**Running title:** Antibiotic induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against *Clostridium difficile*

**Short title:** Alterations of the murine gut microbiota lead to *C. difficile* colonization

Alyxandria M. Schubert, Hamide Sinani, and Patrick D. Schloss\*

\* Correspondence: > Department of Microbiology and Immunology

University of Michigan

1520A Medical Science Research Bldg. I

1500 W. Medical Center Dr.

Ann Arbor, MI 48109

pschloss@umich.edu

734.647.5801

### Abstract

### Introduction

The microbiota, or the diverse community of microorganisms living in and on the body, has an integral role in deterring pathogen colonization and infection [ref]. This native protection by the microbiota from invasive pathogenic species is termed colonization resistance [ref]. It is well established that the gut bacterial microbiota is critical in the hosts’ defense against the pathogen *Clostridium difficile* [ref]. When this indigenous community is perturbed, this often leads to a loss of resistance. This is especially important in many hospital settings where patients are not only exposed to various types and degrees of perturbations, such as antibiotics, diet changes, and chemotherapy, but they are also exposed to *C. difficile* spores from their environment [ref]. *C. difficile* infections (CDI) are the most reported hospital-acquired infection in the United States and are responsible for 14,000 deaths a year [ref].

It is not completely understood how different perturbations to the gut microbiota result in a loss of colonization resistance to *C. difficile*. There is a clear need to better understand the ecology of *C. difficile* and its interactions with members of the microbiota. In mouse models of CDI, the baseline, untreated murine microbiota is completely resistant to *C. difficile* colonization. It was previously shown that C57Bl/6 mice treated with cefoperazone [], clindamycin [], or clindamycin in combination with a five antibiotic cocktail [] had decreased colonization resistance. These studies suggest that a loss of Lachnospiraceae and a bloom of Lactobacilliaceae and Enterobacteriaceae are responsible for the loss of colonization resistance. These results are largely supported by human association studies. We observed significant differences between the gut microbiota of hospitalized individuals with and without *C. difficile* and non-hospitalized controls. In addition, fecal microbiota transplants increase *Bacteroidetes* and decrease *Proteobacteria* levels in recipients, resulting in a successful restoration of colonization resistance in patients [Seekatz et al. 2014]. Precisely how this occurs is not fully understood, but it emphasizes the importance of the gut microbiota in colonization resistance against *C. difficile*.

Because the gut microbiota is a complex community we need tools that can enable us to dissect the interactions within the community and with *C. difficile*. One approach is the use of models to identify associations between members of the microbiota and *C. difficile*. Models have been used to predict *C. difficile* and *Citrobacter* infection [ref], colon cancer [ref], and psoriasis [ref] based on the composition of the gut microbiota [Pamer, psoriases, Schubert; Statnikov et al 2013] [Schubert et al. 2013] [Stein et al. 2013] [Belzer et al 2014]. By using models to explain the relationship between members of the gut microbiota, we hope to identify the subset of the normal murine microbiota that are responsible for colonization resistance.

The purpose of this investigation was to expand our current knowledge of the effects of various perturbations on colonization resistance against *C. difficile*. Through the administration of different antibiotic classes, doses, recovery times we altered the murine gut microbiota and challenged the communities with C. difficile spores and observed differences in colonization resistance. We then used 16S rRNA gene sequencing to identify structural changes within the microbiota that would be predictive of colonization resistance. Using these data, we built a random forest model to predict *C. difficile* colonization levels. Through this analysis, we have identified groups of related bacteria that are associated with *C. difficile* colonization resistance. This model revealed that the interactions that gave rise to colonization resistance were non-linear and context dependent. These findings show we can successfully apply modeling techniques to accurately measure the colonization resistance ability of a given microbiota.

