

1 **CSF metabolites associated with biomarkers of**
2 **Alzheimer's disease pathology**

3 Ruocheng Dong¹, Qiongshi Lu², Hyunseung Kang², Ivonne
4 Suridjan³, Gwendlyn Kollmorgen⁴, Norbert Wild⁴, Yuetiva
5 Deming^{1,5,6}, Carol A. Van Hulle^{5,6}, Rozalyn M. Anderson^{5,7}, Henrik
6 Zetterberg⁸⁻¹², Kaj Blennow^{8,9}, Cynthia M. Carlsson⁵⁻⁷, Sanjay
7 Asthana⁵⁻⁷, Sterling C. Johnson^{5-7,13}, Corinne D. Engelman^{1,6,13*}

8

9 ¹Department of Population Health Sciences, University of
10 Wisconsin-Madison School of Medicine and Public Health,
11 Madison, WI, USA

12 ²Department of Biostatistics and Medical Informatics, University
13 of Wisconsin-Madison School of Medicine and Public Health,
14 Madison, WI, USA

15 ³Roche Diagnostics International Ltd, Rotkreuz, Switzerland

16 ⁴Roche Diagnostics GmbH, Penzberg, Germany

17 ⁵Department of Medicine, School of Medicine and Public Health,
18 University of Wisconsin-Madison, Madison, Wisconsin, USA

19 ⁶Wisconsin Alzheimer's Disease Research Center, University of
20 Wisconsin-Madison School of Medicine and Public Health,
21 Madison, WI, USA

22 ⁷Geriatrics Research Education and Clinical Center, Middleton VA
23 Hospital, Madison, WI, USA

24 ⁸Department of Psychiatry and Neurochemistry, Institute of
25 Neuroscience and Physiology, The Sahlgrenska Academy at
26 University of Gothenburg, Mölndal, S-43180, Sweden

27 ⁹Clinical Neurochemistry Laboratory, Sahlgrenska University

28 Hospital, Mölndal, S-43180, Sweden

29 ¹⁰UK Dementia Research Institute at UCL, London, WC1E6BT, UK

30 ¹¹Department of Neurodegenerative Disease, UCL Institute of

31 Neurology, London, WC1H0AL, UK

32 ¹²Hong Kong Center for Neurodegenerative Diseases, Clear Water

33 Bay, Hong Kong, China

34 ¹³Wisconsin Alzheimer's Institute, University of Wisconsin-

35 Madison School of Medicine and Public Health, Madison, WI, USA

36 *Correspondence:

37 Corinne D. Engelman

38 608-265-5491

39 cengelman@wisc.edu

40 **Abstract**

41 **INTRODUCTION:**

42 Metabolomics technology facilitates studying
43 associations between small molecules and disease processes.
44 Correlating metabolites in cerebrospinal fluid (CSF) with
45 Alzheimer's disease (AD) CSF biomarkers may elucidate
46 additional changes that are associated with early AD
47 pathology and enhance our knowledge of the disease.

48 **METHODS:**

49 The relative abundance of untargeted metabolites was
50 assessed in 161 individuals. A metabolome-wide association
51 study (MWAS) was conducted between 269 CSF metabolites

52 and protein biomarkers reflecting brain amyloidosis, tau
53 pathology, neuronal and synaptic degeneration, and
54 astrocyte or microglial activation and neuroinflammation.
55 Linear mixed-effects regression analyses were performed
56 with random intercepts for sample relatedness and repeated
57 measurements and fixed effects for age, sex, and years of
58 education. The metabolome-wide significance was
59 determined by a false discovery rate threshold of 0.05. The
60 significant metabolites were replicated in 154 independent
61 individuals. Mendelian randomization was performed using
62 genome-wide significant single nucleotide polymorphisms
63 from a CSF metabolites genome-wide association study.

64 **RESULTS:**

65 MWAS results showed several significantly associated
66 metabolites for all the biomarkers except A β _{42/40} and IL-6.
67 Genetic variants associated with metabolites and Mendelian
68 randomization analysis provided evidence for a causal
69 association of metabolites for soluble triggering receptor
70 expressed on myeloid cells 2 (sTREM2), amyloid β (A β ₄₀),
71 α -synuclein, total tau, phosphorylated tau, and neurogranin,
72 for example, palmitoyl sphingomyelin (d18:1/16:0) for
73 sTREM2, and erythritol for A β ₄₀ and α -synuclein.

74 **DISCUSSION:**

75 This study provides evidence that CSF metabolites are
76 associated with AD-related pathology, and many of these
77 associations may be causal.

78

79 **Keywords:** Alzheimer’s disease, metabolomics, CSF

80 NeuroToolKit biomarkers, Mendelian randomization

81

82

83 **1. Introduction**

84 The neuropathological changes of Alzheimer’s disease
85 (AD) consist of extracellular amyloid- β ($A\beta$) plaques and
86 intracellular neurofibrillary tangles of hyperphosphorylated
87 tau proteins in the brain[1]. Well-established core
88 biomarkers that reflect AD pathology and show promising
89 performance in evaluating AD risk and diagnosing AD are
90 the 42 amino acid form $A\beta$ ($A\beta_{42}$), the ratio of $A\beta_{42}/40$,
91 phosphorylated tau (P-tau), and total tau (T-tau) in the
92 cerebrospinal fluid (CSF)[2]. However, it has been suggested
93 that other pathophysiology such as neuroinflammation
94 through glial activation and neuronal and synaptic
95 degeneration also contribute to symptomatic AD, and CSF
96 biomarkers of these may provide valuable information about
97 disease progression[2]. Thus, the NeuroToolKit (NTK), a
98 panel of automated CSF immunoassays, was introduced to
99 complement the established core AD biomarkers[3]. The
100 NTK panel includes S100 calcium-binding protein B
101 (S100b), chitinase-3-like protein 1 (YKL-40), and glial
102 fibrillary acidic protein (GFAP) as markers of astrocyte
103 activation; soluble triggering receptor expressed on myeloid
104 cells 2 (sTREM2) and interleukin-6 (IL-6) as markers of
105 microglial activation and inflammation; and neurofilament

106 light (NfL), neurogranin, and α -synuclein as markers of
107 axonal injury and synaptic dysfunction[4].

108 Untargeted metabolomics technology is a promising
109 approach that can simultaneously identify and quantify a
110 large number of small molecules (<1500 Da, *e.g.*, lipids) in a
111 biological sample[5]. Previous research has shown that
112 metabolomic changes in the human brain and CSF were
113 associated with AD status and AD pathological
114 alterations[6]. For example, Koal *et al.*[6] identified eight
115 metabolites that were significantly increased in the CSF
116 samples with AD-like pathology including an acylcarnitine
117 (C3), two sphingomyelins [SM (d18:1/18:0) and SM
118 (d18:1/18:1)], and five glycerophospholipids (PC aa C32:0,
119 PC aa C34:1, PC aa C36:1, PC aa C38:4, and PC aa C38:6).
120 However, no studies have examined associations between
121 the untargeted CSF metabolome and a broad panel of
122 biomarkers such as the NTK panel. Thus, our study aims to
123 link CSF metabolites with established and developing AD
124 biomarkers with the goals of (1) identifying individual CSF
125 metabolites that are associated with the CSF NTK
126 biomarkers and (2) conducting Mendelian randomization
127 (MR) to determine if the CSF metabolites significantly
128 associated with NTK biomarkers are likely to be in the causal

129 pathway instead of simply changing with, or as a result of,
130 AD biomarker changes.

131 **2. Methods**

132 **2.1 Participants**

133 The Wisconsin Registry for Alzheimer's Prevention
134 (WRAP) began recruitment in 2001 as a prospective cohort
135 study, with initial follow-up four years after baseline and
136 subsequent ongoing follow-up every two years. WRAP is
137 comprised of initially cognitively-unimpaired, asymptomatic,
138 middle-aged (between 40 and 65) adults enriched for
139 parental history of clinical AD[7]. At each visit, the
140 participants undergo comprehensive medical and cognitive
141 evaluations. Additional details of the study design and
142 methods of WRAP have been described previously[7]. From
143 the WRAP cohort, we identified 161 self-reported non-
144 Hispanic white individuals with longitudinal CSF biomarker
145 and metabolomic data. The sample size for other
146 racial/ethnic groups was too small ($n < 10$) to include in the
147 analyses.

148 The Wisconsin Alzheimer's Disease Research Center's
149 (ADRC) clinical core cohorts started in 2009 and are
150 comprised of well-characterized participants who undergo
151 cognitive testing and physical exams every two years[8]. The
152 Wisconsin ADRC has a cohort of initially cognitively-

153 unimpaired, asymptomatic middle-aged (between 45 and 65)
154 adults with a similar study design to WRAP (the
155 **Investigating Memory in Preclinical AD-Causes and**
156 **Treatments [IMPACT] cohort**)[9–11]. From the IMPACT
157 cohort, we identified 154 self-reported non-Hispanic white
158 participants with cross-sectional CSF biomarker and
159 metabolomic data. As with WRAP, the sample size for other
160 racial/ethnic groups was too small ($n < 10$) to include in the
161 analyses.

162 This study was conducted with the approval of the
163 University of Wisconsin Institutional Review Board, and all
164 participants provided signed informed consent before
165 participation.

166 **2.2 CSF sample collection and biomarkers**

167 **quantification**

168 Fasting CSF samples were collected via lumbar
169 puncture using a Sprotte 25- or 24-gauge spinal needle at the
170 L3/4 or L4/5 interspace with gentle extraction into
171 polypropylene syringes. More details can be found in the
172 previous study[9]. The CSF collection for WRAP and the
173 Wisconsin ADRC followed the same protocol, and the
174 lumbar puncture for both studies was performed by the same
175 group of well-trained individuals.

176 All CSF samples were batched together and assayed
177 for the NTK biomarkers at the Clinical Neurochemistry
178 Laboratory, University of Gothenburg, using the same lot of
179 reagents, under strict quality control procedures. The
180 immunoassays of Elecsys® A β (1-42), P-tau(181P) and T-tau,
181 as well as S100b and IL-6, were performed on a cobas e 601
182 analyzer[3]. The remaining NTK panel was assayed on a
183 cobas e 411 analyzer including A β (1-40), α -synuclein, GFAP,
184 YKL-40, sTREM2, NfL, and neurogranin[3].

185 **2.3 CSF metabolomic profiling and quality control**

186 CSF metabolomic analyses and quantification were
187 performed in one batch by Metabolon (Durham, NC) using
188 an untargeted approach, based on Ultrahigh Performance
189 Liquid Chromatography-Tandem Mass Spectrometry
190 platform (UPLC-MS/MS)[12]. Details of the metabolomic
191 profiling were described in an earlier study[13].

