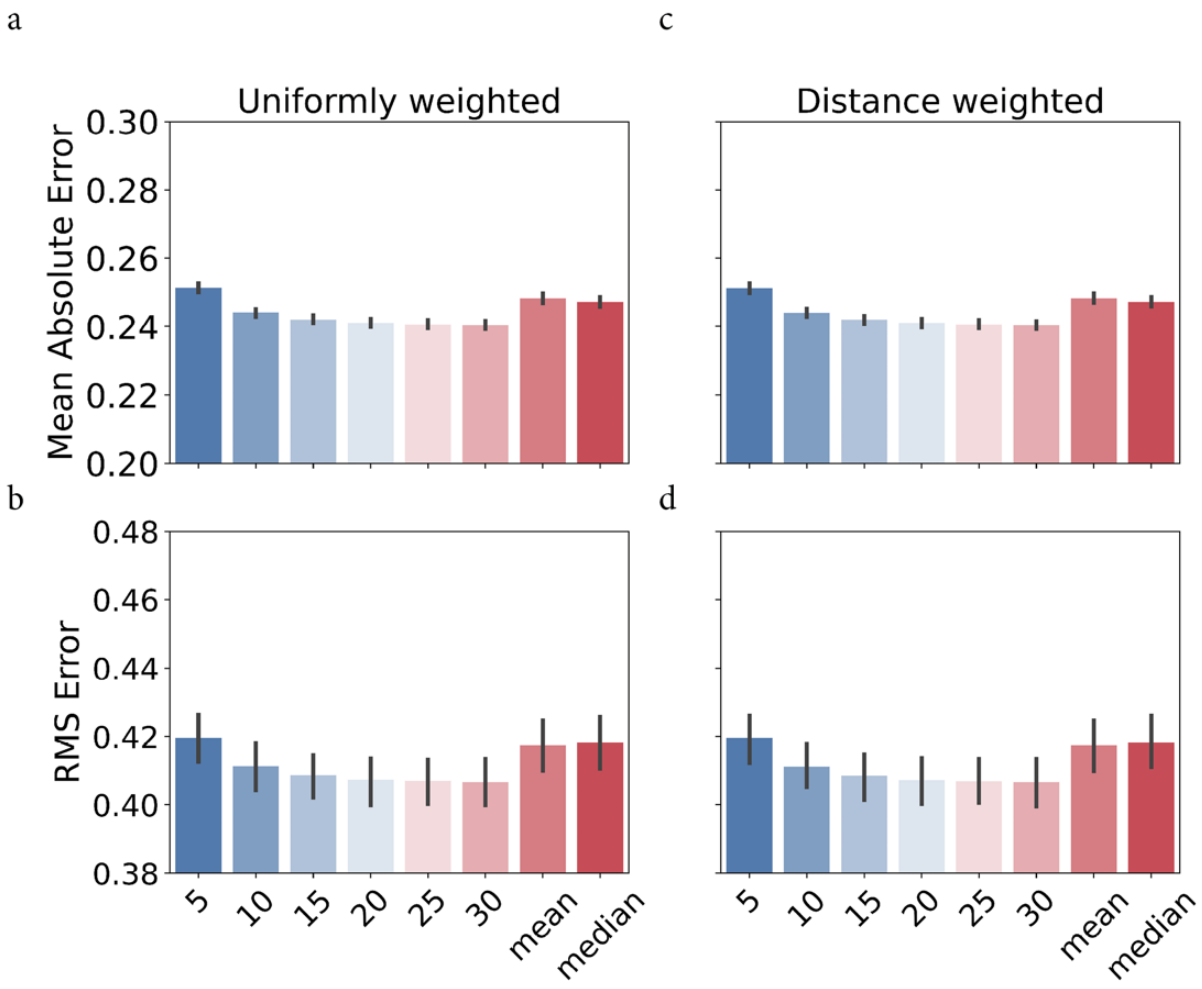
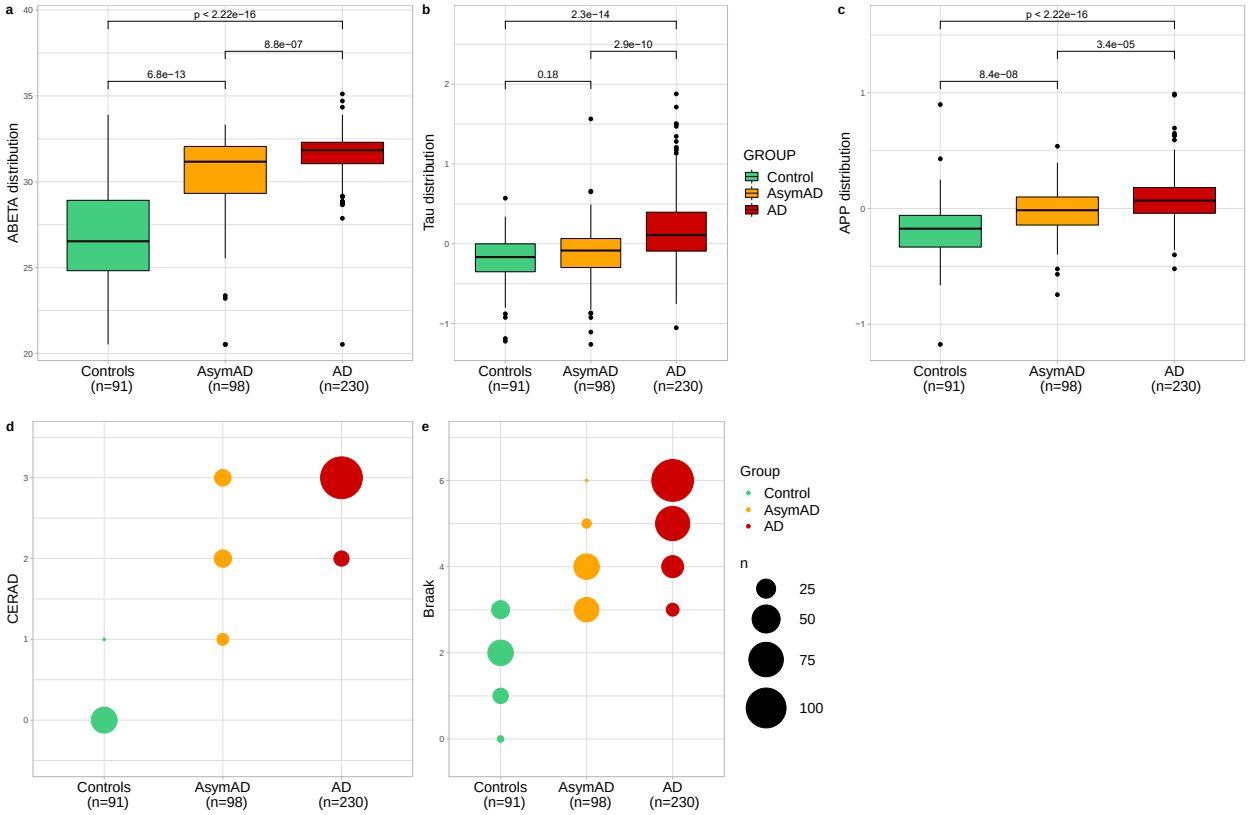
Supplementary Material

