Cross-omics integration shows cortex-wide synaptic dysfunction at later stages of Alzheimer disease

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

[1]Department of Psychiatry, Washington University, Saint Louis, St. Louis, MO, USA

[2]NeuroGenomics and Informatics Center, Washington University, St. Louis, MO, USA

[3]The Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington University, St. Louis, MO, USA

[4]Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, NC, USA

[5]Department of Neurology, Washington University, St. Louis, MO, USA

[6]Department of Pathology and Immunology, Washington University, St. Louis, MO, USA

[7]Hope Center for Neurological Disorders, Washington University, St. Louis, MO, USA

[8]Department of Neurology, Harvard Medical School and Beth Israel Deaconess Medical Center, Boston, MA, USA

#To whom the correspondence should be addressed:

Oscar Harari, PhD

Department of Psychiatry

Washington University School of Medicine

4444 Forest Park Ave., Saint Louis, Missouri, 63110

E-mail: harario@wustl.edu

Abstract

Clinical and molecular heterogeneity of Alzheimer disease (AD) is increasingly recognized as a major factor impacting the diagnosis, the design of therapeutic interventions and clinical trials, and therefore possibly delaying the development of effective treatments for AD. We sought to determine whether cross-omics integration reveals molecular profiles of AD with clinical and biological relevance. By leveraging cross-omics approaches, we integrated high-throughput transcriptomic, proteomic, metabolomic, and lipidomic profiles from AD and control cases from multiple cortical regions and cohorts. We identified four distinct molecular profiles in which a specific profile was associated with significantly higher Clinical Dementia Rating (CDR) at death, shorter survival after onset, more severe neurodegeneration and astrogliosis, and decreased

levels of metabolomic profiles. Its molecular signatures show significant dysregulation of synaptic genes indicating neuron/synapse losses and dysfunction at later stages of AD, and present in multiple cortical regions. Among other molecules, we found that the expression of alpha-synuclein (*SNCA*) and *SNAP25* were downregulated in this profile. Overall, our results suggest that AD has a more prominent synaptic loss and dysfunction associated with worse cognition. These novel molecular findings open the possibility for new biomarkers for the molecular staging of AD and potential therapeutic targets to ameliorate cognitive decline and AD progression.

52 **Main**

53 Alzheimer disease (AD) is a neurodegenerative disorder pathologically characterized by amyloid
54 (Aβ) plaques, neurofibrillary tangles (NFTs), neuroinflammation, and synaptic and neuronal loss.
55 AD is the most common cause of dementia among the elderly, currently affecting approximately
56 6.2 million individuals of age 65 and older in the United States and 50 million worldwide[1]. However,
57 AD shows a large genetic, clinical, and pathological heterogeneity. Similar to the variability of
58 patterns and rates at which amyloid deposition, tauopathy, and neurodegeneration are observed
59 among the affected population[2–4], participants with variable degrees of cognitive impairment at
60 the time of death can exhibit very similar AD associated neuropathology. In addition, AD pathology
61 hallmarks are commonly accompanied by variable burdens of one or more other additional
62 neurodegenerative pathologic features, including cerebrovascular lesions, Lewy body pathology,
63 and TDP-43 proteinopathy[3]. Concordant with these observations, growing evidence suggests that
64 molecular profiles also differ substantially within AD[5,6]. This multifaceted heterogeneity suggests
65 that multiple biological pathways are involved in AD – presenting a major challenge for the
66 development of therapeutic interventions and companion diagnostics. However, the
67 heterogeneity of AD has not been fully considered in the design of therapeutic interventions and
68 clinical trials, possibly because of a detailed map of the AD molecular heterogeneity is still
69 missing. Characterizing this heterogeneity on a biological level – by identifying specific changes
70 in synapse/neuronal population, demyelination, glial activation, for instance – may provide critical
71 insights to develop new peripheral biomarkers, companion diagnostics, and therapeutic
72 interventions, tailor the design of clinical trials, and bring closer precision medicine for AD.
73

74 In recent years, multiple clinical and pathological AD subtypes have been identified based on
75 distinct spatiotemporal trajectories of tau pathology, brain atrophy, postmortem brain
76 transcriptomics profiles, and cerebrospinal fluid proteomics. Differences in characteristics
77 associated with these subtypes including, age at onset, sex, APOE ε4 carrier status, cognitive
78 status, disease duration, and cerebrospinal fluid (CSF) biomarker levels, were previously
79 reported[2,3,5–11]. Prior research efforts have leveraged molecular profiles and biomarkers to
80 characterize the AD population into subgroups, and have provided information on AD progression,
81 that in turn could be used in clinical trials. However, the influence that such AD profiles may have
82 on cognitive decline remains poorly understood and under-studied, representing a barrier to using
83 this information effectively in clinical trials of potential disease-modifying therapies.
84

Previous studies have proposed multiple AD subtypes examining clinical, neuropathological, and molecular features[7]. Recently, distinct spatiotemporal tau deposition trajectories have been identified using tau PET imaging[2], which suggest significant differences between sex, APOE e4, and age. In parallel, three AD subtypes associated with multiple dysregulated pathways, such as susceptibility to tau-mediated neurodegeneration, amyloid-β neuroinflammation, synaptic signaling, immune activity, mitochondria organization, and myelination using a system biology approach to study transcriptomics from multiple cortical regions[5]. Despite the substantial effort in identifying molecular subtypes of AD described in this study, broader conclusions are limited by its nature as a single-omics analyses typically captures changes in a single component of the biological cascade.

Comprehensive studies that aggregate multiple modalities of omics (cross-omics) have proven to be a powerful way to identify molecular subtypes for complex human diseases including cancer[12–17]. Omics layers describe the biological system at different biomolecular levels[18], allowing more comprehensive segregation of AD biology into molecular profiles that the partial analyses of any individual modality might miss. Cross-omics approaches can reveal not only how complex biomolecular profiles change with the disease, but also the relationships and correlations between distinct classes of biological molecules. Their potential for studying neurodegeneration has been previously described[19–28]. However, only recently deep molecular characterization of multiple cortical regions from diverse brain banks and cohorts has only recently been generated[5,6], which now enables a more detailed and comprehensive examination of cross-omics data in AD.

In the current study, we leveraged machine learning, digital deconvolution, and traditional statistical approaches to integrate and analyze multiple high-throughput *omics* from different cortical regions and cohorts. Our cross-omics data integration approach identified distinct molecular profiles of AD associated with cognitive function and clinical attributes. Our results highlight that molecular heterogeneity is a common characteristic in AD brains and by virtue of molecular stratification, cross-omics approaches reveal multiple molecular pathways associated with AD progression. Overall, these results provide a foundation for implementing precision medicine approaches for AD.

Results

Study design

118 We leveraged high-throughput transcriptomic, proteomic, metabolomic and lipidomic data from
119 the parietal cortex of participants with neuropath-confirmed AD from the Knight Alzheimer Disease
120 Research Center (Knight ADRC) for our discovery cohort (see Methods, **Fig. 1** and **Table 1**).
121 Omics data were generated using next-generation sequencing, SomaScan, and Metabolon
122 Precision Metabolomics platforms for transcriptomic (60,0754 transcripts)[29–31], proteomic (1,092
123 proteins)[32], and metabolomic (627 metabolites)[33] respectively (see Methods). Preprocessing,
124 quality control (QC), and feature selection were performed on each omics data type, and only
125 samples that passed the QC criteria in the three omics modalities were retained, resulting in 278
126 samples (255 AD and 23 control). We employed a Bayesian integrative clustering method named
127 iClusterBayes[34] to integrate AD multi-omics data (see Methods) and identify molecular profiles of
128 AD. Control cases were later integrated for comparative analyses.
129
130 To investigate the biological relevance of the AD molecular profiles, we performed multiple
131 association analyses including association with clinical and neuropathological attributes (e.g. sex,
132 age of death, age at onset, post-mortem interval, CDR, Braak, etc.), survival, and differential
133 expression analyses to characterize AD molecular profiles. We leveraged digital deconvolutions
134 approaches to infer the cellular population structure using bulk RNA-seq data[31], and tested for
135 association of these profiles with cell-type abundances. We have also integrated single-nuclei
136 RNA-seq data (snRNA-seq) from the parietal cortex from 67 Knight ADRC participants to
137 determine the specific cell-type effects.
138
139 For replication, we leveraged transcriptomics (RNA-seq) and proteomics (tandem mass tag;TMT)
140 from 116 brains (93 AD and 23 control, **Table 1**) extracted from the parahippocampal gyrus
141 (BM36) from The Mount Sinai Brain Bank (MSBB) study[35] as well as transcriptomics data from
142 237 brains (144 AD and 93 control, **Table 1**) from the dorsolateral prefrontal cortex (DLPFC) from
143 The Religious Orders Study and Memory and Aging Project (ROSMAP) cohort[36] accessed
144 through the AMP-AD Synapse portal. Preprocessing, QC, and feature selection were performed
145 similar to the discovery datasets.
146
147 **Identification of distinct molecular profiles of Alzheimer disease**
148 We sought to investigate how multiple *omic* modalities capture optimally the molecular
149 heterogeneity in AD, without any additional conditioning on additional neuropathological or clinical
150 data. We used Bayesian information criteria (BIC) to evaluate clustering solutions between 2 and
151 10 clusters and identified four clusters (Knight-Clusters, **Fig. 2a**) which maximized compactness

(intra-cluster similarity). Among the four clusters, we observed that the molecular profiles of Cluster4 (named Knight-C4 hereafter, n=42) exhibited more pronounced dysregulation for many molecules compared to AD participants from other clusters. For instance, the transcriptomic profiles of the top 75% quantile genes show significant dysregulation signatures associated with Knight-C4 (**Fig. 2b**).

Cross-omics analyses identified a molecular profile of AD brains associated with worse cognitive function and molecular attributes

We initially evaluated whether the alternative molecular profiles were associated with genetic risk factors, neuropathological status, or clinical and demographic variables. We found that AD cases in Knight-C4 showed more severe cognitive decline, measured by significantly higher Clinical Dementia Rating (CDR) scores at death ($p=5.24x10^{-3}$; **Fig. 2c**), shorter survival time following disease onset (8 months shorter, **Fig. 2d**, $p=4.2\ x10^{-3}$, HR=1.6) and younger age at death (**Fig. 2d**, $p=4.2\ x10^{-3}$, HR=1.7) compared to donors in Knight-C1-3. However, the Knight-C4 cluster was not associated with either AD polygenic risk score (**Supplementary Fig. 2**) or APOE ε4 allele carrier status (**Supplementary Fig. 3a**). We did not detect any significant association between Knight-C4 and either Braak neurofibrillary tangle stage or amyloid-β plaque score (**Supplementary Fig. 3b,c**). It is worth mentioning that these annotations are not complete for all the brains, restraining the statistical power.

Next, we leveraged digital deconvolution methods to infer the cellular population structure from bulk RNA-seq[31] (see Methods). We identified that the parietal cortices from donors in Knight-C4 show significantly higher astrocyte and lower neuronal proportions (**Fig. 2e and Supplementary Table 1**). We then evaluated whether a metabolomic profile that we previously reported associated with AD duration[33] would reveal additional molecular differences among the clusters. Lower values of this profile (or *eigenmetabolite*) were associated with a longer disease duration driven by earlier onset[33]. We found that brains in Knight-C2 and Knight-C4 exhibited significantly lower eigenmetabolite scores (**Fig. 2f**) using ordinal logistic regression ($p=6.0x10^{-03}$).

Altogether, these results show the feasibility of our approach in extracting novel insights from AD subpopulations by identifying a late-stage AD profiles associated with more neuronal loss and astrogliosis that caused more alterations in different molecular processes.

The AD molecular profile associated with worse cognition is represented consistently across multiple cortical regions

We sought to investigate whether the molecular profile associated with worse cognitive outcomes was specific to the parietal cortex or also present in additional cortical regions affected at distinct stages of AD. Our analyses of the transcriptomics and proteomics from AD parahippocampal gyrus (BM36) from the MSBB study[35] (**Table 1**) identified two distinct molecular profiles (**Fig. 3a**). We identified that brains in Cluster1 (MSBB-C1) recapitulate the transcriptomic signatures we had previously identified in Knight-C4. We also ascertained the extent of overlap between the molecules dysregulated in Knight-C4 (**Supplementary Table 2**) and MSBB-C1 (**Supplementary Table 3**) and determined a coincident profile of dysregulated genes (**Supplementary Fig. 4a**, $p < 2.2 \times 10^{-16}$ hypergeometric test, **Supplementary Table 4**). Furthermore, these donors in MSBB-C1 also exhibited more severe dementia at expiration, shown by a significantly higher CDR scores ($p=1.5\times10^{-3}$, **Fig. 3b**), and estimated higher and lower proportions of astrocytes and neurons, respectively ($p=5.5\times10^{-07}$ and 3.2×10^{-09} respectively; **Fig. 3c**, **Supplementary Table 1**) compared to other AD cases in MSBB. Furthermore, MSBB-C1 was associated with higher Braak tau scores ($p=3.9\times10^{-03}$, **Supplementary Fig. 5a**). Although not significant ($p=8.0\times10^{-02}$), these donors show a trend suggesting earlier age of death (**Supplementary Fig. 5b**). In contrast, we did not identify a significant overlap of dysregulated proteins between regions ($p > 0.05$, **Supplementary Fig. 4b**).

Furthermore, we leveraged RNA-seq generated from MSBB donors in additional cortical regions and identified that the distinguishing molecular profile of MSBB-C1 is consistently present in multiple cortical regions including frontal cortex (BM10), superior temporal gyrus (BM22), and inferior frontal gyrus (BM44) (**Supplementary Fig. 5c**).

Similarly, our analyses of the RNA-seq from the dorsolateral prefrontal cortex (DLPFC) from the ROSMAP cohort[36] identified subset of AD donors (**Fig. 3d**). Brains in Cluster1 (ROSMAP-C1) recapitulated the transcriptomic profiles of Knight-C4 ($p < 2.2 \times10^{-16}$ hypergeometric test; **Supplementary Fig. 4c and Supplementary Table 4 and 5**). Donors from ROSMAP-C1 showed evidence of greater cognitive impairment and increased neuropathological AD scores, indicated by significantly lower Mini-Mental State Examination scores (MMSE; $p=4.4\times10^{-02}$; **Fig. 3e**), higher tau tangle densities ($p=1.3\times10^{-02}$) and neurofibrillary tangle burdens (PHF tau tangles, $p=2.1\times10^{-02}$; **Supplementary Fig. 6a,b**). Although not significant ($p > 0.05$), ROSMAP-C1 showed a trend to higher plaque densities—Global burden of AD pathology (**Supplementary Fig. 6c**, see Methods and Bennett *et al.*,2012[36] for details on how these neuropathological variables are identified). Cell proportions from deconvolution analysis (**Supplementary Table 1**) also showed

221 that ROSMAP-C1 has significantly higher astrocyte ($p=3.1x10^{-08}$) and lower neuron ($p=9.7x10^{-14}$)
222 proportions (**Fig. 3f**).
223
224 Altogether, these results indicate that cellular and molecular profiles of AD cases with worse
225 cognitive status are similar across multiple cortical regions, usually affected at different stages of
226 AD pathology.
227
228 **The parietal cortex of subjects with worse cognitive function exhibited a remarkable**
229 **molecular dysregulation**
230 Differential expression (DE) analyses (see Methods, **Supplementary Tables 2,6,7**) showed that
231 Knight-C4 has a ~10-fold increase in the number of significantly differentially expressed genes
232 (DEGs) compared to other AD participants (**Table 2** and **Supplementary Fig. 7a**); furthermore,
233 90% of the 4,661 DEGs were unique to participants in the Knight-C4. (**Table 2** and **Fig. 4a-left**
234 **panel**). Similarly, we observed an increased number of proteins and metabolites (**Table 2** and
235 **Supplementary Fig. 7b,c**) with significant abundance levels in Knight-C4. However, we
236 determined that Knight-C2 has the largest number of dysregulated proteins. We identified 24
237 (34%), 232 (51%), and 250 (52%) significant proteins unique to Knight-C1, 2, and 4, respectively
238 (**Fig. 4b-left panel** and **Table 2**). We did not detect significant proteins dysregulated uniquely in
239 Knight-C3.
240 Knight-C4 was associated with cluster-specific significantly different metabolic reading levels (90
241 hits; **Table 2** and **Supplementary Fig. 7c**) while only one hit was detected in Knight-C2. Out of
242 the 90 hits, 80 (90%) significant metabolites were also significantly different between Knight-C4
243 and other AD cases. These results suggest that cross-omics analyses capture molecular
244 heterogeneity among AD cases that are otherwise missed in non-clustered *omics* specific
245 approaches.
246
247 To quantify the benefits provided by this cross-omics study that clusters AD brain donors using
248 multi-omics profiles, we performed DE analyses comparing all AD cases combined (un-clustered)
249 to control brains. Then, we determined the percentage of significant features detected in each
250 cluster that are not captured by un-clustered analyses (**Table 2**, **Fig. 4a-c-right panel**). We
251 observed that Knight-C4 exhibited the highest percentage of missed hits with 92%, 98%, and
252 100% for transcriptomics, proteomics, and metabolomics respectively (**Fig. 4a-c-right panel**).
253 Notably, a high percentage of hits was also observed with Knight-C2 in proteomics (98%). On the
254 other hand, un-clustered analyses identified a small percentage of significant features that are
255 not captured by clustered analyses. For instance, only 5% and 0.5% of un-clustered significant

hits for transcriptomics and proteomics, respectively were not captured by Knight-C4. No un-clustered significant metabolites were missed by Knight-C4.

To further characterize distinct profiles of AD brains, we performed DE analyses to compare AD brains in each cluster to brains in other clusters (**Supplementary Tables 2,6,7**). Knight-C4 showed the highest percentage of significant hits compared to other AD brains (**Fig. 4d** and **Table 3**). This significant difference was observed in transcriptomics with 6,822 significant hits compared to 678, 22, and 1 for Knight-C1-3 (**Fig. 4d-outer track** and **Table** 3). Similarly, this difference was observed in metabolites with 228 significant hits in Knight-C4 compared to 105, 0, and 0 for Knight-C1-3 respectively (**Fig. 4d-inner track** and **Table 3**). No significant differences in the number of hits were observed between the four clusters in proteomics with all clusters showed relatively similar detections (**Fig. 4d-middle track** and **Table 3**). We also verified that none of the omics dominate this solution, each of them capturing molecular heterogeneity among AD brains that are not necessarily detected by other *omics* modalities.

In summary, these results show that molecular profiles of AD are heterogeneous, and cross-omics integration approach can better identify the substantial differences within AD brains.

The AD molecular profile associated with worse cognitive function is significantly enriched in dysregulated synaptic genes and is associated with AD-related pathways

We performed pathway analyses to characterize the brains with worse cognitive function. We found that the hits in Knight-C4 are significantly enriched in dysregulated synaptic genes (**Fig. 4e,f**, **Supplementary Table 8**) using the annotation from The SynGO consortium[37] ($p=1.0x10^{-15}$ and $p=4.8x10^{-45}$ for Knight-C4 vs Control and Knight-C4 vs Knight-C1-3 sets respectively) and curated list from Fei *et al.,* 2018[38] ($p=2.3x10^{-07}$ and $p=1.1x10^{-21}$ for Knight-C4 vs Control and Knight-C4 vs Knight-C1-3 sets respectively). Coincidently, molecular data from the MSBB and ROSMAP studies replicated these results, as MSBB-C1 and ROSMAP-C1 were significantly enriched in synaptic genes (**Supplementary Fig. 8a,b and Supplementary Table 8**). These results suggest that synaptic and neuronal loss occurs at later stages of AD.

In addition, the pathway analyses of Knight-C4 identified multiple AD-related pathways (**Fig. 4g,h**). KEGG pathways identified intracellular trafficking and signaling pathways (**Fig. 4g**) including Endocytosis ($p=6.5x10^{-04}$), Phagosome ($p=1.2x10^{-06}$), and mTOR signaling pathway ($p=3.3x10^{-02}$), all of which were shown to play major roles in the pathological processes of AD[39–41]. KEGG pathways also identified lysosomes ($p=4.7x10^{-02}$, **Fig. 4g**), a critical component within

291 cells. Growing evidence has shown that the dysfunction of lysosomes plays a major role in
292 pathogenesis of multiple neurodegenerative diseases including the progression of AD[42–46].
293 Another significant example associated with Knight-C4 was Autophagy ($p=1.6x10^{-02}$, **Fig. 4g**), a
294 pathway associated with several neurodegenerative disorders including AD[47–50]. For instance, the
295 dysfunction of autophagy was shown to be linked to the accumulation of misfolded protein
296 aggregates[51]. KEGG pathways results also show that the downregulation of Knight-C4, compared
297 to other AD cases, was associated with pathways related to multiple neurodegenerative diseases
298 (**Fig. 4g**), including AD ($p=2.5x10^{-14}$), Huntington disease ($p=3.3x10^{-14}$), and PD ($p=2.1x10^{-12}$)
299 suggesting that neuronal/synaptic loss and gliosis are common characteristics of most
300 neurodegenerative diseases at late-stages. These analyses also identified synaptic pathways,
301 including dopaminergic synapse ($p=7.0x10^{-03}$) and glutamatergic synapse ($p=1.0x10^{-02}$) (**Fig. 4g**).
302 Furthermore, GO biological process analyses identified significant neuronal differentiation
303 ($p=1.8x10^{-02}$) and development ($p=5.7x10^{-03}$) associated with Knight-C4 (**Fig. 4h**)

304 Compared to the control group, Knight-C4 showed two interesting KEGG pathways Herpes
305 simplex virus 1 infection ($p=4.4x10^{-03}$), a risk factor for AD[52–55] that lead to neuronal damage and
306 induced gliosis[56,57], in addition to synaptic vesicle cycle ($p=4.6x10^{-05}$) shown to play a key role in
307 the pathobiology of AD[58] (**Supplementary Fig. 9a**). GO biological process analyses identified
308 several significant synaptic signaling and transmission pathways (**Supplementary Fig. 9b**).
309 Furthermore, pathway analyses using significant proteins identified AD-related KEGG pathways
310 including PI3K−Akt signaling ($p=6.5x10^{-19}$), Rap1 signaling ($p=5.8x10^{-14}$), Axon guidance
311 ($p=2.1x10^{-12}$), MAPK signaling ($p=1.4x10^{-20}$), and Ras signaling ($p=1.8x10^{-15}$) pathways
312 associated with the downregulation of Knight-C4 (**Supplementary Fig. 10a**). Top GO pathways
313 were related to cell/gene regulatory pathways (**Supplementary Fig. 10b**).
314
315 Altogether, these results show that synapse dysfunction is more pronounced in participants with
316 worse cognitive function (late-stage AD).
317
318 **Molecular heterogeneity among AD clusters established neuropathological biomarkers**
319 Our DE analyses identified several known neurodegenerative disease genes dysregulated in
320 Knight-C4 including *APP, MAPT, CLU, SNAP25, GFAP, SNCA, LRRK2, TARDBP, CHI3L1,*
321 *MMP9,* and *C9ORF72* (**Fig. 5a**). The transcriptomics profiles of Synaptosome Associated Protein
322 25 (*SNAP25*) were significantly downregulated in Knight-C4 (**Fig. 5b**) as has been previously
323 shown in AD[59], supporting the likelihood of impaired synaptic function in neurons[59]. *SNAP25* has

also been shown as a candidate biomarker in the cerebral spinal fluid (CSF) for synapse degeneration in AD patients[60]. *SNAP25* was significantly DE between Knight-C4 and control ($p=5.9x10^{-3}$, **Fig. 5b**) and between Knight-C4 and other AD cases ($p=4.2x10^{-08}$, **Fig. 5b**). However, *SNAP25* was not significantly DE ($p > 0.05$, **Fig. 5b**) in non-clustered AD cases compared to the control, which may suggest that synaptic dysfunction in the parietal cortex is not substantially affected until later stages of AD. The decreased expression of *SNAP25* in Knight-C4 was replicated in MSBB-C1 and ROSMAP-C1 cohorts (**Supplementary Fig. 11**).

We inspected the results for Glial Fibrillary Acidic Protein (*GFAP*), an emerging biomarker in brain disorders[61] usually employed to identify astrocytic reactivity state, a process well documented in AD and neurodegeneration. *GFAP* was overexpressed uniquely in Knight-C4 compared to Knight-C1-3 (**Fig. 5a,c**, $p=1.7x10^{-2}$). In spite of the significant astrogliosis of Knight-C4, *GFAP* was not significantly DE in Knight-C4 compared to controls, nor was it DE in all AD cases compared to the control ($p > 0.05$, **Fig. 5c**). To query the possibility that collinearity between case-control status and cellular population structure introduced spurious results, we repeated these analyses using a simplified model that did not correct for cellular population structure but could not obtain significant association. We obtained similar results when we analyzed *AQP4*, another marker gene of astrocyte reactivity state, as we only detected significant differences in the levels of AQP4 between Knight-C4 and other cases ($p=1.0x10^{-02}$). In addition, we did not identify significant differences in *GFAP* proteomic levels ($p > 0.05$, **Supplementary Fig. 12**), however, proteomic levels of *GFAP* were significantly upregulated in Knight-C2 (**Supplementary Fig. 12**) compared to the control and AD brains in Knight-C1,3,4 ($p=1.0x10^{-4}$ and $1.4x10^{-21}$ respectively). The increased transcriptomic *GFAP* levels were also identified in the subset of brains with worse cognitive function from the replication cohorts, MSBB-C1 and ROSMAP-C1 (**Supplementary Fig. 13**). Altogether these results support heterogeneity in the presence of astrocytic reactivity state and its association with decline of cognition.

