Brain cross-omics integration in Alzheimer's disease

Abdallah M. Eteleeb[1,2], Brenna C. Novotny[1], Carolina Soriano Tarraga[1], Christopher Sohn[1], Eliza Dhungel[3], Logan Brase[1], Aasritha Nallapu[1], Jared Buss[1], Fabiana Farias[1,4], Kristy Bergmann[1,4], Joseph Bradley[1,4], Joanne Norton[1,4], Jen Gentsch[1,4], Fengxian Wang[1,4], Albert A. Davis[5,6], John C. Morris[2,5,6], Celeste M. Karch[1,2,4,6], Richard J. Perrin[2,5,6,7], Bruno A. Benitez[¶8], Oscar Harari[1,2,6*¶].

[1]Department of Psychiatry, Washington University, Saint Louis, St. Louis, Missouri, United States of America

[2]The Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington University, St. Louis, Missouri, United States of America

[3]Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, North Carolina, United States of America

[4]NeuroGenomics and Informatics Center, Washington University, St. Louis, Missouri, United States of America

[5]Department of Neurology, Washington University, St. Louis, Missouri, United States of America

[6]Hope Center for Neurological Disorders, Washington University, St. Louis, Missouri, United States of America

[7]Department of Pathology and Immunology, Washington University, St. Louis, Missouri, United States of America

[8]Department of Neurology, Harvard Medical School and Beth Israel Deaconess Medical Center, Boston, Massachusetts, United States of America

* Corresponding author

E-mail: harario@wustl.edu

[¶] Equally contributing senior author

Abstract

Unbiased data-driven omic approaches are revealing the molecular heterogeneity of Alzheimer disease. Here, we used machine learning approaches to integrate high-throughput bulk and single-nucleus transcriptomic, proteomic, metabolomic, and lipidomic profiles with clinical and neuropathological data from multiple AD cohorts. We discovered four unique multimodal molecular profiles, one showing signs of poor cognitive function, a faster pace of disease progression, shorter survival with the disease, severe neurodegeneration and astrogliosis, and reduced levels of metabolomic profiles. This profile shows similar cellular and molecular profiles in multiple affected cortical regions associated with higher Braak tau scores and significant dysregulation of synapse-related genes and endocytosis, phagosome, mTOR signaling pathways altered in AD early and late stages. The multimodal clusters uncovered cerebrospinal fluid biomarkers to monitor AD progression. AD cross-omics data integration with transcriptomic data from an SNCA mouse model revealed an overlapping signature. Our cross-omics analyses provide novel critical molecular insights into AD.

Introduction

Alzheimer disease (AD) is a heterogeneous multifactorial neurodegenerative disorder pathologically characterized by amyloid (Aβ) plaques, neurofibrillary tangles (NFTs), neuroinflammation, and synaptic and neuronal loss. Recently, multiple clinical and pathological AD subtypes have been identified based on distinct spatiotemporal trajectories of tau pathology, brain atrophy, postmortem brain transcriptomics profiles, or cerebrospinal fluid proteomics. Differences in characteristics associated with these subtypes including age at onset, sex, APOE ε4 carrier status, cognitive status, disease duration, and cerebrospinal fluid (CSF) biomarker levels, were previously reported [1–9]. Related research efforts have leveraged molecular profiles and biomarkers to characterize and cluster the AD population into subgroups and to provide aspects of AD progression that, in turn, could inform clinical trials besides cognitive function or amyloid tau biomarkers. Novel biomarkers for AD biological processes and cognition, in general, can shape new clinical trials and are excruciatingly needed for those therapies targeting amyloid plaques of tau fibrils. However, the influence of such AD profiles on cognitive decline remains poorly understood and under-studied, representing a barrier to using this information effectively in clinical trials of potential disease-modifying therapies. In addition, previous studies have proposed multiple AD subtypes based on molecular features [5]. Despite the substantial effort in identifying molecular subtypes of AD described in this study, broader conclusions are limited by its nature as single-omic analyses typically capture changes in a single component of the biological cascade.

We have previously leveraged high-throughput brain molecular data from AD patients and healthy individuals using transcriptomic [10–12], proteomic [13], metabolomic [14] and single-cell omics [15] to study the signatures for AD. However, we recognize that omics layers are interdependent and interconnected. It has been hypothesized that by combining multiple signals from various

68 molecules, novel significant biological hubs or pathway changes in AD will emerge that are
69 otherwise missed. Previous studies that aggregate multiple modalities of omics (cross-omics)
70 have proven to be a powerful approach to identifying molecular subtypes for complex human
71 diseases, including cancer [16–21]. Different modalities of omics layers describe a biological
72 system at different biomolecular levels [22], allowing more comprehensive disease segregation
73 into molecular profiles that cannot be appreciated through an analysis restricted to a single
74 modality. Cross-omics approaches can reveal how complex biomolecular profiles change in
75 association with a disease and the relationships and correlations between distinct classes of
76 biological molecules. Their potential for studying neurodegeneration has been previously
77 described [23–32]. However, only recently has the deep molecular characterization of multiple
78 cortical regions from diverse brain banks and cohorts been generated [2,3], enabling a more
79 detailed and comprehensive examination of cross-omics data in AD.
80
81 In this study, we sought to investigate whether cross-omics studies of AD brains can provide novel
82 molecular insights into the progression of AD and its staging information. We leveraged machine
83 learning, digital deconvolution, and traditional statistical approaches to integrate and analyze
84 multiple high-throughput *omics* datasets from different cortical regions and cohorts. Our cross-
85 omics data integration approach identified four distinct molecular profiles of AD, one of which was
86 associated with worse cognitive function and neuropathological features, including significantly
87 higher Clinical Dementia Rating (CDR) at death, shorter survival after symptom onset, more
88 severe neurodegeneration and astrogliosis, and decreased levels of metabolomic profiles. The
89 molecular signatures of this profile, present in multiple cortical regions, show a significant
90 dysregulation of synapse-related genes and pathways, suggesting neuron/synapse losses and
91 dysfunction at later stages of AD. Integrating cross-omics approaches with data from established
92 mouse models of neurodegeneration clarified the role of these genes and pathways in the AD
93 pathophysiology. Furthermore, combining cross-omics data with single-cell resolution uncovered

molecular profiles associated with AD features to an unprecedented level of granularity and complexity. Subsequent AD staging and network analyses, as well as survival analyses of CSF proteins related to the synapses, revealed dysregulated of synaptic proteins associated with an increased risk of dementia progression including *IGF1*, *NRXN3*, and *YWHAZ* that can be employed as possible biomarkers for early synaptic dysfunction, cognition, and AD staging.

Our results suggest that cross-omics analyses capture molecular heterogeneity among brains from AD cases that are otherwise missed in non-clustered single-modality *omics* approaches, and demonstrate that, under molecular stratification, cross-omics approaches provide new perspectives of molecular pathways associated with AD that are reflected in associations of cognitive decline progression and peripheral tissue. These novel molecular findings may open the possibility for new biomarkers for the molecular staging of AD, as well as potential therapeutic targets to ameliorate cognitive decline and AD progression that may provide a foundation for implementing precision medicine approaches for AD.

Results

Study design

We leveraged high-throughput transcriptomic (Single-nucleus and bulk RNA-seq), proteomic, metabolomic, and lipidomic data from the parietal cortex of participants with neuropathology-confirmed AD from the Knight Alzheimer Disease Research Center (Knight ADRC) for our discovery cohort (see Methods, **Fig 1** and **S1 Table**). Omics data were generated using next-generation sequencing/10X genomics , SomaScan, and Metabolon Precision Metabolomics platforms for transcriptomics (60,0754 transcripts) [10–12,15], proteomics (1,092 proteins) [13], and metabolomics (627 metabolites) [14], respectively (see Methods). Preprocessing, quality

118 control (QC), and feature selection were performed on each omics data type, and only samples
119 that passed the QC criteria in all three omics modalities were retained, resulting in 278 samples
120 (255 AD and 23 control). We employed a Bayesian integrative clustering method (iClusterBayes
121 [33]) to integrate AD multi-omics data (see Methods) and identify molecular profiles of AD. Control
122 cases were later integrated for comparative analyses.
123
124 **Fig. 1. Study design**. The transcriptomics, proteomics, metabolomics, and lipidomics profiles of
125 postmortem parietal cortex samples from the Knight ADRC participants were used. The three
126 omics shared two hundred seventy-eight subjects (255 sporadic AD and 23 control cases).
127 Expression/reading matrices were prepared for the same set of samples (n=278) and different
128 numbers of features (genes=60,754, proteins=1,092, metabolites=627). Before the integration,
129 feature selection was performed on all omics datasets whose number of features exceeded 1000
130 features by selecting the top and most variant features. A Bayesian integrative clustering method
131 was then employed to integrate the three omics datasets, and the best clustering solution was
132 extracted based on the maximum deviance ratio and the minimum Bayesian information criterion
133 (BIC). AD molecular profiles were then linked to multiple clinical and molecular attributes to
134 examine whether an association with these attributes exists. In addition, multiple analyses,
135 including survival, differential expression, cell-type specific effect, and cell proportions inference
136 analyses, were performed on these profiles. These molecular profiles were then replicated in two
137 independent datasets from MSBB BM36 (AD=93, CO=23) and ROSMAP (AD=144, CO=93).
138 Dysregulated synaptic genes in the Knight-C4 profile were extracted and used for downstream
139 AD staging analyses and in the ROSMAP cohort (early-AD vs. control and late-AD vs. control).
140 Common synaptic genes present in the CSF SOMA dataset and associated with an increased
141 risk of dementia through survival analyses were identified as CSF synaptic biomarkers for AD
142 staging.
143

144 We next performed multiple association analyses, including association with clinical and
145 neuropathological attributes (e.g., sex, age of death, age at onset, post-mortem interval, CDR,
146 Braak amyloid stage, Braak neurofibrillary tangle stage), survival, and differential expression
147 analyses to characterize AD molecular profiles. We leveraged digital deconvolution approaches
148 to infer the cellular population structure using bulk RNA-seq data [12], and tested for association
149 of these profiles with cell-type abundances. We also integrated single-nuclei RNA-seq data
150 (snRNA-seq) from the parietal cortex of 67 Knight ADRC participants to determine the specific
151 cell-type effects.
152
153 For replication, we leveraged transcriptomics (bulk RNA-seq) and proteomics (tandem mass tag;
154 TMT) data from the parahippocampal gyri (BM36) of 116 brains (93 AD and 23 control, **S1 Table**)
155 from The Mount Sinai Brain Bank (MSBB) study [34] cohort, as well as transcriptomics data from
156 237 brains (144 AD and 93 control, **S1 Table**) dorsolateral prefrontal cortex (DLPFC) samples
157 from The Religious Orders Study and Memory and Aging Project (ROSMAP) cohort [35],
158 accessed through the AMP-AD Synapse portal. Preprocessing, QC, and feature selection were
159 performed like that used to generate the discovery datasets.
160
161 Furthermore, we stratified AD cases into early and late-stage AD based on clinical and
162 neuropathological rules from ROSMAP study (see Methods, **Fig 1**) and performed DE and co-
163 expression network analyses to identify the magnitude of the dysregulation of genes at distinct
164 AD stages and provided a deeper understanding of the underlying dynamics of AD and
165 approximate its progression over time (**Fig 1**). Similarly, we evaluated the association of protein
166 levels encoded by selected genes in CSF and disease risk progression (CDR change from 0 to
167 0.5) using survival analyses in CSF proteomic data from the Knight ADRC participants (**Fig 1**).
168

169 # A multimodal profile of AD brains is associated with poor
170 # cognitive function and molecular attributes

171 We initially identified distinct AD molecular profiles by optimizing the number of clusters needed
172 to capture the molecular heterogeneity present in AD brains from the Knight ADRC, independently
173 of any neuropathological or clinical data. (**S1 Fig**). We identified four clusters (**Fig 2A**) in which
174 the molecular profiles of Cluster4 (named Knight-C4 hereafter, n=42) exhibited a pronounced
175 dysregulation of genes, proteins, and metabolites. For instance, the transcriptomic profiles of the
176 top 75% of quantile genes that contribute significantly to the clustering solution show significant
177 dysregulation signatures associated with Knight-C4 (**Fig 2B**).
178
179 **Fig. 2. Cross-omics data integration identified four distinct molecular profiles of AD**
180 **associated with worse clinical outcomes and molecular attributes.** (A) Heatmaps of multi-
181 omics profiles of the top significant features. (B) Transcriptomic profiles **(**scaled) of the top genes
182 contributed significantly to the clustering, showing an overall dysregulation of brains in Knight-C4.
183 (C) The distribution of CDR scores across four clusters with Knight-C4 associated with high CDR
184 scores. (D) Kaplan-Meier plots showing Knight-C4 association with poor outcome for the age of
185 onset and age of death. (E) Boxplots showing the cell proportion from deconvolution analysis
186 across four clusters using bulk RNA-seq from four cell types. Knight-C4 is associated with a
187 significantly higher and lower proportion of astrocytes and neurons, respectively. (F) Boxplot
188 showing each cluster's first principal component of metabolomics profiles, with Knight-C2 and
189 Knight-C4 showing a lower metabolomic signature than other AD cases.
190
191 We then evaluated whether these molecular profiles were associated with genetic risk factors,
192 neuropathological status, or clinical and demographic variables. We found that AD cases in
193 Knight-C4 showed significantly higher Clinical Dementia Rating (CDR) scores at death ($p=5.2x10^{-}$

194 ^{3}; **Fig 2C**), shorter survival time following disease onset (8 months shorter, **Fig 2D**, p=4.2x10^{-3},
195 HR=1.6) and younger age at death (**Fig 2D**, p=4.2 x10^{-3}, HR=1.7) compared to AD cases in
196 Knight-C1-3. The Knight-C4 cluster was not associated with either AD polygenic risk score (**S2**
197 **Fig**), *APOE* ε4 allele carrier status (**S3A Fig**), or Braak stages (neurofibrillary tangle or amyloid)
198 (**S3B and S3C Fig**). However, Knight-C4 included more Braak tau stage VI cases than the other
199 clusters (**S3B Fig**)**.** The parietal cortex is affected in later stages of AD, which could explain the
200 lack of correlation of Knight-C4 cluster with Braak staging.
201
202 Next, we inferred the cellular population structure from bulk RNA-seq [12] leveraging digital
203 deconvolution methods (see Methods). Knight-C4 showed significantly higher astrocyte and lower
204 neuronal proportions (**Fig 2E** and **S2 Table**). We then evaluated whether a metabolomic profile
205 that we previously reported to be associated with AD duration [14] would reveal additional
206 molecular differences among these clusters. Lower values of this metabolomic profile (or
207 *eigenmetabolite*) were associated with a longer disease duration driven by earlier onset [14].
208 Knight-C2 and Knight-C4 exhibited significantly lower eigenmetabolite scores (**Fig 2F**) using
209 ordinal logistic regression (p=6.0x10^{-03}). Our approach identified new molecular profiles from AD
210 subtypes linked to cognitive test scores, the pace of disease progression, survival with the
211 disease, the proportion of neuronal loss, astrocytosis, and metabolic changes.
212

213 Knight-C4 profile is replicated consistently across multiple
214 independent cohorts and brain regions

215 Our analyses of the transcriptomics and proteomics data from AD parahippocampal gyrus (BM36)
216 from the MSBB study [34] (**S1 Table**) identified two distinct molecular profiles (**Fig 3A**). Cluster1
217 (MSBB-C1) recapitulated the transcriptomic signatures identified in Knight-C4. We also
218 ascertained the extent of overlap between the molecules dysregulated in Knight-C4 (**S3 Table**)

219 and MSBB-C1 (**S4 Table**) and determined a coincident profile of dysregulated genes (**Fig 3G-I**,
220 **S4A Fig**, $p < 2.2 \times 10^{-16}$ hypergeometric test and **S5 Table**). Furthermore, MSBB-C1 also exhibited
221 more severe dementia ($p=1.5\times10^{-3}$, **Fig 3B**) and estimated higher and lower proportions of
222 astrocytes ($p=5.5\times10^{-07}$) and neurons (3.2×10^{-09}), respectively (**Fig 3C** and **S2 Table**) compared
223 to other AD cases in MSBB. MSBB-C1 was also associated with higher Braak tau scores
224 ($p=3.9\times10^{-03}$, **S5A Fig**). Although not significant ($p > 0.05$), these donors show a trend suggesting
225 earlier age of death (**S5B Fig**). We did not identify a significant overlap of dysregulated proteins
226 between regions ($p > 0.05$, **S4B Fig**). We also observed that the molecular profile of MSBB-C1,
227 first detected in the parahippocampal gyrus (BM36), is consistently present in multiple cortical
228 regions, including the frontal cortex (BM10), superior temporal gyrus (BM22), and inferior frontal
229 gyrus (BM44) (**S5C Fig**).
230
231 **Fig. 3. Molecular profiles of Knight-C4 are replicated in two independent datasets.** (A)
232 Heatmaps of the transcriptomic and proteomic profiles of the top features from MSBB (BM36)
233 show two distinct clusters. (B) Boxplots showing MSBB-C1 associated with higher CDR scores,
234 replicating Knight-C4 in the knight ADRC. (C) Boxplots showing cell proportions inferred from bulk
235 RNA-seq from MSBB (BM36) using digital deconvolution. MSBB-C1 replicates Knight-C4 by
236 showing an association with significantly higher and lower proportions of astrocytes and neurons,
237 respectively. (D) Heatmap of the transcriptomic profiles of the top features from ROSMAP
238 (DLPFC) showing two distinct clusters. (E) Boxplots showing ROSMAP-C1 associated with lower
239 MMSE30 scores. (F) Boxplots showing cell proportion estimated from bulk RNA-seq from
240 ROSMAP (DLPFC) using digital deconvolution. Like MSBB, ROSMAP-C1 replicates Knight-C4
241 by showing an association with significantly higher and lower proportions of astrocytes and
242 neurons, respectively. (G) Venn diagram showing the common dysregulated genes in the
243 discovery and replicated cohorts. (H) Heatmaps of the mean expression of the common genes
244 across the discovery and replicated cohorts showing a clear, consistent expression pattern. (I)

245 Heatmap of the effect size detected in the discovery and replicated cohorts showing a high effect
246 size similarity across the three cohorts.

247

248 Likewise, our analyses of the RNA-seq data from the dorsolateral prefrontal cortex (DLPFC) from
249 the ROSMAP cohort [35] identified a similar subset of AD brain donors (**Fig 3D**). Relative to those
250 in Cluster2 (ROSMAP-C2), Cluster1 (ROSMAP-C1) recapitulated the transcriptomic profiles of
251 those in Knight-C4 (**Fig 3G-I**, **S4C Fig,** $p < 2.2 \times 10^{-16}$ hypergeometric test and **S5 and S6 Tables**).
252 They also showed higher tau tangle densities ($p=1.3\times10^{-02}$) and neurofibrillary tangle burdens
253 (PHF tau tangles, $p=2.1\times10^{-02}$; **S6A and S6B Fig**), significantly higher astrocyte ($p=3.1\times10^{-08}$) and
254 lower neuron ($p=9.7\times10^{-14}$) proportions (**Fig 3F** and **S2 Table**), and a non-significant ($p > 0.05$),
255 trend to higher plaque densities—Global burden of AD pathology (**S6C Fig**, see Methods and
256 Bennett *et al.*, 2012 [35] for details on how these neuropathological variables are identified).
257 ROSMAP-C1 also showed evidence of greater cognitive impairment and increased
258 neuropathological AD scores (MMSE; $p=4.4\times10^{-02}$; **Fig 3E**). Advanced AD patients with poor
259 cognitive function show similar cellular and molecular profiles in multiple affected cortical regions.
260 The molecular profiles associated with the severity of AD pathology can still vary even in the
261 advanced stages of the disease.

262

263 The Knight-C4 subjects exhibited a remarkable molecular 264 disarray

265 Differential expression (DE) analyses (see Methods, **S3, S7 and S8 Tables**) showed that Knight-
266 C4 has a ~10-fold increase in the number of significantly differentially expressed genes (DEGs)
267 relative to controls, in comparison to Knight-C1-C3 (**S9 Tables** and **S7A Fig**); furthermore, 90%
268 of the 4,661 DEGs were unique to participants in the Knight-C4 (**S9 Table** and **Fig 4A-green**

269 **bars**). Similarly, we observed many proteins and metabolites with significantly different
270 abundance levels in Knight-C4 relative to controls. However, Knight-C2 has the largest
271 dysregulated proteins (**S9 Table**, **S7B** and **S7C Fig**). We identified 24 (34%), 232 (51%), and 250
272 (52%) significant proteins unique to Knight-C1, 2, and 4, respectively (**Fig 4B-green bars** and **S9**
273 **Table**). We did not detect significant proteins dysregulated uniquely in Knight-C3 (**S7B Fig)**.
274 Knight-C4 was associated with many cluster-specific metabolites significantly different from
275 controls (90 hits; **S9 Table** and **S7C Fig**), whereas only one cluster-specific metabolite
276 significantly different from controls was associated with Knight-C2. Of the 90 hits, 80 (90%)
277 significant metabolites significantly differed between Knight-C4 and other AD cases.
278
279 **Fig. 4. The parietal cortex of participants with worse cognitive function exhibited a**
280 **remarkable molecular dysregulation.** (A) Bar plots showing the significant genes detected in
281 each cluster compared to the control. Yellow color represents the total number of genes, green
282 represents the cluster-specific genes, and purple depicts the number of genes missed with an un-
283 clustered approach (all AD cases vs control). (B) Same as panel "a" but for proteins. (C) Same
284 as panel "a" but for metabolites. (D) Race track plot showing the percentage of significant features
285 for each cluster compared to other clusters for the three omics**.** (E) Venn diagrams showing the
286 overlap between significant genes in Knight-C4 and synaptic genes ("synaptosome") from Fei *et*
287 *al.,* 2018[37]. (F) Same as "e" but overlapped with SynGO dataset[36]. The significance of overlap
288 was computed using Fisher's exact test. (G) Top 20 KEGG pathways associated with Knight-C4**.**
289 (H) Top 20 GO biological process pathways associated with Knight-C4.
290
291 We also performed DE analyses comparing all AD cases combined (un-clustered) to control
292 brains. We determined the number of significant cluster-specific features not captured by un-
293 clustered analyses (**S9 Table** and **Fig 4A-4C-purple bars**). We observed that Knight-C4

294 exhibited the highest number of features missed by un-clustered analyses with 4,295 genes
295 (92%), 475 proteins (98%), and 90 metabolites (100%) (**S9 Table** and **Fig 4A-4C-purple bars**),
296 indicating the importance of clustered analyses. Notably, a high percentage of features was also
297 observed with Knight-C2 in proteomics (98%). In contrast, un-clustered analyses identified a small
298 percentage of significant features not captured by clustered analyses. For instance, only 5% and
299 0.5% of un-clustered significant features for transcriptomics and proteomics, respectively, were
300 not captured by Knight-C4. Knight-C4 missed no un-clustered significant metabolites.
301
302 DE analyses comparing brains in each AD cluster to brains in other AD clusters (**S3, S7 and S8**
303 **Tables**) revealed that Knight-C4 showed the highest percentage of significant hits compared to
304 other AD brains (**Fig 4D** and **S10 Table**). This significant difference was observed in
305 transcriptomics, with 6,822 significant hits compared to 678, 22, and 1 for Knight-C1-3 (**Fig 4D-**
306 **outer track** and **S10 Table**). Similarly, this difference was observed in metabolites with 228
307 significant hits in Knight-C4 compared to 105, 0, and 0 for Knight-C1-3, respectively (**Fig 4D-inner**
308 **track** and **S10 Table**). No significant differences in the number of hits were observed between
309 the four clusters in proteomics, with all clusters showing relatively similar detections (**Fig 4D-**
310 **middle track** and **S10 Table**). We also verified that none of the omics dominate this solution,
311 each capturing molecular heterogeneity among AD brains that are not necessarily detected by
312 other *omics* modalities.
313
314 Noticeably, these DE analyses also showed that Knight-C4 is significantly enriched in
315 dysregulated synaptic genes (**Fig 4E** and **4F** and **S11 Table**) using the annotation from The
316 SynGO consortium [36] ($p=1.0x10^{-15}$ and $p=4.8x10^{-45}$ for Knight-C4 vs. Control and Knight-C4 vs.
317 Knight-C1-3 sets respectively) and curated list from Fei *et al.,* 2018 [37] ($p=2.3x10^{-07}$ and
318 $p=1.1x10^{-21}$ for Knight-C4 vs. Control and Knight-C4 vs. Knight-C1-3 sets respectively). MSBB-

319 C1 and ROSMAP-C1 profiles were also significantly enriched in synaptic genes (**S8A** and **S8B**
320 **Fig** and **S11 Table**).
321
322 In summary, these results show that molecular profiles of AD are heterogeneous, and the cross-
323 omics integration approach can better identify the substantial differences between clinically
324 relevant sub-categories of AD brains.
325

326 The Knight-C4 cluster is significantly enriched in AD-related
327 dysregulated pathways.

328 Knight-C4 contains multiple pathways related to different aspects of neurodegeneration (**Fig 4G**
329 and **4H**). For instance, KEGG pathways identified multiple intracellular trafficking and signaling
330 pathways (**Fig 4G**), including (e.g., *VPS29* and *VPS35*) endocytosis ($p=6.5x10^{-04}$), (e.g.,
331 *ATP6V1A* and *RAC1*) phagosome ($p=1.2x10^{-06}$), and (e.g., *GSK3B*, and *IGF1)* mTOR signaling
332 pathway ($p=3.3x10^{-02}$), all of which have been shown to play major roles in the pathological
333 processes of AD [38–40]. Genes (*PPT1*, *GBA*, *HEXB*, and *LGMN)* from lysosomal pathway, which
334 is critical for microglia, and neurons, were also identified ($p=4.7x10^{-02}$, **Fig 4G**). Growing evidence
335 has shown that the dysfunction of lysosomes plays a major role in the pathogenesis of multiple
336 neurodegenerative diseases (ND), including the progression of AD [41–45]. Knight-C4 was also
337 associated with genes (*BECN1* and *MAPK1)* from the autophagy pathway ($p=1.6x10^{-02}$, **Fig 4G**),
338 which is associated with several neurodegenerative disorders, including AD [46–49]. For instance,
339 autophagy dysfunction was shown to be linked to the accumulation of misfolded protein
340 aggregates [50]. The downregulation of Knight-C4 compared to other AD cases identified multiple
341 ND-related pathways, including AD ($p=2.5x10^{-14}$), Huntington disease ($p=3.3x10^{-14}$), and PD
342 ($p=2.1x10^{-12}$) (**Fig 4G**), Knight-C4 also exhibited downregulation of synaptic pathways including
343 dopaminergic synapse ($p=7.0x10^{-03}$) and glutamatergic synapse ($p=1.0x10^{-02}$) (**Fig 4G**), including

344 *GRIA1*, *GRIA2, PPP3CA*, *GRIN2A*, *PLCB1*, *CALY*, and *GLS* genes. Furthermore, GO biological
345 process analyses identified significant neuronal differentiation ($p=1.8x10^{-02}$) and development
346 ($p=5.7x10^{-03}$) associated with Knight-C4 (**Fig 4H**). Most of these pathways were also identified in
347 MSBB-C1 and ROSMAP-C1 (**S9A-S9C Fig**).
348
349 Compared to the control group, Knight-C4 showed additional pathways, including herpes simplex
350 virus 1 infection ($p=4.4x10^{-03}$), a risk factor for AD [51–54] that lead to neuronal damage and
351 induced gliosis [55,56], and synaptic vesicle cycle ($p=4.6x10^{-05}$) shown to play a key role in the
352 pathobiology of AD [57] (**S9D Fig**). Noticeably, herpes simplex virus 1 infection was significantly
353 enriched in Zinc Finger genes, a class of protein-coding genes known to regulate gene expression
354 at the transcriptional and translational levels by binding to bind to certain DNA or RNA molecules.
355 GO biological process analyses identified several significant synaptic signaling and transmission
356 pathways including two well-known AD biomarkers, *SNAP25* and *SYT1*. (**S9E Fig**). Many of these
357 pathways were also present in MSBB-C1 and ROSMAP-C1 (**S10A-S10D Fig**). Pathway analyses
358 using significant proteins identified AD-related pathways including PI3K−Akt signaling ($p=6.5x10^{-}$
359 ^{19}), rap1 signaling ($p=5.8x10^{-14}$), axon guidance ($p=2.1x10^{-12}$), MAPK signaling ($p=1.4x10^{-20}$), and
360 ras signaling ($p=1.8x10^{-15}$) pathways associated with the downregulation of Knight-C4 (**S10E Fig**).
361 Noticeably, *MAPT*, the gene that encodes the microtubule-associated protein tau (MAPT), was
362 among the genes in the MAPK signaling pathway. Top GO pathways, enriched in Knight-C4
363 downregulated proteins (and replicated in MSBB-C1), were related to cell/gene regulatory
364 pathways (**S10F Fig**). Through cross-omics analyses, we discovered numerous known and novel
365 pathways contributing to the cognitive decline in AD patients in multimodal profiles.
366

367 # The multimodal clusters identify potential AD biomarkers

368 Among the top genes/proteins dysregulated in Knight-C4, several AD biomarkers including *APP,*
369 *APOE, CLU, SNAP25, GFAP, SNCA, NOTCH3, TARDBP, GRN, MMP9,* and *C9ORF72* were
370 identified (**S3, S7**, **S8 Tables** and **Fig 5A** and **5D**). Of these, transcript levels of Synaptosome
371 Associated Protein 25 (*SNAP25*), a protein implicated in AD and Down syndrome (DS) [58], were
372 significantly downregulated in Knight-C4 (**Fig 5B**), as it has been previously reported in AD [58].
373 *SNAP25* protein levels were also shown to be reduced in the CSF in AD patients [59]. *SNAP25*
374 was significantly DE between Knight-C4 and control ($p=5.9x10^{-3}$) and between Knight-C4 and
375 other AD cases ($p=4.2x10^{-08}$, **Fig 5B**). However, *SNAP25* was not significantly DE ($p > 0.05$, **Fig**
376 **5B**) in non-clustered AD cases compared to the control. No significant differences were observed
377 in *SNAP25* protein levels in clustered or un-clustered analyses. The decreased expression of
378 *SNAP25* in Knight-C4 was replicated in MSBB-C1 (transcript and protein levels) and ROSMAP-
379 C1 (transcript levels) cohorts (**S11A-S11C Fig**).

380

381 **Fig. 5. Differential expression analyses identified neurodegenerative disease genes**
382 **associated with distinct clusters**. (A) Volcano plot showing the up-and downregulated genes
383 identified between Knight-C4 vs other clusters. Genes in red are examples of known
384 neurodegenerative disease genes. Genes in black are the top 10 significant genes based on the
385 adjusted p-value and genes with log2(fold change) > 1. (B) Boxplots showing the transcriptomic
386 profiles (in FPKM) of *SNAP25* across four clusters (right) and for AD cases combined (left) in the
387 Knight ADRC cohort. (C) Same as "b" but for *GFAP*. (D) Volcano plot showing the up-and
388 downregulated proteins identified between Knight-C2 vs other clusters. (E) Boxplots showing the
389 proteomic profiles of *CLU* across four clusters (right) and for all AD cases (left) in the Knight ADRC
390 cohort. (F) Kaplan-Meier plot showing the association of Clusterin (ApoJ) protein concentrations

391 in CSF with an increased risk of dementia progression using CDR change from 0 to CDR > 0
392 using the CSF dataset from the Knight ADRC participants.

393

394 Glial Fibrillary Acidic Protein (*GFAP*), an emerging fluid biomarker in brain disorders [60] usually
395 employed to identify astrocytic reactivity state, a well-documented process in AD and
396 neurodegeneration, was identified. The increased astrocyte reactivity and overexpression of
397 *GFAP* have been reported in multiple studies [58,60–68]. *GFAP* was overexpressed uniquely in
398 Knight-C4 compared to Knight-C1-3 (**Fig 5C**, $p=1.7x10^{-2}$). However, *GFAP* was not significantly
399 DE in Knight-C4 compared to controls, nor was it DE in all AD cases compared to the control (p
400 > 0.05, **Fig 5C**). We repeated these analyses using a simplified model that did not correct for
401 cellular population structure to examine whether collinearity between case-control status and
402 cellular population structure may introduce spurious results, but we could not obtain a significant
403 association. The same pattern was observed with *AQP4*, another marker gene of astrocyte
404 reactivity state; it was only significantly different between Knight-C4 and Knight-C1-3 ($p=1.0x10^{-}$
405 ^{02}). We did not identify significant differences in *GFAP* protein levels (p > 0.05, **S12 Fig**). However,
406 protein levels of *GFAP* were significantly upregulated in Knight-C2 (**S12 Fig**) compared to the
407 control and AD brains in Knight-C1,3,4 ($p=1.0x10^{-4}$ and $1.4x10^{-21,}$ respectively). The increased
408 transcriptomic *GFAP* levels were also identified in MSBB-C1 and ROSMAP-C1 (**S13A-S13C Fig**)
409 but with less significance level in ROSMAP-C1. Altogether these results support heterogeneity in
410 the presence of astrocytic reactivity state and its association with different AD subtypes.

