

The 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 gut microbiota has a key role in determining susceptibility to *Clostridioides difficile* infections
3 (CDIs). However, much of the mechanistic work examining CDIs in mouse models use animals
4 obtained from a single source. We treated mice from 6 sources (2 University of Michigan colonies
5 and 4 vendors) with 1 clindamycin dose, followed by *C. difficile* challenge 1 day later and then
6 measured *C. difficile* colonization levels through 9 days post-infection. The microbiota were profiled
7 via 16S rRNA gene sequencing to examine the variation across sources and alterations due to
8 clindamycin treatment and *C. difficile* challenge. While all mice were colonized 1-day post-infection,
9 variation emerged from days 3-7 post-infection with animals from some sources colonized with
10 *C. difficile* for longer and at higher levels. We identified bacteria that varied in relative abundance
11 across sources and throughout the experiment. Some bacteria were consistently impacted by
12 clindamycin treatment in all sources of mice including *Lachnospiraceae*, *Ruminococcaceae*, and
13 *Enterobacteriaceae*. To identify bacteria that were most important to colonization regardless of
14 the source, we created logistic regression models that successfully classified mice based on
15 whether they cleared *C. difficile* by 7 days post-infection using baseline, post-clindamycin, and
16 1-day post-infection community composition data. With these models, we identified 4 bacteria that
17 were predictive of whether *C. difficile* cleared. They varied across sources (*Bacteroides*), were
18 altered by clindamycin (*Porphyromonadaceae*), or both (*Enterobacteriaceae* and *Enterococcus*).
19 Microbiota variation across sources better emulates human inter-individual variation and can help
20 identify bacterial drivers of phenotypic variation in the context of CDIs.

21 **Importance**

22 *Clostridioides difficile* is a leading nosocomial infection. Although perturbation to the gut microbiota
23 is an established risk, there is variation in who becomes asymptotically colonized, develops
24 an infection, or has adverse infection outcomes. Mouse models of *C. difficile* infection (CDI) are
25 widely used to answer a variety of *C. difficile* pathogenesis questions. However, the inter-individual
26 variation between mice from the same breeding facility is less than what is observed in humans.
27 Therefore, we challenged mice from 6 different breeding colonies with *C. difficile*. We found that the
28 starting microbial community structures and *C. difficile* persistence varied by the source of mice.

29 Interestingly, a subset of the bacteria that varied across sources were associated with how long *C.*
30 *difficile* was able to colonize. By increasing the inter-individual diversity of the starting communities,
31 we were able to better model human diversity. This provided a more nuanced perspective of *C.*
32 *difficile* pathogenesis.

33 Introduction

34 Antibiotics are a common risk factor for *Clostridioides difficile* infections (CDIs) due to their effect on
35 the intestinal microbiota, but there is variation in who goes on to develop severe or recurrent CDIs
36 after exposure (1, 2). Additionally, asymptomatic colonization, where *C. difficile* is detectable, but
37 symptoms are absent, has been documented in infants and adults (3, 4). The intestinal microbiota
38 has been implicated in asymptomatic colonization (5, 6), susceptibility to CDIs (7), and adverse CDI
39 outcomes (9–12). However, it is not clear how much inter-individual microbiota variation contributes
40 to the range of outcomes observed after *C. difficile* exposure relative to other risk factors.

41 Mouse models of CDIs have been a great tool for understanding *C. difficile* pathogenesis (13). The
42 number of CDI mouse model studies has grown substantially since Chen et al. published their
43 C57BL/6 model in 2008, which disrupted the gut microbiota with antibiotics to enable *C. difficile*
44 colonization and symptoms such as diarrhea and weight loss (14). CDI mouse models have been
45 used to examine translationally relevant questions regarding *C. difficile*, including the role of the
46 microbiota and efficacy of potential therapeutics for treating CDIs (15). However, variation in the
47 microbiota between mice from the same breeding colony is much less than the inter-individual
48 variation observed between humans (16, 17). Studying CDIs in mice with a homogenous microbiota
49 is likely to overstate the importance of individual mechanisms. Using mice that have a more
50 heterogenous microbiota would allow researchers to identify and validate more generalizable
51 mechanisms responsible for CDI.

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

61 **Results**

62 **The variation in the microbiota is high between mice from different sources.** We obtained
63 C57BL/6 mice from 6 different sources: two colonies from the University of Michigan that were
64 split from each other in 2010 (the Young and Schloss lab colonies) and four commercial sources:
65 the Jackson Laboratory, Charles River Laboratories, Taconic Biosciences, and Envigo (which was
66 formerly Harlan). These 4 vendors were chosen because they are commonly used for murine CDI
67 studies (26, 34–40). Two experiments were conducted, approximately 3 months apart.

68 We sequenced the 16S rRNA gene from fecal samples collected from these mice after they
69 acclimated to the University of Michigan animal housing environment. We first examined the
70 alpha diversity across the 6 sources of mice. There was a significant difference in the richness
71 (i.e. number of observed operational taxonomic units (OTUs)), but not Shannon diversity index
72 ($P_{FDR} = 0.03$ and $P_{FDR} = 0.052$, respectively) across the sources of mice (Fig. 1A-B and Tables
73 S1-2). Next, we compared the community structure of mice (Fig. 1C). The source of mice and
74 the interactions between the source and cage effects explained most of the observed variation
75 between fecal communities (PERMANOVA combined $R^2 = 0.90$, $P < 0.001$, Fig. 1C and Table
76 S3). Mice that are co-housed tend to have similar gut microbiotas due to coprophagy (41). Since
77 mice within the same source were housed together, it was not surprising that the cage effect also
78 contributed to the observed community variation. There were some differences between the 2
79 experiments we conducted, as the experiment and cage effects significantly explained the observed
80 community variation for the Schloss and Young lab mouse colonies (Fig. S2A-B and Table S4).
81 However, most of the vendors also clustered by experiment (Fig. S2C-D, F), suggesting there was
82 some community variation between the 2 experiments within each source, particularly for Schloss,
83 Young, and Envigo mice (Fig. S2G-H). After finding differences at the community level, we next
84 identified the bacteria that varied between sources of mice. There were 268 OTUs with relative
85 abundances that were significantly different between the sources (Fig. 1D and Table S5). Though
86 we saw differences between experiments at the community level, there were no OTUs that were
87 significantly different between experiments within Schloss, Young, and Envigo mice (all $P > 0.05$).
88 By using mice from six sources we were able to increase the variation in the starting communities
89 to evaluate in a clindamycin-based CDI model.

90 **Clindamycin treatment renders all mice susceptible to *C. difficile* 630 colonization, but**
91 **clearance time varies across sources.** Clindamycin is frequently implicated with human CDIs
92 (42) and was part of the antibiotic treatment for the frequently cited 2008 CDI mouse model (14). We
93 have previously demonstrated mice are rendered susceptible to *C. difficile*, but clear the pathogen
94 within 9 days when treated with clindamycin alone (21, 43). All mice were treated with 10 mg/kg
95 clindamycin via intraperitoneal injection and one day later challenged with 10^3 *C. difficile* 630
96 spores (Fig. 2A). The day after infection, *C. difficile* was detectable in all mice at a similar level
97 (median CFU range: 2.2e+07-1.3e+08; $P_{FDR} = 0.15$), indicating clindamycin rendered all mice
98 susceptible regardless of source (Fig. 2B). However, between 3 and 7 days post-infection, we
99 observed variation in *C. difficile* levels across sources of mice (all $P_{FDR} \leq 0.019$; Fig. 2B and
100 Table S6). This suggested the mouse source was associated with *C. difficile* clearance. While the
101 colonization dynamics were similar between the two experiments, the Schloss mice took longer to
102 clear in the 1st experiment compared to the 2nd and the Envigo mice took longer to clear in the
103 2nd experiment compared to the 1st (Fig. S2A-B). The change in the mice's weight significantly
104 varied across sources of mice with the most weight loss occurring two days post-infection (Fig.
105 2C and Table S7). There was also one Jackson and one Envigo mouse that died between 1-
106 and 3-days post-infection during the second experiment. Mice obtained from Jackson, Taconic,
107 and Envigo tended to lose more weight, have higher *C. difficile* CFU levels and take longer to
108 clear the infection compared to the other sources of mice (although there was variation between
109 experiments with Schloss and Envigo mice). This was particularly evident 7 days post-infection
110 (Fig. 2B-C, Fig. S2C-D), when 57% of the mice were still colonized with *C. difficile* (Fig. S2E).
111 By 9 days post-infection the majority of the mice from all sources had cleared *C. difficile* (Fig.
112 2B) with the exception of 1 Taconic mouse from the first experiment and 2 Envigo mice from the
113 second experiment. Thus, clindamycin rendered all mice susceptible to *C. difficile* 630 colonization,
114 regardless of source, but there was significant variation in disease phenotype across the sources of
115 mice.

116 **Clindamycin treatment alters bacteria in all sources, but a subset of bacterial differences**
117 **across sources persists.** Given the variation in fecal communities that we observed across
118 breeding colonies, we hypothesized that variation in *C. difficile* clearance would be explained by

119 community variation across the 6 sources of mice. As expected, clindamycin treatment decreased
120 the richness and Shannon diversity across all sources of mice (Fig. 3A-B). Interestingly, significant
121 differences in diversity metrics between sources ($P_{FDR} < 0.05$) emerged after clindamycin treatment,
122 with Charles River mice having higher richness and Shannon diversity than most of the other sources
123 (Fig 3A-B and Tables S1-2). The clindamycin treatment decreased the variation in community
124 structures between sources of mice. The source of mice and the interactions between source and
125 cage effects explained almost all of the observed variation between communities (combined $R^2 =$
126 0.99, $P < 0.001$; Fig. 3C and Table S3). However, there were only 18 OTUs (Fig. 3D and Table S8)
127 with relative abundances that significantly varied between sources. Next, we identified the bacteria
128 that shifted after clindamycin treatment, regardless of source by analyzing paired fecal samples
129 from mice that were collected at baseline and after clindamycin treatment. We identified 153 OTUs
130 that were altered after clindamycin treatment in most mice (Fig. 3E and Table S9). When we
131 compared the list of significant clindamycin impacted bacteria with the bacteria that varied between
132 sources post-clindamycin, we found 4 OTUs (*Enterobacteriaceae* (OTU 1), *Lachnospiraceae* (OTU
133 130), *Lactobacillus* (OTU 6), *Enterococcus* (OTU 23)) overlapped (Fig. 3D-E and Tables S8-9).
134 Importantly, some of the OTUs that varied between sources also shifted with clindamycin treatment.
135 For example, *Proteus* increased after clindamycin treatment (Fig. 3D), but only in Taconic mice.
136 *Enterococcus* was primarily found only in mice purchased from commercial vendors and also
137 increased in relative abundance after clindamycin treatment (Fig. 3D). These findings demonstrate
138 that clindamycin had a consistent impact on the fecal bacterial communities of mice from all sources
139 and only a subset of the OTUs continued to vary between sources.

140 **Microbiota variation between sources is maintained after *C. difficile* challenge.** One day
141 post-infection, significant differences in diversity metrics remained across sources ($P_{FDR} < 0.05$, Fig
142 4A-B and Tables S1-2). Although the Charles River mice had more diverse communities and were
143 also able to clear *C. difficile* faster than the other sources, diversity did not explain the observed
144 variation in *C. difficile* colonization across sources. The Young and Schloss mice had the lowest
145 diversity 1 day post-infection and were able to clear *C. difficile* earlier than Jackson, Taconic and
146 Envigo mice. The source of mice and the interactions between source and cage effects continued
147 to explain most of the observed community variation (combined $R^2 = 0.88$; $P < 0.001$; Fig. 4C and

¹⁴⁸ Table S3). One day after *C. difficile* challenge, there were 44 OTUs (Fig. 4D and Table S10) with
¹⁴⁹ significantly different relative abundances across sources.

