nature aging

Supplementary information

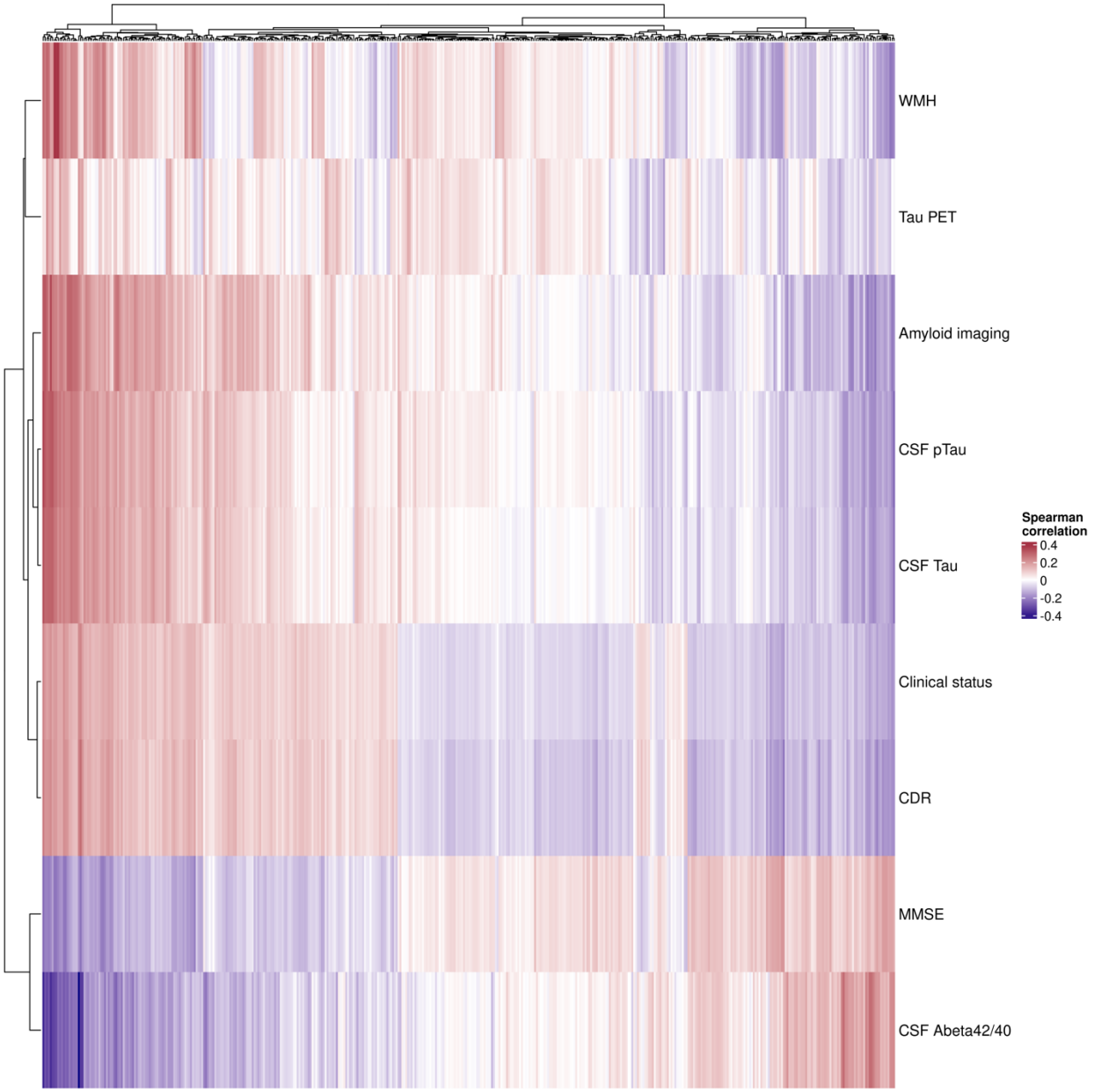
https://doi.org/10.1038/s43587-025-00872-8

Large-scale plasma proteomic profiling unveils diagnostic biomarkers and pathways for Alzheimer's disease

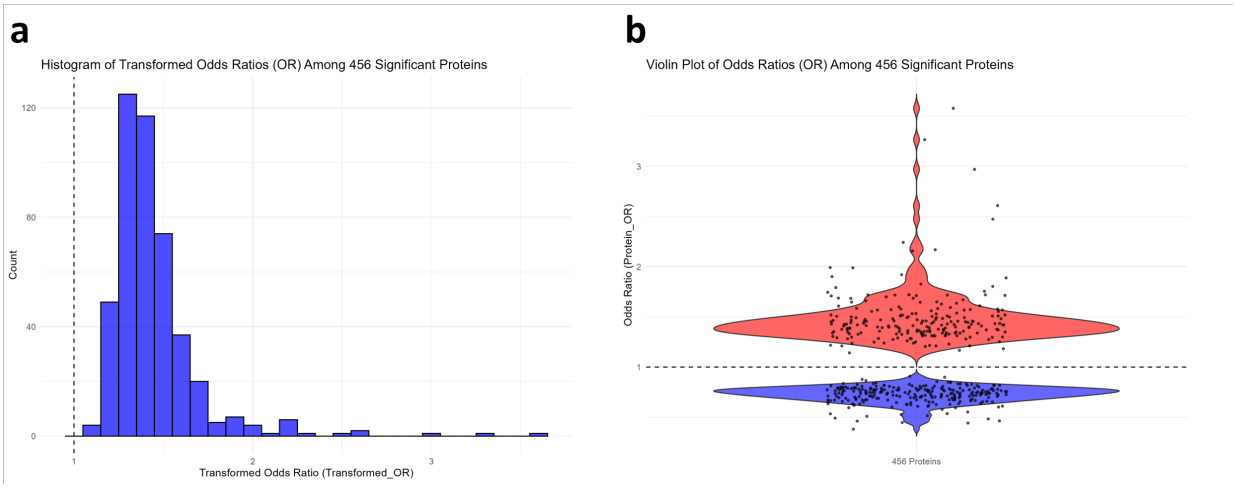
In the format provided by the authors and unedited

