A compendium of *bona fide* reference markers of plant-derived extracellular vesicles and their degree of phylogenetic conservation

Rodríguez de Lope, Miriam M.[1,2]; Sánchez-Pajares, Ibone R.[1,2]; Herranz, Estela[1,2]; López-Vázquez, Cristina M.[3]; González-Moro, Ainara[1,2]; Rivera-Tenorio, Alan[1,2]; González-Sanz Carlos[4,5], Soledad Sacristán[4,5], Chicano-Gálvez, Eduardo[3]; de la Cuesta, Fernando[1,2*].

[1]*Department of Pharmacology, School of Medicine, Universidad Autónoma de Madrid, Spain*

[2]*Instituto de Investigación Hospital Universitario La Paz (IdiPaz), Madrid, Spain*

[3]*IMIBIC Mass Spectrometry and Molecular Imaging Unit, Maimonides Biomedical Research Institute of Cordoba (IMIBIC), Reina Sofia University Hospital, University of Cordoba (UCO), Córdoba, Spain.*

[4]*Centro de Biotecnología y Genómica de Plantas (CBGP, UPM-INIA/CSIC), Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA/CSIC), Madrid, Spain.*

[5]*Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid (UPM), Madrid, Spain.*

Keywords: Plant EVs, PDEVs, markers, apoplastic washing fluid, MISEV, proteomics.

* Corresponding author:

Fernando de la Cuesta

ThEVrapies Lab. Department of Pharmacology

School of Medicine, Universidad Autónoma de Madrid (UAM)

C/ Arzobispo Morcillo 4. 28029 Madrid. SPAIN

e-mail: fernando.delacuesta@uam.es

Abstract

Although the field of plant-derived EVs (PDEVs) is experimenting an exponential growth, rigorous characterization complying with MISEV guidelines has not been yet implemented due to the lack of *bona fide* reference markers. In this work, we have paved the way for the standardization of PDEV markers, providing the most profound proteomic data so far from apoplastic washing fluid-derived EVs, as genuine extracellular vesicles from plant tissue, of two reference plant species: *Arabidopsis thaliana* (Arath-EVs) and *Brassica oleracea* (Braol-EVs). Additionally, we have conducted an exhaustive analysis of the EV proteomic data available so far from any plant species, evaluating current potential markers, together with those found in our proteomic analyses. Our results provide evidence supporting the potential use of the following families as PDEV markers: fasciclin-like arabinogalactan proteins (FLAs), patellins, germin-like proteins, aquaporins, heat shock 70 proteins, 14-3-3 like proteins, Rab proteins, and vacuolar-type ATPase complex subunits. Next, we analyzed the presence of orthologues and their degree of conservation throughout the plant kingdom, as well as in human and mouse species. Their degree of conservation was compared with that of current EV markers used in animal-derived EVs: CD63, CD81 and CD9. Among the protein families with potential to be used as PDEV markers, 3 were found to be plant-specific: FLAs, patellins and germin-like proteins. Among these, fasciclin-like arabinogalactan (FLA) proteins, in particular FLA1, FLA2 and FLA8, were recurrently identified in EVs from different plant sources showing great abundance and thus arise as a novel family of plant-specific GPI-anchored PDEV markers. Furthermore, FLAs stand as ideal PDEV markers since they have orthologues across the different plant subdivisions and display a similar phylogenetic distribution in plants to that of currently used markers in animals.

Introduction

Extracellular vesicles (EVs) from mammalian cell origin have been widely characterized and exhaustive guidelines to follow have been defined by the International Society for Extracellular Vesicles (ISEV), the so-called "Minimal information for studies of extracellular vesicles (MISEV)", its more recent version being published in 2023[1]. Regardless of a constant and exponential growth of publications in the field of plant-derived EVs (PDEVs), to date there is a lack of standardization in this particular scientific area. In fact, a call for rigor in the field of PDEVs was made by Pinedo *et al.* in 2021 in the journal of the ISEV[2], with a special focus on the need for specific markers that could allow to comply with the golden rules of EV characterization. Furthermore, ISEV has recently launched the plant EVs task force to gather experts in the field of PDEVs to put forward initiatives regarding nomenclature, isolation, characterization, standards and standardizations (https://www.isev.org/plant-task-force). To date, several publications have provided data on PDEVs´ cargo and potential markers, but a consensus on *bona fide* markers of PDEVs has not yet been reached and thus, a response to the abovementioned call for rigor is still awaiting. One particular hurdle in this aim is the fact that most of the work in this direction has been performed in *Arabidopsis thaliana* L., a eudicot plant that stands as a reference model in the field, but there is not enough evidence supporting that these markers are applicable to plant EVs from other origin, especially monocots, bryophytes and algae. Studies in *A. thaliana* frequently use two tetraspanins, a protein group well-known to be involved in EV biogenesis and widely used as EV markers, to characterize and monitor PDEVs: tetraspanin 8 (TET8) and tetraspanin 9 (TET9), the former being the most utilized. Another marker arising from *A. thaliana.* is PEN1, also called syntaxin of plants 121 (SY121), a member of the t-SNARE protein superfamily, which is related to the exocytic pathway[3] and whose ortholog in human, syntaxin-2, has been associated with exocytosis in mammalian cells[4]. This protein was also shown to be present in PDEVs from *Nicotiana benthamiana* Domin[5]. Nonetheless, it is unclear whether PEN1 is a marker of all types of EVs or from a specific subcategory[6]. On the other hand, a subset of EVs arising from a characteristic plant biogenesis pathway, exocyst-positive organelle (EXPO)-derived EVs, have

been described in *A. thaliana* and express the specific marker Exocyst complex component EXO70E2. All these markers have proven to be useful to study PDEVs from *A. thaliana* but translation into other species has not been so far demonstrated.

One strategy for defining biomarkers of PDEVs is the search for orthologues of mammalian EV markers. This has partially driven most of the abovementioned findings in *A. thaliana.*, with tetraspanins TET8 and TET9 being structurally similar to mammalian tetraspanin CD63[7], with which they share less than 30% homology[8]. Many other markers belonging to protein families related to EV biogenesis have also been explored: orthologues of Tumor susceptibility gene 101 protein (TSG101), ALG-2 interacting protein X (Alix), and Heat shock 70 kDa protein (Hsp70) have been found to be markers of EVs derived from plants in some studies. Regarding TSG101, called protein ELC-like in *A. thaliana*[9], a homolog in lemon (*Citrus limon* (L.) Burm. f) juice-derived PDEVs has been reported[6,10]. The ortholog of Alix in *A. thaliana* Vacuolar-sorting protein BRO1, has been proven to be involved in the biogenesis of vacuole and multivesicular bodies (MVBs)[11]. Besides, an Alix homolog has also been described in kiwi (*Actinidia chinensis Planch.*) pollen-derived PDEVs, detected with an antibody targeting mammalian Alix[12]. Finally, orthologues of HSP70 have been found in EVs from tangerine[13] and other citrus species[14], as well as in the edible algae *Ecklonia cava* Kjellman[15], the latter also detected using an antibody developed with the human protein as immunogen.

Therefore, there is an urgent need for a consensus in the field of PDEVs to define the validity of current EV-specific markers, considering experimental evidence of abundance in EV isolates and their conservation in the plant kingdom, in order to define their potential use for all plant sources or restricted to certain species. Concerning this, although systematic revision of all the available data in the field is crucial to answer this question, the lack of reproducible and robust data deposited in specialized repositories limits the potential of such approach. Hence, new experimental data may be required to deepen in the proteome of PDEVs, aiming to find novel specific EV markers for plant vesicles and an important level of conservation in the plant kingdom that may allow their use for most, if not all, plant sources. With this goal in mind, we have

conducted an in-depth proteomic analysis of PDEVs from two well-annotated reference plant species: *A. thaliana* and *Brassica oleracea* L. Next, we have exhaustively analyzed the proteomic data published in the field, comparing them with the proteomes obtained in the present study. Our results allow us to provide more rigor and define the utility of the markers described so far. Furthermore, we have defined a family of GPI-anchored proteins as novel PDEV markers with great potential due to their plant specificity and great abundance: fasciclin-like arabinogalactan proteins (FLAs). Furthermore, these proteins have been found repeatedly in previous proteomic data from several types of PDEVs, making them ideal markers with specificity for plant EVs.

Materials and Methods

Plant Material

A. thaliana ecotype Columbia-0 (Col-0) seeds were sown, stratified at 4°C for 3 days in darkness, and moved to a growth chamber at 22°C, 80% relative humidity, under short day photoperiod (10-h light/14-h dark) and light intensity of 110- 120 µE/m2 /s for 5 weeks. Fresh red cabbage (*Brassica oleracea var. capitata F. rubra*) was purchased from a local market.

Collection of apoplastic washing fluid (AWF)

A. thaliana rosettes and cabbage leaves were collected and thoroughly washed, first in 70% ethanol for surface sterilization and then rinsed twice with PBS to remove any residual ethanol. Due to the larger size of cabbage leaves, these were previously cut and washed thoroughly with PBS to avoid contamination with intracellular vesicles. The washed leaves were placed into a 100 mL syringe filled with PBS, and a gentle vacuum was applied to infiltrate the leaves. After the infiltration, the leaves were carefully positioned inside a 5mL plastic pipette tip, which was sealed with parafilm to keep the leaves in place. This setup was then inserted into a syringe without the plunger. The syringe was placed into a 50 mL Falcon tube, ensuring that the leaves stayed in place and did not touch the bottom. The tube was centrifuged for 20 min at 4ºC at 1,000 x g to collect the AWF.

PDEVs isolation

The AWF was centrifuged for 5 min at 4ºC at 15,000 x g to remove cell debris and then filtered through a 0.22 µm filter (Jet Biofil). Next, the supernatants (the clean AWF) were concentrated by ultrafiltration using Amicon Ultra-15 100 kDa filters (Millipore, Merck) at 3500 x g at 4ºC. The concentrated supernatants were then subjected to size exclusion chromatography (SEC) using in-house casted Sepharose CL-2B (Sigma-Aldrich Chemical Co) columns to isolate EVs. Protein quantification of the fractions collected from SEC was performed using the Pierce BCA method (Thermo Fisher Scientific) and fractions containing less than 0,05µg/µL of total soluble protein were further concentrated by ultrafiltration with Amicon Ultra-15 10 kDa filters (Millipore, Merck) at 3,500 x g at 4°C. The isolation and characterization process is illustrated in **Figure 1**.

PDEVs characterization

EVs isolated from *A. thaliana* (Arath-EVs) and *B. oleracea* (Braol-EVs) AWF were characterized by nanoparticle tracking analysis (NTA), using a NanoSight LM10 instrument coupled to an automated Syringe Pump module (Malvern Panalytical), measuring their concentration and size distribution. The same camera level, threshold, and focus were used for all measurements**.**

EVs morphological characteristics were observed by transmission electron microscopy (TEM). For this purpose, they were fixed in 0.1 % PFA. A 15 µL droplet of sample was placed in an electron microscopy Formvar/carbon grid (Ted Pella) for 10 minutes. After 5 minutes in bidistilled water, a negative staining was conducted using 20 µL of uranyl acetate 2% for 30 seconds. The samples were visualized in a JEM 1010 (JEOL Peabody) at the Microscopy Unit, Faculty of Medicine UAM**.**

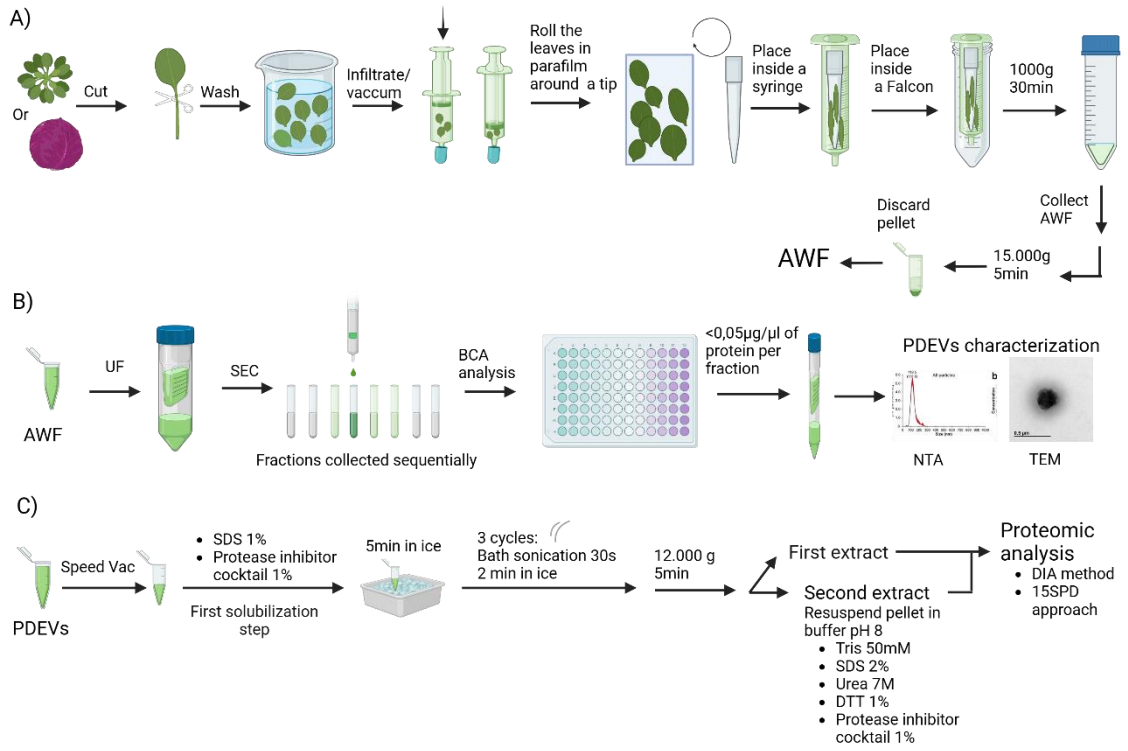


Figure 1. Three-Step workflow for Apoplastic Washing Fluid collection, plant-derived extracellular vesicle isolation, and sample preparation for proteomics*. A) Apoplastic Washing Fluid (AWF) extraction from A. thaliana or B. oleracea. B) Plant-derived EVs (PDEVs) isolation using ultrafiltration (UF) and size exclusion chromatography (SEC). C) PDEVs lysis and sample preparation for proteomics based on a two-step solubilization.*

Proteomic analysis

<u>Sample preparation</u>

EV samples ($1x10^{10}$ particles) were concentrated using a SpeedVac (Thermo Scientific) up to an approximate volume of 50 µl . Samples were then lysed in a first solubilization buffer by adding SDS 1% (Sigma-Aldrich) and a protease inhibitor cocktail 1% (Sigma-Aldrich). The mixture was placed 5 min on ice, followed by 3 cycles of sonication in a water bath sonicator (Fisher Scientific) with each cycle consisting of 30 s of sonication followed by 2 min in an ice bath.

