



OPEN

## Proteomics analysis of serum small extracellular vesicles for the longitudinal study of a glioblastoma multiforme mouse model

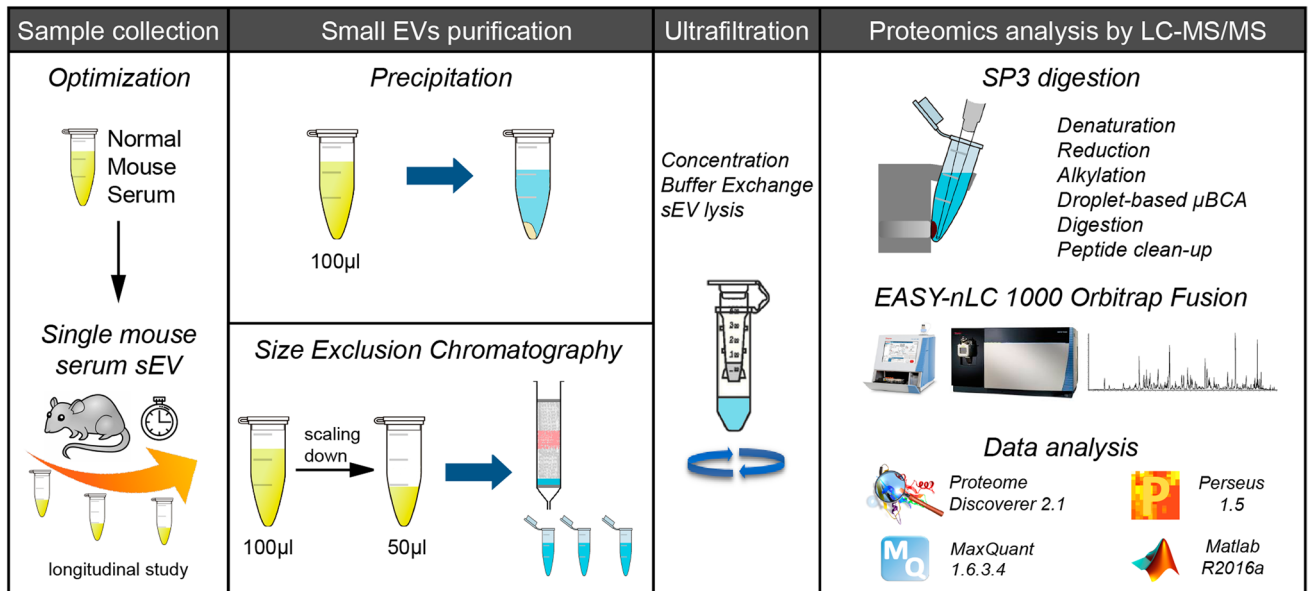
Federica Anastasi<sup>1,2</sup>, Francesco Greco<sup>2,3</sup>, Marialaura Dilillo<sup>2</sup>, Eleonora Vannini<sup>4,5</sup>, Valentina Cappello<sup>6</sup>, Laura Baroncelli<sup>4,7</sup>, Mario Costa<sup>4</sup>, Mauro Gemmi<sup>6</sup>, Matteo Caleo<sup>4,8</sup> & Liam A. McDonnell<sup>1,2</sup>✉

Longitudinal analysis of disease models enables the molecular changes due to disease progression or therapeutic intervention to be better resolved. Approximately 75  $\mu$ l of serum can be drawn from a mouse every 14 days. To date no methods have been reported that are able to analyze the proteome of small extracellular vesicles (sEVs) from such low serum volumes. Here we report a method for the proteomics analysis of sEVs from 50  $\mu$ l of serum. Two sEV isolation procedures were first compared; precipitation based purification (PPT) and size exclusion chromatography (SEC). The methodological comparison confirmed that SEC led to purer sEVs both in terms of size and identified proteins. The procedure was then scaled down and the proteolytic digestion further optimized. The method was then applied to a longitudinal study of serum-sEV proteome changes in a glioblastoma multiforme (GBM) mouse model. Serum was collected at multiple time points, sEVs isolated and their proteins analyzed. The protocol enabled 274 protein groups to be identified and quantified. The longitudinal analysis revealed 25 deregulated proteins in GBM serum sEVs including proteins previously shown to be associated with GBM progression and metastasis (Myh9, Tln-1, Angpt1, Thbs1).

Intercellular communication via the secretion of extracellular vesicles (EVs) is now established as a major pathway in disease progression<sup>1,2</sup>. EVs are a heterogeneous population of vesicles naturally released from cells<sup>3</sup>. EVs are secreted by most cells types, including cancer cells and are present in nearly all body fluids<sup>4,5</sup>. EVs released from cancer cells contain tumour derived material, and thus many studies have focused on their role in the communication between cancer cells, stroma and the immune system, as well as a potential source of biomarkers<sup>6,7</sup>.

It is now well established that small extracellular vesicles (sEVs, diameter < 150 nm) play an important role in transport and communication processes<sup>8</sup>. Small extracellular vesicles offer significant advantages over circulating tumour cells (CTCs) as biomarkers, including their greater abundance and stability<sup>9</sup>. Moreover sEVs have been identified from tumours that did not release detectable CTCs<sup>10</sup>. sEVs represent an opportunity for minimally invasive biomarker research, for instance for cancer diagnosis and monitoring<sup>11</sup>. Glioblastoma multiforme (GBM) is the most frequent and malignant primary tumour of the brain, and has one of the lowest 5-year survival rates of all human cancers<sup>12-14</sup>. One of the reasons for the poor outcome is the late diagnosis, which is based on neuroimaging techniques and tissue biopsies once symptoms are already manifest<sup>15</sup>. sEVs play an important role in the brain tumor environment, regulating the transfer of oncogenic proteins, receptors and soluble molecules that support the tumor progression<sup>16</sup>. Glioblastoma sEVs release angiogenic factors and promote cell proliferation through activation of the PI3K/AKT pathway<sup>17,18</sup>. The prominent role of sEVs in GBM as well as their ability

<sup>1</sup>NEST Laboratories, Scuola Normale Superiore, 56127 Pisa, Italy. <sup>2</sup>Fondazione Pisana per la Scienza ONLUS, 56107 San Giuliano Terme, PI, Italy. <sup>3</sup>Institute of Life Sciences, Sant'Anna School of Advanced Studies, 56127 Pisa, Italy. <sup>4</sup>CNR, Neuroscience Institute, 56124 Pisa, Italy. <sup>5</sup>Fondazione Umberto Veronesi, 20122 Milano, Italy. <sup>6</sup>Istituto Italiano di Tecnologia, Center for Nanotechnology Innovation @NEST, 56127 Pisa, Italy. <sup>7</sup>IRCCS Fondazione Stella Maris, 56018 Calambrone, PI, Italy. <sup>8</sup>Department of Biomedical Sciences, University of Padua, 335122 Padua, Italy. ✉email: liam@fpscience.it



**Figure 1.** Experimental workflow. Comparison between two sEV isolation procedures: precipitation (PPT) and Size Exclusion Chromatography (SEC) using 100  $\mu$ l of mouse serum. SEC was then scaled down on vesicles isolated from 50  $\mu$ l of serum. As a proof of concept, the ultrasensitive microproteomics workflow was then applied to a longitudinal study of a glioblastoma multiforme mouse model. Purified sEVs were concentrated and lysed on protein concentrator spin filters; then extracted, quantified and digested with a modified SP3 protocol. Peptides were analyzed by nLC-MS/MS, and the data analyzed with Proteome Discoverer 2.1 and MaxQuant software. Statistical analysis was performed using Perseus and Matlab.

to cross the blood brain barrier makes them good candidates for the identification of circulating biomarkers to enable earlier diagnosis and treatment<sup>19</sup>.

