

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 university colony or vendor. We treated mice from 6 different sources (2
5 University of Michigan colonies and 4 vendors) with a single clindamycin dose, followed by a *C.*
6 *difficile* challenge 1 day later and then measured *C. difficile* colonization levels through 9 days
7 post-infection. The microbiota was profiled via 16S rRNA gene sequencing to examine the variation
8 across sources and alterations due to clindamycin treatment and *C. difficile* challenge. While all
9 sources of mice were colonized 1-day post-infection, variation emerged from days 3-7 post-infection
10 with animals from some sources colonized with *C. difficile* for longer and at higher levels. We
11 identified bacteria that varied in relative abundance across sources and throughout the experiment.
12 Some bacteria were consistently impacted by clindamycin treatment in all sources of mice including
13 *Lachnospiraceae*, *Ruminococcaceae*, and *Enterobacteriaceae*. To identify bacteria that were most
14 important to colonization regardless of the source, we created logistic regression models that
15 successfully classified mice based on whether they cleared *C. difficile* by 7 days post-infection using
16 baseline, post-clindamycin, and post-infection community composition data. With these models,
17 we identified 4 bacteria that varied across sources (*Bacteroides*), were altered by clindamycin
18 (*Porphyromonadaceae*), or both (*Enterobacteriaceae* and *Enterococcus*). Microbiota variation
19 across sources better emulates human interindividual variation and can help identify bacterial
20 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 has been established as a key risk factor, there is variation in who becomes asymptotically
24 colonized, develops an infection, or has an infection with adverse outcomes. *C. difficile* infection
25 (CDI) mouse models are widely used to answer a variety of *C. difficile* pathogenesis questions.
26 However, the inter-individual variation between mice is less than what is observed in humans,
27 particularly if just one source of mice is used. In this study, we administered clindamycin to mice
28 from 6 different breeding colonies and challenged them with *C. difficile*. Interestingly, only a subset

29 of the bacteria that vary across sources were associated with how long *C. difficile* was able to
30 colonize. Future studies examining the interplay between the microbiota and *C. difficile* should
31 consider using mice from multiple sources to better reflect human interindividual variation.

32 **Introduction**

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

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

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

58 **Results**

59 **Clindamycin treatment renders all mice susceptible to *C. difficile* 630 colonization**
60 **regardless of source.** To test how the microbiotas of mice from different sources impact
61 colonization dynamics after clindamycin exposure, we utilized C57BL/6 mice from 6 different
62 sources: two colonies from the University of Michigan that were split from each other 10 years ago
63 (the Young and Schloss lab colonies) and four commercial source: the Jackson Laboratory, Charles
64 River Laboratories, Taconic Biosciences, and Envigo (which was formerly Harlan). These 4 vendors
65 were chosen because they represent commonly used vendors for CDI studies in mice (26, 34–40).

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

87 By 9 days post-infection the majority of the mice from all sources had cleared *C. difficile* (Fig.
88 S1C) with the exception of 1 Taconic mouse from the first experiment and 2 Envigo mice from the
89 second experiment. Thus, clindamycin rendered all mice susceptible to *C. difficile* 630 colonization,
90 regardless of source, but there was significant variation in disease phenotype across the sources of
91 mice.

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

110 Next, we compared the community structure of mice from the 6 sources over the course of
111 the experiment using principal coordinate analysis (PCoA) of the θ_{YC} distances. Permutational
112 multivariate analysis of variance (PERMANOVA) analysis revealed source was the dominant factor
113 that explained the observed variation across fecal communities ($R^2 = 0.35$, $P = 0.0001$) followed
114 by interactions between cage (mice from the same sources were housed together, primarily at a
115 density of 2 mice per cage) and day of the experiment (Movie S1 and Table S5). Mice that are

116 co-housed have been shown to have similar gut microbiotas due to coprophagy (45) and since
117 mice within the same source were housed together, it is not surprising that cage also contributes
118 to the observed microbiota variation. Since the majority of the perturbations happened over the
119 initial days of the experiment, we decided to focus on the communities at baseline (day -1), after
120 clindamycin treatment (day 0), and post-infection (day 1). For all 3 timepoints, the source and
121 cage interaction significantly explained most of the observed community variation (combined R²
122 = 0.90, 0.99, 0.88, respectively; P = 0.0001; Fig. 3 and Table S6). We also compared baseline
123 communities across the 2 experiments, and found experiment and cage significantly explained the
124 observed variation only for the Schloss and Young lab mouse colonies (Fig. S2A-B and Table S7).
125 However, most of the vendors also clustered by experiment (Fig. S2C-D, F), suggesting there was
126 some community variation between the 2 experiments within each source. Thus, source was the
127 factor that explained the most variation observed in the bacterial communities. Importantly with the
128 exception of the Schloss and Young colonies, the community of each source clustered apart from
129 one another suggesting each community had a unique response to clindamycin treatment and *C.*
130 *difficile* challenge.

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

142 After finding differences at the community level, we next identified the bacteria that varied across
143 sources of mice over the initial days of the experiment. We examined bacterial relative abundances
144 at the operational taxonomic unit (OTU) level. Focusing on the baseline communities first, there were

145 268 OTUs (Table S8) with relative abundances that varied across sources. Clindamycin treatment
146 reduced the number of taxa with relative abundances that varied across sources to 18 OTUs (Table
147 S8). After *C. difficile* challenge, there were 44 OTUs (Table S8) with significantly different relative
148 abundances across sources, as the communities started to recover from antibiotic treatment. In
149 spite of the experimental perturbations that occurred during these 3 timepoints, there were 12 OTUs
150 (Fig 5A-C) with relative abundances that consistently varied across sources. Importantly, some
151 of the OTUs that consistently varied across sources also shifted with clindamycin treatment. For
152 example, *Proteus* increased after clindamycin treatment, but only in Taconic mice. *Enterococcus*
153 was primarily found only in mice purchased from commercial vendors and also increased after
154 clindamycin treatment. In summary, mouse bacterial communities varied significantly according to
155 source throughout the course of the experiment and a consistent subset of bacterial taxa remained
156 different across sources regardless of clindamycin and *C. difficile* challenge.

157 **Clindamycin treatment alters a subset of taxa that were found in all sources.** Although there
158 were bacteria that consistently varied across sources, we also wanted to identify the bacteria that
159 shifted after clindamycin treatment, regardless of source. By analyzing all mice that had sequence
160 data from fecal samples collected at baseline and after clindamycin treatment, we identified
161 153 OTUs that were altered after clindamycin treatment (Fig. 6 and Tables S9). Interestingly,
162 when we compared the list of significant clindamycin impacted bacteria with the bacteria that
163 consistently varied across groups over the initial 3 timepoints of our experiment, we found 3 OTUs
164 (*Lachnospiraceae* (OTU 130), *Lactobacillus* (OTU 6), *Enterococcus* (OTU 23)) overlapped (Fig.
165 5, Fig. 6C-D). These findings demonstrate that clindamycin has a consistent impact on the fecal
166 bacterial communities of mice from all sources and only a subset of the OTUs also varied across
167 sources.

168 **Source-specific and clindamycin impacted bacteria distinguish *C. difficile* colonization**
169 **status in mice.** After identifying taxa that varied by source, changed after clindamycin treatment,
170 or both, we next wanted to determine which taxa were influencing the variation in *C. difficile*
171 colonization at day 7 (Fig. 1D, Fig. S1C). We trained three L2-regularized logistic regression
172 models with input bacterial community data from the baseline, post-clindamycin, and post-infection
173 timepoints of the experiment to predict *C. difficile* colonization status on day 7 (Fig. S3A-B). All

models were better at predicting *C. difficile* colonization status on day 7 than random chance (all $P \leq 5.2e-31$; Table S12). Interestingly, the model based on the post-clindamycin (day 0) community OTU data performed significantly better than all other models with an area under the receiving operator characteristic curve (AUROC) of 0.75 ($P_{FDR} \leq 3.1e-11$ for pairwise comparisons; Table S13). Thus, we were able to use bacterial relative abundance data alone to differentiate mice that had cleared *C. difficile* before day 7 from the mice still colonized with *C. difficile* at that timepoint. Interestingly, the model built with OTU relative abundance data post-clindamycin treatment had the best performance, suggesting how the bacterial community responds to clindamycin treatment has the greatest influence on subsequent *C. difficile* colonization dynamics.

Next, to examine the bacteria that were driving each model's performance, we pulled out the top 20 taxa that had the highest absolute feature weights in each of the 6 models (Tables S14-15). First, we looked at OTUs from the model with the best performance that was based on the post-clindamycin treatment bacterial community data. While most of the 20 OTUs had low relative abundances on day 0, *Enterobacteriace*, *Bacteroides* and *Proteus* had high relative abundances in at least one source of mice and significantly varied across sources (Fig. 7A). Next, the top 20 taxa from each model were compared to the list of taxa that varied across source at the same timepoint (Fig. 5 and Tables S8-9) and the taxa that were altered by clindamycin treatment (Fig. 6 and Table S10-11). We found a subset of OTUs that were important to the model and overlapped with bacteria that varied by either source, clindamycin treatment, or both (Fig. S4, S5A-C). Combining the overall results for the 3 OTU models identified 14 OTUs associated with source, 21 OTUs associated with clindamycin treatment, and 6 OTUs associated with both (Fig. 7B). Several OTUs (*Bacteroides* (OTU 2), *Enterococcus* (OTU 23), *Enterobacteriaceae* (OTU 1), *Porphyromonadaceae* (OTU 7)) appeared across at least 2 models, so we examined how the relative abundances of these key taxa varied over the course of the experiment (Fig. 8). Throughout the experiment, there was at least 1 timepoint where relative abundances of these OTUs significantly varied across sources (Table S13). Interestingly, there were no OTUs that emerged as consistently enriched or depleted in mice that were colonized past 7 days post-infection with *C. difficile* 630, suggesting multiple bacteria influence the time needed to clear the infection. Together, these results suggest the initial bacterial communities and their responses to clindamycin influence the clearance of *C. difficile*.