¹⁵⁰ Throughout the experiment, the source of mice continued to be the dominant factor that explained
¹⁵¹ the observed variation across fecal communities (PERMANOVA $R^2 = 0.35$, $P < 0.001$) followed by
¹⁵² interactions between cage effects and the day of the experiment (Movie S1 and Table S11). Mice
¹⁵³ fecal samples from the same source of mice continued to cluster closely to each other throughout
¹⁵⁴ the experiment. By 7 days post-infection, when approximately 43% mice had cleared *C. difficile*,
¹⁵⁵ most of the mice had not recovered to their baseline community structure (Fig. 4E). The distance to
¹⁵⁶ the baseline community did not explain the variation in *C. difficile* clearance as the Schloss and
¹⁵⁷ Young mice had mostly cleared *C. difficile*, but their communities were a greater distance from
¹⁵⁸ baseline 7 days post-infection compared to the Jackson and Taconic mice that were still colonized.
¹⁵⁹ In summary, mouse bacterial communities varied significantly between sources throughout the
¹⁶⁰ course of the experiment and a consistent subset of bacteria remained different between sources
¹⁶¹ regardless of clindamycin and *C. difficile* challenge.

¹⁶² **Baseline, post-clindamycin, and post-infection community data can predict mice that will**
¹⁶³ **clear *C. difficile* by 7 days post-infection.** After identifying taxa that varied between sources,
¹⁶⁴ changed after clindamycin treatment, or both, we determined which taxa were influencing the
¹⁶⁵ variation in *C. difficile* colonization at day 7 (Fig. 2B, Fig. S2C). We trained three L2-regularized
¹⁶⁶ logistic regression models with input bacterial community data from the baseline (day = -1),
¹⁶⁷ post-clindamycin (day = 0), and post-infection (day = 1) timepoints of the experiment to predict *C.*
¹⁶⁸ *difficile* colonization status on day 7 (Fig. S3A-B). All models were better at predicting *C. difficile*
¹⁶⁹ colonization status on day 7 than random chance (all $P < 0.001$, Table S12). The model based on
¹⁷⁰ the post-clindamycin (AUROC = 0.78) community OTU data performed significantly better than
¹⁷¹ the baseline (AUROC = 0.72) or the post-infection (AUROC = 0.67) models ($P_{FDR} < 0.001$ for
¹⁷² pairwise comparisons, Fig. S3C and Table S13). Thus, we were able to use bacterial relative
¹⁷³ abundance data from the time of *C. difficile* challenge to differentiate mice that had cleared *C.*
¹⁷⁴ *difficile* before day 7 from the mice still colonized with *C. difficile* at that timepoint. This result
¹⁷⁵ suggests the bacterial community's response to clindamycin treatment had the greatest influence
¹⁷⁶ on subsequent *C. difficile* colonization dynamics.

177 To examine the bacteria that were driving each model's performance, we selected the 20 OTUs that
178 had the highest absolute feature weights in each of the 3 models (Table S14). First, we looked at
179 OTUs from the model with the best performance, which was based on the post-clindamycin
180 treatment (day 0) bacterial community data. Out of the 10 highest ranked OTUs, 7 OTUs
181 (*Bacteroides*, *Escherichia/Shigella*, 2 *Lachnospiraceae*, *Lactobacillus*, *Porphyromonadaceae*, and
182 *Ruminococcaceae*) were associated with *C. difficile* colonization 7 days post-infection, while 3 OTUs
183 (*Enterobacteriaceae*, *Lachnospiraceae*, *Porphyromonadaceae*) were associated with clearance
184 (Fig. 5A). Next, we examined whether any of the top 20 ranked OTUs from the post-clindamycin (day
185 0) model were also important in the other 2 classification models based on baseline (day -1) and 1
186 day post-infection community data. We identified 6 OTUs (*Enterobacteriaceae*, *Ruminococcaceae*,
187 *Lactobacillus*, *Bacteroides*, *Porphyromonadaceae*, *Erysipelotrichaceae*) that were important to the
188 post-clindamycin model and either the baseline or 1 day post-infection models (Table S14). Thus, a
189 subset of bacterial OTUs were important for determining *C. difficile* colonization dynamics across
190 multiple timepoints.

191 To determine whether the OTUs driving the classification models also varied between sources,
192 were altered by clindamycin treatment, or both, we identified the OTUs from each model that varied
193 between sources (Fig. 1D, 3D, 4D and Tables S5, S8, S10) or were impacted by clindamycin
194 treatment (Fig. 3E and Table S9; Fig. S4). Comparing the features important to the 3 models
195 identified 14 OTUs associated with source, 21 OTUs associated with clindamycin treatment, and
196 6 OTUs associated with both (Fig. 5B). Together, these results suggest that the initial bacterial
197 communities and their responses to clindamycin influence the clearance of *C. difficile*.

198 Several OTUs (*Bacteroides*, *Enterococcus*, *Enterobacteriaceae*, *Porphyromonadaceae*) that
199 overlapped with our previous analyses appeared across at least 2 models, so we examined
200 how the relative abundances of these OTUs varied over the course of the experiment (Fig. 6).
201 Across the 9 days post-infection, there was at least 1 timepoint when the relative abundances
202 of these OTUs significantly varied between sources (Table S15). Interestingly, there were no
203 OTUs that emerged as consistently enriched or depleted in mice that were colonized past 7 days
204 post-infection, suggesting that multiple bacteria influence *C. difficile* colonization dynamics.

205 **Discussion**

206 Applying our CDI model with mice from 6 different sources, allowed us to identify bacterial taxa that
207 were unique to different sources as well as taxa that were universally impacted by clindamycin. We
208 trained L2 logistic regression models with baseline (day -1), post-clindamycin treatment (day 0),
209 and 1-day post-infection fecal community data that could predict whether mice cleared *C. difficile*
210 by 7 days post-infection better than random chance. We identified *Bacteroides*, *Enterococcus*,
211 *Enterobacteriaceae*, *Porphyromonadaceae* (Fig. 6) as candidate bacteria within these communities
212 that were influencing variation in *C. difficile* colonization dynamics since these bacteria were all
213 important in the logistic regression models and varied by source, were impacted by clindamycin
214 treatment, or both. Overall, our results demonstrated clindamycin was sufficient to render mice from
215 multiple sources susceptible to CDI and only a subset of the inter-individual microbiota variation
216 across mice from different sources was associated with the time needed to clear *C. difficile*.

217 Other studies have taken similar approaches by using mice from multiple sources to identify
218 bacteria that either promote colonization resistance or increase susceptibility to enteric infections
219 (22, 23, 26–30). For example, against *Salmonella* infections, *Enterobacteriaceae* and segmented
220 filamentous bacteria have emerged as protective (22, 27). We found *Enterobacteriaceae* increased
221 in all sources of mice after clindamycin treatment, facilitating *C. difficile* colonization. However,
222 there was also variation in *Enterobacteriaceae* relative abundance levels between sources that was
223 associated with the variation in *C. difficile* colonization dynamics across sources. Thus, bacteria
224 may have differential roles in determining susceptibility depending on the type of bacterial infection.

225 Differences in CDI mouse model studies have been attributed to intestinal microbiota variation
226 across sources. For example, researchers using the same clindamycin treatment and C57BL/6
227 mice had different *C. difficile* outcomes, one having sustained colonization (32), while the other
228 had transient (18), despite both using *C. difficile* VPI 10643. Baseline differences in the microbiota
229 composition have been hypothesized to partially explain the differences in colonization outcomes
230 and overall susceptibility to *C. difficile* after treatment with the same antibiotic (13, 31). A previous
231 study with *C. difficile* identified an endogenous protective *C. difficile* strain LEM1 that bloomed
232 after antibiotic treatment in mice from Jackson or Charles River Laboratories, but not Taconic that

233 protected mice against the more toxigenic *C. difficile* VPI10463 (26). Given that we obtained mice
234 from the same vendors, we checked all mice for endogenous *C. difficile* by plating stool samples that
235 were collected after clindamycin treatment. However, we did not identify any endogenous *C. difficile*
236 strains prior to challenge, suggesting there were no endogenous protective strains in the mice we
237 received and other bacteria mediated the variation in *C. difficile* colonization across sources. The
238 *C. difficile* strain used could also be contributing to the variation in *C. difficile* outcomes seen across
239 different research groups. For example, a group found differential colonization outcomes after
240 clindamycin treatment, with *C. difficile* 630 and M68 infections eventually becoming undetectable
241 while strain BI-7 remained detectable up to 70 days post-treatment (44).

242 The bacterial perturbations induced by clindamycin treatment have been well characterized
243 and our findings agree with previous CDI mouse model work demonstrating *Enterococcus* and
244 *Enterobacteriaceae* were associated with *C. difficile* susceptibility and *Porphyromonadaceae*,
245 *Lachnospiraceae*, *Ruminococcaceae*, and *Turicibacter* were associated with resistance (19, 21, 32,
246 33, 43–46). While we have demonstrated that susceptibility is uniform across sources of mice after
247 clindamycin treatment, there could be different outcomes for either susceptibility or clearance in the
248 case of other antibiotic treatments.

249 We found the time needed to naturally clear *C. difficile* varied across sources of mice implying
250 that at least in the context of the same perturbation, microbiota differences seemed to influence
251 infection outcome more than susceptibility. More importantly, we were able to reduce the variation
252 observed across sources to identify a subset of OTUs that were also important for predicting *C.*
253 *difficile* colonization status 7 days post-infection. Since all but 3 mice eventually cleared *C. difficile*
254 630 by 9 days post-infection and the model built with the post-clindamycin (day 0) OTU relative
255 abundance data had the best performance, our results suggest clindamycin treatment had a larger
256 role in determining *C. difficile* susceptibility and clearance than the source of the mice.

257 Our approach successfully increased the inter-animal variation tested with our clindamycin model of
258 CDI. One alternative approach that has been used in some CDI studies (47–52) is to associate mice
259 with human microbiotas. However, a major caveat to this method is the substantial loss of human
260 microbiota community members upon transfer to mice (53, 54). Additionally with the exception of 2

261 recent studies (47, 48), most of the CDI mouse model studies to date associated mice with just 1
262 types of human microbiota either from a single donor or a single pool from multiple donors (49–52),
263 which does not aid in the goal of modeling the interpersonal variation seen in humans to understand
264 how the microbiota influences susceptibility to CDIs and adverse outcomes. Importantly, our study
265 using mice from 6 different sources increased the variation between groups of mice compared to
266 using 1 source alone, to better reflect the inter-individual microbiota variation observed in humans.

267 Another motivation for associating mice with human microbiotas, is to study the bacteria associated
268 with the disease in humans. Encouragingly, decreased *Bifidobacterium*, *Porphyromonas*,
269 *Ruminococcaceae* and *Lachnospiraceae* and increased *Enterobacteriaceae*, *Enterococcus*,
270 *Lactobacillus*, and *Proteus* have all been associated with human CDIs (7) and were well
271 represented in our study, suggesting most of the mouse sources are suitable for gaining insights
272 into the bacteria influencing *C. difficile* colonization and infections in humans. An important
273 exception was *Enterococcus*, which was primarily absent from University of Michigan colonies and
274 *Proteus*, which was only found in Taconic mice. The fact that some CDI-associated bacteria were
275 only found in a subset of mice has important implications for future CDI mouse model studies.

276 Other microbiota and host factors that were outside the scope of our current study may also
277 contribute to the differences in *C. difficile* colonization dynamics between sources of mice.
278 The microbiota is composed of viruses, fungi, and parasites in addition to bacteria, and these
279 non-bacterial members can also vary across sources of mice (55, 56). While our study focused
280 solely on the bacterial portion, viruses and fungi have also begun to be implicated in the context
281 of CDIs or FMT treatments for recurrent CDIs (35, 57–60). Beyond community composition, the
282 metabolic function of the microbiota also has a CDI signature (20, 46, 61, 62) and can vary across
283 mice from different sources (63). For example, microbial metabolites, particularly secondary
284 bile acids and butyrate production, have been implicated as important contributors to *C. difficile*
285 resistance (33, 44). Interestingly, butyrate has previously been shown to vary across mouse
286 vendors and mediated resistance to *Citrobacter rodentium* infection, a model of enterohemorrhagic
287 and enteropathogenic *Escherichia coli* infections (23). Evidence for immunological toning
288 differences in IgA and Th17 cells across mice from different vendors have also been documented
289 and (64, 65) could influence the host response to CDI (66, 67). The outcome after *C. difficile*

290 exposure depends on a multitude of factors, including age, diet, and immunity; all of which are also
291 influenced by the microbiota.

292 We have demonstrated that the ways baseline microbiotas from different mouse sources respond
293 to clindamycin treatment influence the length of time mice remained colonized with *C. difficile* 630.
294 To better understand the contribution of the microbiota to *C. difficile* pathogenesis and treatments,
295 using multiple sources of mice may yield more insights than a single source. Furthermore, for
296 studies wanting to examine the interplay between particular bacteria such as *Enterococcus* and *C.*
297 *difficile*, these results could serve as a resource for selecting mice to address their question. Using
298 mice from multiple sources helps model the interpersonal microbiota variation among humans to
299 aid our understanding of how the gut microbiota provides colonization resistance to CDIs.