The samples were then centrifugated for 5 min at 4ºC at 12,000 x g and the resulting pellet was resuspended in a different solubilization buffer (50mM Tris, SDS 2%, 7M Urea, DTT 1%, protease inhibitor cocktail 1%, pH 8,0). The amount of protein in both extracts was quantified by

BCA, and samples were sent to the IMIBIC Mass Spectrometry and Molecular Imaging Unit for further analyses (**Fig. 1.C**). Total protein was further quantified in the Mass Spectrometry Unit using the Qubit Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. The samples were then digested using the iST kit (PreOmics, Germany) and peptides were diluted to 10 ng/ul in LC-MS Water containing 0.1% (v/v) formic acid. A pool of samples was formed by combining equal volumes of each sample. Then, 400ng of the peptide solution were loaded onto Evotips (Evosep) following the manufacturer's instructions. Three biological replicates (n=3) were prepared for analysis of samples coming from both *A. thaliana.* and *B. oleracea.* Additionally, to ensure quality control and system equilibration Pierce™ Hela Tryptic Digest Standard (Thermo Scientific) was used. Three technical replicates from each sample were analyzed using Liquid chromatography–mass spectrometry (LC-MS/MS) in DIA-PASEF mode on the EvosepOneTIMSTOF-Flex Clinical Proteomics Platform IMSMI as follows:

Liquid chromatography

LC-purified tryptic digests were separated using the preset 15 SPD method (88-minute gradient time, 400 ng peptides) with an Evosep One LC system. A 10-µm fused silica ID emitter (Bruker Daltonics) was inserted into a nanoelectrospray source (Bruker Daltonics CaptiveSpray). The emitter was linked to a 15-cm × 150-µm reverse-phase column with 1.9-µm C18 beads (PepSep Fifteen columns, Bruker Daltonics), and this column was heated to 60ºC in an oven compartment (Bruker Daltonics). The mobile phases, water and acetonitrile, were buffered with 0.1% formic acid (LC-MS grade, Fisher Scientific).

Mass Spectrometry

LC was connected online to a TIMS Q-TOF apparatus (timsTOF Flex, Bruker Daltonics) via a CaptiveSpray nano-electrospray ion source and DIA-PASEF methods were used[16]. Ion mobility dimension calibration was done with 3 Agilent ESI-L Tuning Mix ions (m/z, 1/K0: 622.0289 Th, 0.9848 Vs cm−2; 922.0097 Th, 1.1895 Vs cm−2; 1221.9906 Th, 1.3820 Vs cm−2). Collision energy was linearly decreased from 59 eV at 1/K0 = 1.6 Vs cm−2 to 20 eV at 1/K0 = 0.6 Vs cm−2.

In addition, a 'long gradient' method was applied (m/z range: 400–1201.0 Th, 1/K0 range: 0.6–1.6 Vs cm−2, diaPASEF windows: 16 × 25 Th).

Data Analysis

Data identification and quantification were conducted using Spectronaut proprietary software. A directDIA approach (Spectronaut 18, Biognosys) analysis was done using the entire sample dataset and QC runs to monitor system performance. By using UniProt (https://www.uniprot.org/) as reference database, a subset of proteins formatted in fasta file format was built for *A. Thaliana* (Taxon 3702) (September, 2023) and *B. Oleracea* (Taxon 109376) (May, 2024). A Global data normalization of label free-quantitation (LFQ) parameter media was performed with Spectronaut18, followed by NaN data imputation by 1/5 limit of detection (LOD) using Metaboanalyst[17]. After normalization protein intensities were filtered by abundance. The raw proteomics data and the minimum information about a proteomics experiment (MIAPE), complying with the Proteomics Standards Initiative (PSI) of the Human Proteome Organization (HuPO) have been deposited to the ProteomeXchange Consortium via the PRIDE[18] partner repository with the dataset identifier PXD058098.

Enrichment analysis

Protein-protein interactions were first analyzed using STRING[19]. Default parameters were applied, using "Medium confidence (0.400)" as the minimum required interaction score and hiding the unconnected nodes. "Functional enrichment visualization" graph was obtained applying default parameters defined by STRING. To perform a specific enrichment analysis of the proteins interacting inside our "Functional enrichments" term of interest, this protein-protein interaction network was downloaded, applying the default parameters except in the case of the minimum required interaction score, which was increased to "High confidence (0.700)". To further mine the results of the enrichment analysis, the results obtained in STRING were implemented in Cytoscape[20]. Then, the "String enrichment" application was used to perform a "Retrieve functional enrichment", followed by filtering of concepts by the category "GO Cellular

Component", from which an Enrichment Map was created by applying a "Connectivity cutoff" of 0.4.

Comparison with other proteomes

To compare the proteomic entries from Arath-EVs with the 3 reference datasets obtained from publications of *A. thaliana* AWF-derived EVs, for some datasets, we had to transform NCBI IDs to UniProt IDs and we had to edit the results from our list of Arath-EVs, since more than one protein could be potentially identified in some cases for a single entry. Therefore, the list was modified with a text editor by separating the proteins that appeared together in the same entry to consider them individually, resulting in a total number of 1445 protein identifiers. This list of proteins was then confronted with the 3 reference datasets, and by creating a Venn diagram in jvenn software (https://jvenn.toulouse.inrae.fr/app/example.html) a final list of common proteins was obtained.

BLASTP tool of Uniprot was used to compare the proteomic entries from Arath-EVs with the 7 datasets obtained from publications of different PDEVs to verify the presence of orthologues in other proteomes. In the case of sunflower, a specific database was used (https://www.heliagene.org/HanXRQ-SUNRISE/)

Orthology search

The identification of orthologues for the proteins of interest was performed using the BLASTP tool of UniProt. For proteins whose orthology was studied primarily within plant kingdom phylogenies, *A. thaliana.,* FASTA sequences were queried against the species listed in **Table 1.A**. All sequences were obtained from UniProt, except for *Picea abies* (L.) H. Karst, which is not well-annotated in this database. Therefore, we obtained its protein sequences and identity percentages using the Evorepro database and the BLAST tool provided by this resource (https://evorepro.sbs.ntu.edu.sg/blast/).

For proteins defined as EV markers in animals, *H. sapiens* FASTA sequences were queried against the species listed in **Table 1.B**.

SPECIES	PHYLOGENY	MNEMONIC NAME	TAXON ID
A			
Arabidopsis thaliana	Plantae, angiosperms, eudicots	ARATH	3702
Brassica oleracea var. oleracea	Plantae, angiosperms, eudicots	BRAOL	109376
Oryza sativa subsp. japonica	Plantae, angiosperms, monocots	ORYSJ	39947
Solanum lycopersicum	Plantae, angiosperms, eudicots	SOLLC	4081
Picea abies	Plantae, gymnosperms, pinopsida	PICAB	3329
***Marchantia polymorfa* (Common liverwort)**	Plantae, marchantiophyta, marchantiopsida (bryophyte)	MARPO	3197
Chlamydomonas reinhardtii	Plantae, chlorophyta, chlorophyceae	CHLRE	3055
Homo sapiens	Animalia, chordata, mammalia	HUMAN	9606
Mus musculus	Animalia, chordata, mammalia	MOUSE	10090
B			
Species	**Phylogeny**	**Mnemonic name**	**Taxon ID**
Homo sapiens	Animalia, chordata, mammalia	HUMAN	9606
Pan troglodytes	Animalia, chordata, mammalia	PANTR	9598
Mus musculus	Animalia, chordata, mammalia	MOUSE	10090
Rattus norvegicus	Animalia, chordata, mammalia	RAT	10116
Oryctolagus cuniculus	Animalia, chordata, mammalia	RABIT	9986
Gallus gallus	Animalia, chordata, birds	CHICK	9031
Xenopus laevis	Animalia, chordata, amphibia	XENLA	8355
Danio rerio	Animalia, chordata, actinopterygii	DANRE	7955
Drosophila melanogaster	Animalia, arthropoda, insecta	DROME	7227
Caenorhabditis elegans	Animalia, nematoda, secernentea	CAEEL	6239

Table 1. Species used for orthology search. *A) Species for orthology search of potential PDEVs markers, B) Species for orthology search of well-defined animal EVs markers.*

For each protein of interest, we identified the putative orthologues from each species with the highest identity score according to the UniProt classification. MEGA 11 software[21] was used to construct phylogenetic trees from the protein sequences found for all considered species. First, the alignment of the FASTA files for each group of putative orthologues was performed with the ClustalW tool and then, phylogenetic trees were generated with the “Neighbour-joining” method.

Results

Isolation and Characterization of Arath-EVs and Braol-EVs

Vesicles were successfully extracted from the AWF of *A. thaliana* and *B. oleracea* leaves. Three batches of each plant species were purified, and the concentration and size distribution were observed by NTA (**Fig. 2.A**). PDEVs from both plant species had a similar size, being the mean for Arath-EVs of 157,82 ± 7,35 nm, and the mean for Braol-EVs 159,13 ± 4,35 nm (**Fig. 2.B**). In all the samples assayed, TEM demonstrated a typical spherical vesicular appearance, with EVs showing integrity and the expected size (**Fig. 2.C**).

Analysis of the proteome of PDEVs

A novel methodology to extract proteins from these PDEVs was developed, based on a two-step solubilization. As part of the method´s optimization, extracts from Arath-EVs coming from both solubilization steps were analyzed separately by proteomics to inspect which one had more proteins as well as to check their complementarity. In the supernatant extract from the first solubilization buffer (first step) 686 proteins were identified, compared to 2,468 proteins identified in the pellet resuspended with the second buffer (second step). An amount of 651 proteins were identified in common, with only 35 proteins identified only in the first step (**Fig. 2.D**). These proteins are shown in **Suppl. Table 1**.

From this moment on, we decided to analyze only the extract after the second step of solubilization in the next samples. Three biological replicates of both types of PDEVs were then analyzed and Venn diagrams were built up in order to focus only on the proteins that were present in all 3 replicates. An amount of 1145 proteins were identified in common for Arath-EVs (**Fig. 2.E**) *versus* 899 for Braol-EVs (**Fig. 2.F**). The proteins identified in each replicate are listed in the **Suppl. Table 2**, organized by abundance, calculated using their normalized LFQ.

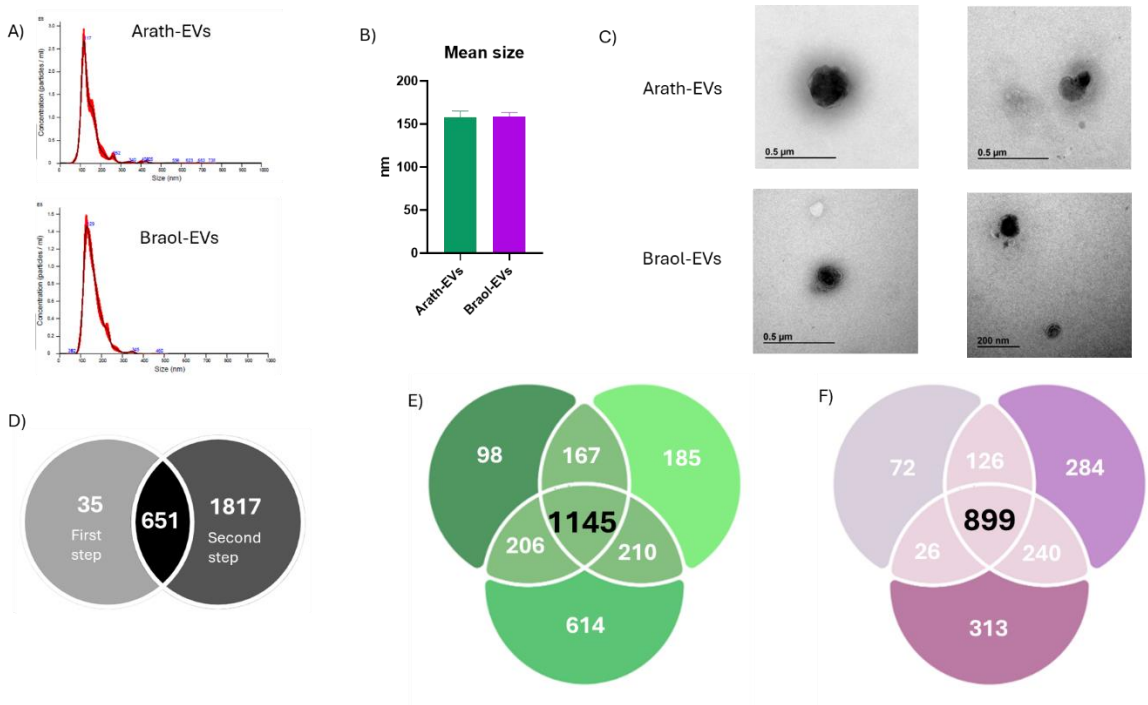


Figure 2. Characterization of PDEVs and proteomics. *A) NTA profile of Arath-EVs and Braol-EVs extracted from apoplastic washing fluid B) Mean size of PDEVs. C) TEM images of the PDEVs. D) Venn diagram of proteins found in extracts from both solubilization steps after lysing Arath-EVs. E) Venn diagram of proteins found in each of the three replicates of the extract from the second step of Arath-EVs. E) Venn diagram of proteins found in each of the three replicates the extract from the second step of Braol-EVs.*

Protein-protein interaction analysis was performed to narrow down the networks of the most relevant proteins with potential utility as PDEV markers in our study. To do so, we analyzed by the software STRING the list of 1145 proteins from the Arath-EVs proteomics. We reviewed the "Functional enrichment visualization" graphs generated and, specifically, the one belonging to the "Cellular component" category, which showed the term "Secretory vesicle" occupying the fifth position in the top 10 enriched terms (**Fig. 3.A**). Hence, we decided to perform a specific functional enrichment analysis with Cytoscape of the proteins within this term, to get a clearer picture of which other cellular compartments may be related to this group of proteins (**Fig. 3.B**). In this network, the proximity of "Secretory vesicle" term to "Plant-type vacuole" and "Vacuole" terms is notable, aligning with the role of vacuoles as key players in the biogenesis of PDEVs. This suggests that protein families enriched in both terms could serve as strong PDEVs markers. Additionally, the "anchored component of membrane" term stood out, containing 6 fasciclin proteins (out of 9 proteins), these six also appearing in secretory vesicle and membrane terms.

