

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

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

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

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

1 Abstract

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

22 Importance

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

29 colony sources and challenged them with *C. difficile*. Interestingly, only a subset of the taxa that
30 vary across sources were associated with how long *C. difficile* was able to colonize. Future studies
31 examining the interplay between the microbiota and *C. difficile* should consider using mice from
32 multiple sources to narrow down the microbes driving the observed phenotypes and reflect human
33 interindividual variation.

34 **Introduction**

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

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

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

60 **Results**

61 **Clindamycin treatment renders all mice susceptible to *C. difficile* 630 colonization**
62 **regardless of colony source.** To test how the microbiotas of mice from different colony sources
63 impact colonization dynamics after clindamycin exposure, we utilized C57BL/6 mice from 6 different
64 sources: two colonies from the University of Michigan (the Young and Schloss lab colonies), the
65 Jackson Laboratory, Charles River Laboratories, Taconic Biosciences, and Envigo (which was
66 formerly Harlan). These 4 vendors were chosen because they represent commonly used vendors
67 for CDI studies in mice (26, 34–40). After a 13-day acclimation period for the mice ordered from
68 vendors, all mice were treated with 10 mg/kg clindamycin via intraperitoneal injection and one day
69 later challenged with 10^3 *C. difficile* 630 spores (Fig. 1A). Clindamycin was chosen because we
70 have previously demonstrated mice are rendered susceptible, but consistently cleared the CDI
71 within 9 days (21, 41), clindamycin is frequently implicated with human CDIs (42), and is also
72 part of the antibiotic treatment for the frequently cited 2008 CDI mouse model (14). The day after
73 infection, *C. difficile* was detectable in all mice at a similar level (Mean CFU range: 3e+07-1.5e+08;
74 $P_{FDR} = 0.15$), indicating clindamycin rendered all mice susceptible regardless of colony source (Fig.
75 1B). Interestingly, variation in *C. difficile* CFU levels across sources of mice emerged from days
76 3-7 post-infection (all $P_{FDR} \leq 0.019$; Fig. 1B and Table S1), suggesting mouse colony source is
77 associated with *C. difficile* clearance. We conducted two experiments approximately 3 months
78 apart and combined the data because the colonization dynamics across sources of mice varied
79 similarly in both experiments (Fig. S1). Although *C. difficile* 630 causes mild symptoms in mice
80 compared to other *C. difficile* strains (43), we also saw that weight change significantly varied
81 across sources of mice with the most weight lost two days post-infection (Fig. 1C and Table S2).
82 Interestingly, mice ordered from Jackson, Taconic, and Envigo tended to lose the most weight,
83 have higher *C. difficile* CFU levels and take longer to clear the infection compared to the other
84 sources of mice, which was particularly evident 7 days post-infection (Fig. 1C and D). Importantly,
85 there was also one Jackson and one Envigo mouse that died between 1- and 3-days post-infection.
86 Thus, clindamycin rendered all mice susceptible to *C. difficile* 630 colonization, regardless of colony
87 source, but variation across sources emerged with 3 out of 6 sources taking longer to clear *C.*
88 *difficile*.

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

105 Next, we compared the bacterial communities from the 6 colonies over the course of the experiment
106 using principal coordinate analysis (PCoA) of the Theta YC distances. Permutational multivariate
107 analysis of variance (PERMANOVA) analysis revealed colony source was the major factor explaining
108 the observed variation across fecal communities ($R^2 = 0.35$, $P = 0.0001$) followed by interactions
109 between cage and day of the experiment (Movie S1 and Table S5). Since, the majority of the
110 perturbations happened over the initial days of the experiment, we decided to focus on the bacterial
111 communities at baseline (day -1), after clindamycin treatment (day 0), and post-infection (day 1). For
112 all 3 timepoints, source and the interaction with cage significantly explained most of the observed
113 community variation (combined $R^2 = 0.90, 0.99, 0.88$, respectively; $P = 0.0001$; Fig. 3 and Table
114 S6). We also compared baseline communities across the 2 experiments, and found experiment
115 and cage significantly explained the observed variation only for the Schloss and Young lab mouse
116 colonies (Fig. S2 and Table S7), suggesting vendor communities were relatively stable between
117 the 2 experiments. Thus, mouse colony source was the factor that explained the most variation

118 observed in the bacterial communities. Importantly with the exception of the 2 University colonies,
119 the community of each source clustered apart from one another suggesting each community had a
120 unique response to clindamycin treatment and *C. difficile* challenge.

121 After finding differences at the community level, we next identified the taxa that varied across
122 sources of mice over the initial days of the experiment. We examined bacterial relative abundances
123 at the operational taxonomic unit (OTU) and family levels, expecting the number of differences to be
124 reduced at the family level due to the nature of bacterial taxonomy (45). Focusing on the baseline
125 communities first, there were 268 OTUs and only 20 families (Table S8-9) with relative abundances
126 that varied across colony sources. Clindamycin treatment reduced the number of taxa with relative
127 abundances that varied across sources to 18 OTUs and 10 families (Table S8-9). After *C. difficile*
128 challenge, there were 44 OTUs and 18 families (Table S8-9) with significantly different relative
129 abundances across sources, as the communities started to recover from antibiotic treatment. In
130 spite of the experimental perturbations that occurred during these 3 timepoints, there were 12
131 OTUs (Fig 4A-C) and 8 families with relative abundances that consistently varied across colony
132 sources (Fig. 4D-F). Importantly, some of the taxa that consistently varied across sources also
133 shifted with clindamycin treatment. For example, *Proteus* increased after clindamycin treatment, but
134 only in Taconic mice. *Enterococcus* was primarily found only in mice purchased from commercial
135 vendors and also increased after clindamycin treatment. In summary, mouse bacterial communities
136 significantly varied according to colony source throughout the course of the experiment and a
137 consistent subset of bacterial taxa remained different across sources regardless of clindamycin
138 and *C. difficile* challenge.

139 **Clindamycin treatment alters a subset of taxa that were found in all colony sources.**
140 Although there were bacteria that consistently varied across colony sources, we also wanted to
141 identify the bacteria that shifted after clindamycin treatment, regardless of colony source. By
142 analyzing all mice that had sequence data from fecal samples collected at baseline and after
143 clindamycin treatment, we identified 153 OTUs and 18 families that were altered after clindamycin
144 treatment (Fig. 5 and Table S10-11). Interestingly, when we compared the list of significant
145 clindamycin impacted bacteria with the bacteria that consistently varied across groups over the
146 initial 3 timepoints of our experiment, we found 3 OTUs (*Lachnospiraceae* (OTU 130), *Lactobacillus*

¹⁴⁷ (OTU 6), *Enterococcus* (OTU 23)) and 3 families (*Porphyromonadaceae*, *Enterococcaceae*,
¹⁴⁸ *Lachnospiraceae*) overlapped (Fig. 4, Fig. 5C-D). These findings demonstrate that clindamycin has
¹⁴⁹ a consistent impact on the fecal bacterial communities of mice from all colony sources and only a
¹⁵⁰ subset of the taxa also varied across colony sources.

¹⁵¹ **Source-specific and clindamycin impacted bacteria distinguish *C. difficile* colonization**
¹⁵² **status in mice.** After identifying taxa that varied by colony source, changed after clindamycin
¹⁵³ treatment, or both, we next wanted to determine which taxa were influencing the variation in
¹⁵⁴ *C. difficile* colonization status at day 7 (Fig. 1D). We trained L2-regularized logistic regression
¹⁵⁵ models with input bacterial community data from the baseline, post-clindamycin, and post-infection
¹⁵⁶ timepoints of the experiment to predict *C. difficile* colonization status on day 7 (Fig. S3). All models
¹⁵⁷ were better at predicting *C. difficile* colonization status on day 7 than random chance (all $P \leq 5e-15$;
¹⁵⁸ Table S12), however the models trained with OTU level data generally performed better than those
¹⁵⁹ trained with family level data with the exception of the models based on the post-infection (day 1)
¹⁶⁰ communities (Fig. 6). Interestingly, the model based on the post-clindamycin (day 0) community
¹⁶¹ OTU data performed significantly better than all other models with an AUROC of 0.75 ($P_{FDR} \leq$
¹⁶² 3.9e-10 for pairwise comparisons; Table S13). Thus, we were able to use community bacterial
¹⁶³ relative abundance data alone to differentiate mice that were still colonized with *C. difficile* on
¹⁶⁴ day 7 from those mice that had cleared the *C. difficile*. Interestingly, the model built with OTU
¹⁶⁵ relative abundance data post-clindamycin treatment had the best performance, suggesting how the
¹⁶⁶ bacterial community responds to clindamycin treatment has the greatest influence on subsequent
¹⁶⁷ *C. difficile* colonization dynamics.

¹⁶⁸ Next, to examine the bacteria that were driving each model's performance, we pulled out the top
¹⁶⁹ 20 taxa that had the highest absolute feature weights in each of the 6 models (Table S14-15).
¹⁷⁰ The top 20 taxa from each model were compared to the list of taxa that varied across colony
¹⁷¹ source (Fig. 4 and Table S8-9) at the same timepoint and the taxa that were altered by clindamycin
¹⁷² treatment (Fig. 5 and Table S10-11). We found a subset of OTUs and families that were important
¹⁷³ to the model and overlapped with bacteria that varied by either source, clindamycin treatment,
¹⁷⁴ or both (Fig. S4, S5, S6). Combining the overall results for the 3 OTU models identified 14
¹⁷⁵ OTUs associated with source, 21 OTUs associated with clindamycin treatment, and 6 OTUs

176 associated with both (Fig. 7A). Combining the overall results for the 3 family models identified
177 18 families associated with source, 14 families associated with clindamycin treatment and 8
178 families associated with both (Fig. 7B). Several OTUs (*Bacteroides* (OTU 2), *Enterococcus* (OTU
179 23), *Enterobacteriaceae* (OTU 1), *Porphyromonadaceae* (OTU 7)) and families (*Bacteroidaceae*,
180 *Deferrribacteraceae*, *Enteroccaceae*, *Lachnospiraceae*, *Bifidobacteriaceae*, *Coriobacteriaceae*,
181 *Ruminococcaceae*, *Verrucomicrobiaceae*) appeared across at least 2 models, so we examined
182 how the relative abundances of these key taxa varied over the course of the experiment (Fig. 8 and
183 Fig. S7). Throughout the experiment, there was at least 1 timepoint where relative abundances
184 of these taxa significantly varied across sources (Table S16-17). Interestingly, there were no taxa
185 that emerged as consistently enriched or depleted in Jackson, Taconic, and Envigo mice that were
186 colonized for longer with *C. difficile* 630, suggesting multiple bacteria influence the time needed
187 to clear the infection. Together, these results suggest the initial bacterial communities and their
188 responses to clindamycin have a large influence on the time needed to clear *C. difficile*.