300 **Acknowledgements**

301 This work was supported by the National Institutes of Health (U01AI124255). ST was supported by
302 the Michigan Institute for Clinical and Health Research Postdoctoral Translation Scholars Program
303 (UL1TR002240). We thank members of the Schloss lab for feedback on planning the experiments
304 and data presentation. In particular, we thank Begüm Topçuoğlu for help with implementing L2
305 logistic regression models, Ana Taylor for help with media preparation and sample collection, and
306 Nicholas Lesniak for his critical feedback on the manuscript. We also thank members of Vincent
307 Young's lab, particularly Kimberly Vendrov, for guidance with the *C. difficile* infection mouse model
308 and donating the mice. We also thank the Unit for Laboratory Animal Medicine at the University of
309 Michigan for maintaining our mouse colony and providing the institutional support for our mouse
310 experiments. Finally, we thank Kwi Kim, Austin Campbell, and Kimberly Vendrov for their help in
311 maintaining the Schloss lab's anaerobic chamber.

312 **Materials and Methods**

313 **(i) Animals.** All experiments were approved by the University of Michigan Animal Care and Use
314 Committee (IACUC) under protocol number PRO00006983. Female C57BL/7 mice were obtained
315 from 6 different sources: The Jackson Laboratory, Charles River Laboratories, Taconic Biosciences,
316 Envigo, and two colonies at the University of Michigan (the Schloss lab colony and the Young lab
317 colony). The Young lab colony was originally established with mice purchased from Jackson, and
318 the Schloss lab colony was established in 2010 with mice donated from the Young lab. The 4
319 groups of mice purchased from vendors were allowed to acclimate to the University of Michigan
320 mouse facility for 13 days prior to starting the experiment. At least 4 female mice (age 5-10 weeks)
321 were obtained per source and mice from the same source were primarily housed at a density of 2
322 mice per cage. The experiment was repeated once, approximately 3 months after the start of the
323 first experiment.

324 **(ii) Antibiotic treatment.** After the 13-day acclimation period and 1 day prior to challenge (Fig.
325 1A), all mice received 10 mg/kg clindamycin (filter sterilized through a 0.22 micron syringe filter
326 prior to administration) via intraperitoneal injection.

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

340 **(iv) 16S rRNA sequencing.** DNA was extracted from -80 °C stored stool samples using the DNeasy
341 Powersoil HTP 96 kit (Qiagen) and an EpMotion 5075 automated pipetting system (Eppendorf).
342 The V4 region was amplified for 16S rRNA with the AccuPrime Pfx DNA polymerase (Thermo
343 Fisher Scientific) using custom barcoded primers, as previously described (68). The ZymoBIOMICS
344 microbial community DNA standards was used as a mock community control (69) and water was
345 used as a negative control per 96-well extraction plate. The PCR amplicons were cleaned up
346 and normalized with the SequalPrep normalization plate kit (Thermo Fisher Scientific). Amplicons
347 were pooled and quantified with the KAPA library quantification kit (KAPA biosystems), prior to
348 sequencing using the MiSeq system (Illumina).

349 **(v) 16S rRNA gene sequence analysis.** mothur (v. 1.43) was used to process all sequences
350 (70) with a previously published protocol (68). Reads were combined and aligned with the SILVA
351 reference database (71). Chimeras were removed with the VSEARCH algorithm and taxonomic
352 assignment was completed with a modified version (v16) of the Ribosomal Database Project
353 reference database (v11.5) (72) with an 80% confidence cutoff. Operational taxonomic units (OTUs)
354 were assigned with a 97% similarity threshold using the opticlus algorithm (73). To account for
355 uneven sequencing across samples, samples were rarefied to 5,437 sequences 1,000 times for
356 alpha and beta diversity analyses, and a single time to generate relative abundances for model
357 training. PCoAs were generated based on θ_{YC} distances. Permutational multivariate analysis
358 of variance (PERMANOVA) was performed on mothur-generated θ_{YC} distance matrices with the
359 adonis function in the vegan package (74) in R (75).

360 **(vi) Classification model training and evaluation.** Models were generated based on mice that
361 were categorized as either cleared or colonized 7 days post-infection and had sequencing data
362 from the baseline (day -1), post-clindamycin (day 0), and post-infection (day 1) timepoints of
363 the experiment. Input bacterial community relative abundance data at the OTU level from the
364 baseline, post-clindamycin, and post-infection timepoints was used to generate 6 classification
365 models that predicted *C. difficile* colonization status 7 days post-infection. The L2-regularized
366 logistic regression models were trained and tested using the caret package (76) in R as previously
367 described (77) with the exception that we used 60% training and 40% testing data splits for the
368 cross-validation of the training data to select the best cost hyperparameter and the testing of

369 the held out test data to measure model performance. The modified training to testing ratio was
370 selected to accommodate the small number of samples in the dataset. Code was modified from
371 https://github.com/SchlossLab/ML_pipeline_microbiome to update the classification outcomes
372 and change the data split ratios. The modified repository to regenerate this analysis is available at
373 https://github.com/tomkose/TomkoseML_pipeline_microbiome.

374 **(vii) Statistical analysis.** All statistical tests were performed in R (v 4.0.2) (75). The Kruskal-Wallis
375 test was used to analyze differences in *C. difficile* CFU, mouse weight change, and alpha
376 diversity across sources with a Benjamini-Hochberg correction for testing multiple timepoints,
377 followed by pairwise Wilcoxon comparisons with Benjamini-Hochberg correction. For taxonomic
378 analysis and generation of logistic regression model input data, *C. difficile* (OTU 20) was removed.
379 Bacterial relative abundances that varied across sources at the OTU level were identified with the
380 Kruskal-Wallis test with Benjamini-Hochberg correction for testing all identified OTUs, followed by
381 pairwise Wilcoxon comparisons with Benjamini-Hochberg correction. The Wilcoxon rank sum test
382 was used to test for OTUs that differed between experiments within the Schloss, Young, and Envigo
383 sources with Benjamini-Hochberg correction for testing all identified OTUs. OTUs impacted by
384 clindamycin treatment were identified using the paired Wilcoxon signed rank test with matched
385 pairs of mice samples for day -1 and day 0. To determine whether classification models had better
386 performance (test AUROCs) than random chance (0.5), we used the one-sample Wilcoxon signed
387 rank test. To examine whether there was an overall difference in predictive performance across the
388 6 classification models we used the Kruskal-Wallis test followed by pairwise Wilcoxon comparisons
389 with Benjamini-Hochberg correction for multiple hypothesis testing. The tidyverse package (v 1.3.0)
390 was used to wrangle and graph data (78).

391 **(viii) Code availability.** Code for all data analysis and generating this manuscript is available at
392 https://github.com/SchlossLab/Tomkovich_Vendor_XXXX_2020.

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

395 **References**

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

- 420 Alimentary Pharmacology & Therapeutics 44:715–727.
- 421 11. Pakpour S, Bhanvadia A, Zhu R, Amarnani A, Gibbons SM, Gurry T, Alm EJ, Martello LA. 2017.
422 Identifying predictive features of *Clostridium difficile* infection recurrence before, during, and after
423 primary antibiotic treatment. Microbiome 5.
- 424 12. Lee AA, Rao K, Limsrivilai J, Gilliland M, Malamet B, Briggs E, Young VB, Higgins PDR. 2020.
425 Temporal gut microbial changes predict recurrent *Clostridioides difficile* infection in patients with
426 and without ulcerative colitis. Inflammatory Bowel Diseases <https://doi.org/10.1093/ibd/izz335>.
- 427 13. Hutton ML, Mackin KE, Chakravorty A, Lyras D. 2014. Small animal models for the study of
428 *Clostridium difficile* disease pathogenesis. FEMS Microbiology Letters 352:140–149.
- 429 14. Chen X, Katchar K, Goldsmith JD, Nanthakumar N, Cheknis A, Gerding DN, Kelly CP. 2008. A
430 mouse model of *Clostridium difficile*-associated disease. Gastroenterology 135:1984–1992.
- 431 15. Best EL, Freeman J, Wilcox MH. 2012. Models for the study of *Clostridium difficile* infection.
432 Gut Microbes 3:145–167.
- 433 16. Baxter NT, Wan JJ, Schubert AM, Jenior ML, Myers P, Schloss PD. 2014. Intra- and
434 interindividual variations mask interspecies variation in the microbiota of sympatric peromyscus
435 populations. Applied and Environmental Microbiology 81:396–404.
- 436 17. Nagpal R, Wang S, Woods LCS, Seshie O, Chung ST, Shively CA, Register TC, Craft S,
437 McClain DA, Yadav H. 2018. Comparative microbiome signatures and short-chain fatty acids in
438 mouse, rat, non-human primate, and human feces. Frontiers in Microbiology 9.
- 439 18. Reeves AE, Theriot CM, Bergin IL, Huffnagle GB, Schloss PD, Young VB. 2011. The interplay
440 between microbiome dynamics and pathogen dynamics in a murine model of *Clostridium difficile*
441 infection 2:145–158.
- 442 19. Schubert AM, Sinani H, Schloss PD. 2015. Antibiotic-induced alterations of the murine gut
443 microbiota and subsequent effects on colonization resistance against *Clostridium difficile*. mBio 6.
- 444 20. Jenior ML, Leslie JL, Young VB, Schloss PD. 2017. *Clostridium difficile* colonizes alternative

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

- 471 AL, Kalinke U, Flavell RA, Erhardt M, Strowig T. 2017. Enhancement of IFNgamma production by
472 distinct commensals ameliorates *Salmonella*-induced disease. *Cell Host & Microbe* 21:682–694.e5.
- 473 29. Rolig AS, Cech C, Ahler E, Carter JE, Ottemann KM. 2013. The degree of *Helicobacter*
474 *pylori*-triggered inflammation is manipulated by preinfection host microbiota. *Infection and Immunity*
475 81:1382–1389.
- 476 30. Ge Z, Sheh A, Feng Y, Muthupalani S, Ge L, Wang C, Kurnick S, Mannion A, Whary MT, Fox
477 JG. 2018. *Helicobacter pylori*-infected C57BL/6 mice with different gastrointestinal microbiota have
478 contrasting gastric pathology, microbial and host immune responses. *Scientific Reports* 8.
- 479 31. Lawley TD, Young VB. 2013. Murine models to study *Clostridium difficile* infection and
480 transmission. *Anaerobe* 24:94–97.
- 481 32. Buffie CG, Jarchum I, Equinda M, Lipuma L, Gobourne A, Viale A, Ubeda C, Xavier J, Pamer
482 EG. 2011. Profound alterations of intestinal microbiota following a single dose of clindamycin results
483 in sustained susceptibility to *Clostridium difficile*-induced colitis. *Infection and Immunity* 80:62–73.
- 484 33. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew
485 M, Viale A, Littmann E, Brink MRM van den, Jenq RR, Taur Y, Sander C, Cross JR, Toussaint
486 NC, Xavier JB, Pamer EG. 2014. Precision microbiome reconstitution restores bile acid mediated
487 resistance to *Clostridium difficile*. *Nature* 517:205–208.
- 488 34. Spinler JK, Brown A, Ross CL, Boonma P, Conner ME, Savidge TC. 2016. Administration of
489 probiotic kefir to mice with *Clostridium difficile* infection exacerbates disease. *Anaerobe* 40:54–57.
- 490 35. Markey L, Shaban L, Green ER, Lemon KP, Mecsas J, Kumamoto CA. 2018. Pre-colonization
491 with the commensal fungus candida albicans reduces murine susceptibility to *Clostridium difficile*
492 infection. *Gut Microbes* 1–13.
- 493 36. McKee RW, Aleksanyan N, Garrett EM, Tamayo R. 2018. Type IV pili promote *Clostridium*
494 *difficile* adherence and persistence in a mouse model of infection. *Infection and Immunity* 86.
- 495 37. Yamaguchi T, Konishi H, Aoki K, Ishii Y, Chono K, Tateda K. 2020. The gut microbiome diversity

