



ELSEVIER

Contents lists available at ScienceDirect

Neurobiology of Disease

journal homepage: www.elsevier.com/locate/ynbdi

A multiomic approach to characterize the temporal sequence in Alzheimer's disease-related pathology

Mikael Marttinen^{a,1}, Jussi Paananen^{a,1}, Antonio Neme^{a,1}, Vikram Mitra^b, Mari Takalo^a, Teemu Natunen^a, Kaisa M.A. Paldanius^a, Petra Mäkinen^a, Michael Bremang^b, Mitja I. Kurki^{c,d,e,f}, Tuomas Rauramaa^g, Ville Leinonen^f, Hilikka Soininen^{h,i}, Annakaisa Haapasalo^j, Ian Pike^b, Mikko Hiltunen^{a,*}

^a Institute of Biomedicine, University of Eastern Finland, Kuopio 70210, Finland

^b Proteome Sciences plc, Cobham, London WC1H 9BB, United Kingdom

^c Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA

^d Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

^e Department of Psychiatry, Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA

^f Neurosurgery of Neuro Center, Kuopio University Hospital, Kuopio 70029, Finland

^g Department of Pathology, Kuopio University Hospital and University of Eastern Finland, Kuopio 70029, Finland

^h Neurology of Neuro Center, Kuopio University Hospital, Kuopio 70029, Finland

ⁱ Institute of Clinical Medicine/Neurology, University of Eastern Finland, Kuopio 70210, Finland

^j A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio 70211, Finland

ARTICLE INFO

Keywords:

Alzheimer's disease
Human brain
Transcriptomics
Phosphoproteomics
Systems biology

ABSTRACT

No single-omic approach completely elucidates the multitude of alterations taking place in Alzheimer's disease (AD). Here, we coupled transcriptomic and phosphoproteomic approaches to determine the temporal sequence of changes in mRNA, protein, and phosphopeptide expression levels from human temporal cortical samples, with varying degree of AD-related pathology. This approach highlighted fluctuation in synaptic and mitochondrial function as the earliest pathological events in brain samples with AD-related pathology. Subsequently, increased expression of inflammation and extracellular matrix-associated gene products was observed. Interaction network assembly for the associated gene products, emphasized the complex interplay between these processes and the role of addressing post-translational modifications in the identification of key regulators. Additionally, we evaluate the use of decision trees and random forests in identifying potential biomarkers differentiating individuals with different degree of AD-related pathology. This multiomic and temporal sequence-based approach provides a better understanding of the sequence of events leading to AD.

1. Introduction

Alzheimer's disease (AD) is a common, progressive neurodegenerative disorder, affecting the majority of individuals above the age of 85. To date, no preventative or curative treatments exist for the disease, which carries high socioeconomic costs (Querfurth and LaFerla, 2010). Current biomarkers and drug trials largely revolve around the hypothesis that amyloid-beta (A β) deposition in the brain is a central force driving AD pathogenesis, triggering the accumulation of tau neurofibrillary tangles and alterations in various biological processes (Selkoe

and Hardy, 2016). Interestingly though, abnormal accumulation of A β and tau and other neuropathological changes in AD begin several years before signs of cognitive decline, indicating a need to elucidate the molecular alterations taking place at the different stages of disease (Sperling et al., 2014). A large amount of transcriptomic studies on human samples have been performed to address this, implementing systems-level analyses tools for identifying key Alzheimer's disease molecular pathways (Aibar et al., 2016; Bossers et al., 2010; Zhang et al., 2013). However, correlation coefficients between mRNA and the corresponding protein expression in brain tissue has reached no higher

Abbreviations: AD, Alzheimer's disease; A β , amyloid- β ; CDK5, cyclin-dependent kinase 5; ECM, extracellular matrix; IL8, interleukin 8; PAK, p21-activated kinase 1/3; RF, random forest; v-ATP, vacuolar-ATP

* Corresponding author.

E-mail address: mikko.hiltunen@uef.fi (M. Hiltunen).

¹ These authors contributed equally to this work

<https://doi.org/10.1016/j.nbd.2018.12.009>

Received 23 August 2018; Received in revised form 15 November 2018; Accepted 13 December 2018

Available online 15 December 2018

0969-9961/ © 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

than 0.45, suggesting translational and post-translation control mechanisms for protein abundance (Sharma et al., 2015).

To get a broader systems-level view of biological processes involved in AD pathogenesis, coupling of transcriptomic and proteomic analyses is expected to be beneficial. However, they do not account for post-translational control mechanisms. Thus, direct integration of datasets from only transcriptomic and proteomic experiments would ultimately limit the identification of key signalling events. In AD, the importance of studying defects in post-translational control mechanisms is evident, as abundant hyperphosphorylation of aggregated-tau is a histopathological hallmark of the disease, and is used as an indicator for the progression of AD (Braak et al., 2006; Šimić et al., 2016). High throughput proteomics platforms, such as liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) have been developed, allowing for routine identification and quantification of thousands of proteins and associated phosphorylation levels (Britton et al., 2014; Dammer et al., 2015; Humphrey et al., 2015). Therefore, coupling of transcriptomic and proteomic techniques, and more specifically those that capture phosphorylation level regulation, is essential to obtain a comprehensive systems-level view of AD pathogenesis.

Here, we coupled microarray-based transcriptomics and MS-based phosphoproteomics assay to determine mRNA, protein, and phosphopeptide expression levels from 71 autopsied temporal cortical samples, with varying degree of AD-related neurofibrillary pathology. With computational analysis, we identified disease-related transcript, protein and phosphopeptide expression patterns, associated with distinct biological processes and cell types. Further observation of the obtained gene product expression patterns allowed for generation of a hypothetical temporal sequence of the changes in the activity of specific processes at different stages of AD-related neurofibrillary pathology. Furthermore, we provide an overview of kinase activity alterations in AD through kinase-substrate coupling and measurement of phosphopeptide expression changes in relation to AD-related neurofibrillary pathology. Lastly, we implement decision trees and random forest (RF) analysis to identify potential disease-stage differentiating biomarkers. Collectively, our findings highlight the advantages of a multi-omic approach in the elucidation of complex transcriptional, translational and post-translational alterations associated with the initiation and progression of AD in relation to neurofibrillary pathology.

2. Materials and methods

2.1. Neuropathological cohort

The neuropathological cohort consists of *post-mortem* temporal cortical brain samples from 71 individuals investigated within research projects in Neuro Center, Kuopio University Hospital. At autopsy, assessment of the extent of AD-related neurofibrillary pathology was performed by immunostaining of paraffin sections with the AT8 antibody (Thermo Fisher Scientific, RRID:AB_223647), detecting hyperphosphorylated tau (Braak et al., 2006). The cohort was divided into seven disease severity stages based on Braak staging: Braak stages 0–VI. 60 of the 71 samples underwent microarray analysis and 36 proteomic analysis (Table S1), with a total of 25 samples undergoing both microarray and proteomic analysis (Braak 0: $n = 3$; Braak I: $n = 0$, Braak II: $n = 5$, Braak III: $n = 3$; Braak IV: $n = 3$; Braak V: $n = 6$; Braak VI: $n = 5$). The study was approved by the Ethics Committee of Kuopio University Hospital, University of Eastern Finland, the Finnish National Supervisory Authority, and the Finnish Ministry of Social Affairs and Health (License number 362/06.01.03.01/2017).

2.2. RNA and protein extraction

RNA extraction from the brain tissue samples has been described in detail previously (Natunen et al., 2013). In short, RNA was extracted from frozen brain samples using Qiagen RNeasy Lipid Tissue Mini Kit

(Qiagen), and RNA integrity number (RIN) was determined using 2100 Bioanalyzer (Agilent). For protein extraction, tissue samples were homogenized, sonicated, and suspended in a lysis buffer (SysQuant® lysis buffer (8 M urea (#108487.1000, Merck), 75 mM NaCl (A1371, Applichem), 50 mM Tris (T2788-1L, Sigma), pH 8.2) containing phosphatase (#04906845001, Roche) and protease inhibitors (#04693124001, Roche)), lysed using a Sonifier microtip (Branson) and centrifuged to remove cellular debris. The aqueous supernatant after lysis and centrifugation was used as the protein extract and the protein concentration of each sample was determined using the Bradford protein assay. For quality control purposes, 10 µg (or available amount) of each sample was visualized on Coomassie stained (Imperial Stain, Pierce) SDS-PAGE 4–20% gradient gels (Criterion, Biorad).

2.3. Soluble Aβ42 levels and α-, β-, and γ-secretase activity measurements

Aβ42 and α-, β-, and γ-secretase activity measurements were determined from a fraction of the obtained temporal cortical tissue sample for each patient, as described in Martiskainen et al., 2015 and Natunen et al., 2013. γ-secretase activity was determined from tissue homogenates as previously described [17]. These fractions were not used for proteomic assessment explained later on.

2.4. Microarray data assembly, pre-processing and normalisation

The expression array (Agilent 8x60K Custom Exon array) was performed at the Finnish Microarray and Sequencing Centre in Turku, Finland, and is described in detail in Martiskainen et al., 2015. Microarray analyses were performed with R statistical software version 3.3.3 and Bioconductor version 2.13. Data import was done with limma package version 3.16.7. The probe signal intensities were background-corrected using “normexp” method in limma package. Quantile normalisation was used to bring the array signal intensities to comparable levels between each other. Average expression levels between duplicate probes were used as the signal for the probe. For transcript expression analysis, 3' untranslated region (UTR) probe signal was used as a measure of the respective transcript's global expression level. If multiple such probes targeted the same transcript, average value of the probes was used. The expression level of each individual exon is affected by the global expression level of the transcript, and to adjust for this in the alternative splicing analysis, each exon signal was normalised by the respective transcripts' global expression level (measured by 3'UTR probe). Probes that did not map to any known transcript were excluded, and in case of multiple probes mapping to a single transcript, the probe with the most interquartile-range (IRQ) variation across the samples, was retained in the analysis.

2.5. TMT® labelling

The tissue samples prior to TMT® labelling, 2 mg protein of each sample was reduced (dithiothreitol), alkylated (iodoacetamide), digested (trypsin) to generate peptides, desalted (SepPak tC18 cartridges (Waters, Milford, MA, USA)) and lyophilised, according to SOP. For TMT® labelling, peptides were re-suspended in TEAB/ACN buffer, mixed with their respective TMT® 10plex and incubated for 1 h at room temperature. Individual TMT® reactions were terminated with hydroxylamine and the labelled digests of each of the 10 samples were pooled into the respective TMT® 10plex and incubated for another hour. Finally, the TMT® 10plex analytical sample was acidified, diluted to an acetonitrile concentration < 5%, divided into two aliquots, desalted (as above, each aliquot separately) and lyophilised to completion.

2.6. SCX-HPLC

Following peptide labelling using TMT, samples were aliquoted into 12 × 4-min fractions by strong cation exchange chromatography. Each

of 12 fractions combined into 6 fractions, which had roughly equal peptide amounts. Approximately 100 µg (~5% of total concentration) of peptide was then taken from each of the six fractions to create the non-enriched branch of SysQuant, while the rest were subjected to phosphopeptide enrichment.

2.7. Fe-NTA (IMAC) and TiO₂

The six SCX fractions of the TMT® 10plex were submitted to phosphopeptide enrichment in two parallel arms - Fe-NTA - IMAC (Thermo Scientific Pierce product code 88300) and TiO₂ (Thermo Scientific Pierce product code 88301) - in accordance with manufacturer's instructions in a batch process yielding six fractions for each enrichment matrix. Complementary IMAC and TiO₂ fractions were merged to create a single set of six phosphopeptide enriched fractions per TMT® 10plex. Following phosphopeptide enrichment, peptides in each fraction were purified using graphite spin columns (Thermo Scientific Pierce product code 88302), according to manufacturer's instructions.