Next, we selected all proteins with greater potential as PDEV markers across these four terminologies and generated a protein-protein interaction network using STRING to highlight these associations (**Fig. 3.C**).

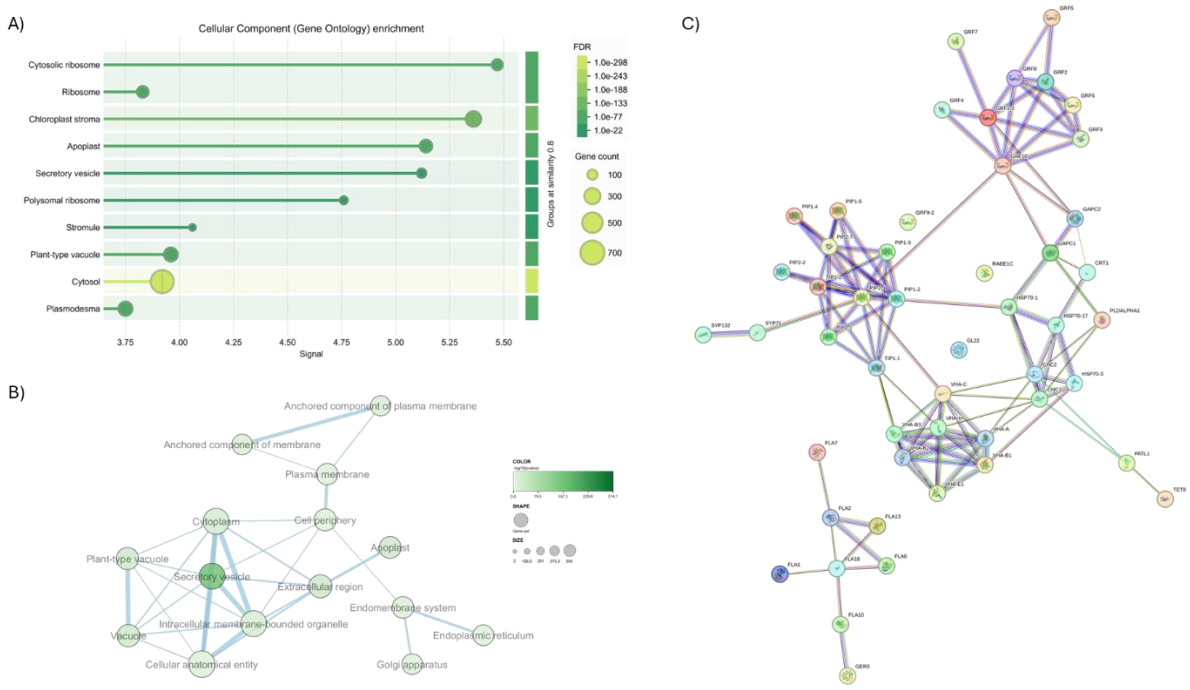


Figure 3. Protein-Protein Interaction and Functional Enrichment Analysis of Arath-EVs Proteins. *A) Cellular Component enrichment analysis after the analysis of the 1145 proteins identified in 3 biological replicates of Arath-EVs. B) Network displaying specific functional enrichment for proteins within the "Secretory vesicle" term, revealing connections to related terms such as "Plant-type vacuole" and "Vacuole," indicating shared biosynthetic pathways relevant to PDEVs. C) Protein-protein interaction network of selected proteins with high potential to be PDEVs markers. Protein-protein interactions were analyzed using STRING (https://string-db.org/). The results obtained in STRING were implemented in Cytoscape*[20].

Focusing on the protein lists obtained after this analysis, we searched for these protein´s orthologues in Braol-EVs, selecting the protein families identified in PDEVs from both species. Among these, 12 families of proteins were chosen as potential markers present in both EVs: fasciclin-like arabinogalactan proteins (FLAs), germin-like protein, patellins, heat shock proteins, aquaporins, calreticulins, clathrins, phospholipase D, Rab proteins, Vacuolar-type ATPase (V-type ATPase) proteins, 14-3-3-like proteins and glyceraldehyde-3-phosphate dehydrogenases (GADPHs) (**Table 2**).

	Arab-EVs proteins ID Uniprot (Entry name)	Braol-EVs proteins ID Uniprot	Identity (%)
PLANT SPECIFIC			
Fasciclin-like arabinogalactan proteins	O22126 (FLA8_ARATH)	A0A0D3BQ40	87,8
	Q9FFH6 (FLA13_ARATH)	A0A0D3C7E8	76,9
	Q9SJ81 (FLA7_ARATH)	A0A0D3BDD6	87,5
	Q9SU13 (FLA2_ARATH)	A0A0D3BG61 A0A0D3A9C4	89,4 90,1
	Q9FM65 (FLA1_ARATH)	A0A0D3EBG9	89,4
	Q93W32 (FLA18_ARATH)	A0A0D3CL34	90,8
	Q9LZX4 (FLA10_ARATH)	A0A0D3DUQ7	88,6
Germin-like proteins	P94072 (GL33_ARATH)	A0A0D3EDR5 A0A0D3AK68	92,4 91,5
	Q9FZ27 (GL22_ARATH)	A0A0D3B244 A0A0D3C734	92,4 91,3
Patellins	Q56WK6 (PATL1_ARATH)	A0A0D3D022 A0A0D3AQ01	81,7 81,7
ALREADY DEFINED AS PLANT MARKERS			
Heat shock proteins	P22953 (HS701_ARATH) O65719 (HSP7C_ARATH) F4JMJ1 (HSP7R_ARATH)	A0A0D3A5X5 A0A0D3BE26 A0A0D3A7J8	93,5 92,8 89,2
HOMOLOGOUS TO MAMMALS			
Aquaporins	P43286 (PIP21_ARATH)	A0A0D3BRA6	90,2
	Q06611 (PIP12_ARATH)	A0A0D3C6I1 A0A0D3B6Y1	99,0 98,6
	P93004 (PIP27_ARATH)	A0A0D3BRA6	94,7
	Q08733 (PIP13_ARATH)	A0A0D3B963	95,5
	Q8LAA6 (PIP15_ARATH)	A0A0D3B963	92,0
	P61837 (PIP11_ARATH)	A0A0D3BQ83	97,6
	P43287 (PIP22_ARATH)	A0A0D3BRA6 A0A0D3B5C8	94,7 90,2
	Q39196 (PIP14_ARATH)	A0A0D3B963	97,9
	P25818 (TIP11_ARATH)	A0A0D3B5A5	94,0
	Q41963 (TIP12_ARATH)	A0A0D2ZPE6	91,3
Calreticulin	O04151 (CALR1_ARATH)	A0A0D3A9V8	92,1
		A0A0D3E5U4	92,3
		A0A0D3DNK8	84,3
Clathrins	Q0WLB5 (CLAH2_ARATH)	A0A0D3AEI6	97,2
	Q0WNJ6 (CLAH1_ARATH)		97,9
Phospholipase D	Q38882 (PLDA1_ARATH)	A0A0D3CJS6	94,6

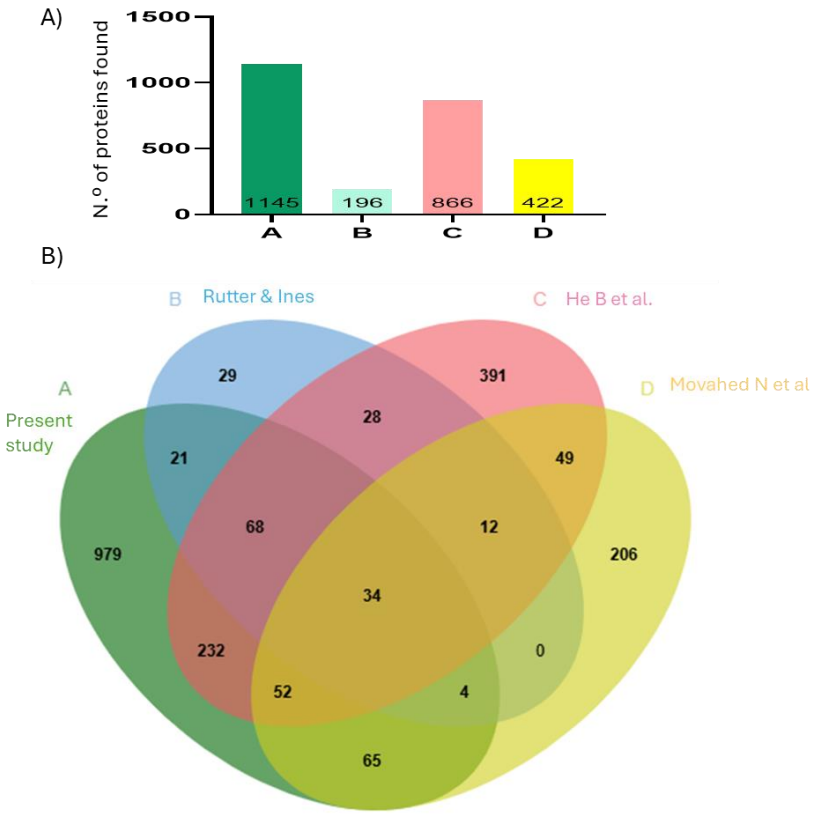
Rab	P28186 (RAE1C_ARATH)	A0A0D2ZPS4	98,6
V-type proton ATPases	P11574 (VATB1_ARATH)	A0A0D3D108	99,0
	Q8W4E2 (VATB3_ARATH)	A0A0D3D108	98,2
	O23654 (VATA_ARATH)	A0A0D3ARK5 A0A0D3CVY0	97,4 97,8
	Q39258 (VATE1_ARATH)	A0A0D3B844 A0A0D3E956	94,3 93,5
	Q9SDS7 (VATC_ARATH)	A0A0D3DPE1 A0A0D3DU49	96,0 96,5
	Q9SZN1 (VATB2_ARATH)	A0A0D3D108	96,1
	Q9LX65 (VATH_ARATH)	A0A0D3DHM0	92,0
14_3_3	P46077 (14334_ARATH)	A0A0D3CW05 A0A0D3ARF8 A0A0D3D1J9	91,4 89,5 90,7
	P42643 (14331_ARATH)	A0A0D3CW05	90,4
	Q96299 (14339_ARATH)	A0A0D3B6E0	92,8
	Q01525 (14332_ARATH)	A0A0D3CW05	96,1
	P48347 (14310_ARATH)	A0A0D3CCE5	95,7
	Q96300 (14337_ARATH)	A0A0D2ZTX3 A0A0D3B9F1	95,0 93,1
	P48349 (14336_ARATH)	A0A0D3BET0 A0A0D3AIF5	93,1 93,1
	P42644 (14333_ARATH)	A0A0D2ZSK7 A0A0D3BZM6	95,3 94,9
	P48348 (14338_ARATH)	A0A0D3BET0 A0A0D3E2C2	96,8 95,5
	P42645 (14335_ARATH)	A0A0D2ZTX3	90,3
GAPDHs	P25858 (G3PC1_ARATH)	A0A0D3C9T6 A0A0D3DPI3	96,7 96,5
	Q9FX54 (G3PC2_ARATH)	A0A0D3AFP0 A0A0D3CMV8	96,7 97,0

Table 2. Putative protein markers. *This table presents 12 protein families selected as potential markers due to their biological relevance. Within these families, proteins identified in both Arath-EVs and Braol-EVs are sorted by abundance based on Arath-EVs´ protein list and their percentage of identity is displayed.*

Comparison with previously published PDEV proteomes

The 12 protein families found in common on the proteomic results of Arath-EVs and Braol-EVs, were analyzed by a second approach to explore their presence in other EVs´ proteomics studies.

As *A. thaliana* is the best annotated species in the UniProt database, we compared the proteomic data obtained in the present study with 3 reference datasets from publications of *A. thaliana* AWF-derived EVs[3,22,23] to compare these protein lists with the results present in this study. The size of each list appears in **Figure 4.A**. In common, 34 proteins were found between the four lists (**Fig. 4.B and C**). Of these 34 proteins, 27 appeared in our data of Braol-EVs (underlined in **Fig. 4.C**). The UniProt IDs and % of identity of these proteins are shown in **Suppl. Table 3**. This list includes the previously mentioned FLAs, heat shock protein, aquaporins, clathrins, phospholipase D, Rab proteins, V-type proton ATPases and 14-3-3 family proteins. Additionally, some unspecific proteins with low evidence to be present in EVs were also identified, which were subsequently excluded from further analyses. Notably, one protein family of common EV-related proteins consistently identified on all Arath-EVs studies but not found in our previous analysis emerged: the syntaxins, which we could not identify in Braol-EVs.

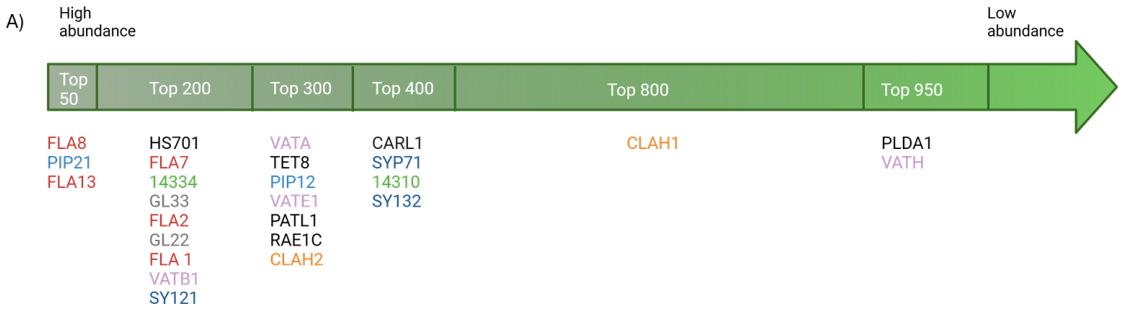


C)

Uniprot ID	Entry Name
O03042	RBL_ARATH
O22126	FLA8_ARATH
O50008	METE1_ARATH
Q8RWV0	TKTC1_ARATH
Q7Y208	GPDL1_ARATH
Q42547	CATA3_ARATH
P22953	HS701_ARATH
Q9SU13	FLA2_ARATH
P93819	MDHC1_ARATH
P48491	TPIS_ARATH
Q9ZSD4	SY121_ARATH
O23654	VATA_ARATH
O04309	JAL35_ARATH
Q06611	PIP12_ARATH
P25696	ENO2_ARATH
P93004	PIP27_ARATH
Q39258	VATE1_ARATH
P28186	RAE1C_ARATH
Q0WLB5	CLAH2_ARATH
Q9ASR1	EF2_ARATH
Q9SF29	SYP71_ARATH
P48347	14310_ARATH
Q8VZU2	SY132_ARATH
Q9XIE2	AB36G_ARATH
P24636	TBB4_ARATH
P53492	ACT7_ARATH
P47998	CYSK1_ARATH
P19456	PMA2_ARATH
Q0WNJ6	CLAH1_ARATH
Q9SPE6	SNAA2_ARATH
Q38882	PLDA1_ARATH
Q9XGM1	VATD_ARATH
Q9LX65	VATH_ARATH
Q9LJI5	VA0D1_ARATH

Figure 4. Comparative Analysis of A. thaliana vesicle proteins. *A) Bar diagram showing the number of vesicle-associated proteins identified in each of the four studies used. B) Venn diagram illustrating the 34 proteins shared across all datasets. C) Table listing these 34 common proteins, with proteins also identified in our Braol-EVs dataset highlighted in bold.*

To assess the potential of these candidates as global markers for PDEVs, we generated a unified list of proteins shared by both approaches. This merged list included 11 out of the 12 families in **Table 2** (excluding the GADPH family due to their unspecifity), plus the syntaxin family, noted in **Figure 4.C**. We also added the tetraspanin family, as TET8 is suggested as a PDEVs marker in multiple studies[6]. TET8 was identified in our Arath-EVs proteomics data, but was absent in two out of the 3 Braol-EVs´ replicates. Additionally, TET8 appeared in two out of the three *A. thaliana* reference studies previously published.



