Reference and Ghost Proteins Identification in Rat C6 Glioma Extracellular Vesicles

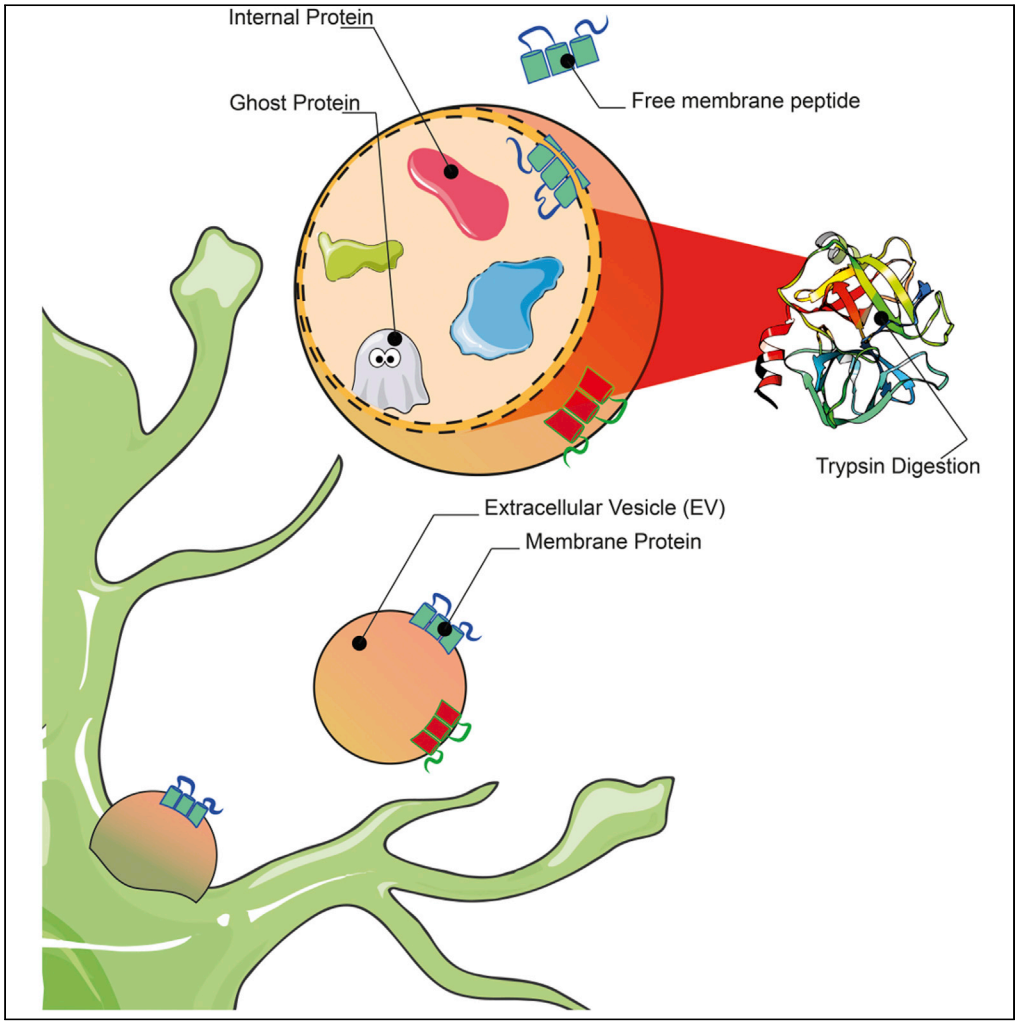
Adriana-Natalia Murgoci, Tristan Cardon, Soulaimane Aboulouard, Marie Duhamel, Isabelle Fournier, Dasa Cizkova, Michel Salzet

iScience

CellPress
OPEN ACCESS

Article

Reference and Ghost Proteins Identification in Rat C6 Glioma Extracellular Vesicles



Adriana-Natalia Murgoci, Tristan Cardon, Soulaimane Aboulouard, Marie Duhamel, Isabelle Fournier, Dasa Cizkova, Michel Salzet

dasa.cizkova@uvlf.sk (D.C.)
michel.salzet@univ-lille.fr (M.S.)

HIGHLIGHTS

Glioma C6 extracellular vesicle protein mapping

Quick steps protocols to map outer/inner membrane EV membrane proteins

C6 glioma cell line EVs contain ghost proteins

Murgoci et al., iScience *23*, 101045
May 22, 2020
https://doi.org/10.1016/j.isci.2020.101045

iScience

CellPress
OPEN ACCESS

Article

Reference and Ghost Proteins Identification in Rat C6 Glioma Extracellular Vesicles

Adriana-Natalia Murgoci,[1,2,4,5] Tristan Cardon,[1,4] Soulaimane Aboulouard,[1] Marie Duhamel,[1] Isabelle Fournier,[1] Dasa Cizkova,[1,2,3,*] and Michel Salzet[1,6,*]

SUMMARY

Extracellular vesicles (EVs) mediate intercellular communication and regulate a broad range of biological processes. Novel therapeutic strategies have emerged based on the use of EVs as biological nanoparticles. To separate isolated EVs from protein aggregates and the external part of EVs membrane proteins, we performed a Trypsin/Lys C digestion treatment of EVs pellets, followed by Amicon filtration. After these steps, all the fractions have been subjected to proteomic analyses. Comparison between 6 h Trypsin/Lys C treatment or non-treated EVs revealed a quantitative variation of the surface proteins. Some surface proteins have been demasked after 6 h enzymatic digestion like CD81, CD82, Ust, Vcan, Lamp 1, Rab43, Annexin A2, Synthenin, and VSP37b. Moreover, six ghost proteins have also been identified and one corresponds to a long noncoding RNA. We thus demonstrate the presence of ghost proteins in EVs produced by glioma cells that can contribute to tumorigenesis.

INTRODUCTION

Extracellular vesicles (EVs) were identified in 1983 (Harding et al., 1983; Pan et al., 1985; Pan and Johnstone, 1983). The authors have shown that culture of immature red blood cells—reticulocytes—labeled transferrin receptors were internalized within the reticulocytes and then were repackaged into small (~50-nm) vesicles and secreted out of the maturing blood reticulocytes into the extracellular space. The name of these vesicles "exosomes" was given by Johnstone et al. (Johnstone et al., 1989). Exosomes belong to a large family of membrane vesicles referred to as EVs, which generally include microvesicles (100–350 nm), apoptotic blebs (500–1,000 nm), and exosomes (30–150 nm) (Johnstone, 2005, 2006). The EVs are involved in many biological processes as intercellular communication messengers (Johnstone, 2005, 2006). Their pathophysiological roles have been characterized by various diseases including cancer (Reclusa et al., 2017; Ruivo et al., 2017; Zhou et al., 2017; Couto et al., 2018; Rajagopal and Harikumar, 2018). Recently, EVs and especially exosomes have been used as therapeutic targets (Harrell et al., 2018; Jing et al., 2018; Kojima et al., 2018; Yamashita et al., 2018; Zhou et al., 2017) or nanotherapeutic agents (Murgoci et al., 2018).

Ultracentrifugation is the most common approach for EVs isolation (Théry et al., 2001; Lane et al., 2017; Momen-Heravi, 2017). Recently, size exclusion chromatography has been introduced after ultracentrifugation (Benedikter et al., 2017) and other separation techniques have been tested, including OptiPrep density-based separation or immune affinity capture using anti-EpCAM-coated magnetic beads (Tauro et al., 2012).Validation of EVs isolation is often based on fluorescent microscopy, scanning electron microscopy, and nanoparticle tracking analysis (NanoSight) (Murgoci et al., 2018). Under an electron microscope, exosomes show characteristic cup-shaped morphology, appearing as flattened spheres with diameters ranging from 30 to 150 nm (Murgoci et al., 2018). To identify the proteins, present at the surface of EVs and the ones within EVs, we performed a proteomic approach after Trypsin/Lys C treatment and Amicon filtration. Using this procedure, we were able to identify plasmatic membranes as well as luminal proteins.

At the same time, we were interested in looking for ghost proteins present in and on the surface of vesicles. Ghost proteins or alternative proteins (AltProts) represent the invisible part of the proteome, because they are not currently identified, annotated or added to the current database using bottom-up proteomics (Delcourt et al., 2018). In fact, mass spectrometric (MS) proteomic strategies are based on protein identification against databases of annotated proteins. However, it has been shown that a large number of proteins have not yet been referenced to these databases, in particular because of the rules used to predict the proteins.

[1]Univ. Lille, Inserm, U-1192 - Laboratoire Protéomique, Réponse Inflammatoire et Spectrométrie de Masse-PRISM, Lille 59000, France

[2]Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava 84510, Slovakia

[3]Department of Anatomy, Histology and Physiology, University of Veterinary Medicine and Pharmacy in Košice, Košice 04181, Slovakia

[4]These authors contributed equally

[5]Present address: Institute of Environmental Medicine, Toxicology Unit, Karolinska Institutet, Stockholm, 17177, Sweden

[6]Lead Contact

*Correspondence: dasa.cizkova@uvlf.sk (D.C.), michel.salzet@univ-lille.fr (M.S.)

https://doi.org/10.1016/j.isci.2020.101045

Thus, AltProts represent proteins of lower molecular weight than reference proteins (RefProts) and are derived from regions of the RNA described as non-coding. AltProts can be translated from 5′ UTR or 3′ UTR or even by shifting on the open reading frame by +2 or +3. Proteins from these regions do not share common sequences with the RefProts produced from the same mRNA and are therefore not isoforms. Some of these AltProts are also described as encoded by non-coding (ncRNA) and long non-coding RNA. Although the role of these proteins remains largely unexplored, some studies have shown their roles in crucial cell processes (Delcourt et al., 2018; Cardon et al., 2019a, 2019b), but as of today they have not been studied in vesicles.

