

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

1 Title: Gut microbiome compositional and functional features associate with 2 Alzheimer's disease pathology

3

4 Authors:

5 Jea Woo Kang^{1*}, Lora A. Khatib^{2,3*}, Margo B. Heston^{1*}, Amanda H. Dilmore², Jennifer S.
6 Labus^{4,5,6}, Yuetiva Deming¹, Leyla Schimmel⁷, Colette Blach⁸, Daniel McDonald², Antonio
7 Gonzalez², MacKenzie Bryant², Karenina Sanders², □ara Schwartz², Tyler K. Ulland^{1,9}, Sterling
8 C. Johnson¹, Sanjay Asthana¹, Cynthia M. Carlsson¹, Nathaniel A. Chin¹, Kaj Blennow¹⁰, Henrik
9 Zetterberg^{10,11,12,13,14,1}, Federico E. Rey¹⁵, Alzheimer Gut Microbiome Project Consortium⁷, Rima
10 Kaddurah-Daouk^{7,16,17}, Rob Knight^{2,18,19,20,21}, and Barbara B. Bendlin^{1,22,#}

11

12 *These three authors contributed equally to this work.

13 #Correspondence

14 Affiliations and addresses:

- 15 1. Wisconsin Alzheimer's Disease Research Center, University of Wisconsin School of
16 Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA
17 Address: 600 Highland Ave, J5/1 Mezzanine, Madison, WI, USA 53792
- 18 2. Department of Pediatrics, University of California San Diego, La Jolla, California, USA
19 Address: 9461 Gilman Dr, La Jolla, CA, USA 92093
- 20 3. Neurosciences Graduate Program, University of California San Diego, La Jolla, California,
21 USA
22 Address: 9500 Gilman Dr, La Jolla, CA, USA 92093
- 23 4. Integrative Biostatistics and Bioinformatics Core (IBBC) at the Goodman-Luskin Microbiome
24 Center
25 Address: 42-210 CHS, Los Angeles, CA, USA 90095
- 26 5. G. Oppenheimer Center for Neurobiology of Stress and Resilience
27 Address: 10833 Le Conte Ave, Los Angeles, CA, USA 90095
- 28 6. UCLA Vatche and Tamar Manoukian Division of Digestive Diseases, David Geffen School of
29 Medicine at UCLA
30 Address: 100 Medical Plaza, Los Angeles, CA, USA 90095
- 31 7. Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA
32 Address: 905 W Main St, Durham, NC, USA 27701
- 33 8. Duke Molecular Physiology Institute, Duke University, Durham, NC, USA
34 Address: 300 N Duke St, Durham, NC, USA 27701
- 35 9. Department of Pathology and Laboratory Medicine, University of Wisconsin-Madison,
36 Madison, WI, USA
37 Address: 1685 Highland Ave, Madison, WI, USA 53705
- 38 10. Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology,
39 the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 40 Address: Blå stråket 15, vån 3 SU/Sahlgrenska 413 45 Göteborg, Sweden
41 **11.** Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
42 Address: Blå stråket 5, 413 45 Göteborg, Sweden
43 **12.** Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square,
44 London, UK
45 Address: Queen Square, London WC1N 3BG, United Kingdom
46 **13.** UK Dementia Research Institute at UCL, London, UK
47 Address: 6th Floor, Maple House, Tottenham Ct Rd, London W1T 7NF, United Kingdom
48 **14.** Hong Kong Center for Neurodegenerative Diseases, Clear Water Bay, Hong Kong, China
49 Address: Units 1501-1502, 1512-1518, 15/F, Building 17W, Hong Kong Science Park,
50 Shatin, N.T., Hong Kong
51 **15.** Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA
52 Address: 1550 Linden Dr, Madison, WI, USA 53706
53 **16.** Duke Institute of Brain Sciences, Duke University, Durham, NC, USA
54 Address: 308 Research Dr, Durham, NC, USA 27710
55 **17.** Department of Medicine, Duke University, Durham, NC, USA
56 Address: 40 Duke Medicine Circle, 124 Davison Building, Durham, NC, USA 27710
57 **18.** Center for Microbiome Innovation, Joan and Irwin Jacobs School of Engineering, University
58 of California San Diego, La Jolla, California, USA
59 Address: Franklin Antonio Hall, Jacobs School of Engineering, 9500 Gilman Dr, La Jolla,
60 CA, USA 92093
61 **19.** Department of Computer Science and Engineering, University of California, San Diego, La
62 Jolla, CA, USA
63 Address: 3235 Voigt Dr, La Jolla, CA, USA 92093
64 **20.** Halicioğlu Data Science Institute, University of California, San Diego, La Jolla, CA, USA
65 Address: 3234 Matthews Ln, La Jolla, CA, USA 92093
66 **21.** Shu Chien-Gene Lay Department of Bioengineering, University of California, San Diego, La
67 Jolla, CA, USA
68 Address: 3223 Voigt Dr, La Jolla, CA, USA 92093
69 **22.** Wisconsin Alzheimer's Institute, School of Medicine and Public Health, University of
70 Wisconsin-Madison, Madison, WI, USA
71 Address: 610 Walnut Street, 9th Floor, Madison, WI, USA 53726

72
73
74
75
76
77
78

79 **Abstract**

80 BACKGROUND: The gut microbiome is a potentially modifiable factor in Alzheimer's disease
81 (AD); however, understanding of its composition and function regarding AD pathology is limited.

82
83 METHODS: Shallow-shotgun metagenomic data was used to analyze fecal microbiome from
84 participants enrolled in the Wisconsin Microbiome in Alzheimer's Risk Study, leveraging clinical
85 data and cerebrospinal fluid (CSF) biomarkers. Differential abundance and ordinary least
86 squares regression analyses were performed to find differentially abundant gut microbiome
87 features and their associations with CSF biomarkers of AD and related pathologies.

88
89 RESULTS: Gut microbiome composition and function differed between people with AD and
90 cognitively unimpaired individuals. The compositional difference was replicated in an
91 independent cohort. Differentially abundant gut microbiome features were associated with CSF
92 biomarkers of AD and related pathologies.

93
94 DISCUSSION: These findings enhance our understanding of alterations in gut microbial
95 composition and function in AD, and suggest that gut microbes and their pathways are linked to
96 AD pathology.

97
98
99

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120

121 **1 BACKGROUND**

122 The human gut microbiome is recognized as an important modifiable factor in health and
123 disease. It is related to overall gut health by maintaining gut barrier integrity and gut immune
124 homeostasis via balanced composition and production of microbial metabolites such as short-
125 chain fatty acids (SCFAs).¹⁻³ However, in certain disease states, including Alzheimer's disease
126 (AD), gut microbiome composition and its metabolic changes may alter and exacerbate the
127 disease.

128 Compositional differences in gut microbiota, including relative abundance and diversity,
129 have been observed between control and AD groups.⁴ Other studies have reported that gut
130 microbiome composition is altered among people with AD dementia, individuals with mild
131 cognitive impairment (MCI), or preclinical AD compared with healthy controls.⁵⁻⁸ To better
132 determine the relationship between gut microbiome and AD pathology, studies have leveraged
133 measures of AD biomarkers obtained via cerebrospinal fluid (CSF) analysis,⁹ positron emission
134 tomography (PET),¹⁰ and plasma.¹¹ Additionally, inflammatory markers have been utilized to find
135 associations with gut inflammation-driven AD pathology.^{12,13}

136 The shift towards functional analysis provides deeper insights into the functional
137 potential of the microbiome, which is important given that multiple lines of evidence indicate that
138 gut microbial pathways and associated metabolites influence disease development, as well as
139 providing the opportunity to identify therapeutic targets.¹⁴⁻¹⁷ A small number of studies have
140 examined gut microbiome composition and function together with markers of AD pathology in
141 humans,^{8,18-20} but many other studies to date are limited to the composition of gut microbes in a
142 single cohort. The Alzheimer Gut Microbiome Project (AGMP) initiative continues to leverage gut
143 microbiome and metabolome to better understand metabolic processes that influence AD
144 pathology.

145 To identify differences in composition and function as well as relationships between gut
146 microbiome and AD pathology, this study collected stool samples from participants enrolled in
147 the Microbiome in Alzheimer's Risk Study (MARS). Differences in microbiome diversity as well
148 as abundance were compared between AD-related groups (diagnosis, amyloid status, and
149 *APOE ε4* status), and the co-occurrence of the common gut microbiota features was analyzed
150 among comparison groups. In addition, validation of gut microbiota composition that was
151 differentially abundant between AD compared to healthy controls was performed in a larger
152 cohort of participants who are part of the AGMP. Furthermore, determination of microbiome
153 functional differences was compared between people with AD dementia and cognitively
154 unimpaired (CU) individuals. Lastly, associations between any differentially abundant gut
155 microbiome features were examined in relation to CSF biomarkers of AD and related
156 pathologies to identify the potential microbes and microbial pathways that relate to AD
157 pathology. We hypothesized that gut microbiome alterations in composition and function would
158 be present among people with AD dementia compared to CU, as well as associate with AD
159 pathology.

160

161 2 METHODS

162 2.1 Participants

163 Participants included in this study were recruited from the Wisconsin Alzheimer's
164 Disease Research Center (ADRC) Clinical Core and the Wisconsin Registry for Alzheimer's
165 Prevention (WRAP).²¹ The WRAP study enrolled participants between the ages of 40–65 years
166 at study entry, and the cohort is enriched for parental history of AD dementia. The Wisconsin
167 ADRC clinical core enrolls participants who span the clinical and biological spectrum of AD,
168 from those who are CU to individuals with mild cognitive impairment (MCI) and AD dementia.
169 Participants underwent *APOE* genotyping using competitive allele-specific PCR-based KASP™
170 genotyping assays (LGC Genomics, Beverly, MA)²² as well as longitudinal assessments of
171 cognition and laboratory tests. Biomarkers of AD determined with CSF collection and PET
172 neuroimaging were collected in a subset of the cohort. Participants underwent fecal sample
173 collection as part of their participation in MARS, which was used to analyze the gut microbiome.
174 Participants completed questionnaires including medical history and diet, at the time of fecal
175 sample collection.

176 Participant diagnosis of AD was determined by a multidisciplinary consensus diagnostic
177 panel and based on the National Institute on Aging–Alzheimer's Association (NIA-AA)
178 criteria.^{23,24} Participants underwent dynamic [¹¹C]Pittsburgh compound B (PiB) PET scans and
179 lumbar puncture for CSF collection to determine their amyloid status. Amyloid positivity on PET
180 imaging was achieved by the visual rating (1.19 or greater)²⁵ from a global PiB distribution
181 volume ratio (DVR) and determined for CSF via the A β ₄₂/A β ₄₀ ratio (less than 0.046).²⁶ The
182 study procedures were approved by the University of Wisconsin Institutional Review Board, and
183 all participants provided signed or oral informed consent.

184 An independent cohort of participants who are part of the AGMP and who were recruited
185 from multiple NIA-funded ADRC across the U.S. (n = 448, **Table S1**) was included for validation
186 of differential abundance analysis. AGMP participants from Wisconsin were excluded to ensure
187 a unique validation sample.
188

189 2.2 Fecal sample collection and metagenomic data sequencing

190 Fecal samples were collected as previously described.⁴ Briefly, participants collected
191 their stool samples at home with provided fecal collection kits. Participants returned their
192 samples in insulated containers which were chilled with frozen gel packs. Returned samples
193 were immediately weighed and scored on the Bristol stool scale. Fecal samples were then
194 subsampled using sterile straws and stored at –80°C until processing.

195 Fecal samples were processed for DNA extraction as previously described.²⁷ Briefly,
196 samples were extracted using the MoBio PowerMag Soil DNA isolation kit with a magnetic bead
197 plate. Extracted genomic DNA (gDNA) was quantified with Quant-iT PicoGreen double-stranded
198 DNA (dsDNA) assay kit (Thermo Fisher Scientific Inc.), and underwent a miniaturized KAPA
199 HyperPlus library preparation using an iTru indexing strategy.^{28,29} Library was normalized
200 pooled based on concentration, PCR cleaned, and then size selected (300-700 bp) on the Sage
201 Science PippinHT. Libraries were sequenced on an Illumina NovaSeq 6000 as a paired-end

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

202 150-cycle run at the University of California San Diego (UCSD) IGM Genomics Center as part of
203 AGMP initiative.³⁰
204

205 2.3 Metagenomic data processing

206 The metagenomic data processing was performed as previously described.³¹ The
207 sequence data were filtered for all adapters known to fastp (version 0.23.4) in paired-end mode
208 by explicitly specifying a known adapters file.³² Fastp also removed sequences shorter than 45
209 nucleotides with -l, a flag to filter the minimum length of each sequence. Each sample was then
210 filtered against each genome in the human pangenome,³³ as well as both T2T-CHM13v2.0³⁴
211 and GRCh38,³⁵ using minimap2³⁶ (version 2.26-r1175) with "-ax sr" for short read mode. The
212 data were first run in paired-end mode, and then run in single-end mode, per genome. Each
213 successive run was converted from SAM to FASTQ using samtools³⁷ (version 1.17) with
214 arguments -f 12 -F 256 -N for paired-end data and -f 4 -F 256 for single-end. The single-end
215 data are repaired using fastq_pair³⁸ (version 1.0) specifying a table size of 50M with -t. Compute
216 support was provided with GNU Parallel³⁹ (version 20180222). Single-end FASTQ output from
217 samtools was split into R1 and R2 with a custom Rust program, with rust-bio for parsing⁴⁰
218 (version 1.4.0). Data were multiplexed with sed and demultiplexed using a custom Python script.
219 Shotgun sequencing data were then uploaded to and processed through Qiita⁴¹ (Study ID
220 13663). Sequence adapter and host filtering were executed using qp-fastp-minimap2 version
221 2022.04. Subsequently, Woltka⁴² version 0.1.4 (qp-woltka 2022.09) with the Web of Life 2
222 database was employed for taxonomic and functional predictions. Genomic coverages were
223 computed, and features with less than 25% coverage were excluded.⁴³ To enhance data quality,
224 a prevalence filter using QIIME 2 v2023.5⁴⁴ was applied, eliminating features present in less
225 than 10% of samples and samples with a sampling depth of less than 500,000 reads to mitigate
226 the inclusion of erroneous and low-quality reads. The resulting feature table was utilized for
227 downstream analysis.

228

229 2.4 Biomarker measurements

230 2.4.1 CSF biomarkers

231 CSF samples were collected via lumbar puncture in the morning after fasting for 8-12
232 hours as previously described.²⁶ CSF biomarkers were measured using the NeuroToolKit
233 (NTK), a panel of exploratory robust prototype assays (Roche Diagnostics International Ltd,
234 Rotkreuz, Switzerland). The following biomarkers were quantified on the Cobas[®] e 601 module
235 (Roche Diagnostics International Ltd, Rotkreuz, Switzerland): A β ₄₂, pTau₁₈₁, tTau, S100 calcium
236 binding protein B (S100B), and interleukin-6 (IL-6), and the remaining biomarkers were assayed
237 on the Cobas[®] e 411 analyzer: A β ₄₀, neurofilament light protein (NfL), neurogranin, α -synuclein,
238 glial fibrillary acidic protein (GFAP), chitinase-3-like protein 1 (YKL-40), and soluble triggering
239 receptor expressed on myeloid cells 2 (sTREM2). Each biomarker was measured as markers of
240 AD and related pathologies which were amyloid pathology (A β ₄₂/A β ₄₀), tau pathophysiology
241 (pTau₁₈₁ and tTau), neurodegeneration (NfL), synaptic dysfunction and injury (neurogranin and
242 α -synuclein), inflammation (IL-6), and glial activation (S100B, GFAP, YKL-40, and sTREM2).

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

243 2.4.2 PiB PET biomarker

244 Dynamic ^{11}C -PiB scans were acquired using a Siemens ECAT EXACT HR+ tomograph
245 as previously described.⁴⁵ DVRs were estimated with Logan graphical analysis and a threshold
246 of 1.19 for global DVR was used to determine PiB status which was used along with CSF
247 $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratio to confirm amyloid status of participants.
248

249 2.5 Statistical analysis

250 Statistical analysis on participant demographics was performed across clinical diagnoses
251 using the Kruskal-Wallis rank sum test for continuous variables and Pearson's Chi-squared test
252 for categorical variables. Analysis with multiple comparisons was corrected for multiple tests
253 employing the Bonferroni correction method. Gut microbiome diversities were calculated using
254 QIIME 2 tools.⁴⁴ Alpha diversity indices were calculated including Shannon,⁴⁶ Evenness,⁴⁷ and
255 Faith's phylogenetic diversity (PD).⁴⁸ Beta diversity indices were calculated including Bray–
256 Curtis dissimilarity,⁴⁹ and Weighted⁵⁰ and Unweighted⁵¹ UniFrac. The Bayesian Inferential
257 Regression for Differential Microbiome Analysis (BIRDMAn) pipeline was used for microbiome
258 differential abundance (DA) analysis.⁵² Microbiome features (composition and function) were
259 outcome variables and each AD-related group (clinical diagnosis, amyloid status, and *APOE* $\epsilon 4$
260 status) was a predictor adjusting for covariates including age, sex, BMI, Bristol type, medication
261 status, and age difference between fecal collection and measurements of each predictor
262 variable. Log ratios of Top and Bottom features of each AD-related group (log[Top
263 features/Bottom features]) were calculated and analyzed using the Mann–Whitney *U* test to
264 identify similarities in microbiome features that differ between AD-related groups. Venn
265 diagrams were created to illustrate any overlapping taxonomies across AD groups in each Top
266 (more abundant in AD-related groups) and Bottom (less abundant in AD-related groups) group
267 at each taxonomic level (phylum, family, genus, and species). Log ratios of Top and Bottom
268 features of each AD group (log[Top features/Bottom features]) were also calculated and
269 analyzed using the Mann–Whitney *U* test in MARS and validation cohorts to identify similarities
270 in microbiome features that differ between AD and CU groups in both cohorts. Log ratios of Top
271 and Bottom functional features (log[Top features/Bottom features]) were calculated and
272 statistical significance was determined by the Mann–Whitney *U* test between AD and CU
273 groups. Differentially abundant KEGG Orthology (KO)⁵³ pathways and their associated species
274 were determined using BIRDMAn.⁵² Robust Aitchison principal component analysis (RPCA)
275 from Gemelli (version 0.0.10) was used to analyze sparse compositional KO⁵³ microbiome
276 pathway features that are separated by sample variations.⁵⁴ RPCA results were visualized with
277 scores plots and biplots. Statistical analysis on RPCA was performed with permutational
278 multivariate analysis of variance (PERMANOVA)⁵⁵ between groups.

279 To explore the relationships between key microbial features and pathways identified via
280 BIRDMAn and CSF biomarkers of AD and related pathologies, we applied an ordinary least
281 squares (OLS) linear regression approach. Prior to fitting the linear regression model, CSF
282 biomarkers were standardized using the StandardScaler (version 0.24.1) from the scikit-learn
283 library.⁵⁶ To address issues with sparse, compositional data, we used the
284 multiplicative_replacement function from the scikit-bio (version 0.5.7) `skbio.stats.composition`
285 module to preprocess the metagenomics data. This function replaces zeros with small positive

286 values, preserving the compositional nature of the data. Subsequently, a centered log ratio
287 (CLR) transformation was applied to the metagenomics data to account for compositionality.
288 Finally, ordinary least squares (OLS) linear regression was performed on the microbial features
289 differentially abundant in AD versus CU, in relation to each CSF biomarker. The results were
290 visualized using heatmaps.

291 All other statistical analyses were performed using Python libraries SciPy (1.13.0),⁵⁷
292 scikit-learn (1.4.2),⁵⁶ and NumPy (version 1.26.4).⁵⁸ All figures were generated using Python
293 libraries Matplotlib (version 3.6.0)⁵⁹ and Seaborn (version 0.11.2).⁶⁰
294

295 3 RESULTS

296 3.1 Participant demographics

297 Participant characteristics are shown by clinical diagnosis in **Table 1**. Participants were
298 aged between 47-93 years. The mean age differed significantly in dementia-AD vs CU. The
299 percent ratio of *APOE* $\epsilon 3/\epsilon 3$ carriers was significantly lower and the percent ratio of *APOE* $\epsilon 4/\epsilon 4$
300 carriers was higher in dementia-AD compared to CU. Amyloid positivity was higher in the AD
301 dementia group.
302

303 3.2 Diversity results in gut microbiota composition

304 Alpha diversity indices (Shannon, Evenness, and Faith's PD) were presented on the y-
305 axis and all AD group categories were presented on the x-axis (**Figure 1A-I**). The only
306 comparison with significant alpha diversity differences was the *APOE* $\epsilon 4$ comparison (**Figure**
307 **1G and H**). All alpha diversity indices showed significant differences or a trend toward
308 differences between *APOE* $\epsilon 4+$ and *APOE* $\epsilon 4-$ groups. Individuals who were *APOE* $\epsilon 4-$ had
309 significantly higher alpha diversity than *APOE* $\epsilon 4+$.

310 Beta diversity indices (Bray-Curtis dissimilarity, and Weighted and Unweighted UniFrac)
311 were visualized with principal coordinates analysis (PCoA) plots for each AD group (**Figure 1J-**
312 **R**). Significant differences between groups based on clinical diagnosis (diagnosis) were
313 observed with each metric (**Figure 1J-L**). Differences in beta diversity between amyloid-positive
314 and amyloid-negative individuals were detected only with the Bray-Curtis dissimilarity index
315 (**Figure 1M-O**). Individuals positive and negative for *APOE* $\epsilon 4$ demonstrated differences with
316 both the Bray-Curtis and Weighted UniFrac metrics, but not with Unweighted UniFrac (**Figure**
317 **1P-R**), suggesting that the relative abundance of major taxa rather than community membership
318 is important in driving these differences.
319

320 3.3 Gut microbiota composition in clinical diagnosis, amyloid status, and *APOE* $\epsilon 4$ 321 status groups

322 A DA analysis on gut microbiota composition was performed based on clinical diagnosis,
323 amyloid status, and *APOE* $\epsilon 4$ status groups using BIRDMAN and visualized as forest plots
324 (**Figure 2**). Gut microbiota taxonomic features that showed the most differences by the effect

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

325 size (log ratio) in each comparison group were displayed up to 20 features in each Top (more
326 abundant) and Bottom (less abundant) group for taxonomic levels including phylum, family,
327 genus, and species. DA analysis between AD and CU in clinical diagnosis showed distinct gut
328 microbiota composition at each taxonomic level (**Figure 2A, D, G, and J**). At the phylum level,
329 the abundance of phylum Firmicutes_A was lower in AD compared to CU (Bottom), and the
330 abundance of phyla Bacteroidota, Patescibacteria, and Fusobacteriota was higher in AD
331 compared to CU (Top) (**Figure 2A**). At the family level, families such as *Clostridiaceae*,
332 *Turicibacteraceae*, *Pasteurellaceae*, *Dialisteraceae*, *Enterococcaceae*, and *Ruminococcaceae*
333 were in the Bottom group and *Fusobacteriaceae*, *Nanogingivalaceae*, *Gemellaceae*, and
334 *Bacteroidaceae* were in the Top group (**Figure 2D**). At the genus level, genera including
335 *Clostridium_P*, *Ruminococcus*, and *Cryptobacteroides* were in the Bottom group and
336 *Fusobacterium_A*, *Fusobacterium*, *Nanogingivalis*, and *Gemella* were in the Top group (**Figure**
337 **2G**). At the species level, species included in the Bottom group were *Cryptobacteroides* spp.,
338 *Clostridium_P perfringens*, *Turicibacter sanguinis*, *Prevotella hominis*, and *Prevotella copri*, and
339 in the Top group were *Fusobacterium_A mortiferum*, *Fusobacterium nucleatum*, *Fusobacterium*
340 *animalis*, *Nanogingivalis gingivitus*, *Collinsella stercoris*, and *Collinsella tanakaei* (**Figure 2J**).

341 DA analysis between A+ and A- in amyloid status showed distinct gut microbiota
342 composition at each taxonomic level (**Figure 2B, E, H, and K**). At the phylum level, the
343 abundance of phylum Firmicutes_C was lower in A+ compared to A- (Bottom), and the
344 abundance of phyla Thermoplasmatota and Campylobacterota was higher in A+ compared to
345 A- (Top) (**Figure 2B**). At the family level, families such as *Neisseriaceae*, *Anaeroplasmataceae*,
346 *Turicibacteraceae*, and *Ruminococcaceae* were in the Bottom group and *Nanogingivalaceae*,
347 *Campylobacteraceae*, and *Coriobacteriaceae* were in the Top group (**Figure 2E**). At the genus
348 level, genera including *Prevotella*, *Eubacterium_R*, and *Lactococcus* were in the Bottom group
349 and *Fusobacterium*, *Nanogingivalis*, and *Acidaminococcus* were in the Top group (**Figure 2H**).
350 At the species level, species included in the Bottom group were *Dialister hominis*, *Prevotella*
351 *copri*, and *Prevotella hominis*, and in the Top group were *Nanogingivalis gingivitus*, *Collinsella*
352 *tanakaei*, and *Fusobacterium animalis* (**Figure 2K**).