B)

			Nº of proteins	Year	Authors	Reference
AWF	**Shorgum**	*Sorghum bicolor*	437	2023	Caplan JL. et al	24
	Sunflower	*Helianthus annuus*	237	2018	Pinedo & Canal	25
Juice	**Pomegranate**	*Punica granatum*	145	2022	Marcilla et al	26
	Citrics	*Citrus sinensis, C. limon, C. paradisi, C. aurantium*	600-800	2018, 2015	Vékey K. et al; Alessandro R. et al	14, 27
Callus	**Pine**	*Picea abies*	557	2024	Kärkönen A.	28
Polen	**Kiwi**	*Actinidia deliciosa*	1170	2023	Del Duca S	12

Figure 5. Potential PDEVs marker proteins: experimental abundance estimation and cross-study comparation. *A) Proteins of interest ordered by their potential abundance, calculated with the normalized LFQ parameter, in the proteomic data from A. thaliana. B) Selected proteomic datasets.*

A certain number of proteins from each family were chosen for further analyses based on two criteria: 1) Being the most abundant in our data from Arath-EVs and 2) Presence among the 34 proteins found in common across our data and the 3 reference publications explored. This list of proteins is ordered by their abundance in **Figure 5.A**. Finally, we conducted a bibliographic revision of all the proteomic studies on EVs from other plant sources to verify the frequency of these proteins as potential markers. The proteomic datasets selected had to fulfil the following criteria: 1) Isolating true extracellular vesicles (intracellular vesicles underrepresented); 2) Depth of > 100 proteins identified; 3) Valid isolation methods according to MISEV guidelines. This

search yielded 8 articles: 3 on AWF from sorghum[24] and sunflower[25], 3 on juice from pomegranate[26] and diverse citrus plants[14,27] 1 on callus secretome from pine[28] and 1 on pollen secretome from kiwi[12]. All of these are shown in **Figure 5. B.**

	SORGHUM	SUNFLOWER	POMEGRANATE	CITRICS	PINE	KIWI
Fascilins	✓	✓	✓	✓	✓	-
FLA8_ARATH/FLA10*	-	-	✓	✓	✓	-
FLA13_ARATH	✓	-	-	-	-	-
FLA7_ARATH	✓	-	✓	-	-	-
FLA2_ARATH*	✓	✓	-	-	-	-
FLA 1_ARATH*	✓	✓	-	✓	-	-
Germin-like proteins	-	✓	-	✓	✓	-
GL33_ARATH*	-	✓	-	✓	-	-
GL22_ARATH	-	✓	-	-	✓	-
Patellins	✓	-	✓	✓	-	-
PATL1_ARATH*	✓	-	✓	✓	-	-
Heat shock proteins	✓	✓	✓	✓	✓	✓
HS701_ARATH*	-	✓	✓	-	✓	✓
Aquaporins	✓	-	✓	✓	-	✓
PIP21_ARATH*	✓	-	✓	✓	-	-
PIP12_ARATH*	-	-	-	✓	✓	-
Calreticulins	-	-	-	✓	-	✓
CALR1_ARATH	-	-	-	✓	-	✓
Clahtrins	✓	✓	-	✓	-	✓
CLAH2/CLAH1_ARATH*	✓	✓	-	✓	-	✓
Phospholipases	✓	-	✓	✓	-	✓
PLDA1_ARATH	✓	-	✓	✓	-	✓
Rab	✓	✓	✓	✓	✓	✓
RAE1C_ARATH*	✓	-	✓	✓	✓	✓
V-type proton ATPases	✓	✓	-	✓	✓	✓
VATB1_ARATH*	-	-	-	✓	-	-
VATA_ARATH*	✓	✓	-	✓	-	-
VATE1_ARATH	-	-	-	✓	-	✓
14_3-3	✓	-	✓	✓	✓	✓
14334_ARATH*	✓	-	✓	✓	✓	✓
14310_ARATH*	✓	-	-	✓	-	✓
Syntaxins	✓	-	-	✓	✓	✓
SY121_ARATH*	-	-	-	✓	-	-
SYP71_ARATH	✓	-	-	✓	-	✓
SY132_ARATH*	✓	-	-	✓	✓	✓
Tetraspanins	✓	-	✓	✓	-	✓
TET8_ARATH*	✓	-	✓	✓	-	✓

Table 3. Presence / absence of selected protein families and family members across various PDEVs. *The table lists each protein family and its corresponding subfamilies, indicating the presence (✓) or absence (-) of each in the analyzed species. Family members written in red come from our Arath-EVs and Braol-EVs data and appear in the 3 A. thaliana datasets. Those writings in green appear in our data of Arath-EVs and Braol-EVs, but do not appear in the 3 A. thaliana datasets. Finally, those written in blue do not appear in our Braol-EVs data, but appear in Arath-EVs and in the datasets. FLA8/FLA10 and CHAH2/CLAH1 were grouped for this analysis due to their similar identity scores.*

FLA8/FLA10 and CHAH2/CLAH1 were grouped for this analysis due to their similar identity scores. All the protein families we have selected appear recurrently in other proteomic studies, which corroborates their presence in different sources of PDEVs **(Table 3).**

Protein orthology and phylogenetic distance

Next, an orthologues′ search was conducted using BLASTP from UniProt, with *A. thaliana* protein sequences queried against those of six representative species from each subdivision of the plant kingdom, together with two mammalian species (human and mouse). All IDs chosen to make the phylogenetic trees appear in **Suppl. Table 4.** The aim was to evaluate the conservation degree of these proteins among representative phylogenies, as well as their preservation between kingdoms. Furthermore, a similar orthology analysis within the animal kingdom was conducted for CD63, CD81, and CD9, which are established EV markers in animals, providing a reference to help validate our proposed proteins as potential PDEV markers.

To illustrate the evolutionary relationships among putative orthologues, phylogenetic trees were generated using MEGA11, with all trees represented using a uniform scale of 0.05. While phylogenetic trees for a selection of proteins listed in **Table 3** (labelled with *) are provided in the supplementary materials, **Figure 6** focuses on representative trees for some proteins from certain specific families. Since more abundant proteins will be more favorable markers, we decided to eliminate the most scarce proteins: clathrin heavy chain 1 (CLAH1), phospholipase D alpha 1 (PLDA1) and V-type proton ATPase subunit H (VATH) for further analyses (**Fig. 5.A**). Also, calreticulin family was discarded since it was only present in 2 out of the 6 datasets of PDEVs analyzed (**Table 3**).

Two proteins already defined as EV markers are included to serve as benchmarks, allowing us to determine whether the evolutionary distances of our proposed markers fall within acceptable ranges: CD63, a specific marker for animal species, and TET8, defined as a specific marker for PDEVs that is also present in animals. The other two proteins highlighted in **Figure 5** correspond to FLA8, the most abundant member of a novel protein family of potential PDEV markers, FLAs, and heat shock 70 kDa protein 1 HSP701, a classical EV marker with potential as PDEV marker expressed across both plant and animal taxonomy.

In the phylogenetic tree for CD63 (**Fig. 6.A**), using *H. sapiens* as reference, the putative orthologues from mammalian species form a distinct cluster in the upper branches, with minimal

evolutionary distances between them. This cluster is followed by a more distant grouping of proteins from amphibians, birds, and fish. At the most distal evolutionary branch, proteins from invertebrate species are observed. Notably, the total evolutionary distance between *H. sapiens* and the most distant species, *C. elegans*, is 0.69. CD9 and CD81 phylogenetic trees display a similar pattern (**Suppl. Fig. S.1**).

For the TET8 phylogenetic tree (**Fig. 6.B**), as expected, the putative orthologues from species closely related to *A. thaliana*, such as other angiosperms and gymnosperms, exhibited the shortest evolutionary distances. No orthologue was found for chlorophytes, and orthologues from the animal kingdom appear on the furthest branches, reflecting the largest evolutionary distances. The distance between *A. thaliana* and the most distant plant orthologue, TET8 from *M. Polymorpha L.,* is 0.47, and in comparison, to *H. sapiens*, is 0.8.

The phylogenetic tree of FLA8 (**Fig. 6.C**) demonstrates a similar evolutionary pattern for plant species to that of TET8, with angiosperms and gymnosperms showing shorter evolutionary distances. However, in this case, the orthologues identified for phylogenetically distant plant species, such as bryophytes and chlorophytes, displayed significantly greater evolutionary distances, particularly in the case of the algae. Excluding the algae, which consistently occupy a distant position in all phylogenetic trees and in which no ortholog for TET8 was found, the next most distant FLA8 orthologue is in *M. polymorpha*, with an evolutionary distance from *A. thaliana* of 0.67. FLA1 and FLA2, as well as germin-like 3 (GER33) and patellin-1 (PATL1) phylogenetic trees (**Suppl. Fig. S.2 and 3**), which we have also proposed as PDEVs specific markers, are similar to that of FLA8. These three proteins are plant-specific and, logically, the evolutionary distances within the plant kingdom for these proteins are the highest among all proteins analyzed, yet remaining within acceptable thresholds when compared to the previously described distances for CD63 and TET8.

Finally, HSP701 (**Fig. 6.D**) displays a similar evolutionary trend for the plant species than the FLA8 tree, though with much shorter phylogenetic distances. For mammalian orthologues, as anticipated, the evolutionary distances are the greatest. This pattern is observed across all potential

PDEVs markers assayed. Also, the evolutionary distance between the orthologues of *A. thaliana* and *M. polymorpha* is 0.09, the shortest distance among all proteins evaluated before. Additionally, the evolutionary distance between *A. thaliana* and *H. sapiens* HSP701 is 0.23, indicating a high level of conservation across species and kingdoms. This conservation pattern, based on evolutionary distances, is similarly observed for Ras-related protein RABE1c (RAE1C), for proteins from the V-type proton ATPase family, specifically the catalytic subunit A (VATA) and subunit B1 (VATB1), as well as for the 14-3-3 protein (**Suppl. Fig. S4, 5 and 6**). However, it is worth noting that for 14334, the distribution pattern of different phylogenies within the tree appears somewhat unclear, possibly due to issues with protein sequence annotation.

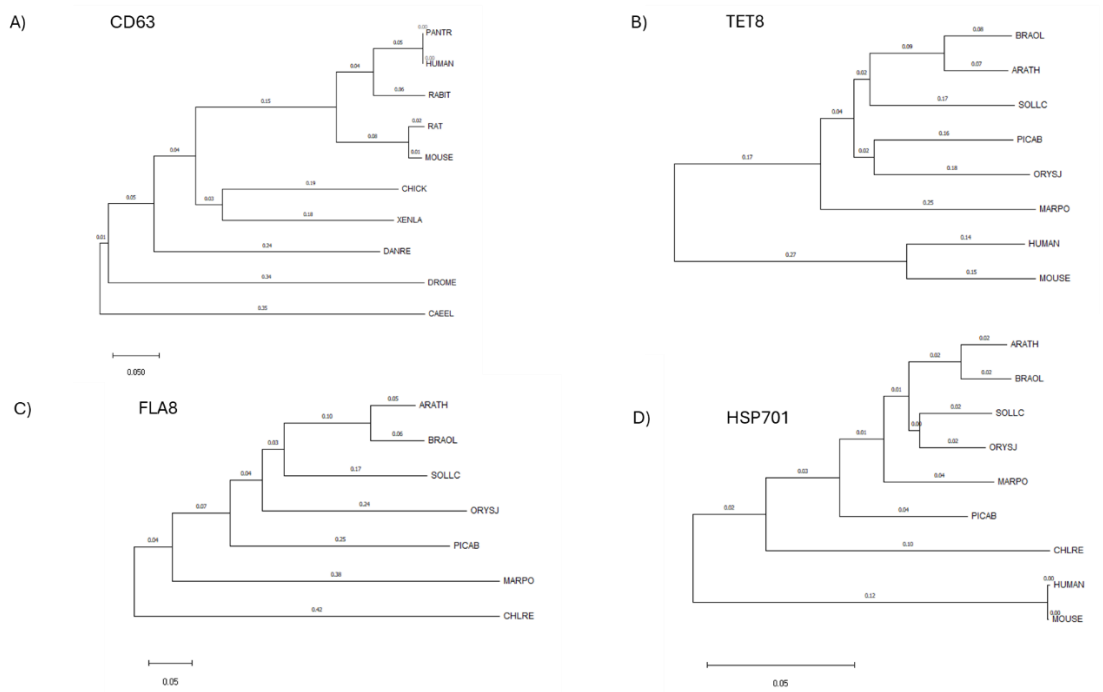


Figure 6. Phylogenetic trees. *Evolutionary distance representation of A) CD63, B) TET8, C) FLA8, D) HS701. The scale bar is set at 0.05. Phylogenetic trees were generated using MEGA11*[21].

Moreover, both aquaporins, clathrin heavy chain 2 (CLAH2) and syntaxins 121 (SY121) and 132 (SY132) (**Suppl. Fig. S7, 8 and 9**) show slightly larger evolutionary distances that remain within a range comparable to those observed for TET8, along with a similar distribution pattern across phylogenies. In some cases, however, the algae ortholog appears more distantly related than orthologous proteins from animal species.

