

**Initial gut microbiota and response to antibiotic perturbation
influence *Clostridioides difficile* colonization in mice**

Sarah Tomkovich¹, Joshua M.A. Stough¹, Lucas Bishop¹, Patrick D. Schloss^{1†}

† To whom correspondence should be addressed: pschloss@umich.edu

¹ Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI 48109

1 Abstract

2 The microbiota plays a key role in determining susceptibility to *Clostridioides difficile* infections
3 (CDIs). However, much of the mechanistic work examining CDIs in mouse models use a single
4 university colony or vendor. We treated mice from 6 different colony sources (2 University of
5 Michigan colonies and 4 vendors) with a single clindamycin dose, followed by *C. difficile* challenge 1
6 day later and measured *C. difficile* colonization levels through 9 days post-infection. The microbiota
7 was profiled via 16S rRNA gene sequencing analysis to examine variation across colony sources
8 and alterations due to clindamycin treatment and *C. difficile* challenge. While all sources of mice
9 were colonized 1-day post-infection, variation in *C. difficile* colonization levels emerged from days
10 3-7 post-infection with some sources colonized with *C. difficile* for slightly longer and at higher levels.
11 We identified bacterial taxa with different relative abundances across colony sources throughout the
12 experiment, as well as taxa that were consistently impacted by clindamycin treatment in all sources
13 of mice. We created logistic regression models that successfully classified mice based on whether
14 they cleared *C. difficile* by 7 days post-infection using baseline, post-clindamycin, and post-infection
15 community composition data. After examining the taxa that were most important to the classification
16 models, we identified a subset of key taxa that varied across colony sources (*Bacteroides*,
17 *Deferrribacteraceae*), were altered by clindamycin (*Porphyromonadaceae*, *Ruminococcaceae*), or
18 both (*Enterobacteriaceae*, *Enterococcus*, *Bifidobacteriaceae*, *Coriobacteriaceae*, *Lachnospiraceae*,
19 and *Verrucomicrobiaceae*). These results suggest the response of the initial gut microbiota to
20 clindamycin treatment influences *C. difficile* 630 colonization dynamics.

21 Importance

22 *Clostridioides difficile* is a leading nosocomial infection. Although the microbiota has been
23 established as a key risk factor, there is variation in who becomes asymptotically colonized,
24 develops an infection, or has an infection with adverse outcomes. *C. difficile* infection (CDI) mouse
25 models are widely used to answer a variety of *C. difficile* pathogenesis questions. However, the
26 inter-individual variation between mice is less than what is observed in humans, particularly if just
27 one source of mice is used. In this study, we administered clindamycin to mice from 6 different
28 colony sources and challenged them with *C. difficile*. Interestingly, only a subset of the taxa that

29 vary across sources were associated with how long *C. difficile* was able to colonize. Future studies
30 examining the interplay between the microbiota and *C. difficile* should consider using mice from
31 multiple sources to narrow down the microbes driving the observed phenotypes and reflect human
32 interindividual variation.

33 **Introduction**

34 Antibiotics are a clear risk factor for *Clostridioides difficile* infections (CDIs), but there is variation in
35 who goes on to develop severe or recurrent CDIs after exposure (1, 2). Additionally, asymptomatic
36 colonization, where *C. difficile* is detectable, but symptoms are absent has been documented
37 in infants and adults (3, 4). The intestinal microbiome has been implicated in asymptomatic
38 colonization (5, 6), susceptibility to CDIs (7), and adverse CDI outcomes (9–12).

39 Mouse models of CDIs have been a great tool for understanding *C. difficile* pathogenesis (13). The
40 number of CDI mouse model studies has grown substantially since Chen et al. published their
41 C57BL/6 model in 2008, which disrupted the gut microbiota with antibiotics to enable *C. difficile*
42 colonization and symptoms such as diarrhea and weight loss (14). CDI mouse models have been
43 used to examine translationally relevant questions regarding *C. difficile*, including the role of the
44 microbiota and efficacy of potential therapeutics for treating CDIs (15). However, microbiome
45 variation between lab mice is much less than the variation observed between humans (16, 17).
46 Additionally, studying the contribution of the microbiota to a particular disease phenotype in one set
47 of lab mice after the same perturbation could yield a number of findings of which only a fraction
48 may be driving the phenotype.

49 In the past, our group has attempted to introduce more microbiome variation into the CDI mouse
50 model by using a variety of antibiotic treatments (18–21). An alternative approach to maximize
51 microbiome variation is to use mice from multiple sources (22, 23). Microbiome differences between
52 different mouse vendors have been well documented and shown to influence susceptibility to a
53 variety of diseases (24, 25), including enteric infections (22, 23, 26–30). Additionally, different
54 research groups have observed different CDI outcomes in mice despite using similar models and
55 the microbiome has been proposed as one factor potentially mediating CDI susceptibility and
56 outcomes (13, 18, 21, 31–33). Here we examined how variations in the baseline microbiome
57 and responses to clindamycin treatment in C57BL/6 mice from six different sources influenced
58 susceptibility to *C. difficile* colonization and the time needed to clear the infection.

59 **Results**

60 **Clindamycin treatment renders all mice susceptible to *C. difficile* 630 colonization**
61 **regardless of colony source.** To test how the microbiotas of mice from different colony sources
62 impact colonization dynamics after clindamycin exposure, we utilized C57BL/6 mice from 6 different
63 sources: two colonies from the University of Michigan (the Young and Schloss lab colonies), the
64 Jackson Laboratory, Charles River Laboratories, Taconic Biosciences, and Envigo (which was
65 formerly Harlan). These 4 vendors were chosen because they represent commonly used vendors
66 for CDI studies in mice (26, 34–40). After a 13-day acclimation period for the mice ordered from
67 vendors, all mice were treated with 10 mg/kg clindamycin via intraperitoneal injection and one day
68 later challenged with 10^3 *C. difficile* 630 spores (Fig. 1A). Clindamycin was chosen because we
69 have previously demonstrated mice are rendered susceptible, but consistently cleared the CDI
70 within 9 days (21, 41), clindamycin is frequently implicated with human CDIs (42), and is also part of
71 the antibiotic treatment for the frequently cited 2008 CDI mouse model (14). The day after infection,
72 *C. difficile* was detectable in all mice at a similar level (median CFU range: 2.2e+07-1.3e+08; P_{FDR}
73 = 0.15), indicating clindamycin rendered all mice susceptible regardless of colony source (Fig.
74 1B). Interestingly, variation in *C. difficile* CFU levels across sources of mice emerged from days
75 3-7 post-infection (all $P_{FDR} \leq 0.019$; Fig. 1B and Table S1), suggesting mouse colony source is
76 associated with *C. difficile* clearance. We conducted two experiments approximately 3 months
77 apart and while the colonization dynamics were similar across most sources of mice, there was
78 some variation between the 2 experiments, particularly for the Schloss and Envigo mice (Fig.
79 S1A-B). Although *C. difficile* 630 causes mild symptoms in mice compared to other *C. difficile*
80 strains (43), we also saw that weight change significantly varied across sources of mice with the
81 most weight lost two days post-infection (Fig. 1C and Table S2). Importantly, there was also one
82 Jackson and one Envigo mouse that died between 1- and 3-days post-infection during the second
83 experiment. Interestingly, mice ordered from Jackson, Taconic, and Envigo tended to lose more
84 weight (although there was variation between experiments with Schloss and Envigo mice), have
85 higher *C. difficile* CFU levels and take longer to clear the infection compared to the other sources of
86 mice, which was particularly evident 7 days post-infection (Fig. 1B-C, Fig. S2C-D), when 57% of
87 the mice were still colonized with *C. difficile* (Fig. S1E). By 9 days post-infection the majority of the

88 mice from all sources had cleared *C. difficile* (Fig. S1C) with the exception of 1 Taconic mouse from
89 the first experiment and 2 Envigo mice from the second experiment. Thus, clindamycin rendered all
90 mice susceptible to *C. difficile* 630 colonization, regardless of colony source, but variation across
91 sources emerged with 3 out of 6 sources taking longer to clear *C. difficile*.

92 **Bacterial communities consistently vary across mouse colony sources despite antibiotic**
93 **and infection perturbations.** Given the well known variation in mouse microbiomes across
94 vendors and university colonies (25), we hypothesized that the variation in *C. difficile* clearance
95 could be explained by microbiota variation across the 6 sources. We used 16S rRNA gene
96 sequencing to characterize the fecal bacterial communities from the mice over the course of
97 the experiment. Since antibiotics and other risk factors of CDIs are associated with decreased
98 microbiota diversity (44), we first examined alpha diversity measures across the 6 sources of
99 mice. Examining the bacterial communities at baseline, prior to clindamycin treatment there was a
100 significant difference in the number of observed OTUs ($P_{\text{FDR}} = 0.03$), but not Shannon diversity
101 index ($P_{\text{FDR}} = 0.052$) across sources of mice (Fig. 2A-B and Table S3). As expected, clindamycin
102 treatment decreased richness and Shannon diversity across all sources of mice, and communities
103 started to recover 1 day post-infection (Fig. 2C-D). Interestingly, significant differences in diversity
104 metrics across sources ($P_{\text{FDR}} < 0.05$) emerged after both clindamycin and *C. difficile* infection, with
105 Charles River mice having higher richness and Shannon Diversity than most of the other sources
106 (Fig 2C-F and Table S4). While Charles River mice had more diverse microbiotas, Young and
107 Schloss lab mice were also able to clear *C. difficile* faster, suggesting microbiota diversity alone
108 does not explain the observed variation in *C. difficile* colonization across vendors.

109 Next, we compared the bacterial communities from the 6 colonies over the course of the experiment
110 using principal coordinate analysis (PCoA) of the Theta YC distances. Permutational multivariate
111 analysis of variance (PERMANOVA) analysis revealed colony source was the major factor explaining
112 the observed variation across fecal communities ($R^2 = 0.35$, $P = 0.0001$) followed by interactions
113 between cage and day of the experiment (Movie S1 and Table S5). Since, the majority of the
114 perturbations happened over the initial days of the experiment, we decided to focus on the bacterial
115 communities at baseline (day -1), after clindamycin treatment (day 0), and post-infection (day 1). For
116 all 3 timepoints, source and the interaction with cage significantly explained most of the observed

117 community variation (combined $R^2 = 0.90, 0.99, 0.88$, respectively; $P = 0.0001$; Fig. 3 and Table
118 S6). We also compared baseline communities across the 2 experiments, and found experiment
119 and cage significantly explained the observed variation only for the Schloss and Young lab mouse
120 colonies (Fig. S2A-B and Table S7). However, most of the vendors also clustered by experiment
121 (Fig. S2C-D, F), suggesting there was some community variation between the 2 experiments
122 within each source. Thus, mouse colony source was the factor that explained the most variation
123 observed in the bacterial communities. Importantly with the exception of the 2 University colonies,
124 the community of each source clustered apart from one another suggesting each community had a
125 unique response to clindamycin treatment and *C. difficile* challenge.

126 Since there was some variation in microbiota communities between experiments at baseline, we
127 next looked at how similar the communities were within the same source and between sources in
128 response to clindamycin treatment and *C. difficile* challenge (Fig. 4) . The baseline communities
129 varied most between experiments for Schloss, Young, and Envigo mice and variation between
130 sources of mice was high (Fig. 4A). Clindamycin treatment reduced variation between experiments
131 within Schloss, Young and Jackson mice and some of the variation between sources diminished,
132 particularly for the University of Michigan and Charles River mice (Fig. 4B). Post-infection, the
133 community variation started to increase within sources of mice and variation between sources of
134 mice started to increase towards previous levels (Fig. 4C). By using mice from multiple sources
135 we were able to increase the number of microbiota communities we tested with the clindamycin *C.*
136 *difficile* colonization mouse model.

137 After finding differences at the community level, we next identified the taxa that varied across
138 sources of mice over the initial days of the experiment. We examined bacterial relative abundances
139 at the operational taxonomic unit (OTU) and family levels, expecting the number of differences to be
140 reduced at the family level due to the nature of bacterial taxonomy (45). Focusing on the baseline
141 communities first, there were 268 OTUs and 20 families (Tables S8-9) with relative abundances
142 that varied across colony sources. Clindamycin treatment reduced the number of taxa with relative
143 abundances that varied across sources to 18 OTUs and 10 families (Tables S8-9). After *C. difficile*
144 challenge, there were 44 OTUs and 18 families (Tables S8-9) with significantly different relative
145 abundances across sources, as the communities started to recover from antibiotic treatment. In

146 spite of the experimental perturbations that occurred during these 3 timepoints, there were 12
147 OTUs (Fig 5A-C) and 8 families with relative abundances that consistently varied across colony
148 sources (Fig. 5D-F). Importantly, some of the taxa that consistently varied across sources also
149 shifted with clindamycin treatment. For example, *Proteus* increased after clindamycin treatment, but
150 only in Taconic mice. *Enterococcus* was primarily found only in mice purchased from commercial
151 vendors and also increased after clindamycin treatment. In summary, mouse bacterial communities
152 significantly varied according to colony source throughout the course of the experiment and a
153 consistent subset of bacterial taxa remained different across sources regardless of clindamycin
154 and *C. difficile* challenge.

155 **Clindamycin treatment alters a subset of taxa that were found in all colony sources.**
156 Although there were bacteria that consistently varied across colony sources, we also wanted to
157 identify the bacteria that shifted after clindamycin treatment, regardless of colony source. By
158 analyzing all mice that had sequence data from fecal samples collected at baseline and after
159 clindamycin treatment, we identified 153 OTUs and 18 families that were altered after clindamycin
160 treatment (Fig. 6 and Tables S10-11). Interestingly, when we compared the list of significant
161 clindamycin impacted bacteria with the bacteria that consistently varied across groups over the
162 initial 3 timepoints of our experiment, we found 3 OTUs (*Lachnospiraceae* (OTU 130), *Lactobacillus*
163 (OTU 6), *Enterococcus* (OTU 23)) and 3 families (*Porphyromonadaceae*, *Enterococcaceae*,
164 *Lachnospiraceae*) overlapped (Fig. 5, Fig. 6C-D). These findings demonstrate that clindamycin has
165 a consistent impact on the fecal bacterial communities of mice from all colony sources and only a
166 subset of the taxa also varied across colony sources.

167 **Source-specific and clindamycin impacted bacteria distinguish *C. difficile* colonization
168 status in mice.** After identifying taxa that varied by colony source, changed after clindamycin
169 treatment, or both, we next wanted to determine which taxa were influencing the variation in *C.*
170 *difficile* colonization at day 7 (Fig. 1D, Fig. S1C). We trained L2-regularized logistic regression
171 models with input bacterial community data from the baseline, post-clindamycin, and post-infection
172 timepoints of the experiment to predict *C. difficile* colonization status on day 7 (Fig. S3A-B). All
173 models were better at predicting *C. difficile* colonization status on day 7 than random chance (all
174 $P \leq 5e-15$; Table S12). However, the models trained with OTU level data generally performed

175 better than those trained with family level data with the exception of the models based on
176 the post-infection (day 1) communities (Fig. S3C-D). Interestingly, the model based on the
177 post-clindamycin (day 0) community OTU data performed significantly better than all other models
178 with an area under the receiving operator characteristic curve (AUROC) of 0.75 ($P_{FDR} \leq 3.9e-10$
179 for pairwise comparisons; Table S13). Thus, we were able to use community bacterial relative
180 abundance data alone to differentiate mice that had cleared *C. difficile* before day 7 from the mice
181 still colonized with *C. difficile* at that timepoint. Interestingly, the model built with OTU relative
182 abundance data post-clindamycin treatment had the best performance, suggesting how the
183 bacterial community responds to clindamycin treatment has the greatest influence on subsequent
184 *C. difficile* colonization dynamics.

