

1 **SARS-CoV-2 virus in Raw Wastewater from Student Residence Halls with concomitant**  
2 **16S rRNA Bacterial Community Structure changes**

3 Y. Li<sup>1</sup>, K. T. Ash<sup>1,2</sup>, D. C. Joyner<sup>1,2</sup>, D. E. Williams<sup>1</sup>, T. C. Hazen<sup>1,2,3,4,5\*</sup>

4  
5 <sup>1</sup>Department of Civil and Environmental Sciences, University of Tennessee, Knoxville, TN, USA

6 <sup>2</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

7 <sup>3</sup>Department of Microbiology, University of Tennessee, Knoxville, TN, USA

8 <sup>4</sup>Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, TN, USA

9 <sup>5</sup>Institute for a Secure and Sustainable Environment, University of Tennessee, Knoxville, TN, USA

10 **\* Correspondence:**

11 Terry C. Hazen

12 [tchazen@utk.edu](mailto:tchazen@utk.edu)

13 **Keywords:** SARS-CoV-2, COVID-19, Raw Sewage, 16S rRNA, Bacterial Community Structure

14 **Word count:**

15 **Figures: 15 (Not include the Figure S1)**

16 **Tables: 2**

17 **Author Contributions:**

18 Writing: YL, TCH; Collection: DEW, DCJ; Lab analysis: YL, KTA, DEW, DCJ; Data Analysis:

19 YL, KTA, TCH; Project Management: TCH, DCJ

## 20 **Abstract**

21 The detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA in  
22 sewage is well-established, but the concomitant changes in microbial compositions during the  
23 pandemic remain insufficiently explored. This study investigates the impact of the SARS-CoV-2  
24 virus on microbial compositions in raw sewage, utilizing 16S rRNA sequencing to analyze  
25 wastewater samples collected from six dormitories over a one-year field trial at the University of  
26 Tennessee, Knoxville. The concentration of SARS-CoV-2 RNA was assessed using a reverse  
27 transcription-quantitative polymerase chain reaction. Significant variations in bacterial  
28 composition were evident across the six dormitories, highlighting the importance of  
29 independently considering spatial differences when evaluating the raw wastewater microbiome.  
30 Positive samples for SARS-CoV-2 exhibited a prominent representation of exclusive species  
31 across all dormitories, coupled with significantly reduced bacterial diversity compared to  
32 negative samples. The correlation observed between the relative abundance of enteric pathogens  
33 and potential pathogens at sampling sites introduces a significant dimension to our understanding  
34 of COVID-19, especially the notable correlation observed in positive SARS-CoV-2 samples.  
35 Furthermore, the significant correlation in the relative abundance of potential pathogens between  
36 positive and negative SARS-CoV-2 raw sewage samples may be linked to the enduring effects of  
37 microbial dysbiosis observed during COVID-19 recovery. These findings provide valuable  
38 insights into the microbial dynamics in raw sewage during the COVID-19 pandemic.

## 39 **Introduction**

40 The interconnection between sewage and the human gut microbiota has garnered significant  
41 interest, revealing a substantial overlap in microbial composition. Newton et al. (2015) noted that  
42 the microbial composition of sewage, primarily originating from the human gut, comprises a  
43 diverse array of both beneficial and pathogenic species, with bacteria and viruses playing central  
44 roles. Robust evidence consistently supports the notable similarity between the microbial profiles  
45 of raw sewage and the human gut. Cai et al. (2014) emphasized that the total abundance of high-  
46 level genera in influent sewage is nearly 50%, similar to that of the human gut, thus highlighting  
47 the human gut as the primary source of bacterial collection in sewage. Newton et al. (2015)  
48 further reported that sequences representing approximately 78% of a stool sample comprised  
49 around 12% of a sewage sample. Extrapolating this ratio to 100%, their estimation suggests that  
50 only 15% of amplicons in a typical sewage sample originate from human stool. However, Fierer  
51 et al. (2022) found that bacteria derived from fecal material constitute a relatively small fraction  
52 of the taxa in collected samples, underscoring the significance of environmental sources in  
53 shaping the sewage microbiome. It is still unclear if raw sewage truly reflects the microbial  
54 composition of the human gut.

55 Crucially, numerous studies have illustrated that respiratory infections associated with COVID-  
56 19 correlate with changes in the composition of the gut microbiota (Gu et al. 2020, Zuo et al.  
57 2020). The dysbiosis of COVID-19 may enhance gut permeability, leading to secondary  
58 infections and organ failure. Simultaneously, disruptions in gut barrier integrity could potentially  
59 facilitate the translocation of SARS-CoV-2 from the lungs to the intestinal lumen (AKTAŞ and  
60 Aslim 2020). Gu et al. (2020) and Zuo et al. (2020) observed that, compared to fecal samples  
61 from healthy people, fecal samples from COVID-19 patients had significantly reduced bacterial  
62 diversity, a significantly higher relative abundance of opportunistic pathogens and a lower  
63 relative abundance of beneficial symbionts. Liu et al. (2022) even found that gut dysbiosis  
64 persisted even after clearance of SARS-CoV-2 at 6 months. Patients with COVID-19 exhibit  
65 significant alterations in fecal microbiomes, suggesting potential changes in the wastewater  
66 microbiome during the pandemic. Currently, research on microbial compositions in wastewater  
67 with positive and negative SARS-CoV-2 samples remains limited, with Gallardo-Escárate et al.  
68 (2021) being the sole study to explore such dynamics across three sampling communities using

69 nanopore technology. Their findings highlighted a robust association between the microbiota of  
70 positive SARS-CoV-2 wastewater samples and enteric bacteria. Notably, integrating the  
71 Wastewater-Based Epidemiology tool with metagenomic analysis, employing 16S rRNA  
72 sequencing technology to investigate changes in sewage microbiota during the COVID-19  
73 pandemic, remains an unexplored avenue that warrants further research.

74 This study employs 16S rRNA sequencing to thoroughly analyze microbial compositions in raw  
75 sewage samples, differentiating between those with positive and negative COVID-19 status. The  
76 primary goal is to identify distinct patterns or shifts in the bacterial community associated with  
77 the presence of the virus. Through the utilization of this technology, the research aims to provide  
78 a nuanced understanding of the dynamics of viral shedding, microbial interactions, and the  
79 overall impact of SARS-CoV-2 on the sewage microbiome over a year-long field trial conducted  
80 in six campus dormitories. Including COVID-19-negative sewage samples as a control allows for  
81 identifying specific changes attributable to viral presence, facilitating the establishment of  
82 correlations between the sewage microbiota and COVID-19 prevalence in human communities.  
83 Essentially, this investigation seeks to address the knowledge gap regarding the interplay  
84 between SARS-CoV-2 and the sewage microbiome, offering valuable insights into the potential  
85 utility of wastewater-based epidemiology for monitoring and assessing COVID-19 prevalence.

## 86 **Materials and Methods**

### 87 **Raw Sewage Sampling and Sample Processing**

88 Raw wastewater was systematically collected from six student residence halls on the University  
89 of Tennessee, Knoxville campus, as illustrated in Figure 1. Each of these residential dormitories  
90 accommodated a population of over 400 students, and a detailed summary of their characteristics  
91 is presented in Table 1. Sampling was from access points to the main sewage pipe in the  
92 basement of the building or at the first access point to a raw sewer manhole immediately outside  
93 the building, specifically before the convergence or merging with other sewer conduits. This  
94 sampling initiative occurred from September 14, 2020, to October 11, 2021.

95 Grab samples (>50 ml) were collected at the manhole using a stainless-steel telescopic rod pole  
96 swivel dipper water swing sampler. Alternatively, samples were obtained from the valve by  
97 submerging a sterile Nalgene bottle into the flowing sewage. Sampling commenced at 8:00 am,  
98 and all collected samples were promptly transported to the BSL-2 laboratory in a cooler with ice.  
99 The transit time was kept to less than 3 h to ensure immediate processing.

100 Upon reaching the laboratory, sewage samples underwent pasteurization for 2 h at 60°C.  
101 Following pasteurization, centrifugation at 5,000 x g for 10 min occurred, and subsequent  
102 filtration was carried out through sequentially sized 0.45 µm and 0.22 µm nitrocellulose filters.  
103 These filters were individually placed in DNA LoBind tubes and stored at -80°C until DNA  
104 extraction. Concentration was achieved using an Amicon Ultra-15 filtration device, with  
105 centrifugation at either 4,000 x g for 30 min (Swing-arm rotor) or 5,000 x g for 20 min (Fixed-  
106 angle rotor) at room temperature. The resulting concentrated solution, approximately 250 µL,  
107 was carefully transferred to 2 mL DNA LoBind tubes.

108 RNA extraction was performed using the Qiagen viral RNA Mini Kit, following the instructions  
109 of the manufacturer, yielding 60 µL of extracted RNA, with a negative control using  
110 DNase/RNase-free water. Subsequently, the RNA samples were stored at -80°C and subjected to  
111 RT-qPCR analysis within 24 h following extraction (Ash et al. 2023, Li et al. 2023).

## 112 **RT-qPCR**

113 To quantify the concentrations of SARS-CoV-2 and PMMoV RNA in each sample, we  
114 employed RT-qPCR. Specifically, we measured SARS-CoV-2 N1 using the TaqPath 1-Step RT-  
115 qPCR Master Mix, CG (Thermo Fisher Scientific) on an Applied Biosystems QuantStudios 7 Pro  
116 Real-Time PCR System instrument. Each 20  $\mu\text{L}$  reaction mixture comprised 5  $\mu\text{L}$  of 4X Master  
117 Mix (Thermo Fisher Scientific), 0.25  $\mu\text{L}$  of a 10  $\mu\text{mol/L}$  probe, 1  $\mu\text{L}$  each of 10  $\mu\text{mol/L}$  forward  
118 and reverse primers, 7.75  $\mu\text{L}$  of nuclease-free water, and 5  $\mu\text{L}$  of nucleic acid extract. After  
119 accurate pipetting of reagents into 96-well plates, a 10-second vortex mixing step followed. The  
120 RT-qPCR cycling conditions included an initial uracil-DNA glycosylase incubation for 2 min at  
121 25°C, reverse transcription for 15 min at 50°C, activation of the Taq enzyme for 2 min at 95°C,  
122 and a two-step cycling process involving 3 sec at 95°C and 30 sec at 55°C, repeated for a total of  
123 45 cycles. A positive test result was determined by the presence of an exponential fluorescent  
124 curve intersecting the threshold within 40 cycles (cycle threshold [Ct] <40).

125 The quantification of PMMoV was executed using the TaqPath 1-Step RT-qPCR Master Mix,  
126 CG (Thermo Fisher Scientific) on a QuantStudios 7 Pro instrument. Each reaction was composed  
127 of 20  $\mu\text{L}$ , including 5  $\mu\text{L}$  of 4X Master Mix from Thermo Fisher Scientific, 0.5  $\mu\text{L}$  of 10  $\mu\text{mol/L}$   
128 probe, 1.8  $\mu\text{L}$  each of 10  $\mu\text{mol/L}$  forward and reverse primers, 8.9  $\mu\text{L}$  of nuclease-free water, and  
129 2  $\mu\text{L}$  of nucleic acid extract. The reagents were meticulously transferred into 96-well plates using  
130 pipettes and subsequently mixed by vortexing for 10 sec. The thermocycling conditions utilized  
131 in this study were as follows: incubation of uracil-DNA glycosylase for 2 min at 25°C, reverse  
132 transcription carried out for 15 min at 50°C, activation of the Taq enzyme for 10 min at 95°C,  
133 and a two-step cycling process consisting of 30 sec at 95°C followed by 1 min at 60°C, repeated  
134 for a total of 40 cycles.

135 In each RT-qPCR run, one positive PMMoV control and negative controls, comprising  
136 Mastermix and DNase/RNase-free water, were incorporated. The RT-qPCR reactions were  
137 carried out in triplicate, and the criteria for classifying a sample as positive included the  
138 requirement that all replicates produced positive results, with each individual replicate falling  
139 within the linear range of the standard curve. The N1 standard curve demonstrated a high level of  
140 efficiency, with a value of 94.669% ( $R^2 = 1$ ). The quantification of SARS-CoV-2 RNA was  
141 determined by calculating the average of three replicates of viral copies. The outputs of RT-

142 qPCR were converted into units of copies per liter. In this study, the detection limit for SARS-  
143 CoV-2 and PMMoV was established at 20 and 10 copies per liter, respectively.

#### 144 **DNA Isolation, 16S rRNA Gene Amplification, Sequencing**

145 Before inclusion in the kit, quarter-sections of 0.45  $\mu\text{m}$  and 0.22  $\mu\text{m}$  nitrocellulose filters were  
146 prepared by flame-sterilizing a blade and using ethanol for sterilization. Genomic DNA  
147 extraction was then performed using the FastDNA Spin Kit for Soil (BIO101, Vista, CA, USA),  
148 strictly following the guidelines of manufacturer. Subsequent DNA purification utilized the  
149 SELECT-A-SIZE DNA Clean & Concentrator Kits (Zymo Research, Irvine, CA). The quality of  
150 the extracted DNA was assessed by determining the 260/280 and 260/230 ratios on a NanoDrop  
151 spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

152 After confirming successful DNA extraction, Polymerase Chain Reaction (PCR) was conducted  
153 on 1–10  $\mu\text{L}$  of the extracted DNA. DNA libraries were prepared following the methodology  
154 outlined by Caporaso et al. (2012). PCR amplification of the V4 region employed Phusion DNA  
155 polymerase (Master Mix; Thermo Fisher Scientific, Waltham, MA) and universal primers 515f  
156 and barcoded 806r, designed to anneal to both bacterial and archaeal sequences. A 12-bp barcode  
157 index on the reverse primer facilitated multiplexing for sequencing analysis.

158 Subsequently, amplicon quality and size were assessed using an Agilent Bioanalyzer (Agilent  
159 Technologies Santa Clara, CA). Following the protocol of manufacturer, the DNA amplicons  
160 were pooled and quantified with a NEBNext Library Quant Kit for Illumina (New England  
161 Biolabs, Ipswich, MA). Sequencing was performed using a MiSeq V2 kit on an Illumina MiSeq  
162 platform (Illumina, San Diego, CA).

163 Digital sequence data from the MiSeq underwent processing through the QIIME2 (v1.9) pipeline  
164 on a Linux Server (Caporaso et al. 2010). DADA2 within QIIME2 was employed for denoising,  
165 and fast-join facilitated the joining of paired-end sequences. Subsequent demultiplexing  
166 excluded sequences with a Phred score below 20, and UCHIME identified and removed chimeric  
167 sequences. Genus-level identification of sequences utilized the Silva database, with operational  
168 taxonomic units (OTUs) determined and sample populations normalized by total sequence count  
169 to ascertain the relative abundance of each OTU.

170 **Data Analysis**

171 Statistical analyses were conducted using R version 4.2.3. Initially, samples were processed by  
172 rarefying OTU tables to the lowest library size across all samples in each student residence hall.  
173 Subsequently, we computed common  $\alpha$ -diversity metrics (Observed, ACE, Shannon, Simpson,  
174 InvSimpson, Fisher, Coverage, and PD) and  $\beta$ -diversity metric (Bray-Curtis) using the R  
175 phyloseq package. To assess differences in  $\alpha$ -diversity metrics between groups, linear regression  
176 was employed, with semesters included as covariates. For the evaluation of differences in  $\beta$ -  
177 diversity metrics between groups, nonmetric-multidimensional scaling (NMDS) was utilized,  
178 and  $p$  values for the comparison between groups were determined using permutational  
179 multivariate ANOVA models, which included semesters as covariates. A Pearson Correlation  
180 analysis was undertaken to explore the correlations between various parameters, including the  
181 relative abundance of enteric pathogen and potential pathogens, the relative abundance of enteric  
182 pathogen and potential pathogen in positive and negative SARS-CoV-2 samples, as well as the  
183 relative abundance of potential pathogen in positive and negative SARS-CoV-2 sewage samples.



## 184 **Results**

### 185 **Concentration of SARS-CoV-2 in Raw Sewage**

186 The 174 raw sewage samples included in this study were collected from 6 different dormitories  
187 in the same sewage network across the University of Tennessee, Knoxville (Figure 1). Figure 2  
188 depicts the concentrations of SARS-CoV-2 from September 2020 to October 2021 within various  
189 high-density student residence halls. Over the sampling period, SARS-CoV-2 concentrations  
190 were consistently measured at different levels in the respective halls:  $3.09 \pm 3.46$  log<sub>10</sub> copies/L  
191 in D1,  $2.02 \pm 2.19$  log<sub>10</sub> copies/L in D2,  $2.80 \pm 3.26$  log<sub>10</sub> copies/L in D3,  $2.97 \pm 3.61$  log<sub>10</sub>  
192 copies/L in D4,  $2.94 \pm 3.30$  log<sub>10</sub> copies/L in D5, and  $2.36 \pm 2.71$  log<sub>10</sub> copies/L in D6.

193 Furthermore, the positive rates, calculated by dividing the number of positive samples by the  
194 total number of samples and multiplying by 100%, varied across the halls. Specifically, the  
195 positive rates were 70% in D1, 39% in D2, 52% in D3, 20% in D4, 68% in D5, and 37% in D6  
196 (Table 2). These results provide valuable insights into the dynamics of SARS-CoV-2  
197 concentrations and positivity rates within high-density student residence halls during the  
198 specified timeframe.

### 199 **Characteristics of the predominant flora in different dormitories.**

#### 200 **Characterization on phylum, family, and genus level.**

201 The sequences extracted from the samples underwent comprehensive analysis, resulting in the  
202 classification of data into 56 phyla, 145 classes, 315 orders, 548 families, and 1170 genera.  
203 Figure 3 illustrates the relative abundances of the top 10 phyla across six dormitories, revealing  
204 notable variations. The phylum Bacteroidetes was identified as the most abundant across all  
205 sampling sites, with relative abundance ranging from 46.1% to 26.9%. Firmicutes emerged as the  
206 second most abundant phylum in dormitories 1, 2, 3, and 4, with relative abundance varying  
207 from 35.2% to 14.8%. Meanwhile, Proteobacteria emerged as the second most abundant phylum  
208 in dormitories 5 and 6, with relative abundance ranging from 41.3% to 25.4%. Additionally, it  
209 was observed that dormitories 2 and 3 have the same four highest abundance phyla of  
210 Bacteroidetes, Firmicutes, Proteobacteria, and Spirochaetota. Similarly, Dormitories 5 and 6  
211 have the same four highest abundance phyla of Bacteroidetes, Proteobacteria, Firmicutes, and  
212 Campilobacterota.

213 Figure 4 highlights the 10 most dominant families, showcasing significant differences among the  
214 six dormitories. Paludibacteraceae and Spirochaetaceae emerged as the most dominant families,  
215 with varying relative abundance in D1, D2, and D3. Interestingly, Bacteroidaceae took  
216 precedence in D4, D5, and D6, with distinct relative abundance up to 12.6%.

217 The relative abundances of the top 50 genera detected in all dormitory samples are shown in  
218 Figure 5. Moreover, many typical gut bacteria were also found at very high levels in the sewage  
219 such as Bacteroides, Acinetobacter, Prevotella, Pseudomonas, Blautia, Faecalibacterium,  
220 Ruminococcus, Dorea, and others, corresponding to ranks 1, 8, 10, 11, and 21 within the top 50  
221 genera in Figure 5 (Furet et al. 2009, Cai et al. 2014, Bäckhed et al. 2015, Do et al. 2019).  
222 Among the top 50 genera, 14 genera (33.52%) were identified as potential pathogens, including  
223 Bacteroides, Arcobacter, Treponema, Aeromonas, Acinetobacter, Prevotella, Pseudomonas,  
224 Erysipelothrix, Faecalibacterium, Flavobacterium, Ruminococcus, Bifidobacterium, Laribacter,  
225 and Streptococcus (Cai and Zhang 2013, Cai et al. 2014, Do et al. 2019, Oluseyi Osunmakinde et  
226 al. 2019, Poopedi et al. 2023).

227 In the examination of the top 50 genera, dormitories D1 to D6 exhibited varying relative  
228 abundance of potential pathogens, with 13 (40%), 12 (23%), 11 (18%), 12 (28%), 14 (45%), and  
229 14 (41%) genera recognized as such, respectively (Figure 5). Notably, a substantial number of  
230 these potential pathogens displayed an increased relative abundance in samples from D1, D5,  
231 and D6 compared to other sites. Among the detected enteric pathogens in the top 50 genera were  
232 *Arcobacter*, *Aeromonas*, and *Laribacter*, with total relative abundances of 13.61%, 2.37%,  
233 5.45%, 5.64%, 20.82%, and 9.01% from D1 to D6, respectively. *Arcobacter* and *Aeromonas*  
234 were identified across all six dormitories, while *Laribacter* was exclusively found in D3, D5, and  
235 D6. Furthermore, an observation revealed a correlation between the relative abundance of enteric  
236 pathogen and potential pathogens (Pearson Correlation = 0.842,  $p = 0.018$ ). Additionally,  
237 *Mycobacterium*, the most prevalent respiratory tract-associated pathogen, contributed from  
238 0.02% to 0.15% of the total bacterial community across all six dormitories, respectively.

### 239 **Diversity of bacterial communities**

240 The analysis of the microbiota communities within the collected wastewater samples revealed  
241 significant distinctions across all sampled locations (Figure 6). At the species level, dormitory 6

242 (D6) exhibited the highest count of exclusive taxa, totaling 1206, while the other dormitories (D1  
243 to D5) displayed varying counts of exclusive taxa, ranging from 546 to 1081 species. It is  
244 noteworthy that a core microbiome consisting of 286 bacterial species was consistently observed  
245 across all sampled dormitories.

246 Considerable distinctions were identified in microbiota communities within the collected  
247 wastewater from all sampled locations, as illustrated in Figure 6. Notably, at the species level,  
248 D6 demonstrated the highest count of exclusive taxa, totaling 1206. Conversely, the other  
249 dormitories (D1 to D5) exhibited diverse counts of exclusive taxa, with 1081, 764, 856, 546, and  
250 755 species, respectively. Notably, a core microbiome was observed with 286 bacterial species.

251 Alpha diversity analysis was employed to assess the diversity and richness of bacterial  
252 communities within the microbiome of six dormitories (Figure 7). A comprehensive comparison  
253 among the six dormitories was conducted using linear regression models, with semester serving  
254 as a covariate. The results showed statistically significant differences in bacterial diversity across  
255 the dormitories.

256 Figure 8 depicts the clustering of beta diversity, assessed through Bray-Curtis distance metrics,  
257 among the six dormitories. A permutational multivariate ANOVA model, which included  
258 semester as a covariate, demonstrated significant differences in the measured  $\beta$ -diversity metrics  
259 between groups ( $p < 0.05$ ).

## 260 **Characteristics of the predominant flora in positive and negative samples**

### 261 **Characterization on phylum, family, and genus level**

262 The microbial composition of various sampling sites was analyzed to determine the abundance  
263 of specific phyla and family, with a focus on the impact of COVID-19 status on the results.  
264 These results showed that the microbial composition of different dormitory locations remains  
265 consistent at the top 10 dominant phyla and family, regardless of the COVID-19 status (Figure 9,  
266 10).

267 The current study, encompassing multiple dormitories, unveiled several noteworthy distinctions  
268 in the relative abundances of microbial families (Figure 11). Dormitory 1 showed a distinction  
269 between the families Lachnospiraceae and Streptococcaceae. Similarly, dormitory 3 exhibited  
270 significant differences in the relative abundances of Arcobacteraceae, Peptostreptococcaceae,

271 Rhodocyclaceae, and V2072-189E03 between COVID-19 positive and negative samples. In  
272 dormitory 4, Peptostreptococcales-Tissierellales displayed significant variation in relative  
273 abundances between the two sample groups. In dormitory 5, Desulfovibrionaceae demonstrated a  
274 significant difference in relative abundances based on COVID-19 status. Lastly, in dormitory 6,  
275 significant differences were observed in the relative abundances of Aeromonadaceae and  
276 Paludibacteraceae between COVID-19 positive and negative samples. These findings underscore  
277 the potential impact of COVID-19 on specific microbial families within the microbiota of  
278 different dormitories, providing valuable insights into the nuanced variations in microbial  
279 composition associated with the viral infection.

280 The LEfSe analysis was employed to discern and differentiate the microbiome composition  
281 between samples that tested positive and negative for COVID-19 across the six dormitories  
282 (Figure S1). Interestingly, no universal biomarkers were identified across all six dormitories.  
283 Instead, distinct biomarkers were exclusively found in dormitories 3, 4, and 5, suggesting unique  
284 microbial signatures associated with COVID-19 status in these specific dormitory environments.

285 The present study aimed to analyze the relative abundance of potential pathogen in both SARS-  
286 CoV-2 positive and negative samples of the top 50 genera (Figure 12). As depicted in Table 2,  
287 the results revealed a noteworthy correlation between the relative abundance of potential  
288 pathogen in positive and negative samples, with a Pearson Correlation coefficient of 0.918 ( $p =$   
289 0.010). Additionally, the study found a significant correlation between the relative abundance of  
290 enteric pathogen and potential pathogen in positive SARS-CoV-2 samples (Pearson Correlation  
291 = 0.817,  $p = 0.024$ ), irrespective of relative abundance of potential pathogen in negative SARS-  
292 CoV-2 samples.

### 293 **Diversity of bacterial communities**

294 Our study conducted a comprehensive analysis of samples from different locations, with a  
295 specific focus on viral quantification to distinguish the differences between exclusive and shared  
296 species in wastewaters tested for SARS-CoV-2. We found that samples testing positive for  
297 SARS-CoV-2 demonstrated a higher diversity of taxa compared to their negative counterparts  
298 (Figure 13). The analysis highlighted that exclusive species were most prominently represented  
299 in positive samples for SARS-CoV-2 collected from D1 at 31.87%, while D3 exhibited the  
300 lowest representation at 18.99%. Conversely, negative samples for SARS-CoV-2 were

301 associated with exclusive bacterial species in wastewater collected from D3 (18.24%), with D1  
302 displaying the lowest representation at 8.49%. Despite the SARS-CoV-2 status, the analysis  
303 further indicated a low representativity for exclusive bacteria found in other dormitories.

304 An observation revealed a correlation between the positive rate of sampling sites and the relative  
305 abundance of exclusive species in positive samples (Pearson Correlation = 0.771,  $p = 0.036$ ).  
306 Additionally, the presence of 1033, 914, 954, 671, 1071, and 917 taxa in both SARS-CoV-2  
307 positive and negative samples from D1 to D6, respectively. These findings collectively  
308 contribute to our understanding of the microbial dynamics associated with SARS-CoV-2 in  
309 wastewater samples across different dormitory locations.

310 The  $\alpha$ -diversity of the microbiome across all locations exhibited a general trend of being higher  
311 in negative samples compared to positive samples. Specifically, the observed species index  
312 showed a significant difference in D5 and D6 ( $p < 0.05$ , Figure 14), as determined through linear  
313 regression models that incorporated semester as a covariate. Notably, significant differences in  
314 the measured  $\beta$ -diversity metrics were discerned in D3 and D6 between groups ( $p < 0.05$  for the  
315 Bray-Curtis indices, using permutational multivariate ANOVA with semester as a covariate), as  
316 illustrated in Figure 15.