Table of contents

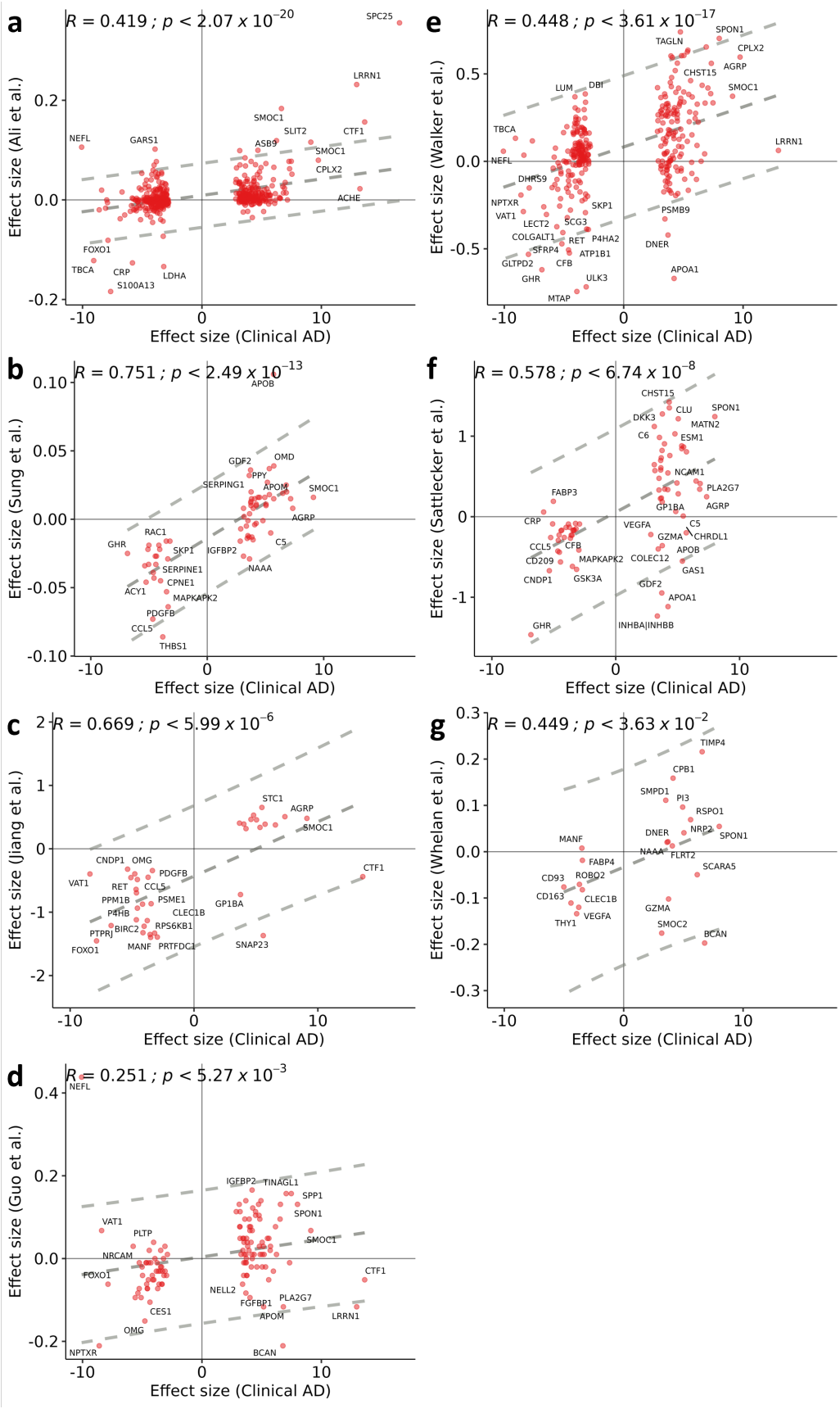
Supplementary figures



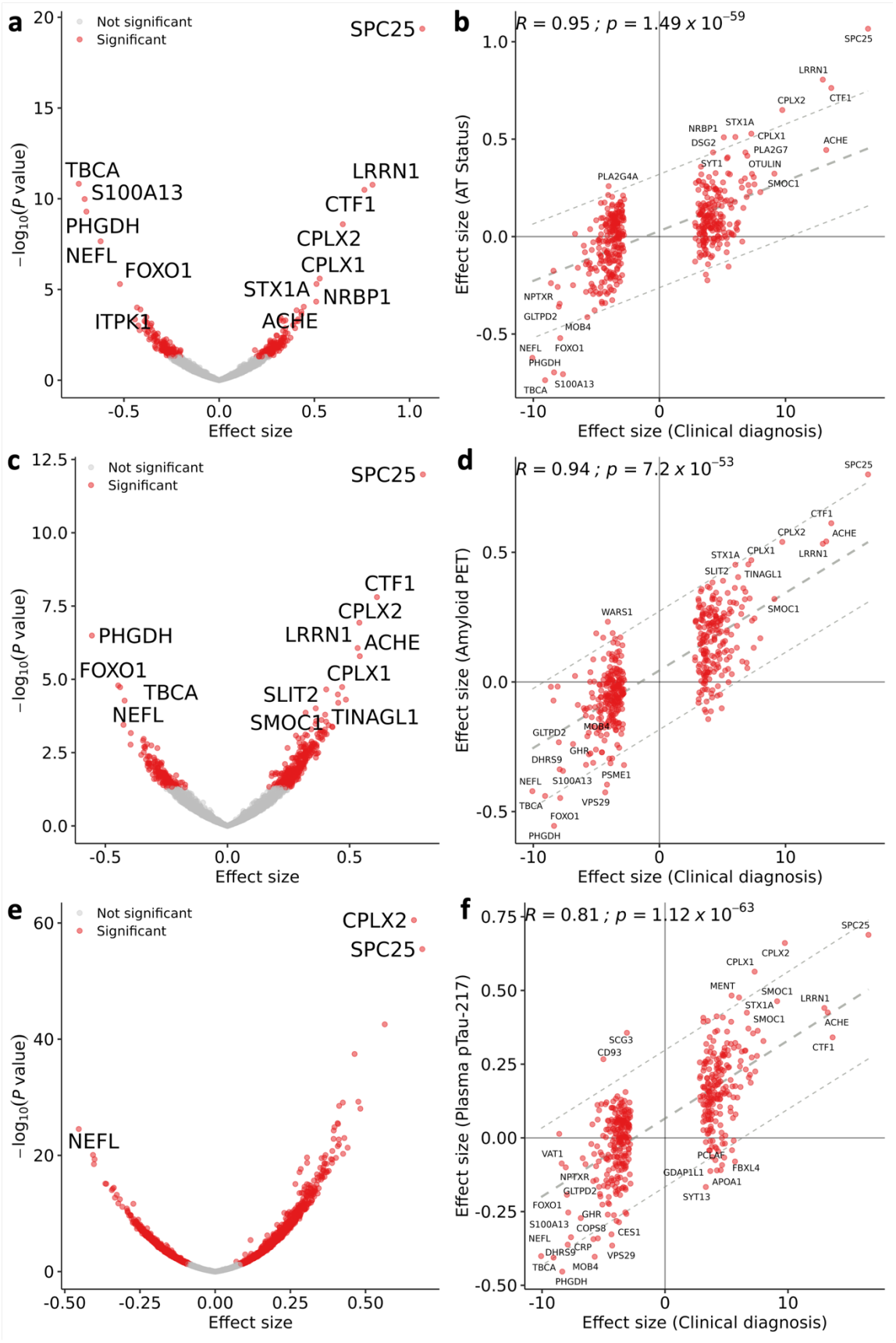
Supplementary Figure 1. Correlation between proteins and AD-related phenotypes. Heatmap showing correlations between 456 significant aptamers and AD-related phenotypes.
Abbreviations: CSF, Cerebrospinal fluid; pTau, Phosphorylated tau; CDR, Clinical Dementia Rating; WMH, White matter hyperintensities; PET, Positron emission tomography; MMSE, Mini-Mental State Examination; Abeta, amyloid beta.



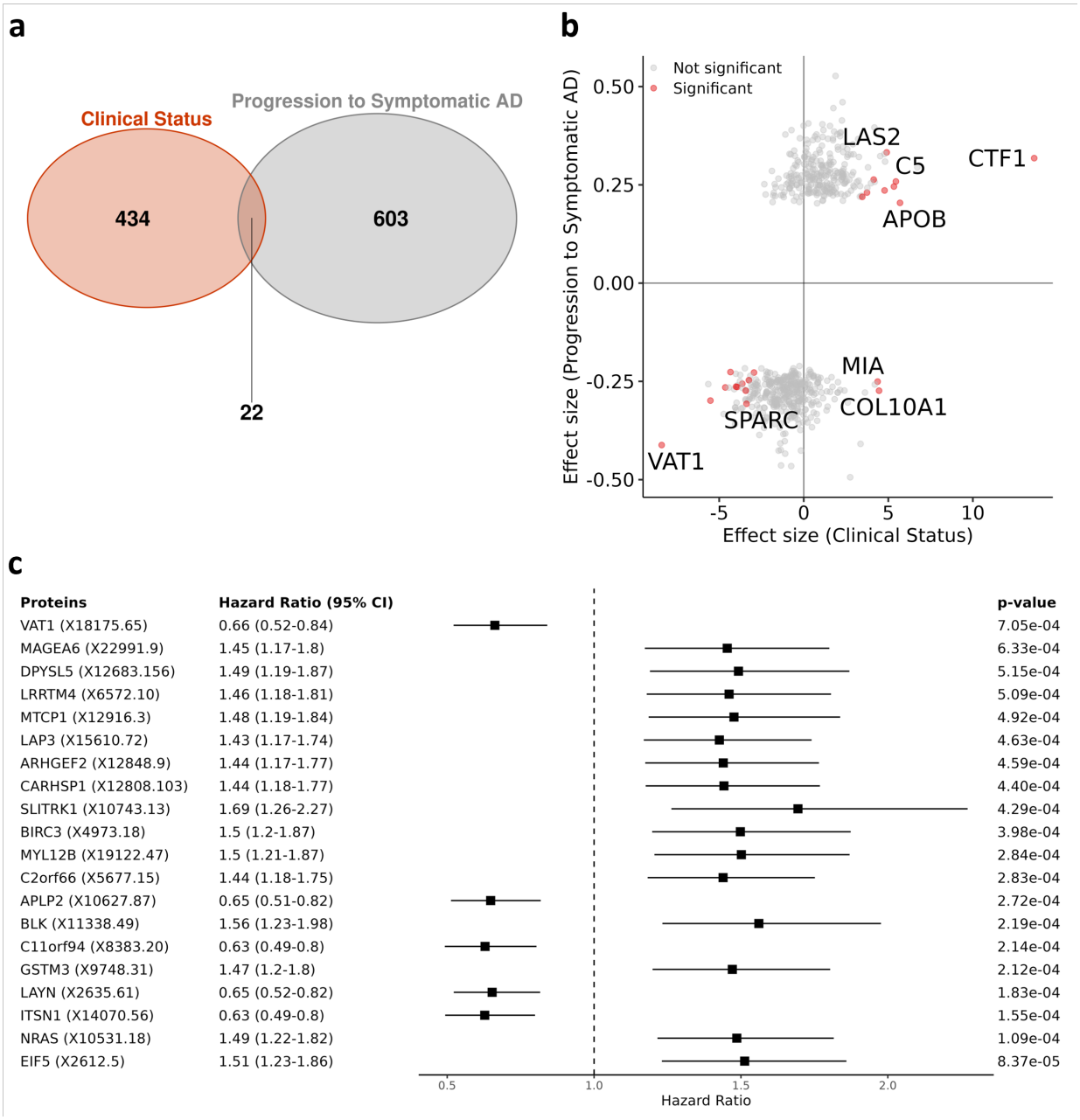
Supplementary Figure 2. Odd Ratio (OR) Analyses of 456 Significant Proteins. (A) Histogram showing the distribution of transformed OR across 456 significant proteins. **(B)** Violin plot of OR for the same 456 significant proteins. The red (upper) and blue (lower) sections highlight proteins with OR > 1 and OR < 1, respectively. The dashed line represents an OR of 1.