185 Next, to examine the bacteria that were driving each model's performance, we pulled out the top 20
186 taxa that had the highest absolute feature weights in each of the 6 models (Tables S14-15). First, we
187 looked at OTUs from the model with the best performance that was based on the post-clindamycin
188 treatment bacterial community data. While most of the 20 OTUs had low relative abundances on day
189 0, *Enterobacteriace*, *Bacteroides* and *Proteus* had high relative abundances in at least one source
190 of mice and significantly varied across sources (Fig. 7A). Next, the top 20 taxa from each model
191 were compared to the list of taxa that varied across colony source at the same timepoint (Fig. 5 and
192 Tables S8-9) and the taxa that were altered by clindamycin treatment (Fig. 6 and Table S10-11).
193 We found a subset of OTUs and families that were important to the model and overlapped with
194 bacteria that varied by either source, clindamycin treatment, or both (Fig. S4, S5A-C). Combining
195 the overall results for the 3 OTU models identified 14 OTUs associated with source, 21 OTUs
196 associated with clindamycin treatment, and 6 OTUs associated with both (Fig. 7B). Combining
197 the overall results for the 3 family models identified 18 families associated with source, 14 families
198 associated with clindamycin treatment and 8 families associated with both (Fig. S5D). Several OTUs
199 (*Bacteroides* (OTU 2), *Enterococcus* (OTU 23), *Enterobacteriaceae* (OTU 1), *Porphyromonadaceae*
200 (OTU 7)) and families (*Bacteroidaceae*, *Deferrribacteraceae*, *Enterococcaceae*, *Lachnospiraceae*,
201 *Bifidobacteriaceae*, *Coriobacteriaceae*, *Ruminococcaceae*, *Verrucomicrobiaceae*) appeared across
202 at least 2 models, so we examined how the relative abundances of these key taxa varied over the
203 course of the experiment (Fig. 8 and Fig. S6). Throughout the experiment, there was at least

204 1 timepoint where relative abundances of these taxa significantly varied across sources (Table
205 S16-17). Interestingly, there were no taxa that emerged as consistently enriched or depleted in
206 mice that were colonized past 7 days post-infection with *C. difficile* 630, suggesting multiple bacteria
207 influence the time needed to clear the infection. Together, these results suggest the initial bacterial
208 communities and their responses to clindamycin influence on the time needed to clear *C. difficile*.

209 Discussion

210 By examining the *C. difficile* colonization dynamics within mice from 6 different colony sources
211 after perturbing the microbiota with clindamycin treatment, we were able to identify bacterial taxa
212 that were unique to sources throughout the experiment as well as taxa that were universally
213 impacted by clindamycin. We built L2 logistic regression models with baseline, post-clindamycin
214 treatment, and post-infection fecal community data that successfully predicted whether mice
215 cleared *C. difficile* by 7 days post-infection better than random chance. We identified *Bacteroides*
216 (OTU 2), *Enterococcus* (OTU 23), *Enterobacteriaceae* (OTU 1), *Porphyromonadaceae* (OTU
217 7), *Bacteroidaceae*, *Deferrribacteraceae*, *Enterococcaceae*, *Lachnospiraceae*, *Bifidobacteriaceae*,
218 *Coriobacteriaceae*, *Ruminococcaceae*, *Verrucomicrobiaceae* (Fig. 8, Fig. S6) as candidate bacteria
219 within these communities that were influencing variation in *C. difficile* colonization dynamics since
220 these bacteria were all important in the logistic regression models and varied by colony source,
221 were impacted by clindamycin treatment, or both. Overall, our results demonstrate clindamycin
222 is sufficient to render mice from multiple sources susceptible to CDI and only a subset of the
223 interindividual microbiota variation across mice from different sources was associated with the time
224 needed to clear *C. difficile*.

225 Other groups have taken similar approaches by using mice from multiple colony sources to identify
226 bacteria that either promote colonization resistance or increase susceptibility to enteric infections
227 (22, 23, 26–30). For example, in the context of *Salmonella* infections, *Enterobacteriaceae* and
228 segmented filamentous bacteria have emerged as protective (22, 27). A previous study with *C.*
229 *difficile* identified an endogenous protective *C. difficile* strain LEM1 that bloomed after antibiotic
230 treatment in mice from Jackson or Charles River Laboratories, but not Taconic that protected
231 mice against the more toxigenic *C. difficile* VPI10463 (26). Given that we ordered mice from the

232 same vendors, we checked all mice for endogenous *C. difficile* by plating stool samples that were
233 collected after clindamycin treatment. However, we did not identify any endogenous *C. difficile*
234 strains prior to challenge, suggesting there were no endogenous protective strains in the mice we
235 received and other bacterial taxa mediated the variation in *C. difficile* colonization across sources.
236 Although all mice were susceptible to *C. difficile* colonization, by following colonization over time
237 we found Jackson, Taconic, and Envigo mice tended to remain colonized through at least 7 days
238 post-infection. We identified a subset of bacteria that were important in predicting whether a mouse
239 was still colonized with *C. difficile* 7 days post-infection. These results suggest a subset of the
240 bacterial community is responsible for determining the length of time needed to clear *C. difficile*
241 colonization.

242 In the past variation between different CDI mouse model studies have been attributed to intestinal
243 microbiome differences in mice across different institutional environments. For example, groups
244 using the same clindamycin treatment and C57BL/6 mice had different *C. difficile* outcomes, one
245 having sustained colonization (32), while the other had transient (18). Baseline differences in the
246 microbiota composition have been hypothesized to partially explain the differences in colonization
247 outcomes and overall susceptibility to *C. difficile* after treatment with the same antibiotic (13,
248 31). We have shown that mice from 6 different sources were all susceptible to *C. difficile* 630,
249 suggesting the microbiota influences *C. difficile* clearance more than susceptibility. Fortunately,
250 the bacterial perturbations induced by clindamycin treatment have been well characterized
251 and our findings agree with previous CDI mouse model work demonstrating *Enterococcus* and
252 *Enterobacteriaceae* were associated with *C. difficile* susceptibility and *Porphyromonadaceae*,
253 *Lachnospiraceae*, *Ruminococcaceae*, and *Turicibacter* were associated with resistance (19, 21, 32,
254 33, 41, 46–48). While we have demonstrated that susceptibility is uniform across vendors after
255 clindamycin treatment, there could be different outcomes for either susceptibility or clearance in
256 the case of other antibiotic treatments. The *C. difficile* strain used could also be contributing to
257 the variation in *C. difficile* outcomes seen across different research groups (47). We found the
258 time needed to naturally clear *C. difficile* varied across sources of mice implying that at least in the
259 context of the same perturbation, microbiota differences seemed to influence infection outcome
260 more than susceptibility. More importantly, we were able to narrow down from all the variation

261 observed across colony sources to a subset of bacterial taxa that were also important for predicting
262 *C. difficile* colonization status 7 days post-infection. Since all but 3 mice eventually cleared *C.*
263 *difficile* 630 by 9 days post-infection and the model built with the post-clindamycin OTU relative
264 abundance data had the best performance, our results suggest clindamycin treatment had a large
265 role in determining *C. difficile* susceptibility and clearance in the mice.

266 Our approach successfully increased the diversity of murine bacterial communities tested in our
267 clindamycin *C. difficile* model. One alternative approach that has been used in some CDI studies
268 (49–54) is to associate mice with human microbiotas. However, a major caveat to this method
269 is the substantial loss of human microbiota community members upon transfer to mice (55, 56).
270 Additionally with the exception of 2 recent studies (49, 50), most of the CDI mouse model studies
271 to date associated mice with just 1 types of human microbiota either from a single donor or a
272 single pool from multiple donors (51–54), which does not aid in the goal of figuring out how a
273 variety of unique microbiotas influence susceptibility to CDIs and adverse outcomes. Importantly,
274 our study using mice from 6 different sources increased the variation between groups of mice
275 compared to using 1 source alone, to better reflect the inter-individual microbiota variation observed
276 in humans. Encouragingly, decreased *Bifidobacterium*, *Porphyromonas*, *Ruminococcaceae*
277 and *Lachnospiraceae* and increased *Enterobacteriaceae*, *Enterococcus*, *Lactobacillus*, and
278 *Proteus* have all been associated with human CDIs (7) and were well represented in our study,
279 suggesting most of the mouse sources are suitable for gaining insights into microbiota associated
280 factors influencing *C. difficile* colonization and infections in humans. An important exception was
281 *Enterococcus*, which was primarily absent from the mice from University of Michigan colonies and
282 *Proteus*, which was only found in Taconic mice. Importantly, the fact that some CDI-associated
283 bacteria were only found in a subset of mice has important implications for future CDI mouse model
284 studies.

285 There are several limitations to our work. The microbiome is composed of viruses, fungi, and
286 parasites in addition to bacteria, and these non-bacterial members can also vary across mouse
287 vendors (57, 58). While our study focused solely on the bacterial portion, viruses and fungi have
288 also begun to be implicated in the context of CDIs or FMT treatments for recurrent CDIs (35, 59–62).
289 Beyond community composition, the metabolic function of the microbiota also has a CDI signature

(20, 48, 63, 64) and can vary across mice from different sources (65). For example, microbial metabolites, particularly secondary bile acids and butyrate production, have been implicated as important contributors to *C. difficile* resistance (33, 47). Although, we only looked at composition, *Ruminococcaceae* and *Lachnospiraceae* both emerged as important taxa for classifying day 7 *C. difficile* colonization status and metagenomes from these bacteria have been shown to contain the bile acid-inducible gene cluster necessary for secondary bile acid formation and ability to produce butyrate (52, 66). Interestingly, butyrate has previously been shown to vary across vendors and mediated resistance to *Citrobacter rodentium* infection, a model of enterohemorrhagic and enteropathogenic *Escherichia coli* infections (23). Evidence for immunological toning differences in IgA and Th17 cells across mice from different vendors have also been documented and (67, 68) may also influence response to CDI, particularly in the context of severe CDIs (69, 70). The outcome after *C. difficile* exposure depends on a multitude of factors, including age, diet, and immunity; all of which are also influenced by the microbiota. We have demonstrated that the ways baseline microbiotas from different mouse colony sources respond to clindamycin treatment influences the length of time mice remained colonized with *C. difficile* 630. For those interested in dissecting the contribution of the microbiome to *C. difficile* pathogenesis and treatments, using multiple sources of mice may yield more insights than a single model alone. Furthermore, for studies wanting to examine the interplay between a particular bacterial taxon such as *Enterococcus* and *C. difficile*, these results could serve as a resource for selecting which mice to order to address the question.

310 **Acknowledgements**

311 This work was supported by the National Institutes of Health (U01AI124255). ST was supported by
312 the Michigan Institute for Clinical and Health Research Postdoctoral Translation Scholars Program
313 (UL1TR002240). We thank members of the Schloss lab for feedback on planning the experiments
314 and data presentation, as well as code tutorials and feedback through Code club. In particular, we
315 want to thank Begüm Topçuoğlu for help with implementing L2 logistic regression models using
316 her pipeline, Ana Taylor for help with some of the media preparation and sample collection, and
317 Nicholas Lesniak for his critical feedback on the manuscript. We also thank members of Vincent
318 Young's lab, particularly Kimberly Vendrov, for guidance with the *C. difficile* infection mouse model
319 and donating the mice. We also want to thank the Unit for Laboratory Animal Medicine at the
320 University of Michigan for maintaining our mouse colony and providing the institutional support for
321 our mouse experiments. Finally, we thank Kwi Kim, Austin Campbell, and Kimberly Vendrov for
322 their help in maintaining the Schloss lab's anaerobic chamber.

323 **Materials and Methods**

324 **(i) Animals.** All experiments were approved by the University of Michigan Animal Care and Use
325 Committee (IACUC) under protocol number PRO00006983. Female C57BL/7 mice were obtained
326 from 6 different colony sources: The Jackson Laboratory, Charles River Laboratories, Taconic
327 Biosciences, Envigo, and two colonies at the University of Michigan (the Schloss lab colony and
328 the Young lab colony). The Young lab colony was originally established with mice purchased from
329 Jackson, and the Schloss lab colony was later founded with mice donated from the Young lab. The
330 4 groups of mice purchased from vendors were allowed to acclimate to the University of Michigan
331 mouse facility for 13 days prior to starting the experiment. At least 4 female mice (age 5-10 weeks)
332 were obtained per source and mice from the same source were primarily housed at a density of 2
333 mice per cage. The experiment was repeated once, approximately 3 months after the start of the
334 first experiment.

335 **(ii) Antibiotic treatment.** After the 13-day acclimation period and 1 day prior to challenge (Fig.
336 1A), all mice received 10 mg/kg clindamycin (filter sterilized through a 0.22 micron syringe filter
337 prior to administration) via intraperitoneal injection.

338 **(iii) *C. difficile* infection model.** Mice were challenged with 10^3 spores of *C. difficile* strain 630
339 via oral gavage post-infection 1 day after clindamycin treatment as described previously (21). Mice
340 weights and stool samples were taken daily through 9 days post-challenge. Collected stool was
341 split for *C. difficile* CFU quantification and 16S rRNA sequencing analysis. *C. difficile* quantification
342 stool samples were transferred to the anaerobic chamber, serially diluted in PBS, plated on
343 taurocholate-cycloserine-cefoxitin-fructose agar (TCCFA) plates, and counted after 24 hours of
344 incubation at 37°C under anaerobic conditions. A sample from the day 0 timepoint (post-clindamycin
345 and prior to *C. difficile* challenge) was also plated on TCCFA to ensure mice were not already
346 colonized with *C. difficile* prior to infection. There were 3 deaths recorded over the course of the
347 experiment, 1 Taconic mouse died prior to *C. difficile* challenge and 1 Jackson and 1 Envigo mouse
348 died between 1- and 3-days post-infection. Mice were categorized as cleared when no *C. difficile*
349 was detected in the first serial dilution (limit of detection: 100 CFU). Stool samples for 16S rRNA
350 sequencing were snap frozen in liquid nitrogen and stored at -80°C until DNA extraction.

351 **(iv) 16S rRNA sequencing.** DNA was extracted from -80 °C stored stool samples using the DNeasy
352 Powersoil HTP 96 kit (Qiagen) and an EpMotion 5075 automated pipetting system (Eppendorf).
353 The V4 region was amplified for 16S rRNA with the AccuPrime Pfx DNA polymerase (Thermo
354 Fisher Scientific) using custom barcoded primers, as previously described (71). The ZymoBIOMICS
355 microbial community DNA standards was used as a mock community control (72) and water was
356 used as a negative control per 96-well extraction plate. The PCR amplicons were cleaned up
357 and normalized with the SequalPrep normalization plate kit (Thermo Fisher Scientific). Amplicons
358 were pooled and quantified with the KAPA library quantification kit (KAPA biosystems), prior to
359 sequencing using the MiSeq system (Illumina).