## 317 **Discussion**

318 The identified variations at the phylum, family, and genus levels across the six dormitories shed  
319 light on the geographic differences in bacterial composition in this study (Figure 3, 4, and 5).

320 The analysis revealed two clusters of community types, as illustrated in Figure 3. Dormitories 2,  
321 3, 1, and 4, organized by the closer relationship of bacterial phyla in each building, exhibited  
322 similar dominance patterns in these phyla, while D5 and D6 exhibited comparable compositions.  
323 The spatial arrangement depicted in the map (Figure 1) highlights that D1, D2, and D4 are in  
324 proximity, D5 and D6 are likewise nearby, and D3 is closer to D5 and D6. This spatial variation  
325 suggests a potential impact of geographic factors on the microbial composition in different  
326 dormitories. This observation aligns with the study by Fierer et al. (2022) finding five clusters of  
327 17 different locations, revealing no strong relationship with the distance between sampling  
328 locations.

329 The significant alpha and beta diversity further underscore pronounced geographical variations  
330 in microbial communities in this study, aligning with Fierer et al. (2022) findings. Their  
331 emphasis on independently considering spatial variations when assessing the wastewater  
332 microbiome highlights the need to account for the influence of location on microbial diversity.  
333 Their research identified geographic variations in bacterial composition unrelated to sewer  
334 material, sewer depth, or resident human population on the campus. They attributed these  
335 variations to sample pH, with total suspended solids concentrations and sample volume playing a  
336 lesser role. This pH correlation aligns with studies by Fujii et al. (2012) and Lindström et al.  
337 (2005), which demonstrated the close association between pH and shifts in bacterial community  
338 composition in aquatic environments. Despite the detected variations in bacterial composition  
339 across dormitories in our study, the pH did not exhibit significant changes. Future research could  
340 explore specific factors such as organic carbon or nutrient concentrations to better understand the  
341 observed geographic variations in microbial communities.

342 The analysis of the microbial community in raw sewage yielded results consistent with previous  
343 research, indicating the influence of the human gut bacterial community on the bacterial profile  
344 in raw sewage. Specifically, the phyla Bacteroidota was identified as the most abundant and  
345 variable across samples, aligning with findings from Arumugam et al. (2011). However, a study  
346 by Cai et al. (2014) reported Firmicutes as the most dominant phylum in influent samples,

347 asserting its alignment with the human microbiome composition. The findings of may clarify this  
348 discrepancy Turnbaugh et al. (2006) and Clemente et al. (2012), indicating that the gut  
349 microbiota typically showcases dominance of bacteria, particularly from the Bacteroidota and  
350 Firmicutes divisions. Furthermore, Huttenhower et al. (2012) revealed that gut microbiota  
351 relationships were characterized by inverse associations with Bacteroidota, varying from  
352 dominance in some subjects to a minority in others with a greater diversity of Firmicutes. These  
353 nuanced observations highlight the intricate dynamics of the human gut microbiota and  
354 underscore the pivotal roles played by Bacteroidota and Firmicutes in shaping microbial profiles  
355 observed in raw sewage.

356 The primary objective of our study was to investigate potential changes in wastewater  
357 microbiomes during the pandemic, considering the influence of the human gut bacterial  
358 community on the bacterial profile in raw sewage. This investigation was prompted by existing  
359 literature highlighting significant alterations in fecal microbiomes among individuals with  
360 COVID-19 (Gu et al. 2020, Zuo et al. 2020, Yeoh et al. 2021). The significant differences in  
361 bacterial composition observed across the six dormitories prompted a recommendation for  
362 separate analyses of the 16S rRNA data for each dormitory. This approach aims to mitigate  
363 biases that may arise when combining data from diverse dormitory settings. The prominent  
364 representation of exclusive species in positive samples for SARS-CoV-2 were found across all  
365 six dormitories supports the findings of Gallardo-Escárate et al. (2021). Moreover, the observed  
366 trend of higher  $\alpha$ -diversity in the microbiome of negative samples compared to positive samples  
367 across some locations echoes the results reported by Gu et al. (2020) and Yeoh et al. (2021), who  
368 documented a significant decrease in gut microbiota diversity and abundance in COVID-19  
369 patients relative to healthy individuals.

370 The observed correlation between the relative abundance of enteric pathogen and potential  
371 pathogens at sampling sites adds a significant layer of understanding in the context of COVID-  
372 19, particularly highlighting the notable association between the relative abundance of enteric  
373 pathogen and potential pathogen in positive SARS-CoV-2 samples. The presence of three enteric  
374 genera, namely, *Arcobacter*, *Aeromonas*, and *Laribacter*, in our study, commonly residing in the  
375 human intestines and potentially utilizing pathogenic mechanisms to induce gastrointestinal tract  
376 infections, emphasizes the relevance of these microbes in the sewage context during the  
377 pandemic. Notably, the *Aeromonas* genus ranked as the third leading cause of diarrhea after

378 *Campylobacter* and *Salmonella* [OB], exhibited a notably high abundance exclusively in D5  
379 (9.09%) and D6 (6.14%) compared to other dormitories, where the abundance ranged from  
380 0.66% to 1.14%. Additionally, two *Arcobacter* species, *A. butzleri*, and *A. cryaerophilus*, are  
381 considered emerging pathogens posing threats to human health, adding depth to discussing  
382 potential pathogenic risks in the sewage microbiome. Additionally, the genus *Laribacter*,  
383 represented by the species *L. hongkongensis*, known for its associations with traveler  
384 gastroenteritis and diarrhea (Beilfuss et al. 2015), further contributes to understanding the  
385 microbial landscape in the context of COVID-19.

386 In the context of the ongoing discourse surrounding COVID-19, an emerging respiratory  
387 infectious disease, the investigation into the presence of *Mycobacterium*, a medically significant  
388 respiratory tract-associated pathogen, within sewage systems has garnered attention. Notably,  
389 this scrutiny extends across six dormitories, revealing a discernibly lower total abundance of  
390 *Mycobacterium* in sewage than the prevalent genera identified in the samples. The quantification  
391 of *Mycobacterium* in our samples aligns with findings from previous studies, providing a basis  
392 for comparative analysis. Cai and Zhang (2013) reported an overall abundance of  
393 *Mycobacterium* in influent and effluent samples that remained below the threshold of 0.02%.  
394 Numberger et al. (2019) the genus *Mycobacterium* was observed exclusively in October effluent  
395 samples with a relative abundance of less than 0.02%. The 16S rRNA gene sequences analysis in  
396 our work determined the presence of the bacterial genera but not species. These genera may  
397 contain both pathogenic and non-pathogenic species. Therefore, the identification of pathogens  
398 requires further study.

399 Our study did not unveil a significant universal biomarker distinguishing positive from negative  
400 SARS-CoV-2 sewage samples across all sampling locations. This contrasts with the findings of  
401 Gu et al. (2020), who identified five biomarkers to differentiate between COVID-19 patients and  
402 healthy individuals. It is essential to note that the absence of SARS-CoV-2 detection in certain  
403 patients may not necessarily signify a complete recovery of their gut microbiota. The restoration  
404 of microbial communities may require an extended period, even when SARS-CoV-2 is not  
405 detectable. This aligns with the observations of Zhang et al. (2023), who documented persistent  
406 dysbiosis for months after the clearance of the virus. Individuals recovered from COVID-19,  
407 when compared to healthy controls, exhibited reduced bacterial diversity and richness at 3  
408 months. This reduction was accompanied by a lower abundance of beneficial commensals and a



409 higher abundance of opportunistic pathogens. Hence, the significant correlation in the relative  
410 abundance of potential pathogen between positive and negative SARS-CoV-2 sewage samples in  
411 our study may be attributed to the lingering effects of microbial dysbiosis observed in COVID-  
412 19 recovery.

## 413 **Conclusion**

414 In conclusion, our study provides valuable insights into the raw sewage microbiota as a  
415 reflection of the gut microbiota during the COVID-19 pandemic and its potential association  
416 with fecal SARS-CoV-2 shedding. The observed significant differences in raw sewage microbial  
417 communities across all sampling sites and the prominent representation of exclusive species in  
418 positive samples for SARS-CoV-2 emphasize the potential of sewage microbiota as an indicator  
419 of viral shedding. Positive samples for SARS-CoV-2 exhibited a significant reduction in  
420 bacterial diversity, highlighting the impact of the virus infection on microbial composition.  
421 These findings introduce a novel and targeted approach for modulating sewage microbiota,  
422 specifically linked to gastrointestinal manifestations, as a strategy for monitoring and predicting  
423 the presence of SARS-CoV-2 in raw sewage.

424 While our analysis did not uncover a significant universal biomarker distinguishing positive and  
425 negative SARS-CoV-2 raw sewage samples, the observed noteworthy correlation in the relative  
426 abundance of potential pathogens between these samples suggests a potential connection to the  
427 enduring effects of microbial dysbiosis during the recovery phase of COVID-19. Moreover, the  
428 identified correlation between the relative abundance of enteric pathogens and potential  
429 pathogens at sampling sites adds a significant dimension to our understanding of COVID-19,  
430 particularly in the context of the substantial correlation in positive SARS-CoV-2 samples. These  
431 findings underscore the importance of monitoring enteric pathogens in raw wastewater  
432 surveillance systems to comprehend the potential spread of COVID-19 and other infectious  
433 diseases. Such insights carry crucial implications for public health monitoring and management  
434 strategies.

435 **Acknowledgment**

436 The authors would like to thank the Office of Research & Engagement of the University of  
437 Tennessee, Knoxville, for providing funding for the project.

## 438 **References**

- 439 AKTAŞ, B. and Aslim, B. (2020). "Gut-lung axis and dysbiosis in COVID-19." *Turkish journal*  
440 *of biology* **44**(7): 265-272 DOI: <https://doi.org/10.3906/biy-2005-102>.
- 441 Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G.  
442 R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L.,  
443 Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K.,  
444 Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H. B., Nielsen, T., Pons, N., Poulain,  
445 J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E. G.,  
446 Wang, J., Guarner, F., Pedersen, O., de Vos, W. M., Brunak, S., Doré, J., Antolín, M.,  
447 Artiguenave, F., Blottiere, H. M., Almeida, M., Brechot, C., Cara, C., Chervaux, C.,  
448 Cultrone, A., Delorme, C., Denariáz, G., Dervyn, R., Foerstner, K. U., Friss, C., van de  
449 Guchte, M., Guedon, E., Haimet, F., Huber, W., van Hylckama-Vlieg, J., Jamet, A.,  
450 Juste, C., Kaci, G., Knol, J., Kristiansen, K., Lakhdari, O., Layec, S., Le Roux, K.,  
451 Maguin, E., Mériéux, A., Melo Minardi, R., M'Rini, C., Muller, J., Oozeer, R., Parkhill,  
452 J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K.,  
453 Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J.,  
454 Ehrlich, S. D., Bork, P. and Meta, H. I. T. C. (2011). "Enterotypes of the human gut  
455 microbiome." *Nature* **473**(7346): 174-180 DOI: <https://10.1038/nature09944>.
- 456 Ash, K. T., Li, Y., Alamilla, I., Joyner, D. C., Williams, D. E., McKay, P. J., Green, B. M., Iler,  
457 C., DeBlander, S. E., North, C. M., Kara-Murdoch, F., Swift, C. M. and Hazen, T. C.  
458 (2023). "SARS-CoV-2 raw wastewater surveillance from student residences on an urban  
459 university campus." *Frontiers in Microbiology* **14** DOI:  
460 <https://doi.org/10.3389/fmicb.2023.1101205>.
- 461 Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y.,  
462 Xie, H. and Zhong, H. (2015). "Dynamics and stabilization of the human gut microbiome  
463 during the first year of life." *Cell host & microbe* **17**(5): 690-703 DOI: [https://doi-](https://doi-org.utk.idm.oclc.org/10.1016/j.chom.2015.04.004)  
464 [org.utk.idm.oclc.org/10.1016/j.chom.2015.04.004](https://doi-org.utk.idm.oclc.org/10.1016/j.chom.2015.04.004).
- 465 Beilfuss, H. A., Quig, D., Block, M. A. and Schreckenberger, P. C. (2015). "Definitive  
466 identification of *Laribacter hongkongensis* acquired in the United States." *Journal of*  
467 *Clinical Microbiology* **53**(7): 2385-2388 DOI: [https://doi-](https://doi-org.utk.idm.oclc.org/10.1128/jcm.00539-15)  
468 [org.utk.idm.oclc.org/10.1128/jcm.00539-15](https://doi-org.utk.idm.oclc.org/10.1128/jcm.00539-15).
- 469 Cai, L., Ju, F. and Zhang, T. (2014). "Tracking human sewage microbiome in a municipal  
470 wastewater treatment plant." *Applied microbiology and biotechnology* **98**: 3317-3326  
471 DOI: <https://doi.org/10.1007/s00253-013-5402-z>.
- 472 Cai, L. and Zhang, T. (2013). "Detecting Human Bacterial Pathogens in Wastewater Treatment  
473 Plants by a High-Throughput Shotgun Sequencing Technique." *Environmental Science &*  
474 *Technology* **47**(10): 5433-5441 DOI: <https://doi.org/10.1021/es400275r>.
- 475 Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K.,  
476 Fierer, N., Peña, A. G., Goodrich, J. K. and Gordon, J. I. (2010). "QIIME allows analysis  
477 of high-throughput community sequencing data." *Nature methods* **7**(5): 335-336 DOI:  
478 <https://doi-org.utk.idm.oclc.org/10.1038/nmeth.f.303>.

- 479 Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens,  
480 S. M., Betley, J., Fraser, L. and Bauer, M. (2012). "Ultra-high-throughput microbial  
481 community analysis on the Illumina HiSeq and MiSeq platforms." *The ISME journal*  
482 **6**(8): 1621-1624 DOI: <https://doi-org.utk.idm.oclc.org/10.1038/ismej.2012.8>.
- 483 Clemente, J. C., Ursell, L. K., Parfrey, L. W. and Knight, R. (2012). "The impact of the gut  
484 microbiota on human health: an integrative view." *Cell* **148**(6): 1258-1270 DOI:  
485 <https://doi-org.utk.idm.oclc.org/10.1016/j.cell.2012.01.035>.
- 486 Do, T. T., Delaney, S. and Walsh, F. (2019). "16S rRNA gene based bacterial community  
487 structure of wastewater treatment plant effluents." *FEMS microbiology letters* **366**(3)  
488 DOI: <https://10.1093/femsle/fnz017>.
- 489 Fierer, N., Holland-Moritz, H., Alexiev, A., Batther, H., Dragone, N. B., Friar, L., Gebert, M. J.,  
490 Gering, S., Henley, J. B. and Jech, S. (2022). "A metagenomic investigation of spatial  
491 and temporal changes in sewage microbiomes across a university campus." *Msystems*  
492 **7**(5): e00651-00622 DOI: <https://doi-org.utk.idm.oclc.org/10.1128/msystems.00651-22>.
- 493 Fujii, M., Kojima, H., Iwata, T., Urabe, J. and Fukui, M. (2012). "Dissolved Organic Carbon as  
494 Major Environmental Factor Affecting Bacterioplankton Communities in Mountain  
495 Lakes of Eastern Japan." *Microbial ecology* **63**(3): 496-508 DOI:  
496 <https://doi.org/10.1007/s00248-011-9983-8>.
- 497 Furet, J.-P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., Doré, J. and  
498 Corthier, G. (2009). "Comparative assessment of human and farm animal faecal  
499 microbiota using real-time quantitative PCR." *FEMS microbiology ecology* **68**(3): 351-  
500 362 DOI: <https://doi.org/10.1111/j.1574-6941.2009.00671.x>.
- 501 Gallardo-Escárate, C., Valenzuela-Muñoz, V., Núñez-Acuña, G., Valenzuela-Miranda, D.,  
502 Benaventel, B. P., Sáez-Vera, C., Urrutia, H., Novoa, B., Figueras, A., Roberts, S.,  
503 Assmann, P. and Bravo, M. (2021). "The wastewater microbiome: A novel insight for  
504 COVID-19 surveillance." *Science of the Total Environment* **764**: 142867 DOI:  
505 <https://doi.org/10.1016/j.scitotenv.2020.142867>.
- 506 Gu, S., Chen, Y., Wu, Z., Chen, Y., Gao, H., Lv, L., Guo, F., Zhang, X., Luo, R. and Huang, C.  
507 (2020). "Alterations of the gut microbiota in patients with coronavirus disease 2019 or  
508 H1N1 influenza." *Clinical Infectious Diseases* **71**(10): 2669-2678 DOI: [https://doi-  
509 org.utk.idm.oclc.org/10.1093/cid/ciaa709](https://doi-org.utk.idm.oclc.org/10.1093/cid/ciaa709).
- 510 Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., Creasy,  
511 H. H., Earl, A. M., FitzGerald, M. G., Fulton, R. S., Giglio, M. G., Hallsworth-Pepin, K.,  
512 Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., Mitreva, M., Muzny, D. M.,  
513 Sodergren, E. J., Versalovic, J., Wollam, A. M., Worley, K. C., Wortman, J. R., Young,  
514 S. K., Zeng, Q., Aagaard, K. M., Abolude, O. O., Allen-Vercoe, E., Alm, E. J., Alvarado,  
515 L., Andersen, G. L., Anderson, S., Appelbaum, E., Arachchi, H. M., Armitage, G., Arze,  
516 C. A., Ayvaz, T., Baker, C. C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser,  
517 M. J., Bloom, T., Bonazzi, V., Paul Brooks, J., Buck, G. A., Buhay, C. J., Busam, D. A.,  
518 Campbell, J. L., Canon, S. R., Cantarel, B. L., Chain, P. S. G., Chen, I. M. A., Chen, L.,  
519 Chhibba, S., Chu, K., Ciulla, D. M., Clemente, J. C., Clifton, S. W., Conlan, S., Crabtree,  
520 J., Cutting, M. A., Davidovics, N. J., Davis, C. C., DeSantis, T. Z., Deal, C., Delehaunty,  
521 K. D., Dewhirst, F. E., Deych, E., Ding, Y., Dooling, D. J., Dugan, S. P., Michael Dunne,

- 522 W., Scott Durkin, A., Edgar, R. C., Erlich, R. L., Farmer, C. N., Farrell, R. M., Faust, K.,  
523 Feldgarden, M., Felix, V. M., Fisher, S., Fodor, A. A., Forney, L. J., Foster, L., Di  
524 Francesco, V., Friedman, J., Friedrich, D. C., Fronick, C. C., Fulton, L. L., Gao, H.,  
525 Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M. Y., Goldberg, J. M., Goll, J.,  
526 Gonzalez, A., Griggs, A., Gujja, S., Kinder Haake, S., Haas, B. J., Hamilton, H. A.,  
527 Harris, E. L., Hepburn, T. A., Herter, B., Hoffmann, D. E., Holder, M. E., Howarth, C.,  
528 Huang, K. H., Huse, S. M., Izard, J., Jansson, J. K., Jiang, H., Jordan, C., Joshi, V.,  
529 Katancik, J. A., Keitel, W. A., Kelley, S. T., Kells, C., King, N. B., Knights, D., Kong, H.  
530 H., Koren, O., Koren, S., Kota, K. C., Kovar, C. L., Kyrpides, N. C., La Rosa, P. S., Lee,  
531 S. L., Lemon, K. P., Lennon, N., Lewis, C. M., Lewis, L., Ley, R. E., Li, K., Liolios, K.,  
532 Liu, B., Liu, Y., Lo, C.-C., Lozupone, C. A., Dwayne Lunsford, R., Madden, T.,  
533 Mahurkar, A. A., Mannon, P. J., Mardis, E. R., Markowitz, V. M., Mavromatis, K.,  
534 McCorrison, J. M., McDonald, D., McEwen, J., McGuire, A. L., McInnes, P., Mehta, T.,  
535 Mihindukulasuriya, K. A., Miller, J. R., Minx, P. J., Newsham, I., Nusbaum, C.,  
536 O’Laughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S. M., Pearson, M.,  
537 Peterson, J., Podar, M., Pohl, C., Pollard, K. S., Pop, M., Priest, M. E., Proctor, L. M.,  
538 Qin, X., Raes, J., Ravel, J., Reid, J. G., Rho, M., Rhodes, R., Riehle, K. P., Rivera, M. C.,  
539 Rodriguez-Mueller, B., Rogers, Y.-H., Ross, M. C., Russ, C., Sanka, R. K., Sankar, P.,  
540 Fah Sathirapongsasuti, J., Schloss, J. A., Schloss, P. D., Schmidt, T. M., Scholz, M.,  
541 Schriml, L., Schubert, A. M., Segata, N., Segre, J. A., Shannon, W. D., Sharp, R. R.,  
542 Sharpton, T. J., Shenoy, N., Sheth, N. U., Simone, G. A., Singh, I., Smillie, C. S., Sobel,  
543 J. D., Sommer, D. D., Spicer, P., Sutton, G. G., Sykes, S. M., Tabbaa, D. G., Thiagarajan,  
544 M., Tomlinson, C. M., Torralba, M., Treangen, T. J., Truty, R. M., Vishnivetskaya, T. A.,  
545 Walker, J., Wang, L., Wang, Z., Ward, D. V., Warren, W., Watson, M. A., Wellington,  
546 C., Wetterstrand, K. A., White, J. R., Wilczek-Boney, K., Wu, Y., Wylie, K. M., Wylie,  
547 T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B. P., Zhang, L., Zhou, Y., Zhu,  
548 Y., Zoloth, L., Zucker, J. D., Birren, B. W., Gibbs, R. A., Highlander, S. K., Methé, B.  
549 A., Nelson, K. E., Petrosino, J. F., Weinstock, G. M., Wilson, R. K., White, O. and The  
550 Human Microbiome Project, C. (2012). "Structure, function and diversity of the healthy  
551 human microbiome." *Nature* **486**(7402): 207-214 DOI:  
552 <https://doi.org/10.1038/nature11234>.
- 553 Li, Y., Ash, K. T., Joyner, D. C., Williams, D. E., Alamilla, I., McKay, P., Iler, C., Green, B.,  
554 Kara-Murdoch, F. and Swift, C. M. (2023). "Decay of enveloped SARS-CoV-2 and non-  
555 enveloped PMMoV RNA in raw sewage from university dormitories." *Frontiers in*  
556 *Microbiology* **14**: 1144026 DOI: <https://doi.org/10.3389/fmicb.2023.1144026>.
- 557 Lindström, E. S., Agterveld, M. P. K.-V. and Zwart, G. (2005). "Distribution of Typical  
558 Freshwater Bacterial Groups Is Associated with pH, Temperature, and Lake Water  
559 Retention Time." *Applied and Environmental Microbiology* **71**(12): 8201-8206 DOI:  
560 <https://doi.org/10.1128/AEM.71.12.8201-8206.2005>.
- 561 Liu, Q., Mak, J. W. Y., Su, Q., Yeoh, Y. K., Lui, G. C.-Y., Ng, S. S. S., Zhang, F., Li, A. Y. L.,  
562 Lu, W., Hui, D. S.-C., Chan, P. K., Chan, F. K. L. and Ng, S. C. (2022). "Gut microbiota  
563 dynamics in a prospective cohort of patients with post-acute COVID-19 syndrome." *Gut*  
564 **71**(3): 544-552 DOI: <https://doi.org/10.1136/gutjnl-2021-325989>.
- 565 Newton, R. J., McLellan, S. L., Dila, D. K., Vineis, J. H., Morrison, H. G., Eren, A. M. and  
566 Sogin, M. L. (2015). "Sewage reflects the microbiomes of human populations." *mBio*

- 567 6(2): 10.1128/mbio.02574-02514 DOI: [https://doi-](https://doi-org.utk.idm.oclc.org/10.1128/mbio.02574-14)  
568 [org.utk.idm.oclc.org/10.1128/mbio.02574-14](https://doi-org.utk.idm.oclc.org/10.1128/mbio.02574-14).
- 569 Numberger, D., Ganzert, L., Zoccarato, L., Mühldorfer, K., Sauer, S., Grossart, H.-P. and  
570 Greenwood, A. D. (2019). "Characterization of bacterial communities in wastewater with  
571 enhanced taxonomic resolution by full-length 16S rRNA sequencing." *Scientific Reports*  
572 **9**(1): 9673 DOI: <https://doi.org/10.1038/s41598-019-46015-z>.
- 573 Oluseyi Osunmakinde, C., Selvarajan, R., Mamba, B. B. and Msagati, T. A. M. (2019).  
574 "Profiling Bacterial Diversity and Potential Pathogens in Wastewater Treatment Plants  
575 Using High-Throughput Sequencing Analysis." *Microorganisms* **7**(11): 506 DOI:  
576 <https://doi-org.utk.idm.oclc.org/10.3390/microorganisms7110506>.
- 577 Poopedi, E., Singh, T. and Gomba, A. (2023). "Potential Exposure to Respiratory and Enteric  
578 Bacterial Pathogens among Wastewater Treatment Plant Workers, South Africa."  
579 *International Journal of Environmental Research and Public Health* **20**(5): 4338 DOI:  
580 <https://doi-org.utk.idm.oclc.org/10.3390/ijerph20054338>.
- 581 Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R. and Gordon, J. I.  
582 (2006). "An obesity-associated gut microbiome with increased capacity for energy  
583 harvest." *Nature* **444**(7122): 1027-1031 DOI: <https://10.1038/nature05414>.
- 584 Yeoh, Y. K., Zuo, T., Lui, G. C.-Y., Zhang, F., Liu, Q., Li, A. Y., Chung, A. C., Cheung, C. P.,  
585 Tso, E. Y., Fung, K. S., Chan, V., Ling, L., Joynt, G., Hui, D. S.-C., Chow, K. M., Ng, S.  
586 S. S., Li, T. C.-M., Ng, R. W., Yip, T. C., Wong, G. L.-H., Chan, F. K., Wong, C. K.,  
587 Chan, P. K. and Ng, S. C. (2021). "Gut microbiota composition reflects disease severity  
588 and dysfunctional immune responses in patients with COVID-19." *Gut* **70**(4): 698-706  
589 DOI: <https://doi.org/10.1136/gutjnl-2020-323020>.
- 590 Zhang, F., Lau, R. I., Liu, Q., Su, Q., Chan, F. K. L. and Ng, S. C. (2023). "Gut microbiota in  
591 COVID-19: key microbial changes, potential mechanisms and clinical applications."  
592 *Nature Reviews Gastroenterology & Hepatology* **20**(5): 323-337 DOI:  
593 <https://doi.org/10.1038/s41575-022-00698-4>.
- 594 Zuo, T., Zhang, F., Lui, G. C., Yeoh, Y. K., Li, A. Y., Zhan, H., Wan, Y., Chung, A. C., Cheung,  
595 C. P. and Chen, N. (2020). "Alterations in gut microbiota of patients with COVID-19  
596 during time of hospitalization." *Gastroenterology* **159**(3): 944-955. e948 DOI:  
597 <https://doi.org/10.1053/j.gastro.2020.05.048>.
- 598

Sampling site	Sampling Point	Gender	Student Number
D1	Direct Dispense from the valve	Male	387-504
D2	Direct Dispense from the valve	Female	469-531
D3	Direct Dispense from the valve	Mix	254-279
D4	Direct Dispense from the valve	Mix	529-637
D5	Direct Dispense from the valve	Mix	10-672
D6	Direct Dispense from the valve	Mix	580-672

599 Table 1. Demography data for D1, D2, D3, D4, D5, and D6.

600



601 Table 2. Raw wastewater data information for D1, D2, D3, D4, D5, and D6.

Dorms	pH	Positive Rate	Relative abundance of potential Pathogen		602
			Positive SARS-CoV-2 Sample	Negative SARS-CoV-2 Sample	
D1	6.71-9.08	70.00%	40.04%	40.35%	
D2	6.83-8.27	39.00%	21.80%	29.37%	
D3	6.51-8.97	52.00%	17.58%	13.88%	
D4	6.27-9.01	20.00%	30.63%	24.11%	
D5	6.38-8.95	68.00%	43.49%	45.17%	
D6	5.72-8.63	37.00%	41.30%	45.64%	

603 **Figure Legends**

604 Figure 1. Map of the sampling locations on the University of Tennessee-Knoxville campus.

605 Figure 2. Experimental design and sampling points/times for microbiome sequencing. SARS-  
606 CoV-2 concentrations are indicated as yellow lines. The yellow points at 10 copies/L represent  
607 negative samples. The sequencing runs are indicated as black points. The virus load was  
608 estimated by qPCR in untreated wastewater from different dormitories: D1, D2, D3, D4, D5, and  
609 D6. The study was conducted from Sep 2020 to Oct 2021.

