Proteomic and Biological Profiling of Extracellular Vesicles from Alzheimer's Disease Human Brain Tissues

Satoshi Muraoka[1*], Annina M. DeLeo[1*], Manveen K. Sethi[2*], Kayo Yukawa-Takamatsu[1], Zijian Yang[3], Jina Ko[3], John D. Hogan[4], Zhi Ruan[1], Yang You[1], Yuzhi Wang[1], Maria Medalla[5], Seiko Ikezu[1], Weiming Xia[1,6], Santi Gorantla[7], Howard E. Gendelman[7], David Issadore[3], Joseph Zaia[2], Tsuneya Ikezu[1,8,9#]

[1]Department of Pharmacology and Experimental Therapeutics, Boston University School of Medicine, Boston, MA 02118, USA, [2]Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118, USA, [3]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA, [4]Program in Bioinformatics, Boston University, Boston, MA 02215, USA, [5]Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, MA 02118, USA,[6]Geriatric Research, Education and Clinical Center, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA 01730, USA, [7]Department of Pharmacology and Experimental Neurosciences, University of Nebraska Medical Center, Omaha, NE 68198, USA, [8]Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA, [9]Center for Systems Neuroscience, Boston University, Boston, MA 2215.

[*]These authors contributed equally to this manuscript

[#]**Correspondence author:** Tsuneya Ikezu, MD, PhD

Boston University School of Medicine, 72 East Concord St, L-606B, Boston, MA 02118, USA. Email: tikezu@bu.edu Phone: 617-358-9575 Fax: 617-358-9574

Keywords: **Alzheimer's disease, cortical grey matter, extracellular vesicles, microtubule-associated protein tau, amyloid beta peptide, proteome, machine learning**

Abstract

Introduction: Extracellular vesicles (EVs) from human Alzheimer's disease (AD) biospecimens contain amyloid-β peptide (Aβ) and tau. While AD EVs are known to affect brain disease pathobiology, their biochemical and molecular characterizations remain ill defined.

Methods: EVs were isolated from the cortical grey matter of 20 AD and 18 control brains. Tau and Aβ levels were measured by immunoassay. Differentially expressed EV proteins were assessed by quantitative proteomics and machine learning.

Results: Levels of pS396 tau and Aβ were significantly elevated in AD EVs. High levels of neuron- and glia- specific factors are detected in control and AD EVs, respectively. Machine learning identified ANXA5, VGF, GPM6A and ACTZ in AD EV compared to controls. They distinguished AD EVs from controls in the test sets with 88% accuracy.

Discussion: In addition to Aβ and tau, ANXA5, VGF, GPM6A and ACTZ are new signature proteins in AD EVs.

1. Background

Alzheimer's disease (AD) is the most common forms of dementia affecting nearly 50 million people worldwide. Neuropathologically, disease is characterized by amyloid plaques formed by extracellular aggregation of amyloid beta peptide (Aβ) and intracellular accumulation of neurofibrillary tangles (NFTs). These are formed in brain tissue by the hyperphosphorylated and misfolded microtuble-associated protein tau [1,2]. As AD progresses, Aβ and tau aggregates spread throughout the brain in a spatiotemporal manner [3,4]. Aβ deposition is most prominent in the frontal and temporal lobes, hippocampus and limbic system. Tau pathology, as classified by Braak and Braak, occurs in six histopathological stages. These correspond to tauopathy stages of AD. In stages I and II, NFTs appear in the entorhinal cortex and hippocampus,while in stage III and IV, higher densities extend beyond the entorhinal cortex and hippocampus to the neocortex. In the final V-VI stages, pathological tau deposits are present in throughout the hippocampus [3].

Extracellular vesicles (EVs), including exosomes (50-150nm), ectosomes/microvesicles (150-1000nm), and apoptotic bodies (1000-5000nm) are released from neurons, glia, and various other neural cells into the extracellular space [5-7]. These vesicles contain multiple forms of nucleic acids (microRNA, ncRNA, mRNA, DNA among others), lipids and proteins that are transferred from cell to cell, and found in blood, urine and cerebrosprinal fluid (CSF) [8,9]. In the central nervous system (CNS), it has been reported that AD-associated pathogenic proteins in brain EVs including tau and Aβ oligomers play important roles in AD pathogenesis [9-11][12,13]. Moreover, it has been reported that inhibition of EV synthesis reduced amyloid plaque deposition in the mouse model of AD, and stimulation of EV secretion increased intracellular transfer of prion protein in AD mouse models [14,15]. EVs are involved in the extracellular

enzymatic degradation of Aβ and promote both Aβ aggregation and clearance by microglia [16,17]. Moreover, models of neuron-to-neuron transfer of tau seeds by EVs were reported [12,18-20]. Our own prior work demonstrated that microglia spread tau by EV secretion and that reducing EV synthesis significantly reduces tau propagation [10]. One mechanism centers around Bridging integrator 1 (BIN1), which is associated with the progression of tau pathology and observed by its abilities to alter tau clearance and by promoting the release of tau-enriched microglial EVs [21].

While EVs recovered from human and mouse brain tissues were examined by morphology, proteomics and RNA analyses [22-25], no comprehensive and quantitative proteomics database have yet been acquired for AD human brain tissues. Herein, we provide the first proteomic profiling of EVs isolated from postmortem AD and cognitively impaired control brain tissues . The analyses were combined with machine learning and quantitation of Aβ and tau species by epitope-specific ELISA. Machine learning identified and distinguished protein signatures of AD brain-derived EVs from controls with high degrees of accuracy.

2. Methods

2.1. Brain sample acquisitions

Two cohorts of brains were used in this study. The first cohort was obtained from the University of Nebraska Medical School (11 AD and 9 control) and the Greater Los Angeles Veteran's Affairs Hospital (9 AD and 9 control) as part of NIH NeuroBioBank, which were matched for age and sex. The second cohort was obtained from the NIH NeuroBioBank (22 AD and 18 control), All were matched for age and sex. The Institutional Review Board at the University of

Nebraska Medical School, Greater Los Angeles Veteran's Affairs Hospital and the NIH NeuroBioBank approved the brain acquisitions provided by informed consent.

2.2. Purification of EVs from human brain samples

0.5 g of largely grey matter tissue from the frontal cortex of deceased AD or control cases were processed for EV extraction based on reported method with some modifications [23]. Briefly, frozen brain tissue was chopped on ice using a razor blade (# 12-640 Fischer Scientific) to generate 2-3mm^3 sections. The sections were transferred to 3mL of Hibernate E solution containing 20 units of Papain (# LK003178 Worthington-biochemical corporation) in Earle's Balanced Salt Solution (EBSS) (# 14155063 Gibco) and then incubated in a water bath at 37°C for 15 minutes by stirring once every 5 minutes. After incubation, the samples were immediately place on ice, and 6mL of ice-cold Hibernate E solution (# A1247601 Gibco). The dissociated brain tissue samples were gently homogenized (20 strokes) with a glass-Teflon homogenizer (# 89026-384 VWR). The homogenized tissue samples were filtered with 40 µm mesh filter (# 22-363-547 Fisher scientific). After filtration, the tissue samples were centrifuged at 300 × *g* for 10 min at 4°C (# 5720R Eppendorf). The supernatant from 300 × *g* was transferred to new 15mL tube and then centrifuged at 2,000 × *g* for 10 min at 4°C (# 5720R Eppendorf). The supernatant from 2,000 × *g* was transferred to 30mL conical tube and then centrifuged at 10,000 × *g* for 10 min at 4°C (# Avanti J-E JA25-50 Beckman Coulter). The supernatant from this spin was filtered through a 0.22-µm mesh filter (# SLGP033RS EMD Millipore) into new a polyallomer ultracentrifuge tube with 13.2mL capacity (# 331372 Beckman Coulter) and then was centrifuged at 100,000 × *g* for 70 min at 4°C (# Optima-XE SW41 Beckman Coulter). After ultracentrifugation, the supernatant was completely discarded, and the pellet was resuspended in 2mL of 0.475M of sucrose solution (# S5-3 Fisher science). All sucrose solutions were diluted in

double-filtered PBS (dfPBS) with 0.22-μm pore size (# SLGP033RS EMD Millipore). The sucrose step gradient was created with six 2-mL steps starting from 2.0 M to 1.5M, 1.0M, 0.825M, 0.65M, and 0.475M (containing the resuspended pellet) in a polyallomer ultracentrifuge tube. Each layer was colored with commercially available food coloring to facilitate capture of the EV-rich interphase present between certain steps. The gradient was centrifuged at 200,000 × *g* for 20 h at 4°C (35,000 rpm with # Optima-XE SW41 Beckman Coulter), after centrifugation, the gradient was collected in 2mL fractions, except for the first and last fractions, which were 1mL each. Fraction III corresponded to the 1mL removed from the tube, and particles in this area had a buoyant density of approximately 1.08 g/cm^3. The area between the second (0.65M) and third (0.825M) steps was collected and corresponded to fraction "V", with a buoyant density of 1.10 - 1.12 g/cm^3, while the interphase between the third and fourth steps corresponded to fraction "VI", with a buoyant density of 1.12 - 1.15 g/cm^3. The V and VI fractions were then diluted separately to a total volume of 12mL with dfPBS and centrifuged at 100,000 × *g* for 70 min at 4°C (# Optima-XE SW41 Beckman Coulter). The final pellet from each fraction was resuspended in 30 μL of dfPBS. For all biochemical analyses, these fractions were combined such that an equal volume of each fraction (V and VI) was used. For proteomics, an equal amount of total protein from each fraction (V and VI) was used.