Discussion

Although the field of plant-derived EVs needs to fulfill the gap of marker determination to comply with the guidelines of the ISEV (MISEV), this has not been accomplished so far due to the lack of a compendium of *bona fide* markers of PDEVs that could be used as a reference. In this work, we have paved the way for the standardization of PDEVs markers, providing in-depth proteomic data of two reference plant species, discovering novel potential markers, as well as conducting an exhaustive analysis of the EVs proteomic data from any plant species available so far to evaluate current potential markers.

In order to overcome the absence of raw proteomic data available in repositories, we hereby provide the most profound proteomic data obtained so far from apoplastic washing fluid-derived EVs from *A. thaliana.* and *B. oleracea.* To accomplish this, we have developed a novel methodology to extract proteins from these EVs, based on a two-step solubilization. Preliminary proteomic analysis of the first extraction step, compared to that of the second one, demonstrated that most proteins from the first extract (651 out of 686, **Suppl. Table 1 and Fig 2.D**) were as well identified in the second extract, which shall additionally be enriched in membrane bound proteins, that may potentially contain more relevant EVs markers. For these reasons, we focused on the second extract, identifying 1145 proteins in *A. thaliana* and 899 in *B. oleracea*, detected in all 3 biological replicates assayed (**Suppl. Table 2, Fig. 2.E and 2.F**). Next, we compared both extracts, searching for orthologues identified in both species with potential as EVs markers, focusing on proteins related to EVs biology, arising from the bioinformatic post-analysis (**Fig. 3**). Twelve protein families were found consistently identified in both EV proteomes: fasciclin-like arabinogalactan proteins, germin-like proteins, patellins, heat shock proteins, aquaporins, calreticulins, clathrins, phospholipase Ds, Rab proteins, V-type proton ATPase subunits, 14-3-3 like proteins and GAPDHs (**Table 2**). All the protein orthologues showed more than 75% of identity, most being above 90%. *A. thaliana* and *B. oleracea* belong to the *Brassicaceae* family of angiosperm eudicot plants and therefore, our analysis will allow to rule out species-specific markers, since those not reproducibly identified in plant material from the same family would

rarely be useful to identify EVs from other subdivisions of the plant kingdom. Interestingly, none of the classical Arath-EVs PDEVs markers were consistently identified in both proteomes: TET8 present in *A. thaliana* but only detected in one replicate of *B. oleracea*; TET9 not identified in any replicate of both; and syntaxin 121 (SY121 or PEN1) present in *A. thaliana.* but not detected in any replicate of *B. oleracea.* Besides, EXO70E2 was not detected in any EV extract, which may imply a low contribution of EXPO-derived EVs to the AWF-EV pool. Concerning orthologues of mammalian EV markers suggested in literature, TSG101 and Alix were not detected in any of the replicates from both plant EVs. The absence of reproducible identification of these markers in Arath-EVs may be explained by a low protein abundance, since there is great evidence in the literature of the presence of these markers in such EVs[6]. This shall be enough reason to consider these as unideal markers, in favor of those that we could consistently identify, since their determination may be complicated due to their low abundance. Anyhow, we can never rule out the possibility that all these proteins′ peptides may not be the most favorable to be detected by LC-MS/MS, which could explain why they were not identified. On the contrary, one member of the Hsp70 family, HS701, was consistently found in both proteomes and arises as a robust marker for PDEVs. The high level of homology of this protein with that of animal sources potentially offers the possibility to use antibodies raised against animal immunogens to target it but has the drawback of lacking plant specificity, as has been the case in one study with algae EVs[15]. Concerning GAPDHs, their consistent identification may allow to use them as housekeeping proteins in certain assays, but we can anticipate a lack of specificity as PDEVs markers due to their ubiquitous expression in cells and tissues.

Since *A. thaliana.* is the best annotated species in UniProt database (the one used in our proteomic analysis) and has the most PDEVs proteomic studies published to date, we selected the 3 reference datasets from AWF-derived Arath-EVs published so far[3,22,23] as to compare with our results. Considering that the raw data was not deposited in repositories in all 3 cases, we compared the list of identified proteins in all 4 studies, finding 34 proteins in common. Among these, fasciclin-like arabinogalactan proteins 2 (FLA2) and 8 (FLA8), heat shock 70 kDa protein 1 (HS701),

aquaporins PIP1-2 and PIP2-7, clathrin heavy-chain 1 (CLAH1) and 2 (CLAH2), phospholipase D alpha 1 (PLDA1), Ras-related protein RABE1c (RAE1C), V-type proton ATPase subunits A, D, D1, E1 and H; 14-3-3-like protein GF14 epsilon (14310) and syntaxin 121 (SY121 or PEN1), 71 (SYP71) and 132 (SY132) were found.

With the goal in mind of assessing the utility of the potential PDEVs markers found, we searched for the appearance of the protein families aforementioned in previous proteomic studies of EVs from other plant sources. The proteomic datasets selected had to fulfill the following criteria: 1) Isolating true extracellular vesicles (intracellular vesicles underrepresented); 2) Depth of > 100 proteins identified; 3) Valid isolation methods according to MISEV guidelines. Concerning the first criterium, we have included fruit juice-derived EVs as a mild homogenization method providing an isolate enriched in EVs, since the potential contribution of intracellular EVs is expected to be minimized, but conclusions extracted from such studies may not be as specific as those from AWF, or else pollen or tissue/cell culture supernatants. All families studied but calreticulins were found at least in 50% of the PDEVs´ datasets analyzed, supporting their recurrent identification in EVs from other plant sources. This reinforces their utility as markers for PDEVs but, for this matter, their specificity for differentiating EVs also needs to be considered. In fact, clathrin heavy chains 1 and 2 are reproducibly identified on PDEVs but may not be ideal markers since they are highly expressed in intracellular vesicles. On the other hand, V-type proton ATPase subunits A, D, D1, E1 and H arise as novel EV markers, all belonging to a vacuolar ATPase protein complex that shall be present in these vesicles´ membrane. This protein complex has not been either proposed as EV marker in mammalian cells, but presence of these in EVs from several animal species is provided in Vesiclepedia[29] (**Suppl. Table 5**).

Among the protein families with potential to be used as PDEVs markers, there are 3 that are plant-specific: patellins[30], germin-like proteins[31] and fasciclin-like arabinogalactan proteins (FLAs)[32]. There are several advantages of using plant-specific markers to identify EVs: 1) Ruling out cross-contamination with animal material; 2) Unequivocally targeting PDEVs on animal *in vitro* and *in vivo* settings; and 3) Possibility to differentiate PDEVs from EVs from bacteria/fungi/oomycetes

in host/pathogen and plant/microbe interaction studies. The presence of patellins[6,33] on PDEVs has been previously reported and, in our data, patellin 1 (PATL1) stands as the more reproducibly identified. This protein has been shown to be involved in membrane-trafficking[30] and so, seems a valid candidate to be used as PDEVs marker. On the other hand, germin-like proteins are extracellular matrix proteins belonging to the apoplast[31]. For this reason, although these are interesting candidates, further analyses will be required for their use as EVs markers, since they could either be co-isolated with EVs or be a part of the protein corona, rather than present within EVs. The case of the FLAs family deserves particular attention. These proteins are plant-specific and most are unequivocally bound to the plasma membrane through a GPI-anchor[34], being FLA18 the only exception among the 7 FLAs found in both proteomes. This is indeed the case for 6 FLAs (1, 2, 7, 8, 10, 13)[32] found in proteomes from *A. thaliana.* and *B. oleracea* EVs. Furthermore, there are no other variants described for these proteins that could be secreted instead of anchored to the plasma membrane[32]. GPI-anchored proteins are included in MISEV2023[1], together with transmembrane proteins, in the first class of protein markers considered EV hallmarks, therefore as *bona fide* markers of EVs. The presence of several members of the FLA family, with great abundance, in the proteome of both analyzed species, as well as the recurrent identification of these in proteomic analyses of PDEVs from other sources, reinforces their great potential to be used as plant-specific markers of EVs. Proteins from this family have a common protein domain, FAS1[32], a cell adhesion domain conserved in plant and animals. Nonetheless, animal proteins containing this domain lack arabinogalactan polysaccharides and a GPI-anchor, which implies alternative subcellular localization and results in completely different functions[35].

From our proteomic data, we chose 3 FLAs presenting a GPI-anchor identified in both types of PDEVs to further analyze the presence of orthologues and their degree of conservation throughout the plant kingdom: FLA8 and FLA2, due to their presence in the three reference datasets of *A. thaliana* AWF-derived EVs[3,22,23]; and FLA1, based on its high frequency of appearance in PDEVs from other sources. Besides, the analysis was extended to other members of the most relevant protein families found to be repeatedly identified in PDEVs: GER33, PATL1, HS701, PIP21 and

PIP12, CLAH2, RAE1C, VATB1 and VATA, 14334 and 14310, and SY132. Furthermore, we decided to include the currently proposed PDEVs markers found in *A. thaliana*: SY121 (PEN1) and TET8. Besides representative species of the different plant taxonomy subdivisions, *Homo sapiens* (human) and *Mus musculus* (mouse) species were added to the analysis to evaluate the global degree of conservation of the analyzed markers. Likewise, in order to compare the degree of conservation of these PDEVs markers with that of the current EVs markers used in animal-derived EVs, the most used: CD63, CD81 and CD9[1], were chosen. The results obtained for the proteins: HS701, RAE1C, VATB1, VATA and both 14-3-3-like proteins point to a high degree of conservation throughout plants and animals. This may imply functional conservation and support their use as markers for both animal EVs and PDEVs, with the drawbacks of not ruling out cross-contamination of EVs from other sources and the impossibility to differentiate EVs from different origins when necessary. A subset of proteins (CLAH2 and aquaporins PIP12 and PIP21) do have orthologues in human and mouse, but these display low homology with those of plants, while all orthologues throughout the plant kingdom show short phylogenetic distance in all species but the algae *C. reinhardtii*. In fact, in the case of RAE1C and the aquaporins, the human and mouse orthologues display shorter phylogenetic distance than the algae ortholog. For the rest of the proteins analyzed, *C. reinhardtii* shows the greatest phylogenetic distance among all plant species and is very close to that of human and mouse orthologues, when these exist. Furthermore, we could not find an annotated ortholog in *C. reinhardtii* for GER33. This seems to imply that most of the hereby proposed PDEVs markers may not be applicable to algae-derived EVs or at least may need further confirmation in such samples. Five proteins were confirmed to be plant specific: FLA1, FLA2, FLA8, PATL1 and GER33, since no orthologues in human nor mouse were found. Their degree of conservation in plants was compared with the degree of conservation of classical markers used in mammalian cell-derived EVs in the animal kingdom. The phylogenetic distances among the same plant family (*Brassicacea*) were analogous to that of human to rabbit and rodent orthologues. The distance of human markers with bird, fish and amphibia orthologues is similar to that of *A. thaliana* proteins to other phanerogams: such as another eudicot family member (*Solanum lycopersicum*) and gymnosperms (*Picea abies*). Besides, human CD63, CD81 and CD9

have a comparable distance to invertebrates *(Drosophila melanogaster* and *Caenorhabditis elegans*) than that of FLA1, FLA2, FLA8 and PATL1 with cryptogams (*bryophyta* or mosses and liverworts; and *cryptophyta* or algae). We could not find an ortholog of GER33 in *C. reinhardtii*, as aforementioned, but the distance with the liverwort *M. polymorfa* is also analogous. These data suggest that these 5 plant-specific PDEV markers possess a similar degree of conservation in the plant kingdom than the most currently used markers in the field of EVs have throughout animal species.

Figure 6 provides an overview of the key aspects considered in this work to evaluate these protein families as potential vesicle markers. Although integrating all these factors is challenging, we have categorized the families into three groups based on our criteria. Proteins ranked at the top are those we believe possess the most favorable characteristics for serving as vesicle markers. Additionally, noteworthy features for each protein are highlighted.

Conclusions

Altogether, the results hereby presented offer a wide list of potential markers for extracellular vesicles derived from plant sources, providing evidence of identification in previous proteomic studies, as well as the existence of orthologues and conservation throughout the plant kingdom, together with human and mouse species, as reference animal proteomes. The data gathered does not strongly support the use of current PDEV markers used in *Arabidopsis thaliana*, since these are not robustly identified in our proteomic data from *A. thaliana* and *B. oleracea*, and this is the case in other PDEV proteomes published so far. In contrast, taking into account the abundance in the proteomic results of Arath-EVs in the present study and the consistency of their presence across diverse datasets of different species PDEVs, we provide evidence supporting the potential of the following protein families as PDEVs markers: fasciclin-like arabinogalactan proteins, patellins, germin-like proteins, heat shock proteins, aquaporins, 14-3-3-like proteins, Rab proteins and V-type proton ATPase subunits.