610 Figure 3. Relative abundances of the top 10 dominant phyla in 6 dormitories.

611 Figure 4. Relative abundances at family levels for six dormitories.

612 Figure 5. Relative abundances of top 50 genera.

613 Figure 6. Venn diagram of exclusives and shared bacteria among the 6 dormitories.

614 Figure 7. Diversity index in 6 dormitories. The box-and-whisker plots show the mean (diamond),  
615 median (middle bar), first quartile (lower bar), third quartile (upper bar), minimum observation  
616 above the lowest fence (lower whisker), and maximum observation below the upper fence (upper  
617 whisker) of common  $\alpha$ -diversity metrics for each group. The P values for the comparison  
618 between groups using linear regression models including semester as covariate is also shown.

619 Figure 8. The scatter plots show each participant's microbial community composition (small  
620 circles) by group, as well as each group's centroid (large circles) and 95% CI ellipses. The  
621 scatter plots were generated using Principal Coordinates Analysis (PCoA) ordination based on  
622 common  $\beta$ -diversity metrics. For ease of visualization, only 2 dimensions were used. The P  
623 values for the comparison between groups using permutational multivariate ANOVA models  
624 including semester as covariate is also show.

625 Figure 9. Relative abundances of the top 10 dominant phyla in 6 dormitories with positive and  
626 negative SARS-CoV-2 samples.

627 Figure 10. Relative abundances of the top 10 dominant family in 6 dormitories with positive and  
628 negative SARS-CoV-2 samples.

629 Figure 11. Significant changes at family levels with positive and negative SARS-CoV-2 samples  
630 in 6 dormitories.

631 Figure 12. Relative abundances of top 50 genera and potential pathogens with positive and  
632 negative SARS-CoV-2 samples in 6 dormitories. The genera are listed from the highest relative  
633 abundance (top) to the least relative abundance (bottom). The pathogens are marked with an  
634 orange box around their name.

635 Figure 13. Venn diagram of exclusives and shared bacteria with positive and negative SARS-  
636 CoV-2 samples in the 6 dormitories.

637 Figure 14. Diversity index with significant difference between the positive and negative SARS-  
638 CoV-2 samples in 6 dormitories. The box-and-whisker plots show the mean (diamond), median  
639 (middle bar), first quartile (lower bar), third quartile (upper bar), minimum observation above the  
640 lowest fence (lower whisker), and maximum observation below the upper fence (upper whisker)  
641 of common  $\alpha$ -diversity metrics just for significant group. The P values for the comparison  
642 between groups using linear regression models including semester as covariate is also shown.

643 Figure 15. The scatter plots show each participant's microbial community composition (small  
644 circles) by D4 and D6, as well as their centroid (large circles) and 95% CI ellipses. The scatter  
645 plots were generated using Principal Coordinates Analysis (PCoA) ordination based on common  
646  $\beta$ -diversity metrics. For ease of visualization, only 2 dimensions were used. The P values for the  
647 comparison between groups using permutational multivariate ANOVA models including  
648 semester as covariate is also shown.

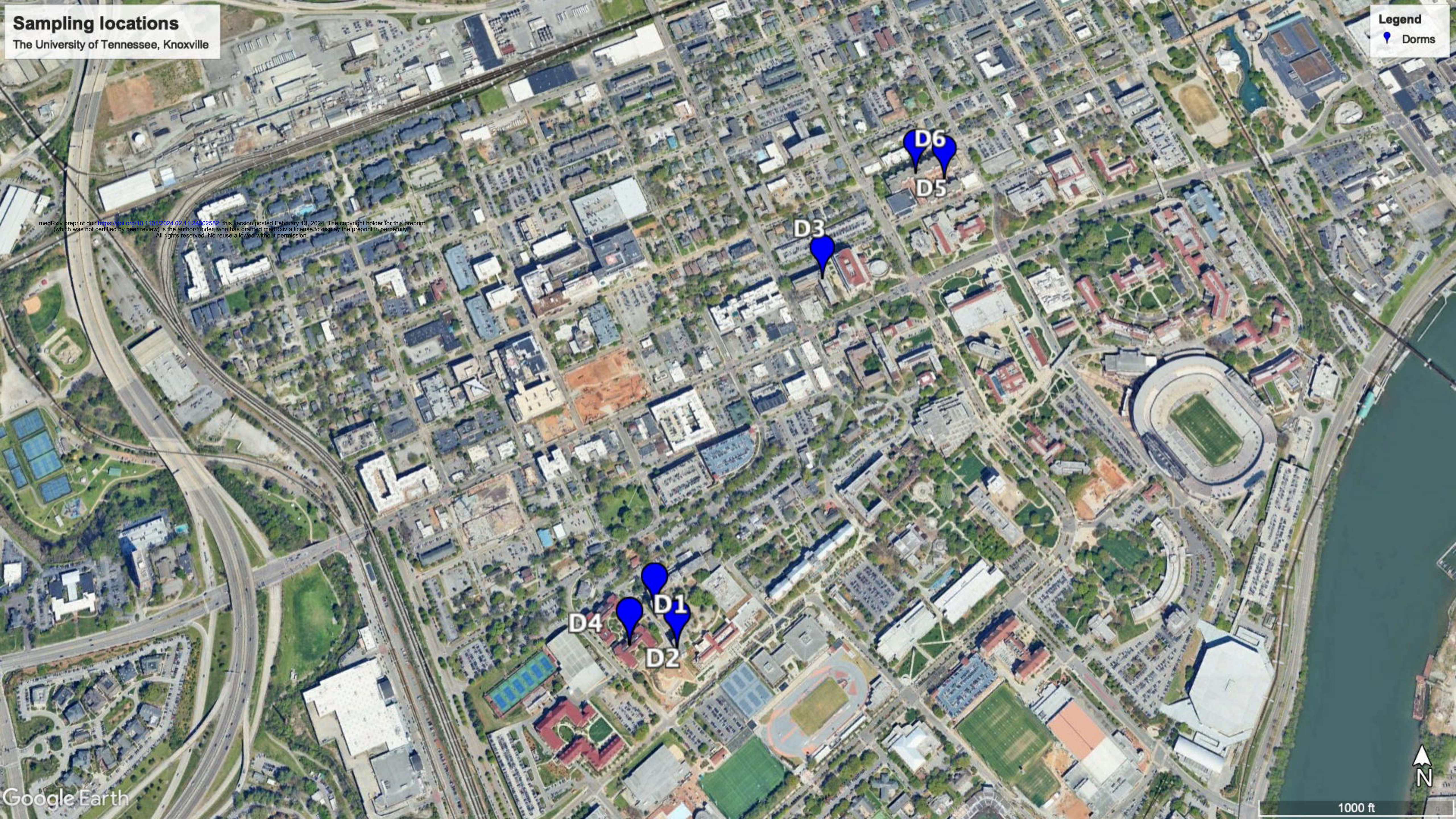
649 Figure S1. Histograms of linear discriminant analysis (LDA) effect size (LEfSe) comparison  
650 between positive and negative SARS-CoV-2 samples microbiota at the genus level in D3, D4  
651 and D5. Log-level changes in LDA score are displayed on the x axis

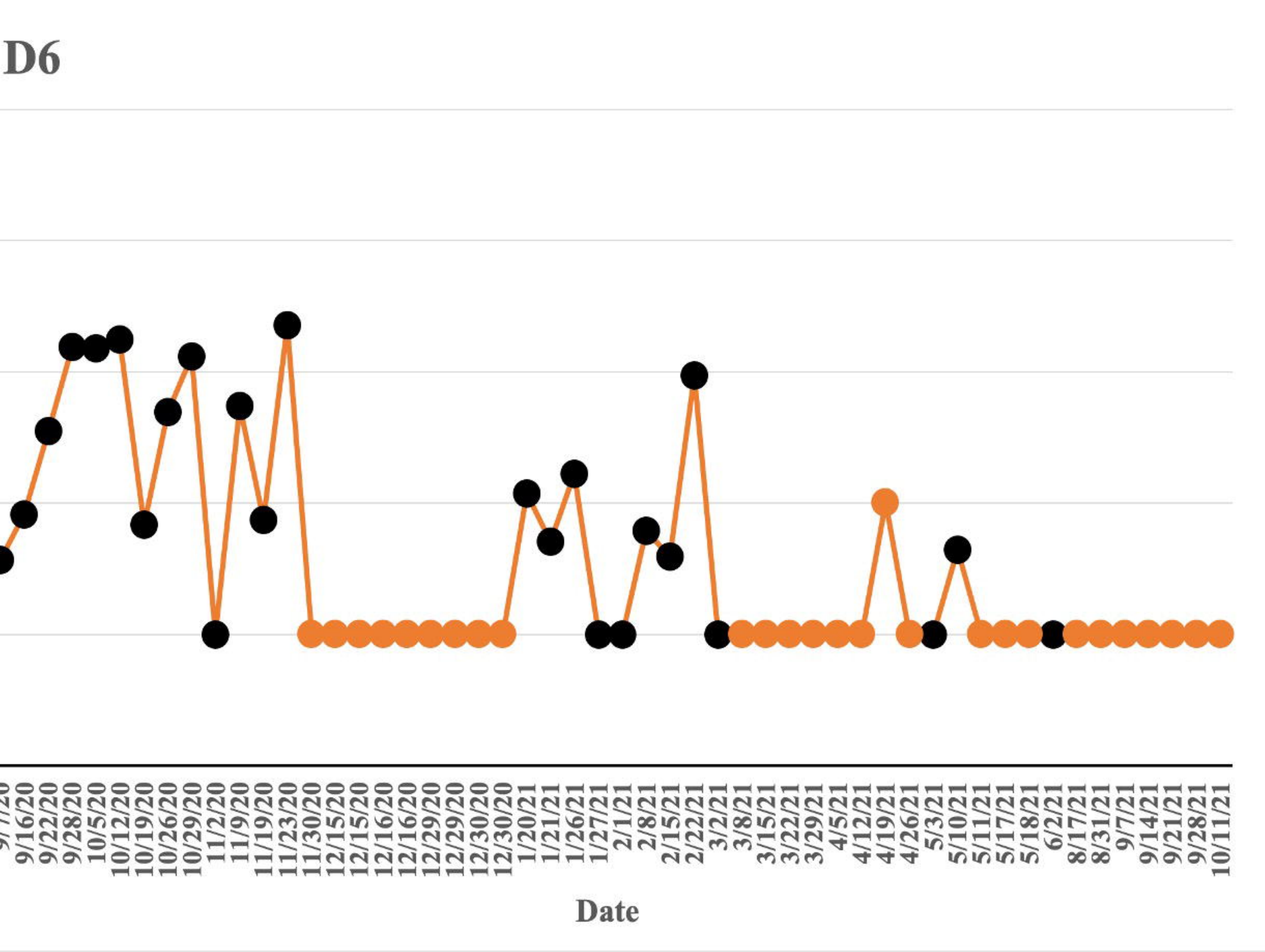
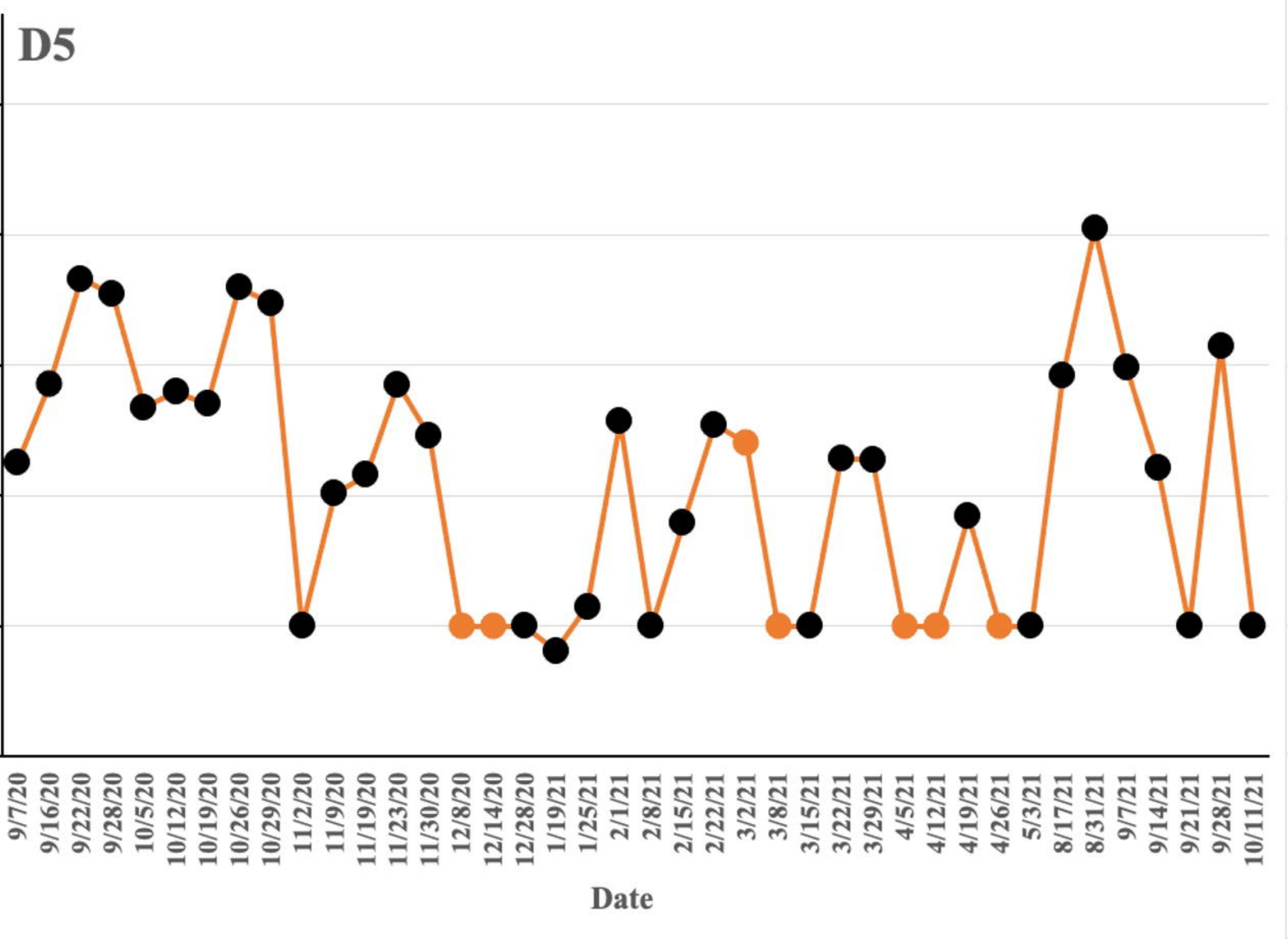
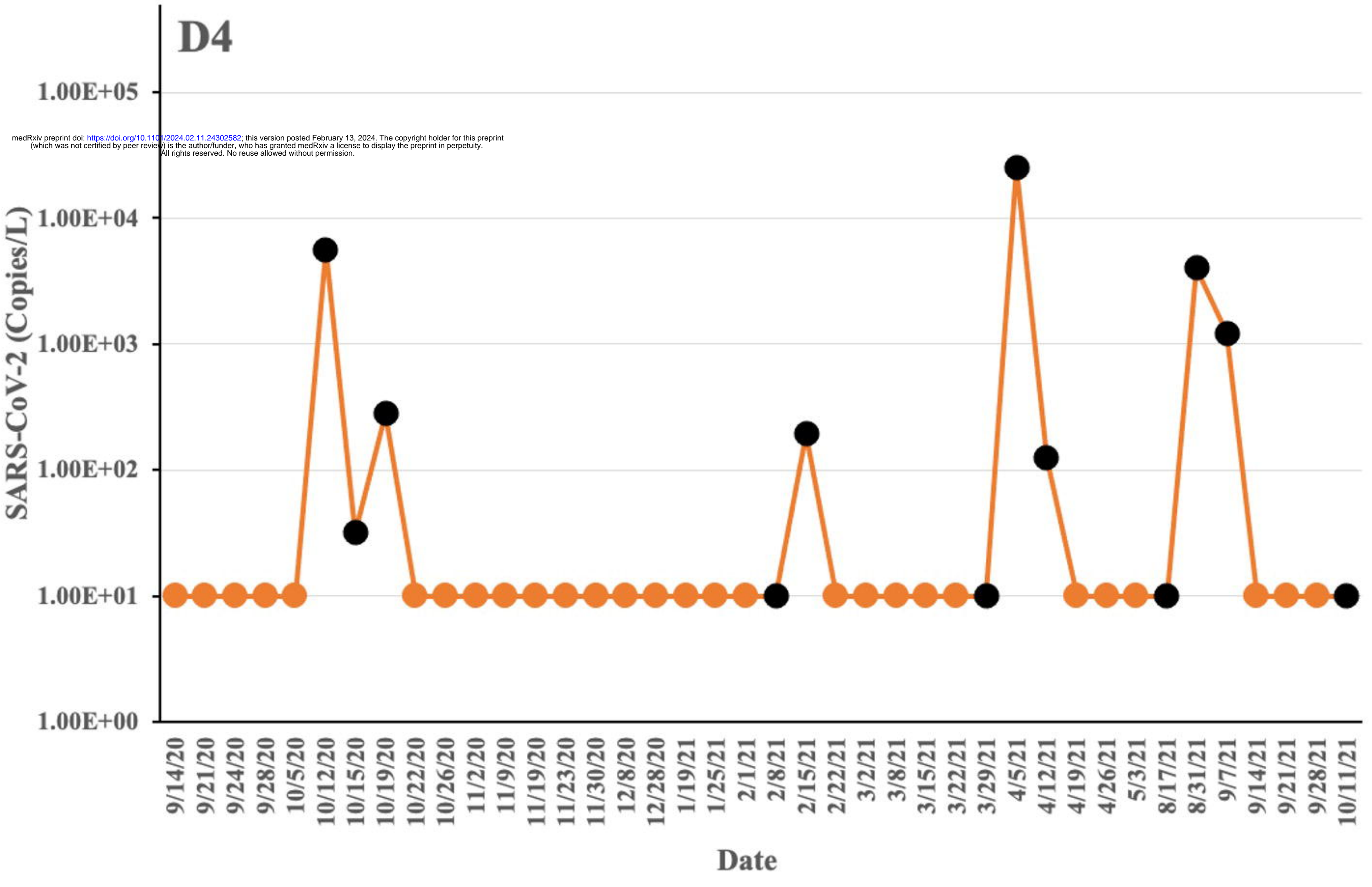
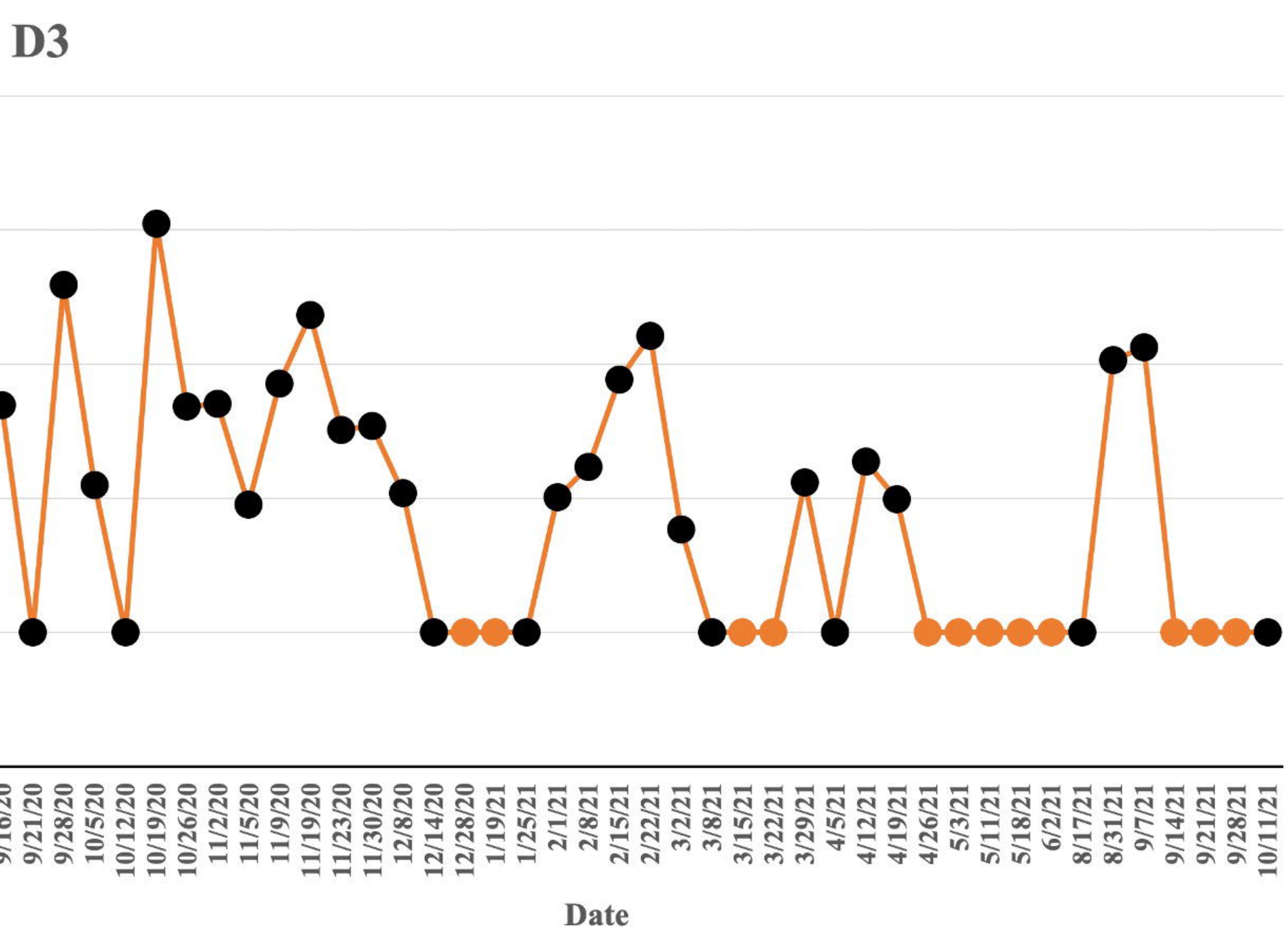
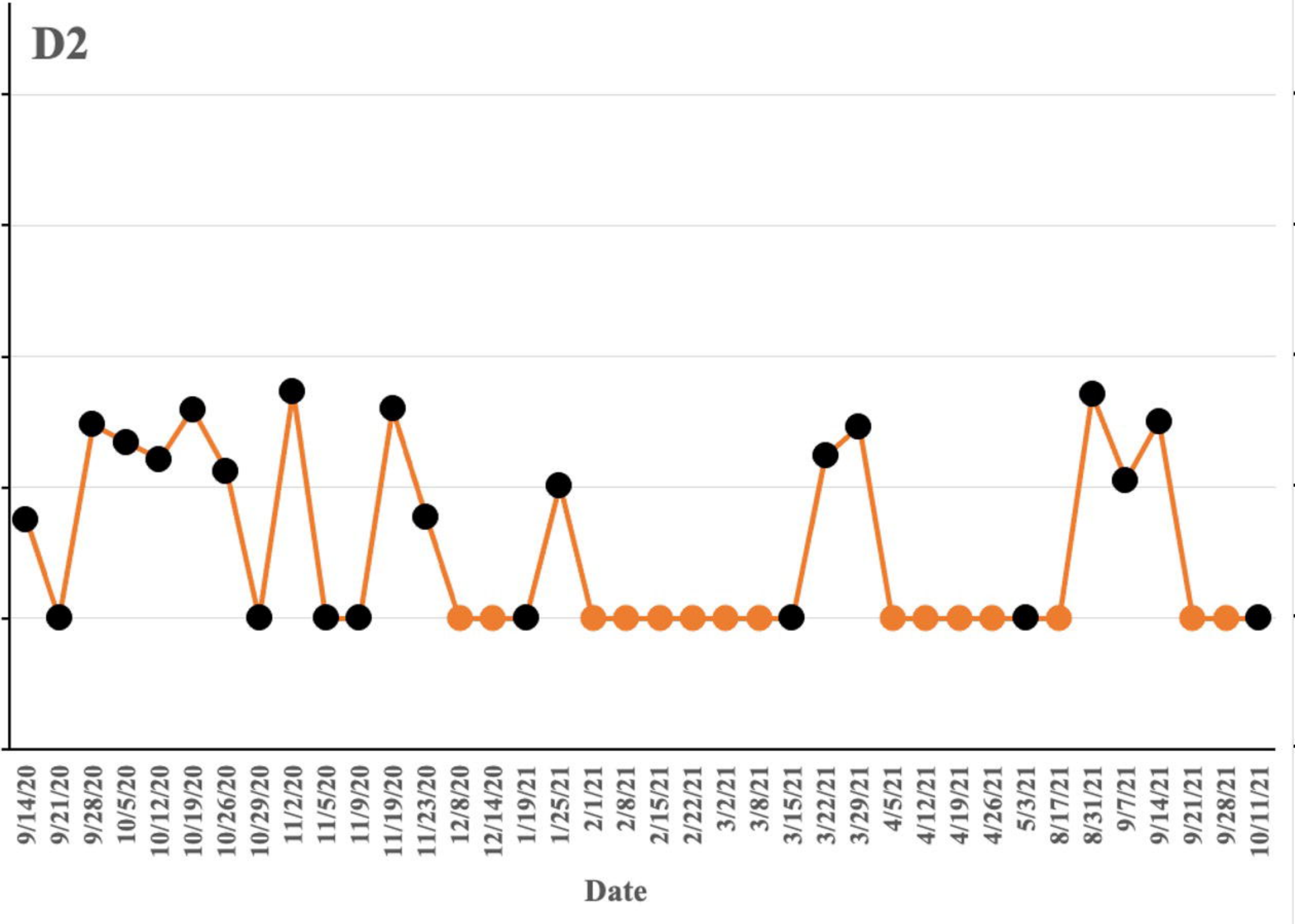
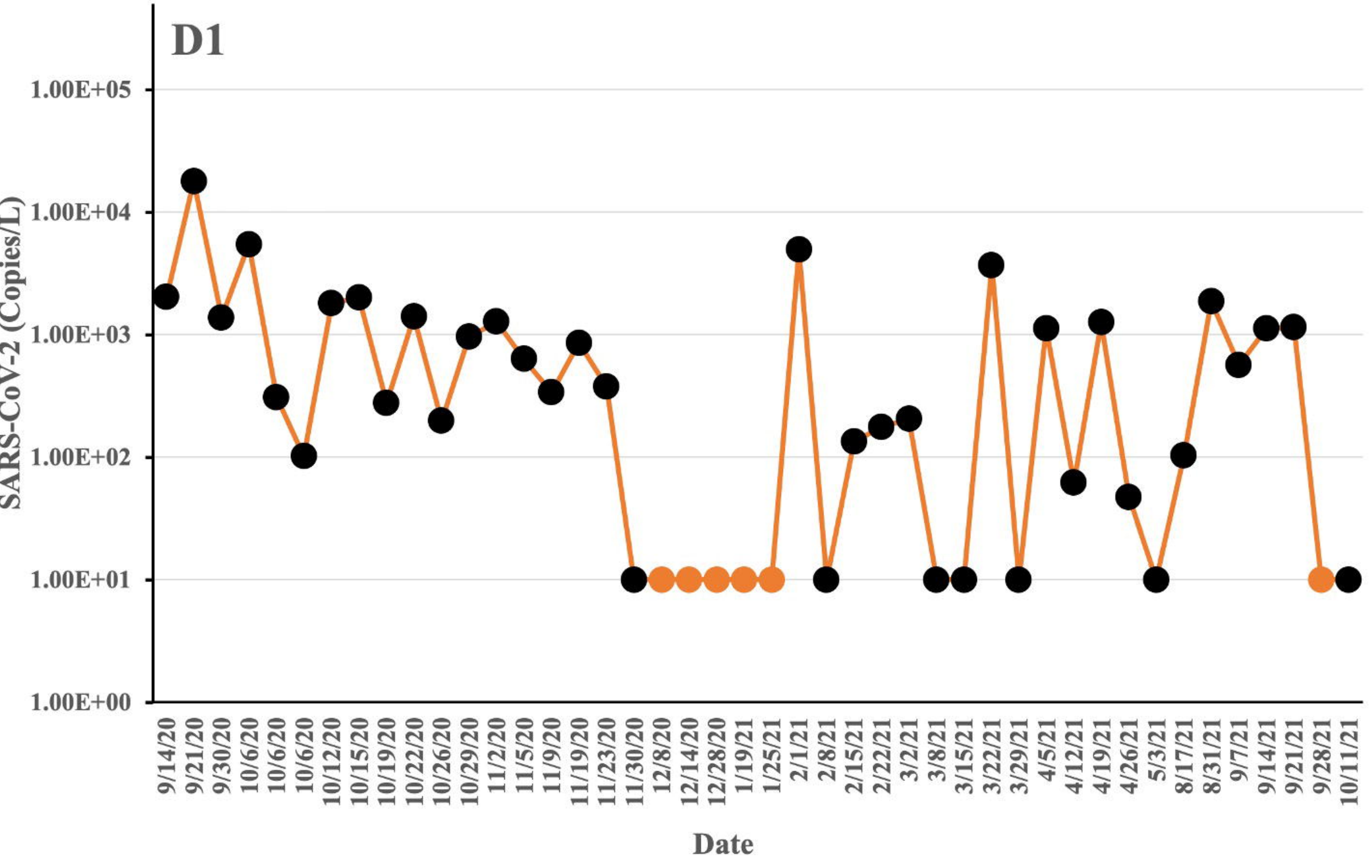
# Sampling locations