2.3. Protein concentrations

The bicinchoninic acid (BCA) assay was used to determine protein concentration for each sample. The Pierce BCA protein assay kit was used (# 23225 Pierce). Due to the limited amount of sample, EVs were diluted 1:10 before loading into the assay, and a 10:80 ratio of sample to reaction components was used. All assays were allowed to incubate at 60°C for 30 min before protein concentration was read in a Biotek SynergyMix at 562nm. For all protein assays, raw

readings were adjusted by dilution of the sample for the assay, and then by the final dilution of the sample and starting weight of the material. The average concentration was then taken for fractions V and VI.

2.4. Measurement of total tau (t-tau), phosphorylated tau (pT181 tau, pS199 tau, pS396 tau), amyloid beta (Aβ1-40, Aβ1-42) and Annexin V (ANXA5)

EVs were diluted 1:500 in TET buffer (50 mM Tris HCl pH 7.5, 2mM EDTA, 1% Triton X-100) supplemented with Pierce HALT inhibitor (# 78425 ThermoFisher) for t-tau and ANXA5 or TENT buffer (50mM Tris HCl pH 7.5, 2mM EDTA, 150mM NaCl, 1% Triton X-100) supplemented with phosphatase inhibitors (either addition of 0.5mM PMSF, 10mM NaF (# S7920 SIGMA), 200mM glycerol-2-phosphate (# 50020 SIGMA), and 1mM Na_3VO_4 (# S6508 SIGMA) or Pierce HALT inhibitor for pT181 tau, pS196 tau, and pS396 tau. Enzyme-linked immunosorbent assays (ELISAs) were performed to assess levels of t-tau, p-tau, and ANXA5. Commercially available kits from ThermoFisher (t-tau: # KHB0042, pT181: # KHO0631, pS199: # KHO0631, pS396: # KHO0631, ANXA5: # BMS252) were used according to manufacturer's instructions. For Aβ1-40 and Aβ1-42, the levels were measured using the Single Molecule Counting (SMC®) Immunoassay Technology using commercially available kits from Millipore / Sigma (Aβ1-40: # 03-0145-00, Aβ1-42: # 03-0146-00).

2.5. Nanosight Tracking Analysis (NTA)

All samples were diluted in dfPBS at least 1:1000 or more to get particles within the target reading range for the Nanosight 300 machine (Malvern Panalytical Inc), which is 10-100 particles per frame. Using a syringe pump infusion system (Harvard Laboratories/Malvern), five 60-second videos were taken for each sample at 21°C. Analysis of particle counts was carried out in the Nanosight NTA 3.2 software (Malvern Panalytical Inc) with a detection threshold of 5.

Particle counts were normalized for dilution on the machine, dilution of the final pellet, and starting material for EVs extraction. The average count was then taken for fractions V and VI.

2.6. Transmission electron microscopy (TEM)

The EV isolated from AD or control brain tissue were analyzed by TEM. 5µl of the EV sample was adsorbed for 1 min to a carbon-coated mesh grid (# CF400-CU EMS www.emsdiasum.com) that had been made hydrophilic by a 20-sec exposure to a glow discharge (25mA). Excess liquid was removed with a filter paper (# 1 Whatman), the grid was then floated briefly on a drop of water (to wash away phosphate or salt), blotted on a filer paper, and then stained with 0.75% uranyl formate (# 22451 EMS) for 30 sec. After removing the excess uranyl formate with a filter paper, the grids were examined and random fields were photographed using a JEOL 1200EX TEM with an AMT 2k CCD camera.

2.7. Mass spectrometry

2.7.1. Sample preparation

The EV samples were mixed with 2,2,2-Trifluoroethanol (TFE) (# 75-89-8 Millipore) to a final concentration of 50%. The samples were sonicated for 2 min on an ice-water bath (VWR Scientific) and then incubated at 60°C for 2 h. After cool down, the 5mM dithiothreitol (# 3483-12-3 Sigma Aldrich) was added to the samples, and then reduced for 30 min at 60°C. Further, the samples were alkylated with 10mM iodoacetamide (# 163-2109 BioRad) for 30 min at room temperature in the dark. The samples were digested with mass spectrometry grade trypsin (# V5280 Promega) in 50mM ammonium bicarbonate (pH 7.5) for protein digestion (1:30 w/w trypsin-to-protein) for 16h at 37°C. The digested peptides were dried by vacuum centrifugation (# SPD1010 Speedvac system, Thermo Savant). The dried samples were resuspended in 2% acetonitrile/water/0.1% trifluoroacetic acid (TFA) and desalted using C-18 spin columns (#

89870 ThermoScientific); the cleaned peptides were eluted using 60% acetonitrile (ACN) /water/0.1% TFA, dried by vacuum centrifugation (# SPD1010 Speedvac system, Thermo Savant) and further resuspended in 1% ACN/water/0.1% formic acid (FA) and analyzed by nano-liquid chromatography and tandem mass-spectrometry (Nano-LC-MS/MS).

2.7.2. Liquid chromatography (LC)- electrospray ionization (ESI) tandem mass-spectroscopy (MS/MS) Analysis

Nano-LC-MS/MS analysis was conducted by a Q-ExactiveHF mass spectrometer (Thermo-Fisher Scientific) equipped with a nano ultra-performance liquid chromatography (UPLC) (Water Technology) Peptides were trapped on a trapping column (180 µm × 20 mm) and separated on a reversed phased C-18 analytical column (BEH C18, 150 µm × 100 mm) (Waters technology). We loaded 1µL onto the column and separation was achieved using a 75 min gradient of 2 to 98% ACN in 0.1% FA at a flow rate of ~500nL/min. Data-dependent acquisition tandem MS was acquired in the positive ionization mode for the top 20 most abundant precursor ions. The scan sequence began with an MS1 spectrum (Orbitrap; resolution 60,000; mass range 300-2000m/z; automatic gain control (AGC) target 1×10^6; maximum injection time 100 ms). MS2 analysis consisted of higher energy collisional dissociation (Orbitrap; resolution 15,000; AGC 1×10^5; normalized collision energy (NCE) 30; maximum injection time 100 ms, dynamic exclusion time of 8 s).

2.7.3. Sequence database

The raw LC-MS/MS data were converted into mZML format using ProteoWizard msConvert [26]. The data were searched using PeaksDB and PeaksPTM using Peaks Studio version 8.0 (Bioinformatics Solutions, Inc., Waterloo, ON, Canada) against the Uniprot/Swissprot database for Homo sapiens with a 0.1% false discovery rate (FDR) and at least two unique peptides. A 10-

ppm error tolerance for the precursor (MS1) and 0.02 Da mass error tolerance for fragment ions (MS2) were specified. A maximum of 3 missed cleavages per peptide was allowed for the database search, permitting non-tryptic cleavage at one end. Trypsin was specified as the enzyme and carbamidomethylation as a fixed modification. A PeaksPTM search was queued after the peaksDB search, using advanced settings of a larger set of variable modifications, including hydroxylation P, oxidation M, hydroxylation K, hydroxylation-Hex K, hydroxylation-Hex-Hex K, HexNAc ST, HexHexNAc ST, phosphorylation STY, ubiquitination K, deamidation N, methoxy K, and nitrotyrosine Y. The final protein list generated was a combination of peaskDB and peaksPTM searches. The label-free quantification was achieved using PEAKS Studio Quantification-label-free module with a setting of mass error tolerance of 10 ppm and a retention time shift tolerance of 2.0 min. For data filtering for label-free peptide quantification following parameters were used: significance 15, fold change 1, quality 0, average area 1E4, charges from 1-10. The protein quantification following settings was used: significance 0, fold change 1, at least 2 unique peptides, significance method PEAKS.

2.8. Statistical analyses

Statistical analysis was conducted using IBM SPSS software ver.25 and GraphPad Prism6. Between-group comparisons were analyzed by Student's t-test, nonparametric Mann-Whitney U, or one-way ANOVA followed by Bonferroni correction for multiple comparisons. The Gene Ontology of identified proteins were elucidated by DAVID Bioinformatics Resources 6.8. (https://david.ncifcrf.gov). The Venn diagram was generated using Venny_2.1 (http://bioinfogp.cnb.csic.es/tools/venny/).

2.9. Machine learning

The protein biomarkers to distinguish patients with Alzheimer's from controls were selected using Least Absolute Shrinkage and Selection Operator (LASSO) on the proteomics data from the training set (n = 21), where each patient's true state is labeled. An ensemble machine learning classifier to evaluate the performance of the selected proteins was developed. The ensemble machine learning classifier consists of five individual machine learning algorithms to mitigate overfitting, including Linear Discriminant Analysis, Logisitic Regression, Naïve Bayes, Support Vector Machine, and K-Nearest-Neighbours [27,28]. The machine learning generated model's performance was evaluated on a separate, user-blinded test set (n = 17).

3. Results

A total of 38 patient samples, composed of 20 AD (15M, 5F, Mean age: 75.0) and 18 control (14M, 4F, Mean age: 75.7) cases were used in study for EV biological and proteomics analyses. A total of 78 patient samples, composed of 42 AD (24M, 17F, Mean age: 81.0) and 36 control (22M, 15F, Mean age: 79.0) cases, were included for validation analysis **(Table 1)**. There were no statistical differences in the demographics between AD and controls with the exception of postmoterm intervals (PMI) of the validation set, which will be discussed in the validation study **(Figure 4E)**.

3.1. Experimental workflow

The experimental workflow is summarized in **Figure 1A**. The EVs were isolated from 20 AD and 18 sex- and age-matched cognitively unimpaired controls using the discontinuous sucrose gradient ultracentrifugation as previously described with modifications (see Materials and Methods) [23].