Xenograft tumour mouse models are widely used for the study of disease progression, biomarker discovery and drug testing, on account of their shorter lifespan and rapid disease progression<sup>20,21</sup>. The increasing requirement to replace, reduce and refine (3Rs) animal research places great emphasis on ensuring maximum information content from each experiment<sup>22</sup>. Longitudinal measurements of individual animals reduce the impact of inter-individual variability by examining the molecular changes for each individual, rather than comparing populations of animals for each time point. The longitudinal analysis of sEVs from individual mice would improve data quality through reducing inter individual variability, as well as better satisfy the 3Rs animal welfare requirement. The longitudinal analysis of circulating sEVs in rodent models is challenging owing to the limited serum/plasma that may be obtained at each time point; for instance just 75  $\mu$ l of serum may be obtained every 14 days for adult mice<sup>23</sup>.

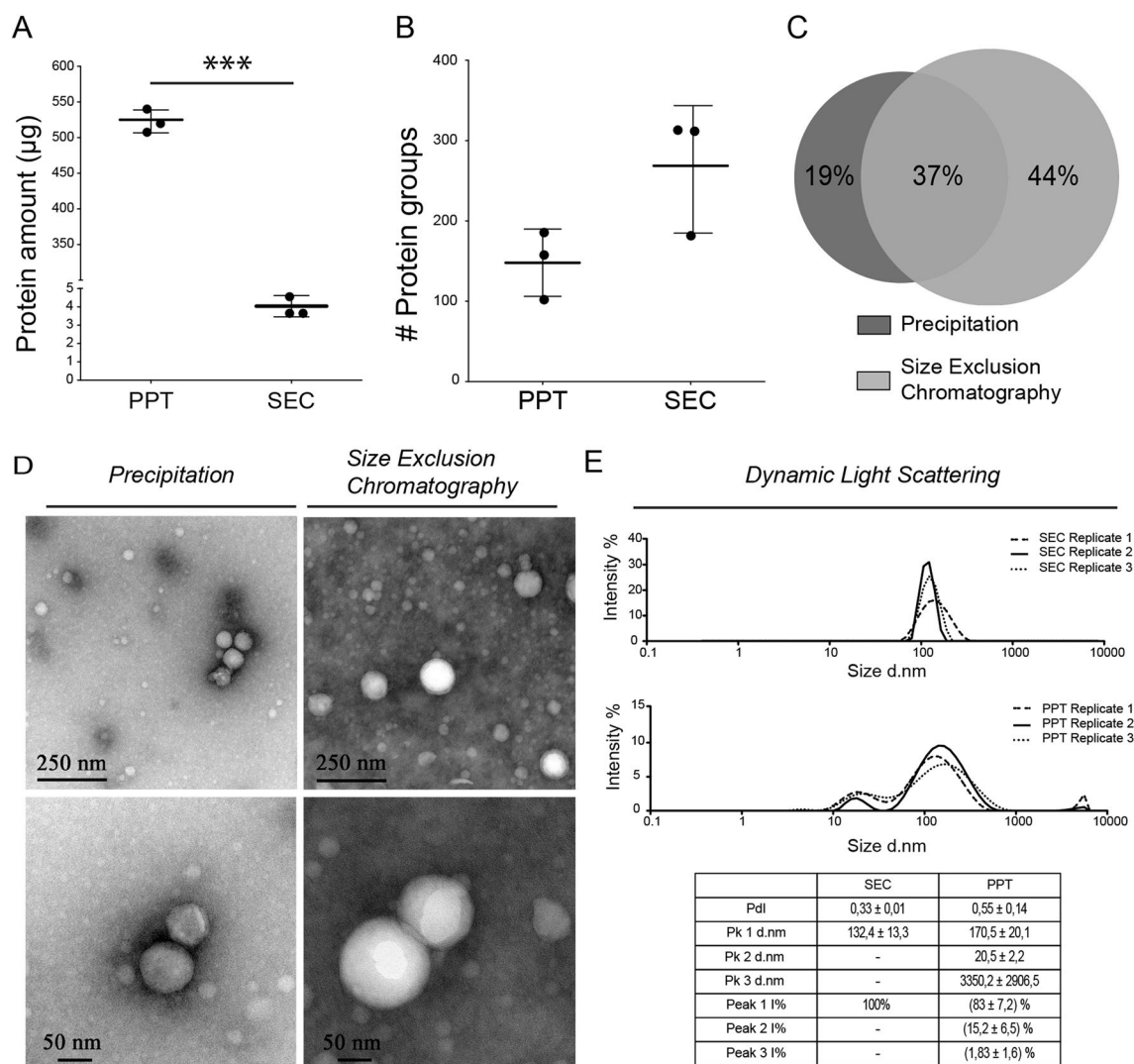
The analysis of the proteome of serum/plasma sEVs requires methods that are able to isolate sEVs from the highly complex chemical background of serum while maintaining sEV integrity. This includes the separation of sEVs from protein aggregates, lipoproteins, and the abundant proteins present in serum/plasma<sup>24,25</sup>. The current gold-standard for sEV isolation is based on differential ultracentrifugation<sup>26</sup>, however the sample volume needed is typically several millilitres<sup>27,28</sup>. Precipitation and SEC based methods have been reported for sEV proteomics studies that use relatively low volume samples<sup>29–31</sup>. Munson et al. used SEC and LC-MS/MS based proteomics to investigate the exosome proteome following asbestos exposure<sup>30</sup> using only 200  $\mu$ l of serum. Here we report the development of a workflow for the proteomics analysis of sEVs from just 50  $\mu$ l of mouse serum and demonstrate its ability to perform longitudinal analysis of GBM development in individual mice.

## Results and discussion

A workflow was developed for the isolation and proteomics analysis of sEVs from low volumes of serum for the longitudinal analysis of GBM development in individual mice. The study design is illustrated in Fig. 1 and described in detail in the “Materials and methods” and Supplementary Information sections.