RESULTS

Rat C6 Glioma EVs Purification

We previously published a protocol based on the ultracentrifugation approach for EVs isolation from microglia cells (Murgoci et al., 2018). Here, we present an updated version of this strategy by including an enzymatic digestion step using Trypsin/Lys C for 6 h validated by shotgun proteomics analysis. We thus performed such an approach on vesicles derived from rat C6 glioma cells. Therefore, for C6 glioma EVs isolation, multiple centrifugation steps were used to eliminate debris and other macrovesicles, with the final ultracentrifugation steps to collect EVs pellet. Once EVs were isolated, we have continued their purification by treatment with Trypsin/Lys C. For non-treated EVs (T-0h), the enzyme activity was stopped immediately after its introduction in the tube, so it was considered as a control group compared with EVs incubated for 6 h with enzymes (T-6h) (Figure 1). After enzymatic treatment, the EVs were controlled

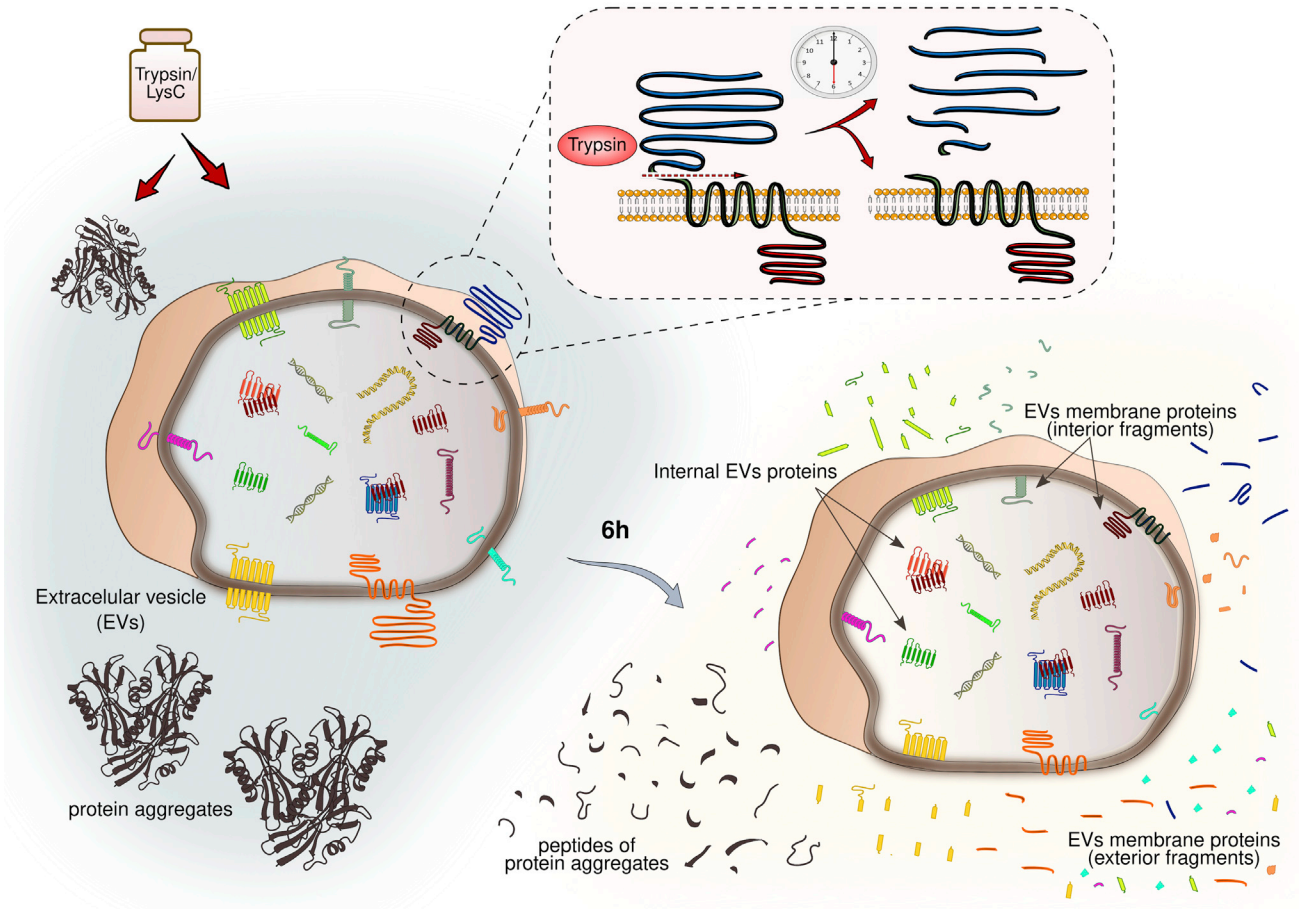


Figure 1. Scheme of Trypsin/LysC Digestion on EVs

The trypsin enzyme cleaves peptide chains at the amino acid residues lysine or arginine, except when either is followed by proline, and lysine-C can cleave lysine followed by proline. After 6 h of Trypsin/Lys C treatment, the exterior parts of exosomes membrane proteins are cleaved as well as protein aggregates in the pellet.

by nanoparticle tracking analysis (Figure S1) and then subjected to an Amicon filtration step to separate EVs from the digested peptides coming from protein aggregates and the external part of EVs membrane proteins. The next step was to extract EVs proteins using radioimmunoprecipitation assay buffer (RIPA) buffer, and by filter aided sample preparation (FASP), the proteins were digested and prepared for tandem MS (MS/MS) analyses (Figure 2).

Two groups of proteins were identified, one corresponding to EVs membrane proteins (external fragment) and protein aggregates and the other corresponding to EVs membrane proteins (internal fragment) and internal EVs proteins (Figure 1). Using this procedure, we can purify EVs released from cells in the cell culture medium and can enrich the list of EVs proteins identified by shotgun proteomics (Data S1, S2, S3, and S4).

Identification of Glioma C6 EVs Proteins by Shotgun Proteomics

A total of 115 proteins have been identified from all EVs released by glioma C6 cells (Table S1). String analysis revealed the presence of two major pathways; the first one is centered on Fn1 and includes, e.g., TGFα1, Notch1, Notch2, Adam10, Syndecan-1, and Syndecan-2, and the second is centered on CD63 and included, e.g., CD81, CD14, CD9, and annexins. Both pathways are connected by CD44, CD9, and RhoA proteins (Figure S2A). All these proteins are known to be implicated in tumorigenesis, therefore the EVs carry a cell-type-specific signature (Mathivanan et al., 2010). KEGGS analyses revealed 11 proteins involved in cancer (CD44, CD63, RhoA, Socd3,Sdc1, Itgb1, Fn1, Tgfb1,Vcam, Rras, Rac1) and 14 in glioblastoma (CD44, FnA, Itgb1, Tgfb1,ECm1,App, Anxa1, HSP90aa1,Flot1, HSP90ab1,Rab7a, Pdcd6ip, Dnja2, Mvp). From the 115 identified proteins, whatever the fractions (internal or external), 15 proteins are unique to EVs not treated with Trypsin/Lys C (T-0h) and 100 proteins are specific to EVs treated for 6 h (T-6h) (Figure S2B).

Among the 15 proteins identified from non-treated EVs, a pathway including GABA(A) receptor-associated protein-like 1 and 2 and their receptors were identified, linked to Tollip, Ran, PSMA1, and PSMA4. By contrast for T-6h-treated EVs, a unique pathway is found centered on the guanine nucleotide-binding protein family. After MaxQuant followed by Perseus analyses on the triplicate experiments conducted with a p value 1%, a heatmap has been constructed (Figure S3). From the Heatmap, good separation can be observed between the isolated