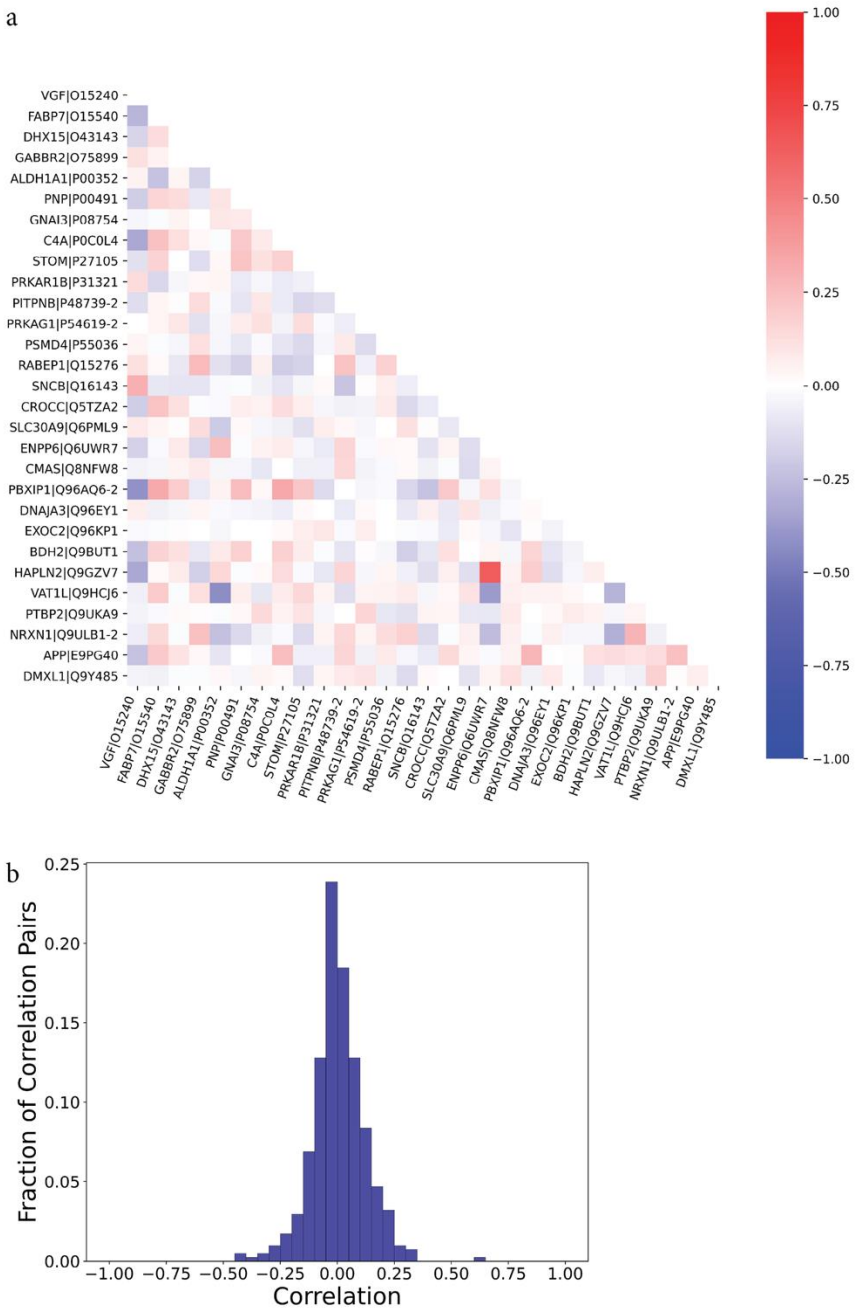
Machine Learning Selection of Most Predictive Brain Proteins Suggests Role of Sugar Metabolism in Alzheimer's Disease



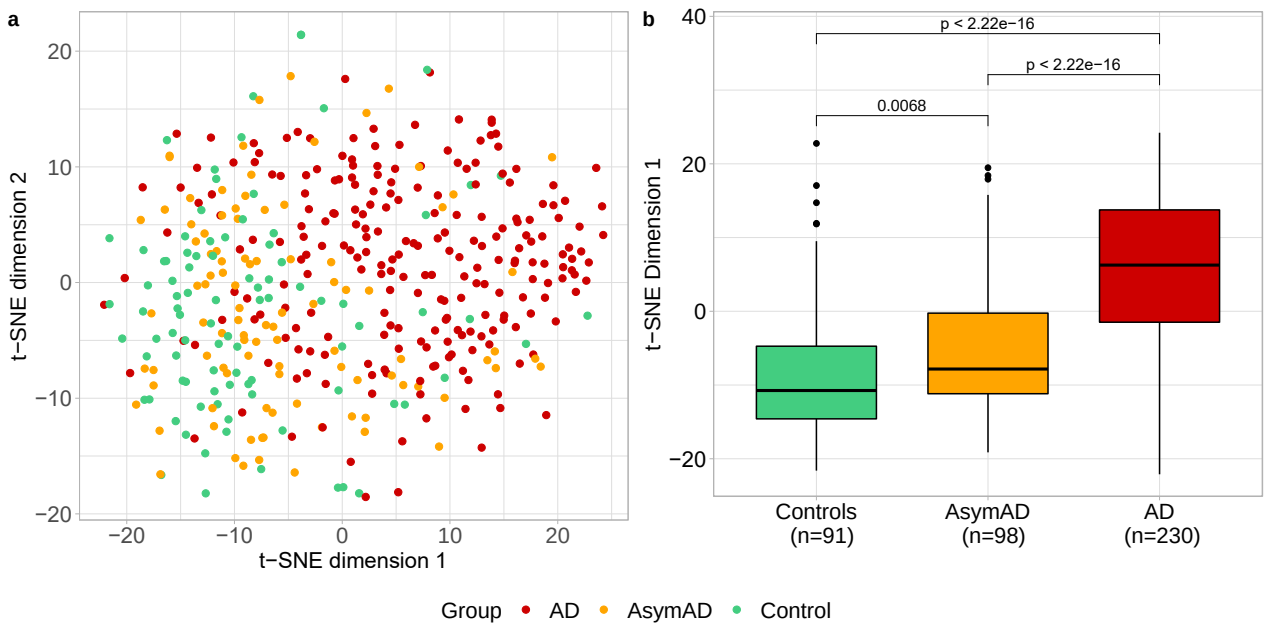
Supplementary Figure 1. Data imputation procedure. A KNN based imputation method is used for missing value estimation by iteratively optimizing the 'K' parameter in the KNN algorithm. First, some randomly chosen observations are deleted in the dataset. Then, these "induced" missing values were estimated by iterating K in the KNN imputation method. Variations of KNN imputation are also tried by using uniformly weighted (a, b) and distance weighted approaches (c, d) to weight the neighbors. Two different performance metrics – Mean Absolute Error (a, c) and RMS error (b, d) are also used to evaluate the imputation performance. In all iterations and conditions, it is found that using 20 neighbors leads to best observed performance. This holds true regardless of the uniformly weighted or distance weighted approaches. Ultimately a uniformly weighted KNN imputation approach, with K=20 (neighbors) is selected for missing value imputation.