2.8. Liquid chromatography mass spectrometry

After peptide fractionation and enrichment there were 12 fractions (6 x non-enriched, 6 x phosphopeptides) derived from each of the three original TMT® 10plex samples. These 12 fractions were each analysed in duplicate by LC-MS/MS using the EASY-nLC 1000 system coupled to an Orbitrap Fusion™ Tribrid™ Mass Spectrometer (both Thermo Scientific) for each of the three TMT 10plex sets. Re-suspended peptides were loaded onto a nanoViper C18 Acclaim PepMap 100 pre-column (Thermo Scientific). This was resolved using an increasing gradient of 0.1% Formic acid in ACN through a 50 cm C18 PepMap RSLC analytical column (Thermo Scientific) at a flow rate of 200 nL/min. Peptide mass spectra were acquired throughout the entire chromatographic run (180 min), using top speed higher collision induced dissociation (HCD) FTMS2 scans at 30000 resolving power @ 400 m/z, following each FTMS scan (120,000 resolving power @ 400 m/z).

2.9. Mass spectrometry data assembly, pre-processing and normalisation

Raw mass spectrometry data was analysed with Proteome Discoverer v1.4 (Thermo Scientific) using the Spectrum Files node. SEQUEST HT node was suitably set up to search data against the human FASTA UniProtKB/Swiss-Prot database (version August 2016). The reporter ions quantifier node was set up to measure the raw intensity values for TMT® 10plex mono-isotopic ions (126, 127 N, 127C, 128 N, 128C, 129 N, 129C, 130 N, 130C, 131). The SEQUEST HT search engine was programmed to search for tryptic peptides (with two missed cleavages) and with static modifications of carbamidomethyl (C), TMT6plex (K), and TMT6plex (N-Term). Dynamic modifications were set to deamidation (N/Q), oxidation (M), and phosphorylation of STY. Precursor mass tolerance was set to 20 ppm and fragment (b and y ions) mass tolerance to 0.02 Da. Spectra were also searched using the PhosphoRS-3.1 (fragment mass tolerance of 20mmu, considering neutral loss peaks for CID and HCD) and Percolator nodes. The PhosphoRS node scored each phosphate group with a local site confidence probability score within each phosphorylated peptide sequence. A score ≥ 75% was considered as an unambiguous site assignment, while scores < 75% were considered ambiguous (Taus et al., 2011). Grouped protein results were filtered at 1% FDR (PSM (peptide spectral matches) level) and 1 x Rank 1 peptide per protein. Protein grouping was performed using Parsimony Principle. In the first stage of data processing the Multi-consensus files for each TMT® 10plex set were filtered to include PSMs that had a signal in at least 1/10 TMT® reporter ion channels. Correction for impurities due to isotopic overlaps of the different reporter ion masses was performed. Next, for every mass spectrometry run, the TMT® reporter ion intensities were normalised using the sum-scaling at the PSM level. Post sum scaling, the ratios of each sample PSM were

related to the reference sample channel. The PSM level ratios to the reference sample were calculated as a log₂ transformation of sample PSM intensity relative to the reference sample PSM intensity: log₂[Sample_{psm}/Reference_{psm}]. PSM level ratios were summarised to peptide level ratios, by averaging PSM level ratios specific to a peptide sequence. To summarize the ratio metric data specific phosphorylation sites, peptides containing ≥ 75% *phosphoRS* score, unique to a gene and uniprot accession number, were averaged using trim mean (with trim factor = 0.2). Peptide features with more than ~ 33% missing quantitative values in a particular Braak stage group were removed from the data set. The remaining missing quantitative values were replaced by values imputed using an iterative PCA method (Josse and Husson, 2016). Following imputation, the data was quantile normalised. To reduce the influence of technical factors, batch effect removal was performed by a linear model in which the TMT® 10plex set (*TMTplex*) and TMT® 10plex channel were used as categorical factors. For protein-level analysis, expression values were computed by averaging (trimmed mean with trim factor = 0.2) peptides that were non-phosphorylated peptides and unique to the gene identifier.

2.10. Clustering expression values

Clustering of the transcripts, proteins and phosphopeptides was performed using the Mfuzz package (version 2.3) (Kumar and Futschik, 2007). Relative expression levels within each Braak stage was calculated by dividing the mean expression level within a stage by the mean expression level through all the stages, and centering the value around zero by log₂-transformation. Optimal number of clusters for different datasets was performed using minimum centroid distance (Schwämmle and Jensen, 2010) for range of 2–20 clusters. Fuzzification parameter was chosen with mestimate-function implementing method proposed by (Schwämmle and Jensen, 2010). Association of clusters to Braak stages and other phenotypic variables were performed using permutation analysis. Variable labels were randomly shuffled 1000 times, and association of expression levels to variables in these randomized sets was analysed by ANOVA (for Braak stages) and linear regression (for other variables). For each gene product, a relative rank within the result list was recorded. The mean rank from the randomized sets within a cluster was compared to observed mean rank within a cluster. This comparison was used to calculate a permuted *P*-value for association of a variable to expression levels from each cluster.

2.11. Cluster enrichment analysis

Enrichment of Gene Ontology terms within clusters was performed using the following Bioconductor packages: KEGG.db, Category and GOstats. Conditional hypergeometric testing was performed to identify enriched biological processes within each cluster. For visualization of the enriched terms, top three terms from each cluster were selected, with a cut-off to exclude biological processes with a *P*.FDR value > 0.05. Both clusters and terms were ordered by hierarchical clustering, and visualized as a heatmap.

2.12. Cell type enrichment

Cell type enrichment was determined by cross-referencing of cluster gene products with lists of genes known to be preferentially expressed in different cell types in mouse brain (Seyfried et al., 2017; Sharma et al., 2015; Zhang et al., 2014). Mouse gene symbols were converted to human gene symbols using R package biomaRt (Durinck et al., 2009). For transcript and protein clusters, all identified transcripts or proteins was used as background, respectively. Cell type specific gene lists were filtered for presence in the transcript and protein datasets prior to cross-referencing. Significance of enrichment of cell type markers with clusters was assessed using one tailed Fisher's exact test.

2.13. Kinase-substrate enrichment analysis

Kinase enrichment analysis for the phosphopeptides of clusters Ph1–6 was performed using the curated kinase-substrate dataset KEA2. Filtering of the input, identification of enrichment, and calculation of significance of enrichment of kinases have been previously described in Lachmann and Ma'ayan, 2009.

2.14. Decision trees and random forest analysis

Decision trees were constructed to determine the gene products capable of assigning individuals to their corresponding Braak stages. This is done in such a way that a binary decision is made based on gene product expression levels, splitting individuals to their corresponding Braak stage. Entropy was used as a splitting metric to determine the best gene products for nodes of the tree. The gene product that minimizes entropy the most is selected as the node of a new branch of the tree. Two branches are generated, and the splitting is applied recursively, considering again all available gene products, until all individuals are correctly assigned to their corresponding Braak stage (Quinlan, 1987). Next, random forests constituting an algorithm developed on top of decision trees were used (Altman and Krzywinski, 2017; Breiman, 2001). Here, an iterative process first randomly modifies the input data by mutating a certain gene product, following by a random selection of a subset of the sample. The subsample is presented as the training data to a decision tree. This process of modification of gene products and subset selection is applied several times, and each time, a new tree is generated. The most relevant gene products are those that are more commonly selected along the individual decision trees. In this contribution, we applied the Python library *scikit* to obtain random forests and decision trees to our data.

2.15. Statistical analysis

Differential expression across the seven Braak stages were analysed using linear regression for linear changes, and with ANOVA for stage specific changes. Association of expression values to linear phenotypes (e.g. secretases and APOE ϵ 4 under additive model) was analysed using linear regression. All *P*-values have been adjusted for multiple testing by Benjamini-Hochberg False Discovery Rate (FDR), and $P\text{-FDR} < 0.05$ was considered as statistically significant (Benjamini and Hochberg, 1995).

2.16. Data availability

The microarray data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar, 2002) and are accessible through GEO Series accession number GSE106241. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD008016 (Vizcaino et al., 2016).

3. Results

3.1. Transcriptomic and phosphoproteomic analysis of human temporal cortical samples

We collected 71 autopsied tissue samples from the temporal cortex of individuals with varying degrees of AD-related neurofibrillary pathology. Based on the extent of AD-related neurofibrillary pathology, individuals were divided into seven Braak stages (Braak 0–VI) (Table S1) (Braak et al., 2006). We have previously shown that β -secretase activity as well as cerebrospinal fluid (CSF) total-tau and p-tau significantly increase with advancing Braak stage in these brain samples (Martiskainen et al., 2015). Concomitantly, a statistically significant decrease in CSF-A β 42 levels was shown. Here, we also show that

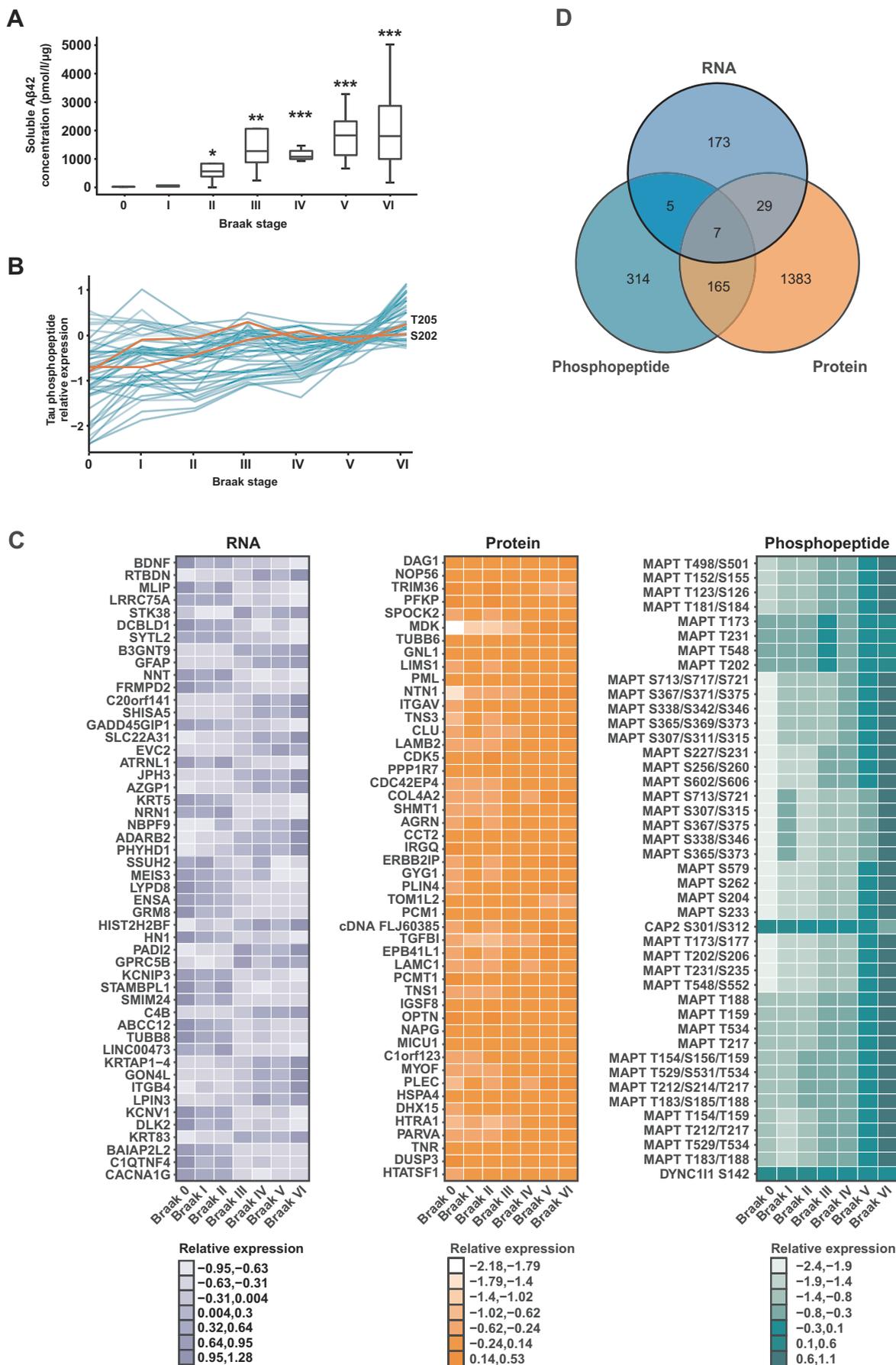
soluble brain A β 42 levels significantly increase with increasing Braak stage (Fig. 1A). Interestingly, no significant alterations in soluble A β 42 levels were observed from Braak stage 0–I. From 60 of these samples, 19,367 unique transcripts were identified by microarray analysis. Additionally, utilizing LC-MS/MS phosphoproteomics, 146,396 peptides mapping to 7516 proteins and 6297 phosphopeptides were identified from 36 samples, giving a comprehensive view of mRNA, protein and phosphopeptide expression levels of our sample set. The mean Pearson's correlation coefficient between mRNA levels and the corresponding protein levels for sample that underwent both microarray and proteomic analysis was 0.10 when collapsed to common gene level (5642 common genes). Linear regression between age and post mortem delay and Braak stage showed no significant associations, and were thus not adjusted for or used as covariates in later analysis.