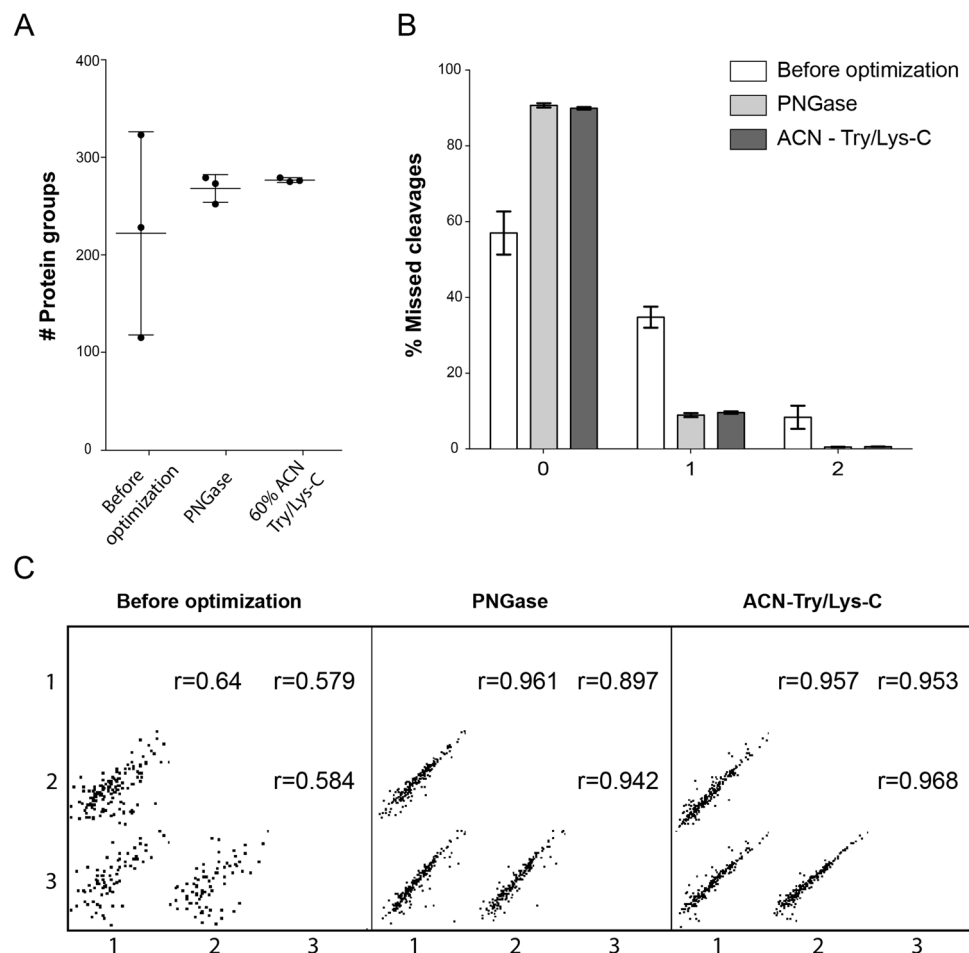
The first objective was to compare two sEV isolation methods that have been reported for serum volumes in the range 200–500  $\mu$ l (still appreciably greater than the 50  $\mu$ l needed for the longitudinal analysis of individual mice), namely SEC and precipitation (PPT). Both procedures were performed in triplicate using 100  $\mu$ l serum aliquots.

The total protein amount that could be extracted from the sEVs isolated by PPT was greater than that isolated by SEC, 523  $\pm$  16  $\mu$ g and 4  $\pm$  0.58  $\mu$ g respectively, Fig. 2A. LC-MS/MS proteomics analysis revealed that the SEC-EV method led to the identification of more proteins than PPT-EV, 334  $\pm$  28 versus 274  $\pm$  21 protein groups respectively (Fig. 2B, Supplementary Dataset SD1). Just 37% of the total number of identified proteins were common to both datasets, 19% were exclusive to PPT-EV and 44% exclusive to SEC-EV, Fig. 2C. A comparison with the proteins that are expected to be present in sEVs (MISEV 2018 guidelines<sup>3</sup>) demonstrated that the



**Figure 2.** Comparison of PPT and SEC sEV isolation methods. **(A)** Extracted protein amount from SEC- and PPT-sEV's. The number of identified protein groups is described in **B**. Welch's test: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . Data are mean  $\pm$  SD. **(B)** Number of identified protein groups of SEC- and PPT-sEV's. Data are mean  $\pm$  SD. **(C)** Venn diagram showing overlap of proteins identified from the SEC and PPT preparations. Dark grey circles indicate PPT IDs, light grey circles indicate SEC IDs. **(D)** Transmission electron microscopy images at low and high magnification of negative staining sEV's obtained by PPT and by SEC. **(E)** Dynamic light scattering analysis and relative dimension distribution obtained for PPT-EV and SEC-EV. Data are mean (nm)  $\pm$  SD.

SEC-EV proteins included a larger number of vesicle markers, which included both membrane proteins (tetraspanins and integrins) and cytosolic proteins (Sdcbp, Hspa8). A comparison with the EV databases ExoCarta Top 100<sup>32</sup> and EVpedia<sup>33</sup> also confirmed the higher purity of the SEC-EVs. The results of these comparisons are summarized in Supplementary Table S2, and the list of ExoCarta Top 100 proteins identified here are reported in Supplementary Table S3. It should be noted that the results reported in Fig. 2 and Supplementary Tables S2 and S3 are conservative, as they were calculated using only those proteins identified in all technical replicates; many other EV proteins were identified in the SEC-EV samples but not in all replicates (e.g., Cd9, Itga2, Anxa4, Anxa5, Anxa7, Vamp8). Finally, a gene ontology analysis of the identified proteins also demonstrated the higher purity of the sEV's isolated by SEC, whereas 'extracellular exosomes' (FDR: 5,8E-110) was identified as the first cellular component category for the SEC-EV proteins, it was the fourth category for the PPT-EV samples and was significantly below the 'extracellular region', 'blood microparticle', and 'extracellular space' categories, Supplementary Table S4. Dynamic Light Scattering (DLS) and Transmission Electron Microscopy (TEM) analyses were also performed to characterize the sEV preparations, and which further supported the LC-MS/MS result that SEC provided purer sEV's. The TEM analysis (Fig. 2D) revealed that both SEC and PPT isolated intact vesicles, the PPT-EVs also included large features that could not be resolved by TEM and that were not present in the SEC-EV samples. The DLS measurements (Fig. 2E) revealed that the SEC-EVs were characterized by a single peak, centred at 130 nm hydrodynamic radius, whereas the PTT-EVs had three distinct populations, centred at 20 nm, 170 nm and 3350 nm hydrodynamic radius.



**Figure 3.** Optimization of proteolytic digestion of sEV proteins obtained from 50  $\mu$ l of serum. **(A)** Number of identified protein groups obtained using an 18 h incubation with a Try/Lys-C mixture, and using two supplemental digestion strategies in which the initial 18 h incubation was changed to a 16 h incubation followed by supplemental addition (2 h. incubation) of the *N*-glycosidase PNGase F or additional Try/Lys-C in 60% acetonitrile. **(B)** Percentage of peptide spectral matches (PSMs) that are zero-missed-cleavage peptides, which are considered an indication of digestion performance. Data are mean  $\pm$  SD. **(C)** Scatter plots between triplicates of different digestion conditions (before optimization, PNGase, and ACN) and relative Pearson correlation coefficients.

**Optimization of proteolysis of SEC-EVs isolated from 50  $\mu$ l of serum.** The comparison of the SEC-EV and PPT-EV isolation methods by LC-MS/MS, DLS, and TEM indicated that the SEC-based system provided sEV samples with less serum protein background and less aggregation. The SEC based isolation method and proteomics analysis were then further optimized to enable serum sEV proteomics analysis from 50  $\mu$ l mouse serum. Initial tests resulted in the extraction of  $2.6 \pm 0.26$   $\mu$ g of protein, which led to the identification of  $222 \pm 104$  protein groups, Fig. 3A. Close examination of the data indicated incomplete proteolysis, as the percentage of zero missed-cleavage peptides (57%) was significantly lower than the 90% threshold we routinely use to quality control the LC-MS/MS data (Fig. 3B).