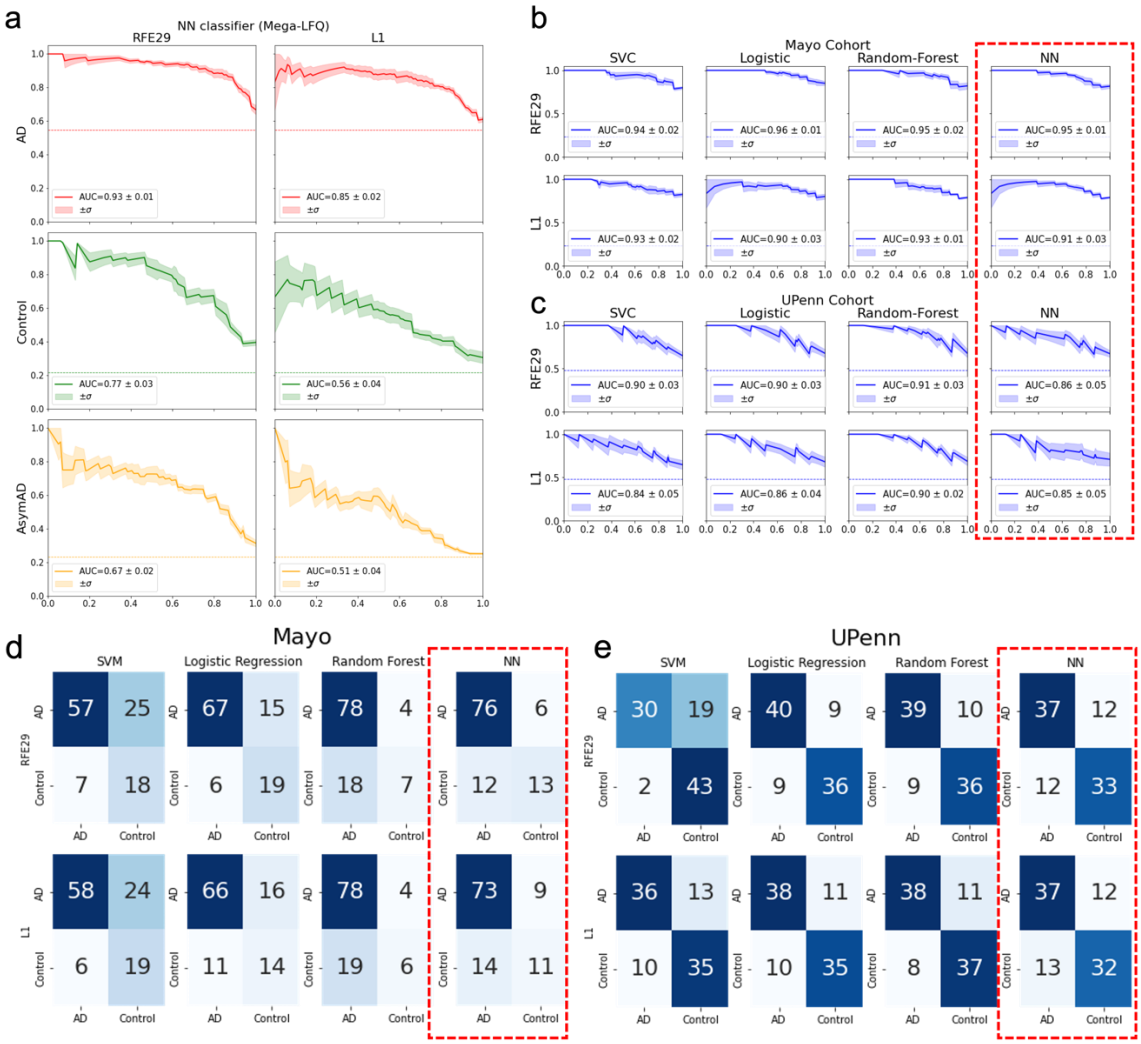
Supplementary Figure 2. Statistical distributions of protein selection cohort.
a-c) Distribution of amyloid-β, Tau, and APP across the subjects in the 4-cohort data (Banner, BLSA, MSSB, ACT) used to identify protein biomarkers. d, e) Distribution of CERAD and Braak scores across the same set of subjects.