353 DA analysis between APOE $\epsilon 4+$ and APOE $\epsilon 4-$ in APOE $\epsilon 4$ status analyses showed
354 distinct gut microbiota composition at each taxonomic level (**Figure 2C, F, I, and L**). At the
355 phylum level, the abundance of phyla Firmicutes_A, Firmicutes_C, Spirochaetota, and
356 Synergistota was lower in APOE $\epsilon 4+$ compared to APOE $\epsilon 4-$ (Bottom), and the abundance of
357 phylum Bacteroidota was higher in APOE $\epsilon 4+$ compared to APOE $\epsilon 4-$ (Top) (**Figure 2C**). At the
358 family level, families such as *Selenomonadaceae*, *Neisseriaceae*, *Pasteurellaceae*,
359 *Turicibacteraceae*, and *Clostridiaceae* were in the Bottom group and *Lactobacillaceae*,
360 *Eubacteriaceae*, and *Bacteroidaceae* were in the Top group (**Figure 2F**). At the genus level,
361 genera including *Ruminococcus* and *Clostridium_P* were in the Bottom group and *Enterobacter*,
362 *Hafnia*, and *Lactobacillus* were in the Top group (**Figure 2I**). At the species level, species
363 included in the Bottom group were *Prevotella hominis*, *Dialister* spp., and *Ruminococcus* spp.,
364 and in the Top group were *Enterobacter hormaechei_A*, *Hafnia proteus*, *Collinsella stercoris*,
365 *Enterobacter cloacae*, and *Collinsella tanakaei* (**Figure 2L**).

366 Log ratios of microbiome counts were calculated between the sum of the Top and
367 Bottom groups from DA analysis to test the overall significant differences of microbiome
368 features between AD conditions (log[sum of Top features/sum of Bottom features]) (**Figure 3**).

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

369 Overall, there were significant differences in each AD-related group (diagnosis, amyloid, and
370 *APOE* $\epsilon 4$) (**Figure 3A, E, and I**). Features that significantly differed between AD and CU also
371 differed between *APOE* $\epsilon 4+$ and *APOE* $\epsilon 4-$ (**Figure 3A and C**). Features that significantly
372 differed between A+ and A- also differed between AD and CU (**Figure 3D and E**). Features that
373 significantly differed between *APOE* $\epsilon 4+$ and *APOE* $\epsilon 4-$ also differed between AD and CU
374 (**Figure 3G and I**).
375

376 3.4 Common gut microbiota features in clinical diagnosis, amyloid status, and *APOE* $\epsilon 4$ 377 status groups

378 Venn diagrams were used to find common microbiome features across different AD
379 conditions for each Bottom and Top group at each taxonomic rank (**Table S2**). At the phylum
380 level, phyla Bacteroidota co-occurred between diagnosis and *APOE* $\epsilon 4$ in the Top group (**Figure**
381 **4A**). Firmicutes_A co-occurred between diagnosis and *APOE* $\epsilon 4$, and Firmicutes_C co-occurred
382 between amyloid and *APOE* $\epsilon 4$ in the Bottom group (**Figure 4B**). In the Top group at the family
383 level, the family *Lactobacillaceae* co-occurred across all conditions, families *Bacteroidaceae*
384 and *Coprobacteraceae* between diagnosis and *APOE* $\epsilon 4$, and families *Nanogingivalaceae* and
385 *Aerococcaceae* co-occurred between diagnosis and amyloid (**Figure 4C**). In the Bottom group,
386 *UBA1829* and *Turicibacteraceae* co-occurred across all conditions (diagnosis, amyloid, and
387 *APOE* $\epsilon 4$), families *CAG-508*, *CAG-74*, *Oscillospiraceae*, *Clostridiaceae*, *Pasteurellaceae*, and
388 *CAG-138* co-occurred between diagnosis and *APOE* $\epsilon 4$, families *Ruminococcaceae*,
389 *Anaeroplasmataceae*, and *CAG-312* co-occurred between diagnosis and amyloid, and family
390 *Neisseriaceae* co-occurred between amyloid and *APOE* $\epsilon 4$ (**Figure 4D**). Multiple genera and
391 species co-occurred across all conditions (**Figure 4E-H**). In the Top group, the genus
392 *Veillonella_A* co-occurred across all conditions, and the following numbers of genera co-
393 occurred between each intersection, i.e., diagnosis and *APOE* $\epsilon 4$: 7, diagnosis and amyloid: 6,
394 and amyloid and *APOE* $\epsilon 4$: 4 (**Figure 4E, Table S2**). In the Bottom group, 13 genera co-
395 occurred across all conditions including *Prevotella* and *Turicibacter*, and the following numbers
396 of genera co-occurred between each intersection, i.e., diagnosis and *APOE* $\epsilon 4$: 30, diagnosis
397 and amyloid: 9, and amyloid and *APOE* $\epsilon 4$: 8 (**Figure 4F, Table S2**). In the Top group at the
398 species level, 12 species including *Bacteroides ovatus*, *Collinsella tanakaei*, *Prevotella corporis*,
399 and more co-occurred across all conditions, and the following numbers of species co-occurred
400 between each intersection, i.e., diagnosis and *APOE* $\epsilon 4$: 18, diagnosis and amyloid: 10, and
401 amyloid and *APOE* $\epsilon 4$: 7 (**Figure 4G, Table S2**). In the Bottom group, 27 species including
402 *Ruminococcus_C callidus*, *Dialister succinatiphilus*, *Prevotella copri*, and more co-occurred
403 across all conditions, and the following numbers of species co-occurred between each
404 intersection, i.e., diagnosis and *APOE* $\epsilon 4$: 42, diagnosis and amyloid: 13, and amyloid and
405 *APOE* $\epsilon 4$: 16 (**Figure 4H, Table S2**).
406

407 3.5 Validation of shallow-shotgun data with the ADRC dataset on gut microbiota 408 composition

409 To validate the features linked with dementia-AD (AD) compared to healthy controls
410 (CU), we tested the log-transformed ratios of features more and less abundant in AD in the
411 MARS cohort against a larger cohort of participants who are part of the AGMP and who were
412 recruited from multiple NIA-funded ADRC across the U.S. (n = 448; **Figure 5**). AGMP
413 participants from Wisconsin were excluded to ensure a unique validation sample. The features
414 that were differentially abundant in the MARS cohort (Kruskal-Wallis: 31.81, *P* value: < .001;
415 **Figure 5A**) were also found to be differentially abundant in the larger validation cohort (Kruskal-
416 Wallis: 5.59, *P* value: .02; **Figure 5B**).
417

418 3.6 Gut microbiome functional pathways in a clinical diagnosis group

419 The DA analysis of gut microbiome functional pathways, stratified by species within each
420 pathway, identified 116 distinct pathways that differ between individuals with AD and CU
421 individuals (**Table S3**). Among 116 distinct pathways, we focused our analysis on pathways that
422 only showed abundance in either the Top (more abundant in AD) or the Bottom (less abundant
423 in AD) group. Among pathway features only with either the Top (15 pathways) or the Bottom (6
424 pathways) group, the log ratios of Top/Bottom features were shown to be significantly different
425 between AD and CU (**Figure 6A**). Furthermore, gut microbiome taxonomic features that were
426 associated with pathway features (36 features) only with either the Top or Bottom group were
427 visualized with each pathway category (**Figure 6B**). For example, a pathway, naphthalene
428 degradation, was one of the Bottom pathways, and microbes associated with this pathway were
429 species *Turicibacter sanguinis*, *Bifidobacterium angulatum*, and *Lactococcus lactis* (**Figure 6B**).
430 Another example in the Top pathway is benzoate degradation which is associated with microbes
431 including *Anaerostipes caccae*, *Bacteroides finegoldii*, and *Bacteroides thetaiotaomicron*
432 (**Figure 6B**).

433 RPCA on microbiome pathway features (36 features from DA analysis) visualized with a
434 biplot indicated a significant separation between the clinical diagnosis group (AD and CU, *P*
435 value = .004) (**Figure 6C**). Microbial features that belonged to the Top group indicated by the
436 red vector directed towards many AD subjects indicated by the orange dots. Microbial features
437 that belonged to the Bottom group indicated by the green vector directed towards many CU
438 subjects indicated by the blue dots. For instance, multiple pathways from species *Bacteroides*
439 *thetaitaomicron* pointed towards the AD group suggesting a potentially stronger association
440 between the Top features and AD group (**Figure 6C**). On the other hand, pathways from
441 species *Turicibacter sanguinis*, *Bifidobacterium angulatum*, and *Lactococcus lactis*, which
442 belonged to the Bottom group, pointed towards the CU group or showed different directions
443 compared to the Top features suggesting a potentially weaker association between the Bottom
444 features and AD group (**Figure 6C**).
445

446 3.7 Associations between gut microbiome compositional and functional features in
447 clinical diagnosis group and CSF biomarkers of AD and related pathologies

448 Associations between gut microbiome features (composition and function) and CSF
449 biomarkers of AD and related pathologies were performed as described in the 'Statistical
450 analysis' section of the 'Methods' (**Figure 7**).

451 Overall, in the association between gut microbiome compositional features and CSF
452 biomarkers of AD and related pathologies, most species that were more abundant in AD
453 compared to CU individuals were positively correlated with CSF biomarkers. Conversely,
454 species that were less abundant in AD were generally negatively associated with CSF
455 biomarkers. CSF biomarkers for AD and related pathologies included in the analysis were
456 amyloid pathology ($A\beta_{42}/A\beta_{40}$), tau pathophysiology (pTau₁₈₁ and tTau), neurodegeneration
457 (NfL), synaptic dysfunction and injury (neurogranin and α -synuclein), inflammation (IL-6), and
458 glial activation (S100B, GFAP, YKL-40, and sTREM2) (**Figure 7A**).

459 Species that were more abundant in AD were generally positively associated with CSF
460 biomarkers. For example, *Nanogingivalis gingivitus*, more abundant in AD, was positively
461 associated with S100B, neurogranin, pTau₁₈₁, and tTau. *Fusobacterium_A mortiferum* was
462 positively correlated with neurogranin, pTau₁₈₁, tTau, and α -synuclein, and negatively
463 associated with CSF amyloid ($A\beta_{42}/A\beta_{40}$). *Fusobacterium animalis* was positively associated
464 with neurogranin, pTau₁₈₁, tTau, and α -synuclein. *CAG-1031 sp000431215*, a species within the
465 Bacteroidetes phylum, was positively correlated with NfL, YKL-40, pTau₁₈₁, tTau, and α -
466 synuclein. *Berrvella sp001552935* was positively associated with YKL-40, sTREM2, tTau, and
467 α -synuclein. Additionally, *Lactobacillus acidophilus* and *Bifidobacterium vaginale* were positively
468 associated with S100B, *CAG-977 sp000434295* was associated with pTau₁₈₁ and tTau,
469 *Fusobacterium nucleatum* was associated with tTau, *Limosilactobacillus vaginalis* and
470 *Collinsella tanakaei* were associated with pTau₁₈₁, and *CAG-177 sp000431775* was associated
471 with IL-6.

472 Species that were less abundant in AD were generally negatively associated with CSF
473 biomarkers. Species *SFMI01 sp004556155*, *Turicibacter sanguinis*, and *Dialister hominis*, all
474 within the Firmicutes phylum, were negatively associated with neurogranin, pTau₁₈₁, tTau, and
475 α -synuclein. *UBA5809 sp002417965*, another Firmicutes species, was also negatively
476 associated with neurogranin, pTau₁₈₁, tTau, and α -synuclein, as well as sTREM2. *UBA11524*
477 *sp000437595*, another Firmicutes species, was negatively associated with NfL.
478 *Cryptobacteroides sp900544195* and *Cryptobacteroides sp000432515*, both species under
479 phylum Bacteroidetes, were negatively associated with S100B.

480 In the association between gut microbiome functional features and CSF biomarkers of
481 AD and related pathologies, multiple microbial pathways more abundant in AD compared to CU
482 showed a tendency to positively correlate with the CSF biomarkers, whereas pathways less
483 abundant in AD compared to CU showed a tendency to have negative associations with the
484 CSF biomarkers (**Figure 7B**). It should be noted that a lower $A\beta_{42}/A\beta_{40}$ ratio is associated with a
485 higher risk of having AD pathology whereas higher levels of the rest of the CSF biomarkers are
486 associated with a higher risk of having AD pathology. The same categories of CSF biomarkers
487 for AD and related pathologies were included in the analysis including $A\beta_{42}/A\beta_{40}$, pTau₁₈₁, tTau,
488 IL-6, NfL, neurogranin, α -synuclein, S100B, GFAP, YKL-40, and sTREM2.

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

489 Microbial functional features in the Top group showed overall positive associations with
490 CSF biomarkers with the exception of $A\beta_{42}/A\beta_{40}$. Multiple *Bacteroides* spp. and their related
491 pathways were positively associated with several CSF biomarkers including NfL, neurogranin,
492 α -synuclein, pTau₁₈₁, and tTau. *Bacteroides thetaiotaomicron* and its associated pathways
493 including benzoate degradation, ubiquinone and other terpenoid-quinone biosynthesis,
494 biosynthesis of various plant secondary metabolites, inositol phosphate metabolism, lipoic acid
495 metabolism, biosynthesis of various antibiotics, beta-alanine metabolism, carbapenem
496 biosynthesis, neomycin, kanamycin and gentamicin biosynthesis, polyketide sugar unit
497 biosynthesis, ascorbate and aldarate metabolism, and taurine and hypotaurine metabolism had
498 generally positive relationship with CSF biomarkers including NfL, YKL-40, neurogranin, α -
499 synuclein, pTau₁₈₁, and tTau, and negative relationship with IL-6. *Collinsella stercoris* and
500 polyketide sugar unit biosynthesis pathway showed positive correlation with GFAP and
501 sTREM2, and negative correlation with $A\beta_{42}/A\beta_{40}$. *Collinsella stercoris* and O-antigen repeat unit
502 biosynthesis pathway showed negative correlation with $A\beta_{42}/A\beta_{40}$.

503 Microbial functional features in the Bottom group showed overall negative associations
504 with CSF biomarkers for AD and related pathologies. Two pathways, ether lipid metabolism and
505 alpha-linolenic acid metabolism, related to *Parabacteroides merdae*, a species more abundant
506 in CU group were associated with lower CSF $A\beta_{42}/A\beta_{40}$. It implies that higher abundance of
507 these pathways of *Parabacteroides merdae* in CU individuals is associated with more brain
508 amyloid. Moreover, bacterial chemotaxis pathway from *Coprococcus eutactus* was negatively
509 associated with NfL and YKL-40, and naphthalene degradation pathway from *Turicibacter*
510 *sanguinis* was negatively associated with α -synuclein, pTau₁₈₁, and tTau.

511

512 4 DISCUSSION

513 In this study, we compared gut microbiome composition and function between several
514 AD-relevant groups, including those with a clinical diagnosis, differential amyloid status, and
515 *APOE* $\epsilon 4$ carrier status. The objective was to determine the association of gut microbiome
516 features that are differentially abundant in AD dementia and determine association with CSF
517 biomarkers of AD and related pathological features, to potentially identify gut microbial features
518 associated with AD.

519 Alpha and beta diversity analysis was performed between groups of each clinical
520 diagnosis, amyloid status, and *APOE* $\epsilon 4$ status groups. Prior studies have found that alpha and
521 beta diversities do not differ between AD vs CU^{61,62} and A+ vs A-⁶³ while other studies showed
522 significant differences in alpha and beta diversity indices in humans^{4,7} and mice.⁶⁴⁻⁶⁶

523 The DA analysis in gut microbiome composition at each taxonomic level using BIRDMAN
524 was performed in AD-related groups. Similar results were reported in other studies^{6,67} while
525 opposite findings were also found, where fewer Bacteroidetes and more Firmicutes were
526 reported in MCI compared to healthy controls and fewer genera *Bacteroides* and *Alistipes* and
527 more genus *Bifidobacterium* were found in AD compared to health controls.^{68,69} A meta-analysis
528 of gut microbiome compositional differences in AD across studies between 2000 to 2021
529 demonstrated similar outcomes measured by overall pooled effect size at each taxonomic
530 level.⁷

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

531 Discrepancies between studies may be due to differences in sample size, population
532 variation, disease heterogeneity, sequencing method, and confounding factors.⁷⁰ Our study
533 addresses these discrepancies through robust methodologies. Firstly, we utilized shotgun
534 metagenomic sequencing, which provides more comprehensive taxonomic and functional
535 profiling of microbial communities compared to 16S rRNA sequencing.⁷¹ Additionally, we
536 employed advanced statistical methodologies that have been shown to be replicable across
537 multiple cohorts.⁵² Lastly, we validated our findings with a larger cohort from the AGMP, which
538 significantly enhances the reliability and generalizability of our result. Our validation of the DA
539 analysis in a larger cohort largely recapitulates what we found in the smaller sample, confirming
540 that alterations in gut microbiome composition are present in AD dementia.

541 To further determine co-occurring gut microbes among clinical diagnosis, amyloid status,
542 and *APOE ε4* status groups, a co-occurrence analysis was performed on differentially abundant
543 microbes in each group. Investigating co-occurring taxonomic features may be useful in the
544 examination of gut microbiota that could potentially coexist and contribute to AD pathology
545 related to amyloid pathology or *APOE ε4* pathology which are phenotypic and genotypic
546 pathological signatures of AD. Studies have shown gut microbiota differences between CU+
547 (Aβ-positive) and CU- (Aβ-negative).^{8,11,72} One of the studies showed that Phylum
548 Bacteroidetes, class Bacteroidia, and order Bacteroidales were enriched in CN+ and phylum
549 Firmicutes, class Clostridia, order Clostridiales, families *Lachnospiraceae* and
550 *Ruminococcaceae*, and genera *Faecalibacterium* and *Bilophila* were enriched in CN-.¹¹ Higher
551 abundance of genera *Faecalibacterium* and *Bilophila* was negatively correlated with the global
552 brain Aβ burden.¹¹ Another study explored gut microbiome taxa which are pro-inflammatory with
553 blood inflammation markers.⁷³ Genus *Escherichia/Shigella* was significantly more abundant in
554 A+ compared with A- individuals. Genus *Escherichia/Shigella* was correlated positively with
555 peripheral inflammatory cytokines in individuals with cognitive impairment and brain
556 amyloidosis.⁷³

557 Taken together, results from our study and other studies suggest that diverse and
558 distinct gut microbiota taxonomic composition is altered in AD dementia, among individuals with
559 preclinical AD, and individuals with genetic risk for AD. However, studies are limited to
560 taxonomic and compositional associations and the determination of microbial functions in AD
561 pathogenesis is needed to better understand the role of specific gut microbes and their
562 functions in the progression of AD.

563 This study further examined the functional pathways of gut microbiome in a clinical
564 diagnosis group between AD and CU. Gut microbiome pathways and associated species that
565 are differentially abundant between AD and CU were determined. We found 116 distinct
566 pathways between AD and CU. Among 116 KO pathway features, 21 pathways had
567 associations either with AD (Top) or CU (Bottom) group. The log ratios of these Top/Bottom
568 microbial pathway features between AD and CU were significantly different. Interestingly,
569 species that were associated with these microbial pathways were mostly *Bacteroides* spp. (*B.*
570 *finnegoldii*, *B. thetaiotaomicron*, and *B. ovatus*) under phylum Bacteroidota in the Top group.
571 Multiple studies have reported the association between *Bacteroides* and AD.^{4,74,75} Administration
572 of *Bacteroides fragilis* to AD mice increased Aβ plaques and inhibition of microglial clearance of
573 Aβ was observed after introduction to *B. fragilis*.⁷⁶ Another study showed the role of *B. fragilis* in

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

574 AD pathology in mice.⁷⁷ Studies have suggested that genus *Bacteroides* to be dominant in older
575 adults compared with healthy and younger controls.^{78,79} However, contrasting findings have
576 been reported related to *Bacteroides*^{8,80} and further strain-specific studies are needed to
577 understand the role of *Bacteroides* in AD pathology.

578 Additionally, RPCA showed a distinct significant separation between AD and CU groups.
579 RPCA is known to handle sparse and high-dimensional datasets and is sensitive to datasets
580 with outliers.⁵⁴ Microbiome datasets are often sparse and zero-inflated, thus we employed
581 RPCA to identify microbiome features that could explain the separation between groups.
582 Consistent with the DA analysis, RPCA showed microbial pathways that are more abundant in
583 AD (Top) were more associated with AD dementia, whereas microbial pathways less abundant
584 in AD (Bottom) were less associated with AD dementia.

585 To determine whether distinct microbial features in AD and CU correlate with CSF
586 biomarkers of AD and related pathologies, the relationship between CSF biomarkers and each
587 compositional and functional gut microbiome feature was explored using the OLS regression
588 model.

589 Key microbiome species, particularly *Fusobacterium nucleatum*, *Fusobacterium*
590 *animalis*, and *Nanogingivalis gingivitus*, identified as commonly more abundant between AD
591 and A+, were associated with more intense tau pathophysiology (pTau₁₈₁ and tTau) and/or
592 synaptic dysfunction and injury (neurogranin and α -synuclein). Interestingly, the species
593 *Fusobacterium nucleatum* is an oral bacteria often associated with cavity and periodontal
594 diseases as well as colorectal cancer.^{81,82} *Fusobacterium nucleatum* produces
595 lipopolysaccharides (LPS) that induce microglial activation with elevated expression of
596 proinflammatory cytokines.⁸³ In prior studies using the 5XFAD mouse model, the mRNA
597 expression levels of the same proinflammatory cytokines as well as numbers of microglia in the
598 mice brain were increased after *Fusobacterium nucleatum* infection.⁸³ Moreover, enhanced A β
599 accumulation, tau protein phosphorylation, and memory impairment were observed in 5XFAD
600 mice compared to controls.⁸³ The oral infection of *Fusobacterium nucleatum* in AD-like
601 periodontitis rats exhibited increased accumulation of A β and pTau₁₈₁ expression in the brain.⁸⁴
602 Although these results propose valuable mechanistic backgrounds for *Fusobacterium*
603 *nucleatum* and AD, further investigation in humans is needed.

604 The species *Dialister hominis*, a microbe that co-occurred between CU and A- displayed
605 a negative correlation with AD pathology (pTau₁₈₁), neuronal damage (tTau), and synaptic
606 dysfunction and injury (neurogranin and α -synuclein). These findings are similar to previous
607 works which identified genus *Dialister* (less in AD) to be more abundant in CU individuals.^{85,86}
608 Another species *Turicibacter sanguinis* which was a microbe that co-occurred between CU and
609 *APOE ϵ 4-* showed a negative correlation with pTau₁₈₁, tTau, neurogranin, and α -synuclein. In
610 animal models for AD, *Turicibacter sanguinis* is reported to be less abundant in AD compared to
611 controls.^{87,88} In humans, *Turicibacter* was observed to be less abundant in individuals with AD.⁴
612 However, studies on the role of species specific to *Dialister hominis* and *Turicibacter sanguinis*
613 in AD pathology are scarce.

614 This result indicates that microbiota compositional features in CU are related to lower
615 levels of AD pathology whereas microbiota features in AD are related to higher levels of AD
616 pathology, suggesting overall microbiota composition in people with AD may be vulnerable to

617 development or progression of AD compared to CU individuals. While these results support a
618 relationship between gut microbiome composition and AD pathology, further investigation into
619 these mechanisms is required to find a causal relationship.

620 Multiple *Bacteroides* spp. and their related functional pathways more abundant in AD
621 were associated with greater AD pathology represented in CSF biomarkers of AD and related
622 pathologies. Biomarkers including neurodegeneration (NfL), synaptic dysfunction and injury
623 (neurogranin and α -synuclein), and tau pathophysiology (pTau₁₈₁ and tTau) showed a positive
624 relationship with *Bacteroides thetaiotaomicron* and their functions. Studies have linked the
625 abundance of *Bacteroides thetaiotaomicron* with AD. The abundance of *B. thetaiotaomicron*
626 was significantly higher in AD mice and was related to poorer spatial learning.⁸⁹ Increased
627 abundance of *B. thetaiotaomicron* was reported in AD participants.^{90,91} However, in a non-AD
628 model, *B. thetaiotaomicron* was suggested to regulate enteric neuronal cell populations and
629 neurogenic function.

630 Although evidence related to these microbial functions is limited, these results suggest
631 that alterations in gut microbiome composition and function are related to AD pathological
632 markers measured in CSF.