411

412 Our cross-omics integration analyses also identified cluster-specific dysregulation of Clusterin
413 (*CLU*), a gene associated with AD [69,70]. Genome-wide association studies have shown that
414 *CLU* (*APOJ*) is the third most associated late-onset AD (LOAD) risk gene after *APOE* and *BIN1*
415 [70–73]. It was also shown to play a neuroprotective role in AD by altering Aβ aggregations and

416 promoting Aβ clearance [73–77]. *CLU* protein levels were significantly downregulated in Knight-
417 C2 compared to the control ($p=6.4x10^{-3}$) and AD cases in Knight-C1,3,4 ($p= 9.8x10^{-7}$; **Fig 5E**).
418 *CLU* was also significantly downregulated in Knight-C4 compared to the control ($p=3.4x10^{-02}$; **Fig**
419 **5E**). We did not observe significant differences in *CLU* transcriptomic and proteomic levels in the
420 MSBB and ROSMAP cohorts. To assess the clinical implications of *CLU* dysregulation, we
421 evaluated its levels in the CSF from 553 participants of the Knight ADRC by performing survival
422 analyses for progression using CDR change from 0 to 0.5. These analyses indicate that lower
423 CSF *CLU* is associated with an increased risk of dementia progression (HR=0.42; $p=1.0x10^{-07}$,
424 **Fig 5F**). Together, results show that cross-omics approach identified potential biomarkers for AD
425 that traditional approaches may miss.
426

427 # Metabolic differential abundance and pathway analyses in
428 # Knight-C4

429 We next looked at the significant metabolites that uniquely characterize brain samples in Knight-
430 C4. To determine whether significant metabolites are replicated in additional cohorts, we
431 accessed metabolomics from the same ROSMAP cohort [35] (**S1 Table**, Methods). We performed
432 differential abundance analyses comparing the levels from donors in the previously identified
433 ROSMAP-C1 cluster to controls and other AD cases (**S6 Table**). We identified 57 significant
434 metabolites (43 increased and 14 decreased) in Knight-C4 compared to controls replicated in the
435 ROSMAP-C1 (**Fig 6A** and **S12 Table**). Similarly, we identified 74 metabolites (70 increased and
436 4 decreased) in the Knight-C4 compared to other AD cases replicated in ROSMAP-C1 (**Fig 6B**
437 and **S12 Table**).
438

439 **Fig. 6. Differential abundance analyses identified significant metabolites and pathways**
440 **associated with Knight-C4**. (A) Venn diagrams show the significantly increased and decreased

441 metabolites shared between the Knight-C4 and ROSMAP-C1 compared to the control. The left
442 and right boxes show the names of shared metabolites. (B) Similar to panel "a" but for metabolites
443 shared between Knight-C4 and ROSMAP-C1 compared to the other cases. The significance of
444 overlap was computed using Fisher's exact test.

445

446 Pathway analyses using these common metabolites identified several pathways significantly
447 related to AD. Among the metabolites decreased in AD (**Fig 6A** and **6B**), gamma-aminobutyrate
448 (GABA) and choline were identified, which involved major neurotransmitter pathways. GABA,
449 synthesized in neurons, was significantly decreased in Knight-C4 ($p= 5.1x10^{-04}$) and ROSMAP-
450 C1 ($p=1.3x10^{-02}$) compared to the control (**S14A Fig**). The same observation was found when
451 these profiles were compared to other AD cases ($p=1.2x10^{-05}$ and $p= 1.6x10^{-02}$ for the Knight-C4
452 and ROSMAP-C1, respectively, **S14A Fig**). As an inhibitory neurotransmitter, GABA has been
453 shown to play a key role in synchronizing the activity of the human cerebral cortex [78,79].
454 Similarly, the metabolomics profiles of choline were decreased in these profiles in both Knight-C4
455 ($p= 4.1x10^{-06}$) and ROSMAP-C1 ($p=1.2x10^{-04}$) compared to controls (**S14B Fig**). These reading
456 levels of choline also show significantly different abundance in these clusters compared to other
457 AD cases ($p= 2.9x10^{-07}$ and $p=4.7x10^{-3}$ for Knight-C4 and ROSMAP-C1, respectively, **S14B Fig**).
458 Choline was also shown to have a critical role in neurotransmitter function because of its effect
459 on the function of acetylcholine and dopamine [80]. However, we did not find significant
460 differences in the reading levels of these active neurotransmitters. Nor did we observe significant
461 differences for additional neurotransmitters such as serotonin and glutamate. We then studied
462 tryptophan, an amino acid essential for metabolites that serve as neurotransmitters and signaling
463 molecules [81]. We found that its levels increased in the Knight-C4 compared to the control and
464 other AD cases ($p=8.3x10^{-3}$ and $3.0x10^{-06,}$ respectively) (**S15A Fig**). Similarly, we found that brains

465 in ROSMAP-C1 showed increased tryptophan levels compared to the control and other AD brains
466 ($p=1.4x10^{-4}$ and $2.8x10^{-07,}$ respectively) (**S15A Fig**).
467
468 Several increased carbohydrate/sugar metabolites were identified in Knight-C4 (**Fig 6A** and **6B**),
469 including glucose, mannitol, sorbitol, ribose, and UDP-glucose. For example, glucose, a
470 metabolite implicated in the pathogenesis of AD [82,83], was significantly increased in Knight-C4
471 and ROSMAP-C1 (**S15B Fig**). Components of glycolysis and the tricarboxylic acid (TCA) cycle
472 were also dysregulated. Specifically, levels of citrate were decreased in both Knight-C4 (p=
473 $9.7x10^{-05}$) and ROSMAP-C1 (p= $1.4x10^{-03}$) compared to other AD cases (**S16A Fig**). NAD+, an
474 important oxidizing cofactor in glucose metabolism and cellular respiration, was also decreased
475 in both Knight-C4 ($p=2.2x10^{-02}$) and ROSMAP-C1 ($p=2.2x10^{-02}$) compared to other AD cases
476 (**S16B Fig**). NAD+ is synthesized from tryptophan, and levels of NAD+ have been shown to
477 decrease with age, potentially contributing to the pathogenesis of AD [84]. These observations
478 align with previous findings that associate increased brain glucose and slowed glucose
479 metabolism with more severe AD [83,85]. In addition, several increased sphingolipid metabolites
480 were identified, including sphingomyelin, sphingosine, and sphingadienine (**Fig 6A** and **6B**).
481 Previous studies showed the implication of sphingolipid metabolism deregulation in AD [86]. For
482 instance, sphingosine, which had higher levels in AD brains [86], showed significantly increased
483 metabolism levels in Knight-C4 and ROSMAP-C1 (**S16C Fig**). Altogether, these results show that
484 cross-omics methods can uncover dysregulated metabolites in AD brains, indicating a link to other
485 conditions like diabetes [85].
486

487 # Integration with single-cell transcriptomic data reveals cell
488 type-specific changes associated with distinct molecular
489 profiles of AD

490 We studied the cell type-specific pathways altered in the AD profiles by leveraging single-nuclei
491 RNA-seq from a subset of the Knight ADRC participants. We used the top 5% of genes that
492 showed the most specific cell-type expression profile in neurons, astrocytes, oligodendrocytes,
493 oligodendrocyte precursor cells (OPC), and microglia. Then, we used this information to
494 determine if any AD molecular profiles showed a stronger signature of cell-type specific
495 dysregulation. These analyses showed that Knight-C4 was enriched in genes overexpressed in
496 astrocytes and endothelial cells (38% and 23% for all hits, respectively; **Fig 7A-left panel** and
497 **S13 Table**). Whereas the first may indicate reactive astrogliosis, the latter may suggest
498 endothelial dysfunction in AD development, providing more evidence of vascular involvement in
499 this disease. Pathways analyses of astrocytic genes identified astrocyte-related pathways (**Fig**
500 **7B**), including axon guidance ($p=2.8x10^{-05}$) and hedgehog signaling pathway ($p=1.6x10^{-06}$).
501 Previous work has shown the role of astrocytes in axon guidance during development and repair
502 [87] and hedgehog signaling pathways involved in neuroinflammation, neuronal cell
503 differentiation, and neuronal death [88]. Examples of genes involved in these pathways are *FYN*,
504 *SMO*, *KIF7*, and *GLI2*. Many of these pathways were also enriched in overexpressed genes in
505 endothelial cells (**S17A Fig**).
506
507 **Fig. 7. Integrating single-nuclei data with cross-omics profiles identified cell-type-specific**
508 **genes and pathways.** (A) Bar plots showing the percentage of up-and downregulated genes
509 (clusters vs. control) identified in each cell type. (B) Top 20 KEGG pathways enriched in astrocytic
510 upregulated genes. (C) Top 20 GO biological process pathways enriched in neuronal

511 downregulated genes. (D) Top KEGG pathways enriched in neuronal downregulated genes. (E)
512 Boxplots show the transcriptomic profiles (in FPKM) of examples of cell-type specific genes
513 identified in the Knight ADRC cohort. (F) The effect size of these examples was generated from
514 single-nuclei data from the Knight ADRC participants. Red bars represent the cell type in which
515 the gene is overexpressed.

516

517 Knight-C4 has also enriched in genes downregulated in microglial and neuronal genes (**Fig 7A-**
518 **right panel** and **S13 Table**). Neuronal genes showed an association with synaptic vesicle cycle
519 ($p=1.0x10^{-06}$) and retrograde endocannabinoid signaling ($p=3.6x10^{-03}$) [89–91], and gap-junction
520 ($p=2.2x10^{-02}$) formation known to play an essential role in intercellular metabolic communication
521 and transmission across electrical synapses [92] (**Fig 7C** and **7D**). Two AD biomarker genes,
522 *SNAP25* and *SYT1,* were included in these pathways. Microglial downregulated genes implicate
523 AD-related pathways, including phagosome ($p=1.0x10^{-06}$), herpes simplex virus 1 infection
524 ($p=2.0x10^{-04}$)—previously associated with AD risk [51–54], gap junction ($p=2.2x10^{-02}$), and AD
525 ($p=5.4x10^{-08}$) (**S17B Fig**). Two interesting examples of genes represented in these pathways are
526 *C3* and *CD74*, which are implicated in AD pathology. Further explorations showed that Knight-C4
527 is significantly enriched in downregulated genes related to homeostatic microglia, including C-X3-
528 C Motif Chemokine Receptor 1 (*CX3CR1;* $p= 5.10x^{-03}$, **Fig 7E** and **7F**) and purinergic receptor
529 P2Y13 (*P2RY13;* $p=1.2x10^{-2}$, **Fig 7E** and **7F**).

530

531 Additionally, we identified genes dysregulated in specific clusters and cell types, including the
532 TP53-induced glycolysis regulatory phosphatase (*TIGAR*) overexpressed in microglia and
533 upregulated in Knight-C1 (**Fig 7E** and **7F**). *TIGAR* is a p53-inducible protein that regulates energy
534 metabolism and oxidative stress. Among the upregulated endothelial-specific genes in Knight-C1,
535 we identified *MYOCD*, *CNN1*, *PLN*, *MYH11*, S*100A4*, *SLC13A4*, *DES*, implicating pathways

536 associated with cell cycle, junction, anion transport, Aβ clearance and actin. In particular, *MYH11*
537 overexpressed in Knight-C4 (**Fig 7E** and **7F**) has been reported to be associated with the risk of
538 dementia [93]. Moreover, oligodendrocyte-specific genes dysregulated in Knight-C2 include *STX4*
539 (Syntaxin-4), *ACTN4*, and *FBXO32* relating to pathways associated with vesicle docking, neural
540 tube, junction, apoptosis, and actin. Of particular interest, *STX4* upregulated in Knight-C2,4 (**Fig**
541 **7E** and **7F**) is critical for vesicle docking and the Wnt frizzled receptor (*FZD9*) and coreceptor
542 (*LRP5*), regulating the synaptic vesicle cycle [30]. *STX4* has also been shown as one of the
543 disease risk genes in AD [94].
544 These results suggest that combining cross-omics data with single-cell resolution allows
545 uncovering of molecular profiles associated with AD features to unprecedented granularity and
546 complexity.
547

548 ## *SNCA* is a top hit in the multimodal cluster Knight-C4

549 Transcriptomic and proteomic assays showed *SNCA* among the top gene/proteins that best
550 discriminate Knight-C4 versus Knight-C1-3 molecular profiles. *SNCA* transcriptomic levels were
551 significantly downregulated in Knight-C4 compared to Knight-C1-3 ($p= 2.0x10^{-05}$) and control
552 brains ($p=2.2x10^{-02}$; **Fig 8A**). Interestingly, we did not observe a significant association ($p > 0.05$)
553 in non-clustered AD compared to the control (**Fig 8A-left panel**). *SNCA* encodes alpha-synuclein
554 (αSyn) protein, whose protein levels were marginally downregulated in brains from Knight-C4 (**Fig**
555 **8B**) compared to controls ($p=5.16x10^{-02}$) and significantly lower in Knight-C2 compared to controls
556 ($p=7.8x10^{-5}$) and Knight-C1,3,4 ($p=6.2x10^{-09}$).
557

558 **Fig. 8. Cross-omics integration identified alpha-synuclein levels downregulated in AD**
559 **participants with worse cognitive function common across multiple brain regions.** (A)
560 Transcriptomic profiles (in FPKM) of *SNCA* across four clusters (right) and for all AD cases (left)

561 in the Knight ADRC cohort. (B) Proteomic profiles of *SNCA* across four clusters (right) and all AD
562 cases (left) in the Knight ADRC cohort. (C) Transcriptomic profiles of *SNCA* across four clusters
563 (right) and all AD cases (left) in the MSBB BM36 cohort. (D) Proteomic (TMT) profiles of *SNCA*
564 across four clusters (right) and all AD cases (left) in the MSBB BM36 cohort. (E) Transcriptomic
565 profiles of *SNCA* across four clusters (right) and all AD cases (left) in ROSMAP DLPFC cohort.

566

567 *SNCA*/*αSyn* is also downregulated in MSBB-C1 (**Fig 8C** and **8D**) compared to the controls
568 ($p=1.8x10^{-02}$ and $p=1.0x10^{-05}$ for transcriptomics and proteomics, respectively). Similarly, *SNCA*
569 levels in MSBB-C1 were significantly lower than other AD ($p=1.0x10^{-02}$ and $p=3.4x10^{-03}$ for
570 transcriptomics and proteomics, respectively, **Fig 8C** and **8D**). However, no significant differences
571 in transcript or protein levels of *SNCA* were observed in MSBB-C2 compared to controls ($p >$
572 0.05, **Fig 8C** and **8D**). Downregulation of *SNCA* transcript in the MSBB-C1 donors was also seen
573 in the frontal pole (BM10), superior temporal gyrus (BM22), and inferior frontal gyrus (BM44) brain
574 regions (**S18A-S18C Fig**). Similarly, *SNCA* transcriptomic levels were significantly downregulated
575 in ROSMAP-C1 (**Fig 8E**) compared to the control ($p=3.5x10^{-02}$) and other AD cases ($p=4.5x10^{-}$
576 ^{03}).

577

578 *αSyn* is the major constituent of Lewy bodies, the histopathologic hallmark of PD and DLB. LB
579 pathology is also present in most advanced AD cases, including autosomal dominant AD and DS
580 [95–101]. In addition, previous work has demonstrated the relationship between *αSyn* and AD
581 pathologies, including Aβ plaques and tau aggregations using different mouse models [102,103]
582 and implicated *αSyn* pathology in neuronal/synaptic dysfunction and death [104,105]. Co-
583 pathology with *αSyn* is present in 50-60% of AD brains [95,96,106]. However, it is not clear if it
584 plays any role in AD pathophysiology. In animal models, it has been shown that adding *αSyn* to
585 an amyloidosis model increased neuron loss that correlated with the progressive decline of

586 cognitive and motor performance, and a "feed-forward" mechanism whereby Aβ plaques enhance
587 endogenous α-syn seeding and spreading over time has been propose[87]. Mutations in *SNCA*,
588 overexpression of wildtype *αSyn*, and injection of pre-formed *aSyn* fibrils (PFFs) are sufficient to
589 cause neuropathology and parkinsonian syndromes in both humans and animal models [107–
590 113].
591
592 We also analyzed NanoString gene expression data (see Methods) generated from a panel of
593 807 gene transcripts measured in brain tissue of A53T *αSyn* transgenic mice crossed onto *APOE*
594 backgrounds [114]. The *SNCA* A53T mutation cause familial Parkinson's disease. Of the 807
595 genes included in the panel, 89 (11%) were classified as synaptic genes (SynGO consortium
596 [36]). We performed DE analysis to compare mice with high pSyn pathology (defined as > 5%
597 coverage by IHC) to mice with low pathology (< 5% coverage by IHC, see Methods). We identified
598 255 DEGs, including 13 synaptic genes. We then interrogated which synaptic genes in DE in
599 Knight-C4 were also dysregulated in the A53T αSyn-APOE transgenic mice (**S19 Fig**). We
600 identified that all 13 synaptic genes identified as dysregulated in the mouse model were also
601 dysregulated in Knight-C4 compared with other AD cases (Fisher $p=3.6x10^{-09}$) (**S19A Fig**). These
602 results suggest that the profile of Knight-C4 is reproducible in mouse models of aggregation and
603 neurodegeneration. Compared to the control, only five synaptic genes were shared between the
604 three datasets (**S19B Fig**). These findings were limited because the nanoString panel contains
605 only 12% of the significant genes in Knight-C4. These results show that integrating data from
606 multimodal clusters in humans with molecular data from an established mouse model of
607 neurodegeneration due to *αSyn* provides new clues about the role of known proteins in AD
608 pathophysiology.
609

610 # Cross-omics integration identified CSF synaptic biomarkers
611 # for the molecular staging of AD

612 Knight-C4 exhibits significant dysregulation of synaptic genes and association with AD pathology
613 but critically, is also associated with poor cognitive performance. Thus, we hypothesized that
614 synaptic changes associated with AD molecular multimodal profiles could reflect multiple stages
615 of AD. To evaluate this, we selected those genes dysregulated in Knight-C4 and ROSMAP-C1,
616 and examined the effect size differences between the early-AD vs. control compared to that of
617 the late-AD vs. control (**Fig 9A**). Interestingly, we observed a concordant change of synaptic gene
618 dysregulation in early- and late-AD compared to controls; however, the effect was more
619 pronounced at later stages. (**Fig 9A and 9B**, R^2 = 0.96, $p < 2.2x10^{-16}$, slope =0.75). This was
620 consistent among 170 synaptic genes, including *IGF1*, *NRXN3*, and *YWHAZ* (**Fig 9C**). Co-
621 expression analyses (see Methods) identified a core module (**S20 Fig, turquoise color**) highly
622 enriched in interacting proteins, including *SNAP25*, *APP*, and *SNCA* (**Fig 9D**). Subsequent
623 pathway analyses identified several AD-related pathways, including 160 protein-protein
624 interaction hub proteins (**Fig 9E** and **S14 Table**). Additional gene interaction analyses showed
625 that those proteins have a strong interaction, including physical and genetic interactions (**Fig 9F-**
626 **9H**).
627

628 **Fig. 9. Synaptic dysregulation at multiple stages of AD identified CSF synaptic biomarkers**
629 **for the molecular staging of AD.** (A) Scatterplot showing the correlation of effect size between
630 early-AD (EAD) and late-AD (AD) for common dysregulated synaptic genes in Knight-C4 and
631 RSOSMAP. (B) Heatmap showing the effect size similarity of the most significant synaptic genes
632 detected in EAD and AD. (C) Boxplots showing a consistent expression (transcriptomic) pattern
633 between control, early-AD, and late-AD for genes *IGF1*, *NRXN3*, and *YWHAZ*. (D) The gene
634 interaction network for the turquoise module shows high interactions between dysregulated

synaptic genes such as *SNAP25*, *SNCA*, and *APP*. (E) Barplot showing the top 25 hub proteins extracted from protein-protein interaction (PPI) hub proteins pathways. (F) Gene interaction network showing the relatedness of the 25 dysregulated hub proteins. The plot demonstrates multiple strong interactions, including physical and genetic interactions among the query genes (25 hub proteins) and other genes. (G) *YWHAB* sub-network shows the genes highly connected with YWHAB, some of which are examples of either synaptic or other AD-related genes. (H) Same as "g" but for *SNCA* sub-network. (**i**) Kaplan-Meier plots showing the association between *IGF1*, *NRXN3*, and *YWHAZ* genes and an increased risk of dementia progression using CSF proteomics data.

Next, we sought to determine whether some of these genes are potential CSF biomarkers for staging AD pathology and monitoring disease progression. To explore this hypothesis, we performed survival analyses, including individuals with CDR change from 0 to 0.5, and using CSF proteomic (Somascan) data from participants from the Knight ADRC. Our results showed that 11 of 13 (84%) synaptic genes were associated with an increased risk of dementia progression, including *IGF1*, *NRXN3*, and *YWHAZ* (**Fig 9I**). Several studies have shown the involvement of *IGF1* in regulating signaling pathways altered in AD and neurodegeneration [115–119] and its potential as an early CSF/plasma biomarker of disease onset [120]. However, the association between *IGF1* and the increased risk of dementia progression in different stages of AD has not been studied. Likewise, the low expression of *NRXN3* is associated with an increased AD risk [121,122]. The *YWHAZ/ 14−3−3* protein, present in CSF patients, has also been shown to be involved in the rapid progression of dementia [123] and identified as a promising therapeutic target for AD intervention [124]. Results show that the multimodal clusters identified by the cross-omics approach in post-mortem tissue could identify molecular CSF biomarkers for monitoring AD progression in live patients.

660

661 # Discussion

662 In the current study, we leveraged machine learning approaches to integrate high-throughput
663 multi-omics data from controls and AD cases in multiple cohorts and brain regions. Our results
664 highlight the utility of integrating multi-omics data and indicate that cross-omics signatures can
665 better capture differences and discriminate molecular variations in complex and heterogenous
666 diseases such as AD. We found a distinct molecular profile associated with worse cognitive
667 function, earlier age at onset/death, and dysregulated neuronal/synaptic pathways in the parietal
668 cortex. Our results suggest that this multi-omics molecular profiling captures novel molecular
669 events not fully captured by β-amyloid or tau staging as previously reported [125,126]. Similar to
670 previous studies[8], we did not find a significant association of *APOE* ε4 or PRS with any molecular
671 profile, suggesting that these molecular profiles are not driven by genetic factors associated with
672 AD risk. Through a rigorous and stringent replication using multiple cohorts (MSBB, ROSMAP),
673 we also determined that this molecular profile replicates in additional cortical regions (parietal
674 cortex, DLPFC, PHG) where it associates with tau pathology. Further, we demonstrated that this
675 molecular profile is associated with poor clinical and cognitive function in these studies, offering
676 new insights into the genes/proteins/metabolites associated with cognition decline and pathways
677 associated with worse prognosis.

678

679 Notably, we observed that these donors with poor cognitive function, earlier age of onset, and
680 younger age at death showed significantly increased cortical cellular proportions of astrocytes
681 and reduced proportions of neurons. This subset of AD-afflicted participants is significantly
682 enriched in dysregulated synapse-related genes and pathways, which may reflect greater
683 synaptic losses, including dopaminergic synapse and glutamatergic pathways. Although the
684 involvement of dopamine in AD is still debatable, several studies have shown an association

685 between AD pathologies (e.g., Aβ deposition) and dopaminergic dysfunction [127–135]. Similarly,
686 increasing evidence has demonstrated an association between glutamate alterations and AD
687 pathology [136–139]. In addition, transcriptomic and proteomic DE analyses identified several
688 genes associated with AD, including *SNAP25*, *GFAP* and *CLU,* that are more noticeably
689 dysregulated in the distinct molecular profiles. We showed that lower levels of CSF *CLU* are
690 associated with an increased risk of dementia progression in an extended cohort from the Knight
691 ADRC. Moreover, metabolic differential abundance analyses identified significant metabolites
692 associated with biologically relevant AD pathways, most of which are related to major
693 neurotransmitter pathways. These brains also exhibited a metabolic profile associated with
694 disease duration, supporting its implication in the later stages of AD. These novel molecular
695 findings can be leveraged to identify new biomarkers and potential therapeutic targets – not
696 necessarily related to amyloid and tau – that might be effective even in the later stages of the
697 disease.
698
699 By integrating single-nuclei RNA-seq, we extracted additional insights supporting cell-type distinct
700 patterns implicated in molecular AD profiles. For instance, we determined that *CX3CR1*, a
701 homeostatic microglia gene, is significantly downregulated in Knight-C4. The lower expression of
702 *CX3CR1* may indicate fewer homeostatic microglia, leading to neuronal damage and loss. The
703 association between homeostatic microglia function and the degree of neuronal cell loss is linked
704 to AD [140]. More recently, the loss of *CX3CR1* was associated with microglial dysfunction
705 illustrated by dampened TGFβ-signaling, increased oxidative stress responses, and dysregulated
706 pro-inflammatory activation[136136]. However, we could not replicate a significant overexpression of
707 microglia-activated genes in additional cortical regions.
708

709 We identified *SNCA/*αSyn as one of the top molecules for discriminating AD profiles. We
710 demonstrated that the transcriptomic/proteomic levels of *SNCA* were under-expressed in a subset
711 of AD cases with more advanced tau pathology and worse cognition and replicated this finding in
712 several cohorts and brain regions. Despite the substantial efforts in studying the association
713 between this protein and AD pathology, the mechanisms by which α*Syn* contributes to the
714 pathogenesis and progression of AD remain elusive. Several regulatory mechanisms that
715 contribute to α*Syn* levels have been proposed. For instance, previous studies based on primary
716 neuronal cultures [142], mouse models [143], and live cell imaging [144] have shown that
717 neuronal or synaptic activities regulate the aggregation of α*Syn*. The downregulation of
718 *SNCA/*αSyn in a subset of AD cases with unfavorable outcomes suggests that *SNCA* might be a
719 marker of neuron/synapse losses that can be used for identifying AD stages.
720
721 We performed AD staging analyses to study the progression of AD and its effect on the synaptic
722 activities using dysregulated synaptic genes in the Knight-C4 and ROSMAP-C1 profiles. Our data
723 suggest that synaptic dysfunction occurs at the early stages of the disease, but it is more
724 pronounced as AD progresses. Survival analyses from CSF proteomics from the Knight ADRC
725 identified several synaptic genes associated with an increased risk of dementia, such as *IGF1*,
726 *NRXN3*, and *YWHAZ*. Those proteins can be used as potential CSF biomarkers to monitor
727 disease progression.
728
729 Although this study provides new insights for understanding the molecular heterogeneity in AD
730 brains, our study has several limitations, highlighting critical future research directions. A limited
731 number of brain metabolomic studies were available, and these were largely underpowered. In
732 addition, the multiple omics data from different brain regions and different cohorts were generated
733 at different times on different samples using different platforms. Thus, future studies that

consistently apply the same platforms on multiple brain regions and cohorts can provide novel and deeper findings. We also had to integrate omics data using different platforms due to the need for more harmonization in the methods used by different groups. Independent replication is needed to validate and better understand the role of key drivers of these profiles using disease models (e.g., iPSC or mouse models), including the role of *SNCA* in AD and its association with worsening cognitive function in AD. Longitudinal data from CSF/plasma from large cohorts could provide novel insights into the role of dysregulated genes/proteins/metabolites in AD progression. Additional multi-omics studies exploring multiple brain regions, particularly resilient areas, would be exceptionally useful to determine the temporal progression of protein changes throughout AD progression. We only used one tool (iClusterBayes) to integrate large multi-omics data; however, developing novel and more statistically robust tools is needed. We plan to continue integrating additional cohorts and omics data (e.g. epigenomics) as they become available. Another promising area for future exploration is integrating cross-omics data with drug repositioning datasets that may lead to identifying novel therapeutic potentials for AD, bringing AD closer to precision medicine.

In summary, we have shown the capability of cross-omics approaches and its superiority over single-omic analyses as a powerful resource that provides molecular insight into the AD pathogenesis and identifies possible biomarkers for early synaptic dysfunction, cognition, and AD staging that may eventually enable precision medicine.

Materials and Methods

Study Cohorts

Discovery (Knight ADRC)

Frozen postmortem parietal lobe tissue samples from Knight Alzheimer Disease Research Center (Knight ADRC) participants, collected with informed consent for research use, were provided by the Knight ADRC Neuropathology Core. The institutional review board of Washington University in St. Louis approved analysis of these samples. Only sporadic AD participants and controls with transcriptomics, proteomics, and metabolomics data available were included in this study, resulting in 255 sporadic AD and 23 controls.

- RNA-seq: The data generation for the RNA-seq dataset has been previously described [10–12]. Briefly, total RNA was extracted from frozen parietal cortex tissue using a Tissue Lyser LT and purified using RNeasy Mini Kits (Qiagen). The Nanodrop 8000 (Thermo Scientific) and TapeStation 4200 (Agilent Technologies) were used to perform quality control of the RNA's concentration and purity, and degradation. The RNA integrity number (RIN) was calculated using an RNA 6000 Pico assay on a Bioanalyzer 2100 and TapeStation 4200 (Agilent Technologies). The software determines the RINe on the Bioanalyzer and TapeStation taking into account the entire electrophoretic trace of the RNA, including the presence or absence of degradation products. The DV200 value is defined as the percentage of nucleotides greater than 200nt. All cDNA libraries were prepared using a TruSeq Stranded Total RNA Sample Prep with Ribo-Zero Gold kit (Illumina) and sequenced on an Illumina HiSeq 4000 using 2 x 151 paired-end reads at the McDonnell Genome Institute at Washington University in St. Louis. The average number of reads per sample was 43,195,080.

780 - Proteomics: Proteomic data from the Knight ADRC were generated on the SomaLogic
781 platform (SomaLogic Operating Co., Inc., Boulder, CO) from frozen parietal cortical tissue
782 samples (~500mg) and CSF samples. The SomaLogic platform is a multiplexed, aptamer-
783 based protein quantification platform [145]. The platform measured 1,305 aptamers in
784 total. Full details of the proteomic data generation have been described previously [13].
785 - Metabolomics: Metabolomic data from the Knight ADRC were generated on the Metabolon
786 Precision Metabolomics™ UPLC-MS/MS platform (Metabolon, Inc., Morrisville, USA) from
787 50mg frozen parietal cortical tissue samples. The platform measured 880 metabolites from
788 nine "super pathways": amino acids, carbohydrates, cofactors and vitamins, energy, lipids,
789 nucleotides, peptides, xenobiotics, and partially characterized molecules. Metabolon
790 provided structural identity annotations for 815 metabolites; the remaining 65 were
791 excluded from the final dataset. Full details of the metabolomics data generation have
792 been described previously [14].
793 - Single-nuclei (snRNA-seq): The data generation of snRNA-seq has been previously
794 described in [146,147]. Briefly, 74 frozen parietal tissues were processed according to the
795 'Nuclei extraction and library preparation' protocol described in [146]. The tissue was
796 homogenized in this protocol, and the nuclei were isolated using a density gradient. The
797 nuclei were then sequenced using the 10X Chromium single-cell Reagent Kit v3, with
798 10,000 cells per sample and 50,000 reads per cell for each of the 74 samples.

799 **Replication (MSBB, BM36)**

800 - RNA-seq: The RNA-seq raw data from The Mount Sinai Brain Bank (MSBB) study was
801 downloaded from the Synapse portal (syn3157743). It is publicly available as part of the
802 Accelerating Medicines Partnership for Alzheimer's Disease (AMP-AD). In brief, these
803 data were generated and sequenced from RNA extracted from four postmortem brain
804 regions, including frontal pole (BM10), superior temporal gyrus (BM22), parahippocampal
805 gyrus (BM36), and inferior frontal gyrus tissue (BM44) from 313 subjects. RNA-seq

806 libraries were prepared using the TruSeq RNA Sample Preparation Kit v2 (Illumina, San
807 Diego, CA). The rRNAs were depleted using the Ribo-Zero rRNA Removal Kit
808 (human/mouse/rat) (Illumina). Single-end non-standard reads of 101bp were generated
809 by Illumina HiSeq 2500 (Illumina) [34]. The average number of reads per sample was
810 32,201,200. RNA-seq data from the region parahippocampal gyrus (BM36) were selected
811 for the cross-omics integration approach because of the availability of proteomics data.
812 Data from the remaining regions were used to explore the identified profiles' expression
813 profiles to determine whether these are consistent across cortical brain regions.