203 **Discussion**

204 By running our CDI model with mice from 6 different sources, we were able to identify bacterial
205 taxa that were unique to sources throughout the experiment as well as taxa that were universally
206 impacted by clindamycin. We trained L2 logistic regression models with baseline, post-clindamycin
207 treatment, and post-infection fecal community data that could predict whether mice cleared *C.*
208 *difficile* by 7 days post-infection better than random chance. We identified *Bacteroides* (OTU
209 2), *Enterococcus* (OTU 23), *Enterobacteriaceae* (OTU 1), *Porphyromonadaceae* (OTU 7) (Fig.
210 8) as candidate bacteria within these communities that were influencing variation in *C. difficile*
211 colonization dynamics since these bacteria were all important in the logistic regression models and
212 varied by source, were impacted by clindamycin treatment, or both. Overall, our results demonstrate
213 clindamycin is sufficient to render mice from multiple sources susceptible to CDI and only a subset
214 of the interindividual microbiota variation across mice from different sources was associated with
215 the time needed to clear *C. difficile*.

216 Other studies have taken similar approaches by using mice from multiple sources to identify bacteria
217 that either promote colonization resistance or increase susceptibility to enteric infections (22, 23,
218 26–30). For example, in the context of *Salmonella* infections, *Enterobacteriaceae* and segmented
219 filamentous bacteria have emerged as protective (22, 27). A previous study with *C. difficile* identified
220 an endogenous protective *C. difficile* strain LEM1 that bloomed after antibiotic treatment in mice
221 from Jackson or Charles River Laboratories, but not Taconic that protected mice against the more
222 toxicogenic *C. difficile* VPI10463 (26). Given that we obtained mice from the same vendors, we
223 checked all mice for endogenous *C. difficile* by plating stool samples that were collected after
224 clindamycin treatment. However, we did not identify any endogenous *C. difficile* strains prior to
225 challenge, suggesting there were no endogenous protective strains in the mice we received and
226 other bacterial taxa mediated the variation in *C. difficile* colonization across sources.

227 Although all mice were susceptible to *C. difficile* colonization, mice from Jackson, Taconic, and
228 Envigo mice tended to remain colonized through at least 7 days post-infection. We identified
229 a subset of bacteria that were important in predicting whether a mouse was still colonized with
230 *C. difficile* 7 days post-infection. These results suggest a subset of the bacterial community is

231 responsible for determining the length of time needed to clear *C. difficile* colonization.

232 Differences in CDI mouse model studies have been attributed to different intestinal microbiotas
233 of mice from different sources. For example, groups using the same clindamycin treatment and
234 C57BL/6 mice had different *C. difficile* outcomes, one having sustained colonization (32), while
235 the other had transient (18), despite both using *C. difficile* VPI 10643. Baseline differences in the
236 microbiota composition have been hypothesized to partially explain the differences in colonization
237 outcomes and overall susceptibility to *C. difficile* after treatment with the same antibiotic (13,
238 31). The bacterial perturbations induced by clindamycin treatment have been well characterized
239 and our findings agree with previous CDI mouse model work demonstrating *Enterococcus* and
240 *Enterobacteriaceae* were associated with *C. difficile* susceptibility and *Porphyromonadaceae*,
241 *Lachnospiraceae*, *Ruminococcaceae*, and *Turicibacter* were associated with resistance (19, 21, 32,
242 33, 41, 46–48). While we have demonstrated that susceptibility is uniform across sources of mice
243 after clindamycin treatment, there could be different outcomes for either susceptibility or clearance
244 in the case of other antibiotic treatments. The *C. difficile* strain used could also be contributing
245 to the variation in *C. difficile* outcomes seen across different research groups (47). We found the
246 time needed to naturally clear *C. difficile* varied across sources of mice implying that at least in the
247 context of the same perturbation, microbiota differences seemed to influence infection outcome
248 more than susceptibility. More importantly, we were able to reduce the variation observed across
249 sources to identify a subset of OTUs that were also important for predicting *C. difficile* colonization
250 status 7 days post-infection. Since all but 3 mice eventually cleared *C. difficile* 630 by 9 days
251 post-infection and the model built with the post-clindamycin OTU relative abundance data had the
252 best performance, our results suggest clindamycin treatment had a large role in determining *C.*
253 *difficile* susceptibility and clearance in the mice.

254 Our approach successfully increased the diversity of murine bacterial communities tested in
255 our clindamycin *C. difficile* model. One alternative approach that has been used in some CDI
256 studies (49–54) is to associate mice with human microbiotas. However, a major caveat to this
257 method is the substantial loss of human microbiota community members upon transfer to mice
258 (55, 56). Additionally with the exception of 2 recent studies (49, 50), most of the CDI mouse model
259 studies to date associated mice with just 1 types of human microbiota either from a single donor

260 or a single pool from multiple donors (51–54), which does not aid in the goal of modeling the
261 interpersonal variation seen in humans to understand how the microbiota influences susceptibility to
262 CDIs and adverse outcomes. Importantly, our study using mice from 6 different sources increased
263 the variation between groups of mice compared to using 1 source alone, to better reflect the
264 inter-individual microbiota variation observed in humans. Encouragingly, decreased *Bifidobacterium*,
265 *Porphyromonas*, *Ruminococcaceae* and *Lachnospiraceae* and increased *Enterobacteriaceae*,
266 *Enterococcus*, *Lactobacillus*, and *Proteus* have all been associated with human CDIs (7) and were
267 well represented in our study, suggesting most of the mouse sources are suitable for gaining insights
268 into microbiota associated factors influencing *C. difficile* colonization and infections in humans. An
269 important exception was *Enterococcus*, which was primarily absent from the mice from University
270 of Michigan colonies and *Proteus*, which was only found in Taconic mice. Importantly, the fact that
271 some CDI-associated bacteria were only found in a subset of mice has important implications for
272 future CDI mouse model studies.

273 There are several limitations to our work. The microbiome is composed of viruses, fungi, and
274 parasites in addition to bacteria, and these non-bacterial members can also vary across mouse
275 vendors (57, 58). While our study focused solely on the bacterial portion, viruses and fungi have
276 also begun to be implicated in the context of CDIs or FMT treatments for recurrent CDIs (35, 59–62).
277 Beyond community composition, the metabolic function of the microbiota also has a CDI signature
278 (20, 48, 63, 64) and can vary across mice from different sources (65). For example, microbial
279 metabolites, particularly secondary bile acids and butyrate production, have been implicated as
280 important contributors to *C. difficile* resistance (33, 47). Although, we only looked at composition,
281 *Ruminococcaceae* and *Lachnospiraceae* both emerged as important taxa for classifying day 7 *C.*
282 *difficile* colonization status and metagenomes from these bacteria have been shown to contain
283 the bile acid-inducible gene cluster necessary for secondary bile acid formation and ability to
284 produce butyrate (52, 66). Interestingly, butyrate has previously been shown to vary across vendors
285 and mediated resistance to *Citrobacter rodentium* infection, a model of enterohemorrhagic and
286 enteropathogenic *Escherichia coli* infections (23). Evidence for immunological toning differences
287 in IgA and Th17 cells across mice from different vendors have also been documented and (67,
288 68) may also influence response to CDI, particularly in the context of severe CDIs (69, 70). The

289 outcome after *C. difficile* exposure depends on a multitude of factors, including age, diet, and
290 immunity; all of which are also influenced by the microbiota.

291 We have demonstrated that the ways baseline microbiotas from different mouse sources respond
292 to clindamycin treatment influences the length of time mice remained colonized with *C. difficile* 630.
293 For those interested in dissecting the contribution of the microbiome to *C. difficile* pathogenesis
294 and treatments, using multiple sources of mice may yield more insights than a single model alone.
295 Furthermore, for studies wanting to examine the interplay between a particular bacterial taxon such
296 as *Enterococcus* and *C. difficile*, these results could serve as a resource for selecting which mice
297 to order to address the question. Using mice from multiple sources helps model the interpersonal
298 microbiota variation among humans to aid our understanding of how the gut microbiota contributes
299 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, as well as code tutorials and feedback through Code club. In particular, we
305 want to thank Begüm Topçuoğlu for help with implementing L2 logistic regression models using her
306 pipeline, Ana Taylor for help with media preparation and sample collection, and Nicholas Lesniak for
307 his critical feedback on the manuscript. We also thank members of Vincent Young's lab, particularly
308 Kimberly Vendrov, for guidance with the *C. difficile* infection mouse model and donating the mice.
309 We also want to thank the Unit for Laboratory Animal Medicine at the University of Michigan for
310 maintaining our mouse colony and providing the institutional support for our mouse experiments.
311 Finally, we thank Kwi Kim, Austin Campbell, and Kimberly Vendrov for their help in maintaining the
312 Schloss lab's anaerobic chamber.