633 The main limitation of our study is the small number of cognitively impaired participants
634 relative to CU participants. The sample size decreased after matching clinical measurements,
635 gut microbiome data, and presence of CSF biomarkers. The resulting low statistical power may
636 have led to losing significance after multiple test corrections for association analyses between
637 gut microbiome features and CSF biomarkers. A similar challenge is the inclusion of cognitively
638 impaired individuals in both the A+ and *APOE* ϵ 4 groups. Excluding these individuals reduced
639 the sample sizes, resulting in low statistical power. Future studies should aim to collect sufficient
640 samples to separate these groups, allowing for a clearer distinction between disease effects and
641 symptom effects. Moreover, due to the cross-sectional approach, it is difficult to capture the
642 longitudinal changes over time considering the progression of AD for each individual. We
643 included the ADRC validation cohort to account for differences in gut microbiome composition
644 across diverse populations, however, due to the limitation of the availability of biomarkers
645 matched with fecal samples, we were not able to test associations with CSF biomarkers in the
646 validation cohort. Additionally, the results are correlational, and further mechanistic studies are
647 needed to find causal relationships between gut microbiome features and biomarkers for AD
648 pathology. Finally, other environmental factors (exposome) which may impact the gut
649 microbiome and which contribute to AD risk require additional study in the future.

650

651 **5 CONCLUSIONS**

652 This study suggests that gut microbiome composition and function differ between people with
653 AD dementia and CU individuals. Beta diversity indices differed among AD-related groups:
654 diagnosis (AD vs CU), amyloid (A+ vs A-), and *APOE* ϵ 4 (*APOE* ϵ 4+ vs *APOE* ϵ 4-) groups,
655 indicating that the gut microbiome diversity varies between each group. Multiple gut microbes at
656 each taxonomic level including phylum, family, genus, and species were differentially abundant
657 across AD groups. Co-occurring gut microbes across AD-related groups were determined,

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

658 many of which showed associations with CSF biomarkers for AD and related pathologies.
659 Microbial functional pathways were differentially abundant between AD and CU, which were
660 correlated with AD pathology markers measured in CSF. These findings identify specific targets
661 for stratifying key gut microbes and microbial pathways that may be related to AD pathology.
662 Further investigation on metabolomic changes as well as exposome and host genome that may
663 be mediating the interconnectome between the gut microbiome and AD pathology is needed.

664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699

700 References

- 701 1. Sonnenburg JL, Bäckhed F. Diet–microbiota interactions as moderators of human
702 metabolism. *Nature*. 2016;535(7610):56-64. doi:10.1038/nature18846
- 703 2. Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota
704 Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell*.
705 2016;167(5):1339-1353.e21. doi:10.1016/j.cell.2016.10.043
- 706 3. Lee WJ, Hase K. Gut microbiota–generated metabolites in animal health and disease. *Nat*
707 *Chem Biol*. 2014;10(6):416-424. doi:10.1038/nchembio.1535
- 708 4. Vogt NM, Kerby RL, Dill-McFarland KA, et al. Gut microbiome alterations in Alzheimer's
709 disease. *Sci Rep*. 2017;7:13537. doi:10.1038/s41598-017-13601-y
- 710 5. Zhuang ZQ, Shen LL, Li WW, et al. Gut Microbiota is Altered in Patients with Alzheimer's
711 Disease. *Journal of Alzheimer's Disease*. 2018;63(4):1337-1346. doi:10.3233/JAD-180176
- 712 6. Liu P, Wu L, Peng G, et al. Altered microbiomes distinguish Alzheimer's disease from
713 amnesic mild cognitive impairment and health in a Chinese cohort. *Brain Behav Immun*.
714 2019;80:633-643. doi:10.1016/j.bbi.2019.05.008
- 715 7. Hung CC, Chang CC, Huang CW, Nouchi R, Cheng CH. Gut microbiota in patients with
716 Alzheimer's disease spectrum: a systematic review and meta-analysis. *Aging (Albany NY)*.
717 2022;14(1):477-496. doi:10.18632/aging.203826
- 718 8. Ferreiro AL, Choi J, Ryou J, et al. Gut microbiome composition may be an indicator of
719 preclinical Alzheimer's disease. *Science Translational Medicine*. 2023;15(700):eabo2984.
720 doi:10.1126/scitranslmed.abo2984
- 721 9. Vogt NM, Romano KA, Darst BF, et al. The gut microbiota-derived metabolite
722 trimethylamine N-oxide is elevated in Alzheimer's disease. *Alzheimer's Research &*
723 *Therapy*. 2018;10(1):124. doi:10.1186/s13195-018-0451-2
- 724 10. Nabizadeh F, Valizadeh P, Fallahi MS, Alzheimer's disease Neuroimaging Initiative. Bile
725 acid profile associated with CSF and PET biomarkers in Alzheimer's disease. *Aging Clin*
726 *Exp Res*. 2024;36(1):62. doi:10.1007/s40520-024-02729-3
- 727 11. Sheng C, Yang K, He B, Du W, Cai Y, Han Y. Combination of gut microbiota and plasma
728 amyloid- β as a potential index for identifying preclinical Alzheimer's disease: a cross-
729 sectional analysis from the SILCODE study. *Alzheimer's Research & Therapy*.
730 2022;14(1):35. doi:10.1186/s13195-022-00977-x
- 731 12. Sochocka M, Donskow-Łysoniewska K, Diniz BS, Kurpas D, Brzozowska E, Leszek J. The
732 Gut Microbiome Alterations and Inflammation-Driven Pathogenesis of Alzheimer's
733 Disease—a Critical Review. *Mol Neurobiol*. 2019;56(3):1841-1851. doi:10.1007/s12035-
734 018-1188-4
- 735 13. Heston MB, Hanslik KL, Zarbock KR, et al. Gut inflammation associated with age and
736 Alzheimer's disease pathology: a human cohort study. *Sci Rep*. 2023;13(1):18924.
737 doi:10.1038/s41598-023-45929-z

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 738 14. Heintz-Buschart A, Wilmes P. Human Gut Microbiome: Function Matters. *Trends in*
739 *Microbiology*. 2018;26(7):563-574. doi:10.1016/j.tim.2017.11.002
- 740 15. Chernikova MA, Flores GD, Kilroy E, Labus JS, Mayer EA, Aziz-Zadeh L. The Brain-Gut-
741 Microbiome System: Pathways and Implications for Autism Spectrum Disorder. *Nutrients*.
742 2021;13(12):4497. doi:10.3390/nu13124497
- 743 16. Kenna JE, Chua EG, Bakeberg M, et al. Changes in the Gut Microbiome and Predicted
744 Functional Metabolic Effects in an Australian Parkinson's Disease Cohort. *Front Neurosci*.
745 2021;15. doi:10.3389/fnins.2021.756951
- 746 17. Ananthakrishnan AN, Luo C, Yajnik V, et al. Gut Microbiome Function Predicts Response to
747 Anti-integrin Biologic Therapy in Inflammatory Bowel Diseases. *Cell Host & Microbe*.
748 2017;21(5):603-610.e3. doi:10.1016/j.chom.2017.04.010
- 749 18. Paley EL. Discovery of Gut Bacteria Specific to Alzheimer's Associated Diseases is a Clue
750 to Understanding Disease Etiology: Meta-Analysis of Population-Based Data on Human Gut
751 Metagenomics and Metabolomics. *Journal of Alzheimer's Disease*. 2019;72(1):319-355.
752 doi:10.3233/JAD-190873
- 753 19. Laske C, Müller S, Preische O, et al. Signature of Alzheimer's Disease in Intestinal
754 Microbiome: Results From the AlzBiom Study. *Front Neurosci*. 2022;16.
755 doi:10.3389/fnins.2022.792996
- 756 20. Li J, Zhu S, Wang Y, et al. Metagenomic association analysis of cognitive impairment in
757 community-dwelling older adults. *Neurobiology of Disease*. 2023;180:106081.
758 doi:10.1016/j.nbd.2023.106081
- 759 21. Johnson SC, Kosciak RL, Jonaitis EM, et al. The Wisconsin Registry for Alzheimer's
760 Prevention: A review of findings and current directions. *Alzheimers Dement (Amst)*.
761 2018;10:130-142. doi:10.1016/j.dadm.2017.11.007
- 762 22. Darst BF, Kosciak RL, Racine AM, et al. Pathway-specific polygenic risk scores as predictors
763 of β -amyloid deposition and cognitive function in a sample at increased risk for Alzheimer's
764 disease. *J Alzheimers Dis*. 2017;55(2):473-484. doi:10.3233/JAD-160195
- 765 23. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to
766 Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's
767 Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's &*
768 *Dementia*. 2011;7(3):270-279. doi:10.1016/j.jalz.2011.03.008
- 769 24. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to
770 Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's
771 Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers*
772 *Dement*. 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005
- 773 25. Betthausen TJ, Kosciak RL, Jonaitis EM, et al. Amyloid and tau imaging biomarkers explain
774 cognitive decline from late middle-age. *Brain*. 2020;143(1):320-335.
775 doi:10.1093/brain/awz378

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 776 26. Van Hulle C, Jonaitis EM, Betthausen TJ, et al. An examination of a novel multipanel of CSF
777 biomarkers in the Alzheimer's disease clinical and pathological continuum. *Alzheimer's &*
778 *Dementia*. 2021;17(3):431-445. doi:10.1002/alz.12204
- 779 27. Marotz C, Amir A, Humphrey G, Gaffney J, Gogul G, Knight R. DNA extraction for
780 streamlined metagenomics of diverse environmental samples. *BioTechniques*.
781 2017;62(6):290-293. doi:10.2144/000114559
- 782 28. Glenn TC, Nilsen RA, Kieran TJ, et al. Adapterama I: universal stubs and primers for 384
783 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext).
784 *PeerJ*. 2019;7:e7755. doi:10.7717/peerj.7755
- 785 29. Sanders JG, Nurk S, Salido RA, et al. Optimizing sequencing protocols for leaderboard
786 metagenomics by combining long and short reads. *Genome Biology*. 2019;20(1):226.
787 doi:10.1186/s13059-019-1834-9
- 788 30. Marotz C, Cavagnero KJ, Song SJ, et al. Evaluation of the Effect of Storage Methods on
789 Fecal, Saliva, and Skin Microbiome Composition. *mSystems*. 2021;6(2):e01329-20.
790 doi:10.1128/mSystems.01329-20
- 791 31. Sepich-Poore GD, McDonald D, Kopylova E, et al. Robustness of cancer microbiome
792 signals over a broad range of methodological variation. *Oncogene*. 2024;43(15):1127-1148.
793 doi:10.1038/s41388-024-02974-w
- 794 32. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
795 *Bioinformatics*. 2018;34(17):i884-i890. doi:10.1093/bioinformatics/bty560
- 796 33. Liao WW, Asri M, Ebler J, et al. A draft human pangenome reference. *Nature*.
797 2023;617(7960):312-324. doi:10.1038/s41586-023-05896-x
- 798 34. Rhie A, Nurk S, Cechova M, et al. The complete sequence of a human Y chromosome.
799 *Nature*. 2023;621(7978):344-354. doi:10.1038/s41586-023-06457-y
- 800 35. Schneider VA, Graves-Lindsay T, Howe K, et al. Evaluation of GRCh38 and de novo
801 haploid genome assemblies demonstrates the enduring quality of the reference assembly.
802 *Genome Res*. 2017;27(5):849-864. doi:10.1101/gr.213611.116
- 803 36. Li H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics*.
804 2021;37(23):4572-4574. doi:10.1093/bioinformatics/btab705
- 805 37. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools.
806 *Gigascience*. 2021;10(2):giab008. doi:10.1093/gigascience/giab008
- 807 38. Edwards JA, Edwards RA. Fastq-pair: efficient synchronization of paired-end fastq files.
808 Published online February 19, 2019;552885. doi:10.1101/552885
- 809 39. Tange O. *Gnu Parallel 2018*. [object Object]; 2018. doi:10.5281/ZENODO.1146014
- 810 40. Köster J. Rust-Bio: a fast and safe bioinformatics library. *Bioinformatics*. 2016;32(3):444-
811 446. doi:10.1093/bioinformatics/btv573

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 812 41. Gonzalez A, Navas-Molina JA, Kosciolk T, et al. Qiita: rapid, web-enabled microbiome
813 meta-analysis. *Nat Methods*. 2018;15(10):796-798. doi:10.1038/s41592-018-0141-9
- 814 42. Zhu Q, Huang S, Gonzalez A, et al. Phylogeny-Aware Analysis of Metagenome Community
815 Ecology Based on Matched Reference Genomes while Bypassing Taxonomy. *mSystems*.
816 2022;7(2):e00167-22. doi:10.1128/msystems.00167-22
- 817 43. Hakim D, Wandro S, Zengler K, et al. Zebra: Static and Dynamic Genome Cover Thresholds
818 with Overlapping References. *mSystems*. 2022;7(5):e0075822.
819 doi:10.1128/msystems.00758-22
- 820 44. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible
821 microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37(8):852-857.
822 doi:10.1038/s41587-019-0209-9
- 823 45. Betthausen TJ, Cody KA, Zammit MD, et al. In Vivo Characterization and Quantification of
824 Neurofibrillary Tau PET Radioligand 18F-MK-6240 in Humans from Alzheimer Disease
825 Dementia to Young Controls. *Journal of Nuclear Medicine*. 2019;60(1):93-99.
826 doi:10.2967/jnumed.118.209650
- 827 46. Shannon CE. A mathematical theory of communication. *The Bell System Technical Journal*.
828 1948;27(3):379-423. doi:10.1002/j.1538-7305.1948.tb01338.x
- 829 47. Pielou EC. The measurement of diversity in different types of biological collections. *Journal*
830 *of Theoretical Biology*. 1966;13:131-144. doi:10.1016/0022-5193(66)90013-0
- 831 48. Faith DP. Conservation evaluation and phylogenetic diversity. *Biological Conservation*.
832 1992;61(1):1-10. doi:10.1016/0006-3207(92)91201-3
- 833 49. Sørensen T, Sørensen T, Biering-Sørensen T, Sørensen T, Sorensen JT. *A Method of*
834 *Establishing Group of Equal Amplitude in Plant Sociobiology Based on Similarity of Species*
835 *Content and Its Application to Analyses of the Vegetation on Danish Commons*. Vol 5.
836 København: I kommission hos E. Munksgaard; 1948.
837 <https://api.semanticscholar.org/CorpusID:135206594>
- 838 50. Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and Qualitative β Diversity
839 Measures Lead to Different Insights into Factors That Structure Microbial Communities.
840 *Applied and Environmental Microbiology*. 2007;73(5):1576-1585. doi:10.1128/AEM.01996-
841 06
- 842 51. Lozupone C, Knight R. UniFrac: a New Phylogenetic Method for Comparing Microbial
843 Communities. *Applied and Environmental Microbiology*. 2005;71(12):8228-8235.
844 doi:10.1128/AEM.71.12.8228-8235.2005
- 845 52. Rahman G, Morton JT, Martino C, et al. BIRDMan: A Bayesian differential abundance
846 framework that enables robust inference of host-microbe associations. *bioRxiv*. Published
847 online February 2, 2023:2023.01.30.526328. doi:10.1101/2023.01.30.526328
- 848 53. Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway
849 identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*.
850 2005;21(19):3787-3793. doi:10.1093/bioinformatics/bti430

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 851 54. Martino C, Morton JT, Marotz CA, et al. A Novel Sparse Compositional Technique Reveals
852 Microbial Perturbations. *mSystems*. 2019;4(1):10.1128/msystems.00016-19.
853 doi:10.1128/msystems.00016-19
- 854 55. Anderson MJ. Permutational Multivariate Analysis of Variance (PERMANOVA). In: *Wiley*
855 *StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd; 2017:1-15.
856 doi:10.1002/9781118445112.stat07841
- 857 56. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python.
858 *Journal of Machine Learning Research*. 2011;12(85):2825-2830. Accessed August 12,
859 2024. <http://jmlr.org/papers/v12/pedregosa11a.html>
- 860 57. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific
861 computing in Python. *Nat Methods*. 2020;17(3):261-272. doi:10.1038/s41592-019-0686-2
- 862 58. van der Walt S, Colbert SC, Varoquaux G. The NumPy Array: A Structure for Efficient
863 Numerical Computation. *Computing in Science & Engineering*. 2011;13(2):22-30.
864 doi:10.1109/MCSE.2011.37
- 865 59. Hunter JD. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*.
866 2007;9(3):90-95. doi:10.1109/MCSE.2007.55
- 867 60. Waskom ML. seaborn: statistical data visualization. *Journal of Open Source Software*.
868 2021;6(60):3021. doi:10.21105/joss.03021
- 869 61. Hou M, Xu G, Ran M, Luo W, Wang H. APOE- ϵ 4 Carrier Status and Gut Microbiota
870 Dysbiosis in Patients With Alzheimer Disease. *Front Neurosci*. 2021;15.
871 doi:10.3389/fnins.2021.619051
- 872 62. Kaiyrykzy A, Kozhakhmetov S, Babenko D, et al. Study of gut microbiota alterations in
873 Alzheimer's dementia patients from Kazakhstan. *Sci Rep*. 2022;12:15115.
874 doi:10.1038/s41598-022-19393-0
- 875 63. Jung JH, Kim G, Byun MS, et al. Gut microbiome alterations in preclinical Alzheimer's
876 disease. *PLOS ONE*. 2022;17(11):e0278276. doi:10.1371/journal.pone.0278276
- 877 64. Zajac DJ, Green SJ, Johnson LA, Estus S. APOE genetics influence murine gut
878 microbiome. *Sci Rep*. 2022;12(1):1906. doi:10.1038/s41598-022-05763-1
- 879 65. Tran TTT, Corsini S, Kellingray L, et al. APOE genotype influences the gut microbiome
880 structure and function in humans and mice: relevance for Alzheimer's disease
881 pathophysiology. *FASEB J*. 2019;33(7):8221-8231. doi:10.1096/fj.201900071R
- 882 66. Xiao-hang Q, Si-yue C, Hui-dong T. Multi-strain probiotics ameliorate Alzheimer's-like
883 cognitive impairment and pathological changes through the AKT/GSK-3 β pathway in
884 senescence-accelerated mouse prone 8 mice. *Brain, Behavior, and Immunity*. 2024;119:14-
885 27. doi:10.1016/j.bbi.2024.03.031
- 886 67. Liu P, Jia XZ, Chen Y, et al. Gut microbiota interacts with intrinsic brain activity of patients
887 with amnesic mild cognitive impairment. *CNS Neuroscience & Therapeutics*.
888 2021;27(2):163-173. doi:10.1111/cns.13451

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 889 68. Li B, He Y, Ma J, et al. Mild cognitive impairment has similar alterations as Alzheimer's
890 disease in gut microbiota. *Alzheimers Dement.* 2019;15(10):1357-1366.
891 doi:10.1016/j.jalz.2019.07.002
- 892 69. Nagpal R, Neth BJ, Wang S, Craft S, Yadav H. Modified Mediterranean-ketogenic diet
893 modulates gut microbiome and short-chain fatty acids in association with Alzheimer's
894 disease markers in subjects with mild cognitive impairment. *eBioMedicine.* 2019;47:529-
895 542. doi:10.1016/j.ebiom.2019.08.032
- 896 70. Li Z, Zhou J, Liang H, et al. Differences in Alpha Diversity of Gut Microbiota in Neurological
897 Diseases. *Front Neurosci.* 2022;16. doi:10.3389/fnins.2022.879318
- 898 71. Jovel J, Patterson J, Wang W, et al. Characterization of the Gut Microbiome Using 16S or
899 Shotgun Metagenomics. *Front Microbiol.* 2016;7. doi:10.3389/fmicb.2016.00459
- 900 72. Verhaar BJH, Hendriksen HMA, de Leeuw FA, et al. Gut Microbiota Composition Is Related
901 to AD Pathology. *Front Immunol.* 2022;12. doi:10.3389/fimmu.2021.794519
- 902 73. Cattaneo A, Cattane N, Galluzzi S, et al. Association of brain amyloidosis with pro-
903 inflammatory gut bacterial taxa and peripheral inflammation markers in cognitively impaired
904 elderly. *Neurobiology of Aging.* 2017;49:60-68. doi:10.1016/j.neurobiolaging.2016.08.019
- 905 74. Haran JP, Bhattarai SK, Foley SE, et al. Alzheimer's Disease Microbiome Is Associated with
906 Dysregulation of the Anti-Inflammatory P-Glycoprotein Pathway. *mBio.*
907 2019;10(3):10.1128/mbio.00632-19. doi:10.1128/mbio.00632-19
- 908 75. Cammann D, Lu Y, Cummings MJ, et al. Genetic correlations between Alzheimer's disease
909 and gut microbiome genera. *Sci Rep.* 2023;13(1):5258. doi:10.1038/s41598-023-31730-5
- 910 76. Wasén C, Beauchamp LC, Vincentini J, et al. Bacteroidota inhibit microglia clearance of
911 amyloid-beta and promote plaque deposition in Alzheimer's disease mouse models. *Nat*
912 *Commun.* 2024;15:3872. doi:10.1038/s41467-024-47683-w
- 913 77. Xia Y, Xiao Y, Wang ZH, et al. Bacteroides Fragilis in the gut microbiomes of Alzheimer's
914 disease activates microglia and triggers pathogenesis in neuronal C/EBP β transgenic mice.
915 *Nat Commun.* 2023;14(1):5471. doi:10.1038/s41467-023-41283-w
- 916 78. Claesson MJ, Cusack S, O'Sullivan O, et al. Composition, variability, and temporal stability
917 of the intestinal microbiota of the elderly. *Proceedings of the National Academy of Sciences.*
918 2011;108(supplement_1):4586-4591. doi:10.1073/pnas.1000097107
- 919 79. Wilmanski T, Diener C, Rappaport N, et al. Gut microbiome pattern reflects healthy ageing
920 and predicts survival in humans. *Nat Metab.* 2021;3(2):274-286. doi:10.1038/s42255-021-
921 00348-0
- 922 80. Guo M, Peng J, Huang X, Xiao L, Huang F, Zuo Z. Gut Microbiome Features of Chinese
923 Patients Newly Diagnosed with Alzheimer's Disease or Mild Cognitive Impairment. *Journal*
924 *of Alzheimer's Disease.* 2021;80(1):299-310. doi:10.3233/JAD-201040

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

- 925 81. Pignatelli P, Nuccio F, Piattelli A, Curia MC. The Role of *Fusobacterium nucleatum* in Oral
926 and Colorectal Carcinogenesis. *Microorganisms*. 2023;11(9):2358.
927 doi:10.3390/microorganisms11092358
- 928 82. Zepeda-Rivera M, Minot SS, Bouzek H, et al. A distinct *Fusobacterium nucleatum* clade
929 dominates the colorectal cancer niche. *Nature*. 2024;628(8007):424-432.
930 doi:10.1038/s41586-024-07182-w
- 931 83. Wu H, Qiu W, Zhu X, et al. The Periodontal Pathogen *Fusobacterium nucleatum*
932 Exacerbates Alzheimer's Pathogenesis via Specific Pathways. *Front Aging Neurosci*.
933 2022;14. doi:10.3389/fnagi.2022.912709
- 934 84. Yan C, Diao Q, Zhao Y, et al. *Fusobacterium nucleatum* infection-induced
935 neurodegeneration and abnormal gut microbiota composition in Alzheimer's disease-like
936 rats. *Front Neurosci*. 2022;16. doi:10.3389/fnins.2022.884543
- 937 85. Ling Z, Zhu M, Yan X, et al. Structural and Functional Dysbiosis of Fecal Microbiota in
938 Chinese Patients With Alzheimer's Disease. *Front Cell Dev Biol*. 2021;8.
939 doi:10.3389/fcell.2020.634069
- 940 86. Rouskas K, Mamalaki E, Ntanasi E, et al. Distinct gut microbiota profiles may characterize
941 amyloid beta pathology and mild cognitive impairment. Published online May 3,
942 2024:2024.05.01.24306673. doi:10.1101/2024.05.01.24306673
- 943 87. Dunham SJB, McNair KA, Adams ED, et al. Longitudinal Analysis of the Microbiome and
944 Metabolome in the 5xfAD Mouse Model of Alzheimer's Disease. *mBio*. 2022;13(6):e01794-
945 22. doi:10.1128/mbio.01794-22
- 946 88. Sun P, Zhu H, Li X, et al. Comparative Metagenomics and Metabolomes Reveals Abnormal
947 Metabolism Activity Is Associated with Gut Microbiota in Alzheimer's Disease Mice.
948 *International Journal of Molecular Sciences*. 2022;23(19):11560.
949 doi:10.3390/ijms231911560
- 950 89. Abraham D, Feher J, Scuderi GL, et al. Exercise and probiotics attenuate the development
951 of Alzheimer's disease in transgenic mice: Role of microbiome. *Experimental Gerontology*.
952 2019;115:122-131. doi:10.1016/j.exger.2018.12.005
- 953 90. Fang Chuang Y. Altered Gut Microbiota in Dementia and Mild Cognitive Impairment in
954 Community-Dwelling Older Adults in Taiwan. *Alzheimer's & Dementia*.
955 2023;19(S12):e075212. doi:10.1002/alz.075212
- 956 91. Martinelli F, Heinken A, Henning AK, et al. Whole-body metabolic modelling reveals
957 microbiome and genomic interactions on reduced urine formate levels in Alzheimer's
958 disease. *Sci Rep*. 2024;14(1):6095. doi:10.1038/s41598-024-55960-3
- 959
960
961
962
963
964

965 **ACKNOWLEDGMENTS**

966 This project was enabled in part by The Alzheimer Gut Microbiome Project (AGMP) and the
967 Alzheimer's Disease Metabolomics Consortium (ADMC) funded wholly or in part by the
968 following NIA grants thereto: U01AG061359 and U19AG063744 awarded to Dr. Kaddurah-
969 Daouk at Duke University in partnership with multiple academic institutions. As such, the
970 investigators within the AGMP and the ADMC, not listed specifically in this publication's author's
971 list, provided data along with its pre-processing and prepared it for analysis, but did not
972 participate in analysis or writing of this manuscript. A listing of AGMP Investigators can be found
973 at alzheimergut.org/meet-the-team/. A complete listing of ADMC investigators can be found
974 at: sites.duke.edu/adnimetab/team/. The IGM S10 grant S10 OD026929 was awarded to Dr.
975 Rob Knight at the University of California San Diego IGM Genomics Center. This publication
976 includes data generated at the University of California San Diego IGM Genomics Center utilizing
977 an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health
978 SIG grant (#S10 OD026929).