189 **Discussion**

190 By examining the *C. difficile* colonization dynamics within mice from 6 different colony sources
191 after perturbing the microbiota with clindamycin treatment, we were able to identify bacterial taxa
192 that were unique to sources throughout the experiment as well as taxa that were universally
193 impacted by clindamycin. We built L2 logistic regression models with baseline, post-clindamycin
194 treatment, and post-infection fecal community data that successfully predicted *C. difficile*
195 colonization status 7 days after infection better than random chance. We identified *Bacteroides*
196 (OTU 2), *Enterococcus* (OTU 23), *Enterobacteriaceae* (OTU 1), *Porphyromonadaceae* (OTU 7),
197 *Bacteroidaceae*, *Deferrribacteraceae*, *Enteroccaceae*, *Lachnospiraceae*, *Bifidobacteriaceae*,
198 *Coriobacteriaceae*, *Ruminococcaceae*, *Verrucomicrobiaceae* (Fig. 8, Fig. S7) as candidate bacteria
199 within these communities that were influencing variation in *C. difficile* colonization dynamics since
200 these bacteria were all important in the logistic regression models and varied by colony source,
201 were impacted by clindamycin treatment, or both. Overall, our results demonstrate clindamycin
202 is sufficient to render mice from multiple sources susceptible to CDI and only a subset of the
203 interindividual microbiota variation across mice from different sources was associated with the time

204 needed to clear *C. difficile*.
205 Other groups have taken similar approaches by using mice from multiple colony sources to identify
206 bacteria that either promote colonization resistance or increase susceptibility to enteric infections
207 (22, 23, 26–30). For example, in the context of *Salmonella* infections, *Enterobacteriaceae* and
208 segmented filamentous bacteria have emerged as protective (22, 27). A previous study with *C.*
209 *difficile* identified an endogenous protective *C. difficile* strain LEM1 that bloomed after antibiotic
210 treatment in mice from Jackson or Charles River Laboratories, but not Taconic that protected mice
211 against the more toxigenic *C. difficile* VPI10463 (26). Given that we ordered mice from the same
212 vendors, we checked all mice for endogenous *C. difficile* by plating stool samples that were collected
213 after clindamycin treatment. However, we did not identify any endogenous *C. difficile* strains prior
214 to challenge, suggesting there were no endogenous protective strains in the mice we received
215 and other bacterial taxa were mediating the variation in *C. difficile* colonization across sources.
216 Although all mice were susceptible to *C. difficile* colonization, by following colonization over time we
217 found Jackson, Taconic, and Envigo mice remained colonized for longer. We identified a subset of
218 bacteria that were important in predicting whether a mouse was still colonized with *C. difficile* status
219 7 days post-infection. These results suggest a subset of the bacterial community is responsible for
220 determining the length of time needed to clear *C. difficile* colonization.

221 In the past variation between different CDI mouse model studies have been attributed to intestinal
222 microbiome differences in mice across different institutional environments. For example, groups
223 using the same clindamycin treatment and C57BL/6 mice had different *C. difficile* outcomes, one
224 having sustained colonization (32), while the other had transient (18). Baseline differences in the
225 microbiota composition have been hypothesized to partially explain the differences in colonization
226 outcomes and overall susceptibility to *C. difficile* after treatment with the same antibiotic (13,
227 31). We have shown that mice from 6 different sources were all susceptible to *C. difficile* 630,
228 suggesting the microbiota influences *C. difficile* clearance more than susceptibility. Fortunately,
229 the bacterial perturbations induced by clindamycin treatment have been well characterized
230 and our findings agree with previous CDI mouse model work demonstrating *Enterococcus* and
231 *Enterobacteriaceae* were associated with *C. difficile* susceptibility and *Porphyromonadaceae*,
232 *Lachnospiraceae*, *Ruminococcaceae*, and *Turicibacter* were associated with resistance (19, 21, 32,

33, 41, 46–48). While we have demonstrated that susceptibility is uniform across vendors after
233 clindamycin treatment, there could be different outcomes for either susceptibility or clearance in the
234 case of other antibiotic treatments. The *C. difficile* strain used could also be contributing to the
235 variation in *C. difficile* outcomes seen across different groups (47). We found the time needed
236 to naturally clear *C. difficile* varied across sources of mice implying that at least in the context
237 of the same perturbation, microbiota differences seemed to influence infection outcome more
238 than susceptibility. More importantly, we were able to narrow down from all the variation observed
239 across colony sources to a subset of bacterial taxa that were also important for predicting *C. difficile*
240 colonization status 7 days post-infection.
241

Our approach successfully increased the diversity of murine bacterial communities tested in
242 our clindamycin *C. difficile* model. One alternative approach that has been used in some CDI
243 studies (49–54) is to use mice associated with human microbiotas. However, a major caveat
244 to this method is the substantial loss of human microbiota community members upon transfer
245 to mice (55, 56). Additionally with the exception of 2 recent studies (49, 50), most of the CDI
246 mouse model studies to date associated mice with just 1 types of human microbiota either from
247 a single donor or a single pool from multiple donors (51–54), which does not aid in the goal
248 of figuring out how a variety of unique microbiotas influence susceptibility to CDIs and adverse
249 outcomes. Encouragingly, decreased *Bifidobacterium*, *Porphyromonas*, *Ruminococcaceae* and
250 *Lachnospiraceae* and increased *Enterobacteriaceae*, *Enterococcus*, *Lactobacillus*, and *Proteus*
251 have all been associated with human CDIs (7) and were well represented in our study, suggesting
252 most of the vendors are suitable for gaining insights into microbiota associated factors influencing
253 *C. difficile* colonization and infections in humans. An important exception was *Enterococcus*, which
254 was primarily absent from the mice from University of Michigan colonies and *Proteus*, which was
255 only found in Taconic mice. Importantly, the fact that some CDI associated bacteria were only found
256 in a subset of mice has important implications for future CDI mouse model studies.
257

There are several limitations to our work. The microbiome is composed of viruses, fungi, and
258 parasites in addition to bacteria, and these non-bacterial members can also vary across mouse
259 vendors (57, 58). While our study focused solely on the bacterial portion, viruses and fungi have
260 also begun to be implicated in the context of CDIs or FMT treatments for recurrent CDIs (35, 59–62).
261

262 Beyond community composition, the metabolic function of the microbiota also has a CDI signature
263 (20, 48, 63, 64) and can vary across mice from different sources (65). For example, microbial
264 metabolites, particularly secondary bile acids and butyrate production, have been implicated as
265 important contributors to *C. difficile* resistance (33, 47). Although, we only looked at composition,
266 *Ruminococcaceae* and *Lachnospiraceae* both emerged as important taxa for classifying day 7 *C.*
267 *difficile* colonization status and metagenomes from these bacteria have been shown to contain
268 the bile acid-inducible gene cluster necessary for secondary bile acid formation and ability to
269 produce butyrate (52, 66). Interestingly, butyrate has previously been shown to vary across vendors
270 and mediates resistance to *Citrobacter rodentium* infection, a model of enterohemorrhagic and
271 enteropathogenic *Escherichia coli* infections (23). Evidence for immunological toning differences in
272 IgA and Th17 cells across mice from different vendors have also been documented and (67, 68),
273 may also influence response to CDI, particularly in the context of severe CDIs (69, 70). The outcome
274 after *C. difficile* exposure depends on a multitude of factors, including age, diet, and immunity; all of
275 which are also influenced by the microbiota. We have demonstrated that the ways different baseline
276 microbiotas from different mouse colony sources respond to clindamycin treatment influences the
277 length of time mice remained colonized with *C. difficile* 630. For those interested in dissecting the
278 contribution of the microbiome to *C. difficile* pathogenesis and treatments, using multiple sources
279 of mice may yield more insights than a single model alone. Furthermore, for studies wanting to
280 examine the interplay between a particular bacterial taxon such as *Enterococcus* and *C. difficile*,
281 these results could serve as a resource for selecting which mice to order to address the question.

282 **Acknowledgements**

283 This work was supported by the National Institutes of Health (U01AI124255). ST was supported by
284 the Michigan Institute for Clinical and Health Research Postdoctoral Translation Scholars Program
285 (UL1TR002240 and KL2TR002241). We thank members of the Schloss lab for feedback on
286 planning the experiments and data presentation, as well as code tutorials and feedback through
287 Code club. In particular, we want to thank Begum Topçuoğlu for help with implementing logistic
288 regression models using her pipeline and Ana Taylor for help with some media preparation and
289 sample collection. We also thank members of Vincent Young's lab, particularly Kimberly Vendrov,
290 for guidance with the *C. difficile* infection mouse model and donating the mice. We also want to
291 thank the Unit for Laboratory Animal Medicine at the University of Michigan for maintaining our
292 mouse colony and providing the institutional support for our mouse experiments. Finally, we thank
293 Kwi Kim, Austin Campbell, and Kimberly Vendrov for their help in maintaining the Schloss lab's
294 anaerobic chamber.

295 **Materials and Methods**

296 **(i) Animals.** All experiments were approved by the University of Michigan Animal Care and Use
297 Committee (IACUC) under protocol number PRO00006983. Female C57BL/7 mice were obtained
298 from 6 different colony sources: The Jackson Laboratory, Charles River Laboratories, Taconic
299 Biosciences, Envigo, and two colonies at the University of Michigan (the Schloss lab colony and
300 the Young lab colony). The Young lab colony was originally established with mice purchased from
301 Jackson, and the Schloss lab colony was later founded with mice donated from the Young lab. The
302 4 groups of mice purchased from vendors were allowed to acclimate to the University of Michigan
303 mouse facility for 13 days prior to starting the experiment. At least 4 female mice (age 5-10 weeks)
304 were obtained per source and mice from the same source were primarily housed at a density of 2
305 mice per cage. The experiment was repeated once, approximately 3 months after the start of the
306 first experiment.

307 **(ii) Antibiotic treatment.** After the 13-day acclimation period and 1 day prior to challenge (Fig.
308 1A), all mice received 10 mg/kg clindamycin (filter sterilized through a 0.22 micron syringe filter
309 prior to administration) via intraperitoneal injection.

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

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

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

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

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

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

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

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

375 **References**

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

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

- 425 nutrient niches during infection across distinct murine gut microbiomes. *mSystems* 2.
- 426 21. Jenior ML, Leslie JL, Young VB, Schloss PD. 2018. *Clostridium difficile* alters the structure and
427 metabolism of distinct cecal microbiomes during initial infection to promote sustained colonization.
428 *mSphere* 3.
- 429 22. Velazquez EM, Nguyen H, Heasley KT, Saechao CH, Gil LM, Rogers AWL, Miller BM, Rolston
430 MR, Lopez CA, Litvak Y, Liou MJ, Faber F, Bronner DN, Tiffany CR, Byndloss MX, Byndloss AJ,
431 Bäumler AJ. 2019. Endogenous enterobacteriaceae underlie variation in susceptibility to salmonella
432 infection. *Nature Microbiology* 4:1057–1064.
- 433 23. Osbelt L, Thiemann S, Smit N, Lesker TR, Schröter M, Gálvez EJC, Schmidt-Hohagen K, Pils
434 MC, Mühlen S, Dersch P, Hiller K, Schlüter D, Neumann-Schaal M, Strowig T. 2020. Variations
435 in microbiota composition of laboratory mice influence citrobacter rodentium infection via variable
436 short-chain fatty acid production. *PLOS Pathogens* 16:e1008448.
- 437 24. Stough JMA, Dearth SP, Denny JE, LeCleir GR, Schmidt NW, Campagna SR, Wilhelm SW.
438 2016. Functional characteristics of the gut microbiome in c57bl/6 mice differentially susceptible to
439 plasmodium yoelii. *Frontiers in Microbiology* 7.
- 440 25. Alegre M-L. 2019. Mouse microbiomes: Overlooked culprits of experimental variability. *Genome
441 Biology* 20.
- 442 26. Etienne-Mesmin L, Chassaing B, Adekunle O, Mattei LM, Bushman FD, Gewirtz AT. 2017.
443 Toxin-positive *clostridium difficile* latently infect mouse colonies and protect against highly
444 pathogenic *c. difficile*. *Gut* 67:860–871.
- 445 27. Lai NY, Musser MA, Pinho-Ribeiro FA, Baral P, Jacobson A, Ma P, Potts DE, Chen Z, Paik D,
446 Soualhi S, Yan Y, Misra A, Goldstein K, Lagomarsino VN, Nordstrom A, Sivanathan KN, Wallrapp A,
447 Kuchroo VK, Nowarski R, Starnbach MN, Shi H, Surana NK, An D, Wu C, Huh JR, Rao M, Chiu IM.
448 2020. Gut-innervating nociceptor neurons regulate peyer's patch microfold cells and SFB levels to
449 mediate salmonella host defense. *Cell* 180:33–49.e22.
- 450 28. Thiemann S, Smit N, Roy U, Lesker TR, Gálvez EJ, Helmecke J, Basic M, Bleich A, Goodman