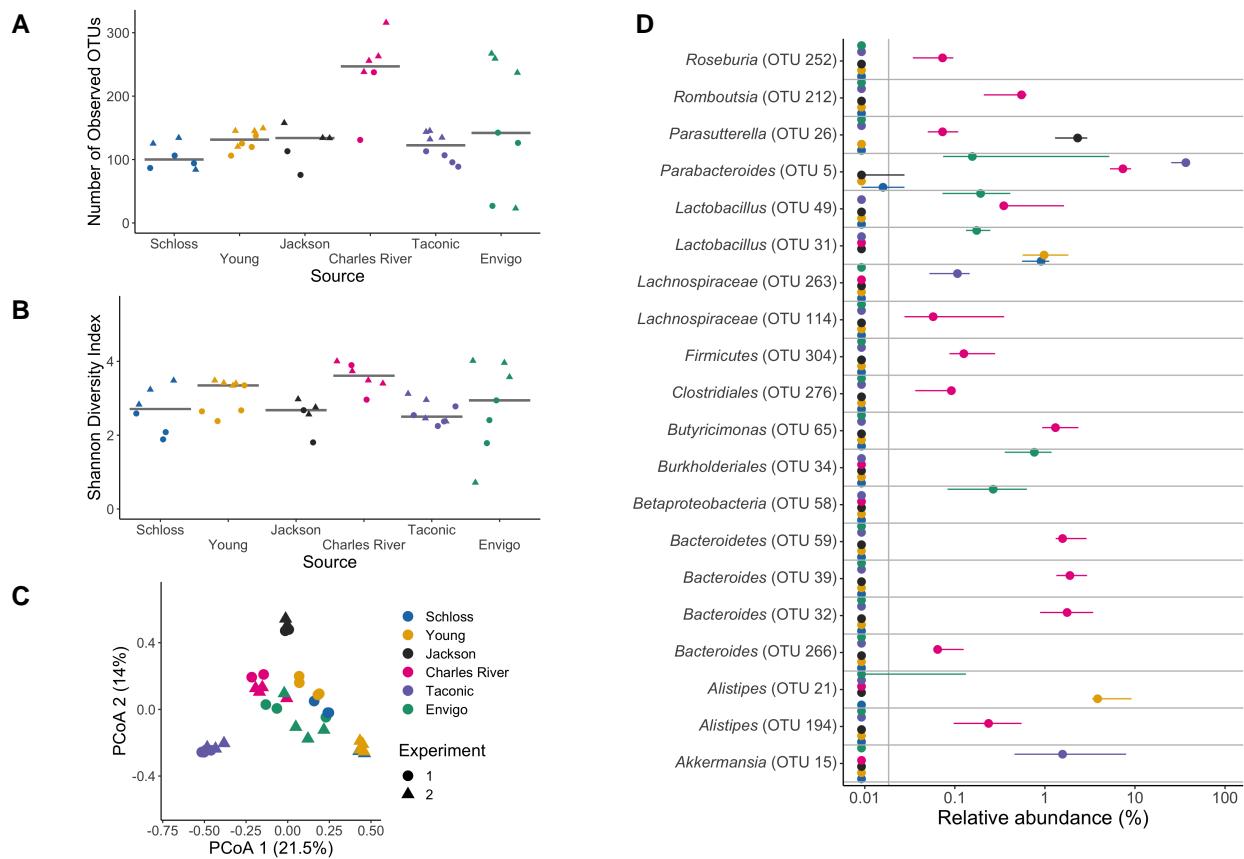
- 496 of *Clostridioides difficile*-inoculated mice treated with vancomycin and fidaxomicin. Journal of
497 Infection and Chemotherapy 26:483–491.
- 498 38. Stroke IL, Letourneau JJ, Miller TE, Xu Y, Pechik I, Savoly DR, Ma L, Sturzenbecker LJ,
499 Sabalski J, Stein PD, Webb ML, Hilbert DW. 2018. Treatment of *Clostridium difficile* infection
500 with a small-molecule inhibitor of toxin UDP-glucose hydrolysis activity. Antimicrobial Agents and
501 Chemotherapy 62.
- 502 39. Quigley L, Coakley M, Alemayehu D, Rea MC, Casey PG, O'Sullivan, Murphy E, Kiely B, Cotter
503 PD, Hill C, Ross RP. 2019. *Lactobacillus gasseri* APC 678 reduces shedding of the pathogen
504 *Clostridium difficile* in a murine model. Frontiers in Microbiology 10.
- 505 40. Mullish BH, McDonald JAK, Pechlivanis A, Allegretti JR, Kao D, Barker GF, Kapila D, Petrof
506 EO, Joyce SA, Gahan CGM, Glegola-Madejska I, Williams HRT, Holmes E, Clarke TB, Thursz
507 MR, Marchesi JR. 2019. Microbial bile salt hydrolases mediate the efficacy of faecal microbiota
508 transplant in the treatment of recurrent *Clostridioides difficile* infection. Gut 68:1791–1800.
- 509 41. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. 2015. How informative is the mouse for human
510 gut microbiota research? Disease Models & Mechanisms 8:1–16.
- 511 42. Guh AY, Kutty PK. 2018. *Clostridioides difficile* infection 169:ITC49.
- 512 43. Tomkovich S, Lesniak NA, Li Y, Bishop L, Fitzgerald MJ, Schloss PD. 2019. The proton
513 pump inhibitor omeprazole does not promote *Clostridioides difficile* colonization in a murine model.
514 mSphere 4.
- 515 44. Lawley TD, Clare S, Walker AW, Stares MD, Connor TR, Raisen C, Goulding D, Rad R,
516 Schreiber F, Brandt C, Deakin LJ, Pickard DJ, Duncan SH, Flint HJ, Clark TG, Parkhill J, Dougan
517 G. 2012. Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy
518 resolves relapsing *Clostridium difficile* disease in mice. PLoS Pathogens 8:e1002995.
- 519 45. Lawley TD, Clare S, Walker AW, Goulding D, Stabler RA, Croucher N, Mastroeni P, Scott P,
520 Raisen C, Mottram L, Fairweather NF, Wren BW, Parkhill J, Dougan G. 2009. Antibiotic treatment of
521 *Clostridium difficile* carrier mice triggers a supershedder state, spore-mediated transmission, and

- 522 severe disease in immunocompromised hosts. *Infection and Immunity* 77:3661–3669.
- 523 46. Jump RLP, Polinkovsky A, Hurless K, Sitzlar B, Eckart K, Tomas M, Deshpande A, Nerandzic
524 MM, Donskey CJ. 2014. Metabolomics analysis identifies intestinal microbiota-derived biomarkers
525 of colonization resistance in clindamycin-treated mice. *PLoS ONE* 9:e101267.
- 526 47. Nagao-Kitamoto H, Leslie JL, Kitamoto S, Jin C, Thomsson KA, Gilliland MG, Kuffa P, Goto Y,
527 Jenq RR, Ishii C, Hirayama A, Seekatz AM, Martens EC, Eaton KA, Kao JY, Fukuda S, Higgins PDR,
528 Karlsson NG, Young VB, Kamada N. 2020. Interleukin-22-mediated host glycosylation prevents
529 *Clostridioides difficile* infection by modulating the metabolic activity of the gut microbiota. *Nature
530 Medicine* 26:608–617.
- 531 48. Battaglioli EJ, Hale VL, Chen J, Jeraldo P, Ruiz-Mojica C, Schmidt BA, Rekdal VM, Till LM, Huq
532 L, Smits SA, Moor WJ, Jones-Hall Y, Smyrk T, Khanna S, Pardi DS, Grover M, Patel R, Chia N,
533 Nelson H, Sonnenburg JL, Farrugia G, Kashyap PC. 2018. *Clostridioides difficile* uses amino acids
534 associated with gut microbial dysbiosis in a subset of patients with diarrhea. *Science Translational
535 Medicine* 10:eaam7019.
- 536 49. Robinson CD, Auchtung JM, Collins J, Britton RA. 2014. Epidemic *Clostridium difficile* strains
537 demonstrate increased competitive fitness compared to nonepidemic isolates. *Infection and
538 Immunity* 82:2815–2825.
- 539 50. Collins J, Auchtung JM, Schaefer L, Eaton KA, Britton RA. 2015. Humanized microbiota mice
540 as a model of recurrent *Clostridium difficile* disease. *Microbiome* 3.
- 541 51. Collins J, Robinson C, Danhof H, Knetsch CW, Leeuwen HC van, Lawley TD, Auchtung JM,
542 Britton RA. 2018. Dietary trehalose enhances virulence of epidemic *Clostridium difficile*. *Nature
543* 553:291–294.
- 544 52. Hryckowian AJ, Treuren WV, Smits SA, Davis NM, Gardner JO, Bouley DM, Sonnenburg JL.
545 2018. Microbiota-accessible carbohydrates suppress *Clostridium difficile* infection in a murine
546 model. *Nature Microbiology* 3:662–669.
- 547 53. Fouladi F, Glenny EM, Bulik-Sullivan EC, Tsilimigras MCB, Sioda M, Thomas SA, Wang Y, Djukic

- 548 Z, Tang Q, Tarantino LM, Bulik CM, Fodor AA, Carroll IM. 2020. Sequence variant analysis reveals
549 poor correlations in microbial taxonomic abundance between humans and mice after gnotobiotic
550 transfer. *The ISME Journal* <https://doi.org/10.1038/s41396-020-0645-z>.
- 551 54. Walter J, Armet AM, Finlay BB, Shanahan F. 2020. Establishing or exaggerating causality for
552 the gut microbiome: Lessons from human microbiota-associated rodents. *Cell* 180:221–232.
- 553 55. Rasmussen TS, Vries L de, Kot W, Hansen LH, Castro-Mejía JL, Vogensen FK, Hansen AK,
554 Nielsen DS. 2019. Mouse vendor influence on the bacterial and viral gut composition exceeds the
555 effect of diet. *Viruses* 11:435.
- 556 56. Mims TS, Abdallah QA, Watts S, White C, Han J, Willis KA, Pierre JF. 2020. Variability in
557 interkingdom gut microbiomes between different commercial vendors shapes fat gain in response
558 to diet. *The FASEB Journal* 34:1–1.
- 559 57. Stewart DB, Wright JR, Fowler M, McLimans CJ, Tokarev V, Amaniera I, Baker O, Wong H-T,
560 Brabec J, Drucker R, Lamendella R. 2019. Integrated meta-omics reveals a fungus-associated
561 bacteriome and distinct functional pathways in *Clostridioides difficile* infection. *mSphere* 4.
- 562 58. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, Grasis JA, Cassidy L, Tholey A,
563 Fickenscher H, Seegert D, Rosenstiel P, Schreiber S. 2017. Efficacy of sterile fecal filtrate transfer
564 for treating patients with *Clostridium difficile* infection. *Gastroenterology* 152:799–811.e7.
- 565 59. Zuo T, Wong SH, Lam K, Lui R, Cheung K, Tang W, Ching JYL, Chan PKS, Chan MCW, Wu
566 JCY, Chan FKL, Yu J, Sung JJY, Ng SC. 2017. Bacteriophage transfer during faecal microbiota
567 transplantation in *Clostridium difficile* infection is associated with treatment outcome. *Gut*
568 *gutjnl*–2017–313952.
- 569 60. Zuo T, Wong SH, Cheung CP, Lam K, Lui R, Cheung K, Zhang F, Tang W, Ching JYL, Wu JCY,
570 Chan PKS, Sung JJY, Yu J, Chan FKL, Ng SC. 2018. Gut fungal dysbiosis correlates with reduced
571 efficacy of fecal microbiota transplantation in *Clostridium difficile* infection. *Nature Communications*
572 9.
- 573 61. Robinson JI, Weir WH, Crowley JR, Hink T, Reske KA, Kwon JH, Burnham C-AD, Dubberke