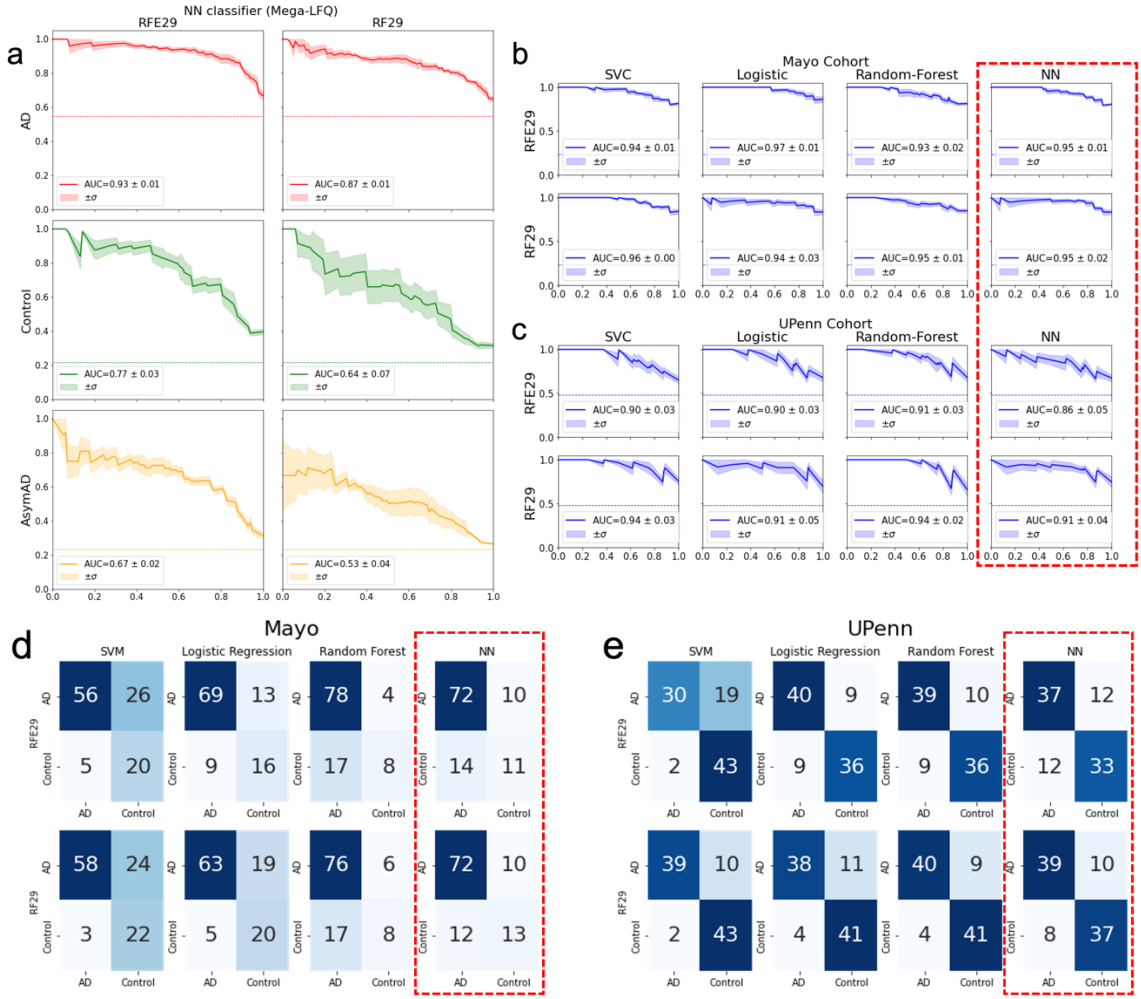
Supplementary Figure 3. Pearson correlation coefficient between the identified protein biomarkers. Pearson correlation coefficients between all pairs of the 29 proteins. The analysis illustrates that the proteins are not highly correlated. Thus, the classification model performance is not simply being driven by one or even a few proteins. The correlation between most (95%) pairs was low (between 0.25 and -0.25) a) Correlation coefficient (Pearson) matrix between all pairs of proteins. b) Distribution of correlation coefficients between all pairs of proteins (406 unique pairs in total)



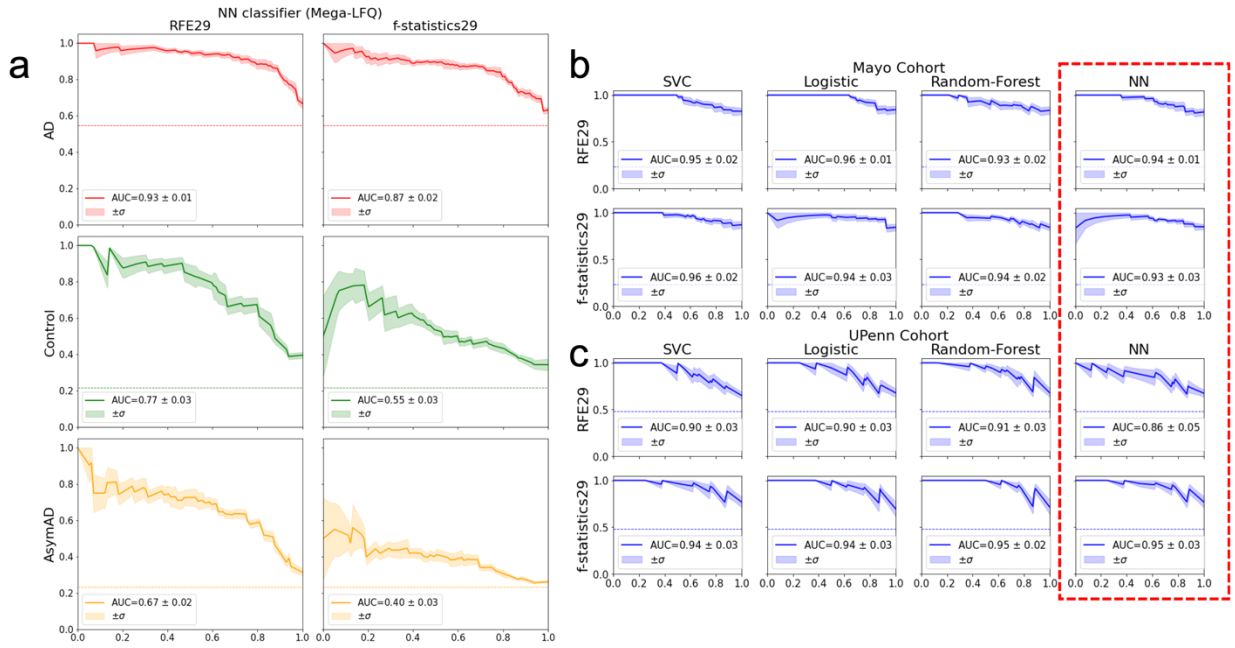
Supplementary Figure 4. tSNE analysis on the 4-cohort data using only proteins included in the biomarker panel (29 proteins). a) tSNE shows some separation between the 3 classes. The AD group is denser towards the right on dimension 1, while the control groups is denser towards the left. Asymptomatic AD seems to be in between the controls and AD groups. b) Box plot showing the distribution of t-SNE dimension-1 scores for each class shows significant differences between all pairs of classes. The result implies that increasing t-SNE dimension-1 score increases the risk of transition from control to AD.