192 A total of 412 CSF metabolites were identified and quality
193 control procedures were performed. First, 46 metabolites
194 missing for at least 80% of the individuals were excluded.
195 Then the values for each of the remaining metabolites were
196 scaled so that the median equaled 1. Two metabolites with an
197 interquartile range (IQR) of zero were excluded and no
198 metabolites had zero variability between individuals. Log₁₀
199 transformation was applied to normalize the data. After

200 quality control, 269 metabolites with known biochemical
201 names remained for this investigation. The missing
202 percentage of each metabolite in WRAP and Wisconsin
203 ADRC is available in Supplemental Table 1.

204 **2.4 Genotyping and quality control**

205 In the WRAP participants, DNA was extracted from
206 whole blood using the PUREGENE® DNA Isolation Kit, and
207 the concentrations were quantified using the Invitrogen™
208 Quant-iT™ PicoGreen™ dsDNA Assay Kit. More details can
209 be found in the previous study[13]. Genotyping data were
210 generated by the University of Wisconsin Biotechnology
211 Center using the Illumina Multi-Ethnic Genotyping Array. In
212 the WRAP genetic data, (1) duplicate samples were used to
213 calculate a concordance rate of 99.99%, and discordant
214 genotypes were set to missing; (2) samples missing
215 genotypes for >5% of the single nucleotide polymorphisms
216 (SNPs) were excluded, while SNPs missing in >5% of
217 individuals were also excluded; (3) samples were excluded if
218 the self-reported and genetic sex were inconsistent; (4) SNPs
219 that were not in Hardy-Weinberg equilibrium (HWE;
220 $p < 3.08E-8$) or were monomorphic were removed; (5)
221 individuals that were not of European ancestry were
222 removed due to small sample sizes of other ancestries; (6)
223 the imputation was performed through the Michigan

224 Imputation Server v1.0.328, using the Haplotype Reference
225 Consortium (HRC)[14–16] and the SNPs with a quality score
226 $R^2 < 0.80$, minor allele frequency (MAF) < 0.001 , or that were
227 out of HWE were excluded; (7) genetic ancestry was assessed
228 by using Principal Components Analysis in Related Samples
229 (PC-AiR) because of the sibling relationships present in the
230 WRAP cohort[13].

231 Genetic data in the Wisconsin ADRC were generated
232 from DNA extracted from blood samples at baseline and
233 genotyped with either the Infinium OmniExpressExome-8
234 Kit or the Infinium Global Screening Array-24 Kit. Genetic
235 data for the Wisconsin ADRC underwent the same quality
236 control (QC) and imputation as the WRAP data except
237 samples and SNPs missing in $> 2\%$ were excluded and HWE
238 threshold was $p < 1e-6$ due to differences in sample sizes and
239 the number of SNPs between the two cohorts.

240 **2.5 Statistical analysis**

241 **2.5.1 Metabolome-wide association study**

242 A metabolome-wide association study (MWAS) was
243 conducted in the WRAP cohort between 269 individual CSF
244 metabolites and 13 CSF NTK biomarkers using linear mixed-
245 effects regression models with random intercepts to account
246 for repeated measures and family relationships (10 families
247 with two or more siblings) and fixed effects for age at CSF

248 collection, sex, and years of education. Replication of each
249 CSF metabolite significantly associated with one or more
250 biomarkers in WRAP was then conducted in the Wisconsin
251 ADRC cohort using linear regression adjusting for the same
252 covariates. Both Bonferroni and false discovery rate (FDR)
253 methods were used to correct the p-values for multiple
254 testing; the FDR corrected q value was used to determine
255 statistical significance in each analysis. Potential functional
256 pathways of the replicated significant metabolites were
257 identified by pathway analyses using the web-based software
258 Metabo-analyst 5.0[17] based on the Kyoto Encyclopedia of
259 Genes and Genomes (KEGG) Homo sapiens pathway. The
260 hypergeometric test and relative-betweenness centrality
261 were employed to evaluate the pathway importance, and the
262 pathways were considered as important if the impact was
263 ≥ 0.1 .

264 **2.5.2 Prediction performance and elastic net** 265 **regression**

266 The variance for each biomarker explained by its
267 corresponding significant metabolites was evaluated using r^2
268 in the combined cohorts of WRAP and the Wisconsin ADRC.
269 For this analysis, we only included the first available
270 measures of independent participants from WRAP. Since the
271 number of significant metabolites for each biomarker was

272 large and some of the metabolites were highly correlated,
273 elastic net regression[18] was employed to select the
274 important independent metabolites. Then the r^2 of elastic
275 net-selected metabolites was re-calculated. For each
276 biomarker, we fit three types of models, the (1) base model,
277 which only included the demographics of age, sex, years of
278 education, and cohort, (2) metabolite model, which included
279 the demographics in the base model plus all the replicated
280 significant metabolites, and (3) elastic net-selected
281 metabolite model, which contained the demographics and
282 elastic-net-selected metabolites.

283 **2.5.3 Mendelian randomization**

284 The genome-wide significant SNPs ($p < 5 \times 10^{-8}$)
285 from a previous genome-wide meta-analysis of CSF
286 metabolites[19] were extracted for each elastic net-selected
287 metabolite (5863 SNPs for 52 metabolites). These SNPs (or
288 the top 100 SNPs if there were more than 100 genome-wide
289 significant SNPs for a metabolite) were used as instrumental
290 variables (IV) for the metabolite in an MR analysis for each
291 elastic net-selected metabolite-NTK biomarker association
292 pair in the combined WRAP and Wisconsin ADRC cohort.
293 For each MR test, we first checked the strength of the IVs
294 using F statistics. Typically, an IV with an F statistic greater
295 than 10 is considered to be strong, while instruments with F

296 statistics below 10 are considered to be weak[20]. Next, the
297 estimated (or less confounded) beta and p values for the
298 effect of the metabolite on the NTK biomarker were
299 calculated using the two-stage least squares method if the
300 IVs were strong, but using the limited information maximum
301 likelihood (LIML) for IVs that were relatively weak. The
302 confidence intervals (CI) of the point estimates from both
303 LIML and another conditional likelihood ratio (CLR)
304 method, which is robust to weak IVs[21], were compared and
305 only significant results with CIs in the same direction and
306 with a similar range of effect size between these two methods
307 were considered as evidence of a causal effect. The
308 Bonferroni corrected p-value < 0.05 based on the number of
309 all MR tests performed was used to determine significance.
310 The MR analysis was conducted using the R package
311 "ivmodel"[22].

312 **3. Results**

313 **3.1 Participant characteristics**

314 Characteristics of the WRAP and Wisconsin ADRC
315 participants can be found in Table 1. Among 161 WRAP
316 participants, the mean baseline age and education level were
317 62.1 and 16.2 years, respectively. The mean age and years of
318 education in the Wisconsin ADRC were 58.1 and 16.2,
319 respectively. Females comprised 65.2% of WRAP

320 participants and 68.8% of the Wisconsin ADRC. The mean
321 values of each biomarker are also listed in Table 1.

322 **3.2 MWAS**

323 The significant MWAS results in WRAP and the
324 Wisconsin ADRC are summarized in Figure 1. In WRAP, a
325 large number of CSF metabolites reached the significance
326 threshold after FDR correction [Figure 1. (a)]. 47 metabolites
327 were associated with P-tau, 56 were associated with T-tau,
328 58 were associated with A β ₄₂, 80 were associated with
329 A β ₄₀, 65 were associated with NfL, and 62 were associated
330 with neurogranin. However, no metabolites were associated
331 with the ratio of A β ₄₂/40 or IL-6. Many of the metabolites
332 that were significant in WRAP were also significant in the
333 Wisconsin ADRC [Figure 1. (b)]. For example, among 47
334 significant metabolites for P-tau in WRAP, 40 metabolites
335 were also significant in the Wisconsin ADRC. Table 2 shows
336 the replication results for the top 10 significant CSF
337 metabolite-biomarker associations (if there were 10 or more
338 significant metabolites) in the Wisconsin ADRC. For
339 example, the top three metabolites associated with P-tau and
340 T-tau were 1-palmitoyl-2-stearoyl-GPC (16:0/18:0), N-
341 acetylneuraminate, and C-glycosyltryptophan. N-
342 acetylneuraminate and 1,2-dipalmitoyl-GPC (16:0/16:0)
343 were the top two metabolites associated with A β ₄₂ and A β ₄₀.

344 The top three metabolites associated with NfL were N-
345 acetylthreonine, N-acetylalanine, and beta-citrylglutamate.
346 N-acetylneuraminate, C-glycosyltryptophan, and N6-
347 succinyladenosine were the top three metabolites for
348 neurogranin. N-acetylneuraminate, 1,2-dipalmitoyl-GPC
349 (16:0/16:0), and stearyl sphingomyelin (d18:1/18:0) were
350 the top three metabolites for YKL40. Stearyl sphingomyelin
351 (d18:1/18:0), 1-stearoyl-2-docosahexaenoyl-GPC (18:0/22:6),
352 and 1-palmitoyl-2-oleoyl-GPC (16:0/18:1) were the top three
353 metabolites associated with S100b. Only six metabolites
354 were associated with GFAP, and the top three were 1,2-
355 dipalmitoyl-GPC (16:0/16:0), 1-palmitoyl-2-stearoyl-GPC
356 (16:0/18:0), and beta-citrylglutamate. For sTREM2, the top
357 metabolites were stearyl sphingomyelin (d18:1/18:0), 1,2-
358 dipalmitoyl-GPC (16:0/16:0), and palmitoyl sphingomyelin
359 (d18:1/16:0). Finally, for α -synuclein, the top three
360 metabolites were 1-palmitoyl-2-stearoyl-GPC (16:0/18:0),
361 1,2-dipalmitoyl-GPC (16:0/16:0), and N-acetylneuraminate.
362 The full results of WRAP and the Wisconsin ADRC can be
363 found in Supplemental Tables 2-25. The association patterns
364 between significant CSF metabolites and NTK biomarkers
365 are provided in the Figure 2. The summary of the number of
366 significant associations and the name of NTK biomarkers
367 that were replicated in the Wisconsin ADRC are presented in

368 Supplemental Table 26. Most of the significant metabolites
369 were lipids, amino acids, and carbohydrates. For example,
370 the lipid, 1,2-dipalmitoyl-GPC (16:0/16:0), the amino acid,
371 beta-citrylglutamate, and the carbohydrate N-
372 acetylneuraminate were strongly associated with almost
373 every CSF NTK biomarker of AD. On the contrary, amino
374 acids like kynurenate and proline were only significantly
375 associated with α -synuclein.