814 - Proteomics (TMT): The proteomics TMT data were obtained from the AMP-AD Synapse
815 portal (syn25006650) processed by Johnson *et al.*, 2022 [148]. The raw proteomic TMT
816 data is publicly available at AMP-AD and can be downloaded from the Synapse portal
817 (syn21347564). In short, before TMT labeling, the 198 samples were randomized by co-
818 variates (protein quality, sample concentration, diagnosis, age, and sex) into 20 batches
819 (10 cases per batch). The digested peptides were resuspended in 50 mM HEPES (pH 8.5)
820 and labeled with the TMT 11-plex kit (Thermo Fisher) according to the manufacturer's
821 protocol. In each batch, GIS samples were labeled using TMT channel 126. All 11
822 channels were mixed equally and desalted with a 100 mg C18 Sep-Pak column (Waters)
823 for the subsequent fractionation [148,149].

824 **Replication (ROSMAP, DLPFC)**

825 - RNA-seq: The RNA-seq raw data from The Religious Orders Study and Memory and
826 Aging Project (ROSMAP) Study were downloaded from the Synapse portal
827 (syn17008934) available from AMP-AD. Briefly, RNA-seq samples were extracted using
828 Qiagen's miRNeasy mini kit (cat. no. 217004) and the RNase-free DNase Set (cat. no.
829 79254) and quantified by Nanodrop. Agilent Bioanalyzer evaluated quality. Library
830 preparation was performed by poly-A selection followed by first strand-specific cDNA
831 synthesis, then dUTP for second strand-specific cDNA synthesis, and fragmentation and

832 Illumina adapter ligation for library construction. Illumina HiSeq generated paired-end read
833 sequences with a length of 101bp. Fifty-seven samples were submitted later to the
834 platform and run on an updated protocol (batch 2 protocol). For consistency, we removed
835 those samples from this cohort. The average number of reads per sample was
836 47,434,600. Only No Cognitive Impairment (NCI) and AD with NO other cause of cognitive
837 impairment cases were used in this study. All cases with "unknown" BraakTau scores were
838 discarded. Plaque density and Neurofibrillary tangle burden (PHF tau tangles) are a
839 summary of AD pathology derived from three AD pathologies: neuritic plaques, diffuse
840 plaques, and neurofibrillary tangles based on five regions: mid-frontal cortex, mid-temporal
841 cortex, inferior parietal cortex, entorhinal cortex, and hippocampus. Tangles are identified
842 by molecularly specific immunohistochemistry based on the mean of eight regions:
843 tangles_hip (hippocampus), tangles_ec (entorhinal cortex), tangles_mf (mid frontal
844 cortex), tangles_it (inferior temporal), tangles_ag (angular gyrus), tangles_calc (calcarine
845 cortex), tangles_cg (anterior cingulate cortex), angles_sf (superior frontal cortex).
846 - Metabolomics: Data from the Religious Orders Study and Memory and Aging Project
847 (ROSMAP) were generated by the Duke Metabolomics and Proteomics Shared Resource,
848 a member of the ADMC, using protocols published previously for blood samples [150–
849 152]. A custom protocol developed for the brain samples can be found on Synapse at
850 syn10235609. The dorsolateral prefrontal cortex (DLPFC) metabolomic data from the
851 ROSMAP studies quantified on the Metabolon Precision Metabolomics platform and
852 preprocessed by the ADMC as described in [153] were downloaded from Synapse in July
853 2021 (syn25878459). The platform measured a total of 1,055 metabolites in the ROSMAP
854 dataset.
855

856 # Quantification and statistical analysis

857 RNA-seq QC, alignment, and gene expression quantification: All RNA-seq datasets from the
858 discovery and replication were processed and aligned using our in-house RNA-seq pipeline
859 (https://github.com/HarariLab/RNA-seq-Pipeline). Genome reference and gene models were
860 selected similar to the TOPmed pipeline (https://github.com/broadinstitute/gtex-
861 pipeline/blob/master/TOPMed_RNAseq_pipeline.md). Reference genome GRCh38 and
862 GENCODE 33 annotation, including the addition of ERCC spike-in annotations, were used. We
863 excluded ALT, HLA, and Decoy contigs from the reference genome due to insufficient RNA-seq
864 tools that properly handle these regions. Before alignment, the quality of raw read sequences for
865 all libraries was assessed using FastQC (v0.11.9) [154]. All raw read sequences were aligned to
866 the human reference genome (GRCh38) using STAR (v.2.7.1a) [155]. The alignment quality was
867 evaluated using sequencing metrics such as reads distribution, ribosomal content, or alignment
868 quality provided by STAR using Picard tools (v.2.8.2) [156]. All samples that failed to pass the QC
869 or were outliers (Knight ADRC=9, MSBB=18, ROSMAP=6) were removed from the downstream
870 analyses. Raw read counts for transcripts and genes were generated using STAR and computed
871 transcript/gene expression levels as normalized in FPKM (Fragments Per Kilobase of transcript
872 per Million mapped reads) format.

873 Metabolomics QC and Quantification**:** Metabolomics data from Knight ADRC and ROSMAP were
874 processed similarly. Full details of the QC process have been previously described [157]. Briefly,
875 188 metabolites with readings missing in over 20% of donors were excluded, readings were
876 log10-transformed, means per metabolite were adjusted to zero, and outlier readings (outside
877 1.5xIQR) were excluded. Missing metabolite values in < 20% of donors were replaced by their
878 mean value. No samples were removed due to the missingness rate. Principal component
879 analysis was performed with the FactoMineR R package [158] to identify outlier samples; four
880 outlier samples were excluded. The final dataset consisted of 627 metabolites. A similar

881 procedure was used to clean the ROSMAP metabolomic dataset. Metabolites without assigned
882 structural identities and metabolites with greater than 20% missing readings were excluded,
883 readings were log10-transformed, the mean of each metabolite's distribution was adjusted to
884 zero, and outlier readings were removed. Two samples missing greater than 20% of readings
885 were excluded. The final dataset consisted of 595 metabolites.

886 Proteomics QC and Quantification**:** Proteomic data for brain and CSF samples from the Knight
887 ADRC were first filtered based on a limit of detection (LOD). Samples with greater than 15%
888 outlier analytes (analyte readings below the LOD) were excluded. Analytes were then filtered
889 based on scale factor difference, meaning the absolute value of the maximum difference between
890 the calibration scale factor per aptamer and the median for each plate run. Analytes were
891 excluded if their scale factor difference exceeded 0.5. Based on log10-transformation of protein
892 readings, samples exceeding 1.5-fold of the interquartile range (IQR) per analyte were considered
893 outliers. Analytes with greater than 15% outlier samples were excluded. Similarly, samples that
894 were outliers for greater than 15% of analytes were excluded. Full details of the proteomic data
895 QC have been described previously [13]. Missing analyte readings and removed outlier readings
896 were imputed using the impute.knn function from the impute R package (v1.56.0) [159]. A plate-
897 wise batch correction was performed using the ComBat function from the sva R package (v3.30.1)
898 [160]. The final brain dataset comprised 1,092 analytes, and the final CSF dataset consisted of
899 713 analytes. Before using these data, analytes with missing expressions in > 20% of donors
900 were excluded, and missing expressions of those in < 20% were replaced with their mean value.
901 The processed data for TMT proteomics from MSBB BM36 was obtained from AMP-AD
902 (syn25006647) and processed by Johnson *et al.*,2022 [148]. In brief, 760 raw files generated from
903 19 TMT 11-plexes were analyzed using the Proteome Discoverer suite (version 2.3, Thermo
904 Fisher Scientific). The UniProtKB human proteome database containing Swiss-Prot and TrEMBL
905 human reference protein sequences was used to search mass spectra. Peptide spectral matches
906 and peptides were filtered to a false discovery rate (FDR) of less than 1% using Percolator

907 software. After spectral assignment, peptides were assembled into proteins and filtered based on
908 the combined probabilities of their constituent peptides to a final FDR of 1% [148]. Unique and
909 razor (that is, parsimonious) peptides were considered for quantification. TMT reporter abundance
910 was transformed into a ratio followed by $\log_2$ transformation. Quality control analysis was
911 performed, and outliers were removed. Similar to the Knight ADRC, analytes with missing
912 expressions in > 20% of donors were excluded, and missing expressions of those in < 20% were
913 replaced with their mean value.
914 Single-nuclei (snRNA-seq) QC, alignment, and gene expression quantification: Full details of the
915 data process of snRNA-seq data have been previously described in [146,147]. In short,
916 CellRanger (v2.1.1 10XGenomics) [161] software was employed to align the sequences and
917 quantify gene expression. Read sequences were aligned to a custom pre-mRNA reference
918 constructed from genome assembly GRCh38 and generated as described by 10X Genomics
919 technical manual. Filtering and QC analyses were performed individually using each subject's
920 Seurat package (v2.20 and 2.30). All nuclei with high mitochondria gene expression were
921 removed, and only genes expressed in > 3 cells were kept for downstream analysis. Cells with <
922 1800 or > 8000 genes expressed in them were excluded. The data were normalized using the
923 *LogNormalize* function that normalizes the gene expression measurements for each cell by the
924 total expression, scales by a factor equal to the median counts of all genes, and log-transforms
925 the expression.
926

927 # Clustering of AD participants

928 To cluster AD participants, iClusterBayes (Integrative clustering of multiple genomic data types)
929 [33] available from iClusterPlus R package (v1.22.0) [162] was employed. For each integration
930 process, shared samples across omics datasets were extracted, and the top 2000 most variant
931 features were selected. iClusterBayes was performed by fitting multiple Bayesian models using

932 different clusters ranging from 1 to 10. We ran 22,000 Markov Chain Monte Carlo (MCMC)
933 iterations for each, of which the first 12,000 were discarded as burn-in. The prior probability for
934 the indicator variable gamma of each data set was set to 0.1, whereas the posterior probability
935 cutoff was set at 0.5. The standard deviation of the random walk proposal for the latent variable
936 was set to 0.5. The remaining parameters were left with the default values. We used the Bayesian
937 information criteria (BIC) and deviance ratio to evaluate clustering solutions and maximizing the
938 compactness of the clusters (intra-cluster similarity).
939

940 # Cell proportion analysis

941 CellMix R package (v1.6.2) [163] employs multiple digital deconvolution methods and was used
942 to estimate cellular population structure from gene expression data (TPM) generated by the
943 Salmon quantification tool. The machine learning model and the gene panel (genes marker
944 expressed highly in specific cell types) were previously described in [12]. Four cell types
945 (astrocytes, microglia, neurons, and oligodendrocytes) were tested using two algorithms,
946 meanProfile [163] and ssNMF (semi-supervised nonnegative matrix factorization) [164]. The
947 significant difference in cell proportion between clusters was computed using a generalized linear
948 model (GLM) test and corrected by sex and age of death (AOD).
949

950 # Differential expression analysis of AD clusters

951 RNA-seq: differential expression (DE) analyses were performed to compare participants in each
952 cluster to the control participants and to other AD clusters using the DESeq2 R package (v.1.22.2)
953 [165]. Our models were adjusted for sex, age at death, and the percentage of astrocytes and
954 neurons. Lowly expressed genes were removed to enhance our confidence in the differentially
955 expressed genes we discovered. Only genes with expression > 0.5 CPM (count per million) in at
956 least 25% of samples in either group being compared were retained for downstream analyses.

957 The false discovery rate (FDR) was estimated using Benjamini-Hochberg (BH) [166]. All genes
958 with FDR < 0.05 were considered differentially expressed genes.
959 Proteomics: differential expression analyses between clusters were carried out with linear models
960 (R function glm) with age at death and sex as covariates. P-values were adjusted using Benjamini-
961 Hochberg (BH) [166].
962 Metabolomics: differential abundance analyses were also carried out with linear models, with age
963 at death, sex, and post-mortem interval (PMI) as covariates. P-values were adjusted using
964 Benjamini-Hochberg (BH) [166].
965 Single-nuclei: differential expression analyses were performed to compare the overexpression of
966 each gene in each cell type relative to other cell types using glmmTMB R Package (v1.1.3). This
967 approach uses a template builder (glmmTMB) to fit generalized linear mixed models. Our model
968 was corrected by sex covariate. To ensure cell-type specific overexpression, for each cell type,
969 only top genes whose effect size is greater than 90% percentile cutoff were retained for
970 downstream integration analyses.
971

972 # Survival analyses

973 Kaplan-Meier survival analyses and Cox proportional hazards models were used to determine the
974 association of different clusters with survival outcomes and also to determine the association with
975 the risk of dementia progression using survival (v3.2-3) [167] and surviplot (v1.1.1) [168] R
976 packages.
977

978 # Significance of overlap

979 The significance of overlap for all analyses was computed using Fisher's exact test available
980 from the R stats package (v4.3.0) [169].
981

982 ## Pathway analysis

983 All pathway analyses for transcriptomics and proteomics were performed using the EnrichR R
984 package (v3.0) [170] using genes/proteins significantly up-and downregulated between clusters
985 vs. control and between distinct clusters. Two widely used databases, gene ontologies (GOs),
986 related to molecular function, biological process, and cellular components, and KEGG databases,
987 were used. All pathways with an adjusted p-value < 0.05 were considered statistically significant.
988 Similar pathways analyses were performed for single-nuclei (snRNA-seq) using EnrichR using
989 significantly up-and downregulated genes in specific cell types. For metabolomics, pathways
990 analyses were performed using MetaboAnalyst (v5.0) software [171] using the metabolites that
991 were increased- and decreased between clusters and the control and across clusters.
992

993 ## Replication analysis

994 Our replication strategy was designed to focus on the molecular profiles identified in the discovery
995 analyses of the Knight ADRC cohort by extracting the significant genes and proteins identified in
996 the Knight ADRC cohort. Using this set of features, iClusterBayes was run on both datasets
997 (MSBB and ROSMAP) using the same parameters used with the discovery cohort.
998

999 ## NanoString gene expression analysis

1000 The data generation and full details of the QC processes have been previously described [114].
1001 Briefly, A53T αSyn-Tg mice were crossed with human *APOE2*, *APOE3*, or *APOE4* knockin mice
1002 or *Apoe* knockout mice to generate A53T mice homozygous for one of the three human *APOE*
1003 alleles or completely null for *Apoe*. Sample sizes and genotypes for NanoString gene expression
1004 analysis were selected based on αSyn pathology and biochemistry results. RNA was isolated
1005 from 12-month-old A53T/EKO (n=10), A53T/E2 (n=2), and A53T/E4 (n=10) mouse midbrain. 793

1006 transcripts were quantified with the NanoString nCounter multiplexed target platform using a
1007 customized chip based on the Mouse Neuroinflammation panel (www.nanostring.com). The
1008 geometric mean of negative control lanes was subtracted from gene transcript counts, and gene
1009 expression data were normalized to the geometric mean of positive control lanes and a list of
1010 housekeeping genes [114]. QC analyses, including principal components analyses, were
1011 performed using nSolver 4.0 and the Advanced Analysis 2.0 plugin (NanoString). For differential
1012 expression analysis, pSyn pathology was expressed as low (< 5% coverage by IHC) or high (>
1013 5% coverage by IHC) for each animal, and APOE genotype and pSyn pathology were selected
1014 as predictor covariates. Differentially expressed analyses were performed using linear regression
1015 models in R with sex as covariates. P-values were adjusted using Benjamini-Hochberg (BH) [166].
1016

1017 AD staging analysis

1018 Clinical and pathological data, including Reagan, Cerad, and Braak scores as well as final
1019 consensus cognitive diagnosis (cogdx) from the ROSMAP cohort, were used to classify AD cases
1020 into early-AD and late-AD as well as other clinical statuses such as presymptomatic, mild
1021 cognitive impairment (MCI), and control. Differential expression analyses were performed to
1022 compare early and late AD cases to control cases using the DESeq2 R package (v.1.22.2) [165].
1023 Dysregulated synaptic genes from the Knight ADRC profile (Knight-C4) overlapped with the
1024 dysregulated synaptic genes from each comparison from ROSMAP, and the difference in effect
1025 size of the two groups (early-AD vs. Control and late-AD vs. Control) was examined.
1026

1027 Co-expression and gene interaction analyses

1028 Co-expression analyses were performed using dysregulated synaptic genes in Knight-C4 using
1029 the WGCNA R package (v1.72-1) [172]. Gene interaction analyses were performed using
1030 GeneMANIA webserver [173].

Data availability

Transcriptomics data from the Knight ADRC are available by request from the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS) with accession number ng00083 (https://www.niagads.org/datasets/ng00083). Proteomics data from the Knight ADRC donors are available at the NIAGADS Knight ADRC collection with accession number ng00102 (https://www.niagads.org/datasets/ng00102). Metabolomics data from the Knight ADRC are available at the NIAGADS Knight ADRC collection with accession number NG00113 (https://dss.niagads.org/datasets/ng00113/). The single nucleus data from the Knight ADRC are available at NIAGADS Knight ADRC collection with accession number NG00108 (https://dss.niagads.org/datasets/ng00108/). MSBB transcriptomics and proteomics data are publicly available at Synapse under Synapse IDs syn3157743 (https://www.synapse.org/#!Synapse:syn3157743) and syn25006650 (https://www.synapse.org/#!Synapse:syn25006650) respectively. ROSMAP transcriptomics and metabolomics data are publicly available at Synapse under Synapse IDs syn17008934 (https://www.synapse.org/#!Synapse:syn17008934) and syn25878459 (https://www.synapse.org/#!Synapse:syn25878459) respectively.

Code availability

The analysis code used for this study will be available at the time of publication.

1052 # Acknowledgments

1053 We thank all the participants, their families, the many involved institutions, and their staff, whose
1054 help and participation made this work possible. This work was supported by access to equipment
1055 made possible by the Hope Center for Neurological Disorders and the Departments of Neurology
1056 and Psychiatry at Washington University School of Medicine.

1057

1058 ROSMAP (RNA-seq)

1059 Study data were provided by the Rush Alzheimer's Disease Center, Rush University
1060 Medical Center, Chicago. Data collection was supported through funding by NIA grants
1061 P30AG10161 (ROS), R01AG15819 (ROSMAP; genomics and RNAseq), R01AG17917
1062 (MAP), R01AG30146, R01AG36042 (5hC methylation, ATACseq), RC2AG036547
1063 (H3K9Ac), R01AG36836 (RNAseq), R01AG48015 (monocyte RNAseq) RF1AG57473
1064 (single nucleus RNAseq), U01AG32984 (genomic and whole exome sequencing),
1065 U01AG46152 (ROSMAP AMP-AD, targeted proteomics), U01AG46161(TMT
1066 proteomics), U01AG61356 (whole genome sequencing, targeted proteomics, ROSMAP
1067 AMP-AD), the Illinois Department of Public Health (ROSMAP), and the Translational
1068 Genomics Research Institute (genomic). Additional phenotypic data can be requested at
1069 www.radc.rush.edu. Study data were provided through NIA grant 3R01AG046171-02S2
1070 awarded to Rima Kaddurah-Daouk at Duke University, based on specimens provided by
1071 the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, where
1072 data collection was supported through funding by NIA grants P30AG10161,
1073 R01AG15819, R01AG17917, R01AG30146, R01AG36836, U01AG32984,

U01AG46152, the Illinois Department of Public Health, and the Translational Genomics Research Institute.

Alzheimer's Disease Metabolomics Consortium (ADMC)

The results published here are in whole or partly based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org). Metabolomics data is provided by the Alzheimer's Disease Metabolomics Consortium (ADMC) and funded wholly or in part by the following grants and supplements: NIA R01AG046171, RF1AG051550, 3U01AG024904-09S4, RF1AG057452, R01AG059093, RF1AG058942, U01AG061359, U19AG063744, and FNIH: #DAOU16AMPA awarded to Dr. Kaddurah-Daouk at Duke University in partnership with a large number of academic institutions. As such, the investigators within the ADMC, not listed specifically in this publication's author's list, provided data along with its pre-processing and prepared it for analysis but did not participate in the analysis or writing of this manuscript. A complete listing of ADMC investigators can be found at: https://sites.duke.edu/adnimetab/team/.

MSBB

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org/). These data were generated from postmortem brain tissue collected through the Mount Sinai VA Medical Center Brain Bank. Dr. Eric Schadt from Mount Sinai School of Medicine provided them. Dr. Levey provided proteomics data from Emory University based on postmortem brain tissue collected through the Mount Sinai VA Medical Center Brain Bank provided by Dr. Eric Schadt from Mount Sinai School of Medicine.

Funding

Research reported in this work was supported by the Knight Alzheimer Disease Research Center at Washington University School of Medicine through the National Institute on Aging (NIA: grant no. P30 AG066444 - JCM), Healthy Aging and Senile Dementia (HASD: grant no. P01AG003991 - JCM), and Antecedent Biomarkers for Alzheimer Disease: The Adult Children Study (ACS: grant no. P01AG026276 - JCM). This work was supported by grants from the National Institutes of Health: R01AG057777 (OH), K01AG046374 (CMK), R56AG067764 (OH), R01NS118146 (BAB), R21NS127211 (BAB), RF1AG071706 (BAB), and K25AG083057 (AME) and by the Chan Zuckerberg Initiative (CMK). O.H. is an Archer Foundation Research Scientist. A.M.E. is a scholar recipient of the Knight ADRC Research Education Component (NIA: grant no. P30 AG066444).

Author's contributions

A.M.E. conciliated the data, supervised the processing of the RNA-seq, optimized the ML learning approach, performed the analyses, interpreted the data, and wrote the manuscript. A.M.E. and B.C.N. processed, QCed, and analyzed the metabolomics and proteomics data. C.S.T. analyzed and interpreted pathway results. A.M.E. and E.D. processed and QCed all RNA-seq data. L.B. processed and QCed single-nuclei data. A.N. performed NanoString gene expression analyses and prepared all supplementary tables for the manuscript. J.B. and C.S. contributed to the computational analyses and phenotype harmonization. F.F. processed genetic data. K.B., J.B., J.N., J.G., F.W. contributed to data collection, processing, quality control, and cleaning from the Knight ADRC. J.C.M. and R.J.P. contributed samples and/or data. B.A.B., R.J.P., A.A.D., and C.M.K. interpreted the results and provided critical feedback and input to the manuscript. O.H.

conceived and designed the study, collected the data, and supervised the project. O.H. and B.A.B. interpreted the results and supervised the drafting of the manuscript. All the authors critically reviewed the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

References

1. Jellinger KA. Pathobiological subtypes of Alzheimer disease. Dement Geriatr Cogn Disord. 2020;49(4):321–33.

2. Tijms BM, Gobom J, Reus L, Jansen I, Hong S, Dobricic V, et al. Pathophysiological subtypes of Alzheimer's disease based on cerebrospinal fluid proteomics. Brain. 2020;143(12):3776–92.

3. Neff RA, Wang M, Vatansever S, Guo L, Ming C, Wang Q, et al. Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets. Sci Adv. 2021;7(2):eabb5398.

4. Vogel JW, Young AL, Oxtoby NP, Smith R, Ossenkoppele R, Strandberg OT, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. Nat Med. 2021;27(5):871–81.

5. Ferreira D, Nordberg A, Westman E. Biological subtypes of Alzheimer disease: A systematic review and meta-analysis. Neurology. 2020;94(10):436–48.

6. Whitwell JL, Dickson DW, Murray ME, Weigand SD, Tosakulwong N, Senjem ML, et al. Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. Lancet Neurol. 2012;11(10):868–77.

7. Whitwell JL, Graff-Radford J, Tosakulwong N, Weigand SD, Machulda M, Senjem ML, et al. [18F] AV-1451 clustering of entorhinal and cortical uptake in Alzheimer's disease. Ann Neurol. 2018;83(2):248–57.

8. Byun MS, Kim SE, Park J, Yi D, Choe YM, Sohn BK, et al. Heterogeneity of regional brain atrophy patterns associated with distinct progression rates in Alzheimer's disease. PloS One. 2015;10(11):e0142756.

1150 9. Choi M, Youn H, Kim D, Lee S, Suh S, Seong JK, et al. Comparison of neurodegenerative
1151 types using different brain MRI analysis metrics in older adults with normal cognition, mild
1152 cognitive impairment, and Alzheimer's dementia. PloS One. 2019;14(8):e0220739.

1153 10. Dube U, Del-Aguila JL, Li Z, Budde JP, Jiang S, Hsu S, et al. An atlas of cortical circular
1154 RNA expression in Alzheimer disease brains demonstrates clinical and pathological
1155 associations. Nat Neurosci. 2019;22(11):1903–12.

1156 11. Chen HH, Eteleeb A, Wang C, Fernandez MV, Budde JP, Bergmann K, et al. Circular RNA
1157 detection identifies circPSEN1 alterations in brain specific to autosomal dominant
1158 Alzheimer's disease. Acta Neuropathol Commun. 2022;10(1):1–11.

1159 12. Li Z, Del-Aguila JL, Dube U, Budde J, Martinez R, Black K, et al. Genetic variants
1160 associated with Alzheimer's disease confer different cerebral cortex cell-type population
1161 structure. Genome Med. 2018;10(1):1–19.

1162 13. Yang C, Farias FH, Ibanez L, Suhy A, Sadler B, Fernandez MV, et al. Genomic atlas of the
1163 proteome from brain, CSF and plasma prioritizes proteins implicated in neurological
1164 disorders. Nat Neurosci. 2021;24(9):1302–12.

1165 14. Novotny BC, Fernandez MV, Wang C, Budde JP, Bergmann K, Eteleeb AM, et al.
1166 Metabolomic and lipidomic signatures in autosomal dominant and late-onset Alzheimer's
1167 disease brains. Alzheimers Dement J Alzheimers Assoc. 2022 Oct 17;

1168 15. Brase L, You SF, D'Oliveira Albanus R, Del-Aguila JL, Dai Y, Novotny BC, et al. Single-
1169 nucleus RNA-sequencing of autosomal dominant Alzheimer disease and risk variant
1170 carriers. Nat Commun. 2023;14(1):2314.

1171 16. Abeshouse A, Adebamowo C, Adebamowo SN, Akbani R, Akeredolu T, Ally A, et al.
1172 Comprehensive and integrated genomic characterization of adult soft tissue sarcomas.
1173 Cell. 2017;171(4):950–65.

1174 17. Wheeler DA, Roberts LR, Network CGAR. Comprehensive and integrative genomic
1175 characterization of hepatocellular carcinoma. Cell. 2017;169(7):1327.

1176 18. Network CGAR. Comprehensive genomic characterization of squamous cell lung cancers.
1177 Nature. 2012;489(7417):519.

1178 19. Levine DA. Integrated genomic characterization of endometrial carcinoma. Nature.
1179 2013;497(7447):67–73.

1180 20. Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, Bernard B, et al.
1181 Comprehensive molecular characterization of gastric adenocarcinoma. Nature.
1182 2014;513(7517):202.

1183 21. Never-smoker NE smoker. Comprehensive molecular profiling of lung adenocarcinoma.
1184 Nature. 2014;511:543–50.

1185 22. Jendoubi T. Approaches to Integrating Metabolomics and Multi-Omics Data: A Primer.
1186 Metabolites. 2021;11(3):184.

1187 23. Clark C, Dayon L, Masoodi M, Bowman GL, Popp J. An integrative multi-omics approach
1188 reveals new central nervous system pathway alterations in Alzheimer's disease.
1189 Alzheimers Res Ther. 2021;13(1):1–19.

1190 24. Nativio R, Lan Y, Donahue G, Sidoli S, Berson A, Srinivasan AR, et al. An integrated multi-
1191 omics approach identifies epigenetic alterations associated with Alzheimer's disease. Nat
1192 Genet. 2020;52(10):1024–35.

1193 25. Hampel H, Nisticò R, Seyfried NT, Levey AI, Modeste E, Lemercier P, et al. Omics
1194 sciences for systems biology in Alzheimer's disease: State-of-the-art of the evidence.
1195 Ageing Res Rev. 2021;69:101346.

1196 26. Byrns CN, Bonini NM. An Integrated Multi-omics Approach Identifies Therapeutic Potential
1197 for ATP6V1A in Late Onset Alzheimer's Disease. Neuron. 2021;109(2):193–4.

1198 27. Sancesario GM, Bernardini S. Alzheimer's disease in the omics era. Clin Biochem.
1199 2018;59:9–16.

1200 28. Jin T, Rehani P, Ying M, Huang J, Liu S, Roussos P, et al. scGRNom: a computational
1201 pipeline of integrative multi-omics analyses for predicting cell-type disease genes and
1202 regulatory networks. Genome Med. 2021;13(1):1–15.

1203 29. Li Z, Jiang X, Wang Y, Kim Y. Applied machine learning in Alzheimer's disease research:
1204 omics, imaging, and clinical data. Emerg Top Life Sci. 2021;5(6):765–77.

1205 30. Canchi S, Raao B, Masliah D, Rosenthal SB, Sasik R, Fisch KM, et al. Integrating gene
1206 and protein expression reveals perturbed functional networks in Alzheimer's disease. Cell
1207 Rep. 2019;28(4):1103–16.

1208 31. Xie L, He B, Varathan P, Nho K, Risacher SL, Saykin AJ, et al. Integrative-omics for
1209 discovery of network-level disease biomarkers: a case study in Alzheimer's disease. Brief
1210 Bioinform. 2021;22(6):bbab121.

1211 32. Ying M, Rehani P, Roussos P, Wang D. Integrative multi-omics analyses identify cell-type
1212 disease genes and regulatory networks across schizophrenia and Alzheimer's disease.

1213 33. Mo Q, Shen R, Guo C, Vannucci M, Chan KS, Hilsenbeck SG. A fully Bayesian latent
1214 variable model for integrative clustering analysis of multi-type omics data. Biostatistics.
1215 2018;19(1):71–86.

1216 34. Wang M, Beckmann ND, Roussos P, Wang E, Zhou X, Wang Q, et al. The Mount Sinai
1217 cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease.
1218 Sci Data. 2018;5(1):1–16.

1219 35. A Bennett D, A Schneider J, Arvanitakis Z, S Wilson R. Overview and findings from the
1220 religious orders study. Curr Alzheimer Res. 2012;9(6):628–45.

1221 36. Koopmans F, van Nierop P, Andres-Alonso M, Byrnes A, Cijsouw T, Coba MP, et al.
1222 SynGO: an evidence-based, expert-curated knowledge base for the synapse. Neuron.
1223 2019;103(2):217–34.

1224 37. Zhu F, Cizeron M, Qiu Z, Benavides-Piccione R, Kopanitsa MV, Skene NG, et al.
1225 Architecture of the mouse brain synaptome. Neuron. 2018;99(4):781–99.

1226 38. Patel KR, Zhu K, Henrion MY, Beckmann ND, Moein S, Alamprese ML, et al. Single cell-
1227 type integrative network modeling identified novel microglial-specific targets for the
1228 phagosome in Alzheimer's disease. bioRxiv. 2020;

1229 39. Wang X, Huang T, Bu G, Xu H. Dysregulation of protein trafficking in neurodegeneration.
1230 Mol Neurodegener. 2014;9(1):1–9.

1231 40. Wang C, Yu JT, Miao D, Wu ZC, Tan MS, Tan L. Targeting the mTOR signaling network for
1232 Alzheimer's disease therapy. Mol Neurobiol. 2014;49(1):120–35.

1233 41. Hwang J, Estick CM, Ikonne US, Butler D, Pait MC, Elliott LH, et al. The role of lysosomes
1234 in a broad disease-modifying approach evaluated across transgenic mouse models of
1235 Alzheimer's disease and Parkinson's disease and models of mild cognitive impairment. Int
1236 J Mol Sci. 2019;20(18):4432.

1237 42. Butler D, Hwang J, Estick C, Nishiyama A, Kumar SS, Baveghems C, et al. Protective
1238 effects of positive lysosomal modulation in Alzheimer's disease transgenic mouse models.
1239 PloS One. 2011;6(6):e20501.

1240 43. Zhang L, Sheng R, Qin Z. The lysosome and neurodegenerative diseases. Acta Biochim
1241 Biophys Sin. 2009;41(6):437–45.

1242 44. Root J, Merino P, Nuckols A, Johnson M, Kukar T. Lysosome dysfunction as a cause of
1243 neurodegenerative diseases: Lessons from frontotemporal dementia and amyotrophic
1244 lateral sclerosis. Neurobiol Dis. 2021;154:105360.

1245 45. Orr ME, Oddo S. Autophagic/lysosomal dysfunction in Alzheimer's disease. Alzheimers
1246 Res Ther. 2013;5(5):1–9.

1247 46. Eshraghi M, Adlimoghaddam A, Mahmoodzadeh A, Sharifzad F, Yasavoli-Sharahi H,
1248 Lorzadeh S, et al. Alzheimer's disease pathogenesis: Role of autophagy and mitophagy
1249 focusing in microglia. Int J Mol Sci. 2021;22(7):3330.

1250 47. Li Q, Liu Y, Sun M. Autophagy and Alzheimer's disease. Cell Mol Neurobiol.
1251 2017;37(3):377–88.

1252 48. Zare-Shahabadi A, Masliah E, Johnson GV, Rezaei N. Autophagy in Alzheimer's disease.
1253 Rev Neurosci. 2015;26(4):385–95.

1254 49. Nixon RA. The role of autophagy in neurodegenerative disease. Nat Med. 2013;19(8):983–
1255 97.