The University of Tennessee, Knoxville

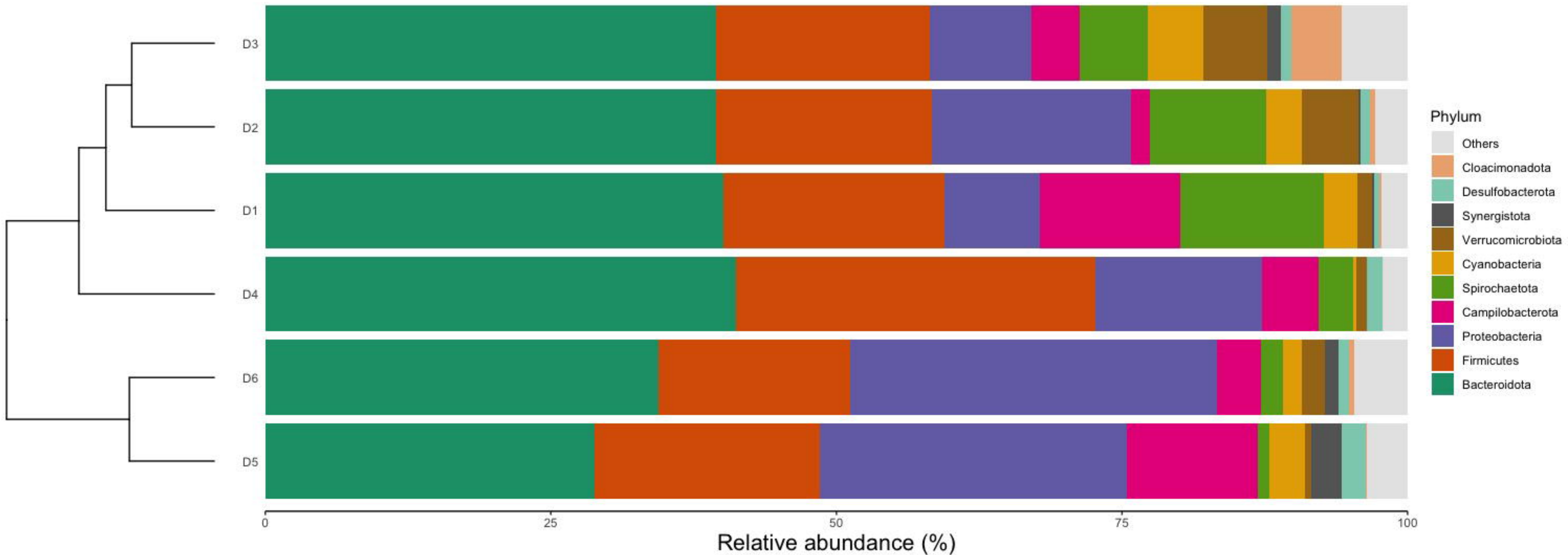
**Legend**  
📍 Dorms

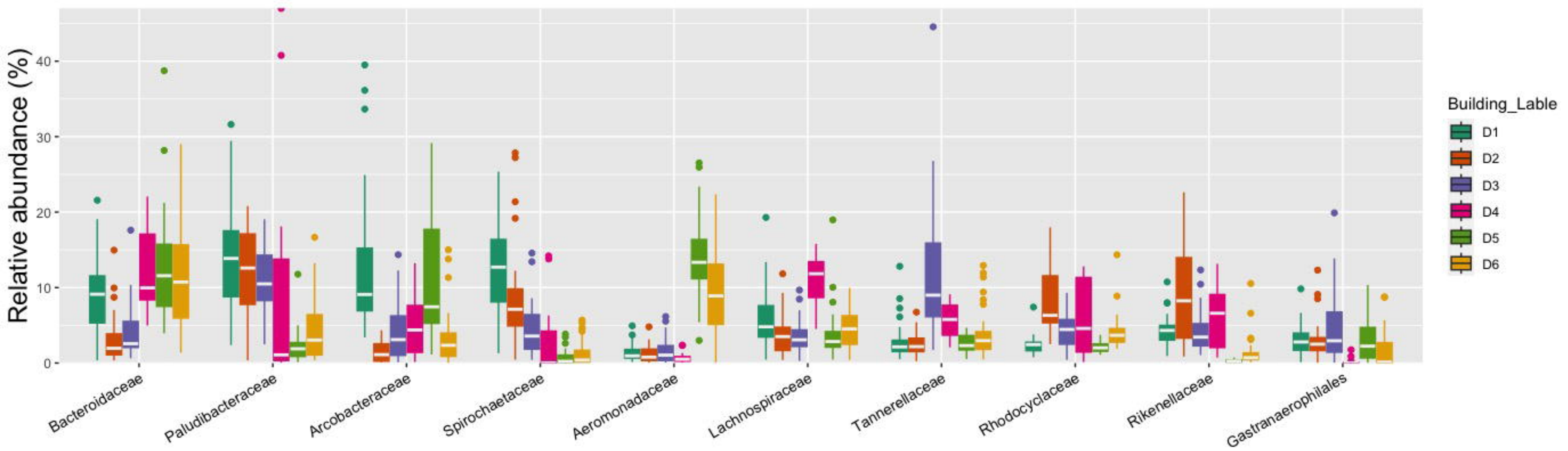
medRxiv preprint doi: <https://doi.org/10.1101/2024.02.11.24102558>; this version posted February 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

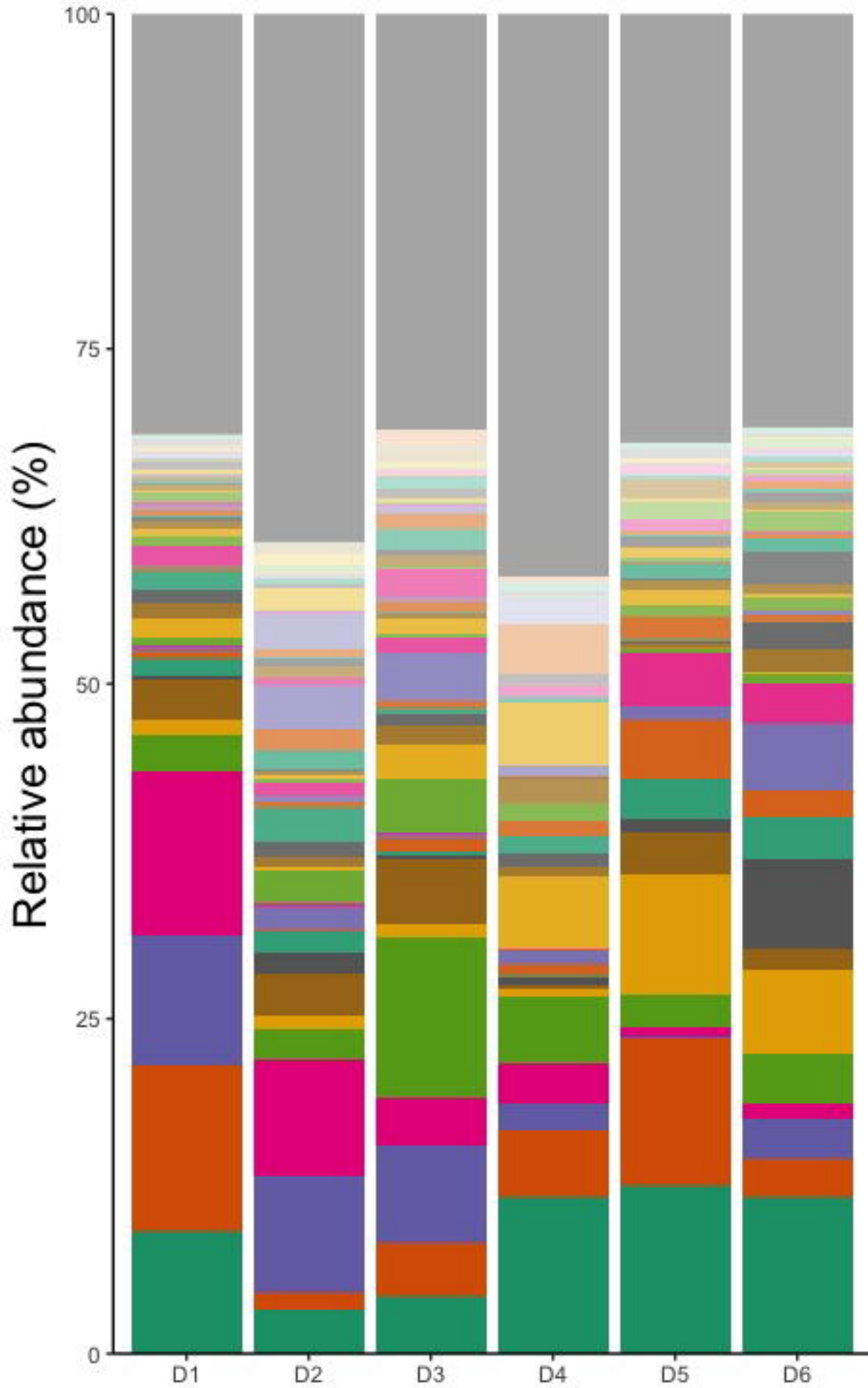




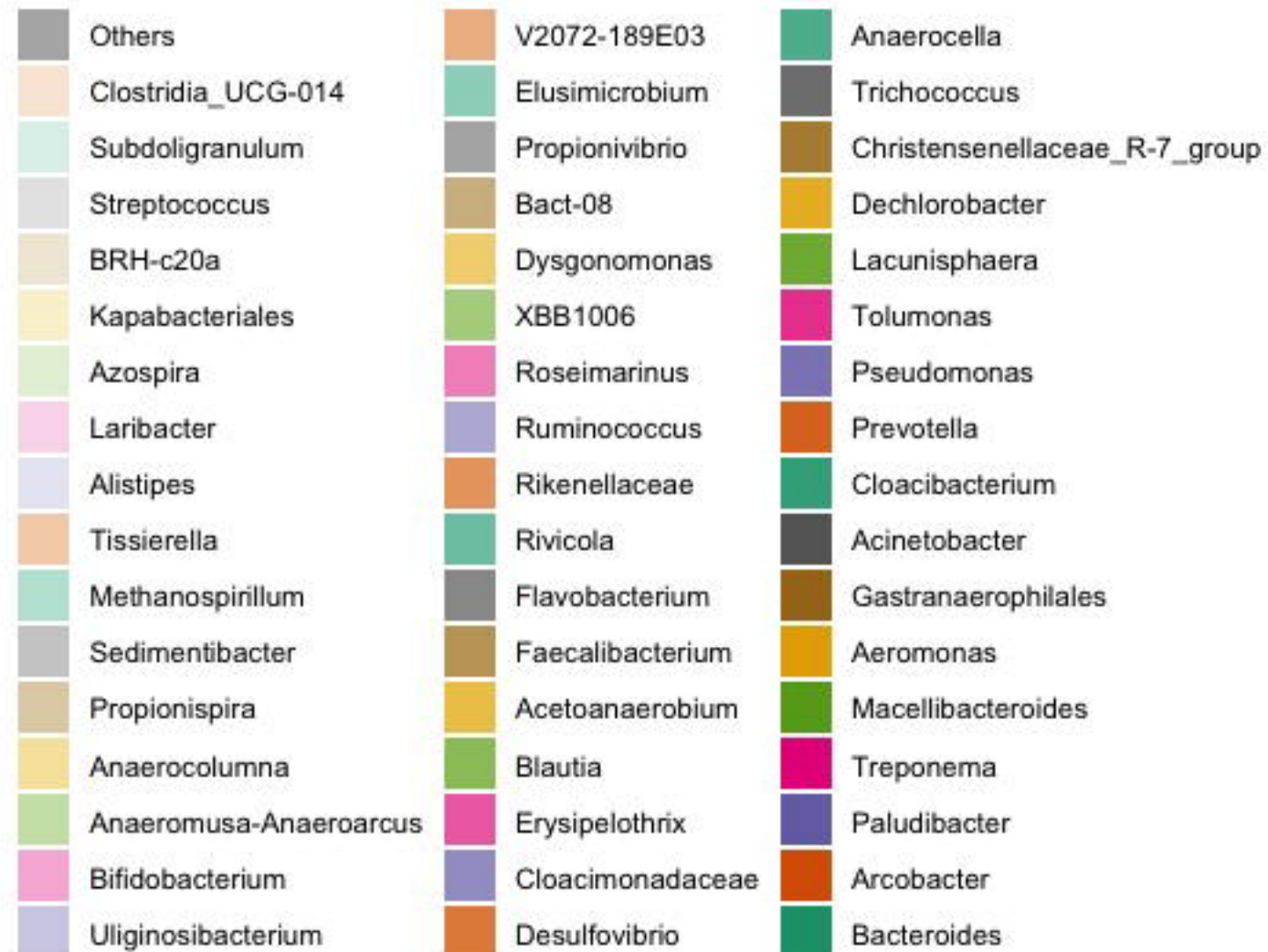
medRxiv preprint doi: <https://doi.org/10.1101/2024.02.11.24302582>; this version posted February 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



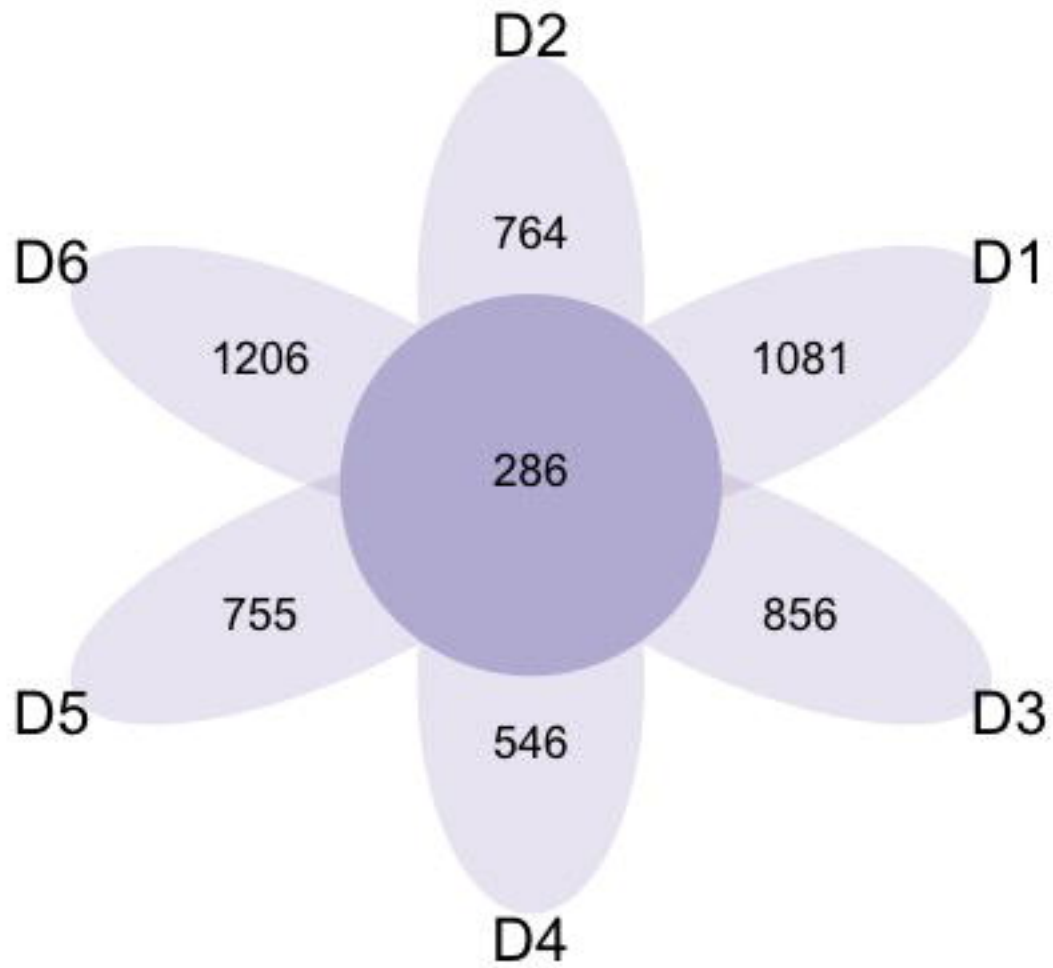


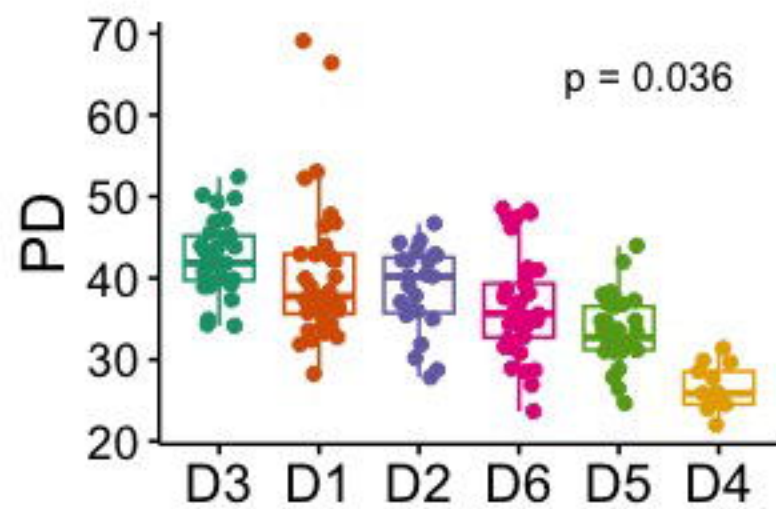
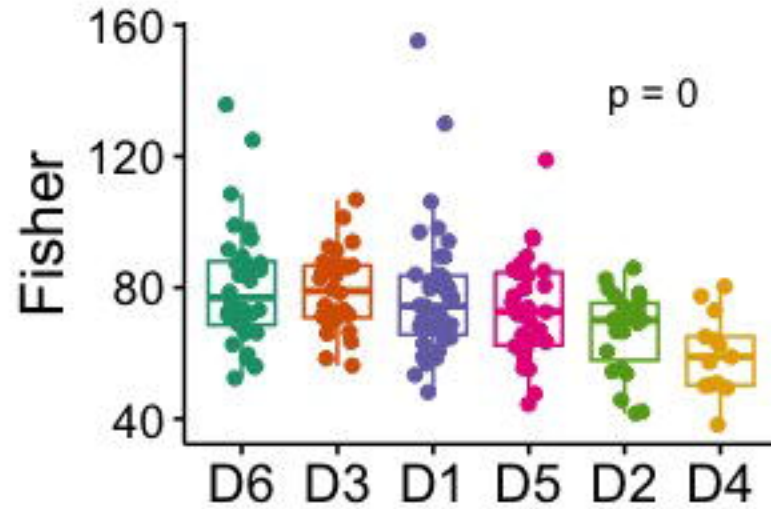
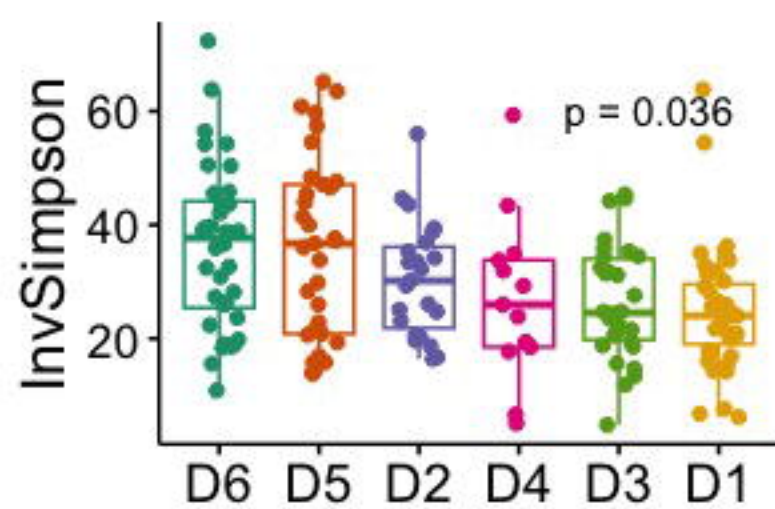
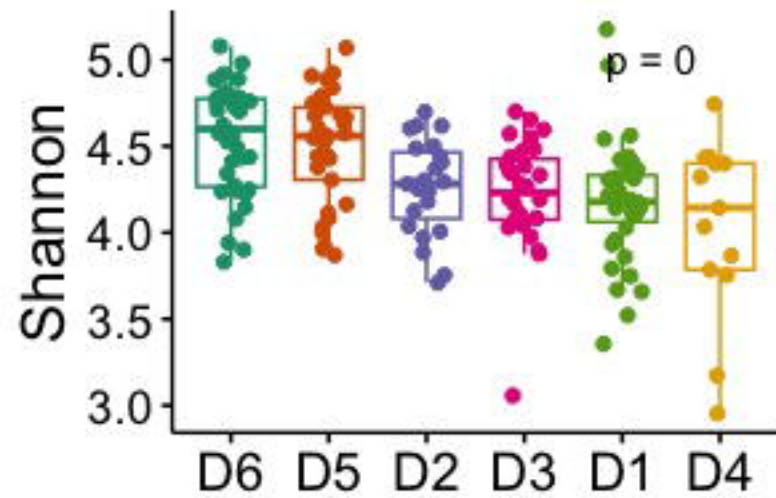
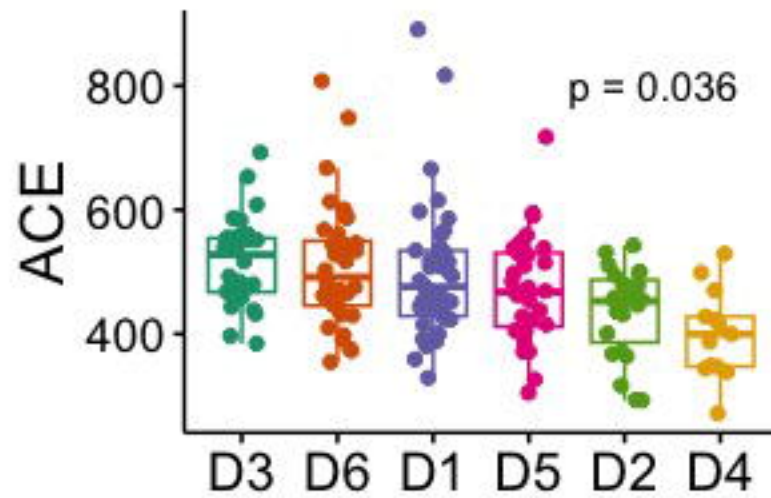
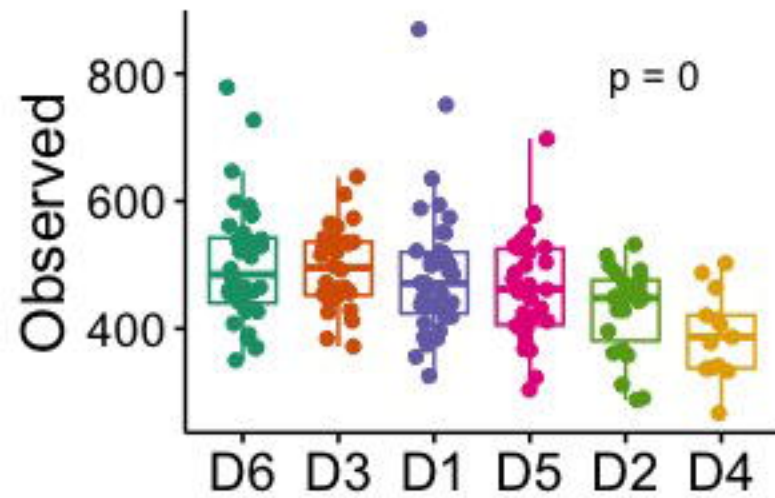