Our cross-omics integration analyses also identified cluster-specific dysregulation of Clusterin (*CLU*), a gene associated with AD[62,63]. Genome-wide associations studies have shown that *CLU* (also known as APOJ) is the third most associated late-onset AD (LOAD) risk gene after *APOE* and *BIN1*[63–66]. Proteomic levels of *CLU* were significantly downregulated in Knight-C2 compared to the control ($p=6.4x10^{-3}$) and AD cases in Knight-C1,3,4 ($p= 9.8x10^{-7}$; **Fig. 5d,e**). *CLU* was also significantly downregulated in Knight-C4 compared to the control ($p=3.4x10^{-02}$; **Fig. 5e**). To assess the clinical implications of *CLU* dysregulation, we evaluated its levels in the cerebrospinal fluid (CSF) from 553 participants of the Knight ADRC by performing survival analyses for

359 progression using CDR. These analyses indicate that lower CSF *CLU* is associated with an
360 increased risk of dementia progression (HR=0.42; $p=1.0x10^{-07}$, **Fig. 5f**). Nonetheless, we did not
361 observe any significant differences in *CLU* transcriptomic and proteomics levels in the MSBB and
362 ROSMAP cohorts.
363
364 The cross-omics approach identified subset of AD brains with unique abundance levels for AD
365 biomarkers. These results show the capability of cross-omics approaches in detecting AD
366 molecular biomarkers that typical non-clustered approaches sometimes miss when AD cases are
367 analyzed altogether. These results provide support for alternative strategies to reposition
368 traditional biomarkers for precision medicine approaches.
369

370 **Metabolic differential abundance analyses identified biological pathways associated with**
371 **synaptic dysregulation and cognitive function**

372 We next looked at the significant metabolites that uniquely characterize brain samples in Knight-
373 C4. To determine whether significant metabolites replicated in additional cohorts, we accessed
374 metabolomics from the same ROSMAP cohort[36] (**Table 1**, Methods) and performed differential
375 abundance analyses comparing the levels from donors in the previously identified ROSMAP-C1
376 cluster to controls and to other AD cases (**Supplementary Table 5**). We identified 57 significant
377 metabolites (43 increased and 14 decreased) in Knight-C4 compared to controls that replicated
378 in the ROSMAP-C1 (**Fig. 6a**, **Supplementary Table 9**). Similarly, we identified 74 metabolites
379 (70 increased and 4 decreased) in the Knight-C4 compared to other AD cases that also replicated
380 in ROSMAP-C1 (**Fig. 6b**, **Supplementary Table 9**).
381
382 Our pathway analyses identified several pathways significantly related to AD. Among the set of
383 metabolites decreased in AD (**Fig. 6a,b**) gamma-aminobutyrate (GABA) and choline, both
384 involved in major neurotransmitter pathways, were identified. GABA, synthesized in neurons, was
385 significantly decreased in Knight-C4 ($p= 5.1x10^{-04}$) and ROSMAP-C1 ($p=1.3x10^{-02}$) compared to
386 the control (**Supplementary Fig. 14a**). The same observation was found when these profiles
387 were compared to other AD cases ($p=1.2x10^{-05}$ and $p= 1.6x10^{-02}$ for the Knight-C4 and ROSMAP-
388 C1 respectively, **Supplementary Fig. 14a**). These results support that those cases with worse
389 cognitive outcomes are associated with dysregulated neuronal or synaptic pathways. As an
390 inhibitory neurotransmitter, GABA has been shown to play a key role in synchronizing the activity
391 of the human cerebral cortex[67,68]. Similarly, the metabolomics profiles of choline were decreased
392 in these profiles in both Knight-C4 ($p= 4.1x10^{-06}$) and ROSMAP-C1 ($p=1.2x10^{-04}$) compared to the

control (**Supplementary Fig. 14b**). These reading levels of choline also show significantly different abundance in these clusters compared to other AD cases ($p= 2.9x10^{-07}$ and $p=4.7x10^{-3}$ for Knight-C4 and ROSMAP-C1 respectively, **Supplementary Fig. 14b**). Choline was also shown to have a critical role in neurotransmitter function because of its effect on the function of acetylcholine and dopamine[69]. However, we did not find significant differences in the reading levels of acetylcholine and dopamine. We also did not observe significant differences for additional neurotransmitters included in the panel (serotonin and glutamate). This indicates that the dysregulation of neurotransmitters is possibly specific to GABA and choline. We then studied tryptophan, an amino acid essential for metabolites that serve as neurotransmitters and signaling molecules[70]. We found that its levels were increased in the Knight-C4 when compared to the control and other AD cases ($p=8.3x10^{-3}$ and $3.0x10^{-06}$ respectively) (**Supplementary Fig. 15a**). Similarly, we found that brains in ROSMAP-C1 showed increased tryptophan levels compared to the control and other AD brains ($p=1.4x10^{-4}$ and $2.8x10^{-07}$ respectively) (**Supplementary Fig. 15a**)

Furthermore, several increased carbohydrate/sugar metabolites were identified in the set of metabolites increased in Knight-C4 (**Fig. 6a,b**) including glucose, mannitol, sorbitol, ribose, and UDP-glucose, indicating a possibility of slower carbohydrate metabolism. As an example, glucose, a metabolite that has been implicated in the pathogenesis of AD[71,72], was significantly increased in the Knight-C4 and ROSMAP-C1 (**Supplementary Fig. 15b**). Components of glycolysis and the tricarboxylic acid (TCA) cycle were also dysregulated in a manner that suggests decreased glucose metabolism and cellular respiration. Specifically, levels of citrate were decreased in both Knight-C4 ($p= 9.7x10^{-05}$) and ROSMAP-C1 ($p= 1.4x10^{-03}$) compared to other AD cases (**Supplementary Fig. 16a**) indicating possible impaired cellular respiration through the TCA cycle, which produces citrate. NAD+, an important oxidizing cofactor in glucose metabolism and cellular respiration, was also decreased in both Knight-C4 ($p=2.2x10^{-02}$) and ROSMAP-C1 ($p=2.2x10^{-02}$) compared to other AD cases (**Supplementary Fig. 16b**). NAD+ is synthesized from tryptophan, and levels of NAD+ have been shown to decrease with age, potentially contributing to the pathogenesis of AD[73]. This decreased availability of NAD+ could also contribute to impaired glucose metabolism. These observations are in line with previous findings that associate increased brain glucose and slowed glucose metabolism with more severe AD[72,74].

In addition, several increased sphingolipid metabolites were identified (**Fig. 6a,b**). Previous studies showed the implication of sphingolipid metabolism deregulation in AD[75]. For instance,

427 sphingosine, found to have higher levels in AD brains[75], showed significantly increased
428 metabolism levels in both Knight-C4 and ROSMAP-C1 (**Supplementary Fig. 16c**).
429
430 Together, these results further characterize the molecular profiles in this subset of AD brains and
431 show synaptic/neuronal dysregulation association and also possible comorbidity between AD and
432 other diseases such as diabetes[74] illustrated by the increased carbohydrate/sugar metabolites in
433 this subset.
434
435 **Single-cell data suggests that pathways in different cell types are more prominently altered**
436 **in distinct molecular profiles of AD**
437 We studied the cell type specific pathways altered in the AD profiles by leveraging single-nuclei
438 RNA-seq from a subset of the Knight ADRC participants. We selected the top 5% of genes that
439 showed the most specific cell-type expression profile in neurons, astrocytes, oligodendrocytes,
440 oligodendrocyte precursor cells (OPC), and microglia. Then, we used this information to
441 determine if any AD molecular profiles showed a stronger signature of cell-type specific
442 dysregulation. These analyses showed that Knight-C4 was enriched in genes overexpressed in
443 astrocytes and endothelial cells (38% and 23% for all hits respectively; **Fig. 7a-left panel** and
444 **Table 4**). KEGG pathways analyses of astrocytic genes identified astrocyte-related pathways
445 (**Fig. 7b**) including Axon guidance ($p=2.8x10^{-05}$), and Hedgehog signaling pathway ($p=1.6x10^{-06}$).
446 Previous work has shown the role of astrocytes in axon guidance during development and repair[76]
447 and hedgehog signaling pathways involved in neuroinflammation, neuronal cell differentiation,
448 and neuronal death[77]. Many of these pathways were also enriched in overexpressed genes in
449 endothelial cells (**Supplementary Fig. 17a**).
450
451 Knight-C4 was also enriched in genes downregulated in microglial and neuronal genes (**Fig. 7a-**
452 **right panel** and **Table 4**). Neuronal genes showed an association with synaptic/neuronal-related
453 pathways (**Fig. 7c-d**), including Synaptic vesicle cycle ($p=1.0x10^{-06}$), Retrograde
454 Endocannabinoid Signaling ($p=3.6x10^{-03}$) key modulators of synaptic function[78–80], and Gap-
455 junction ($p=2.2x10^{-02}$) formation known to play an essential role in intercellular metabolic
456 communication and transmission across electrical synapses[81]. Microglial downregulated genes
457 implicated AD related pathways including Phagosome ($p=1.0x10^{-06}$), Herpes simplex virus 1
458 infection ($p=2.0x10^{-04}$)—previously associated with AD risk[52–55] leading to neuronal damage and
459 induce gliosis[56,57], Gap junction ($p=2.2x10^{-02}$), and AD ($p=5.4x10^{-08}$) (**Supplementary Fig. 17b**).
460 Further explorations showed that Knight-C4 is significantly enriched in downregulated genes

related to homeostatic microglia, including C-X3-C Motif Chemokine Receptor 1 (*CX3CR1;* p= $5.10x^{-03}$, **Fig. 7e,f**) and purinergic receptor P2Y13 (*P2RY13;* p=$1.2x10^{-2}$, **Fig. 7e,f**).

Additionally, we identified genes dysregulated in specific clusters and cell types, including the TP53 induced glycolysis regulatory phosphatase (*TIGAR*) overexpressed in microglia and upregulated in Knight-C1 (**Fig. 7e,f**). TIGAR is a p53-inducible protein that regulates energy metabolism and oxidative stress. Among the upregulated endothelial-specific genes in Knight-C1, we identified *MYOCD*, *CNN1*, *PLN*, *MYH11*, S*100A4*, *SLC13A4*, *DES*, implicating pathways associated with cell cycle, junction, anion transport, Aβ clearance and actin. In particular, *MYH11*, overexpressed in Knight-C4 (**Fig. 7e,f**), has been reported to be associated with the risk of dementia[82]. Moreover, oligodendrocyte-specific genes dysregulated in Knight-C2 include *STX4*, *ACTN4*, and *FBXO32* relating to pathways associated with vesicle docking, neural tube, junction, apoptosis and actin. Particularly, Syntaxin-4 (*STX4*), upregulated in Knight-C2,4 (**Fig. 7e,f**), is critical for vesicle docking and the Wnt frizzled receptor (*FZD9*) and coreceptor (*LRP5*), which regulate the synaptic vesicle cycle[26]. *STX4* has also been shown as one of the disease risk genes in AD[83].

These results suggest that different pathways mediate distinct molecular profiles of AD in multiple cell types.

Integrative cross-omics analyses identified reduced transcript levels of SNCA across multiple brain regions in AD participants with worse cognitive function

Transcriptomic and proteomic assays showed *SNCA* (alpha-synuclein) was among the top molecules that best discriminate Knight-C4 versus Knight-C1-3 molecular profiles. Our results show that *SNCA* transcriptomic levels were significantly downregulated in Knight-C4 compared to Knight-C1-3 (p= $2.0x10^{-05}$) and control brains (p=$2.2x10^{-02}$; **Fig. 8a**). Interestingly, we did not observe a significant association ($p > 0.05$) in non-clustered AD compared to the control (**Fig. 8a-left panel**). This suggests that *SNCA* levels in the parietal cortex may be altered only in late-stage AD. Alpha-synuclein (αSyn) protein levels were marginally downregulated in brains from Knight-C4 (**Fig. 8b**) compared to the control (p=$5.16x10^{-02}$) and significantly lower in Knight-C2 compared to the control (p=$7.8x10^{-5}$) and Knight-C1,3,4 (p=$6.2x10^{-09}$).

We also identified downregulation of *SNCA/*αSyn in MSBB-C1 (**Fig. 8c,d**) compared to the controls (p=$1.8x10^{-02}$ and p=$1.0x10^{-05}$ for transcriptomics and proteomics respectively). Similarly,

495 *SNCA* levels in MSBB-C1 were significantly lower compared to other AD ($p=1.0x10^{-02}$ and
496 $p=3.4x10^{-03}$ for transcriptomics and proteomics respectively, **Fig. 8c,d**). However, no significant
497 differences in transcriptomic or proteomic levels in the additional donors (MSBB-C2) were
498 observed ($p > 0.05$, **Fig. 8c,d**). Moreover, the downregulation of *SNCA* in the MSBB-C1 donors
499 was also seen in the frontal pole (BM10), superior temporal gyrus (BM22), and inferior frontal
500 gyrus (BM44) brain regions from the same study (**Supplementary Fig. 18a-c**). Similarly, *SNCA*
501 transcriptomic levels were significantly downregulated in ROSMAP-C1 (**Fig. 8e**) compared to the
502 control ($p=3.5x10^{-02}$) and other AD cases ($p=4.5x10^{-03}$).
503
504 The downregulation of *SNCA* in Knight-C4 further supports the association of this profile with
505 synaptic dysfunction. S*NCA* encodes α-synuclein (αSyn), an abundant presynaptic protein that
506 plays a role in regulating synaptic vesicle trafficking and subsequent neurotransmitter
507 release[84][PMID:7646890]. αSyn itself is released from neurons via multiple cell biological
508 mechanisms including exocytosis and is detectable in the extracellular space as a monomer and
509 pathologically aggregated species[85–87]. Mutations in *SNCA*, overexpression of wildtype αSyn, and
510 injection of pre-formed aSyn fibrils (PFFs) are sufficient to cause neuropathology and
511 parkinsonian syndromes in both humans and animal models[88–94].
512
513 To further study synapses in a model of pathological αSyn aggregation, we integrated and
514 analyzed NanoString gene expression data (see Methods) generated from a panel of 807 gene
515 transcripts measured in brain tissue of A53T αSyn transgenic mice crossed onto *APOE*
516 backgrounds[95]. Of the 807 genes included in the panel, 89 (11%) were classified as synaptic
517 genes (SynGO consortium[37]). We performed DE analysis to compare mice with high pSyn
518 pathology (defined as > 5% coverage by IHC) to mice with low pathology (< 5% coverage by IHC,
519 see Methods) and identified 255 DEGs, including 13 synaptic genes. We then interrogated which
520 of synaptic genes in DE in Knight-C4 were also dysregulated in the A53T αSyn-APOE transgenic
521 mice (**Supplementary Fig. 19**). We identified that all of the 13 genes identified in the mouse
522 model were also dysregulated in Knight-C4 compared with other AD cases (Fisher $p=3.6x10^{-09}$)
523 (**Supplementary Fig. 19a**). These results suggest that the profile of Knight-C4 might be driven
524 by synaptic dysregulation. When comparing to the control, only five synaptic genes were shared
525 between the three datasets (**Supplementary Fig. 19b**). These findings were limited by the fact
526 that the NanoString panel contains only 12% of the significant genes in Knight-C4.
527

528 Altogether, these results provide more evidence of the association of Knight-C4 with synaptic
529 dysfunction, and suggest that the lower levels of soluble *SNCA* observed in this profile with worse
530 cognitive function in multiple brain regions may be associated with αSyn (Lewy body) pathology
531 which is present in 50-60% of AD brains[96–98].
532
533 **Discussion**
534 In the current study, we leveraged machine learning approaches to integrate high-throughput
535 multi-omics datasets from AD cases in multiple cohorts and brain regions. Our analyses led us to
536 identify multiple molecular profiles, genes/pathways, and cellular population structures affected in
537 specific AD cases. We found a distinct molecular profile associated with worse cognitive function,
538 earlier age at onset/death, and dysregulated neuronal/synaptic pathways in the parietal cortex.
539 Our results suggest that this molecular profile captures additional neuropathological events not
540 fully captured by β-amyloid or tau staging as previously reported[99,100]. Similar to previous studies[8],
541 we did not find a significant association between any molecular profile and APOE ε4 or PRS,
542 suggesting that these molecular profiles are not driven by genetic factors associated with genetic
543 AD risk. We also determined, through an exhaustive replication, that this molecular profile is also
544 present in multiple brain regions from the MSBB and ROSMAP cohorts. Further, we demonstrated
545 that in each of these studies, this molecular profile is associated with worse clinical and cognitive
546 function, offering new insights into the molecules that may wax or wane as cognition declines.
547
548 Notably, we observed that these donors with worse cognitive function showed a significantly
549 increased proportions of astrocytes and reduced proportions of neurons. This might indicate a
550 neuronal loss at later stages of AD, leading to the development of severe cognitive impairment.
551 Consistent with these findings of gliosis and neuronal loss, we found that the molecular profile of
552 this subset of participants is significantly enriched in dysregulated synaptic genes, which may
553 reflect greater synaptic losses in participants with worse cognitive function. Furthermore, we
554 identified dysfunction of several neuronal and synaptic pathways that are significantly associated
555 with this profile including dopaminergic synapse and glutamatergic pathways. Although the
556 involvement of dopamine in AD is still debatable, several studies have shown an association
557 between AD pathologies (e.g. Aβ deposition) and the dopaminergic dysfunction[101–109]. Similarly,
558 glutamatergic synapse is a major excitatory neurotransmitter in the central nervous system (CNS)
559 that regulates several neuronal functions. Increasingly evidence has demonstrated an association
560 between glutamate alterations and AD pathology[110–113]. Moreover, metabolic differential
561 abundance analyses identified significant metabolites associated with biologically relevant AD

562 pathways, most of which are related to major neurotransmitter pathways. This may suggest that
563 donors in this profile have a more profound neuronal or synaptic dysregulation. These brains also
564 exhibit a metabolic profile associated with disease duration, supporting its implication in later
565 stages AD. Altogether, these results suggest that these donors have a more prominent synaptic
566 loss and dysfunction associated with worse cognition and advanced AD indicating the feasibility
567 of our approach in identifying molecular changes that occur at late-stage AD. These novel
568 molecular findings can be leveraged to identify new biomarkers and potentially therapeutic targets
569 – not necessarily related to amyloid and tau – that might be effective in later stages of disease,
570 when therapeutics targeting amyloid seem to be not effective.
571
572 Our subsequent transcriptomic and proteomic differential expression analyses showed a
573 significantly higher number of hits associated with this profile. Further investigations showed that
574 many of the genes detected in this profile were neuronal/synaptic genes. Among these, we
575 identified *SNAP25* downregulated in this profile in the discovery and replication cohorts. *SNAP25*
576 is implicated in several physiological pathways and neuropathologies, including AD and Down
577 syndrome (DS)[59]. Transcriptomic and proteomic DE analyses also identified several genes
578 associated with AD including *GFAP* and *CLU* that are more noticeably dysregulated in the distinct
579 profiles. These two genes were not significantly altered in non-clustered AD cases compared to
580 controls. Increased astrocyte reactivity and overexpression of *GFAP* has been reported in multiple
581 studies[61,114–121]. For instance, Greber *et al*., 1999[59], showed that elevated expression of *GFAP* was
582 found in multiple brain regions including frontal, parietal, temporal and occipital cortex regions of
583 AD patients[59]. In addition, we leveraged CSF proteomics and longitudinal clinical evaluations to
584 show that lower levels of CSF *CLU* are associated with increased risk of dementia progression in
585 an extended cohort from the Knight ADRC. Multiple studies have shown the ability of *CLU* to play
586 a neuroprotective role in AD by altering Aβ aggregations and promote Aβ clearance[66,122–125]. More
587 recently, additional studies have identified *CLU* as mediator regulating Aβ-induced
588 neurotoxicity[66,126,127]. The heterogeneity among AD cases opens the possibility to develop novel
589 approaches using these and additional genes for biomarkers for precision medicine approaches.
590
591 By integrating single-nuclei RNA-seq, we extracted additional insights that support distinct patters
592 of cell types implicated in molecular AD profiles. For instance, we determined that *CX3CR1*, a
593 homeostatic microglia gene, is significantly downregulated in Knight-C4. The lower expression of
594 *CX3CR1* may indicate loss of homeostatic microglia, leading to neuronal damage and loss. The
595 association between homeostatic microglia function and the degree of neuronal cell loss has been

596 shown to be linked to AD[128]. More recently, the loss of *CX3CR1* was shown to be associated with
597 microglial dysfunction illustrated by dampened TGFβ-signaling, increased oxidative stress
598 responses, and dysregulated pro-inflammatory activation[129]. However, we could not determine a
599 significant overexpression of microglia activated genes, nor replicate this in additional cortical
600 regions.
601
602 Our results highlight the utility of integrating multi-omics and indicate that cross-omics signatures
603 can better capture differences and discriminate molecular variations in complex and
604 heterogenous diseases such as AD. An increasing body of evidence endorses broad clinical and
605 neuropathological heterogeneity among individuals with AD. Similarly, there is evidence that
606 molecular profiles differ substantially within AD[5]. Recognizing this heterogeneity is highly
607 important as the *statistical noise* introduced by molecular heterogeneity within a population has
608 potential to conceal dysregulated pathways and risk biomarkers that might be altered in only
609 subsets of the AD general population. Therefore, identifying cross-omics molecular profiles of AD
610 and their molecular signatures, including genes and pathways, may extract new insights into AD
611 and provide a platform to identify novel treatments and biomarkers and eventually be used toward
612 precision medicine.
613
614 We identified that alpha-synuclein is among the top molecules that better discriminate AD profiles.
615 We demonstrated that the transcriptomic/proteomic levels of *SNCA* were under-expressed in a
616 subset of AD cases with more advanced tau pathology and worse cognition and replicated this
617 finding in several cohorts and brain regions. Alpha-synuclein is the major constituent of Lewy
618 bodies, the histopathologic hallmark of PD and DLB. LB pathology is also present in a majority of
619 advanced AD cases, including autosomal dominant AD and DS[96,97,130–134]. In addition, previous
620 work has demonstrated the relationship between alpha-synuclein and AD pathologies including
621 Aβ plaques and tau aggregations using different mouse models[135,136] and implicated alpha-
622 synuclein pathology in neuronal/synaptic dysfunction and death[137,138]. Despite the substantial
623 efforts in studying the association between this protein and AD pathology, the mechanism by
624 which alpha-synuclein contributes to the pathogenesis and progression of AD remains elusive.
625 Several regulatory mechanisms that contribute to *SNCA* levels have been proposed. For instance,
626 previous studies based on primary neuronal cultures[87], mice models[139], and live cell imaging[140]
627 have shown that the neuronal or synaptic activities regulate the aggregation of *SNCA*. The
628 downregulation of *SNCA* in a subset of AD cases with unfavorable outcomes suggests that *SNCA*
629 might be a marker of neuron/synapse losses that can be used for identifying AD stages.

630 Although this work provides new insights for understanding the genetic and molecular
631 heterogeneity in AD brains, additional work is needed to further validate and better understand
632 the role of key drivers of these profiles using model organisms (e.g., mouse models). Therefore,
633 future studies analyzing longitudinal data from mouse models and longitudinal CSF/plasma from
634 large cohorts can certainly provide novel insights into the role of synaptic dysregulation in AD,
635 and the possible implication in additional neurodegenerative diseases. Additional work is also
636 needed to investigate the possible role that alpha-synuclein in AD, and its association with worse
637 cognitive function. In addition, this work has promising implications in drug repositioning and
638 staging biomarkers that may lead to identify novel therapeutic potentials for AD and therefore
639 bring AD a step closer to precision medicine care and treatment.
640

641 **Methods**

642

643 **Study Cohorts**

644

645 **Discovery (Knight ADRC)**

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

652

653 - **RNA-seq**

654 The data generation for the RNA-seq dataset has been previously described[29–31]. Briefly,
655 total RNA was extracted from frozen parietal cortex tissue using a Tissue Lyser LT and
656 purified using RNeasy Mini Kits (Qiagen). The Nanodrop 8000 (Thermo Scientific) and
657 TapeStation 4200 (Agilent Technologies) were used to perform quality control of the
658 RNA's concentration and purity, and degradation. The RNA integrity number (RIN) was
659 calculated using an RNA 6000 Pico assay on a Bioanalyzer 2100 and TapeStation 4200
660 (Agilent Technologies). The software determines the RINe on the Bioanalyzer and
661 TapeStation taking into account the entire electrophoretic trace of the RNA, including the
662 presence or absence of degradation products. The DV200 value is defined as the
663 percentage of nucleotides greater than 200nt. All cDNA libraries were prepared using a
664 TruSeq Stranded Total RNA Sample Prep with Ribo-Zero Gold kit (Illumina) and

665 sequenced on an Illumina HiSeq 4000 using 2 x 151 paired-end reads at the McDonnell
666 Genome Institute at Washington University in St. Louis. The average number of reads per
667 sample was 43,195,080.

668 - **Proteomics**

669 Proteomic data from the Knight ADRC were generated on the SomaLogic platform
670 (SomaLogic Operating Co., Inc., Boulder, CO) from frozen parietal cortical tissue samples
671 (~500mg) and CSF samples. The SomaLogic platform is a multiplexed, aptamer-based
672 protein quantification platform[141]. The platform measured 1,305 aptamers in total. Full
673 details of the proteomic data generation have been described previously[32].

674 - **Metabolomics**

675 Metabolomic data from the Knight ADRC were generated on the Metabolon Precision
676 Metabolomics™ UPLC-MS/MS platform (Metabolon, Inc., Morrisville, USA), from 50mg
677 frozen parietal cortical tissue samples. The platform measured 880 metabolites from nine
678 "super pathways": amino acids, carbohydrates, cofactors and vitamins, energy, lipids,
679 nucleotides, peptides, xenobiotics, and partially characterized molecules. Metabolon
680 provided structural identity annotations for 815 metabolites; the remaining 65 were
681 excluded from the final dataset. Full details of the metabolomics data generation have
682 been described previously[33].

683 - **Single-nuclei (snRNA-seq)**

684 The data generation of snRNA-seq has been previously described in[142,143]. Briefly, 74
685 frozen parietal tissues were processed according to the 'Nuclei extraction and library
686 preparation' protocol described in[142]. In this protocol, the tissue was homogenized, and
687 the nuclei were isolated using a density gradient. The nuclei were then sequenced using
688 the 10X Chromium single cell Reagent Kit v3, with 10,000 cells per sample and 50,000
689 reads per cell for each of the 74 samples.

690 **Replication (MSBB, BM36)**

691 - **RNA-seq**

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

698 prepared using the TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, CA).
699 The rRNAs were depleted using the Ribo-Zero rRNA Removal Kit (human/mouse/rat)
700 (Illumina). Single-end non-standard reads of 101bp were generated by Illumina HiSeq
701 2500 (Illumina)[35]. The average number of reads per sample was 32,201,200. RNA-seq
702 data from region parahippocampal gyrus (BM36) was selected for the cross-omics
703 integration approach because of the availability of proteomics data. Data from the
704 remaining regions were used to explore the identified profiles' expression profiles to
705 determine whether these are consistent across cortical brain regions.

706 - **Proteomics (TMT)**

707 The proteomics TMT data were obtained from the AMP-AD Synapse portal (syn25006650)
708 processed by Johnson *et al.*, 2022[144]. The raw proteomic TMT data is publicly available
709 at AMP-AD and can be downloaded from Synapse portal (syn21347564). In short, before
710 TMT labeling, the 198 samples were randomized by co-variates (protein quality, sample
711 concentration, diagnosis, age, and sex) into 20 batches (10 cases per batch). The
712 digested peptides were resuspended in 50 mM HEPES (pH 8.5) and labeled with the TMT
713 11-plex kit (Thermo Fisher) according to the manufacturer protocol. In each batch, GIS
714 samples were labeled using TMT channel 126. All 11 channels were mixed equally,
715 desalted with a 100 mg C18 Sep-Pak column (Waters) for the subsequent
716 fractionation[144,145].