376 The functional pathways for replicated significant
377 metabolites with known human metabolome database
378 (HMDB) IDs for each CSF NTK biomarker are shown in
379 Supplemental Table 27. Two significant metabolites, 1,2-
380 dipalmitoyl-GPC (16:0/16:0) and 1-oleoyl-GPC (18:1), were
381 enriched in the glycerophospholipid metabolism pathway for
382 most biomarkers. Other pathways such as pyrimidine
383 metabolism (including orotate and orotidine), ascorbate and
384 aldarate metabolism (including gulonate and glucuronate),
385 arginine biosynthesis (including N-acetylglutamate and
386 argininosuccinate), and pentose and glucuronate
387 interconversions (also including gulonate and glucuronate)
388 may also be of interest.

389 **3.3 Prediction performance and elastic net**

390 **regression results**

391 The prediction performance of replicated significant
392 metabolites was measured by r^2 and presented in Table 3.
393 The r^2 of the base models, which only included the
394 demographic variables, ranged from 0.01 to 0.25. Adding the
395 replicated significant metabolites increased the r^2
396 substantially for each biomarker, ranging from 0.13 to 0.94.
397 The elastic net regression further prioritized candidate
398 metabolites associated with each biomarker. For example, 22
399 of the original 40 significant metabolites were selected by the
400 elastic net as important independent metabolites for P-tau.
401 Initially, 40 significant metabolites explained about 72% of
402 the variance in P-tau; the 22 elastic net-selected metabolites
403 still explained 70% of the variance.

404 **3.4 Mendelian randomization**

405 According to the F statistics, we employed the LIML
406 method for MR. The full results of the test statistics are
407 provided in Supplemental Table 28. After checking for
408 consistency of the CIs for the LIML and CLR methods, the
409 significant and consistent MR results are displayed in Table
410 4, showing metabolites with a potential causal effect on the
411 NTK biomarker based on instrumental variables formed by
412 genome-wide significant SNPs. For example, we observed a
413 positive causal association between palmitoyl sphingomyelin
414 (d18:1/16:0) and sTREM2.

415 **4. Discussion**

416 In this analysis, we tested the associations between
417 CSF metabolites and CSF NTK biomarkers representing
418 different pathologies of AD in initially cognitively-
419 unimpaired individuals. Significant metabolites were
420 identified in the WRAP cohort using linear mixed effects
421 regression and most of the metabolites were replicated in the
422 Wisconsin ADRC cohort. The elastic net regression method
423 reduced the number of CSF metabolites by selecting the
424 important and independent metabolites for each CSF
425 biomarker. This provides a smaller, more practical set of
426 metabolites to focus on in future research. The results of the
427 MR analyses suggested several metabolites that may play a
428 causal role in AD pathology. A detailed look into these
429 associations, such as the contributing genes and their
430 corresponding functions, is worth exploring.

431 We have identified and replicated multiple CSF
432 metabolites that were associated with CSF NTK biomarkers
433 for AD pathology; most of these CSF metabolites were lipids,
434 particularly sphingolipids, phosphatidylcholines, and
435 lysophospholipids, which are all types of phospholipids.
436 Phospholipids are a class of lipids that construct the cellular
437 membranes and are involved in many complex activities of
438 membrane proteins, receptors, enzymes, and ion channels in

439 the cell or at the cell surface[23]. In the neurodegenerative
440 brain, *e.g.*, in the AD brain, which has suffered extensive
441 damage, the compromise of the membrane functions is
442 expected, explaining how phospholipids may be involved in
443 AD pathology[24]. Previous studies have demonstrated that
444 various phospholipids such as phosphatidylcholines,
445 sphingolipids, glycerophospholipids, and lysophospholipids
446 have changed in the AD patient's brain, CSF and blood when
447 compared to healthy controls[23,25,26]. For example, a
448 serum metabolomics study conducted by González-
449 Domínguez *et al.*[25] showed that the concentration of
450 numerous phosphatidyl lipids, like 1,2-dipalmitoyl-GPC
451 (16:0/16:0), 1-palmitoyl-2-linoleoyl-GPC (16:0/18:2), and 1-
452 palmitoyl-2-oleoyl-GPC (16:0/18:1), and
453 lysophosphatidylcholines, like 1-palmitoyl-GPC (16:0) and 1-
454 stearoyl-GPC (18:0), were different in AD versus healthy
455 controls. The 1,2-dipalmitoyl-GPC (16:0/16:0)
456 phosphatidylcholine has also been suggested as one of three
457 serum metabolites to predict AD development in MCI
458 individuals[27]. Another brain metabolomics study found
459 that higher levels of palmitoyl sphingomyelin (d18:1/16:0)
460 and sphingomyelin (d18:1/18:1, d18:2/18:0) were associated
461 with the severity of AD pathology at autopsy and AD
462 progression across prodromal and preclinical stages[28].

463 The stearyl sphingomyelin (d18:1/18:0) was also
464 significantly changed in the CSF with “AD-like pathology”
465 that was dichotomized by A β 42, T-tau, and P-tau levels[6].
466 In summary, our results confirmed the importance of the
467 previously identified lipids but also provided novel lipid
468 findings for AD pathologies beyond the major established
469 ones.

470 Another class of metabolites that are of potential
471 interest are several carbohydrates like N-acetylneuraminate,
472 arabitol/xylitol, arabinose, and erythronate. Among them, N-
473 acetylneuraminate, also known as sialic acid, had a
474 significant effect on most NTK biomarkers. In addition to
475 our study, a previous study conducted by Nagata *et al.*[29] in
476 2018 also showed that CSF N-acetylneuraminate was
477 significantly increased in AD when compared to patients
478 with idiopathic normal pressure hydrocephalus and was
479 positively correlated with CSF P-tau ($r=0.55$), as it was in our
480 study. N-acetylneuraminate is an acetyl derivative of the
481 amino sugar neuraminic acid, which occurs in many
482 glycoproteins, glycolipids, and polysaccharides. Specifically,
483 it is a functional and structural component of gangliosides,
484 which are found predominantly in the nervous system and
485 are abundant in the brain, especially in the grey matter[30].
486 Studies have shown that gangliosides play important roles in

487 AD. For example, it has been suggested that GM1-
488 ganglioside binds to A β , and the resulted GA β has the
489 capability to accelerate A β assembly[31] and is the
490 endogenous seed for amyloid fibril in the AD brain[32]. The
491 gangliosides also have important roles in organizing the lipid
492 rafts, which integrate numerous types of lipid proteins
493 involved in cell signaling, cell-cell adhesion, and intracellular
494 vesicular trafficking[29] and contain many AD-associated
495 proteins such as amyloid precursor protein (APP)[33].
496 Furthermore, the gene *CD33*, which belongs to the sialic-
497 acid-binding immunoglobulin-like lectin family, has been
498 reported as a strong genetic locus associated with AD by
499 GWASs[34–36] and has been suggested to impair the
500 microglia-mediated A β clearance[37–39]. Erythronate
501 (erythronic acid) was previously identified as the main
502 hallmark of pentose–phosphate pathway defects[40], and
503 consistent with abnormal function of pentose–phosphate
504 pathway in certain regions of the AD-brain[41], and the
505 upregulation of the pentose–phosphate pathway was
506 reported in a previous study of mild cognitive impairment
507 (MCI) participants that later progressed to AD[42].

508 As mentioned above, a couple of metabolites were
509 common to most of the AD pathologies defined by the CSF
510 NTK biomarkers. On the contrary, some metabolites were

511 unique to specific NTK biomarkers. For example, lipids like
512 1-palmitoyl-2-linoleoyl-GPC (16:0/18:2), 1-stearoyl-2-
513 arachidonoyl-GPC (18:0/20:4), sphingomyelin (d18:1/20:0,
514 d16:1/22:0) and sphingomyelin (d18:1/22:1, d18:2/22:0,
515 d16:1/24:1) were only associated with α -synuclein. These
516 metabolites may be helpful to study synaptic dysfunction and
517 could potentially be used as biomarkers to differentiate AD
518 pathologies.

519 By utilizing Mendelian randomization, we found
520 causal evidence for several of the associations between CSF
521 metabolites and CSF NTK biomarkers. Among these
522 metabolites, most of them were lipids, with some amino
523 acids and cofactors/vitamins, and a xenobiotic metabolite,
524 erythritol. Another metabolite of interest, homocarnosine, is
525 an inhibitory neuromodulator synthesized in the neuron
526 from gamma-aminobutyric acid (GABA) and histidine[43].
527 The level of human CSF homocarnosine declines drastically
528 with age [44] and was suggested to be related to AD through
529 CSF protein glycation[45]. At the same time, GABA also
530 plays an important role in the brain and may be related to
531 AD[46].

532 This study has some limitations. First, the analysis
533 only included non-Hispanic white individuals, so the results
534 may not extrapolate to other racial/ethnic groups. Second,

535 the sample sizes of both the WRAP and Wisconsin ADRC
536 cohorts were relatively small and will need to be replicated in
537 a larger independent sample. The significant associations
538 between a number of metabolites and both A β ₄₂ and A β ₄₀,
539 but not with A β _{42/40} may indicate that the metabolites
540 associated with A β ₄₂ and A β ₄₀ only influence the
541 production of amyloid in general versus clearance of the
542 pathological form, A β ₄₂. In general, the research confirmed
543 that several novel metabolites changed along with AD CSF
544 biomarkers and extended several developing and
545 understudied AD pathologies, e.g., synaptic dysfunction,
546 based on untargeted CSF metabolomics and will expand our
547 knowledge of the biological mechanisms behind AD.

548 **Acknowledgments**

549 The authors especially thank the WRAP and
550 Wisconsin ADRC participants and staff for their
551 contributions to the studies. Without their efforts, this
552 research would not be possible. This study was supported by
553 the National Institutes of Health (NIH) grants [R01AG27161
554 (Wisconsin Registry for Alzheimer Prevention: Biomarkers
555 of Preclinical AD), R01AG054047 (Genomic and
556 Metabolomic Data Integration in a Longitudinal Cohort at
557 Risk for Alzheimer's Disease), R01AG037639 (White Matter
558 Degeneration: Biomarkers in Preclinical Alzheimer's

559 Disease), R21AGO67092 (Identifying Metabolomic Risk
560 Factors in Plasma and Cerebrospinal Fluid for Alzheimer's
561 Disease), and P30AGO62715 (Wisconsin Alzheimer's Disease
562 Research Center Grant)], the Helen Bader Foundation,
563 Northwestern Mutual Foundation, Extencicare Foundation,
564 State of Wisconsin, the Clinical and Translational Science
565 Award (CTSA) program through the NIH National Center for
566 Advancing Translational Sciences (NCATS) grant
567 [UL1TR000427], and the University of Wisconsin-Madison
568 Office of the Vice Chancellor for Research and Graduate
569 Education with funding from the Wisconsin Alumni
570 Research Foundation. Computational resources were
571 supported by core grants to the Center for Demography and
572 Ecology [P2CHDo47873] and the Center for Demography of
573 Health and Aging [P30AGO17266].