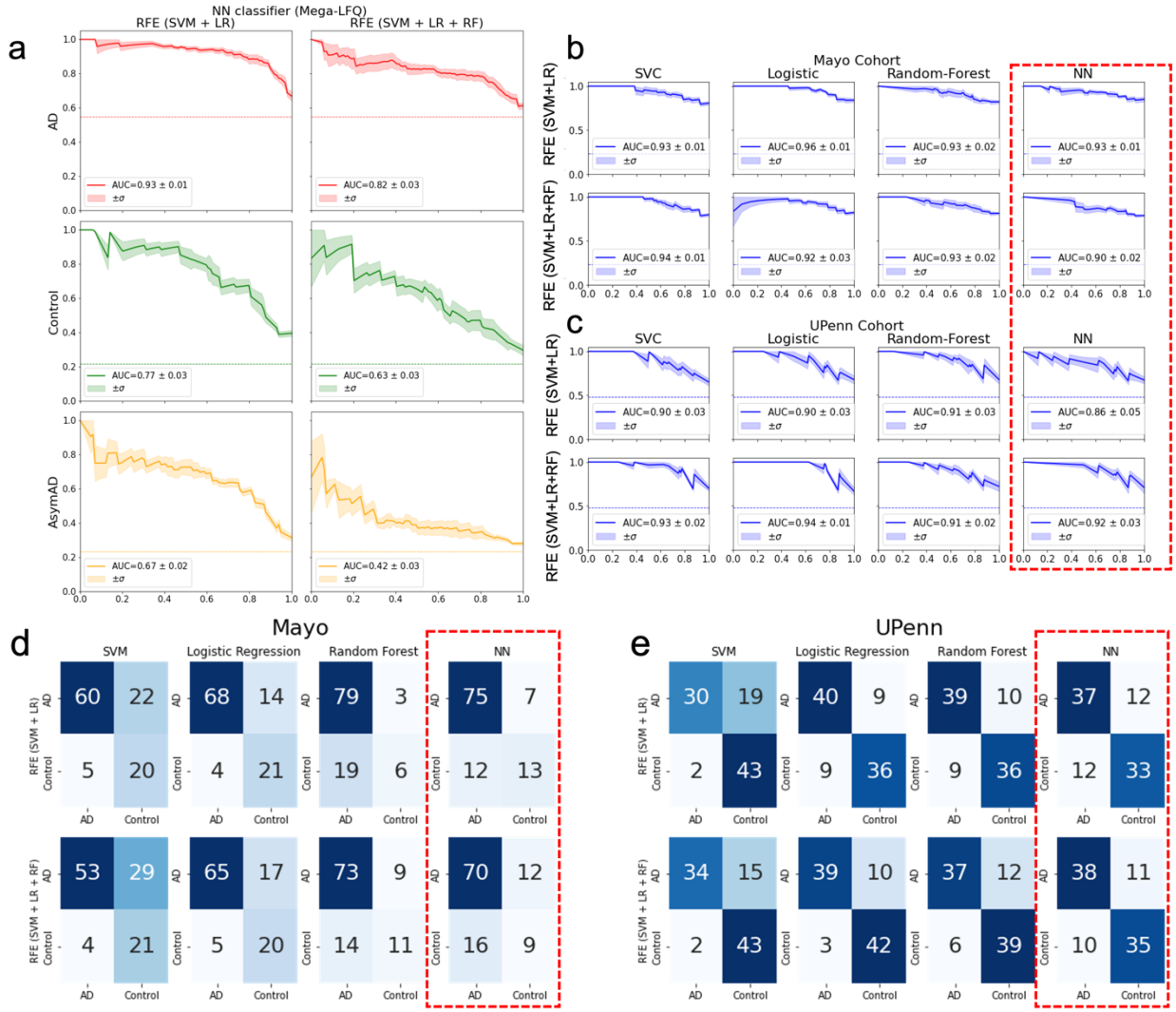
Supplementary Figure 5. Comparing the identified 29 protein biomarkers (RFE29) to another set of 29 selected via penalized lasso regularization within logistic regression (L1). The neural network NN (highlighted in red dotted line box) illustrates classification performance by a separate method not utilized in feature selection. The two sets (RFE29 and L1) had a small overlap of 3 proteins (C4A|P0C0L4, FABP7|O15540, VGF|O15240). However, the selected 29 proteins (via RFE) performed better in correctly identifying control and AsymAD classes in the Mega-LFQ dataset. On the held-out datasets (Mayo and UPenn), the difference in performance is smaller, but that could be due to these datasets having fewer classes and hence being simpler than Mega-LFQ. In summary, RFE performs better and is more robust than penalized lasso regulation within logistic regression. a) AUC of LFQ cohort. b) AUC of Mayo cohort. c) AUC on UPenn cohort. d) Confusion matrix for Mayo cohort. e) Confusion matrix for UPenn.



Supplementary Figure 6. Comparing the identified 29 protein biomarkers (RFE29) to another set of 29 selected via random forest feature importance (RF29). The neural network NN (highlighted in red dotted line box) illustrates classification performance by a separate method not utilized in feature selection. The two sets (RFE29 and RF29) had a small overlap of 4 proteins (C4A|P0C0L4, PBXIP1|Q96AQ6-2, VGF|O15240, APP|E9PG40). However, the selected 29 proteins (via RFE) performed better in correctly identifying control and AsymAD classes in the Mega-LFQ dataset. The two approaches did similarly well on the Mayo dataset, with RF29 doing slightly better on the UPenn dataset. This could be due to these datasets having fewer labels. a) AUC of LFQ cohort. b) AUC of Mayo cohort. c) AUC on UPenn cohort. d) Confusion matrix for Mayo cohort. e) Confusion matrix for UPenn.



Supplementary Figure 7. Comparing the identified 29 protein biomarkers (RFE29) to another set of 29 selected via f-statistics. The two sets (RFE29 and f-statistic29) had a small overlap of 4 proteins (FABP7|O15540, PBXIP1|Q96AQ6-2, VGF|O15240, APP|E9PG40). The neural network NN (highlighted in red dotted line box) illustrates classification performance by a separate method not utilized in feature selection. However, the RFE-selected 29 proteins (RFE29) performed better in correctly identifying control and AsymAD classes in the Mega-LFQ dataset. The performance was similar on the Mayo cohort, with f-statistics based selection (f-statistic29) performing slightly better on the UPenn cohort. a) AUC of LFQ cohort. b) AUC of Mayo cohort. c) AUC on UPenn cohort. d) Confusion matrix for Mayo cohort. e) Confusion matrix for UPenn.



Supplementary Figure 8. Comparing the identified 29 protein biomarkers produced by the intersection of two classifiers, i.e., RFE(SVM+LR) which corresponds to RFE29 biomarker set, versus biomarkers selected by the intersection of three classifiers, i.e., RFE(SVM+LR+RF). In both methods, the 50 top proteins are initially chosen, followed by final selection of the overlapping proteins by different classifiers. The RFE(SVM +LR) method utilized the intersecting proteins selected by SVM and LR to produce the RFE29 set. The RFE (SVM + LR + RF) utilized the intersecting proteins from SVM, LR, and RF (random forest), which reduced the final intersecting protein subset to 8 proteins: (RABEP1|Q15276, VGF|O15240, FABP7|O15540, PBXIP1|Q96AQ6-2, APP|E9PG40, DNAJA3|Q96EY1, NRXN1|Q9ULB1-2, C4A|P0C0L4). However, the RFE29 proteins (29 protein set) performed better in correctly identifying control and AsymAD classes in the Mega-LFQ dataset. The performance for the two approaches was similar on the Mayo cohort. On the UPenn cohort RFE (SVM+LR+RF) performed slightly better. However, the difference in performance on the Mega-LFQ dataset (more difficult with 3 classes) and UPenn dataset (binary classes) is stark with the smaller set of 8 proteins. Thus, while a smaller subset may be sufficient for a binary classification task, it performs subpar on the more complex multi-class classification, namely for classifying AsymAD. a) AUC of LFQ cohort. b) AUC of Mayo cohort. c) AUC on UPenn cohort. d) Confusion matrix for Mayo cohort. e) Confusion matrix for UPenn.