717 **Replication (ROSMAP, DLPFC)**

718 - **RNA-seq**

719 The RNA-seq raw data from The Religious Orders Study and Memory and Aging Project
720 (ROSMAP) Study was downloaded from the Synapse portal (syn17008934) available from
721 AMP-AD. Briefly, RNA-seq samples were extracted using Qiagen's miRNeasy mini kit (cat.
722 no. 217004) and the RNase-free DNase Set (cat. no. 79254) and quantified by Nanodrop.
723 Agilent Bioanalyzer evaluated quality. Library preparation was performed by poly-A
724 selection followed by first strand-specific cDNA synthesis, then dUTP for second strand-
725 specific cDNA synthesis followed by fragmentation and Illumina adapter ligation for library
726 construction. Paired-end read sequences with a length of 101bp were generated by
727 Illumina HiSeq. Fifty-seven samples were submitted later to the platform and run on an
728 updated protocol (batch 2 protocol). For consistency, we removed those samples from this
729 cohort. The average number of reads per sample was 47,434,600. Only No Cognitive
730 Impairment (NCI) and AD with NO other cause of cognitive impairment cases were used

731 in this study. All cases with “unknown” BraakTau scores were discarded. Plaque density
732 and Neurofibrillary tangle burden (PHF tau tangles) are a summary of a AD pathology
733 derived from three AD pathologies: neuritic plaques, diffuse plaques, and neurofibrillary
734 tangles based on five regions: midfrontal cortex, midtemporal cortex, inferior parietal
735 cortex, entorhinal cortex, and hippocampus. Tangles are identified by molecularly specific
736 immunohistochemistry based on the mean of eight regions: tangles_hip (hippocampus),
737 tangles_ec (entorhinal cortex), tangles_mf (midfrontal cortex), tangles_it (inferior
738 temporal), tangles_ag (angular gyrus), tangles_calc (calcarine cortex), tangles_cg
739 (anterior cingulate cortex), angles_sf (superior frontal cortex).

740 - **Metabolomics**

741 Data from the Religious Orders Study and Memory and Aging Project (ROSMAP) were
742 generated by the Duke Metabolomics and Proteomics Shared Resource, a member of the
743 ADMC, using protocols published previously for blood samples[146–148]. A custom protocol
744 developed for the brain samples can be found on Synapse at syn10235609. The
745 dorsolateral prefrontal cortex (DLPFC) metabolomic data from the ROSMAP studies
746 quantified on the Metabolon Precision Metabolomics platform and preprocessed by the
747 ADMC as described in[149] were downloaded from Synapse in July 2021 (syn25878459).
748 The platform measured a total of 1,055 metabolites in the ROSMAP dataset.

749

750 **Data processing**

751 RNA-seq QC, alignment, and gene expression quantification: All RNA-seq datasets from the
752 discovery and replication were processed and aligned using our in-house RNA-seq pipeline
753 (https://github.com/HarariLab/RNA-seq-Pipeline). Genome reference and gene models were
754 selected similar to the TOPmed pipeline (https://github.com/broadinstitute/gtex-
755 pipeline/blob/master/TOPMed_RNAseq_pipeline.md). Reference genome GRCh38 and
756 GENCODE 33 annotation, including the addition of ERCC spike-in annotations, were used. We
757 excluded ALT, HLA, and Decoy contigs from the reference genome due to the lack of RNA-seq
758 tools that properly handle these regions. Before alignment, the quality of raw read sequences for
759 all libraries was assessed using FastQC (v0.11.9)[150]. All raw read sequences were then aligned
760 to the human reference genome (GRCh38) using STAR (v.2.7.1a)[151]. The alignment quality was
761 evaluated using sequencing metrics such as reads distribution, ribosomal content, or alignment
762 quality provided by STAR using Picard tools (v.2.8.2)[152]. All samples that failed to pass the QC or
763 were outliers (Knight ADRC=9, MSBB=18, ROSMAP=6) were removed from the downstream
764 analyses. Raw read counts for transcripts and genes were generated using STAR and computed

765 transcript/gene expression levels as normalized in FPKM (Fragments Per Kilobase of transcript
766 per Million mapped reads) format.

767 Metabolomics QC and Quantification**:** Metabolomics data from Knight ADRC and ROSMAP were
768 processed similarly. Full details of the QC process have been previously described[153]. Briefly, 188
769 metabolites with readings missing in over 20% of donors were excluded, readings were log10-
770 transformed, means per metabolite were adjusted to zero, and outlier readings (outside 1.5xIQR)
771 were excluded. Missing values for metabolites in < 20% of donors were replaced by their mean
772 value. No samples were removed due to the missingness rate. Principal component analysis was
773 performed with the FactoMineR R package[154] to identify outlier samples; four outlier samples were
774 excluded. The final dataset consisted of 627 metabolites. A similar procedure was used to clean
775 the ROSMAP metabolomic dataset. Metabolites without assigned structural identities and
776 metabolites with greater than 20% missing readings were excluded, readings were log10-
777 transformed, the mean of each metabolite's distribution was adjusted to zero, and outlier readings
778 were removed. Two samples missing greater than 20% of readings were excluded. The final
779 dataset consisted of 595 metabolites.

780 Proteomics QC and Quantification**:** Proteomic data for brain and CSF samples from the Knight
781 ADRC were first filtered based on a limit of detection (LOD). Samples with greater than 15%
782 outlier analytes (analyte readings below the LOD) were excluded. Analytes were then filtered
783 based on scale factor difference, meaning the absolute value of the maximum difference between
784 the calibration scale factor per aptamer and the median for each plate run. Analytes were
785 excluded if their scale factor difference exceeded 0.5. Based on log10-transformation of protein
786 readings, samples exceeding 1.5-fold of the interquartile range (IQR) per analyte were considered
787 outliers. Analytes with greater than 15% outlier samples were excluded. Similarly, samples that
788 were outliers for greater than 15% of analytes were excluded. Full details of the proteomic data
789 QC have been described previously[32]. Missing analyte readings and removed outlier readings
790 were imputed using the impute.knn function from the impute R package (v1.56.0)[155]. A plate-wise
791 batch correction was performed using the ComBat function from the sva R package (v3.30.1)[156].
792 The final brain dataset consisted of 1,092 analytes, and the final CSF dataset consisted of 713
793 analytes. Before using this data, analytes with missing expressions in > 20% of donors were
794 excluded, and missing expressions of those in < 20% were replaced with their mean value.

795 The processed data for TMT proteomics from MSBB BM36 was obtained from AMP-AD
796 (syn25006647) and processed by Johnson *et al.*,2022[144]. In brief, 760 raw files generated from
797 19 TMT 11-plexes were analyzed using the Proteome Discoverer suite (version 2.3, Thermo

798 Fisher Scientific). The UniProtKB human proteome database containing Swiss-Prot and TrEMBL
799 human reference protein sequences was used to search mass spectra. Peptide spectral matches
800 and peptides were filtered to a false discovery rate (FDR) of less than 1% using Percolator
801 software. After spectral assignment, peptides were assembled into proteins and were further
802 filtered based on the combined probabilities of their constituent peptides to a final FDR of 1%[144].
803 Unique and razor (that is, parsimonious) peptides were considered for quantification. TMT
804 reporter abundance was transformed into a ratio followed by $\log_2$ transformation. Quality control
805 analysis was performed, and outliers were removed. Similar to the Knight ADRC, analytes with
806 missing expressions in > 20% of donors were excluded, and missing expressions of those in <
807 20% were replaced with their mean value.

808 Single-nuclei (snRNA-seq) QC, alignment, and gene expression quantification: Full details of the
809 data process of snRNA-seq data have been previously described in [142,143]. In short, CellRanger
810 (v2.1.1 10XGenomics)[157] software was employed to align the sequences and quantify gene
811 expression. Read sequences were aligned to a custom pre-mRNA reference constructed from
812 genome assembly GRCh38 and generated as described by 10X Genomics technical manual.
813 Filtering and QC analyses were performed individually using the Seurat package (v2.20 and 2.30)
814 on each subject. All nuclei with high mitochondria gene expression were removed and only genes
815 expressed in > 3 cells were kept for downstream analysis. Cells with < 1800 or > 8000 genes
816 expressed in them were excluded. The data were normalized using the *LogNormalize* function
817 that normalizes the gene expression measurements for each cell by the total expression, scales
818 by a factor equal to the median counts of all genes, and log-transforms the expression.

819 **Clustering of AD participants**

820 To cluster AD participants, iClusterBayes (Integrative clustering of multiple genomic data types)[34]
821 available from iClusterPlus R package (v1.22.0)[158] was employed. For each integration process,
822 shared samples across omics datasets were extracted, and the top 2000 most variant features
823 were selected. iClusterBayes was performed by fitting multiple Bayesian models using a different
824 number of clusters ranging from 1 to 10. W ran 22,000 Markov Chain Monte Carlo (MCMC)
825 iterations for each, of which the first 12,000 were discarded as burn-in. The prior probability for
826 the indicator variable gamma of each data set was set to 0.1, whereas the posterior probability
827 cutoff was set at 0.5. The standard deviation of the random walk proposal for the latent variable
828 was set to 0.5. The remaining parameters were left with the default values.

829 **Cell proportion analysis**

830 CellMix R package (v1.6.2)[159], employed multiple digital deconvolution methods, was used to
831 estimate cellular population structure from gene expression data (TPM) generated by the Salmon
832 quantification tool. The machine learning model and the gene panel (genes marker expressed
833 highly in specific cell types) previously described in[31]. Four cell types, astrocytes, microglia,
834 neurons, and oligodendrocytes, were tested using two algorithms, meanProfile[159] and ssNMF
835 (semi-supervised nonnegative matrix factorization)[160]. The significant difference in cell proportion
836 between clusters was computed using a generalized linear model (GLM) test and corrected by
837 sex and age of death (AOD).
838
839 **Differential expression analysis of AD clusters**
840 RNA-seq: differential expression (DE) analyses were performed to compare subjects in each
841 cluster to the control subjects as well as other AD subjects using the DESeq2 R package
842 (v.1.22.2)[161]. Our models were adjusted for sex, age at death, and the percentage of astrocytes
843 and neurons. To enhance our confidence in the differentially expressed genes we discovered,
844 lowly expressed genes were removed. Only genes with expression > 0.5 CPM (count per million)
845 in at least 25% of samples in either group being compared were retained for downstream
846 analyses. The false discovery rate (FDR) was estimated using Benjamini-Hochberg (BH)[162]. All
847 genes with FDR < 0.05 were considered differentially expressed genes.

848 Proteomics: differential expression analyses between clusters were carried out with linear models
849 (R function glm) with age at death and sex as covariates. P-values were adjusted using Benjamini-
850 Hochberg (BH)[162].

851 Metabolomics: differential abundance analyses were also carried out with linear models, with age
852 at death, sex, and post-mortem interval (PMI) as covariates. P-values were adjusted using
853 Benjamini-Hochberg (BH)[162].

854 Single-nuclei: differential expression analyses were performed to compare the overexpression of
855 each gene in each cell-type relative to other cell types using glmmTMB R Package (v1.1.3). This
856 approach fits generalized linear mixed models using a template builder (glmmTMB). Our model
857 was corrected by sex covariate. To ensure cell-type specific overexpression, for each cell type,
858 only top genes whose effect size is greater than 90% percentile cutoff were retained for
859 downstream integration analyses.
860
861 **Survival analyses**

Kaplan-Meier survival analyses and Cox proportional hazards models were used to determine the association of different clusters with survival outcomes and also to determine the association with the risk of dementia progression using survival (v3.2-3)[163] and surviplot (v1.1.1)[164] R packages.

Significance of overlap

The significance of overlap for all analyses was computed using Fisher's exact test available from the R stats package (v4.3.0)[165].

Pathway Analysis

All pathway analyses for transcriptomics and proteomics were performed using the EnrichR R package (v3.0)[166] using genes/proteins significantly up-and downregulated between clusters vs. control and between distinct clusters. Two widely used databases, gene ontologies (GOs), related to molecular function, biological process, and cellular components, and KEGG databases were used. All pathways with an adjusted p-value < 0.05 were considered statistically significant. Similar pathways analyses were performed for single-nuclei (snRNA-seq) using EnrichR using significantly up-and downregulated genes in specific cell types. For metabolomics, pathways analyses were performed using MetaboAnalyst (v5.0) software[167] using the metabolites that were increased- and decreased between clusters and the control and across clusters.

Replication analysis

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

NanoString gene expression analysis

The data generation and full details of the QC processes have been previously described[95]. Briefly, A53T αSyn-Tg mice were crossed with human *APOE2*, *APOE3*, or *APOE4* knockin mice or *Apoe* knockout mice to generate A53T mice homozygous for one of the three human alleles or completely null for *Apoe*. Sample sizes and genotypes for NanoString gene expression analysis were selected on the basis of αSyn pathology and biochemistry results. RNA was isolated from 12-month old A53T/EKO (n=10), A53T/E2 (n=2), and A53T/E4 (n=10) mouse midbrain. 793 transcripts were quantified with the NanoString nCounter multiplexed target platform using a customized chip based on the Mouse Neuroinflammation panel (www.nanostring.com). The

geometric mean of negative control lanes was subtracted from gene transcript counts, and gene expression data were normalized to the geometric mean of positive control lanes and a list of housekeeping genes[95]. QC analyses including principal components analysis were performed using nSolver 4.0 and the Advanced Analysis 2.0 plugin (NanoString). For differential expression analysis, pSyn pathology was expressed as low (< 5% coverage by IHC) or high (> 5% coverage by IHC) for each animal, and APOE genotype and pSyn pathology were selected as predictor covariates. Differentially expressed analyses were performed using linear regression models in R with sex as covariates. P-values were adjusted using Benjamini-Hochberg (BH)[162].

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 pf publication.

References

1. Alzheimer's A. 2021 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2021;17:327-406.

2. Vogel JW, Young AL, Oxtoby NP, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat Med*. 2021;27(5):871-881.

930 3. Jellinger KA. Pathobiological subtypes of Alzheimer disease. *Dement Geriatr Cogn Disord*.
931 2020;49(4):321-333.

932 4. Collij LE, Salvadó G, Wottschel V, et al. Spatial-temporal patterns of β-amyloid accumulation:
933 a subtype and stage inference model analysis. *Neurology*. 2022;98(17):e1692-e1703.

934 5. Neff RA, Wang M, Vatansever S, et al. Molecular subtyping of Alzheimer's disease using
935 RNA sequencing data reveals novel mechanisms and targets. *Sci Adv*. 2021;7(2):eabb5398.

936 6. Tijms BM, Gobom J, Reus L, et al. Pathophysiological subtypes of Alzheimer's disease
937 based on cerebrospinal fluid proteomics. *Brain*. 2020;143(12):3776-3792.

938 7. Ferreira D, Nordberg A, Westman E. Biological subtypes of Alzheimer disease: A systematic
939 review and meta-analysis. *Neurology*. 2020;94(10):436-448.

940 8. Whitwell JL, Dickson DW, Murray ME, et al. Neuroimaging correlates of pathologically
941 defined subtypes of Alzheimer's disease: a case-control study. *Lancet Neurol*.
942 2012;11(10):868-877.

943 9. Whitwell JL, Graff-Radford J, Tosakulwong N, et al. [18F] AV-1451 clustering of entorhinal
944 and cortical uptake in Alzheimer's disease. *Ann Neurol*. 2018;83(2):248-257.

945 10. Byun MS, Kim SE, Park J, et al. Heterogeneity of regional brain atrophy patterns associated
946 with distinct progression rates in Alzheimer's disease. *PloS One*. 2015;10(11):e0142756.

947 11. Choi M, Youn H, Kim D, et al. Comparison of neurodegenerative types using different brain
948 MRI analysis metrics in older adults with normal cognition, mild cognitive impairment, and
949 Alzheimer's dementia. *PloS One*. 2019;14(8):e0220739.

950 12. Abeshouse A, Adebamowo C, Adebamowo SN, et al. Comprehensive and integrated
951 genomic characterization of adult soft tissue sarcomas. *Cell*. 2017;171(4):950-965.

952 13. Wheeler DA, Roberts LR, Network CGAR. Comprehensive and integrative genomic
953 characterization of hepatocellular carcinoma. *Cell*. 2017;169(7):1327.

954 14. Network CGAR. Comprehensive genomic characterization of squamous cell lung cancers.
955 *Nature*. 2012;489(7417):519.

956 15. Levine DA. Integrated genomic characterization of endometrial carcinoma. *Nature*.
957 2013;497(7447):67-73.

958 16. Bass AJ, Thorsson V, Shmulevich I, et al. Comprehensive molecular characterization of
959 gastric adenocarcinoma. *Nature*. 2014;513(7517):202.

960 17. Never-smoker NE smoker. Comprehensive molecular profiling of lung adenocarcinoma.
961 *Nature*. 2014;511:543-550.

962 18. Jendoubi T. Approaches to Integrating Metabolomics and Multi-Omics Data: A Primer.
963 *Metabolites*. 2021;11(3):184.

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

967 20. Nativio R, Lan Y, Donahue G, et al. An integrated multi-omics approach identifies epigenetic
968 alterations associated with Alzheimer's disease. *Nat Genet*. 2020;52(10):1024-1035.

969 21. Hampel H, Nisticò R, Seyfried NT, et al. Omics sciences for systems biology in Alzheimer's
970 disease: State-of-the-art of the evidence. *Ageing Res Rev*. 2021;69:101346.

971 22. Byrns CN, Bonini NM. An Integrated Multi-omics Approach Identifies Therapeutic Potential
972 for ATP6V1A in Late Onset Alzheimer's Disease. *Neuron*. 2021;109(2):193-194.

973 23. Sancesario GM, Bernardini S. Alzheimer's disease in the omics era. *Clin Biochem*.
974 2018;59:9-16.

975 24. Jin T, Rehani P, Ying M, et al. scGRNom: a computational pipeline of integrative multi-omics
976 analyses for predicting cell-type disease genes and regulatory networks. *Genome Med*.
977 2021;13(1):1-15.

978 25. Li Z, Jiang X, Wang Y, Kim Y. Applied machine learning in Alzheimer's disease research:
979 omics, imaging, and clinical data. *Emerg Top Life Sci*. 2021;5(6):765-777.

980 26. Canchi S, Raao B, Masliah D, et al. Integrating gene and protein expression reveals
981 perturbed functional networks in Alzheimer's disease. *Cell Rep*. 2019;28(4):1103-1116.

982 27. Xie L, He B, Varathan P, et al. Integrative-omics for discovery of network-level disease
983 biomarkers: a case study in Alzheimer's disease. *Brief Bioinform*. 2021;22(6):bbab121.

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

986 29. Dube U, Del-Aguila JL, Li Z, et al. An atlas of cortical circular RNA expression in Alzheimer
987 disease brains demonstrates clinical and pathological associations. *Nat Neurosci*.
988 2019;22(11):1903-1912.

989 30. Chen HH, Eteleeb A, Wang C, et al. Circular RNA detection identifies circPSEN1 alterations
990 in brain specific to autosomal dominant Alzheimer's disease. *Acta Neuropathol Commun*.
991 2022;10(1):1-11.

992 31. Li Z, Del-Aguila JL, Dube U, et al. Genetic variants associated with Alzheimer's disease
993 confer different cerebral cortex cell-type population structure. *Genome Med*. 2018;10(1):1-
994 19.

995 32. Yang C, Farias FH, Ibanez L, et al. Genomic atlas of the proteome from brain, CSF and
996 plasma prioritizes proteins implicated in neurological disorders. *Nat Neurosci*.
997 2021;24(9):1302-1312.

998 33. Novotny BC, Fernandez MV, Wang C, et al. Metabolomic and lipidomic signatures in
999 autosomal dominant and late-onset Alzheimer's disease brains. *Alzheimers Dement J*
1000 *Alzheimers Assoc*. Published online October 17, 2022. doi:10.1002/alz.12800

1001 34. Mo Q, Shen R, Guo C, Vannucci M, Chan KS, Hilsenbeck SG. A fully Bayesian latent
1002 variable model for integrative clustering analysis of multi-type omics data. *Biostatistics*.
1003 2018;19(1):71-86.

1004 35. Wang M, Beckmann ND, Roussos P, et al. The Mount Sinai cohort of large-scale genomic,
1005 transcriptomic and proteomic data in Alzheimer's disease. *Sci Data*. 2018;5(1):1-16.

1006 36. A Bennett D, A Schneider J, Arvanitakis Z, S Wilson R. Overview and findings from the
1007 religious orders study. *Curr Alzheimer Res*. 2012;9(6):628-645.

1008 37. Koopmans F, van Nierop P, Andres-Alonso M, et al. SynGO: an evidence-based, expert-
1009 curated knowledge base for the synapse. *Neuron*. 2019;103(2):217-234.

1010 38. Zhu F, Cizeron M, Qiu Z, et al. Architecture of the mouse brain synaptome. *Neuron*.
1011 2018;99(4):781-799.

1012 39. Patel KR, Zhu K, Henrion MY, et al. Single cell-type integrative network modeling identified
1013 novel microglial-specific targets for the phagosome in Alzheimer's disease. *bioRxiv*.
1014 Published online 2020.

1015 40. Wang X, Huang T, Bu G, Xu H. Dysregulation of protein trafficking in neurodegeneration.
1016 *Mol Neurodegener*. 2014;9(1):1-9.

1017 41. Wang C, Yu JT, Miao D, Wu ZC, Tan MS, Tan L. Targeting the mTOR signaling network for
1018 Alzheimer's disease therapy. *Mol Neurobiol*. 2014;49(1):120-135.

1019 42. Hwang J, Estick CM, Ikonne US, et al. The role of lysosomes in a broad disease-modifying
1020 approach evaluated across transgenic mouse models of Alzheimer's disease and
1021 Parkinson's disease and models of mild cognitive impairment. *Int J Mol Sci*.
1022 2019;20(18):4432.

1023 43. Butler D, Hwang J, Estick C, et al. Protective effects of positive lysosomal modulation in
1024 Alzheimer's disease transgenic mouse models. *PloS One*. 2011;6(6):e20501.

1025 44. Zhang L, Sheng R, Qin Z. The lysosome and neurodegenerative diseases. *Acta Biochim*
1026 *Biophys Sin*. 2009;41(6):437-445.

1027 45. Root J, Merino P, Nuckols A, Johnson M, Kukar T. Lysosome dysfunction as a cause of
1028 neurodegenerative diseases: Lessons from frontotemporal dementia and amyotrophic
1029 lateral sclerosis. *Neurobiol Dis*. 2021;154:105360.

1030 46. Orr ME, Oddo S. Autophagic/lysosomal dysfunction in Alzheimer's disease. *Alzheimers Res*
1031 *Ther*. 2013;5(5):1-9.

1032 47. Eshraghi M, Adlimoghaddam A, Mahmoodzadeh A, et al. Alzheimer's disease pathogenesis:
1033 Role of autophagy and mitophagy focusing in microglia. *Int J Mol Sci*. 2021;22(7):3330.

1034 48. Li Q, Liu Y, Sun M. Autophagy and Alzheimer's disease. *Cell Mol Neurobiol*.
1035 2017;37(3):377-388.

1036 49. Zare-Shahabadi A, Masliah E, Johnson GV, Rezaei N. Autophagy in Alzheimer's disease.
1037 *Rev Neurosci*. 2015;26(4):385-395.

1038 50. Nixon RA. The role of autophagy in neurodegenerative disease. *Nat Med*. 2013;19(8):983-
1039 997.

1040 51. Funderburk SF, Marcellino BK, Yue Z. Cell "self-eating"(autophagy) mechanism in
1041 Alzheimer's disease. *Mt Sinai J Med J Transl Pers Med J Transl Pers Med*. 2010;77(1):59-
1042 68.

1043 52. Dobson CB, Itzhaki RF. Herpes simplex virus type 1 and Alzheimer's disease. *Neurobiol*
1044 *Aging*. 1999;20(4):457-465.

1045 53. Itzhaki RF. Herpes simplex virus type 1 and Alzheimer's disease: increasing evidence for a
1046 major role of the virus. *Front Aging Neurosci*. 2014;6:202.

1047 54. Wozniak MA, Mee AP, Itzhaki RF. Herpes simplex virus type 1 DNA is located within
1048 Alzheimer's disease amyloid plaques. *J Pathol J Pathol Soc G B Irel*. 2009;217(1):131-138.

1049 55. Itzhaki RF, Lin WR, Shang D, Wilcock GK, Faragher B, Jamieson GA. Herpes simplex virus
1050 type 1 in brain and risk of Alzheimer's disease. *The Lancet*. 1997;349(9047):241-244.

1051 56. Cairns DM, Rouleau N, Parker RN, Walsh KG, Gehrke L, Kaplan DL. A 3D human brain–like
1052 tissue model of herpes-induced Alzheimer's disease. *Sci Adv*. 2020;6(19):eaay8828.

1053 57. Duarte LF, Farías MA, Álvarez DM, Bueno SM, Riedel CA, González PA. Herpes simplex
1054 virus type 1 infection of the central nervous system: insights into proposed interrelationships
1055 with neurodegenerative disorders. *Front Cell Neurosci*. 2019;13:46.

1056 58. Ovsepian SV, O'Leary VB, Zaborszky L, Ntziachristos V, Dolly JO. Synaptic vesicle cycle
1057 and amyloid β: Biting the hand that feeds. *Alzheimers Dement*. 2018;14(4):502-513.

1058 59. Greber S, Lubec G, Cairns N, Fountoulakis M. Decreased levels of synaptosomal
1059 associated protein 25 in the brain of patients with Down syndrome and Alzheimer's disease.
1060 *Electrophor Int J*. 1999;20(4-5):928-934.

1061 60. Brinkmalm A, Brinkmalm G, Honer WG, et al. SNAP-25 is a promising novel cerebrospinal
1062 fluid biomarker for synapse degeneration in Alzheimer's disease. *Mol Neurodegener*.
1063 2014;9(1):1-13.

1064 61. Abdelhak A, Foschi M, Abu-Rumeileh S, et al. Blood GFAP as an emerging biomarker in
1065 brain and spinal cord disorders. *Nat Rev Neurol*. Published online 2022:1-15.

1066 62. Vranová HP, Hényková E, Mareš J, et al. Clusterin CSF levels in differential diagnosis of
1067 neurodegenerative disorders. *J Neurol Sci*. 2016;361:117-121.

1068 63. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies
1069 variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*.
1070 2009;41(10):1088-1093.

1071 64. Lambert JC, Heath S, Even G, et al. Genome-wide association study identifies variants at
1072 CLU and CR1 associated with Alzheimer's disease. *Nat Genet*. 2009;41(10):1094-1099.