574 HZ is a Wallenberg Scholar supported by grants from
575 the Swedish Research Council (#2018-02532), the European
576 Research Council (#681712), Swedish State Support for
577 Clinical Research (#ALFGBG-720931), the Alzheimer Drug
578 Discovery Foundation (ADDF), USA (#201809-2016862),
579 the AD Strategic Fund and the Alzheimer's Association
580 (#ADSF-21-831376-C, #ADSF-21-831381-C and #ADSF-21-
581 831377-C), the Olav Thon Foundation, the Erling-Persson
582 Family Foundation, Stiftelsen för Gamla Tjänarinnor,

583 Hjärnfonden, Sweden (#FO2019-0228), the European
584 Union's Horizon 2020 research and innovation programme
585 under the Marie Skłodowska-Curie grant agreement No
586 860197 (MIRIADE), European Union Joint Program for
587 Neurodegenerative Disorders (JPND2021-00694), and the
588 UK Dementia Research Institute at UCL.

589 KB is supported by the Swedish Research Council
590 (#2017-00915), ADDF, USA [#RDAPB-201809-2016615],
591 the Swedish Alzheimer Foundation [#AF-742881],
592 Hjärnfonden, Sweden [#FO2017-0243], the Swedish state
593 under the agreement between the Swedish government and
594 the County Councils, the ALF-agreement [#ALFGBG-
595 715986], and European Union Joint Program for
596 Neurodegenerative Disorders [JPND2019-466-236], and the
597 Alzheimer's Association 2021 Zenith Award (ZEN-21-
598 848495).

599 We thank the University of Wisconsin Madison
600 Biotechnology Center Gene Expression Center for providing
601 Illumina Infinium genotyping services and Roche for
602 providing the NTK kits for this study. COBAS, COBAS E and
603 ELECSYS are trademarks of Roche.

604 **Conflicts of Interest**

605 HZ has served at scientific advisory boards and/or as
606 a consultant for Abbvie, Alector, Annexon, Artery

607 Therapeutics, AZTherapies, CogRx, Denali, Eisai, Nervgen,
608 Pinteon Therapeutics, Red Abbey Labs, Passage Bio, Roche,
609 Samumed, Siemens Healthineers, Triplet Therapeutics, and
610 Wave, has given lectures in symposia sponsored by
611 Cellectricon, Fujirebio, Alzecure, Biogen, and Roche, and is a
612 co-founder of Brain Biomarker Solutions in Gothenburg AB
613 (BBS), which is a part of the GU Ventures Incubator Program
614 (outside submitted work).

615 KB has served as a consultant, at advisory boards, or
616 at data monitoring committees for Abcam, Axon, Biogen,
617 JOMDD/Shimadzu. Julius Clinical, Lilly, MagQu, Novartis,
618 Pharmatrophix, Prothena, Roche Diagnostics, and Siemens
619 Healthineers, and is a co-founder of Brain Biomarker
620 Solutions in Gothenburg AB (BBS), which is a part of the GU
621 Ventures Incubator Program, all unrelated to the work
622 presented in this paper.

623 Gwendlyn Kollmorgen and Norbert Wild are full-time
624 employees of Roche Diagnostics GmbH. Ivonne Suridjan is a
625 full-time employee of Roche Diagnostics International Ltd
626 and holds non-voting equities in F. Hoffmann-La Roche.

627 **Reference**

- 628 [1] Anoop A, Singh PK, Jacob RS, Maji SK. CSF Biomarkers
629 for Alzheimer's Disease Diagnosis. *Int J Alzheimers Dis*
630 2010;2010. <https://doi.org/10.4061/2010/606802>.
- 631 [2] Blennow K, Zetterberg H, Fagan AM. Fluid Biomarkers
632 in Alzheimer Disease. *Cold Spring Harb Perspect Med*
633 2012;2:a006221–a006221.
634 <https://doi.org/10.1101/cshperspect.a006221>.
- 635 [3] Hulle CV, Jonaitis EM, Betthausen TJ, Batrla R, Wild N,
636 Kollmorgen G, et al. An examination of a novel
637 multipanel of CSF biomarkers in the Alzheimer's
638 disease clinical and pathological continuum. *Alzheimers*
639 *Dement* 2021;17:431–45.
640 <https://doi.org/10.1002/alz.12204>.
- 641 [4] Fluid AD Biomarkers Link P-Tau to Synapses,
642 Inflammation | ALZFORUM n.d.
643 [https://www.alzforum.org/news/conference-](https://www.alzforum.org/news/conference-coverage/fluid-ad-biomarkers-link-p-tau-synapses-inflammation)
644 [coverage/fluid-ad-biomarkers-link-p-tau-synapses-](https://www.alzforum.org/news/conference-coverage/fluid-ad-biomarkers-link-p-tau-synapses-inflammation)
645 [inflammation](https://www.alzforum.org/news/conference-coverage/fluid-ad-biomarkers-link-p-tau-synapses-inflammation) (accessed June 5, 2021).
- 646 [5] Hasin Y, Seldin M, Lusic A. Multi-omics approaches to
647 disease. *Genome Biol* 2017;18:83.
648 <https://doi.org/10.1186/s13059-017-1215-1>.
- 649 [6] Koal T, Klavins K, Seppi D, Kemmler G, Humpel C.
650 Sphingomyelin SM(d18:1/18:0) is Significantly

- 651 Enhanced in Cerebrospinal Fluid Samples
652 Dichotomized by Pathological Amyloid- β 42, Tau, and
653 Phospho-Tau-181 Levels. *J Alzheimers Dis*
654 2015;44:1193–201. [https://doi.org/10.3233/JAD-](https://doi.org/10.3233/JAD-142319)
655 142319.
- 656 [7] Johnson SC, Kosciak RL, Jonaitis EM, Clark LR, Mueller
657 KD, Berman SE, et al. The Wisconsin Registry for
658 Alzheimer’s Prevention: A review of findings and
659 current directions. *Alzheimers Dement Diagn Assess*
660 *Dis Monit* 2018;10:130–42.
661 <https://doi.org/10.1016/j.dadm.2017.11.007>.
- 662 [8] Bettcher BM, Johnson SC, Fitch R, Casaletto KB,
663 Heffernan KS, Asthana S, et al. Cerebrospinal Fluid and
664 Plasma Levels of Inflammation Differentially Relate to
665 CNS Markers of Alzheimer’s Disease Pathology and
666 Neuronal Damage. *J Alzheimers Dis* 2018;62:385–97.
667 <https://doi.org/10.3233/JAD-170602>.
- 668 [9] Darst BF, Kosciak RL, Racine AM, Oh JM, Krause RA,
669 Carlsson CM, et al. Pathway-specific polygenic risk
670 scores as predictors of β -amyloid deposition and
671 cognitive function in a sample at increased risk for
672 Alzheimer’s disease. *J Alzheimers Dis JAD*
673 2017;55:473–84. <https://doi.org/10.3233/JAD-160195>.

- 674 [10] Racine AM, Kosciak RL, Berman SE, Nicholas CR, Clark
675 LR, Okonkwo OC, et al. Biomarker clusters are
676 differentially associated with longitudinal cognitive
677 decline in late midlife. *Brain* 2016;139:2261–74.
678 <https://doi.org/10.1093/brain/aww142>.
- 679 [11] Vogt NM, Romano KA, Darst BF, Engelman CD,
680 Johnson SC, Carlsson CM, et al. The gut microbiota-
681 derived metabolite trimethylamine N-oxide is elevated
682 in Alzheimer’s disease. *Alzheimers Res Ther*
683 2018;10:124. [https://doi.org/10.1186/s13195-018-0451-](https://doi.org/10.1186/s13195-018-0451-2)
684 2.
- 685 [12] Bridgewater BR EA. High Resolution Mass
686 Spectrometry Improves Data Quantity and Quality as
687 Compared to Unit Mass Resolution Mass Spectrometry
688 in High-Throughput Profiling Metabolomics. *J*
689 *Postgenomics Drug Biomark Dev* 2014;04.
690 <https://doi.org/10.4172/2153-0769.1000132>.
- 691 [13] Darst BF, Lu Q, Johnson SC, Engelman CD. Integrated
692 analysis of genomics, longitudinal metabolomics, and
693 Alzheimer’s risk factors among 1,111 cohort participants.
694 *Genet Epidemiol* 2019;43:657–74.
695 <https://doi.org/10.1002/gepi.22211>.
- 696 [14] Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong
697 A, et al. Next-generation genotype imputation service

- 698 and methods. *Nat Genet* 2016;48:1284–7.
- 699 <https://doi.org/10.1038/ng.3656>.
- 700 [15] Loh P-R, Danecek P, Palamara PF, Fuchsberger C, A
701 Reshef Y, K Finucane H, et al. Reference-based phasing
702 using the Haplotype Reference Consortium panel. *Nat*
703 *Genet* 2016;48:1443–8.
704 <https://doi.org/10.1038/ng.3679>.
- 705 [16] McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood
706 AR, Teumer A, et al. A reference panel of 64,976
707 haplotypes for genotype imputation. *Nat Genet*
708 2016;48:1279–83. <https://doi.org/10.1038/ng.3643>.
- 709 [17] Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L,
710 Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap
711 between raw spectra and functional insights. *Nucleic*
712 *Acids Res* 2021;49:W388–96.
713 <https://doi.org/10.1093/nar/gkab382>.
- 714 [18] Zou H, Hastie T. Regularization and variable selection
715 via the elastic net. *J R Stat Soc Ser B Stat Methodol*
716 2005;67:301–20. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-9868.2005.00503.x)
717 [9868.2005.00503.x](https://doi.org/10.1111/j.1467-9868.2005.00503.x).
- 718 [19] Panyard DJ, Kim KM, Darst BF, Deming YK, Zhong X,
719 Wu Y, et al. Cerebrospinal fluid metabolomics identifies
720 19 brain-related phenotype associations. *Commun Biol*