360 **(v) 16S rRNA gene sequence analysis.** mothur (v. 1.43) was used to process all sequences
361 (73) with a previously published protocol (71). Reads were combined and aligned with the SILVA
362 reference database (74). Chimeras were removed with the VSEARCH algorithm and taxonomic
363 assignment was completed with a modified version (v16) of the Ribosomal Database Project
364 reference database (v11.5) (75) with an 80% cutoff. Operational taxonomic units (OTUs) were
365 assigned with a 97% similarity threshold using the opticlus algorithm (76). To account for uneven
366 sequencing across samples, samples were rarefied to 5,437 sequences 1,000 times for alpha and
367 beta diversity analyses. PCoAs were generated based on Theta YC distances. Permutational
368 multivariate analysis of variance (PERMANOVA) was performed on mothur-generated Theta YC
369 distance matrices with the adonis function in the vegan package (77) in R (78).

370 **(vi) Classification model training and evaluation.** Models were generated based on mice that
371 were categorized as either cleared or colonized 7 days post-infection and had sequencing data
372 from the baseline (day -1), post-clindamycin (day 0), and post-infection (day 1) timepoints of the
373 experiment. Input bacterial community relative abundance data at either the OTU or family level from
374 the baseline, post-clindamycin, and post-infection timepoints was used to generate 6 classification
375 models that predicted *C. difficile* colonization status 7 days post-infection. The L2-regularized
376 logistic regression models were trained and tested using the caret package (79) in R as previously
377 described (80) with the exception that we used 60% training and 40% testing data splits for the
378 cross-validation of the training data to select the best cost hyperparameter and the testing of
379 the held out test data to measure model performance. The modified training to testing ratio was

380 selected to accommodate the small number of samples in the dataset. Code was modified from
381 https://github.com/SchlossLab/ML_pipeline_microbiome to update the classification outcomes and
382 change the data split ratios. The modified repository to regenerate this analysis is available at
383 https://github.com/tomkosev/ML_pipeline_microbiome.

384 **(vii) Statistical analysis.** All statistical tests were performed in R (v 3.5.2) (78). The Kruskal-Wallis
385 test was used to analyze differences in *C. difficile* CFU, mouse weight change, and alpha diversity
386 across vendors with a Benjamini-Hochberg correction for testing multiple timepoints, followed by
387 pairwise Wilcoxon comparisons with Benjamini-Hochberg correction. For taxonomic analysis and
388 generation of logistic regression model input data, *C. difficile* (OTU 20) was removed. Bacterial
389 relative abundances that varied across sources at the OTU and family taxonomic levels were
390 identified with the Kruskal-Wallis test with Benjamini-Hochberg correction for testing all identified
391 taxa at each level, followed by pairwise Wilcoxon comparisons with Benjamini-Hochberg correction.
392 Taxa impacted by clindamycin treatment were identified using the Wilcoxon signed rank test with
393 matched pairs of mice samples for day -1 and day 0. To determine whether classification models had
394 better performance (test AUROCs) than random chance (0.5), we used the one-sample Wilcoxon
395 signed rank test. To examine whether there was an overall difference in predictive performance
396 across the 6 classification models we used the Kruskal-Wallis test followed by pairwise Wilcoxon
397 comparisons with Benjamini-Hochberg correction for multiple hypothesis testing. The tidyverse
398 package was used to wrangle and graph data (v 1.3.0) (81).

399 **(viii) Code availability.** Code for all data analysis and generating this manuscript is available at
400 https://github.com/SchlossLab/Tomkovich_vendor_difs_XXXX_2020.

401 **(ix) Data availability.** The 16S rRNA sequencing data have been deposited in the National Center
402 for Biotechnology Information Sequence Read Archive (BioProject Accession no. PRJNA608529).

403 **References**

- 404 1. Teng C, Reveles KR, Obodozie-Ofoegbu OO, Frei CR. 2019. *Clostridium difficile* infection
405 risk with important antibiotic classes: An analysis of the FDA adverse event reporting system.
406 International Journal of Medical Sciences 16:630–635.
- 407 2. Kelly C. 2012. Can we identify patients at high risk of recurrent *Clostridium difficile* infection?
408 Clinical Microbiology and Infection 18:21–27.
- 409 3. Zacharioudakis IM, Zervou FN, Pliakos EE, Ziakas PD, Mylonakis E. 2015. Colonization with
410 toxinogenic *C. difficile* upon hospital admission, and risk of infection: A systematic review and
411 meta-analysis. American Journal of Gastroenterology 110:381–390.
- 412 4. Crobach MJT, Vernon JJ, Loo VG, Kong LY, Péchiné S, Wilcox MH, Kuijper EJ. 2018.
413 Understanding *Clostridium difficile* colonization. Clinical Microbiology Reviews 31.
- 414 5. Zhang L, Dong D, Jiang C, Li Z, Wang X, Peng Y. 2015. Insight into alteration of gut microbiota
415 in *Clostridium difficile* infection and asymptomatic *c. difficile* colonization. Anaerobe 34:1–7.
- 416 6. VanInsberghe D, Elsherbini JA, Varian B, Poutahidis T, Erdman S, Polz MF. 2020. Diarrhoeal
417 events can trigger long-term *Clostridium difficile* colonization with recurrent blooms. Nature
418 Microbiology 5:642–650.
- 419 7. Mancabelli L, Milani C, Lugli GA, Turroni F, Cocconi D, Sinderen D van, Ventura M. 2017.
420 Identification of universal gut microbial biomarkers of common human intestinal diseases by
421 meta-analysis. FEMS Microbiology Ecology 93.
- 422 8. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. 2017. Meta-analysis of gut microbiome
423 studies identifies disease-specific and shared responses. Nature Communications 8.
- 424 9. Seekatz AM, Rao K, Santhosh K, Young VB. 2016. Dynamics of the fecal microbiome in patients
425 with recurrent and nonrecurrent *Clostridium difficile* infection. Genome Medicine 8.
- 426 10. Khanna S, Montassier E, Schmidt B, Patel R, Knights D, Pardi DS, Kashyap PC. 2016. Gut
427 microbiome predictors of treatment response and recurrence in primary *Clostridium difficile* infection.

- 428 Alimentary Pharmacology & Therapeutics 44:715–727.
- 429 11. Pakpour S, Bhanvadia A, Zhu R, Amarnani A, Gibbons SM, Gurry T, Alm EJ, Martello LA. 2017.
430 Identifying predictive features of *Clostridium difficile* infection recurrence before, during, and after
431 primary antibiotic treatment. Microbiome 5.
- 432 12. Lee AA, Rao K, Limsrivilai J, Gilliland M, Malamet B, Briggs E, Young VB, Higgins PDR. 2020.
433 Temporal gut microbial changes predict recurrent *Clostridioides difficile* infection in patients with
434 and without ulcerative colitis. Inflammatory Bowel Diseases.
- 435 13. Hutton ML, Mackin KE, Chakravorty A, Lyras D. 2014. Small animal models for the study of
436 *Clostridium difficile* disease pathogenesis. FEMS Microbiology Letters 352:140–149.
- 437 14. Chen X, Katchar K, Goldsmith JD, Nanthakumar N, Cheknis A, Gerding DN, Kelly CP. 2008. A
438 mouse model of *Clostridium difficile*-associated disease. Gastroenterology 135:1984–1992.
- 439 15. Best EL, Freeman J, Wilcox MH. 2012. Models for the study of *Clostridium difficile* infection.
440 Gut Microbes 3:145–167.
- 441 16. Baxter NT, Wan JJ, Schubert AM, Jenior ML, Myers P, Schloss PD. 2014. Intra- and
442 interindividual variations mask interspecies variation in the microbiota of sympatric peromyscus
443 populations. Applied and Environmental Microbiology 81:396–404.
- 444 17. Nagpal R, Wang S, Woods LCS, Seshie O, Chung ST, Shively CA, Register TC, Craft S,
445 McClain DA, Yadav H. 2018. Comparative microbiome signatures and short-chain fatty acids in
446 mouse, rat, non-human primate, and human feces. Frontiers in Microbiology 9.
- 447 18. Reeves AE, Theriot CM, Bergin IL, Huffnagle GB, Schloss PD, Young VB. 2011. The interplay
448 between microbiome dynamics and pathogen dynamics in a murine model of *Clostridium difficile*
449 infection 2:145–158.
- 450 19. Schubert AM, Sinani H, Schloss PD. 2015. Antibiotic-induced alterations of the murine gut
451 microbiota and subsequent effects on colonization resistance against *Clostridium difficile*. mBio 6.
- 452 20. Jenior ML, Leslie JL, Young VB, Schloss PD. 2017. *Clostridium difficile* colonizes alternative

- 453 nutrient niches during infection across distinct murine gut microbiomes. *mSystems* 2.
- 454 21. Jenior ML, Leslie JL, Young VB, Schloss PD. 2018. *Clostridium difficile* alters the structure and
455 metabolism of distinct cecal microbiomes during initial infection to promote sustained colonization.
456 *mSphere* 3.
- 457 22. Velazquez EM, Nguyen H, Heasley KT, Saechao CH, Gil LM, Rogers AWL, Miller BM, Rolston
458 MR, Lopez CA, Litvak Y, Liou MJ, Faber F, Bronner DN, Tiffany CR, Byndloss MX, Byndloss
459 AJ, Bäumler AJ. 2019. Endogenous Enterobacteriaceae underlie variation in susceptibility to
460 *Salmonella* infection. *Nature Microbiology* 4:1057–1064.
- 461 23. Osbelt L, Thiemann S, Smit N, Lesker TR, Schröter M, Gálvez EJC, Schmidt-Hohagen K, Pils
462 MC, Mühlen S, Dersch P, Hiller K, Schlüter D, Neumann-Schaal M, Strowig T. 2020. Variations in
463 microbiota composition of laboratory mice influence *Citrobacter rodentium* infection via variable
464 short-chain fatty acid production. *PLOS Pathogens* 16:e1008448.
- 465 24. Stough JMA, Dearth SP, Denny JE, LeCleir GR, Schmidt NW, Campagna SR, Wilhelm SW.
466 2016. Functional characteristics of the gut microbiome in C57BL/6 mice differentially susceptible to
467 *Plasmodium yoelii*. *Frontiers in Microbiology* 7.
- 468 25. Alegre M-L. 2019. Mouse microbiomes: Overlooked culprits of experimental variability. *Genome
469 Biology* 20.
- 470 26. Etienne-Mesmin L, Chassaing B, Adekunle O, Mattei LM, Bushman FD, Gewirtz AT. 2017.
471 Toxin-positive *Clostridium difficile* latently infect mouse colonies and protect against highly
472 pathogenic *C. difficile*. *Gut* 67:860–871.
- 473 27. Lai NY, Musser MA, Pinho-Ribeiro FA, Baral P, Jacobson A, Ma P, Potts DE, Chen Z, Paik D,
474 Soualhi S, Yan Y, Misra A, Goldstein K, Lagomarsino VN, Nordstrom A, Sivanathan KN, Wallrapp A,
475 Kuchroo VK, Nowarski R, Starnbach MN, Shi H, Surana NK, An D, Wu C, Huh JR, Rao M, Chiu IM.
476 2020. Gut-innervating nociceptor neurons regulate peyer's patch microfold cells and SFB levels to
477 mediate *Salmonella* host defense. *Cell* 180:33–49.e22.
- 478 28. Thiemann S, Smit N, Roy U, Lesker TR, Gálvez EJ, Helmecke J, Basic M, Bleich A, Goodman

- 479 AL, Kalinke U, Flavell RA, Erhardt M, Strowig T. 2017. Enhancement of IFNgamma production by
480 distinct commensals ameliorates *Salmonella*-induced disease. *Cell Host & Microbe* 21:682–694.e5.
- 481 29. Rolig AS, Cech C, Ahler E, Carter JE, Ottemann KM. 2013. The degree of *Helicobacter*
482 *pylori*-triggered inflammation is manipulated by preinfection host microbiota. *Infection and Immunity*
483 81:1382–1389.
- 484 30. Ge Z, Sheh A, Feng Y, Muthupalani S, Ge L, Wang C, Kurnick S, Mannion A, Whary MT, Fox
485 JG. 2018. *Helicobacter pylori*-infected C57BL/6 mice with different gastrointestinal microbiota have
486 contrasting gastric pathology, microbial and host immune responses. *Scientific Reports* 8.
- 487 31. Lawley TD, Young VB. 2013. Murine models to study *Clostridium difficile* infection and
488 transmission. *Anaerobe* 24:94–97.
- 489 32. Buffie CG, Jarchum I, Equinda M, Lipuma L, Gobourne A, Viale A, Ubeda C, Xavier J, Pamer
490 EG. 2011. Profound alterations of intestinal microbiota following a single dose of clindamycin results
491 in sustained susceptibility to *Clostridium difficile*-induced colitis. *Infection and Immunity* 80:62–73.
- 492 33. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew
493 M, Viale A, Littmann E, Brink MRM van den, Jenq RR, Taur Y, Sander C, Cross JR, Toussaint
494 NC, Xavier JB, Pamer EG. 2014. Precision microbiome reconstitution restores bile acid mediated
495 resistance to *Clostridium difficile*. *Nature* 517:205–208.
- 496 34. Spinler JK, Brown A, Ross CL, Boonma P, Conner ME, Savidge TC. 2016. Administration of
497 probiotic kefir to mice with *Clostridium difficile* infection exacerbates disease. *Anaerobe* 40:54–57.
- 498 35. Markey L, Shaban L, Green ER, Lemon KP, Mecsas J, Kumamoto CA. 2018. Pre-colonization
499 with the commensal fungus candida albicans reduces murine susceptibility to *Clostridium difficile*
500 infection. *Gut Microbes* 1–13.
- 501 36. McKee RW, Aleksanyan N, Garrett EM, Tamayo R. 2018. Type IV pili promote *Clostridium*
502 *difficile* adherence and persistence in a mouse model of infection. *Infection and Immunity* 86.
- 503 37. Yamaguchi T, Konishi H, Aoki K, Ishii Y, Chono K, Tateda K. 2020. The gut microbiome diversity