313 **Materials and Methods**

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

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

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

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

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

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 (79) in R as previously
367 described (80) 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 and
372 change the data split ratios. The modified repository to regenerate this analysis is available at
373 https://github.com/tomkosev/ML_pipeline_microbiome.

374 **(vii) Statistical analysis.** All statistical tests were performed in R (v 3.5.2) (78). The Kruskal-Wallis
375 test was used to analyze differences in *C. difficile* CFU, mouse weight change, and alpha
376 diversity across vendors 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
381 by pairwise Wilcoxon comparisons with Benjamini-Hochberg correction. OTUs impacted by
382 clindamycin treatment were identified using the Wilcoxon signed rank test with matched pairs
383 of mice samples for day -1 and day 0. To determine whether classification models had better
384 performance (test AUROCs) than random chance (0.5), we used the one-sample Wilcoxon signed
385 rank test. To examine whether there was an overall difference in predictive performance across the
386 6 classification models we used the Kruskal-Wallis test followed by pairwise Wilcoxon comparisons
387 with Benjamini-Hochberg correction for multiple hypothesis testing. The tidyverse package was
388 used to wrangle and graph data (v 1.3.0) (81).

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

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

393 **References**

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

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

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

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

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

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

- 545 Britton RA. 2018. Dietary trehalose enhances virulence of epidemic *Clostridium difficile*. Nature
546 553:291–294.
- 547 54. Hryckowian AJ, Treuren WV, Smits SA, Davis NM, Gardner JO, Bouley DM, Sonnenburg JL.
548 2018. Microbiota-accessible carbohydrates suppress *Clostridium difficile* infection in a murine
549 model. *Nature Microbiology* 3:662–669.
- 550 55. Fouladi F, Glenny EM, Bulik-Sullivan EC, Tsilimigras MCB, Sioda M, Thomas SA, Wang Y, Djukic
551 Z, Tang Q, Tarantino LM, Bulik CM, Fodor AA, Carroll IM. 2020. Sequence variant analysis reveals
552 poor correlations in microbial taxonomic abundance between humans and mice after gnotobiotic
553 transfer. *The ISME Journal*.
- 554 56. Walter J, Armet AM, Finlay BB, Shanahan F. 2020. Establishing or exaggerating causality for
555 the gut microbiome: Lessons from human microbiota-associated rodents. *Cell* 180:221–232.
- 556 57. Rasmussen TS, Vries L de, Kot W, Hansen LH, Castro-Mejía JL, Vogensen FK, Hansen AK,
557 Nielsen DS. 2019. Mouse vendor influence on the bacterial and viral gut composition exceeds the
558 effect of diet. *Viruses* 11:435.
- 559 58. Mims TS, Abdallah QA, Watts S, White C, Han J, Willis KA, Pierre JF. 2020. Variability in
560 interkingdom gut microbiomes between different commercial vendors shapes fat gain in response
561 to diet. *The FASEB Journal* 34:1–1.
- 562 59. Stewart DB, Wright JR, Fowler M, McLimans CJ, Tokarev V, Amaniera I, Baker O, Wong H-T,
563 Brabec J, Drucker R, Lamendella R. 2019. Integrated meta-omics reveals a fungus-associated
564 bacteriome and distinct functional pathways in *Clostridioides difficile* infection. *mSphere* 4.
- 565 60. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, Grasis JA, Cassidy L, Tholey A,
566 Fickenscher H, Seegert D, Rosenstiel P, Schreiber S. 2017. Efficacy of sterile fecal filtrate transfer
567 for treating patients with *Clostridium difficile* infection. *Gastroenterology* 152:799–811.e7.
- 568 61. Zuo T, Wong SH, Lam K, Lui R, Cheung K, Tang W, Ching JYL, Chan PKS, Chan MCW, Wu
569 JCY, Chan FKL, Yu J, Sung JJY, Ng SC. 2017. Bacteriophage transfer during faecal microbiota
570 transplantation in *Clostridium difficile* infection is associated with treatment outcome. *Gut*

- 571 gutjnl–2017–313952.
- 572 62. Zuo T, Wong SH, Cheung CP, Lam K, Lui R, Cheung K, Zhang F, Tang W, Ching JYL, Wu JCY,
573 Chan PKS, Sung JJY, Yu J, Chan FKL, Ng SC. 2018. Gut fungal dysbiosis correlates with reduced
574 efficacy of fecal microbiota transplantation in *Clostridium difficile* infection. Nature Communications
575 9.
- 576 63. Robinson JI, Weir WH, Crowley JR, Hink T, Reske KA, Kwon JH, Burnham C-AD, Dubberke
577 ER, Mucha PJ, Henderson JP. 2019. Metabolomic networks connect host-microbiome processes to
578 human *Clostridioides difficile* infections. Journal of Clinical Investigation 129:3792–3806.
- 579 64. Fletcher JR, Erwin S, Lanzas C, Theriot CM. 2018. Shifts in the gut metabolome and *Clostridium*
580 *difficile* transcriptome throughout colonization and infection in a mouse model. mSphere 3.
- 581 65. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, Liu C, Fang
582 Z, Chou J, Glanville J, Hao Q, Kotowska D, Colding C, Licht TR, Wu D, Yu J, Sung JJY, Liang Q, Li
583 J, Jia H, Lan Z, Tremaroli V, Dworzynski P, Nielsen HB, Bäckhed F, Doré J, Chatelier EL, Ehrlich
584 SD, Lin JC, Arumugam M, Wang J, Madsen L, Kristiansen K. 2015. A catalog of the mouse gut
585 metagenome. Nature Biotechnology 33:1103–1108.
- 586 66. Vital M, Rud T, Rath S, Pieper DH, Schlüter D. 2019. Diversity of bacteria exhibiting bile
587 acid-inducible 7alpha-dehydroxylation genes in the human gut. Computational and Structural
588 Biotechnology Journal 17:1016–1019.
- 589 67. Fransen F, Zagato E, Mazzini E, Fosso B, Manzari C, Aidy SE, Chiavelli A, D'Erchia AM,
590 Sethi MK, Pabst O, Marzano M, Moretti S, Romani L, Penna G, Pesole G, Rescigno M. 2015.
591 BALB/c and C57BL/6 mice differ in polyreactive IgA abundance, which impacts the generation of
592 antigen-specific IgA and microbiota diversity. Immunity 43:527–540.
- 593 68. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee
594 CA, Lynch SV, Tanoue T, Imaoka A, Itoh K, Takeda K, Umesaki Y, Honda K, Littman DR. 2009.
595 Induction of intestinal th17 cells by segmented filamentous bacteria. Cell 139:485–498.
- 596 69. Azrad M, Hamo Z, Tkhawkho L, Peretz A. 2018. Elevated serum immunoglobulin a levels in

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

621 package.

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

624 79. Kuhn M. 2008. Building predictive models inRUsing thecaretPackage. Journal of Statistical
625 Software 28.

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

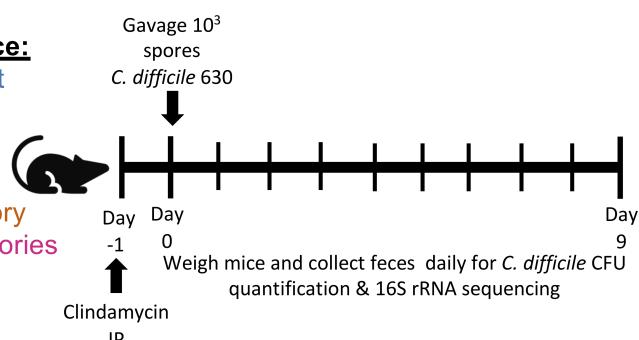
628 81. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes
629 A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D,
630 Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the
631 tidyverse. Journal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

632 **Figures**

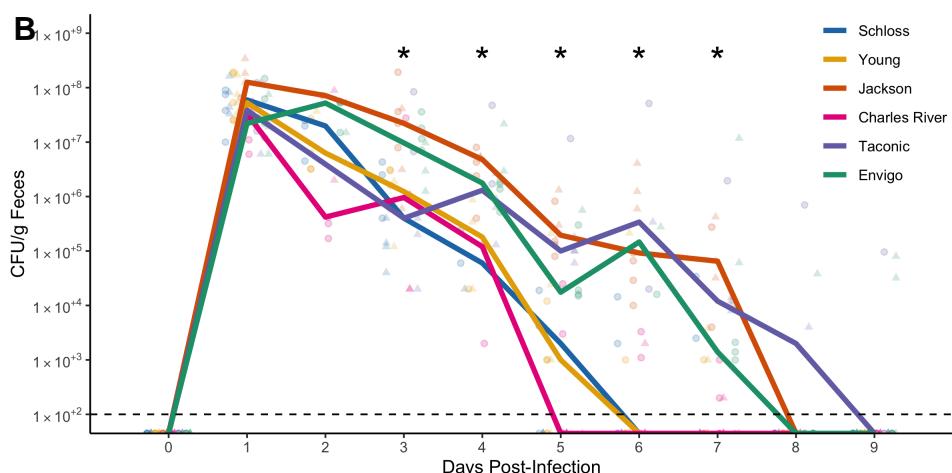
A

Sources of C57BL/6 mice:

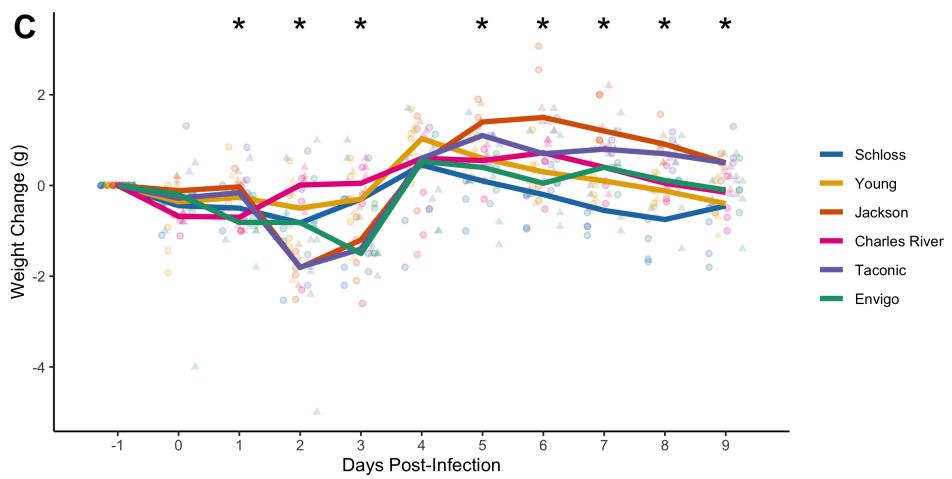
1. Schloss Lab Colony at University of Michigan
2. Young Lab Colony at University of Michigan
3. The Jackson Laboratory
4. Charles River Laboratories
5. Taconic Biosciences
6. Envigo



B



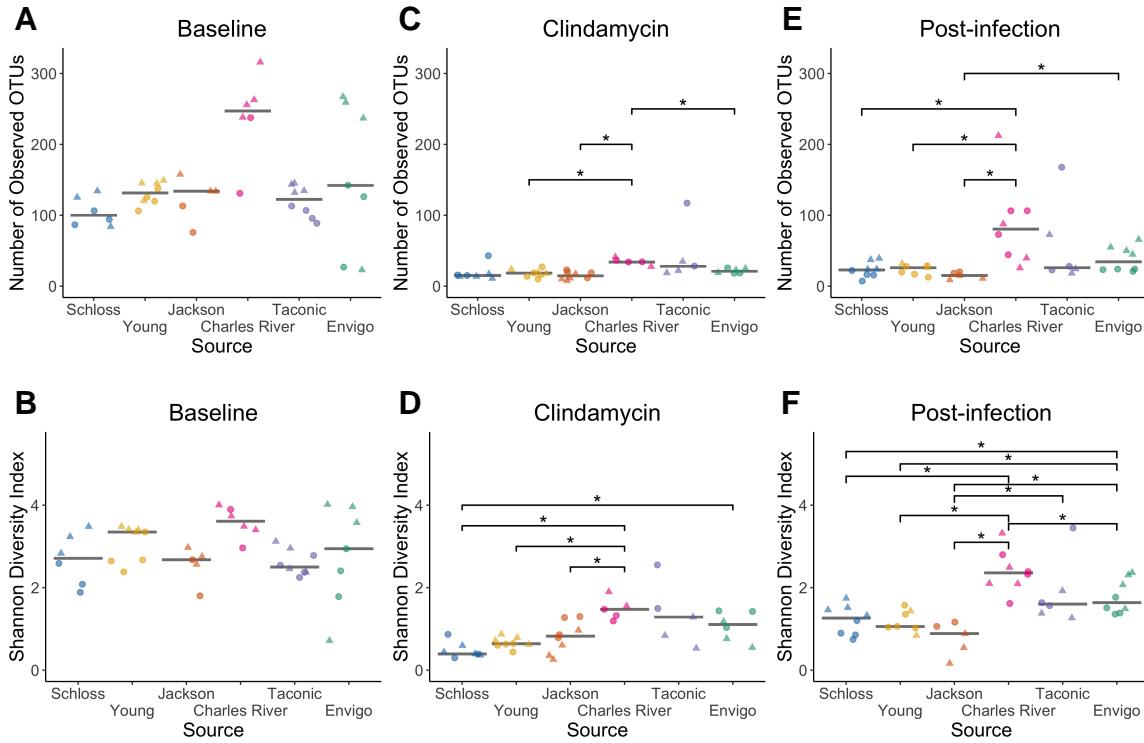
C



633 **Figure 1. Clindamycin**

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

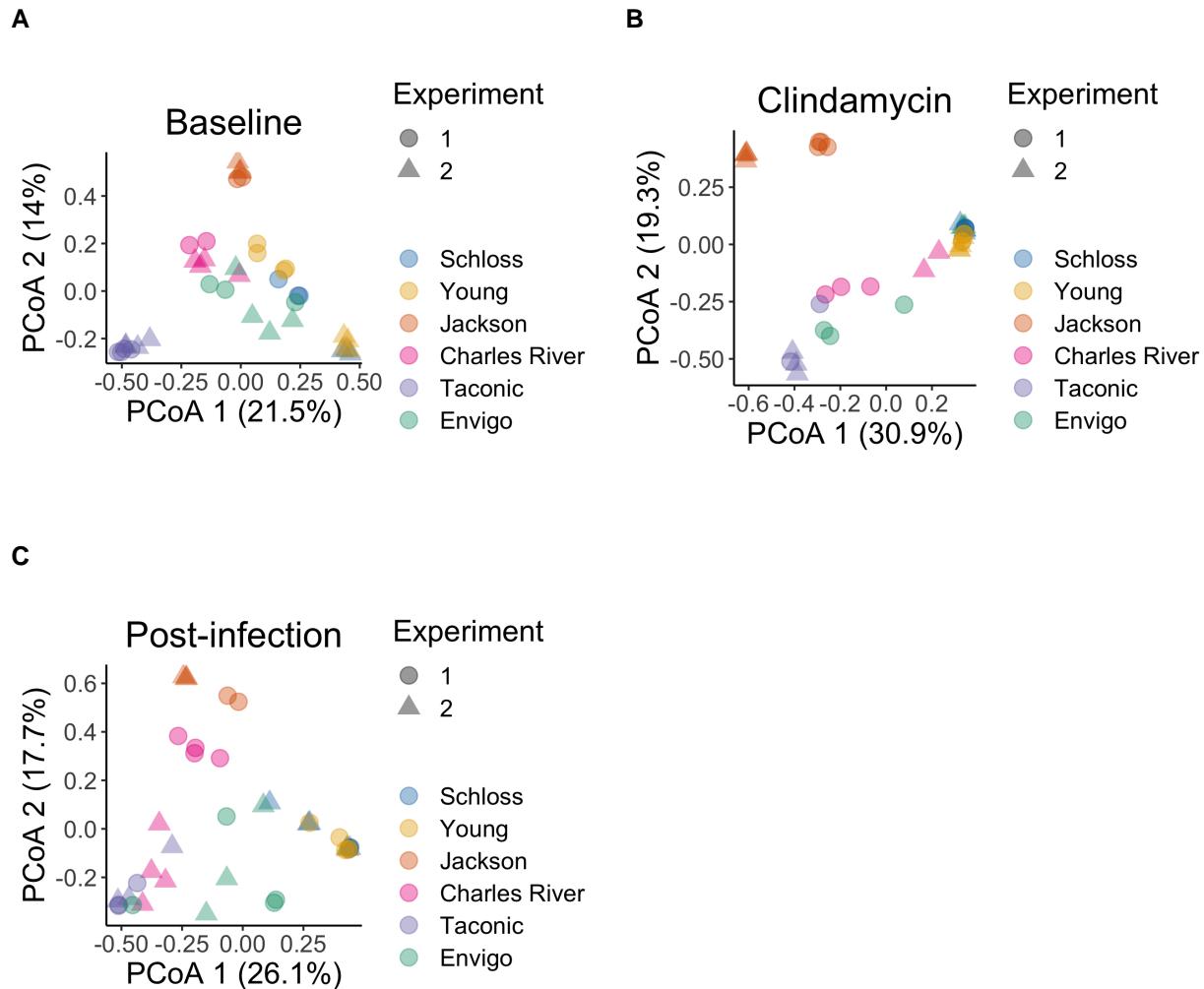
637 the University of Michigan, the Jackson Laboratory (N = 8), Charles River Laboratory (N = 8),
638 Taconic Biosciences (N = 8), and Envigo (N = 8). All mice were administered 10 mg/kg clindamycin
639 intraperitoneally (IP) 1 day before challenge with *C. difficile* 630 spores on day 0. Mice were
640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