- 451 AL, Kalinke U, Flavell RA, Erhardt M, Strowig T. 2017. Enhancement of IFNgamma production by
452 distinct commensals ameliorates salmonella-induced disease. *Cell Host & Microbe* 21:682–694.e5.
- 453 29. Rolig AS, Cech C, Ahler E, Carter JE, Ottemann KM. 2013. The degree of helicobacter
454 pylori-triggered inflammation is manipulated by preinfection host microbiota. *Infection and Immunity*
455 81:1382–1389.
- 456 30. Ge Z, Sheh A, Feng Y, Muthupalani S, Ge L, Wang C, Kurnick S, Mannion A, Whary MT, Fox
457 JG. 2018. Helicobacter pylori-infected c57bl/6 mice with different gastrointestinal microbiota have
458 contrasting gastric pathology, microbial and host immune responses. *Scientific Reports* 8.
- 459 31. Lawley TD, Young VB. 2013. Murine models to study clostridium difficile infection and
460 transmission. *Anaerobe* 24:94–97.
- 461 32. Buffie CG, Jarchum I, Equinda M, Lipuma L, Gobourne A, Viale A, Ubeda C, Xavier J, Pamer
462 EG. 2011. Profound alterations of intestinal microbiota following a single dose of clindamycin results
463 in sustained susceptibility to clostridium difficile-induced colitis. *Infection and Immunity* 80:62–73.
- 464 33. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew
465 M, Viale A, Littmann E, Brink MRM van den, Jenq RR, Taur Y, Sander C, Cross JR, Toussaint
466 NC, Xavier JB, Pamer EG. 2014. Precision microbiome reconstitution restores bile acid mediated
467 resistance to clostridium difficile. *Nature* 517:205–208.
- 468 34. Spinler JK, Brown A, Ross CL, Boonma P, Conner ME, Savidge TC. 2016. Administration of
469 probiotic kefir to mice with clostridium difficile infection exacerbates disease. *Anaerobe* 40:54–57.
- 470 35. Markey L, Shaban L, Green ER, Lemon KP, Mecsas J, Kumamoto CA. 2018. Pre-colonization
471 with the commensal fungus candida albicans reduces murine susceptibility to clostridium difficile
472 infection. *Gut Microbes* 1–13.
- 473 36. McKee RW, Aleksanyan N, Garrett EM, Tamayo R. 2018. Type IV pili promote clostridium
474 difficile adherence and persistence in a mouse model of infection. *Infection and Immunity* 86.
- 475 37. Yamaguchi T, Konishi H, Aoki K, Ishii Y, Chono K, Tateda K. 2020. The gut microbiome

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

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

- 527 Britton RA. 2018. Dietary trehalose enhances virulence of epidemic clostridium difficile. Nature
528 553:291–294.
- 529 54. Hryckowian AJ, Treuren WV, Smits SA, Davis NM, Gardner JO, Bouley DM, Sonnenburg JL.
530 2018. Microbiota-accessible carbohydrates suppress clostridium difficile infection in a murine model.
531 Nature Microbiology 3:662–669.
- 532 55. Fouladi F, Glenny EM, Bulik-Sullivan EC, Tsilimigras MCB, Sioda M, Thomas SA, Wang Y, Djukic
533 Z, Tang Q, Tarantino LM, Bulik CM, Fodor AA, Carroll IM. 2020. Sequence variant analysis reveals
534 poor correlations in microbial taxonomic abundance between humans and mice after gnotobiotic
535 transfer. The ISME Journal.
- 536 56. Walter J, Armet AM, Finlay BB, Shanahan F. 2020. Establishing or exaggerating causality for
537 the gut microbiome: Lessons from human microbiota-associated rodents. Cell 180:221–232.
- 538 57. Rasmussen TS, Vries L de, Kot W, Hansen LH, Castro-Mejía JL, Vogensen FK, Hansen AK,
539 Nielsen DS. 2019. Mouse vendor influence on the bacterial and viral gut composition exceeds the
540 effect of diet. Viruses 11:435.
- 541 58. Mims TS, Abdallah QA, Watts S, White C, Han J, Willis KA, Pierre JF. 2020. Variability in
542 interkingdom gut microbiomes between different commercial vendors shapes fat gain in response
543 to diet. The FASEB Journal 34:1–1.
- 544 59. Stewart DB, Wright JR, Fowler M, McLimans CJ, Tokarev V, Amaniera I, Baker O, Wong H-T,
545 Brabec J, Drucker R, Lamendella R. 2019. Integrated meta-omics reveals a fungus-associated
546 bacteriome and distinct functional pathways in clostridioides difficile infection. mSphere 4.
- 547 60. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, Grasis JA, Cassidy L, Tholey A,
548 Fickenscher H, Seegert D, Rosenstiel P, Schreiber S. 2017. Efficacy of sterile fecal filtrate transfer
549 for treating patients with clostridium difficile infection. Gastroenterology 152:799–811.e7.
- 550 61. Zuo T, Wong SH, Lam K, Lui R, Cheung K, Tang W, Ching JYL, Chan PKS, Chan MCW,
551 Wu JCY, Chan FKL, Yu J, Sung JJY, Ng SC. 2017. Bacteriophage transfer during faecal
552 microbiota transplantation inClostridium difficile infection is associated with treatment outcome. Gut

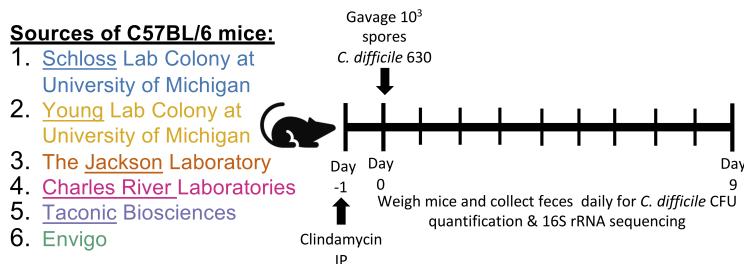
- 553 gutjnl–2017–313952.
- 554 62. Zuo T, Wong SH, Cheung CP, Lam K, Lui R, Cheung K, Zhang F, Tang W, Ching JYL, Wu JCY,
555 Chan PKS, Sung JJY, Yu J, Chan FKL, Ng SC. 2018. Gut fungal dysbiosis correlates with reduced
556 efficacy of fecal microbiota transplantation in clostridium difficile infection. *Nature Communications*
557 9.
- 558 63. Robinson JI, Weir WH, Crowley JR, Hink T, Reske KA, Kwon JH, Burnham C-AD, Dubberke
559 ER, Mucha PJ, Henderson JP. 2019. Metabolomic networks connect host-microbiome processes to
560 human *Clostridioides difficile* infections. *Journal of Clinical Investigation* 129:3792–3806.
- 561 64. Fletcher JR, Erwin S, Lanzas C, Theriot CM. 2018. Shifts in the gut metabolome and *Clostridium*
562 *difficile* transcriptome throughout colonization and infection in a mouse model. *mSphere* 3.
- 563 65. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, Liu C, Fang
564 Z, Chou J, Glanville J, Hao Q, Kotowska D, Colding C, Licht TR, Wu D, Yu J, Sung JJY, Liang Q, Li
565 J, Jia H, Lan Z, Tremaroli V, Dworzynski P, Nielsen HB, Bäckhed F, Doré J, Chatelier EL, Ehrlich
566 SD, Lin JC, Arumugam M, Wang J, Madsen L, Kristiansen K. 2015. A catalog of the mouse gut
567 metagenome. *Nature Biotechnology* 33:1103–1108.
- 568 66. Vital M, Rud T, Rath S, Pieper DH, Schlüter D. 2019. Diversity of bacteria exhibiting bile
569 acid-inducible 7alpha-dehydroxylation genes in the human gut. *Computational and Structural
570 Biotechnology Journal* 17:1016–1019.
- 571 67. Fransen F, Zagato E, Mazzini E, Fosso B, Manzari C, Aidy SE, Chiavelli A, D'Erchia AM, Sethi
572 MK, Pabst O, Marzano M, Moretti S, Romani L, Penna G, Pesole G, Rescigno M. 2015. BALB/c and
573 c57bl/6 mice differ in polyreactive IgA abundance, which impacts the generation of antigen-specific
574 IgA and microbiota diversity. *Immunity* 43:527–540.
- 575 68. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee
576 CA, Lynch SV, Tanoue T, Imaoka A, Itoh K, Takeda K, Umesaki Y, Honda K, Littman DR. 2009.
577 Induction of intestinal th17 cells by segmented filamentous bacteria. *Cell* 139:485–498.
- 578 69. Azrad M, Hamo Z, Tkhawkho L, Peretz A. 2018. Elevated serum immunoglobulin a levels in

- 579 patients with clostridium difficile infection are associated with mortality. *Pathogens and Disease* 76.
- 580 70. Saleh MM, Frisbee AL, Leslie JL, Buonomo EL, Cowardin CA, Ma JZ, Simpson ME, Scully KW,
581 Abhyankar MM, Petri WA. 2019. Colitis-induced th17 cells increase the risk for severe subsequent
582 clostridium difficile infection. *Cell Host & Microbe* 25:756–765.e5.
- 583 71. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a
584 dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the
585 MiSeq illumina sequencing platform. *Applied and Environmental Microbiology* 79:5112–5120.
- 586 72. Sze MA, Schloss PD. 2019. The impact of DNA polymerase and number of rounds of
587 amplification in PCR on 16S rRNA gene sequence data. *mSphere* 4.
- 588 73. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
589 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Horn DJV, Weber CF.
590 2009. Introducing mothur: Open-source, platform-independent, community-supported software
591 for describing and comparing microbial communities. *Applied and Environmental Microbiology*
592 75:7537–7541.
- 593 74. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012.
594 The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools.
595 *Nucleic Acids Research* 41:D590–D596.
- 596 75. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR,
597 Tiedje JM. 2013. Ribosomal database project: Data and tools for high throughput rRNA analysis.
598 *Nucleic Acids Research* 42:D633–D642.
- 599 76. Westcott SL, Schloss PD. 2017. OptiClust, an improved method for assigning amplicon-based
600 sequence data to operational taxonomic units. *mSphere* 2.
- 601 77. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB,
602 Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2018. Vegan: Community ecology

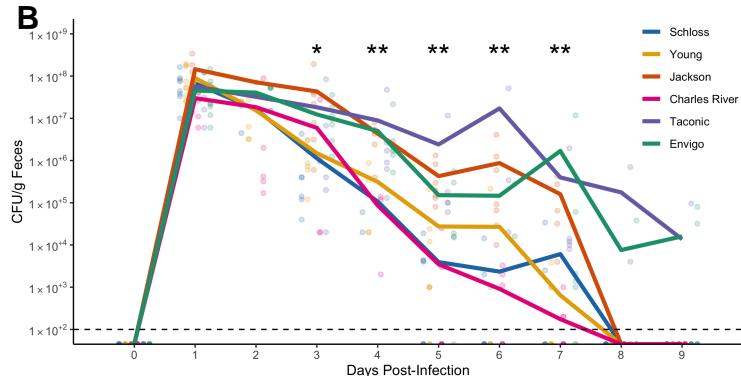
- 603 package.
- 604 78. R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for
605 Statistical Computing, Vienna, Austria.
- 606 79. Kuhn M. 2008. Building predictive models inRUsing thecaretPackage. Journal of Statistical
607 Software 28.
- 608 80. Topçuoğlu BD, Lesniak NA, Ruffin M, Wiens J, Schloss PD. 2019. A framework for effective
609 application of machine learning to microbiome-based classification problems.
- 610 81. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes
611 A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D,
612 Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the
613 tidyverse. Journal of Open Source Software 4:1686.