1256 50. Funderburk SF, Marcellino BK, Yue Z. Cell "self-eating"(autophagy) mechanism in
1257 Alzheimer's disease. Mt Sinai J Med J Transl Pers Med J Transl Pers Med.
1258 2010;77(1):59–68.

1259 51. Dobson CB, Itzhaki RF. Herpes simplex virus type 1 and Alzheimer's disease. Neurobiol
1260 Aging. 1999;20(4):457–65.

1261 52. Itzhaki RF. Herpes simplex virus type 1 and Alzheimer's disease: increasing evidence for a
1262 major role of the virus. Front Aging Neurosci. 2014;6:202.

1263 53. Wozniak MA, Mee AP, Itzhaki RF. Herpes simplex virus type 1 DNA is located within
1264 Alzheimer's disease amyloid plaques. J Pathol J Pathol Soc G B Irel. 2009;217(1):131–8.

1265 54. Itzhaki RF, Lin WR, Shang D, Wilcock GK, Faragher B, Jamieson GA. Herpes simplex virus
1266 type 1 in brain and risk of Alzheimer's disease. The Lancet. 1997;349(9047):241–4.

1267 55. Cairns DM, Rouleau N, Parker RN, Walsh KG, Gehrke L, Kaplan DL. A 3D human brain–
1268 like tissue model of herpes-induced Alzheimer's disease. Sci Adv. 2020;6(19):eaay8828.

1269 56. Duarte LF, Farías MA, Álvarez DM, Bueno SM, Riedel CA, González PA. Herpes simplex
1270 virus type 1 infection of the central nervous system: insights into proposed
1271 interrelationships with neurodegenerative disorders. Front Cell Neurosci. 2019;13:46.

1272 57. Ovsepian SV, O'Leary VB, Zaborszky L, Ntziachristos V, Dolly JO. Synaptic vesicle cycle
1273 and amyloid β: Biting the hand that feeds. Alzheimers Dement. 2018;14(4):502–13.

1274 58. Greber S, Lubec G, Cairns N, Fountoulakis M. Decreased levels of synaptosomal
1275 associated protein 25 in the brain of patients with Down syndrome and Alzheimer's
1276 disease. Electrophor Int J. 1999;20(4–5):928–34.

1277 59. Brinkmalm A, Brinkmalm G, Honer WG, Frölich L, Hausner L, Minthon L, et al. SNAP-25 is
1278 a promising novel cerebrospinal fluid biomarker for synapse degeneration in Alzheimer's
1279 disease. Mol Neurodegener. 2014;9(1):1–13.

1280 60. Abdelhak A, Foschi M, Abu-Rumeileh S, Yue JK, D'Anna L, Huss A, et al. Blood GFAP as
1281 an emerging biomarker in brain and spinal cord disorders. Nat Rev Neurol. 2022;1–15.

1282 61. Colangelo AM, Alberghina L, Papa M. Astrogliosis as a therapeutic target for
1283 neurodegenerative diseases. Neurosci Lett. 2014;565:59–64.

1284 62. Abu-Rumeileh S, Steinacker P, Polischi B, Mammana A, Bartoletti-Stella A, Oeckl P, et al.
1285 CSF biomarkers of neuroinflammation in distinct forms and subtypes of neurodegenerative
1286 dementia. Alzheimers Res Ther. 2020;12(1):1–15.

1287 63. Ishiki A, Kamada M, Kawamura Y, Terao C, Shimoda F, Tomita N, et al. Glial fibrillar acidic
1288 protein in the cerebrospinal fluid of Alzheimer's disease, dementia with Lewy bodies, and
1289 frontotemporal lobar degeneration. J Neurochem. 2016;136(2):258–61.

1290 64. Jesse S, Steinacker P, Cepek L, Arnim CV, Tumani H, Lehnert S, et al. Glial fibrillary acidic
1291 protein and protein S-100B: different concentration pattern of glial proteins in
1292 cerebrospinal fluid of patients with Alzheimer's disease and Creutzfeldt-Jakob disease. J
1293 Alzheimers Dis. 2009;17(3):541–51.

1294 65. Oeckl P, Halbgebauer S, Anderl-Straub S, Steinacker P, Huss AM, Neugebauer H, et al.
1295 Glial fibrillary acidic protein in serum is increased in Alzheimer's disease and correlates
1296 with cognitive impairment. J Alzheimers Dis. 2019;67(2):481–8.

1297 66. Wallin A, Blennow K, Rosengren LE. Glial fibrillary acidic protein in the cerebrospinal fluid
1298 of patients with dementia. Dement Geriatr Cogn Disord. 1996;7(5):267–72.

1299 67. Cicognola C, Janelidze S, Hertze J, Zetterberg H, Blennow K, Mattsson-Carlgren N, et al.
1300 Plasma glial fibrillary acidic protein detects Alzheimer pathology and predicts future
1301 conversion to Alzheimer dementia in patients with mild cognitive impairment. Alzheimers
1302 Res Ther. 2021;13(1):1–9.

1303 68. Chatterjee P, Pedrini S, Stoops E, Goozee K, Villemagne VL, Asih PR, et al. Plasma glial
1304 fibrillary acidic protein is elevated in cognitively normal older adults at risk of Alzheimer's
1305 disease. Transl Psychiatry. 2021;11(1):1–10.

1306 69. Vranová HP, Hényková E, Mareš J, Kaiserová M, Menšíková K, Vaštík M, et al. Clusterin
1307 CSF levels in differential diagnosis of neurodegenerative disorders. J Neurol Sci.
1308 2016;361:117–21.

1309 70. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-
1310 wide association study identifies variants at CLU and PICALM associated with Alzheimer's
1311 disease. Nat Genet. 2009;41(10):1088–93.

1312 71. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide
1313 association study identifies variants at CLU and CR1 associated with Alzheimer's disease.
1314 Nat Genet. 2009;41(10):1094–9.

1315 72. Moon HJ, Herring SK, Zhao L. Clusterin: a multifaceted protein in the brain. Neural Regen
1316 Res. 2021;16(7):1438.

1317 73. Foster EM, Dangla-Valls A, Lovestone S, Ribe EM, Buckley NJ. Clusterin in Alzheimer's
1318 disease: mechanisms, genetics, and lessons from other pathologies. Front Neurosci.
1319 2019;13:164.

1320 74. DeMattos RB, Cirrito JR, Parsadanian M, May PC, O'Dell MA, Taylor JW, et al. ApoE and
1321 clusterin cooperatively suppress Aβ levels and deposition: evidence that ApoE regulates
1322 extracellular Aβ metabolism in vivo. Neuron. 2004;41(2):193–202.

1323 75. Zandl-Lang M, Fanaee-Danesh E, Sun Y, Albrecher NM, Gali CC, Čančar I, et al.
1324 Regulatory effects of simvastatin and apoJ on APP processing and amyloid-β clearance in
1325 blood-brain barrier endothelial cells. Biochim Biophys Acta BBA-Mol Cell Biol Lipids.
1326 2018;1863(1):40–60.

1327 76. Nuutinen T, Huuskonen J, Suuronen T, Ojala J, Miettinen R, Salminen A. Amyloid-β 1–42
1328 induced endocytosis and clusterin/apoJ protein accumulation in cultured human
1329 astrocytes. Neurochem Int. 2007;50(3):540–7.

1330 77. Yeh FL, Wang Y, Tom I, Gonzalez LC, Sheng M. TREM2 binds to apolipoproteins,
1331 including APOE and CLU/APOJ, and thereby facilitates uptake of amyloid-beta by
1332 microglia. Neuron. 2016;91(2):328–40.

1333 78. McCormick DA. GABA as an inhibitory neurotransmitter in human cerebral cortex. J
1334 Neurophysiol. 1989;62(5):1018–27.

1335 79. Jiménez-Balado J, Eich TS. GABAergic dysfunction, neural network hyperactivity and
1336 memory impairments in human aging and Alzheimer's disease. In: Seminars in Cell &
1337 Developmental Biology. Elsevier; 2021.

1338 80. Pillsbury L, Oria M, Erdman J. Nutrition and traumatic brain injury: improving acute and
1339 subacute health outcomes in military personnel. 2011;

1340 81. Platten M, Nollen EA, Röhrig UF, Fallarino F, Opitz CA. Tryptophan metabolism as a
1341 common therapeutic target in cancer, neurodegeneration and beyond. Nat Rev Drug
1342 Discov. 2019;18(5):379–401.

1343 82. Yan X, Hu Y, Wang B, Wang S, Zhang X. Metabolic dysregulation contributes to the
1344 progression of Alzheimer's disease. Front Neurosci. 2020;14:1107.

1345 83. An Y, Varma VR, Varma S, Casanova R, Dammer E, Pletnikova O, et al. Evidence for brain
1346 glucose dysregulation in Alzheimer's disease. Alzheimers Dement. 2018;14(3):318–29.

1347 84. Wang X, He HJ, Xiong X, Zhou S, Wang WW, Feng L, et al. NAD+ in Alzheimer's Disease:
1348 Molecular Mechanisms and Systematic Therapeutic Evidence Obtained in vivo. Front Cell
1349 Dev Biol. 2021;9:668491.

1350 85. González A, Calfío C, Churruca M, Maccioni RB. Glucose metabolism and AD: evidence
1351 for a potential diabetes type 3. Alzheimers Res Ther. 2022;14(1):1–11.

1352 86. He X, Huang Y, Li B, Gong CX, Schuchman EH. Deregulation of sphingolipid metabolism in
1353 Alzheimer's disease. Neurobiol Aging. 2010;31(3):398–408.

1354 87. Vickland H, Silver J. The role of astrocytes in axon guidance during development and
1355 repair. Glial Cell Dev Basic Princ Clin Relev Jessen KR Richardson WD Eds. 1996;197–
1356 208.

1357 88. Cheng J, Wang S, Dong Y, Yuan Z. The role and regulatory mechanism of hippo signaling
1358 components in the neuronal system. Front Immunol. 2020;11:281.

1359 89. Südhof TC. The synaptic vesicle cycle. Annu Rev Neurosci. 2004;27:509–47.

1360 90. Chanaday NL, Cousin MA, Milosevic I, Watanabe S, Morgan JR. The synaptic vesicle cycle
1361 revisited: new insights into the modes and mechanisms. J Neurosci. 2019;39(42):8209–
1362 16.

1363 91. Castillo PE, Younts TJ, Chávez AE, Hashimotodani Y. Endocannabinoid signaling and
1364 synaptic function. Neuron. 2012;76(1):70–81.

1365 92. Dong A, Liu S, Li Y. Gap junctions in the nervous system: probing functional connections
1366 using new imaging approaches. Front Cell Neurosci. 2018;12:320.

1367 93. Blue EE, Thornton TA, Kooperberg C, Liu S, Wactawski-Wende J, Manson J, et al. Non-
1368 coding variants in MYH11, FZD3, and SORCS3 are associated with dementia in women.
1369 Alzheimers Dement. 2021;17(2):215–25.

1370 94. Ou YN, Yang YX, Deng YT, Zhang C, Hu H, Wu BS, et al. Identification of novel drug
1371 targets for Alzheimer's disease by integrating genetics and proteomes from brain and
1372 blood. Mol Psychiatry. 2021;1–9.

1373 95. Parkkinen L, Soininen H, Alafuzoff I. Regional distribution of α-synuclein pathology in
1374 unimpaired aging and Alzheimer disease. J Neuropathol Exp Neurol. 2003;62(4):363–7.

1375 96. Lippa CF, Fujiwara H, Mann DM, Giasson B, Baba M, Schmidt ML, et al. Lewy bodies
1376 contain altered α-synuclein in brains of many familial Alzheimer's disease patients with
1377 mutations in presenilin and amyloid precursor protein genes. Am J Pathol.
1378 1998;153(5):1365–70.

1379 97. Hulette C, Mirra S, Wilkinson W, Heyman A, Fillenbaum G, Clark C. The Consortium to
1380 Establish a Registry for Alzheimer's Disease (CERAD): Part IX. A prospective
1381 cliniconeuropathologic study of Parkinson's features in Alzheimer's disease. Neurology.
1382 1995;45(11):1991–5.

1383 98. Hansen LA. The Lewy body variant of Alzheimer disease. In: Dementia in Parkinsonism.
1384 Springer; 1997. p. 83–93.

1385 99. Hamilton RL. Lewy bodies in Alzheimer's disease: a neuropathological review of 145 cases
1386 using α-synuclein immunohistochemistry. Brain Pathol. 2000;10(3):378–84.

1387 100. Arai Y, Yamazaki M, Mori O, Muramatsu H, Asano G, Katayama Y. α-Synuclein-positive
1388 structures in cases with sporadic Alzheimer's disease: morphology and its relationship to
1389 tau aggregation. Brain Res. 2001;888(2):287–96.

1390 101. Cairns NJ, Perrin RJ, Franklin EE, Carter D, Vincent B, Xie M, et al. Neuropathologic
1391 assessment of participants in two multi-center longitudinal observational studies: The A
1392 lzheimer D isease N euroimaging I nitiative (ADNI) and the D ominantly I nherited A
1393 lzheimer N etwork (DIAN). Neuropathology. 2015;35(4):390–400.

1394 102. Bassil F, Brown HJ, Pattabhiraman S, Iwasyk JE, Maghames CM, Meymand ES, et al.
1395 Amyloid-beta (Aβ) plaques promote seeding and spreading of alpha-synuclein and tau in a
1396 mouse model of Lewy body disorders with Aβ pathology. Neuron. 2020;105(2):260–75.

1397 103. Bassil F, Meymand ES, Brown HJ, Xu H, Cox TO, Pattabhiraman S, et al. α-Synuclein
1398 modulates tau spreading in mouse brains. J Exp Med. 2021;218(1).

1399 104. Henderson MX, Trojanowski JQ, Lee VMY. α-Synuclein pathology in Parkinson's disease
1400 and related α-synucleinopathies. Neurosci Lett. 2019;709:134316.

1401 105. Volpicelli-Daley LA, Luk KC, Patel TP, Tanik SA, Riddle DM, Stieber A, et al. Exogenous
1402 α-synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron
1403 death. Neuron. 2011;72(1):57–71.

1404 106. Lippa CF, Schmidt ML, Lee VMY, Trojanowski JQ. Antibodies to α-synuclein detect Lewy
1405 bodies in many Down's syndrome brains with Alzheimer's disease. Ann Neurol Off J Am
1406 Neurol Assoc Child Neurol Soc. 1999;45(3):353–7.

1407 107. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, et al. Mutation in
1408 the alpha-synuclein gene identified in families with Parkinson's disease. Science. 1997
1409 Jun 27;276(5321):2045–7.

1410 108. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-
1411 synuclein in Lewy bodies. Nature. 1997 Aug 28;388(6645):839–40.

1412 109. Masliah E, Rockenstein E, Veinbergs I, Mallory M, Hashimoto M, Takeda A, et al.
1413 Dopaminergic loss and inclusion body formation in alpha-synuclein mice: implications for
1414 neurodegenerative disorders. Science. 2000 Feb 18;287(5456):1265–9.

1415 110. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, et al. alpha-
1416 Synuclein locus triplication causes Parkinson's disease. Science. 2003 Oct
1417 31;302(5646):841.

1418 111. Giasson BI, Duda JE, Quinn SM, Zhang B, Trojanowski JQ, Lee VMY. Neuronal alpha-
1419 synucleinopathy with severe movement disorder in mice expressing A53T human alpha-
1420 synuclein. Neuron. 2002 May 16;34(4):521–33.

1421 112. Lee MK, Stirling W, Xu Y, Xu X, Qui D, Mandir AS, et al. Human α-synuclein-harboring
1422 familial Parkinson's disease-linked Ala-53→ Thr mutation causes neurodegenerative
1423 disease with α-synuclein aggregation in transgenic mice. Proc Natl Acad Sci.
1424 2002;99(13):8968–73.

1425 113. Luk KC, Kehm V, Carroll J, Zhang B, O'Brien P, Trojanowski JQ, et al. Pathological α-
1426 synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice.
1427 Science. 2012 Nov 16;338(6109):949–53.

1428 114. Davis AA, Inman CE, Wargel ZM, Dube U, Freeberg BM, Galluppi A, et al. APOE
1429 genotype regulates pathology and disease progression in synucleinopathy. Sci Transl
1430 Med. 2020;12(529):eaay3069.

1431 115. Ostrowski PP, Barszczyk A, Forstenpointner J, Zheng W, Feng ZP. Meta-analysis of
1432 serum insulin-like growth factor 1 in Alzheimer's disease. PLoS One.
1433 2016;11(5):e0155733.

1434 116. Gubbi S, Quipildor GF, Barzilai N, Huffman DM, Milman S. IGF-1: the Jekyll & Hyde of the
1435 aging brain. J Mol Endocrinol. 2018;61(1):T171.

1436 117. Galle SA, Van der Spek A, Drent ML, Brugts MP, Scherder EJ, Janssen JA, et al.
1437 Revisiting the role of insulin-like growth factor-I receptor stimulating activity and the
1438 apolipoprotein E in Alzheimer's disease. Front Aging Neurosci. 2019;11:20.

1439 118. Gasparini L, Xu H. Potential roles of insulin and IGF-1 in Alzheimer's disease. Trends
1440 Neurosci. 2003;26(8):404–6.

1441 119. Salehi Z, Mashayekhi F, Naji M. Insulin like growth factor-1 and insulin like growth factor
1442 binding proteins in the cerebrospinal fluid and serum from patients with Alzheimer's
1443 disease. Biofactors. 2008;33(2):99–106.

1444 120. Trueba-Saiz A, Cavada C, Fernandez AM, Leon T, Gonzalez DA, Fortea Ormaechea J, et
1445 al. Loss of serum IGF-I input to the brain as an early biomarker of disease onset in
1446 Alzheimer mice. Transl Psychiatry. 2013;3(12):e330–e330.

1447 121. Hishimoto A, Pletnikova O, Lang DL, Troncoso JC, Egan JM, Liu QR. Neurexin 3
1448 transmembrane and soluble isoform expression and splicing haplotype are associated with
1449 neuron inflammasome and Alzheimer's disease. Alzheimers Res Ther. 2019;11(1):1–15.

1450 122. Zheng JJ, Li WX, Liu JQ, Guo YC, Wang Q, Li GH, et al. Low expression of aging-related
1451 NRXN3 is associated with Alzheimer disease: a systematic review and meta-analysis.
1452 Medicine (Baltimore). 2018;97(28).

1453 123. Huang N, Marie SK, Livramento JA, Chammas R, Nitrini R. 14-3-3 protein in the CSF of
1454 patients with rapidly progressive dementia. Neurology. 2003;61(3):354–7.

1455 124. Lu Y. Early increase of cerebrospinal fluid 14-3-3ζ protein in the alzheimer's disease
1456 continuum. Front Aging Neurosci. 2022;14.

1457 125. Sekiya M, Wang M, Fujisaki N, Sakakibara Y, Quan X, Ehrlich ME, et al. Integrated
1458 biology approach reveals molecular and pathological interactions among Alzheimer's
1459 Aβ42, Tau, TREM2, and TYROBP in Drosophila models. Genome Med. 2018;10(1):1–20.

1460 126. Iacono D, Resnick SM, O'Brien R, Zonderman AB, An Y, Pletnikova O, et al. Mild cognitive
1461 impairment and asymptomatic Alzheimer disease subjects: equivalent β-amyloid and tau
1462 loads with divergent cognitive outcomes. J Neuropathol Exp Neurol. 2014;73(4):295–304.

1463 127. Martorana A, Koch G. Is dopamine involved in Alzheimer's disease? Front Aging Neurosci.
1464 2014;6:252.

1465 128. Koch G, Di Lorenzo F, Bonnì S, Giacobbe V, Bozzali M, Caltagirone C, et al.
1466 Dopaminergic modulation of cortical plasticity in Alzheimer's disease patients.
1467 Neuropsychopharmacology. 2014;39(11):2654–61.

1468 129. Nobili A, Latagliata EC, Viscomi MT, Cavallucci V, Cutuli D, Giacovazzo G, et al.
1469 Dopamine neuronal loss contributes to memory and reward dysfunction in a model of
1470 Alzheimer's disease. Nat Commun. 2017;8(1):1–14.

1471 130. Cordella A, Krashia P, Nobili A, Pignataro A, La Barbera L, Viscomi MT, et al. Dopamine
1472 loss alters the hippocampus-nucleus accumbens synaptic transmission in the Tg2576
1473 mouse model of Alzheimer's disease. Neurobiol Dis. 2018;116:142–54.

1474 131. Ceyzériat K, Gloria Y, Tsartsalis S, Fossey C, Cailly T, Fabis F, et al. Alterations in
1475 dopamine system and in its connectivity with serotonin in a rat model of Alzheimer's
1476 disease. Brain Commun. 2021;3(2):fcab029.

1477 132. Koch G, Esposito Z, Codecà C, Mori F, Kusayanagi H, Monteleone F, et al. Altered
1478 dopamine modulation of LTD-like plasticity in Alzheimer's disease patients. Clin
1479 Neurophysiol. 2011;122(4):703–7.

1480 133. Guzmán-Ramos K, Moreno-Castilla P, Castro-Cruz M, McGaugh JL, Martínez-Coria H,
1481 LaFerla FM, et al. Restoration of dopamine release deficits during object recognition

memory acquisition attenuates cognitive impairment in a triple transgenic mice model of Alzheimer's disease. Learn Mem. 2012;19(10):453–60.

134. Perez SE, Lazarov O, Koprich JB, Chen EY, Rodriguez-Menendez V, Lipton JW, et al. Nigrostriatal dysfunction in familial Alzheimer's disease-linked APPswe/PS1ΔE9 transgenic mice. J Neurosci. 2005;25(44):10220–9.

135. Ambrée O, Richter H, Sachser N, Lewejohann L, Dere E, de Souza Silva MA, et al. Levodopa ameliorates learning and memory deficits in a murine model of Alzheimer's disease. Neurobiol Aging. 2009;30(8):1192–204.

136. Bukke VN, Archana M, Villani R, Romano AD, Wawrzyniak A, Balawender K, et al. The dual role of glutamatergic neurotransmission in Alzheimer's disease: from pathophysiology to pharmacotherapy. Int J Mol Sci. 2020;21(20):7452.

137. Conway ME. Alzheimer's disease: targeting the glutamatergic system. Biogerontology. 2020;21(3):257–74.

138. Rodriguez-Perdigon M, Tordera RM, Gil-Bea FJ, Gerenu G, Ramirez MJ, Solas M. Down-regulation of glutamatergic terminals (VGLUT1) driven by Aβ in Alzheimer's disease. Hippocampus. 2016;26(10):1303–12.

139. Cheng YJ, Lin CH, Lane HY. Involvement of cholinergic, adrenergic, and glutamatergic network modulation with cognitive dysfunction in Alzheimer's disease. Int J Mol Sci. 2021;22(5):2283.

140. Sobue A, Komine O, Hara Y, Endo F, Mizoguchi H, Watanabe S, et al. Microglial gene signature reveals loss of homeostatic microglia associated with neurodegeneration of Alzheimer's disease. Acta Neuropathol Commun. 2021;9(1):1–17.

141. Puntambekar SS, Moutinho M, Lin PBC, Jadhav V, Tumbleson-Brink D, Balaji A, et al. CX3CR1 deficiency aggravates amyloid driven neuronal pathology and cognitive decline in Alzheimer's disease. Mol Neurodegener. 2022 Jun 28;17(1):47.

142. Yamada K, Iwatsubo T. Extracellular α-synuclein levels are regulated by neuronal activity. Mol Neurodegener. 2018;13(1):1–8.

143. Wu Q, Shaikh MA, Meymand ES, Zhang B, Luk KC, Trojanowski JQ, et al. Neuronal activity modulates alpha-synuclein aggregation and spreading in organotypic brain slice cultures and in vivo. Acta Neuropathol (Berl). 2020;140(6):831–49.

144. Fortin DL, Nemani VM, Voglmaier SM, Anthony MD, Ryan TA, Edwards RH. Neural activity controls the synaptic accumulation of α-synuclein. J Neurosci. 2005;25(47):10913–21.

145. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody E, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. Nat Preced. 2010;1–1.

146. Del-Aguila JL, Li Z, Dube U, Mihindukulasuriya KA, Budde JP, Fernandez MV, et al. A single-nuclei RNA sequencing study of Mendelian and sporadic AD in the human brain. Alzheimers Res Ther. 2019;11(1):1–16.

147. Brase L, You SF, Del-Aguila JL, Dai Y, Novotny BC, Sariano-Tarraga C, et al. A landscape of the genetic and cellular heterogeneity in Alzheimer disease. medRxiv. 2021;

148. Johnson EC, Carter EK, Dammer EB, Duong DM, Gerasimov ES, Liu Y, et al. Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong proteomic disease-related changes not observed at the RNA level. Nat Neurosci. 2022;25(2):213–25.

149. Bai B, Wang X, Li Y, Chen PC, Yu K, Dey KK, et al. Deep multilayer brain proteomics identifies molecular networks in Alzheimer's disease progression. Neuron. 2020;105(6):975–91.

150. Arnold M, Nho K, Kueider-Paisley A, Massaro T, Huynh K, Brauner B, et al. Sex and APOE ε4 genotype modify the Alzheimer's disease serum metabolome. Nat Commun. 2020;11(1):1–12.

151. Toledo JB, Arnold M, Kastenmueller G, Chang R, Baillie RA, Han X, et al. Metabolic network failures in Alzheimer's disease: a biochemical road map. Alzheimers Dement. 2017;13(9):965–84.

152. John-Williams S, Blach C, Toledo JB, Rotroff DM, Kim S, Klavins K, et al. Targeted metabolomics and medication classification data from participants in the ADNI1 cohort. Sci Data. 2017;4(1):1–11.

153. Batra R, Arnold M, Wörheide MA, Allen M, Wang X, Blach C, et al. The landscape of metabolic brain alterations in Alzheimer's disease. bioRxiv. 2021;

154. Andrews S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom; 2010.

155. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

156. Picard Toolkit." 2019. Broad Institute, GitHub Repository. https://broadinstitute.github.io/picard/; Broad Institute.

157. Novotny BC, Fernandez MV, Bahena J, Budde JP, Bergmann K, Eteleeb A, et al. Metabolomic and lipidomic signatures in autosomal dominant and late-onset Alzheimer disease brains. medRxiv. 2021;

158. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008;25:1–18.

159. Hastie T, Tibshirani R, Narasimhan B, Chu G. impute: impute: Imputation for microarray data. R package version; 2018.

160. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–3.

161. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8(1):1–12.

162. Mo Q, Shen R. iClusterPlus: integrative clustering of multiple genomic data sets. 2013;

163. Gaujoux R, Seoighe C. CellMix: a comprehensive toolbox for gene expression deconvolution. Bioinformatics. 2013;29(17):2211–2.

164. Gaujoux R, Seoighe C. Semi-supervised Nonnegative Matrix Factorization for gene expression deconvolution: a case study. Infect Genet Evol. 2012;12(5):913–21.

165. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.

166. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol. 1995;57(1):289–300.

167. Therneau T. A Package for Survival Analysis in R. R package version 3.2-3 [Internet]. 2020. Available from: https://CRAN.R-project.org/package=survival

168. Eklund A. surviplot: Draw Survival Curves with Number-at-Risk. R package version 1.1.1 [Internet]. 2017. Available from: http://www.cbs.dtu.dk/~eklund/surviplot/

169. Team RDC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL Httpwww Rproject Org. 2008;

170. Jawaid W. enrichR: Provides an R Interface to “Enrichr”. R package version 3.0. https://CRAN.R-project.org/package=enrichR. 2021;

171. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021;49(W1):W388–96.

172. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):1–13.

173. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;38(suppl_2):W214–20.

Supporting information

S1 Fig. The maximum deviance ratio (left) and the minimum Bayesian information criterion—BIC (right) results across cluster solutions. The optimal solution is observed with cluster number *K=3* dividing our samples into four clusters (*k+1*).

1587

1588 **S2 Fig. Knight-C4 is not associated with polygenic risk scores (PRS)**. Boxplots showing the
1589 polygenic risk scores (PRS) across four clusters using four different p-value thresholds showing
1590 no significant association between Knight-C4 and PRS scores.

1591

1592 **S3 Fig. Knight-C4 is not associated with genetic or neuropathological effects.** (**a**) The
1593 distribution of the APOE ε4 allele across four clusters shows no association with Knight-C4. (**b**)
1594 The distribution of Braak scores for tau across four clusters shows no significant correlation
1595 between Knight-C4 and Braak staging for tau. However, Knight-C4 exhibited more stage VI cases
1596 exceeding the mean distribution of Braak stage VI (dashed line) across all clusters. (**d**) The
1597 distribution of Braak scores for amyloid-β across four clusters showing no association with Knight-
1598 C4.

1599

1600 **S4 Fig. Dysregulated genes are significantly common across cohorts from multiple brain**
1601 **regions.** (**a**) Venn diagrams depicting the overlap between significant genes in Knight-C4 and
1602 MSBB-C1 cohorts, showing a replication of the molecular signature associated with worse
1603 cognitive outcomes. (**b**) Same as "a" but for significant proteins. (**c**) Venn diagrams depicting the
1604 overlap between significant genes in Knight-C4 and ROSMAP-C1 cohorts, showing a replication
1605 of the molecular signature associated with worse cognitive outcomes.