Figure

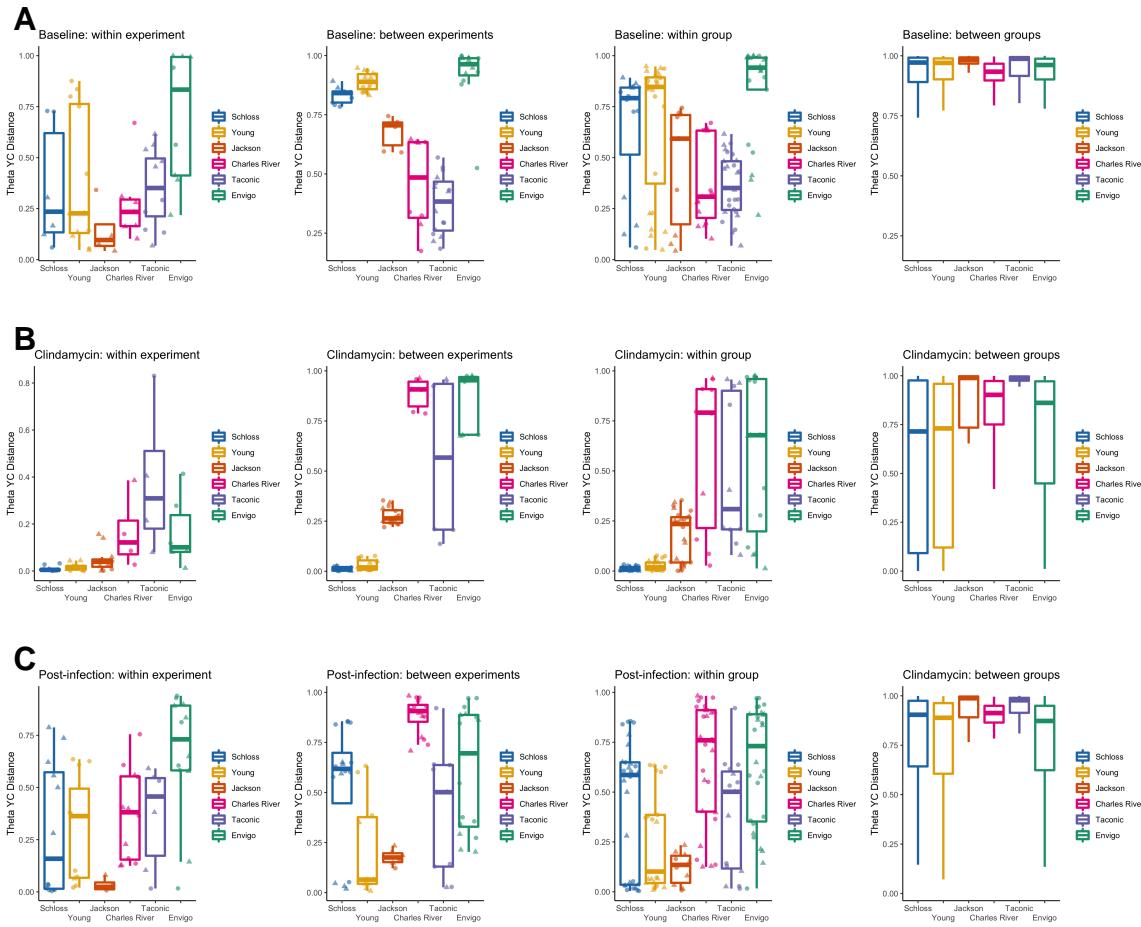
652

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



663

664 **Figure 3. Mouse source is the variable that explains most of the variation observed in**
 665 **the baseline, post-clindamycin, and post-infection bacterial communities.** A-C. Principal
 666 Coordinates Analysis of θ_{YC} distances from stools collected at baseline (A), post-clindamycin
 667 (B), and post-infection (C) timepoints of the experiment. Each symbol represents a stool sample
 668 from an individual mouse, with circles representing experiment 1 mice and triangles representing
 669 experiment 2 mice. PERMANOVA analysis demonstrated that source and the interaction between
 670 source and cage explained most of the variation observed in the baseline (combined $R^2 = 0.90$),
 671 post-clindamycin (combined $R^2 = 0.99$), and post-infection (combined $R^2 = 0.88$) communities (all
 672 $P = 0.0001$, see Table S6).

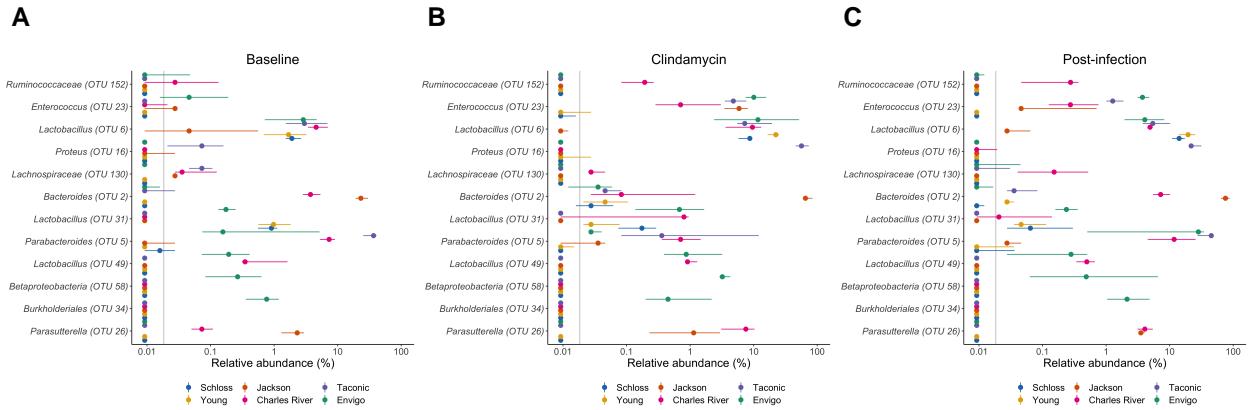


Figure

673

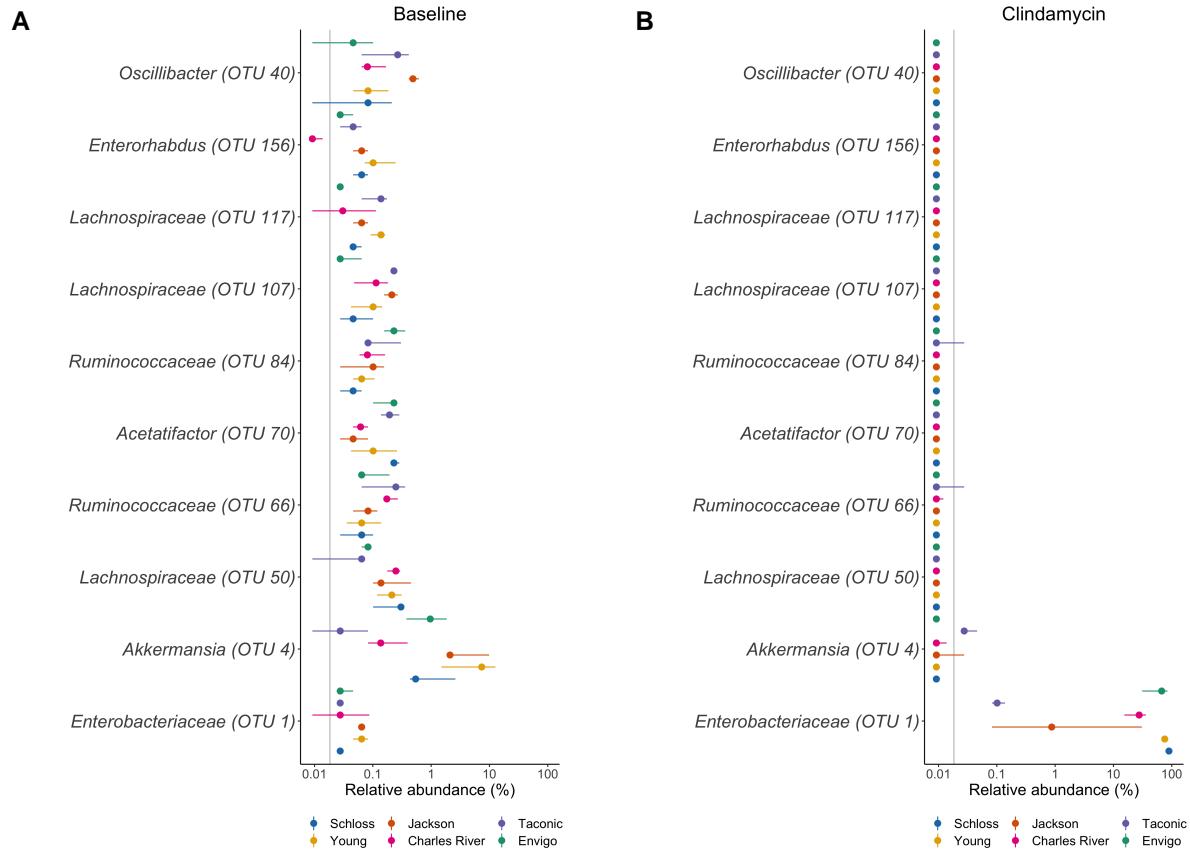
674 **4. High inter-group variation across mouse sources is diminished by clindamycin treatment**

675 A-C. Boxplots of the θ_{YC} distances of the 6 sources of mice relative to mice within the same
 676 source and experiment, mice within the same source and between experiments, mice within the
 677 same source, and mice from other groups at the baseline (A), after clindamycin treatment (B),
 678 and post-infection (C) timepoints. For comparisons within mice from the same source, symbols
 679 represent individual mouse samples: circles for experiment 1 and triangles for experiment 2.