The degree of proteolytic digestion was increased by the supplemental addition of more enzyme (see Supplementary Table S5 for experimental conditions). The supplemental addition of more trypsin/lys-c increased the number of identified proteins and reduced variability (Fig. 3C), with  $277 \pm 2$  protein groups identified from the sEVs isolated from 50  $\mu$ l mouse serum. Similar results were obtained following the supplemental addition of the glycosidase PNGase F (Supplementary Dataset SD1). The supplemental addition of enzyme increased the number of identified proteins, increased the degree of proteolysis, and increased technical reproducibility. We then applied the SEC-EV workflow on 50  $\mu$ l serum and 50  $\mu$ l plasma from the same mice ( $n = 3$ ). The serum-EV derived proteins exhibited greater reproducibility than the plasma-EV derived proteins; a principal component analysis (PCA) revealed that the main difference was due to the presence of fibrinogen proteins in the plasma-EV samples. The results of this experiment are provided in the Supplementary Information (Supplementary Results SR1, Supplementary Fig. S2).



proteins were not identified. Nevertheless, there was evidence for the endosomal origin of the sEVs, specifically the proteins Ifitm3, Lamp1, Egfr, Trf, ApoE, Anxa2, Hspa8 are associated with late endosomes (GO0005770) and Pigr, Itgb1, Tfrc, Cltc, Plin3, Lrp1, Hrg are associated with endosomes (GO0005768). Note: In this study the term sEV was used instead of exosome because, according to guidelines published by the International Society of Extracellular Vesicles, the term exosome should only be used if it can be explicitly demonstrated that they have a late endosomal origin<sup>3,35</sup>.

**Longitudinal sampling better resolves molecular changes associated with GBM development.** A comparison of the abundances of the sEV proteins from symptomatic GBM (T2) mice with those from independent control animals using a Student' t-test resulted in 16 protein groups with a p-value < 0.05, but which were no longer significant after an FDR correction because of the small number of animals (Supplementary Dataset SD3). Such group-wise analyses compare mean protein intensities for each class, and large numbers of animals/patients are needed to counter the substantial biological variation. The longitudinal sampling of individuals enables the changes in protein expression to be differentiated from inter-subject variability. The longitudinal GBM data were analysed using a mixed-effects model (Supplementary Methods S5), in which the model was individually fitted to each protein. The model better accounts for inter-subject variability since it includes an effect for each time-point and a variable intercept for each mouse, i.e. the baseline level for each mouse is allowed to vary but the changes associated with disease progression should be consistent. The model led to the identification 25 significantly deregulated protein groups (Supplementary Dataset SD4). Proteomics data of the deregulated proteins such as accession number, sequence coverage, peptide score and MW are available in Supplementary Table S6. The proteins that exhibited a statistically significant time effects are reported in Fig. 5 with their relative p-value and regression coefficients for time points T1 and T2. Note that since the data were log<sub>2</sub> transformed during data pre-processing (see Supplementary Methods S5) the regression coefficients can be considered as fold changes with respect to the baseline. 20 out of the 25 significantly deregulated proteins are annotated as glycoproteins (KW-0325); it is well known that sEV-proteome is highly glycosylated<sup>36</sup>, glycans play a major role in sEV cellular recognition and in the uptake of sEVs by recipient cells<sup>37,38</sup>. Among the 25 proteins, Lrp1, Cpn1, Mhy9, and Tln1 are associated to the plasma membrane (GO0098590). The low-density lipoprotein receptor-related protein 1 (Lrp1), positively regulates tau protein uptake and spread between neurons<sup>39</sup>, the tau protein is a known mediator of microtubule-dependent migration of glioblastoma cells<sup>40</sup>. Myosin IIA (Myh9) is involved in cell migration and is required for tumour invasion and metastasis<sup>41,42</sup>. Talin-1 (Tln1) has previously been reported in neuroblastoma derived exosomes<sup>43</sup>, is highly expressed in glioma tissues<sup>44,45</sup> and promotes the motility of glioma cells<sup>46</sup>. Thbs1, Itga2b and Tfrc belong to the external side of plasma membrane (GO0009897). Thrombospondin (Thbs1) is involved in the tumour microenvironment and is known to be more highly expressed in glioblastoma<sup>47,48</sup>. Moreover, Thbs1 was found to be upregulated in hypoxic GBM-exosomes<sup>49</sup>. Transferrin receptor protein 1 (Tfrc, CD71) is known to be expressed on the surface of neuroblastoma exosomes<sup>43</sup>, and has been reported as a marker for radio-resistant GBM cells<sup>50</sup>. Angpt1, Vtn, Vcan and Lamb1 are known mediators of cell adhesion (GO0007155) and are all glycoproteins that likely contribute to sEV-interaction and cellular uptake. Moreover, Itga2b, Vtn, Lamb1, Thbs1 and Angpt1 are positive regulators of cell proliferation, angiogenesis, and cell survival through the extracellular activation of the PI3K/AKT pathway and were found to be significantly increased during GBM progression. Figure 5 shows that some of the deregulated proteins (Apoc4, C1ra, C1sa, Lrp1, Itih1, Apoc3, Angpt1, Itga2b, Itih3, Vcan) were significantly deregulated at both T1 and T2, meaning that protein changes were also detected at a presymptomatic stage; these proteins could be considered as candidate biomarkers for earlier detection but will first require further validation.

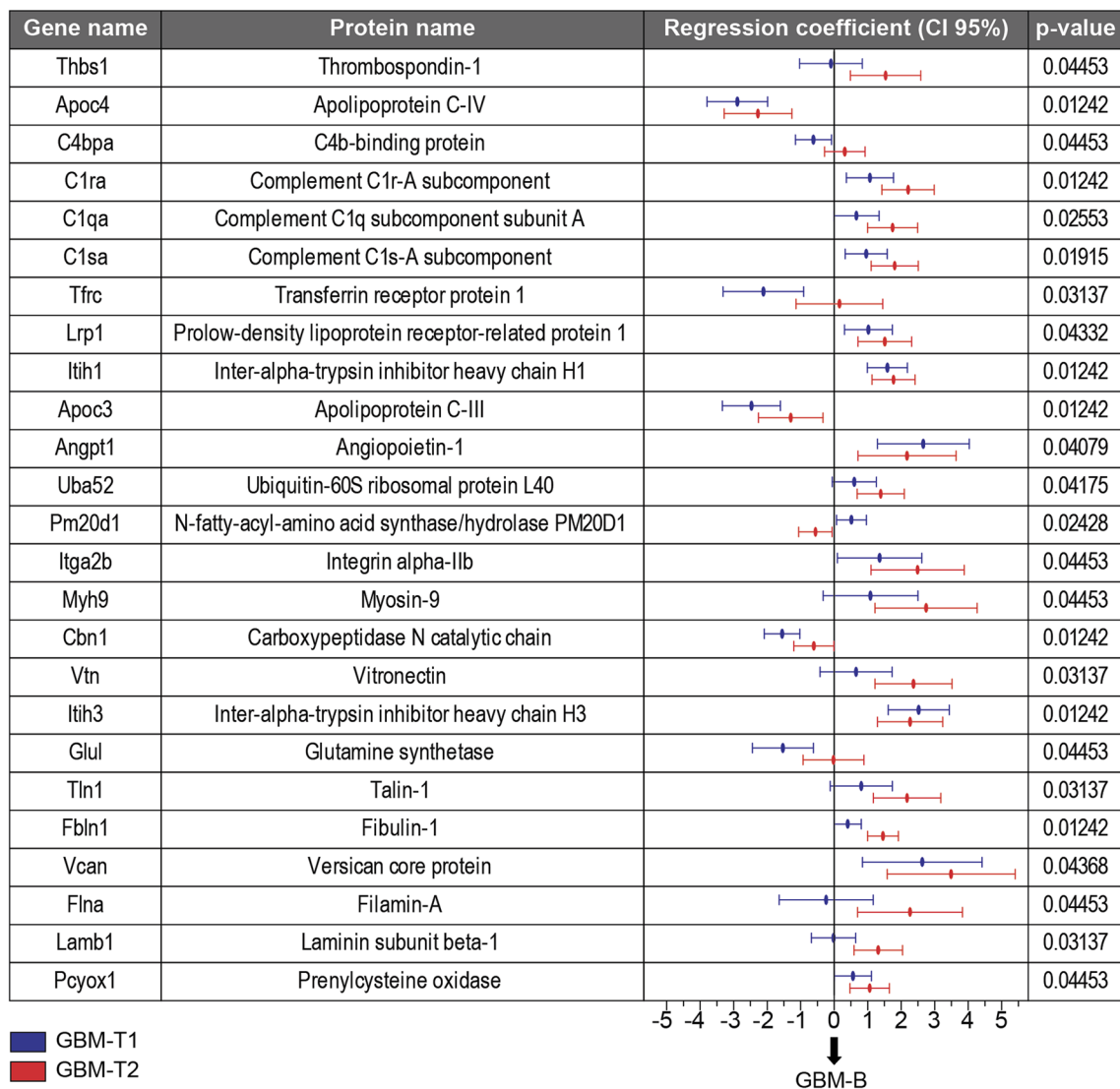
**Concluding remarks.** These results demonstrate that the combination of SEC-EV isolation with SP3 bottom-up proteomics enabled the longitudinal analysis of sEV proteins in individual mouse models of GBM, and enabled the identification of proteins associated with tumor development and candidate biomarkers for the presymptomatic stage.

