuates Acute Myocardial Infarction-Induced Myocardial Damage. *Mol Ther Nucleic Acids* **2018**, *11*, 103-115, doi:10.1016/j.omtn.2018.01.010.

76. Zhou, L.Y.; Zhang, F.W.; Tong, J.; Liu, F. MiR-191-5p inhibits lung adenocarcinoma by repressing SATB1 to inhibit Wnt pathway. *Mol Genet Genomic Med* **2020**, *8*, e1043, doi:10.1002/mgg3.1043.

77. Liu, W.; Kang, L.; Han, J.; Wang, Y.; Shen, C.; Yan, Z.; Tai, Y.; Zhao, C. miR-342-3p suppresses hepatocellular carcinoma proliferation through inhibition of IGF-1R-mediated Warburg effect. *Oncotargets Ther* **2018**, *11*, 1643-1653, doi:10.2147/ott.S161586.

78. Haase, A.D.; Jaskiewicz, L.; Zhang, H.; Laine, S.; Sack, R.; Gatignol, A.; Filipowicz, W. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep* **2005**, *6*, 961-967, doi:10.1038/sj.embo.7400509.

79. Chendrimada, T.P.; Gregory, R.I.; Kumaraswamy, E.; Norman, J.; Cooch, N.; Nishikura, K.; Shiekhattar, R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **2005**, *436*, 740-744, doi:10.1038/nature03868.