614 **Figures**

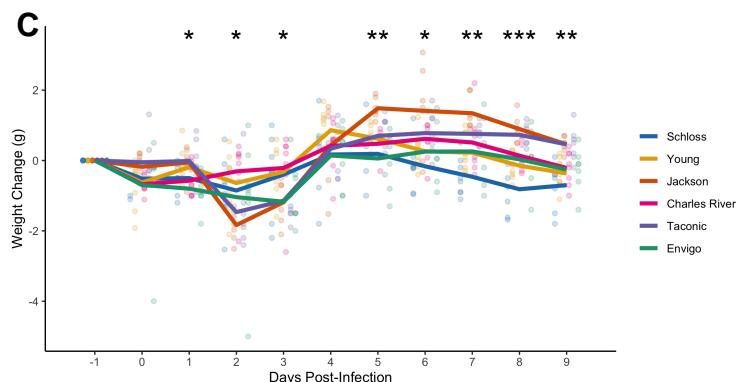
A



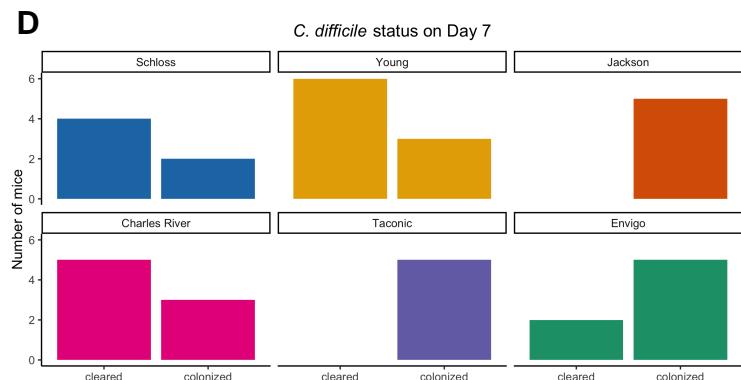
B



C



D

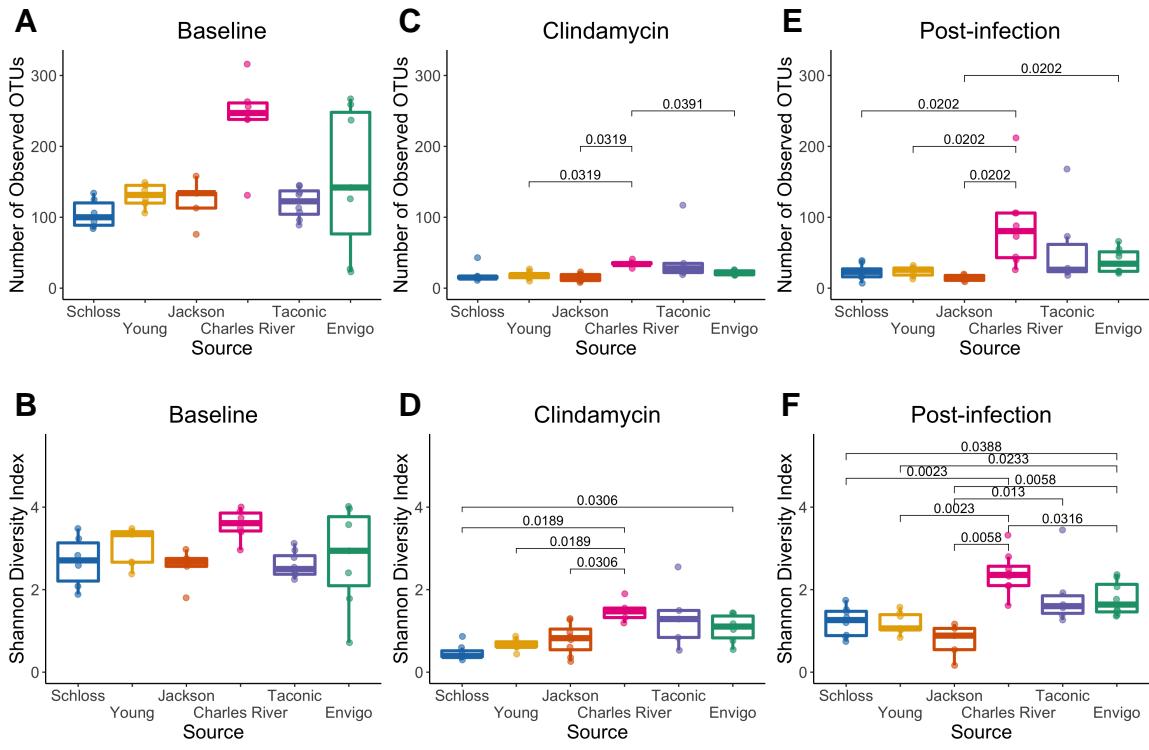


615

Figure 1. Clindamycin is sufficient

to promote *C. difficile* colonization regardless of colony source, but clearance time varies

617 **across sources of C57BL/6 mice.** A. Setup of the experimental timeline. Mice for the experiments
618 were obtained from 6 different sources: the Schloss (N = 8) and Young lab (N = 9) colonies at
619 the University of Michigan, the Jackson Laboratory (N = 8), Charles River Laboratory (N = 8),
620 Taconic Biosciences (N = 8), and Envigo (N = 8). All mice were administered 10 mg/kg clindamycin
621 intraperitoneally (IP) 1 day before challenge with *C. difficile* 630 spores on day 0. Mice were
622 weighed and feces was collected daily through the end of the experiment (9 days post-infection).
623 Note: 3 mice died during course of experiment. 1 Taconic mouse prior to infection and 1 Jackson
624 and 1 Envigo mouse between 1- and 3-days post-infection. B. *C. difficile* CFU/gram stool measured
625 over time (N = 20-49 mice per timepoint) via serial dilutions. The black line represents the limit of
626 detection for the first serial dilution. CFU quantification data was not available for each mouse due to
627 early deaths, stool sampling difficulties, and not plating all of the serial dilutions. C. Mouse weight
628 change measured in grams over time (N = 45-49 mice per timepoint), all mice were normalized to
629 the weight recorded 1 day before infection. For B-C: timepoints where differences across sources
630 of mice were statistically significant by Kruskal-Wallis test with Benjamini-Hochberg correction for
631 testing across multiple days (Table S1 and Table S2) are reflected by the asterisk(s) above each
632 timepoint (*, P < 0.05, **, P < 0.01, ***, P < 0.001). Lines represent the mean for each source and
633 circles represent individual mice. D. *C. difficile* CFU status on Day 7 within each mouse colony
634 source. Mice were classified as colonized or cleared (not detectable at the limit of detection of 100
635 CFU). N = 5-9 mice per group.



Figure

636

637 **2. Differences in microbial richness and diversity across mouse colony sources emerge**
 638 **after clindamycin treatment and infection.** A-B. Boxplots of the number of observed OTUs
 639 and Shannon diversity index values at baseline: day -1 (A-B), after clindamycin: day 0 (C-D)
 640 and post-infection: day 1 (E-F) timepoints of the experiment. Each circle represents the value
 641 for a stool sample from an individual mouse. Data were analyzed by Kruskal-Wallis test with
 642 Benjamini-Hochberg correction for testing each day of the experiment and the adjusted P value
 643 was < 0.05 for all panels except for B (Table S3). Significant P values from the pairwise Wilcoxon
 644 comparisons between sources with Benjamini-Hochberg correction are shown (Table S4).

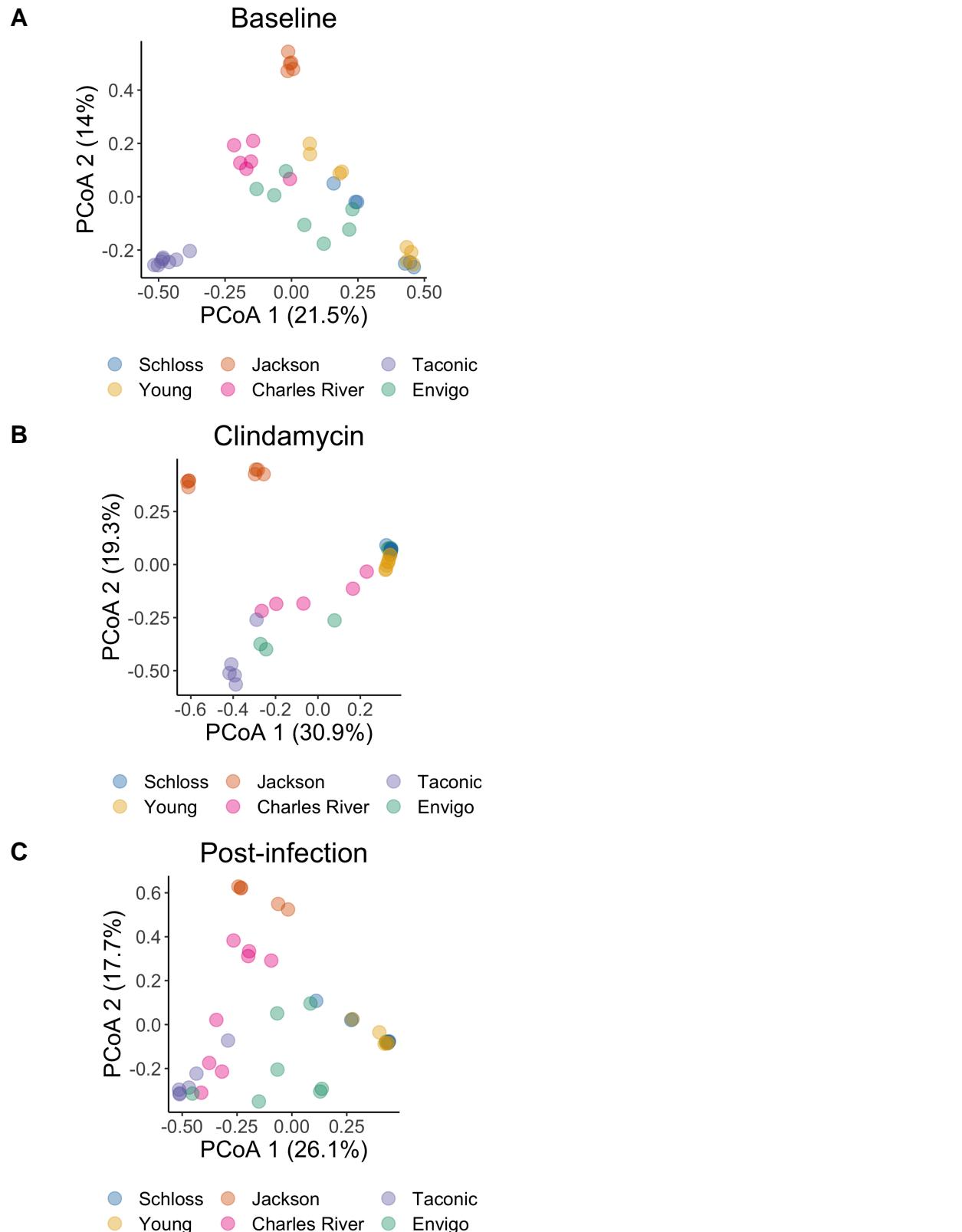
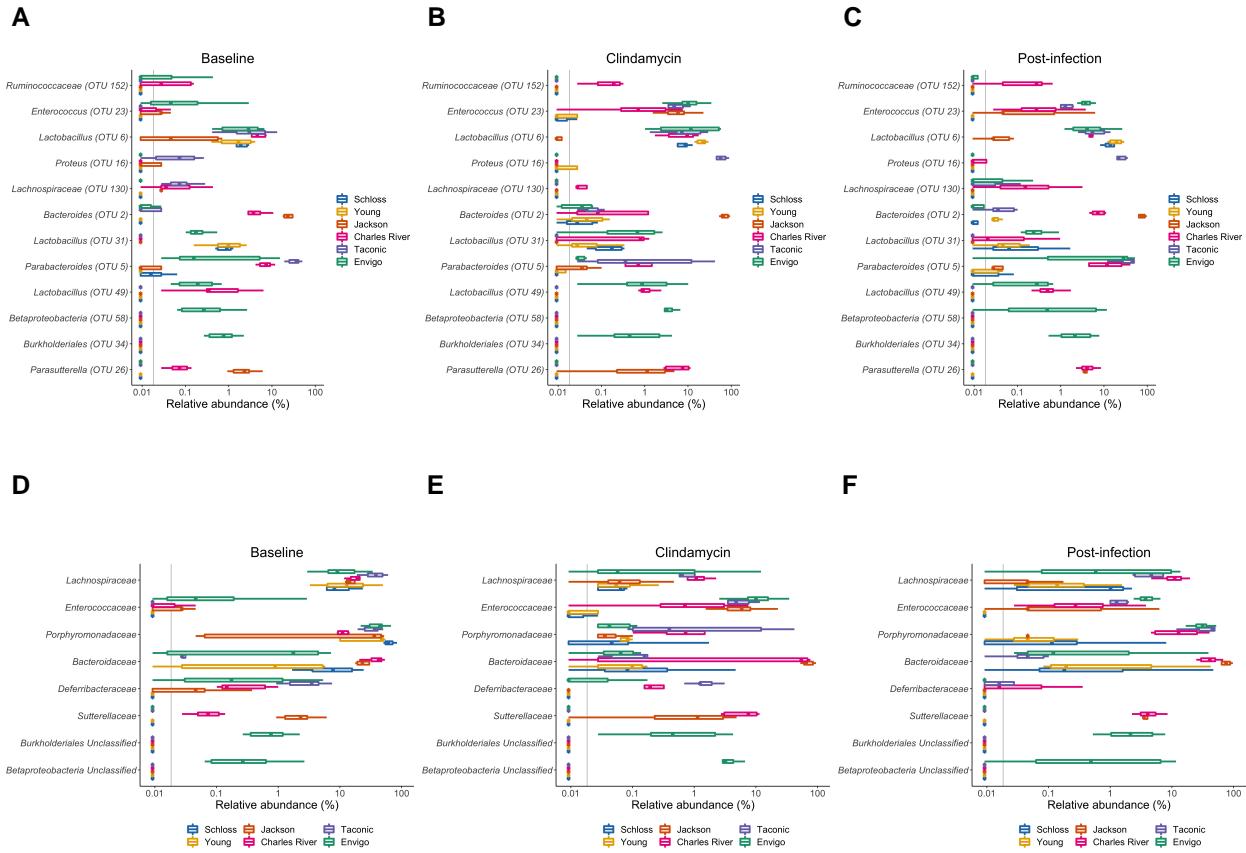