- 721 2021;4:1–11. <https://doi.org/10.1038/s42003-020->
722 01583-z.
- 723 [20] Stock JH, Wright JH, Yogo M. A Survey of Weak
724 Instruments and Weak Identification in Generalized
725 Method of Moments. *J Bus Econ Stat* 2002;20:518–29.
726 <https://doi.org/10.1198/073500102288618658>.
- 727 [21] Moreira MJ. A Conditional Likelihood Ratio Test for
728 Structural Models. *Econometrica* 2003;71:1027–48.
729 <https://doi.org/10.1111/1468-0262.00438>.
- 730 [22] Kang H, Jiang Y, Zhao Q, Small DS. ivmodel: An R
731 Package for Inference and Sensitivity Analysis of
732 Instrumental Variables Models with One Endogenous
733 Variable. *ArXiv200208457 Stat* 2020.
- 734 [23] Kosicek M, Hecimovic S. Phospholipids and Alzheimer’s
735 Disease: Alterations, Mechanisms and Potential
736 Biomarkers. *Int J Mol Sci* 2013;14:1310–22.
737 <https://doi.org/10.3390/ijms14011310>.
- 738 [24] Wong MW, Braidy N, Poljak A, Pickford R, Thambisetty
739 M, Sachdev PS. Dysregulation of lipids in Alzheimer’s
740 disease and their role as potential biomarkers.
741 *Alzheimers Dement* 2017;13:810–27.
742 <https://doi.org/10.1016/j.jalz.2017.01.008>.
- 743 [25] González-Domínguez R, García-Barrera T, Gómez-Ariza
744 JL. Combination of metabolomic and phospholipid-

- 745 profiling approaches for the study of Alzheimer's
746 disease. *J Proteomics* 2014;104:37–47.
747 <https://doi.org/10.1016/j.jprot.2014.01.014>.
- 748 [26] Kao Y-C, Ho P-C, Tu Y-K, Jou I-M, Tsai K-J. Lipids and
749 Alzheimer's Disease. *Int J Mol Sci* 2020;21:1505.
750 <https://doi.org/10.3390/ijms21041505>.
- 751 [27] Orešič M, Hyötyläinen T, Herukka S-K, Sysi-Aho M,
752 Mattila I, Seppänen-Laakso T, et al. Metabolome in
753 progression to Alzheimer's disease. *Transl Psychiatry*
754 2011;1:e57–e57. <https://doi.org/10.1038/tp.2011.55>.
- 755 [28] Varma VR, Oommen AM, Varma S, Casanova R, An Y,
756 Andrews RM, et al. Brain and blood metabolite
757 signatures of pathology and progression in Alzheimer
758 disease: A targeted metabolomics study. *PLOS Med*
759 2018;15:e1002482.
760 <https://doi.org/10.1371/journal.pmed.1002482>.
- 761 [29] Nagata Y, Hirayama A, Ikeda S, Shirahata A, Shoji F,
762 Maruyama M, et al. Comparative analysis of
763 cerebrospinal fluid metabolites in Alzheimer's disease
764 and idiopathic normal pressure hydrocephalus in a
765 Japanese cohort. *Biomark Res* 2018;6:5.
766 <https://doi.org/10.1186/s40364-018-0119-x>.
- 767 [30] Palmano K, Rowan A, Guillermo R, Guan J, Mc Jarrow
768 P. The Role of Gangliosides in Neurodevelopment.

- 769 Nutrients 2015;7:3891–913.
770 <https://doi.org/10.3390/nu7053891>.
- 771 [31] Yanagisawa K, Odaka A, Suzuki N, Ihara Y. GM1
772 ganglioside-bound amyloid β -protein ($A\beta$): A possible
773 form of preamyloid in Alzheimer's disease. Nat Med
774 1995;1:1062–6. <https://doi.org/10.1038/nm1095-1062>.
- 775 [32] Hayashi H, Kimura N, Yamaguchi H, Hasegawa K,
776 Yokoseki T, Shibata M, et al. A Seed for Alzheimer
777 Amyloid in the Brain. J Neurosci 2004;24:4894–902.
778 <https://doi.org/10.1523/JNEUROSCI.0861-04.2004>.
- 779 [33] Eehalt R, Keller P, Haass C, Thiele C, Simons K.
780 Amyloidogenic processing of the Alzheimer β -amyloid
781 precursor protein depends on lipid rafts. J Cell Biol
782 2003;160:113–23.
783 <https://doi.org/10.1083/jcb.200207113>.
- 784 [34] Bertram L, Lange C, Mullin K, Parkinson M, Hsiao M,
785 Hogan MF, et al. Genome-wide Association Analysis
786 Reveals Putative Alzheimer's Disease Susceptibility Loci
787 in Addition to APOE. Am J Hum Genet 2008;83:623–
788 32. <https://doi.org/10.1016/j.ajhg.2008.10.008>.
- 789 [35] Hollingworth P, Harold D, Sims R, Gerrish A, Lambert
790 J-C, Carrasquillo MM, et al. Common variants at
791 ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP

- 792 are associated with Alzheimer's disease. *Nat Genet*
793 2011;43:429–35. <https://doi.org/10.1038/ng.803>.
- 794 [36] Naj AC, Jun G, Beecham GW, Wang L-S, Vardarajan
795 BN, Buross J, et al. Common variants at
796 MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are
797 associated with late-onset Alzheimer's disease. *Nat*
798 *Genet* 2011;43:436–41. <https://doi.org/10.1038/ng.801>.
- 799 [37] Bradshaw EM, Chibnik LB, Keenan BT, Ottoboni L, Raj
800 T, Tang A, et al. CD33 Alzheimer's disease locus: altered
801 monocyte function and amyloid biology. *Nat Neurosci*
802 2013;16:848–50. <https://doi.org/10.1038/nn.3435>.
- 803 [38] Jiang T, Yu J-T, Hu N, Tan M-S, Zhu X-C, Tan L. CD33
804 in Alzheimer's Disease. *Mol Neurobiol* 2014;49:529–35.
805 <https://doi.org/10.1007/s12035-013-8536-1>.
- 806 [39] Griciuc A, Serrano-Pozo A, Parrado AR, Lesinski AN,
807 Asselin CN, Mullin K, et al. Alzheimer's Disease Risk
808 Gene CD33 Inhibits Microglial Uptake of Amyloid Beta.
809 *Neuron* 2013;78:631–43.
810 <https://doi.org/10.1016/j.neuron.2013.04.014>.
- 811 [40] Engelke UFH, Zijlstra FSM, Mochel F, Valayannopoulos
812 V, Rabier D, Kluijtmans LAJ, et al. Mitochondrial
813 involvement and erythronic acid as a novel biomarker in
814 transaldolase deficiency. *Biochim Biophys Acta BBA -*

- 815 Mol Basis Dis 2010;1802:1028–35.
816 <https://doi.org/10.1016/j.bbadis.2010.06.007>.
- 817 [41] Xu J, Begley P, Church SJ, Patassini S, Hollywood KA,
818 Jüllig M, et al. Graded perturbations of metabolism in
819 multiple regions of human brain in Alzheimer’s disease:
820 Snapshot of a pervasive metabolic disorder. Biochim
821 Biophys Acta BBA - Mol Basis Dis 2016;1862:1084–92.
822 <https://doi.org/10.1016/j.bbadis.2016.03.001>.
- 823 [42] Orešič M, Hyötyläinen T, Herukka S-K, Sysi-Aho M,
824 Mattila I, Seppänen-Laakso T, et al. Metabolome in
825 progression to Alzheimer’s disease. Transl Psychiatry
826 2011;1:e57–e57. <https://doi.org/10.1038/tp.2011.55>.
- 827 [43] Gujar SK, Maheshwari S, Björkman-Burtscher I,
828 Sundgren PC. Magnetic Resonance Spectroscopy. J
829 Neuroophthalmol 2005;25:217–26.
830 <https://doi.org/10.1097/O1.WNO.0000177307.21081.81>.
- 831 [44] Jansen EEW, Gibson KM, Shigematsu Y, Jakobs C,
832 Verhoeven NM. A novel, quantitative assay for
833 homocarnosine in cerebrospinal fluid using stable-
834 isotope dilution liquid chromatography–tandem mass
835 spectrometry. J Chromatogr B 2006;830:196–200.
836 <https://doi.org/10.1016/j.jchromb.2005.10.053>.
- 837 [45] Hipkiss AR. Could Carnosine or Related Structures
838 Suppress Alzheimer’s Disease? J Alzheimers Dis

839 2007;11:229–40. <https://doi.org/10.3233/JAD-2007->
840 11210.

841 [46] Govindpani K, Calvo-Flores Guzmán B, Vinnakota C,
842 Waldvogel HJ, Faull RL, Kwakowsky A. Towards a
843 Better Understanding of GABAergic Remodeling in
844 Alzheimer’s Disease. *Int J Mol Sci* 2017;18:1813.
845 <https://doi.org/10.3390/ijms18081813>.

846

847 **Table 1. Sample characteristics of WRAP and Wisconsin**

848 **ADRC participants.**

	WRAP*		Wisconsin	
	N=161		ADRC	
	Mean	SD	Mean	SD
Age	62.1	6.5	58.1	5.5
Years of education	16.2	2.2	16.2	2.3
P-tau	17.5	6.4	15.9	5.8
T-tau	200.6	67.3	184.5	69.3
Aβ42	895.4	369.4	942.6	363.1
Aβ40	14336.1	4497.9	13897.4	4742.3
NfL	87.2	38.0	83.3	80.7
Neurogranin	798.6	307.5	728.2	286.7
YKL-40	144.6	48.3	128.7	39.1
S100b	1.1	0.3	1.2	0.3
GFAP	8.6	2.9	8.6	3.3
sTREM2	7.9	2.4	7.6	2.1
IL-6	4.3	2.6	5.2	3.7
α-synuclein	157.4	65.1	146.2	64.1
	N	%	N	%
Female	105	65.2	106	68.8
Male	56	34.8	48	31.2

849 * The summary statistics of WRAP were based on the baseline

850 measures.