Supplementary Table 1. Identification and descriptions for best 29-protein subset. These proteins were selected by RFE to be most predictive for AD, AsymAD, or Control classification.

UniqueID	Mod.	Function and its role in context to AD	Ref
PNP | P00491 purine nucleoside phosphorylase	M8 pink	Role in neurotransmission, neuromodulation, trophic factor release, apoptosis, and inflammatory responses; Associated with a faster rate of cognitive decline in AD patients, highlighting the important role of purine metabolism	[1]
SNCB | Q16143 synuclein beta	M6 red	SNCB genotypes are associated with development of Lewy body diseases (Parkinson's disease, dementia with Lewy bodies and AD).	[2]
STOM | P27105 stomatin	M5 green	Lipid metabolism, regulates ion channel activity and transmembrane ion transport. Involved in lipid rafts.	[3]
PBXIP1 | Q96AQ6-2 PBX Homeobox Interacting Protein 1	M4 yellow	Neuropeptide signaling; PBX transcription factors in midbrain dopaminergic neurons plays a role in neurodegenerative diseases. Modulates many cancers, particularly leukemia and breast cancer.	[4]
FABP7 | O15540 fatty acid binding protein 7	M4 yellow	Transports unknown hydrophobic ligand for CNS development; required for radial glial fiber system in developing brain and migration of immature neurons to establish cortical layers; ApoE4 disrupts interaction of sortilin with FAB7 essential for lipid signaling.	[5]
C4A | P0C0L4 complement 4 amide	M4 yellow	Lipid metabolism; responsible for effective binding to form amide bonds with immune aggregates or protein antigens; increased C4A is found in AD patients, indicating role of C4A copy number variants in the risk of developing AD.	[6]
CROCC | Q5TZA2 Ciliary Rootlet Coiled-Coil, Rootletin	M4 yellow	Cillium biogenesis/degradation; required for centrosome cohesion; CROCC implicated in metabolic syndrome tied to insulin resistance, obesity, and type 2 diabetes	[7]
BDH2 | Q9BUT1 3-Hydroxybutyrate Dehydrogenase 2	M4 yellow	Regulation of lipid metabolism; plays a role in susceptibility to bacterial infection by providing an assimilable source of iron exploited by pathogenic bacteria; genes of butanoate metabolism pathway upregulated in AD. Downregulated in lupus, upregulated in cancers due to impact on iron.	[8] [9] [10]
APP | E9PG40 amyloid precursor protein	M4 yellow	Functions as a cell surface receptor and performs physiological functions on the surface of neurons relevant to neurite growth, neuronal adhesion and axon genesis; APP proteolysis is the crucial step in development of AD.	[11]
HAPLN2 | Q9GZV7 hyaluronan and proteoglycan link protein 2	M2 blue	HAPLN2 deficiency leads to abnormal expression of extracellular matrix proteins and dysfunctional neuronal conductivity. Target for multiple neurologic diseases, including Parkinson's and Alzheimer's.	[12]
ENPP6 | Q6UWR7 ectonucleotide pyrophosphatase/ phosphodiesterase 6	M2 blue	Expressed in new differentiating oligodendrocytes; component of early synaptic phases of motor learning; expressed on the myelin membrane and is soluble extracellularly; two loci in ENPP6 are significantly associated with AD + psychosis.	[13]
VGF | O15240 vascular endothelial growth factor	M1 turquoise	Strongly associated with cognitive trajectory; involved in synaptic functions; independent of amyloid-beta plaques and neurofibrillary tangles; protects against AD pathogenesis	[14, 15]
GABBR2 | O75899 gamma-aminobutyric acid type B receptor subunit 2	M1 turquoise	GABBR2 encodes the $GABA_B$ receptor 2 subunit – an important GABA signaling component. $GABA_B$ subunit (a metabotropic receptor) is downregulated in the post-mortem human middle temporal gyrus in AD.	[16]
VAT1L | Q9HCJ6 vesicle amine transport 1	M1 turquoise	Plays a role in neuronal maintenance, neurotransmission and calcium signaling. It is associated with more rapid decline on AD assessment scale-cognitive subscale.	[17]
ALDHA1 | P00352 aldehyde dehydrogenase 1 family, member A1	Grey	Increases with AD severity; catalyzes the conversion of retinal to retinoic acid (RA) to regulate RA signaling, which is essential for normal brain homeostasis.	[18]
PSMD4 | P55036 proteasome 26S subunit ubiquitin receptor, non-ATPase 4	Grey	Downregulated in AD patients; plays a key role in maintenance of protein homeostasis by removing misfolded or damaged proteins. Also involved in hypertension and hypercholesterolemia.	[19] [20]
CMAS | Q8NFW8 cytidine monophosphate n-acetylneuraminic acid synthetase	Grey	Catalyzes activation of N-acetylneuraminic acid (NeuNAc) to CMP-NeuNAc, a substrate required for the addition of sialic acid to form sialylated glycoprotein and glycolipid. Potential therapeutic target for AD.	[21]