Figure 3. Mouse colony source is the

646 **variable that explains most of the variation observed in the baseline, post-clindamycin,**

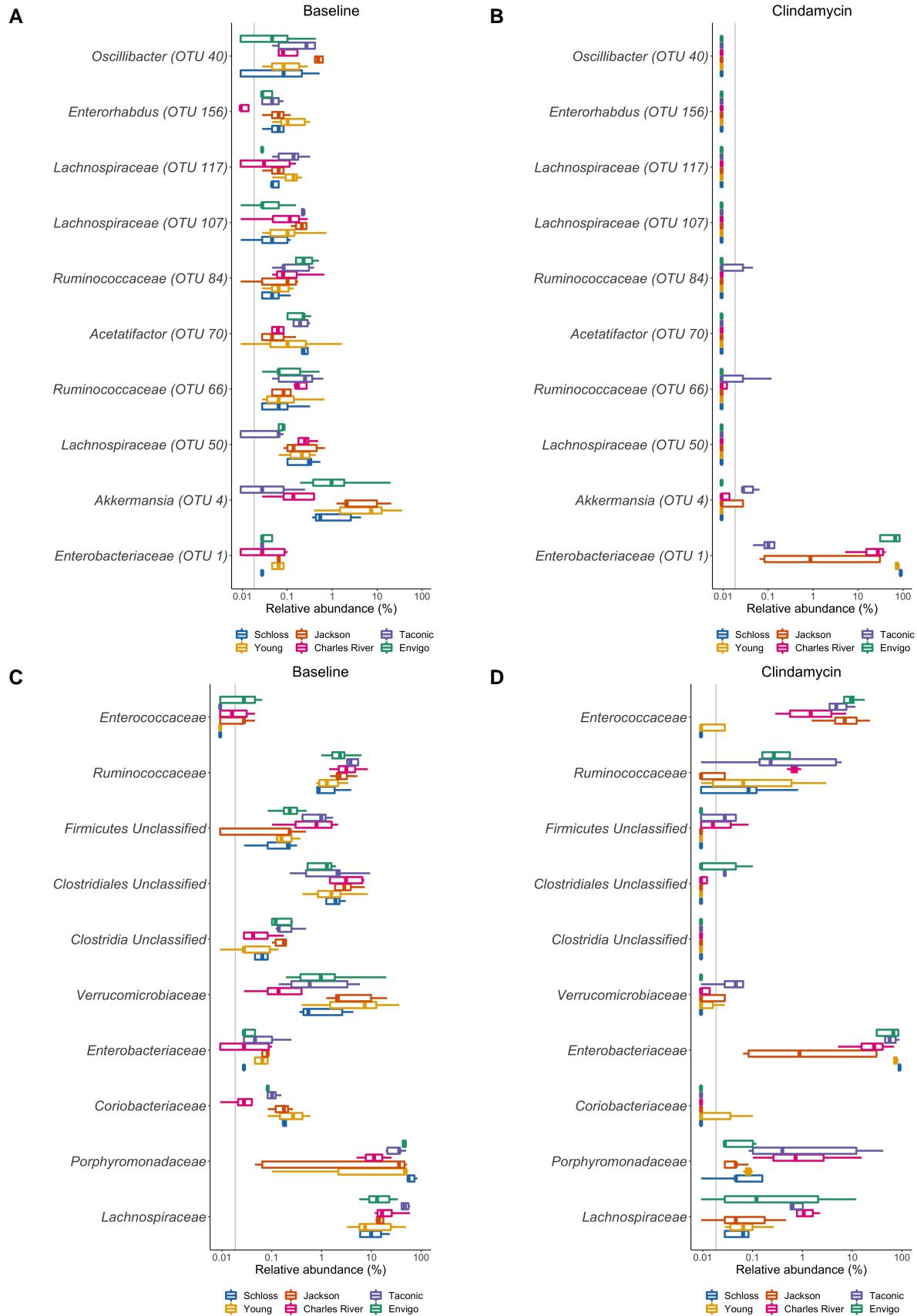
645

647 **and post-infection bacterial communities.** A-C. Principal Coordinates Analysis of Theta YC
648 distances from stools collected at baseline (A), post-clindamycin (B), and post-infection (C)
649 timepoints of the experiment. Each circle represents a stool sample from an individual mouse.
650 PERMANOVA analysis demonstrated that source and the interaction between source and cage
651 explained most of the variation observed in the baseline (combined $R^2 = 0.90$), post-clindamycin
652 (combined $R^2 = 0.99$), and post-infection (combined $R^2 = 0.88$) communities (all $P = 0.0001$, see
653 Table S6).



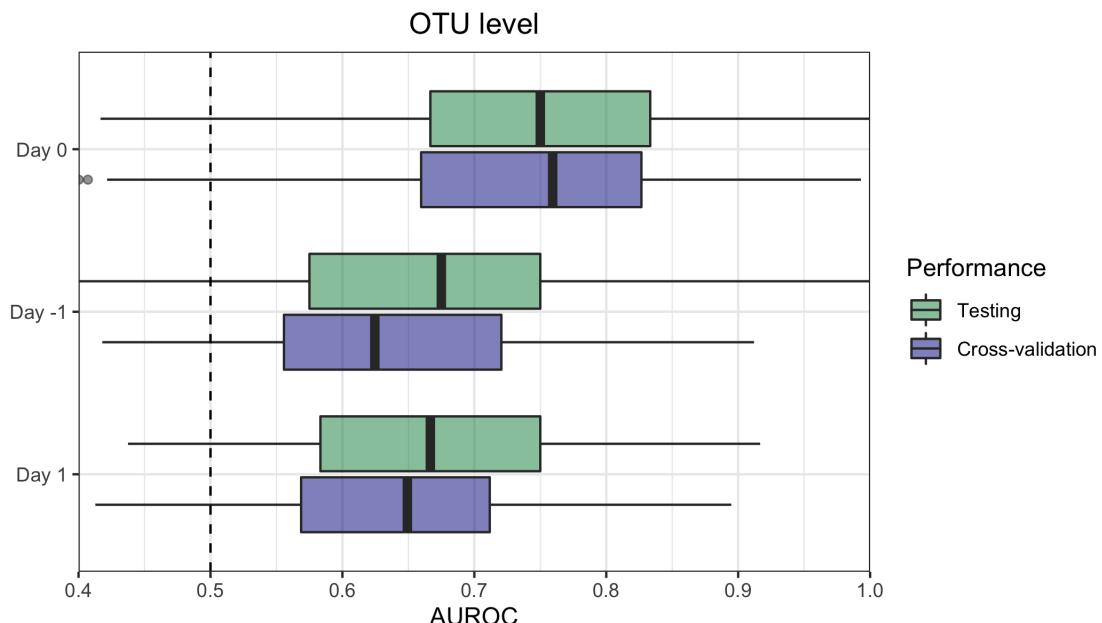
654

655 **Figure 4. A subset of bacterial consistently vary across mouse colony sources despite**
 656 **clindamycin perturbation and *C. difficile* challenge.** A-C. Boxplots of the relative abundances
 657 for the 12 OTUs that consistently varied across sources of mice at the baseline (A), post-clindamycin
 658 (B), and post-infection (C) timepoints of the experiment. D-F. Boxplots of the relative abundances for
 659 the 8 families that consistently varied across sources of mice at the baseline (D), post-clindamycin
 660 (E), and post-infection (F) timepoints of the experiment. For each timepoint bacteria with differential
 661 relative abundances across sources of mice were identified by Kruskal-Wallis test at the family and
 662 OTU level with Benjamini-Hochberg correction for testing all identified taxa at the respective level
 663 (Table S8-9). The grey vertical line indicates the limit of detection for A-F.