3.2. Biochemical characterization of brain-derived EVs

To determine the purity of the EV preparations, NTA was deployed to determine the size of suspended particles based on their Brownian motion. Two cohorts of brain EVs (AD and control) and analyzed their size and number by NTA. The concentration of EVs derived from AD and control brains were not significantly different ($p = 0.6075$). The mode size distribution for EVs was significantly different and peaked at 122 nm for AD and 131 nm for controls ($p = 0.0095$) **(Figure 1B)**. The EVs isolated from frozen brain tissue showed cap-shaped morphology by transmission electron microscopy (TEM) **(Figure 1C)**. To characterize the AD-related proteins in brain-derived EVs, we measured the concentration of total tau (t-tau) and p-tau at threonine 181 (pT181 tau), serine 199 (pS199 tau) and serine 396 (pS396 tau) in lysed EVs by ELISA. The levels of t-tau, pT181 tau, and pS199 tau showed no significant differences between AD and controls (t-tau: $p = 0.398$, pT181 tau: $p = 0.7235$, and pS199 tau: $p = 0.4384$) **(Figure 1D)**. Conversely, pS396 tau was significantly increased in AD-brain derived EVs over controls (pS396 tau: $p = 0.0375$) **(Figure 1D)**. Moreover, we observed a significant increase in both Aβ1-40 and Aβ1-42 in AD-derived EVs over controls (Aβ1-40: $p = 0.0021$ and Aβ1-42: $p < 0.0001$) **(Figure 1E)**.

3.3. Proteomic profiling of brain-derived EVs

We performed a label-free Nano-LC-MS/MS analysis of 38 EV samples for proteomic profiling. Across both cohorts, a total of 1,088 proteins were identified with at least two unique peptides **(Figure 2A** and **Supplementary Table 1 and 2)**. There were 940 proteins identified in control EVs and 1,000 proteins identified in AD EVs. Among them, 852 proteins were detected in both

groups, with 88 proteins unique to the controls and 148 proteins unique to the AD group **(Figure 2A)**. The common, AD-unique and control-unique proteins were tested for properties pertaining to the 'cellular component' and 'pathway' ontologies by Gene Ontology analysis in the Database for Annotation, Visualization and Integrated Discovery (DAVID). Among the 852 shared proteins, 60.9% were found to be included in the extracellular exosome ontology **(Figure 2B)**. The 148 proteins unique to the AD group were linked to mitochondria metabolism known to be dysfunctional in AD brain [29] **(Figure 2B)**. Interestingly, in pathway analysis by DAVID, neurodegenerative disorders, including AD, Parkinson's and Huntington's diseases were enriched in common and unique proteins **(Figure 2C). Figure 2D** shows the peptides identified by Nano-LC-MS/MS, which covered 55.1% of tau (1-441), 9.1% of amyloid-beta precursor protein (APP 1-770), 74.3% of alpha-synuclein (SNCA 1-140) and 6.3% of apolipoprotein E (APOE 1-317). Post-translational modification (PTM) analysis detected six phosphorylation sites (pT181, pS198, pS199, pS202, pT231, and pS404) on tau **(Figure 2D,** red bold letters**)**. Notably, Aβ sequence was identified in APP fragments. **Figure 2E** shows the AD pathway from KEGG pathway analysis based on 68 proteins identified in the AD group, which are designated with red stars. The list of AD pathway-related proteins is provided in **Supplementary Table 3**. Proteins known to play an important role in AD pathogenesis, such as APP, APOE, tau and NADH-ubiquinone oxidoreductase (Cx I-V), were all identified in the AD group, although they were not unique to this group.

3.4. Analysis of label-free quantitative proteomics comparison of AD and control brain-derived EVs

Label-free quantitative proteomics analysis was performed using PEAKS studio software. A total of 949 proteins were quantified **(Figure 3A, Supplementary Table 4)**. The 934 quantified proteins were common between AD and control groups. Between these groups, 3 proteins were uniquely expressed in the controls, while 12 proteins were uniquely expressed in the AD group. The principal component analysis (PCA) showed a marginal separation of the two groups **(Figure 3B)**. **Figure 3C** shows the volcano plot of the common 289 proteins which were detected in more than 50% of the group (AD: $n > 10$ and controls: $n > 9$). Among these proteins, 15 proteins were significantly up-regulated and 3 proteins were significantly down-regulated in AD compared to the control group (as determined by $p < 0.05$ and log2 fold change threshold of >1 or <-1) **(Figure 3C, and Table 2)**. The expression levels of 18 molecules identified in the AD group relative to the control group are displayed in a heatmap **(Figure 3D)**. We next searched for brain cell-type specific molecules within the EV proteomics dataset using the mouse proteomics dataset as a reference [30]. The top 100 ranked cell type-specific molecules, which have at least 5-fold change in concentration in the cell type of interest over the other cell types, were tested with our EV proteomics dataset **(Figure 3E)**. The distribution of these markers indicate that in the human brain, neuron-derived EV represent 49% of EVs, while the other 50% of EVs are derived from glial cell types, including microglia, astrocytes and oligodendrocytes. Moreover, using label-free quantitative value, differences in the expression of cell type-specific marker molecules between AD and controls were seen **(Figure 3F)**. Interestingly, neuron-specific molecules were enriched in the control group (**Figure 3F, blue**), while glia-specific marker molecules were enriched in the AD group (**Figure 3F, red**). These results suggest that glia enhance their generation of EVs in AD brains, which may play a role in spreading AD pathologies.

3.5. Machine learning to identify distinctive AD brain-derived EV proteins

To discover a combination of protein molecules that can accurately distinguish AD EVs from controls, the label-free quantitative proteomics dataset was analyzed using a machine learning method **(Figure 4A).** For this purpose, we split the proteomics dataset into an unblinded training subset (AD: n = 11; controls: n = 10) and a blinded testing subset (AD: n = 9; CTRL: n = 8). The ensemble machine learning model was built using only the data from the training subset, and then the accuracy of the diagnosis was determined using only the blinded testing set. We found that a panel that included annexin-A5 (ANXA5), NGF-induced growth factor (VGF), neuronal membrane glycoprotein M6-a (GPM6A), and alpha-centractin (ACTZ), selected by the LASSO algorithm, resulted in an area under the ROC curve (AUC) of 0.95 within the training set (**Figure 4B** and **Supplementary Table 5).** We then examined the accuracy in the independent blinded test set using the 4 proteins in the dotted green box **(Figure 4B)**. Using this model, we achieved an 88% accuracy (AUC = 0.97) in identifying AD patients using the panel that consisted of ANXA5, VGF, GPM6A, and ACTZ **(Figure 4C).** Further, we ran two control experiments that randomly selected 4 proteins from a total possible 949 proteins to form the diagnosis panel (repeated 20 times, AUC = 0.58, Accuracy = 55%) and shuffling the true labels of the subjects within the training set (AUC = 0.47, Accuracy = 48%). The control's AUC was significantly worse than using the 4-protein panel's AUC ($P<0.001$) **(Figure 4C)**. **Figure 4D** shows the scatter plot of selected 4 proteins from our panel individually, which were each significantly differentially expressed between AD and control groups **(Supplementary Table 6)**. Next, we assessed the correlation of expression of these candidate molecules to the levels of phosphorylated tau (p-tau), Aβ1–40, and Aβ1–42 in EVs by Pearson's correlation analysis.

There was a significantly positive correlation between GPM6A and pS396 tau, and between GPM6A and Aβ1-42 levels (pS396 tau: r = 0.380, *p* = 0.019; Aβ1-42: r = 0.387, *p* = 0.016), and a significantly negative correlation between VGF and Aβ1-42 levels (r = -0.538, *p* = 0.002) **(Figure 4E)**. We also examined the first and second cohorts of AD and control cases to validate ANXA5 expression level by ELISA. ANXA5 expression level was significantly increased in AD brain-derived EVs compared to the control group (*p* = 0.0042, 42 AD and 36 control cases, **Figure 4F**). Although there was statistical difference in PMI between AD and control cases in the validation cohort (**Table 1**), Pearson's correlation analysis of PMI and ANXA5 levels showed no significant correlation (r = -0.165, *p* = 0.149). Thus, this correlation is not due to the difference in PMI between the two groups. Interestingly, ANXA5 expression level shows a tendency to increase along with Braak stages in an AD-dependent manner **(Figure 4F)**. Therefore, ANXA5 is a potential EV molecule for both distinguishing AD and control EVs and as a surrogate marker for Braak stage.

4. Discussion

The biophysical properties of EVs isolated from unfixed postmortem human brain tissues, quantitative analysis of tau and Aβ species, and label-free quantitative proteomic profiling by Nano-LC-MS/MS analyses were performed. pS396 tau, Aβ1-40, and Aβ1-42 levels were significantly increased in AD brain-derived EVs compared to controls. A total of 1,088 unique proteins from brain-derived EVs, were found to be enriched as extracellular exosomes molecules. We also quantified 949 proteins by label-free quantitative proteomic analysis, which were enriched in neuron-specific molecules in controls and glial cell type-specific molecules in the AD group. Differentially expressed EV proteins between AD and control brain samples

included ANXA5, VGF, GPM6A and ACTZ as determined by machine learning approach. Using the validation cohort with the larger sample size, the increased protein level of ANXA5 in the AD group, as compared to controls was confirmed by ELISA.