- 574 ER, Mucha PJ, Henderson JP. 2019. Metabolomic networks connect host-microbiome processes to
575 human *Clostridioides difficile* infections. *Journal of Clinical Investigation* 129:3792–3806.
- 576 62. Fletcher JR, Erwin S, Lanzas C, Theriot CM. 2018. Shifts in the gut metabolome and *Clostridium*
577 *difficile* transcriptome throughout colonization and infection in a mouse model. *mSphere* 3.
- 578 63. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, Liu C, Fang
579 Z, Chou J, Glanville J, Hao Q, Kotowska D, Colding C, Licht TR, Wu D, Yu J, Sung JJY, Liang Q, Li
580 J, Jia H, Lan Z, Tremaroli V, Dworzynski P, Nielsen HB, Bäckhed F, Doré J, Chatelier EL, Ehrlich
581 SD, Lin JC, Arumugam M, Wang J, Madsen L, Kristiansen K. 2015. A catalog of the mouse gut
582 metagenome. *Nature Biotechnology* 33:1103–1108.
- 583 64. Fransen F, Zagato E, Mazzini E, Fosso B, Manzari C, Aidy SE, Chiavelli A, D'Erchia AM,
584 Sethi MK, Pabst O, Marzano M, Moretti S, Romani L, Penna G, Pesole G, Rescigno M. 2015.
585 BALB/c and C57BL/6 mice differ in polyreactive IgA abundance, which impacts the generation of
586 antigen-specific IgA and microbiota diversity. *Immunity* 43:527–540.
- 587 65. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee
588 CA, Lynch SV, Tanoue T, Imaoka A, Itoh K, Takeda K, Umesaki Y, Honda K, Littman DR. 2009.
589 Induction of intestinal th17 cells by segmented filamentous bacteria. *Cell* 139:485–498.
- 590 66. Azrad M, Hamo Z, Tkawkho L, Peretz A. 2018. Elevated serum immunoglobulin a levels in
591 patients with *Clostridium difficile* infection are associated with mortality. *Pathogens and Disease* 76.
- 592 67. Saleh MM, Frisbee AL, Leslie JL, Buonomo EL, Cowardin CA, Ma JZ, Simpson ME, Scully KW,
593 Abhyankar MM, Petri WA. 2019. Colitis-induced th17 cells increase the risk for severe subsequent
594 *Clostridium difficile* infection. *Cell Host & Microbe* 25:756–765.e5.
- 595 68. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a
596 dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the
597 MiSeq illumina sequencing platform. *Applied and Environmental Microbiology* 79:5112–5120.
- 598 69. Sze MA, Schloss PD. 2019. The impact of DNA polymerase and number of rounds of
599 amplification in PCR on 16S rRNA gene sequence data. *mSphere* 4.

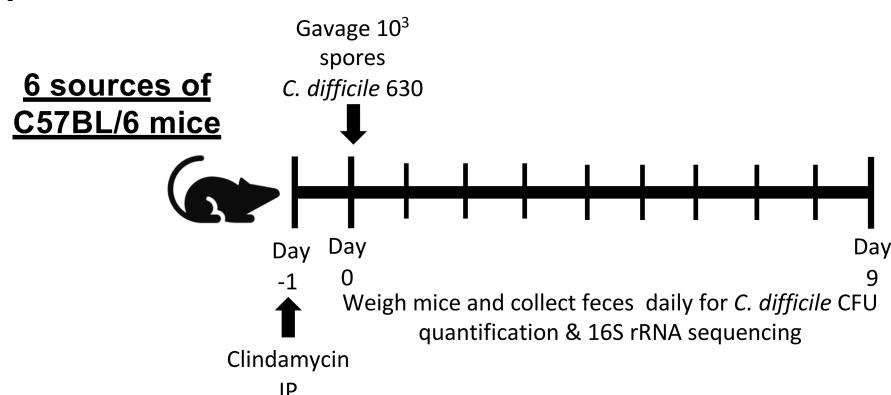
- 600 70. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
601 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Horn DJV, Weber CF.
602 2009. Introducing mothur: Open-source, platform-independent, community-supported software
603 for describing and comparing microbial communities. *Applied and Environmental Microbiology*
604 75:7537–7541.
- 605 71. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012.
606 The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools.
607 *Nucleic Acids Research* 41:D590–D596.
- 608 72. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR,
609 Tiedje JM. 2013. Ribosomal database project: Data and tools for high throughput rRNA analysis.
610 *Nucleic Acids Research* 42:D633–D642.
- 611 73. Westcott SL, Schloss PD. 2017. OptiClust, an improved method for assigning amplicon-based
612 sequence data to operational taxonomic units. *mSphere* 2.
- 613 74. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB,
614 Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2018. Vegan: Community ecology
615 package.
- 616 75. R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for
617 Statistical Computing, Vienna, Austria.
- 618 76. Kuhn M. 2008. Building predictive models inRUsing thecaretPackage. *Journal of Statistical*
619 *Software* 28.
- 620 77. Topçuoğlu BD, Lesniak NA, Ruffin MT, Wiens J, Schloss PD. 2020. A framework for effective
621 application of machine learning to microbiome-based classification problems. *mBio* 11.
- 622 78. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes
623 A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D,
624 Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the
625 tidyverse. *Journal of Open Source Software* 4:1686.

626 **Figures**

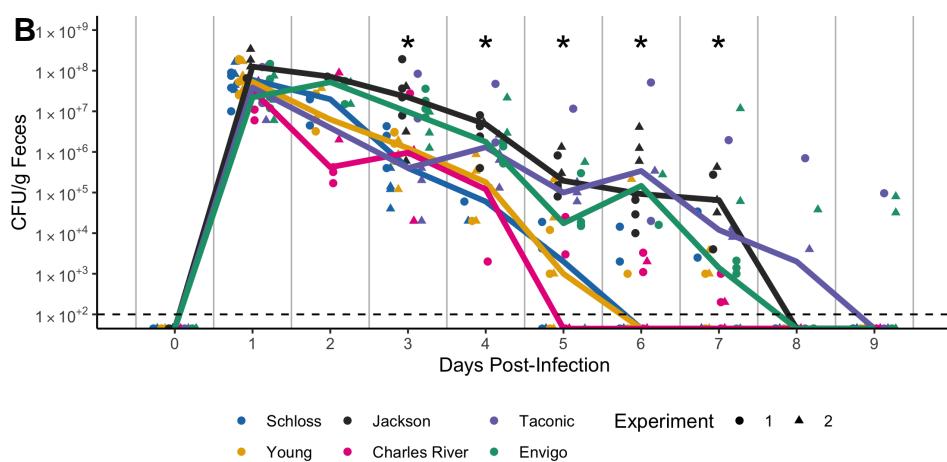
627

Figure 1. Microbiota variation is high between mice from different sources. A-B. Number of observed OTUs (A) and Shannon diversity index values (B) across sources of mice at baseline (day -1 of the experiment). Differences between sources were analyzed by Kruskal-Wallis test with Benjamini-Hochberg correction for testing each day of the experiment and the adjusted P value was < 0.05 for panel A (Table S1). None of the P values from pairwise Wilcoxon comparisons between sources were significant after Benjamini-Hochberg correction (Table S2). Gray lines represent the median values for each source of mice. C. Principal Coordinates Analysis (PCoA) of θ_{YC} distances of baseline stool samples. Source and the interaction between source and cage effects explained most of the variation (PERMANOVA combined $R^2 = 0.90$, $P < 0.001$, see Table S3). For A-C: each symbol represents the value for a stool sample from an individual mouse, circles represent experiment 1 mice and triangles represent experiment 2 mice. D. The median (point) and interquartile range (colored lines) of the relative abundances for the 20 most significant OTUs out of the 268 OTUs that varied across sources at baseline (Table S5).