1073 65. Moon HJ, Herring SK, Zhao L. Clusterin: a multifaceted protein in the brain. *Neural Regen*
1074 *Res*. 2021;16(7):1438.

1075 66. Foster EM, Dangla-Valls A, Lovestone S, Ribe EM, Buckley NJ. Clusterin in Alzheimer's
1076 disease: mechanisms, genetics, and lessons from other pathologies. *Front Neurosci*.
1077 2019;13:164.

1078 67. McCormick DA. GABA as an inhibitory neurotransmitter in human cerebral cortex. *J*
1079 *Neurophysiol*. 1989;62(5):1018-1027.

1080 68. Jiménez-Balado J, Eich TS. GABAergic dysfunction, neural network hyperactivity and
1081 memory impairments in human aging and Alzheimer's disease. In: *Seminars in Cell &*
1082 *Developmental Biology*. Elsevier; 2021.

1083 69. Pillsbury L, Oria M, Erdman J. Nutrition and traumatic brain injury: improving acute and
1084 subacute health outcomes in military personnel. Published online 2011.

1085 70. Platten M, Nollen EA, Röhrig UF, Fallarino F, Opitz CA. Tryptophan metabolism as a
1086 common therapeutic target in cancer, neurodegeneration and beyond. *Nat Rev Drug Discov*.
1087 2019;18(5):379-401.

1088 71. Yan X, Hu Y, Wang B, Wang S, Zhang X. Metabolic dysregulation contributes to the
1089 progression of Alzheimer's disease. *Front Neurosci*. 2020;14:1107.

1090 72. An Y, Varma VR, Varma S, et al. Evidence for brain glucose dysregulation in Alzheimer's
1091 disease. *Alzheimers Dement*. 2018;14(3):318-329.

1092 73. Wang X, He HJ, Xiong X, et al. NAD+ in Alzheimer's Disease: Molecular Mechanisms and
1093 Systematic Therapeutic Evidence Obtained in vivo. *Front Cell Dev Biol*. 2021;9:668491.

1094 74. González A, Calfío C, Churruca M, Maccioni RB. Glucose metabolism and AD: evidence for
1095 a potential diabetes type 3. *Alzheimers Res Ther*. 2022;14(1):1-11.

1096 75. He X, Huang Y, Li B, Gong CX, Schuchman EH. Deregulation of sphingolipid metabolism in
1097 Alzheimer's disease. *Neurobiol Aging*. 2010;31(3):398-408.

1098 76. Vickland H, Silver J. The role of astrocytes in axon guidance during development and repair.
1099 *Glial Cell Dev Basic Princ Clin Relev Jessen KR Richardson WD Eds*. Published online
1100 1996:197-208.

1101 77. Cheng J, Wang S, Dong Y, Yuan Z. The role and regulatory mechanism of hippo signaling
1102 components in the neuronal system. *Front Immunol*. 2020;11:281.

1103 78. Südhof TC. The synaptic vesicle cycle. *Annu Rev Neurosci*. 2004;27:509-547.

1104 79. Chanaday NL, Cousin MA, Milosevic I, Watanabe S, Morgan JR. The synaptic vesicle cycle
1105 revisited: new insights into the modes and mechanisms. *J Neurosci*. 2019;39(42):8209-
1106 8216.

1107 80. Castillo PE, Younts TJ, Chávez AE, Hashimotodani Y. Endocannabinoid signaling and
1108 synaptic function. *Neuron*. 2012;76(1):70-81.

1109 81. Dong A, Liu S, Li Y. Gap junctions in the nervous system: probing functional connections
1110 using new imaging approaches. *Front Cell Neurosci*. 2018;12:320.

1111 82. Blue EE, Thornton TA, Kooperberg C, et al. Non-coding variants in MYH11, FZD3, and
1112 SORCS3 are associated with dementia in women. *Alzheimers Dement*. 2021;17(2):215-225.

1113 83. Ou YN, Yang YX, Deng YT, et al. Identification of novel drug targets for Alzheimer's disease
1114 by integrating genetics and proteomes from brain and blood. *Mol Psychiatry*. Published
1115 online 2021:1-9.

1116 84. Carnazza KE, Komer LE, Xie YX, et al. Synaptic vesicle binding of α-synuclein is modulated
1117 by β-and γ-synucleins. *Cell Rep*. 2022;39(2):110675.

1118 85. Vekrellis K, Rideout HJ, Stefanis L. Neurobiology of alpha-synuclein. *Mol Neurobiol*.
1119 2004;30(1):1-21. doi:10.1385/MN:30:1:001

1120 86. Lee HJ, Patel S, Lee SJ. Intravesicular localization and exocytosis of alpha-synuclein and its
1121 aggregates. *J Neurosci Off J Soc Neurosci*. 2005;25(25):6016-6024.
1122 doi:10.1523/JNEUROSCI.0692-05.2005

1123 87. Yamada K, Iwatsubo T. Extracellular α-synuclein levels are regulated by neuronal activity.
1124 *Mol Neurodegener*. 2018;13(1):1-8.

1125 88. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene
1126 identified in families with Parkinson's disease. *Science*. 1997;276(5321):2045-2047.
1127 doi:10.1126/science.276.5321.2045

1128 89. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein
1129 in Lewy bodies. *Nature*. 1997;388(6645):839-840. doi:10.1038/42166

1130 90. Masliah E, Rockenstein E, Veinbergs I, et al. Dopaminergic loss and inclusion body
1131 formation in alpha-synuclein mice: implications for neurodegenerative disorders. *Science*.
1132 2000;287(5456):1265-1269. doi:10.1126/science.287.5456.1265

1133 91. Singleton AB, Farrer M, Johnson J, et al. alpha-Synuclein locus triplication causes
1134 Parkinson's disease. *Science*. 2003;302(5646):841. doi:10.1126/science.1090278

1135 92. Giasson BI, Duda JE, Quinn SM, Zhang B, Trojanowski JQ, Lee VMY. Neuronal alpha-
1136 synucleinopathy with severe movement disorder in mice expressing A53T human alpha-
1137 synuclein. *Neuron*. 2002;34(4):521-533. doi:10.1016/s0896-6273(02)00682-7

1138 93. Lee MK, Stirling W, Xu Y, et al. Human α-synuclein-harboring familial Parkinson's disease-
1139 linked Ala-53→ Thr mutation causes neurodegenerative disease with α-synuclein
1140 aggregation in transgenic mice. *Proc Natl Acad Sci*. 2002;99(13):8968-8973.

1141 94. Luk KC, Kehm V, Carroll J, et al. Pathological α-synuclein transmission initiates Parkinson-
1142 like neurodegeneration in nontransgenic mice. *Science*. 2012;338(6109):949-953.
1143 doi:10.1126/science.1227157

1144 95. Davis AA, Inman CE, Wargel ZM, et al. APOE genotype regulates pathology and disease
1145 progression in synucleinopathy. *Sci Transl Med*. 2020;12(529):eaay3069.

1146 96. Parkkinen L, Soininen H, Alafuzoff I. Regional distribution of α-synuclein pathology in
1147 unimpaired aging and Alzheimer disease. *J Neuropathol Exp Neurol*. 2003;62(4):363-367.

1148 97. Lippa CF, Fujiwara H, Mann DM, et al. Lewy bodies contain altered α-synuclein in brains of
1149 many familial Alzheimer's disease patients with mutations in presenilin and amyloid
1150 precursor protein genes. *Am J Pathol*. 1998;153(5):1365-1370.

1151 98. Lippa CF, Schmidt ML, Lee VMY, Trojanowski JQ. Antibodies to α-synuclein detect Lewy
1152 bodies in many Down's syndrome brains with Alzheimer's disease. *Ann Neurol Off J Am*
1153 *Neurol Assoc Child Neurol Soc*. 1999;45(3):353-357.

1154 99. Sekiya M, Wang M, Fujisaki N, et al. Integrated biology approach reveals molecular and
1155 pathological interactions among Alzheimer's Aβ42, Tau, TREM2, and TYROBP in
1156 Drosophila models. *Genome Med*. 2018;10(1):1-20.

1157 100. Iacono D, Resnick SM, O'Brien R, et al. Mild cognitive impairment and asymptomatic
1158 Alzheimer disease subjects: equivalent β-amyloid and tau loads with divergent cognitive
1159 outcomes. *J Neuropathol Exp Neurol*. 2014;73(4):295-304.

1160 101. Martorana A, Koch G. Is dopamine involved in Alzheimer's disease? *Front Aging Neurosci*.
1161 2014;6:252.

1162 102. Koch G, Di Lorenzo F, Bonnì S, et al. Dopaminergic modulation of cortical plasticity in
1163 Alzheimer's disease patients. *Neuropsychopharmacology*. 2014;39(11):2654-2661.

1164 103. Nobili A, Latagliata EC, Viscomi MT, et al. Dopamine neuronal loss contributes to memory
1165 and reward dysfunction in a model of Alzheimer's disease. *Nat Commun*. 2017;8(1):1-14.

1166 104. Cordella A, Krashia P, Nobili A, et al. Dopamine loss alters the hippocampus-nucleus
1167 accumbens synaptic transmission in the Tg2576 mouse model of Alzheimer's disease.
1168 *Neurobiol Dis*. 2018;116:142-154.

1169 105. Ceyzériat K, Gloria Y, Tsartsalis S, et al. Alterations in dopamine system and in its
1170 connectivity with serotonin in a rat model of Alzheimer's disease. *Brain Commun*.
1171 2021;3(2):fcab029.

1172 106. Koch G, Esposito Z, Codecà C, et al. Altered dopamine modulation of LTD-like plasticity in
1173 Alzheimer's disease patients. *Clin Neurophysiol*. 2011;122(4):703-707.

1174 107. Guzmán-Ramos K, Moreno-Castilla P, Castro-Cruz M, et al. Restoration of dopamine
1175 release deficits during object recognition memory acquisition attenuates cognitive
1176 impairment in a triple transgenic mice model of Alzheimer's disease. *Learn Mem*.
1177 2012;19(10):453-460.

1178 108. Perez SE, Lazarov O, Koprich JB, et al. Nigrostriatal dysfunction in familial Alzheimer's
1179 disease-linked APPswe/PS1ΔE9 transgenic mice. *J Neurosci*. 2005;25(44):10220-10229.

1180 109. Ambrée O, Richter H, Sachser N, et al. Levodopa ameliorates learning and memory
1181 deficits in a murine model of Alzheimer's disease. *Neurobiol Aging*. 2009;30(8):1192-1204.

1182 110. Bukke VN, Archana M, Villani R, et al. The dual role of glutamatergic neurotransmission in
1183 Alzheimer's disease: from pathophysiology to pharmacotherapy. *Int J Mol Sci*.
1184 2020;21(20):7452.

1185 111. Conway ME. Alzheimer's disease: targeting the glutamatergic system. *Biogerontology*.
1186 2020;21(3):257-274.

1187 112. Rodriguez-Perdigon M, Tordera RM, Gil-Bea FJ, Gerenu G, Ramirez MJ, Solas M. Down-
1188 regulation of glutamatergic terminals (VGLUT1) driven by Aβ in Alzheimer's disease.
1189 *Hippocampus*. 2016;26(10):1303-1312.

1190 113. Cheng YJ, Lin CH, Lane HY. Involvement of cholinergic, adrenergic, and glutamatergic
1191 network modulation with cognitive dysfunction in Alzheimer's disease. *Int J Mol Sci*.
1192 2021;22(5):2283.

1193 114. Colangelo AM, Alberghina L, Papa M. Astrogliosis as a therapeutic target for
1194 neurodegenerative diseases. *Neurosci Lett*. 2014;565:59-64.

1195 115. Abu-Rumeileh S, Steinacker P, Polischi B, et al. CSF biomarkers of neuroinflammation in
1196 distinct forms and subtypes of neurodegenerative dementia. *Alzheimers Res Ther*.
1197 2020;12(1):1-15.

1198 116. Ishiki A, Kamada M, Kawamura Y, et al. Glial fibrillar acidic protein in the cerebrospinal
1199 fluid of Alzheimer's disease, dementia with Lewy bodies, and frontotemporal lobar
1200 degeneration. *J Neurochem*. 2016;136(2):258-261.

1201 117. Jesse S, Steinacker P, Cepek L, et al. Glial fibrillary acidic protein and protein S-100B:
1202 different concentration pattern of glial proteins in cerebrospinal fluid of patients with
1203 Alzheimer's disease and Creutzfeldt-Jakob disease. *J Alzheimers Dis*. 2009;17(3):541-
1204 551.

1205 118. Oeckl P, Halbgebauer S, Anderl-Straub S, et al. Glial fibrillary acidic protein in serum is
1206 increased in Alzheimer's disease and correlates with cognitive impairment. *J Alzheimers*
1207 *Dis*. 2019;67(2):481-488.

1208 119. Wallin A, Blennow K, Rosengren LE. Glial fibrillary acidic protein in the cerebrospinal fluid
1209 of patients with dementia. *Dement Geriatr Cogn Disord*. 1996;7(5):267-272.

1210 120. Cicognola C, Janelidze S, Hertze J, et al. Plasma glial fibrillary acidic protein detects
1211 Alzheimer pathology and predicts future conversion to Alzheimer dementia in patients with
1212 mild cognitive impairment. *Alzheimers Res Ther*. 2021;13(1):1-9.

1213 121. Chatterjee P, Pedrini S, Stoops E, et al. Plasma glial fibrillary acidic protein is elevated in
1214 cognitively normal older adults at risk of Alzheimer's disease. *Transl Psychiatry*.
1215 2021;11(1):1-10.

122. DeMattos RB, Cirrito JR, Parsadanian M, et al. ApoE and clusterin cooperatively suppress Aβ levels and deposition: evidence that ApoE regulates extracellular Aβ metabolism in vivo. *Neuron*. 2004;41(2):193-202.

123. Zandl-Lang M, Fanaee-Danesh E, Sun Y, et al. Regulatory effects of simvastatin and apoJ on APP processing and amyloid-β clearance in blood-brain barrier endothelial cells. *Biochim Biophys Acta BBA-Mol Cell Biol Lipids*. 2018;1863(1):40-60.

124. Nuutinen T, Huuskonen J, Suuronen T, Ojala J, Miettinen R, Salminen A. Amyloid-β 1–42 induced endocytosis and clusterin/apoJ protein accumulation in cultured human astrocytes. *Neurochem Int*. 2007;50(3):540-547.

125. Yeh FL, Wang Y, Tom I, Gonzalez LC, Sheng M. TREM2 binds to apolipoproteins, including APOE and CLU/APOJ, and thereby facilitates uptake of amyloid-beta by microglia. *Neuron*. 2016;91(2):328-340.

126. Robbins JP, Perfect L, Ribe EM, et al. Clusterin is required for β-amyloid toxicity in human iPSC-derived neurons. *Front Neurosci*. 2018;12:504.

127. Killick R, Ribe EM, Al-Shawi R, et al. Clusterin regulates β-amyloid toxicity via Dickkopf-1-driven induction of the wnt–PCP–JNK pathway. *Mol Psychiatry*. 2014;19(1):88-98.

128. Sobue A, Komine O, Hara Y, 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.

129. Puntambekar SS, Moutinho M, Lin PBC, et al. CX3CR1 deficiency aggravates amyloid driven neuronal pathology and cognitive decline in Alzheimer's disease. *Mol Neurodegener*. 2022;17(1):47. doi:10.1186/s13024-022-00545-9

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

131. Hansen LA. The Lewy body variant of Alzheimer disease. In: *Dementia in Parkinsonism*. Springer; 1997:83-93.

132. Hamilton RL. Lewy bodies in Alzheimer's disease: a neuropathological review of 145 cases using α-synuclein immunohistochemistry. *Brain Pathol*. 2000;10(3):378-384.

133. Arai Y, Yamazaki M, Mori O, Muramatsu H, Asano G, Katayama Y. α-Synuclein-positive structures in cases with sporadic Alzheimer's disease: morphology and its relationship to tau aggregation. *Brain Res*. 2001;888(2):287-296.

134. Cairns NJ, Perrin RJ, Franklin EE, et al. Neuropathologic assessment of participants in two multi-center longitudinal observational studies: The A lzheimer D isease N euroimaging I nitiative (ADNI) and the D ominantly I nherited A lzheimer N etwork (DIAN). *Neuropathology*. 2015;35(4):390-400.

1253 135. Bassil F, Brown HJ, Pattabhiraman S, et al. Amyloid-beta (Aβ) plaques promote seeding
1254 and spreading of alpha-synuclein and tau in a mouse model of Lewy body disorders with
1255 Aβ pathology. *Neuron*. 2020;105(2):260-275.

1256 136. Bassil F, Meymand ES, Brown HJ, et al. α-Synuclein modulates tau spreading in mouse
1257 brains. *J Exp Med*. 2021;218(1).

1258 137. Henderson MX, Trojanowski JQ, Lee VMY. α-Synuclein pathology in Parkinson's disease
1259 and related α-synucleinopathies. *Neurosci Lett*. 2019;709:134316.

1260 138. Volpicelli-Daley LA, Luk KC, Patel TP, et al. Exogenous α-synuclein fibrils induce Lewy
1261 body pathology leading to synaptic dysfunction and neuron death. *Neuron*. 2011;72(1):57-
1262 71.

1263 139. Wu Q, Shaikh MA, Meymand ES, et al. Neuronal activity modulates alpha-synuclein
1264 aggregation and spreading in organotypic brain slice cultures and in vivo. *Acta*
1265 *Neuropathol (Berl)*. 2020;140(6):831-849.

1266 140. Fortin DL, Nemani VM, Voglmaier SM, Anthony MD, Ryan TA, Edwards RH. Neural
1267 activity controls the synaptic accumulation of α-synuclein. *J Neurosci*. 2005;25(47):10913-
1268 10921.

1269 141. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for
1270 biomarker discovery. *Nat Preced*. Published online 2010:1-1.

1271 142. Del-Aguila JL, Li Z, Dube U, et al. A single-nuclei RNA sequencing study of Mendelian and
1272 sporadic AD in the human brain. *Alzheimers Res Ther*. 2019;11(1):1-16.

1273 143. Brase L, You SF, Del-Aguila JL, et al. A landscape of the genetic and cellular
1274 heterogeneity in Alzheimer disease. *medRxiv*. Published online 2021.

1275 144. Johnson EC, Carter EK, Dammer EB, et al. Large-scale deep multi-layer analysis of
1276 Alzheimer's disease brain reveals strong proteomic disease-related changes not observed
1277 at the RNA level. *Nat Neurosci*. 2022;25(2):213-225.

1278 145. Bai B, Wang X, Li Y, et al. Deep multilayer brain proteomics identifies molecular networks
1279 in Alzheimer's disease progression. *Neuron*. 2020;105(6):975-991.

1280 146. Arnold M, Nho K, Kueider-Paisley A, et al. Sex and APOE ε4 genotype modify the
1281 Alzheimer's disease serum metabolome. *Nat Commun*. 2020;11(1):1-12.

1282 147. Toledo JB, Arnold M, Kastenmueller G, et al. Metabolic network failures in Alzheimer's
1283 disease: a biochemical road map. *Alzheimers Dement*. 2017;13(9):965-984.

1284 148. John-Williams S, Blach C, Toledo JB, et al. Targeted metabolomics and medication
1285 classification data from participants in the ADNI1 cohort. *Sci Data*. 2017;4(1):1-11.

1286 149. Batra R, Arnold M, Wörheide MA, et al. The landscape of metabolic brain alterations in
1287 Alzheimer's disease. *bioRxiv*. Published online 2021.

150. Andrews S. FastQC: a quality control tool for high throughput sequence data. Published online 2010.

151. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.

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

153. Novotny BC, Fernandez MV, Bahena J, et al. Metabolomic and lipidomic signatures in autosomal dominant and late-onset Alzheimer disease brains. *medRxiv*. Published online 2021.

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

155. Hastie T, Tibshirani R, Narasimhan B, Chu G. impute: impute: Imputation for microarray data. Published online 2018.

156. 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):882-883.

157. Zheng GX, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8(1):1-12.

158. Mo Q, Shen R. iClusterPlus: integrative clustering of multiple genomic data sets. Published online 2013.

159. Gaujoux R, Seoighe C. CellMix: a comprehensive toolbox for gene expression deconvolution. *Bioinformatics*. 2013;29(17):2211-2212.

160. Gaujoux R, Seoighe C. Semi-supervised Nonnegative Matrix Factorization for gene expression deconvolution: a case study. *Infect Genet Evol*. 2012;12(5):913-921.

161. 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.

162. 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.

163. Therneau T. A Package for Survival Analysis in R. R package version 3.2-3. Published online 2020. https://CRAN.R-project.org/package=survival

164. Eklund A. surviplot: Draw Survival Curves with Number-at-Risk. R package version 1.1.1. Published online 2017. http://www.cbs.dtu.dk/~eklund/surviplot/

165. Team RDC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *URL Httpwww Rproject Org*. Published online 2008.

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

167. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021;49(W1):W388-W396.

Acknowledgements

We thank all the participants and their families, as well as the many involved institutions and their staff, whose help and participation made this work possible. 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: NIA R01AG057777 (OH), K01AG046374 (CMK), R56AG067764 (OH), 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). This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

ROSMAP (RNA-seq)

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

data collection was supported through funding by NIA grants P30AG10161, 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 and were provided by Dr. Eric Schadt from Mount Sinai School of Medicine. Proteomics data were provided by Dr. Levey 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.

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary Figures 1-19

Supplementary Tables 1–9

1384 **Tables**

1385

1386 **Table 1**: **Demographic characteristics of AD and control subjects.**

Study	N	Tissue[a]	Data[b]	AD			Control		
				Female (%)	Age (mean,sd)	APOE ε4+ (%)	Female (%)	Age (mean,sd)	APOE ε4+ (%)
Knight ADRC	278	PC	T,P,M	60	83.8(8.6)	58.04	73.9	88.6(8.8)	0
MSBB	217	BM36	T,P	67.1	84.4(6.7)	19.29	66.6	80.6(8.3)	18.5
		BM10	T	67.0	84.2(6.8)	20.73	69.7	81.3(7.0)	15.1
		BM22	T	62.9	83.6(6.9)	23.38	70	81(8.0)	16.6
		BM44	T	66.6	84.3(6.7)	20.92	64.2	80.5(8.3)	17.8
ROSMAP	237	DLPFC	T,M	70.8	90.7(5.9)	39.58	53.7	84.1(6.8)	10.7

1387 [a]PC (parietal cortex), BM10 (frontal pole), BM22 (superior temporal gyrus), BM44 (inferior frontal gyrus), DLPFC
1388 (dorsolateral prefrontal cortex).
1389 [b]Data availability, T, transcriptomics; P, proteomics; M, metabolomics.

1390 **Table 2: Summary of the significant features detected between each cluster and the**
1391 **control.**

Cluster	RNA-seq			Proteomics			Metabolomics		
	Total	Specific[a]	Missed[b]	Total	Specific	Missed	Total	Specific	Missed
1	385	67 (17%)	108 (28%)	71	24 (34%)	59(83%)	0	0(0%)	0(0%)
2	560	102 (18%)	312 (56%)	456	232 (51%)	448 (98%)	1	1(100%)	1(100%)
3	356	36 (10%)	137 (38%)	4	0 (0%)	1(25%)	0	0(0%)	0(0%)
4	4661	4210 (90%)	4295 (92%)	485	250 (52%)	475(98%)	90	90(100%)	90(100%)

1392 [a] features detected only in the specific cluster.
1393 [b] indicates the number and percentage of features missed when un-clustered AD cases were compared to the
1394 control.

1395 **Table 3: Number of significant features between each cluster and the other clusters.**

Cluster	Transcriptomics	Proteomics	Metabolomics
1	678	499	105
2	22	683	0
3	1	413	0
4	6822	562	228

1396

1397 **Table 4: Number and percentage of significant features specific to each cell type.**

Cell-type	Upregulated				Downregulated			
	C1	C2	C3	C4	C1	C2	C3	C4
Astrocyte	0(0%)	1(3%)	0(0%)	214(38%)	6(40%)	21 (75%)	5(38%)	11 (3%)
Microglia	3(19%)	5(14%)	0(0%)	56(10%)	1(7%)	1 (4%)	2(15%)	112 (31%)
Neuron	3(19%)	2(6%)	2(22%)	25(4%)	1(7%)	1 (4%)	0 (0%)	108 (30%)
Oligodendrocyte	3(19%)	21(60%)	5(56%)	85(15%)	0(0%)	1 (4%)	0 (0%)	77 (21%)
OPC	0(0%)	3(9%)	1(11%)	50(9%)	6(40%)	2 (7%)	1(8%)	48 (13%)
Endothelial	7(44%)	3(9%)	1(11%)	130(23%)	1(7%)	2 (7%)	5(38%)	8 (2%)

1398

Figure Legends

Fig. 1: Study design. The transcriptomics, proteomics, metabolomics and lipidomics profiles of postmortem parietal cortex samples from the Knight ADRC participants were used. 278 subjects (255 sporadic AD and 23 control cases) were shared across the three omics. Expression/reading matrices were prepared for the same set of samples (n=278) and different number of features (genes=60,754, proteins=1,092, metabolites=627). Prior to the integration, feature selection was performed on all omics datasets whose number of features exceeded 1000 features by selecting the top most variant features. A Bayesian integrative clustering method (iClusterBayes) was then employed to integrate the three omics datasets and the best clustering solution was extracted based on the maximum deviance ratio and the minimum Bayesian information criterion (BIC). AD molecular profiles were then linked to multiple clinical and molecular attributes to examine whether an association with these attributes exists. In addition, multiple analyses including survival, differential expression, cell-type specific effect, and cell proportions inference analyses were performed on these profiles. These molecular profiles were then replicated in two independent datasets from MSBB BM36 (AD=93, CO=23) and ROSMAP (AD=144, CO=93).

Fig. 2: Cross-omics data integration identified four distinct molecular profiles of AD associated with worse clinical outcomes and molecular attributes. (a) Heatmaps of multi-omics profiles of the top significant features. **(b)** Transcriptomic profiles **(**scaled) of the top genes that contributed significantly to the clustering, showing an overall dysregulation of brains in Knight-C4. (**c**) The distribution of CDR scores across four clusters with Knight-C4 associated with high CDR scores. (**d**) Kaplan-Meier plots showing Knight-C4 association with poor outcome for age of onset and age of death. (**e**) Boxplots showing the cell proportion from deconvolution analysis across four clusters using bulk RNA-seq from four cell types. Knight-C4 is associated with significantly higher and lower proportion of astrocytes and neurons respectively. (**f**) Boxplot showing the first principal component of metabolomics profiles for each cluster with Knight-C2 and Knight-C4 showing a lower metabolomic signature than other AD cases.