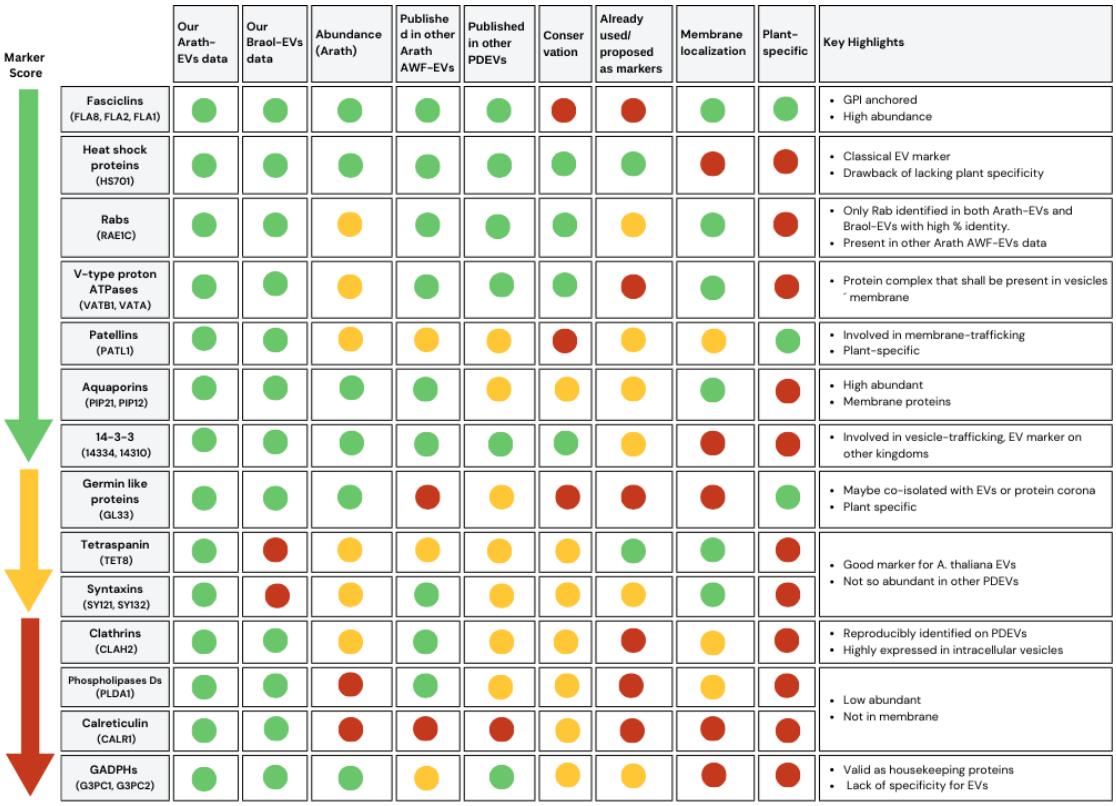


Figure 6. PDEV marker score based on the criteria used to evaluate each protein. *The table organizes proteins based on various evaluation criteria. Each column displays a color code (in order of greater value: green, orange, red) to reflect specific thresholds for each criterion. 1) Presence in our datasets of Arath-EVs: green = present; red = absent. 2) Presence in our datasets of Braol-EVs: green = present; red = absent. 3) Abundance ranking (based on our data from Arath-EVs): green = Top 200; orange = 200 - 400; red = below 400. 4) Published in other AWF-Arath-EVs datasets: green = present in all 3; orange = 2; red = 1 or none. 5) Presence in other PDEVs´ datasets: green = present in 5 - 6; orange = 3 - 4; red = 1 - 2. 6) Conservation: green = very high; orange = high; red = acceptable (plant-specific). 7) Already used/proposed as markers: green = well-established role as markers; orange= proposed in some studies; red = novel markers. 8) Membrane localization: green = transmembrane or GPI-anchored; orange = potentially associated with the membrane (e.g., through lipid interactions); red = membrane localization unlikely. 9) Plant-specificity: green = no orthologues in animals; red = orthologues found in animals.*

Besides, we hereby provide comprehensive data to define a novel family of PDEV markers: fasciclin-like arabinogalactan proteins (FLAs), as well as to sustain the use, in particular of FLA1, FLA2 and FLA8 as ideal markers for PDEVs, since they are plant-specific, possess a GPI-anchor, have orthologues across the different plant subdivisions and display a similar phylogenetic distribution in plants to that of currently used markers in animals.

References

1. Welsh JA, Goberdhan DCI, O'Driscoll L, et al. Minimal information for studies of extracellular vesicles (MISEV2023): From basic to advanced approaches. *Journal of Extracellular Vesicles*. 2024;13(2):e12404. doi:10.1002/jev2.12404

2. Pinedo M, de la Canal L, de Marcos Lousa C. A call for Rigor and standardization in plant extracellular vesicle research. *Journal of Extracellular Vesicles*. 2021;10(6):e12048. doi:10.1002/jev2.12048

3. Rutter BD, Innes RW. Extracellular Vesicles Isolated from the Leaf Apoplast Carry Stress-Response Proteins. *Plant Physiology*. 2017;173(1):728-741. doi:10.1104/pp.16.01253

4. Wojnacki J, Lujan AL, Brouwers N, et al. Tetraspanin-8 sequesters syntaxin-2 to control biphasic release propensity of mucin granules. *Nat Commun*. 2023;14(1):3710. doi:10.1038/s41467-023-39277-9

5. Zhang J, Qiu Y, Xu K. Characterization of GFP-AtPEN1 as a marker protein for extracellular vesicles isolated from *Nicotiana benthamiana* leaves. *Plant Signaling & Behavior*. 2020;15(9):1791519. doi:10.1080/15592324.2020.1791519

6. Holland S, Roth R. Extracellular vesicles in the arbuscular mycorrhizal symbiosis: current understanding and future perspectives. *Molecular Plant-Microbe Interactions*. 2023;36(4). Accessed August 23, 2024. https://ora.ox.ac.uk/objects/uuid:c712f69f-88eb-4455-9967-2cb03965a2f2

7. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes | Science. Accessed August 23, 2024. https://www.science.org/doi/10.1126/science.aar4142

8. Cai Q, Qiao L, Wang M, et al. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science*. 2018;360(6393):1126-1129. doi:10.1126/science.aar4142

9. Spitzer C, Schellmann S, Sabovljevic A, et al. The Arabidopsis elch mutant reveals functions of an ESCRT component in cytokinesis. *Development*. 2006;133(23):4679-4689. doi:10.1242/dev.02654

10. Raimondo S, Naselli F, Fontana S, et al. Citrus limon -derived nanovesicles inhibit cancer cell proliferation and suppress CML xenograft growth by inducing TRAIL-mediated cell death. *Oncotarget*. 2015;6(23):19514-19527. doi:10.18632/oncotarget.4004

11. Kalinowska K, Nagel MK, Goodman K, et al. Arabidopsis ALIX is required for the endosomal localization of the deubiquitinating enzyme AMSH3. *Proceedings of the National Academy of Sciences*. 2015;112(40):E5543-E5551. doi:10.1073/pnas.1510516112

12. Suanno C, Tonoli E, Fornari E, et al. Small extracellular vesicles released from germinated kiwi pollen (pollensomes) present characteristics similar to mammalian exosomes and carry a plant homolog of ALIX. *Front Plant Sci*. 2023;14:1090026. doi:10.3389/fpls.2023.1090026

13. Rabienezhad Ganji N, Urzì O, Tinnirello V, et al. Proof-of-Concept Study on the Use of Tangerine-Derived Nanovesicles as siRNA Delivery Vehicles toward Colorectal Cancer Cell Line SW480. *International Journal of Molecular Sciences*. 2024;25(1):546. doi:10.3390/ijms25010546

14. Pocsfalvi G, Turiák L, Ambrosone A, et al. Protein biocargo of citrus fruit-derived vesicles reveals heterogeneous transport and extracellular vesicle populations. *J Plant Physiol*. 2018;229:111-121. doi:10.1016/j.jplph.2018.07.006

15. Batsukh S, Oh S, Lee JM, Joo JHJ, Son KH, Byun K. Extracellular Vesicles from Ecklonia cava and Phlorotannin Promote Rejuvenation in Aged Skin. *Marine Drugs*. 2024;22(5):223. doi:10.3390/md22050223

16. Meier F, Brunner AD, Frank M, et al. diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods*. 2020;17(12):1229-1236. doi:10.1038/s41592-020-00998-0

17. Pang Z, Xu L, Viau C, et al. MetaboAnalystR 4.0: a unified LC-MS workflow for global metabolomics. *Nat Commun*. 2024;15(1):3675. doi:10.1038/s41467-024-48009-6

18. Perez-Riverol Y, Bai J, Bandla C, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022;50(D1):D543-D552. doi:10.1093/nar/gkab1038

19. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023;51(D1):D638-D646. doi:10.1093/nar/gkac1000

20. Shannon P, Markiel A, Ozier O, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*. 2003;13(11):2498-2504. doi:10.1101/gr.1239303

21. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*. 2021;38(7):3022-3027. doi:10.1093/molbev/msab120

22. Movahed N, Cabanillas DG, Wan J, Vali H, Laliberté JF, Zheng H. Turnip Mosaic Virus Components Are Released into the Extracellular Space by Vesicles in Infected Leaves. *Plant Physiol*. 2019;180(3):1375-1388. doi:10.1104/pp.19.00381

23. He B, Cai Q, Qiao L, et al. RNA-binding proteins contribute to small RNA loading in plant extracellular vesicles. *Nat Plants*. 2021;7(3):342-352. doi:10.1038/s41477-021-00863-8

24. Chaya T, Banerjee A, Rutter BD, et al. The extracellular vesicle proteomes of Sorghum bicolor and Arabidopsis thaliana are partially conserved. *Plant Physiology*. 2024;194(3):1481-1497. doi:10.1093/plphys/kiad644

25. de la Canal L, Pinedo M. Extracellular vesicles: a missing component in plant cell wall remodeling. *Journal of Experimental Botany*. 2018;69(20):4655-4658. doi:10.1093/jxb/ery255

26. M. Sánchez-López C, Cruz Manzaneque-López M, Pérez-Bermúdez P, Soler C, Marcilla A. Characterization and bioactivity of extracellular vesicles isolated from pomegranate. *Food & Function*. 2022;13(24):12870-12882. doi:10.1039/D2FO01806C

27. Raimondo S, Naselli F, Fontana S, et al. Citrus limon-derived nanovesicles inhibit cancer cell proliferation and suppress CML xenograft growth by inducing TRAIL-mediated cell death. *Oncotarget*. 2015;6(23):19514-19527. doi:10.18632/oncotarget.4004

28. Kankaanpää S, Väisänen E, Goeminne G, et al. Extracellular vesicles of Norway spruce contain precursors and enzymes for lignin formation and salicylic acid. *Plant Physiol*. 2024;196(2):788-809. doi:10.1093/plphys/kiae287

29. Chitti SV, Gummadi S, Kang T, et al. Vesiclepedia 2024: an extracellular vesicles and extracellular particles repository. *Nucleic Acids Res*. 2024;52(D1):D1694-D1698. doi:10.1093/nar/gkad1007

30. Peterman TK, Ohol YM, McReynolds LJ, Luna EJ. Patellin1, a Novel Sec14-Like Protein, Localizes to the Cell Plate and Binds Phosphoinositides. *Plant Physiology*. 2004;136(2):3080-3094. doi:10.1104/pp.104.045369

31. Membré N, Bernier F, Staiger D, Berna A. Arabidopsis thaliana germin-like proteins: common and specific features point to a variety of functions. *Planta*. 2000;211(3):345-354. doi:10.1007/s004250000277

32. Johnson KL, Jones BJ, Bacic A, Schultz CJ. The Fasciclin-Like Arabinogalactan Proteins of Arabidopsis. A Multigene Family of Putative Cell Adhesion Molecules. *Plant Physiology*. 2003;133(4):1911-1925. doi:10.1104/pp.103.031237

33. Nemati M, Singh B, Mir RA, et al. Plant-derived extracellular vesicles: a novel nanomedicine approach with advantages and challenges. *Cell Commun Signal*. 2022;20(1):69. doi:10.1186/s12964-022-00889-1

34. Seifert GJ. Fascinating Fasciclins: A Surprisingly Widespread Family of Proteins that Mediate Interactions between the Cell Exterior and the Cell Surface. *IJMS*. 2018;19(6):1628. doi:10.3390/ijms19061628

35. Johnson KL, Kibble NAJ, Bacic A, Schultz CJ. A Fasciclin-Like Arabinogalactan-Protein (FLA) Mutant of Arabidopsis thaliana, fla1, Shows Defects in Shoot Regeneration. *PLOS ONE*. 2011;6(9):e25154. doi:10.1371/journal.pone.0025154

Figure 1.

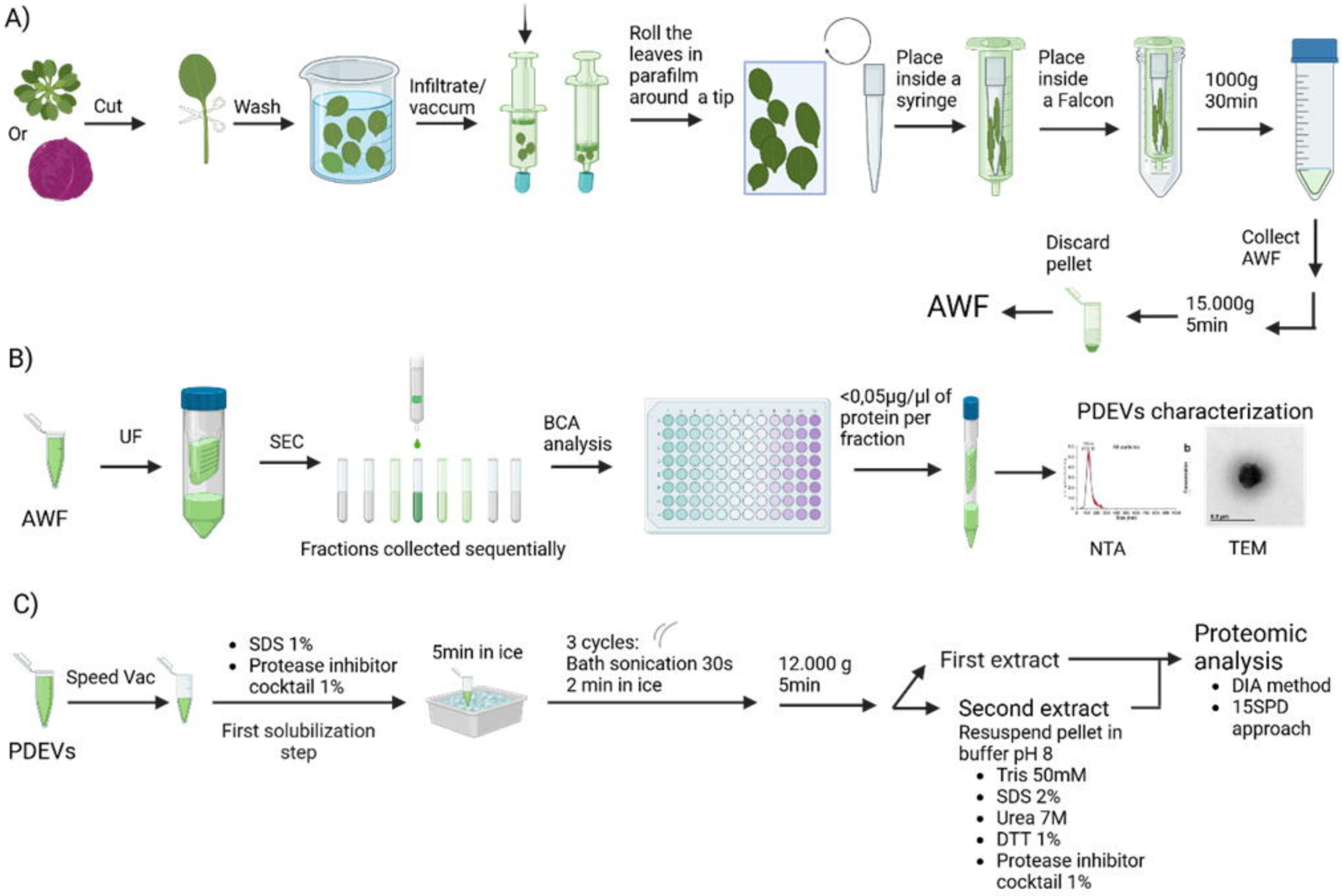

A)
Or
Cut
Wash
Infiltrate/
vaccum
Roll the
leaves in
parafilm
around a tip
Place
inside a
syringe
Place
inside
a Falcon
1000g
30min
Collect
AWF
15.000g
5min
Discard
pellet
AWF
B)
AWF
UF
SEC
Fractions collected sequentially
BCA
analysis
<0,05µg/µl of
protein per
fraction
PDEVs characterization
NTA
TEM
C)
PDEVs
Speed Vac
• SDS 1%
• Protease inhibitor
cocktail 1%
First solubilization
step
5min in ice
3 cycles:
Bath sonication 30s
2 min in ice
12.000 g
5min
First extract
Second extract
Resuspend pellet in
buffer pH 8
• Tris 50mM
• SDS 2%
• Urea 7M
• DTT 1%
• Protease inhibitor
cocktail 1%
Proteomic
analysis
• DIA method
• 15SPD
approach

Figure 2.