We initially determined differentially expressed gene products across Braak stages by linear regression, identifying a total of 214 transcripts, 1584 proteins, and 1400 phosphopeptides differentially expressed ($P\text{-FDR} < 0.1$) in relation to Braak stage (Fig. 1C). Of these, 7 genes showed differential expression at the RNA, protein, and phosphopeptides levels, including previously AD-associated genes such as *GFAP*, *PSD-95 (DLG4)*, and *CAMK4* (Fig. 1D). Likewise, 29 genes showed differential expression at the RNA and protein levels, 165 genes between protein and phosphopeptides levels, and 5 between RNA and phosphopeptides levels (Fig. 1D, Table S2). While tau mRNA and protein levels did not show significant changes in their expression in relation to increasing Braak stage (data not shown), we observed that the expression of 135 tau phosphopeptides (including AT8-specific S202 and T205 epitopes) significantly increased from Braak stages 0–VI (Fig. 1B), confirming the use of Braak staging as a method for assessment of disease progression. Additionally, *BDNF* mRNA levels showed a prominent decrease in expression with increasing Braak stage (Linear regression, $P\text{-FDR} = 0.06$) (Supplementary Fig. S1A), which is known to decrease with increased AD severity (Phillips et al., 1991). Furthermore, an increase in complement/inflammatory factors, such as *C4A* and *GFAP* (Linear regression, mRNA: $P\text{-FDR} = 0.08$ and $P\text{-FDR} = 0.06$, Protein: $P\text{-FDR} = 0.02$ and $P\text{-FDR} = 0.003$, respectively) (Supplementary Fig. S1B), was detected in relation to increasing Braak stage, which complies with the well-established increase in complement/inflammatory response accompanied with AD progression. These changes can be considered as positive controls validating the brain sample set and its classification according to Braak staging (Heppner et al., 2015; Zorzetto et al., 2016).

3.2. Determination of expression profiles associated with AD-related neurofibrillary pathology

Given the progressive nature of AD, we implemented fuzzy c-means clustering to identify distinct expression patterns across the seven Braak stages. Optimal number of clusters for each dataset was identified, resulting in nine clusters for transcripts and six for proteins and phosphopeptides (Fig. 2A). Clusters were ranked based on their relevance to Braak stage and denoted by T, Pr, or Ph, for transcriptomic, proteomic and phosphopeptide clusters, respectively. Therefore, for example cluster T1 represents the most neurofibrillary pathology-associated transcriptomic cluster. In total 1000 random permutations were performed to determine the statistical significance for each cluster (Table S3).

Owing to the nine clusters for the transcriptomic data, and six clusters for the proteomic and phosphopeptide data, 10,522 transcripts, 4622 proteins, and 2993 phosphopeptides were not clearly grouped to any cluster. The obtained clusters contain a range of 611–1590, 284–609, and 289–640 genes for mRNA, protein, and phosphopeptide data, respectively (Supplementary Table S3). To identify potential clustering of cell-type specific gene products, we determined the overlap of gene products for each cluster with previously characterized transcriptomes and proteomes of isolated neurons, microglia, and



(caption on next page)

Fig. 1. Differential abundance in A β , tau phosphorylation, transcripts, proteins, and phosphopeptides in relation to increasing AD-related neurofibrillary pathology. A) A β 42 measurements by ELISA from the soluble fraction of human temporal cortical samples show a significant increase in relation to increased AD-related neurofibrillary pathology. ANOVA, post-hoc LSD, $n = 66$. B) To validate the use of Braak staging as a measure for disease severity, tau phosphopeptide levels were investigated. Several tau phosphopeptides, including AT8 specific MAPT Serine 202 and Threonine 205 (highlighted in red), show a significant increase in expression from Braak 0–VI, as expected. Phosphopeptide expression is plotted as mean expression at each Braak stage. Linear regression with FDR correction for multiple comparisons by the Benjamin-Hochberg method, $n = 36$, 138 MAPT phosphopeptides $p < .05$. C) 30 most significantly altered transcripts (P.FDR: 0.06–0.07), proteins (P.FDR: 3.3×10^{-08} – 9.7×10^{-05}), and phosphopeptides (P.FDR: 1.1×10^{-05} – 1.9×10^{-05}) identified in linear regression analysis. mRNA: $n = 60$. Protein and Phosphopeptide: $n = 36$. D) A Venn diagram representing the overlap between differentially expressed gene products (P.FDR < 0.1) for transcripts, proteins, and phosphopeptides. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

astrocytes (Fig. 2B) (Sharma et al., 2015). This revealed a significant enrichment in neuronal markers for clusters T1, Pr2, Ph2, Ph5, and Ph6. Respectively, significant enrichment for microglial markers was observed in clusters T3 and Pr1. Interestingly, enrichment for astrocytic markers was also evident in clusters T3 ($p = .08$) and Pr1 ($p = 4.04 \times 10^{-08}$). Next, we investigated possible associations between the clusters and α -, β -, and γ -secretase activities, soluble A β 42 levels in brain tissue samples, and APOE genotype. Even with the known drastic impact of APOE genotype on the individual susceptibility to AD, there were no clusters that showed clear association with the APOE genotype (Supplementary Fig. S2) (Liu et al. 2013). In turn, cluster T1 showed increased association with β - and γ -secretase activities, and soluble A β 42 levels. However, most clusters showed low association to soluble A β 42 levels or APOE genotype, indicating the need to look beyond these key factors affecting AD pathogenesis or risk and identify their possible upstream and downstream effectors.

3.3. Integration of transcriptomics and phosphoproteomics

To determine the biological processes related to each cluster and to allow identification of the sequential biological stages of AD, we performed enrichment analysis for overrepresented Gene Ontology terms in each cluster, highlighting the top three enriched terms from each cluster (Fig. 2C). Gene products associated with the enriched biological process are listed in Supplementary Table S4 and Supplementary file 1. To visualise similarities between transcript, protein and phosphopeptide enriched Gene Ontology terms, we hierarchically clustered the top three terms of each cluster from all data levels, with a cut-off to exclude biological processes with a P.FDR value > 0.05. This unravelled several shared biological processes between data levels and between clusters inside the data levels.

We found that at early-phases of AD (Braak 0–I), there is a substantial decrease in the transcription of genes involved in biological processes, such as synaptic transmission, cell part morphogenesis, neurogenesis, establishment of synaptic vesicle localization, inorganic ion transmembrane transport, mitochondrial translational initiation/elongation/termination, mitochondrial electron transport (NADH to ubiquinone), and hydrogen ion transmembrane transport (T1, T2) (Fig. 2). Furthermore, transcript cluster T1 is significantly enriched for neuronal markers (Fig. 2B). This initial decrease is followed by a modest increase in expression at Braak stage II and then ultimately a decrease at later stages. Interestingly, proteins associated with cell part morphogenesis, neurogenesis, inorganic ion transmembrane transport, mitochondrial electron transport (NADH to ubiquinone), and hydrogen ion transmembrane transport are enriched at the protein level, sharing a similar expression pattern from Braak stages 0–III (down-up-down) (Pr2). Similar to transcript cluster T1, protein cluster Pr2 was significantly enriched for neuronal markers. Protein cluster Pr3, enriched in proteins related to lipid biosynthesis and phosphatidylinositol phosphorylation, also shows a similar type of down-up-down expression profile as clusters T1/T2, and Pr2 (Fig. 2). Previously, it has been described that phospholipids potentially play a role in synaptic function and learning (Kim et al., 2017; Ohno et al., 2014).

Contrary to the previous profiles, transcription of genes related to Rho protein signal transduction, positive regulation of phosphatidylinositol 3-kinase (PI3K) signalling, and positive regulation of interleukin

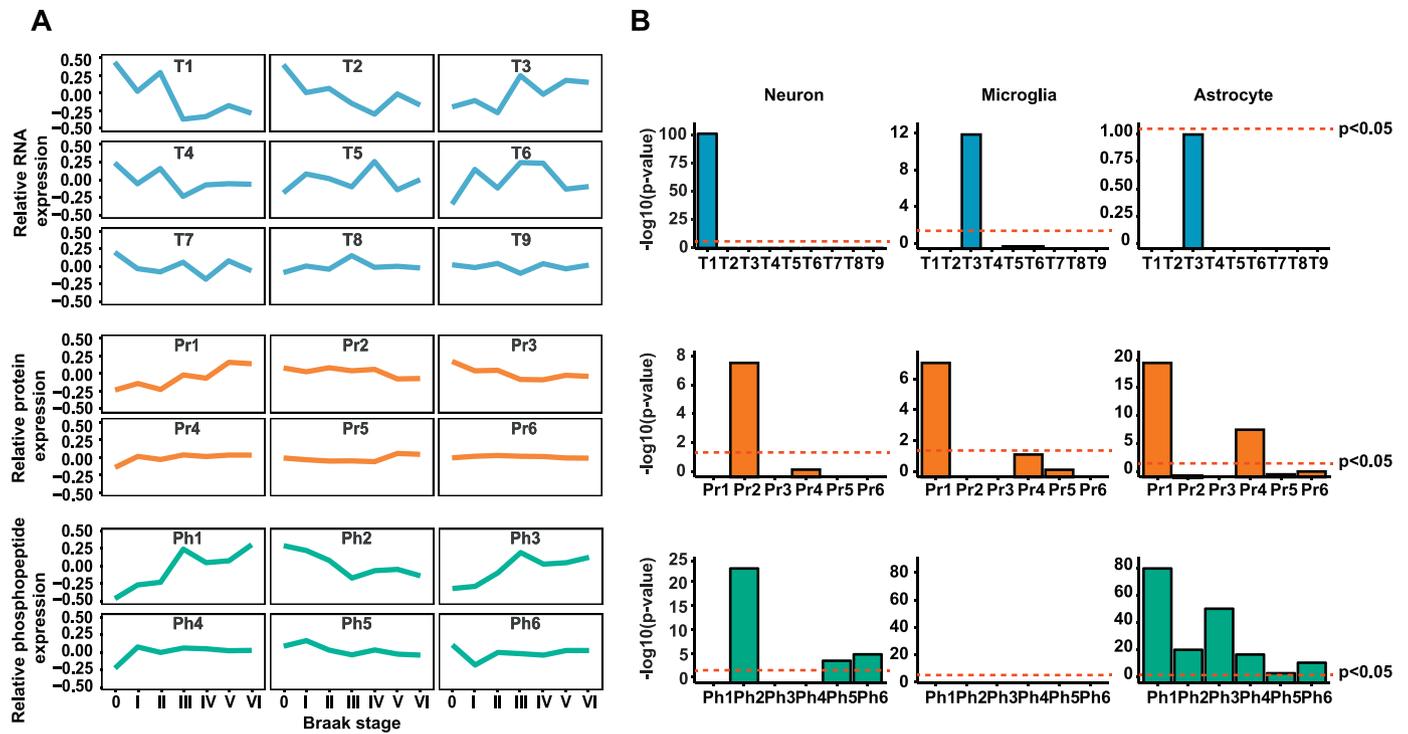
8 (IL8) secretion showed an opposite profile, with a clear increase of expression at Braak stages II–III (T3) (Fig. 2). Concomitant with the increase in the transcription of genes related to inflammation, an increase in the abundance of proteins related to extracellular matrix (ECM) organization, cell-matrix adhesion, and platelet aggregation was apparent (Pr1). Given that clusters T3 and Pr1 are enriched in microglial and astrocytic markers, this could indicate the activation of the innate immune system. Moreover, this may represent a response to or an indication of initial neurodegeneration.