1606

1607 **S5 Fig. Molecular profiles of MSBB-C1 are associated with high Braak scores for tau and**
1608 **replicated in additional brain regions from the MSBB study.** (**a**) Boxplots showing MSBB-C1
1609 are associated with high BraakTau scores. (**b**) Boxplots show a clear pattern of an association
1610 with the early age of death for cases in MSBB-C1. (**c**) Heatmaps showing the replication of
1611 transcriptomic profiles of MSBB-C1 from the BM36 region in additional regions, including frontal

1612 pole (BM10), superior temporal gyrus (BM22), and inferior frontal gyrus (BM44) brain regions from
1613 the MSBB study.
1614
1615 **S6 Fig. Molecular profiles of ROSMAP-C1 are associated with multiple neuropathological**
1616 **variables. (a)** Boxplots showing ROSMAP-C1 association with high tangle densities. (**b**) Boxplots
1617 showing ROSMAP-C1 association with neurofibrillary tangle burdens. (**c**) Boxplots showing
1618 higher plaque densities for AD cases in ROSMAP-C1.
1619
1620 **S7 Fig. Knight-C4 is associated with more significant hits across multi-omics datasets**. (**a**)
1621 Venn diagrams show the differentially expressed genes detected in each cluster (cluster vs
1622 control). (**b**) Same as panel "a" but for proteomics. (**c**) Same as panel "a" but for metabolomics.
1623
1624 **S8 Fig. Significant enrichment of dysregulated synaptic genes in Knight-C4 is replicated**
1625 **in additional brain regions.** (**a**) Venn Diagrams showing the overlap between significant genes
1626 in MSBB-C1 and synaptic genes from Fei *et al.,* 2018 [37] and SynGO [36] datasets. (**b**) Same
1627 as "a" but overlaps with significant genes in ROSMAP-C1.
1628
1629 **S9 Fig. Top significant pathways dysregulated in AD molecular profiles from multiple**
1630 **cortical brain regions and AD cohorts.** (**a**). Top common KEGG pathways dysregulated in
1631 Knight-C4 and MSBB-C1 (Cluster vs other AD cases). (**b**) Top common KEGG pathways
1632 dysregulated in Knight-C4 and ROSMAP-C1 (Cluster vs other AD cases). (**c**) Top common GO
1633 biological process pathways dysregulated in Knight-C4 and ROSMAP-C1 (Cluster vs other AD
1634 cases). (**d**) Top 20 KEGG pathways enriched in Knight-C4 downregulated genes (Knight-C4 vs
1635 control). (**e**) Top 20 GO biological process pathways enriched in Knight-C4 downregulated genes
1636 (Knight-C4 vs control).
1637

1638 **S10 Fig. Top significant pathways dysregulated in AD molecular profiles from multiple**
1639 **cortical brain regions and AD cohorts compared to control cases.** (**a**) Top common KEGG
1640 pathways dysregulated in Knight-C4 and MSBB-C1. (**b**) Top common KEGG pathways
1641 dysregulated in Knight-C4 and ROSMAP-C1. (**c**) Top common GO biological process pathways
1642 dysregulated in Knight-C4 and MSBB-C1. (**d**) Top common GO biological process pathways
1643 dysregulated in Knight-C4 and ROSMAP-C1. (**e**) Top 20 KEGG pathways associated with the
1644 downregulation of proteins in Knight-C4. (**f**) Top 20 GO biological process pathways enriched in
1645 Knight-C4 downregulated proteins.
1646
1647 **S11 Fig. The decreased expression of *SNAP25* in Knight-C4 is replicated in MSBB and**
1648 **ROSMAP cohorts**. (**a**) Boxplots showing transcriptomic profiles of *SNAP25* across the two
1649 clusters (right) and all ADs (left) in the MSBB (BM36) cohort. (**b**) Boxplots showing proteomic
1650 (TMT) profiles of *SNAP25* across the two clusters as mentioned in "a." (**c**) Boxplots showing
1651 transcriptomic profiles of *SNAP25* across the two clusters (right) and all ADs (left) in ROSMAP
1652 (DLPFC) cohort.
1653
1654 **S12 Fig. *GFAP* proteomic levels are dysregulated across distinct clusters.** *GFAP* protomeric
1655 levels are significantly increased in Knight-C2 compared to the control and other AD cases (right).
1656
1657 **S13 Fig. The increased level of *GFAP* is replicated in MSBB and ROSMAP datasets**. (**a**)
1658 Boxplots showing the transcriptomic profiles of *GFAP* across the two clusters (right) and all AD
1659 cases (left) in the MSBB (BM36) cohort. (**b**) Same as "a" but for the proteomic profiles (TMT). (**c**)
1660 Same as "a" but in ROSMAP (DLPFC) cohort.
1661
1662 **S14 Fig. Differential abundance and pathways analyses identified metabolites decreased**
1663 **in Knight-C4 and ROSMAP-C1. (a**) Boxplots showing the decrease of metabolomics profiles of

1664 gamma-aminobutyrate (GABA) in Knight-C4 (left) and ROSMAP-C1 (right). (**b**) Same as "a" but
1665 for choline metabolite.
1666
1667 **S15 Fig. Differential abundance and pathway analyses identified metabolites increased in**
1668 **Knight-C4 and ROSMAP-C1**. **(a**) Boxplots showing the increase of metabolomic profiles of
1669 tryptophan in Knight-C4 (left) and ROSMAP-C1 (right). (**b**) Same as "a" but for glucose metabolite.
1670
1671 **S16 Fig. Metabolic differential abundance and pathway analyses identified multiple**
1672 **dysregulated carbohydrate/sugar and sphingolipid metabolites. (a)** Boxplots showing the
1673 decrease of metabolomic profiles of citrate in Knight-C4 (left) and ROSMAP-C1. (**b**) Boxplots
1674 showing the decrease of metabolomic profiles of NAD+ in both Knight-C4 (left) and ROSMAP-C1
1675 (right). (**d**) Boxplots showing the increased metabolism levels of sphingosine in both Knight-C4
1676 (left) and ROSMAP-C1 (right).
1677
1678 **S17 Fig. Integrating single-nuclei data with cross-omics profiles identified cell-type**
1679 **specific pathways. (a)** Top 20 KEGG pathways enriched in endothelial upregulated genes
1680 (Knight-C4 vs Control). (**b**) Top 20 KEGG pathways enriched in microglia downregulated genes
1681 (Knight-C4 vs Control).
1682
1683 **S18 Fig. Downregulation of *SNCA* is replicated in additional brain regions from the MSBB**
1684 **study.** (**a**) Boxplot showing the transcriptomic profiles of *SNCA* in MSBB-C1 from the frontal pole
1685 (BM10) brain region. (**b**) Same as "a" but from superior temporal gyrus (BM22). (**d**) Same as "a"
1686 and "b" but from the inferior frontal gyrus (BM44) brain region from MSBB.
1687
1688 **S19 Fig. The synaptic dysregulation in Knight-C4 is partially recapitulated in mouse**
1689 **models. (a)** Venn diagram showing the overlap between significant genes in Knight-C4 compared

1690 to other clusters, significant genes from NanoString gene expression data from A53T αSyn mouse
1691 models, and synaptic genes from SynGO dataset. (b) Same as "a" but using significant genes in
1692 Knight-C4 compared to the control.
1693
1694 **S20 Fig. Gene dendrogram and module colors generated by WGCNA showing two**
1695 **distinct gene networks (gray=64, turquoise=209).**
1696
1697 **S1 Table.** Demographic characteristics of AD and control cases.
1698 **S2 Table.** Cell proportions for the Knight ADRC, MSBB (BM36), and ROSMAP (DLPFC)
1699 cohorts.
1700 **S3 Table.** Differentially expressed genes in the Knight ADRC cohort.
1701 **S4 Table.** Differentially expressed genes/proteins in the MSBB (BM36) cohort.
1702 **S5 Table.** Overlap of significant genes between discovery cohort (Knight ADRC) and replication
1703 cohorts (MSBB and ROSMAP cohorts).
1704 **S6 Table.** Differentially abundant genes/metabolites in ROSMAP (DLPFC) cohort.
1705 **S7 Table.** Differentially expressed proteins in the Knight ADRC control.
1706 **S8 Table.** Differentially abundant metabolites in Knight ADRC cohort.
1707 **S9 Table.** Summary of the significant features detected between each cluster and the control.
1708 **S10 Table.** Number of significant features between each cluster and the others.
1709 **S11 Table.** Overlap between significant genes in Knight ADRC, MSBB (BM36), ROSMAP
1710 (DLPFC) and synaptic genes.
1711 **S12 Table.** Overlap between differentially abundant metabolites in Knight ADRC and ROSMAP
1712 cohorts.
1713 **S13 Table.** Number and percentage of significant features specific to each cell type across four
1714 Knight clusters.
1715 **S14 Table.** List of protein-protein interaction (PPI) hub proteins.

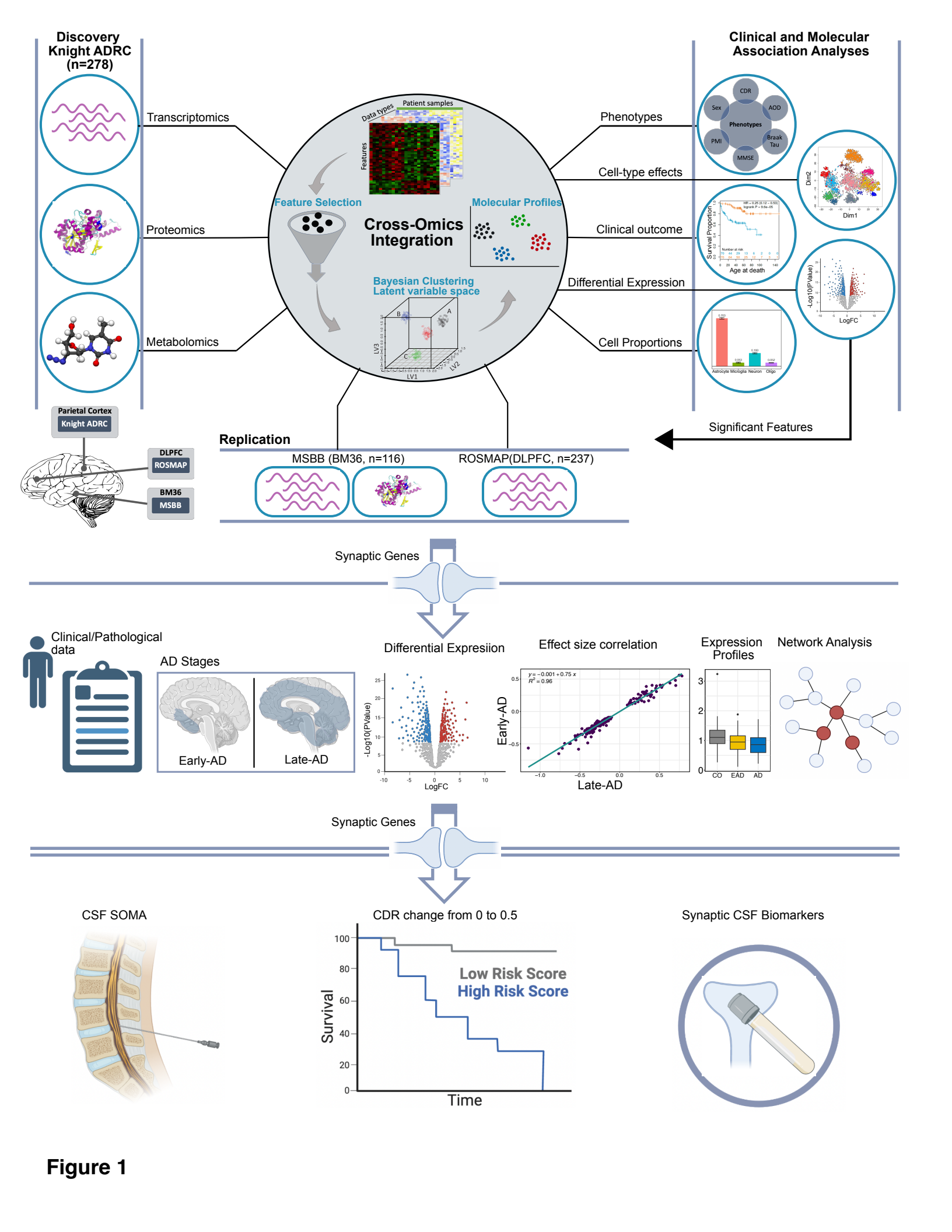

Discovery
Knight ADRC
(n=278)
Transcriptomics
Proteomics
Metabolomics
Patient samples
Data types
Features
Feature Selection
Cross-Omics
Integration
Molecular Profiles
Bayesian Clustering
Latent variable space
LV1
LV2
LV3
Clinical and Molecular
Association Analyses
Phenotypes
Cell-type effects
Clinical outcome
Differential Expression
Cell Proportions
CDR
Sex
AOD
Phenotypes
PMI
Braak
Tau
MMSE
Dim1
Dim2
Survival Proportion
Age at death
-Log10(PValue)
LogFC
Astrocyte
Microglia
Neuron
Oligo
Significant Features
Parietal Cortex
Knight ADRC
DLPFC
ROSMAP
BM36
MSBB
Replication
MSBB (BM36, n=116)
ROSMAP(DLPFC, n=237)
Synaptic Genes
Clinical/Pathological
data
AD Stages
Early-AD
Late-AD
Differential Expresiion
-Log10(PValue)
LogFC
Effect size correlation
Early-AD
Late-AD
Expression
Profiles
CO
EAD
AD
Network Analysis
Synaptic Genes
CSF SOMA
CDR change from 0 to 0.5
Survival
Time
Low Risk Score
High Risk Score
Synaptic CSF Biomarkers

Figure 1

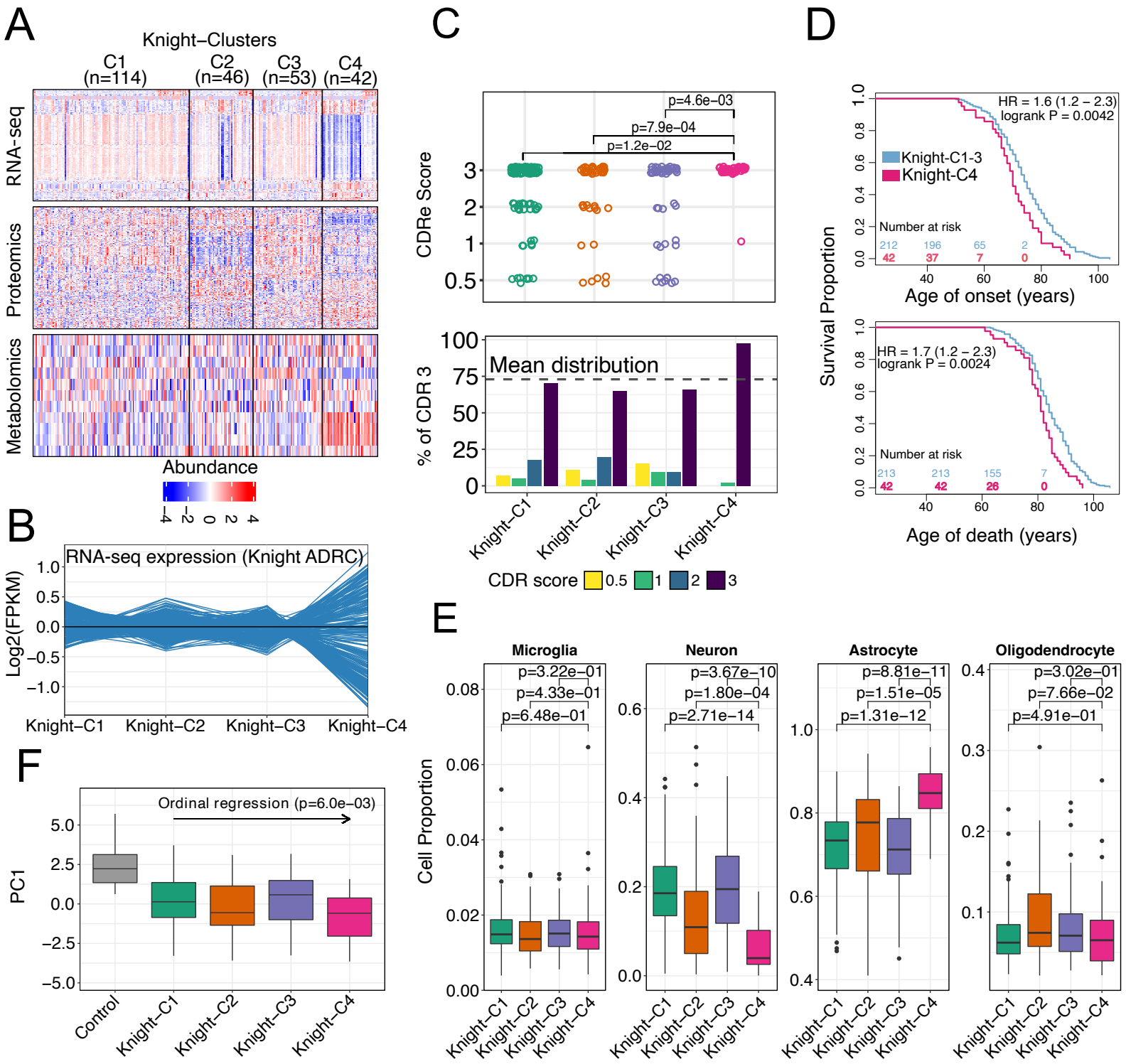

A
Knight−Clusters
C1 (n=114)
C2 (n=46)
C3 (n=53)
C4 (n=42)
RNA-seq
Proteomics
Metabolomics
Abundance
B
RNA-seq expression (Knight ADRC)
Log2(FPKM)
Knight−C1
Knight−C2
Knight−C3
Knight−C4
C
CDRe Score
p=4.6e−03
p=7.9e−04
p=1.2e−02
% of CDR 3
Mean distribution
CDR score
D
HR = 1.6 (1.2 − 2.3)
logrank P = 0.0042
Knight-C1-3
Knight-C4
Number at risk
Survival Proportion
Age of onset (years)
HR = 1.7 (1.2 − 2.3)
logrank P = 0.0024
Age of death (years)
E
Microglia
Neuron
Astrocyte
Oligodendrocyte
Cell Proportion
p=3.22e−01
p=4.33e−01
p=6.48e−01
p=3.67e−10
p=1.80e−04
p=2.71e−14
p=8.81e−11
p=1.51e−05
p=1.31e−12
p=3.02e−01
p=7.66e−02
p=4.91e−01
F
Ordinal regression (p=6.0e−03)
PC1
Control
Knight−C1
Knight−C2
Knight−C3
Knight−C4

Figure 2

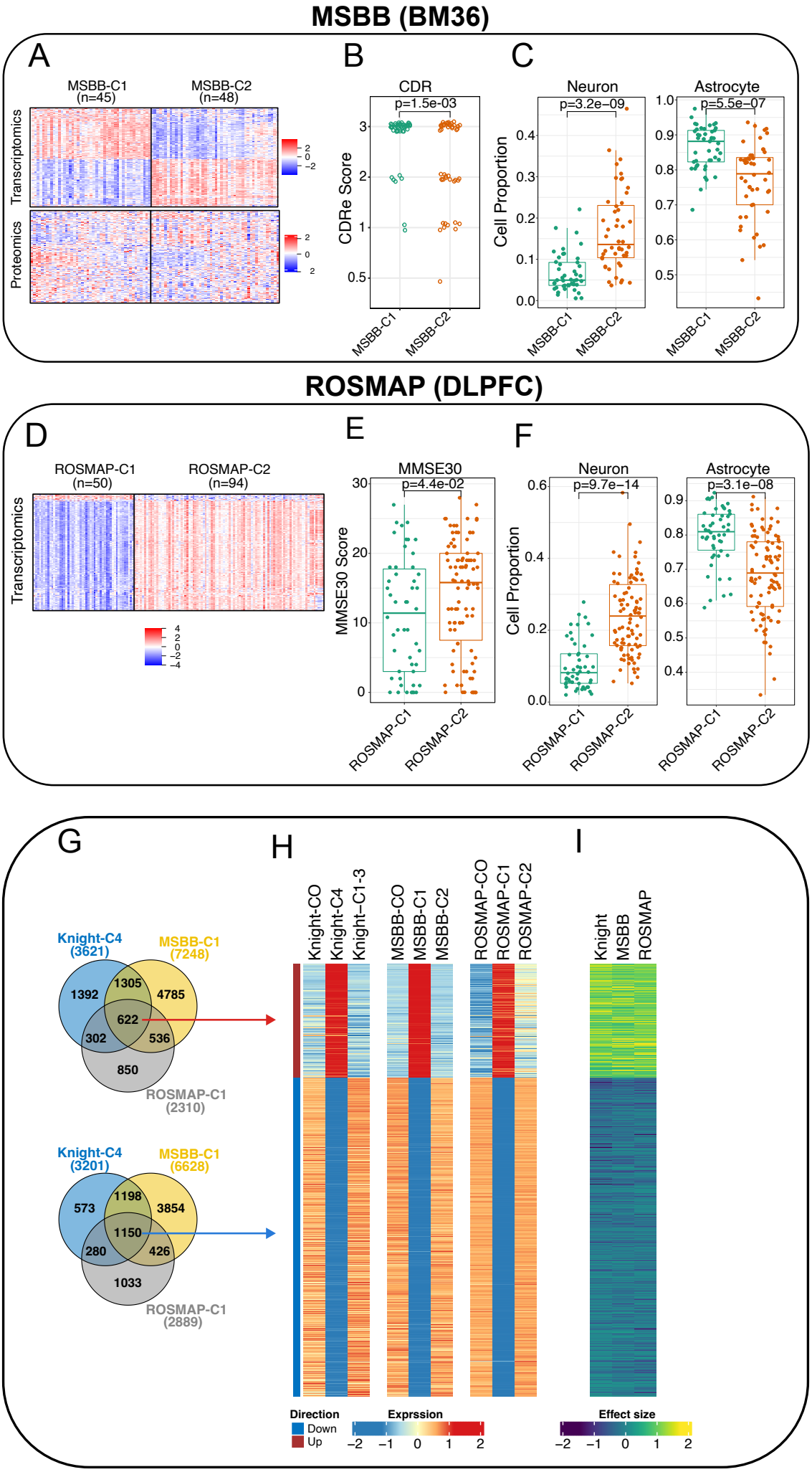

MSBB (BM36)
A
MSBB-C1 (n=45)
MSBB-C2 (n=48)
Transcriptomics
Proteomics
B
CDR
p=1.5e-03
CDRe Score
MSBB-C1
MSBB-C2
C
Neuron
p=3.2e−09
Astrocyte
p=5.5e−07
Cell Proportion
ROSMAP (DLPFC)
D
ROSMAP-C1 (n=50)
ROSMAP-C2 (n=94)
E
MMSE30
p=4.4e-02
MMSE30 Score
ROSMAP-C1
ROSMAP-C2
F
Neuron
p=9.7e−14
Astrocyte
p=3.1e−08
G
Knight-C4 (3621)
MSBB-C1 (7248)
ROSMAP-C1 (2310)
1392
1305
4785
622
302
536
850
Knight-C4 (3201)
MSBB-C1 (6628)
ROSMAP-C1 (2889)
573
1198
3854
1150
280
426
1033
H
Knight-CO
Knight-C4
Knight–C1-3
MSBB-CO
MSBB-C1
MSBB-C2
ROSMAP-CO
ROSMAP-C1
ROSMAP-C2
I
Knight
MSBB
ROSMAP
Direction
Down
Up
Exprssion
Effect size

Figure 3

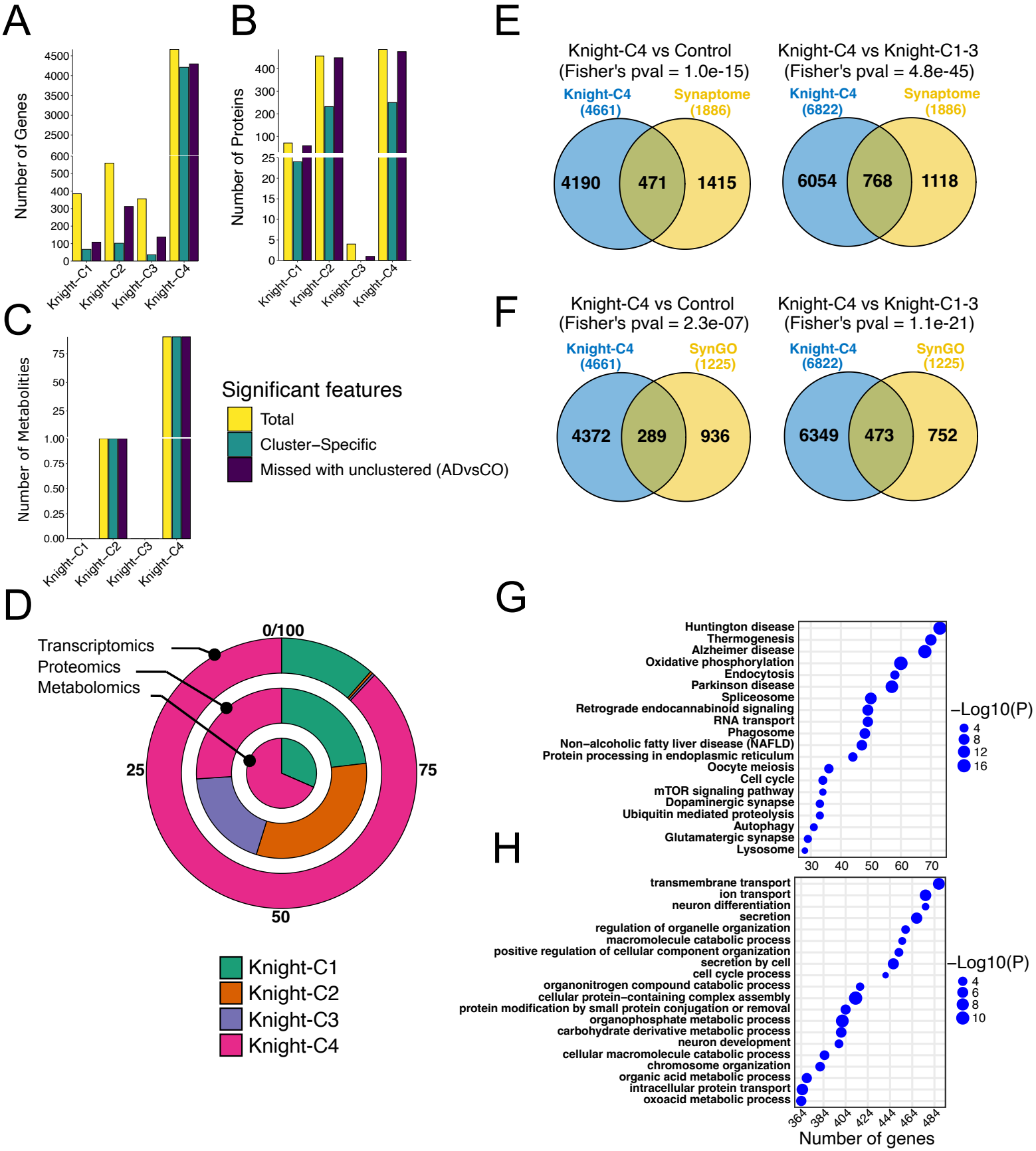

A
Number of Genes
Knight–C1
Knight–C2
Knight–C3
Knight–C4
B
Number of Proteins
C
Number of Metabolities
Significant features
Total
Cluster–Specific
Missed with unclustered (ADvsCO)
D
Transcriptomics
Proteomics
Metabolomics
0/100
25
75
50
Knight-C1
Knight-C2
Knight-C3
Knight-C4
E
Knight-C4 vs Control
(Fisher's pval = 1.0e-15)
Knight-C4 (4661)
Synaptome (1886)
4190
471
1415
Knight-C4 vs Knight-C1-3
(Fisher's pval = 4.8e-45)
Knight-C4 (6822)
Synaptome (1886)
6054
768
1118
F
Knight-C4 vs Control
(Fisher's pval = 2.3e-07)
Knight-C4 (4661)
SynGO (1225)
4372
289
936
Knight-C4 vs Knight-C1-3
(Fisher's pval = 1.1e-21)
Knight-C4 (6822)
SynGO (1225)
6349
473
752
G
Huntington disease
Thermogenesis
Alzheimer disease
Oxidative phosphorylation
Endocytosis
Parkinson disease
Spliceosome
Retrograde endocannabinoid signaling
RNA transport
Phagosome
Non–alcoholic fatty liver disease (NAFLD)
Protein processing in endoplasmic reticulum
Oocyte meiosis
Cell cycle
mTOR signaling pathway
Dopaminergic synapse
Ubiquitin mediated proteolysis
Autophagy
Glutamatergic synapse
Lysosome
–Log10(P)
H
transmembrane transport
ion transport
neuron differentiation
secretion
regulation of organelle organization
macromolecule catabolic process
positive regulation of cellular component organization
secretion by cell
cell cycle process
organonitrogen compound catabolic process
cellular protein–containing complex assembly
protein modification by small protein conjugation or removal
organophosphate metabolic process
carbohydrate derivative metabolic process
neuron development
cellular macromolecule catabolic process
chromosome organization
organic acid metabolic process
intracellular protein transport
oxoacid metabolic process
Number of genes

Figure 4

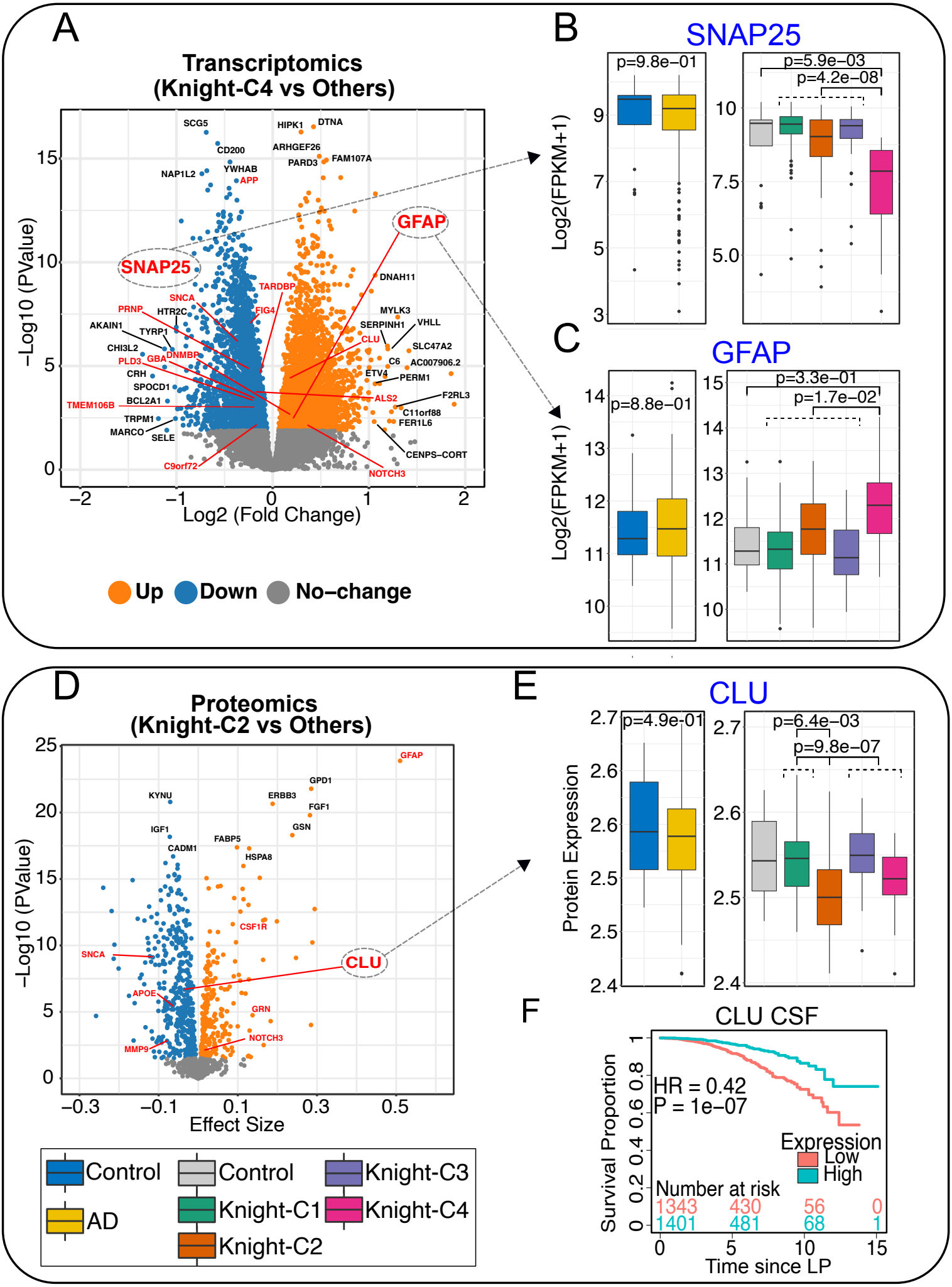
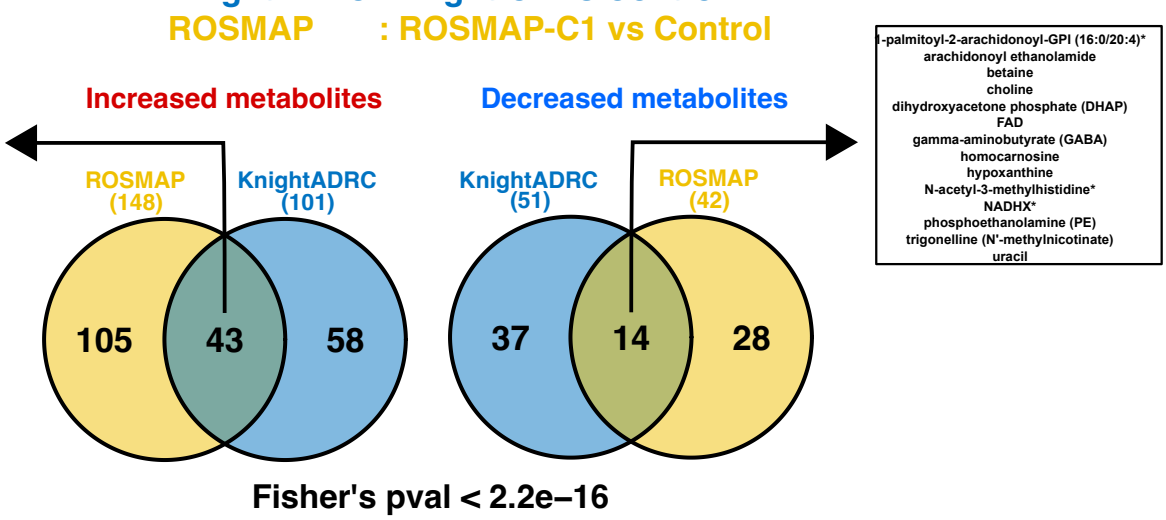

A
Transcriptomics
(Knight-C4 vs Others)
−Log10 (PValue)
Log2 (Fold Change)
SNAP25
GFAP
Up
Down
No−change
B
SNAP25
p=9.8e−01
p=5.9e−03
p=4.2e−08
Log2(FPKM+1)
C
GFAP
p=8.8e−01
p=3.3e−01
p=1.7e−02
Log2(FPKM+1)
D
Proteomics
(Knight-C2 vs Others)
−Log10 (PValue)
Effect Size
GFAP
CLU
SNCA
E
CLU
p=4.9e−01
p=6.4e−03
p=9.8e−07
Protein Expression
F
CLU CSF
HR = 0.42
P = 1e−07
Survival Proportion
Expression
Low
High
Number at risk
1343 430 56 0
1401 481 68 1
Time since LP
Control
AD
Control
Knight-C1
Knight-C2
Knight-C3
Knight-C4

Figure 5

A

1-(1-enyl-palmitoyl)-GPE (P-16:0)*
1-carboxyethylphenylalanine
1-methylnicotinamide
1-stearoyl-GPS (18:0)*
2'-AMP
2'-deoxyguanosine
2'-deoxyinosine
2R,3R-dihydroxybutyrate
3-hydroxypalmitoylcarnitine
5-methyluridine (ribothymidine)
6-oxopiperidine-2-carboxylate
arginine
creatinine
cysteinylglycine disulfide*
dehydroascorbate
dimethylarginine (ADMA + SDMA)
erythronate*
glucose
glycerophosphoethanolamine
glycerophosphorylcholine (GPC)
glycylvaline
inositol 1-phosphate (I1P)
mannitol/sorbitol
methionine sulfoxide
N-acetyl-glucosamine 1-phosphate
N-formylanthranilic acid
N-formylphenylalanine
N-monomethylarginine
N2,N2-dimethylguanosine
N6-acetyllysine
phenylalanine
phenylalanylglycine
ribose
saccharopine
sphingadienine
sphingosine
threonate
tryptophan
tyrosine
tyrosylglycine
UDP-glucose
urea