a
R = 0.419 ; p < 2.07 x 10−20
Effect size (Ali et al.)
Effect size (Clinical AD)
b
R = 0.751 ; p < 2.49 x 10−13
Effect size (Sung et al.)
Effect size (Clinical AD)
c
R = 0.669 ; p < 5.99 x 10−6
Effect size (Jiang et al.)
Effect size (Clinical AD)
d
R = 0.251 ; p < 5.27 x 10−3
Effect size (Guo et al.)
Effect size (Clinical AD)
e
R = 0.448 ; p < 3.61 x 10−17
Effect size (Walker et al.)
Effect size (Clinical AD)
f
R = 0.578 ; p < 6.74 x 10−8
Effect size (Sattlecker et al.)
Effect size (Clinical AD)
g
R = 0.449 ; p < 3.63 x 10−2
Effect size (Whelan et al.)
Effect size (Clinical AD)

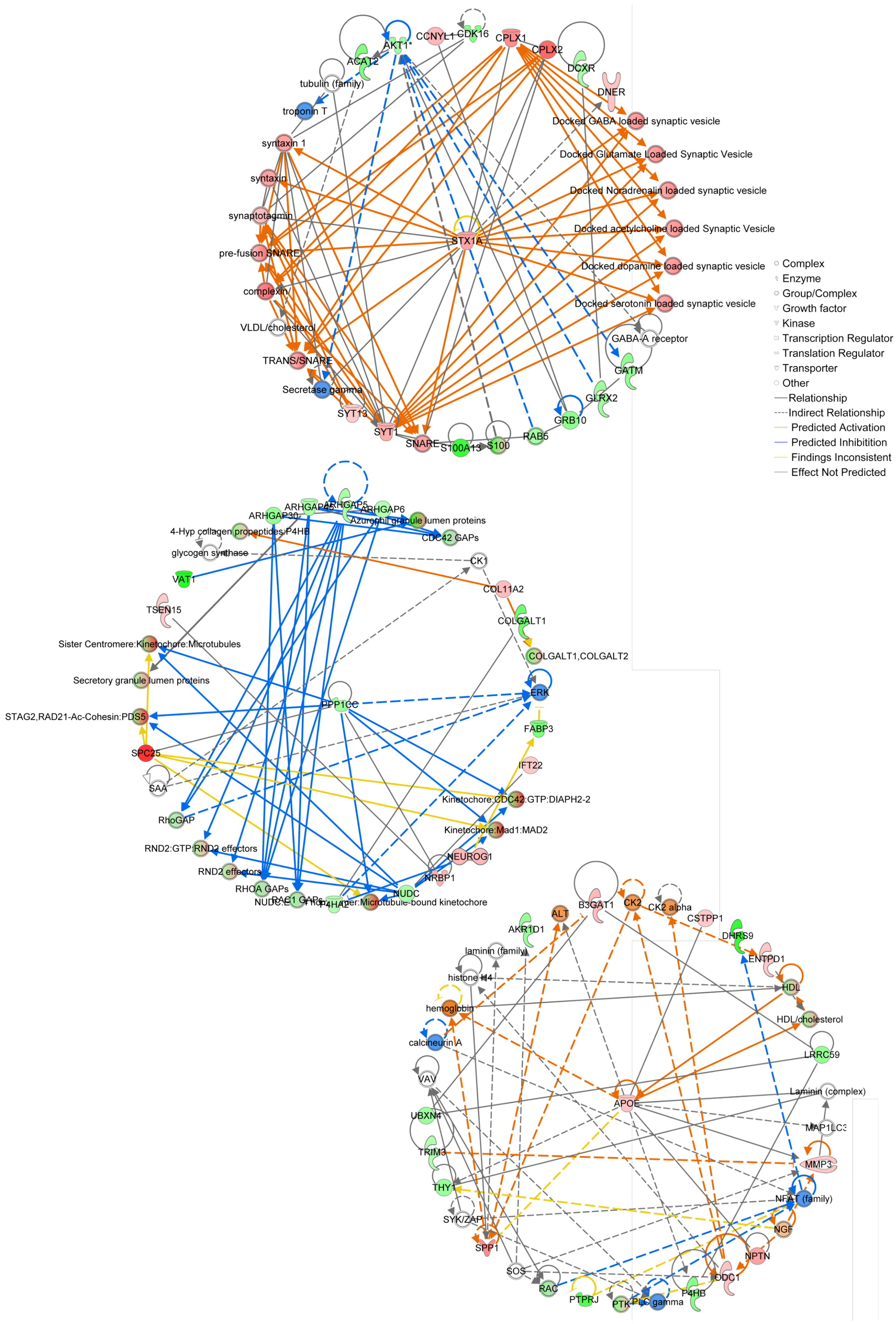
Supplementary Figure 3. Effect Size Comparisons with Other Studies.
Effect size comparisons between the current study and previous studies are shown for overlapping proteins. Pearson correlation coefficients and p-values across studies are noted in each panel. The central dotted line in each plot represents the correlation pattern, while the outer lines mark the 95% confidence interval.
(A) Comparison with Ali*, et al.*[1]
(B) Comparison with Walker*, et al.*[2] – Walker*, et al.*[2] analyzed SomaScan 5K data from 4,110 samples (including 428 dementia cases) and identified 38 proteins significantly associated with incident dementia after multiple testing correction. Of the 456 aptamers (416 proteins) identified in our study, 319 overlapped with their findings. Among these, three aptamers were significant after correction, 52 showed nominal association, and 176 (55%) exhibited a consistent direction between the two studies. The 52 nominally significant aptamers showed an effect size correlation with a Pearson coefficient of 0.872 ($p = 4.16\times10^{-17}$). Notably, their study followed cognitively normal individuals, some of whom later developed dementia over a follow-up period of 5 to 25 years, yet the correlation between the results remains high.
(C) Comparison with Sung*, et al.*[3] – Sung*, et al.*[3] analyzed SomaScan 1.3K data from 105 AD cases and 254 controls, identifying 26 proteins associated with AD after Bonferroni correction. Among the 456 aptamers identified in our study, 67 overlapped with their findings. Of these, two were significant after multiple testing correction, 34 were nominally significant, and 52 (78%) showed a consistent direction. The 34 nominally significant aptamers exhibited a strong correlation (Pearson = 0.805, $p = 9.42\times10^{-9}$).
(D) Comparison with Sattlecker*, et al.*[4] – Sattlecker*, et al.*[4] analyzed SomaScan data measuring 1,001 proteins in 331 AD cases and 211 controls, identifying 138 nominally significant proteins and four significant proteins after FDR correction. Of the 456 aptamers in our study, 76 (84%) overlapped with their findings. Among them, two were significant after multiple testing correction, 22 were nominally significant, and 64 exhibited a consistent direction. The effect size correlation for the 22 nominally significant aptamers was Pearson = 0.811 ($p = 4.72\times10^{-9}$).
(E) Comparison with Jiang*, et al.*[5] – Jiang*, et al.*[5] analyzed plasma proteomics data using the Olink platform, measuring 1,160 proteins in 106 AD cases and 74 controls. They identified 429 proteins significantly associated with AD after FDR correction. In our study, 37 proteins overlapped with their findings, of which 34 (92%) showed a consistent direction and passed multiple testing correction. The correlation between effect sizes in the two studies was highly significant (Pearson = 0.669).
(F) Comparison with Whelan*, et al.*[6] – Whelan*, et al.*[6] analyzed Olink proteomics data from a cohort of 557 cognitively normal individuals (Aβ+ and Aβ-) and 161 Aβ+ AD cases, identifying 33 proteins significantly associated with AD after FDR correction. Among the 416 proteins (456 aptamers) identified in our study, 22 overlapped with their results, showing a significant effect size correlation ($p = 3.63\times10^{-2}$). Among these, four proteins passed multiple testing correction.
(G) Comparison with Guo*, et al.*[7] – Guo*, et al.*[7] leveraged the Olink platform to measure 1,463 proteins in 51,228 controls and 691 incident AD cases from the UK Biobank. They identified 16 proteins associated with dementia risk after FDR correction. Similar to Walker et al., they employed a Cox regression model for survival analysis. In our study, 122 proteins overlapped with their findings, among which three passed multiple testing correction. Notably, 89 proteins (73%) showed a consistent direction ($p = 4.04\times10^{-7}$).
Abbreviations: R, correlation.


a
Not significant
Significant
SPC25
TBCA
S100A13
PHGDH
NEFL
FOXO1
ITPK1
LRRN1
CTF1
CPLX2
CPLX1
STX1A
NRBP1
ACHE
−log10(P value)
Effect size
b
R = 0.95 ; p = 1.49 x 10−59
Effect size (AT Status)
Effect size (Clinical diagnosis)
c
Not significant
Significant
SPC25
CTF1
CPLX2
PHGDH
LRRN1
ACHE
FOXO1
CPLX1
TBCA
SLIT2
NEFL
SMOC1
TINAGL1
−log10(P value)
Effect size
d
R = 0.94 ; p = 7.2 x 10−53
Effect size (Amyloid PET)
Effect size (Clinical diagnosis)
e
Not significant
Significant
CPLX2
SPC25
NEFL
−log10(P value)
Effect size
f
R = 0.81 ; p = 1.12 x 10−63
Effect size (Plasma pTau-217)
Effect size (Clinical diagnosis)

Supplementary Figure 4. Analysis of AD Biomarkers. Volcano plots and scatter plots illustrate the differences in protein abundances across key AD biomarkers. For each biomarker analysis, volcano plots highlight proteins with significantly different abundances between comparison groups, and scatter plots compare effect sizes with those from the meta-analysis of clinical status of AD. The central dotted line in the plots represents the correlation pattern, while the outer lines mark the 95% confidence interval. Color represents the significance (red, p-value < 0.05 without multiple testing correction; gray, not significant). **(A)** CSF-based AT status: Proteins with differing abundances between AT- and AT+ groups. **(B)** Amyloid PET imaging: Proteins with differing abundances between A- and A+ groups. **(C)** Plasma p-tau217: Proteins with differing abundances between T- and T+ groups.
Abbreviations: CSF, Cerebrospinal fluid; pTau, Phosphorylated tau; R, Correlation; PET, Positron emission tomography; AT, Amyloid-Tau.