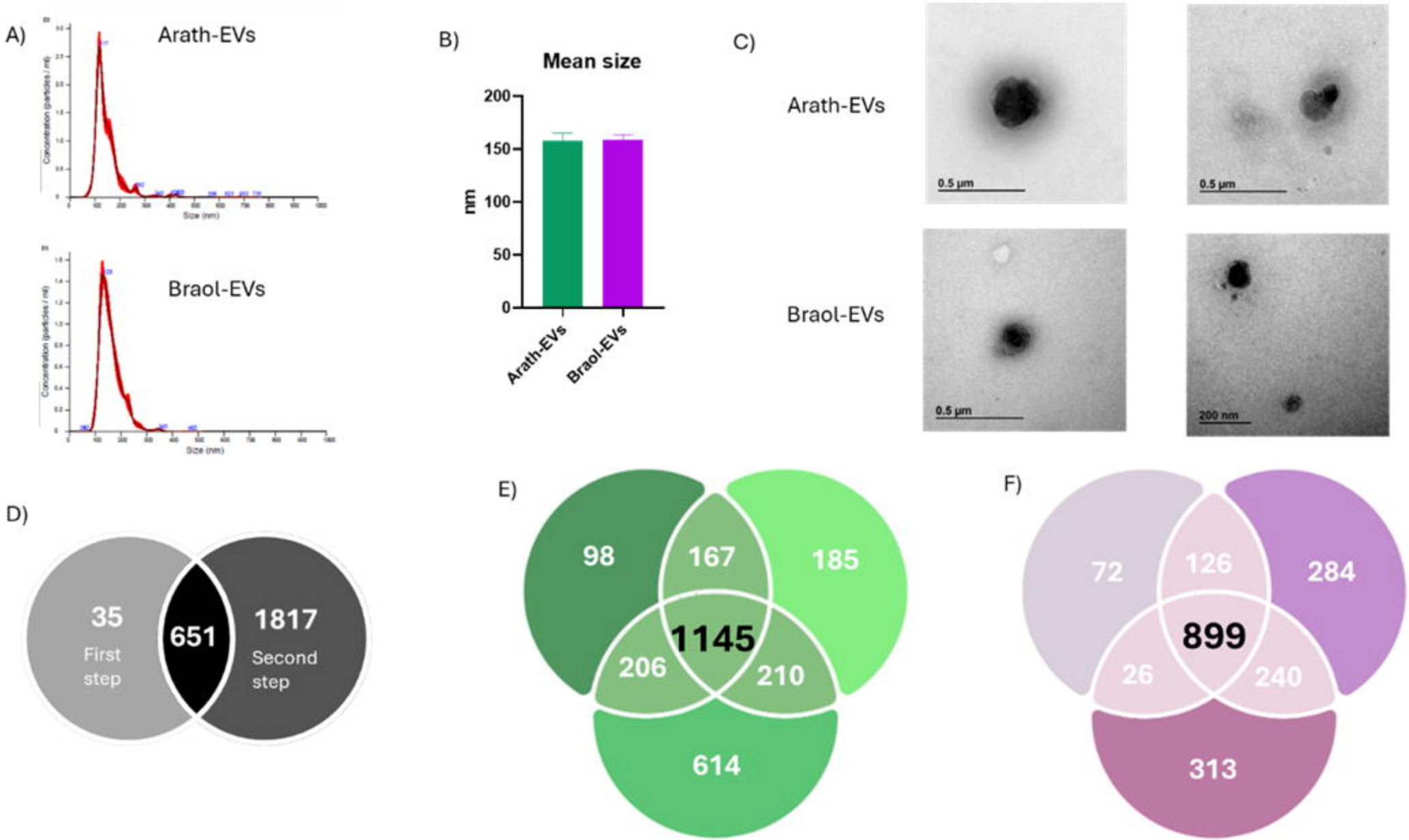
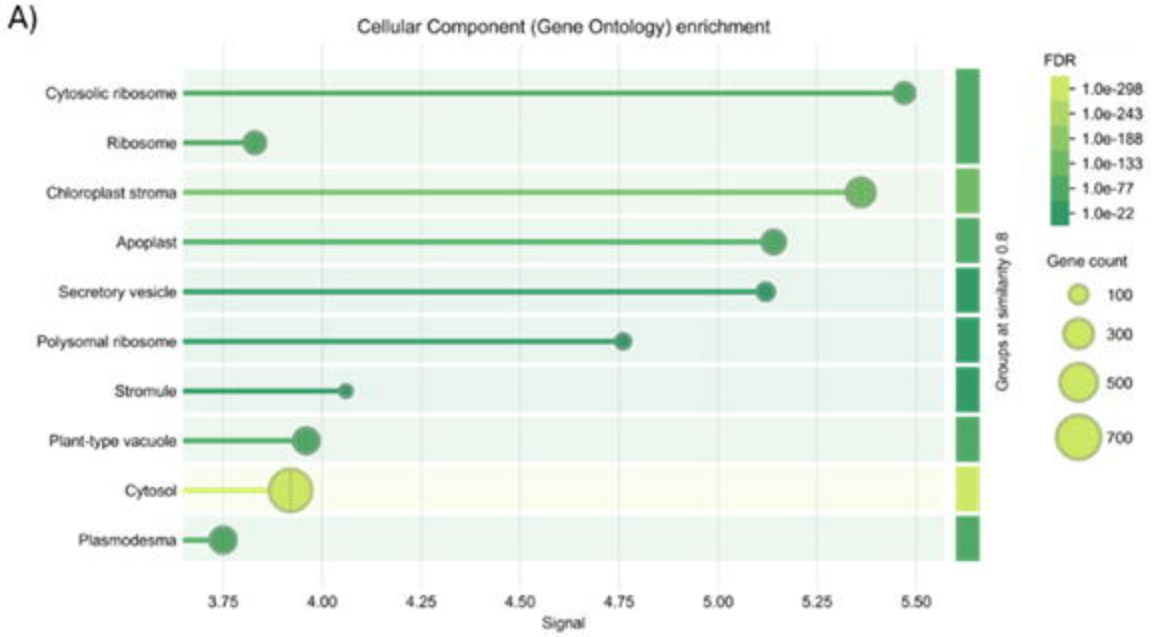
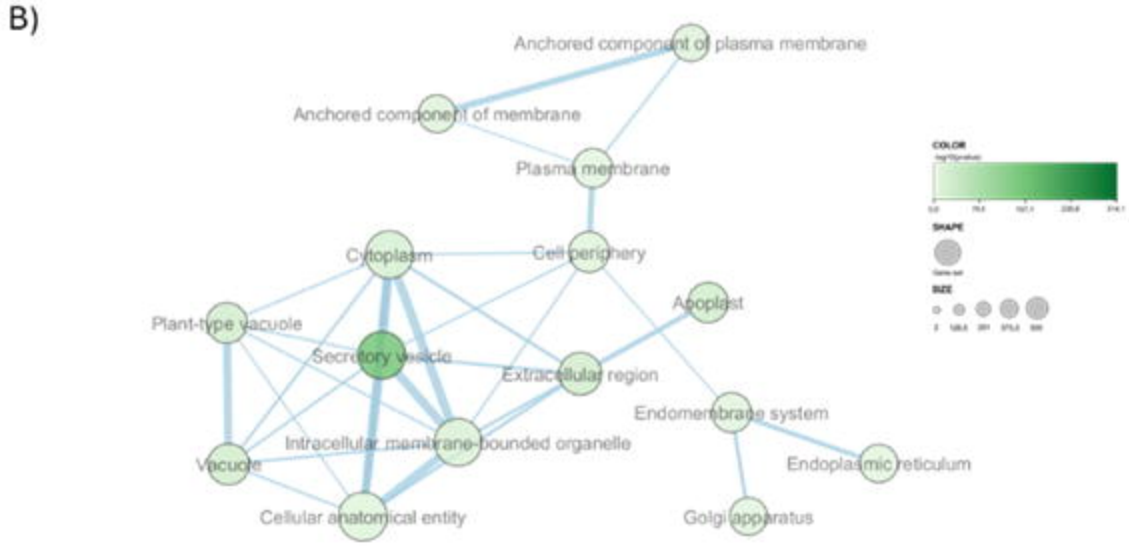

A)
Arath-EVs
Braol-EVs
B)
Mean size
nm
200
150
100
50
0
Arath-EVs
Braol-EVs
C)
Arath-EVs
Braol-EVs
0.5 µm
200 nm
D)
35
First step
651
1817
Second step
E)
98
167
185
1145
206
210
614
F)
72
126
284
899
26
240
313

Figure 3.


A)
Cellular Component (Gene Ontology) enrichment
Cytosolic ribosome
Ribosome
Chloroplast stroma
Apoplast
Secretory vesicle
Polysomal ribosome
Stromule
Plant-type vacuole
Cytosol
Plasmodesma
3.75
4.00
4.25
4.50
4.75
5.00
5.25
5.50
Signal
Groups at similarity 0.8
FDR
1.0e-298
1.0e-243
1.0e-188
1.0e-133
1.0e-77
1.0e-22
Gene count
100
300
500
700


B)
Anchored component of plasma membrane
Anchored component of membrane
Plasma membrane
Cytoplasm
Cell periphery
Apoplast
Plant-type vacuole
Secretory vesicle
Extracellular region
Endomembrane system
Intracellular membrane-bounded organelle
Vacuole
Endoplasmic reticulum
Cellular anatomical entity
Golgi apparatus
COLOR
SHAPE
SIZE

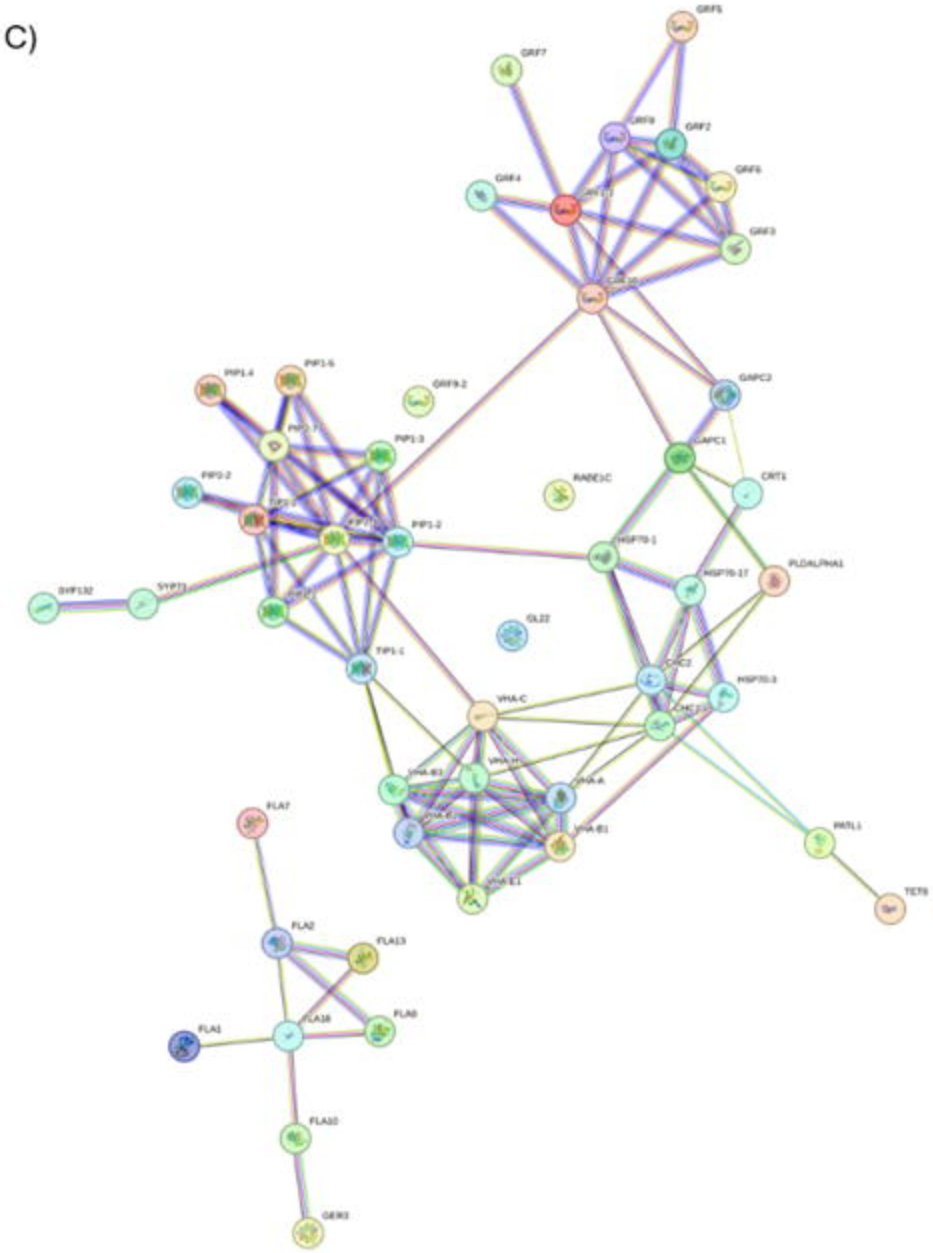
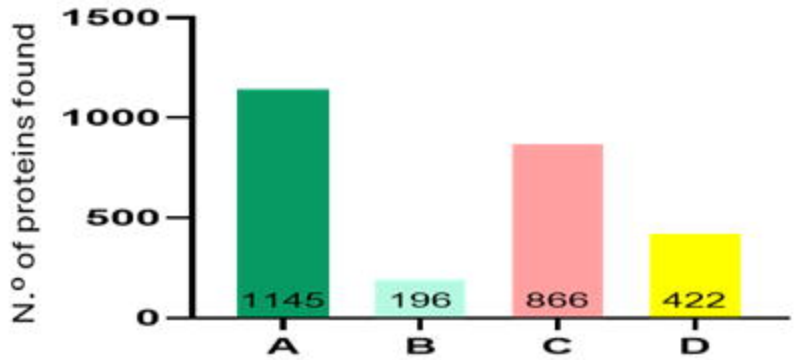
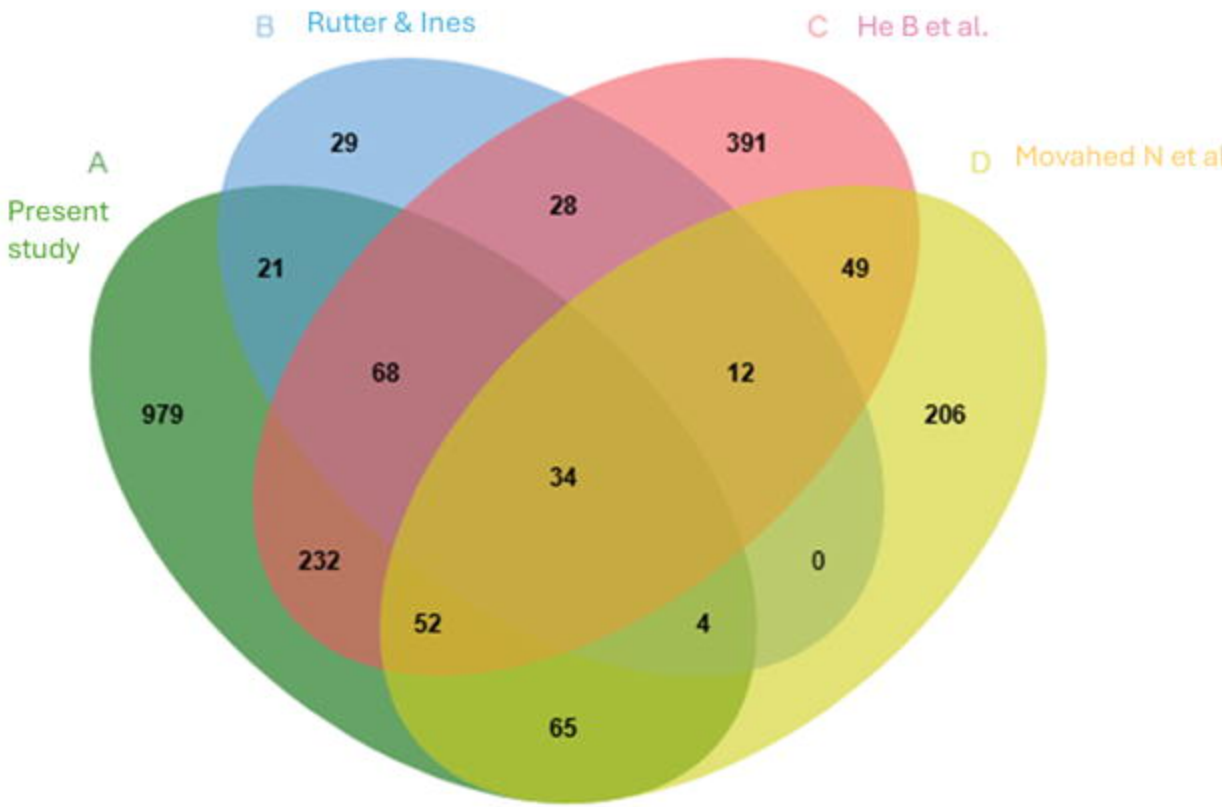