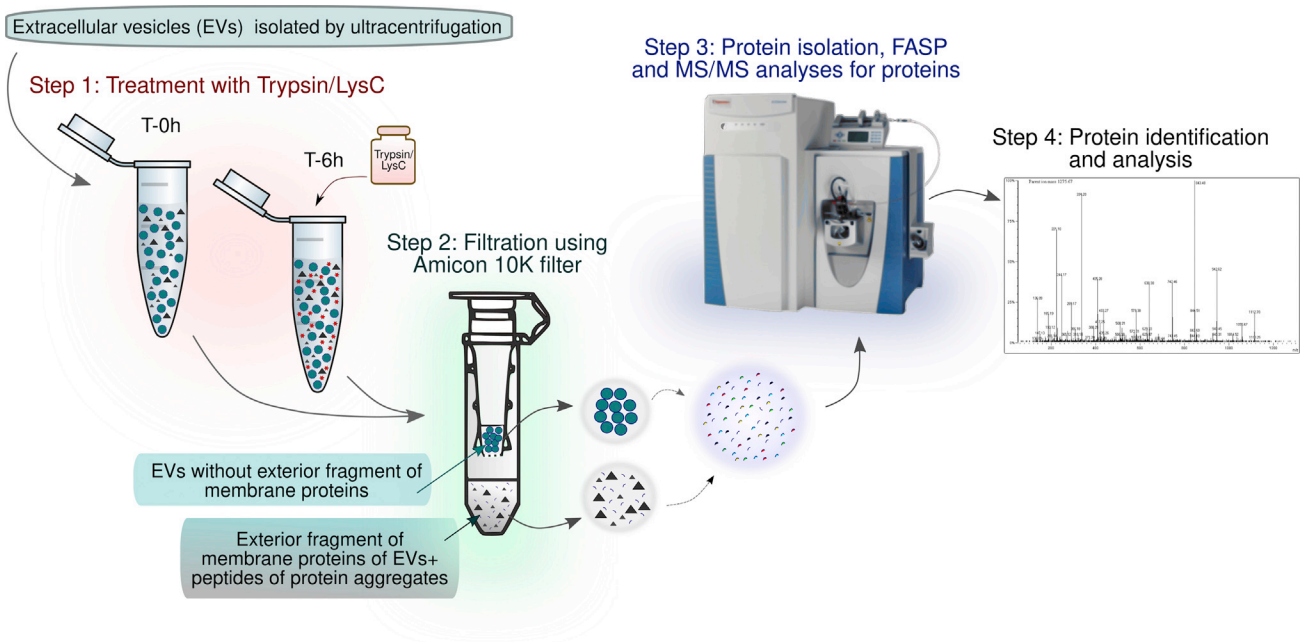


Figure 2. Scheme of EVs Treatment with Trypsin/LysC

After ultracentrifugation at 100,000 × *g* for 70 min, the EVs were incubated for 6 h with 0.2 μg/mL of Trypsin/Lys C (T-6h) or the enzyme activity was immediately stopped (T-0h). Then the vesicles were filtered through Amicon filter 10 kDa, peptides digested by trypsin were recovered in collection tube, and the treated EVs remained on the filter. The filter was transferred into a new tube, proteins were extracted from EVs, and FASP (trypsin digestion) was performed overnight. In the end all the peptides were analyzed by MS/MS.

EVs and the secretome (proteins that passed the Amicon filter, EV). Two branches separating the EVs treated (T-6h) or non-treated (T-0h) are observable (Figure S3). One of the two branches separates the secretome proteins from the EVs T-6h proteins. The secretome proteins is then separated by two sub-branches between T-0h and T-6h. However, some T-0h EVs proteins are masked by the T-0h secretome proteins but can be separated from them in subsidiary branches. This indicates that trypsin digestion is necessary to better separate the EVs from the rest of the cellular secreted factors (Data S1, S2, S3, and S4). We then decided to separate the EVs proteins that are at the membrane and the ones that are extracellular to those that are inside. ANOVA tests were performed with non-supervised clustering of samples (Figure 3A). Three clear clusters were highlighted. The first is vertical separating samples between membrane proteins and internal proteins; the second and the third clusters are horizontal describing proteins groups. Cluster 1 regroups 29 proteins, Cluster 2 groups 17 proteins, and Cluster 3 groups only eight proteins (Table 1). Based on the identified proteins in the different samples, schemes of EVs protein contents treated or non-treated have been drawn (Figure 4). Enrichment analyses of each of these three clusters showed that cluster 1 corresponds to proteins involved in cell proliferation, growth, differentiation, adhesion, migration, and apoptosis (Figure 4). Cluster 2 corresponds to proteins involved in cell transformation, tumor microenvironment, and cancer. Cluster 3 regroups proteins involved in cell migration, internalization, proliferation, and tumor microenvironment in the C6 glioma (Figure 3B and Data S5, S6, and S7). Comparison between clusters suggests a strong pro-tumoral potential of C6 glioma EVs (Figure S2 and Data S5, S6, and S7). Pre-treated and non-pre-treated EVs revealed that the pre-clearance of extravesicular proteins enhanced our ability to identify a broader range of EVs-specific protein species. As a matter of example, we identified a variation of expression of the EVs surface proteins CD9, CD63, Mac2 BP, lactadherin, poly-ubiquitin b, neurensin, Celsr1, and Denn5b in non-treated EVs, whereas in treated EVs, we detected CD81, CD82,Ust, Vcan, Lamp 1, rab43, Cspg4, Annexin A2, Synthenin 1, and VSP37b. Similarly, Actbl2, P-deshydrogenase, Acan, albumin, osteopotin, Sec14-like2, and pdcd6ip are found only at T-6h (Figure 4). These results establish evolution of surface protein nature in EVs produced by glioma cells over time.

Ghost Protein Identification

The internal fraction of proteomic analysis shows some AltProts in the internal part of the EVs. In total six AltProts are identified with minimum one unique peptide (Table 2), showing different expression after trypsin digestion (Table 3). IP_2634467 (AltOdf3l1) is identified in both conditions, treated and non-treated EVs; other AltProts like IP_2656216 (AltKlhl29) are recovered only after a 6 h digestion. Finally, the majority of AltProts are identified at 0 h, e.g., IP_2659453, an AltProt formed from an ncRNA (LOC103695286) specifically identified in MS/MS. This AltProt has some homology with the predicted, but not observed,

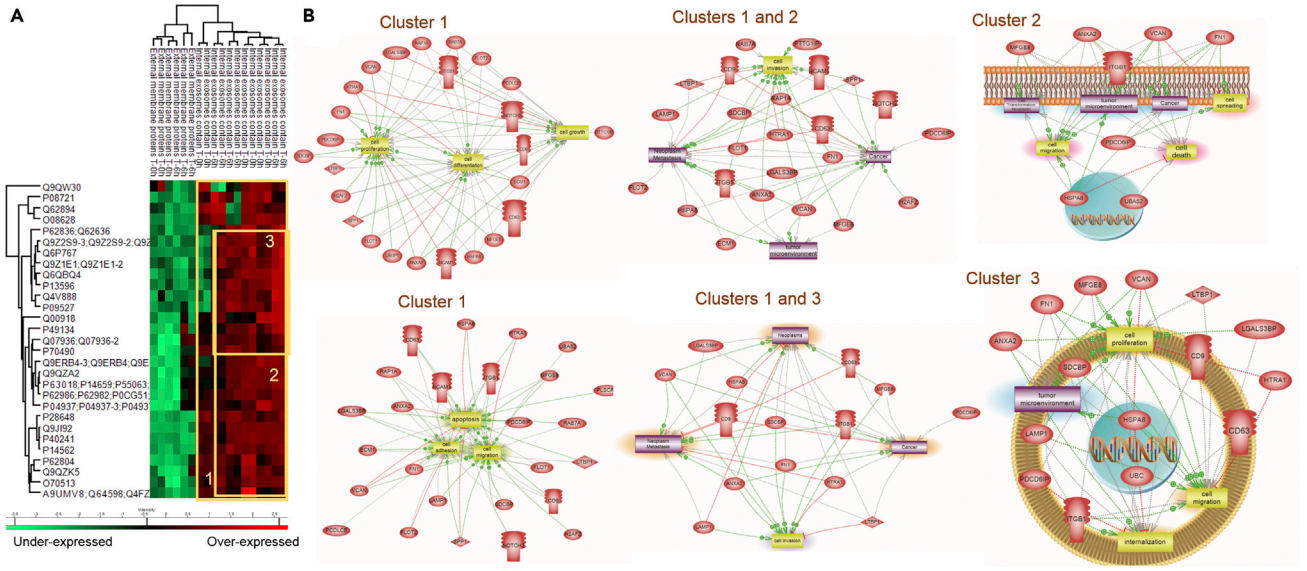


Figure 3. Proteomic Analyses of rat Glioma C6 EVs

(A) Heatmap of the FASP sample obtained according to the different steps described in Figure 2; external samples are the first elution representing the membrane digested protein at 6 h and some free proteins, internal samples are the proteins extracted from the EVs (top part of the Amicon).

(B) Systemic biology analyses. The pathways from systemic biology analyses of clusters 1, 2, and 3 from the heatmap.

Cluster 1	Cluster 2	Cluster 3
H2afj; H2afz	H2afj; H2afz	Fn1
Pcolce	Lgals3bp	Uba52; Rps27a; Ubb; Ubc
Lgals3bp	Fn1	Itgb1
Fn1	Uba52; Rps27a; Ubb; Ubc	Hspa8
Spp1	Lamp1	Mfge8
Rab7a	Cd63	Anxa2
Uba52; Rps27a; Ubb; Ubc	Cd9	Vcan
Ncam1	Itgb1	Pdcd6ip
Lamp1	Hist1h4b	
Cd63	Hspa8	
Cd9	Mfge8	
Itgb1	Ltbp1	
Hist1h4b	Anxa2	
Rap1a; Rap1b	Vcan	
Hspa8	Sdcbp	
Mfge8	Pdcd6ip	
Ltbp1	Htra1	
Anxa2		
Tmem55a		
Ecm1		
Pttg1ip		
Plscr3		
Vcan		
Sdcbp		
Notch2		
Pdcd6ip		
Htra1		
Flot1		
Flot2		

Table 1. Clusters of Specific Proteins Identified in the Internal Part of Glioma C6 EVs in the Time Course of Trypsin/Lys C Digestion after Permutation-Based FDR 0.01

protein rCG63645. This homology can drive the structure and function prediction of this AltProt. IP_2613134 (AltOtud7b), IP_2683599 (AltThoc7), and IP_2577320 (AltMilr1) are specifically identified in non-treated EVs. This suggests that some AltProts might be secreted directly in the media, whereas others are secreted into EVs.