A marked decrease in transcripts related to de novo posttranslational protein folding, RNA biosynthetic process, regulation of nitrogen compounds metabolic process, RNA splicing via transesterification reactions, and nucleic acid-, heterocycle-, and cellular nitrogen compound metabolic process (T2, T4, T7) could be detected already at very early Braak stages (Fig. 2). Conversely, proteins in cluster Pr5, enriched for proteins related to RNA biosynthetic process, regulation of nitrogen compounds metabolic process, RNA splicing via transesterification reactions, and cellular nitrogen compound metabolic process show little expressional changes, with only a slight decrease from Braak stages 0–IV, and an increase in Braak stages V–VI. Changes in RNA transcription and splicing related to AD have not been extensively studied, whereas in other neurodegenerative diseases, such as amyotrophic lateral sclerosis or frontotemporal dementia, RNA transcription and splicing have been identified as a key feature of disease pathogenesis (Bai et al., 2013; Scotti and Swanson, 2015). Taken together, gene products related to RNA transcription splicing arose in several clusters in our analysis, indicating that this could be a significant uncharted area in AD pathology.

3.4. Mapping of the AD-related kinome

To establish a relationship between clustered phosphopeptides (Fig. 2A) and relevant kinases, we performed a kinase enrichment analysis for the phosphopeptides in the Ph1–6 clusters, utilizing the kinase-substrate dataset KEA2 (Fig. 3). From the kinase-substrate dataset we identified 12 significantly enriched kinases; CDK5, GSK3 β , PRKACA, MAPK12, PKN1, CSNK2A2, MAP2K1, CSNK1E, CDK1, MAPK8, CSNK2A1, and MAPK10 (P-FDR < 0.05) (Fig. 3) (Table 1). Based on the phosphopeptide cluster profile, it is possible to delineate whether a kinase has increased or decreased activity in relation of Braak stage.

In line with previous findings, our analyses indicated increased activity of CDK5 and GSK3 β , as a majority of their substrate phosphopeptides showed increased expression in relation to advancing Braak stage (Mazanetz and Fischer, 2007). An increase in activity was indicated also for PRKACA and MAPK12, which have been shown to regulate factors related to inflammation (Shi and Gibson, 2007; Sullivan et al., 2012). Finally, we also observed an increase in PKN1 activity, a kinase that is susceptible to cleavage into truncated active forms by caspase-3 as result of glutamate excitotoxicity (Manser et al., 2008). Interestingly, substrates of PRKACA and MAPK12 were found in clusters Ph1 and Ph3, which show a similar increase in expression levels from Braak stage II onwards, as in inflammation-associated clusters T3 and Pr1. From the kinase-substrate interactions identified, it is also worth highlighting the decreased levels of BIN1 phosphopeptides representing phosphorylation sites S296 and S298 (Table 1). BIN1 has



(caption on next page)

Fig. 2. Gene expression patterns and biological processes associated with AD-related neurofibrillary pathology. A) Clustering of transcripts, proteins, and phosphopeptides identified 9 expression profiles for transcripts, and 6 for proteins and phosphopeptides. Clusters were ranked based on their relevance to Braak stage (1–9 for transcriptomics, 1–6 for proteins and phosphopeptides) and denoted by T, Pr, or Ph, for transcriptomic, proteomic and phosphopeptide clusters, respectively. B) Cell type enrichment was determined by cross-referencing of cluster gene products with lists of genes known to be preferentially expressed in different cell types. Dotted red line indicates P.FDR = 0.05 threshold. Fisher's Exact Test, with FDR correction for multiple comparisons by the Benjamin-Hochberg method. C) Visual comparison of the top 3 enriched biological processes in each cluster at each data level, with a cut-off to exclude biological processes with a P.FDR value > 0.05. Rows represent biological processes and columns represent clusters. Biological processes and clusters are sorted by hierarchical clustering to highlight transcript, protein and phosphopeptide clusters associated with similar processes. Color intensity signifies the significance of enrichment by conditional hypergeometric testing for the biological process in the corresponding cluster with FDR correction for multiple comparisons by the Benjamin-Hochberg method. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

been identified as a highly important susceptibility gene in late-onset AD, but little is known about the functional effects of *BIN1* on AD pathogenesis (Chapuis et al., 2013). *BIN1* has previously been linked to tau hyperphosphorylation, endocytosis/trafficking, inflammation, and calcium homeostasis, but there is no previous research on the effects of changes in *BIN1* phosphorylation, highlighting a valid future research focus (Tan et al., 2013).

To evaluate if the changes in kinase activity can simply be explained by changes in their mRNA or protein levels in relation to Braak stage, we extracted the RNA and protein expression profiles available for the identified kinases (Supplementary Fig. S3). Apart from *CDK5* and *CSNK2A2* showing significant reduction in protein levels in relation to increasing Braak staging (ANOVA, P-FDR = 0.7×10^{-3} and P-FDR = 0.02, respectively), the identified kinases did not show statistically significant alterations in mRNA or protein levels in relation to Braak staging. The increase in *CSNK2A2* and decrease in *CDK5* protein levels from Braak stages 0-III was not an expected finding, as the expression of their substrates showed an inverse expression profile with increasing Braak stage. This inverse relation between kinase protein expression levels and substrate phosphorylation status was especially evident for several *MAPT* phosphopeptides and *CDK5* protein expression levels (Supplementary Fig. S4). Similarly, a large proportion of the identified *CSNK2A2* phosphopeptides showed an inverse correlation with *CSNK2A2* protein expression levels (Supplementary Fig. S5). This suggests that *CDK5* and *CSNK2A2* activity is potentially modulated via specific post-translational mechanisms. This phenomenon is well-described for *CDK5* in AD, where $A\beta$ is suggested to increase the conversion of p35 to p25, which in turn causes prolonged activation of *CDK5* (Lee et al., 2000).

3.5. Molecular interaction networks of pathophysiological changes in relation to AD-related pathology

The clustered expression profiles and cross-comparison of associated biological processes at different data levels suggest that synaptic and mitochondrial dysfunction represent the pathophysiological alterations taking place at the MCI/preclinical phase (Braak 0-II). Subsequently, inflammatory activation is perceived at later Braak stages (Braak II-III), potentially underlining the initiation of

neurodegeneration and transition from MCI to AD/dementia (Fig. 4). To elaborate the potential interplay between the different processes, we mapped out the molecular interactions between both associated gene products (Fig. 5A-B). To do this, we constructed two lists of gene products identified in our clustering analysis associated with (1) cell part morphogenesis, neurogenesis, hydrogen ion transmembrane transport, mitochondrial electron transport (NADH to ubiquinone), and (2) Rho protein signal transduction, positive regulation of PI3K signalling, positive regulation of IL8 secretion, ECM organization, cell-matrix adhesion, and platelet aggregation for which data at both the transcript and protein level was available. By utilizing the molecular interaction database IntAct we constructed two networks, one consisting of synaptic and mitochondria associated gene products (1) (Fig. 5A), and the other consisting of inflammation/ECM associated gene products (2) (Fig. 5B).

This approach highlighted the complex interplay between synaptic and mitochondrial functions. Given the well-established association of *CDK5* with AD, we highlighted an interesting link between *CDK5*, *PAK1*, *DBN1*, *HSP90AA1*, and *ATP6V1A*, and visually integrated the phosphopeptide expression changes found in our data. By doing this, we detected a significant decrease in *PAK1* phosphopeptide expression levels (S155, T185). This is possibly due to altered activity of *CDK5*. Once again, this provides controversial data on the activity of *CDK5*, warranting further research to understand the changes in *CDK5* functions associated with AD. It should be noted that *PAK1* protein levels are significantly decreased potentially explaining the decrease in *PAK1* phosphopeptide expression or alternatively represent a result of aberrant phosphorylation. Mining of the interaction database also revealed an interactor between *PAK1* and *CDK5*; *HSP90AA1*. *HSP90AA1* has been described to regulate the interaction between *CDK5* and p35, altering *CDK5* activity and levels of tau phosphorylation (Luo et al., 2007). Interestingly, *ATP6V1A*, and several other subunits of v-ATPase, were found to interact directly or indirectly with proteins, such as *DBN1* and *AP2M1*. v-ATPase is necessary for several functions relevant to neurodegenerative diseases, such as neurotransmitter priming to synaptic vesicles and regulation of lysosomal acidification (Bagh et al., 2017; Wang and Hiesinger, 2013). Recent findings by Bagh et al., [54] indicate that v-ATPase trafficking to lysosomes is mediated by interaction with *AP2*, and that reduced localization of v-ATPase to lysosomes results in elevated lysosomal pH. Given that autophagy and

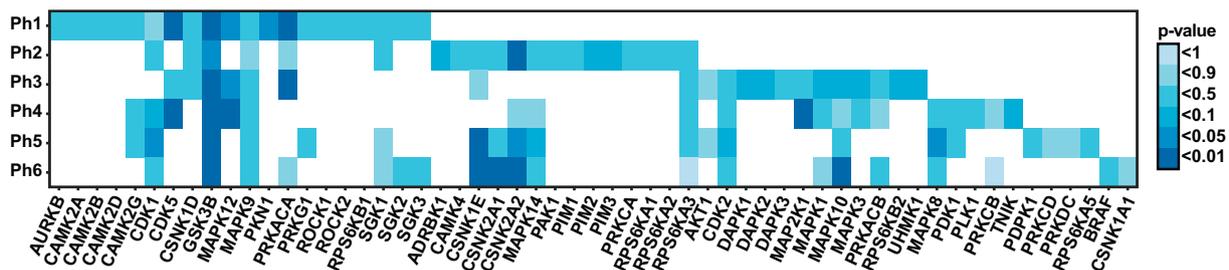


Fig. 3. Kinase enrichment analysis for phosphopeptide clusters. Kinase-substrate enrichment analysis (KSEA) was performed for the clustered phosphopeptides (Ph1–6), utilizing the curated kinase-substrate database KEA2. Rows represent phosphopeptide clusters and columns represent kinases identified in the KSEA. Color intensity signifies the significance of enrichment for the kinase in the corresponding cluster, determined by Fisher's Exact Test, with FDR correction for multiple comparisons by the Benjamin-Hochberg method.

Table 1
Kinase-substrate enrichment analysis of clustered phosphopeptides.