B

1-(1-enyl-palmitoyl)-GPE (P-16:0)*
1-carboxyethylisoleucine
1-carboxyethylleucine
1-carboxyethylphenylalanine
1-carboxyethyltyrosine
1-carboxyethylvaline
1-methyladenosine
1-methylnicotinamide
1-stearoyl-GPS (18:0)*
15-KETE
2-hydroxybutyrate/2-hydroxyisobutyrate
2'-AMP
2'-deoxyguanosine
2'-deoxyinosine
2S,3R-dihydroxybutyrate
3-formylindole
6-oxopiperidine-2-carboxylate
7-methylguanine
alanine
arginine
asparagine
carnitine
dehydroascorbate
dimethylarginine (ADMA + SDMA)
fructose
fructosyllysine
glucose
glycerate
glycerophosphoserine*
glycylvaline
inositol 1-phosphate (I1P)
isobutyrylcarnitine (C4)
isoleucine
lactate
leucine
lysine
mannitol/sorbitol
methionine
methionine sulfoxide
N-acetyl-glucosamine 1-phosphate
N-acetylalanine
N-acetylmethionine sulfoxide
N-formylanthranilic acid
N-formylphenylalanine
N-monomethylarginine
N1-methylinosine
N2,N2-dimethylguanosine
N6-acetyllysine
N6,N6,N6-trimethyllysine
palmitoyl sphingomyelin (d18:1/16:0)
phenylalanine
phenylalanylglycine
phytosphingosine
pyridoxamine
ribose
sedoheptulose
serine
spermidine
sphingadienine
sphingomyelin (d18:2/18:1)*
sphingosine
threonine
tryptophan
tyrosine
tyrosylglycine
UDP-glucose
uridine
valine
xanthine
xanthosine

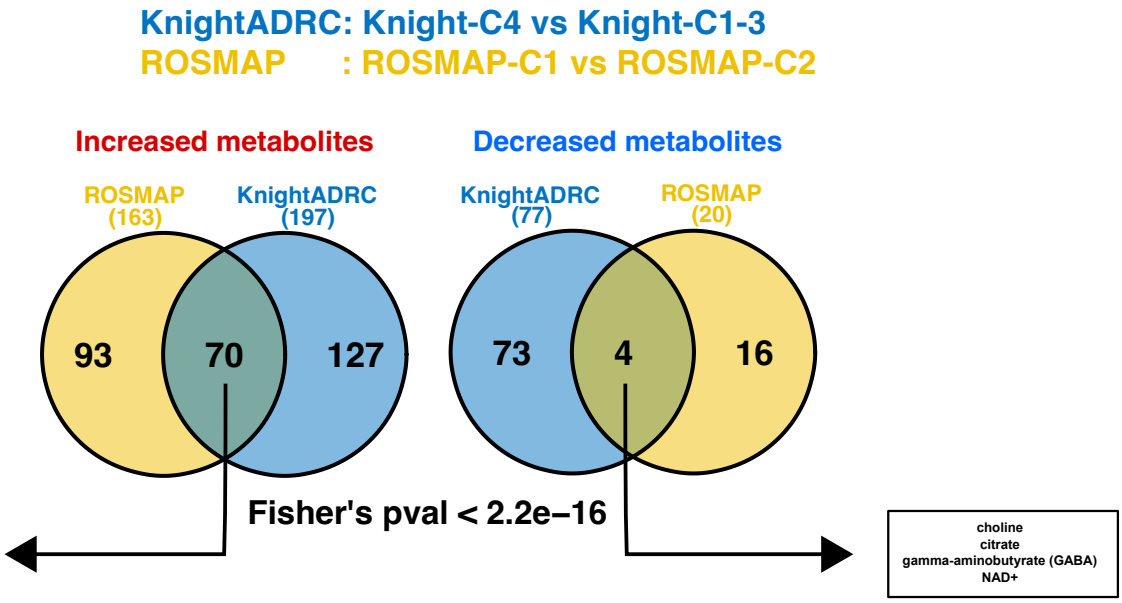


Figure 6

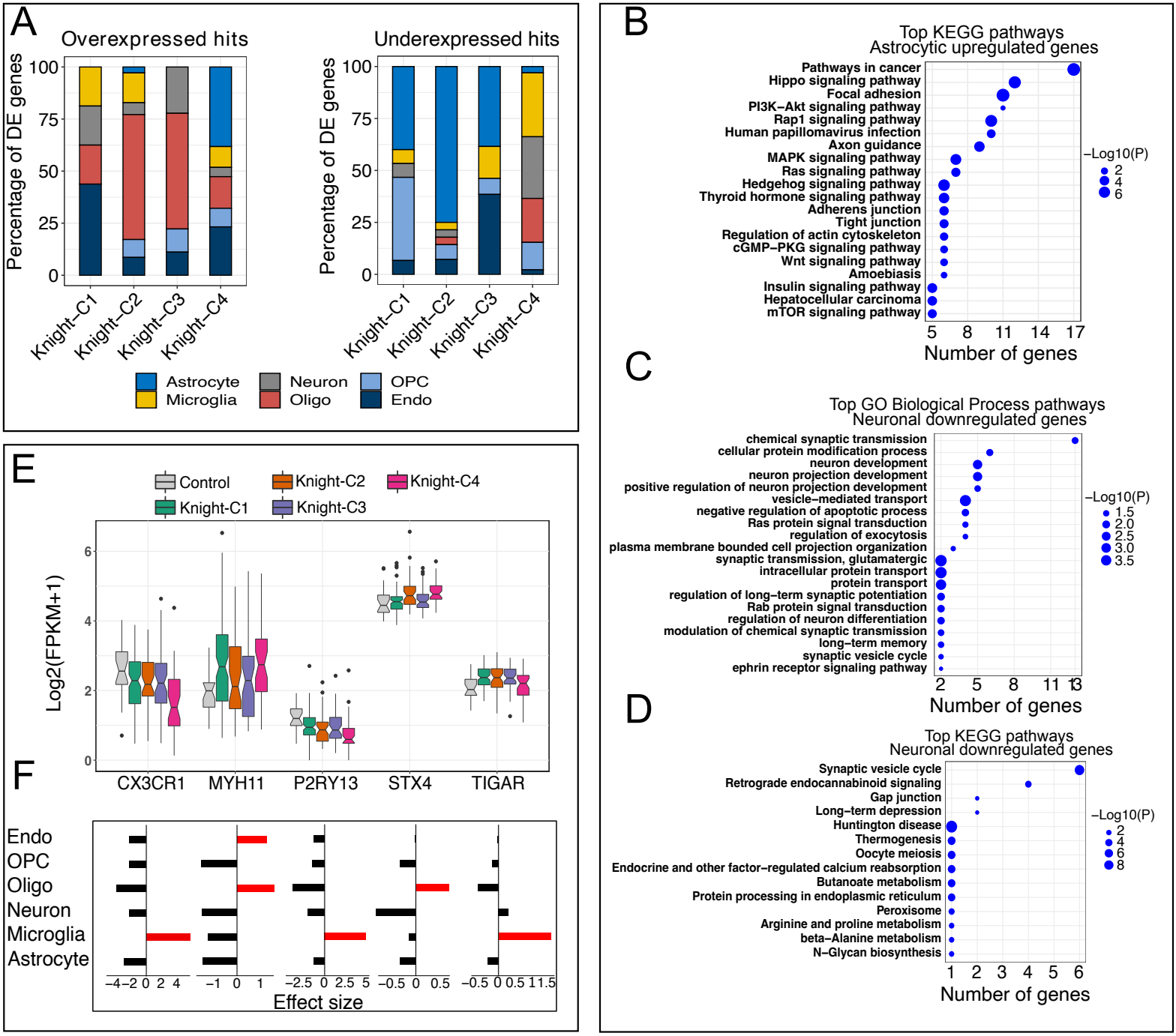

A
Overexpressed hits
Underexpressed hits
Percentage of DE genes
Knight–C1
Knight–C2
Knight–C3
Knight–C4
Astrocyte
Microglia
Neuron
Oligo
OPC
Endo
B
Top KEGG pathways
Astrocytic upregulated genes
Pathways in cancer
Hippo signaling pathway
Focal adhesion
PI3K–Akt signaling pathway
Rap1 signaling pathway
Human papillomavirus infection
Axon guidance
MAPK signaling pathway
Ras signaling pathway
Hedgehog signaling pathway
Thyroid hormone signaling pathway
Adherens junction
Tight junction
Regulation of actin cytoskeleton
cGMP–PKG signaling pathway
Wnt signaling pathway
Amoebiasis
Insulin signaling pathway
Hepatocellular carcinoma
mTOR signaling pathway
–Log10(P)
Number of genes
C
Top GO Biological Process pathways
Neuronal downregulated genes
chemical synaptic transmission
cellular protein modification process
neuron development
neuron projection development
positive regulation of neuron projection development
vesicle–mediated transport
negative regulation of apoptotic process
Ras protein signal transduction
regulation of exocytosis
plasma membrane bounded cell projection organization
synaptic transmission, glutamatergic
intracellular protein transport
protein transport
regulation of long–term synaptic potentiation
Rab protein signal transduction
regulation of neuron differentiation
modulation of chemical synaptic transmission
long–term memory
synaptic vesicle cycle
ephrin receptor signaling pathway
–Log10(P)
Number of genes
D
Top KEGG pathways
Neuronal downregulated genes
Synaptic vesicle cycle
Retrograde endocannabinoid signaling
Gap junction
Long–term depression
Huntington disease
Thermogenesis
Oocyte meiosis
Endocrine and other factor–regulated calcium reabsorption
Butanoate metabolism
Protein processing in endoplasmic reticulum
Peroxisome
Arginine and proline metabolism
beta–Alanine metabolism
N–Glycan biosynthesis
–Log10(P)
Number of genes
E
Control
Knight-C1
Knight-C2
Knight-C3
Knight-C4
Log2(FPKM+1)
CX3CR1
MYH11
P2RY13
STX4
TIGAR
F
Endo
OPC
Oligo
Neuron
Microglia
Astrocyte
Effect size

Figure 7

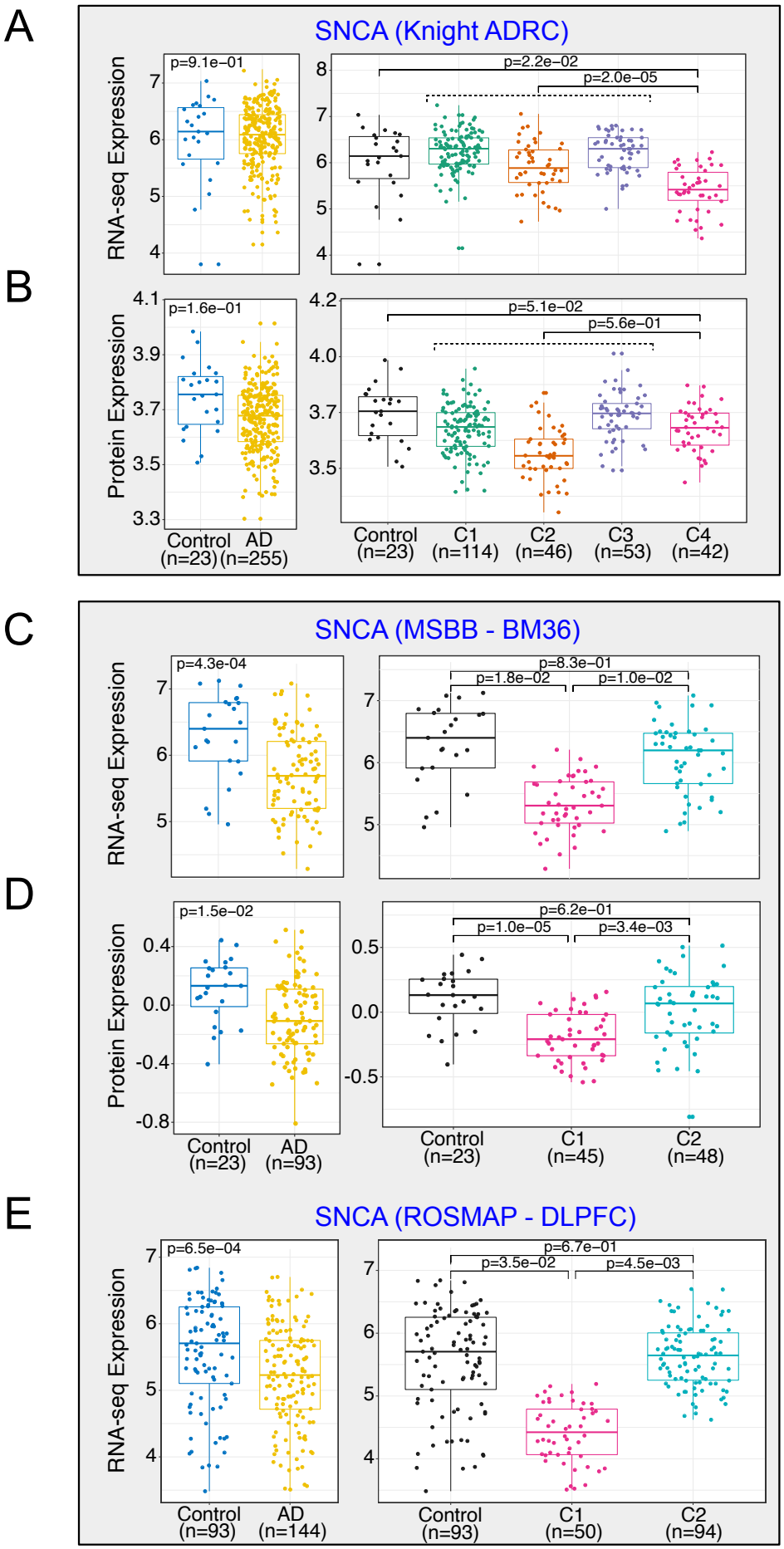

A
SNCA (Knight ADRC)
RNA-seq Expression
p=9.1e−01
p=2.2e−02
p=2.0e−05
B
Protein Expression
p=1.6e−01
p=5.1e−02
p=5.6e−01
Control (n=23)
AD (n=255)
Control (n=23)
C1 (n=114)
C2 (n=46)
C3 (n=53)
C4 (n=42)
C
SNCA (MSBB - BM36)
RNA-seq Expression
p=4.3e-04
p=8.3e−01
p=1.8e−02
p=1.0e−02
D
Protein Expression
p=1.5e−02
p=6.2e−01
p=1.0e−05
p=3.4e−03
Control (n=23)
AD (n=93)
Control (n=23)
C1 (n=45)
C2 (n=48)
E
SNCA (ROSMAP - DLPFC)
RNA-seq Expression
p=6.5e−04
p=6.7e−01
p=3.5e−02
p=4.5e−03
Control (n=93)
AD (n=144)
Control (n=93)
C1 (n=50)
C2 (n=94)

Figure 8

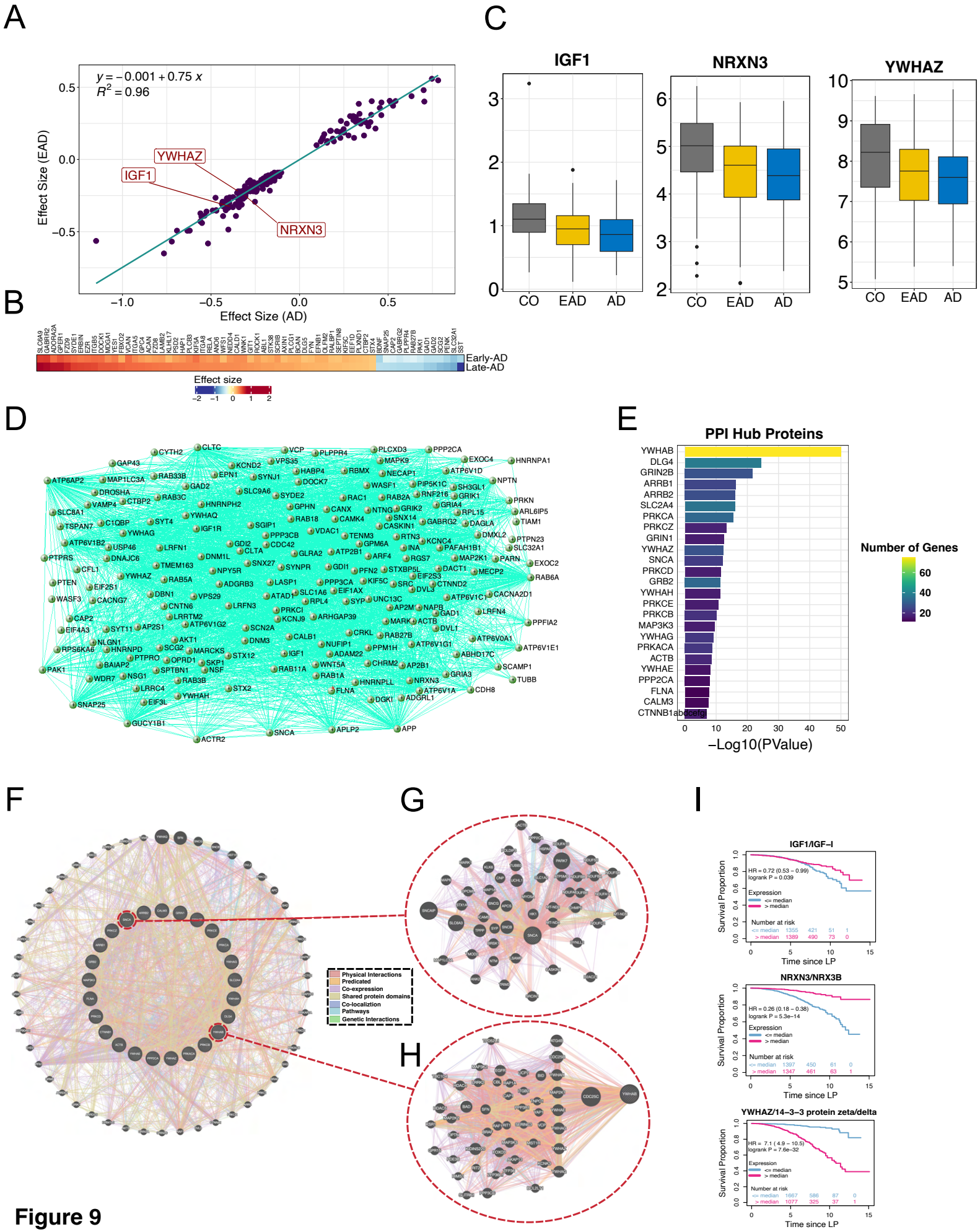
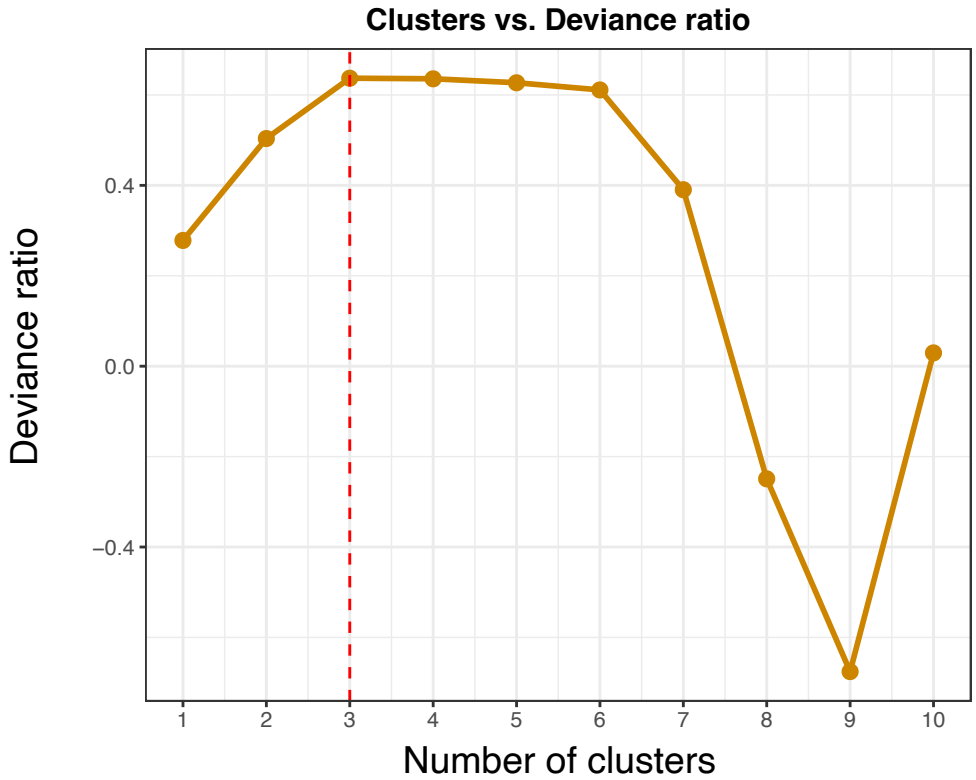
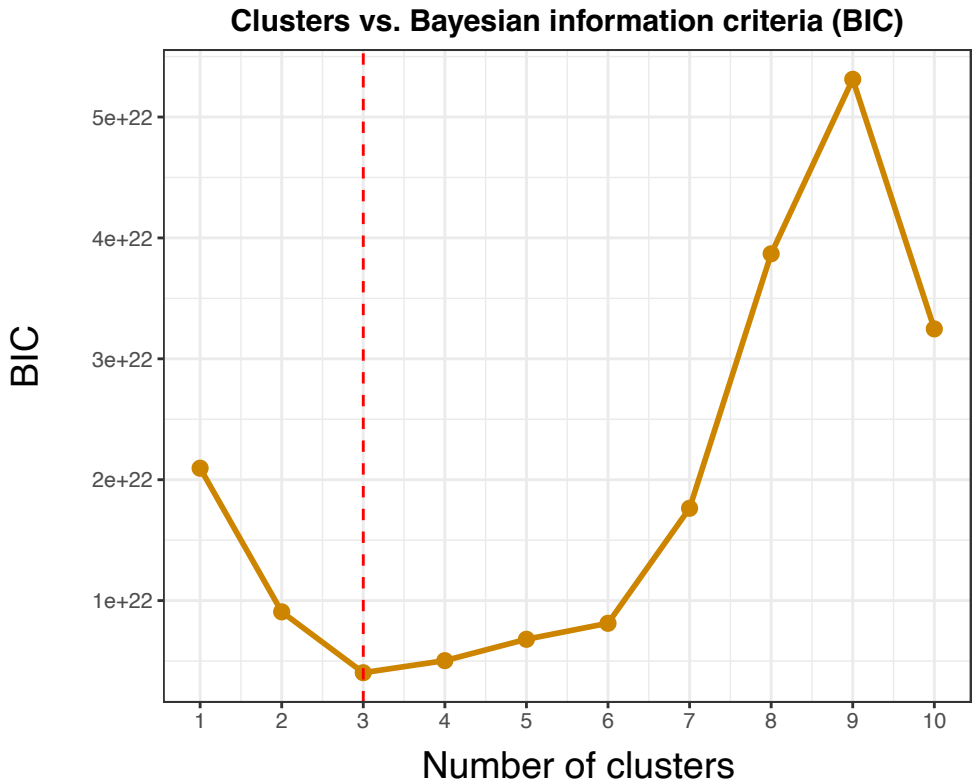
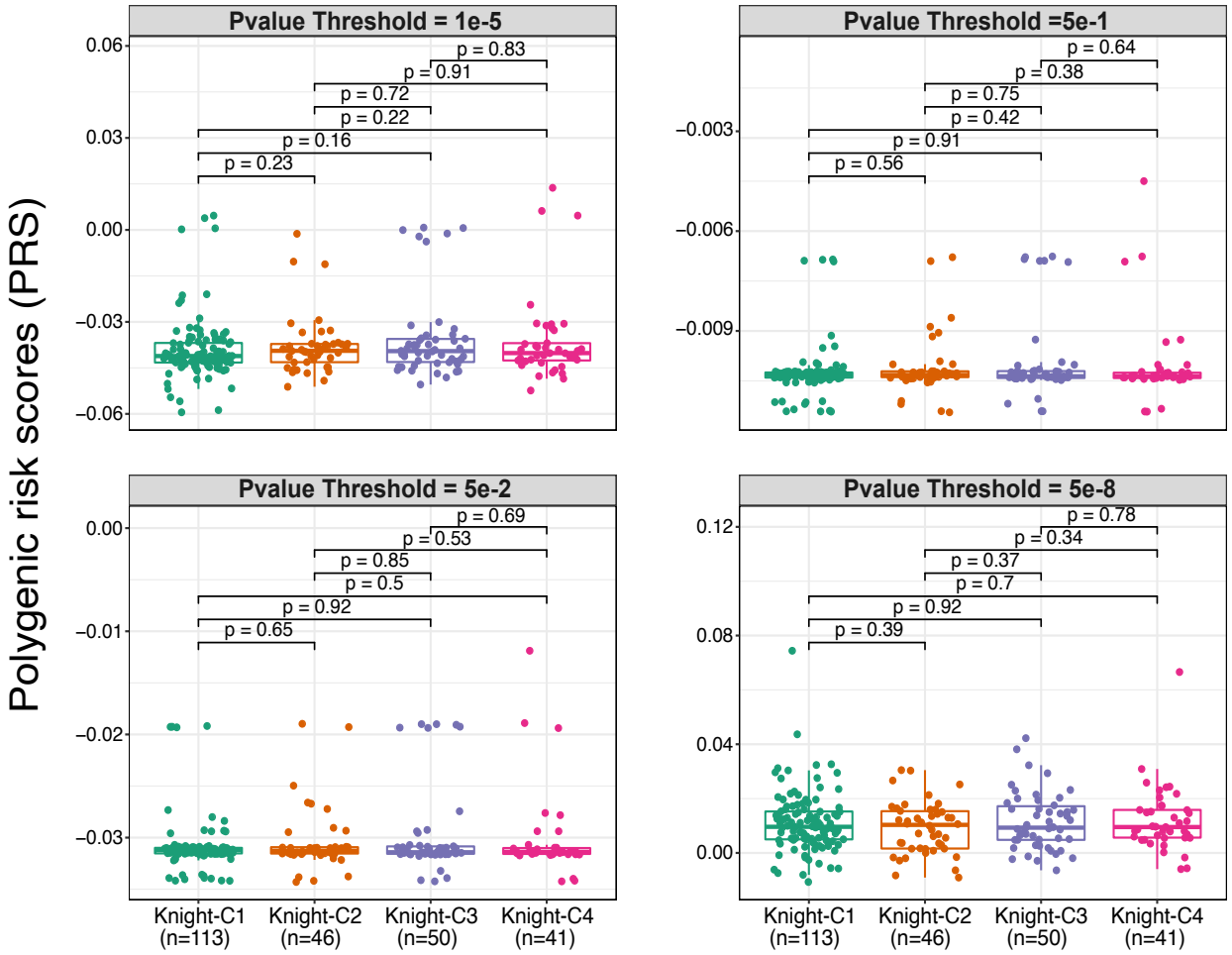

A
y = −0.001 + 0.75 x
R² = 0.96
Effect Size (EAD)
Effect Size (AD)
YWHAZ
IGF1
NRXN3
B
Early-AD
Late-AD
Effect size
C
IGF1
NRXN3
YWHAZ
CO
EAD
AD
D
E
PPI Hub Proteins
YWHAB
DLG4
GRIN2B
ARRB1
ARRB2
SLC2A4
PRKCA
PRKCZ
GRIN1
YWHAZ
SNCA
PRKCD
GRB2
YWHAH
PRKCE
PRKCB
MAP3K3
YWHAG
PRKACA
ACTB
YWHAE
PPP2CA
FLNA
CALM3
CTNNB1
Number of Genes
−Log10(PValue)
F
G
H
I
Physical Interactions
Predicated
Co-expression
Shared protein domains
Co-localiztion
Pathways
Genetic Interactions
IGF1/IGF−I
HR = 0.72 (0.53 − 0.99)
logrank P = 0.039
NRXN3/NRX3B
HR = 0.26 (0.18 − 0.38)
logrank P = 5.3e−14
YWHAZ/14−3−3 protein zeta/delta
HR = 7.1 (4.9 − 10.5)
logrank P = 7.6e−32
Survival Proportion
Time since LP
Expression
<= median
> median
Number at risk