## Methods

**Materials.** Normal Mouse Serum (10 mL), microBCA protein assay kit, Pierce Concentrator 3 K MWCO 0.5 mL, and Total Exosome Isolation from plasma were purchased from Thermo Fisher Scientific (Rockford, IL). Trypsin/Lys-C mix Mass Spec grade was purchased from Promega (Madison, WI) and PNGase F PRIME from N-Zyme Scientifics (Doylestown, PA). Size Exclusion Chromatography columns (qEV/70 nm single) were purchased from IZON (Christchurch, New Zealand) and 0.22 µm centrifugal filters were purchased from Merck (Darmstadt, Germany). All other reagents and solvents were purchased from Sigma-Aldrich (St. Louis, MO).

**Mice and tumour induction.** Adult C57BL/6J mice were bred at the CNR Institute of Neuroscience animal facility and housed in a 12 h' light/dark cycle, with food and water available ad libitum. All experimental procedures were performed in conformity to the European Communities Council Directive 86/609/EEC and were approved by the Italian Ministry of Health (DLSG 26/2014 Authorization No. 260/2016-PR). The murine glioma GL261 cell line was grown as specified in Vannini et al.<sup>51</sup>. Tumour cell inoculation was performed by stereotaxic guided injection; more details are available in the Supplementary Information (Supplementary Methods S1).

**Mouse serum and plasma collection.** The retro-orbital vein was punctured with gentle pressure and twisting motion with a needle at the sinus level<sup>52</sup>, blood was collected in 1.5 ml Eppendorf tubes. Blood was collected from 4 mice at three different time points: GBM-baseline (15 days before GL261 injection), GBM-T1



**Figure 5.** Linear mixed effects analysis of the longitudinal analysis datasets. List of the significantly deregulated proteins with relative p-values and regression coefficients. The regression coefficients (GBM-T1 in blue and GBM-T2 in red) are plotted with a confidence interval of 95%.

(12 days after tumour induction) and GBM-T2 (21 days after GL261 implantation). T1 is a pre-symptomatic stage and T2 as a post-symptomatic stage of the disease<sup>53</sup>. Blood was also collected from three control mice.

Serum was prepared by allowing the blood to clot at room temperature for 30 min, followed by centrifugation at  $2000 \times g$  for 20 min at  $4^\circ\text{C}$  and collecting the supernatant. Approximately 70  $\mu\text{l}$  of serum were obtained from each blood sample. Plasma samples were prepared as following: blood was collected in a refrigerated 1.5 ml tube containing 5  $\mu\text{l}$  EDTA solution (0.5 M, pH 8), centrifuged at  $2000 \times g$  for 20 min at  $4^\circ\text{C}$ ; the supernatant was then collected. Approximately 100  $\mu\text{l}$  of plasma were obtained from each blood sample. The serum and plasma samples were then stored at  $-80^\circ\text{C}$ .

**Sample pre-treatment.** Commercial Normal Mouse Serum (Thermo Fischer Scientific) was used to compare sEV purification procedures and for all optimization steps. Aliquots of Normal Mouse Serum (100  $\mu\text{l}$  and 50  $\mu\text{l}$ ) were prepared and stored at  $-20^\circ\text{C}$  until use. All other serum and plasma samples were collected as described above. All serum and plasma samples used in this study were pre-treated by centrifugation at  $4000 \times g$  for 30 min at  $4^\circ\text{C}$  and filtered through 0.22  $\mu\text{m}$  spin filters (centrifugation at  $16,000 \times g$  for 1 min at  $4^\circ\text{C}$ ) to remove cell debris and larger EVs, respectively.

**sEV isolation by precipitation.** sEV purification by precipitation was performed using the Total Exosome Isolation reagent. 20  $\mu\text{l}$  of the reagent were added to 100  $\mu\text{l}$  of Normal Mouse Serum and mixed by vortexing until a homogeneous solution was obtained. The solution was left to incubate at  $4^\circ\text{C}$  for 30 min, and then centrifuged at  $10,000 \times g$  for 10 min at  $4^\circ\text{C}$ . The supernatant was discarded and the resulting pellet, which contains the

vesicles, was suspended in 100  $\mu$ l phosphate-buffered saline (PBS) for downstream analysis. PBS was previously filtered through 0.22  $\mu$ m membrane.

**sEV purification by size exclusion chromatography.** SEC columns were equilibrated with PBS (previously filtered with a 0.22  $\mu$ m membrane and sonicated to remove air bubbles). 100  $\mu$ l of pre-treated serum were loaded on to the SEC columns (50  $\mu$ l serum/plasma mixed with 50  $\mu$ l of PBS for the scaled down procedure). Vesicles were eluted using PBS buffer. Fractions of 200  $\mu$ l were collected. The column void volume corresponds to the first eluted millilitre (fractions 1 to 5), followed by sEV's and then proteins.

Preliminary LC-MS/MS measurements were used to establish a suitable balance between a sufficient amount of protein for LC-MS/MS analysis and sEV purity as described in detail in the Supplementary Information (Supplementary Methods S2, Supplementary Fig. S1, Supplementary Table S1).