Cluster	Kinase	p-value (FDR)	Genes in Gene Set	Genes in Input	Genes Intersected	Intersecting genes
Ph1	CDK5	3.33E-06	118	17	6	MAPT_T548, MAPT_S713, MAPT_S721, MAPT_S214, MAPT_S156, MAPT_T181
Ph1	GSK3B	1.02E-04	992	17	10	MAPT_T548, MAPT_S713, MAPT_S721, MAPT_S516, MAPT_S184, MAPT_T181, MAPT_S150, MAPT_S141, MAP1B_S1793, SYN3_S483
Ph1	PRKACA	1.07E-04	412	17	7	MAPT_T548, MAPT_S214, MAPT_S156, GFAP_S13, MAPT_S516, MAPT_S141, PPP2R5D_S573
Ph1	MAPK12	0.0157	37	17	2	MAPT_T548, MAPT_T181
Ph1	PKN1	0.0157	39	17	2	MAPT_S214, MAPT_S156
Ph2	CSNK2A2	7.38E-03	412	34	8	HMGAL_S103, VAMP4_S30, STX1A_S14, RPLP2_S105, TMX1_S247, HDGF_S133, HDGF_S132, EEF1B2_S106
Ph2	GSK3B	0.0269	992	34	11	MAPT_T50, MAPT_S46, HNRNPD_S83, RPS3_T221, SRRMI_S393, SRRMI_S389, SPTBN1_S2165, BIN1_S296, BIN1_S298, SRRM2_S353, SRRM2_S351
Ph3	PRKACA	4.12E-03	412	24	7	CAMKK1_S475, ACLY_S455, MAPT_S515, MAPT_S198, MAPT_S140, MAPT_S141, MBP_S146
Ph3	GSK3B	4.40E-03	992	24	10	MAPT_S515, MAPT_S202, MAPT_S144, MAPT_S141, MAPT_S198, MAPT_S140, CCDC6_S244, CTTN_T401, CTTN_S405, KCNAB2_S20
Ph3	MAPK12	0.0471	37	24	2	MAPT_S202, MAPT_S144
Ph4	GSK3B	1.26E-09	992	38	22	MAPT_S721, MAPT_T498, MAPT_S315, MAPT_S346, MAPT_T181, MAPT_T123, NDEL1_S231, SNRNP70_S226, COPA_S173, CTTN_S418, ABLIM2_S294, SPTBN1_S2165, SPTBN1_S2164, SPTBN1_S2169, ADD1_S358, NUDT11_S148, SMARCC2_S347, SORBS1_S945, BIN1_S296, BIN1_S298, BIN1_S303, SNW1_S224
Ph4	CDK5	3.09E-03	118	38	5	MAPT_S721, MAPT_S315, MAPT_S346, NDEL1_S231, MAPT_T181
Ph4	MAP2K1	4.94E-03	37	38	3	CTTN_S418, CTTN_S381, GSK3B_Y216
Ph5	GSK3B	1.32E-04	992	50	19	MAP1B_S1260, ABLIM2_S294, DYNC1L1_S516, PSMA5_S56, SRRMI_S781, STAT1_S727, SPTBN1_S2165, SPTBN1_S2164, SPTBN1_S2138, ADD1_S355, ADD1_S358, STUB1_S23, STUB1_S19, KCNMA1_S924, PRPF4B_S578, PSMD1_T311, PSMD1_S315, RBMX_S208, SRRM2_S954
Ph5	CSNK1E	2.28E-03	186	50	7	PDAP1_S60, NPW1_S125, OSBP_S190, AP3B2_S272, HDAC1_S421, MFAP1_S132, MFAP1_S133
Ph5	CDK1	0.0158	737	50	12	DHX29_S200, SNAP91_S313, MAP1B_S1260, SRRMI_S781, NFI_S864, SPTBN1_S2165, SPTBN1_S2164, STUB1_S23, STUB1_S19, SMARCC2_S283, PRPF4B_S580, PRPF4B_S578
Ph5	CSNK2A2	0.0267	412	50	8	HMGAL_S99, HMGAL_S88, HDAC1_S421, OSBP_S190, SIN3A_S832, AP3B2_S272, MFAP1_S132, MFAP1_S133
Ph5	MAPK8	0.0318	349	50	7	CTNND2_S32, MAP1B_S1260, NFI_S864, DYNC1L1_S194, KCNMA1_S924, TP53BP1_S500, RBMX_S208
Ph6	CSNK2A2	1.31E-04	412	49	12	HMGAL_S103, HMGAL_S102, HMGAL_S92, HMGAL_S91, HDAC1_S421, HDAC1_S423, PPP1R2_S121, PPP1R2_S122, HDAC2_S394, GAP43_S203, RERE_S656, STK4_S320
Ph6	CSNK2A1	1.31E-04	435	49	12	EIF5_S389, EIF5_S390, HMGAL_S103, HMGAL_S102, HMGAL_S92, HMGAL_S91, HDAC1_S421, HDAC1_S423, PPP1R2_S121, PPP1R2_S122, HDAC2_S394, GAP43_S203
Ph6	MAPK10	2.88E-03	153	49	6	ARHGAP12_S240, EFS15L1_S229, ADD1_S12, CRKL_S107, SRRM2_T1003, SRRM2_S398
Ph6	GSK3B	4.14E-03	992	49	15	SFRS14_S277, ARHGAP12_S240, ARHGAP12_T231, SRRMI_S628, SPTBN1_S2165, EFS15L1_S229, ADD1_S12, CRKL_S107, BIN1_S298, BIN1_S303, ADCY6_S54, SRRM2_S2426, SRRM2_S377, CHD3_S1601, PDE4B_S717
Ph6	CSNK1E	4.67E-03	186	49	6	PDAP1_S63, PDAP1_S60, STK4_S320, HDAC2_S394, HDAC1_S423, HDAC1_S421

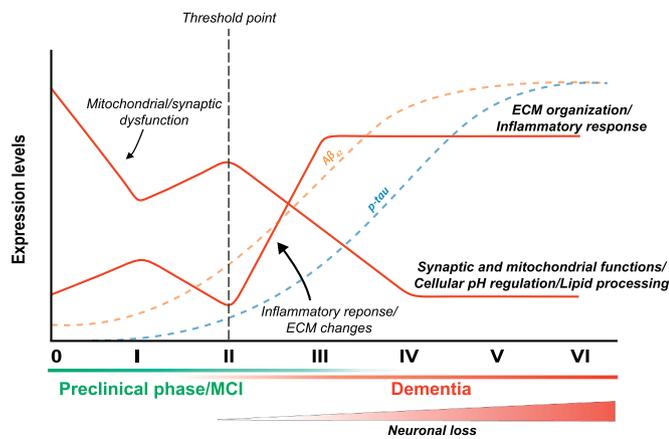


Fig. 4. Hypothetical timeline of the alterations in biological processes in AD. Based on our integrative time-course-based multi-omic approach, we suggest a temporal sequence providing insight on the changes in expression of gene products associated with specific biological processes in relation to AD progression.

lysosome-mediated protein degradation are thought to be impaired in AD, this underscores the potential of these pathways as early-phase intervention targets (Feng et al., 2017; Tian et al., 2013).

The interaction network for gene products associated with inflammation/ECM reveal evident expressional changes at both the transcript and protein level for gene products specifically associated with ECM organization, cell-matrix adhesion, and platelet aggregation. Gene products associated with Rho protein signal transduction, positive regulation of PI3K signalling, and positive regulation of IL8 secretion show moderate expression changes at the transcript level in relation to AD-related neurofibrillary pathology, which are not reflected at the protein level. From the network it can be seen that platelet aggregation associated gene products HSPB1 and FLNA are central, impacting on a large number of downstream targets. Additionally, platelet associated integrin ITGB4 and ITGA6 are found in the network showing significant expressional changes. Recently Donner et al., describe how A β -mediated platelet activation triggers a self-perpetuating cycle resulting in further A β aggregation and thus platelet activation. More specifically, monomeric A β induces integrin-mediated downstream activation of SYK and PLC γ 2, releasing clusterin and adenosine diphosphate from activated platelets (Donner et al., 2016). Taken together, mapping interactions between the identified gene products related to specific biological processes and integration of transcript, protein, and phosphopeptide expression levels provides a wide array of potential biomarker and therapeutic target candidates for further studies.

3.6. Transcript-, protein-, and phosphopeptide-based Braak stage classification

Given the trend of progressive activation of different biological processes with increasing AD-related neurofibrillary pathology, we asked if individuals could be differentiated between Braak stages based on the expression levels of a small set of transcripts, proteins, or phosphopeptides. To assess this, we implemented machine learning to generate decision trees aiming to differentiate individuals between Braak stages. This approach aims to identify a set of genes with similar expression between individuals within a Braak stage, but with differential expression between individuals at different Braak stages, allowing for the differentiation of individuals between Braak stages. In Fig. 6 are represented the decision trees with the best classification success for transcripts, proteins and phosphopeptides. At the transcript level, expression data of 12 transcripts were required to produce a decision tree capable of completely differentiating individuals between Braak stages (Fig. 6A). Equally, at the protein and phosphopeptide

level, expression levels of 8 proteins and 7 phosphopeptides were required, to differentiate individuals between Braak stages (Fig. 6B–C). To further determine the most relevant gene products to differentiate individuals, we performed random forest (RF) analysis. Cross-referencing of the top 30 selected transcripts, proteins, and phosphopeptides in the RF analysis (Supplementary Table S5) to the gene products identified in our clustering approach showed moderate enrichment for protein and phosphopeptides found in clusters Pr1 (8/30) and Pr2 (9/30), and Ph2 (7/30), respectively. Interestingly, the top 30 selected proteins were enriched for proteins associated with the extracellular space and integrin complexes (P-FDR = 0.09), whereas the top 30 phosphopeptides show enrichment for proteins associated with dendrites (P-FDR = 0.051). No clear associations were found for the top 30 transcripts. Taken together, these machine learning approaches provide a set of genes with the potential to differentiate between Braak stages, warranting further validation as potential biomarkers for different stages of AD. Furthermore, they highlight similar functions as found in the clustering approach, emphasizing the potential importance of these biological processes in AD.

4. Discussion

The complexity and progressive nature of AD requires global profiling of alterations in the brain at different stages of disease. Here, we combined gene expression microarray and phosphoproteomic techniques to analyse *post-mortem* brain samples from 71 patients, divided into 7 disease stages based on AD-related neurofibrillary pathology, providing a timeline of changes in biological processes in AD. In summary, our integrative approach included the following aspects: (1) Clustering of distinct gene product expression patterns; (2) Ranking clustered gene product expression patterns based on association to Braak stage; (3) Identifying enriched biological processes for clustered gene products; and (4) Revealing relationships between transcripts, proteins and phosphopeptides. Furthermore, we performed an enrichment analysis for kinases responsible for alterations in the phosphopeptide clusters. Additionally, we utilized machine learning to identify genes capable of classifying individuals to Braak stages based on transcript, protein or phosphopeptides expression levels. An equally comprehensive comparison of various biological levels in human brain in relation to AD-related neurofibrillary pathology has not previously been explored, giving novel insights to the pathomechanisms of AD.

Our approach highlighted that clusters enriched for neuronal markers and gene products related to synaptic- and mitochondrial functions and lipid metabolism are strongly associated with AD-related neurofibrillary pathology at both the mRNA and protein level. Furthermore, observation of the cluster expression profiles revealed similar expressional patterns (down-up-down) at the RNA and protein level, with marked changes taking place at the very earliest Braak stages. This is in line with previous research highlighting these biological processes to be central in AD pathogenesis (Forner et al., 2017; Jones et al., 2015, 2010; Kim et al., 2018; Lin and Beal, 2006). Conversely, clusters enriched for astrocytic and microglial markers showed up-down-up expression profiles for proteins associated with inflammation, ECM organization, and platelet aggregation. Surprisingly, similar biological processes are not enriched between these transcript and protein clusters, which could partially be due to the lack of proteome depths. Nevertheless, the data point to an increase in immune response from Braak stage II onwards. Previous independent network-based analysis of large transcriptomic and proteomic datasets identified microglial gene modules to be strongly associated with neurofibrillary pathology and cognitive decline, proposing a relation between disease progression and inflammatory activation (Seyfried et al., 2017; Zhang et al., 2013). Thus, synaptic and mitochondrial dysfunction may represent pathophysiological alterations at the preclinical phase (Braak 0-II) occurring prior to symptom onset, followed by initiation of inflammation (Braak II-Braak III) and transition to symptomatic AD. To summarize, we