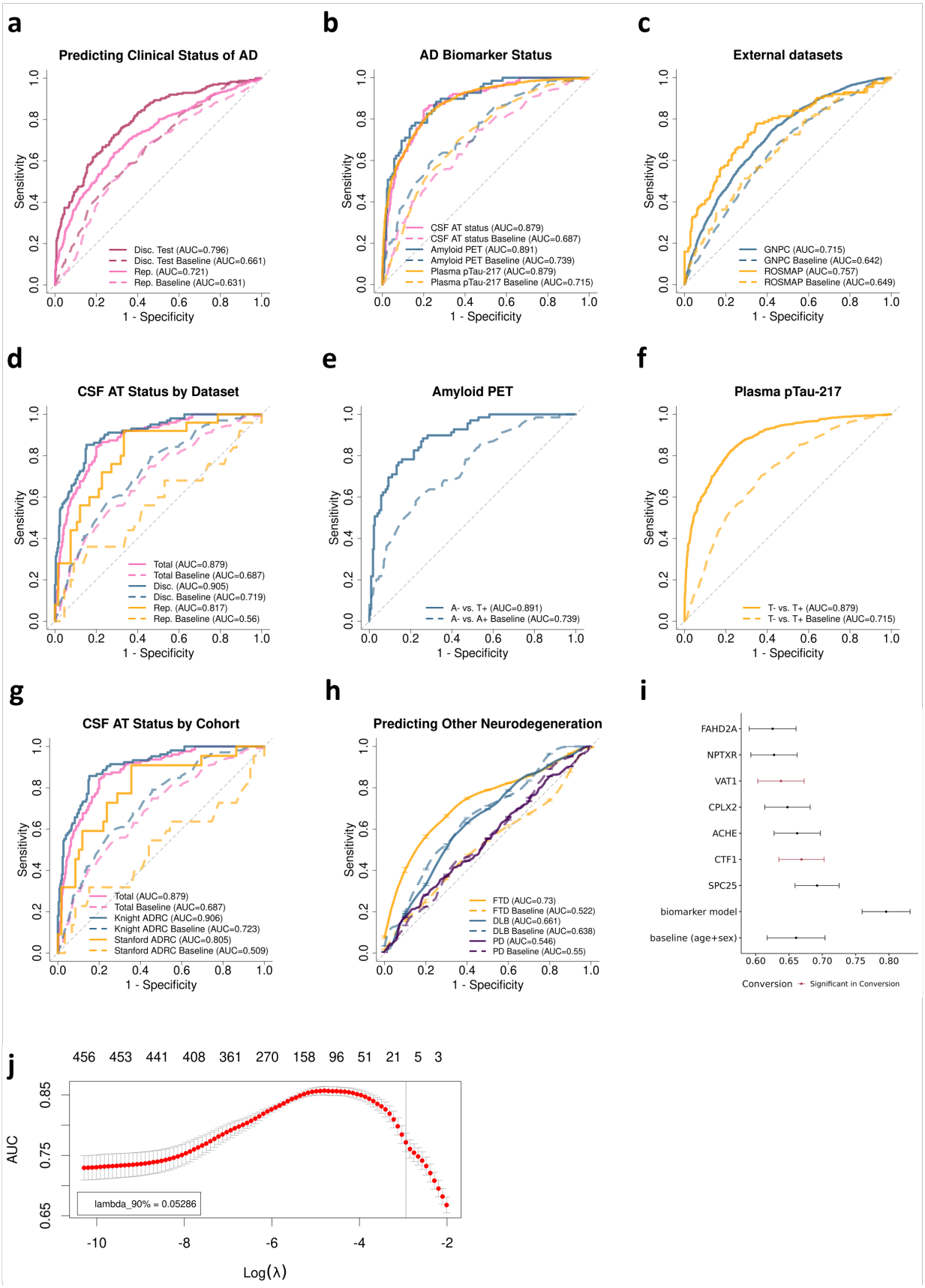


Supplementary Figure 5. Analysis of progression to symptomatic Alzheimer's Disease with single aptamers. (A) Venn diagram showing the overlap between proteins significant in the clinical diagnosis of AD (orange) and those significant in progression to symptomatic AD (gray). **(B)** Scatter plot comparing effect sizes between the meta-analysis (x-axis) and the analysis of progression to symptomatic AD (y-axis) for proteins significant in the progression to symptomatic AD analysis. Color represents the significance (red, FDR < 0.05 in meta-analysis and p < 0.05 in AD risk analysis; gray, not significant). **(C)** Forest plot showing the hazard ratios of the top 20 aptamers from the progression to symptomatic AD with Cox proportional hazards model.

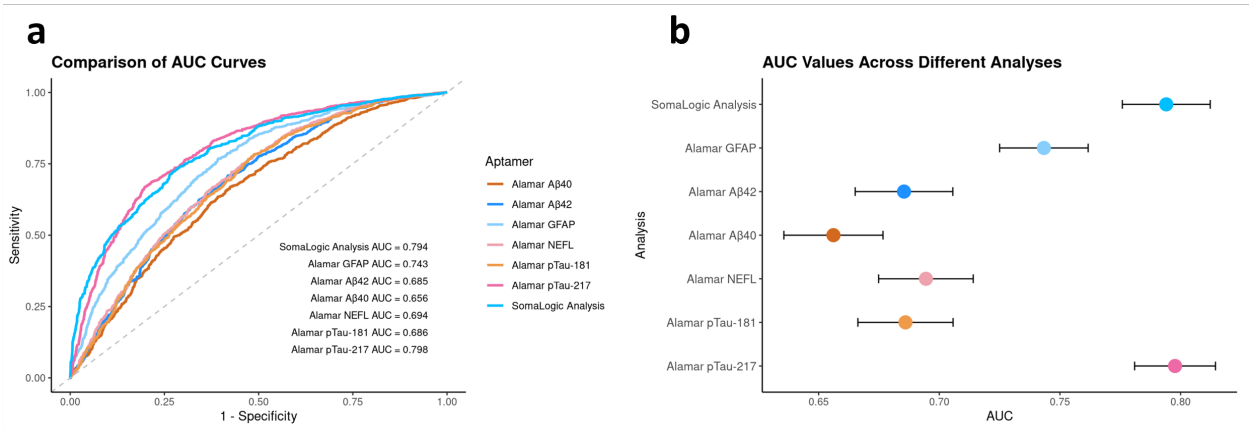


AKT1*
CCNYL1
CDK16
CPLX1
CPLX2
ACAT2
DCXR
tubulin (family)
DNER
troponin T
Docked GABA loaded synaptic vesicle
syntaxin 1
Docked Glutamate Loaded Synaptic Vesicle
syntaxin
Docked Noradrenalin loaded synaptic vesicle
synaptotagmin
Docked acetylcholine loaded Synaptic Vesicle
STX1A
pre-fusion SNARE
Docked dopamine loaded synaptic vesicle
complexin
Docked serotonin loaded synaptic vesicle
VLDL/cholesterol
GABA-A receptor
TRANS/SNARE
GATM
Secretase gamma
GLRX2
SYT13
GRB10
SYT1
RAB5
SNARE
S100A13
S100
Complex
Enzyme
Group/Complex
Growth factor
Kinase
Transcription Regulator
Translation Regulator
Transporter
Other
Relationship
Indirect Relationship
Predicted Activation
Predicted Inhibitition
Findings Inconsistent
Effect Not Predicted
ARHGAP5
ARHGAP30
ARHGAP35
ARHGAP6
Azurophil granule lumen proteins
4-Hyp collagen propeptides:P4HB
CDC42 GAPs
glycogen synthase
CK1
VAT1
COL11A2
TSEN15
COLGALT1
Sister Centromere:Kinetochore:Microtubules
COLGALT1,COLGALT2
Secretory granule lumen proteins
ERK
PPP1CC
STAG2,RAD21-Ac-Cohesin:PDS5
FABP3
SPC25
IFT22
SAA
Kinetochore:CDC42:GTP:DIAPH2-2
RhoGAP
Kinetochore:Mad1:MAD2
RND2:GTP:RND2 effectors
NEUROG1
NRBP1
RND2 effectors
RHOA GAPs
NUDC
NUDCL2
RAC1 GAPs
P4HA2
Trimer:Microtubule-bound kinetochore
B3GAT1
CK2
CK2 alpha
ALT
CSTPP1
AKR1D1
DHRS9
laminin (family)
ENTPD1
histone H4
HDL
hemoglobin
HDL/cholesterol
calcineurin A
LRRC59
VAV
APOE
Laminin (complex)
UBXN4
MAP1LC3
TRIM3
MMP3
THY1
NFAT (family)
SYK/ZAP
NGF
SPP1
NPTN
SOS
ODC1
RAC
P4HB
PTPRJ
PTK
PLC gamma