**Ultrafiltration, buffer exchange and on-filter lysis.** sEV-suspensions obtained by both precipitation (100  $\mu$ l) and SEC (600  $\mu$ l) were charged on 3 kDa MWCO spin filters. The samples were concentrated and rinsed two times with PBS (14,000  $\times$ g at 4  $^{\circ}$ C). Buffer exchange was then performed by three additions (2  $\times$  200  $\mu$ l then 1  $\times$  100  $\mu$ l) of an MS-compatible lysis buffer (LB) containing 1% sodium dodecyl sulfate (SDS), 5 mM ethylenediaminetetraacetic acid (EDTA), 5 mM ethylene glycol-bis( $\beta$ -amino-ethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), 10 mM 4-(2-hydroxy-ethyl)-1-piperazineethanesulfonic acid (HEPES) pH 8.5 and protease inhibitor (cOmplete, Mini, EDTA-free inhibitor Mixture). Note that addition of the lysis buffer and subsequent centrifugation steps were performed at 15  $^{\circ}$ C to avoid SDS precipitation. The buffer-exchanged samples were then concentrated to a final volume of 50  $\mu$ l and the sEV's lysed on the filter by sonication using a Bioruptor Pico (Diagenode, Seraing, Belgium; 10 cycles of 30 s ON and 30 s OFF, 4  $^{\circ}$ C). The sEV-lysate was then recovered and stored at -20  $^{\circ}$ C in 0.5 ml LoBind tubes (Eppendorf) for next day LC-MS/MS analysis.

**Proteomics sample preparation.** Protein digestion was performed using a modified single-pot, solid-phase-enhanced (SP3) sample preparation protocol<sup>54-56</sup>. A detailed description of the protocol is available in the Supplementary Information (Supplementary Methods S3). Briefly, the sEV lysate was thawed and mixed with 1:1 trifluoroethanol together with paramagnetic beads. The lysate-bead mixture was then sonicated using the Bioruptor Pico and proteins denatured by incubation at 95  $^{\circ}$ C for 5 min. The protein mixture was then reduced, alkylated and washed, and the protein concentration determined using a modified micro BCA assay<sup>54</sup>. Proteolytic digestion was performed overnight at 37  $^{\circ}$ C using a Trypsin/Lys-C mixture (Promega) and a protease:protein ratio of 1:25. Further improvements to the digestion were investigated by adding a second digestion step using (i) Try/Lys-C in a protease:protein ratio of 1:75 in 60% acetonitrile, and (ii) PNGase F in a glycosidase:protein ratio of 1:20. An overview of the different digestion conditions is provided in the Supplementary Information (Supplementary Table S5). Finally, the proteolytic peptides were desalted, purified and stored at -20  $^{\circ}$ C.

**nLC-MS/MS analysis.** LC-MS/MS experiments were performed using an Easy-nLC 1000 coupled to an Orbitrap Fusion mass spectrometer (both Thermo Fisher Scientific, Bremen, Germany). A detailed description of the nLC-MS/MS methods is available in the Supplementary Information (Supplementary Methods S4). Briefly, the peptide digest was diluted 1:1 with 10% formic acid prior to injection. The peptides were separated using an EASY-Spray analytical column (ES803: 75  $\mu$ m  $\times$  50 cm, C18, 2  $\mu$ m, 100  $\text{\AA}$ ; Thermo Scientific) using a flow rate of 300 nl/min. A 145-min LC gradient (1  $\mu$ g of peptide digest injected) was used for the comparison of PPT-EV and SEC-EV purification methods and a 75-min gradient (0.5  $\mu$ g of peptide digest injected) was used for all the SEC-EV proteolysis optimizations and the SEC-EV GBM longitudinal study. The Orbitrap Fusion was operated in data dependent Top Speed mode, with a 3 s cycle-time. MS1 scans were performed in the Orbitrap (375 to 1500  $m/z$  at 120 K resolution), ions with charge states from 2+ to 7+ and intensity greater than 5e3 were selected for HCD fragmentation and MS2 scans were acquired in the ion trap using a 1.6  $m/z$  isolation window.

**Protein identification.** Raw data files were processed using Proteome Discoverer 2.1 (Thermo Scientific). The LC-MS/MS data were searched against the UniProt *Mus Musculus* protein database (January 2018, 16,945 entries), supplemented with a home-made common contaminant database (250 sequences). MS/MS spectra were searched with SequestHT search engine<sup>57</sup> using the following settings: 10 ppm precursor mass tolerance and 0.6 Da fragment mass tolerance; up to 2 missed cleavages; minimum peptide length 7 amino acids; methionine oxidation (+15.995 Da) and acetyl (+42.01 Da, N termini) as dynamic modifications; cysteine carbamidomethylation (+57.021 Da) as fixed modification. The search engine results were then filtered for 1% false discovery rate (FDR) using the Percolator algorithm<sup>58</sup> and filtered for a minimum peptide Xcorr score of 1.8. At least one unique peptide was required for definitive protein identification.

MaxQuant software (version 1.6.3.4)<sup>59,60</sup> was used for the longitudinal label-free analysis of serum sEV's from GBM-bearing mice. MS/MS spectra were searched using the Andromeda search engine using the following settings: trypsin digestion; up to 2 missed cleavages; cysteine carbamidomethylation as fixed modification; methionine oxidation and protein N-term acetylation as variable modifications; mass tolerance 20 ppm and 4.5 ppm for the first and main search respectively. Match between runs was enabled using a 0.7-min retention-time alignment. Peptides were filtered to a minimum length of 7 amino acids and 1% FDR. Raw intensities extracted from MaxQuant were used to compare sEV protein levels.

**Data analysis.** Lists of protein groups identified in the PPT-sEV and SEC-sEV datasets were compared using Venn diagrams (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). The identified protein groups were



also compared with the EV markers present in the ExoCarta Top100 database<sup>32,61</sup>, the EVpedia Mouse database<sup>33</sup>, and the 2018 guidelines of the International Society of Extracellular Vesicles<sup>3</sup> to verify the effectiveness of the sEV purification methods. This comparison was performed by string search of the protein names. The results are reported as the number of proteins identified that are in common with the database (ExoCarta Top100), the percentage coverage (EVpedia), and the gene names of the common proteins (MISEV2018).

Gene ontology (GO) analysis was performed using Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 with the whole *mus musculus* genome as statistical background, and which involves a Fisher's exact test followed by an FDR multiple testing correction<sup>62</sup>. All proteomics data was analyzed using Perseus 1.5 software<sup>63</sup>, Microsoft Excel, GraphPad Prism v5 for Windows (GraphPad Software, La Jolla California USA, [www.graphpad.com](http://www.graphpad.com)) and Matlab (ver. R2016a, The MathWorks Inc: Natick, MA, USA, 2010).

Note that for all analyses missing values were not replaced. Common contaminant proteins were eliminated prior to data normalization. Raw protein intensities were log<sub>2</sub> transformed and median normalized. Principal component analysis (PCA) was performed using Perseus to investigate the principal source of variances between the datasets, and the Pearson correlation calculated to assess reproducibility. Protein groups identified in the longitudinal study of the GBM mouse model were first filtered by eliminating serum albumin and immunoglobulins, data were then log<sub>2</sub> transformed and normalized by median subtraction, a linear mixed effect (LME) model was then applied for the three different time points. The LME model is described in detail in the Supplementary Methods S5. Error bars in graphical data represent mean ± standard deviation. All experiments were performed in biological triplicate. Statistical significance was determined by mean comparisons, using a Student's t-test ( $\alpha = 0.05$ ) when the data (or residual) was normally distributed, verified with the Shapiro–Wilk test, or a Mann–Whitney test ( $\alpha = 0.05$ ) when the data did not exhibit a normal distribution.