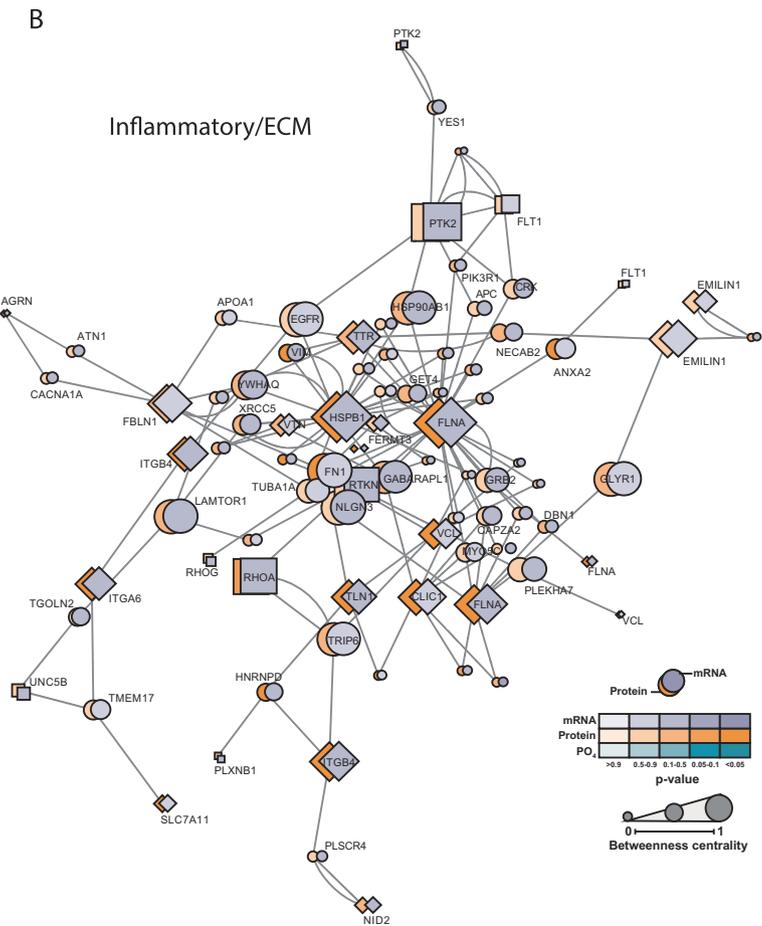
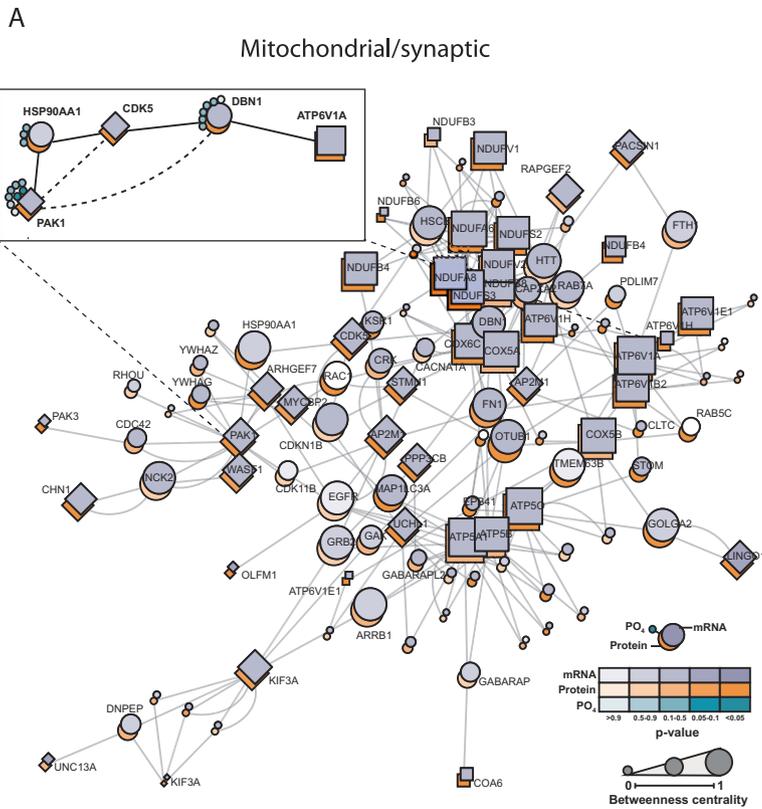
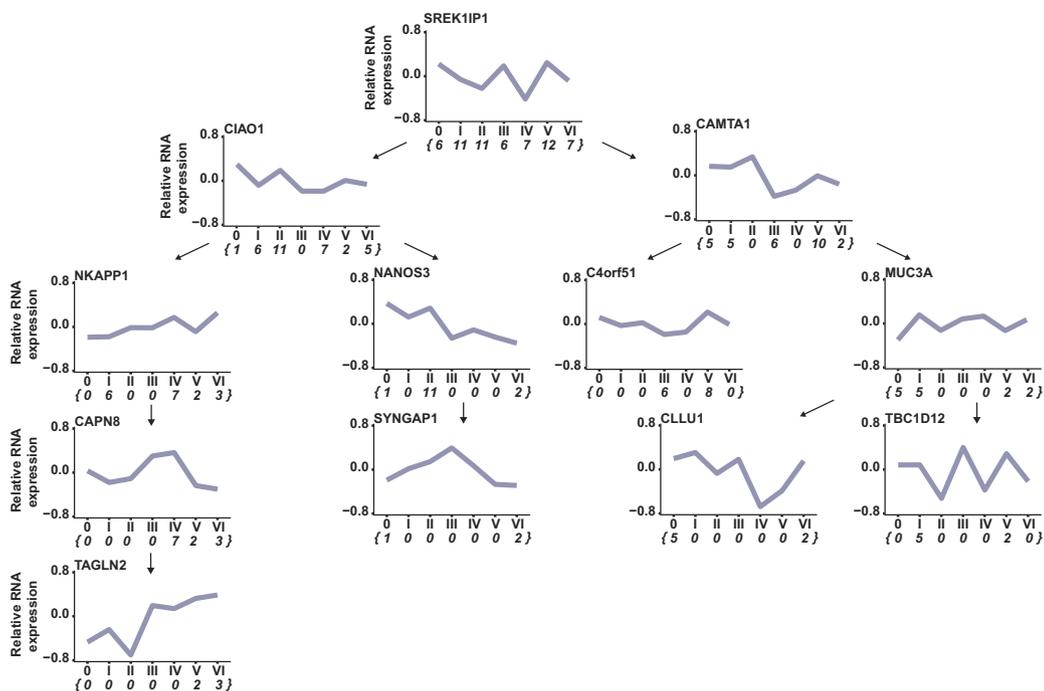
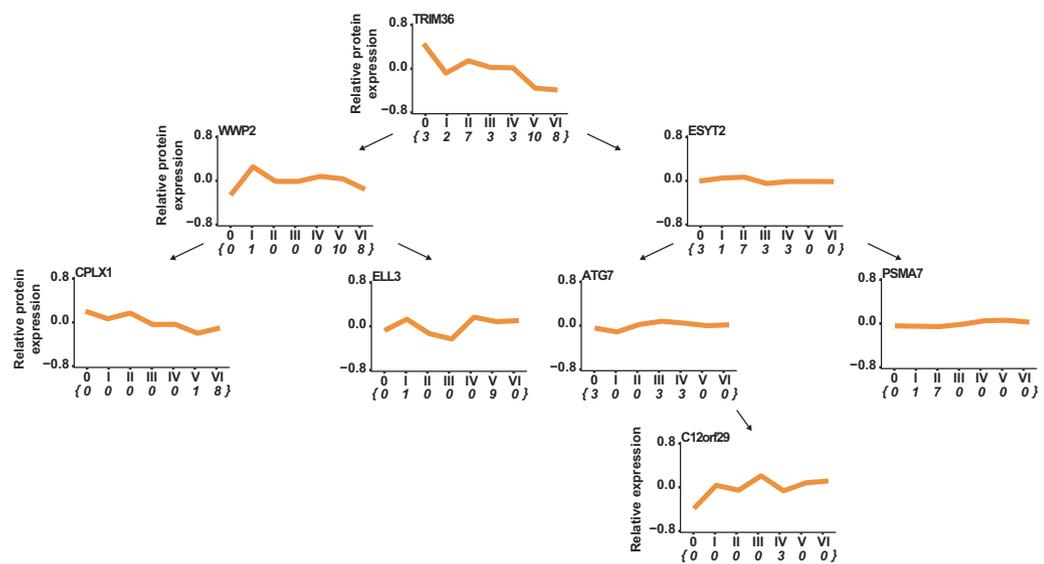


Fig. 5. Protein interaction networks for synaptic- and mitochondria-associated and inflammation- and extracellular matrix-associated gene products. A) Human molecular interactions found in the IntAct database for clustered transcripts and proteins associated with cell part morphogenesis (diamond shape), neurogenesis (diamond shape), hydrogen ion transmembrane transport (square shape), or mitochondrial electron transport, NADH to ubiquinone (square shape). To integrate phosphopeptide data to the network, we highlighted a link between CDK5, PAK1, DBN1, HSP90AA1, and ATP6V1A, and visually integrated phosphopeptide expressional changes as small green circles. Each circle represents a unique phosphopeptide. Color intensity indicates the relevance to Braak stage determined by ANOVA analysis. B) Human molecular interactions found in the IntAct database for clustered transcripts and proteins associated with Rho protein signal transduction (square shape), positive regulation of PI3K signalling (square shape), positive regulation of IL8 secretion (square shape), ECM organization (diamond shape), cell-matrix adhesion (diamond shape), and platelet aggregation (diamond shape). Additional genes curated from the IntAct database, not identified in our clustering analysis, are shown as round shapes. Size of shapes reflects the amount of control that the gene exerts over other genes of the network (*betweenness centrality*). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

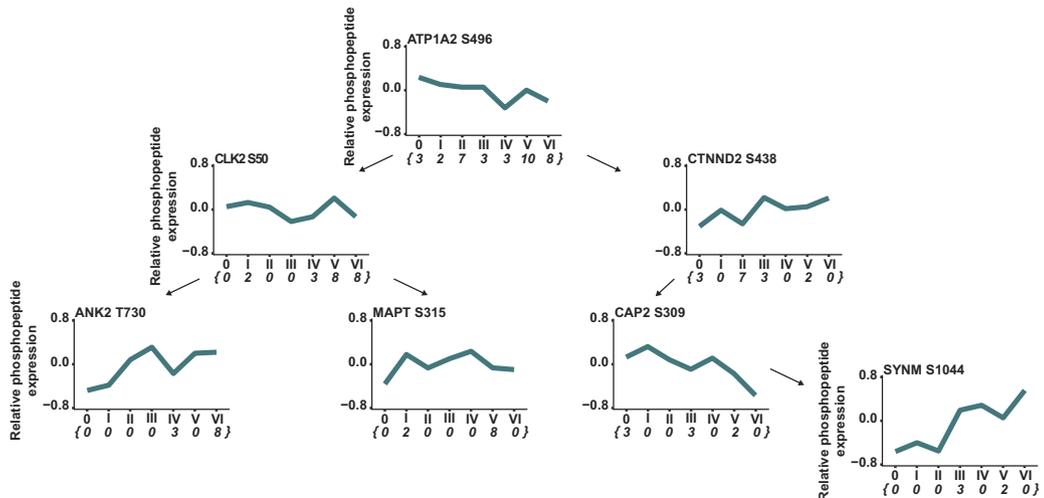
A



B



C



(caption on next page)

Fig. 6. Transcript-, protein-, and phosphopeptide-based decision trees. Expression profiles of 12 transcripts (A), 8 proteins (B), and 7 phosphopeptides (C) required to construct decision trees capable of differentiating individuals at different Braak stages. Y-axis indicates the relative expression of transcripts, proteins, or phosphopeptides. X-axis indicates the Braak stage. Below the X-axis is shown the individuals per Braak stage at the root node and how they are partitioned based on the decision nodes.

provide here a schematic timeline providing an association of specific pathogenic events to different phases in AD pathogenesis (Fig. 4). This highlights Braak stage II as a potential threshold point from where initial changes in synaptic and mitochondrial functions are amplified and inflammatory response and neuronal loss potentially are triggered.

One of the pathological hallmarks of AD is the abnormal phosphorylation of tau, indicating an imbalance in kinase activities (Dolan and Johnson, 2010). By utilizing our phosphopeptide expression data, we mapped kinase activity changes in relation to Braak stage. These analyses highlighted increases in the activity of CDK5, GSK3 β , PRKACA, PKN1 and MAPK12. However, lack of sufficient information on kinase-substrate relations allowed us to only match 7.7% of clustered peptides with their corresponding kinases, still leaving a large part of the AD-related kinome uncharted. Despite lack of availability of comprehensive databases, the available information allowed us to identify novel phosphorylation changes, such as the decreased levels of BIN1 phosphopeptides. Such findings may possibly provide novel candidate targets for the development of new biomarkers or therapeutic approaches. In the attempt to elucidate possible biological processes affected by alterations in phosphorylation, we found out that clustering of phosphopeptides does not provide clusters enriched for specific Gene Ontology terms. Additionally, cell-type enrichment analysis of phosphopeptide clusters revealed a much more heterogenous result than for transcript and protein clusters. This could be in part because a large portion of phosphorylation changes are non-functional and thus cause noise in the data. Furthermore, phosphorylation of one protein may alter the function of an entire signalling pathway, which cannot be detected in our clustering approach (Landry et al., 2009). Therefore, we suggest that the phosphorylation data should be used to complement protein and transcriptomic data and to help understanding shifts in expression changes for specific biological processes, as we show in our network approach.