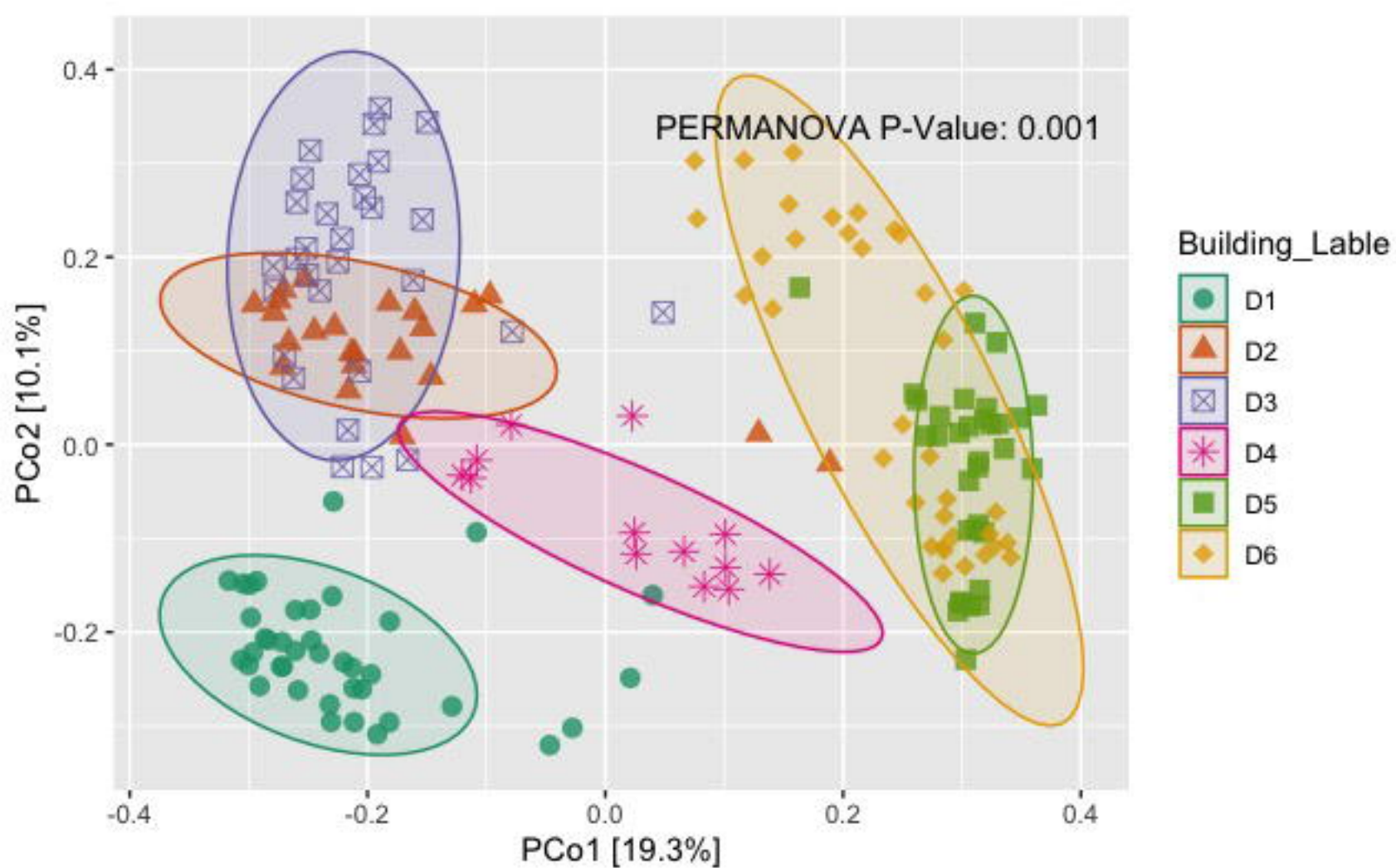
### Genus

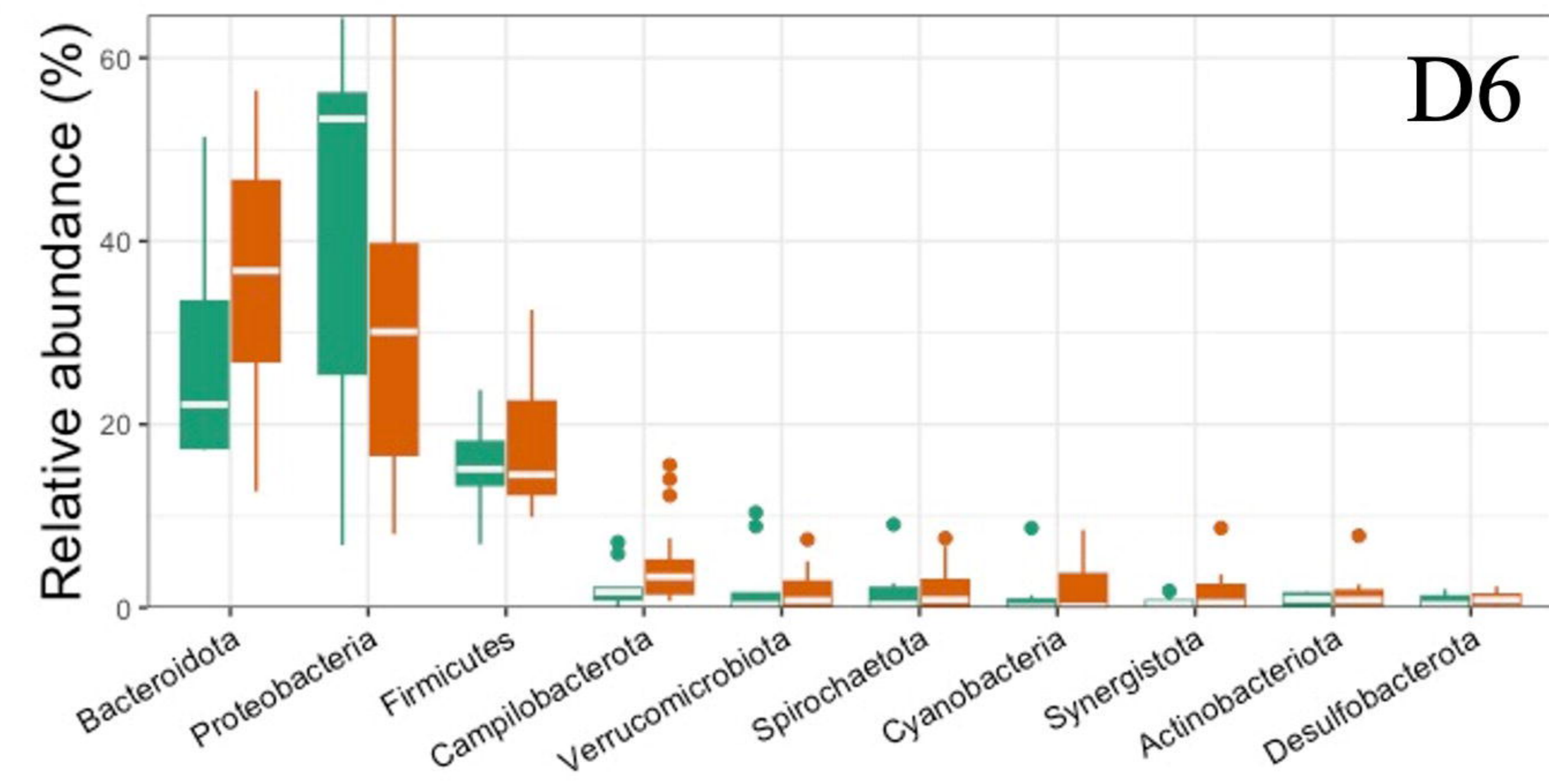
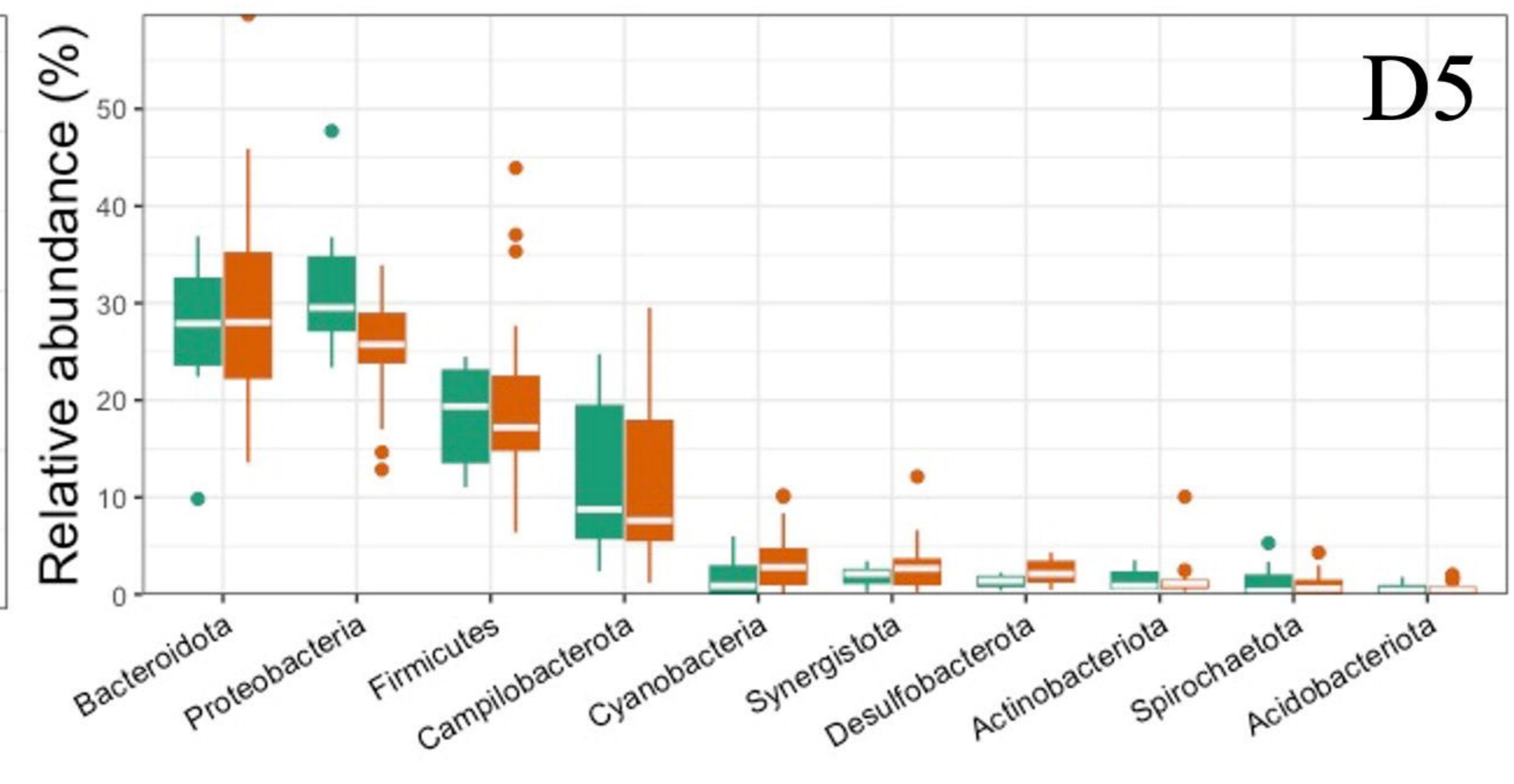
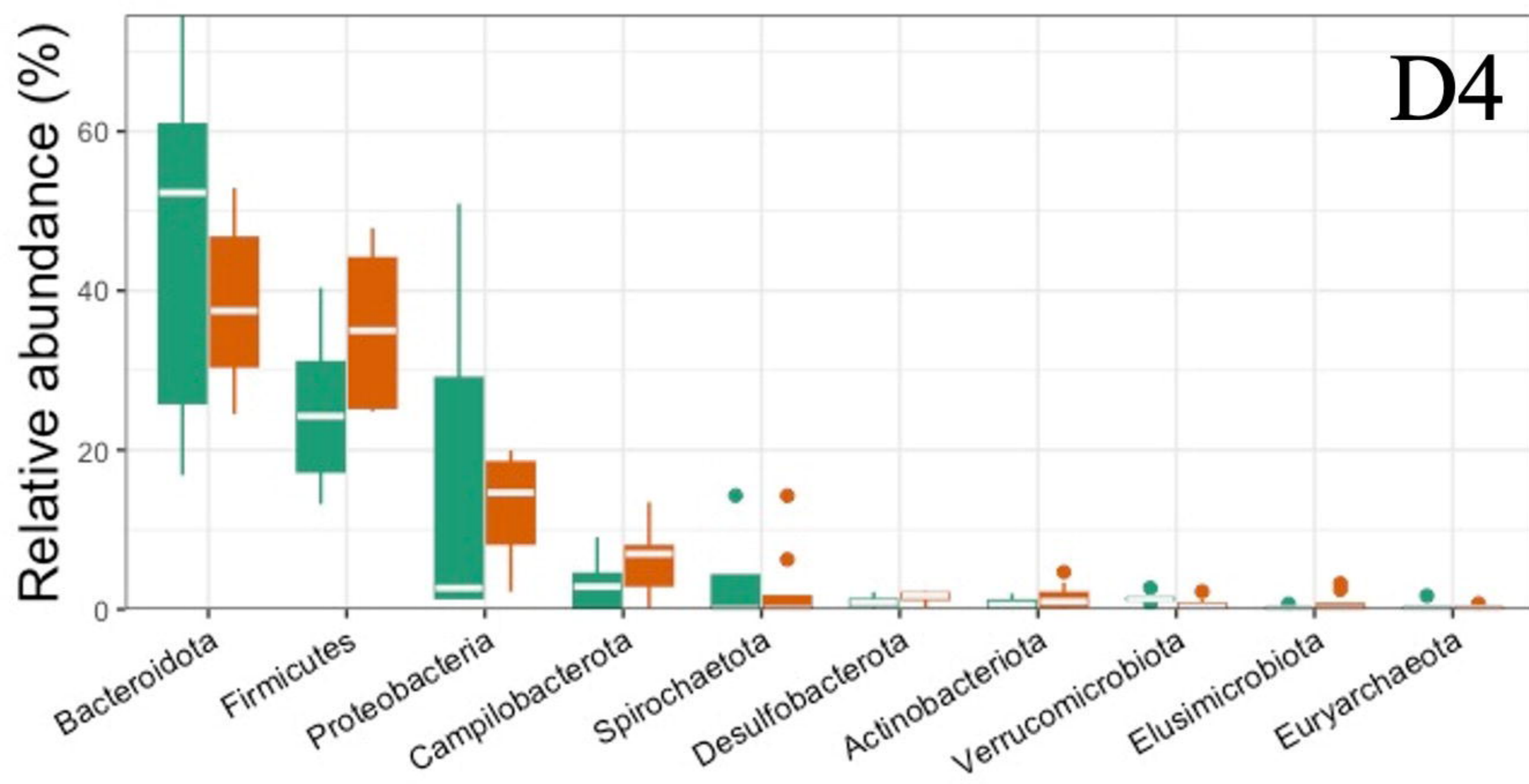
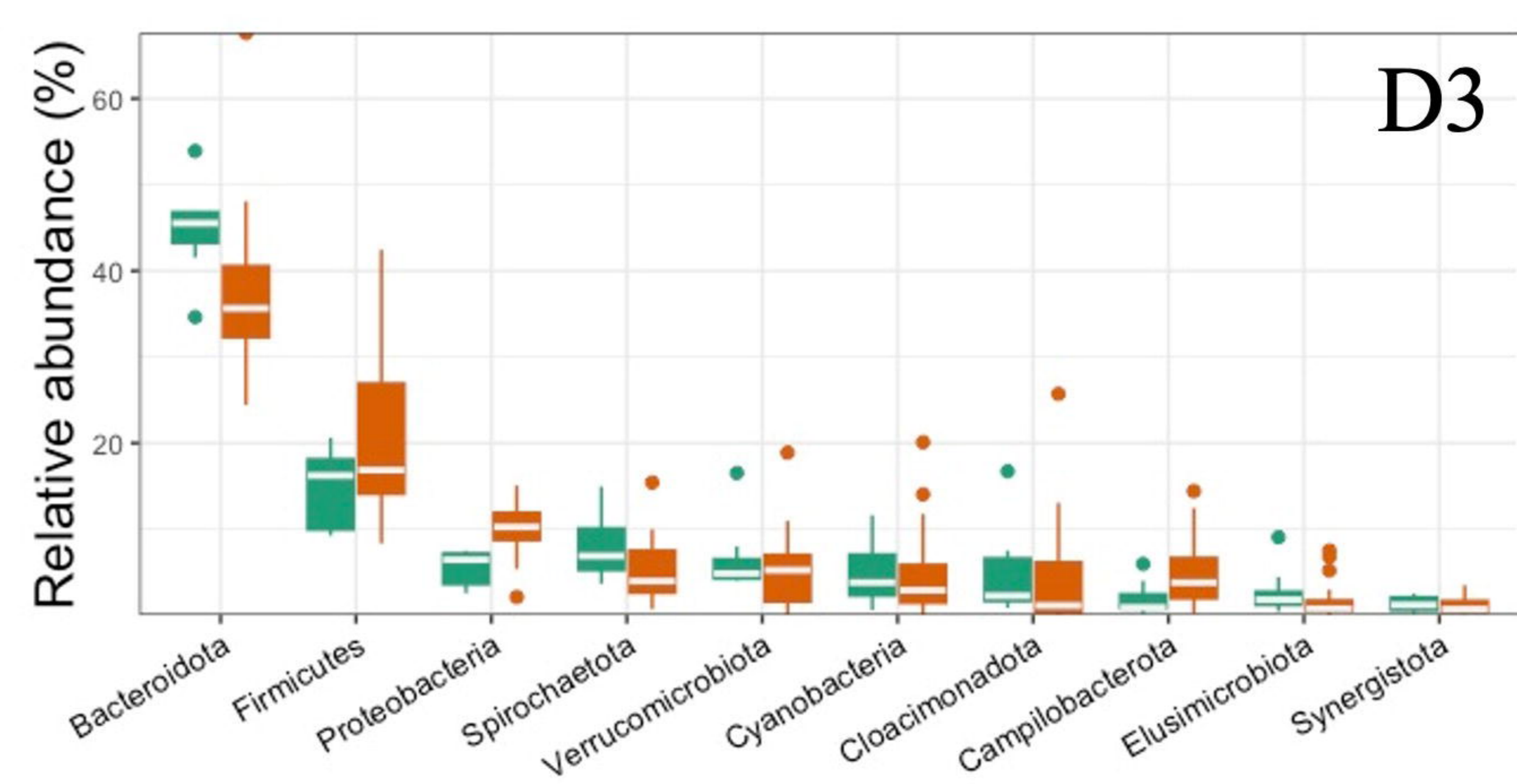
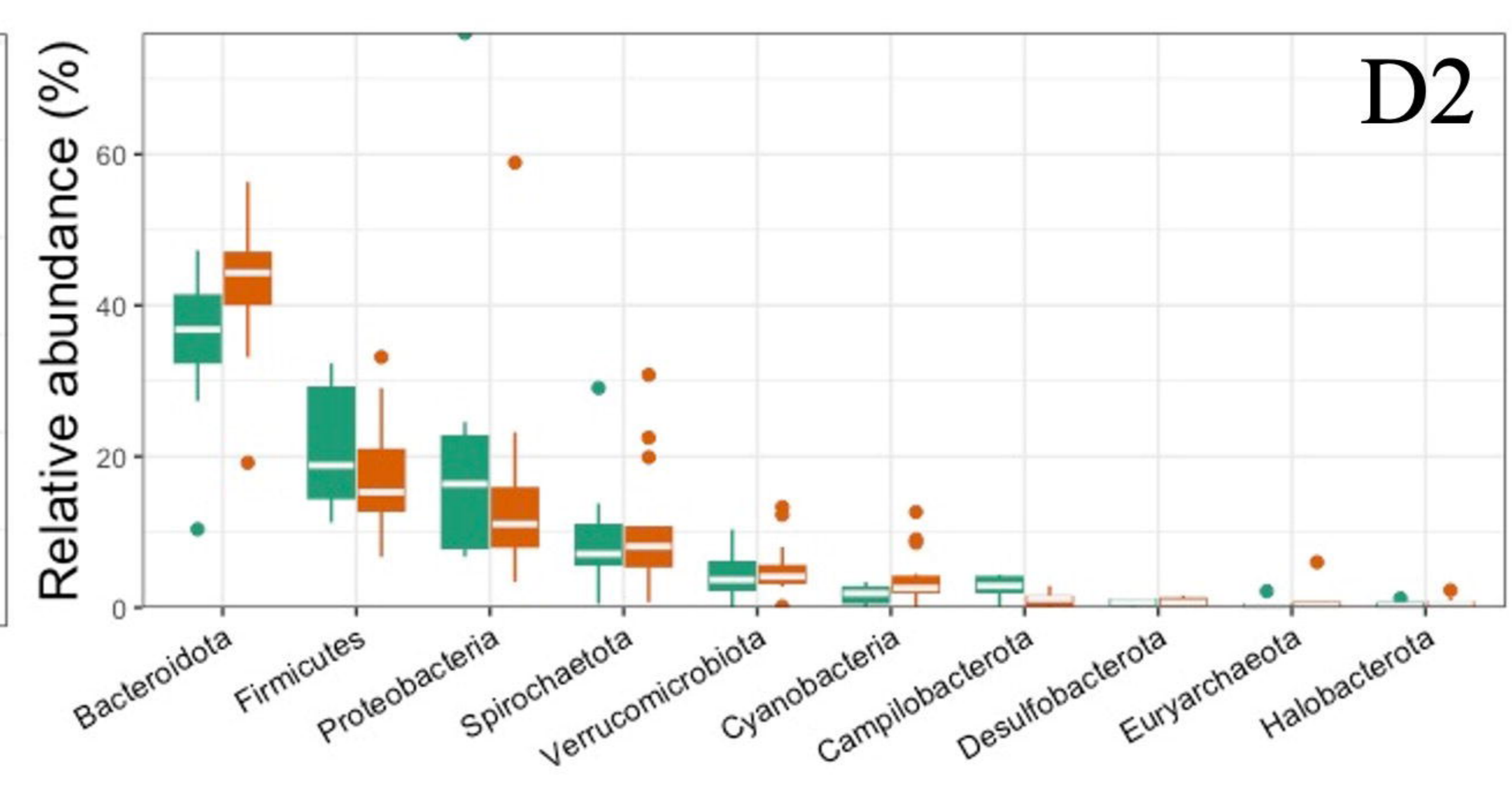
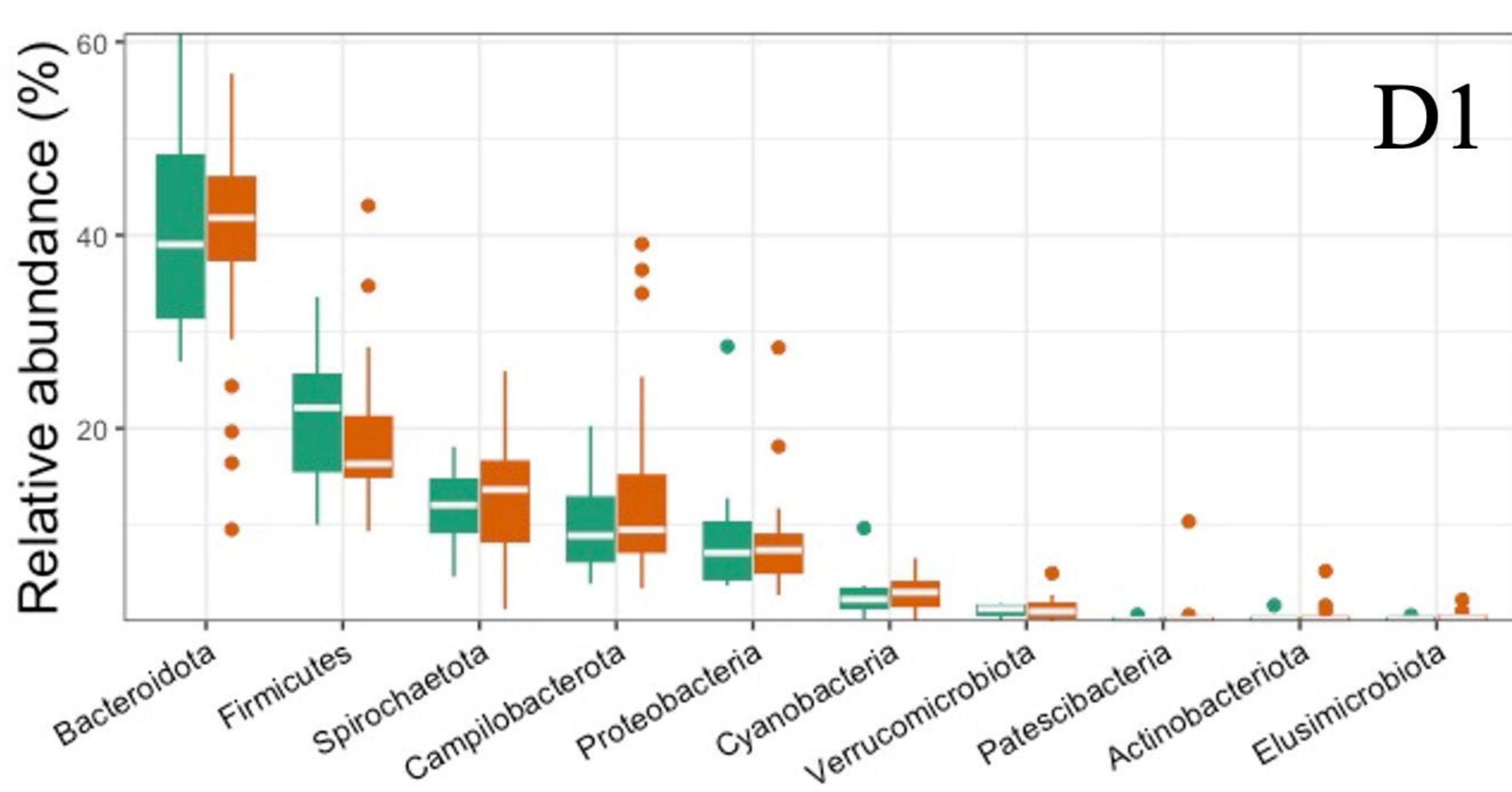