### Results

***Antibiotics differentially alter the structure of the microbiota and their colonization resistance to* C. difficile*.*** We selected a panel of seven antibiotics from six classes with the goal of differentially altering the microbiota and assessing their resistance to *C. difficile* colonization (**Table 1**). On the day after the mice were challenged with *C. difficile* we enumerated the density of *C. difficile* in the animals' feces. We observed reproducibly high levels of *C. difficile* colonization in mice treated with cefoperazone, metronidazole, and streptomycin (**Figure 1**). We observed variable levels of *C. difficile* colonization in mice treated with amipicilin. Only one of six mice receiving vancomycin was colonized with *C. difficile* and none of the mice that received ciprofloxacin were colonized. We sequenced the 16S rRNA genes from the fecal communities of treated and untreated mice prior to *C. difficile* challenge to identify populations within the microbiota that conferred colonization resistance. The antibiotic treatments resulted in distinct and reproducible changes to the structure of the microbiota relative to the untreated animals (**Figure 1** and **Figure S1**; AMOVA, P<0.05). Comparisons of the microbiota between antibiotic classes indicated that their structures were significantly different from each other (AMOVA, XP<0.36X). The community structure of the mice receiving beta-lactams (i.e. cefoperazone and ampicillin) were not significantly different from each other (AMOVA, P<X0.36X). These results indicate that perturbing the gut microbiota with antibiotics resulted in non-overlapping community structures that yielded significant variation in susceptibility to colonization when challenged with *C. difficile*.

***Reduced perturbations result in altered levels of colonization.*** Based on the *C. difficile* colonization levels in our seven antibiotic treatments, we hypothesized that titrating the dose of antibiotics that the mice received would result in smaller perturbations to the microbiota and a reduced sensitivity to *C. difficile*. In addition to the previous treatments, we treated mice with lower concentrations of cefoperazone (0.3 and 0.1 mg/mL), streptomycin (0.5 and 0.1 mg/mL), and vancomycin (0.3 and 0.1 mg/mL; Figure S2). These antibiotics were selected because they are thought to target a broad spectrum of bacteria (i.e. cefoperazone), Gram-negative (i.e. streptomycin), and Gram-positive (i.e. vancomycin) bacteria. As expected, colonization levels decreased significantly in all mice receiving titrated doses of cefoperazone (P<0.02; Figure 2). Titrating the dose of cefoperazone in the animals' drinking water resulted in significant decreases in the relative abundance of an OTU associated with the genus *Escherichia* (OTU 4) and increases in the relative abundances of an OTU associated with the family *Porphyromonadaceae* (OTU 9) and an OTU associated with the genus *Pseudomonas* (OTU 65; Figure 2). The dose response for these three OTUs qualitatively followed what we had expected based on the correlation-based analysis. Reducing the dose of streptomycin significantly reduced the colonization levels (P<0.01; Figure 2). Titrating the dose of streptomycin in the drinking water resulted in significant changes in the relative abundance of OTUs associated with the *Porphyromonadaceae* (OTUs 2, 3, 5, 9, 10, 13), *Allistipes* (OTU 11), and *Bacteroidales* (OTU 17). In addition to its anti-Gram-positive activity, Vancomycin was also selected because although the community was quite different from untreated mice, we observed high levels of colonization in only one mouse. We anticipated that lower doses might result in a community structure that would result in colonization. In fact, the 0.3 and 0.1 mg/mL doses of vancoymicin resulted in similarly high levels of colonization (P=0.96). Seven OTUs were differentially represented across the three doses of vancomycin. Surprisingly, even though the colonization levels did not significantly differ between the mice receiving 0.1 and 0.3 mg/mL of vancomycin in their drinking water, four of the OTUs that had significantly different relative abundances were only found in the lower dose. Three of these were affiliated with members of the *Porphyromonadaceae* (OTUs 2, 3, and 5) and one was affiliated with a member of the genus *Bacteroides* (OTU 1). Two OTUs affiliated with the *Akkermansia* (OTU 6) and *Lactobacillus* (OTU 8) genera increased with increasing dose and a third OTU affiliated with *Escherichia* (OTU 4) had a mixed response to the dose level. These results suggest that the context of the microbiota is important in determining the overall resistance to *C. difficile*. For example, the relationship between the *Bacteroides* (OTU 1) and *C. difficile* colonization is positive in streptomycin-treated mice and it is negative in cefoperazone-treated mice. In addition, cefoperazone and streptomycin-treated mice have high levels of *C. difficile* although the former have significantly higher levels of *Escherichia* (OTU 4), which are absent in the streptomycin-treated mice. Together, these results suggest that individual populations were not sufficient to consistently predict colonization resistance. In light of such results, resistance is likely a product of the overall composition of the community.