As our data suggests a stagewise activation and suppression of biological processes in AD, identification of relevant biomarkers to differentiate individuals between Braak stages are essential. Thus, to reduce the complex datasets and extract relevant transcripts, proteins, and phosphopeptides for the differentiation of individuals at different Braak stages, we performed decision tree and RF analysis. The approach showed how a small number of transcripts (12), proteins (8), or phosphopeptides (7) could provide the necessary information to separate individuals at different Braak stages. Furthermore, proteins and phosphopeptides associated with the ECM and dendrites, respectively, were enriched in the RF analysis, emphasizing the importance of further investigating gene products associated with these functions in AD. This highlights the potential of using machine learning applications in conjunction with modern analysis methods to aid in the detection and classification of individuals with complex diseases such as AD. However, results presented here require extensive research to evaluate their biomarker potential.

In conclusion, our data give insights into the relationships between biological processes altered at the different stages of AD-related neurofibrillary pathology, highlighting several interconnected and altered pathways, which warrant further research as potential therapeutic targets. Additionally, utilizing the information of expression changes taking place at different Braak stages might prove useful for better classification of AD patients into different subtypes, allowing for options for patient-specific optimization of treatments in the future. Ultimately, our integrative multi-omic temporal sequence-based approach offers a comprehensive overview of transcriptional, translational and post-translational changes in relation to AD-related

neurofibrillary pathology, providing new insights for biomarker and drug discovery research in AD.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nbd.2018.12.009>.

Acknowledgements

Not applicable.

Competing interests

No financial or non-financial competing interests to declare.

Funding

This study was funded by Academy of Finland (Nos. 307866, 315459, and 288659), Sigrid Jusélius Foundation, the Strategic Neuroscience Funding of the University of Eastern Finland, EADB project in the JPND-CO-FUND program (no 301220), Grant Agreement no 601055, VPH Dementia Research Enabled by IT VPH-DARE@IT, and SynaNet (No 692340).

References

- Aibar, S., Aباigar, M., Campos-Laborie, F.J., Sánchez-Santos, J.M., Hernandez-Rivas, J.M., De Las Rivas, J., 2016. Identification of expression patterns in the progression of disease stages by integration of transcriptomic data. *BMC Bioinformatics* 17, 432. <https://doi.org/10.1186/s12859-016-1290-4>.
- Altman, N., Krzywinski, M., 2017. Points of significance: ensemble methods: bagging and random forests. *Nat. Methods* 14, 933–934. <https://doi.org/10.1038/nmeth.4438>.
- Bagh, M.B., Peng, S., Chandra, G., Zhang, Z., Singh, S.P., Pattabiraman, N., Liu, A., Mukherjee, A.B., 2017. Misrouting of v-ATPase subunit V0a1 dysregulates lysosomal acidification in a neurodegenerative lysosomal storage disease model. *Nat. Commun.* 8, 14612. <https://doi.org/10.1038/ncomms14612>.
- Bai, B., Hales, C.M., Chen, P.-C., Gozal, Y., Dammer, E.B., Fritz, J.J., Wang, X., Xia, Q., Duong, D.M., Street, C., Cantero, G., Cheng, D., Jones, D.R., Wu, Z., Li, Y., Diner, I., Heilman, C.J., Rees, H.D., Wu, H., Lin, L., Szulwach, K.E., Gearing, M., Mufson, E.J., Bennett, D. a., Montine, T.J., Seyfried, N.T., Wingo, T.S., Sun, Y.E., Jin, P., Hanfelt, J., Willcock, D.M., Levey, A., Lah, J.J., Peng, J., 2013. U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* 110, 6–11. <https://doi.org/10.1073/pnas.1310249110>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300. <https://doi.org/10.2307/2346101>.
- Bossers, K., Wirz, K.T.S., Meerhoff, G.F., Essing, A.H.W., Van Dongen, J.W., Houba, P., Kruse, C.G., Verhaagen, J., Swaab, D.F., 2010. Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease. *Brain* 133, 3699–3723. <https://doi.org/10.1093/brain/awq258>.
- Braak, H., Alafuzoff, I., Arzberger, T., Kretschmar, H., Tredici, K., 2006. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol.* 112, 389–404. <https://doi.org/10.1007/s00401-006-0127-z>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Britton, D., Zen, Y., Quaglia, A., Selzer, S., Mitra, V., Löbner, C., Jung, S., Böhm, G., Schmid, P., Prefot, P., Hoehle, C., Koncarevic, S., Gee, J., Nicholson, R., Ward, M., Castellano, L., Stebbing, J., Zucht, H.D., Sarker, D., Heaton, N., Pike, I., 2014. Quantification of pancreatic cancer proteome and phosphorylome: Indicates molecular events likely contributing to cancer and activity of drug targets. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0090948>.
- Chapuis, J., Hansmannel, F., Gistelinc, M., Mounier, A., Van Cauwenbergh, C., Kolen, K.V., Geller, F., Sottejeau, Y., Harold, D., Dourlen, P., Grenier-Boley, B., Kamatani, Y., Delepine, B., Demiautte, F., Zelenika, D., Zommer, N., Hamdane, M., Bellenguez, C., Dartigues, J.-F., Hauw, J.-J., Letronne, F., Ayrat, A.-M., Slegers, K., Schellens, A., Broeck, L.V., Engelborghs, S., De Deyn, P.P., Vandenberghe, R., O'Donovan, M., Owen, M., Epelbaum, J., Mercken, M., Karran, E., Bantscheff, M., Drewes, G., Joberty, G., Campion, D., Octave, J.-N., Berr, C., Lathrop, M., Callaerts, P., Mann, D., Williams, J., Buée, L., Dewachter, L., Van Broeckhoven, C., Amouyel, P., Moechars, D., Deraut, B., Lambert, J.-C., 2013. Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. *Mol. Psychiatry* 18, 1225–1234. <https://doi.org/10.1038/mp.2013.1>.