851 **Table 2. Top 10 significant CSF metabolites associated with each NTK biomarker in WRAP and replicated in the**
 852 **Wisconsin ADRC.**

NTK Biomarker	Biochemical names	Compound ID	Beta	P	Adjusted P	FDR q	Super pathway	Sub-pathway
P-tau	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	24.10	9.89E-23	4.65E-21	4.65E-21	Lipid	Phosphatidylcholine (PC)
	N-acetylneuramate	32377	29.31	2.66E-21	1.25E-19	6.25E-20	Carbohydrate	Aminosugar Metabolism
	C-glycosyltryptophan	48782	29.26	1.41E-20	6.62E-19	2.21E-19	Amino Acid	Tryptophan Metabolism
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	27.21	6.95E-16	3.27E-14	8.17E-15	Lipid	Phosphatidylcholine (PC)
	stearoyl sphingomyelin (d18:1/18:0)	19503	26.20	9.38E-16	4.41E-14	8.81E-15	Lipid	Sphingolipid Metabolism
	arabitol/xylitol	48885	28.10	1.50E-15	7.04E-14	1.17E-14	Carbohydrate	Pentose Metabolism
	beta-citrylglutamate	54923	23.47	9.84E-15	4.62E-13	5.82E-14	Amino Acid	Glutamate Metabolism
	N-acetylserine	37076	35.24	9.90E-15	4.65E-13	5.82E-14	Amino Acid	Glycine, Serine and Threonine Metabolism
	sphingomyelin (d18:1/18:1, d18:2/18:0)	37529	24.14	4.10E-14	1.93E-12	2.09E-13	Lipid	Sphingolipid Metabolism
	N6-succinyladenosine	48130	22.14	4.44E-14	2.09E-12	2.09E-13	Nucleotide	Purine Metabolism, Adenine containing
T-tau	N-acetylneuramate	32377	349.51	2.25E-21	1.26E-19	1.26E-19	Carbohydrate	Aminosugar Metabolism
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	296.17	1.53E-19	8.54E-18	4.27E-18	Lipid	Phosphatidylcholine (PC)
	C-glycosyltryptophan	48782	336.75	7.77E-19	4.35E-17	1.45E-17	Amino Acid	Tryptophan Metabolism
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	343.88	3.72E-18	2.09E-16	5.21E-17	Lipid	Phosphatidylcholine (PC)
	N-acetylthreonine	33939	372.61	1.37E-17	7.68E-16	1.54E-16	Amino Acid	Glycine, Serine and Threonine

								Metabolism
	arabitol/xylitol	48885	340.87	1.61E-16	9.00E-15	1.50E-15	Carbohydrate	Pentose Metabolism
	stearoyl sphingomyelin (d18:1/18:0)	19503	314.01	6.07E-16	3.40E-14	4.85E-15	Lipid	Sphingolipid Metabolism
	N6-succinyladenosine	48130	275.91	1.19E-15	6.64E-14	8.30E-15	Nucleotide	Purine Metabolism, Adenine containing
	erythronate*	42420	608.03	1.48E-15	8.30E-14	9.23E-15	Carbohydrate	Aminosugar Metabolism
	beta-citrylglutamate	54923	281.29	6.24E-15	3.49E-13	3.49E-14	Amino Acid	Glutamate Metabolism
Aβ42	N-acetylneuraminate	32377	1831.96	7.62E-20	4.42E-18	2.65E-18	Carbohydrate	Aminosugar Metabolism
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	1922.65	9.15E-20	5.31E-18	2.65E-18	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	52461	2019.63	4.04E-19	2.34E-17	7.50E-18	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	1608.22	5.17E-19	3.00E-17	7.50E-18	Lipid	Phosphatidylcholine (PC)
	1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	19258	1779.94	5.29E-18	3.07E-16	6.13E-17	Lipid	Phosphatidylcholine (PC)
	stearoyl sphingomyelin (d18:1/18:0)	19503	1735.41	7.98E-17	4.63E-15	7.72E-16	Lipid	Sphingolipid Metabolism
	N-acetylserine	37076	2220.93	1.96E-15	1.14E-13	1.63E-14	Amino Acid	Glycine, Serine and Threonine Metabolism
	arabitol/xylitol	48885	1732.68	2.00E-14	1.16E-12	1.45E-13	Carbohydrate	Pentose Metabolism
	N-acetylthreonine	33939	1832.87	3.07E-14	1.78E-12	1.98E-13	Amino Acid	Glycine, Serine and Threonine Metabolism
Aβ40	1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)*	52470	1793.19	3.46E-14	2.01E-12	2.01E-13	Lipid	Phosphatidylcholine (PC)
	N-acetylneuraminate	32377	27190.93	2.92E-30	2.34E-28	2.34E-28	Carbohydrate	Aminosugar Metabolism

	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	27204.24	2.46E-26	1.96E-24	9.51E-25	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	23086.44	3.57E-26	2.85E-24	9.51E-25	Lipid	Phosphatidylcholine (PC)
	1-stearoyl-2-oleoyl-GPC (18:0/18:1)	52438	27203.88	1.22E-23	9.72E-22	2.43E-22	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	52461	27594.67	4.59E-23	3.67E-21	7.35E-22	Lipid	Phosphatidylcholine (PC)
	N-acetylserine	37076	34639.61	5.99E-23	4.79E-21	7.99E-22	Amino Acid	Glycine, Serine and Threonine Metabolism
	stearoyl sphingomyelin (d18:1/18:0)	19503	24933.03	8.41E-23	6.73E-21	9.61E-22	Lipid	Sphingolipid Metabolism
	arabitol/xylitol	48885	26504.66	6.06E-22	4.85E-20	5.64E-21	Carbohydrate	Pentose Metabolism
	1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	19258	24513.57	6.34E-22	5.07E-20	5.64E-21	Lipid	Phosphatidylcholine (PC)
	erythronate*	42420	47037.89	2.43E-20	1.94E-18	1.94E-19	Carbohydrate	Aminosugar Metabolism
nfl	N-acetylthreonine	33939	194.81	4.51E-04	2.93E-02	2.09E-02	Amino Acid	Glycine, Serine and Threonine Metabolism
	N-acetylalanine	1585	262.23	9.53E-04	6.19E-02	2.09E-02	Amino Acid	Alanine and Aspartate Metabolism
	beta-citrylglutamate	54923	150.23	9.66E-04	6.28E-02	2.09E-02	Amino Acid	Glutamate Metabolism
	arabitol/xylitol	48885	165.78	1.70E-03	1.10E-01	2.26E-02	Carbohydrate	Pentose Metabolism
	1-palmitoyl-GPC (16:0)	33955	142.17	1.74E-03	1.13E-01	2.26E-02	Lipid	Lysophospholipid
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	154.06	2.68E-03	1.74E-01	2.53E-02	Lipid	Phosphatidylcholine (PC)
	1-oleoyl-GPC (18:1)	48258	136.99	2.72E-03	1.77E-01	2.53E-02	Lipid	Lysophospholipid
	stearoyl sphingomyelin (d18:1/18:0)	19503	145.43	3.35E-03	2.18E-01	2.72E-02	Lipid	Sphingolipid Metabolism

	orotidine	35172	129.73	4.56E-03	2.96E-01	3.00E-02	Nucleotide	Pyrimidine Metabolism, Orotate containing
	cysteine	1868	187.12	4.98E-03	3.24E-01	3.00E-02	Amino Acid	Methionine, Cysteine, SAM and Taurine Metabolism
Neurogranin	N-acetylneuraminate	32377	1574.23	6.97E-26	4.32E-24	4.32E-24	Carbohydrate	Aminosugar Metabolism
	C-glycosyltryptophan	48782	1442.13	9.29E-20	5.76E-18	2.88E-18	Amino Acid	Tryptophan Metabolism
	N6-succinyladenosine	48130	1234.78	3.19E-18	1.98E-16	6.60E-17	Nucleotide	Purine Metabolism, Adenine containing
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	1212.87	1.38E-17	8.53E-16	1.89E-16	Lipid	Phosphatidylcholine (PC)
	arabitol/xylitol	48885	1470.51	1.53E-17	9.48E-16	1.89E-16	Carbohydrate	Pentose Metabolism
	N-acetylthreonine	33939	1549.35	1.83E-17	1.13E-15	1.89E-16	Amino Acid	Glycine, Serine and Threonine Metabolism
	erythronate*	42420	2626.14	1.19E-16	7.40E-15	1.06E-15	Carbohydrate	Aminosugar Metabolism
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	1370.31	5.60E-16	3.47E-14	4.34E-15	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-GPC (16:0)	33955	1202.61	1.09E-15	6.78E-14	7.54E-15	Lipid	Lysophospholipid
N-acetylserine	37076	1719.96	2.67E-15	1.65E-13	1.65E-14	Amino Acid	Glycine, Serine and Threonine Metabolism	
YKL-40	N-acetylneuraminate	32377	162.38	7.62E-17	1.37E-15	1.37E-15	Carbohydrate	Aminosugar Metabolism
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	169.92	1.71E-16	3.08E-15	1.54E-15	Lipid	Phosphatidylcholine (PC)
	stearoyl sphingomyelin (d18:1/18:0)	19503	160.90	8.86E-16	1.60E-14	5.32E-15	Lipid	Sphingolipid Metabolism
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	140.55	1.05E-14	1.89E-13	4.65E-14	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	52461	171.39	1.29E-14	2.32E-13	4.65E-14	Lipid	Phosphatidylcholine (PC)

	arabitol/xylitol	48885	162.00	7.35E-14	1.32E-12	2.21E-13	Carbohydrate	Pentose Metabolism
	1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	19258	147.07	4.57E-13	8.22E-12	1.17E-12	Lipid	Phosphatidylcholine (PC)
	N6-succinyladenosine	48130	121.32	2.81E-11	5.06E-10	6.32E-11	Nucleotide	Purine Metabolism, Adenine containing
	cysteine	1868	181.51	6.78E-11	1.22E-09	1.36E-10	Amino Acid	Methionine, Cysteine, SAM and Taurine Metabolism
	1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)*	52470	148.93	1.16E-10	2.09E-09	2.09E-10	Lipid	Phosphatidylcholine (PC)
S100b	stearoyl sphingomyelin (d18:1/18:0)	19503	0.81	8.99E-07	2.34E-05	2.34E-05	Lipid	Sphingolipid Metabolism
	1-stearoyl-2-docosaheptaenoyl- GPC (18:0/22:6)	52611	0.54	3.45E-05	8.97E-04	4.49E-04	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	52461	0.73	6.95E-05	1.81E-03	6.03E-04	Lipid	Phosphatidylcholine (PC)
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	0.67	1.07E-04	2.79E-03	6.96E-04	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-docosaheptaenoyl- GPC (16:0/22:6)	52610	0.50	1.59E-04	4.14E-03	8.27E-04	Lipid	Phosphatidylcholine (PC)
	sphingomyelin (d18:1/18:1, d18:2/18:0)	37529	0.60	2.07E-04	5.39E-03	8.99E-04	Lipid	Sphingolipid Metabolism
	erythronate*	42420	1.18	3.28E-04	8.52E-03	1.22E-03	Carbohydrate	Aminosugar Metabolism
	palmitoyl sphingomyelin (d18:1/16:0)	37506	0.55	7.47E-04	1.94E-02	2.43E-03	Lipid	Sphingolipid Metabolism
1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	0.50	9.45E-04	2.46E-02	2.73E-03	Lipid	Phosphatidylcholine (PC)	
sphingomyelin (d18:2/16:0,	42459	0.44	1.53E-03	3.97E-02	3.96E-03	Lipid	Sphingolipid Metabolism	