Fig. 3: Molecular profiles of Knight-C4 are replicated in two independent datasets. (**a**) Heatmaps of the transcriptomic and proteomic profiles of the top features from MSBB (BM36) showing two distinct clusters. (**b**) Boxplots showing MSBB-C1 associated with higher CDR scores, replicating Knight-C4 in the knight ADRC. (**c**) Boxplots showing cell proportions inferred from bulk RNA-seq from MSBB (BM36) using digital deconvolution. MSBB-C1 replicates Knight-C4 by showing an association with significantly higher and lower proportions of astrocytes and neurons

respectively. (**d**) Heatmap of the transcriptomic profiles of the top features from ROSMAP (DLPFC) showing two distinct clusters. (**e**) Boxplots showing ROSMAP-C1 associated with lower MMSE30 scores. (**f**) Boxplots showing cell proportion estimated from bulk RNA-seq from ROSMAP (DLPFC) using digital deconvolution. Similar to MSBB, ROSMAP-C1 replicates Knight-C4 by showing an association with significantly higher and lower proportions of astrocytes and neurons respectively.

Fig. 4: The parietal cortex of subjects with worse cognitive function exhibited a remarkable molecular dysregulation. (**a**) Bar plots showing the significant genes detected in each cluster when compared to the control. Yellow color represents the total number of genes, green represents the cluster-specific genes, and the purple depicts the number of genes missed with un-clustered approach (all AD cases vs control). (**b**) Same as panel "a" but for proteins. (**c**) Same as panel "a" but for metabolites. (**d**) Race track plot showing the percentage of significant features for each cluster when compared to other clusters for the three omics**.** (**e**) Venn diagrams showing the overlap between significant genes in Knight-C4 and synaptic genes from Fei *et al.,* 2018[38]. (**f**) Same as "e" but overlapped with SynGO dataset[37]. The significance of overlap was computed using Fisher's exact test. (**g**) Top 20 KEGG pathways associated with Knight-C4**.** (**h**) Top 20 GO biological process pathways associated with Knight-C4.

Fig. 5: Differential expression analyses identified neurodegenerative disease genes associated with distinct clusters. (**a**) Volcano plot showing the up-and downregulated genes identified between Knight-C4 vs other clusters. Genes in red are examples of known neurodegenerative disease genes. Genes in black are the top 10 significant genes based on the adjusted p-value and genes with log2(fold change) > 1. (**b**) Boxplots showing the transcriptomic profiles (in FPKM) of *SNAP25* across four clusters (right) and for AD cases combined (left) in the Knight ADRC cohort. (**c**) Same as "b" but for *GFAP*. (**d**) Volcano plot showing the up-and downregulated proteins identified between Knight-C2 vs other clusters. (**e**) Boxplots showing the proteomic profiles of *CLU* across four clusters (right) and for all AD cases (left) in the Knight ADRC cohort. (**f**) Kaplan-Meier plot showing the association of *CLU* downregulation with an increased risk of dementia progression using CDR change over time using CSF dataset from the Knight ADRC participants.

Fig. 6: Differential abundance analyses identified significant metabolites and pathways associated with Knight-C4. (**a**) Venn diagrams showing the significant increased and decreased

metabolites shared between the Knight-C4 and ROSMAP-C1 when compared to the control. Left and right boxes show the names of shared metabolites. (**b**) Similar to panel "a" but for metabolites shared between Knight-C4 and ROSMAP-C1 when compared to the other cases. The significant of overlap was computed using Fisher's exact test.

Fig. 7: The integration of single-nuclei data with cross-omics profiles identified cell-type specific genes and pathways. (**a**) Bar plots showing the percentage of up-and downregulated genes (clusters vs control) identified in each cell type. (**b**) Top 20 KEGG pathways enriched in astrocytic upregulated genes. (**c**) Top 20 GO biological process pathways enriched in neuronal downregulated genes. (**d**) Top KEGG pathways enriched in neuronal downregulated genes. (**e**) Boxplots showing the transcriptomic profiles (in FPKM) of examples of cell-type specific genes identified in the Knight ADRC cohort. (**f**) The effect size of these examples generated from single-nuclei data from the Knight ADRC participants. Red bars represent the cell type in which the gene is overexpressed.

Fig. 8: Cross-omics integration identified alpha-synuclein levels downregulated in AD participants with worse cognitive function common across multiple brain regions. (**a**) Transcriptomic profiles (in FPKM) of *SNCA* across four clusters (right) and for all AD cases (left) in the Knight ADRC cohort. (**b**) Proteomic profiles of *SNCA* across four clusters (right) and all AD cases (left) in Knight ADRC cohort. (**c**) Transcriptomic profiles of *SNCA* across four clusters (right) and all AD cases (left) in MSBB BM36 cohort. (**d**) Proteomic (TMT) profiles of *SNCA* across four clusters (right) and all AD cases (left) in MSBB BM36 cohort. (**e**) Transcriptomic profiles of *SNCA* across four clusters (right) and all AD cases (left) in ROSMAP DLPFC cohort.

1492 **Supplementary Figures**

1493

1494 **Supplementary Fig. 1: The maximum deviance ratio (left) and the minimum Bayesian**
1495 **information criterion—BIC (right) results across cluster solutions.** The optimal solution is
1496 observed with cluster number *K=3* dividing our samples into four clusters (*k+1*).

1497

1498 **Supplementary Fig. 2: Knight-C4 is not associated with polygenic risk scores (PRS)**.
1499 Boxplots showing the polygenic risk scores (PRS) across four clusters using four different p-value
1500 thresholds showing no significant association between Knight-C4 and PRS scores.

1501

1502 **Supplementary Fig. 3: Knight-C4 is not associated with genetic or neuropathological**
1503 **effects.** (**a**) The distribution of APOE ε4 allele across four clusters showing no association with
1504 Knight-C4. (**b**) The distribution of Braak scores for tau across four clusters showing no significant
1505 correlation between Knight-C4 and Braak staging for tau. However, Knight-C4 exhibited more 6
1506 stage cases exceeding the mean distribution of Braak stage 6 (dashed line) across all clusters.
1507 (**d**) The distribution of Braak scores for amyloid-β across four clusters showing no association
1508 with Knight-C4.

1509

1510 **Supplementary Fig. 4: Dysregulated genes are significantly common across cohorts from**
1511 **multiple brain regions.** (**a**) Venn diagrams depicting the overlap between significant genes in
1512 Knight-C4 and MSBB-C1 cohorts, showing a replication of the molecular signature with worse
1513 cognitive outcomes. (**b**) Same as "a" but for significant proteins. (**c**) Venn diagrams depicting the
1514 overlap between significant genes in Knight-C4 and ROSMAP-C1 cohorts, showing a replication
1515 of the molecular signature with worse cognitive outcomes.

1516

1517 **Supplementary Fig. 5: Molecular profiles of MSBB-C1 are associated with high Braak**
1518 **scores for tau and replicated in additional brain regions from MSBB study.** (**a**) Boxplots
1519 showing MSBB-C1 is associated with high BraakTau scores. (**b**) Boxplots showing a clear pattern
1520 of an association with early age of death for cases in MSBB-C1. (**c**) Heatmaps showing the
1521 replication of transcriptomic profiles of MSBB-C1 from BM36 region in additional regions including
1522 frontal pole (BM10), superior temporal gyrus (BM22), and inferior frontal gyrus (BM44) brain
1523 regions from MSBB study.

1524

1525

1526 **Supplementary Fig. 6: Molecular profiles of ROSMAP-C1 are associated with multiple**
1527 **neuropathological variables. (a)** Boxplots showing ROSMAP-C1 association with high tangle
1528 densities. (**b**) Boxplots showing ROSMAP-C1 association with neurofibrillary tangle burdens. (**c**)
1529 Boxplots showing higher plaque densities for AD cases in ROSMAP-C1.

1530

1531 **Supplementary Fig. 7: Knight-C4 is associated with higher number of significant hits**
1532 **across multi-omics datasets**. (**a**) Venn diagrams showing the number of differentially expressed
1533 genes detected in each cluster (cluster vs control). (**b**) Same as panel "a" but for proteomics. (**c**)
1534 Same as panel "a" but for metabolomics.

1535

1536 **Supplementary Fig. 8: Significant enrichment of dysregulated synaptic genes in Knight-C4**
1537 **is replicated in additional brain regions.** (**a**) Venn Diagrams showing the overlap between
1538 significant genes in MSBB-C1 and synaptic genes from Fei *et al.,* 2018[38] and SynGO[37] datasets.
1539 (**b**) Same as "a" but overlapped with significant genes in ROSMAP-C1.

1540

1541 **Supplementary Fig. 9: Top significant pathways associated with Knight-C4 when brains**
1542 **are compared to the control**. (**a**) Top 20 KEGG pathways enriched in Knight-C4 downregulated
1543 genes. (**b**) Top 20 GO biological process pathways enriched in Knight-C4 downregulated genes.

1544

1545 **Supplementary Fig. 10: Top significant pathways associated with Knight-C4 that are**
1546 **enriched in downregulated proteins**. (**a**) Top 20 KEGG pathways associated with the
1547 downregulation of proteins in Knight-C4. (**b**) Top 20 GO biological process pathways enriched in
1548 Knight-C4 downregulated proteins.

1549

1550 **Supplementary Fig. 11: The decreased expression of *SNAP25* in Knight-C4 is replicated in**
1551 **MSBB and ROSMAP cohorts**. (**a**) Boxplots showing transcriptomic profiles of *SNAP25* across
1552 the two clusters (right) and all ADs (left) in MSBB (BM36) cohort. (**b**) Boxplots showing proteomic
1553 (TMT) profiles of *SNAP25* across the two clusters as mentioned in "a". (**c**) Boxplots showing
1554 transcriptomic profiles of *SNAP25* across the two clusters (right) and all ADs (left) in ROSMAP
1555 (DLPFC) cohort.

1556

1557 **Supplementary Fig. 12: *GFAP* proteomic levels are dysregulated across distinct clusters.**
1558 *GFAP* protomeric levels are significantly increased in Knight-C2 when compared to the control
1559 and other AD cases (right).
1560
1561 **Supplementary Fig. 13: The increased level of *GFAP* is replicated in MSBB and ROSMAP**
1562 **datasets**. (**a**) Boxplots showing the transcriptomic profiles of *GFAP* across the two clusters (right)
1563 and all AD cases (left) in MSBB (BM36) cohort. (**b**) Same as "a" but for the proteomic profiles
1564 (TMT). (**c**) Same as "a" but in ROSMAP (DLPFC) cohort.
1565
1566 **Supplementary Fig. 14: Differential abundance and pathways analyses identified**
1567 **metabolites decreased in Knight-C4 and ROSMAP-C1. (a**) Boxplots showing the decrease of
1568 metabolomics profiles of gamma-aminobutyrate (GABA) in Knight-C4 (left) and ROSMAP-C1
1569 (right). (**b**) Same as "a" but for choline metabolite.
1570
1571 **Supplementary Fig. 15: Differential abundance and pathway analyses identified**
1572 **metabolites increased in Knight-C4 and ROSMAP-C1**. **(a**) Boxplots showing the increase of
1573 metabolomic profiles of tryptophan in Knight-C4 (left) and ROSMAP-C1 (right). (**b**) Same as "a"
1574 but for glucose metabolite.
1575
1576 **Supplementary Fig. 16: Metabolic differential abundance and pathway analyses identified**
1577 **multiple dysregulated carbohydrate/sugar and sphingolipid metabolites. (a)** Boxplots
1578 showing the decrease of metabolomic profiles of citrate in Knight-C4 (left) and ROSMAP-C1. (**b**)
1579 Boxplots showing the decrease of metabolomic profiles of NAD+ in both Knight-C4 (left) and
1580 ROSMAP-C1 (right). (**d**) Boxplots showing the increased metabolism levels of sphingosine in both
1581 Knight-C4 (left) and ROSMAP-C1 (right).
1582
1583 **Supplementary Fig. 17: The integration of single-nuclei data with cross-omics profiles**
1584 **identified cell-type specific pathways. (a)** Top 20 KEGG pathways enriched in endothelial
1585 upregulated genes (Knight-C4 vs Control). (**b**) Top 20 KEGG pathways enriched in microglia
1586 downregulated genes (Knight-C4 vs Control).
1587
1588 **Supplementary Fig. 18: Downregulation of *SNCA* is replicated in additional brain regions**
1589 **from MSBB study.** (**a**) Boxplot showing the transcriptomic profiles of *SNCA* in MSBB-C1 from

1590 frontal pole (BM10) brain region. (**b**) Same as “a” but from superior temporal gyrus (BM22). (**d**)
1591 Same as “a” and “b” but from inferior frontal gyrus (BM44) brain region from MSBB.

1592

1593 **Supplementary Fig. 19: The synaptic dysregulation in Knight-C4 is recapitulated in some**
1594 **degree in mouse models. (a)** Venn diagram showing the overlap between significant genes in
1595 Knight-C4 when compared to other clusters, significant genes from NanoString gene expression
1596 data from A53T αSyn mouse models, and synaptic genes from SynGO dataset. (b) Same as “a”
1597 but using significant genes in Knight-C4 when compared to the control.

Fig. 1

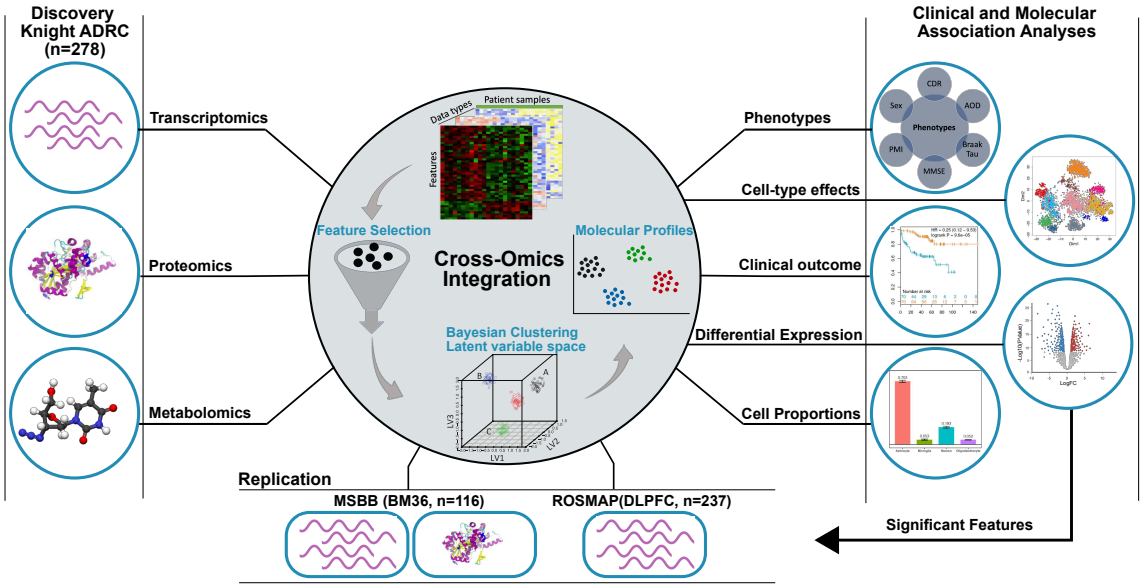

Discovery
Knight ADRC
(n=278)
Transcriptomics
Proteomics
Metabolomics
Data types
Patient samples
Features
Feature Selection
Cross-Omics
Integration
Molecular Profiles
Bayesian Clustering
Latent variable space
LV1
LV2
LV3
Clinical and Molecular
Association Analyses
Phenotypes
Cell-type effects
Clinical outcome
Differential Expression
Cell Proportions
CDR
Sex
AOD
Phenotypes
PMI
Braak
Tau
MMSE
Replication
MSBB (BM36, n=116)
ROSMAP(DLPFC, n=237)
Significant Features

Fig. 2

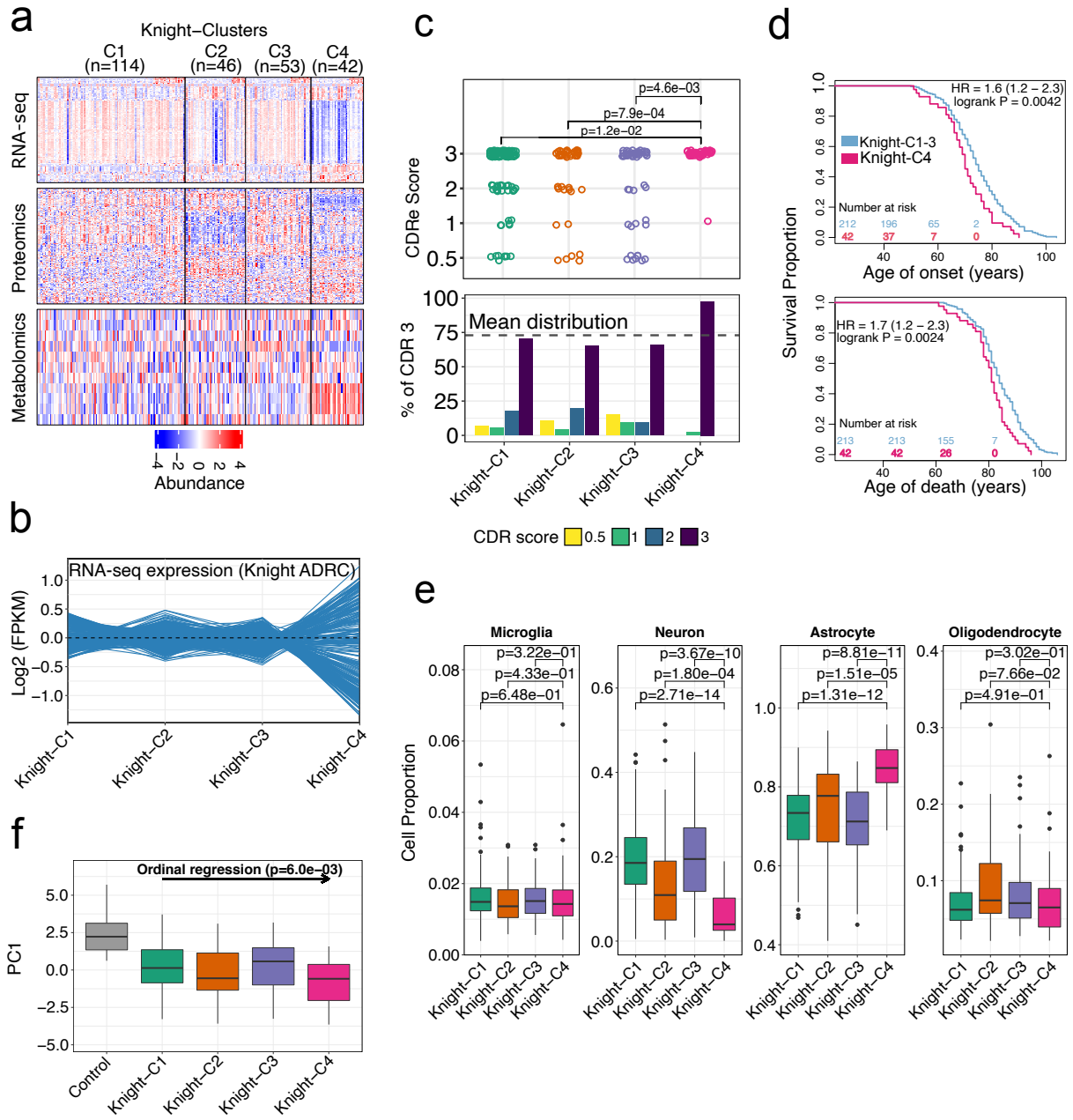

a
Knight−Clusters
C1 (n=114)
C2 (n=46)
C3 (n=53)
C4 (n=42)
RNA-seq
Proteomics
Metabolomics
Abundance
b
RNA-seq expression (Knight ADRC)
Log2 (FPKM)
Knight−C1
Knight−C2
Knight−C3
Knight−C4
c
CDRe Score
p=4.6e−03
p=7.9e−04
p=1.2e−02
% of CDR 3
Mean distribution
CDR score 0.5 1 2 3
d
HR = 1.6 (1.2 − 2.3)
logrank P = 0.0042
Knight-C1-3
Knight-C4
Number at risk
Age of onset (years)
Survival Proportion
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

Fig. 3

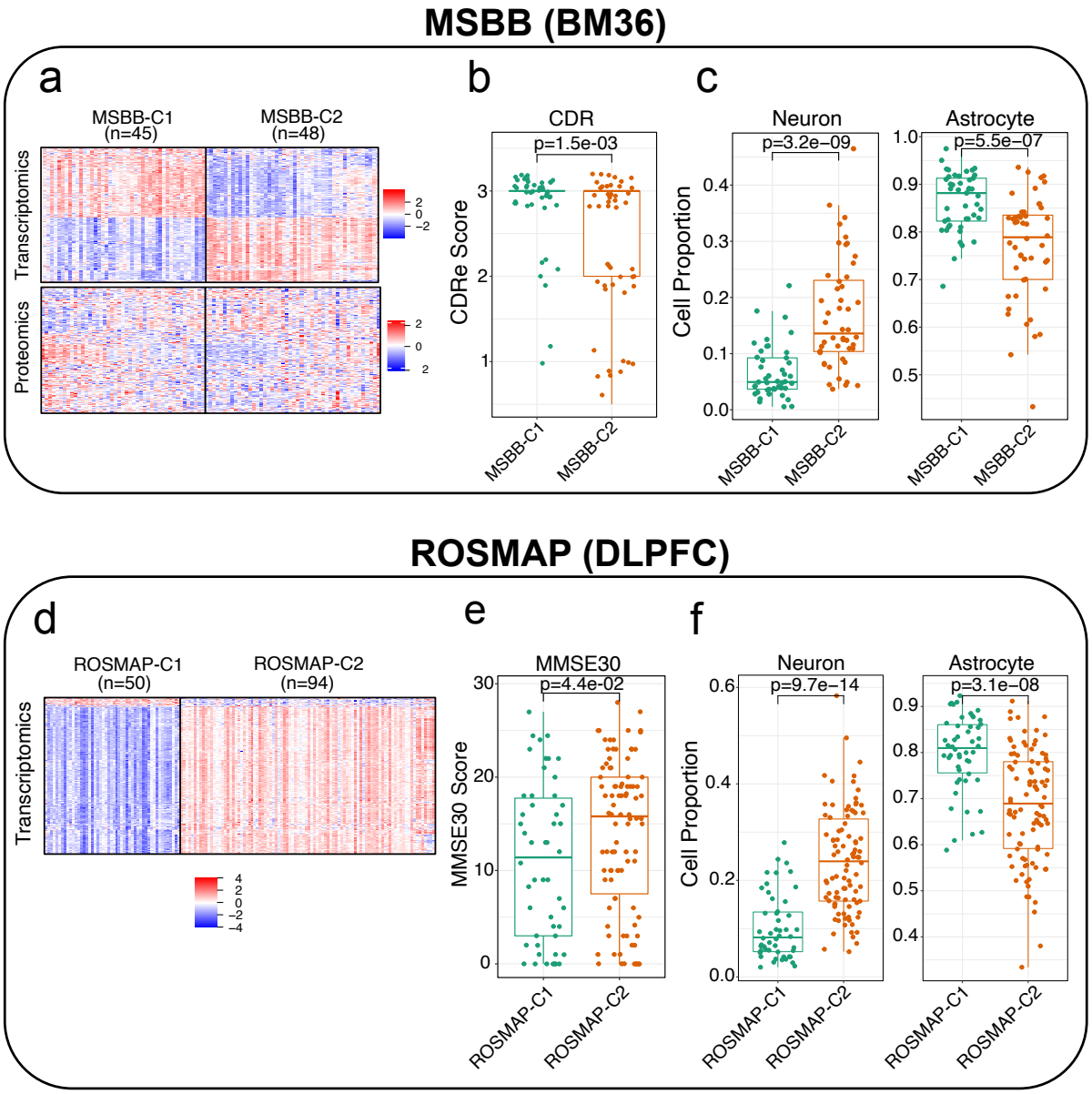

MSBB (BM36)
a
MSBB-C1
(n=45)
MSBB-C2
(n=48)
Transcriptomics
Proteomics
b
CDR
p=1.5e-03
CDRe Score
MSBB-C1
MSBB-C2
c
Neuron
p=3.2e−09
Astrocyte
p=5.5e−07
Cell Proportion
MSBB-C1
MSBB-C2
ROSMAP (DLPFC)
d
ROSMAP-C1
(n=50)
ROSMAP-C2
(n=94)
Transcriptomics
e
MMSE30
p=4.4e-02
MMSE30 Score
ROSMAP-C1
ROSMAP-C2
f
Neuron
p=9.7e−14
Astrocyte
p=3.1e−08
Cell Proportion
ROSMAP-C1
ROSMAP-C2