***Allowing recovery of the microbiota restores colonization resistance.*** In the experiments we have described thus far, we have allowed the gut microbiota to recover for 24 hours before challenging them with *C. difficile*. Several studies have demonstrated that perturbed communities can return to a "healthy" state in which resistance to *C. difficile* is restored [refs]. To test the effect of recovery on colonization and gain greater insights into the populations that confer colonization resistance, we allowed the microbiota of the mice that received the full metronidazole and ampicillin treatment to recover for an additional five days (Figure S3). Among the metronidazole-treated mice, those with extended recovery had a 1.86x10^6-fold reduction in colonization (P<0.001; Figure 4). In addition, of the 14 mice that experienced the longer recovery period, 7 had no detectable *C. difficile* 24 hours after challenge. We detected six OTUs that were differentially represented in the two sets of metronidazole-treated mice (Figure 3). Most notable among these was a member of the *Barnesiella* (OTU 3) and the *Escherichia* (OTU 4). The relative abundance of the former increased with the delay and the relative abundance of the latter decreased. Similar to the metronidazole-treated mice, the ampicillin-treated mice that were allowed to recover an additional five days before challenge had a significant decrease in colonization (P=0.03). Again, we observed a similar increase and decrease in relative abundance for the *Barnesiella* (OTU 3) and *Escherichia* (OTU 4). These results confirm the results of the titration experiments by suggesting that colonization resistance is context dependent. Specifically, during recovery the relative abundance of *Barnesiella* (OTU 3) and *Turicibacter* (OTU 12) increase and *Escherichia* (OTU 4) decrease with recovery; however in untreated mice, the relative abundance of *Barnesiella* (OTU 3) and *Turicibacter* (OTU 12) are significantly lower [XX check on Turicibacter XX].

***Correlation analysis reveals potentially protective bacteria.*** To identify bacterial taxa that could be associated with resistance or susceptibility to *C. difficile*, we measured the correlation between the relative abundance of each OTU that had an average abundance greater than 0.1% on the day of inoculation with the level of *C. difficile* colonization 24 hours later (Figure 2). OTUs associated with providing resistance against *C. difficile* (N=46) outnumbered those with associated associated with susceptibility (N=6). Among various bacterial families, three were consistently associated with *C. difficile* resistance: *Porphyromonadaceae* (ρaverage=-0.48, N=13 OTUs), *Lachnospiraceae* (ρaverage=-0.4, N=11 OTUs), and *Ruminococcaceae* (ρaverage=-0.38, N=7 OTUs). Two OTUs from the *Proteobacteria* had a significant association with *C. difficile* colonization. These included OTUs associated with the genera *Pseudomonas* (ρ=0.26) and *Escherichia* (ρ=0.51). Overall, these results suggest that several large groups of related OTUs, particularly within the *Firmicutes* and *Bacteroidetes*, may participate in colonization resistance against *C. difficile*, while members of the *Proteobacteria* phylum may be involved in susceptibility.

***The composition of the microbiota on day 0 is predictive of* C. difficile *colonization levels on day 1.***

Correlation analysis suck because it does not take into account the non-linearity and statistical interactions between populations.

The three sets of experiments demonstrated that in certain contexts individual OTUs could be associated with C. difficile colonization, but in other contexts they were not.

This suggests that colonization is a phenotype that is driven by multiple populations that act independently and possibly in concert to resist colonization. Therefore, we used a regression-based random forest machine learning algorithm to predict the level of *C. difficile* colonization observed in the three sets of experiments based on the composition of the microbiota at the time of challenge. The model explained 80.8% of the variation in the observed *C. difficile* colonization levels (Figures 5 and S4). When we only included the top 12 OTUs based on the percent increase in the mean squared error when each OTU was removed, the resulting model explained 81.5% of the variation in the observed *C. difficile* colonization levels. Many of the OTUs that contributed the most to the quality of the fit included members of the *Porphyromonadaceae*, *Alistipes*, *Lachnospiraceae*, *Lactobacillus*, and *Escherichia*. These results further validate the observations from the correlation-based analysis. Together these results suggest that colonization resistance is likely conferred by *Porphyromonadaceae* (OTU 13, 5, 19, 10, 9), *Alistipes* (OTU 11), and *Lachnospiraceae* (OTU 24) and a loss in these bacteria, concurrently with a gain of *Escherichia* (OTU 4) *and Streptococcus* (OTU 113), can result in increased susceptibility to infection (Figure 6). Interestingly, the relationship between these OTUs and colonization resistance was not a simple dichotomous or linear relationship. For example, among the streptomycin-treated mice, which altogether lacked the *Escherichia* (OTU 4), *Porphyromonadaceae* (OTU 13) and *Lachnospiraceae* (OTU 24), were the two most important features in accurately predicting subsequent *C. difficile* levels