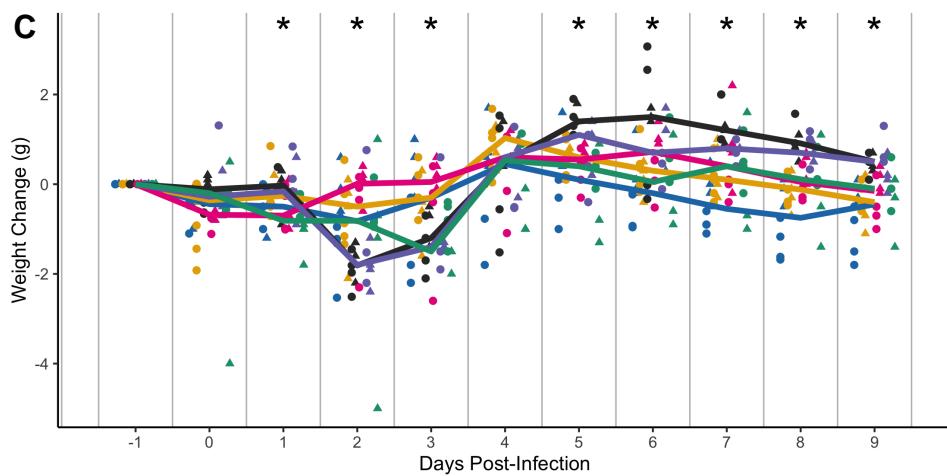
A



B



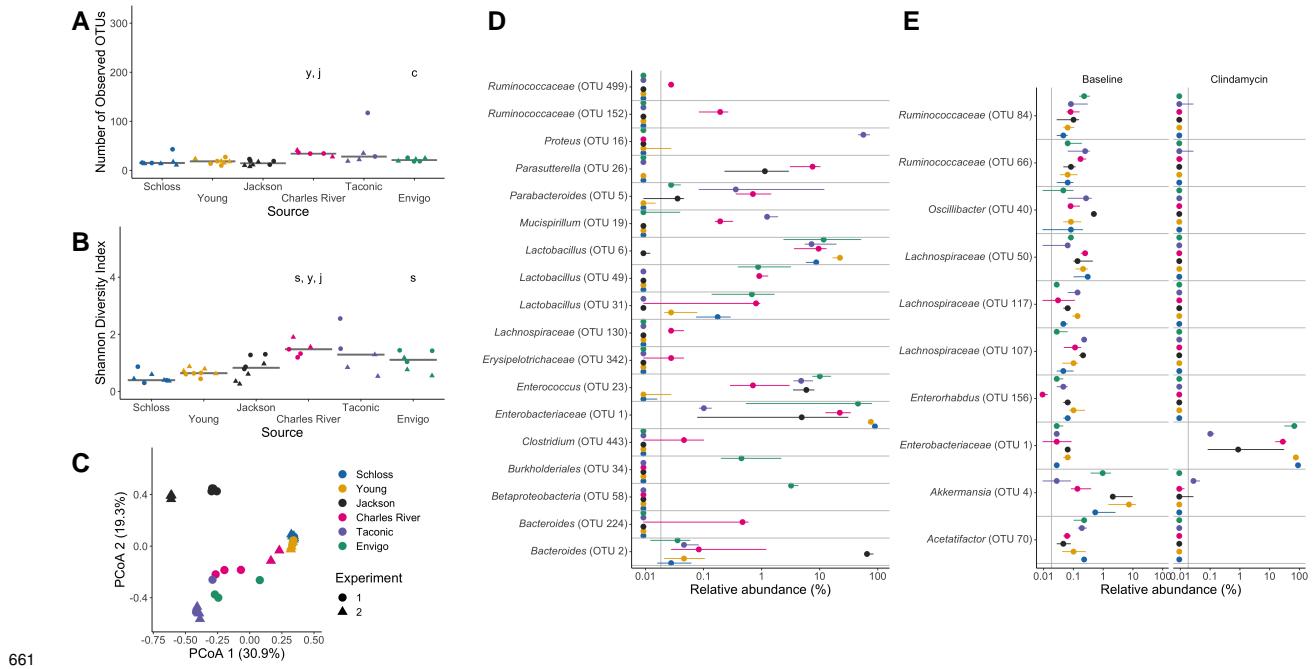
C



641

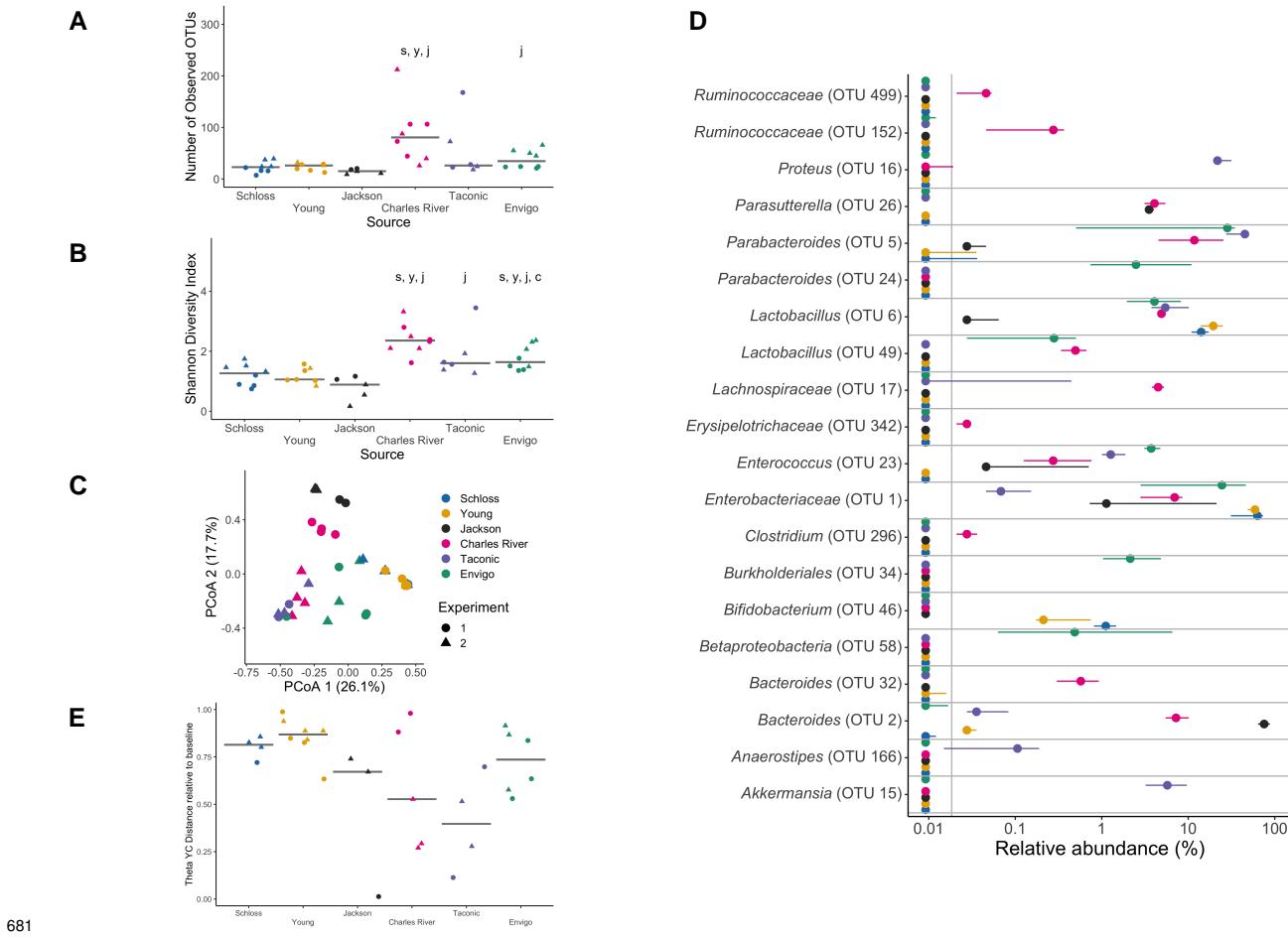
642 **Figure 2. Clindamycin is sufficient to promote *C. difficile* colonization in all mice, but**
643 **clearance time varies across sources.** A. Setup of the experimental timeline. Mice for the
644 experiments were obtained from 6 different sources: the Schloss (N = 8) and Young lab (N = 9)
645 colonies at the University of Michigan, the Jackson Laboratory (N = 8), Charles River Laboratory

646 (N = 8), Taconic Biosciences (N = 8), and Envigo (N = 8). All mice were administered 10 mg/kg
647 clindamycin intraperitoneally (IP) 1 day before challenge with *C. difficile* 630 spores on day 0.
648 Mice were weighed and feces was collected daily through the end of the experiment (9 days
649 post-infection). Note: 3 mice died during course of experiment. 1 Taconic mouse prior to infection
650 and 1 Jackson and 1 Envigo mouse between 1- and 3-days post-infection. B. *C. difficile* CFU/gram
651 stool measured over time (N = 20-49 mice per timepoint) via serial dilutions. The black line
652 represents the limit of detection for the first serial dilution. CFU quantification data was not available
653 for each mouse due to early deaths, stool sampling difficulties, and not plating all of the serial
654 dilutions. C. Mouse weight change measured in grams over time (N = 45-49 mice per timepoint),
655 all mice were normalized to the weight recorded 1 day before infection. For B-C: timepoints
656 where differences between sources of mice were statistically significant by Kruskal-Wallis test
657 with Benjamini-Hochberg correction for testing across multiple days (Table S6 and Table S7) are
658 reflected by the asterisk above each timepoint (*, $P < 0.05$). Lines represent the median for each
659 source and circles represent individual mice from experiment 1 while triangles represent mice from
660 experiment 2.



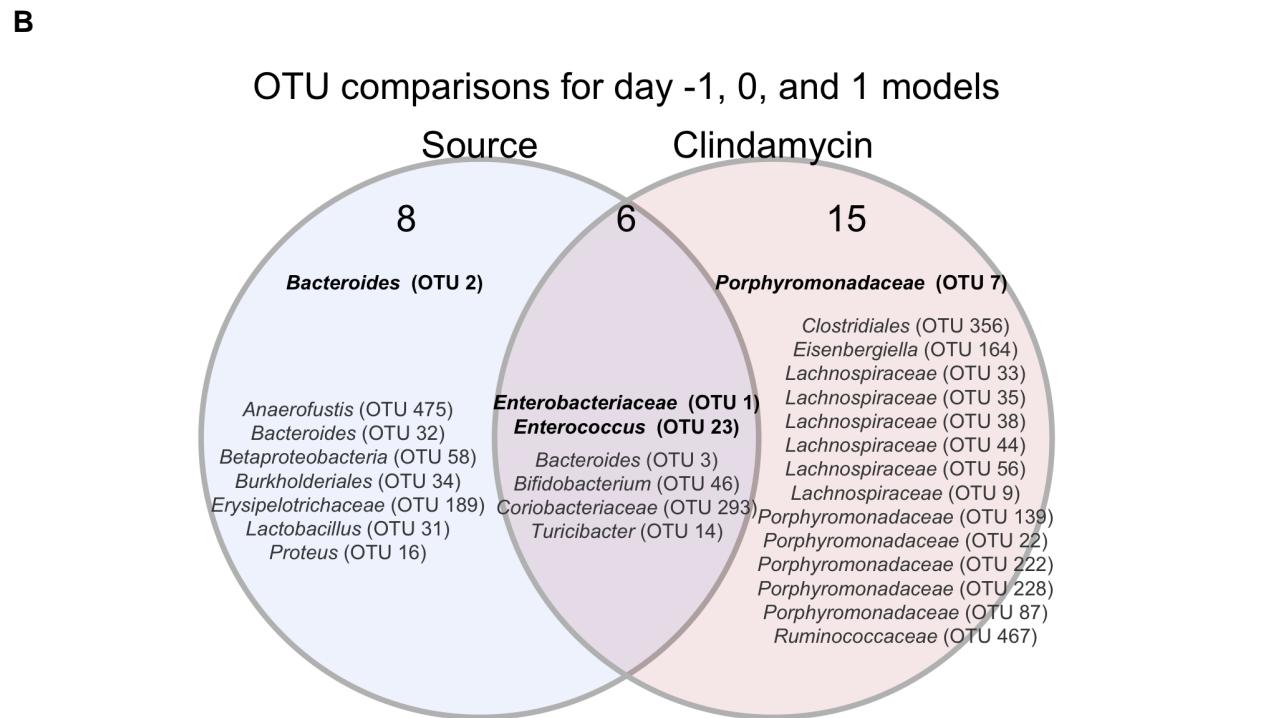
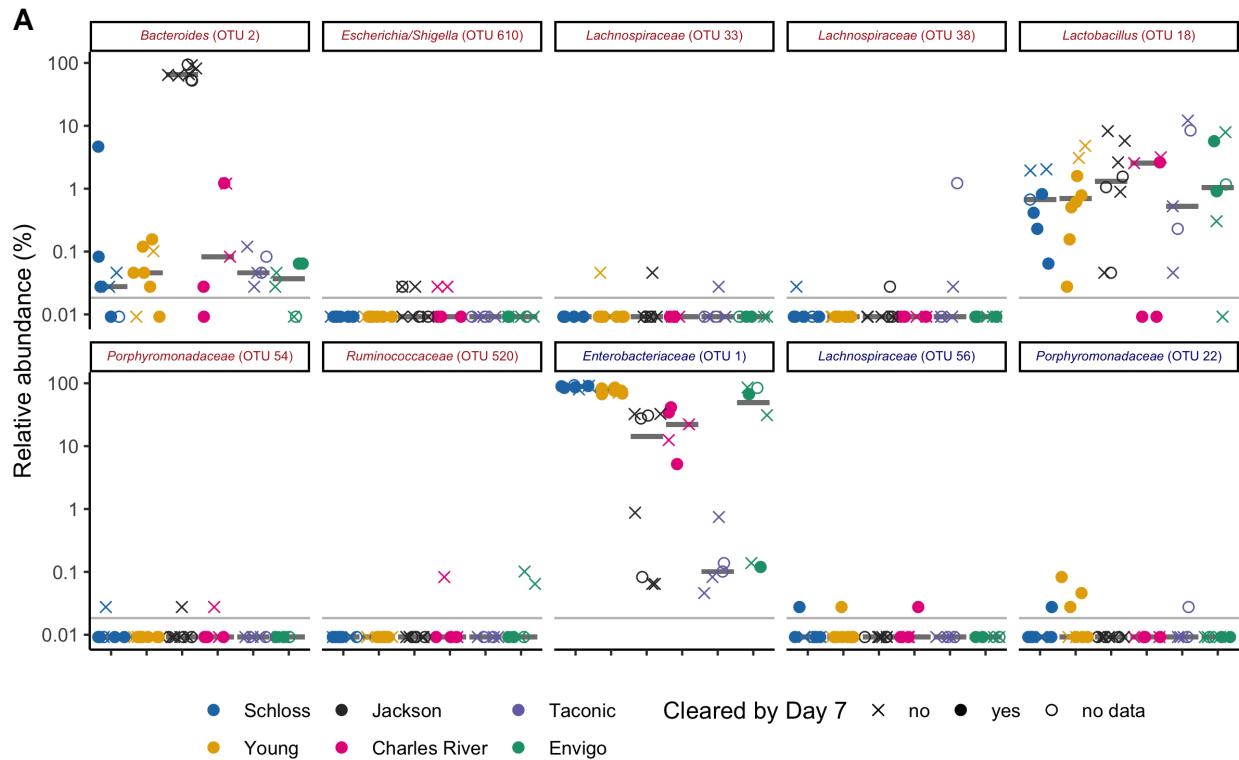
661 **Figure 3. Clindamycin treatment alters bacteria in all sources, but a subset of bacterial**
662 **differences across sources persists.** A-B. Number of observed OTUs (A) and Shannon diversity
663 index values (B) across sources of mice after clindamycin treatment (day 0). Differences between
664 sources were analyzed by Kruskal-Wallis test with Benjamini-Hochberg correction for testing each
665 day of the experiment and the adjusted P value was < 0.05 (Table S1). Significant P values from
666 the pairwise Wilcoxon comparisons between sources with Benjamini-Hochberg correction are
667 displayed as the first initial of each group compared to the group that they are listed above (Table
668 S2). C. PCoA of θ_{YC} distances from stools collected post-clindamycin. Source and the interaction
669 between source and cage effects explained most of the variation observed post-clindamycin
670 (PERMANOVA combined $R^2 = 0.99$, $P < 0.001$, see Table S3). For A-C, each symbol represents a
671 stool sample from an individual mouse, with circles representing experiment 1 mice and triangles
672 representing experiment 2 mice. D. The median (point) and interquartile range (colored lines) of
673 the relative abundances for the 18 OTUs (Table S8) that varied between sources after clindamycin
674 treatment (day 0). E. The median (point) and interquartile range (colored lines) of the top 10 most
675 significant OTUs out of 153 with relative abundances that changed because of the clindamycin
676 treatment (adjusted P value < 0.05). Data were analyzed by paired Wilcoxon signed rank test of
677 mice that had paired sequence data for baseline (day -1) and post-clindamycin (day 0) timepoints
678 ($N = 31$), with Benjamini-Hochberg correction for testing all identified OTUs (Table S9). The gray
679

⁶⁸⁰ vertical line indicates the limit of detection.