Fig. 4

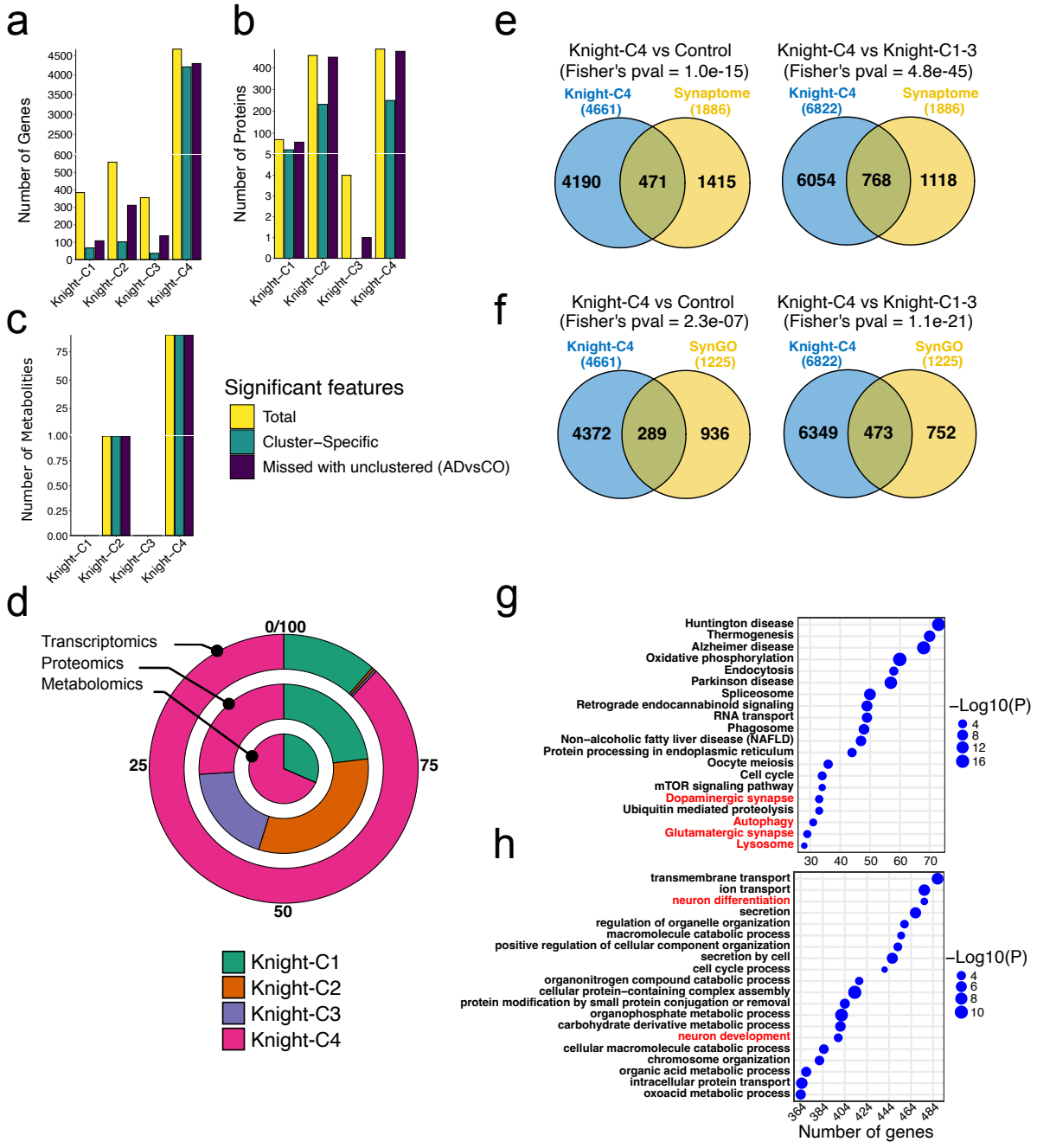

a
Number of Genes
Knight−C1
Knight−C2
Knight−C3
Knight−C4
b
Number of Proteins
c
Number of Metabolities
Significant features
Total
Cluster−Specific
Missed with unclustered (ADvsCO)
d
Transcriptomics
Proteomics
Metabolomics
0/100
25
75
50
Knight-C1
Knight-C2
Knight-C3
Knight-C4
e
Knight-C4 vs Control
(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
−Log10(P)
Number of genes

Fig. 5

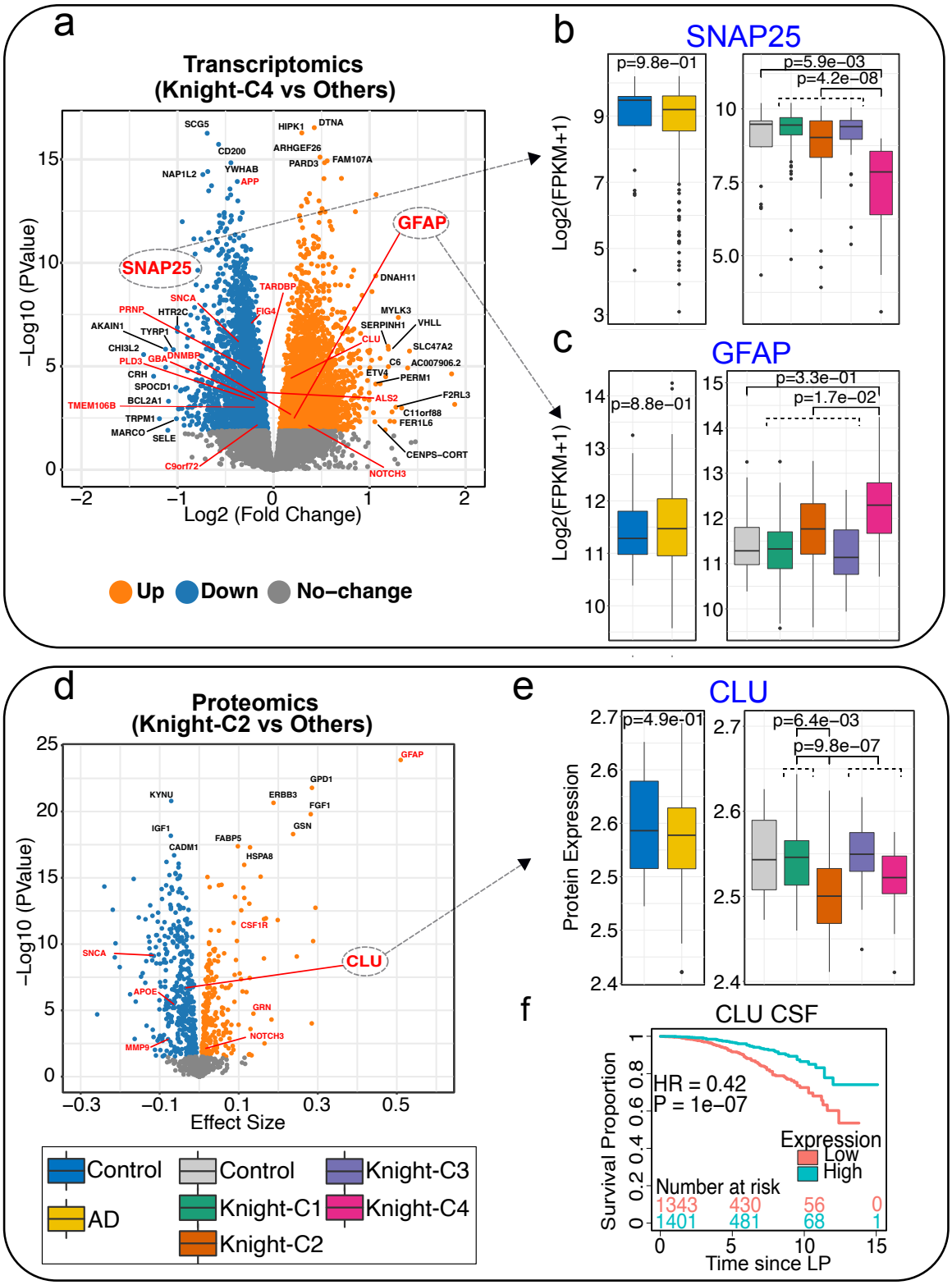

a
Transcriptomics
(Knight-C4 vs Others)
−Log10 (PValue)
Log2 (Fold Change)
SNAP25
GFAP
Up
Down
No−change
b
SNAP25
p=9.8e−01
p=5.9e−03
p=4.2e−08
Log2(FPKM+1)
c
GFAP
p=8.8e−01
p=3.3e−01
p=1.7e−02
d
Proteomics
(Knight-C2 vs Others)
Effect Size
CLU
e
CLU
p=4.9e-01
p=6.4e−03
p=9.8e−07
Protein Expression
f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

Fig. 6

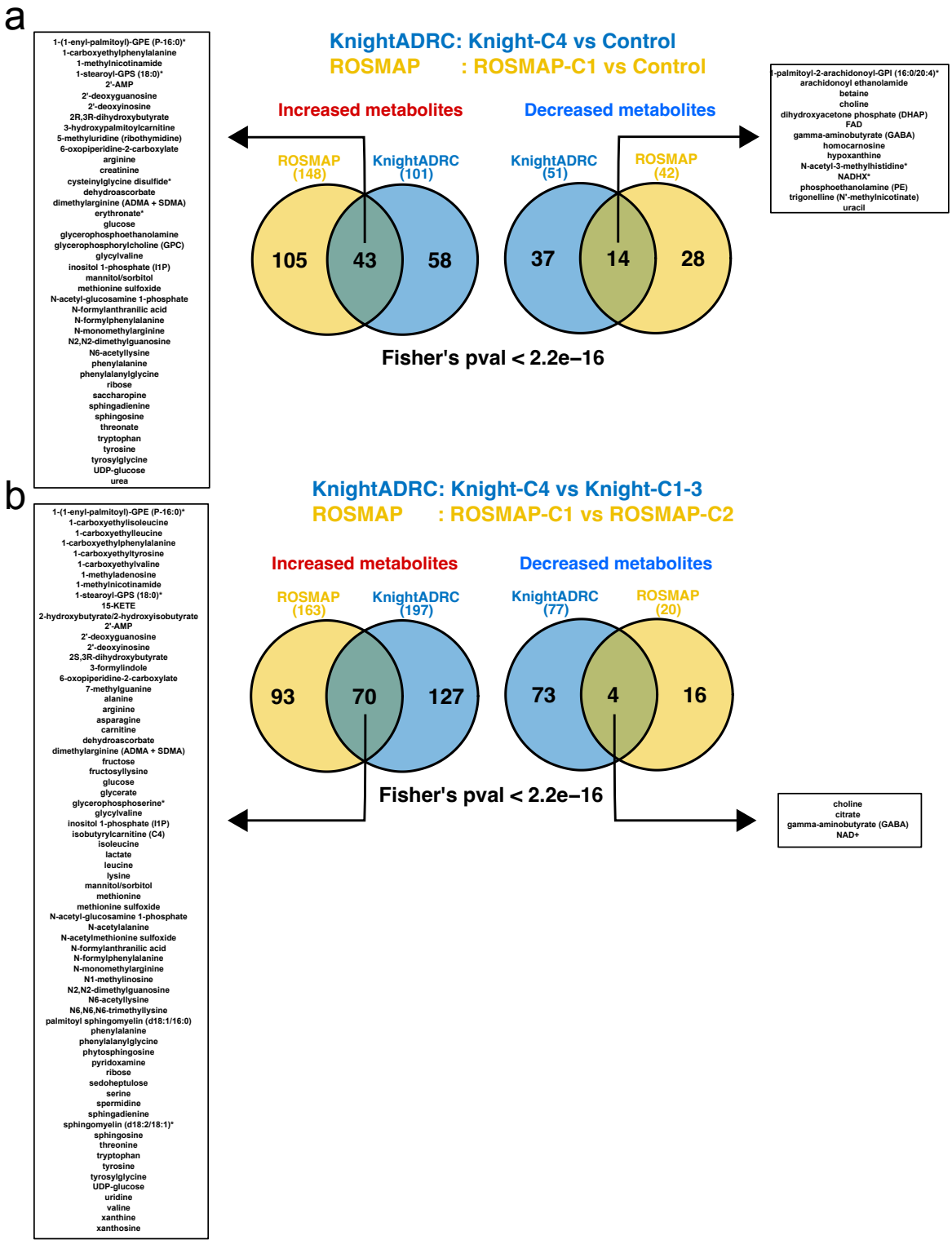

a
1-(1-enyl-palmitoyl)-GPE (P-16:0)*
1-carboxyethylphenylalanine
1-methylnicotinamide
1-stearoyl-GPS (18:0)*
2'-AMP
2'-deoxyguanosine
2'-deoxyinosine
2R,3R-dihydroxybutyrate
3-hydroxypalmitoylcarnitine
5-methyluridine (ribothymidine)
6-oxopiperidine-2-carboxylate
arginine
creatinine
cysteinylglycine 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
KnightADRC: Knight-C4 vs Control
ROSMAP : ROSMAP-C1 vs Control
Increased metabolites
Decreased metabolites
ROSMAP (148)
KnightADRC (101)
105
43
58
KnightADRC (51)
ROSMAP (42)
37
14
28
Fisher's pval < 2.2e−16
1-palmitoyl-2-arachidonoyl-GPI (16:0/20:4)*
arachidonoyl ethanolamide
betaine
choline
dihydroxyacetone phosphate (DHAP)
FAD
gamma-aminobutyrate (GABA)
homocarnosine
hypoxanthine
N-acetyl-3-methylhistidine*
NADHX*
phosphoethanolamine (PE)
trigonelline (N'-methylnicotinate)
uracil
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
KnightADRC: Knight-C4 vs Knight-C1-3
ROSMAP : ROSMAP-C1 vs ROSMAP-C2
Increased metabolites
Decreased metabolites
ROSMAP (163)
KnightADRC (197)
93
70
127
KnightADRC (77)
ROSMAP (20)
73
4
16
Fisher's pval < 2.2e−16
choline
citrate
gamma-aminobutyrate (GABA)
NAD+

Fig. 7

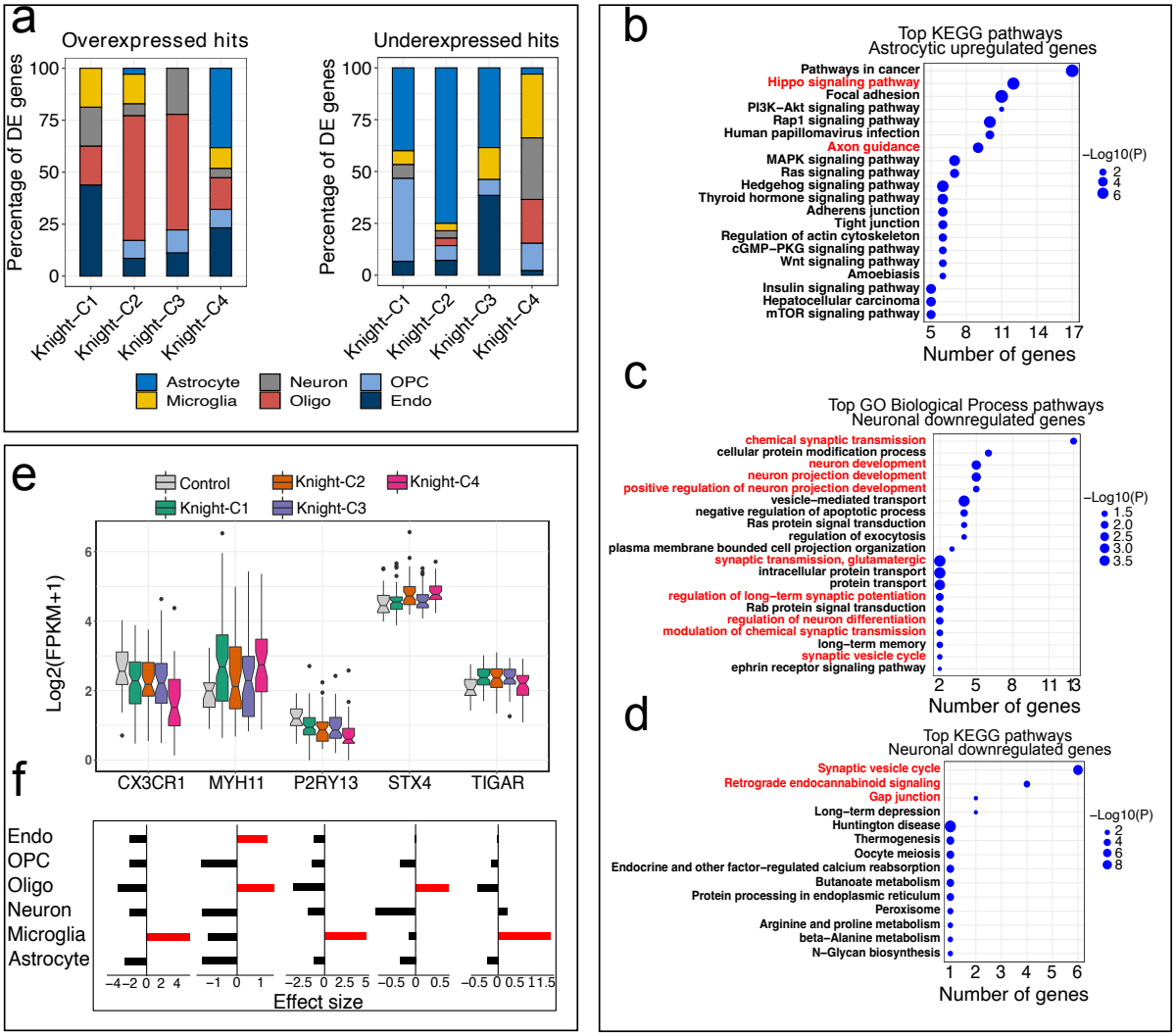
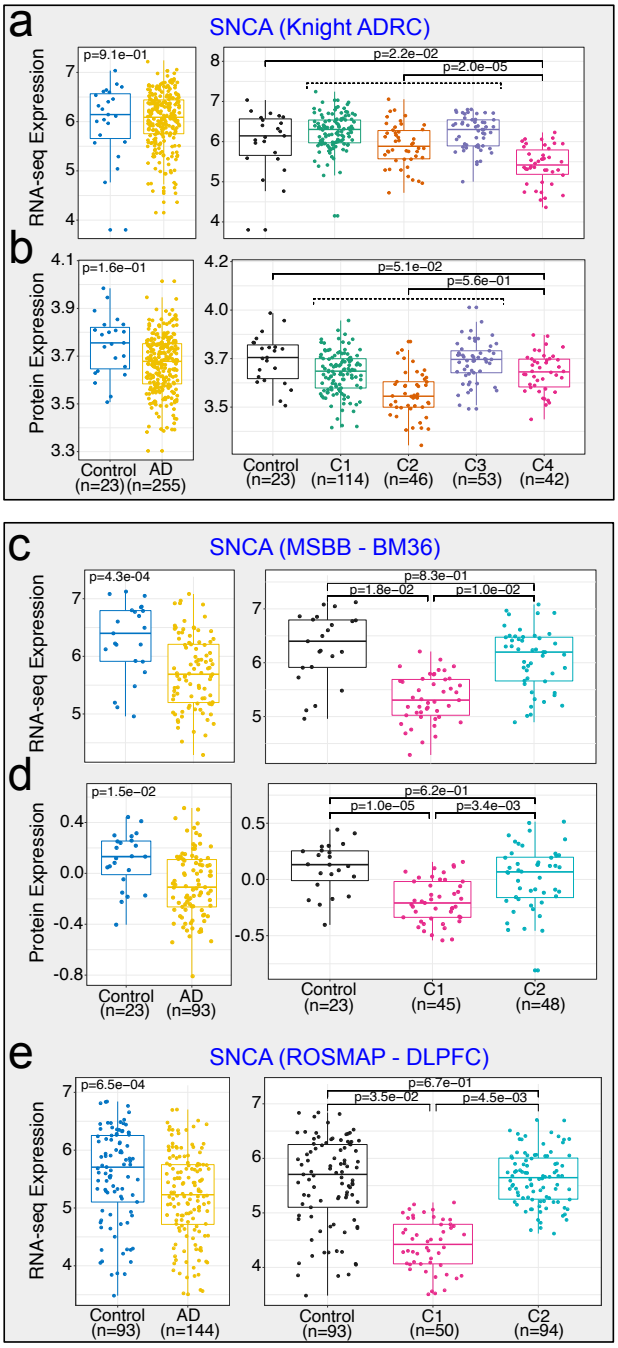
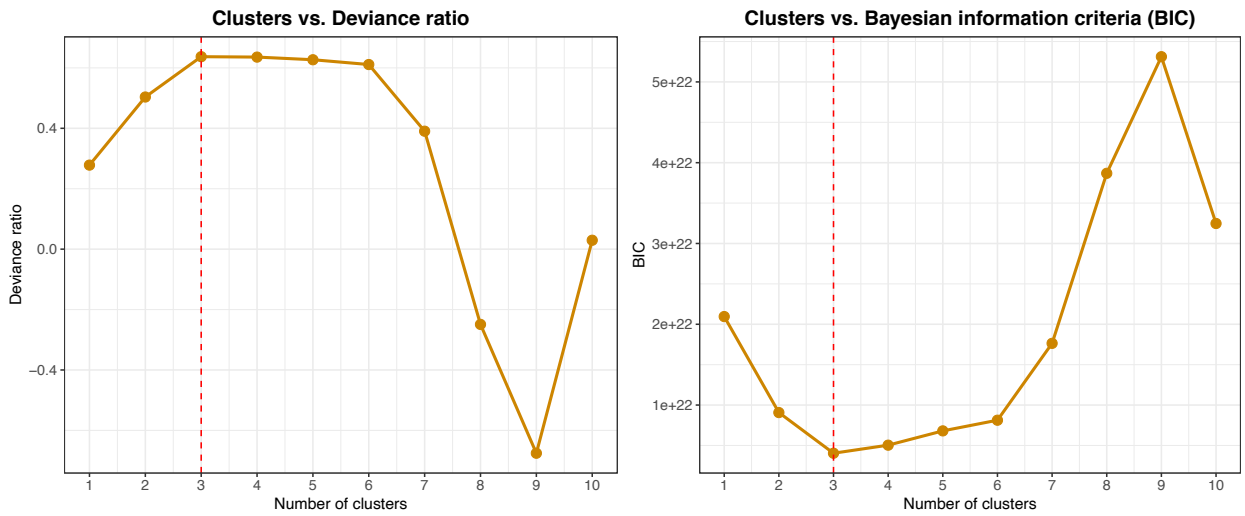

a
Overexpressed hits
Underexpressed hits
Percentage of DE genes
Knight−C1
Knight−C2
Knight−C3
Knight−C4
Astrocyte
Microglia
Neuron
Oligo
OPC
Endo
b
Top KEGG pathways
Astrocytic upregulated genes
Pathways in cancer
Hippo signaling pathway
Focal adhesion
PI3K−Akt signaling pathway
Rap1 signaling pathway
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

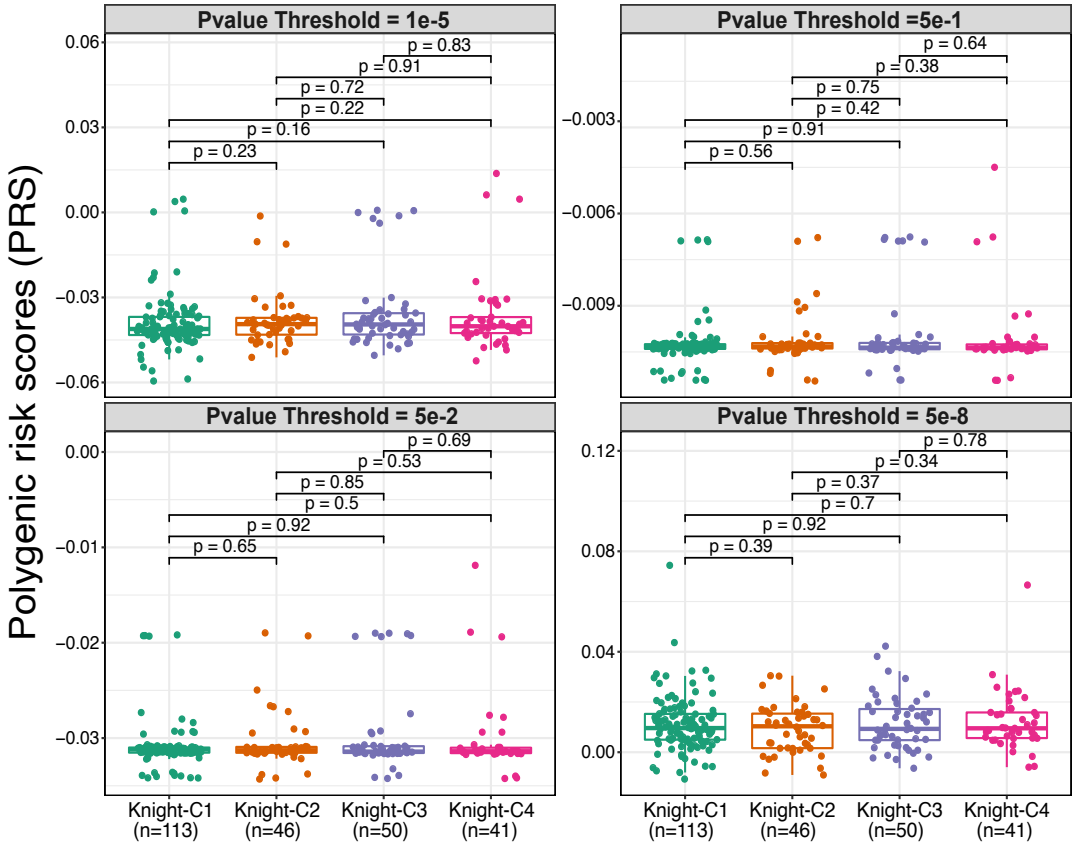
Fig. 8


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)