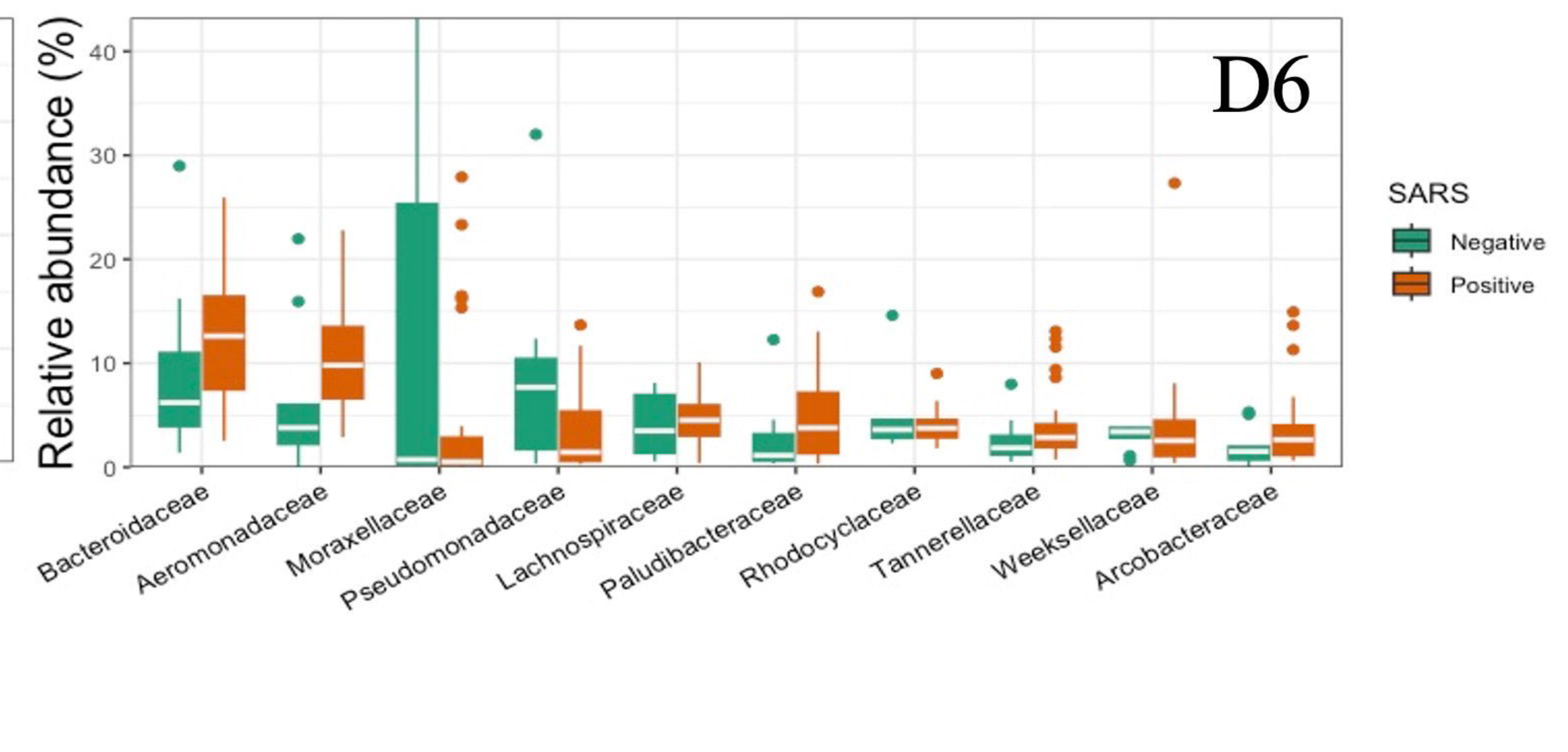
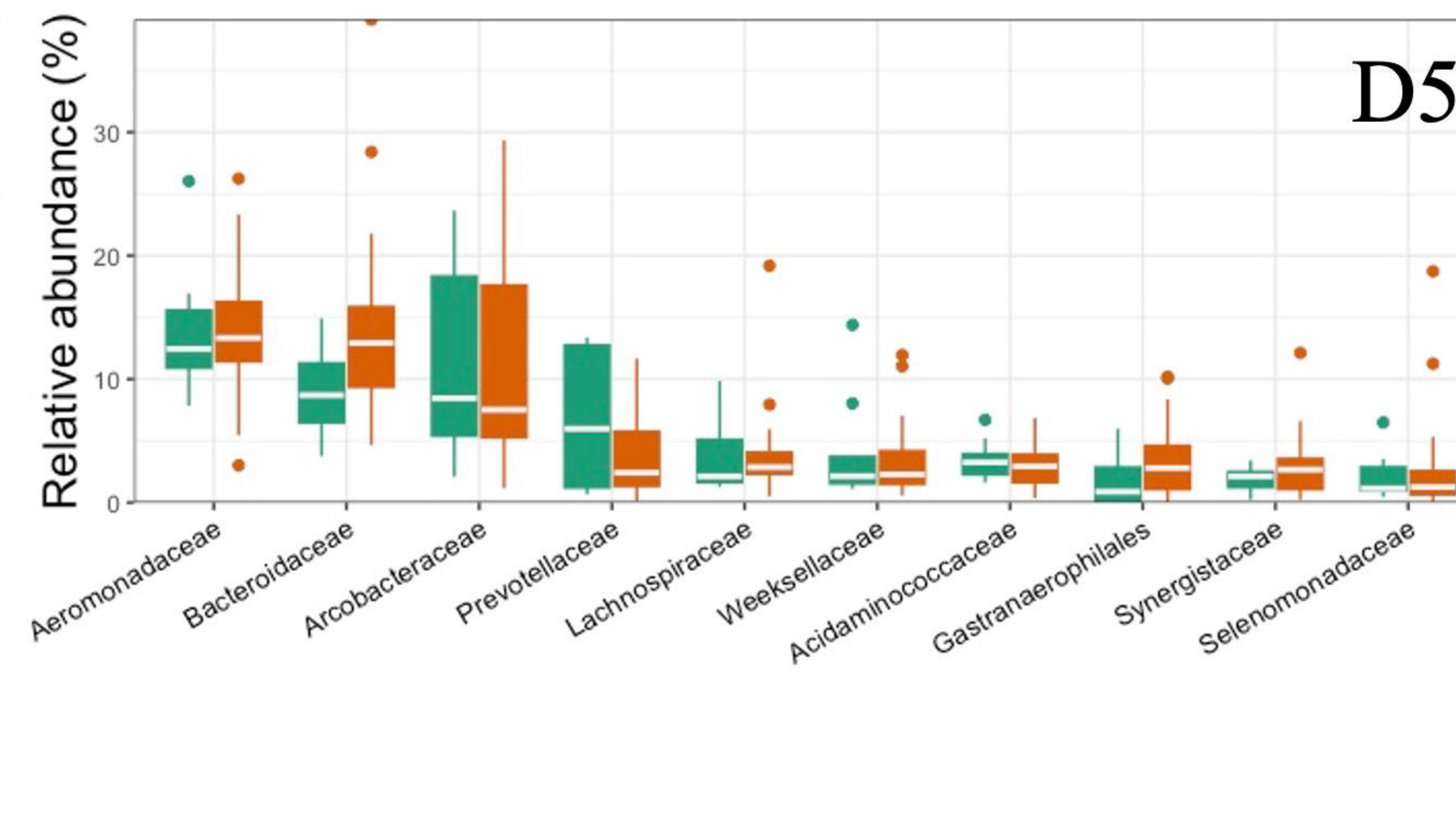
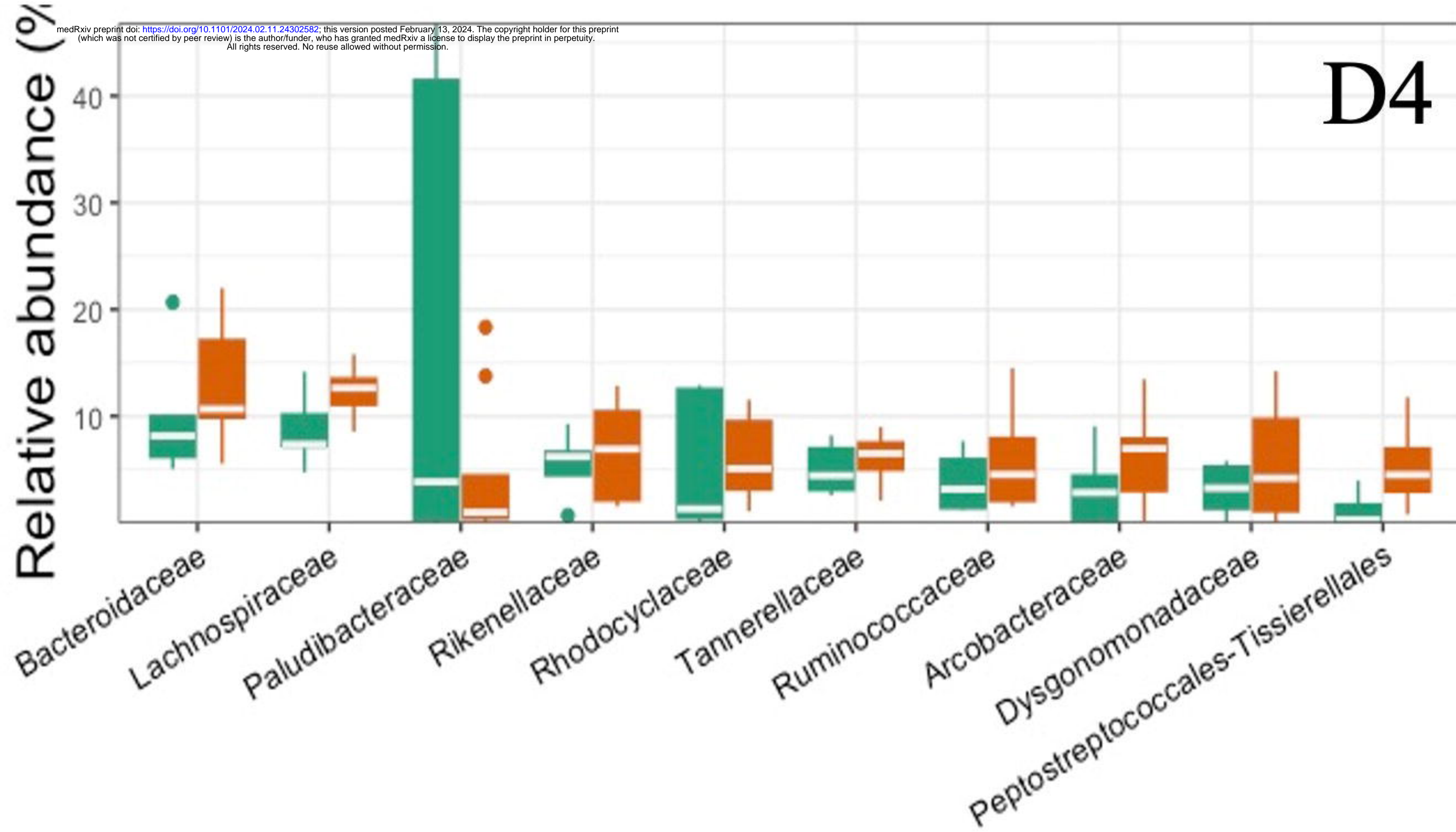
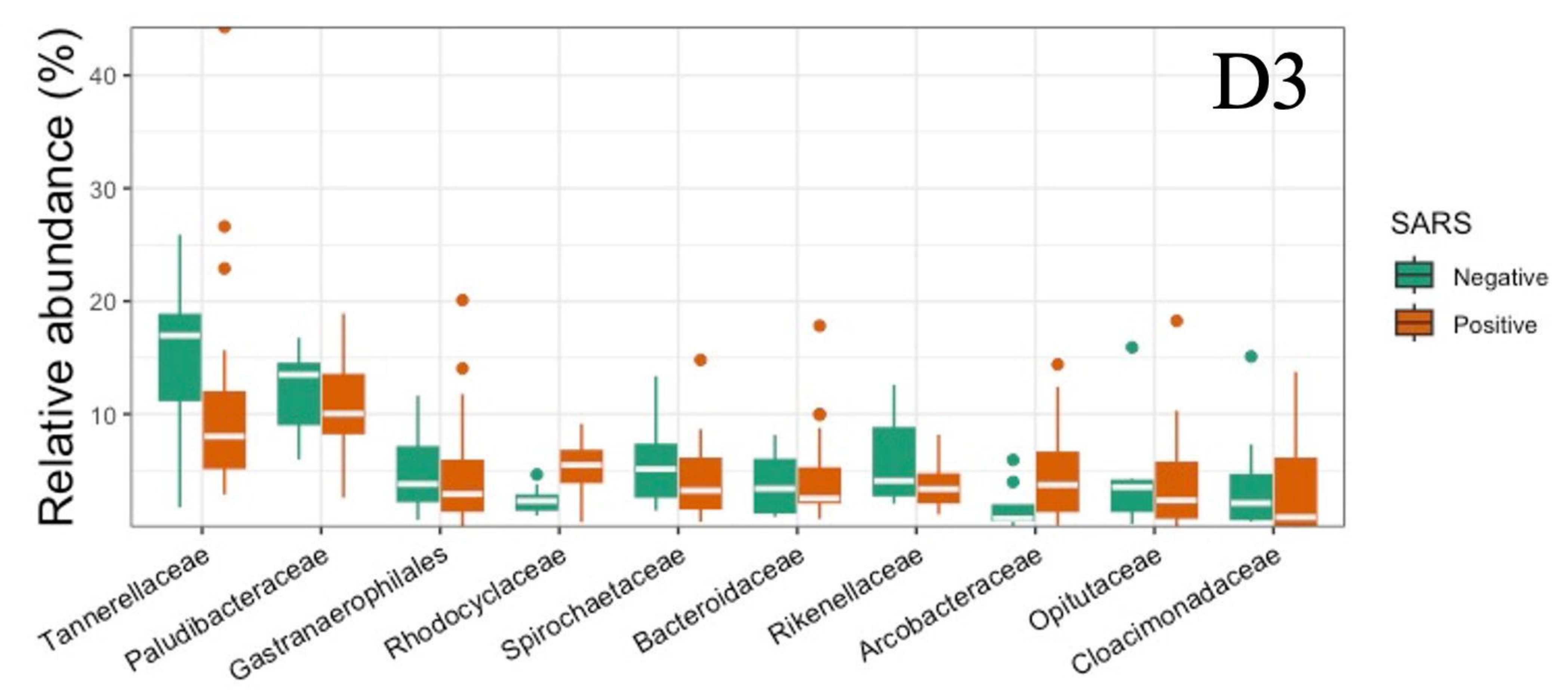
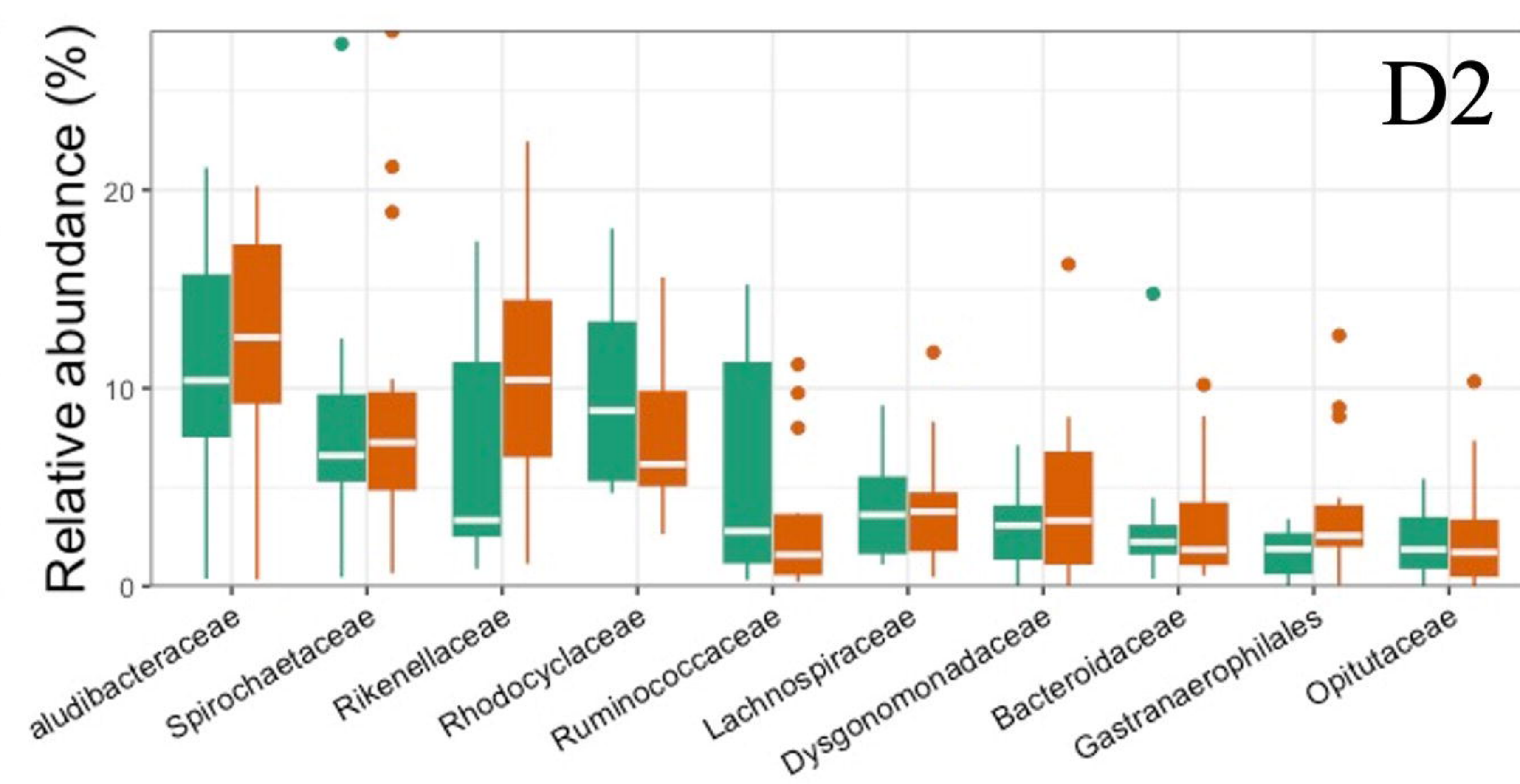
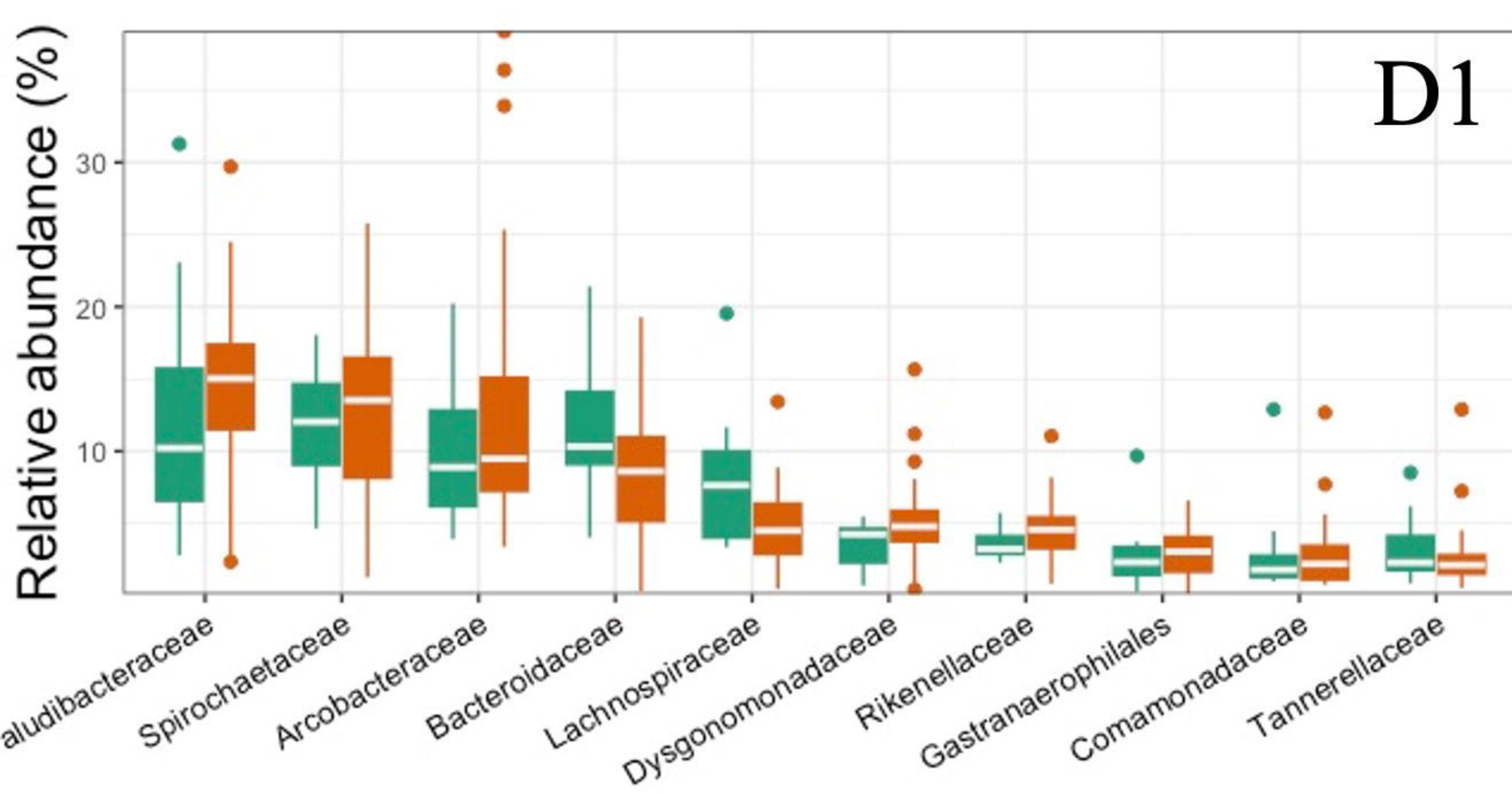