Supplementary Figure 6. Interaction Networks Identified by IPA Analysis. Three interaction networks identified using Ingenuity Pathway Analysis (IPA). The color-coded nodes represent various types of molecules and edges indicate predicted interactions between them.


a
Predicting Clinical Status of AD
Sensitivity
1 - Specificity
Disc. Test (AUC=0.796)
Disc. Test Baseline (AUC=0.661)
Rep. (AUC=0.721)
Rep. Baseline (AUC=0.631)
b
AD Biomarker Status
CSF AT status (AUC=0.879)
CSF AT status Baseline (AUC=0.687)
Amyloid PET (AUC=0.891)
Amyloid PET Baseline (AUC=0.739)
Plasma pTau-217 (AUC=0.879)
Plasma pTau-217 Baseline (AUC=0.715)
c
External datasets
GNPC (AUC=0.715)
GNPC Baseline (AUC=0.642)
ROSMAP (AUC=0.757)
ROSMAP Baseline (AUC=0.649)
d
CSF AT Status by Dataset
Total (AUC=0.879)
Total Baseline (AUC=0.687)
Disc. (AUC=0.905)
Disc. Baseline (AUC=0.719)
Rep. (AUC=0.817)
Rep. Baseline (AUC=0.56)
e
Amyloid PET
A- vs. T+ (AUC=0.891)
A- vs. A+ Baseline (AUC=0.739)
f
Plasma pTau-217
T- vs. T+ (AUC=0.879)
T- vs. T+ Baseline (AUC=0.715)
g
CSF AT Status by Cohort
Total (AUC=0.879)
Total Baseline (AUC=0.687)
Knight ADRC (AUC=0.906)
Knight ADRC Baseline (AUC=0.723)
Stanford ADRC (AUC=0.805)
Stanford ADRC Baseline (AUC=0.509)
h
Predicting Other Neurodegeneration
FTD (AUC=0.73)
FTD Baseline (AUC=0.522)
DLB (AUC=0.661)
DLB Baseline (AUC=0.638)
PD (AUC=0.546)
PD Baseline (AUC=0.55)
i
FAHD2A
NPTXR
VAT1
CPLX2
ACHE
CTF1
SPC25
biomarker model
baseline (age+sex)
Conversion
Significant in Conversion
j
456 453 441 408 361 270 158 96 51 21 5 3
AUC
lambda_90% = 0.05286
Log(λ)

Supplementary Figure 7. Evaluation of the Model's Predictive Performance. (A) Receiver Operating Characteristic (ROC) curves evaluating the predictive power in classifying clinical Alzheimer's disease (AD). The following datasets were used: test set from the Discovery dataset with the proteomic model (solid black) and baseline model (dashed black), and the Replication dataset with the proteomic model (solid orange), baseline model (dashed orange), and baseline model including APOE ε4 genotype (pink).
(B) ROC curves evaluating the model's predictive power in classifying AD biomarkers. The biomarkers included: CSF-based AT status with the proteomic model (solid orange) and baseline model (dashed orange), amyloid PET imaging with the proteomic model (solid pink) and baseline model (dashed pink), and plasma p-tau217 with the proteomic model (solid blue) and baseline model (dashed blue).
(C) ROC curves evaluating the model's predictive power external datasets. ROSMAP (solid orange) and baseline model (dashed orange), GNPC (solid blue) and baseline model (dashed blue).
(D) ROC curves evaluating the model's predictive power for CSF-based AT status (AT- vs. AT+) by dataset. The following datasets were used: both datasets with the proteomic model (solid orange) and baseline model (dashed orange), discovery dataset with the proteomic model (solid pink) and baseline model (dashed pink), and replication dataset with the proteomic model (solid blue) and baseline model (dashed blue).
(E) ROC curves evaluating the model's predictive power in classifying amyloid PET imaging status (A- vs. A+) with the proteomic model (solid pink) and baseline model (dashed pink).
(F) ROC curves assessing the model's predictive power in classifying plasma p-tau217 status (T- vs. T+) with the proteomic model (solid blue) and baseline model (dashed blue).
(G) ROC curves evaluating the model's predictive power in classifying CSF-based AT status (AT- vs. AT+) by cohort. The following cohorts were analyzed: both cohorts with the proteomic model (solid orange) and baseline model (dashed orange), Knight ADRC with the proteomic model (solid pink) and baseline model (dashed pink), and Stanford ADRC with the proteomic model (solid blue) and baseline model (dashed blue).
(H) ROC curves evaluating the predictive power in classifying other neurodegenerative diseases. The following diseases were analyzed: DLB with the proteomic model (solid orange) and baseline model (dashed orange), FTD with the proteomic model (solid pink) and baseline model (dashed pink), and PD with the proteomic model (solid blue) and baseline model (dashed blue).
(I) Whisker plot illustrating the Area Under the Curve (AUC) for 7 aptamers used in the predictive model, along with the AUCs for both the baseline and proteomic models. Error bars indicate the 95% confidence interval of the AUC for each model.
(J) Plot illustrating the change in mean AUC from cross-validation as the lambda value for lasso regression changes. The numbers on the upper x-axis indicate the number of selected variables in lasso regression, with the selected lambda value marked by a dotted line. Error bars indicate standard deviation for the AUC across 10-fold cross-validation.
Abbreviations: AD, Alzheimer's disease; Disc, Discovery; Rep, Replication; CSF, Cerebrospinal fluid; PET, Positron emission tomography; pTau, Phosphorylated tau; FTD, Frontotemporal dementia; DLB, Dementia with Lewy bodies; PD, Parkinson's disease; Lambda_90%, 90th percentile of the lambda distribution.

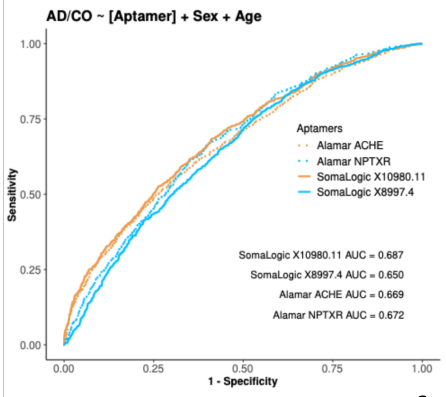


Supplementary Figure 8. AUCs with prediction models using different biomarkers. (A) The plot provides a comparison of Receiver Operating Characteristic curves for various aptamers in Alamar platform (pTau-217, pTau-181, NEFL, Aβ40, Aβ42, and GFAP) in distinguishing Alzheimer's Disease (AD) from cognitively normal (CO) controls. The curves were generated by fitting logistic regression models that include age and sex as covariates. The SomaLogic Analysis represents the model used in our study. **(B)** The whisker plot displays the AUC values for different aptamers, showing the performance of each in terms of classification ability (as evaluated by AUC). Horizontal lines (whiskers) indicate the confidence intervals for each AUC value. pTau-217 and our model both show relatively high AUC values (0.798 and 0.794, respectively), suggesting strong classification performance.

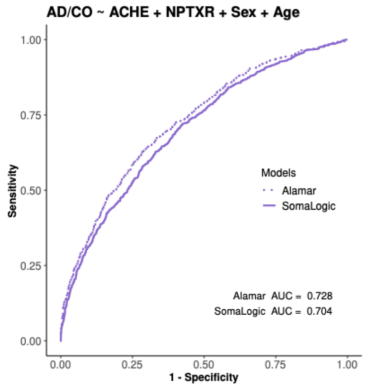
a

Platform	Group	#Samples	Age(SD)	% male	% APOE4+
Knight ADRC	CO	1595	72.94 (±10.73)	40.00	31.10
	AD	1166	77.63 (±8.81)	43.48	57.03
Stanford ADRC	CO	596	71.93 (±7.35)	41.43	23.49
	AD	94	70.67 (±11.19)	44.68	63.83