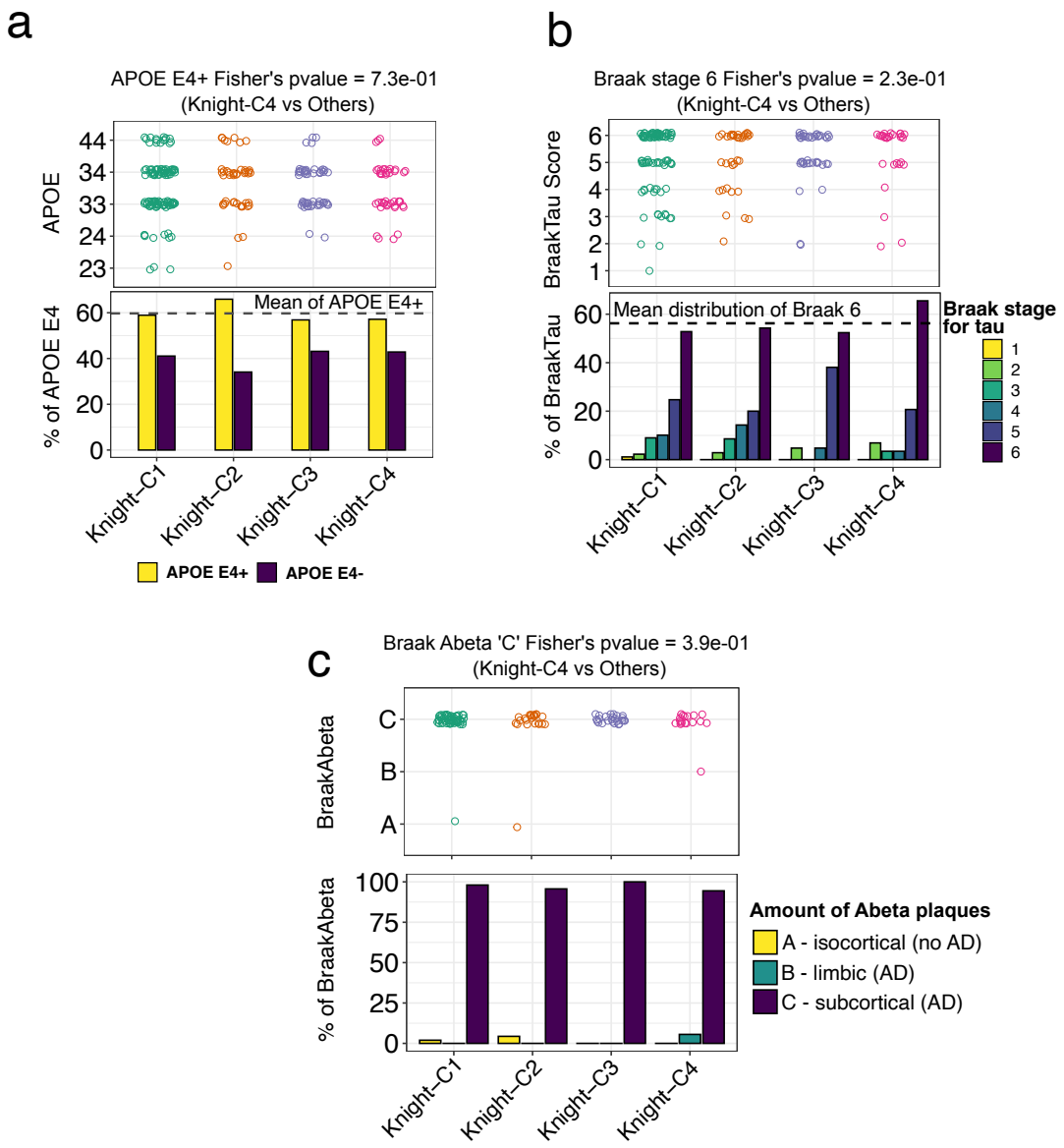
Supplementary Fig. 1



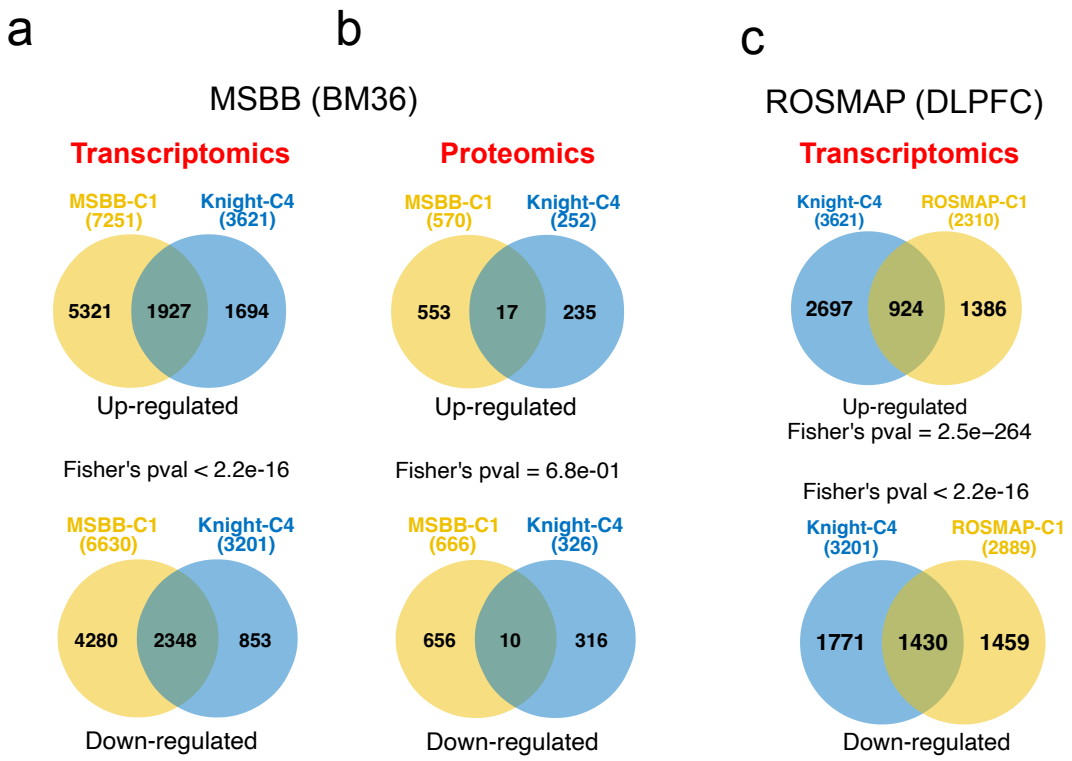
Supplementary Fig. 2



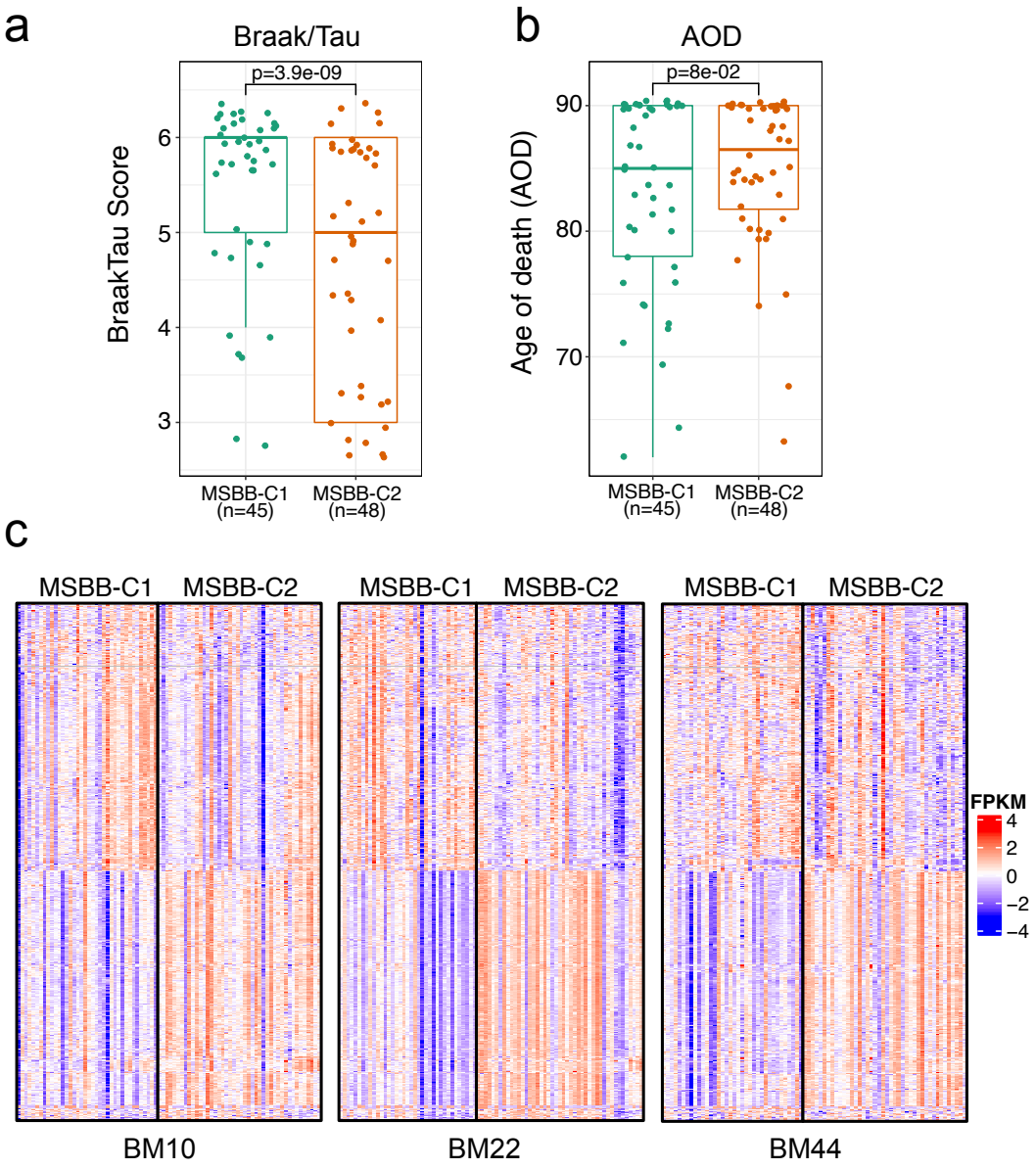
Supplementary Fig. 3



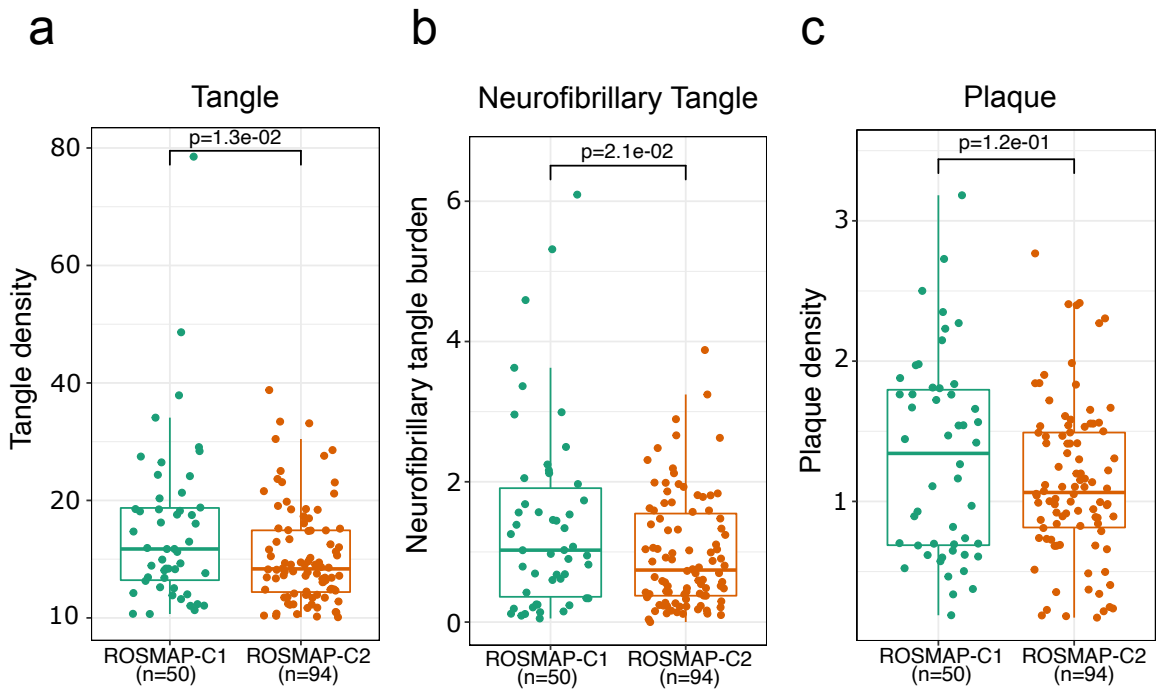
Supplementary Fig. 4



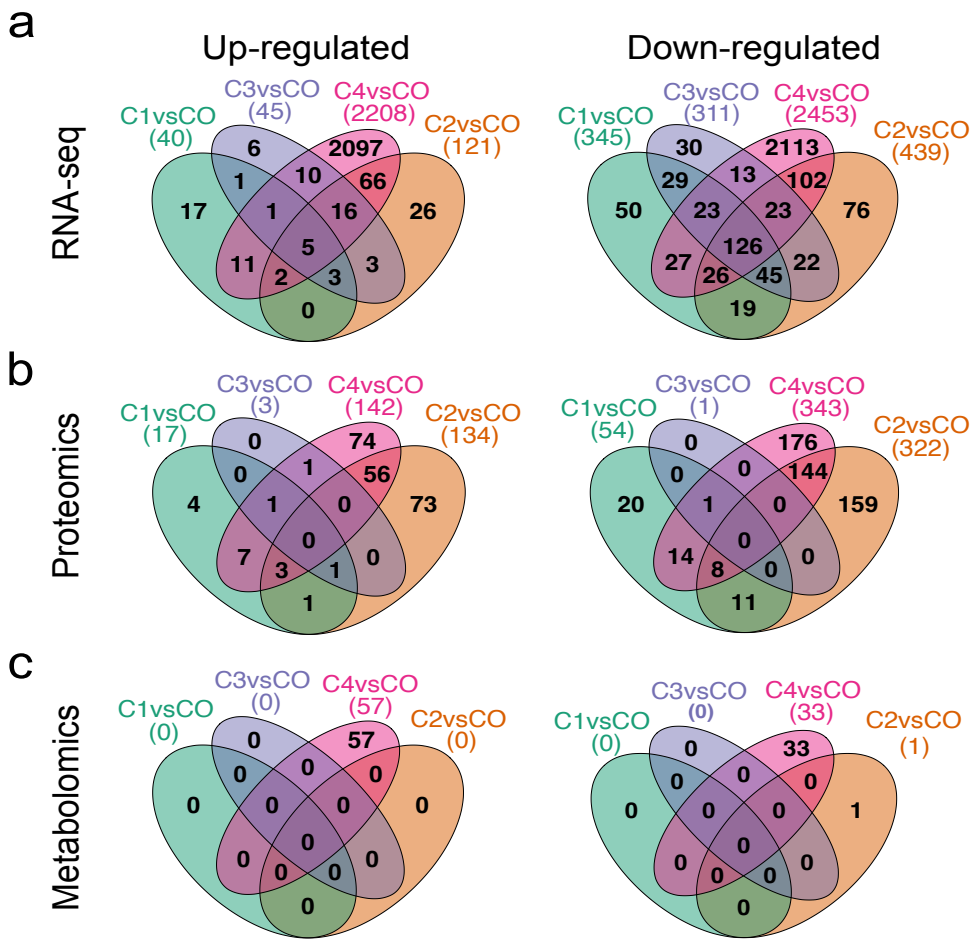
Supplementary Fig. 5



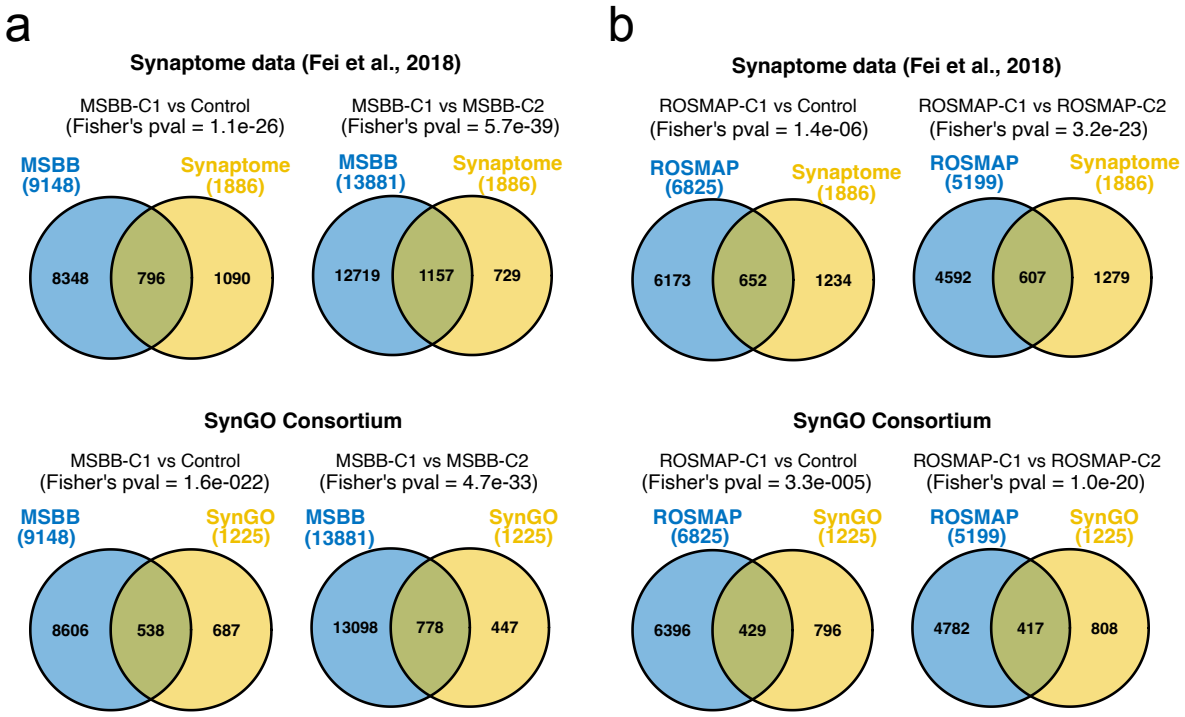
Supplementary Fig. 6



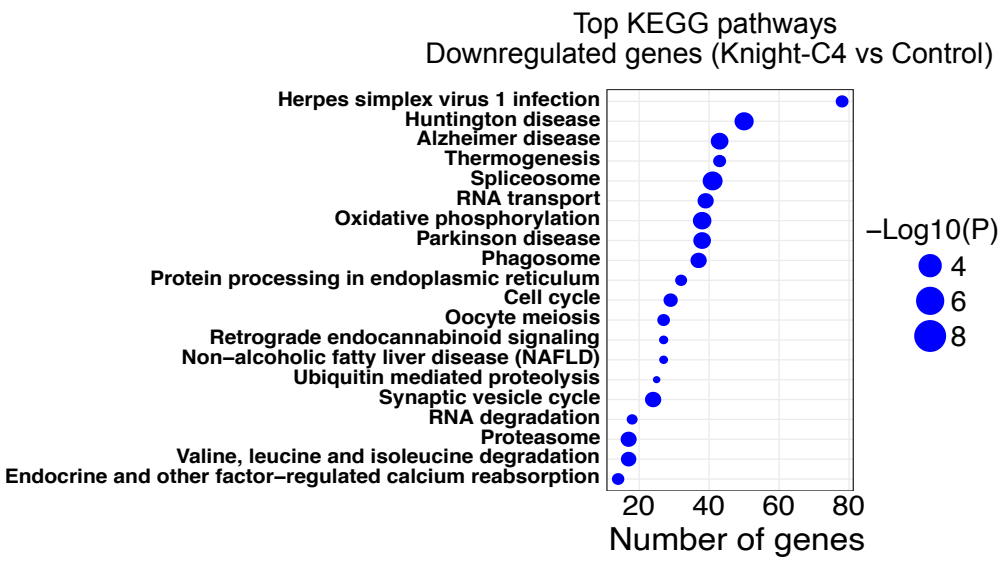
Supplementary Fig. 7

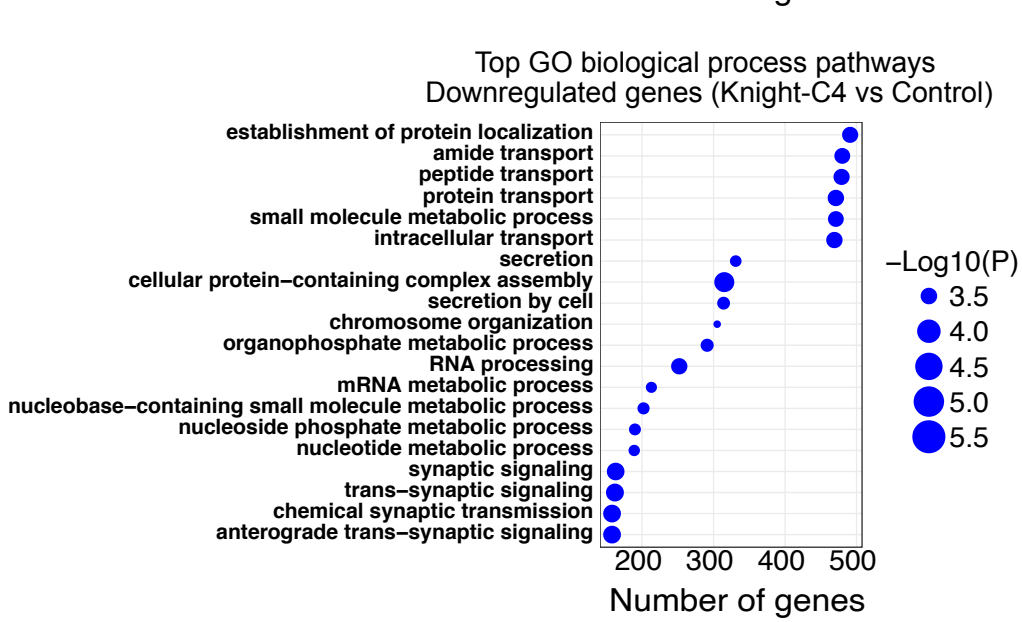


Supplementary Fig. 8



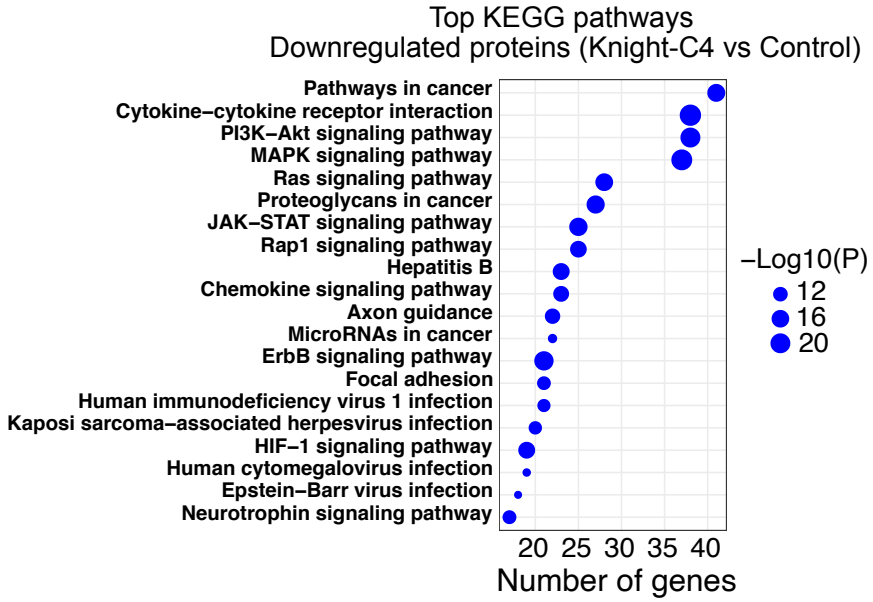
Supplementary Fig. 9



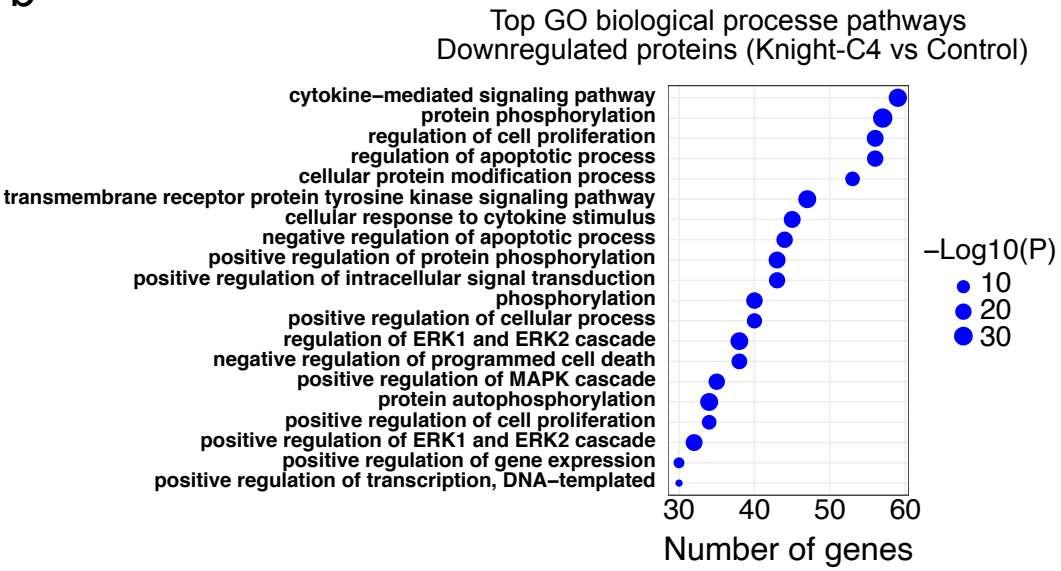


Supplementary Fig. 10