When characterizing brain-derived EVs isolated from AD and control samples by tau-specific ELISA, the pS396 tau levels were found to be significantly increased in the AD group as compared to the control group. Previous reports indicate pT181 tau to be an early PTM associated with AD, pS199 tau modification is thought to promote tau accumulation, and pS396 tau modification is associated with tau seeding activity and aggregation [31,32]. The enrichment of pS396 in EVs derived from AD brain tissue may indicate their spread and aggregation of tau in the AD brain. We also observed tau fragments from the mid-region (156-406), which is inclusive of proline-rich domains (151-240), and microtubule binding repeat domains (244-369). Tau can be cleaved by various proteases including calpain-1 and 2 (at R230), thrombin, cathepsins, asparagine endopeptidase (at D255, N368), and caspase-2 (at D314) and -3 (at D421) [33-36]. However, many tau fragments identified in AD patients have not been well characterized, and the proteases responsible for their generation have not all been identified. In CSF, 25-35 kDa fragments of tau have been used as an early marker of AD; however, the protease responsible for this cleavage event is unknown [37,38]. A further detailed investigation is needed to determine if tau is truncated by proteases and modified by kinases, and are enriched in AD brain-derived EVs.

In the present study, we observed that the size of EVs derived from AD brain samples was smaller than from control EVs. We propose two possible factors for this finding. First, cholesterol level might be related to EV size; the CNS is known to contain 25% of the body's cholesterol and accumulation of cholesterol is reported to be associated with AD [39]. High

cholesterol levels increase Aβ in cellular and animal models of AD, and the chemical that inhibits cholesterol synthesis reduces Aβ in this model [39-43]. Moreover, it has been reported that small EVs contain significantly higher cholesterol than large EVs [40,41]. Therefore, AD brain cells that accumulate high cholesterol might secrete smaller EVs than controls. The second possibility is that the cell of origin determines EV size. Our study found that human brain-derived EVs were enriched in cell-type specific molecules corresponding to 50% neuronal and 50% glial cell origins. Moreover, the AD brain-derived EVs were enriched with more glial-specific molecules compared to the control EVs. Thus, there is the possibility that the difference in EV size between AD and control groups is due to an increase in smaller, glia-produced EVs among AD patients.

Herein, we report for the first time, the quantitative proteomic analysis of a large-scale brain-derived EV samples isolated from AD and control patients. A total of 59.5 % of the identified 1,088 proteins were part of the extracellular exosome ontology, and similar enrichment rate of EV-specific molecules was previously reported [22]. Interestingly, in AD EVs, glia-specific marker molecules were enriched, while neuron-specific molecules were enriched in controls. Thus, there is a possibility that glia-derived EVs act as a vehicle to transfer pathogenic proteins into neurons in AD brains, which are then spreading and propagating tau throughout the brain.

In summary, we have quantified tau and Aβ levels and investigated proteomic profiles in brain-derived EVs from AD, found enriched glia-specific molecules in AD brain-derived EVs, and identified ANXA5, VGF, GPM6A and ACTZ as potential candidate molecules for monitoring the progression of AD. Zhang *et al.* have recently reported that *ANXA5* is associated with familial late-onset AD by GWAS [44]. Further, Sohma *et al*. have reported a significantly

higher plasma level of ANXA5 in AD patients than in control groups [45]. A number of previous studies in CSF and brain tissue have reported markedly lowered concentration of VGF in AD patient groups compared to controls [46-49]. GPM6A expression level has been shown to be downregulated in the hippocampus of transgenic mice modeling AD [50,51]. It has been reported that GPM6A with palmitoylation is involved in the clustering of lipid rafts, which are themselves enriched in sphingolipids and cholesterol [52]. Thus, GPM6A might be loaded into EVs with high cholesterol levels in AD. In our study, ANXA5 was detected in brain-derived EVs from one brain tissue cohort by commercial ELISA, but VGF, GPM6A, and ACTZ were undetectable. Thus, assays with higher sensitivity detection are needed for the quantification of these molecules. Finally, the combination of cell type-specific molecules from brain cells, including ANAX5, VGF, GPM6A or ACTZ may serve as potential biomarker candidate molecules in AD patient blood samples.

Acknowledgments

The author thank Maria Ericsson (Electron Microscopy Facility, Harvard Medical School) for electron microscopic imaging services, participating laboratories of NIH NeuroBioBank for providing frozen human brain tissue specimens, and Li Wu (University of Nebraska Medical Center) for providing human brain tissue specimens.

Funding: This work is in part funded by Alzheimer's Association AARF-9550302678 (AMD & SM), DVT-14-320835 (TI), BrightFocus Foundation (A2016551S), Cure Alzheimer's Fund (TI), NIH RF1AG054199 (TI), NIH R56AG057469 (TI), NIH R01AG054672 (TI), NIH R21NS104609 (TI).

Conflicts of Interest: TI collaborates with Abbvie Inc. (USA), Aethlone Medical, Inc. (USA), Eisai (Japan/USA) and Ono Pharmaceutical (Japan) and consults Takeda (Japan/USA).

Author Contributions: S.M., A.M.D., and T.I. designed research; S.M., A.M.D., M.K.S., K.Y-T., S.A.D., Z.R., Y.Y., and Y.K.W. performed research; S.M., A.M.D., M.K.S., Z.Y., J.K., J.D.H., W.X., J.Z., and T.I. analyzed data; S.G., and H.E.G. provided brain samples; S.M., and T.I. wrote the paper; S.M., A.M.D., M.K.S., S.I., J.K., Z.Y., M.M., D.I., W.X., J.Z., and T.I. edited the paper.

Supplementary Data

The Supplementary data related to this article can be found at:

References

[1] Goedert M, Wischik CM, Crowther RA, Walker JE, Klug A. Cloning and sequencing of the cDNA encoding a core protein of the paired helical filament of Alzheimer disease: identification as the microtubule-associated protein tau. Pnas 1988;85:4051–5. doi:10.1073/pnas.85.11.4051.

[2] Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. Nat Rev Dis Primers 2015;1:15056. doi:10.1038/nrdp.2015.56.

[3] Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathologica 1991;82:239–59. doi:10.1007/BF00308809.

[4] Thal DR, Rüb U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 2002;58:1791–800. doi:10.1212/wnl.58.12.1791.

[5] Thery C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. J Extracell Vesicles 2018;7:1535750. doi:10.1080/20013078.2018.1535750.

[6] You Y, Ikezu T. Emerging roles of extracellular vesicles in neurodegenerative disorders. Neurobiol Dis 2019;130:104512. doi:10.1016/j.nbd.2019.104512.

[7] Delpech J-C, Herron S, Botros MB, Ikezu T. Neuroimmune Crosstalk through Extracellular Vesicles in Health and Disease. Trends in Neurosciences 2019;42:361–72. doi:10.1016/j.tins.2019.02.007.

[8] Becker A, Thakur BK, Weiss JM, Kim HS, Peinado H, Lyden D. Extracellular Vesicles in Cancer: Cell-to-Cell Mediators of Metastasis. Cancer Cell 2016;30:836–48. doi:10.1016/j.ccell.2016.10.009.

[9] DeLeo AM, Ikezu T. Extracellular Vesicle Biology in Alzheimer's Disease and Related Tauopathy. J Neuroimmune Pharmacol 2018;13:292–308. doi:10.1007/s11481-017-9768-z.

[10] Asai H, Ikezu S, Tsunoda S, Medalla M, Luebke J, Haydar T, et al. Depletion of microglia and inhibition of exosome synthesis halt tau propagation. Nat Neurosci 2015;18:1584–93. doi:10.1038/nn.4132.

[11] Rajendran L, Honsho M, Zahn TR, Keller P, Geiger KD, Verkade P, et al. Alzheimer's disease beta-amyloid peptides are released in association with exosomes. Proc Natl Acad Sci USa 2006;103:11172–7. doi:10.1073/pnas.0603838103.

[12] Wang Y, Balaji V, Kaniyappan S, Krüger L, Irsen S, Tepper K, et al. The release and trans-synaptic transmission of Tau via exosomes. Mol Neurodegener 2017;12:5. doi:10.1186/s13024-016-0143-y.

[13] Sardar Sinha M, Ansell-Schultz A, Civitelli L, Hildesjö C, Larsson M, Lannfelt L, et al. Alzheimer's disease pathology propagation by exosomes containing toxic amyloid-beta oligomers. Acta Neuropathologica 2018;136:41–56. doi:10.1007/s00401-018-1868-1.

[14] Guo BB, Bellingham SA, Hill AF. Stimulating the Release of Exosomes Increases the Intercellular Transfer of Prions. J Biol Chem 2016;291:5128–37. doi:10.1074/jbc.M115.684258.

[15] Dinkins MB, Dasgupta S, Wang G, Zhu G, Bieberich E. Exosome reduction in vivo is associated with lower amyloid plaque load in the 5XFAD mouse model of Alzheimer's disease. Neurobiol Aging 2014;35:1792–800. doi:10.1016/j.neurobiolaging.2014.02.012.

[16] Bulloj A, Leal MC, Xu H, Castaño EM, Morelli L. Insulin-degrading enzyme sorting in exosomes: a secretory pathway for a key brain amyloid-beta degrading protease. J Alzheimers Dis 2010;19:79–95. doi:10.3233/JAD-2010-1206.

[17] Yuyama K, Sun H, Mitsutake S, Igarashi Y. Sphingolipid-modulated exosome secretion promotes clearance of amyloid-β by microglia. J Biol Chem 2012;287:10977–89. doi:10.1074/jbc.M111.324616.

[18] Dujardin S, Bégard S, Caillierez R, Lachaud C, Delattre L, Carrier S, et al. Ectosomes: a new mechanism for non-exosomal secretion of tau protein. PLoS ONE 2014;9:e100760. doi:10.1371/journal.pone.0100760.