- 504 of *Clostridioides difficile*-inoculated mice treated with vancomycin and fidaxomicin. Journal of
505 Infection and Chemotherapy 26:483–491.
- 506 38. Stroke IL, Letourneau JJ, Miller TE, Xu Y, Pechik I, Savoly DR, Ma L, Sturzenbecker LJ,
507 Sabalski J, Stein PD, Webb ML, Hilbert DW. 2018. Treatment of *Clostridium difficile* infection
508 with a small-molecule inhibitor of toxin UDP-glucose hydrolysis activity. Antimicrobial Agents and
509 Chemotherapy 62.
- 510 39. Quigley L, Coakley M, Alemayehu D, Rea MC, Casey PG, O'Sullivan, Murphy E, Kiely B, Cotter
511 PD, Hill C, Ross RP. 2019. *Lactobacillus gasseri* APC 678 reduces shedding of the pathogen
512 *Clostridium difficile* in a murine model. Frontiers in Microbiology 10.
- 513 40. Mullish BH, McDonald JAK, Pechlivanis A, Allegretti JR, Kao D, Barker GF, Kapila D, Petrof
514 EO, Joyce SA, Gahan CGM, Glegola-Madejska I, Williams HRT, Holmes E, Clarke TB, Thursz
515 MR, Marchesi JR. 2019. Microbial bile salt hydrolases mediate the efficacy of faecal microbiota
516 transplant in the treatment of recurrent *Clostridioides difficile* infection. Gut 68:1791–1800.
- 517 41. Tomkovich S, Lesniak NA, Li Y, Bishop L, Fitzgerald MJ, Schloss PD. 2019. The proton
518 pump inhibitor omeprazole does not promote *Clostridioides difficile* colonization in a murine model.
519 mSphere 4.
- 520 42. Guh AY, Kutty PK. 2018. *Clostridioides difficile* infection 169:ITC49.
- 521 43. Theriot CM, Koumpouras CC, Carlson PE, Bergin II, Aronoff DM, Young VB. 2011.
522 Cefoperazone-treated mice as an experimental platform to assess differential virulence of
523 *Clostridium difficile* strains. Gut Microbes 2:326–334.
- 524 44. Ross CL, Spinler JK, Savidge TC. 2016. Structural and functional changes within the gut
525 microbiota and susceptibility to *Clostridium difficile* infection. Anaerobe 41:37–43.
- 526 45. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. 2015. How informative is the mouse for human
527 gut microbiota research? Disease Models & Mechanisms 8:1–16.
- 528 46. Lawley TD, Clare S, Walker AW, Goulding D, Stabler RA, Croucher N, Mastroeni P, Scott P,

- 529 Raisen C, Mottram L, Fairweather NF, Wren BW, Parkhill J, Dougan G. 2009. Antibiotic treatment of
530 *Clostridium difficile* carrier mice triggers a supershedder state, spore-mediated transmission, and
531 severe disease in immunocompromised hosts. *Infection and Immunity* 77:3661–3669.
- 532 47. Lawley TD, Clare S, Walker AW, Stares MD, Connor TR, Raisen C, Goulding D, Rad R,
533 Schreiber F, Brandt C, Deakin LJ, Pickard DJ, Duncan SH, Flint HJ, Clark TG, Parkhill J, Dougan
534 G. 2012. Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy
535 resolves relapsing *Clostridium difficile* disease in mice. *PLoS Pathogens* 8:e1002995.
- 536 48. Jump RLP, Polinkovsky A, Hurless K, Sitzlar B, Eckart K, Tomas M, Deshpande A, Nerandzic
537 MM, Donskey CJ. 2014. Metabolomics analysis identifies intestinal microbiota-derived biomarkers
538 of colonization resistance in clindamycin-treated mice. *PLoS ONE* 9:e101267.
- 539 49. Nagao-Kitamoto H, Leslie JL, Kitamoto S, Jin C, Thomsson KA, Gilliland MG, Kuffa P, Goto Y,
540 Jenq RR, Ishii C, Hirayama A, Seekatz AM, Martens EC, Eaton KA, Kao JY, Fukuda S, Higgins PDR,
541 Karlsson NG, Young VB, Kamada N. 2020. Interleukin-22-mediated host glycosylation prevents
542 *Clostridioides difficile* infection by modulating the metabolic activity of the gut microbiota. *Nature
543 Medicine* 26:608–617.
- 544 50. Battaglioli EJ, Hale VL, Chen J, Jeraldo P, Ruiz-Mojica C, Schmidt BA, Rekdal VM, Till LM, Huq
545 L, Smits SA, Moor WJ, Jones-Hall Y, Smyrk T, Khanna S, Pardi DS, Grover M, Patel R, Chia N,
546 Nelson H, Sonnenburg JL, Farrugia G, Kashyap PC. 2018. *Clostridioides difficile* uses amino acids
547 associated with gut microbial dysbiosis in a subset of patients with diarrhea. *Science Translational
548 Medicine* 10:eaam7019.
- 549 51. Robinson CD, Auchtung JM, Collins J, Britton RA. 2014. Epidemic *Clostridium difficile* strains
550 demonstrate increased competitive fitness compared to nonepidemic isolates. *Infection and
551 Immunity* 82:2815–2825.
- 552 52. Collins J, Auchtung JM, Schaefer L, Eaton KA, Britton RA. 2015. Humanized microbiota mice
553 as a model of recurrent *Clostridium difficile* disease. *Microbiome* 3.
- 554 53. Collins J, Robinson C, Danhof H, Knetsch CW, Leeuwen HC van, Lawley TD, Auchtung JM,

- 555 Britton RA. 2018. Dietary trehalose enhances virulence of epidemic *Clostridium difficile*. Nature
556 553:291–294.
- 557 54. Hryckowian AJ, Treuren WV, Smits SA, Davis NM, Gardner JO, Bouley DM, Sonnenburg JL.
558 2018. Microbiota-accessible carbohydrates suppress *Clostridium difficile* infection in a murine
559 model. *Nature Microbiology* 3:662–669.
- 560 55. Fouladi F, Glenny EM, Bulik-Sullivan EC, Tsilimigras MCB, Sioda M, Thomas SA, Wang Y, Djukic
561 Z, Tang Q, Tarantino LM, Bulik CM, Fodor AA, Carroll IM. 2020. Sequence variant analysis reveals
562 poor correlations in microbial taxonomic abundance between humans and mice after gnotobiotic
563 transfer. *The ISME Journal*.
- 564 56. Walter J, Armet AM, Finlay BB, Shanahan F. 2020. Establishing or exaggerating causality for
565 the gut microbiome: Lessons from human microbiota-associated rodents. *Cell* 180:221–232.
- 566 57. Rasmussen TS, Vries L de, Kot W, Hansen LH, Castro-Mejía JL, Vogensen FK, Hansen AK,
567 Nielsen DS. 2019. Mouse vendor influence on the bacterial and viral gut composition exceeds the
568 effect of diet. *Viruses* 11:435.
- 569 58. Mims TS, Abdallah QA, Watts S, White C, Han J, Willis KA, Pierre JF. 2020. Variability in
570 interkingdom gut microbiomes between different commercial vendors shapes fat gain in response
571 to diet. *The FASEB Journal* 34:1–1.
- 572 59. Stewart DB, Wright JR, Fowler M, McLimans CJ, Tokarev V, Amaniera I, Baker O, Wong H-T,
573 Brabec J, Drucker R, Lamendella R. 2019. Integrated meta-omics reveals a fungus-associated
574 bacteriome and distinct functional pathways in *Clostridioides difficile* infection. *mSphere* 4.
- 575 60. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, Grasis JA, Cassidy L, Tholey A,
576 Fickenscher H, Seegert D, Rosenstiel P, Schreiber S. 2017. Efficacy of sterile fecal filtrate transfer
577 for treating patients with *Clostridium difficile* infection. *Gastroenterology* 152:799–811.e7.
- 578 61. Zuo T, Wong SH, Lam K, Lui R, Cheung K, Tang W, Ching JYL, Chan PKS, Chan MCW, Wu
579 JCY, Chan FKL, Yu J, Sung JJY, Ng SC. 2017. Bacteriophage transfer during faecal microbiota
580 transplantation in *Clostridium difficile* infection is associated with treatment outcome. *Gut*

- 581 gutjnl–2017–313952.
- 582 62. Zuo T, Wong SH, Cheung CP, Lam K, Lui R, Cheung K, Zhang F, Tang W, Ching JYL, Wu JCY,
583 Chan PKS, Sung JJY, Yu J, Chan FKL, Ng SC. 2018. Gut fungal dysbiosis correlates with reduced
584 efficacy of fecal microbiota transplantation in *Clostridium difficile* infection. *Nature Communications*
585 9.
- 586 63. Robinson JI, Weir WH, Crowley JR, Hink T, Reske KA, Kwon JH, Burnham C-AD, Dubberke
587 ER, Mucha PJ, Henderson JP. 2019. Metabolomic networks connect host-microbiome processes to
588 human *Clostridioides difficile* infections. *Journal of Clinical Investigation* 129:3792–3806.
- 589 64. Fletcher JR, Erwin S, Lanzas C, Theriot CM. 2018. Shifts in the gut metabolome and *Clostridium*
590 *difficile* transcriptome throughout colonization and infection in a mouse model. *mSphere* 3.
- 591 65. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, Liu C, Fang
592 Z, Chou J, Glanville J, Hao Q, Kotowska D, Colding C, Licht TR, Wu D, Yu J, Sung JJY, Liang Q, Li
593 J, Jia H, Lan Z, Tremaroli V, Dworzynski P, Nielsen HB, Bäckhed F, Doré J, Chatelier EL, Ehrlich
594 SD, Lin JC, Arumugam M, Wang J, Madsen L, Kristiansen K. 2015. A catalog of the mouse gut
595 metagenome. *Nature Biotechnology* 33:1103–1108.
- 596 66. Vital M, Rud T, Rath S, Pieper DH, Schlüter D. 2019. Diversity of bacteria exhibiting bile
597 acid-inducible 7alpha-dehydroxylation genes in the human gut. *Computational and Structural
598 Biotechnology Journal* 17:1016–1019.
- 599 67. Fransen F, Zagato E, Mazzini E, Fosso B, Manzari C, Aidy SE, Chiavelli A, D'Erchia AM,
600 Sethi MK, Pabst O, Marzano M, Moretti S, Romani L, Penna G, Pesole G, Rescigno M. 2015.
601 BALB/c and C57BL/6 mice differ in polyreactive IgA abundance, which impacts the generation of
602 antigen-specific IgA and microbiota diversity. *Immunity* 43:527–540.
- 603 68. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee
604 CA, Lynch SV, Tanoue T, Imaoka A, Itoh K, Takeda K, Umesaki Y, Honda K, Littman DR. 2009.
605 Induction of intestinal th17 cells by segmented filamentous bacteria. *Cell* 139:485–498.
- 606 69. Azrad M, Hamo Z, Tkhawkho L, Peretz A. 2018. Elevated serum immunoglobulin a levels in

- 607 patients with *Clostridium difficile* infection are associated with mortality. *Pathogens and Disease* 76.
- 608 70. Saleh MM, Frisbee AL, Leslie JL, Buonomo EL, Cowardin CA, Ma JZ, Simpson ME, Scully KW,
609 Abhyankar MM, Petri WA. 2019. Colitis-induced th17 cells increase the risk for severe subsequent
610 *Clostridium difficile* infection. *Cell Host & Microbe* 25:756–765.e5.
- 611 71. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a
612 dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the
613 MiSeq illumina sequencing platform. *Applied and Environmental Microbiology* 79:5112–5120.
- 614 72. Sze MA, Schloss PD. 2019. The impact of DNA polymerase and number of rounds of
615 amplification in PCR on 16S rRNA gene sequence data. *mSphere* 4.
- 616 73. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
617 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Horn DJV, Weber CF.
618 2009. Introducing mothur: Open-source, platform-independent, community-supported software
619 for describing and comparing microbial communities. *Applied and Environmental Microbiology*
620 75:7537–7541.
- 621 74. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012.
622 The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools.
623 *Nucleic Acids Research* 41:D590–D596.
- 624 75. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR,
625 Tiedje JM. 2013. Ribosomal database project: Data and tools for high throughput rRNA analysis.
626 *Nucleic Acids Research* 42:D633–D642.
- 627 76. Westcott SL, Schloss PD. 2017. OptiClust, an improved method for assigning amplicon-based
628 sequence data to operational taxonomic units. *mSphere* 2.
- 629 77. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB,
630 Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2018. Vegan: Community ecology

631 package.

632 78. R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for
633 Statistical Computing, Vienna, Austria.

634 79. Kuhn M. 2008. Building predictive models inRUsing thecaretPackage. Journal of Statistical
635 Software 28.

636 80. Topçuoğlu BD, Lesniak NA, Ruffin MT, Wiens J, Schloss PD. 2020. A framework for effective
637 application of machine learning to microbiome-based classification problems. mBio 11.

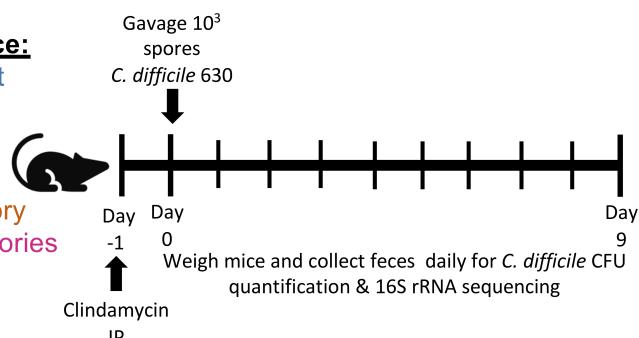
638 81. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes
639 A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D,
640 Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the
641 tidyverse. Journal of Open Source Software 4:1686.

642 **Figures**

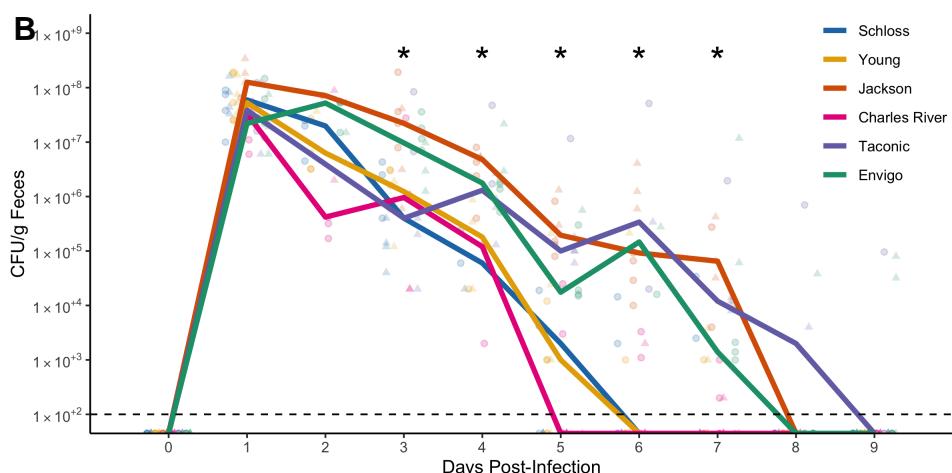
A

Sources of C57BL/6 mice:

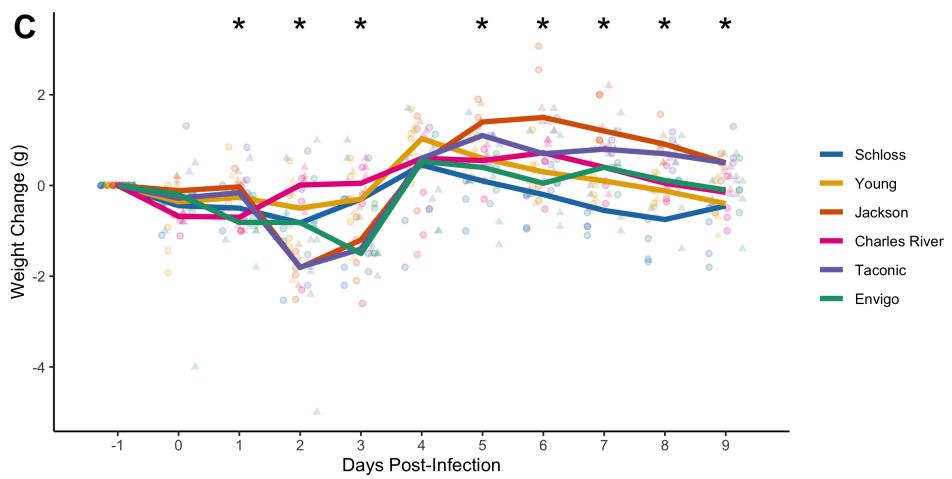
1. Schloss Lab Colony at University of Michigan
2. Young Lab Colony at University of Michigan
3. The Jackson Laboratory
4. Charles River Laboratories
5. Taconic Biosciences
6. Envigo