C)

Figure 4.

A)



B)

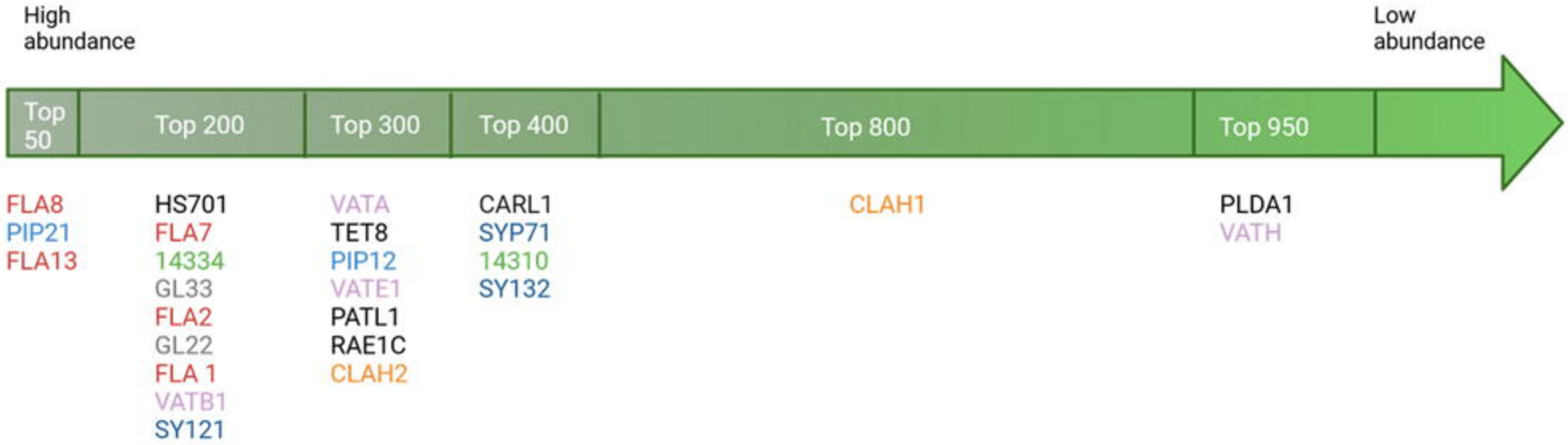


C)

Uniprot ID	Entry Name
O03042	RBL_ARATH
O22126	FLA8_ARATH
O50008	METE1_ARATH
Q8RWV0	TKTC1_ARATH
Q7Y208	GPDL1_ARATH
Q42547	CATA3_ARATH
P22953	HS701_ARATH
Q9SU13	FLA2_ARATH
P93819	MDHC1_ARATH
P48491	TPIS_ARATH
Q9ZSD4	SY121_ARATH
O23654	VATA_ARATH
O04309	JAL35_ARATH
Q06611	PIP12_ARATH
P25696	ENO2_ARATH
P93004	PIP27_ARATH
Q39258	VATE1_ARATH
P28186	RAE1C_ARATH
Q0WLB5	CLAH2_ARATH
Q9ASR1	EF2_ARATH
Q9SF29	SYP71_ARATH
P48347	14310_ARATH
Q8VZU2	SY132_ARATH
Q9XIE2	AB36G_ARATH
P24636	TBB4_ARATH
P53492	ACT7_ARATH
P47998	CYSK1_ARATH
P19456	PMA2_ARATH
Q0WNJ6	CLAH1_ARATH
Q9SPE6	SNAA2_ARATH
Q38882	PLDA1_ARATH
Q9XGM1	VATD_ARATH
Q9LX65	VATH_ARATH
Q9LJI5	VA0D1_ARATH

Figure 5.

A)



B)

			Nº of proteins	Year	Authors	Reference
AWF	**Shorgum**	*Sorghum bicolor*	437	2023	Caplan JL. et al	24
	Sunflower	*Helianthus annuus*	237	2018	Pinedo & Canal	25
Juice	**Pomegranate**	*Punica granatum*	145	2022	Marcilla et al	26
	Citrics	*Citrus sinensis, C. limon, C. paradisi, C. aurantium*	600-800	2018, 2015	Vékey K. et al; Alessandro R. et al	14, 27
Callus	**Pine**	*Picea abies*	557	2024	Kärkönen A.	28
Polen	**Kiwi**	*Actinidia deliciosa*	1170	2023	Del Duca S	12

Figure 6.

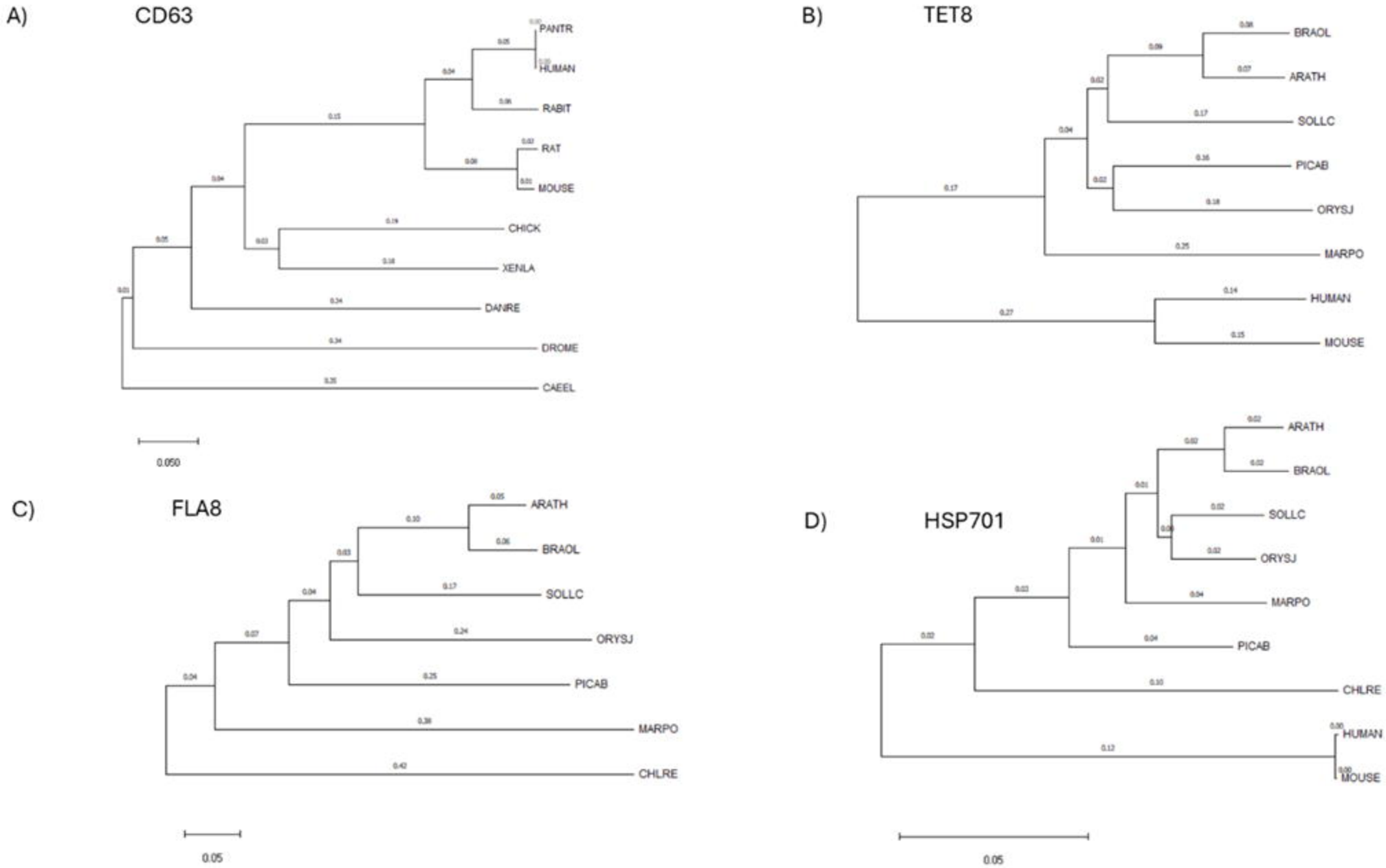

A)
CD63
PANTR
HUMAN
RABIT
RAT
MOUSE
CHICK
XENLA
DANRE
DROME
CAEEL
0.050
B)
TET8
BRAOL
ARATH
SOLLC
PICAB
ORYSJ
MARPO
HUMAN
MOUSE
C)
FLA8
ARATH
BRAOL
SOLLC
ORYSJ
PICAB
MARPO
CHLRE
0.05
D)
HSP701
ARATH
BRAOL
SOLLC
ORYSJ
MARPO
PICAB
CHLRE
HUMAN
MOUSE
0.05

Figure 7.

	Our Arath-EVs data	Our Braol-EVs data	Abundance (Arath)	Published in other Arath AWF-EVs	Published in other PDEVs	Conservation	Already used/ proposed as markers	Membrane localization	Plant-specific	Key Highlights
Fasciclins (FLA8, FLA2, FLA1)	🟢	🟢	🟢	🟢	🟢	🔴	🔴	🟢	🟢	• GPI anchored • High abundance
Heat shock proteins (HS701)	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🔴	🔴	• Classical EV marker • Drawback of lacking plant specificity
Rabs (RAE1C)	🟢	🟢	🟡	🟢	🟢	🟢	🟡	🟢	🔴	• Only Rab identified in both Arath-EVs and Braol-EVs with high % identity. • Present in other Arath AWF-EVs data
V-type proton ATPases (VATB1, VATA)	🟢	🟢	🟡	🟢	🟢	🟢	🔴	🟢	🔴	• Protein complex that shall be present in vesicles ´ membrane
Patellins (PATL1)	🟢	🟢	🟡	🟡	🟡	🔴	🟡	🟡	🟢	• Involved in membrane-trafficking • Plant-specific
Aquaporins (PIP21, PIP12)	🟢	🟢	🟢	🟢	🟡	🟡	🟡	🟢	🔴	• High abundant • Membrane proteins
14-3-3 (14334, 14310)	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🔴	🔴	• Involved in vesicle-trafficking, EV marker on other kingdoms
Germin like proteins (GL33)	🟢	🟢	🟢	🔴	🟡	🔴	🔴	🔴	🟢	• Maybe co-isolated with EVs or protein corona • Plant specific
Tetraspanin (TET8)	🟢	🔴	🟡	🟡	🟡	🟡	🟢	🟢	🔴	• Good marker for A. thaliana EVs • Not so abundant in other PDEVs
Syntaxins (SY121, SY132)	🟢	🔴	🟡	🟢	🟡	🟡	🟡	🟢	🔴	
Clathrins (CLAH2)	🟢	🟢	🟡	🟢	🟡	🟡	🔴	🟡	🔴	• Reproducibly identified on PDEVs • Highly expressed in intracellular vesicles
Phospholipases Ds (PLDA1)	🟢	🟢	🔴	🟢	🟡	🟡	🔴	🟡	🔴	• Low abundant • Not in membrane
Calreticulin (CALR1)	🟢	🟢	🔴	🔴	🔴	🟡	🔴	🔴	🔴	
GADPHs (G3PC1, G3PC2)	🟢	🟢	🟢	🟡	🟢	🟡	🟡	🔴	🔴	• Valid as housekeeping proteins • Lack of specificity for EVs