[19] Polanco JC, Li C, Durisic N, Sullivan R, Götz J. Exosomes taken up by neurons hijack the endosomal pathway to spread to interconnected neurons. Acta Neuropathol Commun 2018;6:10. doi:10.1186/s40478-018-0514-4.

[20] Polanco JC, Scicluna BJ, Hill AF, Götz J. Extracellular Vesicles Isolated from the Brains of rTg4510 Mice Seed Tau Protein Aggregation in a Threshold-dependent Manner. J Biol Chem 2016;291:12445–66. doi:10.1074/jbc.M115.709485.

[21] Crotti A, Sait HR, McAvoy KM, Estrada K, Ergun A, Szak S, et al. BIN1 favors the spreading of Tau via extracellular vesicles. Sci Rep 2019;9:9477. doi:10.1038/s41598-019-45676-0.

[22] Vella LJ, Scicluna BJ, Cheng L, Bawden EG, Masters CL, Ang C-S, et al. A rigorous method to enrich for exosomes from brain tissue. J Extracell Vesicles 2017;6:1348885. doi:10.1080/20013078.2017.1348885.

[23] Perez-Gonzalez R, Gauthier SA, Kumar A, Levy E. The exosome secretory pathway transports amyloid precursor protein carboxyl-terminal fragments from the cell into the brain extracellular space. J Biol Chem 2012;287:43108–15. doi:10.1074/jbc.M112.404467.

[24] Hurwitz SN, Sun L, Cole KY, Ford CR, Olcese JM, Meckes DG. An optimized method for enrichment of whole brain-derived extracellular vesicles reveals insight into neurodegenerative processes in a mouse model of Alzheimer's disease. J Neurosci Methods 2018;307:210–20. doi:10.1016/j.jneumeth.2018.05.022.

[25] Gallart-Palau X, Serra A, Sze SK. Enrichment of extracellular vesicles from tissues of the central nervous system by PROSPR. Mol Neurodegener 2016;11:41. doi:10.1186/s13024-016-0108-1.

[26] Adusumilli R, Mallick P. Data Conversion with ProteoWizard msConvert. Methods Mol Biol 2017;1550:339–68. doi:10.1007/978-1-4939-6747-6_23.

[27] Ko J, Baldassano SN, Loh P-L, Kording K, Litt B, Issadore D. Machine learning to detect signatures of disease in liquid biopsies - a user's guide. Lab Chip 2018;18:395–405. doi:10.1039/c7lc00955k.

[28] Ko J, Bhagwat N, Yee SS, Ortiz N, Sahmoud A, Black T, et al. Combining Machine Learning and Nanofluidic Technology To Diagnose Pancreatic Cancer Using Exosomes. ACS Nano 2017;11:11182–93. doi:10.1021/acsnano.7b05503.

[29] Saman S, Lee NCY, Inoyo I, Jin J, Li Z, Doyle T, et al. Proteins recruited to exosomes by tau overexpression implicate novel cellular mechanisms linking tau secretion with Alzheimer's disease. J Alzheimers Dis 2014;40 Suppl 1:S47–70. doi:10.3233/JAD-132135.

[30] Sharma K, Schmitt S, Bergner CG, Tyanova S, Kannaiyan N, Manrique-Hoyos N, et al. Cell type- and brain region-resolved mouse brain proteome. Nat Neurosci 2015;18:1819–31. doi:10.1038/nn.4160.

[31] Hanger DP, Anderton BH, Noble W. Tau phosphorylation: the therapeutic challenge for neurodegenerative disease. Trends Mol Med 2009;15:112–9. doi:10.1016/j.molmed.2009.01.003.

[32] Takeda S, Wegmann S, Cho H, DeVos SL, Commins C, Roe AD, et al. Neuronal uptake and propagation of a rare phosphorylated high-molecular-weight tau derived from Alzheimer's disease brain. Nat Commun 2015;6:8490. doi:10.1038/ncomms9490.

[33] Rissman RA, Poon WW, Blurton-Jones M, Oddo S, Torp R, Vitek MP, et al. Caspase-cleavage of tau is an early event in Alzheimer disease tangle pathology. J Clin Invest 2004;114:121–30. doi:10.1172/JCI20640.

[34] Zhang Z, Song M, Liu X, Kang SS, Kwon I-S, Duong DM, et al. Cleavage of tau by asparagine endopeptidase mediates the neurofibrillary pathology in Alzheimer's disease. Nat Med 2014;20:1254–62. doi:10.1038/nm.3700.

[35] Kurbatskaya K, Phillips EC, Croft CL, Dentoni G, Hughes MM, Wade MA, et al. Upregulation of calpain activity precedes tau phosphorylation and loss of synaptic proteins in Alzheimer's disease brain. Acta Neuropathol Commun 2016;4:34. doi:10.1186/s40478-016-0299-2.

[36] Zhao X, Kotilinek LA, Smith B, Hlynialuk C, Zahs K, Ramsden M, et al. Caspase-2 cleavage of tau reversibly impairs memory. Nat Med 2016;22:1268–76. doi:10.1038/nm.4199.

[37] Johnson GV, Seubert P, Cox TM, Motter R, Brown JP, Galasko D. The tau protein in human cerebrospinal fluid in Alzheimer's disease consists of proteolytically derived fragments. J Neurochem 1997;68:430–3. doi:10.1046/j.1471-4159.1997.68010430.x.

[38] Portelius E, Hansson SF, Tran AJ, Zetterberg H, Grognet P, Vanmechelen E, et al. Characterization of tau in cerebrospinal fluid using mass spectrometry. J Proteome Res 2008;7:2114–20. doi:10.1021/pr7008669.

[39] Torres M, Busquets X, Escribá PV. Brain Lipids in the Pathophysiology and Treatment of Alzheimer’s Disease. Update on Dementia, InTech; 2016. doi:10.5772/64757.

[40] Durcin M, Fleury A, Taillebois E, Hilairet G, Krupova Z, Henry C, et al. Characterisation of adipocyte-derived extracellular vesicle subtypes identifies distinct protein and lipid signatures for large and small extracellular vesicles. J Extracell Vesicles 2017;6:1305677. doi:10.1080/20013078.2017.1305677.

[41] Pfrieger FW, Vitale N. Cholesterol and the journey of extracellular vesicles. J Lipid Res 2018;59:2255–61. doi:10.1194/jlr.R084210.

[42] Puglielli L, Tanzi RE, Kovacs DM. Alzheimer's disease: the cholesterol connection. Nat Neurosci 2003;6:345–51. doi:10.1038/nn0403-345.

[43] Evangelisti E, Cecchi C, Cascella R, Sgromo C, Becatti M, Dobson CM, et al. Membrane lipid composition and its physicochemical properties define cell vulnerability to aberrant protein oligomers. Journal of Cell Science 2012;125:2416–27. doi:10.1242/jcs.098434.

[44] Zhang X, Zhu C, Beecham G, Vardarajan BN, Ma Y, Lancour D, et al. A rare missense variant of CASP7 is associated with familial late-onset Alzheimer's disease. Alzheimers Dement 2019;15:441–52. doi:10.1016/j.jalz.2018.10.005.

[45] Sohma H, Imai S-I, Takei N, Honda H, Matsumoto K, Utsumi K, et al. Evaluation of annexin A5 as a biomarker for Alzheimer's disease and dementia with lewy bodies. Front Aging Neurosci 2013;5:15. doi:10.3389/fnagi.2013.00015.

[46] Cocco C, D'Amato F, Noli B, Ledda A, Brancia C, Bongioanni P, et al. Distribution of VGF peptides in the human cortex and their selective changes in Parkinson“s and Alzheimer”s diseases. J Anat 2010;217:683–93. doi:10.1111/j.1469-7580.2010.01309.x.

[47] Llano DA, Devanarayan P, Devanarayan V, Alzheimer’s Disease Neuroimaging Initiative (ADNI). VGF in Cerebrospinal Fluid Combined With Conventional Biomarkers Enhances Prediction of Conversion From MCI to AD. Alzheimer Dis Assoc Disord 2019;Publish Ahead of Print:1. doi:10.1097/WAD.0000000000000328.

[48] Hölttä M, Minthon L, Hansson O, Holmén-Larsson J, Pike I, Ward M, et al. An integrated workflow for multiplex CSF proteomics and peptidomics-identification of candidate cerebrospinal fluid biomarkers of Alzheimer's disease. J Proteome Res 2015;14:654–63. doi:10.1021/pr501076j.

[49] Brinkmalm G, Sjödin S, Simonsen AH, Hasselbalch SG, Zetterberg H, Brinkmalm A, et al. A Parallel Reaction Monitoring Mass Spectrometric Method for Analysis of Potential CSF Biomarkers for Alzheimer's Disease. Proteomics Clin Appl 2018;12:1700131. doi:10.1002/prca.201700131.

[50] Xu P-T, Li Y-J, Qin X-J, Scherzer CR, Xu H, Schmechel DE, et al. Differences in apolipoprotein E3/3 and E4/4 allele-specific gene expression in hippocampus in Alzheimer disease. Neurobiol Dis 2006;21:256–75. doi:10.1016/j.nbd.2005.07.004.

[51] Woo J-M, Park SJ, Kang HI, Kim BG, Shim SB, Jee SW, et al. Characterization of changes in global gene expression in the brain of neuron-specific enolase/human Tau23 transgenic mice in response to overexpression of Tau protein. International Journal of Molecular Medicine 2010;25:667–75. doi:10.3892/ijmm_00000390.