	d18:1/16:1)*							
GFAP	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	10.08	1.76E-07	1.06E-06	1.06E-06	Lipid	Phosphatidylcholine (PC)
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	8.07	9.91E-07	5.95E-06	2.97E-06	Lipid	Phosphatidylcholine (PC)
	beta-citrylglutamate	54923	7.74	6.92E-06	4.15E-05	1.38E-05	Amino Acid	Glutamate Metabolism
	N-acetylneuraminate	32377	8.00	1.59E-05	9.56E-05	2.39E-05	Carbohydrate	Aminosugar Metabolism
	gulonate*	46957	6.71	2.75E-05	1.65E-04	3.30E-05	Cofactors and Vitamins	Ascorbate and Aldarate Metabolism
	arabinose	575	9.80	3.71E-04	2.22E-03	3.71E-04	Carbohydrate	Pentose Metabolism
sTREM2	stearoyl sphingomyelin (d18:1/18:0)	19503	9.59	5.72E-16	1.60E-14	1.60E-14	Lipid	Sphingolipid Metabolism
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	9.81	1.59E-15	4.45E-14	2.23E-14	Lipid	Phosphatidylcholine (PC)
	palmitoyl sphingomyelin (d18:1/16:0)	37506	9.18	2.52E-15	7.05E-14	2.35E-14	Lipid	Sphingolipid Metabolism
	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	52461	10.20	9.68E-15	2.71E-13	6.78E-14	Lipid	Phosphatidylcholine (PC)
	cholesterol	63	9.08	2.08E-14	5.83E-13	1.17E-13	Lipid	Sterol
	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	7.89	2.11E-13	5.90E-12	9.83E-13	Lipid	Phosphatidylcholine (PC)
	sphingomyelin (d18:2/16:0, d18:1/16:1)*	42459	7.19	3.53E-13	9.88E-12	1.41E-12	Lipid	Sphingolipid Metabolism
	C-glycosyltryptophan	48782	8.51	2.09E-12	5.86E-11	7.33E-12	Amino Acid	Tryptophan Metabolism
	N-acetylneuraminate	32377	8.35	2.55E-12	7.13E-11	7.92E-12	Carbohydrate	Aminosugar Metabolism
1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	19258	8.45	3.18E-12	8.90E-11	8.90E-12	Lipid	Phosphatidylcholine (PC)	

α-synuclein	1-palmitoyl-2-stearoyl-GPC (16:0/18:0)	52616	306.42	1.42E-22	1.42E-20	1.42E-20	Lipid	Phosphatidylcholine (PC)
	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	337.75	5.02E-20	5.02E-18	2.51E-18	Lipid	Phosphatidylcholine (PC)
	N-acetylneuraminate	32377	315.50	2.29E-19	2.29E-17	7.64E-18	Carbohydrate	Aminosugar Metabolism
	stearoyl sphingomyelin (d18:1/18:0)	19503	316.41	1.31E-18	1.31E-16	3.28E-17	Lipid	Sphingolipid Metabolism
	1-stearoyl-2-oleoyl-GPC (18:0/18:1)	52438	334.85	8.69E-18	8.69E-16	1.74E-16	Lipid	Phosphatidylcholine (PC)
	C-glycosyltryptophan	48782	304.78	3.58E-17	3.58E-15	5.82E-16	Amino Acid	Tryptophan Metabolism
	palmitoyl sphingomyelin (d18:1/16:0)	37506	297.65	4.24E-17	4.24E-15	5.82E-16	Lipid	Sphingolipid Metabolism
	N-acetylthreonine	33939	343.71	4.65E-17	4.65E-15	5.82E-16	Amino Acid	Glycine, Serine and Threonine Metabolism
	N6-succinyladenosine	48130	266.82	9.85E-17	9.85E-15	1.09E-15	Nucleotide	Purine Metabolism, Adenine containing
	1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	19258	300.75	1.17E-16	1.17E-14	1.17E-15	Lipid	Phosphatidylcholine (PC)

853 * Indicates a compound that has not been confirmed based on a standard, but Metabolon was confident in its identity.

854 **Table 3. Prediction performance (r^2) of metabolites in the combined cohort of WRAP and Wisconsin ADRC.**

NTK biomarkers	Base model r^2 (sample size*)	Number of input metabolites	Metabolite model r^2 (sample size*)	Number of elastic net-selected metabolites	Elastic net-selected metabolites model r^2 (sample size*)
P-tau	0.11 (n=295)	40	0.72 (n=226)	22	0.70 (n=232)
T-tau	0.10 (n=296)	44	0.71 (n=236)	17	0.69 (n=281)
Aβ42	0.01 (n=296)	54	0.55 (n=220)	43	0.54 (n=222)
Aβ40	0.04 (n=293)	73	0.85 (n=162)	70	0.83 (n=166)
NfL	0.10 (n=297)	23	0.23(n=280)	13	0.22 (n=290)
Neurogranin	0.07 (n=297)	51	0.78 (n=245)	32	0.76 (n=267)
YKL-40	0.25 (n=297)	15	0.57 (n=281)	10	0.56 (n=288)
S100	0.05 (n=296)	20	0.22 (n=281)	4	0.19 (n=293)
GFAP	0.17 (n=297)	6	0.35 (n=282)	5	0.35(n=288)
sTREM2	0.05 (n=297)	27	0.52 (n=276)	14	0.50 (n=283)
α-synuclein	0.03 (n=297)	90	0.97 (n=105)	27	0.68 (n=218)

855 Variables included in the base model were age, sex, years of education and cohort.

856 Variables included in the metabolite model were age, sex, years of education, cohort and all replicated significant metabolites for
857 each biomarker.

858 Variables included in the metabolite model were age, sex, years of education, cohort and elastic net-selected metabolites.

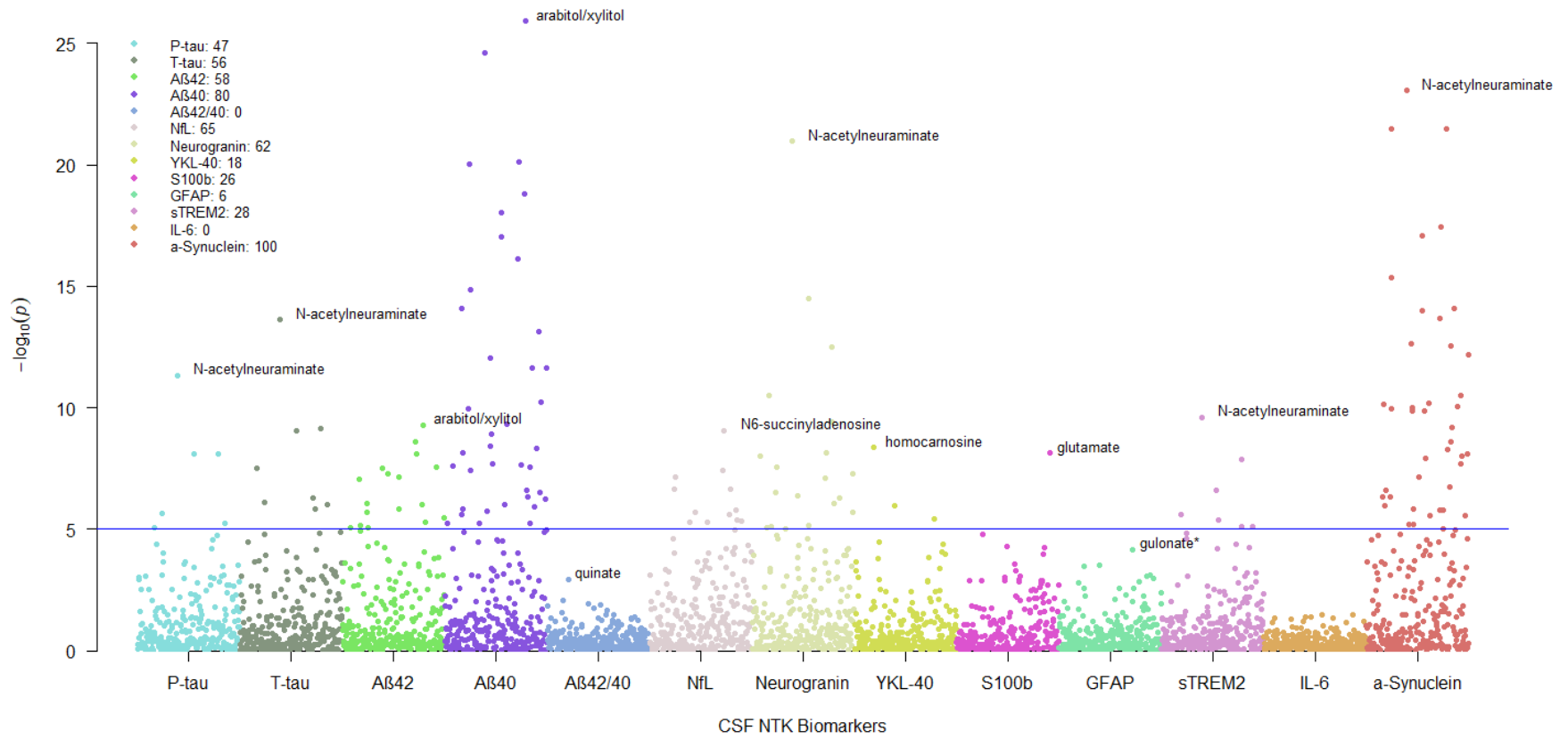
859 * The sample sizes were different because of the missingness of metabolites.