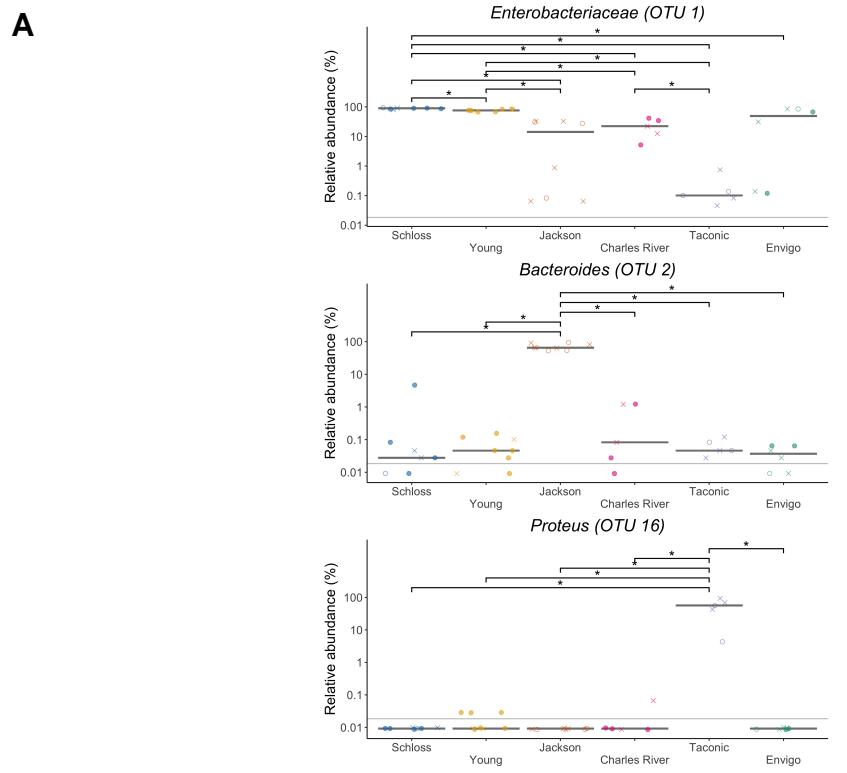
680

681 **Figure 5. A subset of bacteria consistently vary across sources despite clindamycin**
 682 **perturbation and *C. difficile* challenge.** A-C: plots highlighting the median (point) and
 683 interquantile range (colored lines) of the relative abundances for the 12 OTUs that consistently
 684 varied across sources of mice at the baseline (A), post-clindamycin (B), and post-infection (C)
 685 timepoints of the experiment. For each timepoint OTUs with differential relative abundances across
 686 sources of mice were identified by Kruskal-Wallis test with Benjamini-Hochberg correction for
 687 testing all identified OTUs (Table S8). The grey vertical line indicates the limit of detection.



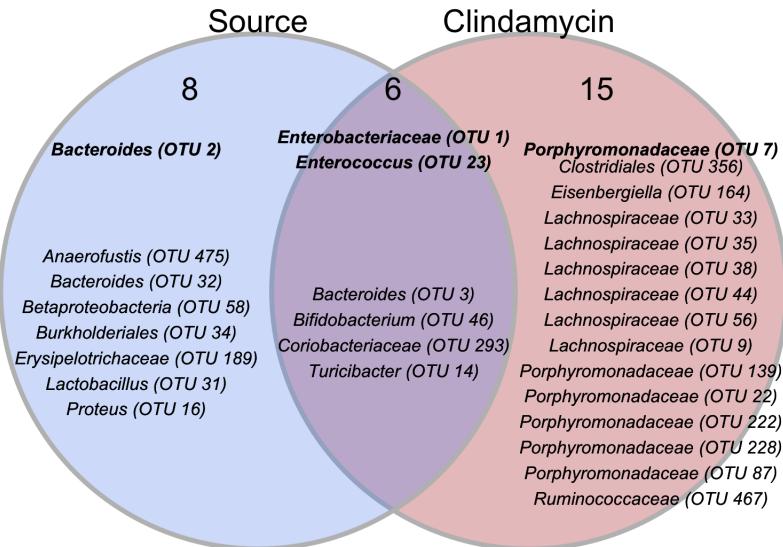
688

689 **Figure 6. Clindamycin treatment has the same effects on a subset of taxa regardless of**
 690 **source.** A-B: plots highlighting the median (point) and interquartile range (colored lines) of the
 691 top 10 most significant (adjusted P value < 0.05) OTUs with relative abundances that changed
 692 after clindamycin treatment. Data were analyzed by Wilcoxon signed rank test limited to mice that
 693 had paired sequence data for day -1 and 0 ($N = 31$). Tests were performed at the OTU level with
 694 Benjamini-Hochberg correction for testing all identified OTUs. See Table S9 for complete list of
 695 OTUs significantly impacted by clindamycin treatment. The grey vertical line indicates the limit of
 696 detection.



B

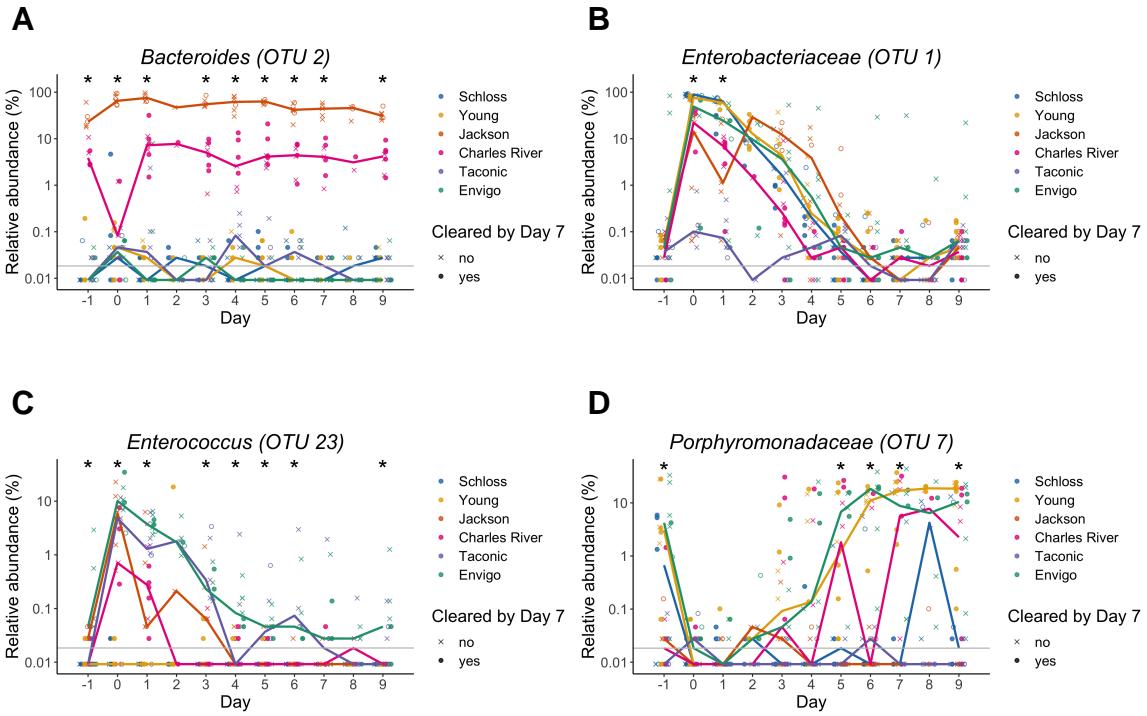
Key taxa comparisons for day -1, 0, and 1 models