979 This work was also supported by the National Institute on Aging Grant R01AG070973 (B.B.B.,
980 F.E.R., T.K.U.), National Institute on Aging Grant R01AG083883 (T.K.U., B.B.B., F.E.R.), Vilas
981 Early-Career Investigator Award (T.K.U.), and National Institute on Aging Grant P30AG062715
982 (which in part supported B.B.B., T.K.U.).

983 Dr. Zetterberg is a Wallenberg Scholar and a Distinguished Professor at the Swedish Research
984 Council supported by grants from the Swedish Research Council (#2023-00356; #2022-01018
985 and #2019-02397), the European Union's Horizon Europe research and innovation programme
986 under grant agreement No 101053962, Swedish State Support for Clinical Research
987 (#ALFGBG-71320), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-
988 2016862), the AD Strategic Fund and the Alzheimer's Association (#ADSF-21-831376-C,
989 #ADSF-21-831381-C, #ADSF-21-831377-C, and #ADSF-24-1284328-C), the European
990 Partnership on Metrology, co-financed from the European Union's Horizon Europe Research
991 and Innovation Programme and by the Participating States (NEuroBioStand, #22HLT07), the
992 Bluefield Project, Cure Alzheimer's Fund, the Olav Thon Foundation, the Erling-Persson Family
993 Foundation, Familjen Rönströms Stiftelse, Stiftelsen för Gamla Tjänarinnor, Hjärtfonden,
994 Sweden (#FO2022-0270), the European Union's Horizon 2020 research and innovation
995 programme under the Marie Skłodowska-Curie grant agreement No 860197 (MIRIADE), the
996 European Union Joint Programme – Neurodegenerative Disease Research (JPND2021-00694),
997 the National Institute for Health and Care Research University College London Hospitals
998 Biomedical Research Centre, and the UK Dementia Research Institute at UCL (UKDRI-1003).
999 COBAS is a trademark of Roche. All other product names and trademarks are the property of
1000 their respective owners. The NeuroToolKit is a panel of exploratory prototype assays designed
1001 to robustly evaluate biomarkers associated with key pathologic events characteristic of AD and
1002 other neurological disorders, used for research purposes only and not approved for clinical use
1003 (Roche Diagnostics International Ltd, Rotkreuz, Switzerland).

1004
1005
1006
1007
1008

1009 **CONFLICT OF INTEREST STATEMENT**

1010 Dr. Kaddurah-Daouk is an inventor on a series of patents on use of metabolomics for the
1011 diagnosis and treatment of CNS diseases and holds equity in Metabolon Inc., Chymia LLC and
1012 PsyProtix.

1013 Dr. Rob Knight is a scientific advisory board member, and consultant for BiomeSense, Inc., has
1014 equity and receives income. He is a scientific advisory board member and has equity in
1015 GenCirq. He is a consultant for DayTwo, and receives income. He has equity in and acts as a
1016 consultant for Cybele. He is a co-founder of Biota, Inc., and has equity. He is a cofounder of
1017 Micronoma, and has equity and is a scientific advisory board member. The terms of these
1018 arrangements have been reviewed and approved by the University of California, San Diego in
1019 accordance with its conflict of interest policies.

1020 Dr. Zetterberg has served at scientific advisory boards and/or as a consultant for Abbvie,
1021 Acumen, Alector, Alzinova, ALZPath, Amylyx, Annexon, Apellis, Artery Therapeutics,
1022 AZTherapies, Cognito Therapeutics, CogRx, Denali, Eisai, LabCorp, Merry Life, Nervgen, Novo
1023 Nordisk, Optoceutics, Passage Bio, Pinteon Therapeutics, Prothena, Red Abbey Labs,
1024 reMYND, Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given
1025 lectures in symposia sponsored by Alzecure, Biogen, Celectricon, Fujirebio, Lilly, Novo Nordisk,
1026 and Roche, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is
1027 a part of the GU Ventures Incubator Program (outside submitted work).

1028 Daniel McDonald is a consultant for, and has equity in, BiomeSense, Inc. The terms of this
1029 arrangement has been reviewed and approved by the University of California, San Diego in
1030 accordance with its conflict of interest policies.

1031
1032

1033 **CONSENT STATEMENT**

1034 All human subjects provided informed consent to participate in this study.

1035
1036

1037 **DATA AVAILABILITY**

1038 Samples were provided by the University of Wisconsin Alzheimer's Disease Research Center.
1039 Clinical data can be requested from the National Alzheimer's Coordinating Center
1040 (naccddata.org/).

1041 Data will be available in the Synapse AD Knowledge Portal.

1042 Gut Microbiome data is stored and accessible via the University of California, San Diego Qiita
1043 platform (qiita.ucsd.edu/).

1044
1045

1046 **Keywords**

1047 Alzheimer's disease; Gut microbiome; Composition; Function; Cerebrospinal fluid; Biomarkers;
1048 Differential abundance; Pathology

1049

1050

1051 **Figure legends**

1052

1053 **Figure 1**

1054 **Alpha and beta diversity metrics across AD groups.** (A-C) Shannon, Evenness, and Faith's
1055 PD metrics for individuals categorized by clinical diagnosis (CU vs. Dementia-AD). (D-F) Metrics
1056 for amyloid status (Negative vs. Positive). (G-I) Metrics across *APOE* $\epsilon 4$ status (Negative vs.
1057 Positive). Each box plot is overlaid with individual data points, enhancing visualization of the
1058 data distribution within each group. Kruskal-Wallis test was used to determine statistical
1059 significance. (J-L) Differences in beta diversity metrics (Bray Curtis, Weighted UniFrac, and
1060 Unweighted UniFrac, respectively) for individuals categorized by clinical diagnosis (CU vs.
1061 Dementia-AD). (M-O) Metrics for amyloid status (Negative vs. Positive). (P-R) Metrics across
1062 *APOE* $\epsilon 4$ status (Negative vs. Positive). Principal coordinates (PC)1 and PC2 axes represent
1063 the most variance in data. Each plot is color-coded by the respective group, highlighting the
1064 spatial distribution and clustering based on the dissimilarity indices. PERMANOVA was used to
1065 determine statistical significance.

1066

1067 **Figure 2**

1068 **Differential abundance (DA) across AD groups.** Forest plots illustrating the DA of microbial
1069 features associated with AD groups. (A, D, G, and J) Contrasts in the abundance of various
1070 bacterial taxa at (A) phylum, (D) family, (G) genus, and (J) species levels between AD dementia
1071 and CU. (B, E, H, and K) The differences in abundance at these taxonomic levels between A+
1072 and A- individuals. (C, F, I, and L) The microbial features differentially abundant between *APOE*
1073 $\epsilon 4+$ and *APOE* $\epsilon 4-$ groups. The x-axes quantify the log ratio of presence between groups, with
1074 values above one indicating a higher abundance in the first-mentioned group. Circles denote
1075 "Top" features, indicating a positive association with AD groups (dementia diagnosis, amyloid
1076 positivity, and *APOE* $\epsilon 4$ positivity), whereas triangles denote "Bottom" features, indicating a
1077 negative association. The lines are color-coded by unique phylum as labeled in the legend. DA
1078 analysis was conducted using BIRDMAn.

1079

1080 **Figure 3**

1081 **Comparative analysis of top and bottom features across AD groups.** Box plots comparing
1082 the distribution of log-transformed ratios of differentially abundant microbial species in relation to
1083 diagnosis, amyloid status, and *APOE* $\epsilon 4$ status. (A-C) The log-transformed ratios of microbes for
1084 diagnosis groups (AD vs CU). (D-F) The log-transformed ratios of amyloid-related microbes (A+
1085 vs A-). (G-I) The log-transformed ratios of *APOE* $\epsilon 4$ -related microbes (*APOE* $\epsilon 4+$ vs *APOE*
1086 $\epsilon 4-$). Each column compares the CU and AD groups (A, D, and G), A+ and A- groups (B, E,
1087 and H), and *APOE* $\epsilon 4+$ and *APOE* $\epsilon 4-$ groups (C, F, and I). Each panel includes a Kruskal-

1088 Wallis test statistic and associated P value, indicating the statistical significance of the
1089 differences observed.

1090

1091 **Figure 4**

1092 **Venn-diagram of co-occurrence of microbial features across AD groups.** The diagrams on
1093 the left column (A, C, E, and G) depict the Top (positively-associated) differentially abundant
1094 features, while those on the right column (B, D, F, and H) show the Bottom (negatively-
1095 associated) differentially abundant features. (A and B) Top and Bottom microbial phyla,
1096 respectively. These diagrams identify unique and shared phyla associated with each of the
1097 three AD groups. (C and D) Top and Bottom microbial families, respectively. These diagrams
1098 highlight the family-level microbial differences that correlate with AD diagnosis, amyloid
1099 presence, and *APOE* $\epsilon 4$ genotype presence. (E and F) Top and Bottom microbial genera,
1100 respectively. These diagrams provide insight into the genus-level microbial composition
1101 influenced by the specified AD groups. (G and H) Top and Bottom microbial species,
1102 respectively. These diagrams detail the number of species that are unique and shared across
1103 the three AD groups. Each diagram contains colored regions representing intersections
1104 between the groups: red for dementia, green for amyloid, and blue for *APOE* $\epsilon 4$. The numbers
1105 within each segment of the diagrams indicate the count of microbial features unique to or
1106 shared between the conditions. Specific microbial features are listed in **Table S2**.

1107

1108 **Figure 5**

1109 **Comparison of log-transformed dementia biomarker ratios in CU and AD dementia**
1110 **across two cohorts.** (A) The results from the MARS cohort. Box plots show the distribution of
1111 log-transformed ratios of top dementia biomarkers to bottom dementia biomarkers for CU
1112 individuals (light blue) and AD dementia (dark blue) (Kruskal-Wallis = 31.81, P value < .001). (B)
1113 The results from a larger validation cohort ($n = 448$). Box plots present the distribution of log-
1114 transformed ratios of top dementia biomarkers to bottom dementia biomarkers found in the
1115 MARS cohort for CU individuals (light blue) and people with AD dementia (dark blue) (Kruskal-
1116 Wallis = 5.59, P value = .02). Each point represents an individual sample, with the boxes
1117 indicating the interquartile range (IQR) and the whiskers extending to 1.5 times the IQR. The
1118 horizontal line within each box denotes the median value.

1119

1120 **Figure 6**

1121 **Differentially abundant microbial pathways between AD and CU.** (A) The distribution of the
1122 log ratios of Top/Bottom pathway features between AD (orange) and CU (blue) was shown in a
1123 box plot. Mann-Whitney U test was performed to determine statistical significance. Asterisks
1124 indicate a significant difference between AD (4.90) and CU (2.74) groups in the median of the
1125 log ratios of Top/Bottom pathway features (P value < .001). (B) A total of 36 differentially
1126 abundant features of microbial species and their corresponding pathways between AD and CU
1127 were displayed in a forest plot. Circles denote "Top" features, indicating a positive association
1128 with AD, whereas triangles denote "Bottom" features, indicating a negative association. The
1129 lines are color-coded by unique species and their corresponding pathways. DA analysis was
1130 conducted using BIRDMan. (C) RPCA on the clinical diagnosis group and a biplot of
1131 microbiome pathway features and their corresponding species. Each point represents an

Gut Microbiome Compositional and Functional Features Associate with Alzheimer's Disease Pathology

1132 individual sample color-coded by the respective group, with CU colored in blue and AD colored
1133 in orange. Vectors represent the direction (arrows) and magnitude (length) of the contribution of
1134 feature variables to the principal components (PCs). Vectors in red indicate Top features and
1135 vectors in green indicate Bottom features. PC1 and PC2 axes represent the most variance in
1136 data. Statistical analysis on RPCA was performed with PERMANOVA between AD and CU
1137 groups.

1138

1139 **Figure 7**

1140 **Heatmap illustrating the associations between gut microbiome compositional and**
1141 **functional features and CSF biomarkers in AD and related pathologies.** (A) This heatmap
1142 represents the coefficients of regression analysis between the top and bottom 20 gut microbial
1143 species linked to dementia and CSF biomarkers in two groups: Top (more abundant in AD,
1144 denoted by the pink bar) and Bottom (less abundant in AD, denoted by the green bar). The color
1145 scale indicates the strength and direction of the associations, with red representing positive
1146 associations and blue representing negative associations. The intensity of the color corresponds
1147 to the magnitude of the coefficient. Listed on the left are the gut microbiome species that were
1148 identified as more or less abundant in dementia-AD through BIRDMA. (B) The heatmap
1149 depicts the coefficients of regression analysis between the gut microbial pathways and CSF
1150 biomarkers. Coefficients are scaled by colors indicating the strength and direction of the
1151 associations, with green representing positive associations and pink representing negative
1152 associations. The intensity of the color corresponds to the magnitude (strength) of the
1153 coefficient. Microbial species and their associated pathway features are listed on the left of the
1154 plot and two groups (Top: more abundant in AD, denoted by the light pink bar; and Bottom: less
1155 abundant in AD or more abundant in CU, denoted by the light green bar) from DA analysis using
1156 BIRDMA are displayed on the right of the plot.

1157 The biomarkers listed along the bottom include amyloid pathology ($A\beta_{42}/A\beta_{40}$), tau
1158 pathophysiology (pTau₁₈₁ and tTau), neurodegeneration (NfL), synaptic dysfunction and injury
1159 (neurogranin and α -synuclein), inflammation (IL-6), and glial activation (S100B, GFAP, YKL-40,
1160 and sTREM2). Asterisks indicate the level of statistical significance of the associations: *** $P <$
1161 $.001$, ** $P <$ $.01$, and * $P <$ $.05$ (uncorrected).

Table 1. Participant demographics at fecal sample collection by clinical diagnosis.

Variable	N	Overall, N = 232 [†]	Dementia-AD, N = 24 [†]	CU, N = 208 [†]	P value [‡]
Age	232	67 (±7)	71 (±7) ^{§**}	66 (±7)	.002
Sex	232				.7
Female		142 (61%)	16 (67%)	126 (61%)	
Male		90 (39%)	8 (33%)	82 (39%)	
Race	232				>0.9
Black or African American		10 (4.3%)	1 (4.2%)	9 (4.3%)	
White		222 (96%)	23 (96%)	199 (96%)	
APOE genotype	227				<.001
ε2ε3		22 (9.7%)	0 (0%)	22 (11%)	
ε3ε3		120 (53%)	5 (21%) ^{§**}	115 (57%)	
ε2ε4		5 (2.2%)	0 (0%)	5 (2.5%)	
ε3ε4		64 (28%)	11 (46%)	53 (26%)	
ε4ε4		16 (7.0%)	8 (33%) ^{§****}	8 (3.9%)	
APOE ε4 genotype	227				<.001
Negative (non-carrier)		142 (63%)	5 (21%)	137 (67%)	
Positive (carrier)		85 (37%)	19 (79%) ^{§****}	66 (33%)	
Bristol stool type	232				.031
1		16 (6.9%)	0 (0%)	16 (7.7%)	
2		19 (8.2%)	3 (13%)	16 (7.7%)	
3		35 (15%)	9 (38%) ^{§*}	26 (13%)	
4		105 (45%)	6 (25%)	99 (48%)	
5		37 (16%)	4 (17%)	33 (16%)	
6		19 (8.2%)	2 (8.3%)	17 (8.2%)	
7		1 (0.4%)	0 (0%)	1 (0.5%)	
BMI	232	28.2 (±5.4)	26.0 (±4.7) ^{§*}	28.4 (±5.4)	.030
Amyloid status	145				<.001
0 (A-)		104 (72%)	0 (0%)	104 (78%)	
1 (A+)		41 (28%)	12 (100%) ^{§****}	29 (22%)	
Medication status	232				.7
medicated		213 (92%)	23 (96%)	190 (91%)	
non-medicated		19 (8.2%)	1 (4.2%)	18 (8.7%)	

Abbreviations: AD, Alzheimer's disease; CU, cognitively unimpaired; APOE, apolipoprotein E; BMI, body mass index; A, amyloid status.

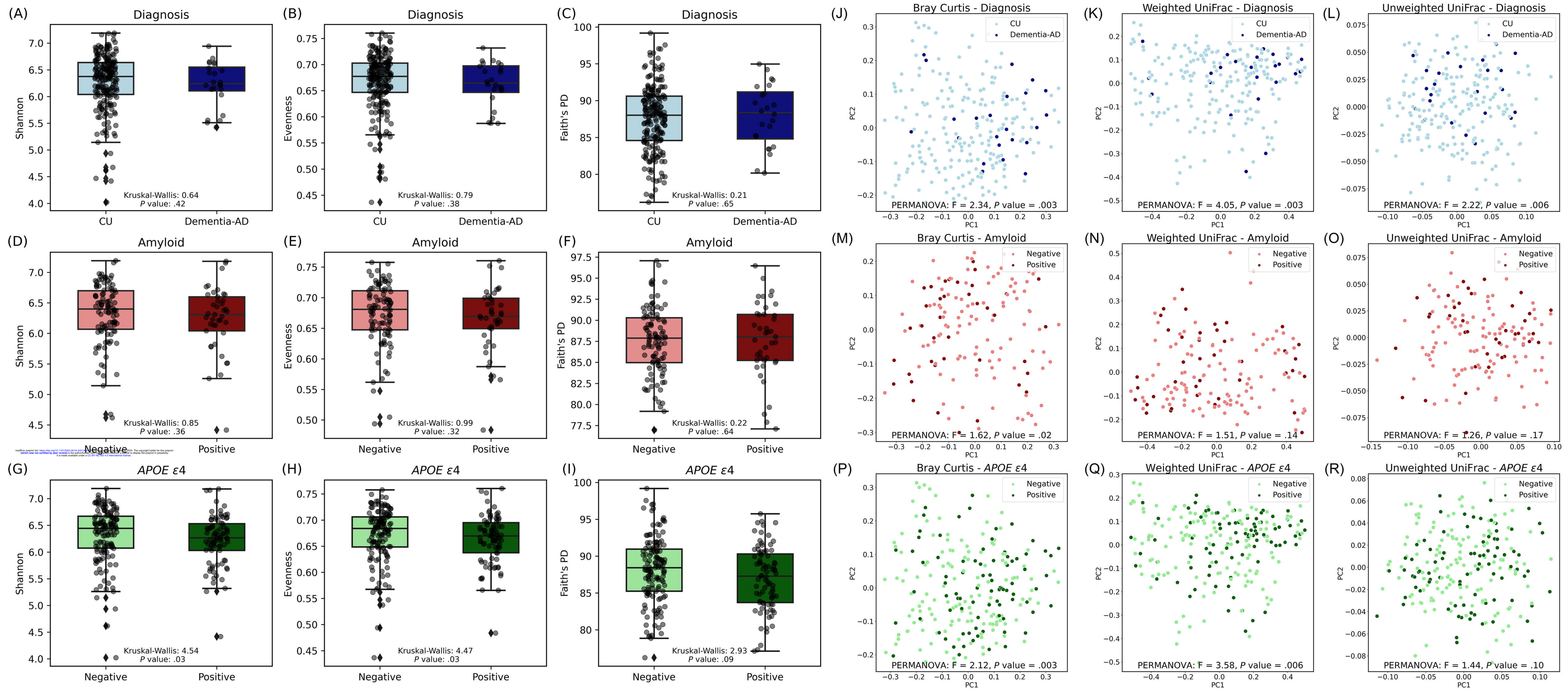
NOTE. The Bristol stool types are classification tools for the diagnosis of human feces form.