[52] Ito Y, Honda A, Igarashi M. Glycoprotein M6a as a signaling transducer in neuronal lipid rafts. Neuroscience Research 2018;128:19–24. doi:10.1016/j.neures.2017.11.002.

Figure legends

Figure 1. Biophysical and biochemical characterization of EVs isolated from AD and Control brain tissues:

A) Schematic of extracellular vesicle isolation protocol from human frozen brain tissue. Frozen human gray matter cortex was chopped with a razor blade on ice to generate 2-3mm^3 sections. The cut sections are dissociated in papain in Hibernate-E at 37°C for 15min. The tissue was homogenized with glass-Teflon homogenizer. After differential centrifugation, the EVs were purified by Sucrose step gradient Ultracentrifugation for 20 hours and then resuspended in double-filtered PBS. **B)** Left: Particle numbers of brain-derived EV fraction from control (CTRL) or AD by Nanoparticle tracking analysis. AD-brain derived EVs had no difference in number when adjusted for starting material than CTRL EVs (p = 0.6075 by Mann-Whitney test). Right: Particle size of brain-derived EV fraction. Average mode for particle size was smaller in AD brain compared with non-demented control brains (p = 0.0095 by Mann-Whitney test). **C)** TEM image of frozen human brain-derived EVs. EVs purified from AD, or non-demented CTRL brains were resuspended in double-filtered PBS and loaded onto formvar/carbon-coated copper mesh grids. The grids were then rinsed three times and negatively stained with 0.75% uranyl acetate. Scale bar is 100 nm. Left: CTRL, Right: AD. **D)** Comparison of total tau and tau phosphorylated at threonine 181, serine 199, and serine 396 in EVs. Tau is thought to be a major pathological factor in AD, especially phosphorylated, misfolded, and oligomerized tau. ELISA revealed that there is no difference in total tau to total protein (p = 0.398 by Mann-Whitney test). There was no difference in Tau phosphorylated at threonine181 to total tau (p = 0.7235) and at serine 199 to total tau (p = 0.4384). Tau phosphorylated at serine 396 to total tau was significantly different (p = 0.0375). **E)** Comparison of Amyloid beta 1-40 or 1-42 in EVs. ELISA

revealed that there is a significant difference in amyloid beta 1-40 or 1-42 to total protein (p = 0.0259, $p < 0.0001$ by Mann-Whitney test).

Figure 2. Proteomics profiling of EVs isolated from AD and CTRL brains:

A) Venn diagram representing the number of EV proteins differentially identified in CTRL and AD. **B)** Gene Ontology (GO) analysis using DAVID Bioinformatics Resources 6.8. The GO term of Top5 Cellular Component with -log10 (FDR *p*-value). **(C)** The GO term of Top5 Pathway Ontology with -log10 (FDR *p*-value). **D)** Sequence coverage of identified tryptic fragment peptide from Alzheimer's disease-related protein (Tau, APP, SNCA, APOE) by LC-MS/MS analysis. Identified peptides are shown in black bold. Identified phosphorylation sites are indicated in red bold. **E)** KEGG pathway of Alzheimer Disease. The 68 proteins identified in the AD group are highlighted by red stars.

Figure 3. Label-free quantitative proteomics comparison of AD brain-derived EVs and CTRL brain-derived EVs:

A) Venn diagram representing the number of EV proteins differentially expressed in CTRL and AD. **B)** The quantitative proteome profiles obtained by the label-free analysis were used to inform a principal component analysis (PCA). CTRL samples are defined by blue symbols, AD samples by red symbols. **C)** Volcano plot showing a degree of differential expression of EV proteins in AD compared with CTRL. The x-axis indicates log transformed fold change in expression. The y-axis shows log-transformed *p*-values. The grey dot line shows the 0.1 *p*-values and 1- or -1-fold change cutoff. **D)** Heat map representation of the up- and down-regulated proteins in AD. Red shows up-regulated proteins, and Green shows down-regulated proteins. **E)**

Enrichment of brain cell type-specific markers in brain-derived EV proteins. Yellow: Neuron, Purple: Microglia, Green: Astrocytes, Orange: Oligodendrocytes. The parentheses show the number and percentage of identified cell type-specific protein. **F)** Comparison of the cell type-specific protein in AD brain-derived EV and CTRL EV. The red bar shows higher expression in AD. Blue bar indicates higher expression in CTRL.

Figure 4. Machine learning model to identify AD brain-specific EV biomarker molecules: **A)** Workflow for Machine learning approach. The training set is fed into LDA to generate LDA vectors, which are applied to the blinded test set for classification. The predicted molecule are then validated from the validation cohort by a commercial ELISA. **B)** The best performing panel based on the area under the ROC curve using the Lasso algorithm was selected in the training set. The y-axis shows the area under the ROC curve. The x-axis indicates the number of proteins selected by the Lasso algorithm in the training set. The green box shows the selected protein with High AUC for the blinded test set. **C)** The Accuracy for 4 proteins with high AUC chose to generate the label of the test set, and note on their outcome to generate the prediction on the test set. Randomly selected control: Accuracy = 55%, AUC = 0.58., Shuffling control: Accuracy = 47.6%, AUC=0.47. **D)** A scatter plot of log2 (normalized ratio) as measured by proteomics per selected candidate protein in Machine Learning. (ANXA5: p = 4.58E-06, log2 fold change = 1.1, VGF: p = 5.30E-04, log2 fold change = -1.6, GPM6A: p = 5.37E-04, log2 fold change = 0.6, ACTZ: p = 4.44E-03, log2 fold change = -1.5). **E)** Scattered plot of candidate molecules and AD pathogenic molecule in brain-derived EVs. Left: GPM6A and pS396 tau (r = 0.380, p = 0.019). Center: GPM6A and Aβ1-42 (r = 0.387, p = 0.016). Right: VGF and Aβ1-42 (r = -0.538, p = 0.002). **F)** Left: Significant difference in total ANXA5 to total EV protein between AD and

CTRL group by ELISA (p = 0.0042 by Mann-Whitney test) (AD = 42, CTRL = 36). Right: Braak stage-dependent increase in the ANXA5 expression level to total EV protein in AD dependent manner.

Table 1. Patients information

Discovery set

For proteomics	AD (n=20)	Control (n=18)	*p*-value[a]
Age, mean	75.0 ± 9.12	75.7 ± 8.88	0.997
Gender (male, female)	15M, 5F	14M, 4F	
PMI, mean	12.5 ± 8.68	17.8 ± 7.57	0.0675

Validation set

For proteomics	AD (n=42)	Control (n=36)	*p*-value
Age, mean	81.0 ± 11.64	79.0 ± 11.59	0.1533
Gender (male, female)	24M, 17F	22M, 15F	
PMI, mean	9.0 ± 7.25	17.3 ± 8.28	0.0069

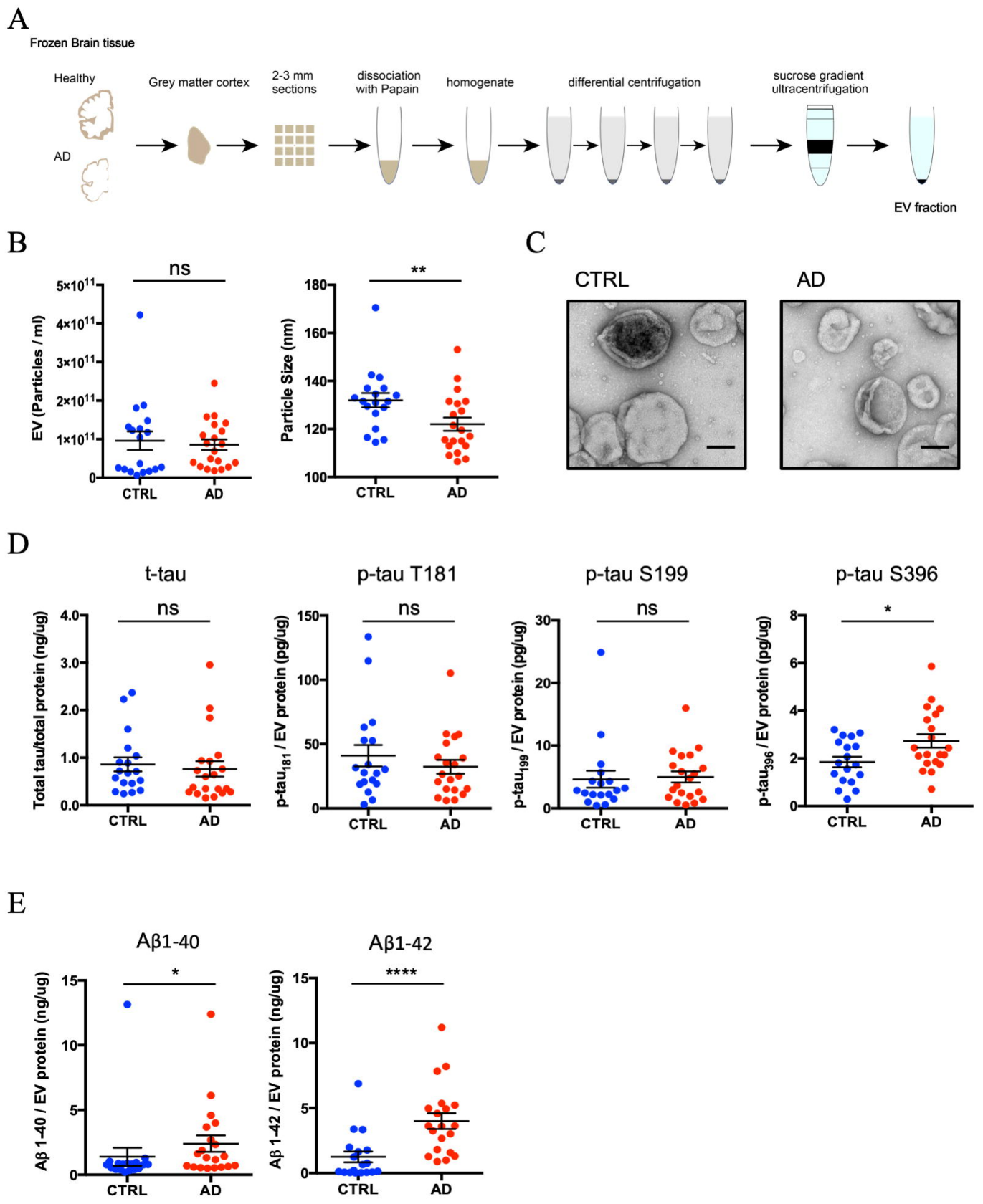
[a] The statistical significance of the differences were calculated using Mann-Whitney test.