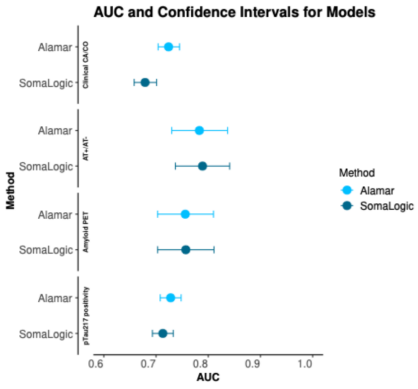
b



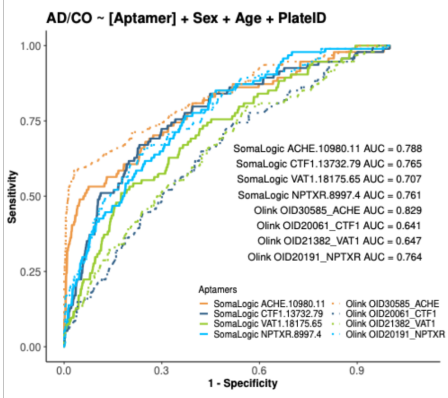
c



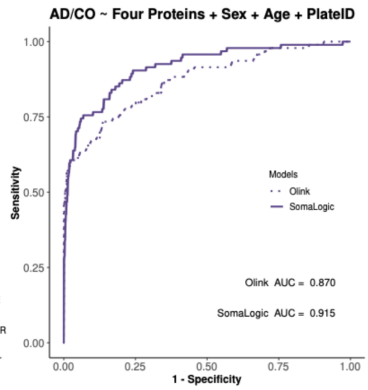
d



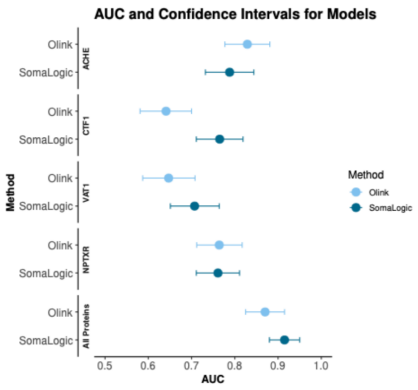
e



f



g



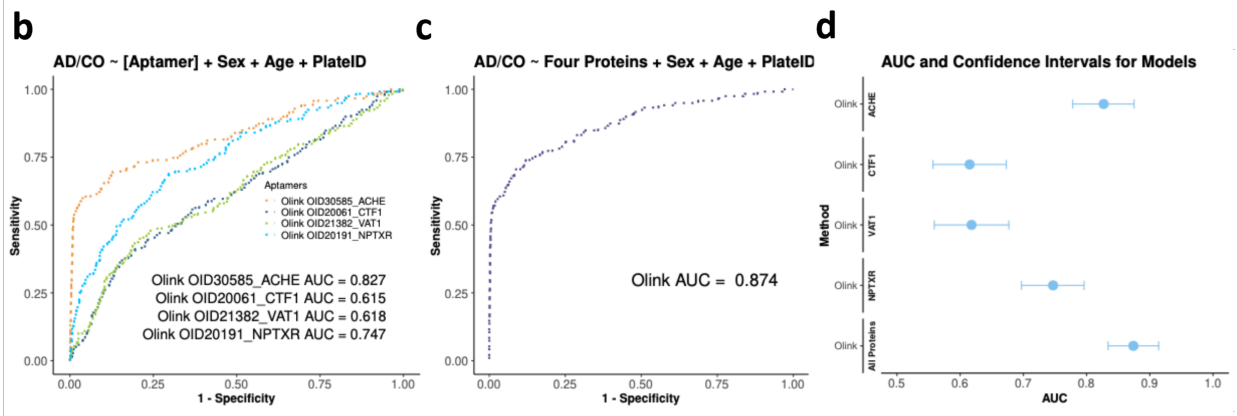
h

Cohort	Platform	Aptamer	AUC	PPV	NPV	Specificity	Sensitivity	Accuracy
Knight ADRC	SomaLogic	ACHE (X10980.11)	0.687	0.527	0.736	0.588	0.685	0.627
		NPTXR (X8997.4)	0.650	0.511	0.720	0.473	0.749	0.590
		ACHE + NPTXR	0.704	0.561	0.728	0.670	0.627	0.653
	Alamar	ACHE	0.669	0.532	0.711	0.702	0.543	0.641
		NPTXR	0.672	0.549	0.711	0.604	0.663	0.628
		ACHE + NPTXR	0.728	0.581	0.754	0.719	0.625	0.683
Stanford ADRC	SomaLogic	ACHE (10980.11)	0.788	0.532	0.925	0.925	0.532	0.871
		CTF1 (13732.79)	0.765	0.317	0.936	0.769	0.670	0.782
		VAT1 (18175.65)	0.707	0.318	0.914	0.825	0.511	0.744
		NPTXR (8997.4)	0.761	0.247	0.944	0.628	0.766	0.786
		All Four Proteins	0.915	0.640	0.960	0.932	0.755	0.908
	Olink	ACHE (OID30585)	0.829	0.753	0.936	0.969	0.585	0.917
		CTF1 (OID20061)	0.641	0.197	0.912	0.579	0.649	0.723
		VAT1 (OID21382)	0.647	0.185	0.932	0.440	0.798	0.802
		NPTXR (OID20191)	0.764	0.300	0.933	0.754	0.660	0.873
		All Four Proteins	0.870	0.463	0.953	0.864	0.734	0.846

Supplementary Figure 9. Orthogonal validation of the SomaLogic-based model on the Alamar and Olink platforms. (A) Demographic information for the Knight ADRC and Stanford ADRC cohorts, including details on age (with standard deviation), percentage of males, and proportion of APOE4+ individuals for each group. **(B–D)** Comparison of the predictive power of proteins between SomaLogic and Alamar platforms. **(B)** ROC curves showing the predictive power of single-protein models, adjusted for sex and age. **(C)** ROC curves for two-protein models, adjusted for sex and age. **(D)** Whisker plot showing AUC values for two-protein models in predicting CSF AT status, amyloid PET status, and pTau217 positivity, with adjustments for sex and age. The whisker plot includes 95% confidence intervals. **(E–G)** Comparison of the predictive power of proteins between SomaLogic and Olink platforms. **(E)** ROC curves comparing single-protein models, adjusted for sex, age, and plate ID. **(F)** ROC curves for four-protein models, adjusted for sex, age, and plate ID. **(G)** Whisker plot showing AUC values for single-protein models, adjusted for sex, age, and plate ID, with 95% confidence intervals. **(H)** Statistical metrics from the replication analysis, including AUC, positive predictive value (PPV), negative predictive value (NPV), specificity, sensitivity, and accuracy, are provided for each model. Abbreviations: AUC, Area under the curve; PPV, Positive predictive value; NPV, Negative predictive value.

a

Platform	Group	#Samples	Age(SD)	% male	% APOE4+
Olink	CO	725	72.10 (±7.50)	40.67	20.83
	AD	120	71.00 (±11.13)	44.17	59.17



e

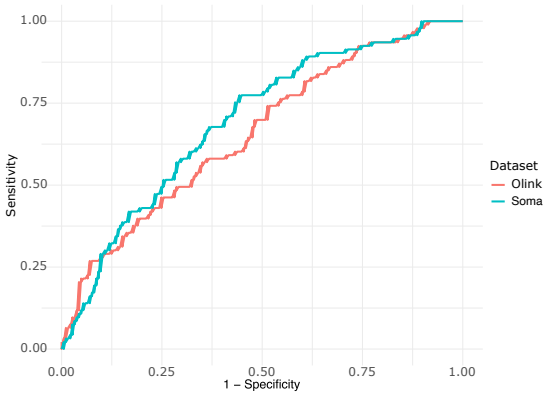
Cohort	Platform	Aptamer	AUC	PPV	NPV	Specificity	Sensitivity	Accuracy
Stanford ADRC	Olink	ACHE (OID30585)	0.827	0.477	0.945	0.872	0.697	0.847
		CTF1 (OID20061)	0.615	0.291	0.890	0.850	0.370	0.781
		VAT1 (OID21382)	0.618	0.265	0.896	0.790	0.454	0.742
		NPTXR (OID20191)	0.747	0.280	0.931	0.704	0.689	0.710
		All Four Proteins	0.874	0.506	0.953	0.879	0.739	0.859

Supplementary Figure 10. Orthogonal validation of the SomaLogic-based model on the Olink platform in Stanford ADRC cohort. (A) Demographic information for the Olink platform data in the Stanford ADRC cohort, including age (with standard deviation), percentage of males, and proportion of APOE4+ individuals for each group. **(B)** ROC curves of individual aptamers in distinguishing AD from CO, with all models adjusted for sex, age, and Plate ID. **(C)** ROC curves of a four-protein model on the Olink platform, adjusted for sex, age, and Plate ID. **(D)** Whisker plot showing AUCs of individual-protein and four-protein models across platforms for predicting clinical status of AD with 95% confidence intervals. **(E)** Statistical metrics from the replication analysis, including AUC, positive predictive value (PPV), negative predictive value (NPV), specificity, sensitivity, and accuracy, are provided for each model. Abbreviations: AUC, Area under the curve; PPV, Positive predictive value; NPV, Negative predictive value.

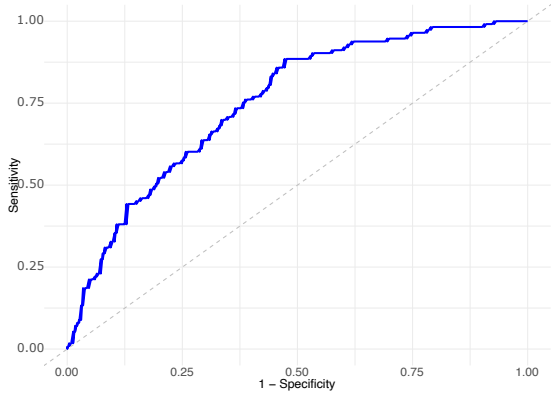
a

Cohort	Platform	CSF A Status	#Samples	Age(SD)	% male
BBRC	Olink	A-	260	60.47 (±4.47)	38.1
		A+	132	62.13 (±4.98)	40.2