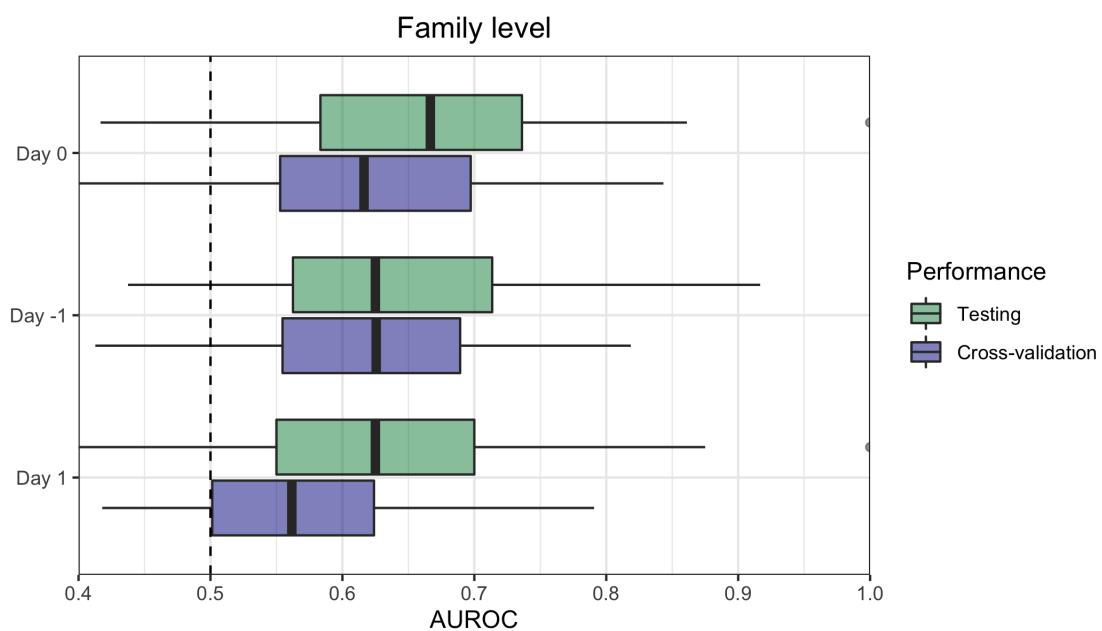


665 **Figure 5. Clindamycin treatment has the same effects on a subset of taxa regardless of**
666 **colony source.** A-B. Boxplots of the top 10 most significant (adjusted *P* value < 0.05) OTUs with
667 relative abundances that changed post clindamycin treatment. C-D. Boxplots of the top 10 most
668 significant families with relative abundances that changed post clindamycin treatment. Data were
669 analyzed by Wilcoxon signed rank test limited to mice that had paired sequence data for day -1
670 and 0 (*N* = 31). Tests were performed at the OTU and family levels with Benjamini-Hochberg
671 correction for testing all identified OTUs and families. See Table S10-11 for complete list of OTUs
672 and families significantly impacted by clindamycin treatment. The grey vertical line indicates the
673 limit of detection for A-D.

A



B



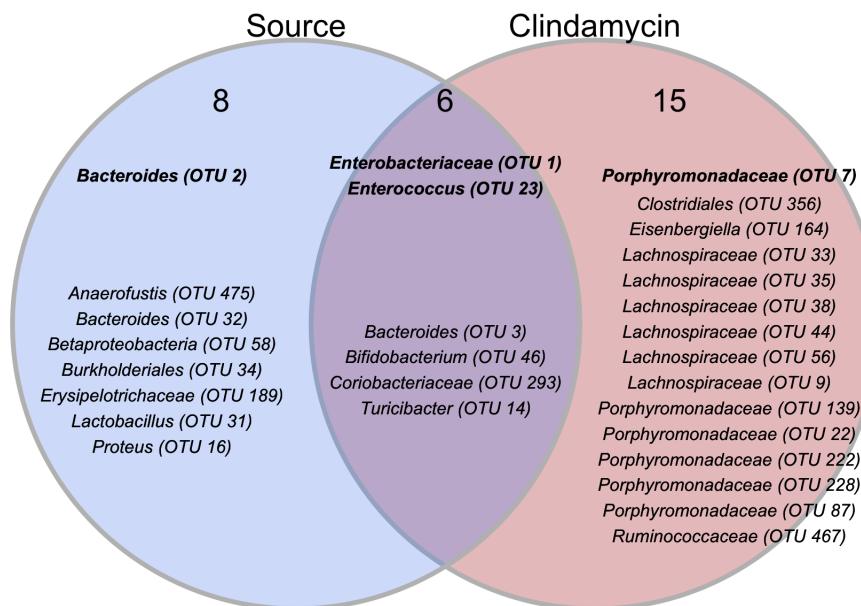
674

Figure

675 **6. Bacterial community composition before, after clindamycin perturbation, and**
676 **post-infection can predict *C. difficile* colonization status 7 days post-challenge. A-B.**
677 Logistic regression classification model AUROCs to predict *C. difficile* CFU on D7 (Fig. 1D, Fig.

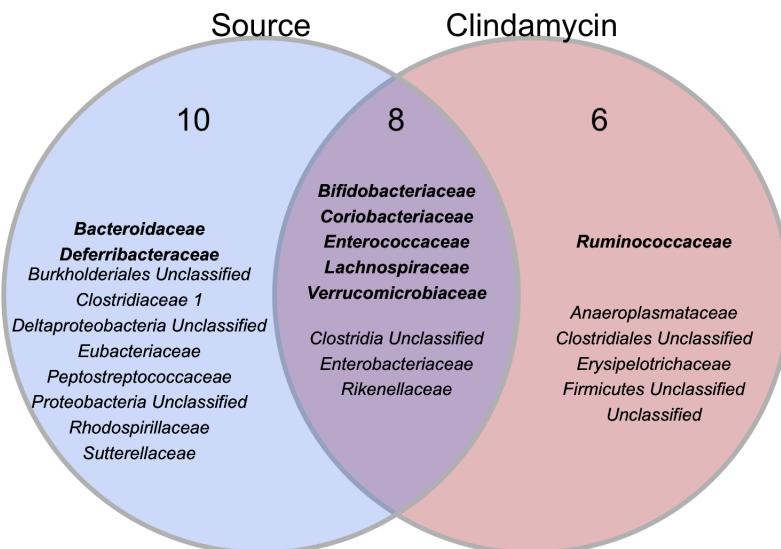
678 S3) based on the community relative abundances at baseline (day -1), post-clindamycin (day 0),
679 and post-infection (day 1) at either the OTU (A) or family (B) level. All models performed better
680 than random chance (AUROC = 0.5), see (all $P \leq 5e-15$; Table S12) and the model built with
681 post-clindamycin treated bacterial OTU relative abundances had the best performance ($P_{FDR} \leq$
682 3.9e-10 for pairwise comparisons; Table S13). A List of the 20 taxa that were ranked as most
683 important to each model are listed in Table S14-15.

A



B

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



Figure

- 684
685 **7. Key taxa that influence *C. difficile* colonization status vary by mouse colony source,**
686 **change after clindamycin treatment, or both.** A-B. Venn diagrams that combine Fig. S4-6
687 summaries of taxa that were important to the day -1, 0, and 1 classification models (Table S14-15)

688 and either overlapped with taxa that varied across vendors at the same timepoint, were impacted
689 by clindamycin treatment, or both. See Fig. S4-6 for separate comparisons of taxa from the day
690 -1, 0, and 1 classification models. Bold taxa signify bacteria that were important to more than 1
691 classification model.

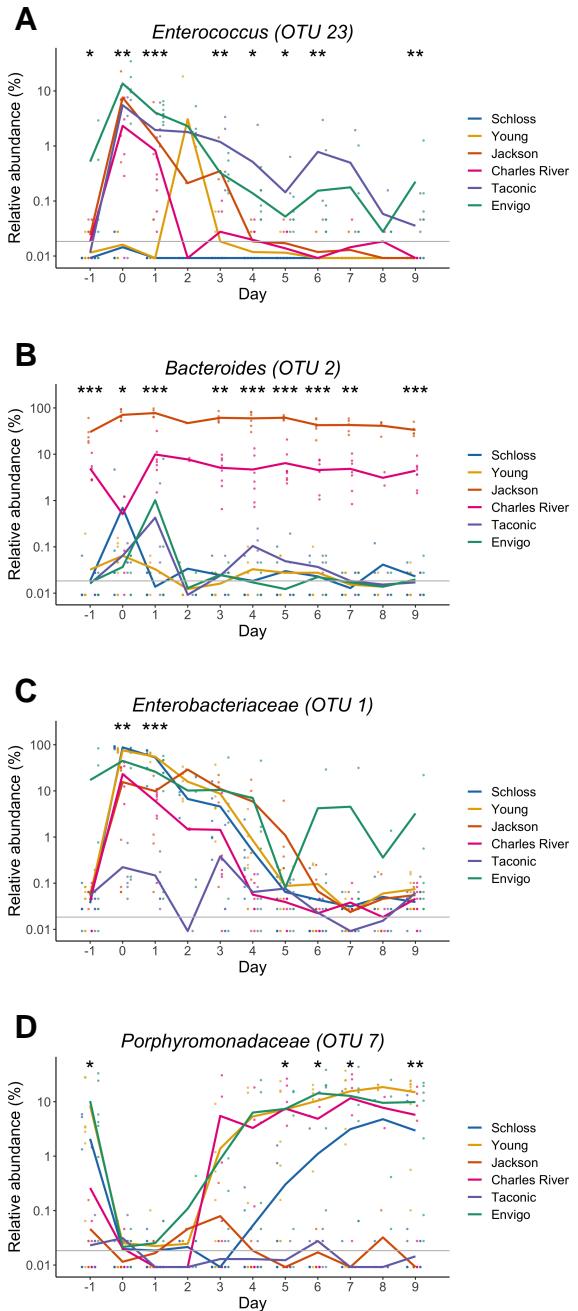


Figure 8: Key OTUs vary across sources

throughout the experiment. A-C. Relative abundances of bold OTUs from Fig. 7A that were important for at least two classification models are shown over time. A. *Enterococcus* (OTU 23), which significantly varied across sources and was impacted by clindamycin treatment. B. *Bacteroides* (OTU 2), which varied across sources throughout the experiment. C. *Enterobacteriaceae* (OTU 1) and *Porphyromonadaceae* (OTU 7) were significantly impacted by clindamycin treatment and examining relative abundance dynamics over the course of the

699 experiment indicated timepoints where relative abundances also significantly varied across sources
700 of mice. Each point represents the relative abundance of an individual mouse stool sample and
701 colored lines represent the mean relative abundances for each source of mice. The gray horizontal
702 line represents the limit of detection. Timepoints where differences across sources of mice were
703 statistically significant by Kruskal-Wallis test with Benjamini-Hochberg correction for testing across
704 multiple days (Table S16) are identified by the asterisk(s) above each timepoint (*, P < 0.05, **, P <
705 0.01, ***, P < 0.001).

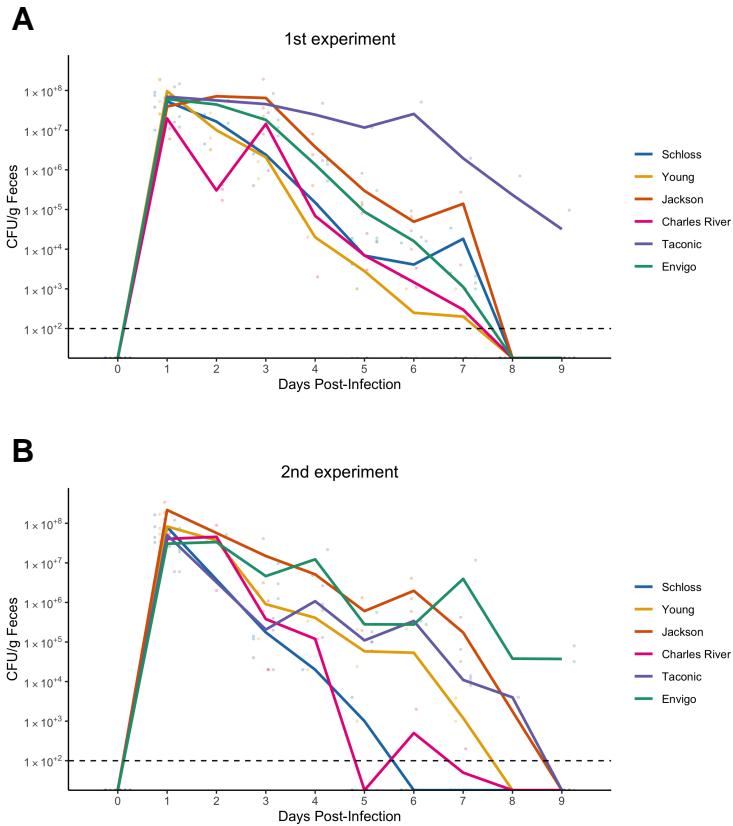


Figure S1. *C. difficile* CFU variation

706
 707 **across vendors is similar between the 2 experiments.** (A-B). *C. difficile* CFU/gram of stool
 708 quantification over time for experiment 1 (A) and 2 (B). Experiments were conducted approximately
 709 3 months apart. Lines represent the mean CFU for each source and circles represent individual
 710 mice and the black line represents the limit of detection.