Figure

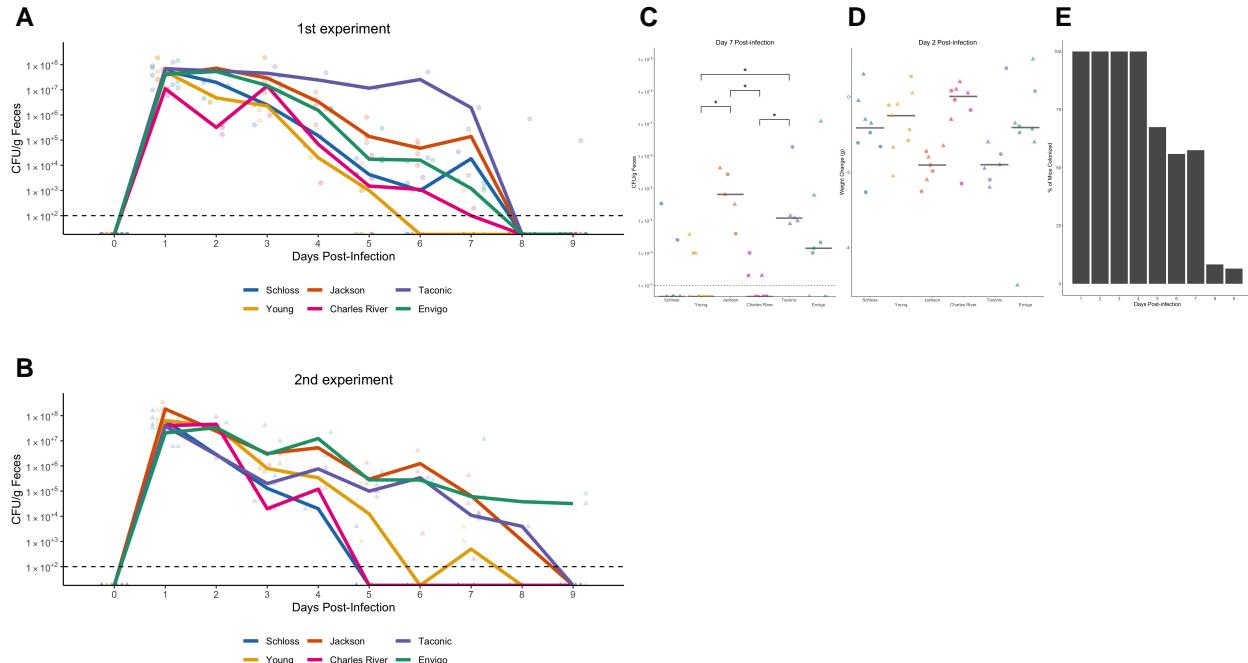
697
698 **7. Key OTUs that influence whether mice cleared *C. difficile* by day 7.** A. Baseline relative
699 abundance data for 3 of the OTUs from the classification model based on day 0 OTU relative
700 abundances that significantly varied across sources of mice and had high relative abundances in

701 the community. Symbols represent the relative abundance data for an individual mouse, circles
702 represent mice that cleared *C. difficile* by day 7, X-shapes represent mice that were still colonized
703 with *C. difficile*, and open circles represent mice that did not have *C. difficile* CFU counts for day 7
704 post-infection. Gray lines indicate the median relative abundances for each source. Asterisks are
705 shown for pairwise Wilcoxon comparisons with Benjamini-Hochberg correction where $P < 0.05$. B.
706 Venn diagram that combines Fig. S4 summaries of OTUs that were important to the day -1, 0, and
707 1 classification models (Table S14) and either overlapped with taxa that varied across vendors at
708 the same timepoint, were impacted by clindamycin treatment, or both. See Fig. S4 for separate
709 comparisons of taxa from the day -1, 0, and 1 classification models. Bold OTUs signify OTUs that
710 were important to more than 1 classification model.



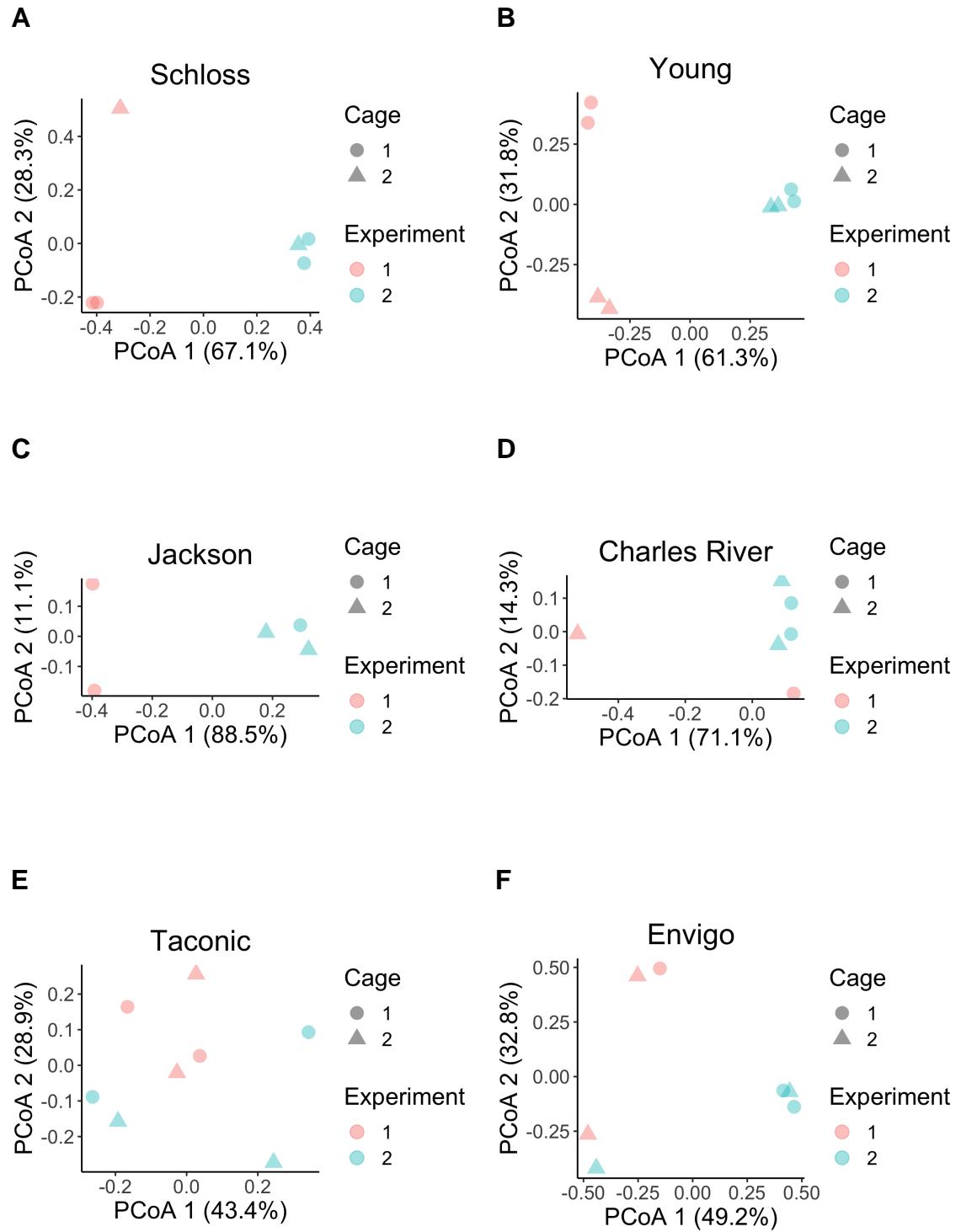
Figure

711
 712 **8: Key OTUs vary across sources throughout the experiment.** A-D. Relative abundances of
 713 bold OTUs from Fig. 7A that were important for at least two classification models are shown
 714 over time. A. *Bacteroides* (OTU 2), which varied across sources throughout the experiment. B-C.
 715 *Enterobacteriaceae* (B) and *Enterococcus* (C), which significantly varied across sources and were
 716 impacted by clindamycin treatment. D. *Porphyromonadaceae* (OTU 7), which was significantly
 717 impacted by clindamycin treatment and examining relative abundance dynamics over the course
 718 of the experiment, revealed timepoints where relative abundances also significantly varied across
 719 sources of mice. Symbols represent the relative abundance data for an individual mouse, circles
 720 represent mice that cleared *C. difficile* by day 7, X-shapes represent mice that were still colonized
 721 with *C. difficile*, and open circles represent mice that did not have *C. difficile* CFU counts for day 7
 722 post-infection. Colored lines indicate the median relative abundances for each source. The gray
 723 horizontal line represents the limit of detection. Timepoints where differences across sources of
 724 mice were statistically significant by Kruskal-Wallis test with Benjamini-Hochberg correction for
 725 testing across multiple days (Table S13) are identified by the asterisk(s) above each timepoint (*, P
 726 < 0.05).



727

728 **Figure S1. *C. difficile* CFU variation across vendors varies slightly across the 2**
 729 **experiments.** A-B. *C. difficile* CFU/gram of stool quantification over time for experiment 1
 730 (A) and 2 (B). Experiments were conducted approximately 3 months apart. Lines represent the
 731 median CFU for each source, symbols represent individual mice and the black line represents the
 732 limit of detection. C. *C. difficile* CFU/gram stool on day 7 post-infection across sources of mice
 733 with asterisks for pairwise Wilcoxon comparisons with Benjamini-Hochberg correction where $P <$
 734 0.05. D. Mouse weight change 2 days post-infection across sources of mice, no pairwise Wilcoxon
 735 comparisons were significant after Benjamini-Hochberg correction. For C-D: circles represent
 736 experiment 1 mice, triangles represent experiment 2 mice and gray lines indicate the median
 737 values for each group. E. Percent of mice that were colonized with *C. difficile* over the course of the
 738 experiment. Each day the percent is calculated based on the mice where *C. difficile* CFU was
 739 quantified for that particular day. Total N for each day: day 1 (N = 42), day 2 (N = 20), day 3 (N =
 740 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 =
 741 46).