B



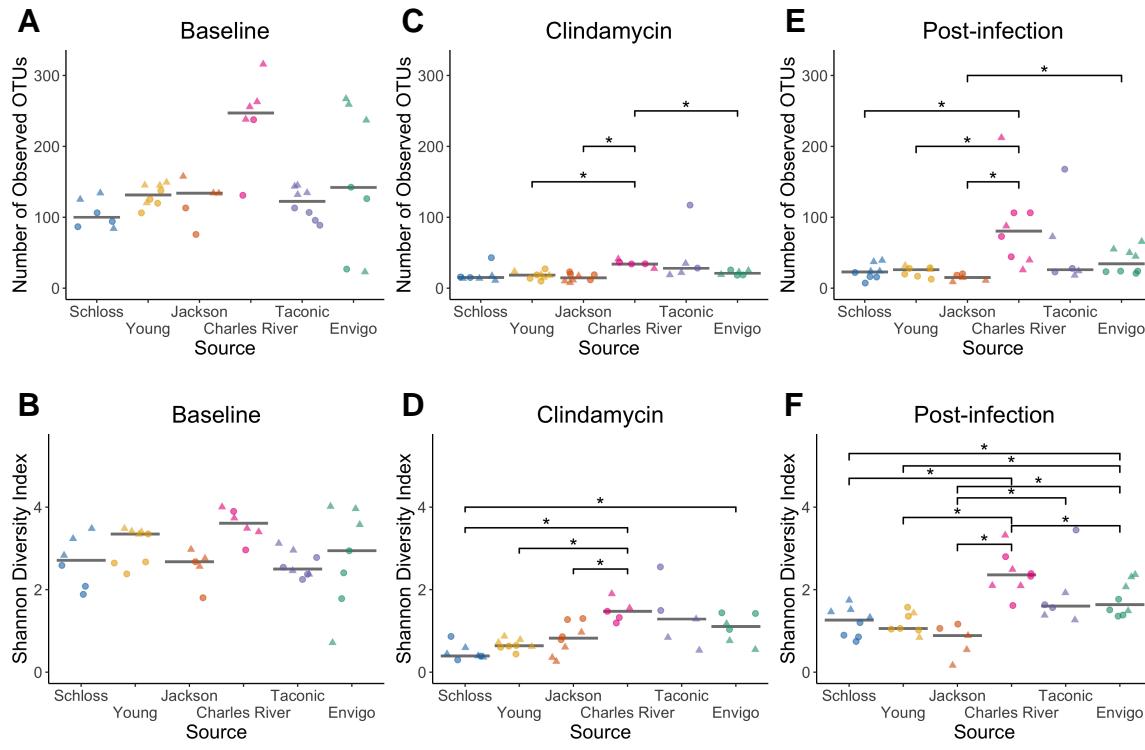
C



643 **Figure 1. Clindamycin**

644 **is sufficient to promote *C. difficile* colonization in all mice, but clearance time varies across**
 645 **sources of C57BL/6 mice.** A. Setup of the experimental timeline. Mice for the experiments
 646 were obtained from 6 different sources: the Schloss ($N = 8$) and Young lab ($N = 9$) colonies at

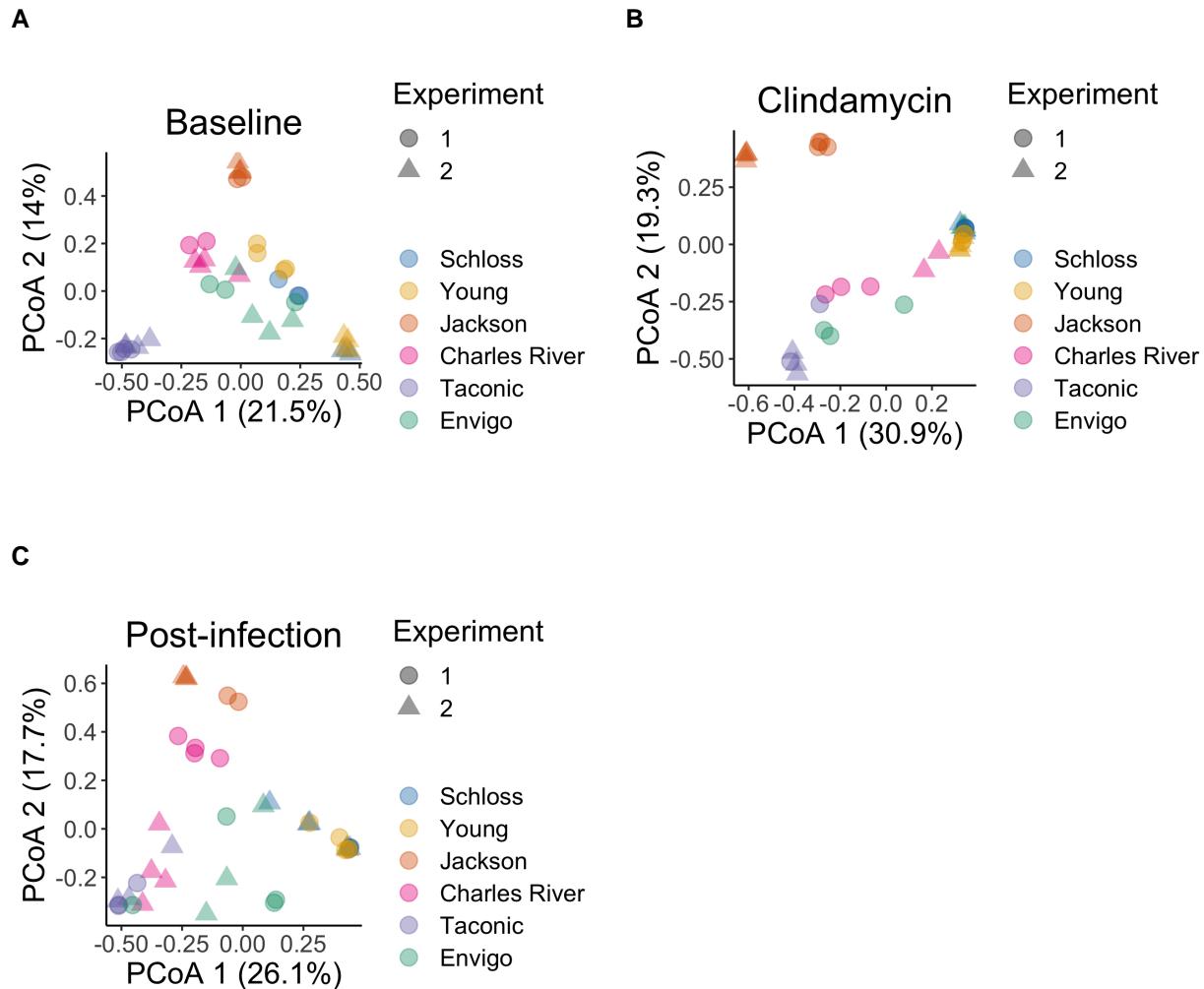
647 the University of Michigan, the Jackson Laboratory (N = 8), Charles River Laboratory (N = 8),
648 Taconic Biosciences (N = 8), and Envigo (N = 8). All mice were administered 10 mg/kg clindamycin
649 intraperitoneally (IP) 1 day before challenge with *C. difficile* 630 spores on day 0. Mice were
650 weighed and feces was collected daily through the end of the experiment (9 days post-infection).
651 Note: 3 mice died during course of experiment. 1 Taconic mouse prior to infection and 1 Jackson
652 and 1 Envigo mouse between 1- and 3-days post-infection. B. *C. difficile* CFU/gram stool measured
653 over time (N = 20-49 mice per timepoint) via serial dilutions. The black line represents the limit of
654 detection for the first serial dilution. CFU quantification data was not available for each mouse due
655 to early deaths, stool sampling difficulties, and not plating all of the serial dilutions. C. Mouse weight
656 change measured in grams over time (N = 45-49 mice per timepoint), all mice were normalized to
657 the weight recorded 1 day before infection. For B-C: timepoints where differences across sources
658 of mice were statistically significant by Kruskal-Wallis test with Benjamini-Hochberg correction for
659 testing across multiple days (Table S1 and Table S2) are reflected by the asterisk(s) above each
660 timepoint (*, $P < 0.05$). Lines represent the median for each source and circles represent individual
661 mice from experiment 1 while triangles represent mice from experiment 2.



Figure

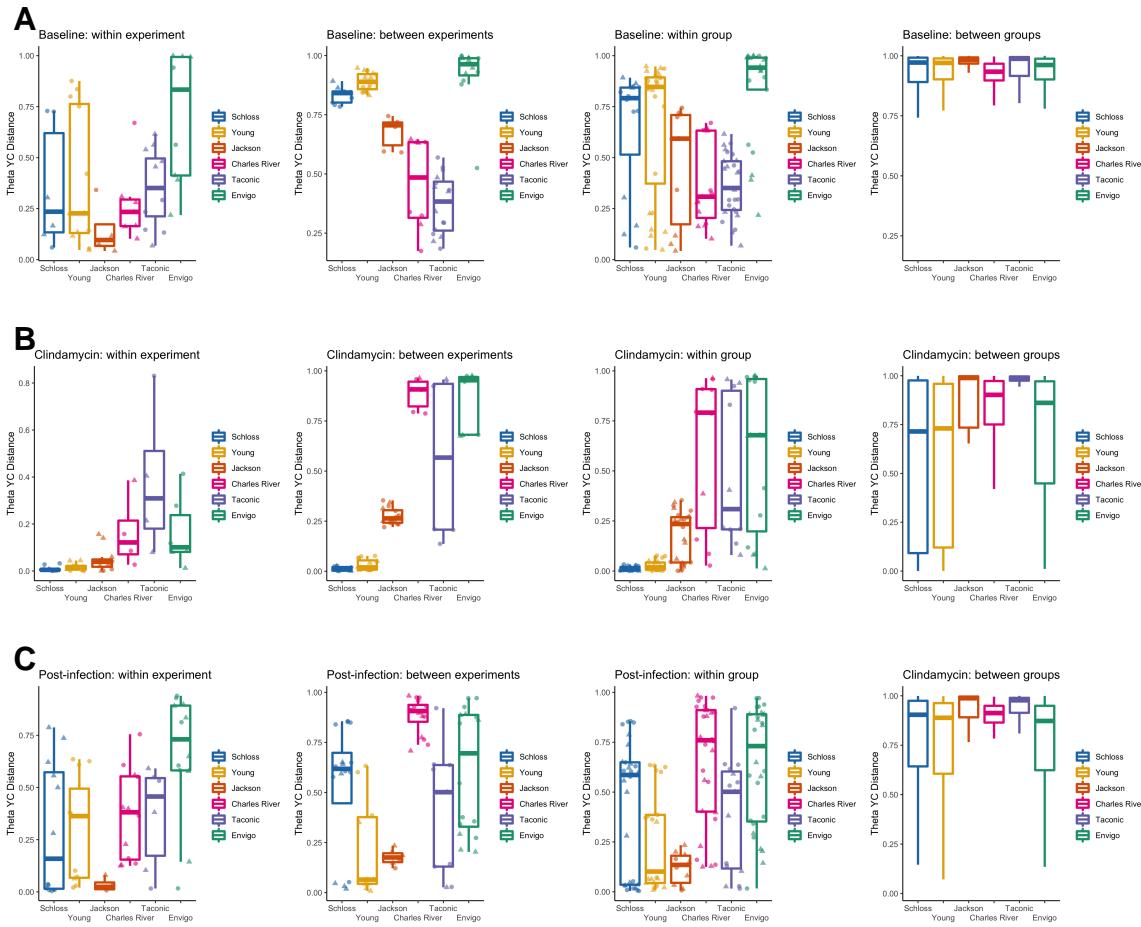
662

663 **2. Differences in microbial richness and diversity across mouse colony sources emerge**
 664 **after clindamycin treatment and infection.** A-F. Number of observed OTUs and Shannon
 665 diversity index values at baseline: day -1 (A-B), after clindamycin: day 0 (C-D) and post-infection:
 666 day 1 (E-F) timepoints of the experiment. Data were analyzed by Kruskal-Wallis test with
 667 Benjamini-Hochberg correction for testing each day of the experiment and the adjusted *P* value
 668 was < 0.05 for all panels except for B (Table S3). Significant *P* values from the pairwise Wilcoxon
 669 comparisons between sources with Benjamini-Hochberg correction are shown (Table S4). For
 670 A-F: circles represent experiment 1 mice, while triangles represent experiment 2 mice with each
 671 symbol representing the value for a stool sample from an individual mouse. Gray lines represent
 672 the median values for each source of mice.



673

674 **Figure 3. Mouse colony source is the variable that explains most of the variation observed**
 675 **in the baseline, post-clindamycin, and post-infection bacterial communities.** A-C. Principal
 676 Coordinates Analysis of Theta YC distances from stools collected at baseline (A), post-clindamycin
 677 (B), and post-infection (C) timepoints of the experiment. Each symbol represents a stool sample
 678 from an individual mouse, with circles representing experiment 1 mice and triangles representing
 679 experiment 2 mice. PERMANOVA analysis demonstrated that source and the interaction between
 680 source and cage explained most of the variation observed in the baseline (combined $R^2 = 0.90$),
 681 post-clindamycin (combined $R^2 = 0.99$), and post-infection (combined $R^2 = 0.88$) communities (all
 682 $P = 0.0001$, see Table S6).