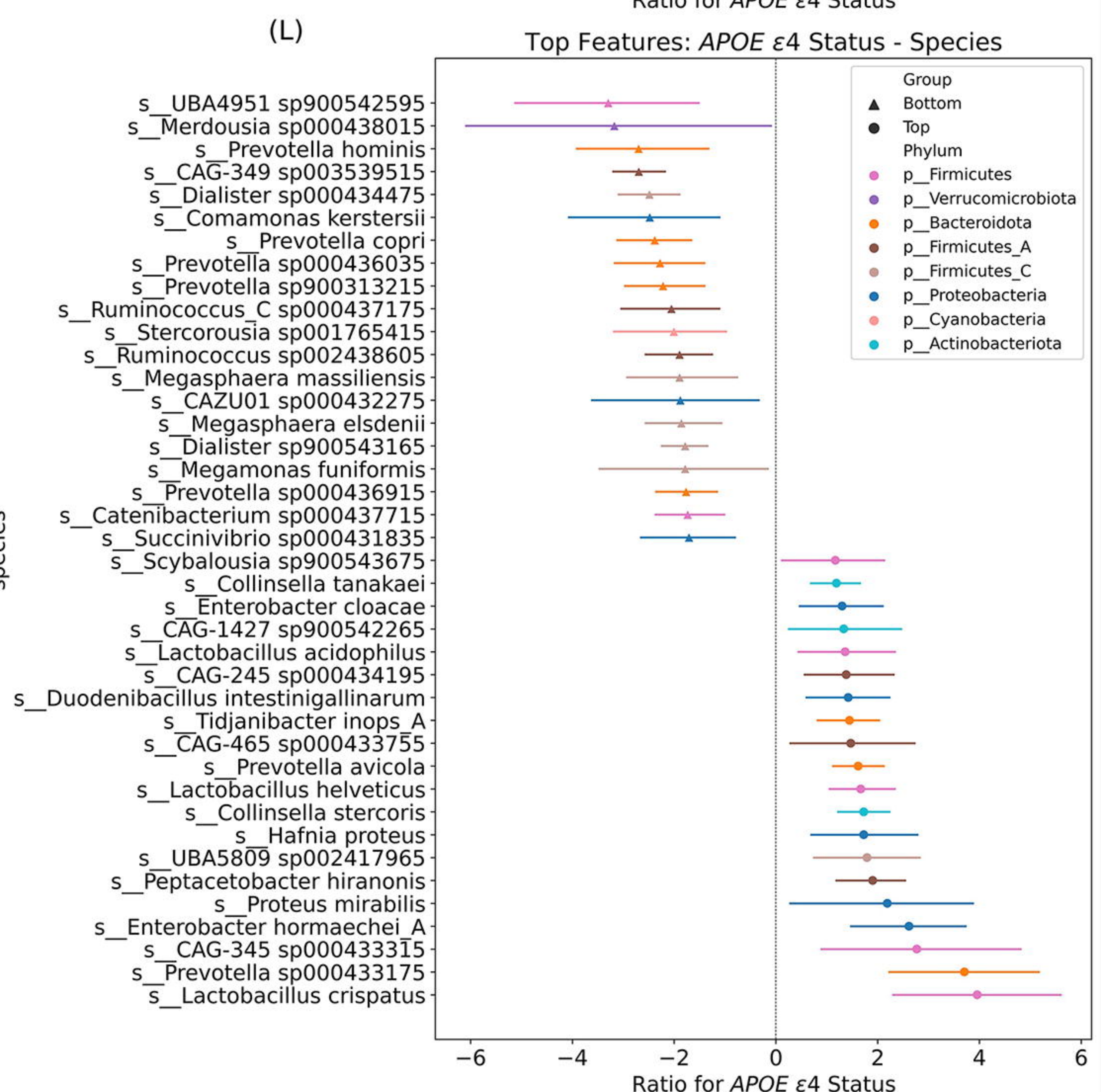
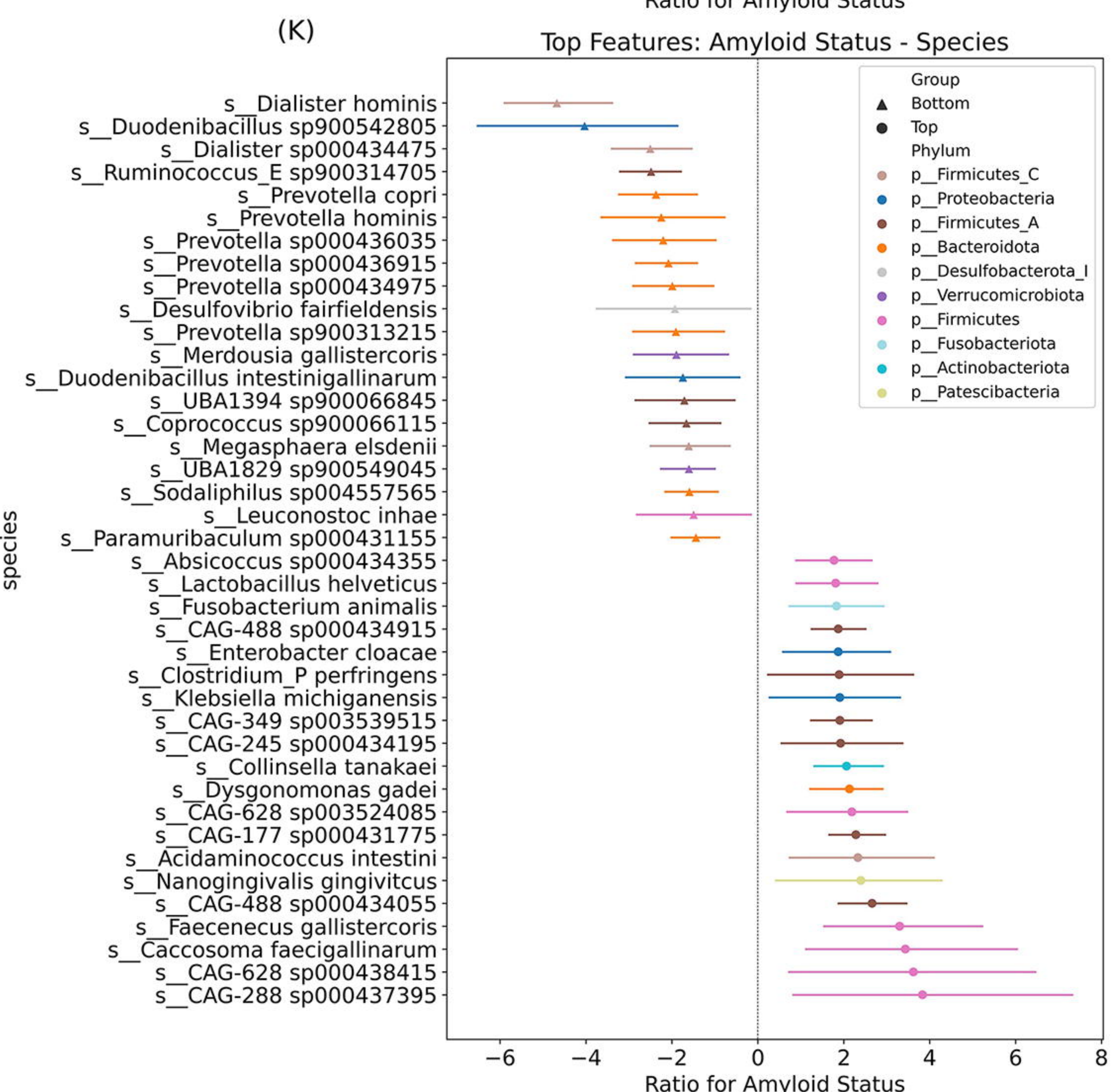
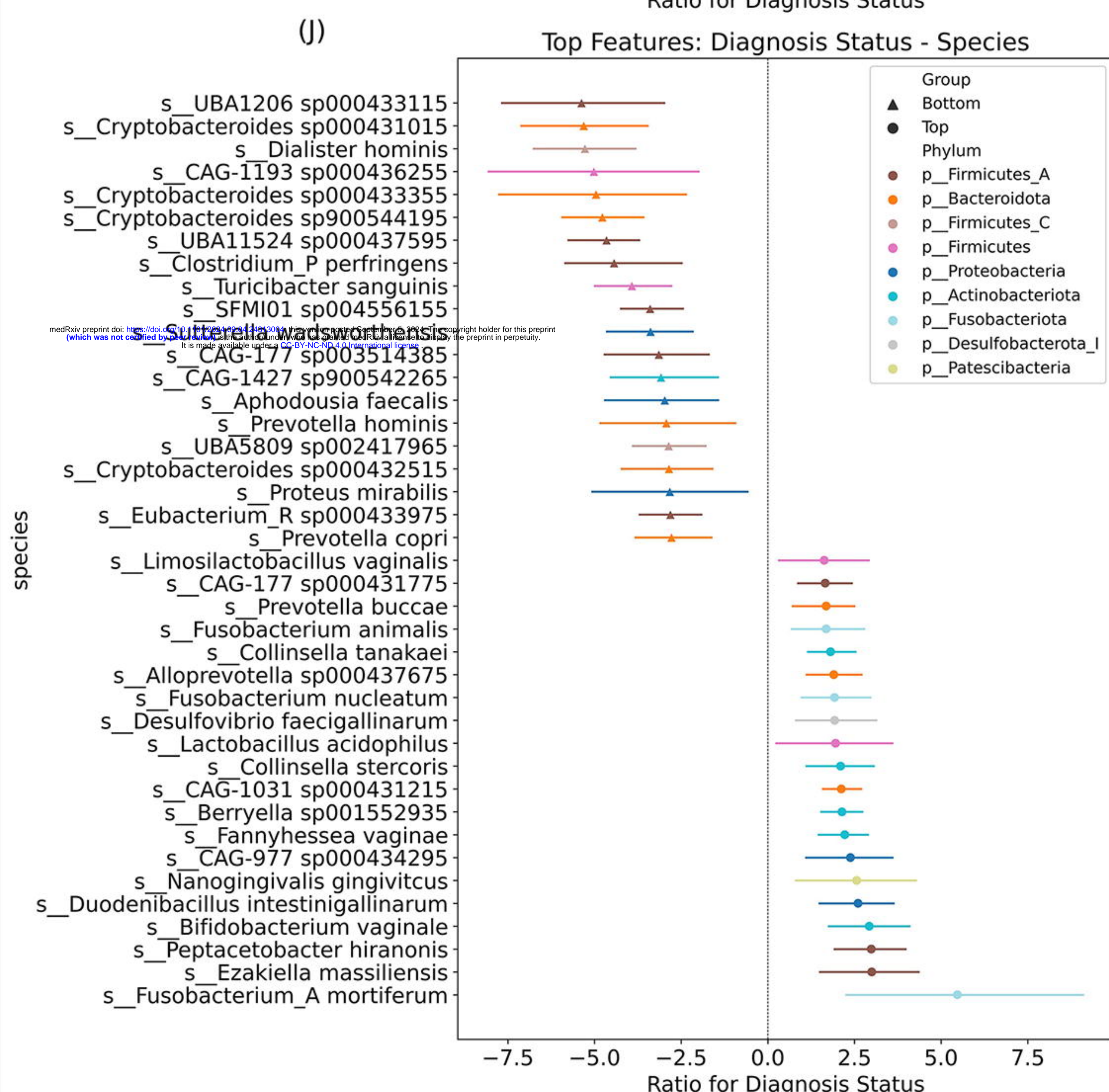
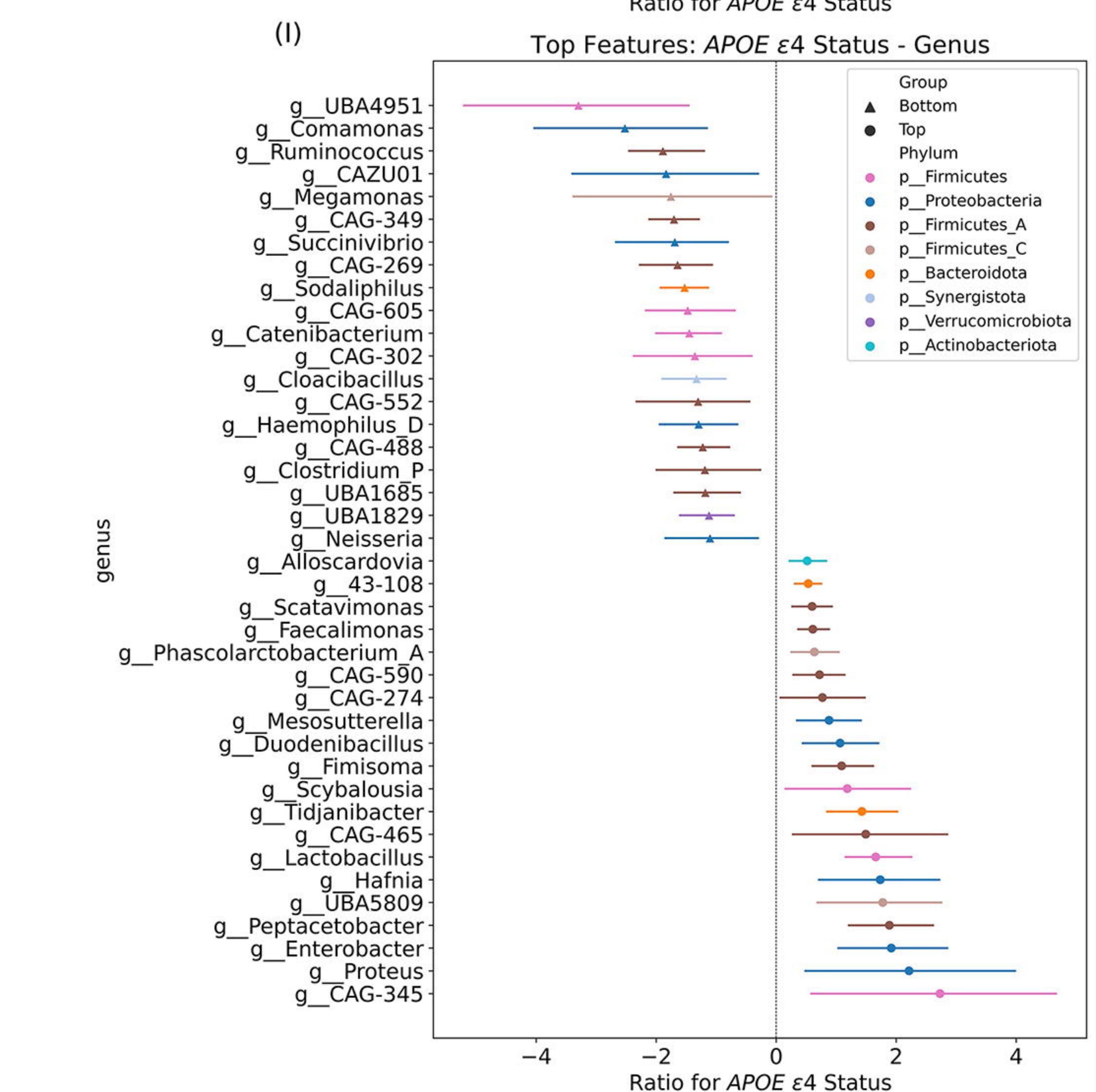
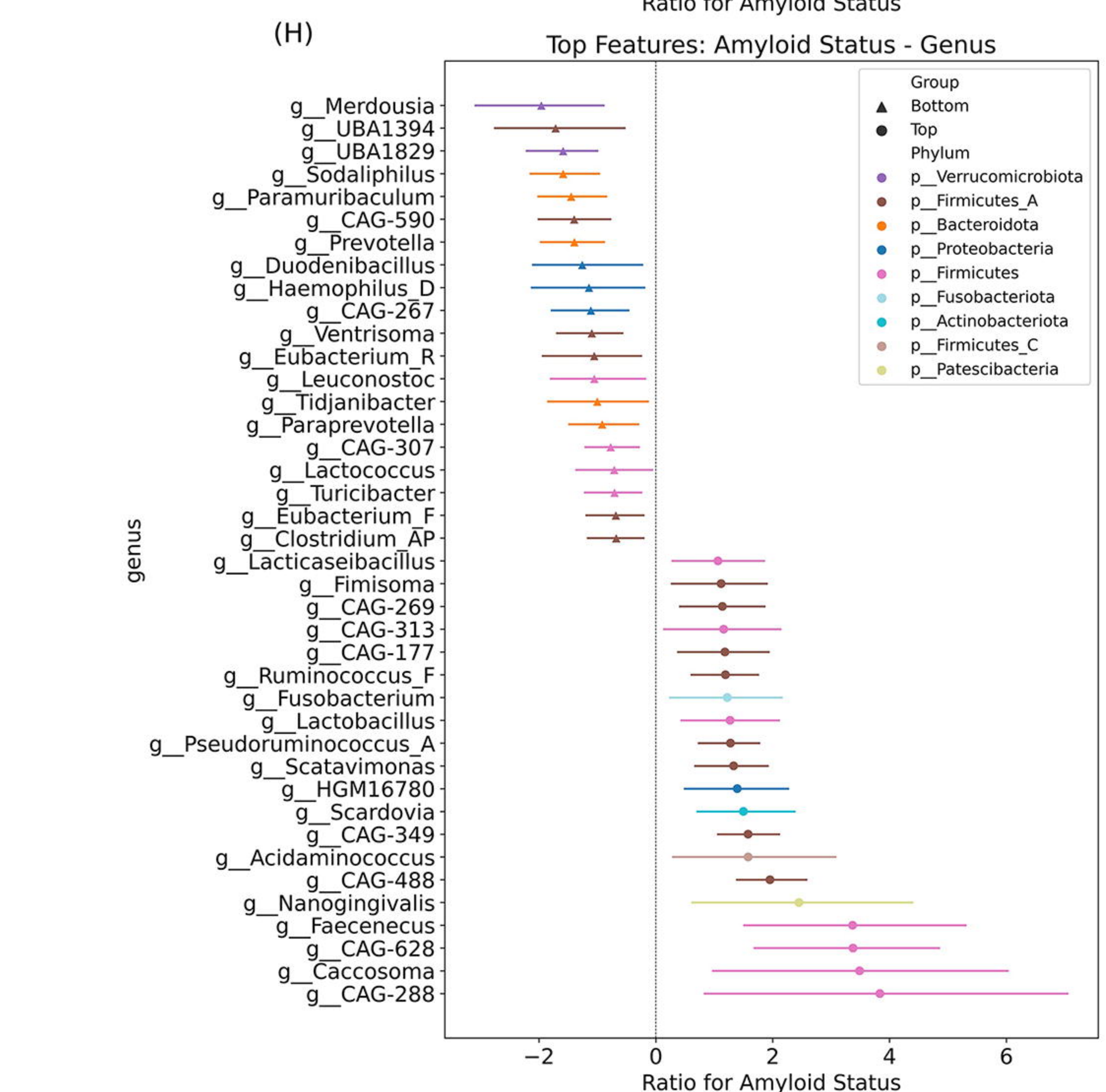
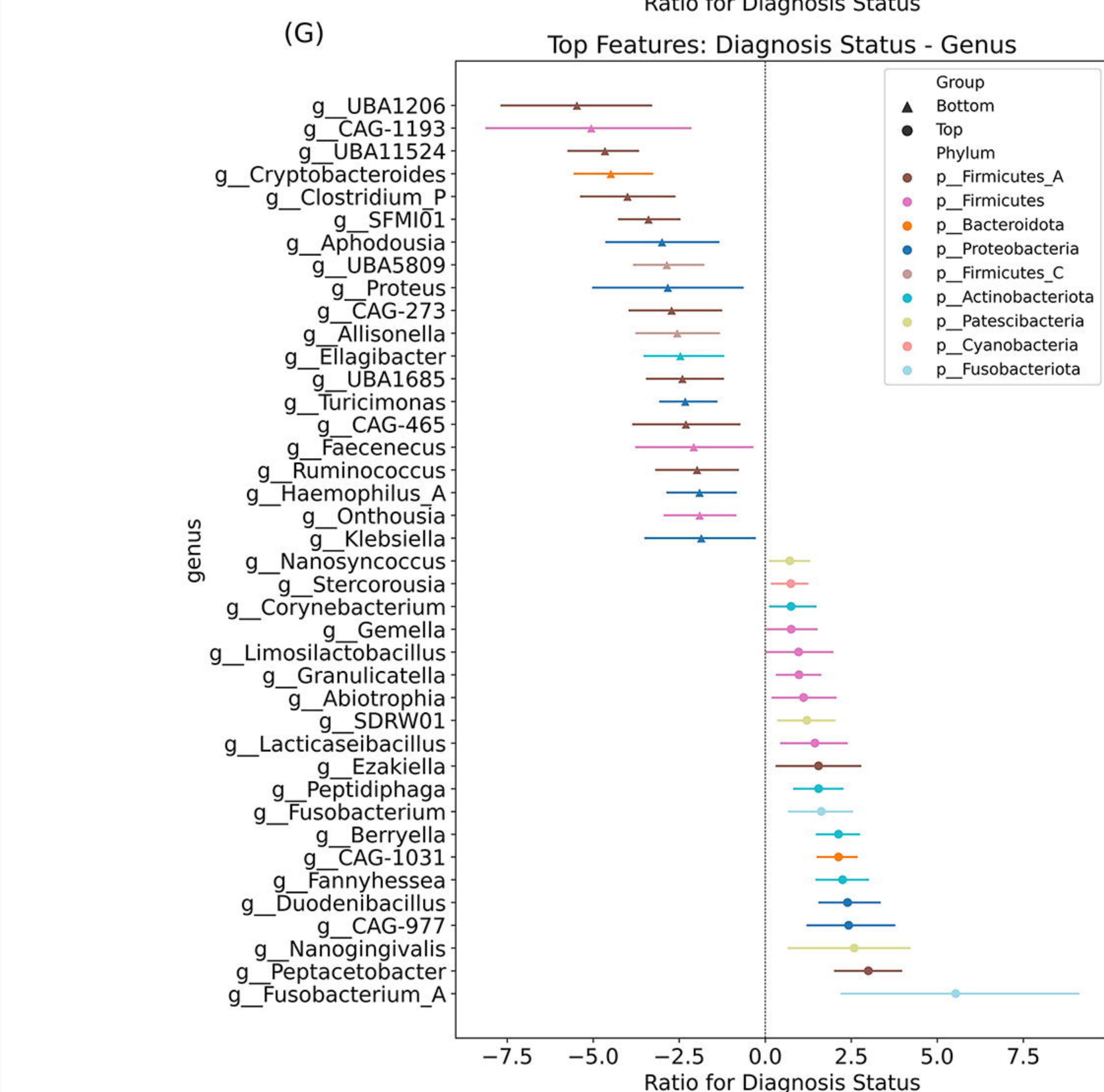
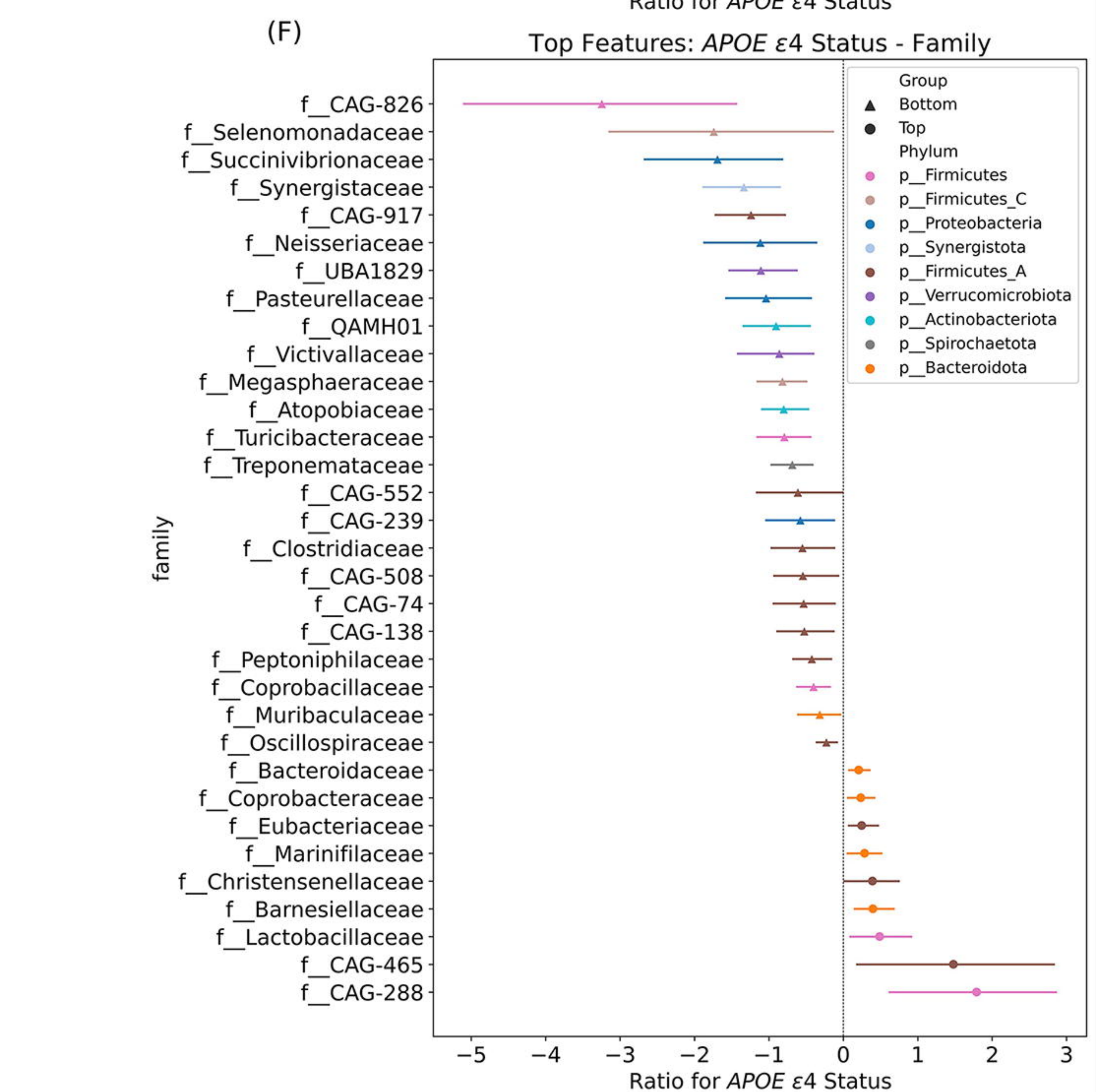
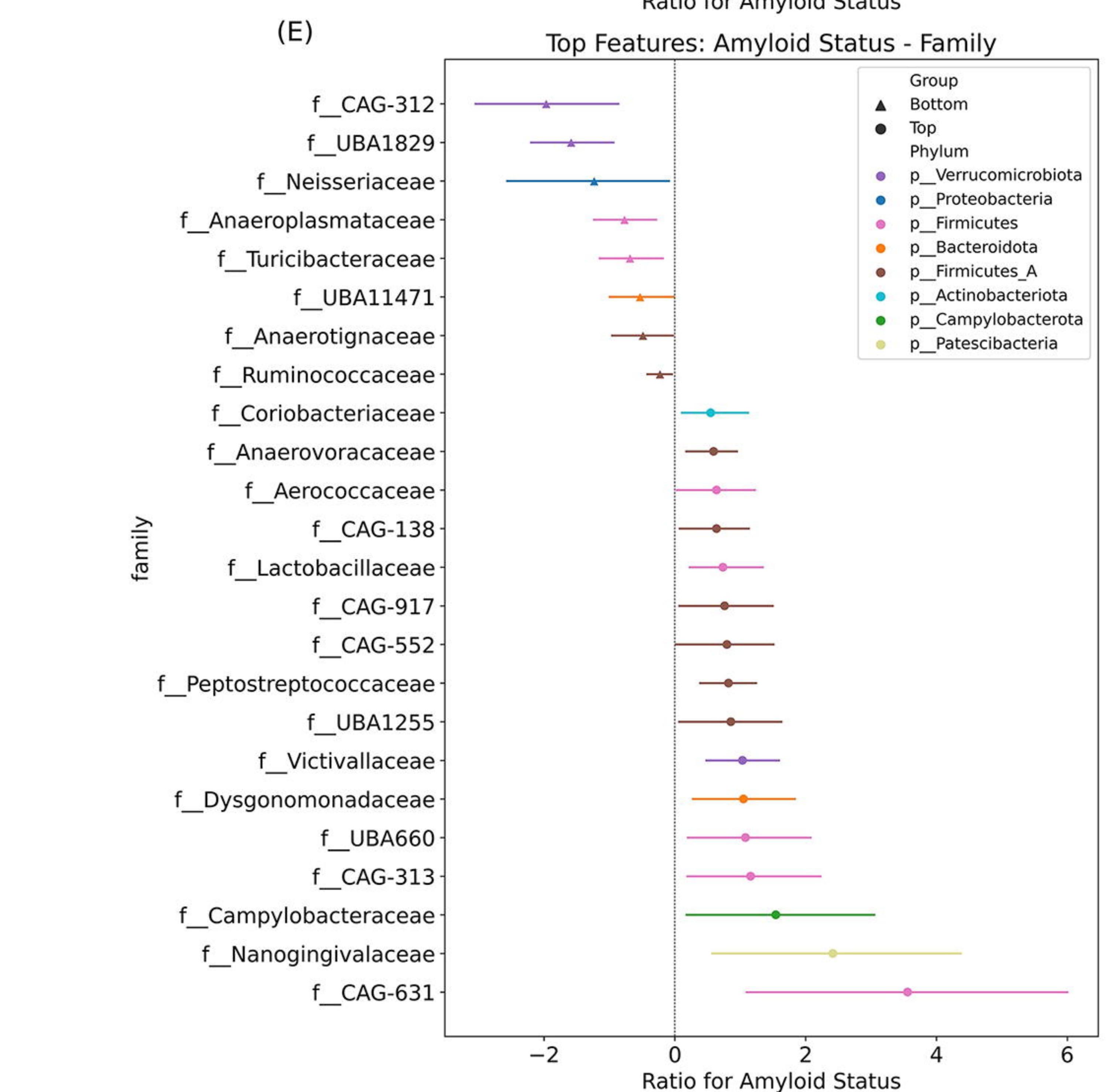
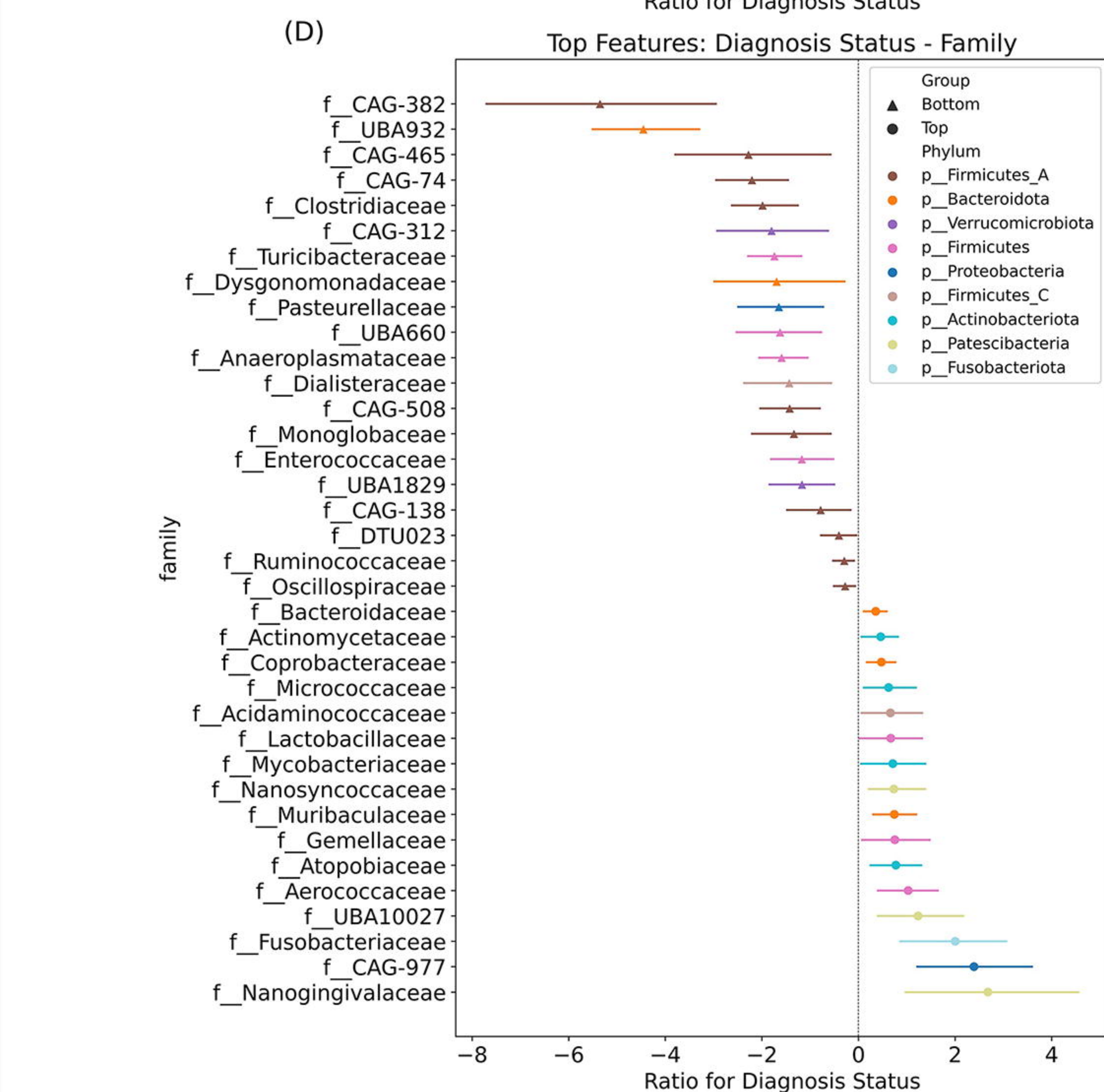
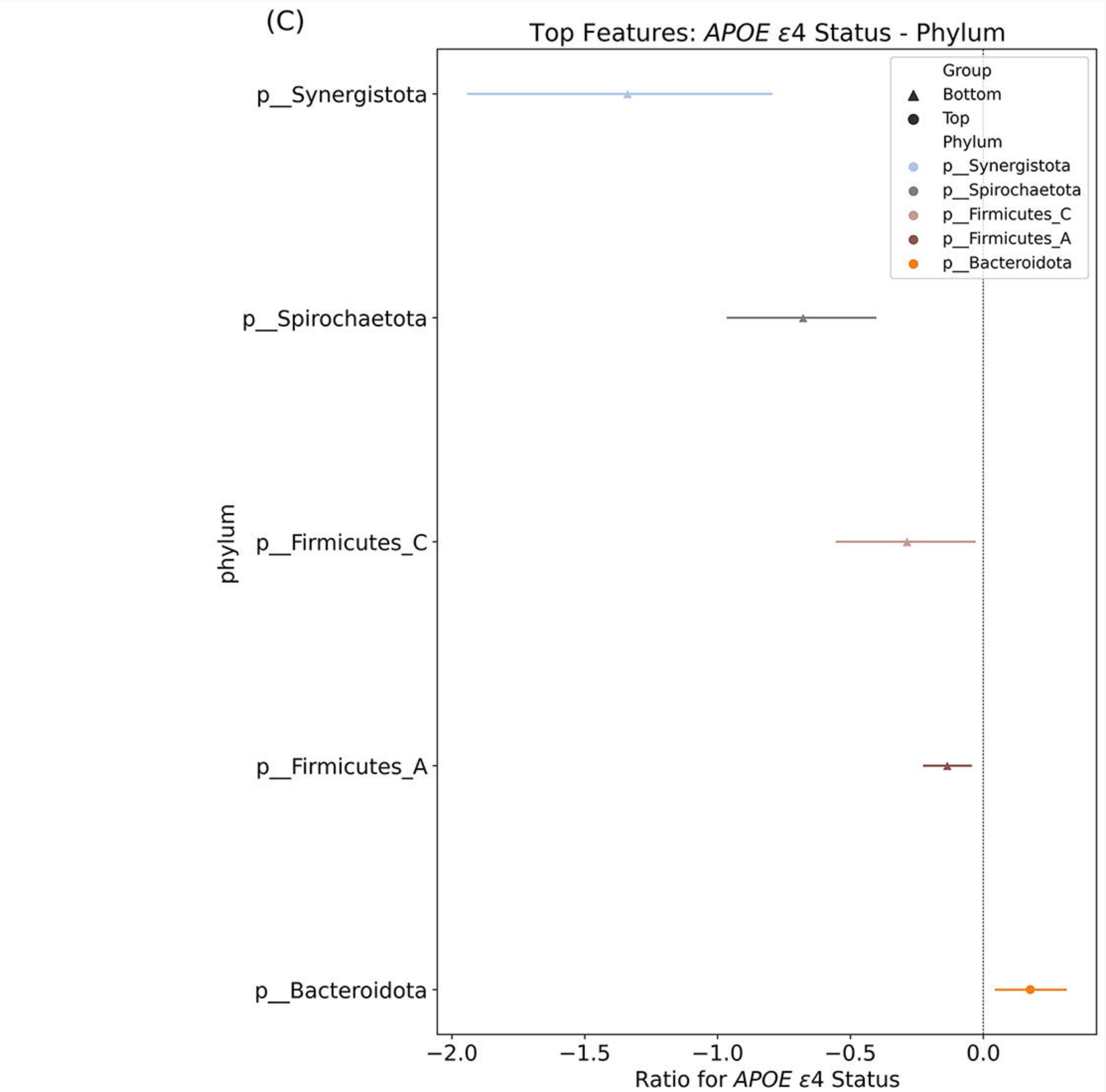
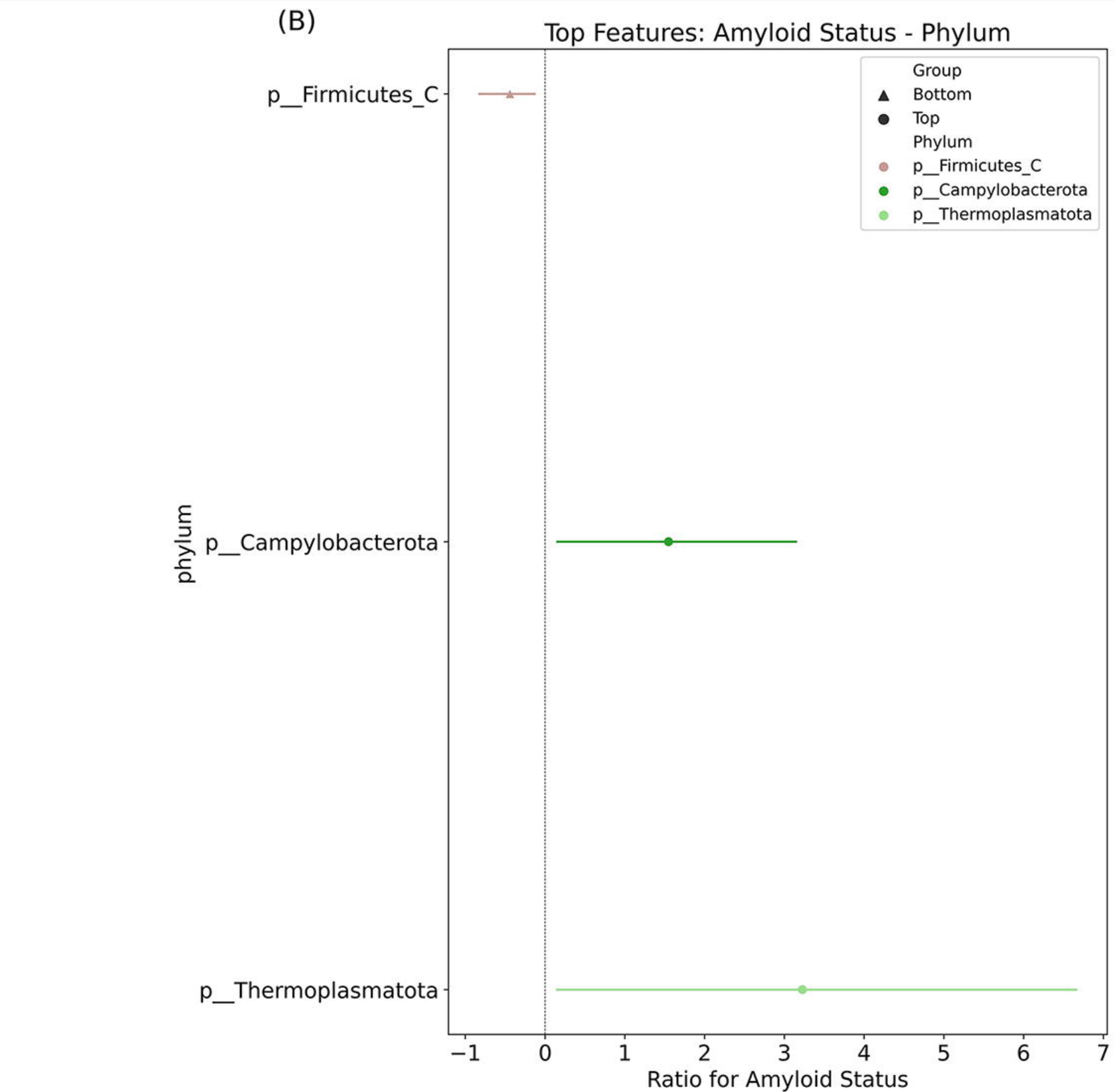
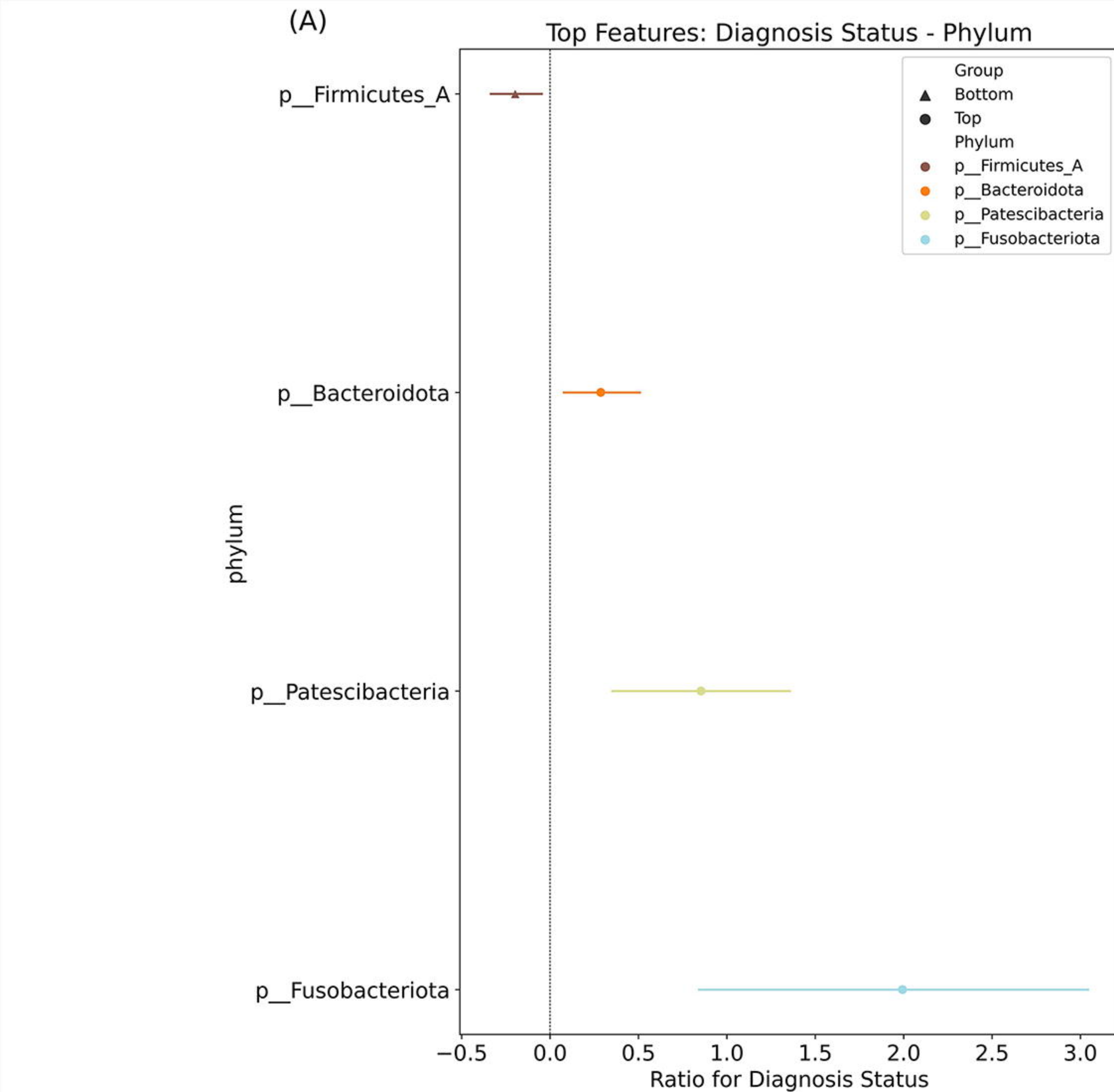
[†]Mean (±SD); n (%)