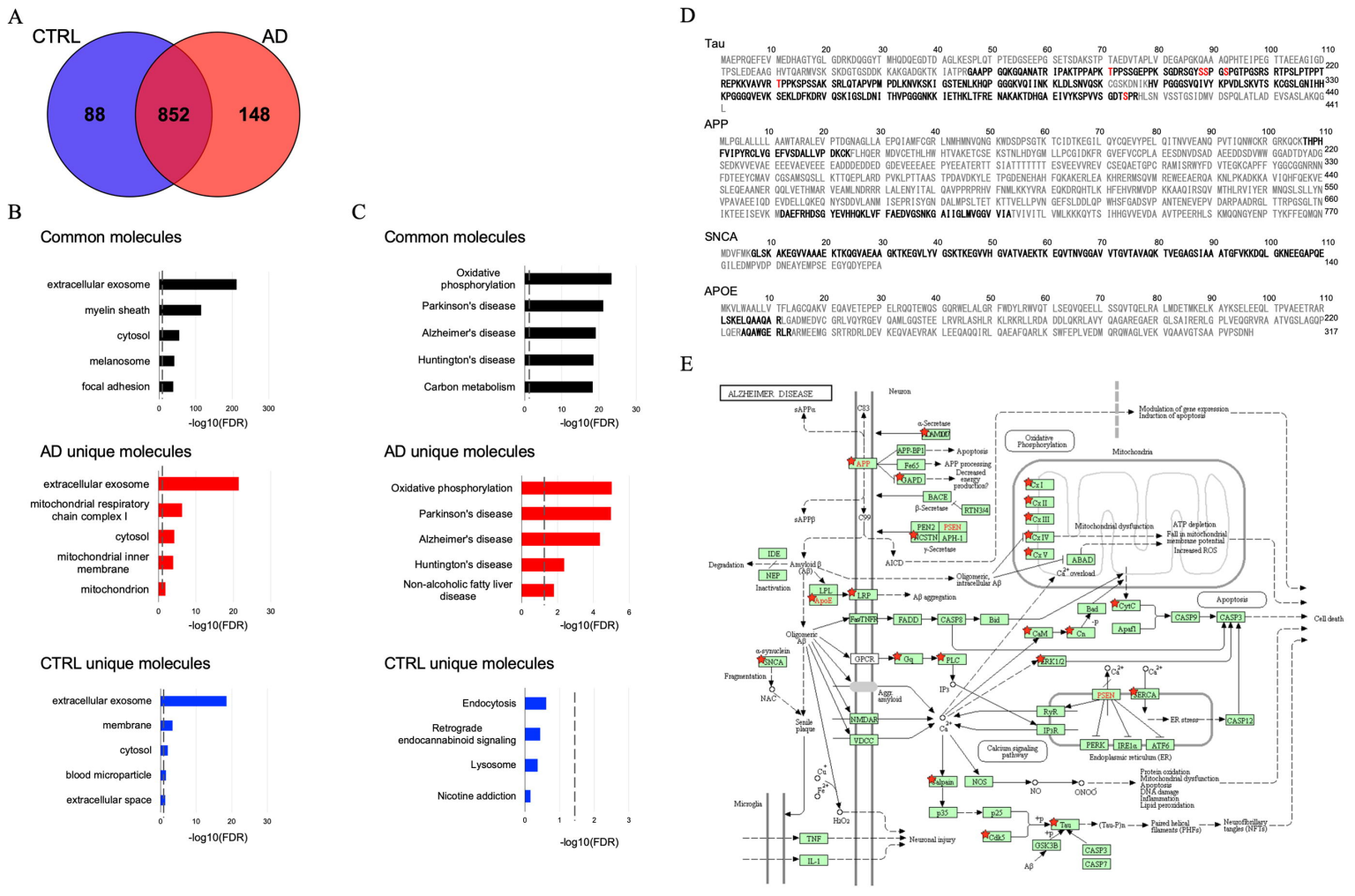
Table 2. Up- and down- regulated protein in expression between AD and CTRL

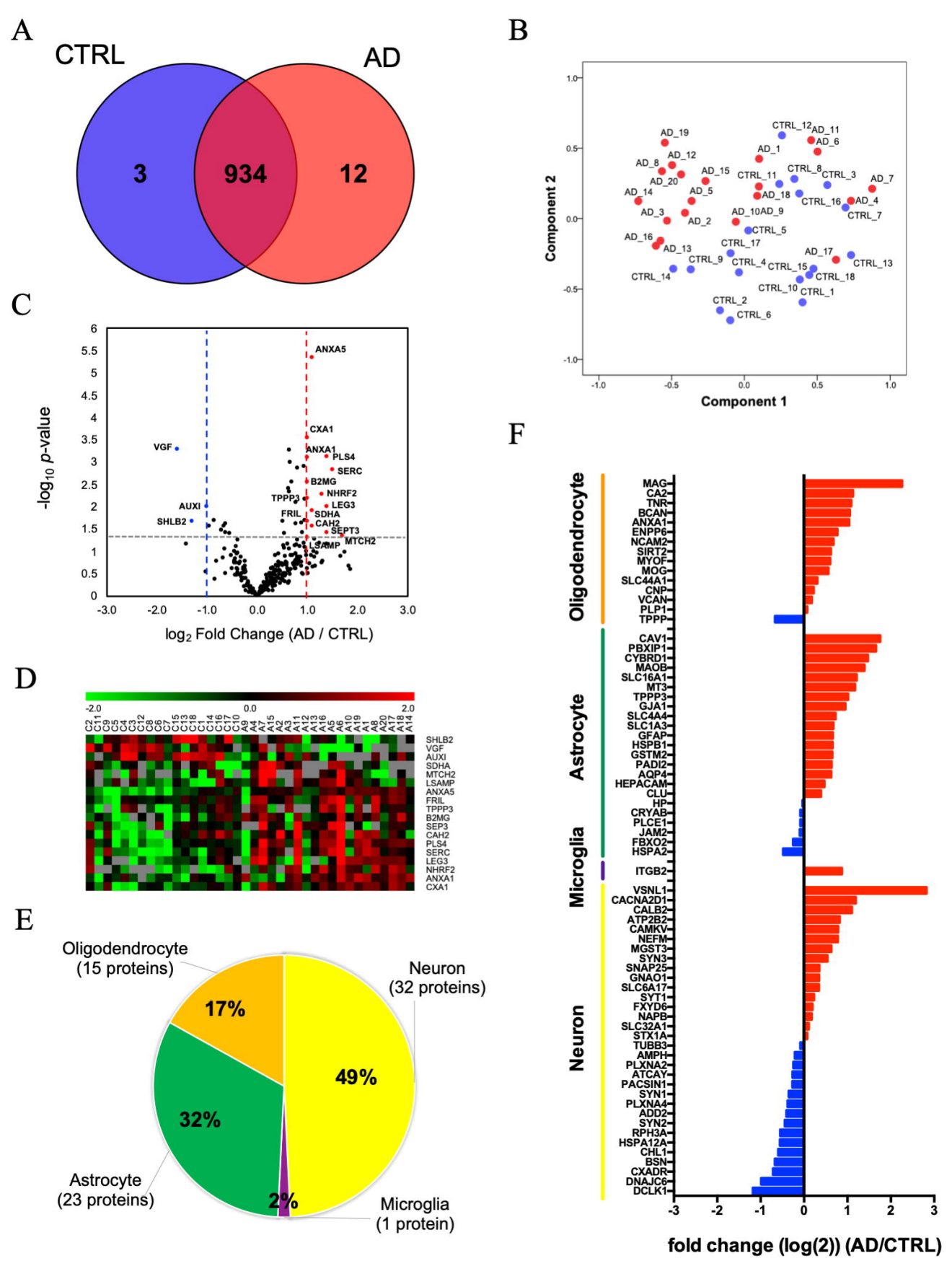
UniProt ID	Protein name	log2 (Fold Change) (AD/ CTRL)[a]	-log10 (*p*-value)[a]
P08758	ANXA5	1.1	5.34
P17302	CXA1	1.0	3.54
Q9NRQ2	PLS4	1.4	3.11
P04083	ANXA1	1.0	3.10
Q9Y617	SERC	1.5	2.83
P61769	B2MG	1.0	2.55
Q15599	NHRF2	1.3	2.27
Q9BW30	TPPP3	1.0	2.18
P17931	LEG3	1.4	2.00
P31040	SDHA	1.1	1.90
P02792	FRIL	1.0	1.67
P00918	CAH2	1.1	1.55
Q9UH03	SEPT3	1.4	1.42
Q9Y6C9	MTCH2	1.7	1.33
Q13449	LSAMP	1.0	1.30
O15240	VGF	-1.6	3.28
Q9NR46	SHLB2	-1.3	1.67
O75061	AUXI	-1.0	2.00

[a]The log2 fold change were calculated by the protein intensity which was analyzed by PEAKS software.
[b]The statistical significance of the differences were calculated using Student's t test.