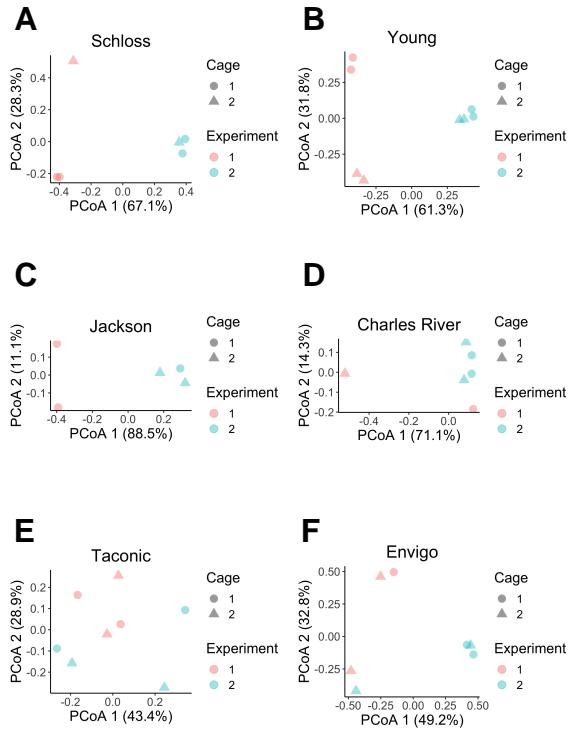


Figure S2. Only bacterial communities from

711 **University of Michigan mice significantly vary between experiments.** A-F. PCoA of Theta YC
 712 distances for the baseline fecal bacterial communities within each source of mice. Each symbol
 713 represents a stool sample from an individual mouse with color corresponding to experiment and
 714 shape representing cage mates. PERMANOVA was performed within each group to examine the
 715 contributions of experiment and cage to observed variation. Experiment number and cage only
 716 significantly explained observed variation for mice from the Schloss (combined $R^2 = 0.99$; $P \leq$
 717 0.033) and Young (combined $R^2 = 0.95$; $P \leq 0.029$) lab colonies (Table S7).

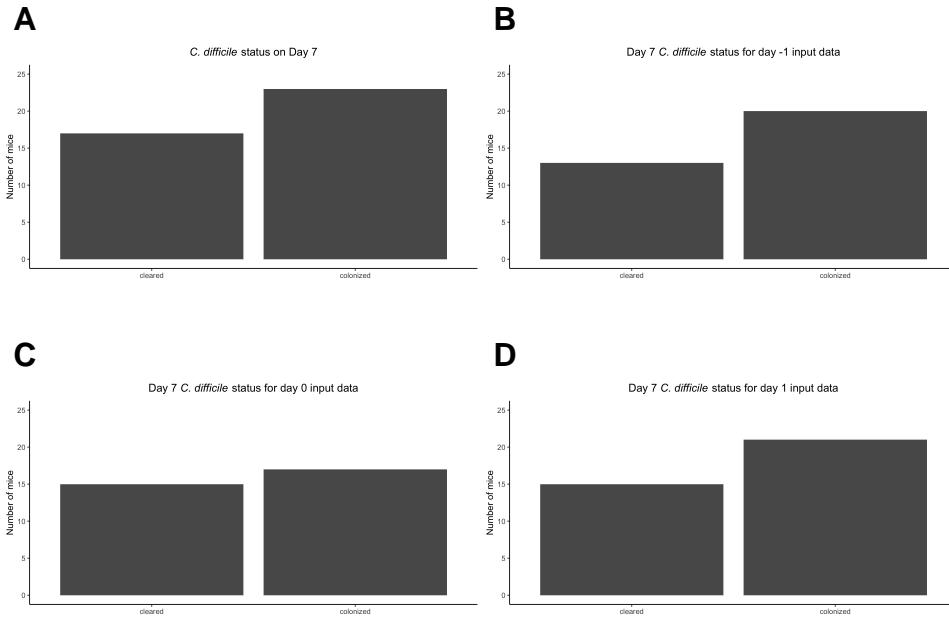
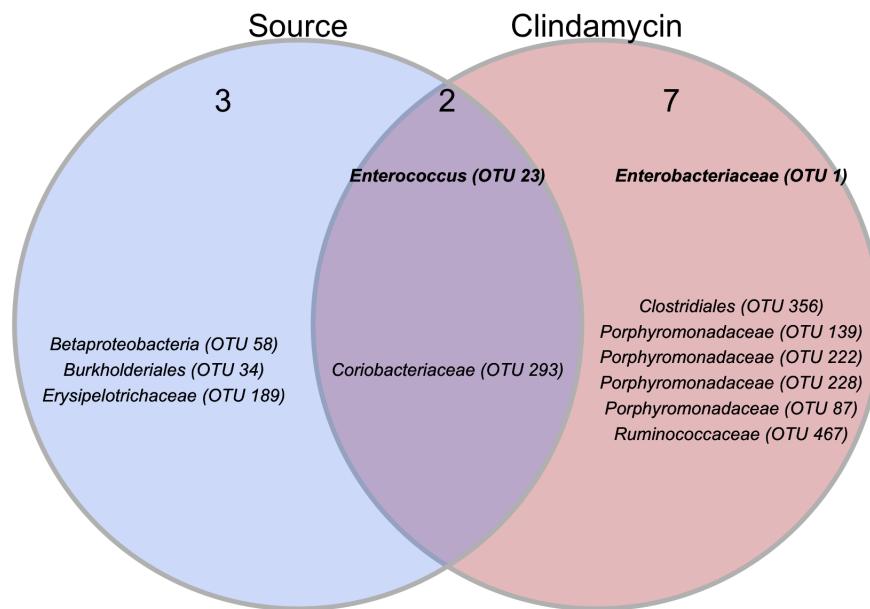


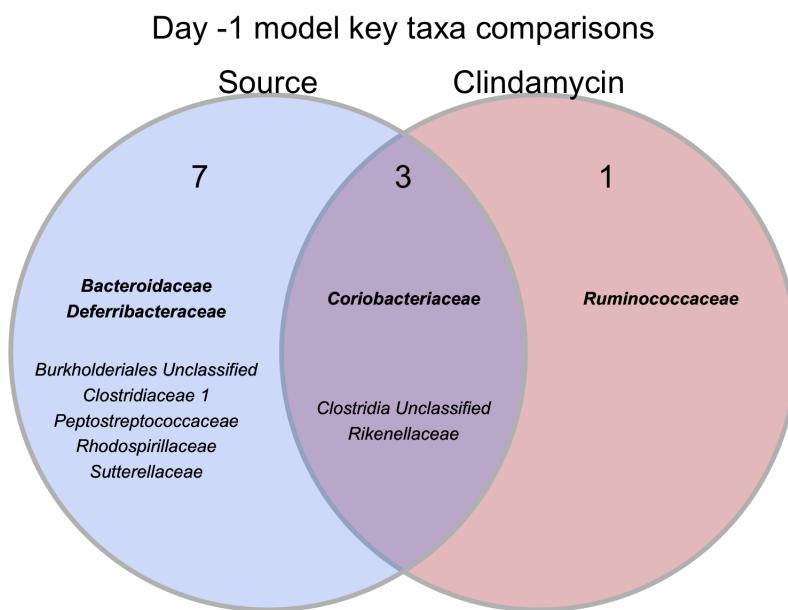
Figure S3. C.

719
 720 ***C. difficile* colonization status 7 days post-infection varies across mice.** A. Bar graph
 721 visualizations of overall day 7 *C. difficile* colonization status that were used as classification
 722 outcomes to build logistic regression models. Mice were classified as colonized or cleared
 723 (not detectable at the limit of detection of 100 CFU) based on CFU g/stool data from 7 days
 724 post-infection. B-D. Classification input data used for models based on day -1 (B), day 0 (C),
 725 and day 1 (D) community relative abundance data. Only mice with day 7 *C. difficile* status and
 726 community relative abundance data for each respective timepoint were used as input data for the
 727 classification models.

A



B



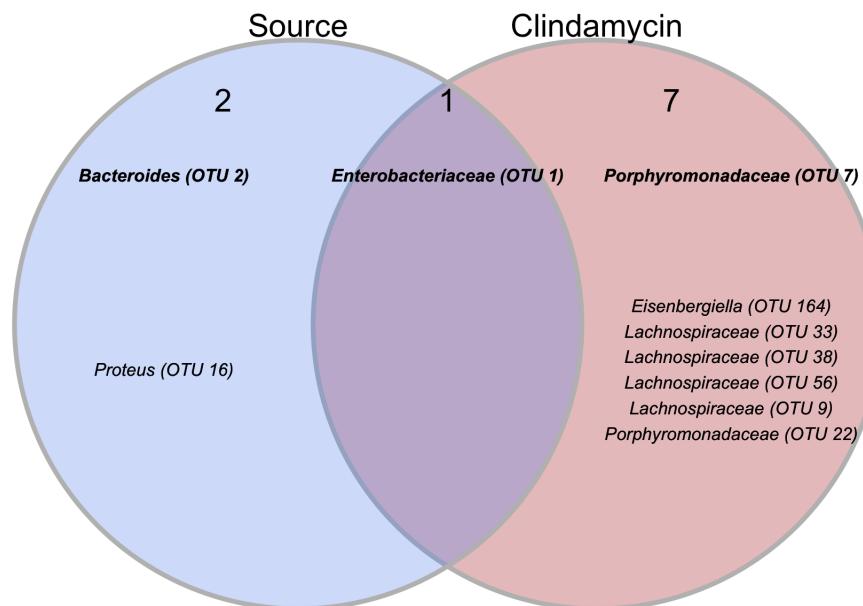
Figure

728

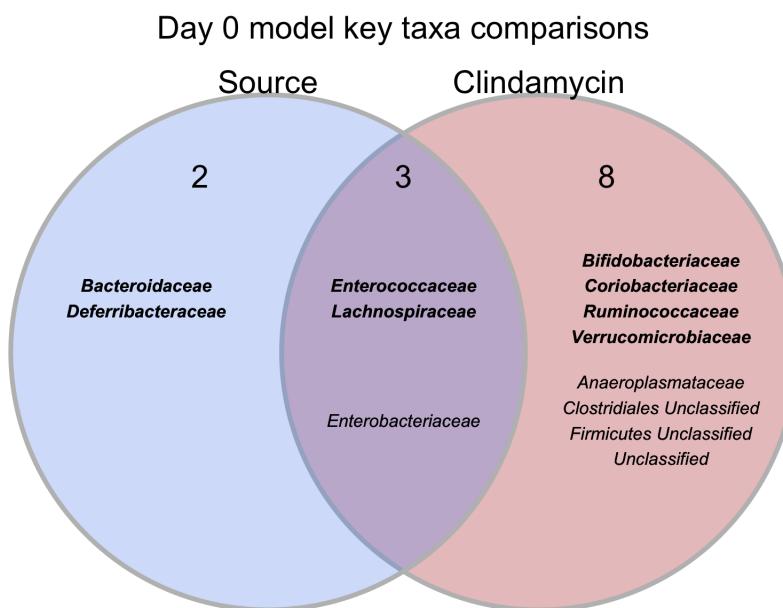
729 **S4. Key taxa from classification model based on baseline community data vary by mouse**
730 **colony source, clindamycin treatment, or both.** A-B. Venn diagrams of top 20 important OTUs
731 (A) and families (B) from day -1 classification models (Table S14-15) that overlapped with taxa that

⁷³² varied across vendors at baseline, were impacted by clindamycin treatment, or both. Bold taxa
⁷³³ signify bacteria that were important to more than 1 classification model.