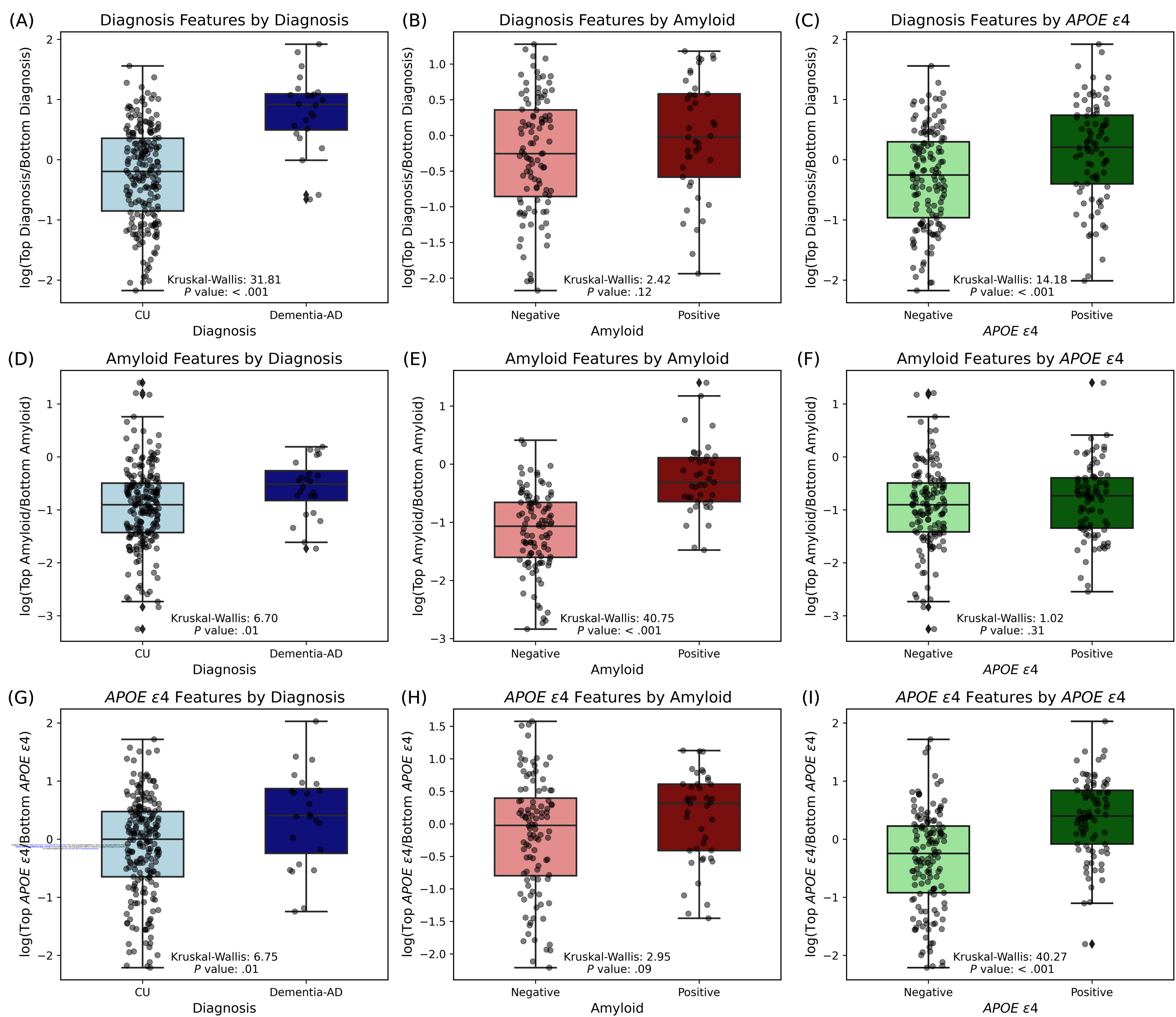
‡Kruskal-Wallis rank sum test; Pearson's Chi-squared test

§Significantly different Dementia-AD vs CU

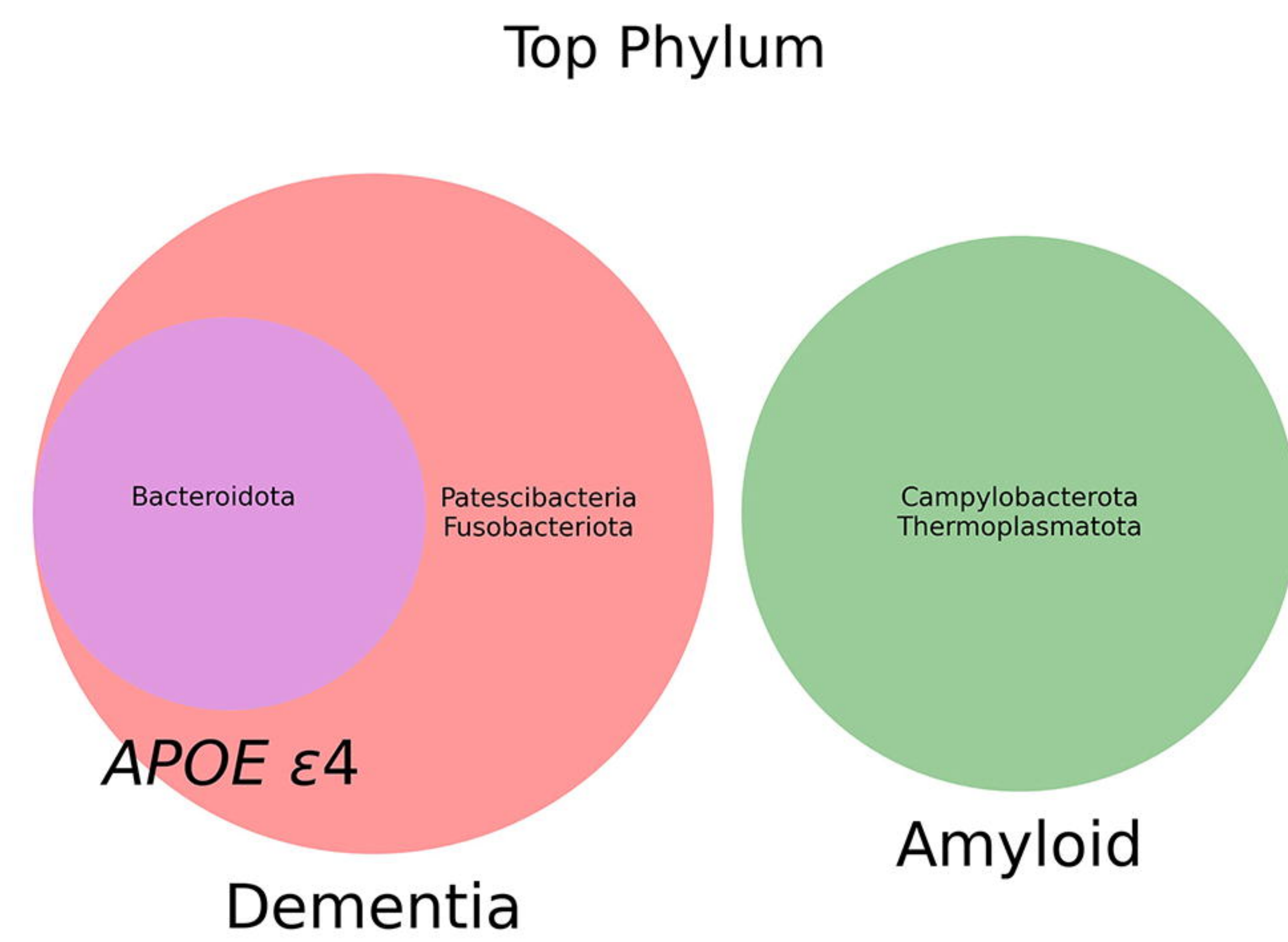
* $P < .05$, ** $P < .01$, *** $P < .001$, **** $P < .0001$ (P values are Bonferroni test corrected)