DISCUSSION

Taken together, our results confirm that C6 glioma EVs carry a majority of proteins that are involved in tumor transformation and proliferation. Specific proteins identified in EVs are involved in tumor progression

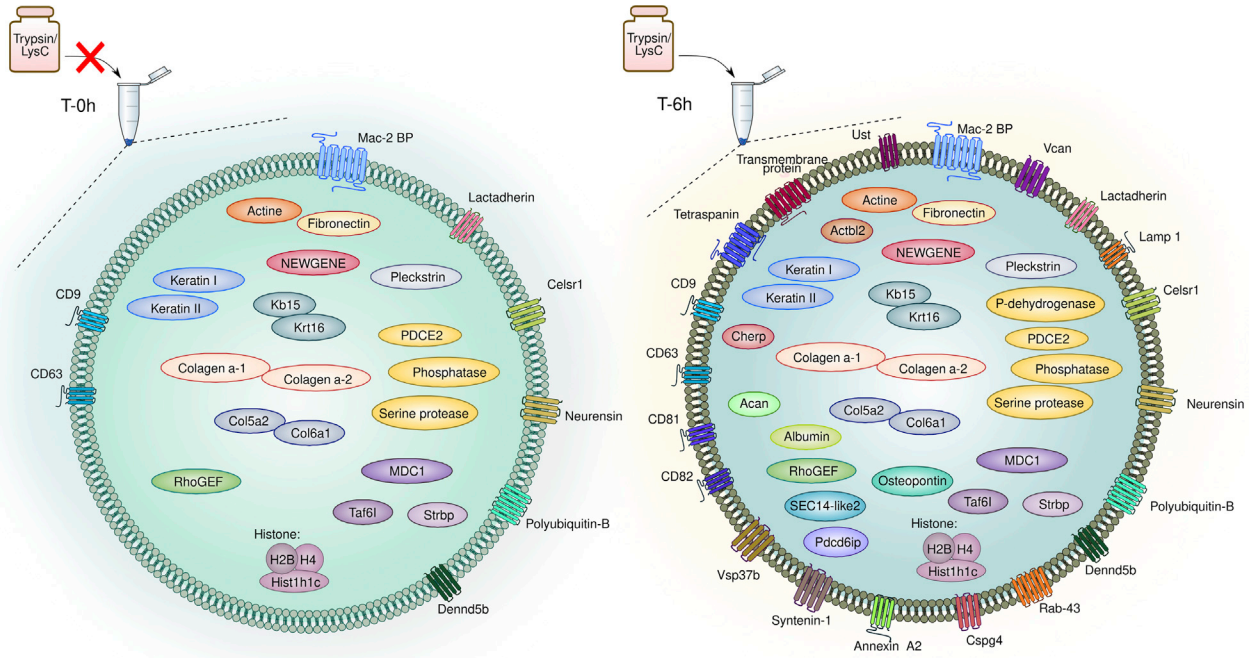


Figure 4. Proteins of Glioma EVs

Schema of proteins identified from C6 glioma EVs non-treated (T-0h) and treated 6h with Trypsin/Lys C (T-6h).

and bad prognosis. For example, Mac-2 BP binds with galectin-1, galectin-3, and galectin-7 and is associated with shorter survival, occurrence of metastasis, and/or a reduced response to chemotherapy in patients with different types of malignancy (Grassadonia et al., 2002). Similarly, Notch1, Notch2, and osteopontin are also bad prognostic factors in glioma (Saito et al., 2015; Zhao et al., 2015).

We also identified Dennd5b. Dennd5b and Dennd5a are markers of glioma according to the Brain Atlas (https://www.proteinatlas.org/). This protein family is also known to be epileptic marker (Han et al., 2016). These results can explain why some patients present epileptic behavior linked to glioma development. Finally, we detected six AltProts in glioma EVs. One corresponds to a non-coding RNA. The other five have their RNA also coding for reference proteins known to be involved in glioma: Kelch like family member 29 (Klhl29), OTU domain-containing protein 7B (Otud7b) known to act as a negative regulator of the B-cell response, outer dense fiber protein 3-like protein 1 (Odf3l1) marker of glioma hypoxia, THO complex subunit 7 homolog (Thoc7), and Milr1 (allergin-1), which plays an inhibitory role in the degranulation of mast cells. Taken together, we demonstrate and describe for the first time the presence of AltProts, which can be specific biomarkers and may be useful to target glioma EVs in body fluids for diagnosis.

LIMITATIONS OF THE STUDY

In agreement with the fact that the AltProts are not referenced, it is complicated at present to monitor and quantify the AltProts by classical approach such as the western blot based on the use of antibodies. The MS strategy remains the best method for identifying this type of protein. The expression of RNA transcripts can also provide information on the expression capacity of these proteins, although the presence of an RNA does not guarantee the production of the protein. The MS strategy uses AltProt prediction; if the databases are getting richer and safer, certain models like the RAT still present a not well-annotated genome. Thus the AltProt database based on Rnor_6.0 genomic data needs to evolve. Finally, if we demonstrate the presence of AltProt in variable amounts under different conditions, their present function remains. Studies of the AltProt interactome have previously been undertaken (Cardon et al., 2019a, 2019b) and have given some clues and showed that they may be involved in phenotypic changes, and in the signaling pathways of gene expression (Chen et al., 2020).

AltProt Accession	Gene	Transcript Accession	Unique Peptide	PSM	Protein Sequence	DNA Sequence	Type	Coverage%
IP_2656216	Klhl29	XM_008764550.2,XM_017594070.1	1	2	MGQAVWRCVSWTC SYHSCVTMTCRRLELSLENRV	ATGGGACAAGCTG TATGGAGATGTGTGAG CTGGACTTGTTCTTAC CATTCCTGTGTCACGA TGACATGCCGAAGGC TGGAGCTGTCCCTAG AGAACAGGGTGTGA	5′UTR	24
IP_2659453	LOC103695286	XR_001838245.1	1	1	MFCFCFVLFCFFLELRT EPRALCLLGKRSTTELPPIRDF	ATGTTTTGTTTTTGTT TTGTTTTGTTTTGTT TTTTTCTGGAGCTGAGG ACCGAACCCAGGGCTTT GTGCTTGCTAGGCAAGCG CTCTACCACTGAACTACC CCCGATTAGGGACTTTTAA	ncRNA	41
IP_2613134	Otud7b	XM_006232957.3	1	2	MIKPLTSVLTIPSIFSVPT PQDTKEQKEALRRKV VERRKVLISLHTLFQKTLR CVVGWWWSAAFRTS TIRLLGLLASKYIPVCSRLCSC	ATGATCAAACCACTTA CCTCCGTTCTGACCATT CCCTCTATATTCTCCGT TCCCACACCTCAAGAC ACTAAAGAACAGAAGG AGGCTCTTAGAAGAAA AGTAGTGGAGAGAAGA AAAGTACTAATTTCCTTA CATACTTTATTCCAGAAG ACACTGAGGTGTGTGG TGGGGTGGTGGTGGTC AGCAGCCTTTAGGACTT CTACAATCAGGCTTCTA GGCCTGCTTGCTAGCAA GTACATACCTGTTTGTTCC AGACTCTGCTCATGCTAG	3′UTR	9

Table 2. Description of the AltProt Identified in the EVs

(Continued on next page)

AltProt Accession	Gene	Transcript Accession	Unique Peptide	PSM	Protein Sequence	DNA Sequence	Type	Coverage%
IP_2634467	Odf3l1	XM_008766272.2,XM_017595650.1	1	5	MTTDRGWGLKLCRLLLQQDSR RGHSQAPSPPPGSKEVPVGS	ATGACCACAGACAGAGG ATGGGGTTTGAAGTTGTG CAGGCTTCTGCTCCAGC AGGACTCAAGGAGGGG CCATTCCCAGGCCCCCT CCCCACCCCCAGGCTC TAAGGAAGTGCCT GTGGGATCCTGA	5′UTR	20
IP_2683599	Thoc7	XM_017599648.1	1	1	MIIKRSLKIFSRTLVWWY LVPCCFLLHLCLTSHM GCKLSDLGAESRTW VLRKSRMHS	ATGATCATAAAGCGAA GTTTAAAGATTTTTAGT CGCACATTGGTGTGG TGGTATTTAGTCCCTTGC TGTTTCCTACTGCATTT ATGTTTGACATCCCACAT GGGCTGTAAACTCTCAGA CTTGGGTGCTGAGAGCC GAACTTGGGTCCTCAGG AAGAGCAGAATGCATTCTTAA	5′UTR	16
IP_2577320	Milr1	XM_017597497.1,XM_008768388.2	1	1	MCMGQHMPRYVHA GRRTSWGSQLSPYS MRPWHRTKVIKFGG QHGWTPSQHT	ATGTGCATGGGCCAAC ATATGCCAAGGTATG TGCACGCAGGTCGGAGG ACGAGTTGGGGGAGTCA GCTTTCTCCTTACTCCA TGAGACCCTGGCATAGA ACTAAGGTCATCAAGTTT GGTGGCCAGCACGGTTGG ACCCCGAGCCAACACACCTGA	3′UTR	18

Table 2. *Continued*

AltProts identification: accession number according to OpenProt annotation correlated to the gene produced the AltProt; column "type" lists the kind of AltProt based on the location in the RNA: AltProts come from the 5′ and 3′ UTRs, by a shift in the CDS, or from ncRNA. The transcript accession is also mentioned permitting the identification of the RNA and gene; in the AltProt sequence in amino acid and nucleic acid, the sequence of the unique peptide identified has been underlined.