**Characterization of sEV dimensions and morphology.** Transmission Electron Microscopy (TEM) and Dynamic Light Scattering (DLS) measurements were performed on the sEV's isolated by SEC and PPT. sEV's were purified by SEC/PPT in duplicate from different 100 µl aliquots of serum and then concentrated on 3 kDa spin filters. The sEV samples were then washed 3 times (50% PBS/H<sub>2</sub>O) on the filter to remove any contaminants and decrease salt concentrations and stored at – 20 °C overnight.

**TEM.** Samples were prepared using a two-step protocol for negative staining<sup>64</sup>. The sEV suspensions were adsorbed for 30 min onto carbon-coated 300 mesh copper grids (Electron Microscope Science, Hatfield, PA, USA), washed three times with pure water, and then stained for 30 s with an uranium free staining solution<sup>65</sup>. The grids were then paper-drained and directly analysed with a Libra 120 Plus transmission electron microscope, operating at 120 kV and equipped with an in-column omega filter and 16-bit CCD camera (Zeiss, Oberkochen, Germany). Samples were analysed with ImageJ software (NIH).

**DLS.** 50 µl of sEV suspension were diluted to 1 ml with PBS buffer, mixed well and loaded into a polystyrene cuvette. Each sample was analysed in triplicate at 25 °C using a Zetasizer Nano S (ZEN 1600, Malvern Instruments Ltd, UK) and a scattering angle detection of 173°. The instrument was calibrated using polymer latex spheres (Malvern).

## Data availability

We have submitted all relevant experimental data to the EV-TRACK knowledgebase (EV-TRACK ID: EV190099)<sup>66</sup>. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository<sup>67</sup> with the dataset identifier PXD016473. All identified proteins and deregulated proteins are available in the Supplementary Datasets SD1–SD4. Any further questions or requests should be made to the corresponding author.

Received: 23 July 2020; Accepted: 9 November 2020

Published online: 24 November 2020

## References

- György, B. *et al.* Membrane vesicles, current state-of-the-art: Emerging role of extracellular vesicles. *Cell. Mol. Life Sci.* **68**, 2667–2688 (2011).
- Roy, S., Hochberg, F. H. & Jones, P. S. Extracellular vesicles: The growth as diagnostics and therapeutics; A survey. *J. Extracell. Vesicles* **7**, 1438720 (2018).
- 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).
- Théry, C., Zitvogel, L. & Amigorena, S. Exosomes: Composition, biogenesis and function. *Nat. Rev. Immunol.* **2**, 569–579 (2002).
- Raposo, G. & Stoorvogel, W. Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol.* **200**, 373–383 (2013).
- Maas, S. L. N., Brakefield, X. O. & Weaver, A. M. Extracellular vesicles: Unique intercellular delivery vehicles. *Trends Cell Biol.* **27**, 172–188 (2017).
- Tkach, M. & Théry, C. Communication by extracellular vesicles: Where we are and where we need to go. *Cell* **164**, 1226–1232 (2016).
- An, T. *et al.* Exosomes serve as tumour markers for personalized diagnostics owing to their important role in cancer metastasis. *J. Extracell. Vesicles* **4**, 27522 (2015).
- Poudineh, M., Sargent, E. H., Pantel, K. & Kelley, S. O. Profiling circulating tumour cells and other biomarkers of invasive cancers. *Nat. Biomed. Eng.* **2**, 72–84 (2018).
- Shao, H., Chung, J. & Issadore, D. Diagnostic technologies for circulating tumour cells and exosomes. *Biosci. Rep.* **36**, e00292 (2016).
- Lässer, C. Exosomes in diagnostic and therapeutic applications: Biomarker, vaccine and RNA interference delivery vehicle. *Expert. Opin. Biol. Ther.* **15**, 103–117 (2015).
- Aliferis, C. & Trafalis, D. T. Glioblastoma multiforme: Pathogenesis and treatment. *Pharmacol. Ther.* **152**, 63–82 (2015).