PITPNB | P48739-2 phosphatidylinositol transfer protein beta isoform	Grey	Catalyzes the transfer of phosphatidylinositol and phosphatidylcholine between membrane; involved in protein-protein interaction network of ciliary proteins indicating association between ciliary protein dysfunction and neuropsychiatric disorders	[22]
PTBP2 | Q9UKA9 polypyrimidine tract binding protein 2	Grey	RNA-binding protein which binds to intronicpolypyrimidine tracts and mediates negative regulation of exons splicing. There is an increase in PTB dependent splicing in AD.	[23]
DHX15 | O43143 putative pre-mRNA-splicing factor ATP-dependent RNA helicase	Grey	Splicing regulator with significant disease-related changes in transcript levels in AD. Has important roles in natural killer cell homeostasis.	[24] [25]
GNAI3 | P08754 guanine nucleotide-binding protein G(i) subunit alpha-3	Grey	Transducers of G-protein-coupled receptors in numerous signaling cascades; negative correlation with G-proteins and Src family of tyrosine kinases with AD phenotypes. Also involved in depression and Parkinson's.	[26]
PRKAG1 | P54619-2 5'-AMP-activated protein kinase subunit gamma-1	Grey	ATP binding subunit of AMP-activate protein kinase, an energy sensor protein kinase that plays a key role in regulating cellular energy metabolism.	[27]
EXOC2 | Q96KP1 exocyst complex component 2	Grey	EXOC2 has been reported for nominal association with AD age of onset modifier gene through a whole exome study. EXOC2 is also involved in skin pigment and vitamin D, where vitamin D deficiency has been tied to AD risk.	[28] [29]
NRXN1 | Q9ULB1-2 neurexin 1	Grey	Neuronal cell surface protein involved in cell recognition and cell adhesion. Forms intracellular junctions through binding neuroligins and interacting with neurexin. Neuroligin-neurexin pathway associated with AD.	[30]
DMXL1 | Q9Y485 DmX-like protein 1	Grey	A member of WD repeat superfamily of proteins, which have regulatory functions. Identified in GWAS studies for AD.	[31]
SLC30A9 | Q6PML9 solute carrier family 30 (zinc transporter), member 9	Grey	Zinc transporter involved in intracellular zinc homeostasis. An autosomal recessive cerebrorenal syndrome is known to be associated with pathogenic variants in SLC30A9.	[32]
PRKAR1B | P31321 protein kinase CAMP-dependent type I regulatory subunit beta	Grey	Regulatory subunit of cAMP-dependent protein kinases involved in cAMP signaling in cells. A pathogenic mutation found in gene coding for PRKAR1B protein is associated with aggregates of intermediate filaments seen in AD and PD.	[33]
DNAJA3 | Q96EY1 DnaJ heat shock protein family	Grey	Modulates apoptotic signal transduction or effector structures within the mitochondrial matrix. Role in neuromuscular junction development as an effector of MUSK signaling. Extracellular heat shock protein involved in neurodegenerative diseases, including AD.	[34]
RABEP1 | Q15276 rab GTPase-binding effector protein 1	Grey	Encodes RAB5 effector protein required for early endosome membrane fusions and phagosome biogenesis; AD risk enhancer in AD GWAS and myeloid epigenomic datasets.	[35]

REFERENCES

[1] Tumini E, Porcellini E, Chiappelli M, Conti CM, Beraudi A, Poli A, Caciagli F, Doyle R, Conti P, Licastro F (2007) The G51S purine nucleoside phosphorylase polymorphism is associated with cognitive decline in Alzheimer's disease patients. *Hum Psychopharmacol* **22**, 75-80.

[2] Gamez-Valero A, Canet-Pons J, Urbizu A, Anillo A, Santos C, Ariza A, Beyer K (2018) INDEL length and haplotypes in the beta-synuclein gene: a key to differentiate dementia with Lewy bodies? *J Alzheimers Dis* **65**, 207-219.

[3] Rungaldier S, Umlauf E, Mairhofer M, Salzer U, Thiele C, Prohaska R (2017) Structure-function analysis of human stomatin: A mutation study. *PLoS One* **12**, e0178646.

[4] Villaescusa JC, Li B, Toledo EM, Rivetti di Val Cervo P, Yang S, Stott SR, Kaiser K, Islam S, Gyllborg D, Laguna-Goya R, Landreh M, Lonnerberg P, Falk A, Bergman T, Barker RA, Linnarsson S, Selleri L, Arenas E (2016) A PBX1 transcriptional network controls dopaminergic neuron development and is impaired in Parkinson's disease. *EMBO J* **35**, 1963-1978.

[5] Asaro A, Sinha R, Bakun M, Kalnytska O, Carlo-Spiewok AS, Rubel T, Rozeboom A, Dadlez M, Kaminska B, Aronica E, Malik AR, Willnow TE (2021) ApoE4 disrupts interaction of sortilin with fatty acid-binding protein 7 essential to promote lipid signaling. *J Cell Sci* **134**, jcs258894.

[6] Zorzetto M, Datturi F, Divizia L, Pistono C, Campo I, De Silvestri A, Cuccia M, Ricevuti G (2017) Complement C4A and C4B gene copy number study in Alzheimer's disease patients. *Curr Alzheimer Res* **14**, 303-308.

[7] Mindikoglu AL, Abdulsada MM, Jain A, Jalal PK, Devaraj S, Wilhelm ZR, Opekun AR, Jung SY (2020) Intermittent fasting from dawn to sunset for four consecutive weeks induces anticancer serum proteome response and improves metabolic syndrome. *Sci Rep* **10**, 18341.

[8] Tsartsalis S, Fancy N, Smith AM, Khozoie C, Yang X, Davey K, Willumsen N, McGarry A, Muirhead RCJ, Debette S, Owen DR, Matthews PM (2021) Single nuclear transcriptional signatures of dysfunctional brain vascular homeostasis in Alzheimer's disease. *bioRxiv*, 2021.2010.2027.465860.

[9] Ansari AA, Tipu HN, Ahmad D, Farhan M (2020) Impact of homozygous C4A deficiency on clinical presentation of systemic lupus erythematosus. *J Coll Physicians Surg Pak* **30**, 790-795.

[10] Yang WC, Lin SF, Wang SC, Tsai WC, Wu CC, Wu SC (2020) The effects of human BDH2 on the cell cycle, differentiation, and apoptosis and associations with leukemia transformation in myelodysplastic syndrome. *Int J Mol Sci* **21**, 3033.

[11] O'Brien RJ, Wong PC (2011) Amyloid precursor protein processing and Alzheimer's disease. *Annu Rev Neurosci* **34**, 185-204.

[12] Wang Q, Wang C, Ji B, Zhou J, Yang C, Chen J (2019) Hapln2 in neurological diseases and its potential as therapeutic target. *Front Aging Neurosci* **11**, 60.

[13] DeMichele-Sweet MAA, Klei L, Creese B, Harwood JC, Weamer EA, McClain L, Sims R, Hernandez I, Moreno-Grau S, Tarraga L, Boada M, Alarcon-Martin E, Valero S, Nia-Load Family Based Study Consortium AsDGC, Liu Y, Hooli B, Aarsland D, Selbaek G, Bergh S, Rongve A, Saltvedt I, Skjellegrind HK, Engdahl B, Stordal E, Andreassen OA, Djurovic S, Athanasiu L, Seripa D, Borroni B, Albani D, Forloni G, Mecocci P, Serretti A, De Ronchi D, Politis A, Williams J, Mayeux R, Foroud T, Ruiz A, Ballard C, Holmans P, Lopez OL, Kamboh MI, Devlin B, Sweet RA (2021) Genome-wide association identifies the first risk loci for psychosis in Alzheimer disease. *Mol Psychiatry* **26**, 5797-5811.

[14] Wingo AP, Dammer EB, Breen MS, Logsdon BA, Duong DM, Troncosco JC, Thambisetty M, Beach TG, Serrano GE, Reiman EM, Caselli RJ, Lah JJ, Seyfried NT, Levey AI, Wingo TS (2019) Large-scale proteomic analysis of human brain identifies proteins associated with cognitive trajectory in advanced age. *Nat Commun* **10**, 1619.