682 **Figure 4. Microbiota variation across sources is maintained after *C. difficile* challenge.** A-B.
 683 Number of observed OTUs (A) and Shannon diversity index values (B) across sources of mice
 684 1-day post-infection. Data were analyzed by Kruskal-Wallis test with Benjamini-Hochberg correction
 685 for testing each day of the experiment and the adjusted *P* value was < 0.05 (Table S1). Significant
 686 *P* values from the pairwise Wilcoxon comparisons between sources with Benjamini-Hochberg
 687 correction are displayed as the first initial of each group compared to the group that they are listed
 688 above (Table S2). PCoA of θ_{YC} distances of 1-day post-infection stool samples. Source and
 689 the interaction between source and cage effects explained most of the variation between fecal
 690 communities (PERMANOVA combined $R^2 = 0.88$, $P < 0.001$, Table S3). For A-C: each symbol
 691 represents the value for a stool sample from an individual mouse, circles represent experiment 1
 692 mice and triangles represent experiment 2 mice. D. The median (point) and interquartile range
 693 (colored lines) of the relative abundances for the top 20 most significant OTUs out of the 44 OTUs
 694 that varied between sources 1-day post-infection. The gray vertical line indicates the limit of

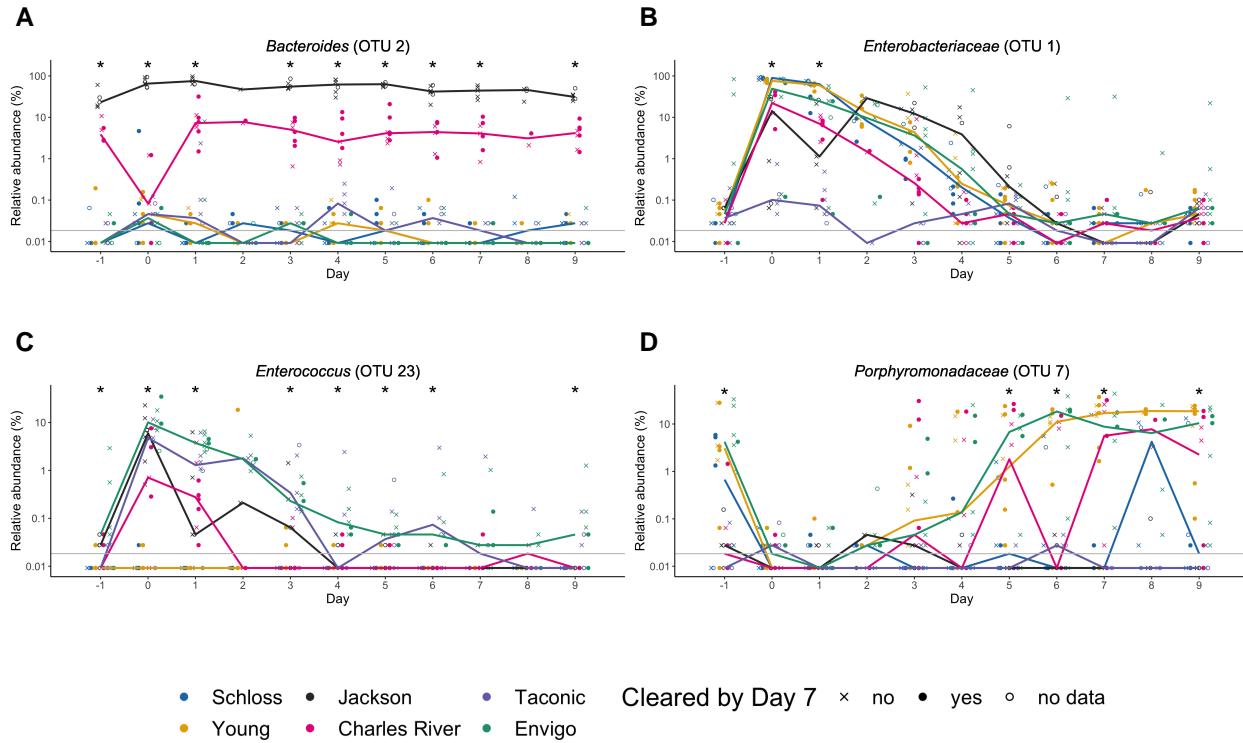
695 detection. For each timepoint OTUs with differential relative abundances across sources of mice
696 were identified by Kruskal-Wallis test with Benjamini-Hochberg correction for testing all identified
697 OTUs (Table S10). E. θ_{YC} distances of fecal samples collected 7-days post-infection relative to the
698 baseline (day -1) sample for each mouse. Each symbol represents an individual mouse. Gray lines
699 represent the median for each source.



700

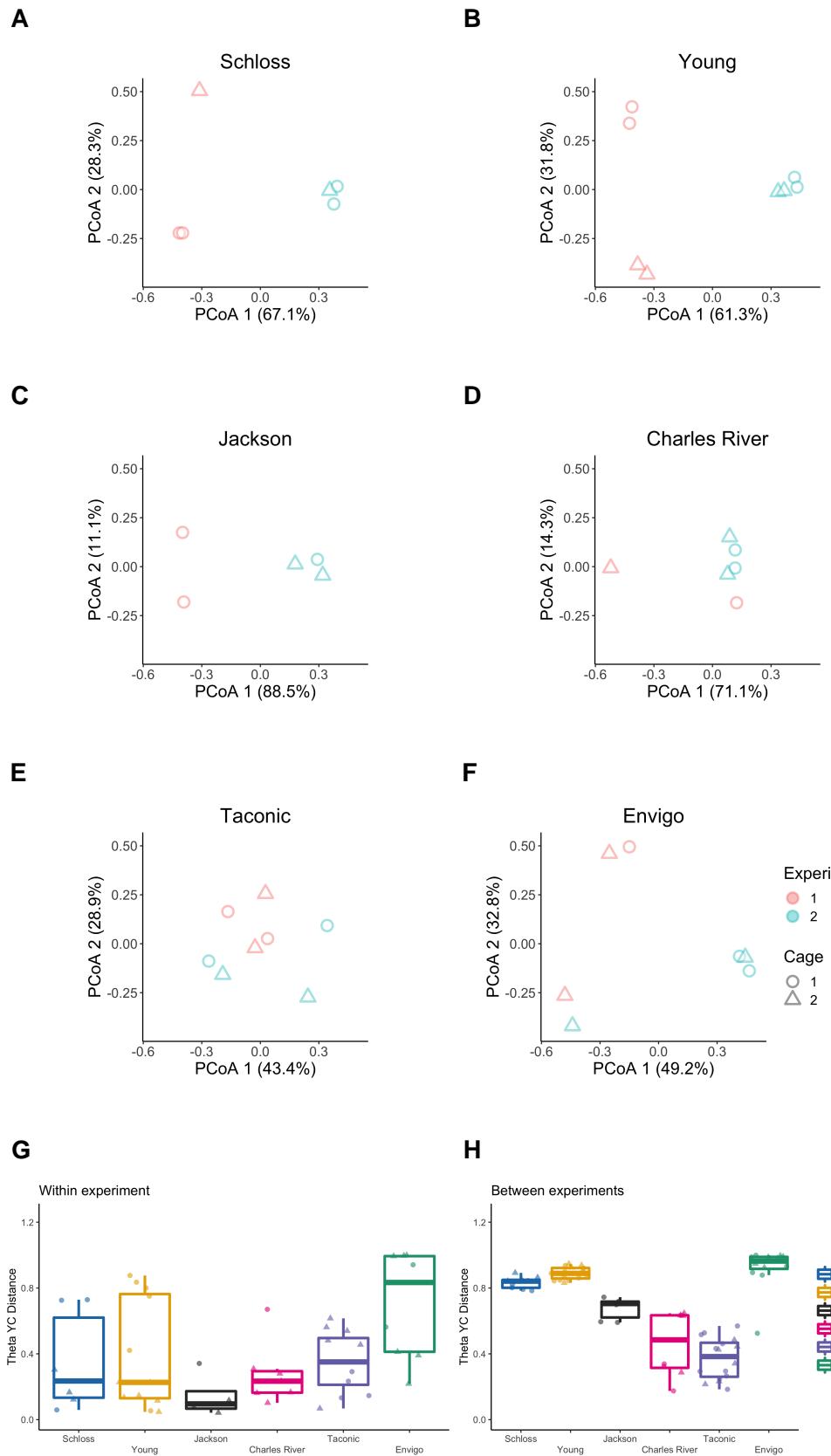
701 **Figure 5. Bacteria that influence whether mice cleared *C. difficile* by day 7.** A.
702 Post-clindamycin (day 0) relative abundance data for the 10 OTUs with the highest rankings

703 based on feature weights in the post-clindamycin (day 0) classification model. Red font represents
704 OTUs that correlated with *C. difficile* colonization and blue font represents OTUs that correlated
705 with clearance. Symbols represent the relative abundance data for an individual mouse. Gray
706 bars indicate the median relative abundances for each source. B. Venn diagram that combines
707 OTUs that were important to the day -1, 0, and 1 classification models (Fig. S4, Table S14) and
708 either overlapped with taxa that varied across sources at the same timepoint, were impacted by
709 clindamycin treatment, or both. Bold OTUs were important to more than 1 classification model.

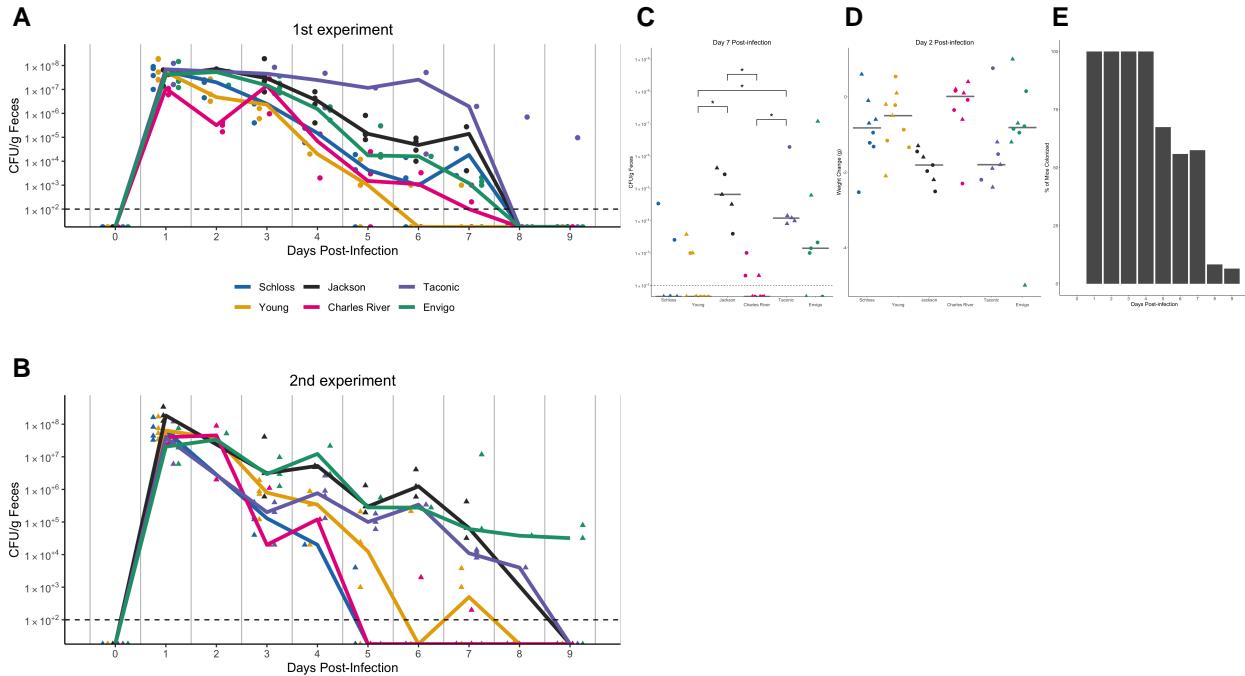


710

711 **Figure 6: OTUs associated with *C. difficile* colonization dynamics vary across sources**
712 **throughout the experiment.** A-D. Relative abundances of bold OTUs from Fig. 5B that were
713 important in at least two classification models are shown over time. A. *Bacteroides* (OTU 2), which
714 varied across sources throughout the experiment. B-C. *Enterobacteriaceae* (B) and *Enterococcus*
715 (C), which significantly varied across sources and were impacted by clindamycin treatment. D.
716 *Porphyromonadaceae* (OTU 7), which was significantly impacted by clindamycin treatment and
717 after examining relative abundance dynamics over the course of the experiment was found to
718 also significantly vary between sources of mice on days -1, 5, 6, 7, and 9 of the experiment.
719 Symbols represent the relative abundance data for an individual mouse. Colored lines indicate
720 the median relative abundances for each source. The gray horizontal line represents the limit of
721 detection. Timepoints where differences between sources of mice were statistically significant by
722 Kruskal-Wallis test with Benjamini-Hochberg correction for testing across multiple days (Table S15)
723 are identified by the asterisk above each timepoint (*, P < 0.05).