860 **Table 4. Significant Mendelian randomization results after Bonferroni correction.**

NTK biomarkers	Metabolite	Compound ID	nSNPs	nRegions (1Mbps)	F statistics	LIML estimate	LIML 95% CI		LIML p	Adjusted p
sTREM2	palmitoyl sphingomyelin (d18:1/16:0)	37506	67	21	1.97	12.97	9.14	16.80	2.09E-10	2.50E-08
A β 40	erythritol	20699	100	34	6.09	12499.02	7539.22	17458.82	1.40E-06	1.61E-04
α -synuclein	homocarnosine	1633	43	14	2.61	-123.04	-172.61	-73.47	1.96E-06	2.23E-04
T-tau	1-palmitoyl-2-stearoyl- GPC (16:0/18:0)	52616	19	8	2.81	319.63	187.99	451.28	3.21E-06	3.59E-04
α -synuclein	erythritol	20699	100	34	5.95	172.93	99.70	246.16	5.69E-06	6.31E-04
Neurogranin	1-palmitoyl-2-stearoyl- GPC (16:0/18:0)	52616	19	8	2.82	1433.89	799.21	2068.57	1.37E-05	1.49E-03
A β 40	1-myristoyl-2-palmitoyl- GPC (14:0/16:0)	19258	38	20	2.48	18666.39	10385.64	26947.14	1.43E-05	1.54E-03
T-tau	1-palmitoyl-2-stearoyl- GPC (16:0/18:0)	52616	19	8	2.79	28.06	15.50	40.62	1.70E-05	1.82E-03
α -synuclein	1-palmitoyl-2-stearoyl- GPC (16:0/18:0)	52616	19	8	2.82	268.34	137.78	398.90	7.16E-05	7.59E-03
A β 40	1,2-dipalmitoyl-GPC (16:0/16:0)	19130	5	2	6.78	25973.09	12768.09	39178.09	1.41E-04	1.48E-02

A β 40	homocarnosine	1633	43	14	2.55	-6548.41	-9903.63	-3193.19	1.58E-04	1.65E-02
Neurogranin	homocarnosine	1633	43	14	2.61	-444.47	-675.36	-213.58	1.93E-04	1.99E-02
A β 40	gulonate*	46957	25	13	6.34	12658.31	6029.03	19287.58	2.17E-04	2.21E-02

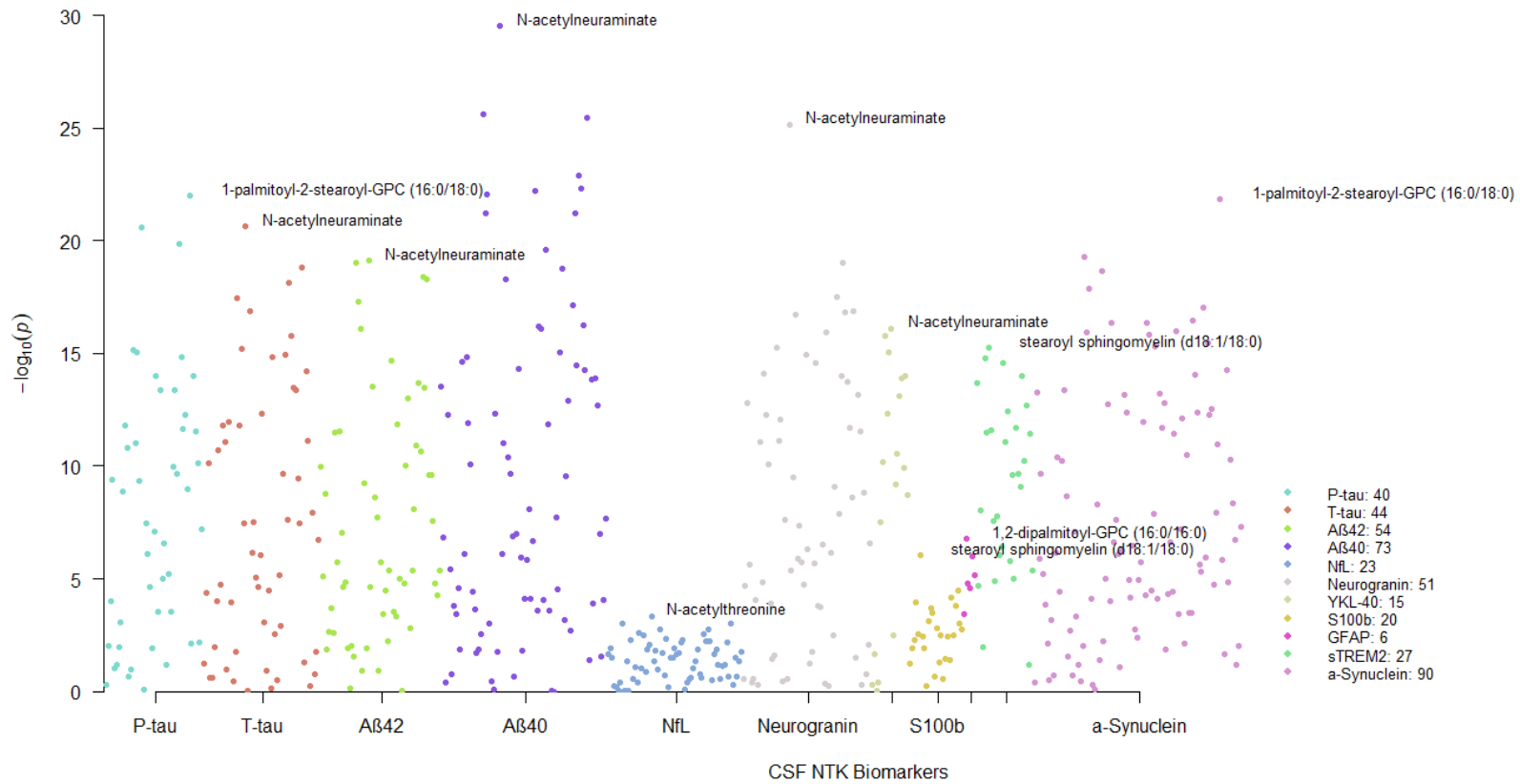
861 The nSNPs refers to the number of SNPs in each IV, and the nRegions is the approximate number of regions defined by up to a 1Mbps
862 of the SNPs. The F statistic represent the strength of the IV (strong IV F statistic >10). The estimate of beta, confidence interval and p
863 values were all based on the limited information maximum likelihood (LIML) method.



(a)

864

865



(b)

866

867

868 **Figure 1. Manhattan plots of MWAS results between CSF metabolites and CSF NTK biomarkers.** Each dot represents a

869 metabolite and the different colors represent the CSF NTK biomarkers (x-axis) in (a) WRAP and (b) the Wisconsin ADRC (only

870 significant metabolites after FDR correction in WRAP were included). The $-\log_{10}(\text{p-value})$ is shown on the y-axis. The legend box
871 indicates the number of metabolites that were significant after FDR correction for each NTK biomarker.

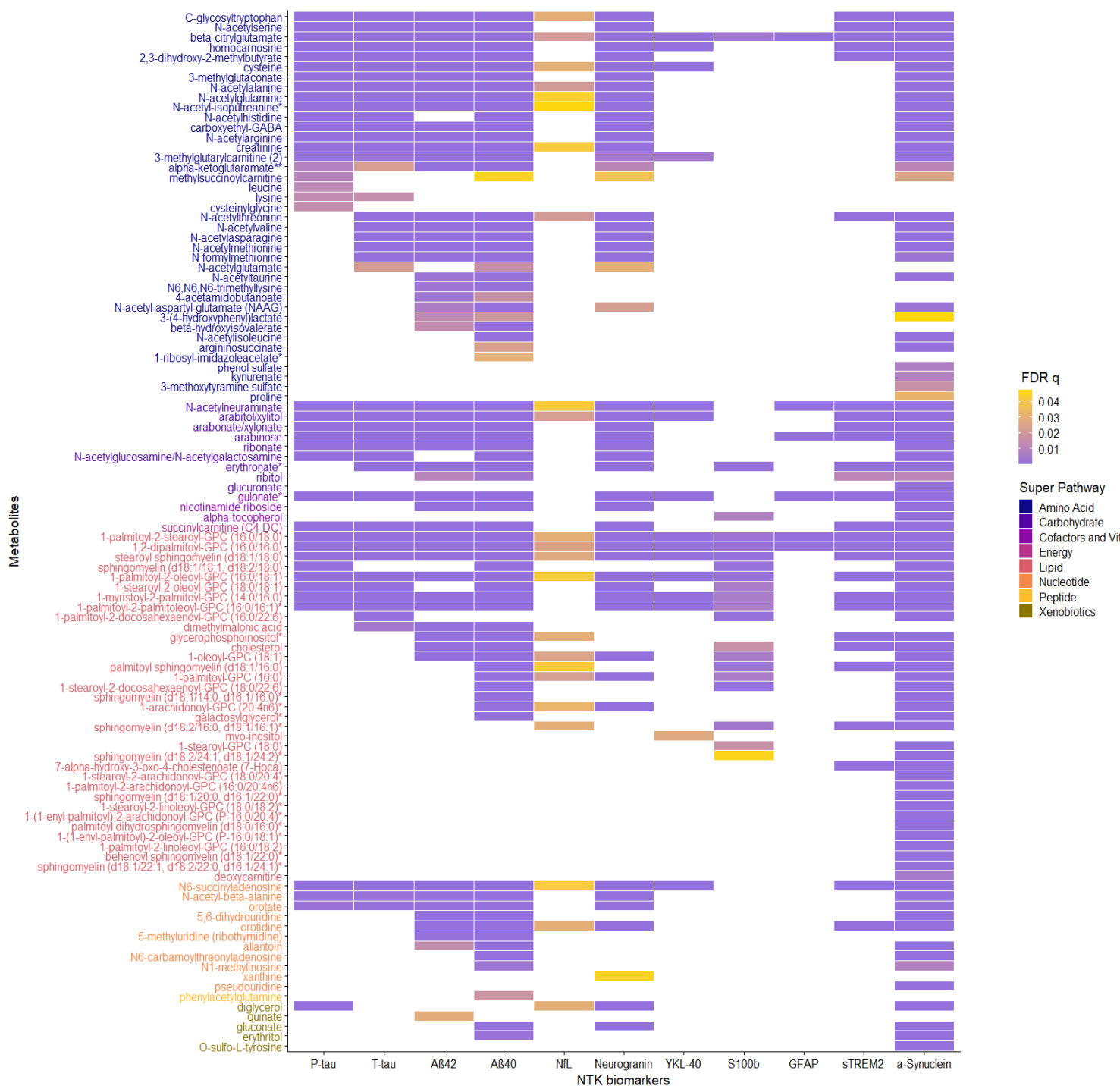


Figure 2. The association patterns between significant CSF metabolites and CSF NTK

biomarkers in Wisconsin-ADRC. Each cell represents the association of a CSF metabolites with a

biomarker. The color scale indicates the magnitude of the FDR q values. The metabolites are also grouped and

colored based on their super pathway.