AltProt Accession	Internal EVs 0 h	External Membrane 0 h	Internal EVs 6 h	External Membrane 6 h
IP_2656216	0	0	2	0
IP_2659453	1	0	0	0
IP_2613134	2	0	0	0
IP_2634467	3	0	2	0
IP_2683599	1	0	0	0
IP_2577320	1	0	0	0

Table 3. Repartition of AltProts Identification on the Samples Condition

Number of identifications in the different conditions: internal/external at 0 h or 6 h of treatment by trypsin. The external parts are related to the first elution after treatment of the intact EVs by trypsin (n = 3), the internal part combined a triplicate passing the filter after the digestion and the top part of the Amicon (n = 6).

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101045.

ACKNOWLEDGMENTS

This research was supported by a collaboration between the PRISM and grants from Ministère de L'Education Nationale, L'Enseignement Supérieur et de la Recherche, Inserm, France (M.S., I.F.), and University de Lille 1, France (A.-N.M.), APVV 15-0613, Slovakia (D.C.), STEFANIK, France and Slovakia (M.S., D.C.), VEGA 1/0376/20, Slovakia (D.C.), ERA-NET AxonRepair, Netherlands (D.C.).

AUTHOR CONTRIBUTIONS

A.-N.M. performed all EVs purification, S.A. performed the proteomic studies. M.S. and T.C. analyzed proteomic data. M.S., I.F., and D.C. obtained funding for the project. A.-N.M. and M.S. designed the study and wrote the manuscript with contributions from all authors. A.-N.M. drew Figures 1, 2, and 4.

DECLARATION OF INTERESTS

The authors declare no competing interest.

Received: February 27, 2020
Revised: March 23, 2020
Accepted: April 3, 2020
Published: May 22, 2020

REFERENCES

Benedikter, B.J., Bouwman, F.G., Vajen, T., Heinzmann, A.C.A., Grauls, G., Mariman, E.C., Wouters, E.F.M., Savelkoul, P.H., Lopez-Iglesias, C., Koenen, R.R., et al. (2017). Ultrafiltration combined with size exclusion chromatography efficiently isolates extracellular vesicles from cell culture media for compositional and functional studies. Sci. Rep. *7*, 15297.

Cardon, T., Salzet, M., Franck, J., and Fournier, I. (2019a). Nuclei of HeLa cells interactomes unravel a network of ghost proteins involved in proteins translation. Biochim. Biophys. Acta *1863*, 1458–1470.

Cardon, T., Hervé, F., Delcourt, V., Roucou, X., Salzet, M., Franck, J., and Fournier, I. (2019b). Optimized sample preparation workflow for improved identification of ghost proteins. Anal. Chem. https://doi.org/10.1021/acs.analchem.9b04188.

Chen, J., Brunner, A.D., Cogan, J.Z., Nuñez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D., and Weissman, J.S. (2020). Pervasive functional translation of noncanonical human open reading frames. Science *367*, 1140–1146.

Couto, N., Caja, S., Maia, J., Strano Moraes, M.C., and Costa-Silva, B. (2018). Exosomes as emerging players in cancer biology. Biochimie *155*, 2–10.

Delcourt, V., Staskevicius, A., Salzet, M., Fournier, I., and Roucou, X. (2018). Small proteins encoded by unannotated ORFs are rising stars of the proteome, confirming shortcomings in genome annotations and current vision of an mRNA. Proteomics *18*, e1700058.

Grassadonia, A., Tinari, N., Iurisci, I., Piccolo, E., Cumashi, A., Innominato, P., D'Egidio, M., Natoli, C., Piantelli, M., and Iacobelli, S. (2002). 90K (Mac-

2 BP) and galectins in tumor progression and metastasis. Glycoconjugate J. *19*, 551–556.

Han, C., Alkhater, R., Froukh, T., Minassian, A.G., Galati, M., Liu, R.H., Fotouhi, M., Sommerfeld, J., Alfrook, A.J., Marshall, C., et al. (2016). Epileptic encephalopathy caused by mutations in the guanine nucleotide exchange factor DENND5A. Am. J. Hum. Genet. *99*, 1359–1367.

Harding, C., Heuser, J., and Stahl, P. (1983). Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. J. Cell Biol. *97*, 329–339.

Harrell, C.R., Simovic Markovic, B., Fellabaum, C., Arsenijevic, A., Djonov, V., Arsenijevic, N., and Volarevic, V. (2018). Therapeutic potential of mesenchymal stem cell-derived exosomes in the treatment of eye diseases. Adv. Exp. Med. Biol. *1089*, 47–57.

Jing, H., He, X., and Zheng, J. (2018). Exosomes and regenerative medicine: state of the art and perspectives. Translational Res. J. Lab. Clin. Med. *196*, 1–16.

Johnstone, R.M. (2005). Revisiting the road to the discovery of exosomes. Blood Cell Mol. Dis. *34*, 214–219.

Johnstone, R.M. (2006). Exosomes biological significance: a concise review. Blood Cell Mol. Dis. *36*, 315–321.

Johnstone, R.M., Bianchini, A., and Teng, K. (1989). Reticulocyte maturation and exosome release: transferrin receptor containing exosomes shows multiple plasma membrane functions. Blood *74*, 1844–1851.

Kojima, R., Bojar, D., Rizzi, G., Hamri, G.C., El-Baba, M.D., Saxena, P., Ausländer, S., Tan, K.R., and Fussenegger, M. (2018). Designer exosomes produced by implanted cells intracerebrally deliver therapeutic cargo for Parkinsons disease treatment. Nat. Commun. *9*, 1305.

Lane, R.E., Korbie, D., Trau, M., and Hill, M.M. (2017). Purification protocols for extracellular vesicles. Methods Mol. Biol. *1660*, 111–130.

Mathivanan, S., Ji, H., and Simpson, R.J. (2010). Exosomes: extracellular organelles important in intercellular communication. J. Proteomics *73*, 1907–1920.

Momen-Heravi, F. (2017). Isolation of extracellular vesicles by ultracentrifugation. Methods Mol. Biol. *1660*, 25–32.

Murgoci, A.-N., Cizkova, D., Majerova, P., Petrovova, E., Medvecky, L., Fournier, I., and Salzet, M. (2018). Brain-cortex microglia-derived exosomes: nanoparticles for Glioma therapy. Chemphyschem. https://doi.org/10.1002/cphc.201701198.

Pan, B.-T., and Johnstone, R.M. (1983). Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: selective externalization of the receptor. Cell *33*, 967–978.

Pan, B.T., Teng, K., Wu, C., Adam, M., and Johnstone, R.M. (1985). Electron microscopic evidence for externalization of the transferrin receptor in vesicular form in sheep reticulocytes. J. Cell Biol. *101*, 942–948.

Rajagopal, C., and Harikumar, K.B. (2018). The origin and functions of exosomes in cancer. Front. Oncol. *8*, 66.

Reclusa, P., Teng, K., Wu, C., Adam, M., and Johnstone, R.M. (2017). Exosomes as diagnostic and predictive biomarkers in lung cancer. J. Thorac. Dis. *9* (*Suppl 13*), S1373–S1382.

Ruivo, C.F., Adem, B., Silva, M., and Melo, S.A. (2017). The biology of cancer exosomes: insights and new perspectives. Cancer Res. *77*, 6480–6488.

Saito, N., Aoki, K., Hirai, N., Fujita, S., Iwama, J., Hiramoto, Y., Ishii, M., Sato, K., Nakayama, H., Harashina, J., et al. (2015). Effect of Notch expression in glioma stem cells on therapeutic response to chemo-radiotherapy in recurrent glioblastoma. Brain Tumor Pathol. *32*, 176–183.

Tauro, B.J., Greening, D.W., Mathias, R.A., Hong, J., Mathivanan, S., Scott, A.M., and Simpson, R.J. (2012). Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. Methods *56*, 293–304.

Théry, C., Amigorena, S., Raposo, G., and Clayton, A. (2001). Isolation and characterization of exosomes from cell culture supernatants and biological fluids. In Current Protocols in Cell Biology (John Wiley & Sons, Inc.). http://onlinelibrary.wiley.com/doi/10.1002/0471143030.cb0322s30/abstract.

Yamashita, T., Takahashi, Y., and Takakura, Y. (2018). Possibility of exosome-based therapeutics and challenges in production of exosomes eligible for therapeutic application. Biol. Pharm. Bull. *41*, 835–842.

Zhao, M., Xu, H., Liang, F., He, J., and Zhang, J. (2015). Association of osteopontin expression with the prognosis of glioma patient: a meta-analysis. Tumour Biol. *36*, 429–436.