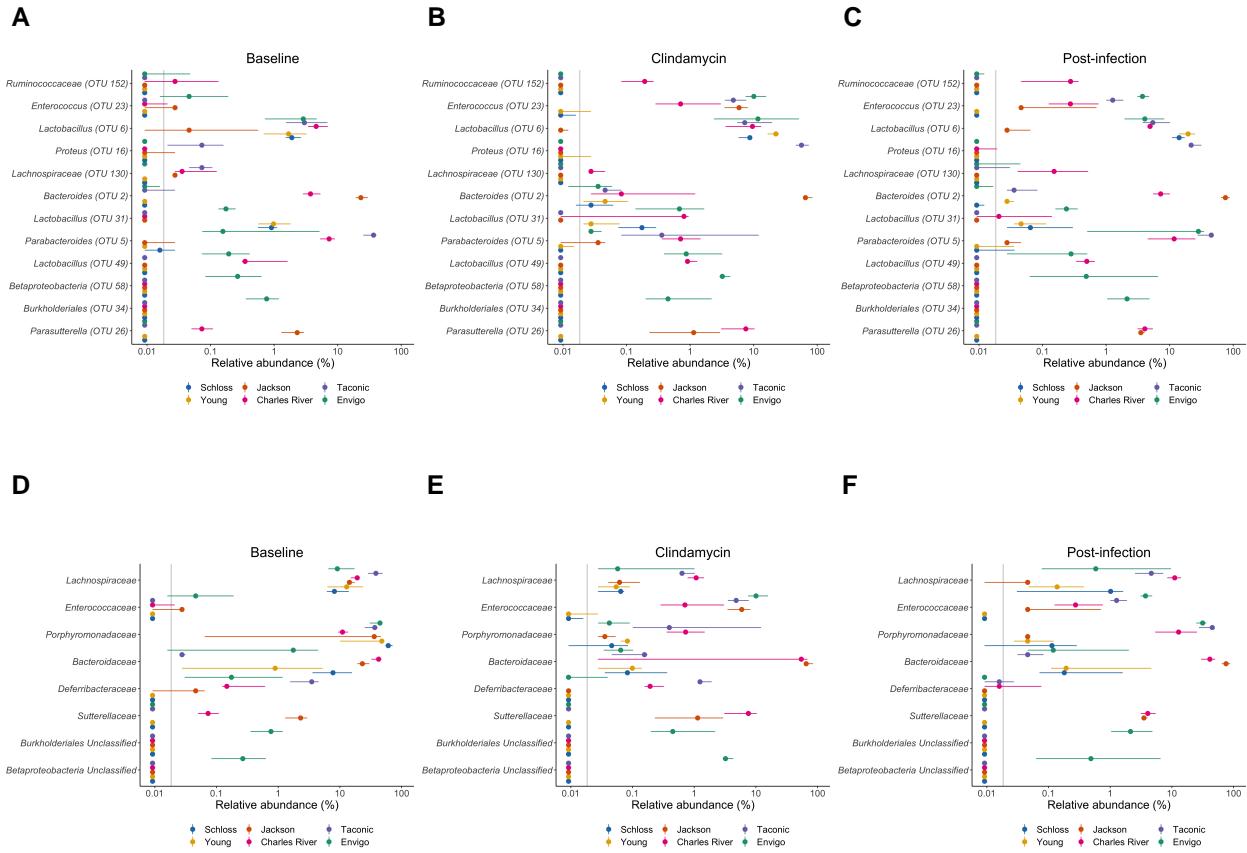


Figure

683

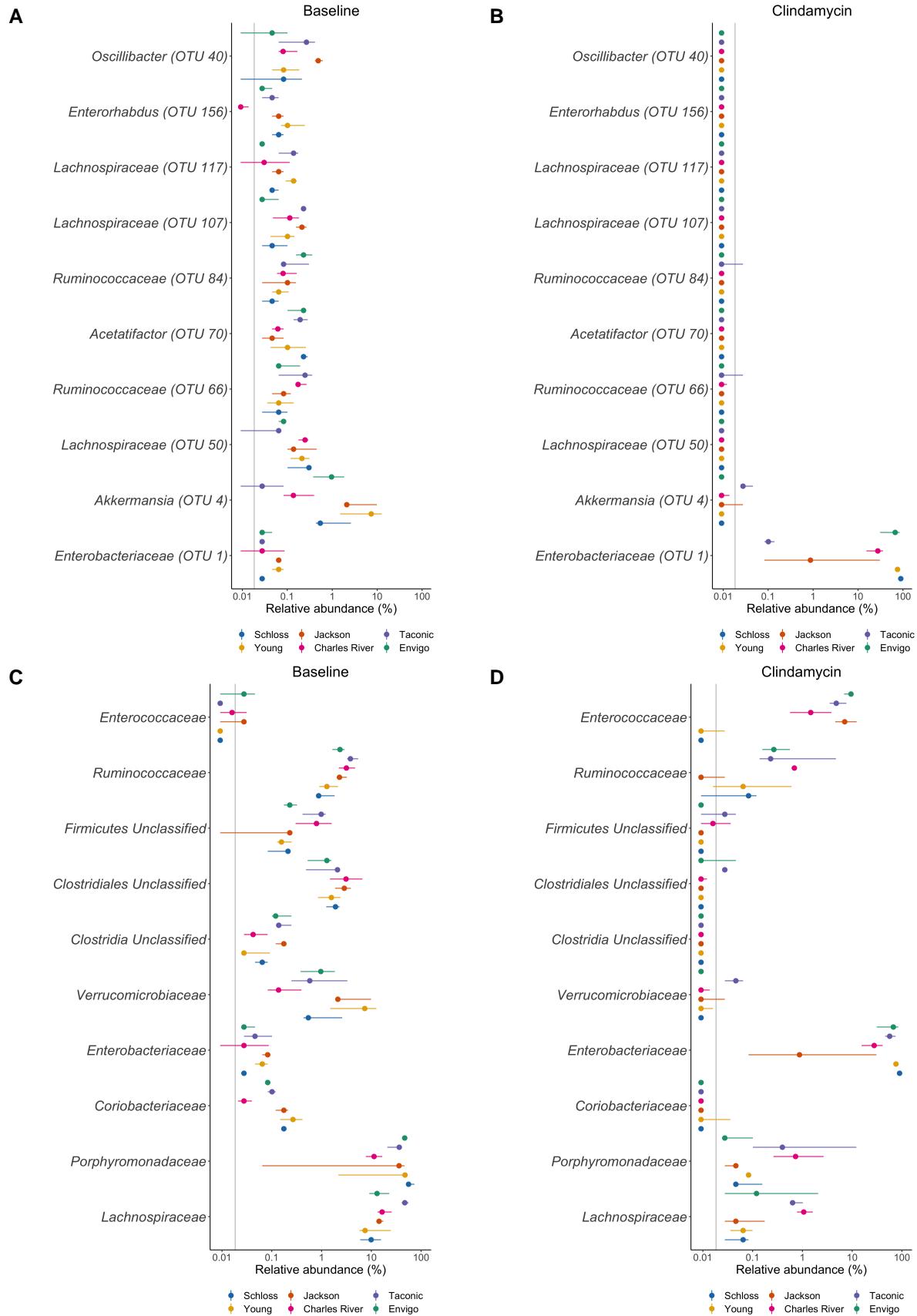
684 **4. High inter-group variation across mouse sources is diminished by clindamycin treatment**

685 A-C. Boxplots of the Theta YC distances of the 6 sources of mice relative to mice within the same
 686 source and experiment, mice within the same source and between experiments, mice within the
 687 same source, and mice from other groups at the baseline (A), after clindamycin treatment (B),
 688 and post-infection (C) timepoints. For comparisons within mice from the same source, symbols
 689 represent individual mouse samples: circles for experiment 1 and triangles for experiment 2.

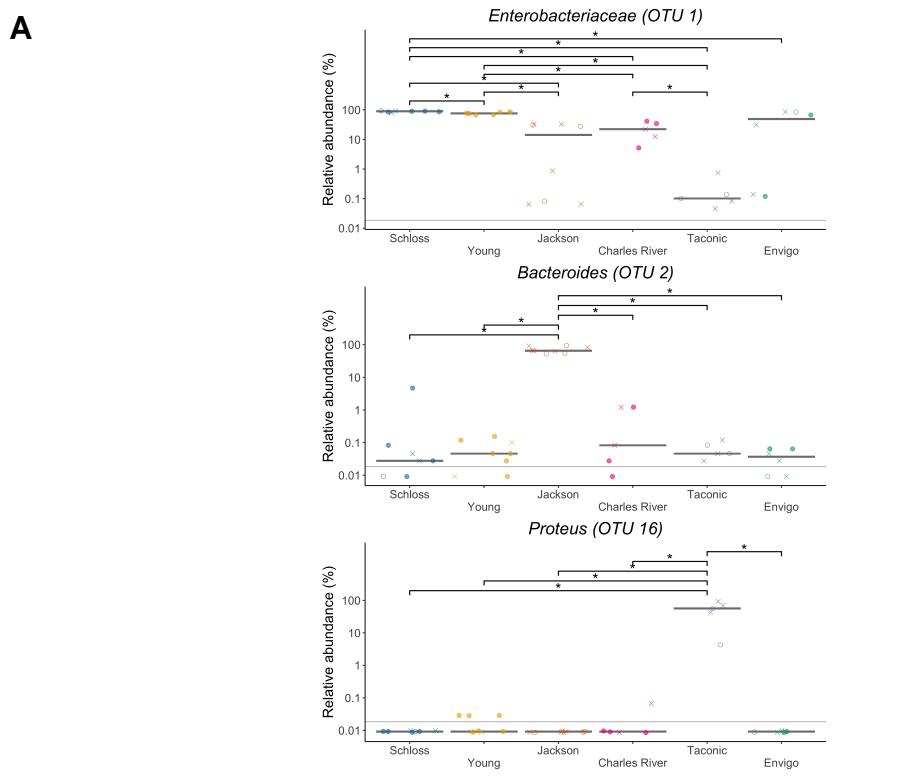


690

691 **Figure 5. A subset of bacteria consistently vary across mouse colony sources despite**
 692 **clindamycin perturbation and *C. difficile* challenge.** A-C: plots highlighting the median (point)
 693 and interquartile range (colored lines) of the relative abundances for the 12 OTUs that consistently
 694 varied across sources of mice at the baseline (A), post-clindamycin (B), and post-infection (C)
 695 timepoints of the experiment. D-F: plots highlighting the median (point) and interquartile range
 696 (colored lines) of the relative abundances for the 8 families that consistently varied across sources
 697 of mice at the baseline (D), post-clindamycin (E), and post-infection (F) timepoints of the experiment.
 698 For each timepoint bacteria with differential relative abundances across sources of mice were
 699 identified by Kruskal-Wallis test at the family and OTU level with Benjamini-Hochberg correction for
 700 testing all identified taxa at the respective level (Table S8-9). The grey vertical line indicates the
 701 limit of detection for A-F.

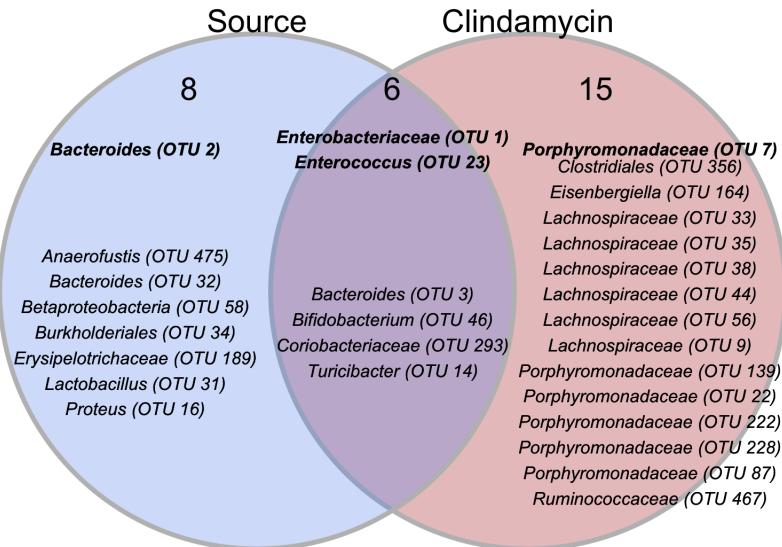


703 **Figure 6. Clindamycin treatment has the same effects on a subset of taxa regardless of**
704 **colony source.** A-B: plots highlighting the median (point) and interquartile range (colored lines) of
705 the top 10 most significant (adjusted P value < 0.05) OTUs with relative abundances that changed
706 after clindamycin treatment. C-D: plots highlighting the median (point) and interquartile range
707 (colored lines) of the top 10 most significant families with relative abundances that changed after
708 clindamycin treatment. Data were analyzed by Wilcoxon signed rank test limited to mice that had
709 paired sequence data for day -1 and 0 ($N = 31$). Tests were performed at the OTU and family levels
710 with Benjamini-Hochberg correction for testing all identified OTUs and families. See Table S10-11
711 for complete list of OTUs and families significantly impacted by clindamycin treatment. The grey
712 vertical line indicates the limit of detection for A-D.



B

Key taxa comparisons for day -1, 0, and 1 models

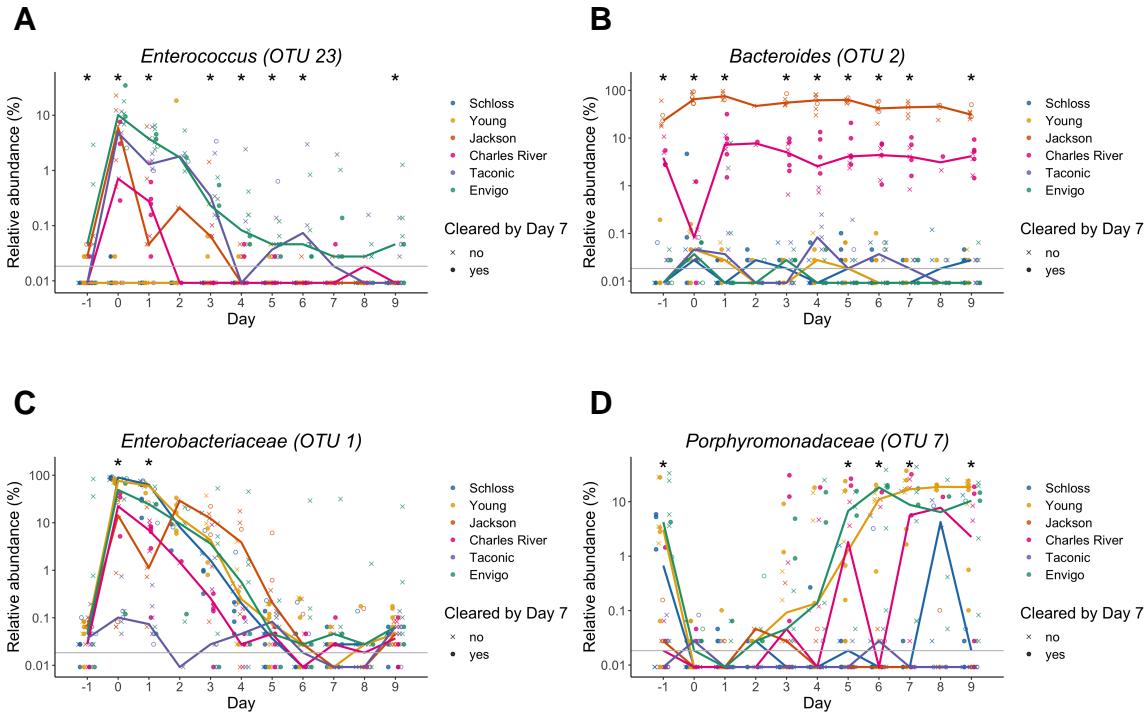


713

Figure

714 **7. Key OTUs that influence whether mice cleared *C. difficile* by day 7.** A. Baseline relative
 715 abundance data for 3 of the OTUs from the classification model based on day 0 OTU relative
 716 abundances that significantly varied across sources of mice and had high relative abundances in