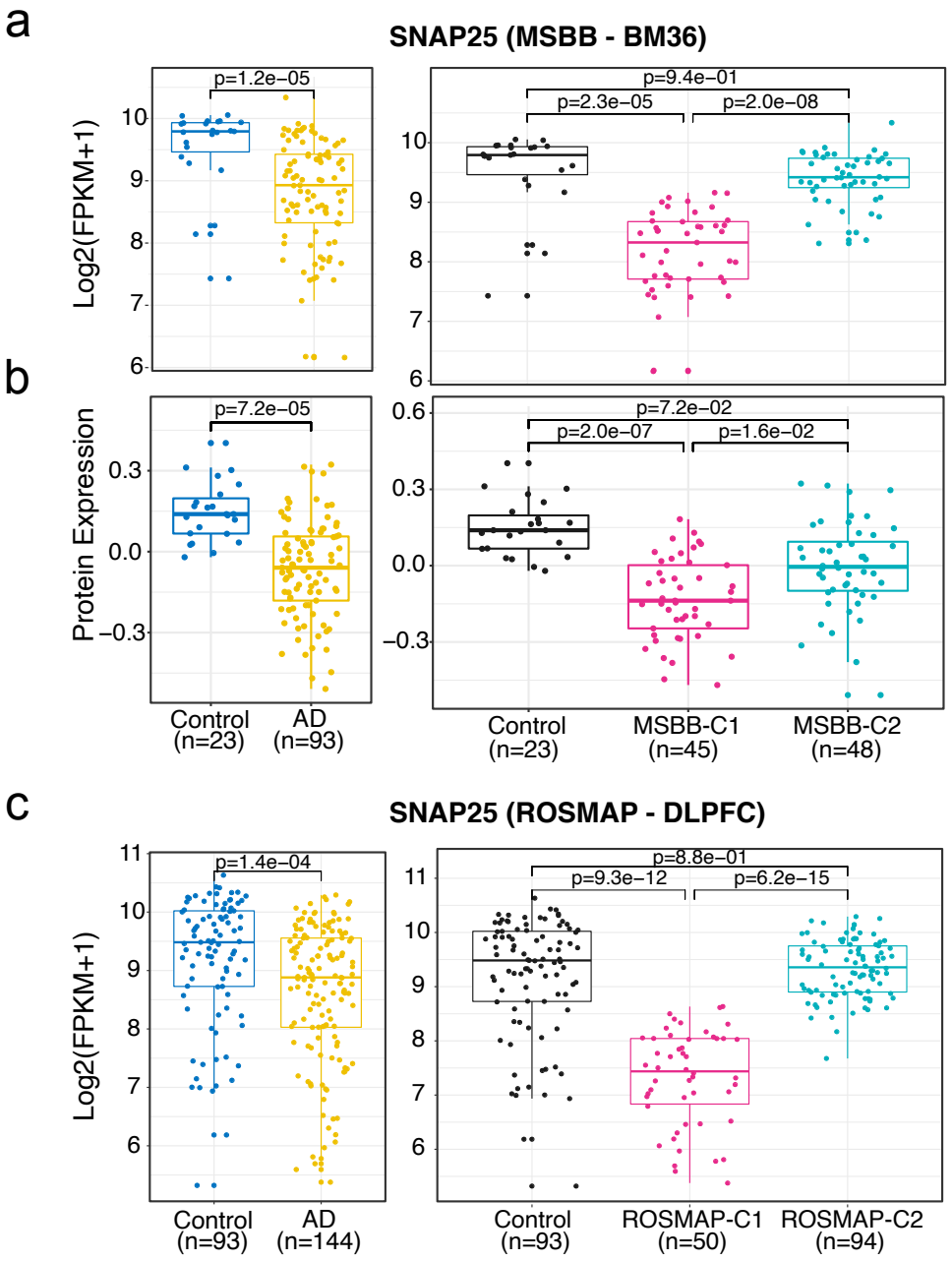
a



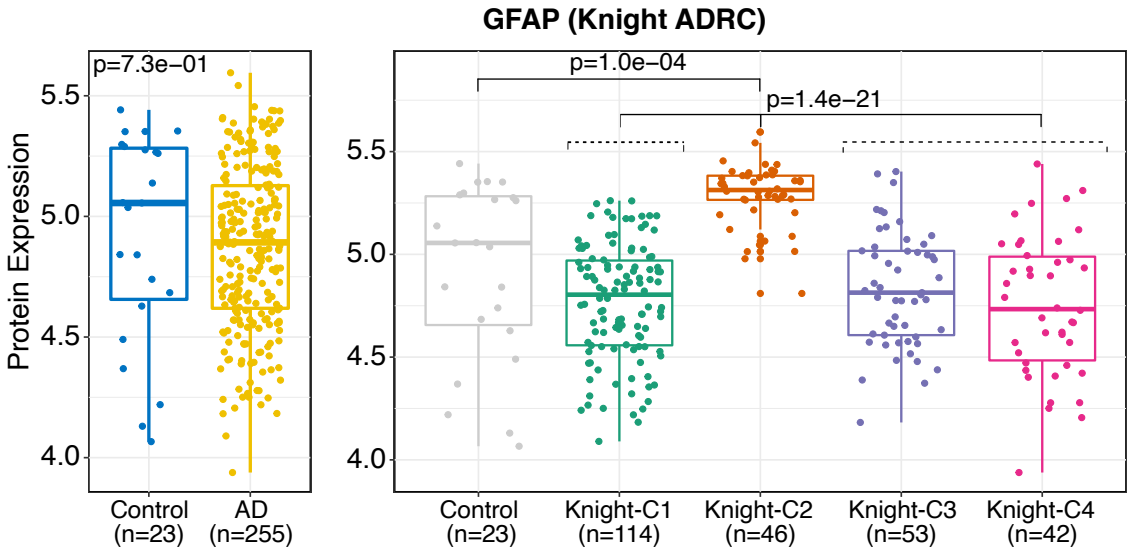
b



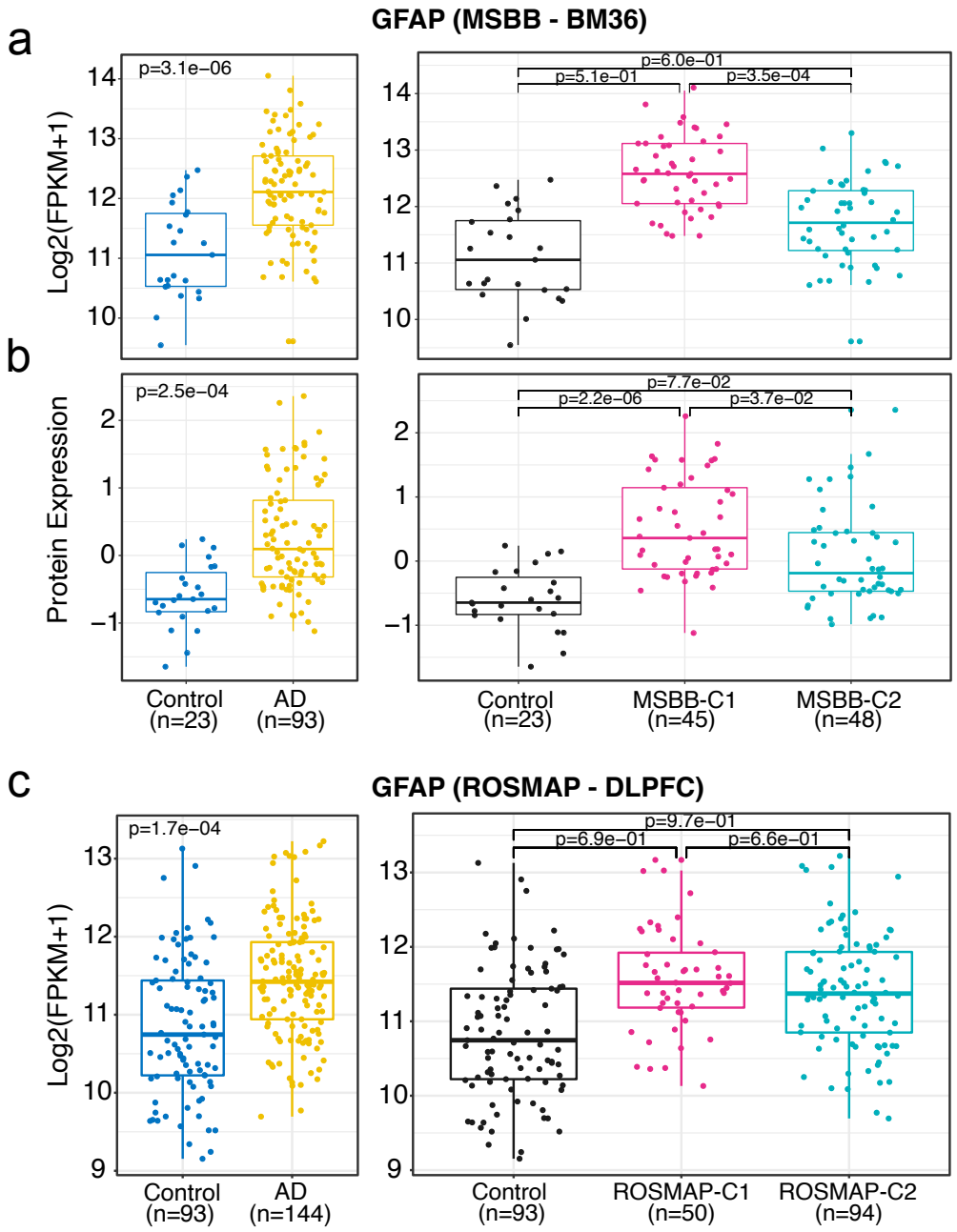
Supplementary Fig. 11



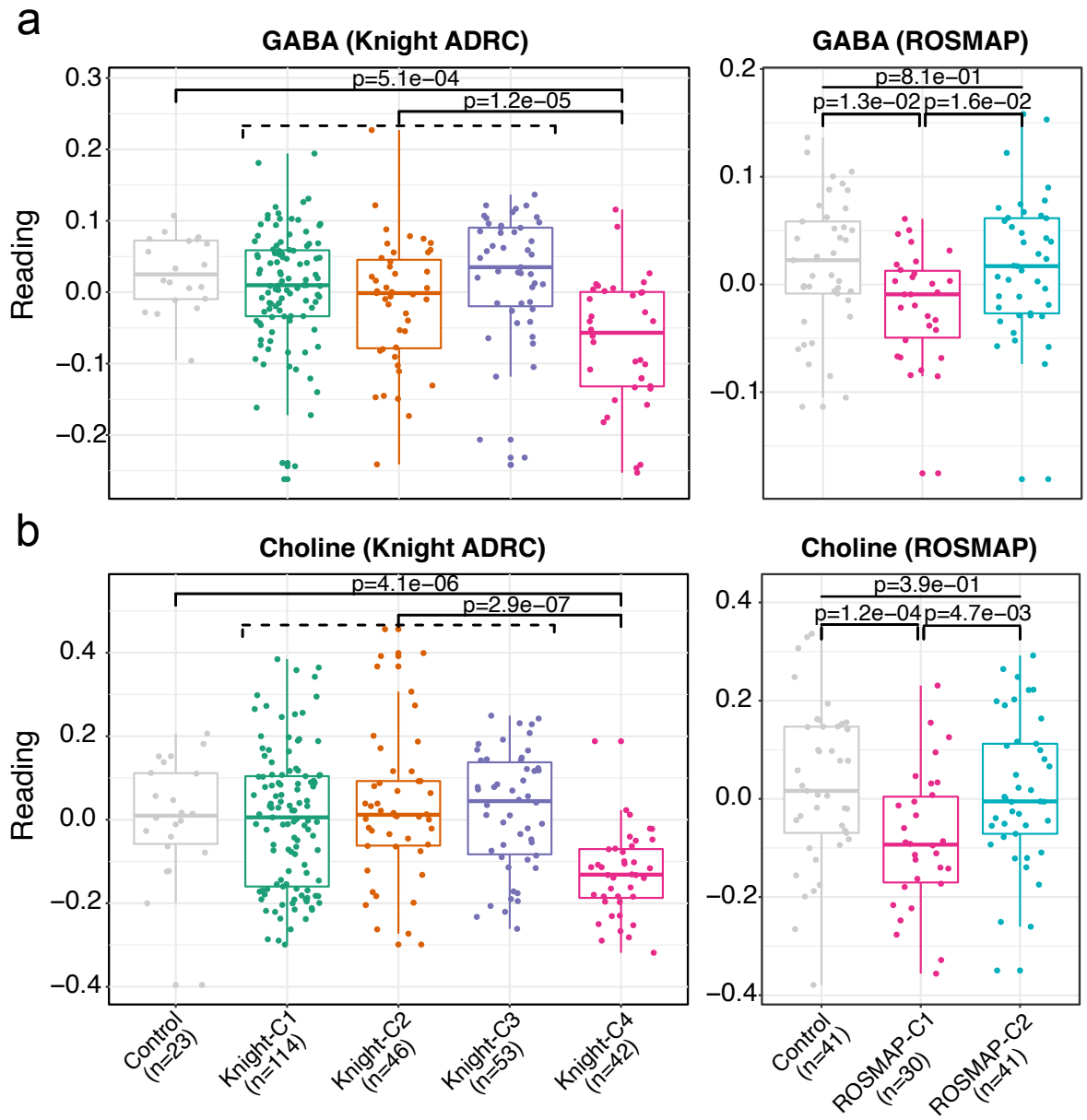
Supplementary Fig. 12



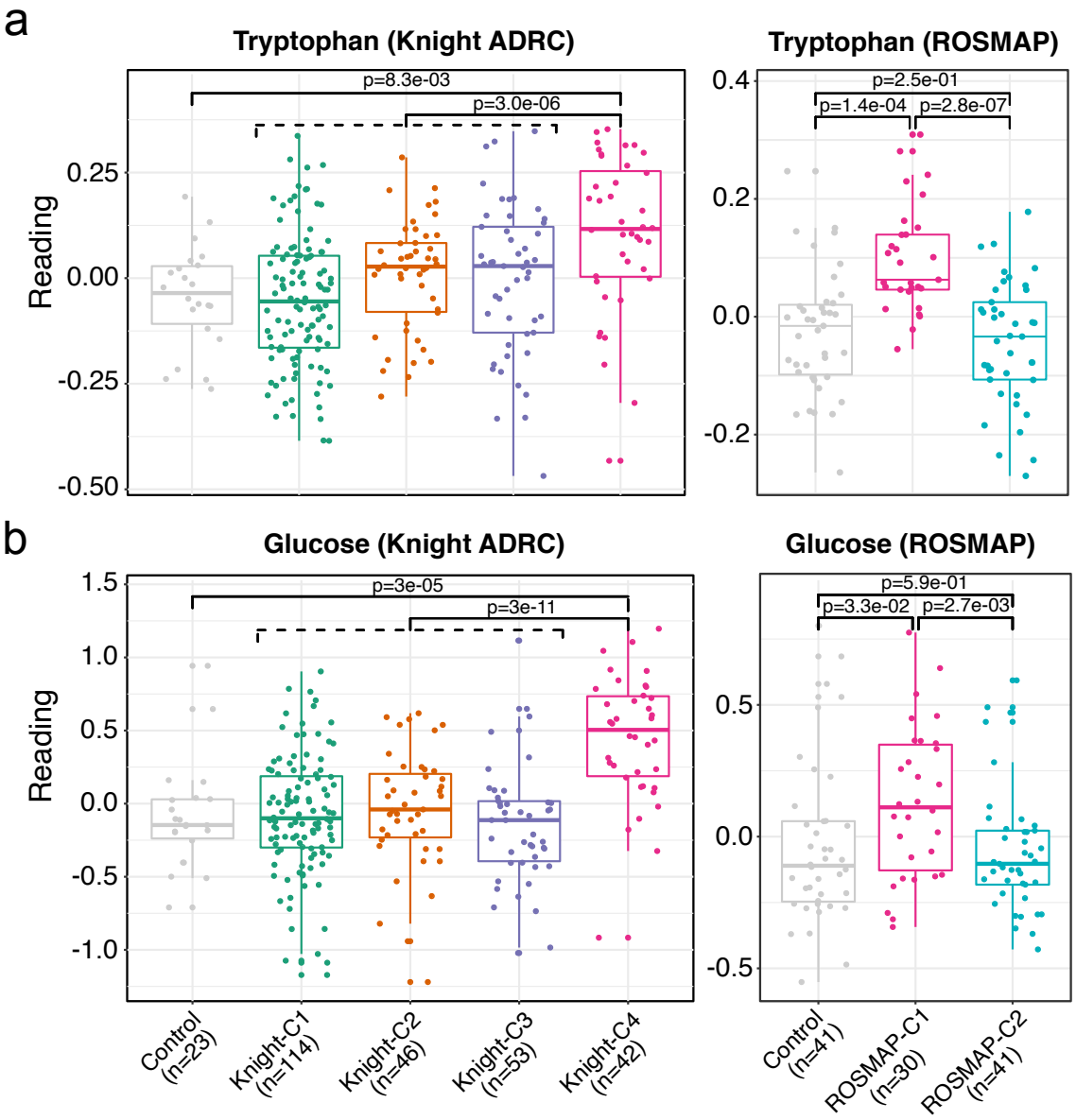
Supplementary Fig. 13



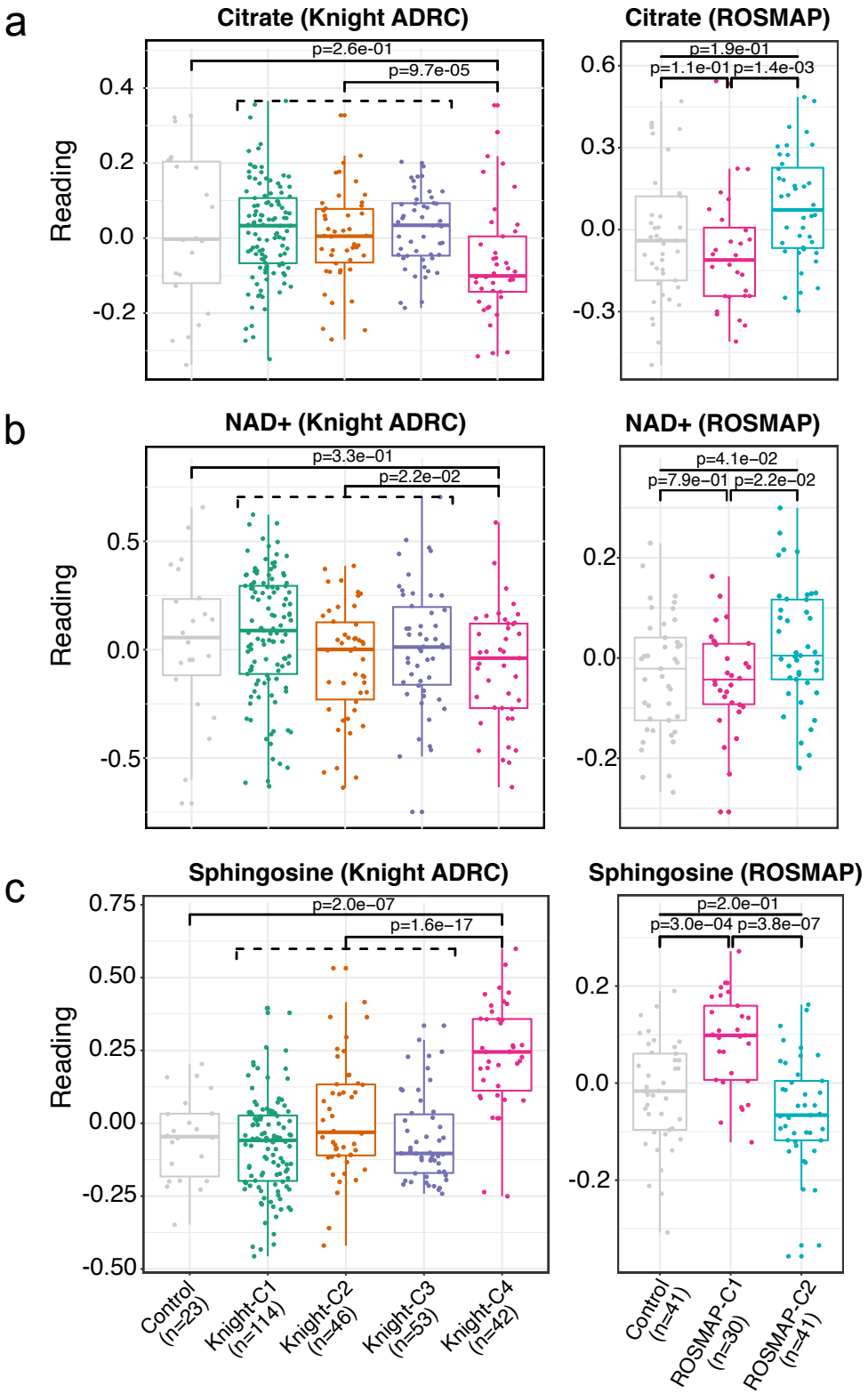
Supplementary Fig. 14



Supplementary Fig. 15

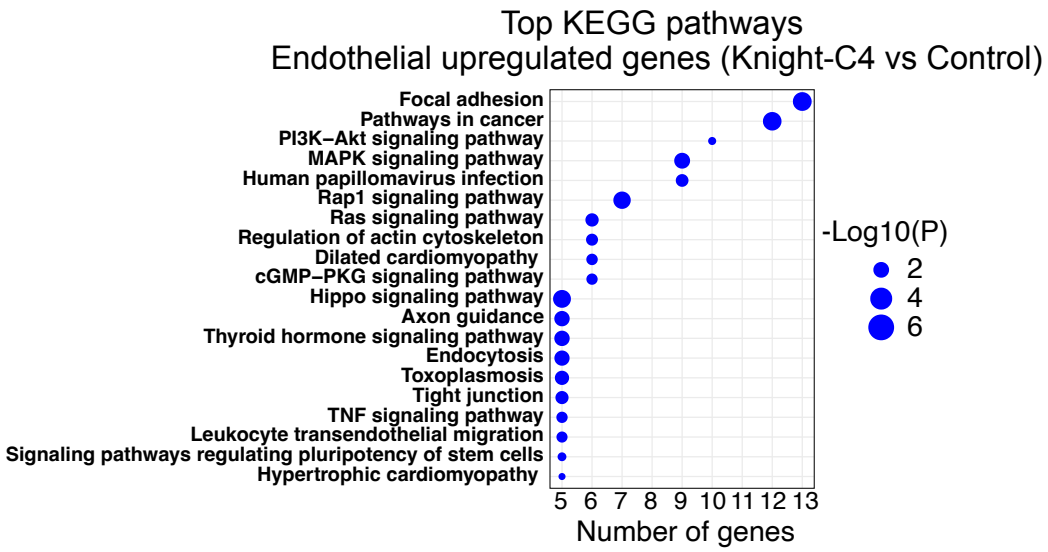


Supplementary Fig. 16

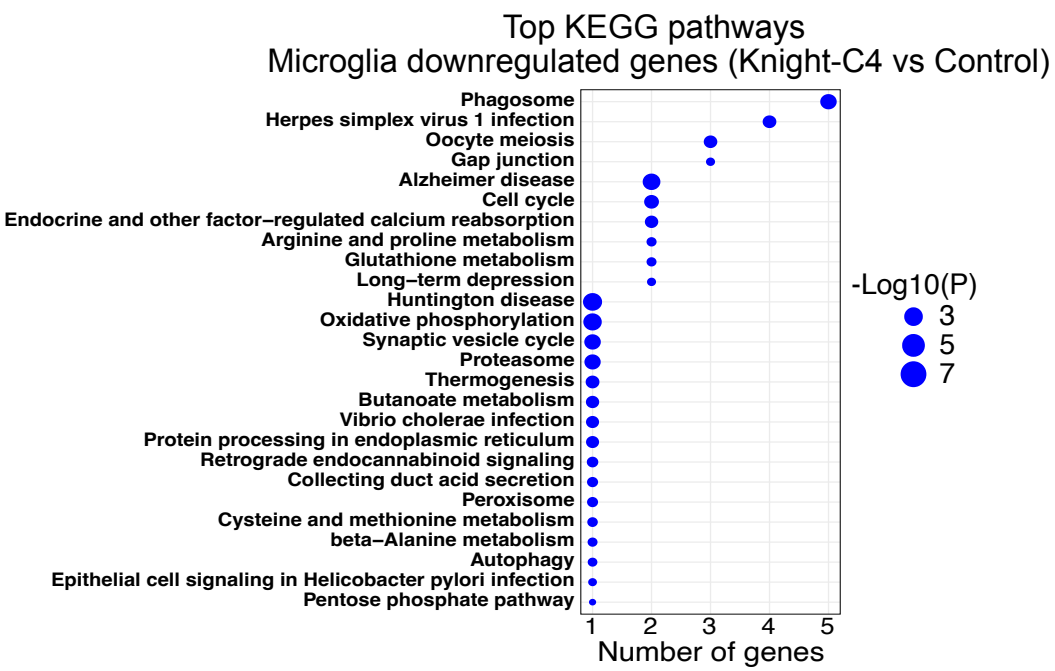


Supplementary Fig. 17

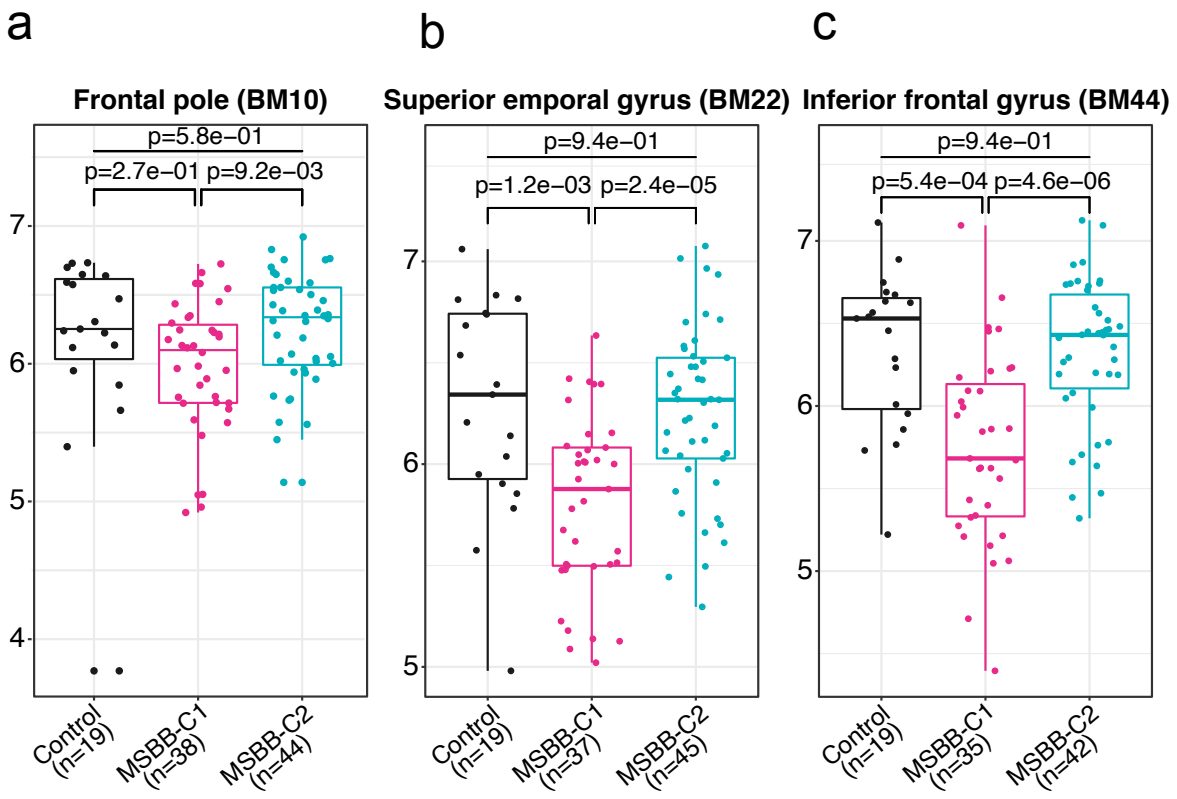
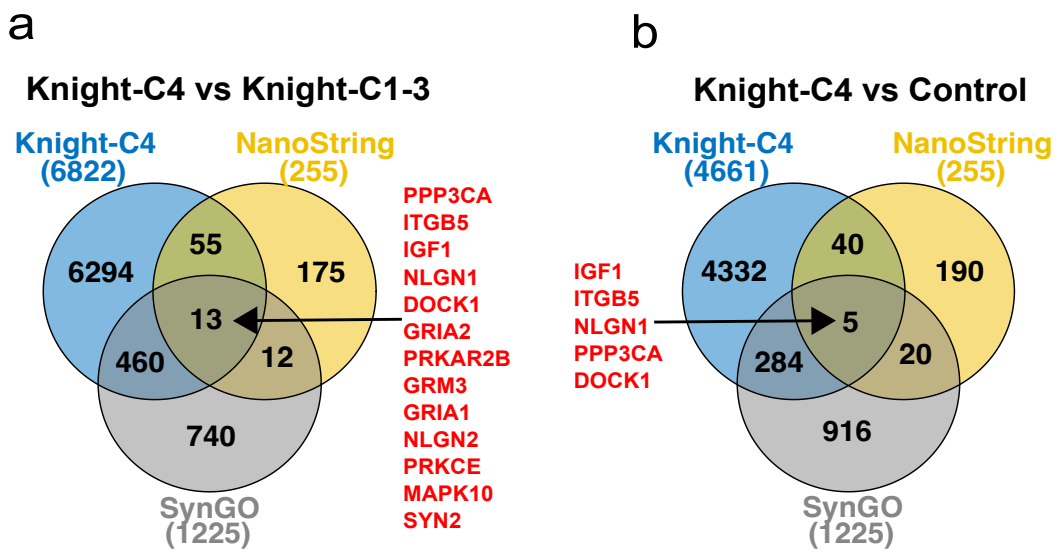
a



b



Supplementary Fig. 18



Supplementary Fig. 19