Zhou, J., Li, X.L., Chen, Z.R., and Chng, W.J. (2017). Tumor-derived exosomes in colorectal cancer progression and their clinical applications. Oncotarget *8*, 100781–100790.

iScience, Volume *23*

Supplemental Information

Reference and Ghost Proteins Identification in Rat C6 Glioma Extracellular Vesicles

Adriana-Natalia Murgoci, Tristan Cardon, Soulaimane Aboulouard, Marie Duhamel, Isabelle Fournier, Dasa Cizkova, and Michel Salzet

SUPPLEMENTAL INFORMATION

SUPPLEMENTAL FIGURES

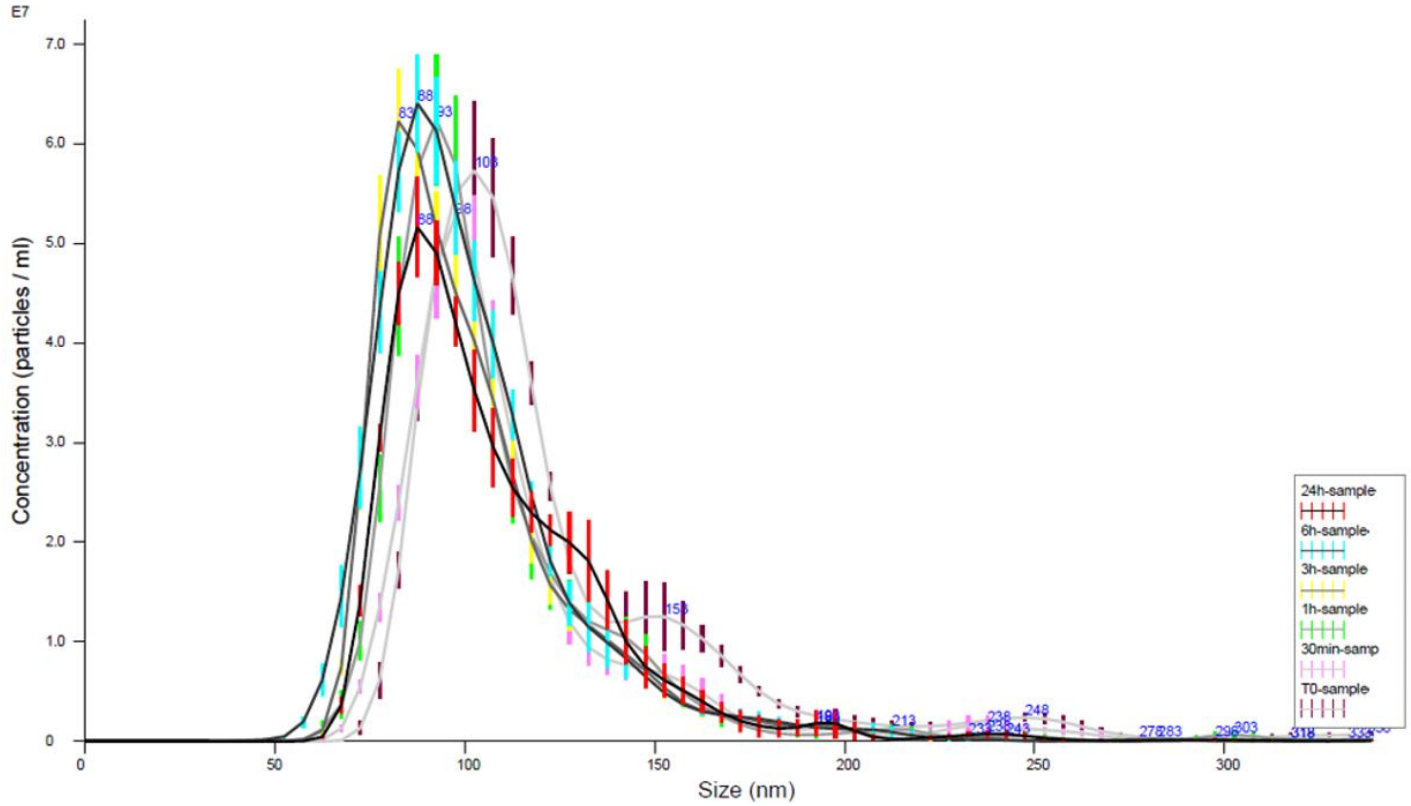


Figure S1. NanoSight analyzes, related to Figure 1. Isolated vesicles were analyzed using the NanoSight to characterize size and concentration. After the last ultracentrifugation step followed by Tryspin/Lys C digestion, the pellet obtained was diluted in particle-free PBS (1:100). To analyze the particles, 5 videos of 60 s for each sample were recorded. A monochromatic laser beam at 488nm was used for analyses. Particle movement was investigated with NTA software (version 3.2, NanoSight). NTA post-acquisition settings were kept constant between samples.

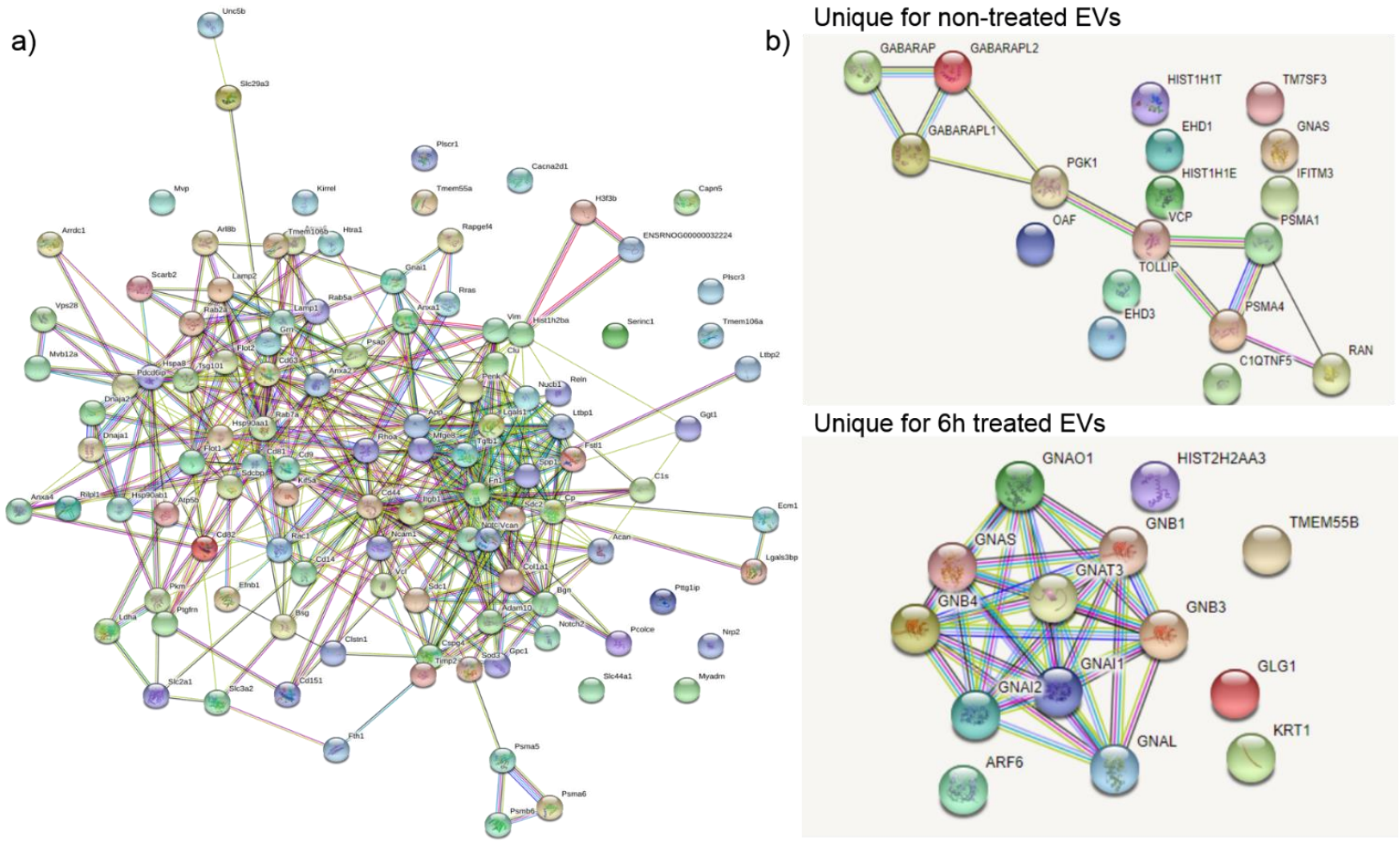


Figure S2. String analysis for glioma EVs proteins, related to Figure 2 and Figure 3. (a) String analysis for all proteins identified for glioma EVs by shotgun proteomics. (b) String analysis results for unique proteins identified from EVs not treated with trypsin/Lys C (T-0h) and for EVs that where 6h treated (T-6h).