Figure 9

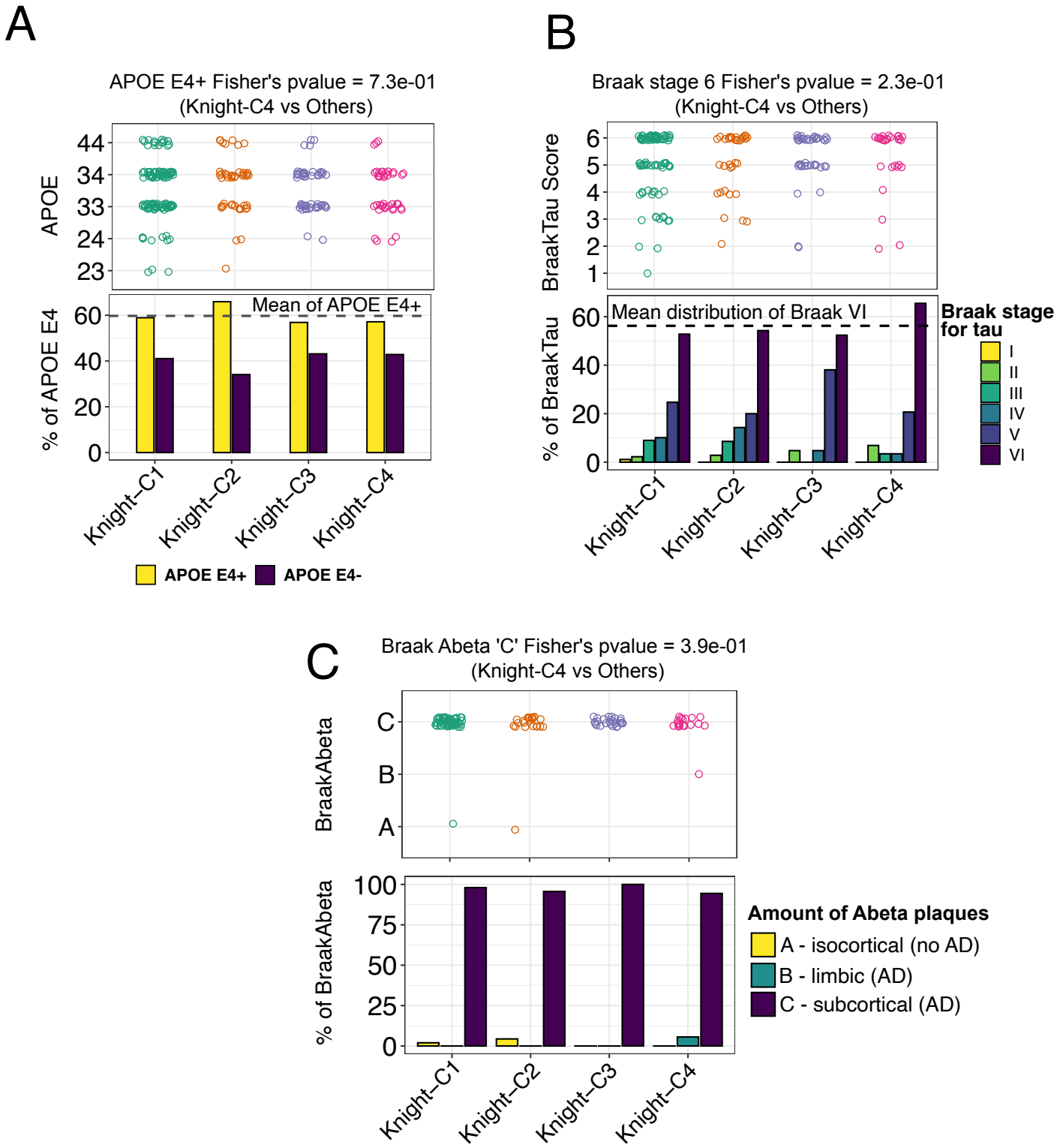




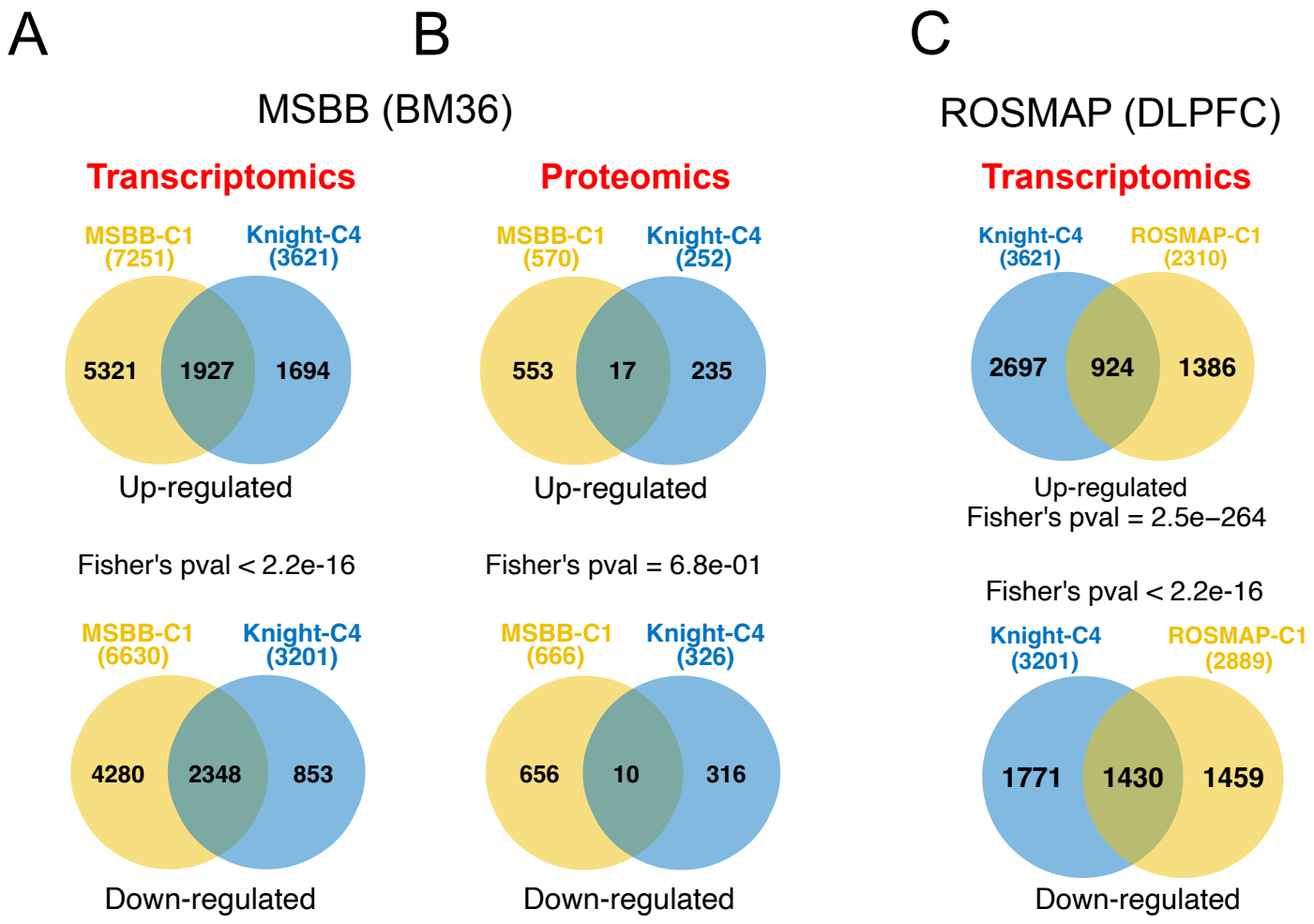
Supplementary Figure 1



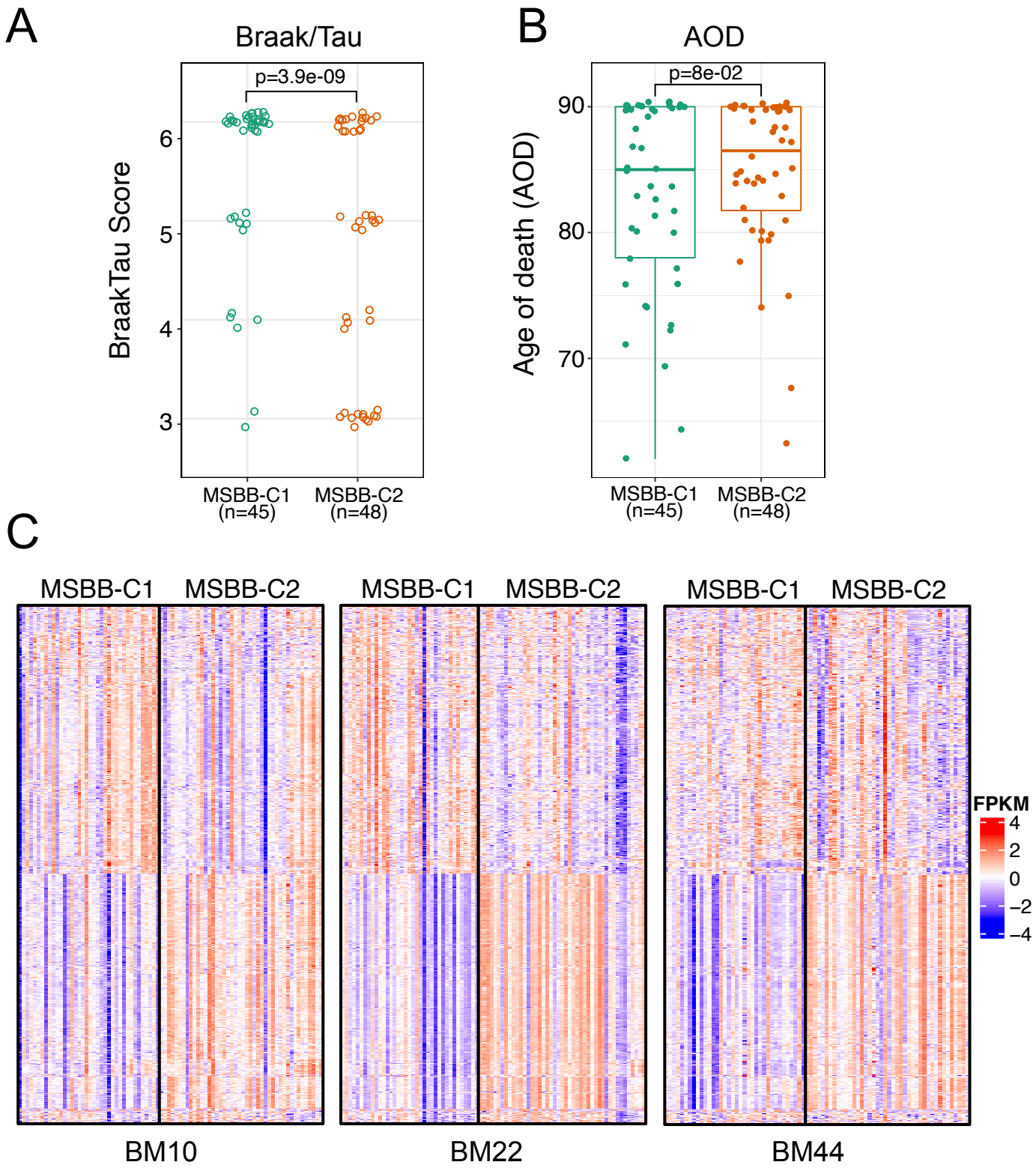
Supplementary Figure 2



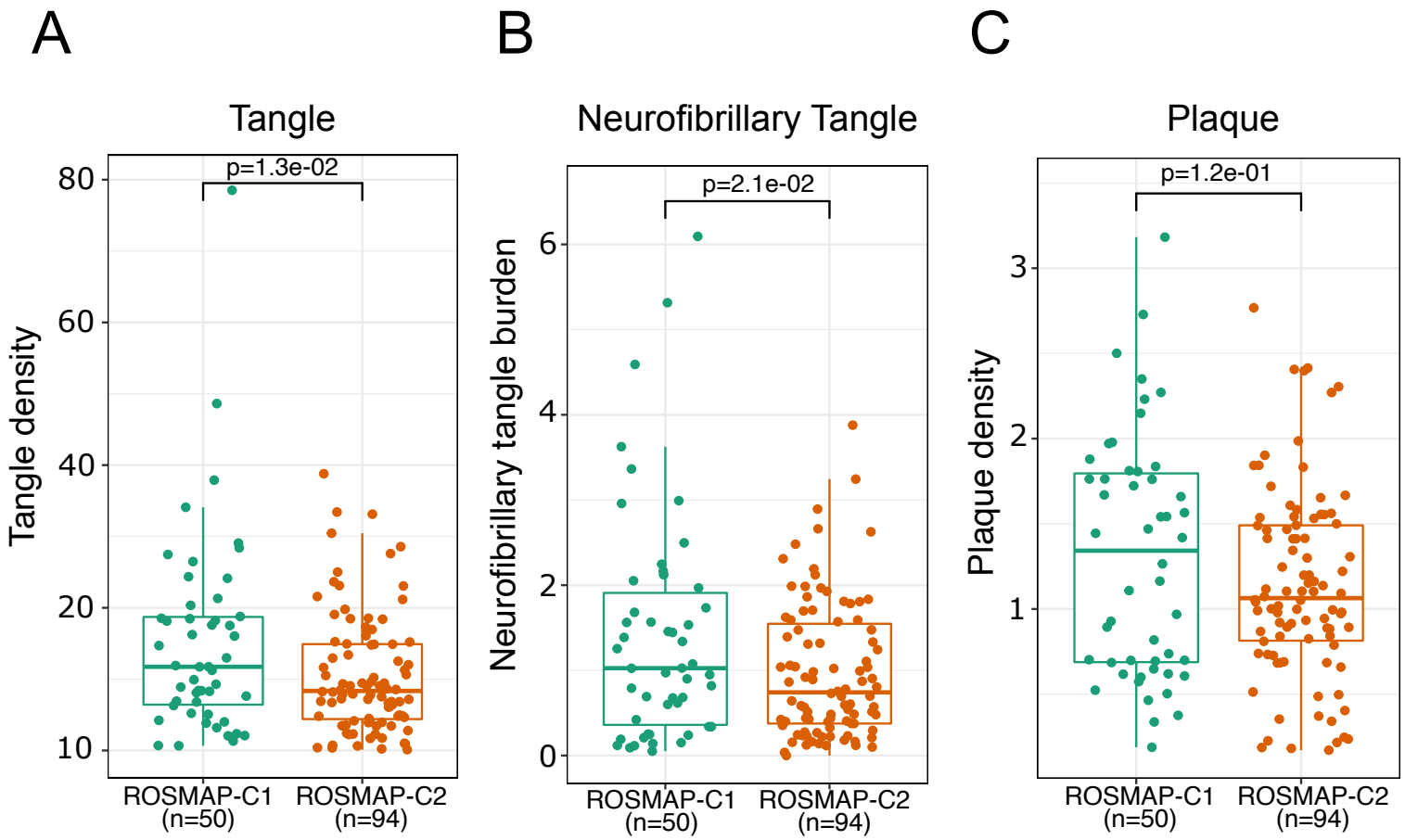
Supplementary Figure 3



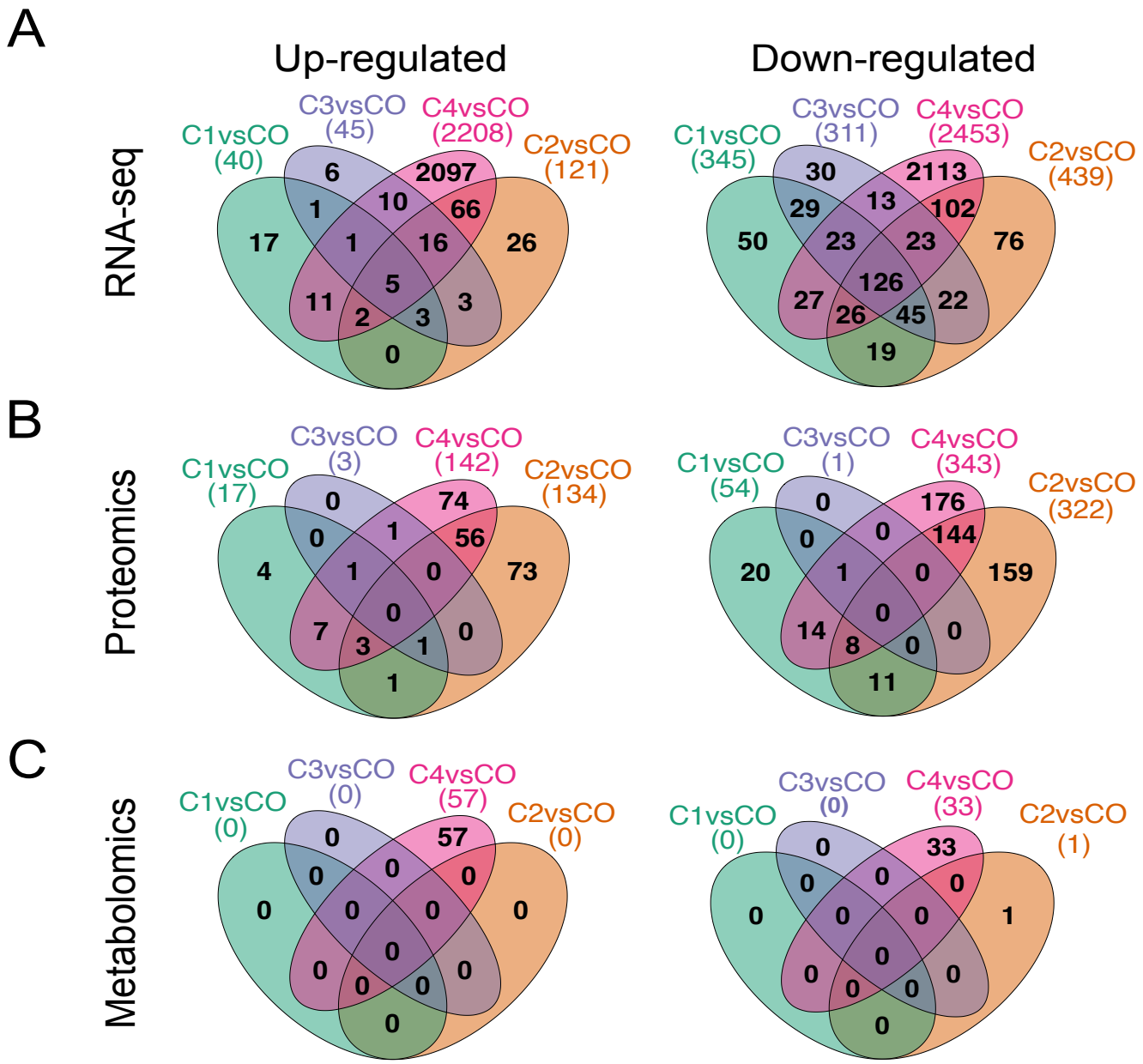
Supplementary Figure 4



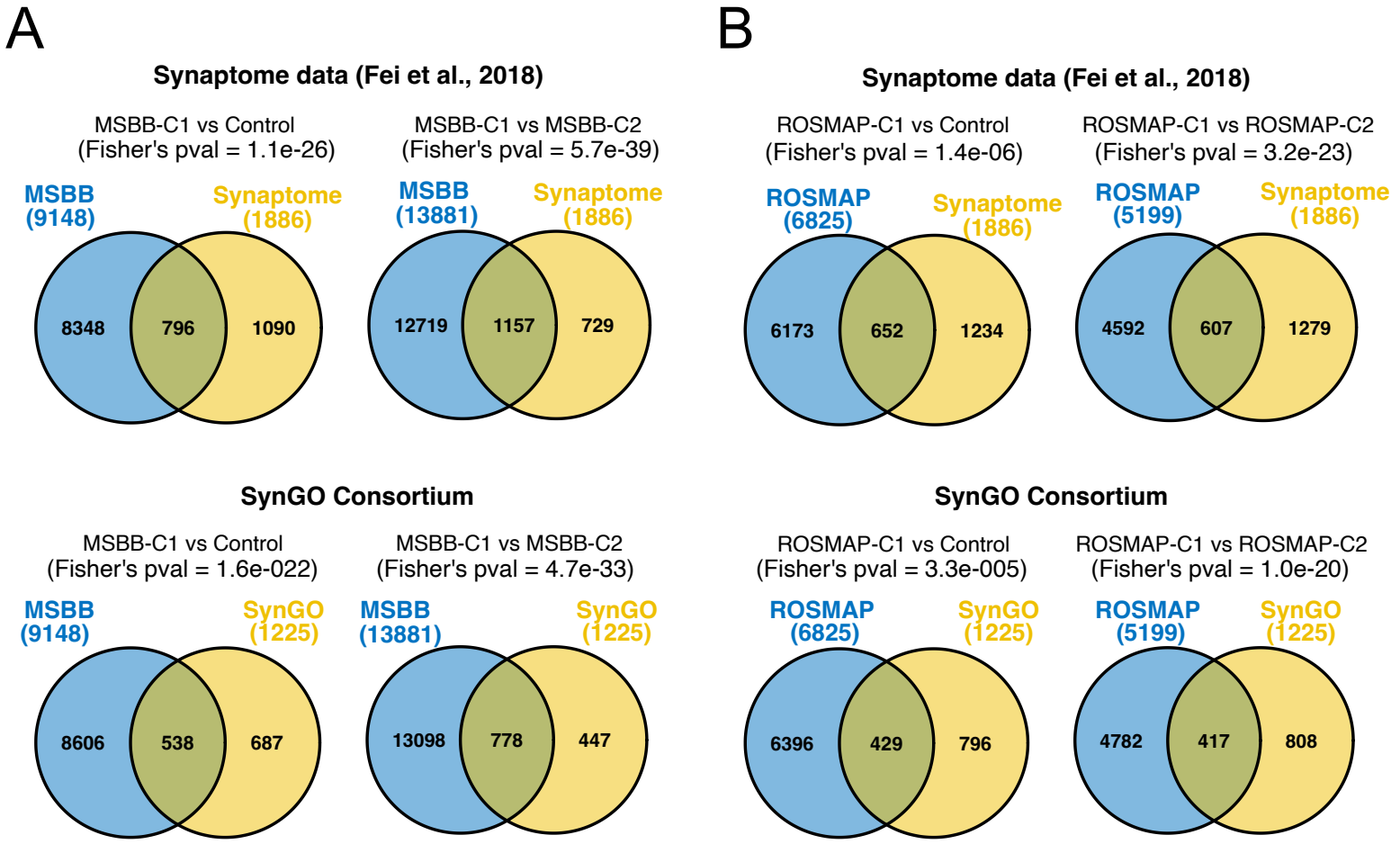
Supplementary Figure 5



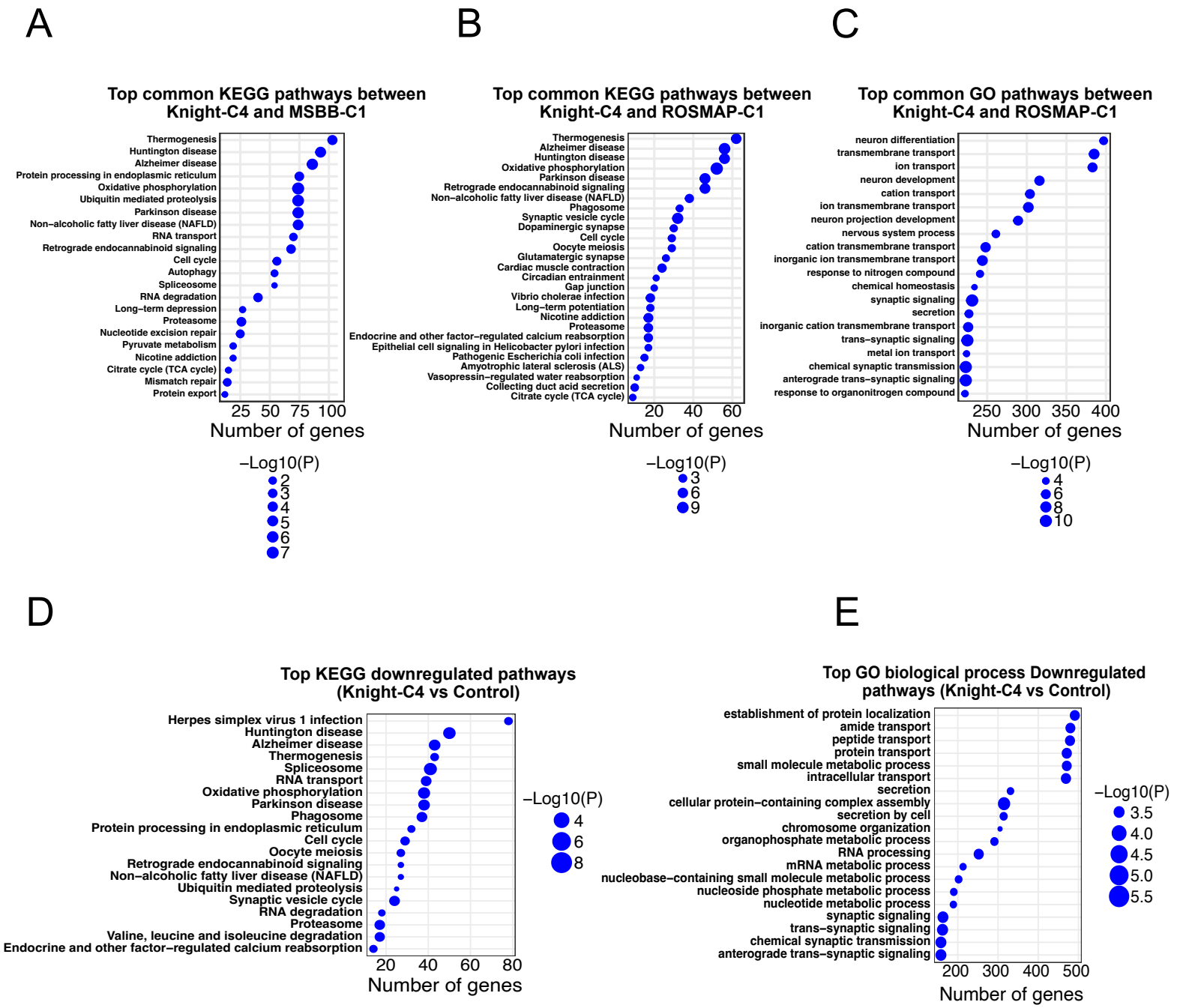
Supplementary Figure 6



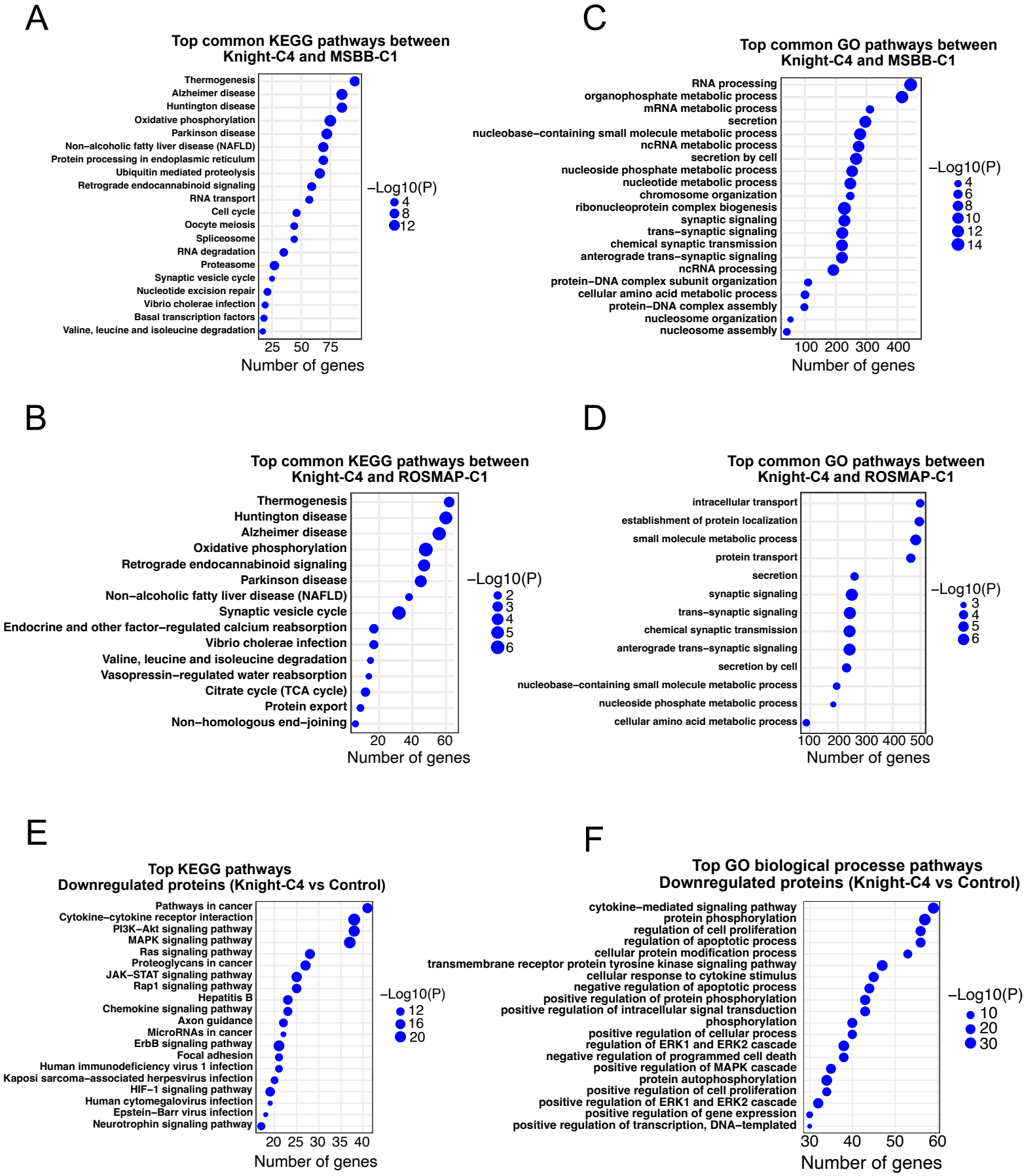
Supplementary Figure 7



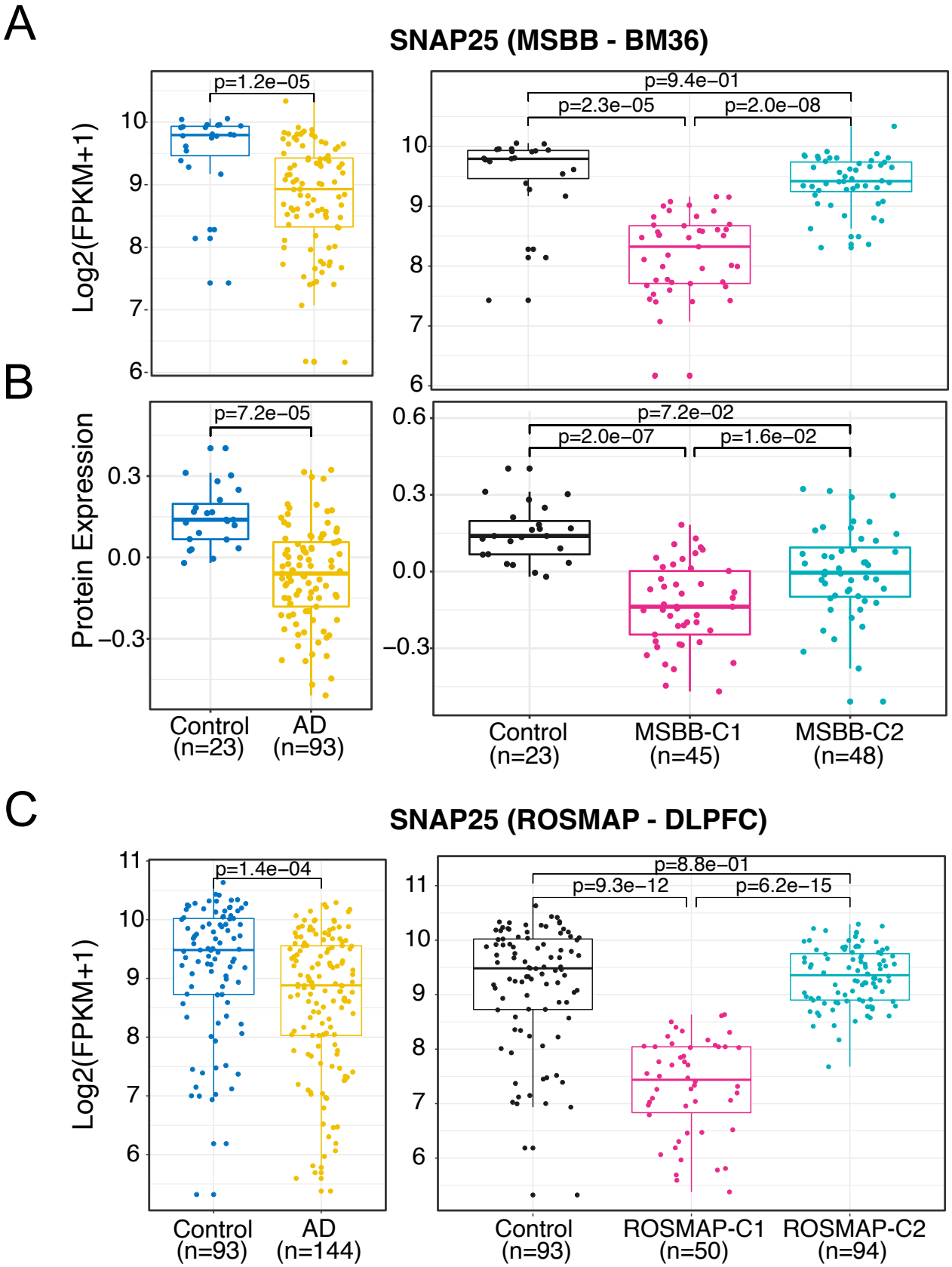
Supplementary Figure 8



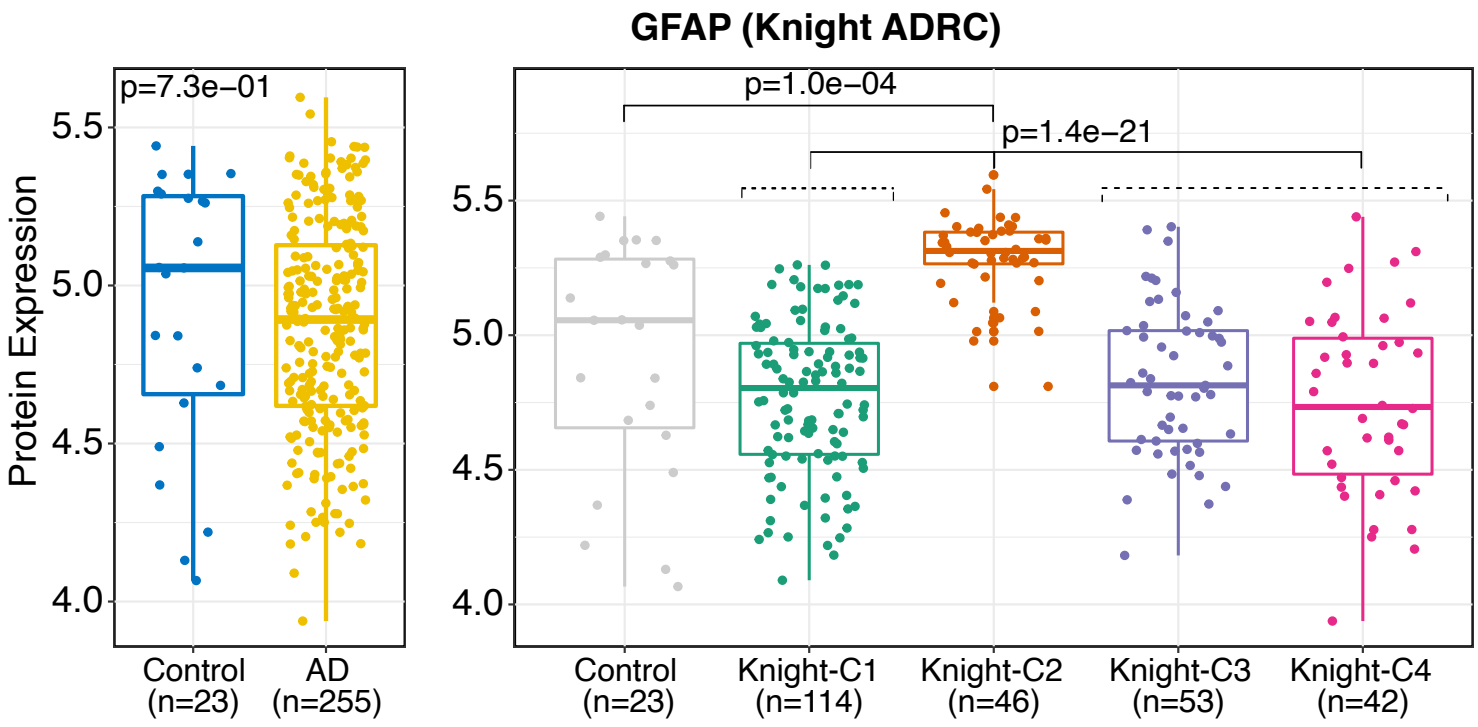
Supplementary Figure 9