b



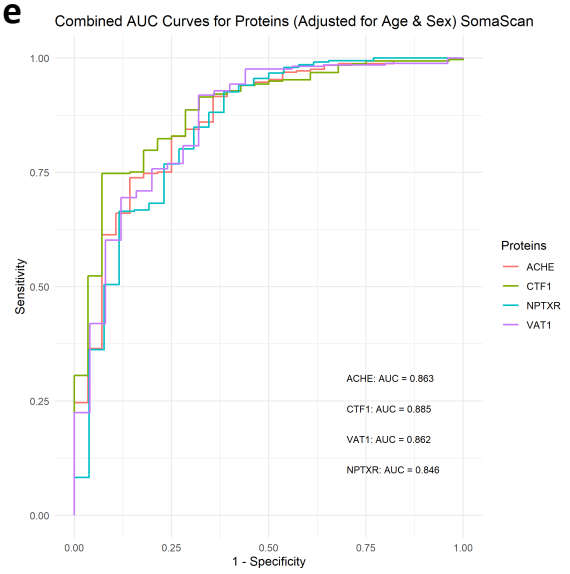
c



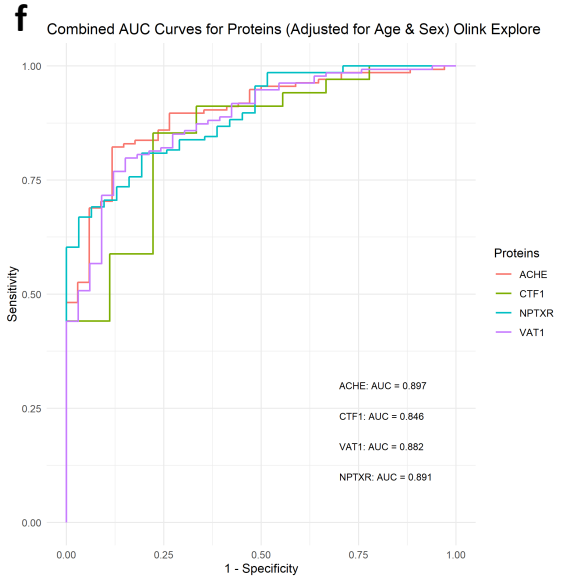
d

Cohort	Platform	Group	#Samples	Age(SD)	% male
ACE	SomaScan	CO	44	65.2 (±7.66)	31.8
		Dementia	388	74.9 (±7.45)	35.1
		MCI	363	72.8 (±8.43)	44.4
	Olink	CO	42	66.3 (±6.30)	47.6
		Dementia	4	72.9 (±13.0)	50.0
		MCI	460	72.8 (±7.62)	50.0

e



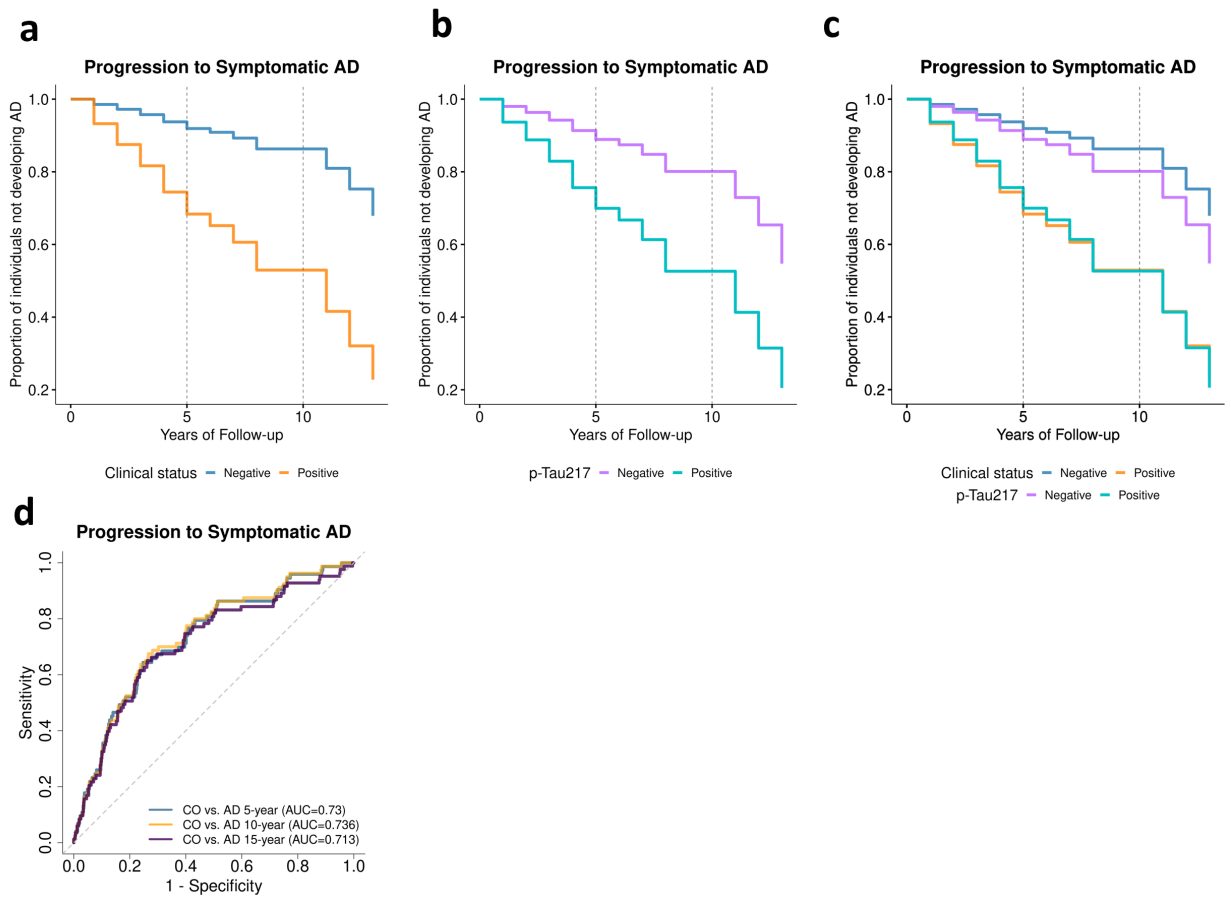
f



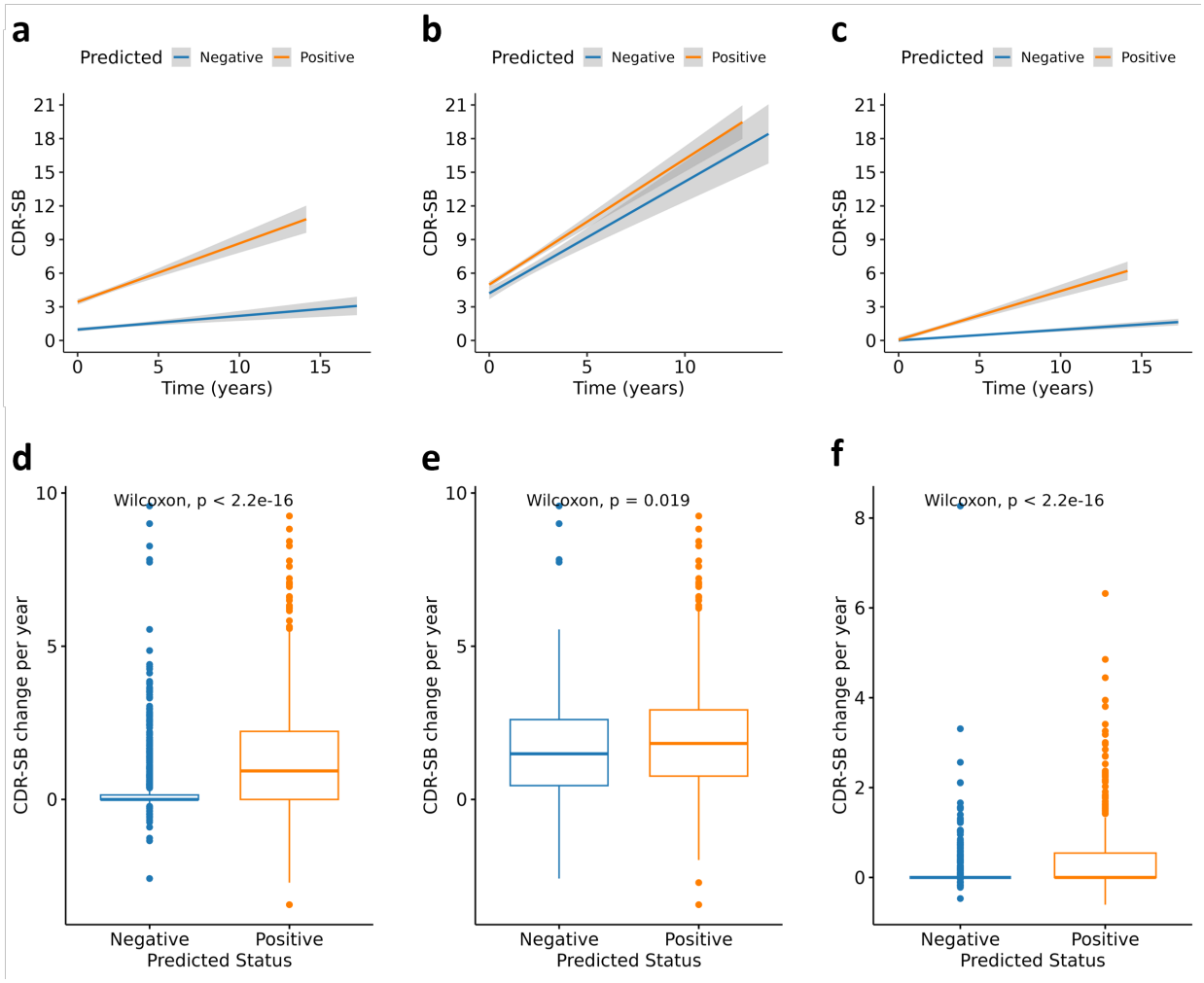
g

Cohort	Platform	Aptamer	AUC	PPV	NPV	Specificity	Sensitivity	Accuracy
ACE	Olink	ACHE	0.897	0.965	0.556	0.882	0.822	0.834
		CTF1	0.846	0.935	0.583	0.778	0.853	0.837
		VAT1	0.882	0.963	0.483	0.879	0.769	0.790
		NPTXR	0.891	0.989	0.400	0.968	0.669	0.725
	Soma	ACHE	0.863	0.983	0.222	0.857	0.738	0.748
		CTF1	0.885	0.992	0.245	0.929	0.748	0.762
		VAT1	0.862	0.975	0.386	0.680	0.919	0.903
		NPTXR	0.846	0.987	0.169	0.885	0.665	0.680