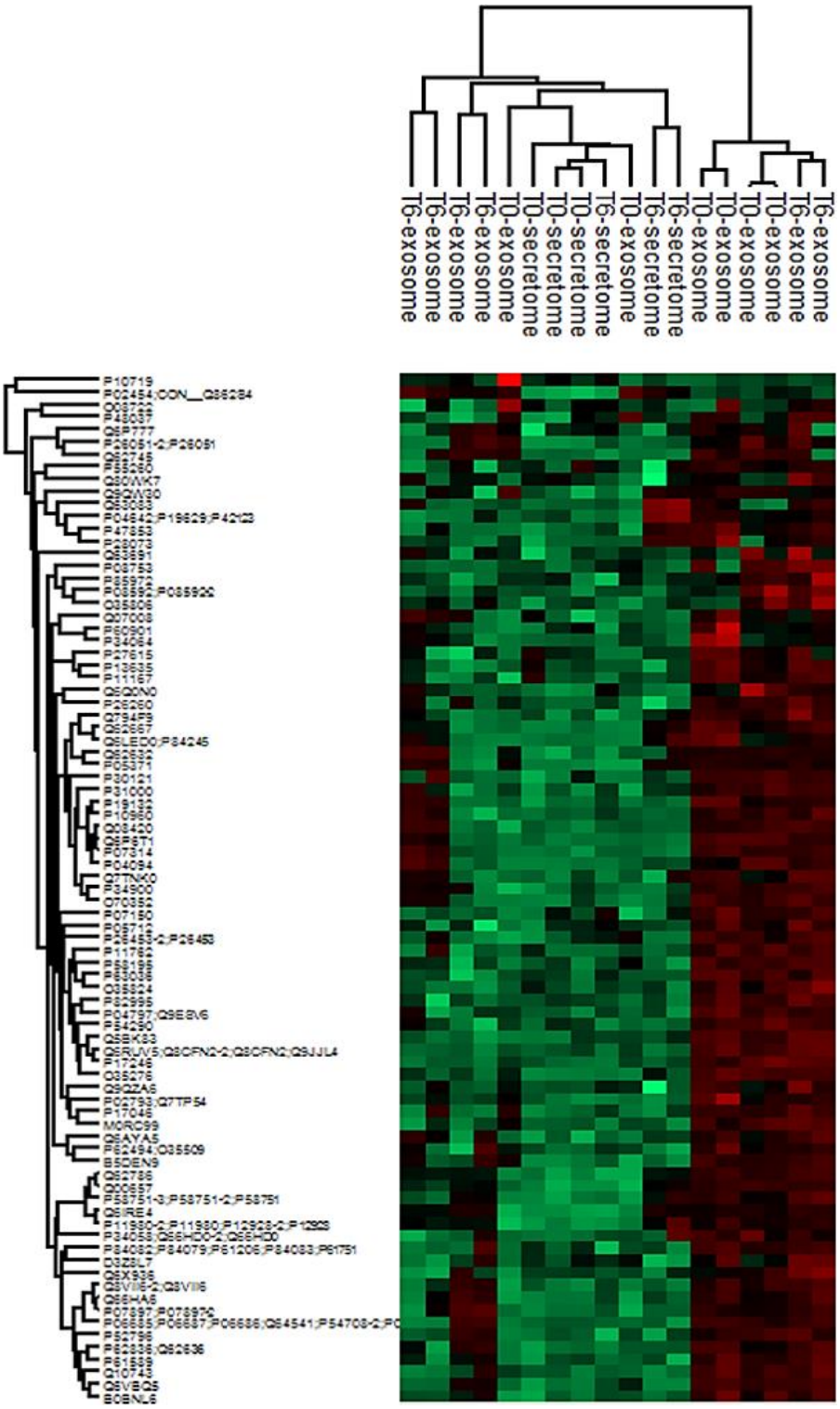


Figure S3. Heatmap from Shot gun proteomics, related to Figure 2 and Figure 3. For identified proteins of EVs treated or not with Trypsin/Lys C was performed quantitative proteomics analyses using MaxQuant software, after ANOVA tests with a p value >0.05,

SUPPLEMENTAL TABLE

Table S1. List of the statistically validated identified proteins by shot gun proteomics from purified glioma C6 EVs proteins, related to Figure 2.

Protein ID	Protein names	Gene names
B5DEN9	Vacuolar protein sorting-associated protein 28 homolog	Vps28
D3Z8L7	Ras-related protein R-Ras	Rras
M0RC99	Ras-related protein Rab-5A	Rab5a
O35276	Neuropilin-2	Nrp2
P62494;O35509	Ras-related protein Rab-11A;Ras-related protein Rab-11B	Rab11a;Rab11b
O35806	Latent-transforming growth factor beta-binding protein 2	Ltbp2
O35824	DnaJ homolog subfamily A member 2	Dnaja2
O70352	CD82 antigen	Cd82
P02793;Q7TP54	Ferritin light chain 1;Protein FAM65B	Ftl1;Fam65b
P04094	Proenkephalin-A	Penk
P04642;P19629;P42123	L-lactate dehydrogenase A chain	Ldha
P04797;Q9ESV6	Glyceraldehyde-3-phosphate dehydrogenase	Gapdh;Gapdhs
P05371	Clusterin;Clusterin beta chain;Clusterin alpha chain	Clu
P05712	Ras-related protein Rab-2A	Rab2a
P06685;P06687;P06686;Q64541;P54708-2;P09626;P54708	Sodium/potassium-transporting ATPase subunit alpha 1	Atp1a1;Atp1a3;Atp1a2
P07150	Annexin A1	Anxa1
P07314	Gamma-glutamyltranspeptidase 1	Ggt1
P07897;P07897-2	Aggrecan core protein	Acan
P08592;P08592-2	Amyloid beta A4 protein	App
P08753	Guanine nucleotide-binding protein G(k) subunit alpha	Gnai3
P10960	Sulfated glycoprotein 1	Psap
P11167	Solute carrier family 2, facilitated glucose transporter member 1	Slc2a1
P11762	Galectin-1	Lgals1
P11980-2;P11980;P12928-2;P12928	Pyruvate kinase PKM	Pkm
P13635	Ceruloplasmin	Cp
P17046	Lysosome-associated membrane glycoprotein 2	Lamp2
P17246	Transforming growth factor beta-1;Latency-associated peptide	Tgfb1
P19132	Ferritin heavy chain;Ferritin heavy chain, N-terminally processed	Fth1
P26051-2;P26051	CD44 antigen	Cd44
P26260	Syndecan-1	Sdc1
P26453-2;P26453	Basigin	Bsg
P27615	Lysosome membrane protein 2	Scarb2
P28073	Proteasome subunit beta type-6	Psmb6
P30121	Metalloproteinase inhibitor 2	Timp2
P31000	Vimentin	Vim
P34058;Q66HD0-2;Q66HD0	Heat shock protein HSP 90-beta	Hsp90ab1
P34064	Proteasome subunit alpha type-5	Psma5
P34900	Syndecan-2	Sdc2

P47853	Biglycan	Bgn
P52796	Ephrin-B1	Efnb1
P54290	Voltage-dependent calcium channel subunit alpha-2	Cacna2d1
P55260	Annexin A4	Anxa4
P58195	Phospholipid scramblase 1	Plscr1
P58751-3;P58751-2;P58751	Reelin	Reln
P60901	Proteasome subunit alpha type-6	Psma6
P84082;P84079;P61206;P84083;P61751	ADP-ribosylation factors	Arf2;Arf1;Arf3;Arf5;Arf4
P61589	Transforming protein RhoA	Rhoa
P62836;Q62636	Ras-related protein Rap-1A;Ras-related protein Rap-1b	Rap1a;Rap1b
P63036	DnaJ homolog subfamily A member 1	Dnaja1
P82995	Heat shock protein HSP 90-alpha	Hsp90aa1
Q6LED0;P84245	Histone H3.1;Histone H3.3	H3f3b
P85972	Vinculin	Vcl
Q00657	Chondroitin sulfate proteoglycan 4	Cspg4
Q07008	Notch 1 intracellular domain	Notch1
Q08420	Extracellular superoxide dismutase [Cu-Zn]	Sod3
Q10743	Disintegrin and metalloproteinase domain-containing protein 10	Adam10
Q5BK83	Transmembrane protein 106A	Tmem106a
Q62632	Follistatin-related protein 1	Fstl1
Q62667	Major vault protein	Mvp
Q62745	CD81 antigen	Cd81
Q62786	Prostaglandin F2 receptor negative regulator	Ptgfrn
Q63083	Nucleobindin-1	Nucb1
Q63691	Monocyte differentiation antigen CD14	Cd14
Q66HA6	ADP-ribosylation factor-like protein 8B	Arl8b
Q6AYA5	Transmembrane protein 106B	Tmem106b
Q6IRE4	Tumor susceptibility gene 101 protein	Tsg101
Q6P6T1	Complement C1s subcomponent	C1s
Q6P777	Multivesicular body subunit 12A	Mvb12a
Q6Q0N0	Calsyntenin-1;Soluble Alc-alpha;CTF1-alpha	Clstn1
Q6RUV5;Q8CFN2-2;Q8CFN2;Q9JJL4	Ras-related C3 botulinum toxin substrate 1	Rac1
Q6VBQ5	Myeloid-associated differentiation marker	Myadm
Q6X936	Kin of IRRE-like protein 1	Kirrel
Q794F9	4F2 cell-surface antigen heavy chain	Slc3a2
Q7TNK0	Serine incorporator 1	Serinc1
Q80WK7	Equilibrative nucleoside transporter 3	Slc29a3
Q8VII6-2;Q8VII6	Choline transporter-like protein 1	Slc44a1
Q9QW30	Notch 2 intracellular domain	Notch2
Q9QZA6	CD151 antigen	Cd151

TRANSPARENT METHODS

Experimental design and statistical rationale

Shotgun proteomics experiments on Amicon filtered glioma EVs, FASP and Trypsin/Lys C digestion were conducted in biological triplicates.

Reagents

The rat C6 glioma cell line was kindly provided by Prof. Dr. Bernd Kaina (Institute of Toxicology, University Medical Center, Mainz, Germany). Ham's F12K, puromycin, phosphate buffer saline (PBS), fetal bovine serum (FBS), Trypsin/Lys C were all obtained from Promega (USA).

EVs purification

Glioma C6 cells were cultured in high-glucose Dulbecco's Modified Eagle's Medium (DMEM) and supplemented with 10% heat-inactivated fetal bovine serum, 1% L-glutamine (2 mM) and 1% gentamicin (50 units per ml), all from Sigma-Aldrich. Confluent population of cells were incubated without FBS for 24h to obtain cells conditioned media (CM) free of FBS. Afterwards, CM was cleared of cells and debris by centrifugation at 350×g for 10 min at 4°C, followed by filtration with nylon filter membranes, pore size 0.2 µm. Membranes and debris were discarded from the CM by centrifugation for 30 min at 2 000×g at 4°C. Then the withdrawn supernatant was centrifuged at 10 000×g for 30 min at 4°C to remove larger vesicles, followed by ultracentrifugation (Beckman Optima XPN80 Ultracentrifuge, USA) at 100 000×g for 70 min, 4°C. The pellet was washed in PBS in order to eliminate contaminating proteins and re-ultracentrifuged at 100 000×g for 120 min, 4°C.