A



B



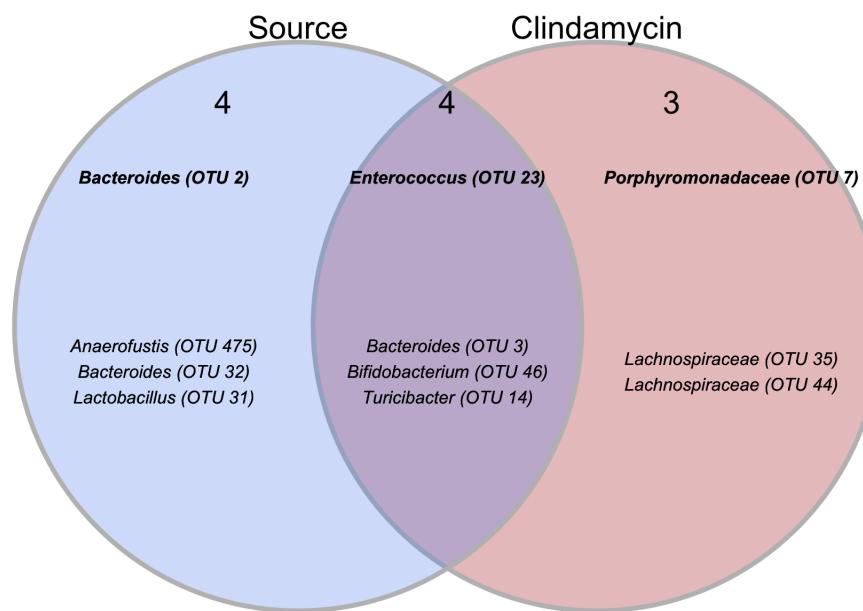
Figure

734

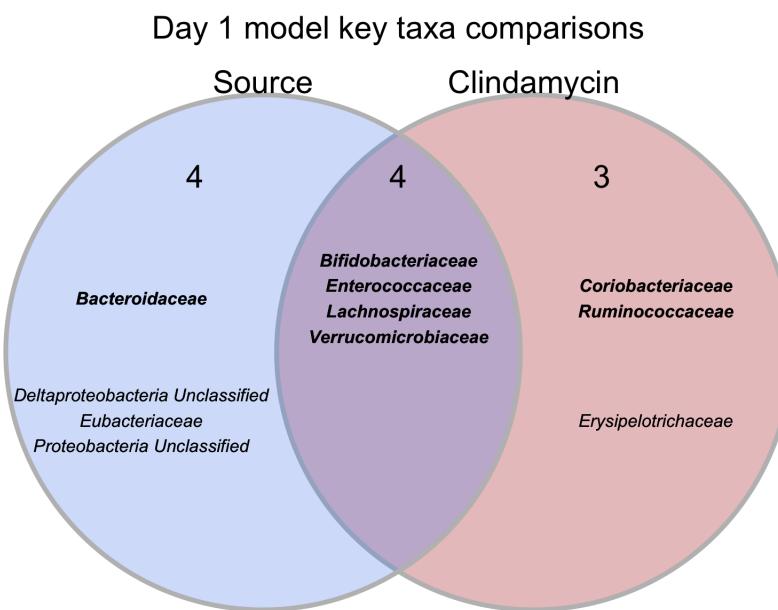
735 **S5. Key taxa from classification model based on post-clindamycin treatment community**
736 **data vary by mouse colony source, clindamycin treatment, or both.** A-B. Venn diagrams
737 of top 20 important OTUs (A) and families (B) from day 0 classification models (Table S14-15)

738 that overlapped with taxa that still varied across vendors after clindamycin, were impacted by
739 clindamycin treatment, or both. Bold taxa signify bacteria that were important to more than 1
740 classification model.

A



B



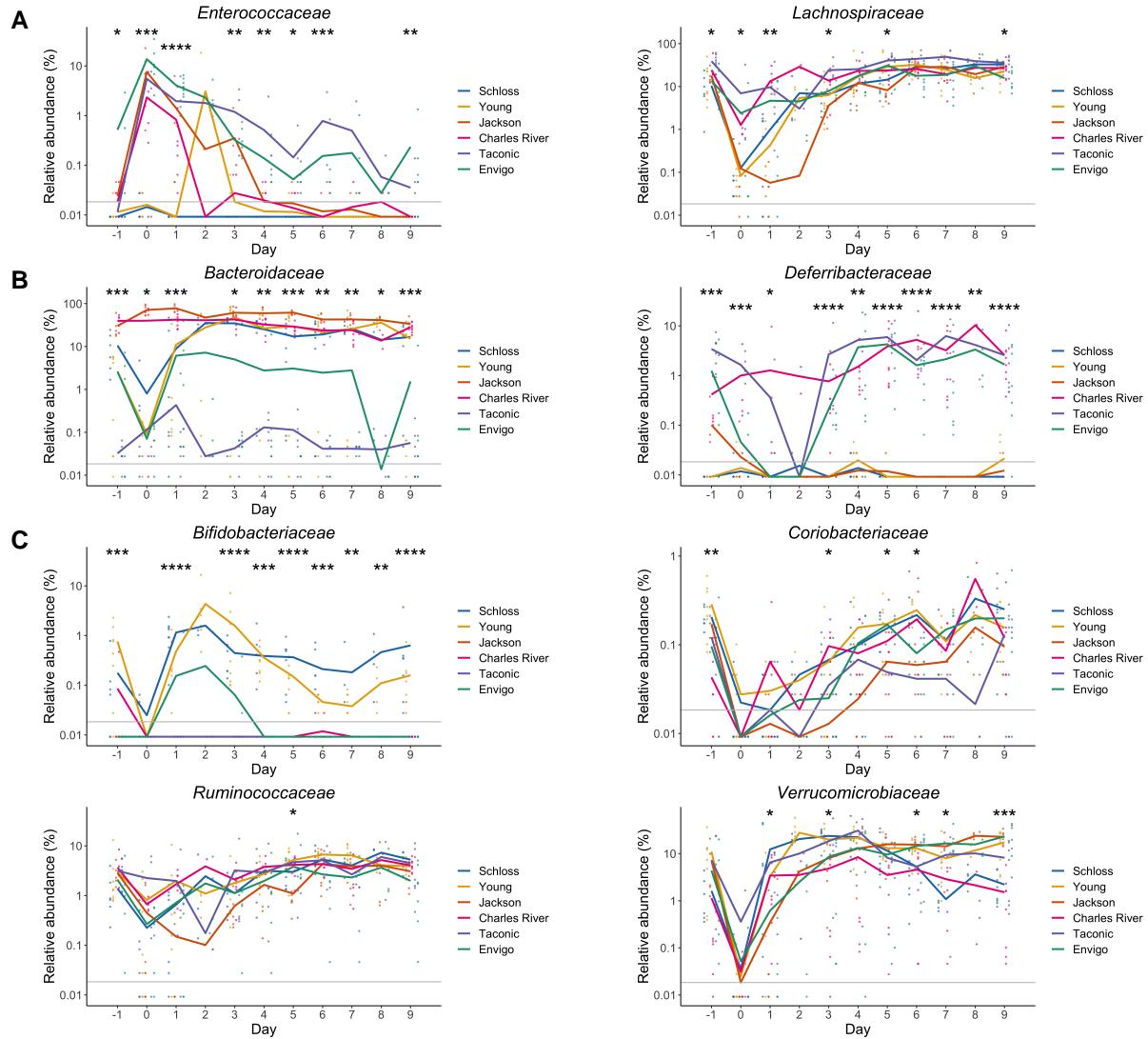
Figure

741

742 **S6. Key taxa from classification model based on post-infection community data vary by**
743 **mouse colony source, clindamycin treatment, or both.** A-B. Venn diagrams of top 20 important
744 OTUs (A) and families (B) from day 1 classification models (Table S14-15) that overlapped with

⁷⁴⁵ taxa that varied across vendors post-infection, were impacted by clindamycin treatment, or both.

⁷⁴⁶ Bold taxa signify bacteria that were important to more than 1 classification model.



747

748 **Figure S7. Key families vary across sources throughout experiment.** Relative abundances of
 749 bold families from Fig. 7B that were important for at least two classification models are shown
 750 over time. A. *Enterococcaceae* and *Lachnospiraceae*, which significantly varied across sources
 751 and were impacted by clindamycin treatment. B. *Bacteroidaceae* and *Deferrribacteraceae*, which
 752 varied across sources throughout the experiment. C. *Bifidobacteriaceae*, *Coriobacteriaceae*,
 753 *Ruminococcaceae*, and *Verrucomicrobiaceae* were significantly impacted by clindamycin treatment.
 754 Examining the relative abundance dynamics throughout the experiment, identified timepoints where
 755 relative abundances also significantly varied across sources of mice. Each point represents the
 756 relative abundance of an individual mouse stool sample and colored lines represent the mean
 757 relative abundances for each source of mice. The gray horizontal line represents the limit of

758 detection. Timepoints where differences across sources of mice were statistically significant by
759 Kruskal-Wallis test with Benjamini-Hochberg correction for testing across multiple days (Table S17)
760 are identified by the asterisk(s) above each timepoint (*, P < 0.05, **, P < 0.01, ***, P < 0.001).

761 **Supplementary Tables and Movie**

762 All supplemental material is available at: https://github.com/SchlossLab/Tomkovich_vendor_difs_XXXX_2020/submission.

764 **Movie S1. Large shifts in bacterial community structure occurred after clindamycin and *C.***
765 ***difficile* infection.** PCoA of Theta YC distances animated from 0 through 9 days post-infection.
766 PERMANOVA analysis indicated colony source was the variable that explained the most observed
767 variation across fecal communities (source $R^2 = 0.35$, $P = 0.0001$) followed by interactions between
768 cage and day of the experiment. Transparency of the circle corresponds to the day of the experiment,
769 each circle represents a sample from an individual mouse at a specific timepoint. See Table S5 for
770 PERMANOVA results).

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

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

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

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

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

776 **Table S6. PERMANOVA results for all mice, all timepoints.**

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

778 **Table S8. OTUs with relative abundances that significantly vary across sources at baseline,**
779 **post-clindamycin, or post-infection timepoints.**

780 **Table S9. Families with relative abundances that significantly vary across sources at**
781 **baseline, post-clindamycin, or post-infection timepoints.**

782 **Table S10. OTUs with relative abundances that significantly changed after clindamycin**
783 **treatment.**

784 **Table S11.** Families with relative abundances that significantly changed after clindamycin
785 treatment. **Table S12.** Statistical results of logistic regression model performances
786 compared to random chance.

787 **Table S13.** Pairwise Wilcoxon results of comparing all 6 logistic regression model
788 performances.

789 **Table S14.** Top 20 most important OTUs for each of the 3 logistic regression models based
790 on OTU relative abundance data.

791 **Table S15.** Top 20 most important families for each of the 3 logistic regression models
792 based on OTU relative abundance data.

793 **Table S16.** OTUs with relative abundances that significantly varied across sources of mice
794 on at least 1 day of the experiment.

795 **Table S17.** Families with relative abundances that significantly varied across sources of mice
796 on at least 1 day of the experiment.