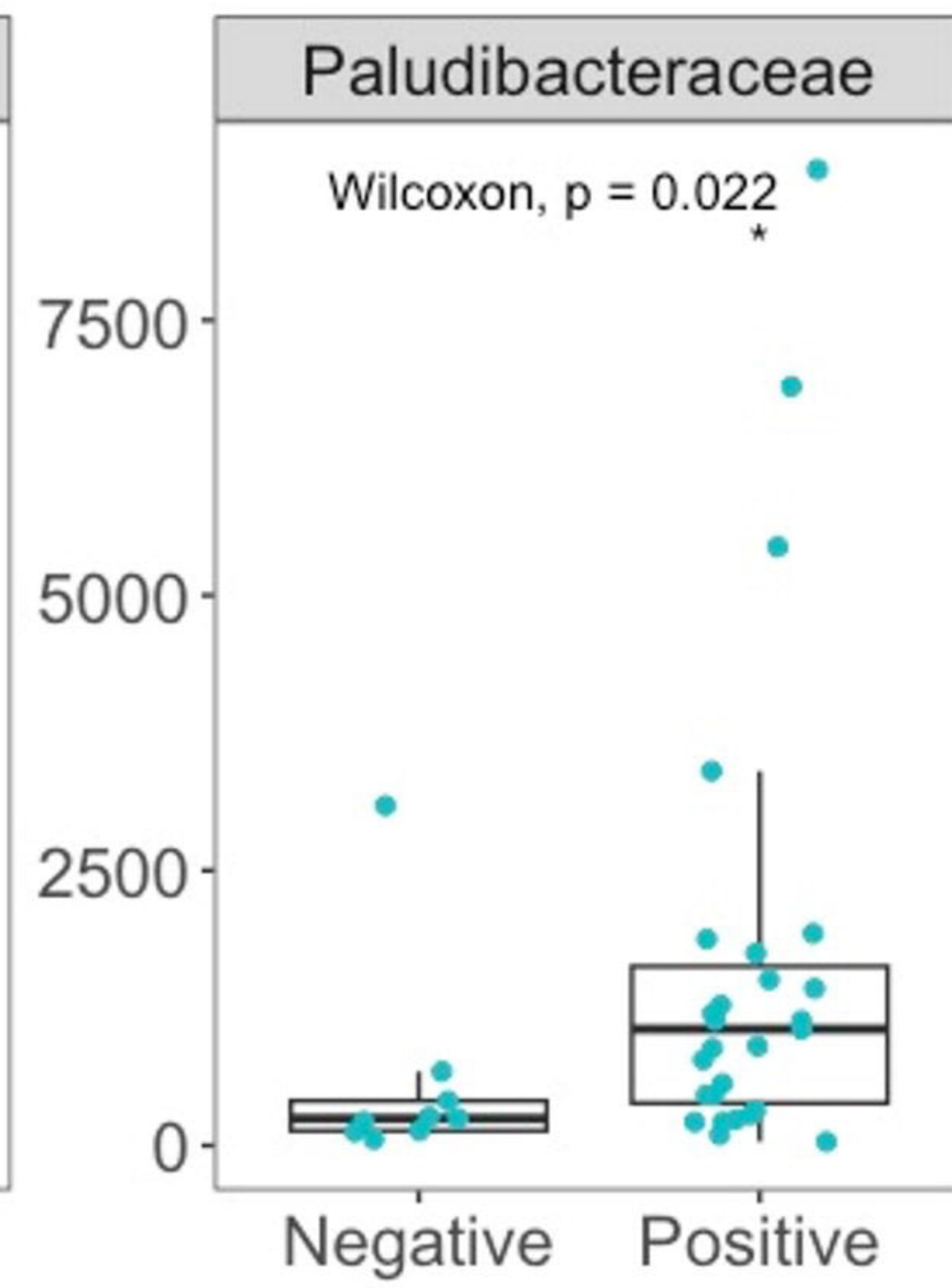
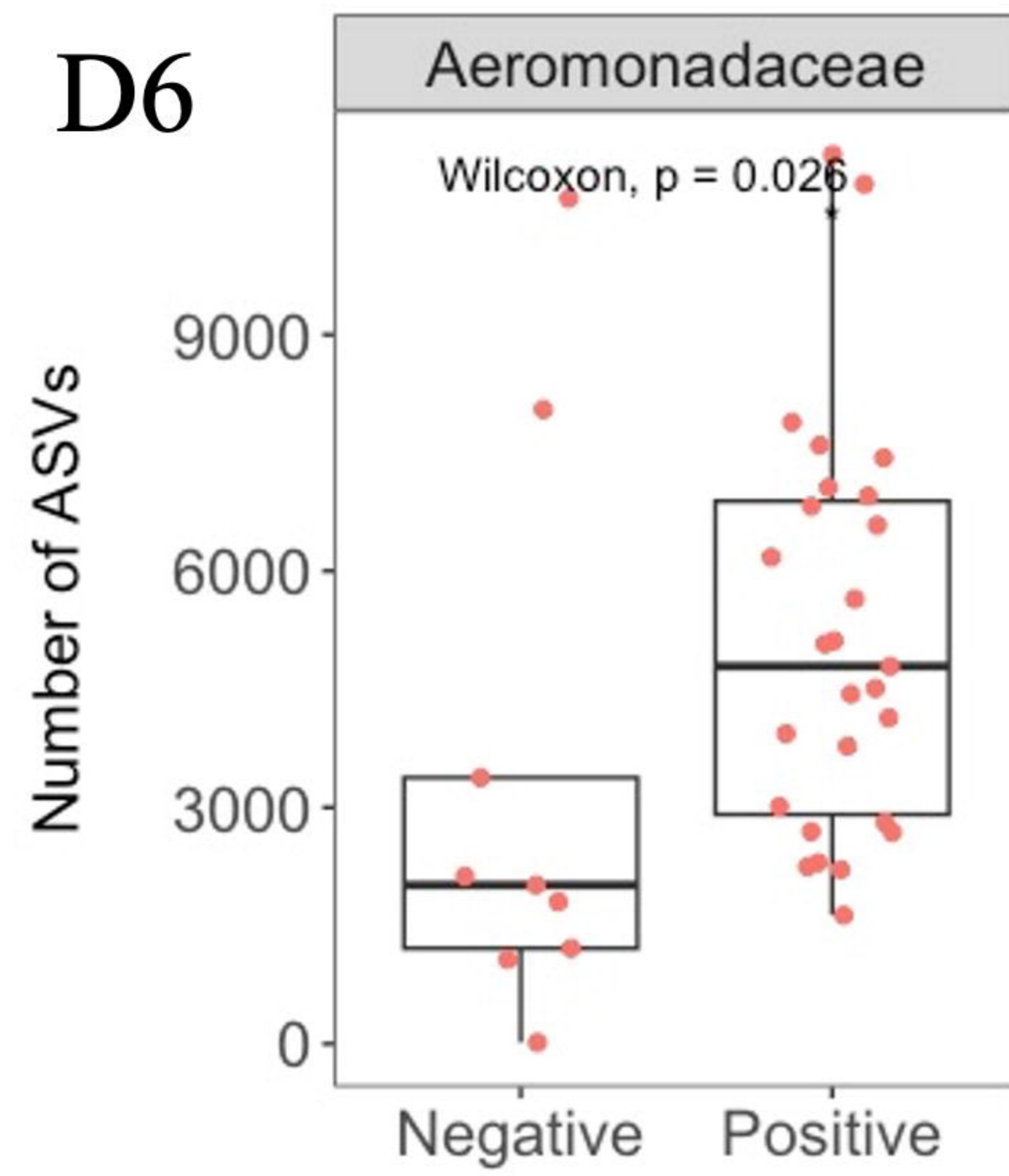
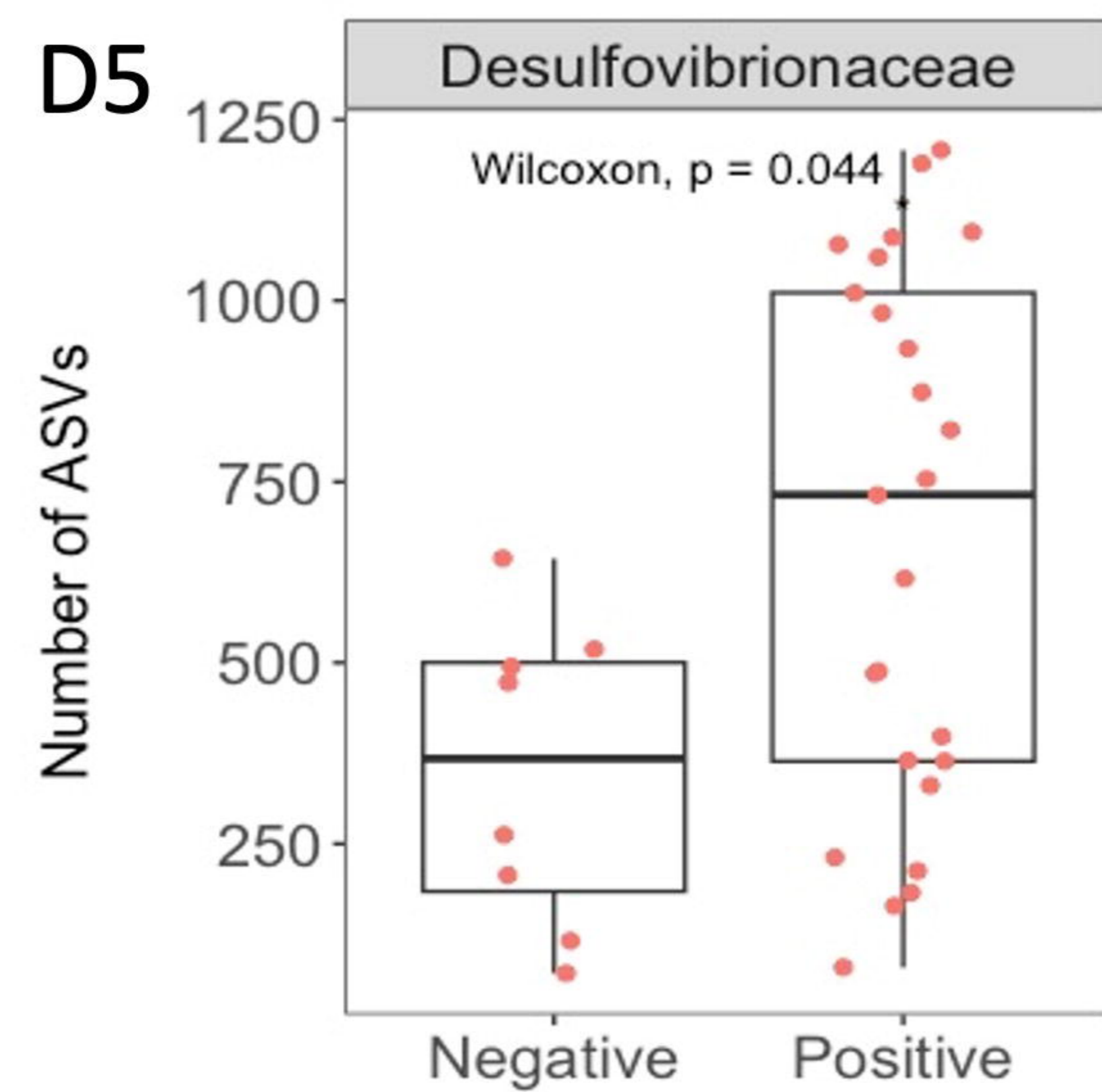
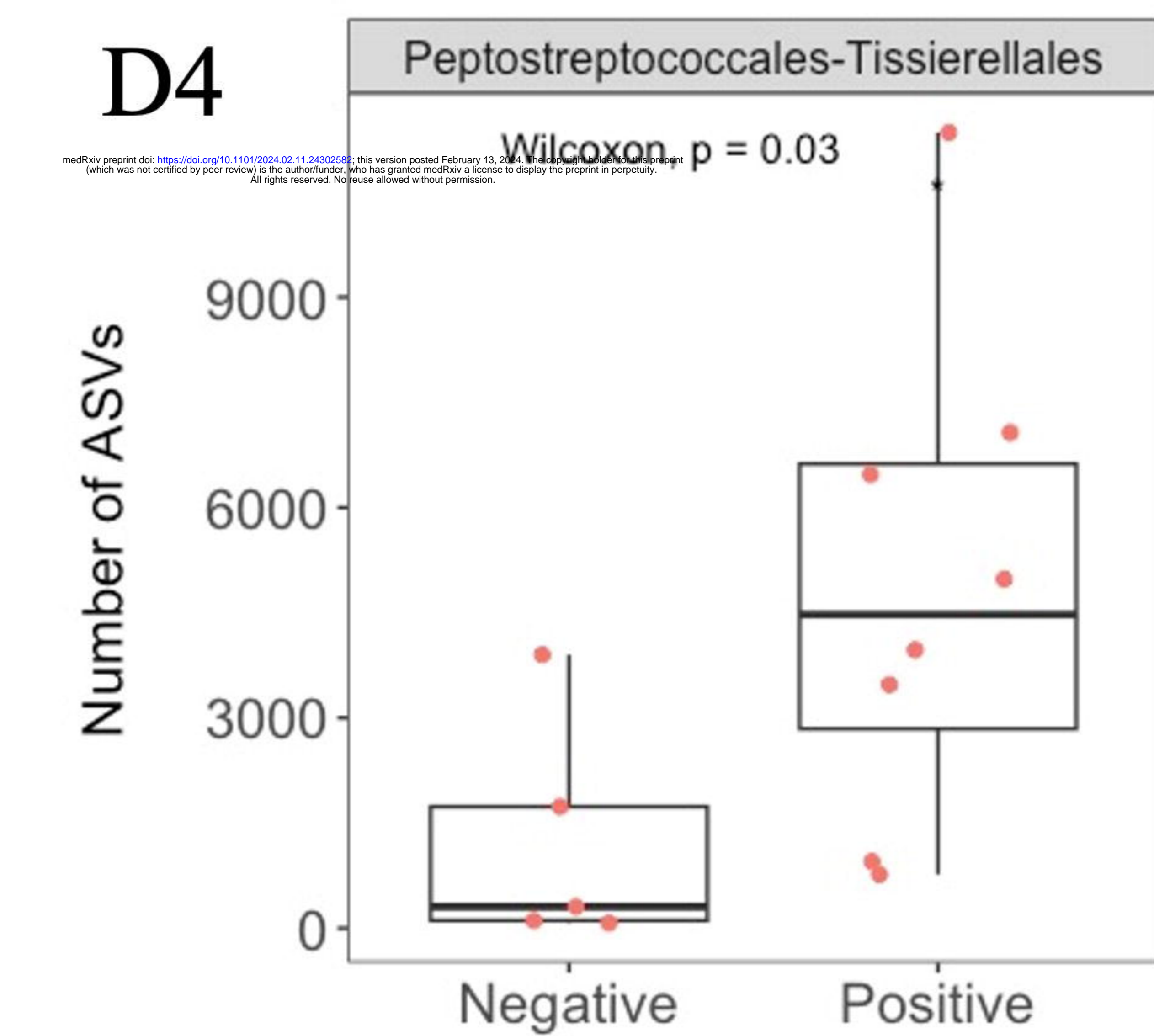
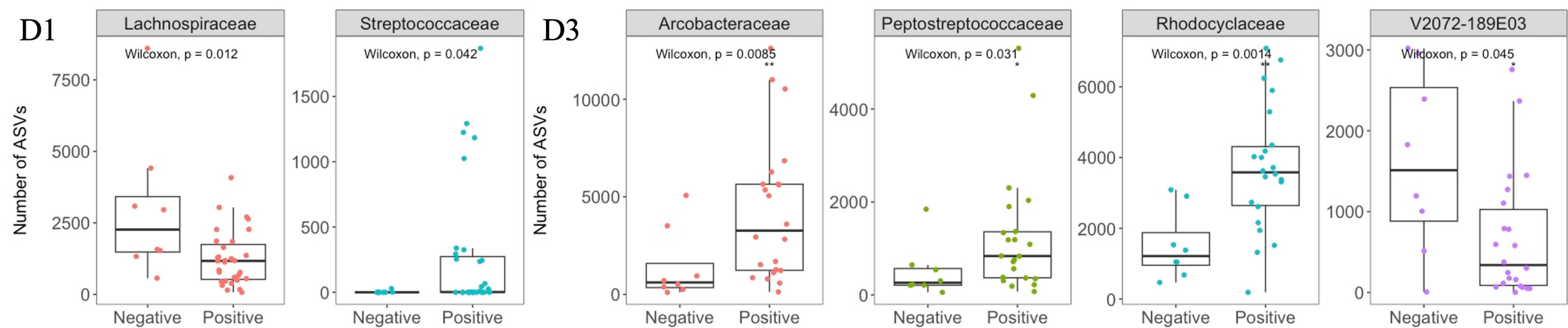


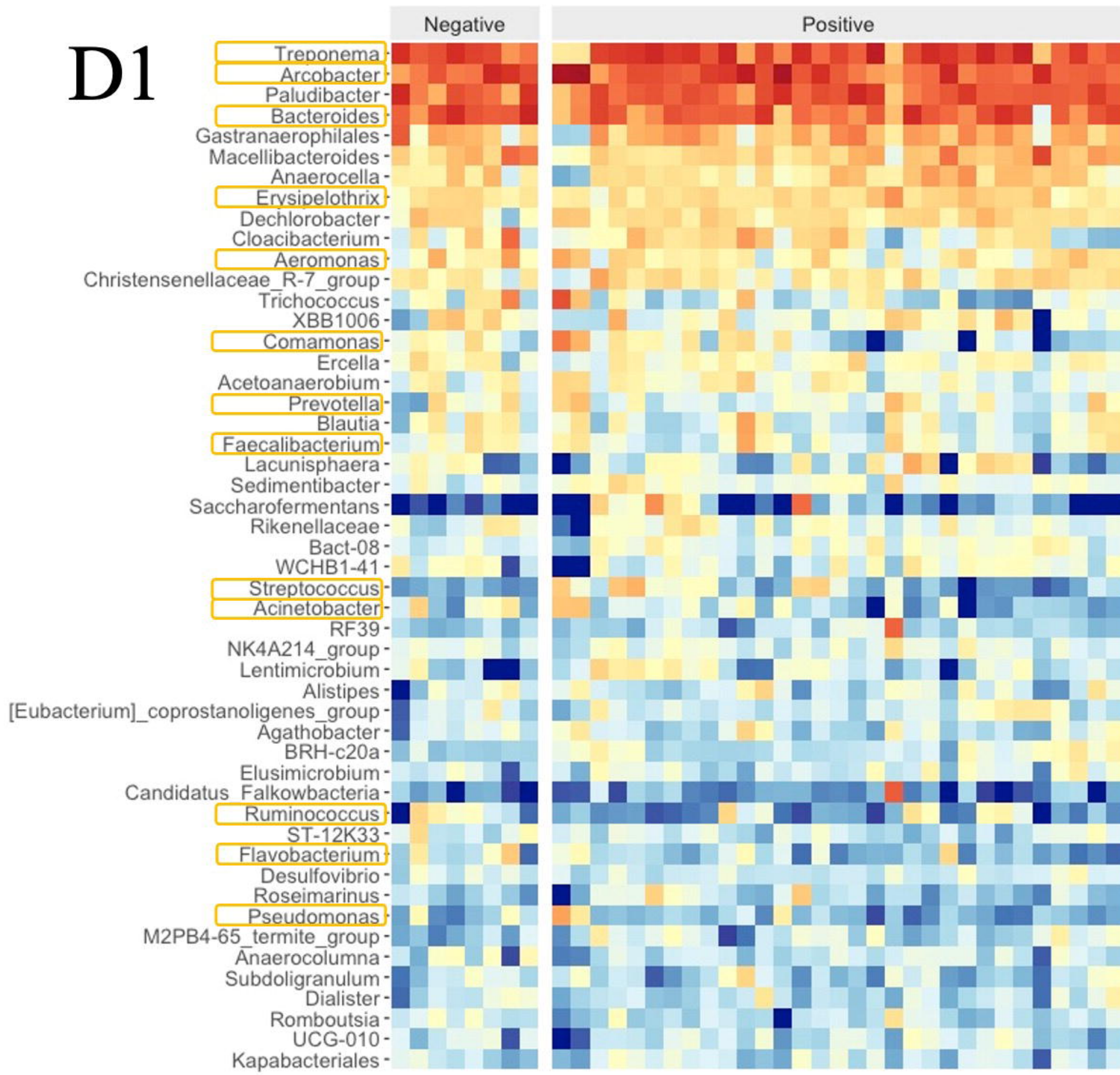
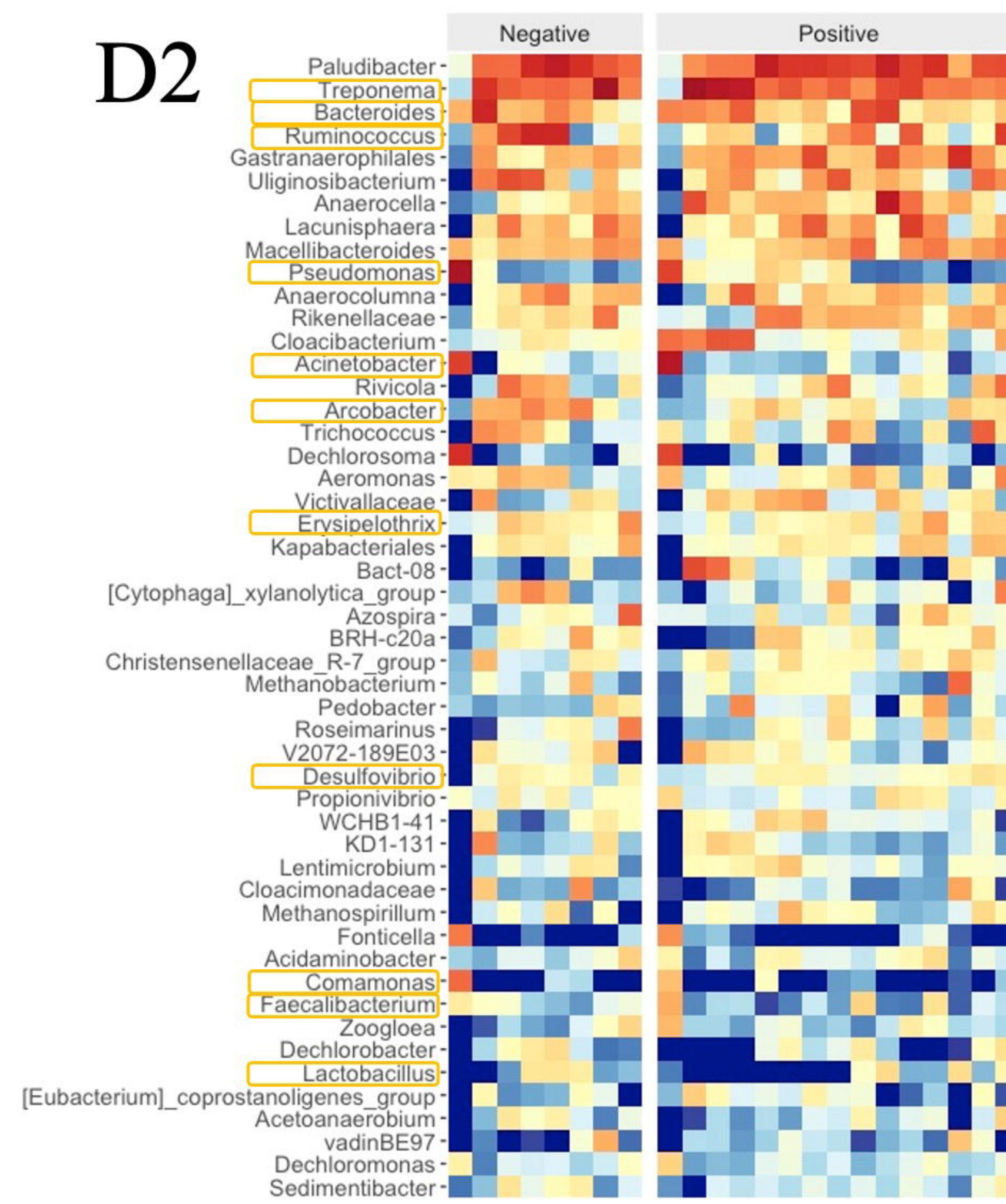
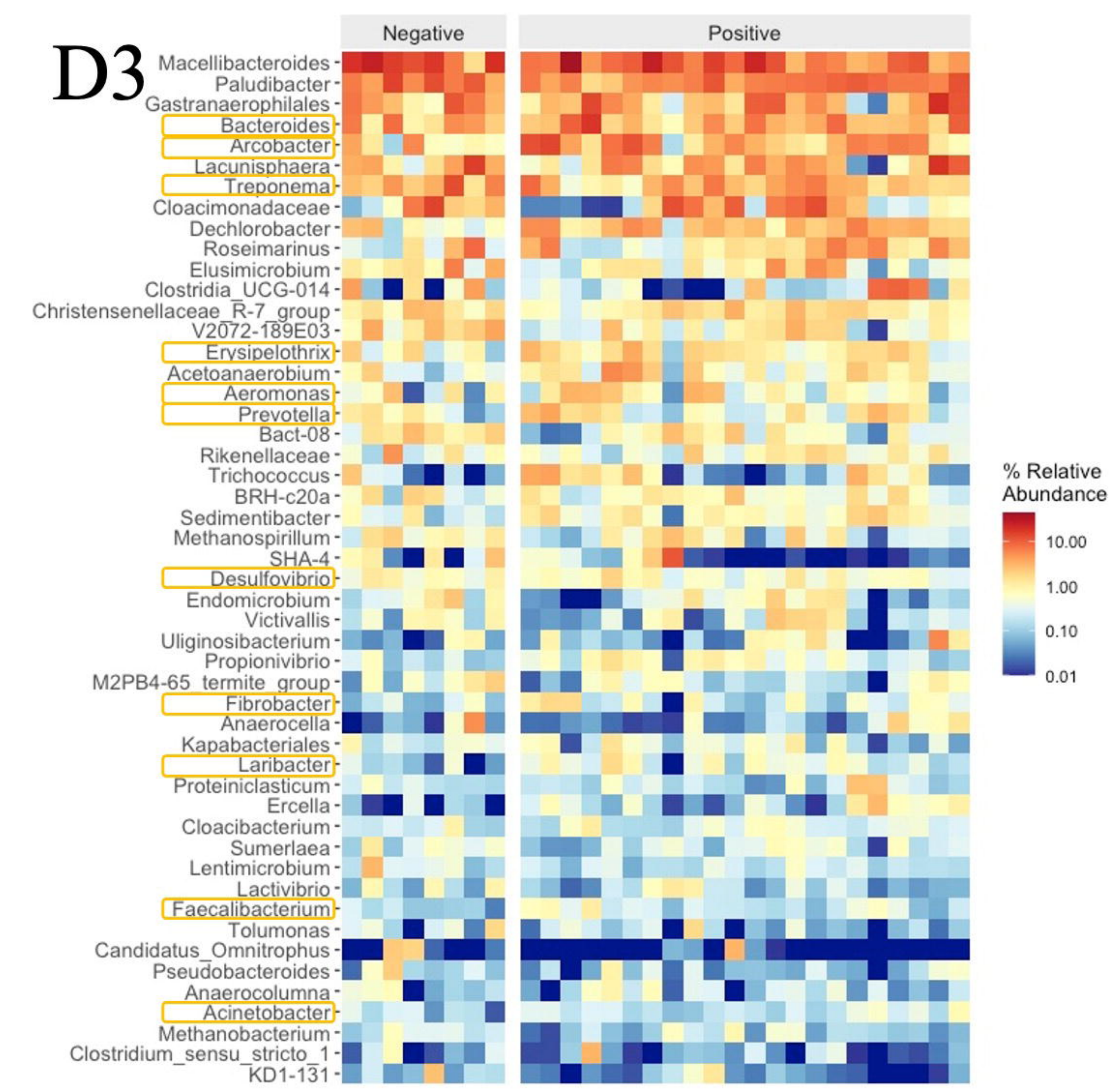
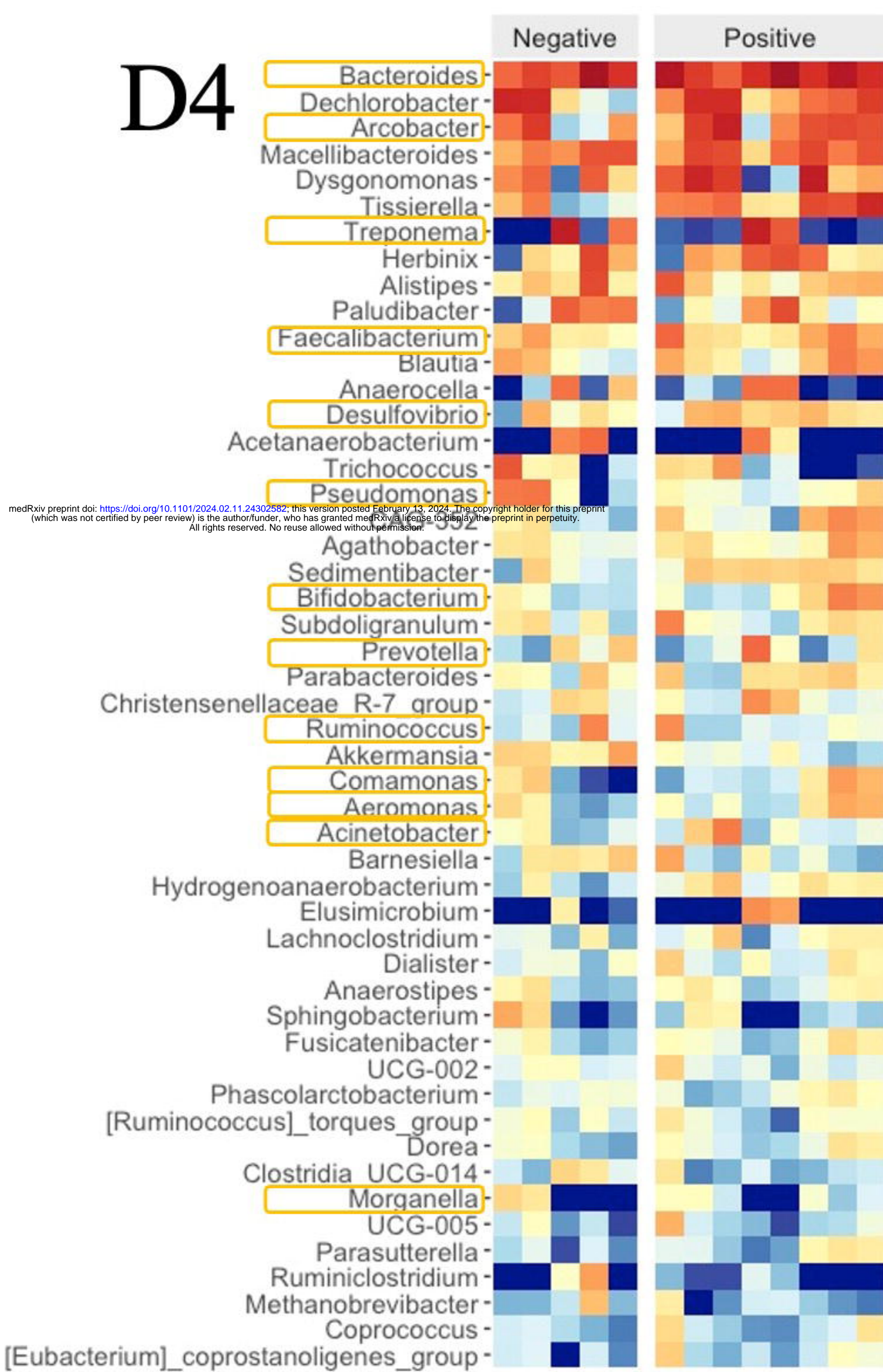
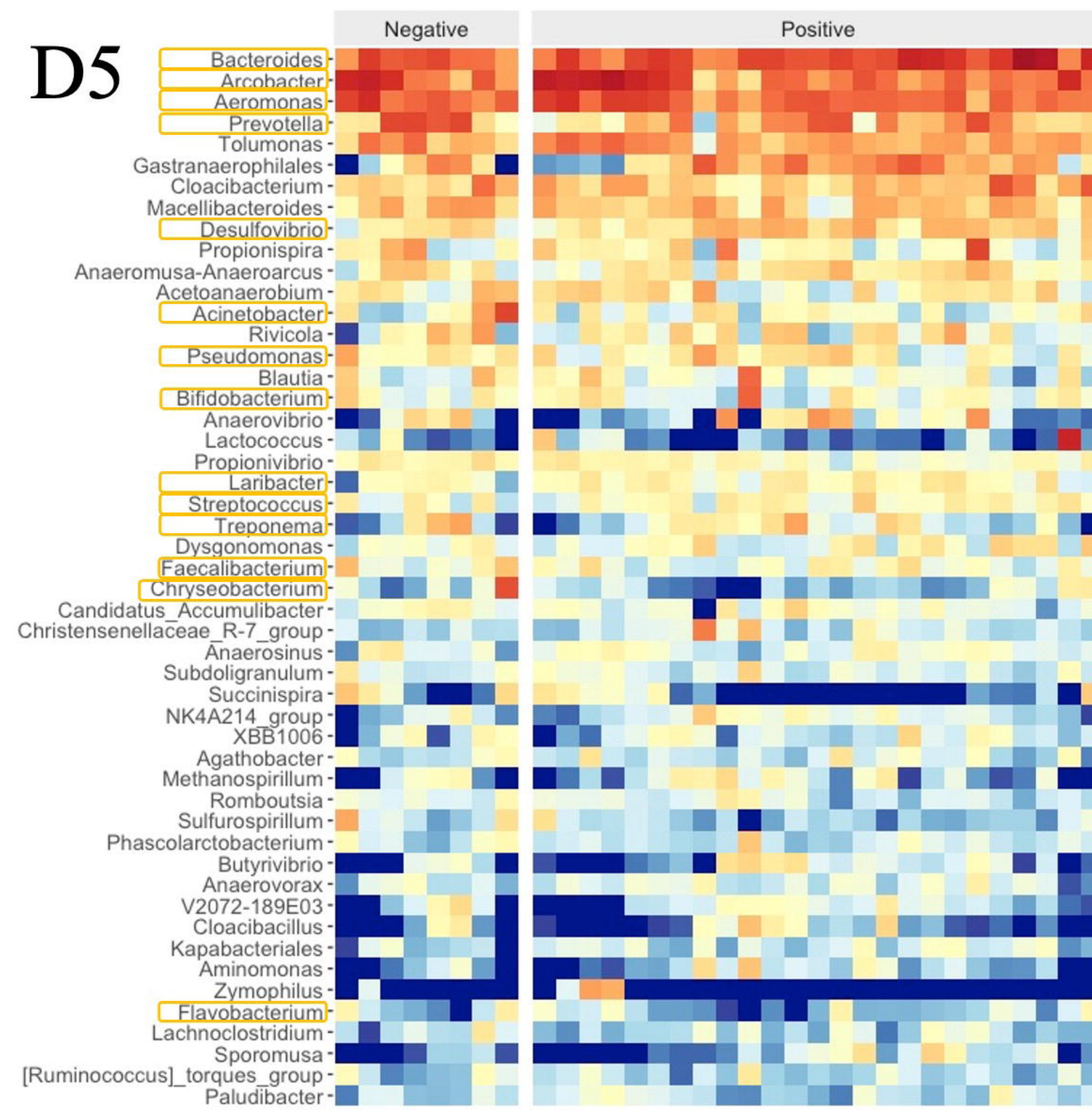
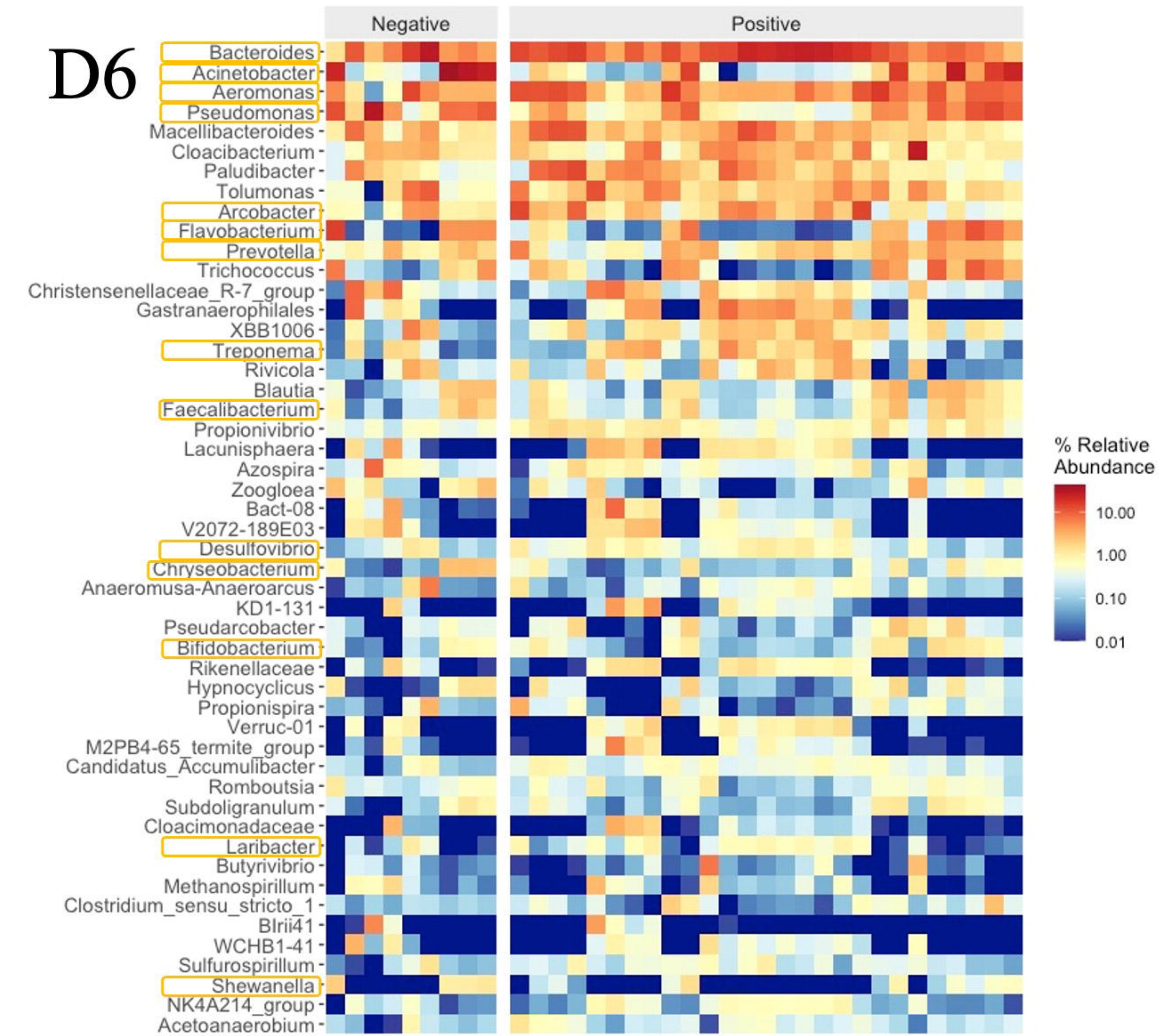