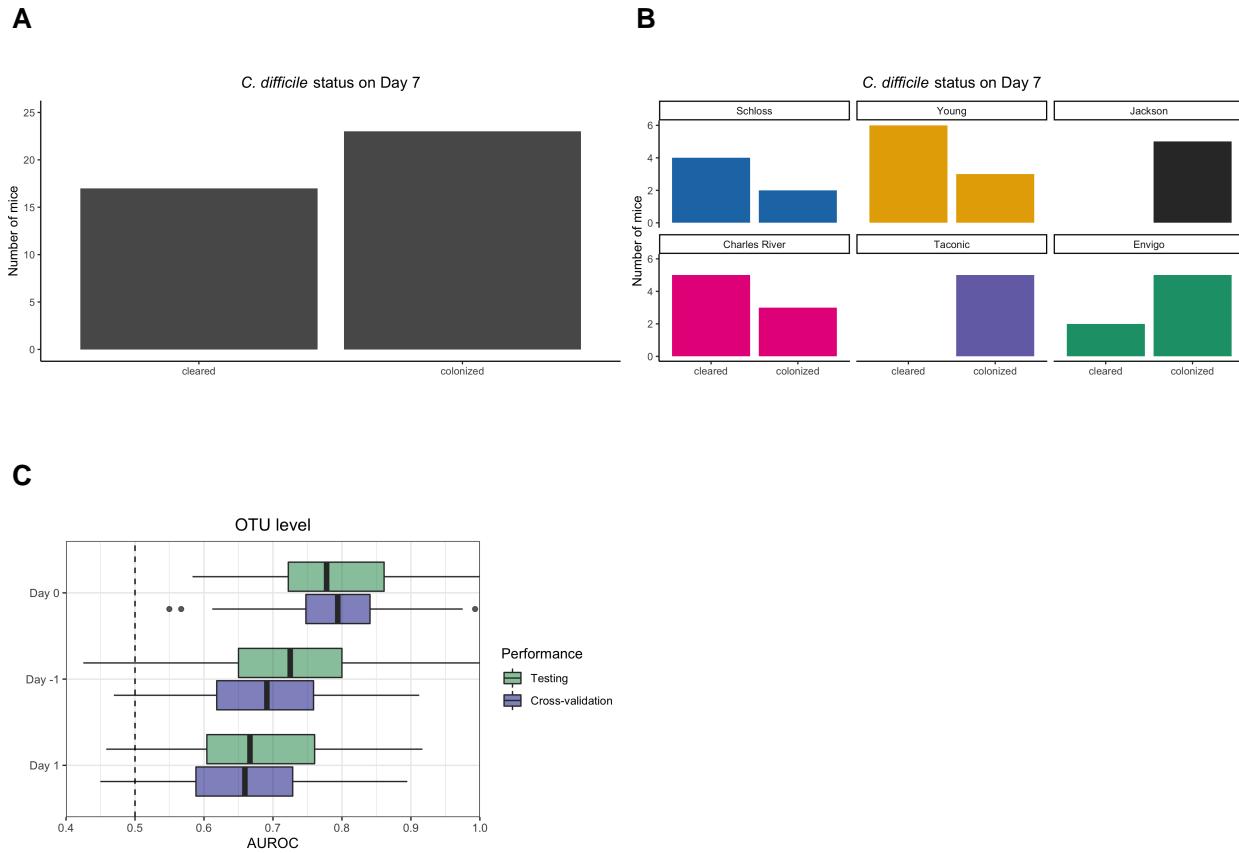
**Figure S1.**

725 **Bacterial communities vary between experiments for some sources.** A-F. PCoA of θ_{YC}
726 distances for the baseline fecal bacterial communities within each source of mice. Each symbol
727 represents a stool sample from an individual mouse with color corresponding to experiment and
728 shape representing cage mates. Experiment number and cage effects explained most of the
729 observed variation for samples from the Schloss (PERMANOVA combined $R^2 = 0.99$; $P \leq 0.033$)
730 and Young (combined $R^2 = 0.95$; $P \leq 0.03$) mice (Table S4). G-H: Boxplots of the θ_{YC} distances of
731 the 6 sources of mice relative to mice within the same source and experiment (G) or mice within
732 the same source and between experiments (H) at baseline (day -1). Symbols represent individual
733 mouse samples: circles for experiment 1 and triangles for experiment 2.



734

735 **Figure S2. *C. difficile* CFU variation across sources varies slightly between the 2**
 736 **experiments.** A-B. *C. difficile* CFU/gram of stool quantification over time for experiment 1 (A) and
 737 (B). Experiments were conducted approximately 3 months apart. Lines represent the median
 738 CFU for each source, symbols represent individual mice and the black line represents the limit
 739 of detection. C. *C. difficile* CFU/gram stool 7-days post-infection across sources of mice with an
 740 asterisk for pairwise Wilcoxon comparisons with Benjamini-Hochberg correction where $P < 0.05$.
 741 D. Mouse weight change 2-days post-infection across sources of mice, no pairwise Wilcoxon
 742 comparisons were significant after Benjamini-Hochberg correction. For C-D: circles represent
 743 experiment 1 mice, triangles represent experiment 2 mice and gray lines indicate the median
 744 values for each group. E. Percent of mice that were colonized with *C. difficile* over the course of the
 745 experiment. Each day the percent is calculated based on the mice where *C. difficile* CFU was
 746 quantified for that particular day. Total N for each day: day 1 (N = 42), day 2 (N = 20), day 3 (N =
 747 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 =
 748 46).

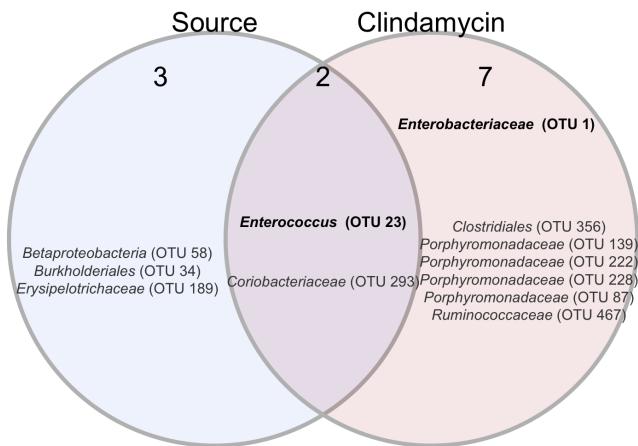


749

750 **Figure S3. Bacterial community composition before, after clindamycin perturbation, and**
 751 **post-infection can predict *C. difficile* colonization status 7 days post-challenge.** A. Bar
 752 graph visualizations of overall 7-days post-infection *C. difficile* colonization status that were used as
 753 classification outcomes to build L2-regularized logistic regression models. Mice were classified as
 754 colonized or cleared (not detectable at the limit of detection of 100 CFU) based on CFU g/stool data
 755 from 7 days post-infection. B. *C. difficile* CFU status on Day 7 within each mouse source. N = 8-9
 756 mice per group. C. L2-regularized logistic regression classification model area under the receiving
 757 operator characteristic curve (AUROCs) to predict *C. difficile* CFU on day 7 post-infection (Fig. 2B,
 758 Fig. S2C) based on the OTU community relative abundances at baseline (day -1), post-clindamycin
 759 (day 0), and 1-day post-infection. All models performed better than random chance (AUROC =
 760 0.5, all $P < 0.001$, Table S12) and the model built with post-clindamycin bacterial OTU relative
 761 abundances had the best performance ($P_{FDR} < 0.001$ for all pairwise comparisons, Table S13).
 762 See Table S14 for list of the 20 OTUs that were ranked as most important to each model.

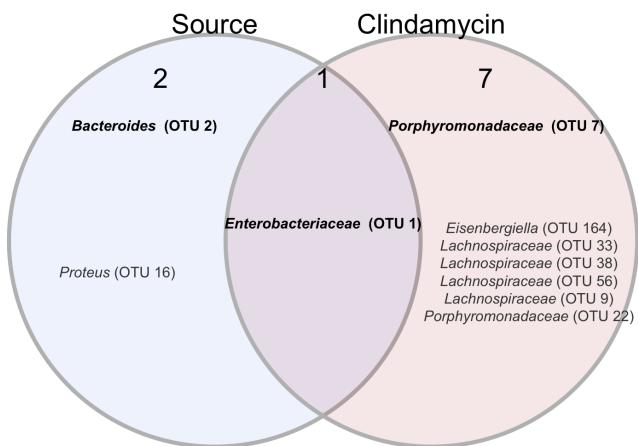
A

Day -1 model OTU comparisons



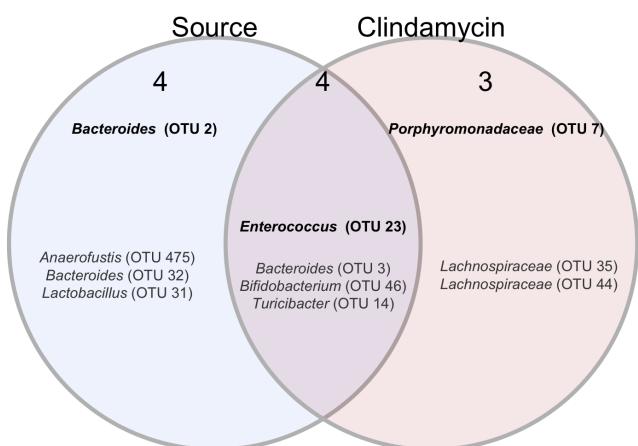
B

Day 0 model OTU comparisons



C

Day 1 model OTU comparisons

**Figure**

763

- 764 **S4. OTUs from classification models based on baseline, post-clindamycin treatment, or**
 765 **1-day post-infection community data vary by source, clindamycin treatment, or both. A-C.**
 766 Venn diagrams of OTUs from the top 20 OTUs from the baseline (A), post-clindamycin treatment (B),

⁷⁶⁷ and 1-day post-infection (C) classification models (Table S14) that overlapped with OTUs that varied
⁷⁶⁸ across sources at the corresponding timepoint (Tables S5, 8, 10), were impacted by clindamycin
⁷⁶⁹ treatment (Table S9), or both. Bold OTUs were important to more than 1 classification model.

770 **Supplementary Tables and Movie**

771 **Movie S1. Large shifts in bacterial community structures occurred after clindamycin and**
772 ***C. difficile* infection.** PCoA of θ_{YC} distances animated from days -1 through 9 of the experiment.
773 Source was the variable that explained the most observed variation across fecal communities
774 (PERMANOVA source $R^2 = 0.35$, $P = 0.0001$, Table S11) followed by interactions between cage
775 effects and day of the experiment. Transparency of the symbol corresponds to the day of the
776 experiment, each symbol represents a sample from an individual mouse at a specific timepoint.
777 Circles represent mice from experiment 1 and triangles represent mice from expeirment 2.

778 **Tables S1-S15. Excel workbook of Tables S1-S15.**

779 **Table S1. Alpha diversity metrics Kruskal-Wallis statistical results.**

780 **Table S2. Alpha diversity metrics pairwise Wilcoxon statistical results.**

781 **Table S3. PERMANOVA results for mice at baseline (day -1), post-clindamycin (day 0), and**
782 **post-infection (day 1).**

783 **Table S4. PERMANOVA results for each source of mice at baseline (day -1).**

784 **Table S5. OTUs with relative abumdances that significantly vary between sources at**
785 **baseline (day -1).**

786 **Table S6. *C. difficile* CFU statistical results.**

787 **Table S7. Mouse weight change statistical results.**

788 **Table S8. OTUs with relative abundances that significantly vary between sources**
789 **post-clindamycin (day 0).**

790 **Table S9. OTUs with relative abundances that significantly changed after clindamycin**
791 **treatment.**

792 **Table S10. OTUs with relative abundances that significantly vary between sources 1-day**
793 **post-infection.**

794 **Table S11.** PERMANOVA results for mice across all timepoints.

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

797 **Table S13.** Pairwise comparisons of L2-regularized logistic regression model performances.

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

800 **Table S15.** OTUs with relative abundances that significantly varied between sources of mice
801 on at least 1 day of the experiment by Kruskal-Wallis test.