681


A
Frozen Brain tissue
Healthy
AD
Grey matter cortex
2-3 mm sections
dissociation with Papain
homogenate
differential centrifugation
sucrose gradient ultracentrifugation
EV fraction
B
ns
EV (Particles / ml)
5×10^11
4×10^11
3×10^11
2×10^11
1×10^11
0
CTRL
AD
**
Particle Size (nm)
180
160
140
120
100
C
CTRL
AD
D
t-tau
ns
Total tau/total protein (ng/ug)
4.0
3.0
2.0
1.0
0.0
p-tau T181
ns
p-tau181 / EV protein (pg/ug)
150
100
50
0
p-tau S199
ns
p-tau199 / EV protein (pg/ug)
30
20
10
0
p-tau S396
*
p-tau396 / EV protein (pg/ug)
8
6
4
2
0
E
Aβ1-40
*
Aβ 1-40 / EV protein (ng/ug)
15
10
5
0
Aβ1-42
****
Aβ 1-42 / EV protein (ng/ug)
15
10
5
0


A
CTRL
AD
88
852
148
B
Common molecules
extracellular exosome
myelin sheath
cytosol
melanosome
focal adhesion
-log10(FDR)
AD unique molecules
extracellular exosome
mitochondrial respiratory chain complex I
cytosol
mitochondrial inner membrane
mitochondrion
-log10(FDR)
CTRL unique molecules
extracellular exosome
membrane
cytosol
blood microparticle
extracellular space
-log10(FDR)
C
Common molecules
Oxidative phosphorylation
Parkinson's disease
Alzheimer's disease
Huntington's disease
Carbon metabolism
-log10(FDR)
AD unique molecules
Oxidative phosphorylation
Parkinson's disease
Alzheimer's disease
Huntington's disease
Non-alcoholic fatty liver disease
-log10(FDR)
CTRL unique molecules
Endocytosis
Retrograde endocannabinoid signaling
Lysosome
Nicotine addiction
-log10(FDR)
D
Tau
APP
SNCA
APOE
E
ALZHEIMER DISEASE
Neuron
Mitochondria
Oxidative Phosphorylation
Apoptosis
Cell death
Microglia
Endoplasmic reticulum (ER)
Calcium signaling pathway
Neuronal injury


A
CTRL
AD
3
934
12
B
Component 1
Component 2
C
ANXA5
CXA1
VGF
ANXA1
PLS4
SERC
B2MG
NHRF2
TPPP3
LEG3
AUXI
FRIL
SDHA
SHLB2
CAH2
SEPT3
MTCH2
LSAMP
-log10 p-value
log2 Fold Change (AD / CTRL)
D
SHLB2
VGF
AUXI
SDHA
MTCH2
LSAMP
ANXA5
FRIL
TPPP3
B2MG
SEP3
CAH2
PLS4
SERC
LEG3
NHRF2
ANXA1
CXA1
E
Oligodendrocyte
(15 proteins)
17%
Neuron
(32 proteins)
49%
32%
Astrocyte
(23 proteins)
2%
Microglia
(1 protein)
F
Oligodendrocyte
Astrocyte
Microglia
Neuron
MAG
CA2
TNR
BCAN
ANXA1
ENPP6
NCAM2
SIRT2
MYOF
MOG
SLC44A1
CNP
VCAN
PLP1
TPPP
CAV1
PBXIP1
CYBRD1
MAOB
SLC16A1
MT3
TPPP3
GJA1
SLC4A4
SLC1A3
GFAP
HSPB1
GSTM2
PADI2
AQP4
HEPACAM
CLU
HP
CRYAB
PLCE1
JAM2
FBXO2
HSPA2
ITGB2
VSNL1
CACNA2D1
CALB2
ATP2B2
CAMKV
NEFM
MGST3
SYN3
SNAP25
GNAO1
SLC6A17
SYT1
FXYD6
NAPB
SLC32A1
STX1A
TUBB3
AMPH
PLXNA2
ATCAY
PACSIN1
SYN1
PLXNA4
ADD2
SYN2
RPH3A
HSPA12A
CHL1
BSN
CXADR
DNAJC6
DCLK1
fold change (log(2)) (AD/CTRL)

A

Discovery step (proteomics analysis)

Validation step (ELISA analysis)

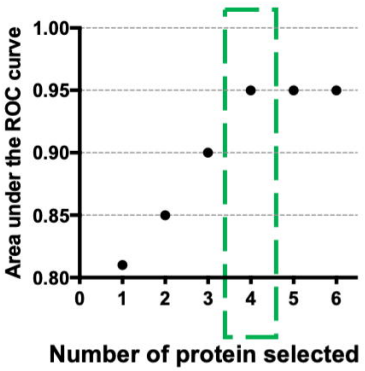
Machine Learning

Development
Training set
10 Control
11 AD

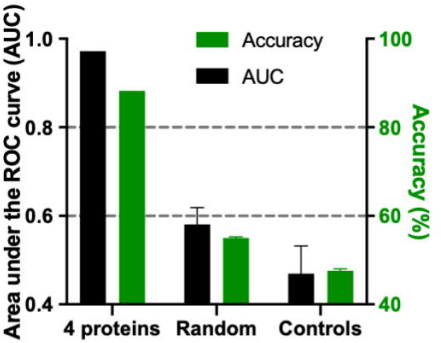
Evaluation
Testing set
8 Control
9 AD

ELISA assay
36 Control
42 AD

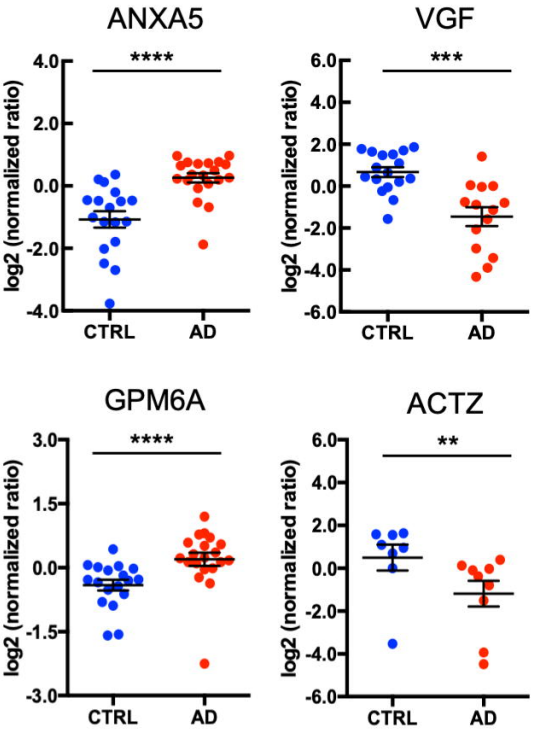
B



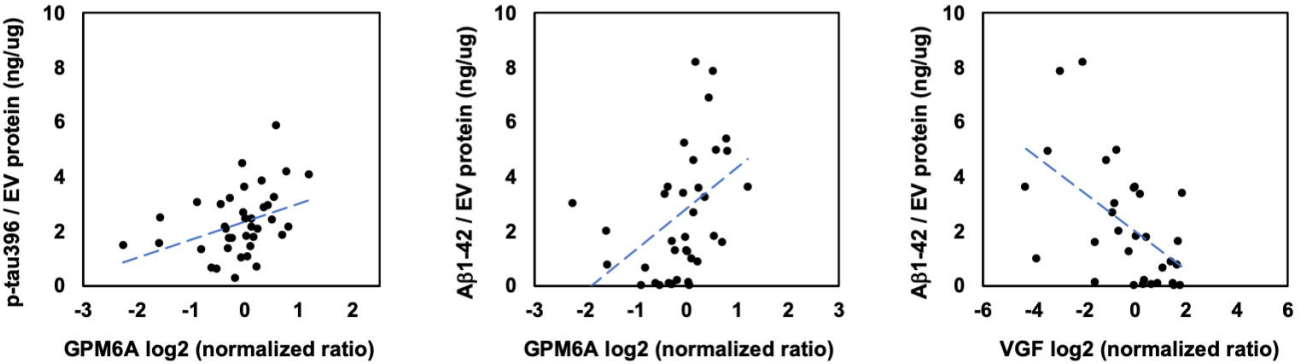
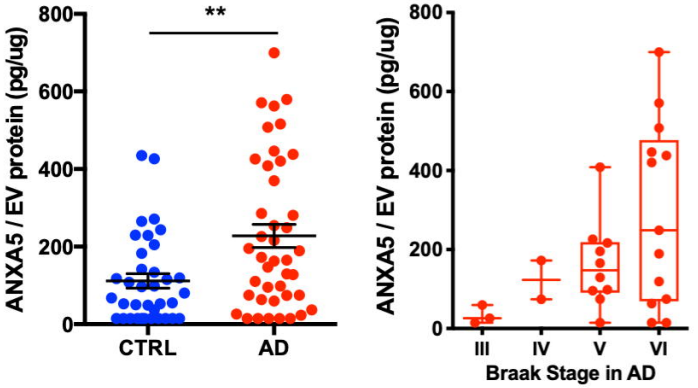
C



D



E



F