Figure

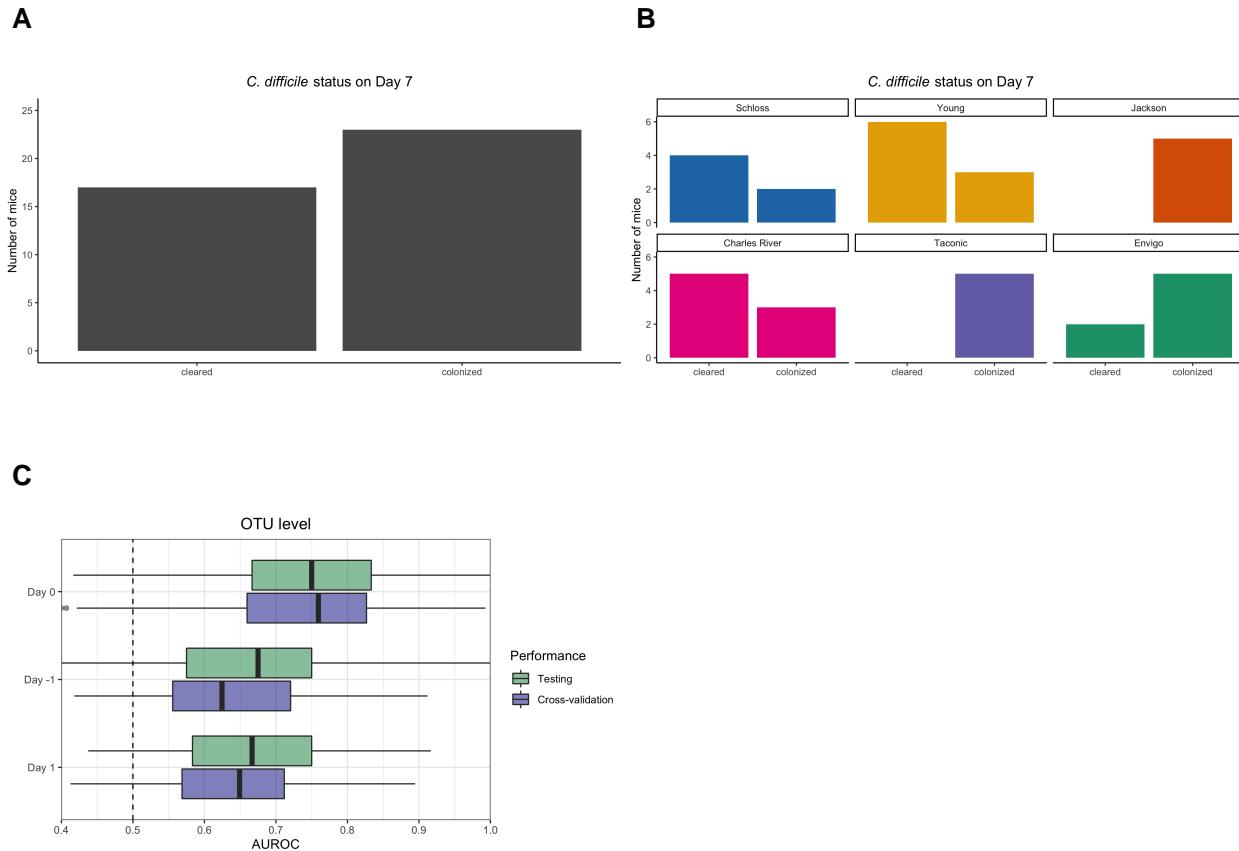
742

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

744 **experiments.** A-F. PCoA of θ_{YC} distances for the baseline fecal bacterial communities within each

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

⁷⁴⁶ corresponding to experiment and shape representing cage mates. PERMANOVA was performed
⁷⁴⁷ within each group to examine the contributions of experiment and cage to observed variation.
⁷⁴⁸ Experiment number and cage only significantly explained observed variation for mice from the
⁷⁴⁹ Schloss (combined $R^2 = 0.99$; $P \leq 0.033$) and Young (combined $R^2 = 0.95$; $P \leq 0.027$) lab colonies
⁷⁵⁰ (Table S7).

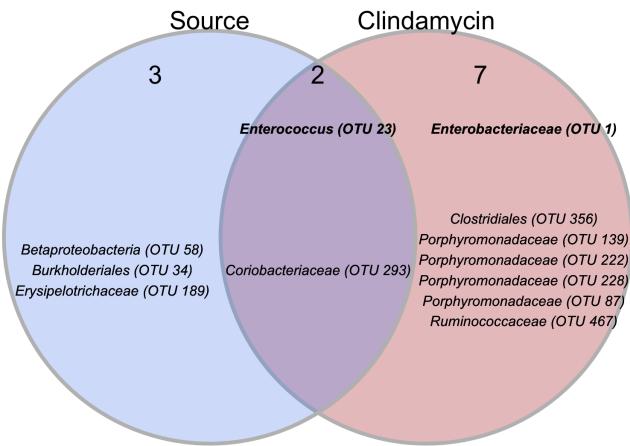


751

752 **Figure S3. Bacterial community composition before, after clindamycin perturbation, and**
 753 **post-infection can predict *C. difficile* colonization status 7 days post-challenge.** A. Bar
 754 graph visualizations of overall day 7 *C. difficile* colonization status that were used as classification
 755 outcomes to build L2-regularized logistic regression models. Mice were classified as colonized or
 756 cleared (not detectable at the limit of detection of 100 CFU) based on CFU g/stool data from 7 days
 757 post-infection. B. *C. difficile* CFU status on Day 7 within each mouse source. N = 5-9 mice per
 758 group. C. L2-regularized logistic regression classification model area under the receiving operator
 759 characteristic curve (AUROCs) to predict *C. difficile* CFU on day 7 post-infectoin (Fig. 1D, Fig. S3)
 760 based on the OTU community relative abundances at baseline (day -1), post-clindamycin (day 0),
 761 and post-infection (day 1). All models performed better than random chance (AUROC = 0.5; all P
 762 $\leq 5.2\text{e-}31$; Table S12) and the model built with post-clindamycin treated bacterial OTU relative
 763 abundances had the best performance ($(P_{\text{FDR}} \leq 3.1\text{e-}11$ for all pairwise comparisons; Table S11).
 764 For list of the 20 OTUs that were ranked as most important to each model, see Table S12.

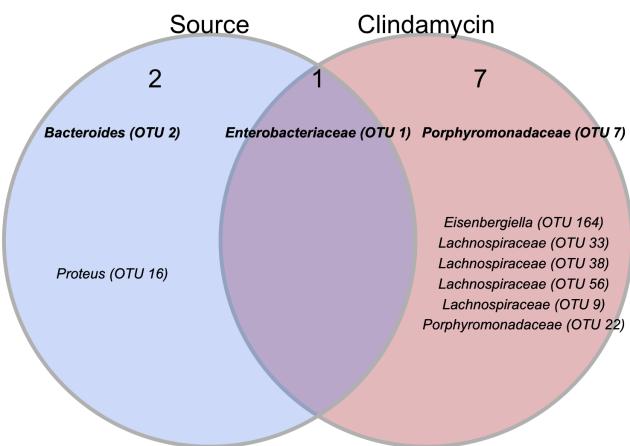
A

Day -1 model key OTU comparisons



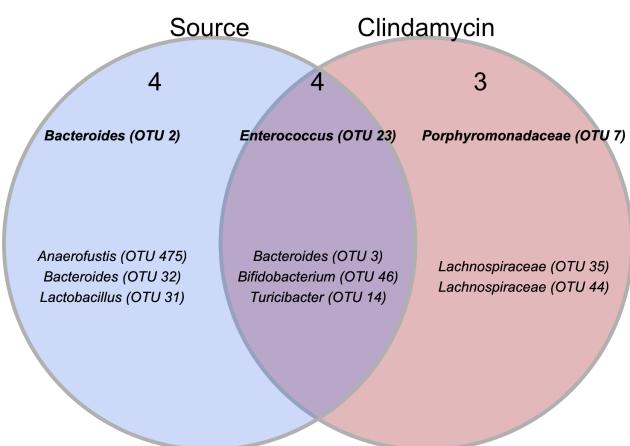
B

Day 0 model key OTU comparisons



C

Day 1 model key OTU comparisons



Figure

765

766 **S4. Key OTUs from classification models based on baseline, post-clindamycin treatment,**
767 **or post-infection community data vary by source, clindamycin treatment, or both.** A-C.
768 Venn diagrams of top 20 important OTUs from baseline (A), post-clindamycin treatment (B), and

⁷⁶⁹ post-infection (C) classification models (Table S12) that overlapped with OTUs that varied across
⁷⁷⁰ vendors at the corresponding timepoint, were impacted by clindamycin treatment, or both. **Bold**
⁷⁷¹ OTUs signify OTUs that were important to more than 1 classification model.

772 **Supplementary Tables and Movie**

773 All supplemental material is available at: https://github.com/SchlossLab/Tomkovich_Vendor_XXXX_
774 2020/submission.

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

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

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

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

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

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

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

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

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

793 **Table S9. OTUs with relative abudances that significantly changed after clindamycin**
794 **treatment.**

795 **Table S10. Statistical results of L2-regularized logistic regression model performances**

796 compared to random chance.

797 **Table S11.** Pairwise Wilcoxon results comparing all 3 L2-regularized logistic regression
798 model performances.

799 **Table S12.** Top 20 most important OTUs for each of the 3 L2-regularized logistic regression
800 models based on OTU relative abundance data.

801 **Table S13.** OTUs with relative abundances that significantly varied across sources of mice
802 on at least 1 day of the experiment by Kruskal-Wallis test.