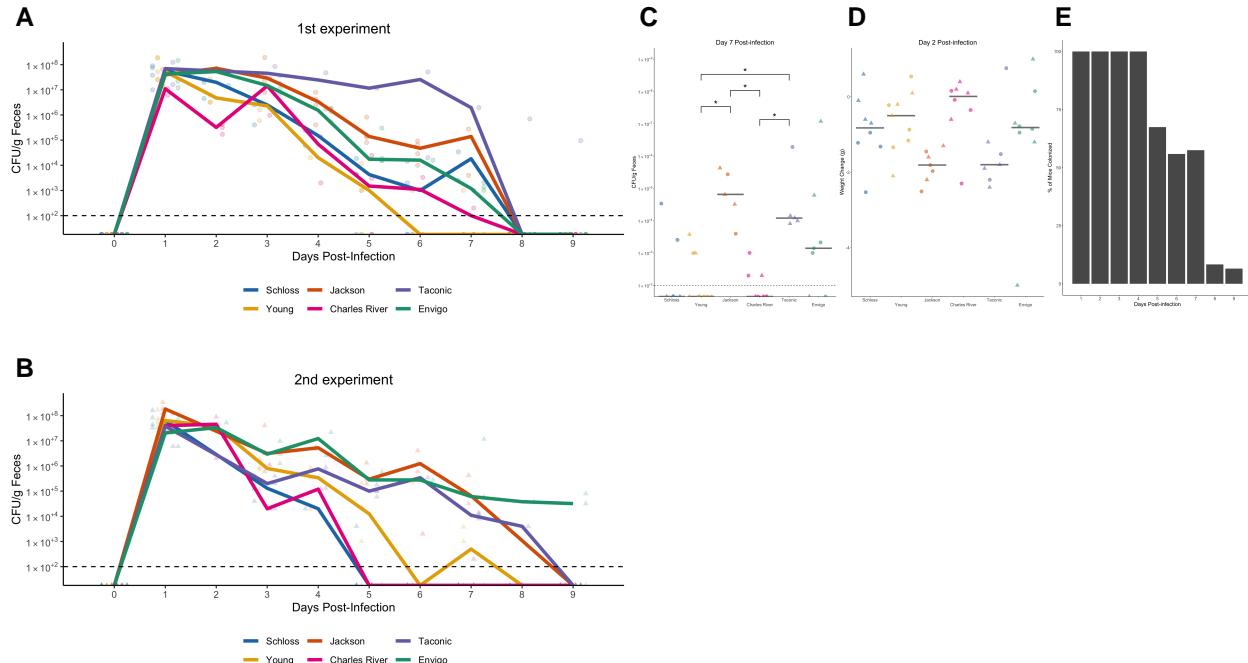
717 the community. Symbols represent the relative abundance data for an individual mouse, circles
718 represent mice that cleared *C. difficile* by day 7, X-shapes represent mice that were still colonized
719 with *C. difficile*, and open circles represent mice that did not have *C. difficile* CFU counts for day 7
720 post-infection. Gray lines indicate the median relative abundances for each source. Asterisks are
721 shown for pairwise Wilcoxon comparisons with Benjamini-Hochberg correction where $P < 0.05$. B.
722 Venn diagram that combines Fig. S4 summaries of OTUs that were important to the day -1, 0, and
723 1 classification models (Table S14) and either overlapped with taxa that varied across vendors at
724 the same timepoint, were impacted by clindamycin treatment, or both. See Fig. S4 for separate
725 comparisons of taxa from the day -1, 0, and 1 classification models. Bold OTUs signify OTUs that
726 were important to more than 1 classification model.



Figure

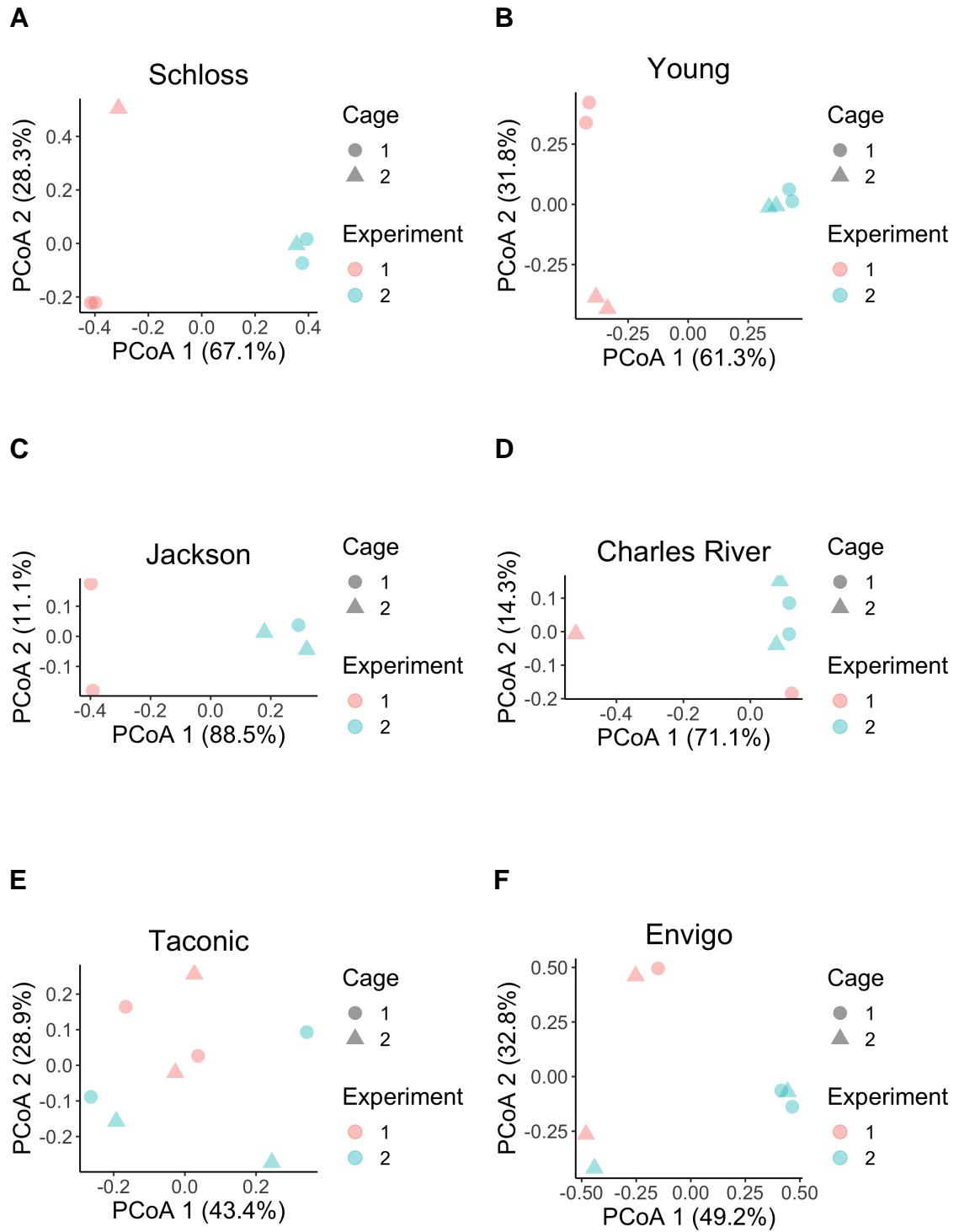
727

728 **8: Key OTUs vary across sources throughout the experiment.** A-C. Relative abundances of
 729 bold OTUs from Fig. 7A that were important for at least two classification models are shown over
 730 time. A. *Enterococcus* (OTU 23), which significantly varied across sources and was impacted
 731 by clindamycin treatment. B. *Bacteroides* (OTU 2), which varied across sources throughout the
 732 experiment. C. *Enterobacteriaceae* (OTU 1) and *Porphyromonadaceae* (OTU 7) were significantly
 733 impacted by clindamycin treatment and examining relative abundance dynamics over the course
 734 of the experiment indicated timepoints where relative abundances also significantly varied across
 735 sources of mice. Symbols represent the relative abundance data for an individual mouse, circles
 736 represent mice that cleared *C. difficile* by day 7, X-shapes represent mice that were still colonized
 737 with *C. difficile*, and open circles represent mice that did not have *C. difficile* CFU counts for day 7
 738 post-infection. Colored lines indicate the median relative abundances for each source. The gray
 739 horizontal line represents the limit of detection. Timepoints where differences across sources of
 740 mice were statistically significant by Kruskal-Wallis test with Benjamini-Hochberg correction for
 741 testing across multiple days (Table S16) are identified by the asterisk(s) above each timepoint (*, P
 742 < 0.05).



743

744 **Figure S1. *C. difficile* CFU variation across vendors varies slightly across the 2**
 745 **experiments.** A-B. *C. difficile* CFU/gram of stool quantification over time for experiment 1
 746 (A) and 2 (B). Experiments were conducted approximately 3 months apart. Lines represent the
 747 median CFU for each source, symbols represent individual mice and the black line represents the
 748 limit of detection. C. *C. difficile* CFU/gram stool on day 7 post-infection across sources of mice
 749 with asterisks for pairwise Wilcoxon comparisons with Benjamini-Hochberg correction where $P <$
 750 0.05. D. Mouse weight change 2 days post-infection across sources of mice, no pairwise Wilcoxon
 751 comparisons were significant after Benjamini-Hochberg correction. For C-D: circles represent
 752 experiment 1 mice, triangles represent experiment 2 mice and gray lines indicate the median
 753 values for each group. E. Percent of mice that were colonized with *C. difficile* over the course of the
 754 experiment. Each day the percent is calculated based on the mice where *C. difficile* CFU was
 755 quantified for that particular day. Total N for each day: day 1 (N = 42), day 2 (N = 20), day 3 (N =
 756 39), day 4 (N = 29), day 5 (N = 43), day 6 (N = 34), day 7 (N = 40), day 8 (N = 36), and day 9 (N =
 757 46).



Figure

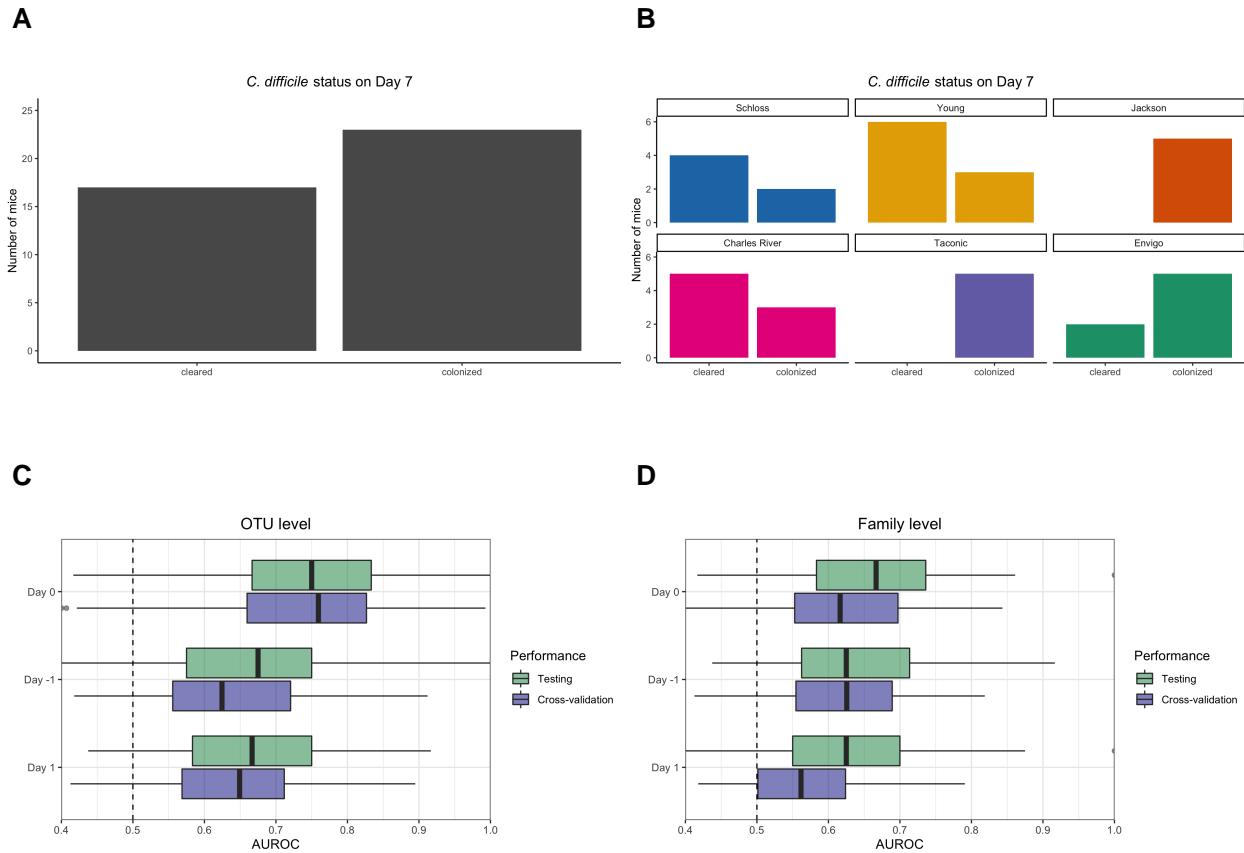
758

759 **S2. Only bacterial communities from University of Michigan mice significantly vary between**

760 **experiments.** A-F. PCoA of Theta YC distances for the baseline fecal bacterial communities within

761 each source of mice. Each symbol represents a stool sample from an individual mouse with color

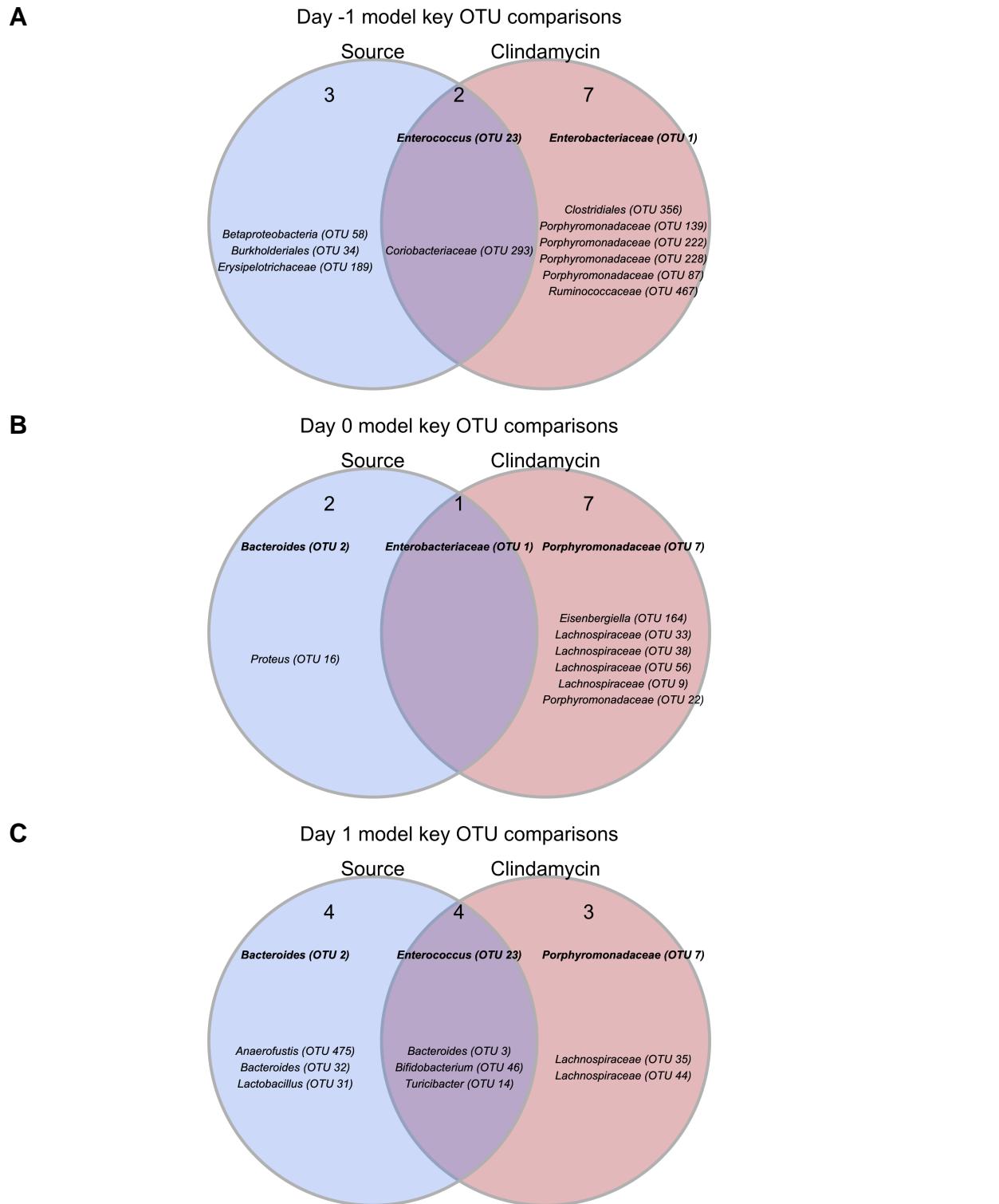
⁷⁶² corresponding to experiment and shape representing cage mates. PERMANOVA was performed
⁷⁶³ within each group to examine the contributions of experiment and cage to observed variation.
⁷⁶⁴ Experiment number and cage only significantly explained observed variation for mice from the
⁷⁶⁵ Schloss (combined $R^2 = 0.99$; $P \leq 0.033$) and Young (combined $R^2 = 0.95$; $P \leq 0.027$) lab colonies
⁷⁶⁶ (Table S7).