Trypsin/Lys C digestion treatment

Isolated EVs were resuspended in Trypsin/Lys C 0.2 µg/mL solution diluted in NH4HCO3 50mM. For T0 condition (non-treated), the enzyme activity was immediately stopped with TFA 0.5%, for rest times points, the EVs were incubated at 37°C for 6h, then TFA 0.5% was added to stop enzyme activity.

Nanoparticle tracking analysis (NTA)

Isolated vesicles were analyzed using the NanoSight LM 10 instrument (Merkel technologies LTD., UK) to characterize size and concentration. After the last ultracentrifugation step followed by Tryspin/Lys digestion, the pellet obtained was diluted in particle-free PBS (1:100). To analyze the particles, 5 videos of 60 s for each sample were recorded. A monochromatic laser beam at 488nm was used for analyses. Particle movement was investigated with NTA software (version 3.2, NanoSight). NTA post-acquisition settings were kept constant between samples.

Samples preparation for mass spectrometry analysis

EVs digested with Trypsin/Lys C for 6h (T-6h) or non-digested samples (T-0h), were subjected to a filtration step using Amicon Ultra-0.5 mL Centrifugal Filters 10 KDa for 20min for 14 000xg at 4°C. In the Amicon "Filtrate" tube, peptides from protein aggregates and from external part of EV membrane proteins were recovered. The vesicles were recovered from Amicon filter by 4°C centrifugation for 5min at 1 000xg. EV proteins were extracted using RIPA buffer. Samples were prepared for mass spectrometry analysis by Filter-Aided Sample Preparation (FASP) as described in Duhamel *et al.*, 2018. Briefly, proteins from all fractions were denatured with 2 M urea in 10 mM HEPES, pH 8.0 by sonication. Then they were reduced with 10 mM DTT for 40 min at 56°C followed by alkylation with 55 mM iodoacetamide for 40 min in the dark. The iodoacetamide was quenched with 100 mM thiourea. The proteins were digested in an Amicon 30KD (Millipore) overnight using 1

µg Trypsin/LysC mixture at 37°C. In a new tube, the Amicon is centrifuge at 14000 Xg to liberate the peptides then the digestion was stopped with 0.5% TFA and dry. The peptides were desalted with a Millipore ZipTip-C18 device. The solution was then dried using the SpeedVac. Dried samples were solubilized in 2% ACN / 98% of 0.1% formic acid in water, before nLC MS/MS analysis.

LC MS/MS analysis

Samples were separated by online reversed-phase chromatography using a Thermo Scientific Proxeon Easy-nLC system equipped with a Proxeon trap column (100 µm ID x 2 cm, Thermo Scientific) and a C18 packed-tip column (75 µm ID x 50 cm, Thermo Scientific). Peptides were separated using an increasing amount of acetonitrile (5–35% for 100 min) at a flow rate of 300 nL/min. The LC eluent was electrosprayed directly from the analytical column and a voltage of 1.7 kV was applied via the liquid junction of the nanospray source. The chromatography system was coupled with a Thermo Scientific Q Exactive mass spectrometer programmed to acquire using the data dependent Top 10 method. Survey scans were acquired at a resolution of 70 000 at m/z 400.

Data analyses

For the Alternative Proteins (AltProt), Proteome Discoverer V2.3 (Thermo Scientific) was used, the protein database was downloaded from Openprot (https://openprot.org/) and includes RefProts, novel isoforms and AltProts predicted from both Ensembl and RefSeq annotations (Rnor_6.0.84, Rnor_6.0) for a total of 293303 entries. The following processing and consensus parameters are used with: Trypsin as enzyme, 2 missed cleavages, methionine oxidation as variable modification and carbamidomethylation of cysteines as static modification, Precursor Mass Tolerance: 10 ppm and Fragment mass tolerance: 0.1 Da. The validation was performed using a Percolator with a protein strict FDR set to 0.001%, according to Brunet et al., for AltProt database, the FDR needs to be more restrictive to remove maximum of false positive (Brunet *et al.*, 2019). A consensus workflow was then applied for the statistical arrangement, using the high confidence protein identification. Results are filtered to keep master protein and high confidence protein FDR. For the Reference Proteins (RefProts) the MS data was processed with MaxQuant (version 1.5.8.3)(Cox and Mann, 2008) using the Andromeda search engine. Proteins were identified by searching MS and MS/MS data against the Decoy version of the complete proteome for *Rattus norvegicus* of the UniProt database(UniProt Consortium, 2012) (Release June 2017, 8022 entries reviewed). Trypsin specificity was used for the digestion mode with N-terminal acetylation and methionine oxidation selected as the variable. Carbamydomethylation of cysteines was set as a fixed modification and we allowed up to two missed cleavages. For MS spectra, an initial mass accuracy of 6 ppm was selected, and the MS/MS tolerance was set to 20 ppm for HCD data. For identification, the FDR at the peptide spectrum matches (PSMs) and protein level was set to 0.01. Relative, label-free quantification of proteins was performed using the MaxLFQ algorithm(Cox *et al.*, 2014) integrated into MaxQuant with the default parameters. Analysis of the proteins identified was performed using Perseus software (http://www.perseus-framework.org/) (version 1.6.0.7). The file containing this information from identification was used with hits to the reverse database, and proteins only identified with modified peptides and potential contaminants removed. Then, the LFQ intensity was transformed by the log2[x] function. Categorical annotation of rows was used to define different groups depending on the following criteria: 1) localization on the exosome and 2) time of enzyme treatment. Multiple-sample tests were performed using an ANOVA test with a p value > 0.05% and preserved grouping in randomization. To determine enrichment of categorical annotations (Gene Ontology terms and KEGG pathway), a Fisher's exact test was used, taking in account the results of the ANOVA test for each group. Normalization was achieved using a Z-score with matrix access by rows. Only proteins significant by the ANOVA tests were used for statistical analysis. A hierarchical clustering was first performed using the Euclidean parameter for distance calculation and an average option for linkage in row and column trees using a maximum of 300 clusters. To quantify fold changes of proteins across samples, we used MaxLFQ. To visualize these fold changes in the context of individual protein abundance in the proteome, we projected them onto the summed

peptide intensities normalized by the number of theoretically observable peptides. Datasets including MaxQuant files and annotated MS/MS datasets, were uploaded to ProteomeXchange Consortium via the PRIDE database, and was assigned the dataset identifier PXD016944 (Username: reviewer95445@ebi.ac.uk, Password: wHS6rIOg). Functional annotation and characterization of identified proteins were obtained using PANTHER software (version 9.0, http://www.pantherdb.org) and STRING (version 9.1, http://string-db.org). The Elsevier's Pathway Studio version 9.0 (Ariadne Genomics/Elsevier) was used to deduce relationships among differentially expressed protein candidates using the Ariadne ResNet database(Bonnet *et al.*, 2009; Yuryev, Kotelnikova and Daraselia, 2009). "Subnetwork Enrichment Analysis" (SNEA) algorithm was selected to extract statistically significant altered biological and functional pathways pertaining to each identified set of protein hits. SNEA utilizes Fisher's statistical test used to determine if there are non-randomized associations between two categorical variables organized by specific relationships. SNEA starts by creating a central "seed" from all relevant entities in the database, and retrieving associated entities based on their relationship with the "seed" (*i.e*. binding partners, expression targets, protein modification targets, regulation). The algorithm compares the sub-network distribution to the background distribution using one-sided Mann-Whitney U-Test and calculates a *p*-value indicating the statistical significance of the difference between two distributions.

SUPPLEMENTAL REFERENCES

Bonnet, A. *et al.* (2009) 'Pathway results from the chicken data set using GOTM, Pathway Studio and Ingenuity softwares', *BMC proceedings*, 3 Suppl 4, p. S11. doi: 10.1186/1753-6561-3-S4-S11.

Brunet, M. A. *et al.* (2019) 'OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes', *Nucleic Acids Research*, 47(D1), pp. D403–D410. doi: 10.1093/nar/gky936.

Cox, J. *et al.* (2014) 'Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ', *Molecular & Cellular Proteomics*, 13(9), pp. 2513–2526. doi: 10.1074/mcp.M113.031591.

Cox, J. and Mann, M. (2008) 'MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification', *Nature Biotechnology*, 26(12), pp. 1367–1372. doi: 10.1038/nbt.1511.

Duhamel, M. *et al.* (2018) 'Paclitaxel treatment and PC1/3 knockdown in macrophages is a promising anti-glioma strategy as revealed by proteomics and cytotoxicity studies', *Molecular & cellular proteomics: MCP*. doi: 10.1074/mcp.RA117.000443.

UniProt Consortium (2012) 'Reorganizing the protein space at the Universal Protein Resource (UniProt)', *Nucleic Acids Research*, 40(D1), pp. D71–D75. doi: 10.1093/nar/gkr981.

Yuryev, A., Kotelnikova, E. and Daraselia, N. (2009) 'Ariadne's ChemEffect and Pathway Studio knowledge base', *Expert Opinion on Drug Discovery*, 4(12), pp. 1307–1318. doi: 10.1517/17460440903413488.