13. Tamimi, A. F. & Juweid, M. *Epidemiology and Outcome of Glioblastoma*. *Glioblastoma* (Codon Publications, 2017). <https://doi.org/10.15586/codon.glioblastoma>.
14. Koshy, M. *et al.* Improved survival time trends for glioblastoma using the SEER 17 population-based registries. *J. Neurooncol.* **107**, 207–212 (2012).
15. Hishii, M., Matsumoto, T. & Arai, H. Diagnosis and treatment of early-stage glioblastoma. *Asian J. Neurosurg.* **14**, 589 (2019).
16. Whitehead, C. A. *et al.* Extracellular vesicles and their role in glioblastoma. *Crit. Rev. Clin. Lab. Sci.* **57**, 227–252 (2020).
17. Kucharzewska, P. *et al.* Exosomes reflect the hypoxic status of glioma cells and mediate hypoxia-dependent activation of vascular cells during tumor development. *Proc. Natl. Acad. Sci.* **110**, 7312–7317 (2013).
18. Yekula, A. *et al.* Extracellular Vesicles in Glioblastoma Tumor Microenvironment. *Frontiers in Immunology* **10**, 3137 (2020).
19. Müller Bark, J., Kulasinghe, A., Chua, B., Day, B. W. & Punyadeera, C. Circulating biomarkers in patients with glioblastoma. *Br. J. Cancer* **122**, 295–305 (2020).
20. Rosenthal, N. & Brown, S. The mouse ascending: Perspectives for human-disease models. *Nat. Cell Biol.* **9**, 993–999 (2007).
21. Cekanova, M. & Rathore, K. Animal models and therapeutic molecular targets of cancer: Utility and limitations. *Drug Des. Dev. Ther.* **8**, 1911 (2014).
22. Prescott, M. J. & Lidster, K. Improving quality of science through better animal welfare: The NC3Rs strategy. *Lab. Anim. (NY)* **46**, 152–156 (2017).
23. Animal Research Advisory Committee—Office of Animal Care. Guidelines for Blood Collection in Mice and Rats. Available at: [https://oacu.oir.nih.gov/sites/default/files/uploads/araac-guidelines/b2\\_blood\\_collection\\_in\\_mice\\_and\\_rats.pdf](https://oacu.oir.nih.gov/sites/default/files/uploads/araac-guidelines/b2_blood_collection_in_mice_and_rats.pdf). Accessed 15 June 2020.
24. Anderson, N. L. & Anderson, N. G. The human plasma proteome: History, character, and diagnostic prospects. *Mol. Cell Proteom.* **11**, 845–867 (2002).
25. Wildes, D. & Wells, J. A. Sampling the N-terminal proteome of human blood. *Proc. Natl. Acad. Sci.* **107**, 4561–4566 (2010).
26. Théry, C., Amigorena, S., Raposo, G. & Clayton, A. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. *Curr. Protoc. Cell Biol.* **30**, 3.22.1–3.22.29 (2006).
27. Ramirez, M. I. *et al.* Technical challenges of working with extracellular vesicles. *Nanoscale* **10**, 881–906 (2018).
28. Kim, J., Tan, Z. & Lubman, D. M. Exosome enrichment of human serum using multiple cycles of centrifugation. *Electrophoresis* <https://doi.org/10.1002/elps.201500131> (2015).
29. de Menezes-Neto, A. *et al.* Size-exclusion chromatography as a stand-alone methodology identifies novel markers in mass spectrometry analyses of plasma-derived vesicles from healthy individuals. *J. Extracell. Vesicles* **4**, 27378 (2015).
30. Munson, P., Lam, Y. W., MacPherson, M., Beuschel, S. & Shukla, A. Mouse serum exosomal proteomic signature in response to asbestos exposure. *J. Cell. Biochem.* **119**, 6266–6273 (2018).
31. Smolarz, M., Pietrowska, M., Matysiak, N., Mielańczyk, Ł. & Widlak, P. Proteome profiling of exosomes purified from a small amount of human serum: The problem of co-purified serum components. *Proteomes* **7**, 18 (2019).
32. Keerthikumar, S. *et al.* ExoCarta: A web-based compendium of exosomal cargo. *J. Mol. Biol.* **428**, 688–692 (2016).
33. EVPedia Extracellular Vesicles Database (MOUSE). Available at: [www.evpedia.info](http://www.evpedia.info). Accessed 17 Jan 2019.
34. Pettersen Hessvik, N. & Llorente, A. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci* **75**, 193–208 (2018).
35. Witwer, K. W. & Théry, C. Extracellular vesicles or exosomes? On primacy, precision, and popularity influencing a choice of nomenclature. *J. Extracell. Vesicles* **8**, 1648167 (2019).
36. Freitas, D. *et al.* Different isolation approaches lead to diverse glycosylated extracellular vesicle populations. *J. Extracell. Vesicles* **8**, 1621131 (2019).
37. Christianson, H. C., Svensson, K. J., Van Kuppevelt, T. H., Li, J. P. & Belting, M. Cancer cell exosomes depend on cell-surface heparan sulfate proteoglycans for their internalization and functional activity. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17380–17385 (2013).
38. Williams, C. *et al.* Assessing the role of surface glycans of extracellular vesicles on cellular uptake. *Sci. Rep.* **9**, 1–14 (2019).
39. Rauch, J. N. *et al.* LRP1 is a master regulator of tau uptake and spread. *Nature* **580**, 381–385 (2020).
40. Breuzard, G. *et al.* Tau regulates the microtubule-dependent migration of glioblastoma cells via the Rho-ROCK signaling pathway. *J. Cell Sci.* **132**, jcs222851 (2019).
41. Medjkane, S., Perez-Sanchez, C., Gaggioli, C., Sahai, E. & Treisman, R. Myocardin-related transcription factors and SRF are required for cytoskeletal dynamics and experimental metastasis. *Nat. Cell Biol.* **11**, 257–268 (2009).
42. Beadle, C. *et al.* The role of myosin II in glioma invasion of the brain. *Mol. Biol. Cell* **19**, 3357–3368 (2008).
43. Marimpietri, D. *et al.* Proteome profiling of neuroblastoma-derived exosomes reveal the expression of proteins potentially involved in tumor progression. *PLoS ONE* **8**, e75054 (2013).
44. Human Protein Atlas. Available at: <http://www.proteinatlas.org>. Accessed 15 June 2020.
45. Uhlen, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
46. Sen, S., Ng, W. P. & Kumar, S. Contributions of talin-1 to glioma cell—Matrix tensional homeostasis. *J. R. Soc. Interface* **9**, 1311–1317 (2012).
47. Huang, T., Sun, L., Yuan, X. & Qiu, H. Thrombospondin-1 is a multifaceted player in tumor progression. *Oncotarget* **8**, 84546–84558 (2017).
48. Daubon, T. *et al.* Deciphering the complex role of thrombospondin-1 in glioblastoma development. *Nat. Commun.* **10**, (2019).
49. Kore, R. A. *et al.* Hypoxia-derived exosomes induce putative altered pathways in biosynthesis and ion regulatory channels in glioblastoma cells. *Biochem. Biophys. Rep.* **14**, 104–113 (2018).
50. Kang, M. K. *et al.* Potential identity of multi-potential cancer stem-like subpopulation after radiation of cultured brain glioma. *BMC Neurosci.* **9**, 15 (2008).
51. Vannini, E. *et al.* The bacterial protein toxin, cytotoxic necrotizing factor 1 (CNF1) provides long-term survival in a murine glioma model. *BMC Cancer* **14**, 449 (2014).
52. Hoggatt, J. *et al.* Rapid mobilization reveals a highly engraftable hematopoietic stem cell. *Cell* **172**, 191–204.e10 (2018).
53. Vannini, E. *et al.* Progression of motor deficits in glioma-bearing mice: Impact of CNF1 therapy at symptomatic stages. *Oncotarget* **8**, 23539–23550 (2017).
54. De Graaf, E. L., Pellegrini, D. & McDonnell, L. A. Set of novel automated quantitative microproteomics protocols for small sample amounts and its application to kidney tissue substructures. *J. Proteome Res.* **15**, 4722–4730 (2016).
55. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).
56. Pellegrini, D. *et al.* Quantitative microproteomics based characterization of the central and peripheral nervous system of a mouse model of Krabbe disease. *Mol. Cell. Proteom.* **18**, 1227–1241 (2019).
57. Washburn, M. P. The H-Index of 'an approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database'. *J. Am. Soc. Mass Spectrom.* **26**, 1799–1803 (2015).
58. Brosch, M., Yu, L., Hubbard, T. & Choudhary, J. Accurate and sensitive peptide identification with mascot percolator. *J. Proteome Res.* **8**, 3176–3181 (2009).
59. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
60. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteom.* **13**, 2513–2526 (2014).

61. ExoCarta: Top 100 Exosome markers. Available at: [http://exocarta.org/exosome\\_markers\\_new](http://exocarta.org/exosome_markers_new). Accessed 17 Jan 2019.
62. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
63. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
64. Forzan, M. *et al.* Generation of virus like particles for epizootic hemorrhagic disease virus. *Res. Vet. Sci.* **107**, 116–122 (2016).
65. Moscardini, A. *et al.* Uranium-free X solution: A new generation contrast agent for biological samples ultrastructure. *Sci. Rep.* **10**, 11540 (2020).
66. Van Deun, J. *et al.* EV-TRACK: Transparent reporting and centralizing knowledge in extracellular vesicle research. *Nat. Methods* **14**, 228–232 (2017).
67. Vizcaino, J. A. *et al.* The Proteomics Identifications (PRIDE) database and associated tools: Status in 2013. *Nucleic Acids Res.* **41**, D1063–D1069 (2012).

## Acknowledgements

This work was performed with the support of Regione Toscana under Grant PAR FAS 2007–2013—“Gliomics: proteomics/genomics/metabolomics for the identification of biomarkers and the development of an ultrasensitive sensing platform for peripheral body fluids: applied to glioblastoma multiforme”, and Fondazione Pisa grant RST 148/16—“Nanobiomarker: nanotechnology for tumour molecular fingerprinting and early diagnosis”. Eleonora Vannini acknowledges support from Fondazione Veronesi. The authors acknowledge Davide Pellegrini for help with the modified SP3 protocol, Sergio Colangiuli for technical support, Francesca Biondi for animal care, and Emilia Bramanti for the help with the DLS measurements.

## Author contributions

L.A.M. and F.A. designed research and wrote the paper. F.A. performed experiments and analyzed data. F.G. analyzed the data. M.D. contributed to the experimental work and analytical tools. V.C. and M.G. performed TEM analysis. M.C. designed animal research and provided the samples. E.V., L.B. and M.C. performed the animal experiment. All authors read and approved the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-77535-8>.

**Correspondence** and requests for materials should be addressed to L.A.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020