767

768 **Figure S3. Bacterial community composition before, after clindamycin perturbation, and**
 769 **post-infection can predict *C. difficile* colonization status 7 days post-challenge.** A. Bar
 770 graph visualizations of overall day 7 *C. difficile* colonization status that were used as classification
 771 outcomes to build L2-regularized logistic regression models. Mice were classified as colonized or
 772 cleared (not detectable at the limit of detection of 100 CFU) based on CFU g/stool data from 7 days
 773 post-infection. B. *C. difficile* CFU status on Day 7 within each mouse colony source. N = 5-9 mice
 774 per group. C-D. L2-regularized logistic regression classification model area under the receiving
 775 operator characteristic curve (AUROCs) to predict *C. difficile* CFU on day 7 post-infectoin (Fig.
 776 1D, Fig. S3) based on the community relative abundances at baseline (day -1), post-clindamycin
 777 (day 0), and post-infection (day 1) at either the OTU (C) or family (D) level. All models performed
 778 better than random chance (AUROC = 0.5; all $P \leq 5e-15$; Table S12) and the model built with
 779 post-clindamycin treated bacterial OTU relative abundances had the best performance ($P_{FDR} \leq$
 780 3.9e-10 for all pairwise comparisons; Table S13). For lists of the 20 taxa that were ranked as most

⁷⁸¹ important to each model, see Tables S14-15.



Figure

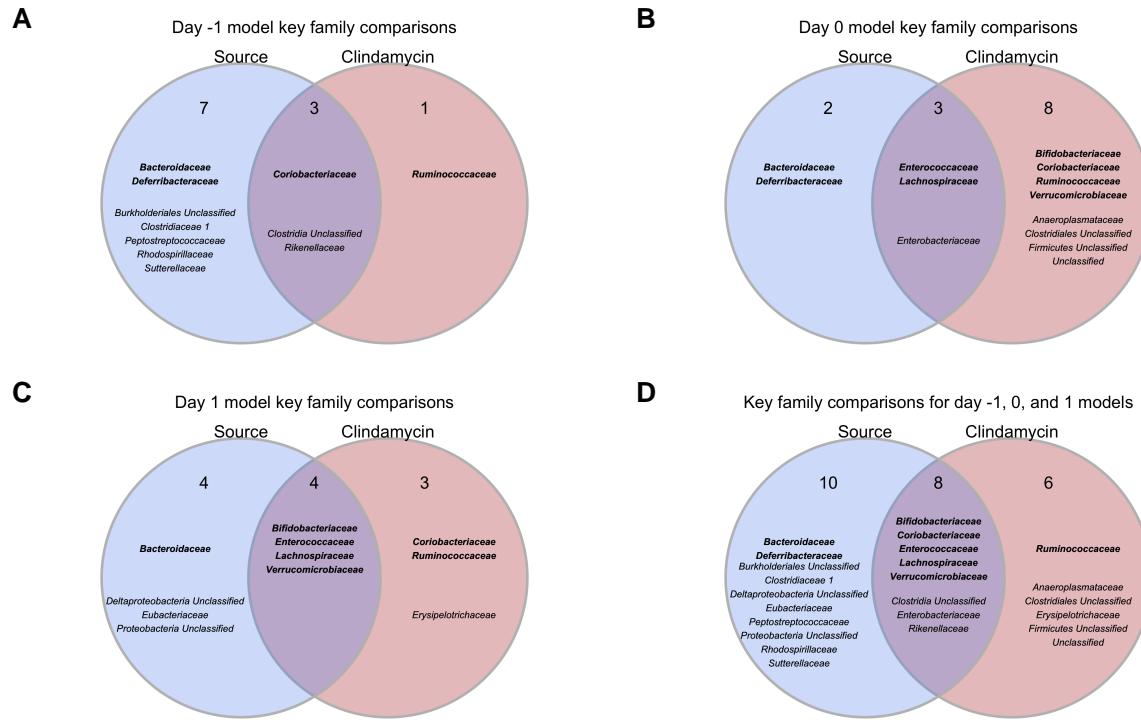
782

783 **S4. Key OTUs from classification models based on baseline, post-clindamycin treatment,**

784 **or post-infection community data vary by mouse colony source, clindamycin treatment, or**

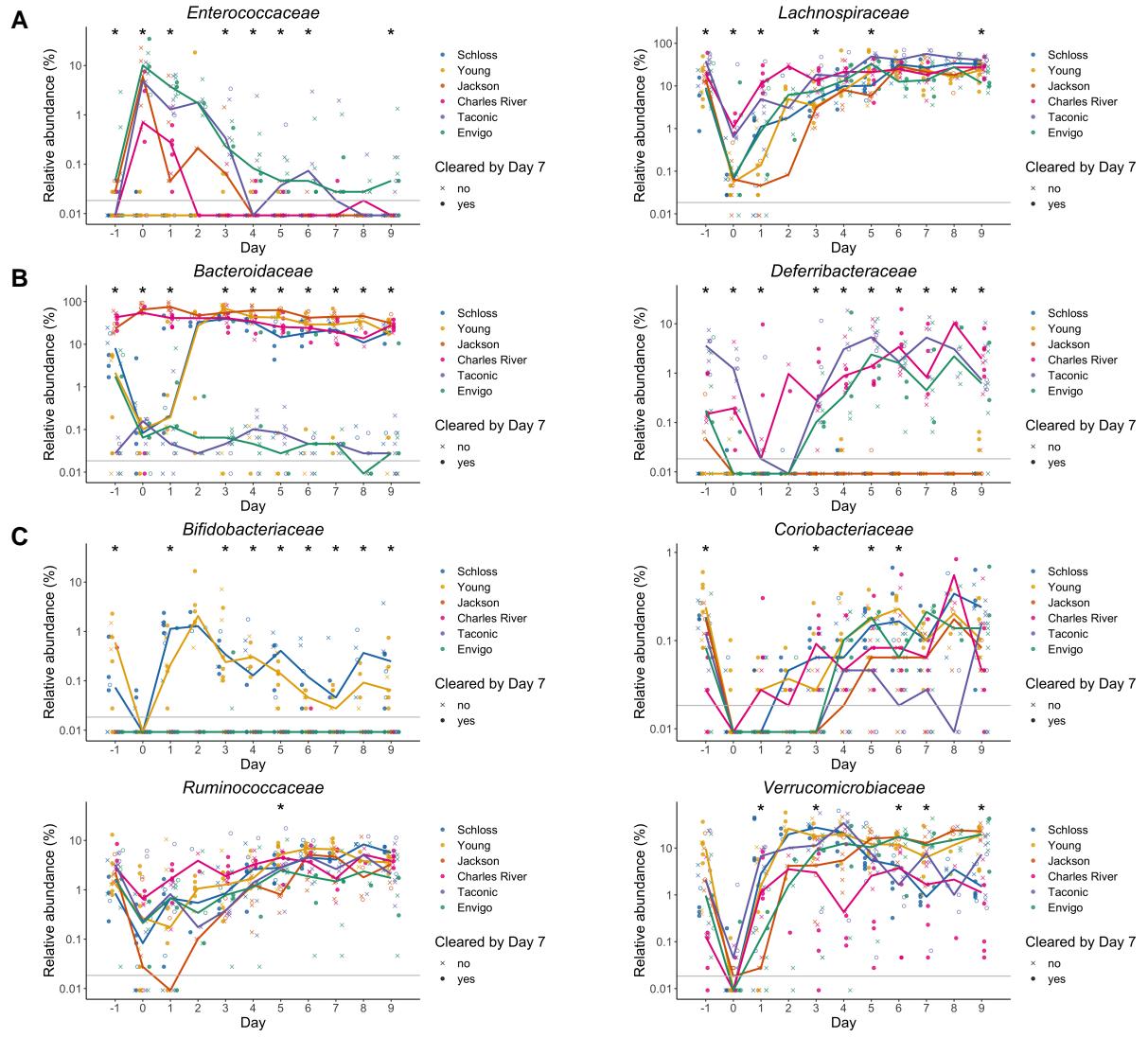
785 **both.** A-C. Venn diagrams of top 20 important OTUs from baseline (A), post-clindamycin treatment

⁷⁸⁶ (B), and post-infection (C) classification models (Table S14) that overlapped with OTUs that varied
⁷⁸⁷ across vendors at baseline, were impacted by clindamycin treatment, or both. Bold OTUs signify
⁷⁸⁸ OTUs that were important to more than 1 classification model.



789

790 **Figure S5. Key families from classification models based on baseline, post-clindamycin**
 791 **treatment, or post-infection community data vary by mouse colony source, clindamycin**
 792 **treatment, or both.** A-C. Venn diagrams of top 20 important families from baseline (A),
 793 post-clindamycin treatment (B), and post-infection (C) classification models (Table S15) that
 794 overlapped with families that varied across vendors after clindamycin, were impacted by clindamycin
 795 treatment, or both. D. Venn diagrams that combines A-C summaries of families that were important
 796 to the day -1, 0, and 1 classification models (Table S15) and either overlapped with families that
 797 varied across vendors at the same timepoint, were impacted by clindamycin treatment, or both.
 798 Bold families signify families that were important to more than 1 classification model.



799

800 **Figure S6. Key families vary across sources throughout experiment.** Relative abundances of
801 bold families from Fig. S5D that were important for at least two classification models are shown
802 over time. A. *Enterococcaceae* and *Lachnospiraceae*, which significantly varied across sources
803 and were impacted by clindamycin treatment. B. *Bacteroidaceae* and *Deferribacteraceae*, which
804 varied across sources throughout the experiment. C. *Bifidobacteriaceae*, *Coriobacteriaceae*,
805 *Ruminococcaceae*, and *Verrucomicrobiaceae* were significantly impacted by clindamycin treatment.
806 Examining the relative abundance dynamics throughout the experiment, identified timepoints
807 where relative abundances also significantly varied across sources of mice. Symbols represent the
808 relative abundance data for an individual mouse, circles represent mice that cleared *C. difficile* by
809 day 7, X-shapes represent mice that were still colonized with *C. difficile*, and open circles represent

810 mice that did not have *C. difficile* CFU counts for day 7 post-infection. Colored lines indicate the
811 median relative abundances for each source. The gray horizontal line represents the limit of
812 detection. Timepoints where differences across sources of mice were statistically significant by
813 Kruskal-Wallis test with Benjamini-Hochberg correction for testing across multiple days (Table S17)
814 are identified by the asterisk(s) above each timepoint (*, P < 0.05).

815 **Supplementary Tables and Movie**

816 All supplemental material is available at: https://github.com/SchlossLab/Tomkovich_vendor_difs_
817 XXXX_2020/submission.

818 **Movie S1. Large shifts in bacterial community structure occurred after clindamycin and *C.***
819 ***difficile* infection.** PCoA of Theta YC distances animated from 0 through 9 days post-infection.
820 PERMANOVA analysis indicated colony source was the variable that explained the most observed
821 variation across fecal communities (source $R^2 = 0.35$, $P = 0.0001$) followed by interactions between
822 cage and day of the experiment. Transparency of the circle corresponds to the day of the experiment,
823 each circle represents a sample from an individual mouse at a specific timepoint. See Table S5
824 for PERMANOVA results). Circles represent mice from experiment 1 and triangles represent mice
825 from expeirment 2.

826 **Table S1. *C. difficile* CFU statistical results.**

827 **Table S2. Mouse weight change statistical results.**

828 **Table S3. Diversity metrics Kruskal-Wallis statistical results.**

829 **Table S4. Diversity metrics pairwise Wilcoxon statistical results.**

830 **Table S5. PERMANOVA results for all mice, all timepoints.**

831 **Table S6. PERMANOVA results for all mice at baseline, post clindamycin, and post-infection**
832 **timepoints.**

833 **Table S7. PERMANOVA results of baseline communities within each source.**

834 **Table S8. OTUs with relative abudances that significantly vary across sources at baseline,**
835 **post-clindamycin, or post-infection timepoints.**

836 **Table S9. Families with relative abudances that significantly vary across sources at**
837 **baseline, post-clindamycin, or post-infection timepoints.**

838 **Table S10. OTUs with relative abudances that significantly changed after clindamycin**

839 treatment.

840 **Table S11.** Families with relative abundances that significantly changed after clindamycin
841 treatment.

842 **Table S12.** Statistical results of L2-regularized logistic regression model performances
843 compared to random chance.

844 **Table S13.** Pairwise Wilcoxon results comparing all 6 L2-regularized logistic regression
845 model performances.

846 **Table S14.** Top 20 most important OTUs for each of the 3 L2-regularized logistic regression
847 models based on OTU relative abundance data.

848 **Table S15.** Top 20 most important families for each of the 3 L2-regularized logistic
849 regression models based on OTU relative abundance data.

850 **Table S16.** OTUs with relative abundances that significantly varied across sources of mice
851 on at least 1 day of the experiment by Kruskal-Wallis test.

852 **Table S17.** Families with relative abundances that significantly varied across sources of mice
853 on at least 1 day of the experiment by Kruskal-Wallis test.