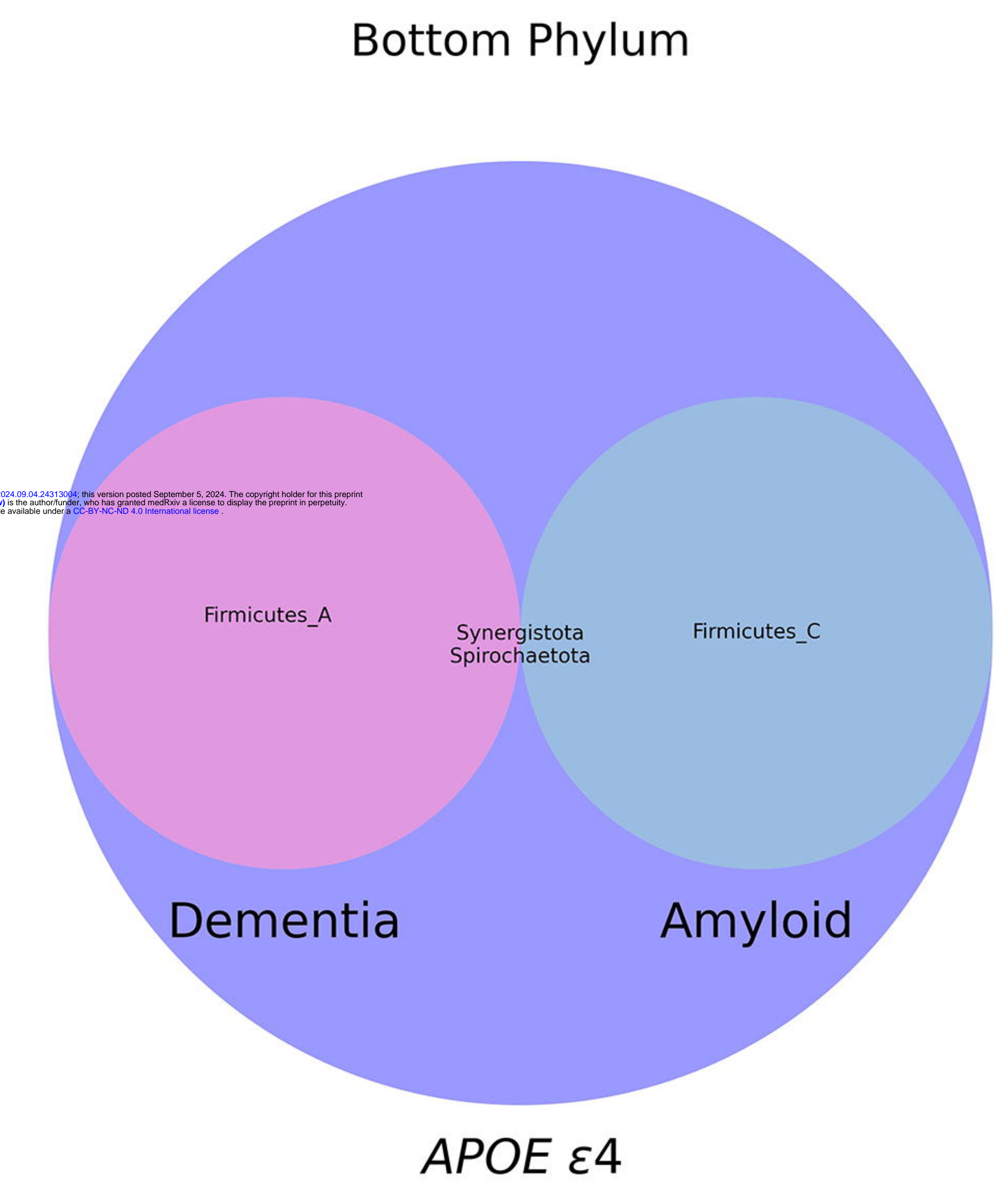




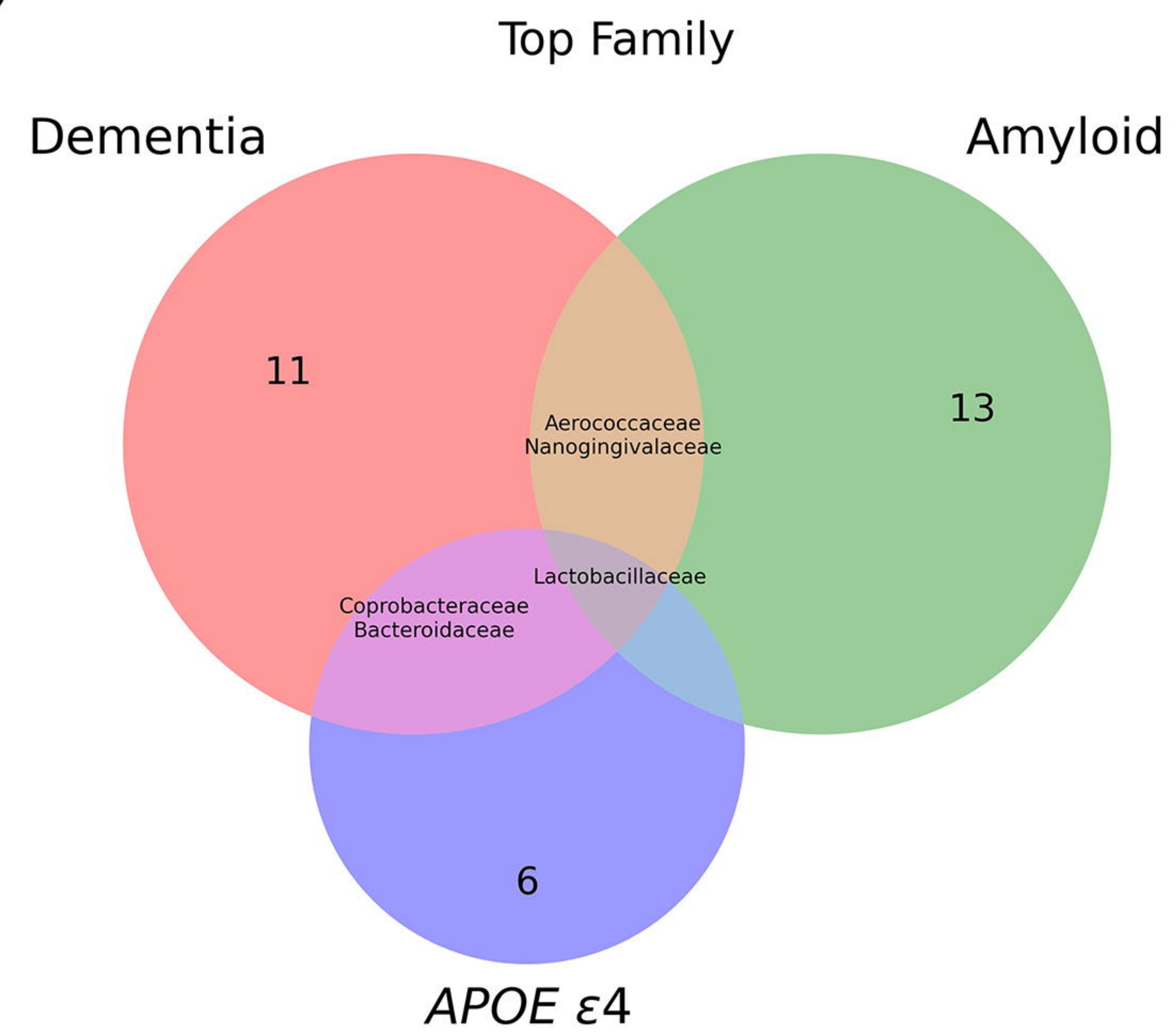
(A)



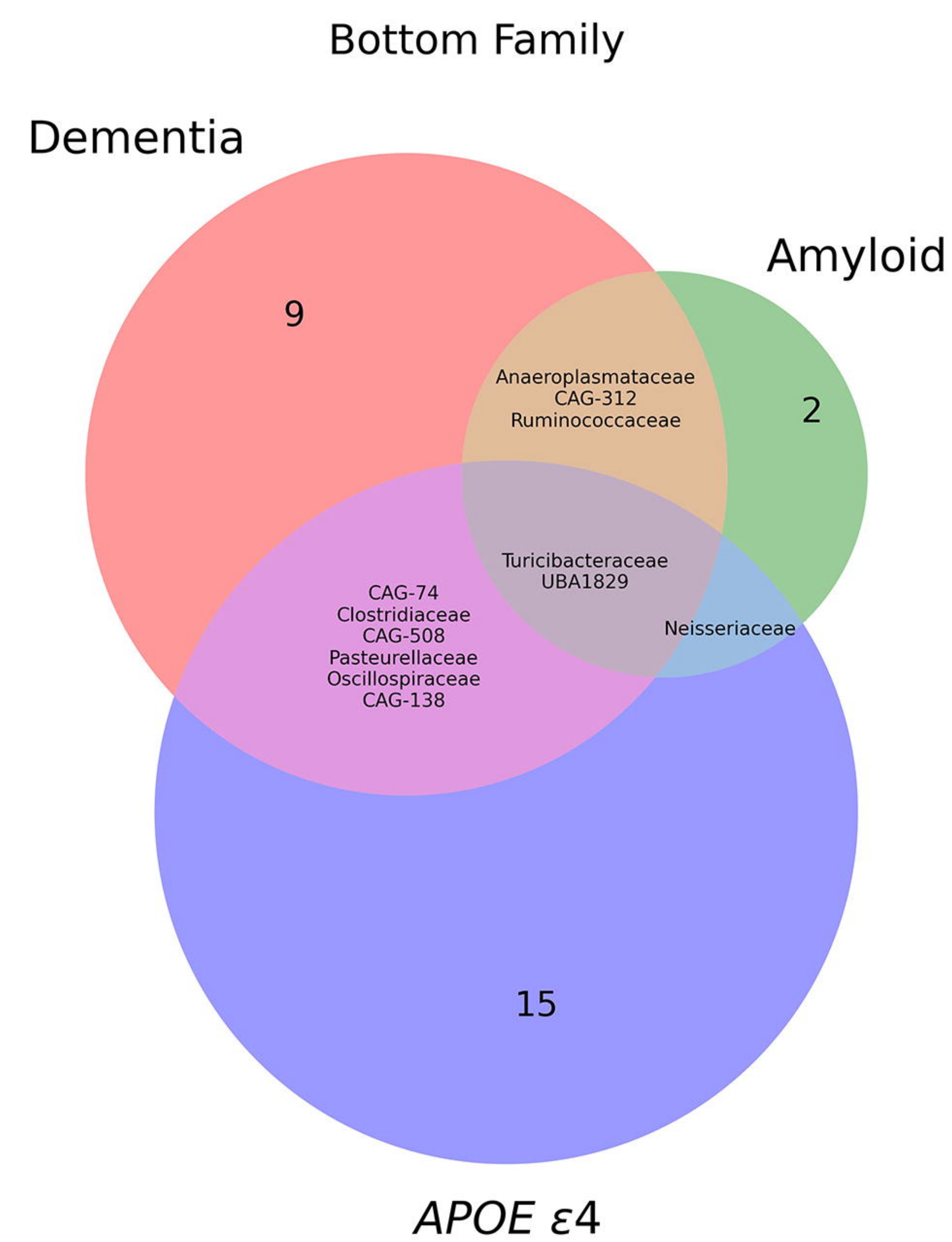
(B)



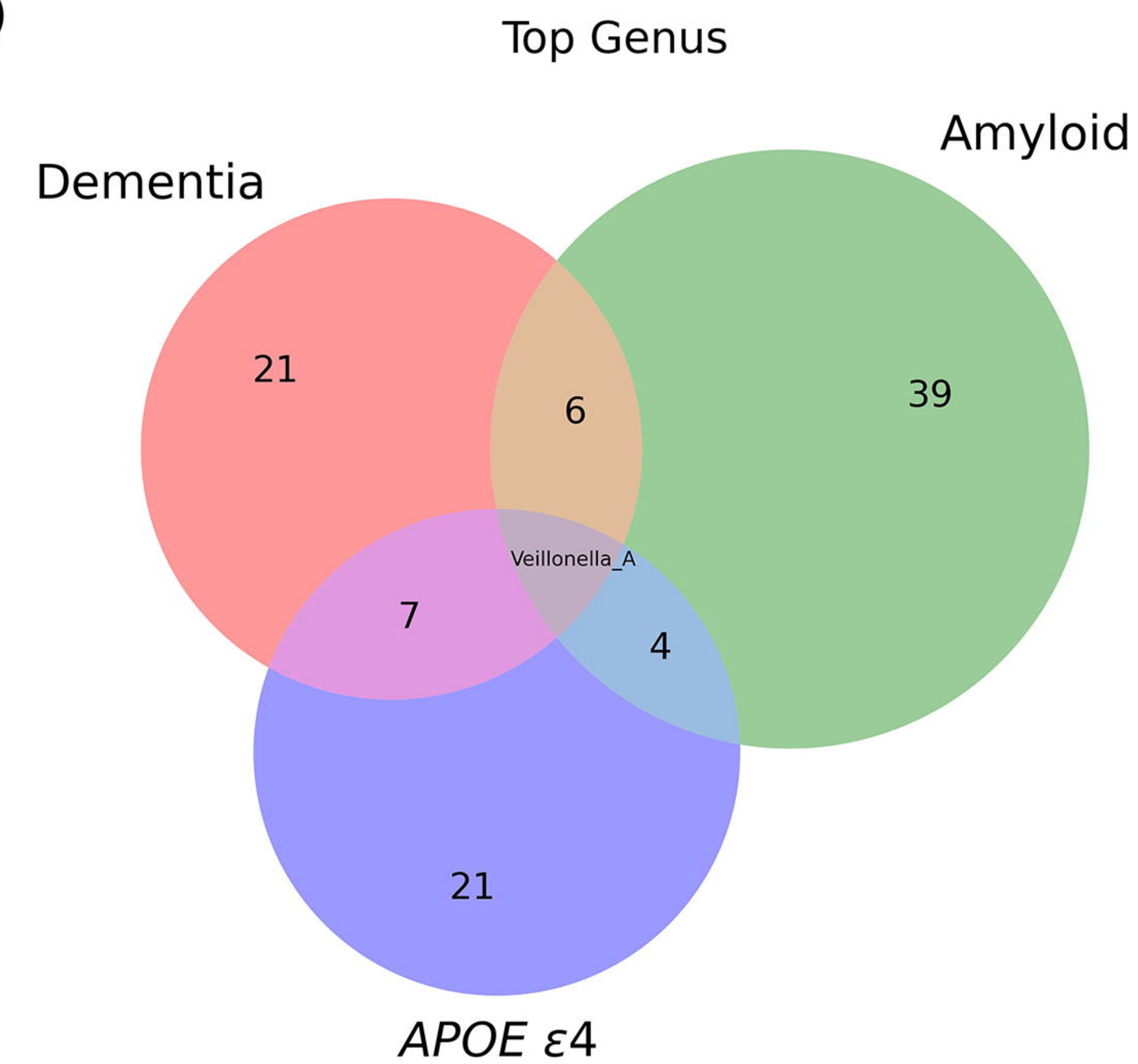
(C)



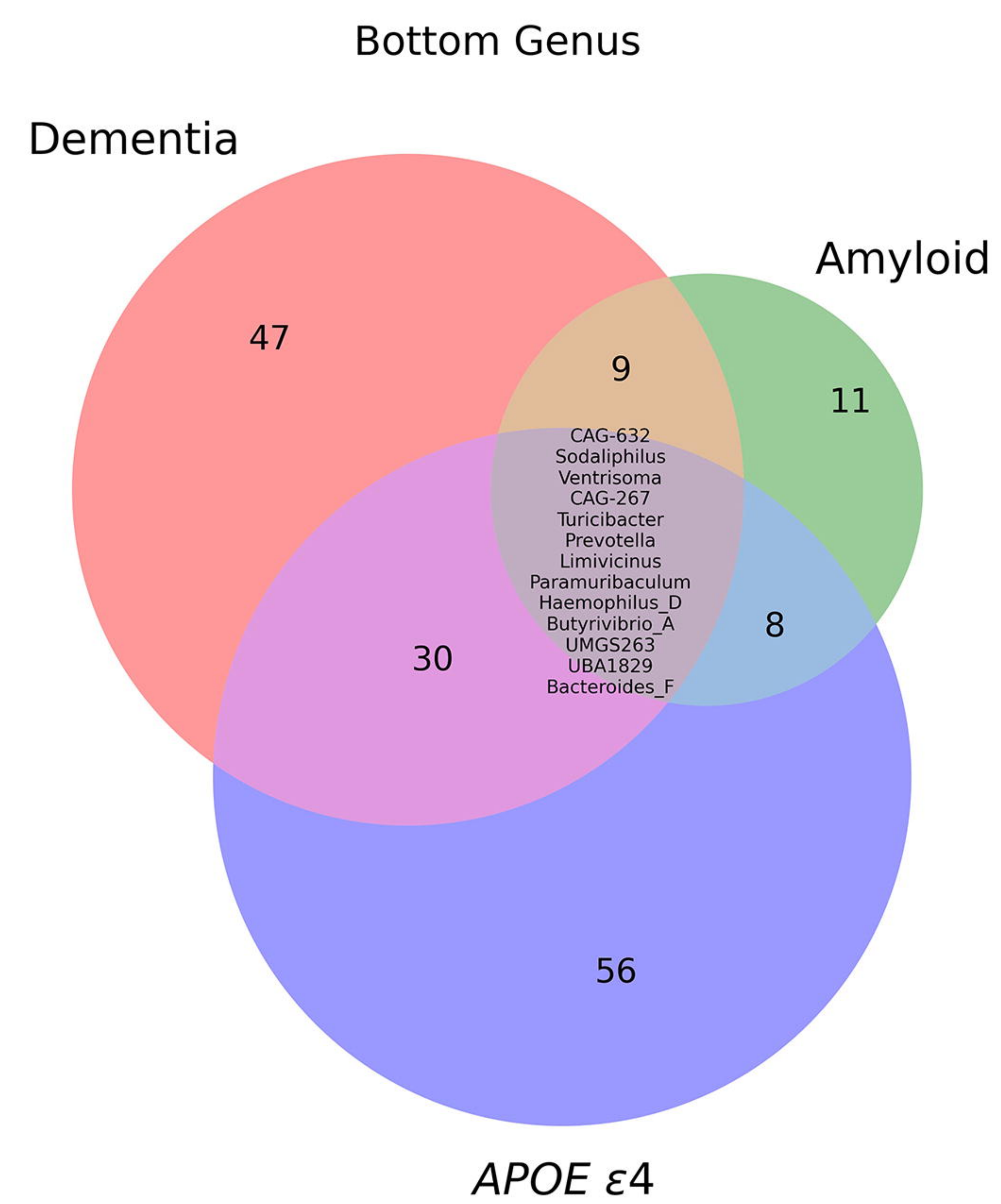
(D)



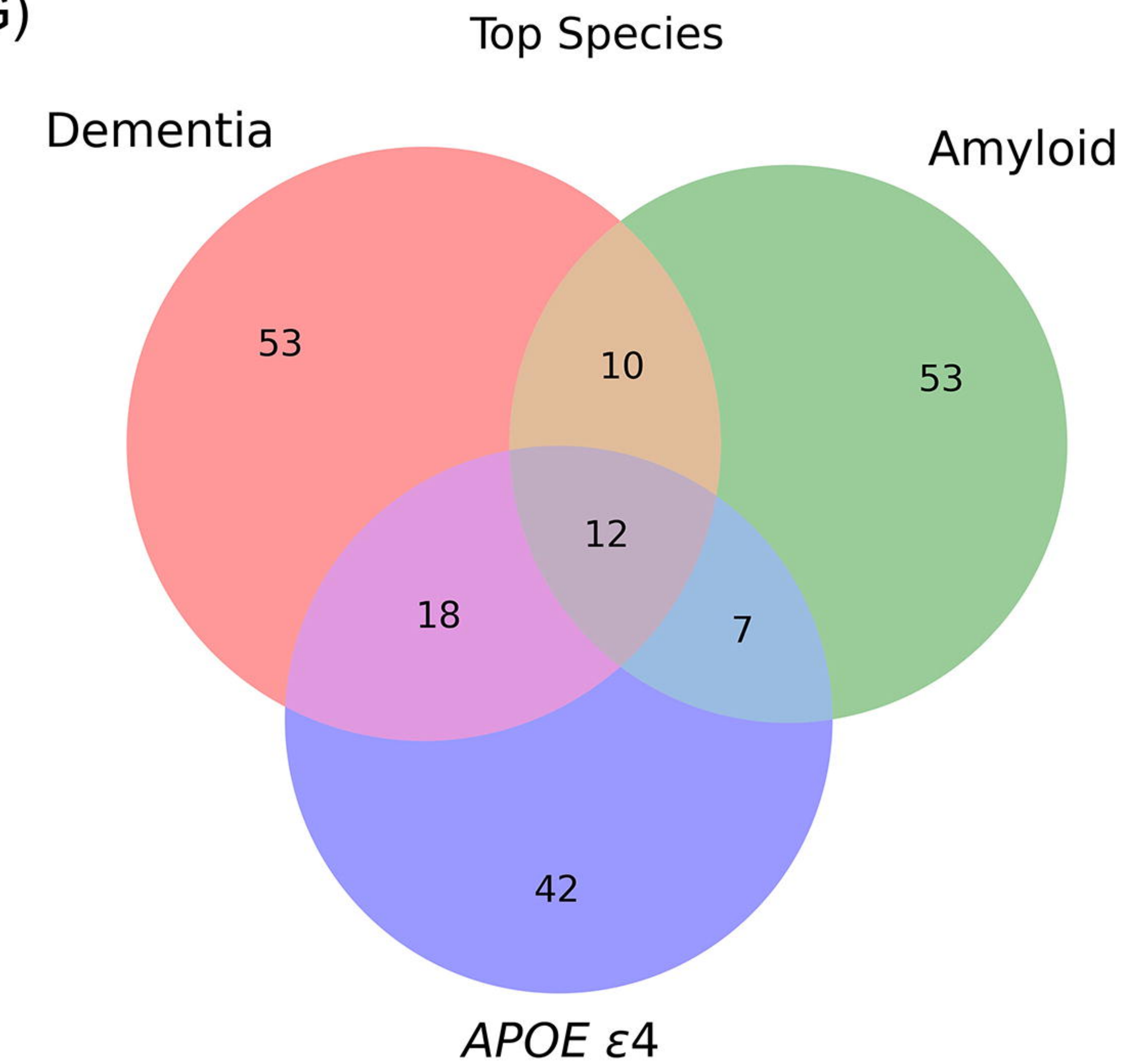
(E)