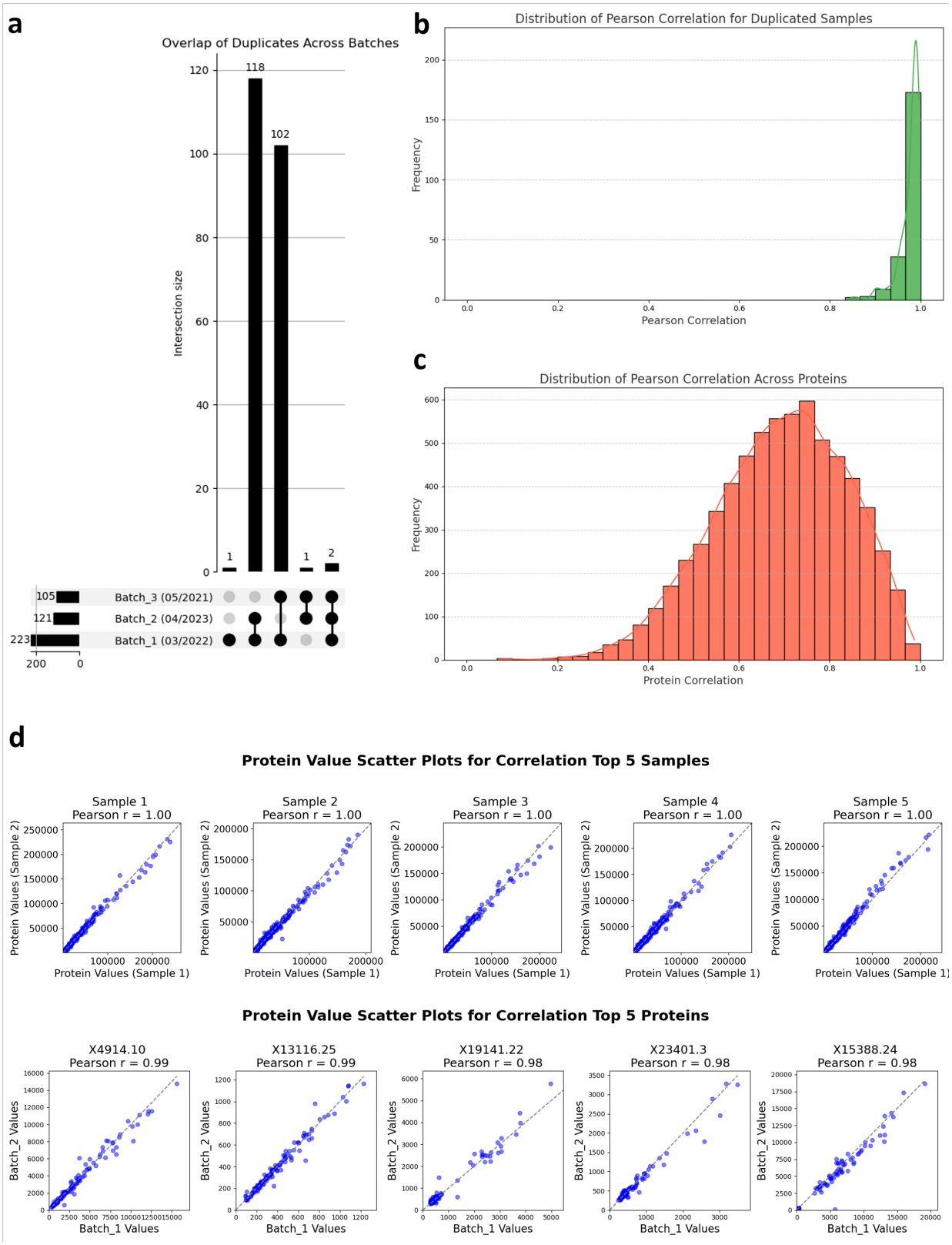
Supplementary Figure 11. Orthogonal validation of the SomaLogic-based model on the Olink platform in BBRC & ACE cohort. (A) Demographic information for the Olink platform data in the BBRC cohort, including age (with standard deviation) and percentage of males. **(B)** ROC curves of a four-protein model on both the Soma and Olink platform, adjusted for sex and age. **(C)** ROC curves of a seven-protein model on the Soma platform, adjusted for sex and age. **(D)** Demographic information for the Olink platform data in the ACE cohort, including age (with standard deviation) and percentage of males. **(E, F)** ROC curves of single-protein model on both the Soma and Olink platform, adjusted for sex and age. **(G)** Statistical metrics from the replication analysis, including AUC, positive predictive value (PPV), negative predictive value (NPV), specificity, sensitivity, and accuracy, are provided for each model. Abbreviations: AUC, Area under the curve; PPV, Positive predictive value; NPV, Negative predictive value.



Supplementary Figure 12. Progression Analysis Based on Different Biomarkers. Kaplan-Meier survival curves illustrate the progression to AD after the initial blood draw based on different biomarker groups, illustrating the proportion of participants from each group remaining cognitively normal over the follow-up period. **(A)** Predicted negative group (blue) and predicted positive group (orange). **(B)** T- group (purple) and T+ group (turquoise) from plasma pTau-217. **(C)** A merged plot displaying all four groups (predicted negative, predicted positive, T-, and T+) for a comprehensive comparison of progression patterns. **(D)** ROC curves evaluating the model's predictive power in classifying progression to symptomatic AD, with three follow-up periods: 5-year (orange), 10-year (pink), and 15-year (blue). Abbreviations: pTau, Phosphorylated tau.



Supplementary Figure 13. Analysis of progression in cognitive impairment and progression to symptomatic AD with predictive mode. (A-C) Changes in Cognitive Dementia Rating-Sum of Boxes (CDR-SB) scores (y-axis) over time (x-axis) in **(A)** both groups, **(B)** the Alzheimer's disease (AD) group, and **(C)** the cognitively normal group, as categorized by the prediction model. The shaded region represents the 95% confidence interval (CI) around the fitted linear regression line. **(D-F)** Boxplots showing the rate of CDR-SB changes per year in **(D)** both groups, **(E)** the AD group, and **(F)** the cognitively normal group, categorized by the prediction model. The center line represents the median, the box boundaries indicate the interquartile range (IQR, 25th–75th percentile), and the whiskers extend to 1.5×IQR from the first and third quartiles. Data points outside this range are considered outliers and are displayed as individual dots.
Abbreviations: CDR-SB, Clinical Dementia Rating Scale-Sum of Boxes; CO, Control; AD, Alzheimer's disease.



Supplementary Figure 14. Validation of technical and biological variance. (A) Upset plot showing overlap across multiple batches (Batch_1, Batch_2, Batch_3). **(B)** Distribution of correction factors for

duplicated samples. **(C)** Distribution of correction factors across 7k proteins. **(D)** Top 5 significant correction factors at both the duplicated sample level and protein level.

Reference:

1 Ali, M. *et al.* Multi-cohort cerebrospinal fluid proteomics identifies robust molecular signatures across the Alzheimer disease continuum. *Neuron* (2025). https://doi.org:10.1016/j.neuron.2025.02.014

2 Walker, K. A. *et al.* Large-scale plasma proteomic analysis identifies proteins and pathways associated with dementia risk. *Nat Aging* **1**, 473-489 (2021). https://doi.org:10.1038/s43587-021-00064-0

3 Sung, Y. J. *et al.* Proteomics of brain, CSF, and plasma identifies molecular signatures for distinguishing sporadic and genetic Alzheimer's disease. *Sci Transl Med* **15**, eabq5923 (2023). https://doi.org:10.1126/scitranslmed.abq5923

4 Sattlecker, M. *et al.* Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* **10**, 724-734 (2014). https://doi.org:10.1016/j.jalz.2013.09.016

5 Jiang, Y. *et al.* Large-scale plasma proteomic profiling identifies a high-performance biomarker panel for Alzheimer's disease screening and staging. *Alzheimers Dement* **18**, 88-102 (2022). https://doi.org:10.1002/alz.12369

6 Whelan, C. D. *et al.* Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease. *Acta Neuropathol Commun* **7**, 169 (2019). https://doi.org:10.1186/s40478-019-0795-2

7 Guo, Y. *et al.* Plasma proteomic profiles predict future dementia in healthy adults. *Nat Aging* **4**, 247-260 (2024). https://doi.org:10.1038/s43587-023-00565-0