[15] Beckmann ND, Lin WJ, Wang M, Cohain AT, Charney AW, Wang P, Ma W, Wang YC, Jiang C, Audrain M, Comella PH, Fakira AK, Hariharan SP, Belbin GM, Girdhar K, Levey AI, Seyfried NT, Dammer EB, Duong D, Lah JJ, Haure-Mirande JV, Shackleton B, Fanutza T, Blitzer R, Kenny E, Zhu J, Haroutunian V, Katsel P, Gandy S, Tu Z, Ehrlich ME, Zhang B, Salton SR, Schadt EE (2020) Multiscale causal networks identify VGF as a key regulator of Alzheimer's disease. *Nat Commun* **11**, 3942.

[16] Govindpani K, Turner C, Waldvogel HJ, Faull RLM, Kwakowsky A (2020) Impaired expression of GABA signaling components in the Alzheimer's disease middle temporal gyrus. *Int J Mol Sci* **21**, 8704.

[17] Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, Harvey D, Jack CR, Jr., Jagust W, Morris JC, Petersen RC, Saykin AJ, Shaw LM, Toga AW, Trojanowski JQ, Alzheimer's Disease Neuroimaging Initiative (2017) Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials. *Alzheimers Dement* **13**, e1-e85.

[18] Nikhil K, Viccaro K, Shah K (2019) Multifaceted regulation of ALDH1A1 by Cdk5 in Alzheimer's disease pathogenesis. *Mol Neurobiol* **56**, 1366-1390.

[19] Peng YS, Tang CW, Peng YY, Chang H, Chen CL, Guo SL, Wu LC, Huang MC, Lee HC (2020) Comparative functional genomic analysis of Alzheimer's affected and naturally aging brains. *PeerJ* **8**, e8682.

[20] Loke SY, Wong PT, Ong WY (2017) Global gene expression changes in the prefrontal cortex of rabbits with hypercholesterolemia and/or hypertension. *Neurochem Int* **102**, 33-56.

[21] Wu ATH, Lawal B, Wei L, Wen YT, Tzeng DTW, Lo WC (2021) Multiomics identification of potential targets for Alzheimer disease and antrocin as a therapeutic candidate. *Pharmaceutics* **13**, 1555.

[22] Karunakaran KB, Chaparala S, Lo CW, Ganapathiraju MK (2020) Cilia interactome with predicted protein-protein interactions reveals connections to Alzheimer's disease, aging and other neuropsychiatric processes. *Sci Rep* **10**, 15629.

[23] Li D, McIntosh CS, Mastaglia FL, Wilton SD, Aung-Htut MT (2021) Neurodegenerative diseases: a hotbed for splicing defects and the potential therapies. *Transl Neurodegener* **10**, 16.

[24] Tollervey JR, Wang Z, Hortobagyi T, Witten JT, Zarnack K, Kayikci M, Clark TA, Schweitzer AC, Rot G, Curk T, Zupan B, Rogelj B, Shaw CE, Ule J (2011) Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. *Genome Res* **21**, 1572-1582.

[25] Wang G, Xiao X, Wang Y, Chu X, Dou Y, Minze LJ, Ghobrial RM, Zhang Z, Li XC (2022) The RNA helicase DHX15 is a critical regulator of natural killer-cell homeostasis and functions. *Cell Mol Immunol* **19**, 687-701.

[26] Zhang Q, Ma C, Gearing M, Wang PG, Chin LS, Li L (2018) Integrated proteomics and network analysis identifies protein hubs and network alterations in Alzheimer's disease. *Acta Neuropathol Commun* **6**, 19.

[27] Caberlotto L, Lauria M, Nguyen TP, Scotti M (2013) The central role of AMP-kinase and energy homeostasis impairment in Alzheimer's disease: a multifactor network analysis. *PLoS One* **8**, e78919.

[28] Velez JI, Rivera D, Mastronardi CA, Patel HR, Tobon C, Villegas A, Cai Y, Easteal S, Lopera F, Arcos-Burgos M (2016) A mutation in DAOA modifies the age of onset in PSEN1 E280A Alzheimer's disease. *Neural Plast* **2016**, 9760314.

[29] Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, Hankinson SE, Hu FB, Duffy DL, Zhao ZZ, Martin NG, Montgomery GW, Hayward NK, Thomas G, Hoover RN, Chanock S, Hunter DJ (2008) A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* **4**, e1000074.

[30] Martinez-Mir A, Gonzalez-Perez A, Gayan J, Antunez C, Marin J, Boada M, Lopez-Arrieta JM, Fernandez E, Ramirez-Lorca R, Saez ME, Ruiz A, Scholl FG, Real LM (2013) Genetic study of neurexin and neuroligin genes in Alzheimer's disease. *J Alzheimers Dis* **35**, 403-412.

[31] Sherva R, Tripodis Y, Bennett DA, Chibnik LB, Crane PK, de Jager PL, Farrer LA, Saykin AJ, Shulman JM, Naj A, Green RC, Consortium G, Alzheimer's Disease Neuroimaging Initiative, Alzheimer's Disease Genetics Consortium (2014) Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimers Dement* **10**, 45-52.

[32] Kleyner R, Arif M, Marchi E, Horowitz N, Haworth A, King B, Gavin M, Amble K, Velinov M, Lyon GJ (2022) Autosomal recessive SLC30A9 variants in a proband with a cerebrorenal syndrome and no parental consanguinity. *Cold Spring Harb Mol Case Stud* **8**, a006137.

[33] Wong TH, Chiu WZ, Breedveld GJ, Li KW, Verkerk AJ, Hondius D, Hukema RK, Seelaar H, Frick P, Severijnen LA, Lammers GJ, Lebbink JH, van Duinen SG, Kamphorst W, Rozemuller AJ, Netherlands Brain B, Bakker EB, International Parkinsonism Genetics Network, Neumann M, Willemsen R, Bonifati V, Smit AB, van Swieten J (2014) PRKAR1B mutation associated with a new neurodegenerative disorder with unique pathology. *Brain* **137**, 1361-1373.

[34] Beck JS, Mufson EJ, Counts SE (2016) Evidence for mitochondrial UPR gene activation in familial and sporadic Alzheimer's disease. *Curr Alzheimer Res* **13**, 610-614.

[35] Novikova G, Kapoor M, Tcw J, Abud EM, Efthymiou AG, Chen SX, Cheng H, Fullard JF, Bendl J, Liu Y, Roussos P, Bjorkegren JL, Liu Y, Poon WW, Hao K, Marcora E, Goate AM (2021) Integration of Alzheimer's disease genetics and myeloid genomics identifies disease risk regulatory elements and genes. *Nat Commun* **12**, 1610.