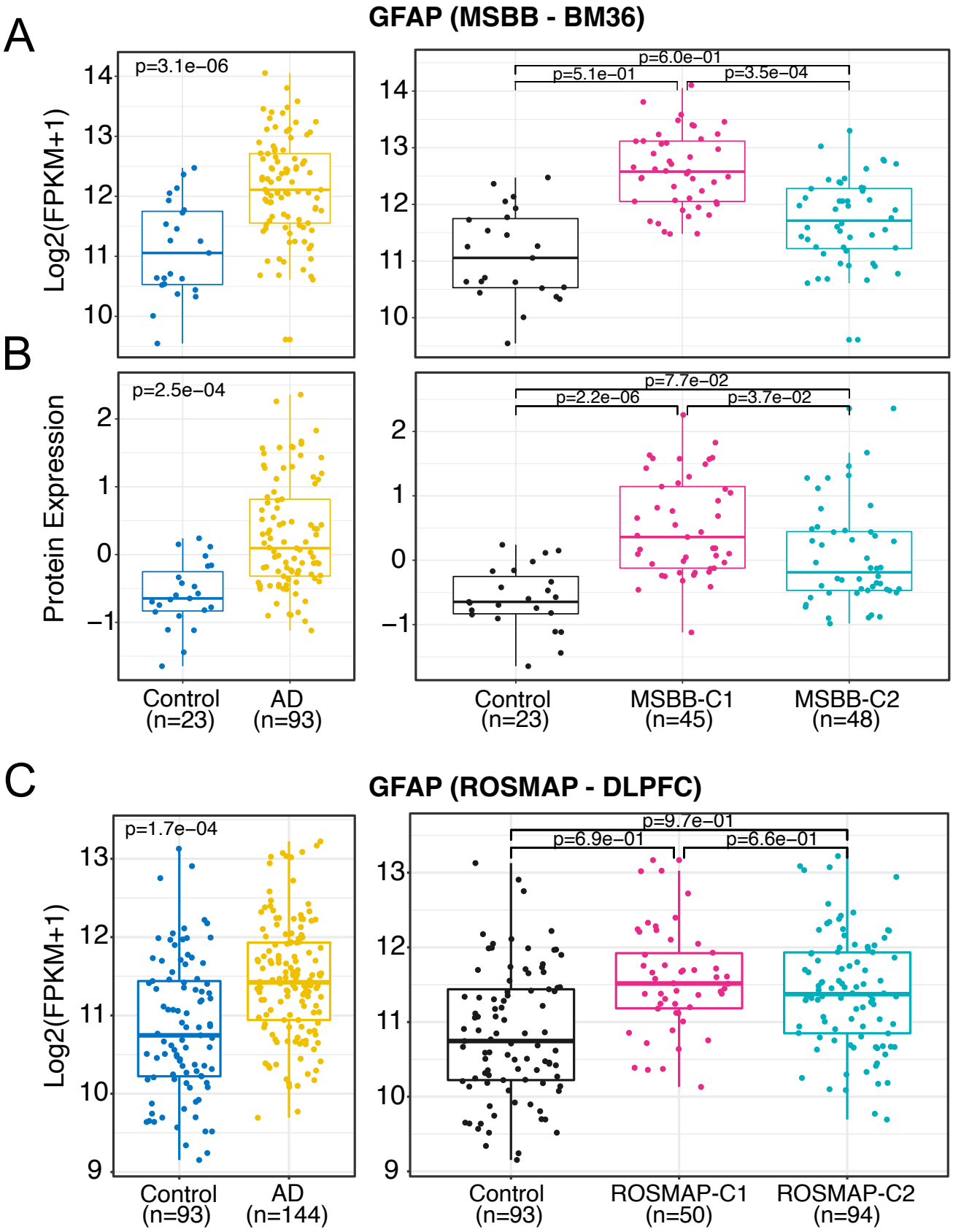
Supplementary Figure 10



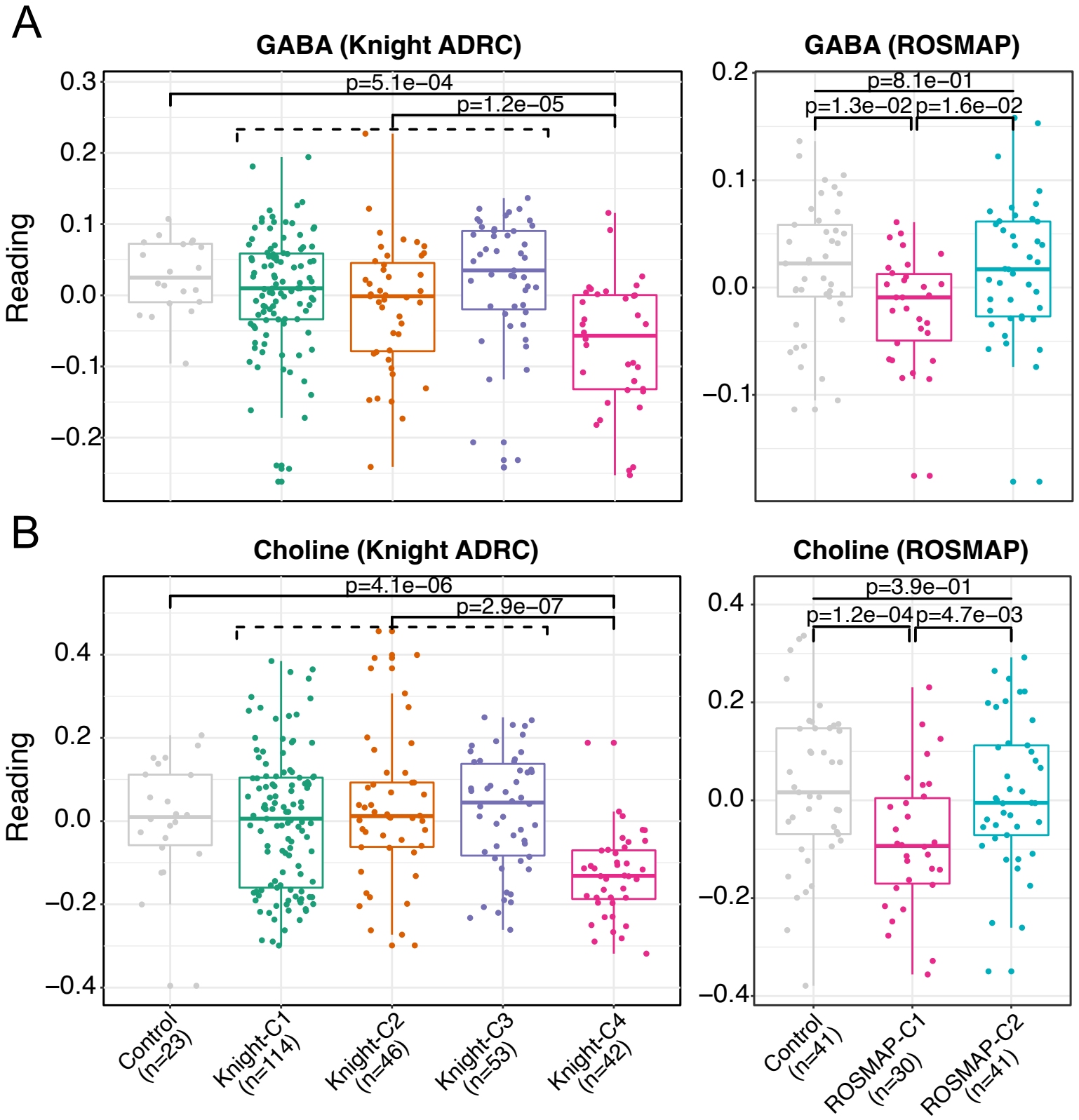
Supplementary Figure 11



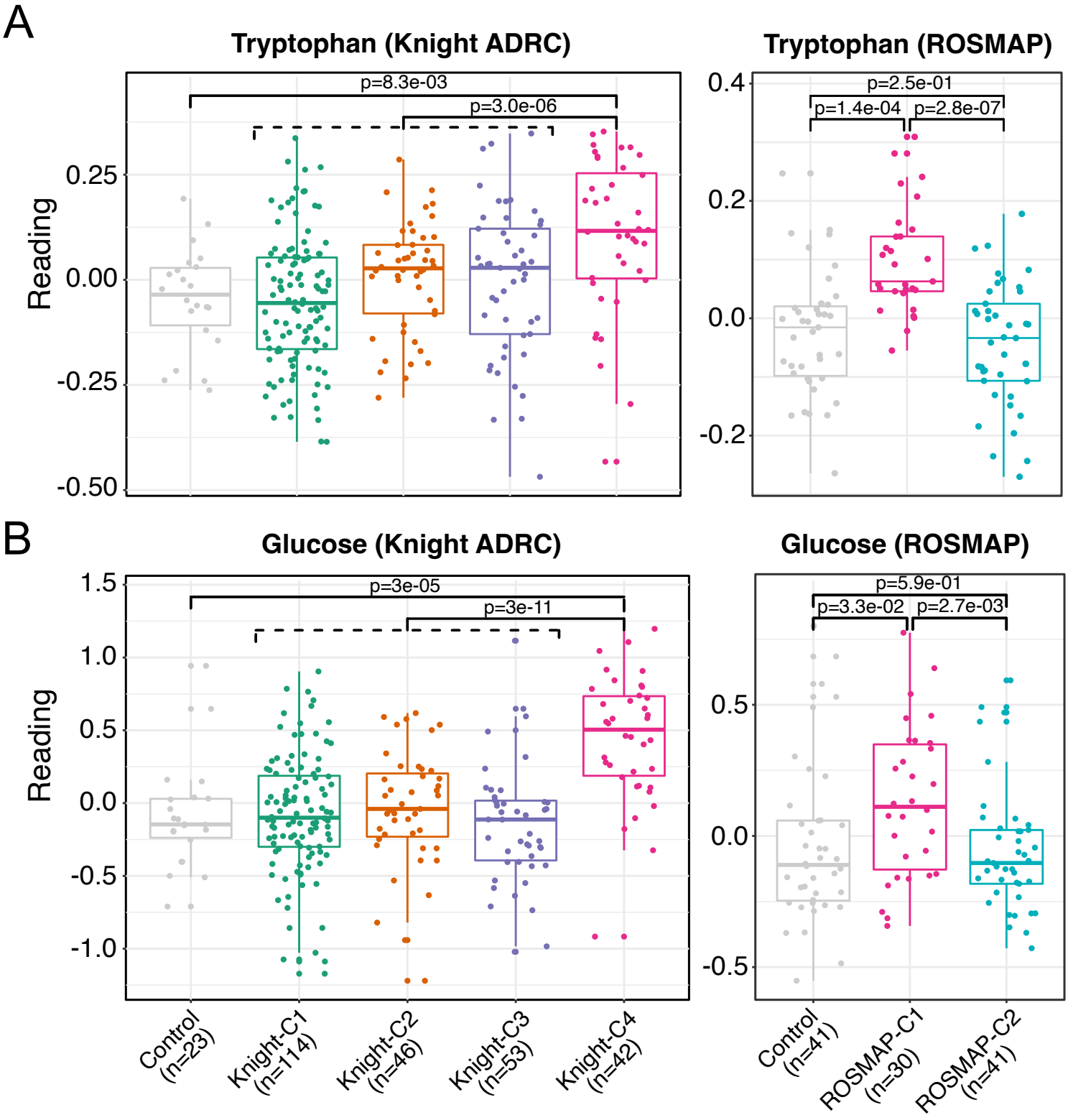
Supplementary Figure 12



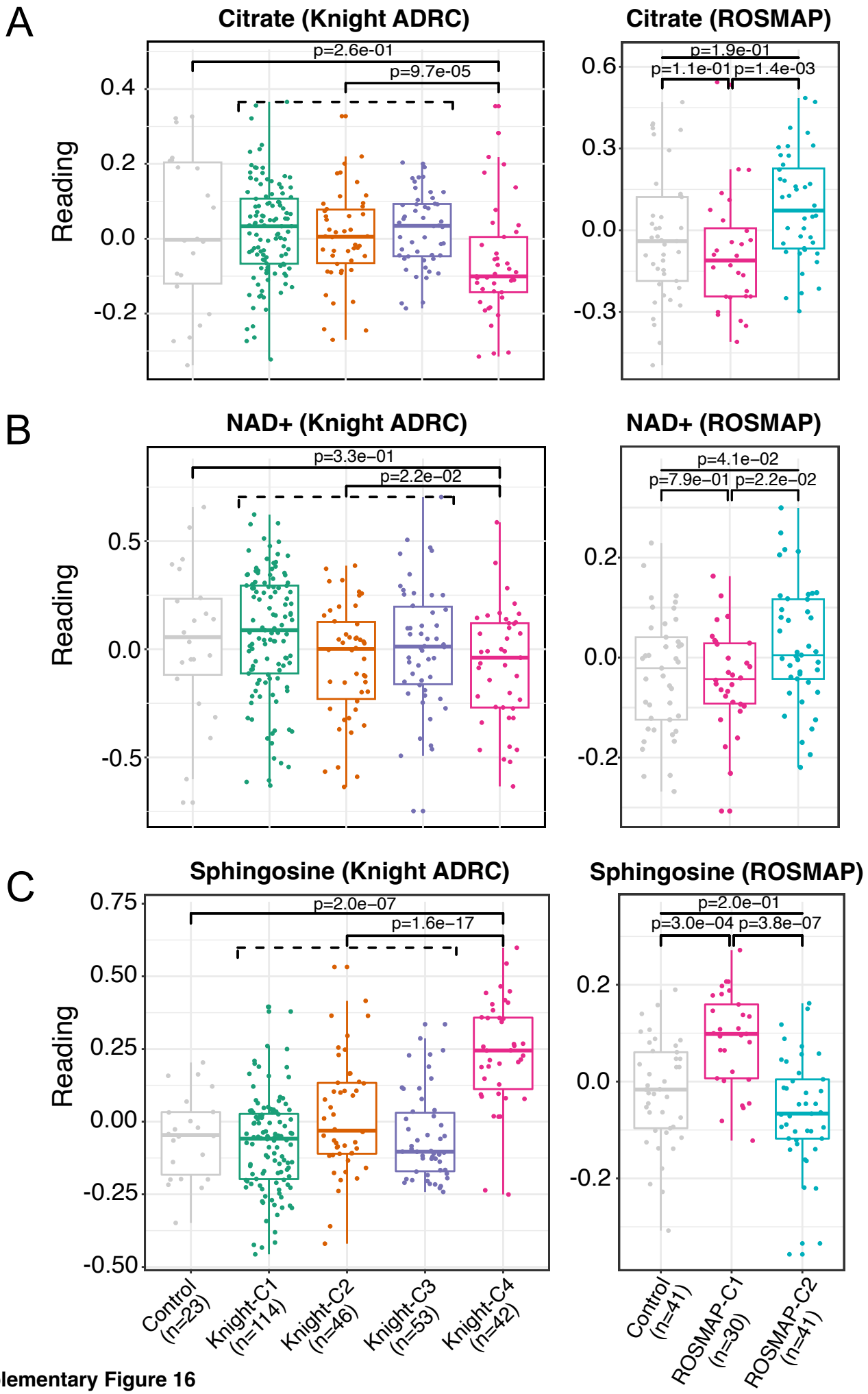
Supplementary Figure 13



Supplementary Figure 14

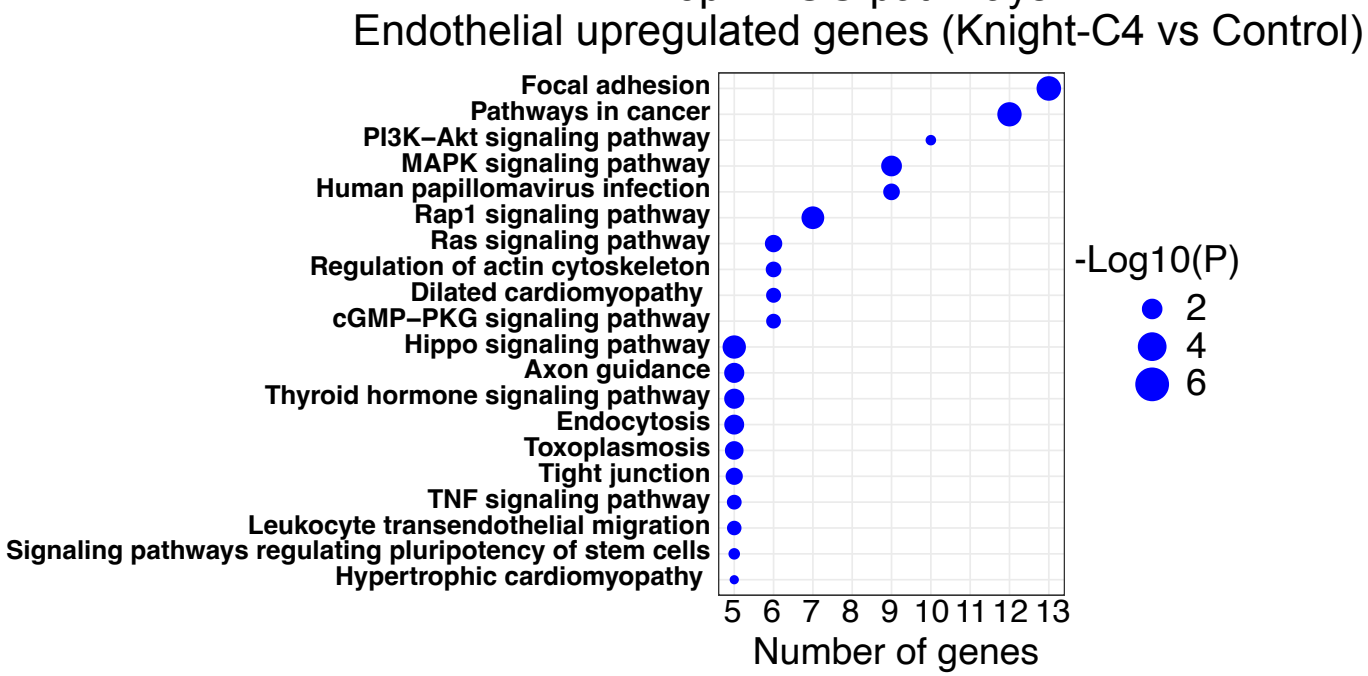


Supplementary Figure 15

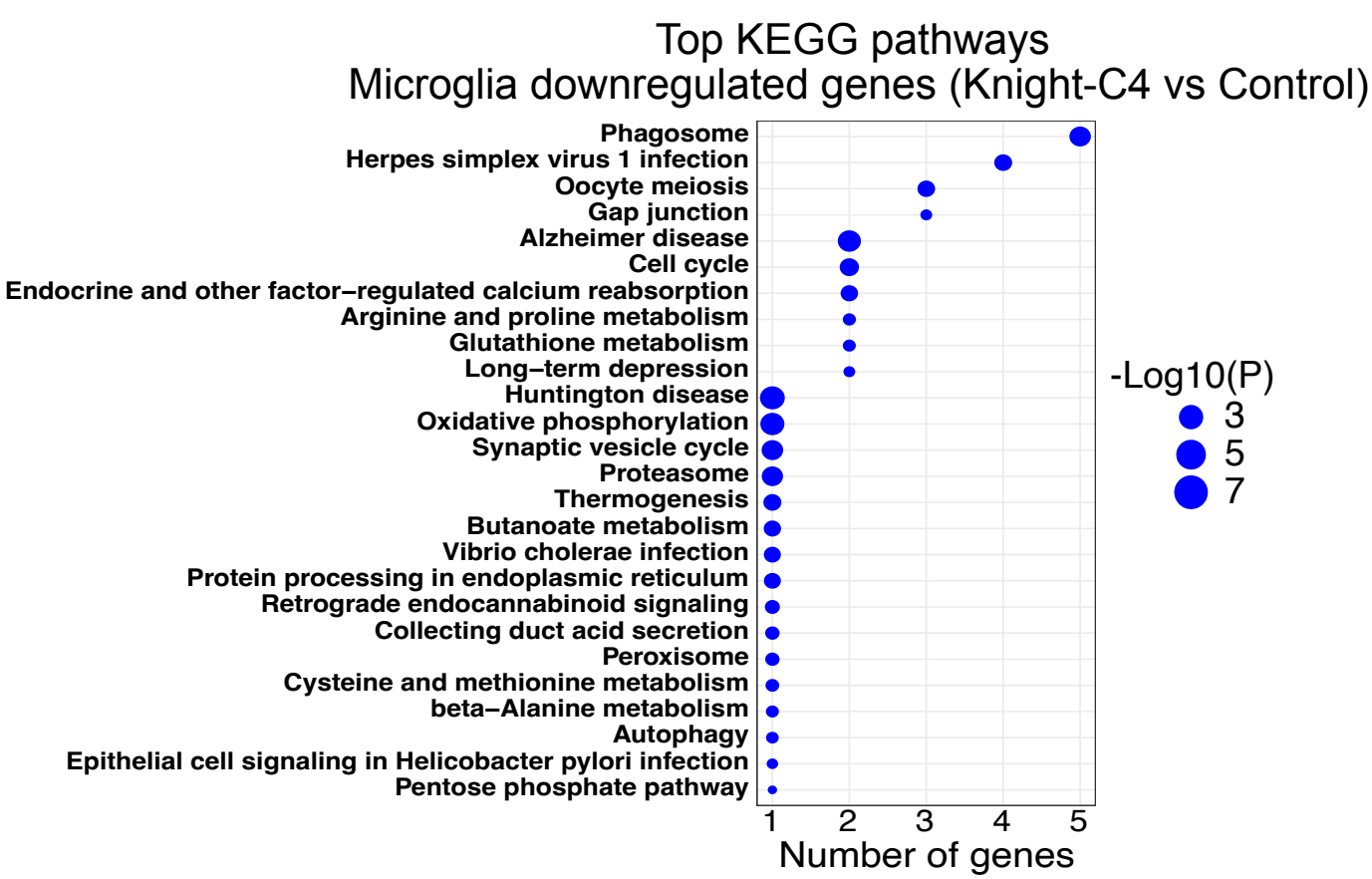


Supplementary Figure 16

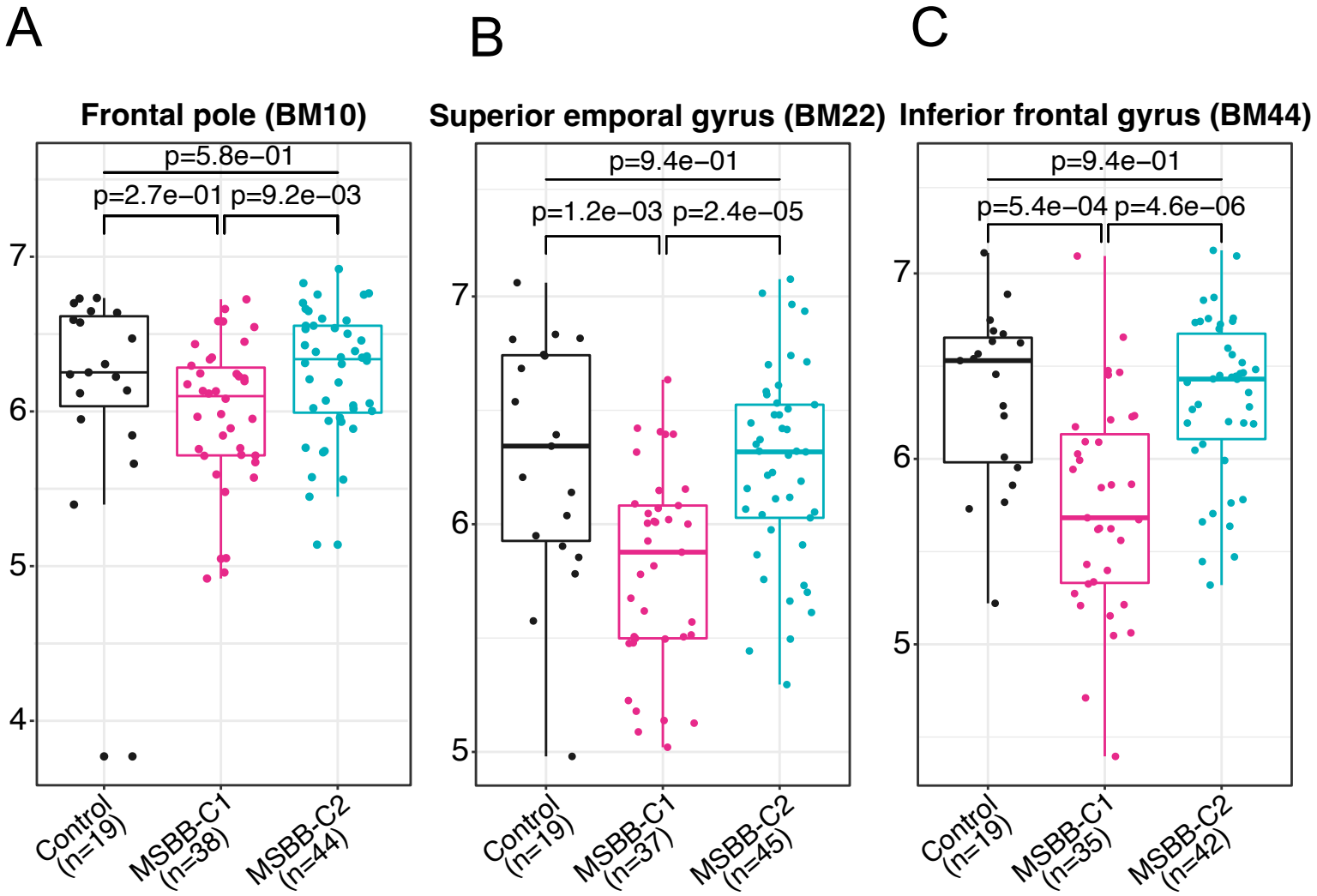
A



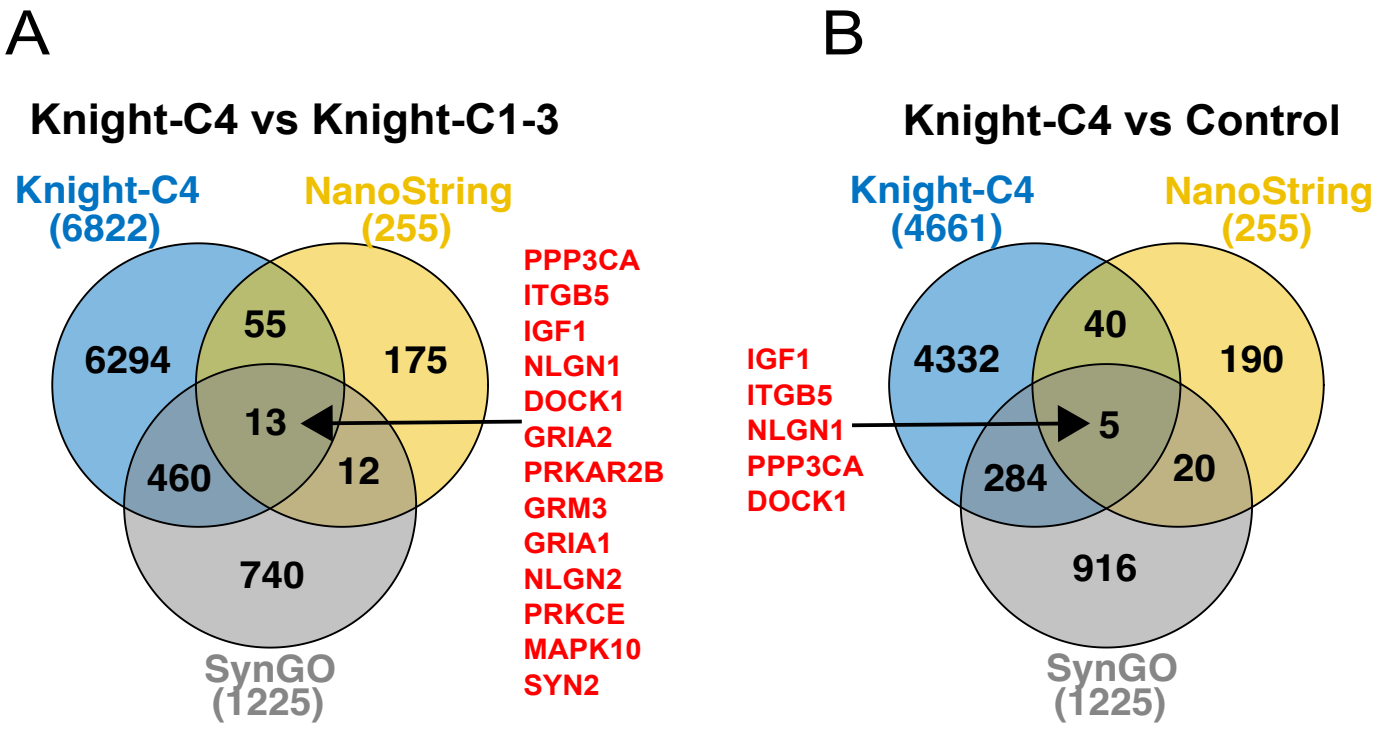
B



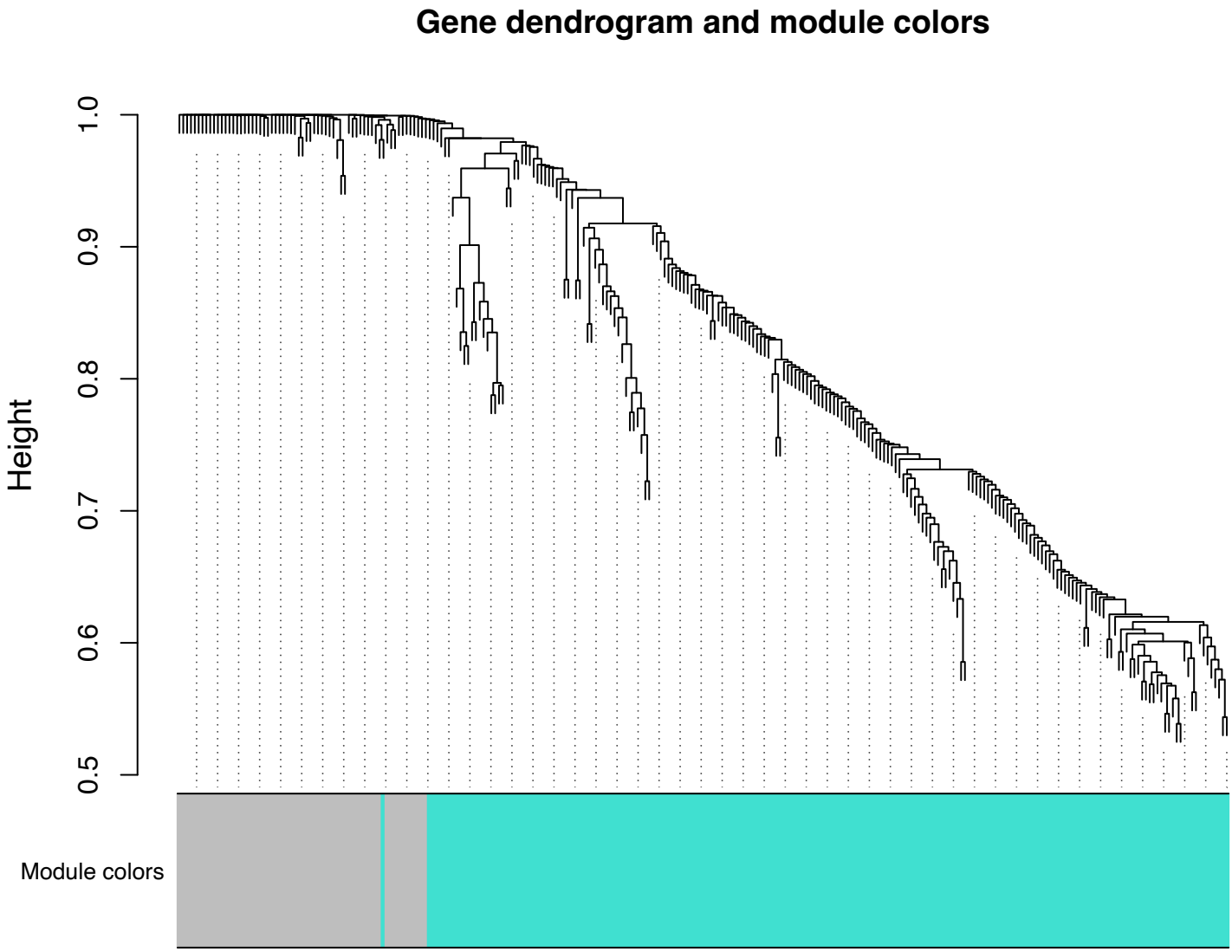
Supplementary Figure 17



Supplementary Figure 18



Supplementary Figure 19



Supplementary Figure 20