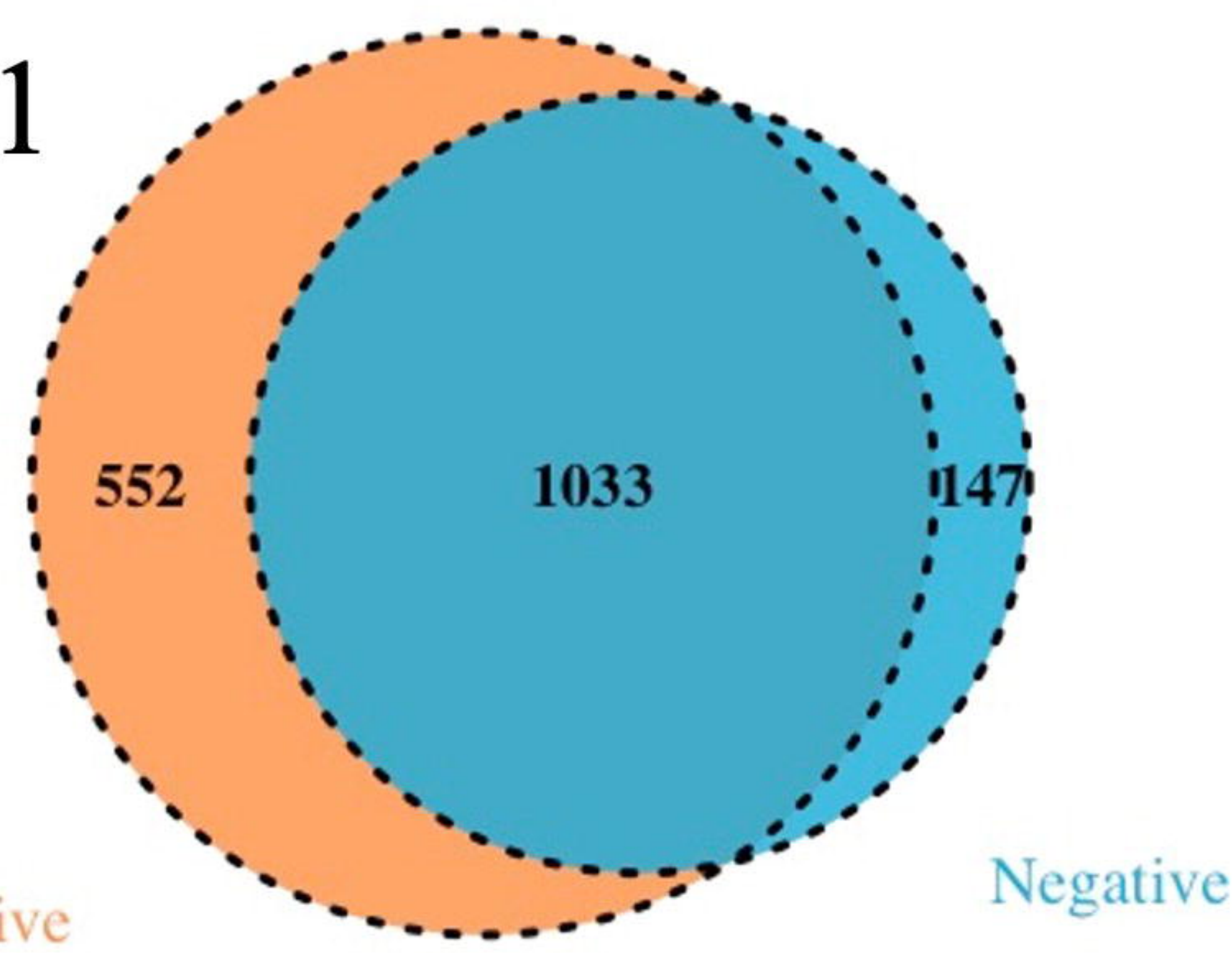




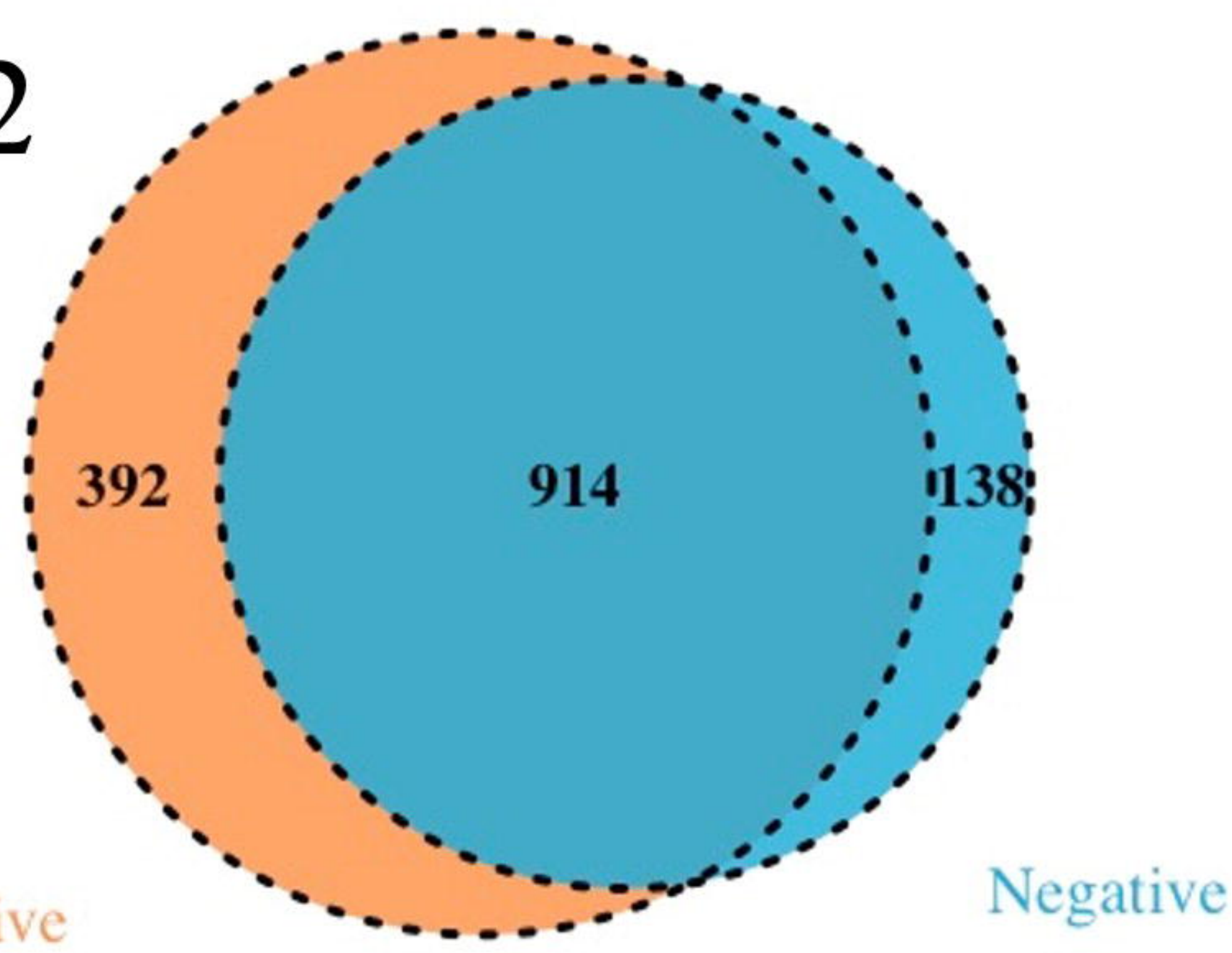


**D1****D2****D3****D4****D5****D6**

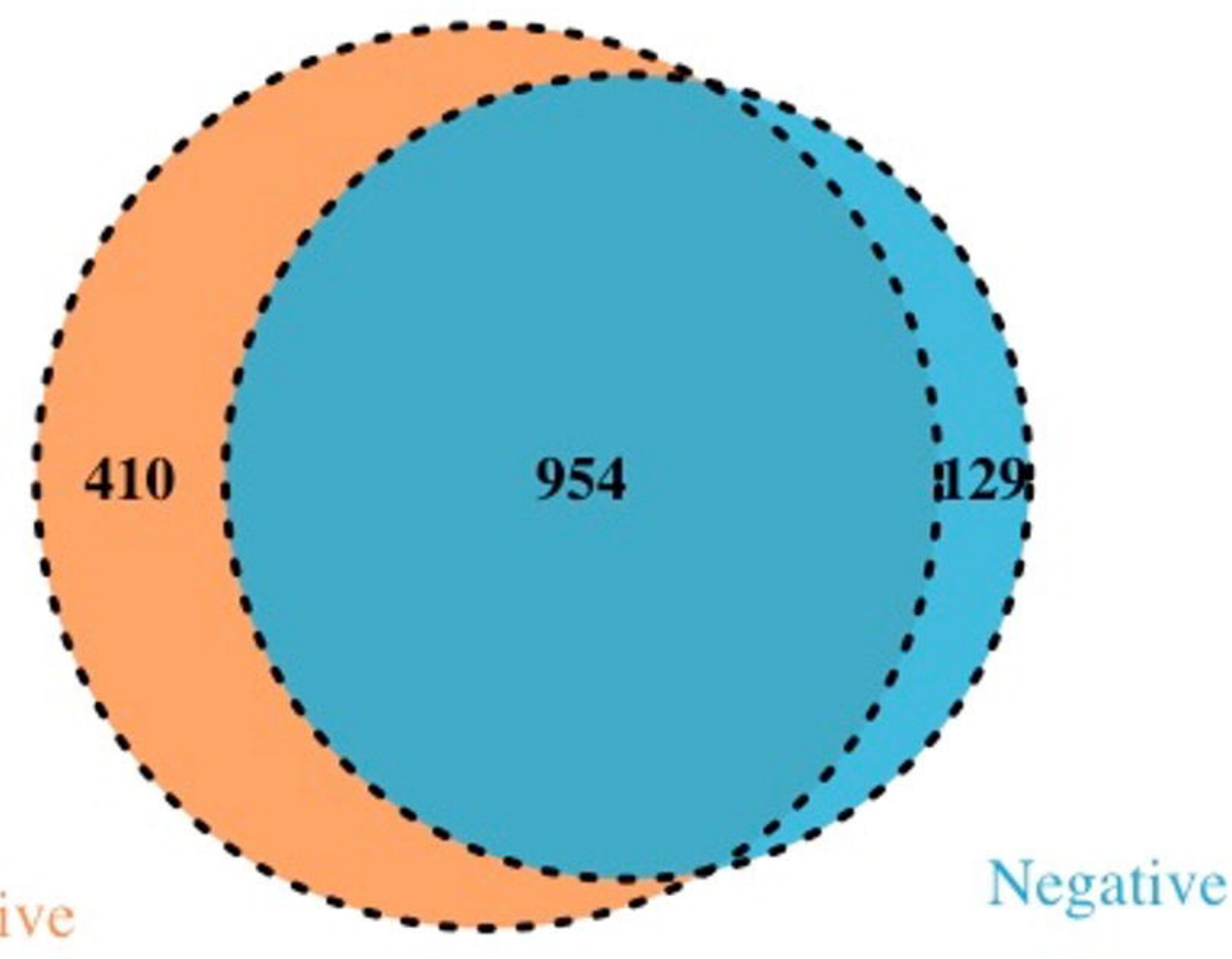
D1



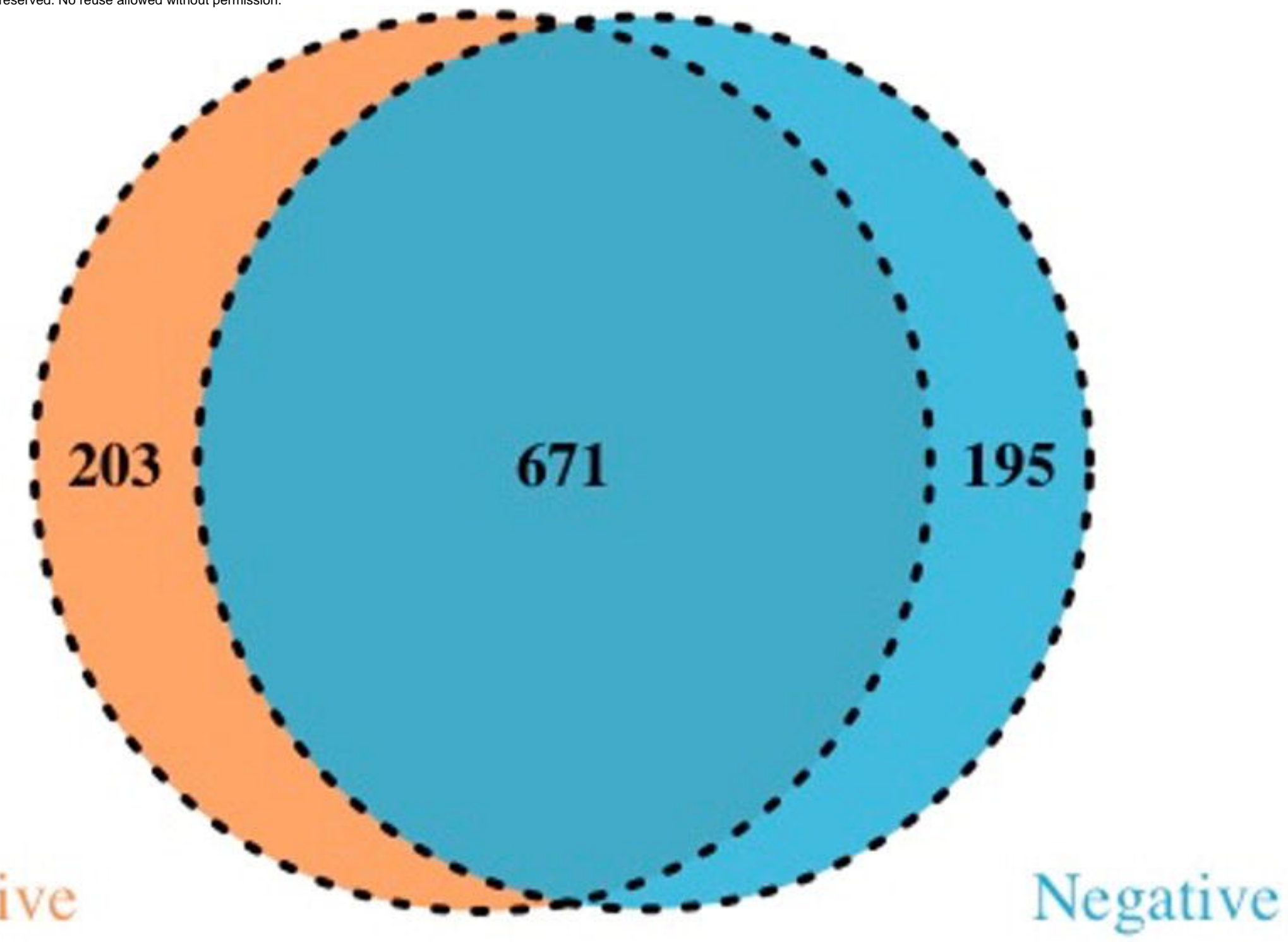
D2



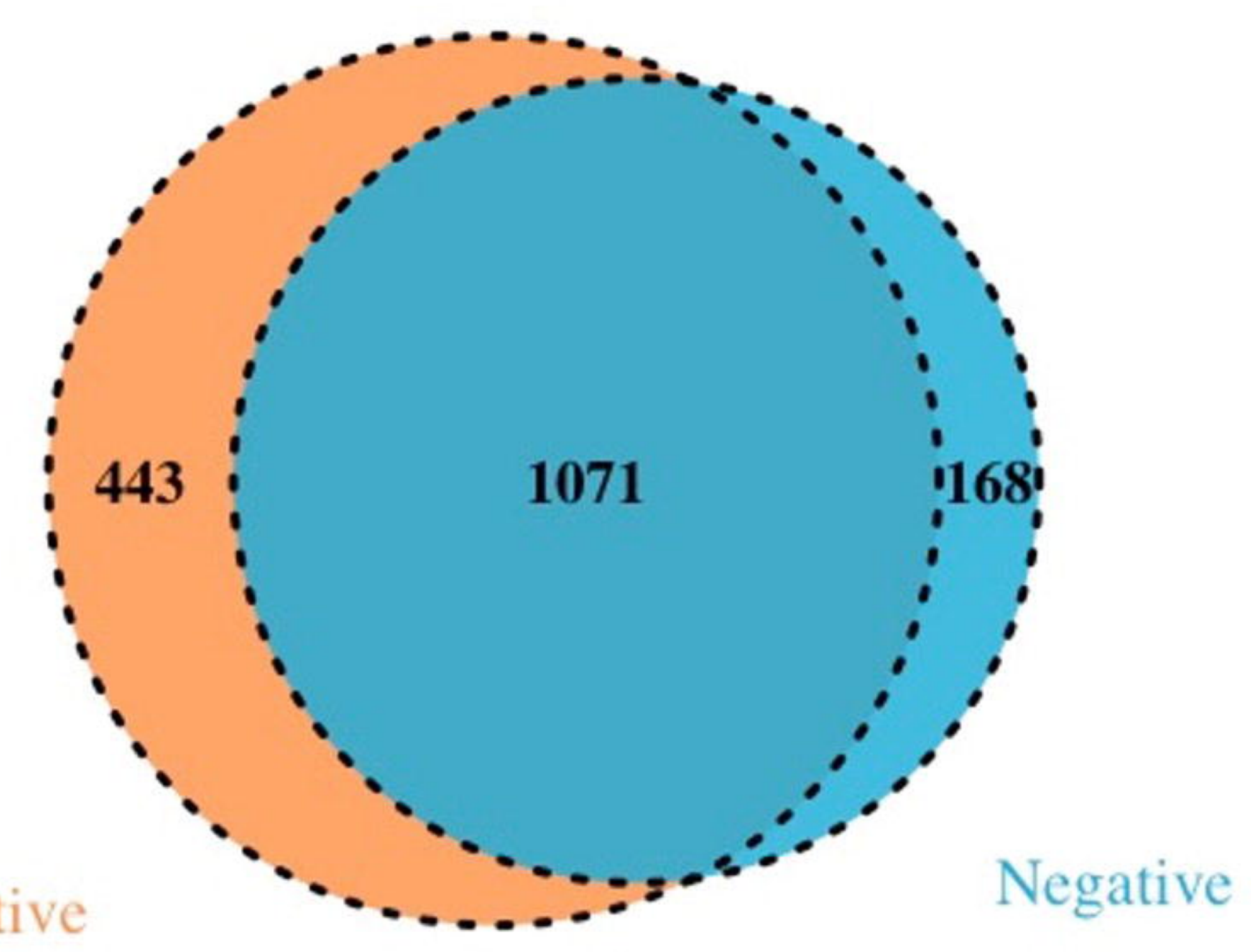
D3



D4



D5



D6