- Dammer, E.B., Lee, A.K., Duong, D.M., Gearing, M., Lah, J.J., Levey, A.I., Seyfried, N.T., 2015. Quantitative phosphoproteomics of Alzheimer's disease reveals cross-talk between kinases and small heat shock proteins. *Proteomics* 15, 508–519. <https://doi.org/10.1002/pmic.201400189>.
- Dolan, P.J., Johnson, G.V.W., 2010. The role of tau kinases in Alzheimer's disease. *Curr. Opin. Drug Discov. Devel.* 13, 595–603.
- Donner, L., Fälker, K., Gremer, L., Klinker, S., Pagani, G., Ljungberg, L.U., Lothmann, K., Rizzati, F., Schaller, M., Gohlke, H., Willbold, D., Grenegard, M., Elvers, M., 2016. Platelets contribute to amyloid- β aggregation in cerebral vessels through integrin $\alpha_{IIb}\beta_3$ -induced outside-in signaling and clusterin release. *Sci. Signal* 9 <https://doi.org/10.1126/scisignal.aaf6240>. ra52-ra52.
- Durink, S., Spellman, P.T., Birney, E., Huber, W., 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* 4, 1184–1191. <https://doi.org/10.1038/nprot.2009.97>.
- Edgar, R., 2002. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210. <https://doi.org/10.1093/nar/30.1.207>.
- Feng, T., Tammineni, P., Agrawal, C., Jeong, Y.Y., Cai, Q., 2017. Autophagy-mediated regulation of BACE1 protein trafficking and degradation. *J. Biol. Chem.* 292, 1679–1690. <https://doi.org/10.1074/jbc.M116.766584>.
- Forner, S., Baglietto-Vargas, D., Martini, A.C., Trujillo-Estrada, L., Laferla, F.M., 2017. Synaptic impairment in Alzheimer's disease: a dysregulated symphony. *Trends Neurosci.* 40, 347–357. <https://doi.org/10.1016/j.tins.2017.04.002>.
- Heppner, F.L., Ransohoff, R.M., Becher, B., 2015. Immune attack: the role of inflammation in Alzheimer disease. *Nat. Rev. Neurosci.* 16, 358–372. <https://doi.org/10.1038/nrn3880>.
- Humphrey, S.J., Azimifar, S.B., Mann, M., 2015. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nat. Biotechnol.* 33, 990–995. <https://doi.org/10.1038/nbt.3327>.
- Jones, L., Holmans, P.A., Hamshere, M.L., Harold, D., Moskva, V., Ivanov, D., Pocklington, A., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Pahwa, J.S., Jones, N., Stretton, A., Morgan, A.R., Lovestone, S., Powell, J., Proitsis, P., Lupton, M.K., Brayne, C., Rubinsztein, D.C., Gill, M., Lawlor, B., Lynch, A., Morgan, K., Brown, K.S., Passmore, P.A., Craig, D., McGuinness, B., Todd, S., Holmes, C., Mann, D., Smith, A.D., Love, S., Kehoe, P.G., Mead, S., Fox, N., Rossor, M., Collinge, J., Maier, W., Jessen, F., Schürmann, B., van den Bussche, H., Heuser, I., Peters, O., Kornhuber, J., Wiltfang, J., Dichgans, M., Frölich, L., Hampel, H., Hüll, M., Rujescu, D., Goate, A.M., Kawe, J.S.K., Cruchaga, C., Nowotny, P., Morris, J.C., Mayo, K., Livingston, G., Bass, N.J., Gurling, H., McQuillin, A., Gwilliam, R., Deloukas, P., Al-Chalabi, A., Shaw, C.E., Singleton, A.B., Guerreiro, R., Mühleisen, T.W., Nöthen, M.M., Moebus, S., Jöckel, K.-H., Klopp, N., Wichmann, H.-E., Ruther, E., Carrasquillo, M.M., Pankratz, V.S., Younkin, S.G., Hardy, J., O'Donovan, M.C., Owen, M.J., Williams, J., 2010. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. *PLoS One* 5, e13950. <https://doi.org/10.1371/journal.pone.0013950>.
- Jones, L., Lambert, J.-C., Wang, L.-S., Choi, S.-H., Harold, D., Vedernikov, A., Escott-Price, V., Stone, T., Richards, A., Bellenguez, C., Ibrahim-Verbaas, C.A., Naj, A.C., Sims, R., Gerrish, A., Jun, G., Destefano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., Russo, G., Thornton-Wells, T.A., Jones, N., Smith, A.V., Chouraki, V., Thomas, C., Ikram, M.A., Zelenika, D., Vardarajan, B.N., Kamatani, Y., Lin, C.-F., Schmidt, H., Kunkle, B.W., Dunstan, M.L., Ruiz, A., Bihoreau, M.-T., Reitz, C., Pasquier, F., Hollingworth, P., Hanon, O., Fitzpatrick, A.L., Buxbaum, J.D., Campion, D., Crane, P.K., Becker, T., Gudnason, V., Cruchaga, C., Craig, D., Amin, N., Berr, C., Lopez, O.L., De Jager, P.L., Deramecourt, V., Johnston, J.A., Evans, D., Lovestone, S., Letteneur, L., Kornhuber, J., Tarraga, L., Rubinsztein, D.C., Eiriksdottir, G., Sleegers, K., Goate, A.M., Fievet, N., Huentelman, M.J., Gill, M., Emilsson, V., Brown, K., Kamboh, M.I., Keller, L., Barberger-Gateau, P., McGuinness, B., Larson, E.B., Myers, A.J., Dufouil, C., Todd, S., Wallon, D., Love, S., Kehoe, P., Rogava, E., Gallacher, J., St George-Hyslop, P., Clarimon, J., Lleó, A., Bayer, A., Tsuang, D.W., Yu, L., Tsolaki, M., Bossù, P., Spalletta, G., Proitsis, P., Collinge, J., Sorbi, S., Sanchez Garcia, F., Fox, N., Hardy, J., Deniz Naranjo, M.C., Razquin, C., Bosco, P., Clarke, R., Brayne, C., Galimberti, D., Mancuso, M., Moebus, S., Mecocci, P., del Zompo, M., Maier, W., Hampel, H., Pilotto, A., Bullido, M., Panza, F., Caffarra, P., Nacmias, B., Gilbert, J.R., Mayhaus, M., Jessen, F., Dichgans, M., Lannfelt, L., Hakonarson, H., Pichler, S., Carrasquillo, M.M., Ingelsson, M., Beekly, D., Alavarez, V., Zou, F., Valladares, O., Younkin, G.G., Coto, E., Hamilton-Nelson, K.L., Mateo, I., Owen, M.J., Faber, K.M., Jonsson, P.V., Combarros, O., O'Donovan, M.C., Cantwell, L.B., Soininen, H., Blacker, D., Mead, S., Mosley, T.H., Bennett, D.A., Harris, T.B., Fratiglioni, L., Holmes, C., de Bruijn, R.F.A.G., Passmore, P., Montine, T.J., Bettens, K., Rotter, J.L., Brice, A., Morgan, K., Foroud, T.M., Kukull, W.A., Hannequin, D., Powell, J.F., Nalls, M.A., Ritchie, K., Lunetta, K.L., Kawe, J.S.K., Boerwinkle, E., Riemschneider, M., Boada, M., Hiltunen, M., Martin, E.R., Pastor, P., Schmidt, R., Rujescu, D., Dartigues, J.-F., Mayeux, R., Tzourio, C., Hofman, A., Nöthen, M.M., Graff, C., Psaty, B.M., Haines, J.L., Lathrop, M., Pericak-Vance, M.A., Launer, L.J., Farrer, L.A., van Duijn, C.M., Van Broeckhoven, C., Ramirez, A., Schellenberg, G.D., Seshadri, S., Amouyel, P., Williams, J., Holmans, P.A., 2015. Convergent genetic and expression data implicate immunity in Alzheimer's disease. *Alzheimer's Dement* 11, 658–671. <https://doi.org/10.1016/j.jalz.2014.05.1757>.
- Josse, J., Husson, F., 2016. *missMDA*: a package for handling missing values in multivariate data analysis. *J. Stat. Softw.* 70, 1–31. <https://doi.org/10.18637/jss.v070.i01>.
- Kim, S.-J., Jeong, M.-J., Jo, H.-J., Jung, J.H., Kaang, B.-K., Choi, Y.-B., Kim, J.-H., 2017. Identification of postsynaptic phosphatidylinositol-4,5-bisphosphate (PIP2) roles for synaptic plasticity using chemically induced dimerization. *Sci. Rep.* 7, 3351. <https://doi.org/10.1038/s41598-017-03520-3>.
- Kim, Y., Zheng, X., Ansari, Z., Bunnell, M.C., Herdy, J.R., Traxler, L., Lee, H., Paquola, A.C.M., Blithikioti, C., Ku, M., Schlachetki, J.C.M., Winkler, J., Edenhofer, F., Glass, C.K., Paucar, A.A., Jaeger, B.N., Pham, S., Boyer, L., Campbell, B.C., Hunter, T., Mertens, J., Gage, F.H., 2018. Mitochondrial aging defects emerge in directly reprogrammed human neurons due to their metabolic profile. *Cell Rep.* 23, 2550–2558. <https://doi.org/10.1016/j.celrep.2018.04.105>.
- Kumar, L., Futschik, M.E., 2007. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* 2, 5–7. <https://doi.org/10.6026/97320630002005>.
- Lachmann, A., Ma'ayan, A., 2009. KEA: Kinase enrichment analysis. *Bioinformatics* 25, 684–686. <https://doi.org/10.1093/bioinformatics/btp026>.
- Landry, C.R., Levy, E.D., Michnick, S.W., 2009. Weak functional constraints on phosphoproteomes. *Trends Genet.* 25, 193–197. <https://doi.org/10.1016/j.tig.2009.03.003>.
- Lee, M., Kwon, Y.T., Li, M., Peng, J., Friedlander, R.M., Tsai, L.-H., 2000. Neurotoxicity induces cleavage of p35 to p25 by calpain. *Nature* 405, 360–364. <https://doi.org/10.1038/35012636>.
- Lin, M.T., Beal, M.F., 2006. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 443, 787–795. <https://doi.org/10.1038/nature05292>.
- Liu, C.-C., Kanekiyo, T., Xu, H., Bu, G., 2013. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat. Rev. Neurol.* 9, 106–118. <https://doi.org/10.1038/nrneuro.2012.263>.
- Luo, W., Dou, F., Rodina, A., Chip, S., Kim, J., Zhao, Q., Moulick, K., Aguirre, J., Wu, N., Greengard, P., Chiosis, G., 2007. Roles of heat-shock protein 90 in maintaining and facilitating the neurodegenerative phenotype in tauopathies. *Proc. Natl. Acad. Sci.* 104, 9511–9516. <https://doi.org/10.1073/pnas.0710055104>.
- Manser, C., Stevenson, A., Banner, S., Davies, J., Tudor, E.L., Ono, Y., Nigel Leigh, P., McLoughlin, D.M., Shaw, C.E., Miller, C.C.J., 2008. Deregulation of PKN1 activity disrupts neurofilament organisation and axonal transport. *FEBS Lett.* 582, 2303–2308. <https://doi.org/10.1016/j.febslet.2008.05.034>.
- Martiskainen, H., Viswanathan, J., Nykänen, N.P., Kurki, M., Helisalme, S., Natunen, T., Sarajärvi, T., Kurkinen, K.M.A., Pursiheimo, J.P., Rauramaa, T., Alafuzoff, I., Jääskeläinen, J.E., Leinonen, V., Soininen, H., Haapasalo, A., Huttunen, H.J., Hiltunen, M., 2015. Transcriptomics and mechanistic elucidation of Alzheimer's disease risk genes in the brain and invitro models. *Neurobiol. Aging* 36, 1221.e15–1221.e28. <https://doi.org/10.1016/j.neurobiolaging.2014.09.003>.
- Mazanetz, M.P., Fischer, P.M., 2007. Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases. *Nat. Rev. Drug Discov.* 6, 464–479. <https://doi.org/10.1038/nrd2111>.
- Natunen, T., Martiskainen, H., Sarajärvi, T., Helisalme, S., Pursiheimo, J.-P., Viswanathan, J., Laitinen, M., Mäkinen, P., Kauppinen, T., Rauramaa, T., Leinonen, V., Alafuzoff, I., Haapasalo, A., Soininen, H., Hiltunen, M., 2013. Effects of NR1H3 genetic variation on the expression of liver X receptor α and the progression of Alzheimer's disease. *PLoS One* 8, e80700. <https://doi.org/10.1371/journal.pone.0080700>.
- Ohno, H., Kato, S., Naito, Y., Kunitomo, H., Tomioka, M., Iino, Y., 2014. Role of synaptic phosphatidylinositol 3-kinase in a behavioral learning response in *C. elegans*. *Science* Vol. 345 (80), 313–317. <https://doi.org/10.1126/science.1250709>.
- Phillips, H.S., Hains, J.M., Armanini, M., Laramee, G.R., Johnson, S.A., Winslow, J.W., 1991. BDNF mRNA is decreased in the hippocampus of individuals with Alzheimer's disease. *Neuron* 7, 695–702. [https://doi.org/10.1016/0896-6273\(91\)90273-3](https://doi.org/10.1016/0896-6273(91)90273-3).
- Querfurth, H.W., Laferla, F.M., 2010. Alzheimer's disease. *N. Engl. J. Med.* 362, 329–344. <https://doi.org/10.1056/NEJMr0909142>.
- Quinlan, J.R., 1987. Simplifying decision trees. *Int. J. Man. Mach. Stud.* 27, 221–234. [https://doi.org/10.1016/S0020-7373\(87\)80053-6](https://doi.org/10.1016/S0020-7373(87)80053-6).
- Schwämmle, V., Jensen, O.N., 2010. A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. *Bioinformatics* 26, 2841–2848. <https://doi.org/10.1093/bioinformatics/btq534>.
- Scotti, M.M., Swanson, M.S., 2015. RNA mis-splicing in disease. *Nat. Rev. Genet.* 17, 19–32. <https://doi.org/10.1038/nrg.2015.3>.
- Selkoe, D.J., Hardy, J., 2016. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* 8, 595–608. <https://doi.org/10.15252/emmm.201606210>.
- Seyfried, N.T., Dammer, E.B., Swarup, V., Nandakumar, D., Duong, D.M., Yin, L., Deng, Q., Nguyen, T., Hales, C.M., Wingo, T., Glass, J., Gearing, M., Thambisetty, M., Troncoso, J.C., Geschwind, D.H., Lah, J.J., Levey, A.I., 2017. A multi-network approach identifies protein-specific co-expression in asymptomatic and symptomatic Alzheimer's disease. *Cell Syst* 4, 60–72. e4. <https://doi.org/10.1016/j.cels.2016.11.006>.
- Sharma, K., Schmitt, S., Bergner, C.G., Tyanova, S., Kannaiyan, N., Manrique-Hoyos, N., Kongi, K., Cantuti, L., Hanisch, U.-K., Phillips, M.-A., Rossner, M.J., Mann, M., Simons, M., 2015. Cell type- and brain region-resolved mouse brain proteome. *Nat. Neurosci.* 18, 1819–1831. <https://doi.org/10.1038/nn.4160>.
- Shi, Q., Gibson, G.E., 2007. Oxidative stress and transcriptional regulation in Alzheimer disease. *Alzheimer Dis. Assoc. Disord.* 21, 276–291. <https://doi.org/10.1097/WAD.0b013e31815721c3>.
- Šimić, G., Babić Leko, M., Wray, S., Harrington, C., Delalle, I., Jovanov-Milošević, N., Bažadona, D., Buée, L., de Silva, R., Giovanni, G. D., Wischik, C., Hof, P.R., 2016. Tau protein hyperphosphorylation and aggregation in Alzheimer's disease and other tauopathies, and possible neuroprotective strategies. *Biomol. Ther.* 6, 2–28. <https://doi.org/10.3390/biom6010006>.
- Sperling, R., Mormino, E., Johnson, K., 2014. The evolution of preclinical Alzheimer's disease: Implications for prevention trials. *Neuron* 84, 608–622. <https://doi.org/10.1016/j.neuron.2014.10.038>.
- Sullivan, S.M., Sullivan, R.K.P., Miller, S.M., Ireland, Z., Björkman, S.T., Pow, D.V., Colditz, P.B., 2012. Phosphorylation of GFAP is associated with injury in the neonatal pig hypoxic-ischemic brain. *Neurochem. Res.* 37, 2364–2378. <https://doi.org/10.1007/s11064-012-0774-5>.
- Tan, M.S., Yu, J.T., Tan, L., 2013. Bridging integrator 1 (BIN1): form, function, and Alzheimer's disease. *Trends Mol. Med.* 19, 594–603. <https://doi.org/10.1016/j.molmed.2013.06.004>.
- Taus, T., Köcher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., Mechtler, K., 2011.

- Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* 10, 5354–5362. <https://doi.org/10.1021/pr200611n>.
- Tian, Y., Chang, J.C., Fan, E.Y., Flajolet, M., Greengard, P., 2013. Adaptor complex AP2/PICALM, through interaction with LC3, targets Alzheimer's APP-CTF for terminal degradation via autophagy. *Proc. Natl. Acad. Sci.* 110, 17071–17076. <https://doi.org/10.1073/pnas.1315110110>.
- Vizcaíno, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.W., Wang, R., Hermjakob, H., 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44, D447–D456. <https://doi.org/10.1093/nar/gkv1145>.
- Wang, D., Hiesinger, P.R., 2013. The vesicular ATPase: a missing link between acidification and exocytosis. *J. Cell Biol.* 203, 171–173. <https://doi.org/10.1083/jcb.201309130>.
- Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keefe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., Deng, S., Liddelow, S.A., Zhang, C., Daneman, R., Maniatis, T., Barres, B.A., Wu, J.Q., 2014. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* 34, 11929–11947. <https://doi.org/10.1523/JNEUROSCI.1860-14.2014>.
- Zhang, B., Gaiteri, C., Bodea, L.G., Wang, Z., McElwee, J., Podtelezchnikov, A.A., Zhang, C., Xie, T., Tran, L., Dobrin, R., Fluder, E., Clurman, B., Melquist, S., Narayanan, M., Suver, C., Shah, H., Mahajan, M., Gillis, T., Mysore, J., MacDonald, M.E., Lamb, J.R., Bennett, D.A., Molony, C., Stone, D.J., Gudnason, V., Myers, A.J., Schadt, E.E., Neumann, H., Zhu, J., Emilsson, V., 2013. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 153, 707–720. <https://doi.org/10.1016/j.cell.2013.03.030>.
- Zorzetto, M., Datturi, F., Divizia, L., Pistono, C., Campo, I., De Silvestri, A., Cuccia, M., Ricevuti, G., 2016. Complement C4A and C4B gene copy number study in Alzheimer's disease patients. *Curr. Alzheimer Res.* 13, 1–6. <https://doi.org/10.2174/15672050136661610130>.