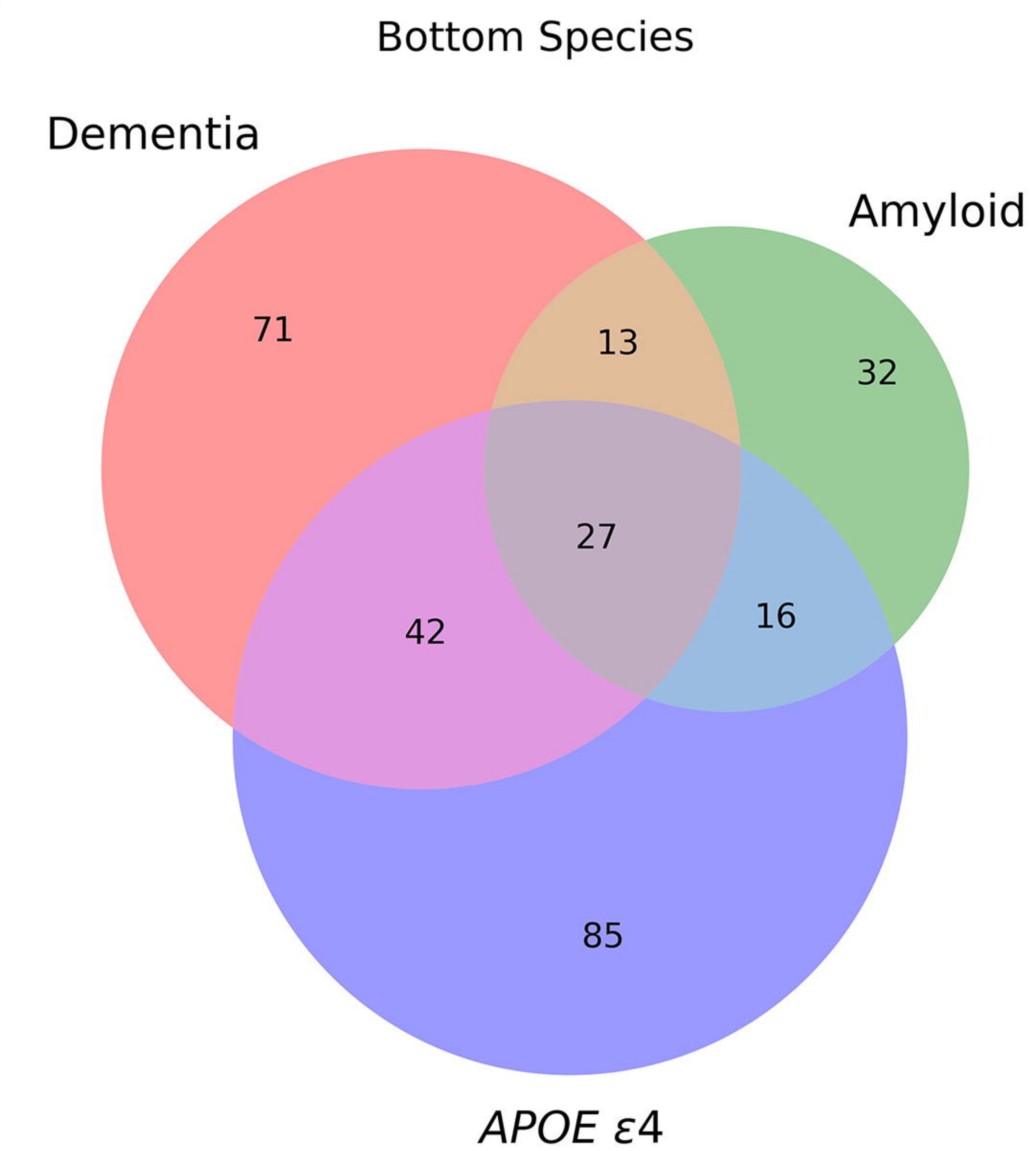
(F)

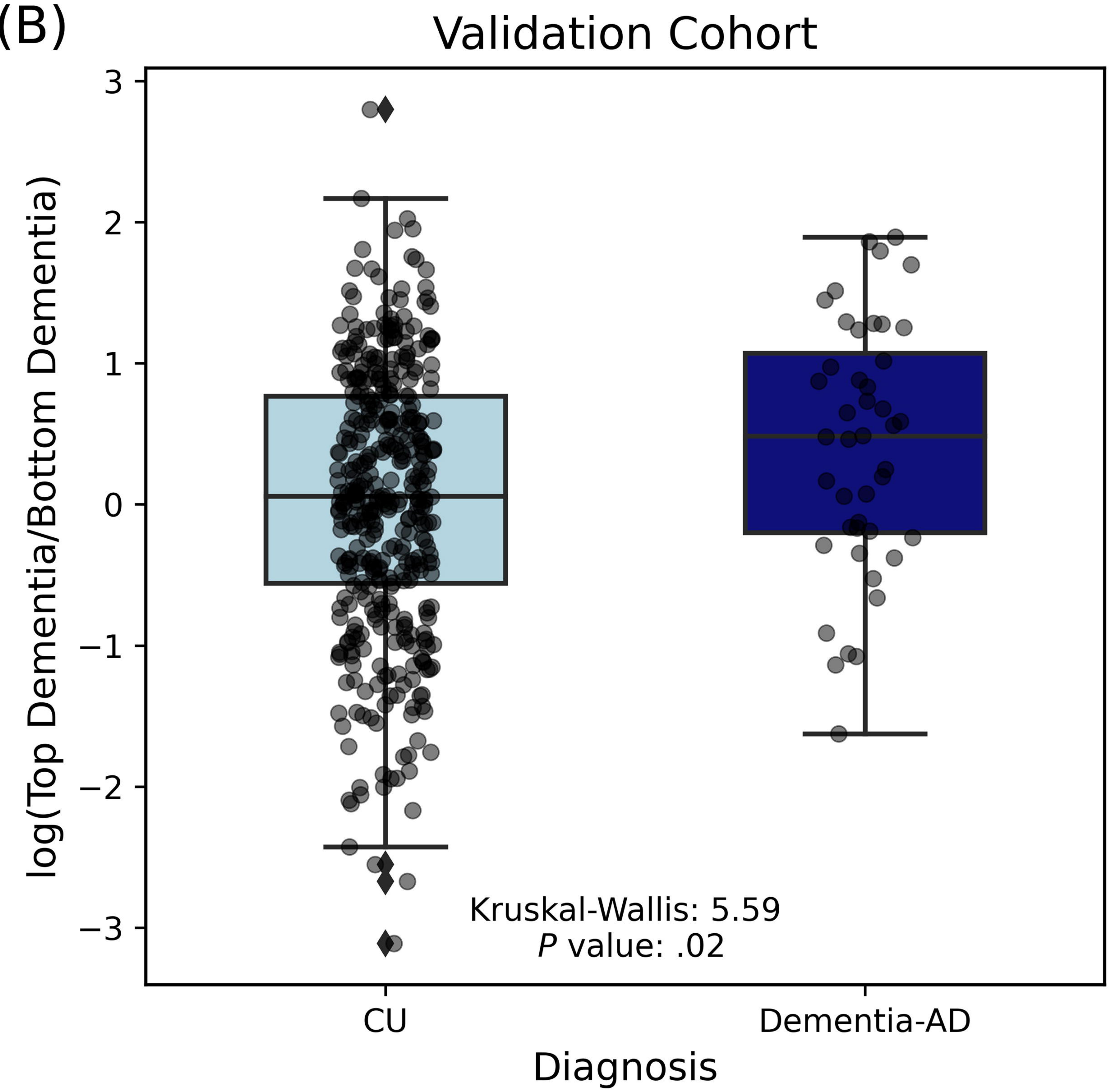
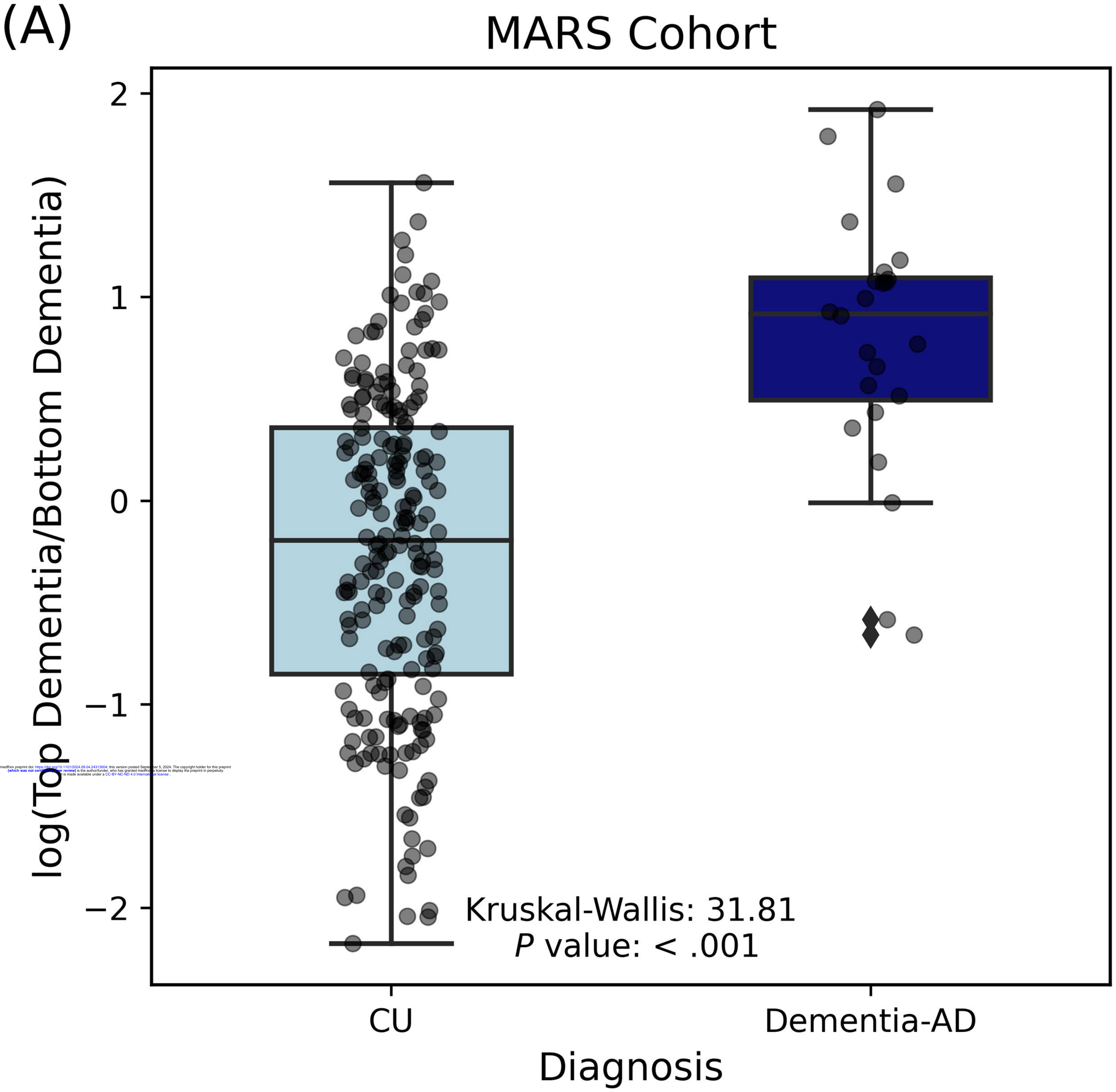


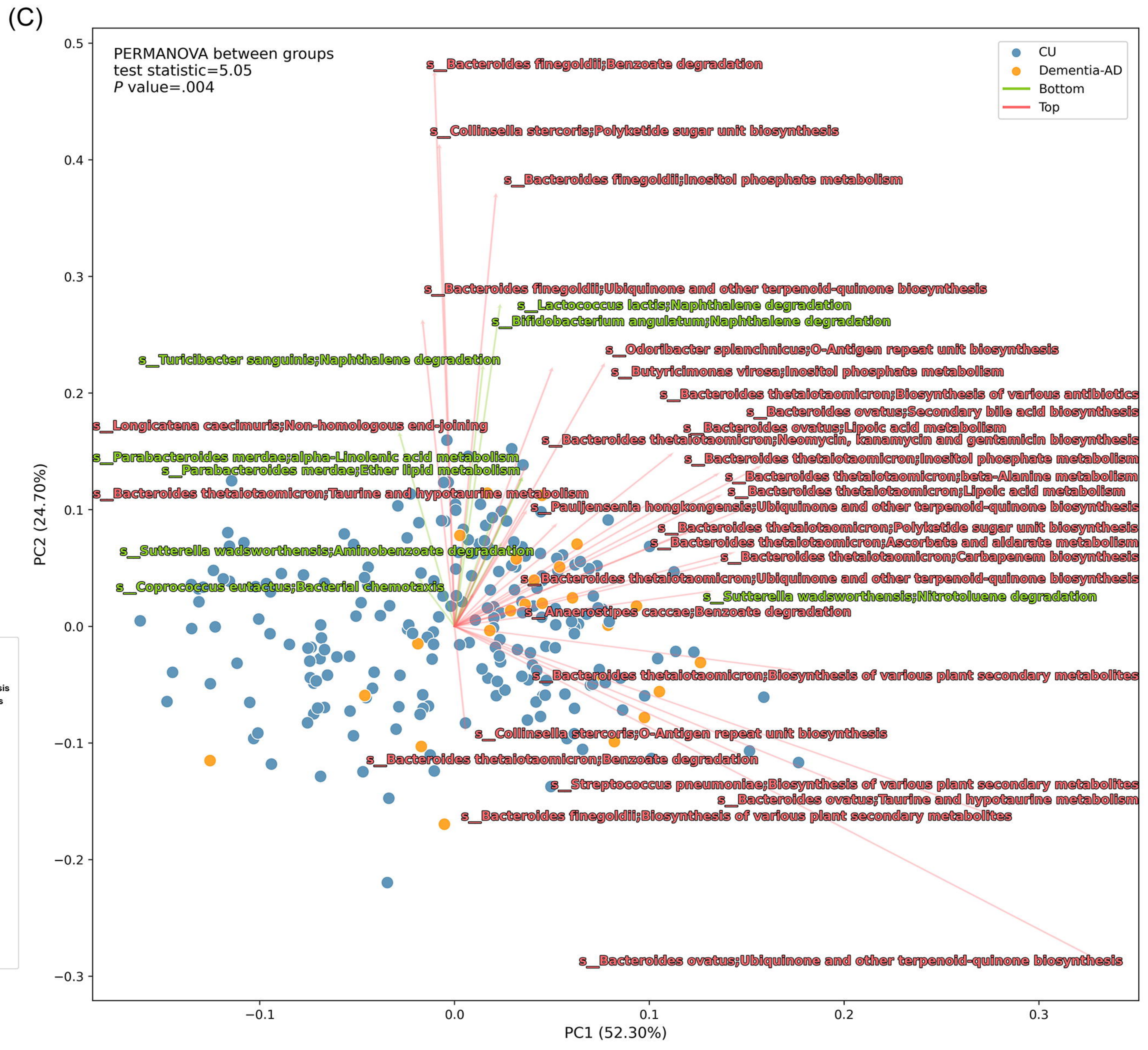
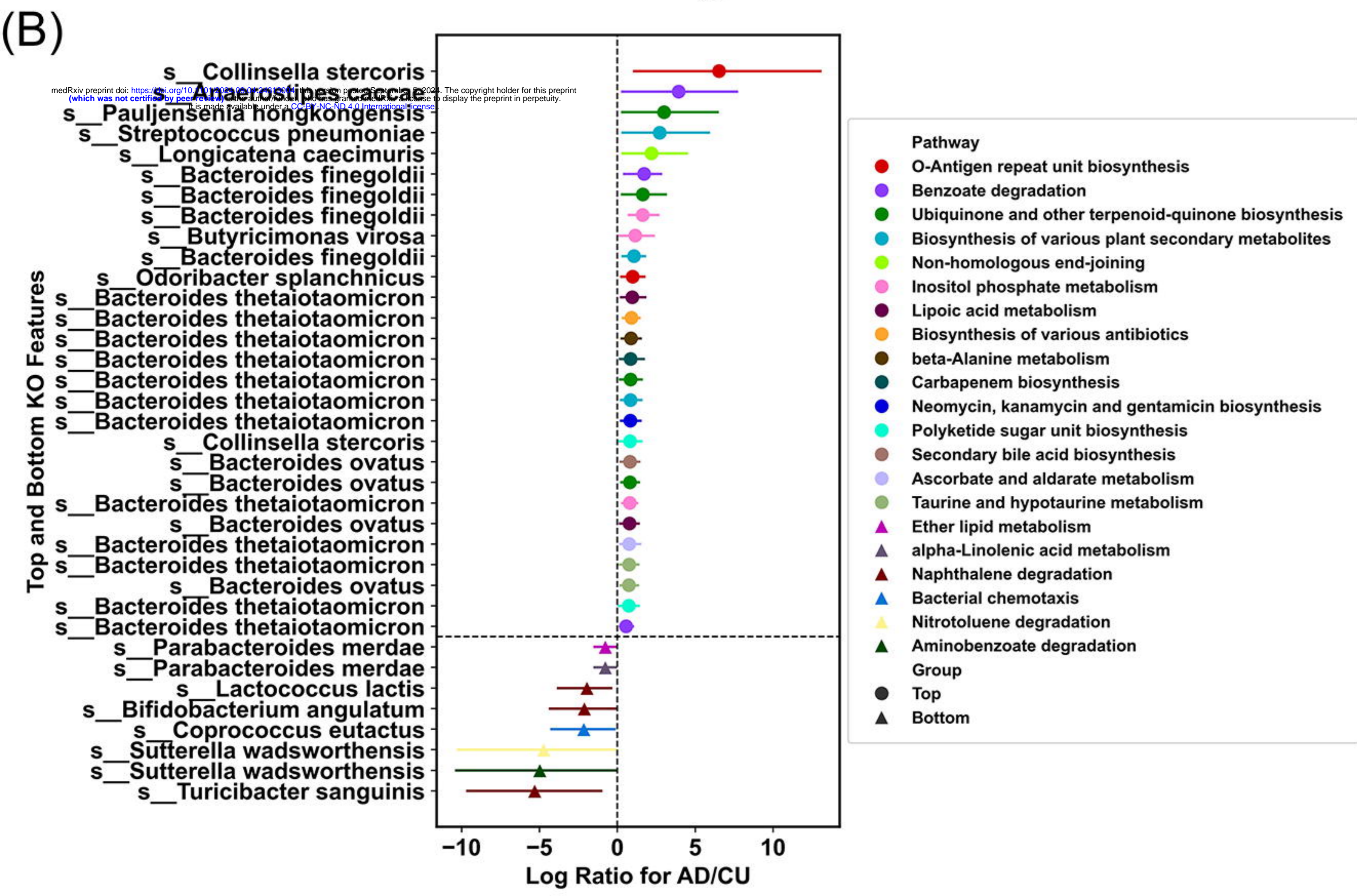
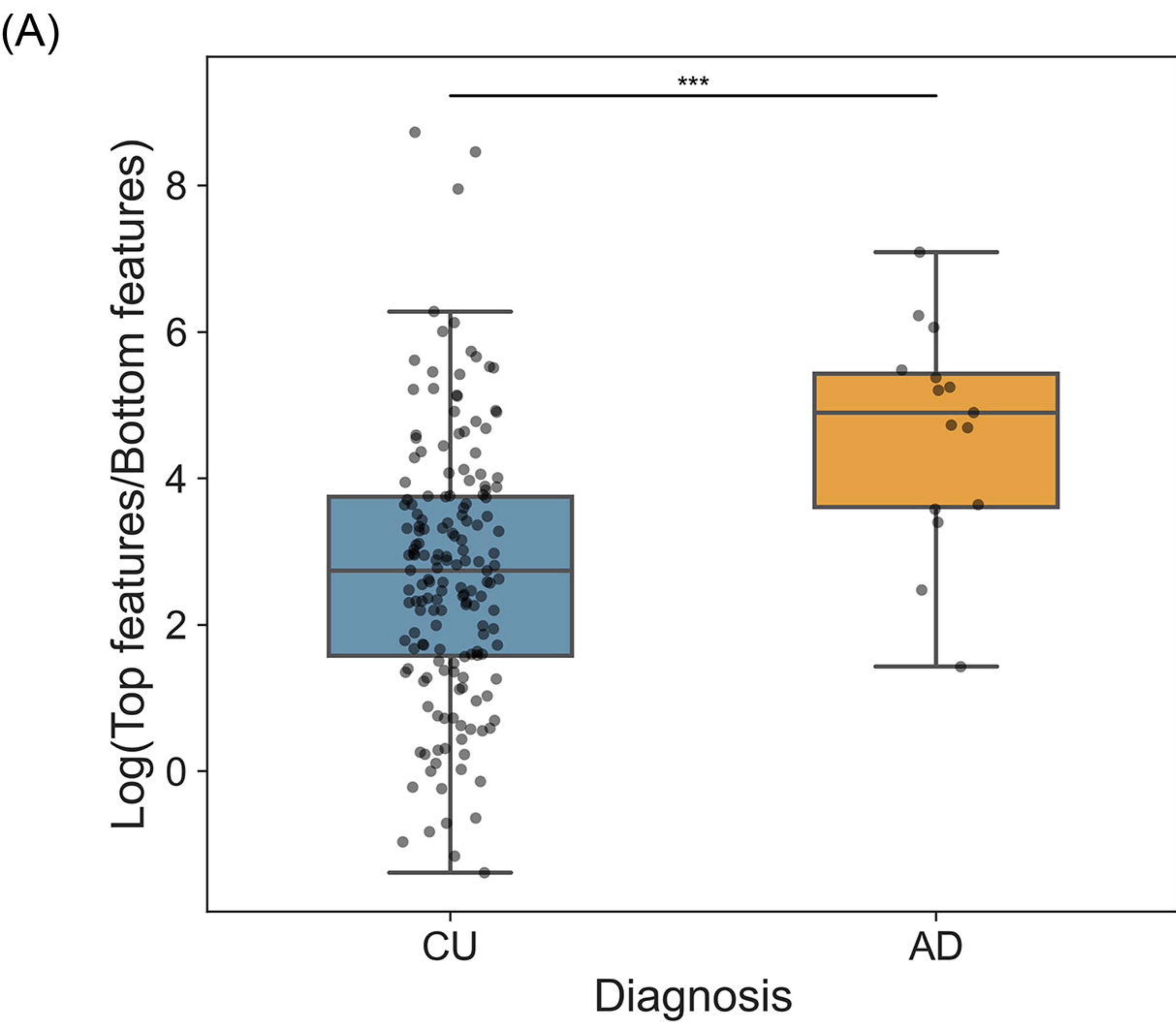
(G)



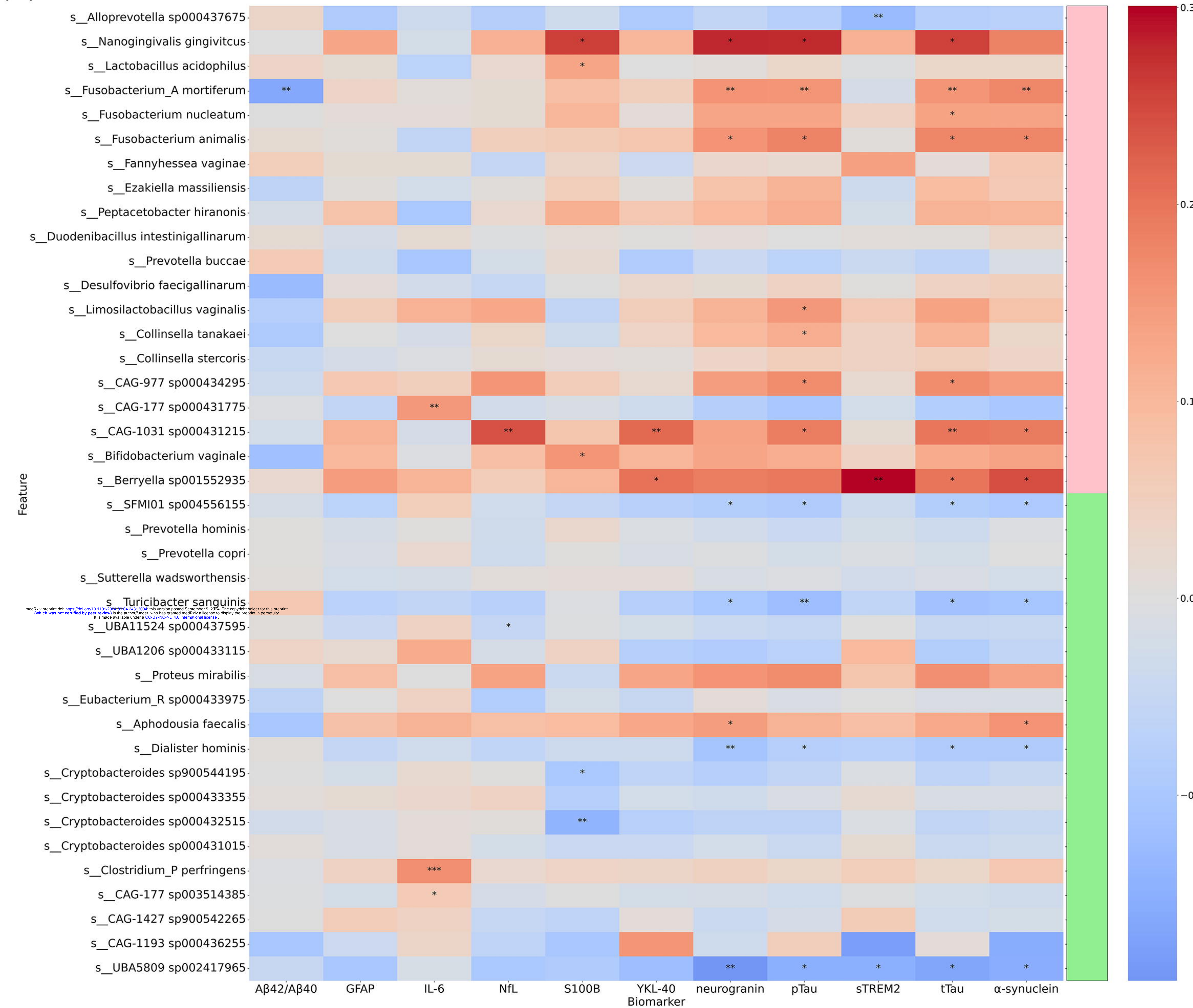
(H)







(A)



(B)