### Discussion

Individual probiotics have not worked with complete success [dendukuri 2005]. In fact, several studies using probiotic mixtures of several bacteria find these members often work better together than by themselves [lawley, Buffie CG et al 2015, stein, hickson 2007]. Thus, studying *C. difficile* in the context of the community it is exposed to is essential for building a complete picture of the variables at play in its ability to colonize. One way of capturing these variables is through statistical modeling. Using models of the microbiome, we can not only identify bacteria with positive and negative roles in *C. difficile* colonization but also predict levels of colonization based on the structure of the microbiota that *C. difficile* encountered.

Through our antibiotic perturbation experiments we identified a panel of bacteria with either strong positive or negative relationships with subsequent levels of *C. difficile* colonization. A large number of *Porphyromonadaceae*, *Lachnospiraceae*, *Ruminococcaceae*, other *Clostridia*, and *Alistipes* were found to have strong potential resistance abilities against *C. difficile*. Conversely, several members of the *Proteobacteria* phylum, including *Pseudomonas*, *Streptococcus*, and especially *Escherichia* have strong potential positive influences on *C. difficile* colonization. These results were supported by their significant correlation values across several antibiotic regimens. By either reducing the amount of microbiota perturbation or allowing recovery following perturbation, we were able to capture microbiota structures which had increased resistance to *C. difficile* colonization. Using this collective microbiota data spanning a wide range of colonization resistance levels, we moved beyond correlation to prediction of *C. difficile* colonization based on the composition of the microbiota. Through modeling we identified bacteria within the community which contributed greatest to accurately predicting *C. difficile* levels. This is the first comprehensive analysis both measuring and predicting colonization resistance levels based on the microbiota structure.

Correlation analysis identified major families with significant negative or positive relationships with *C. difficile*. Many of the correlaions we observed were not surprising given other studies looking at the relationship between *C. difficile* and the microbiome. For instance, monocolonization of germfree mice with either a *Lachnospiraceae* species or *E. coli* were tested for their resistance to *C. difficile* colonization [Reeves]. Pre-colonization with *Lachnospiraceae* was able to decrease the levels of *C. difficile*, while *E. coli* pre-colonization and completely germfree mice had similar *C. difficile* levels. Association studies in humans with and without CDI in hospitals have also highlighted several bacteria found in this study, including *Lachnospiraceae*, *Rumicoccaceae*, *Clostridia*, and *Alistipes* as decreased and *E. coli* and *Streptococcus* as increased in subjects with CDI [Schubert, Antharam]. While we observe overlaps in potentially protective bacteria between mice and humans at broad taxonomy levels, there are notable differences at finer taxonomic levels. Two different *Bacteroidales* species, a *Porphyromonadaceae* and a *Bacteroides*, have the strongest negative correlations with *C. difficile* in mice and humans, respectively [Schubert]. *Porphyromonadaceae* is more commonly found in mice than in humans, indicating that there are host specific bacteria in colonization resistance.

Modeling of the microbiome allows us to characterize *C. difficile* colonization in terms of a community, which is necessary to capture these complex relationships among its members. Correlation analysis alone was insufficient in identifying all informative bacteria. Random forest modeling allowed us to identify several bacterial OTUs that did not have significant correlations with *C. difficile* yet improved model accuracy. *Turicibacter* and *Akkermansia*, for example, both had non-significant correlations with *C .difficile* yet were more important for model accuracy over OTUs with stronger correlations (*Figure 5BC*). While these bacteria may not have important direct effects on *C. difficile*, they may interact positively or negatively with bacteria that do have strong relationships with *C. difficile*.

Throughout our analysis, we observed examples of the significance of the microbiome as a whole over any individual bacterial species in contributing to resistance. No single bacteria was able to consistently explain the subsequent level of *C. difficile* observed. For example, the *Porphyromonadaceae* OTU3, along with the *Lachnospiraceae* OTU39, were the two most important OTUs in the model with strong negative correlations with *C. difficile*. The majority of microbiotas which had over 1% relative abundance of OTU3 had no detectable *C. difficile* (*Figure 5C*). However, a small group of microbiotas with greater than 1% OTU3 abundance were not fully protected by the presence of this OTU alone and ranged in *C. difficile* levels from 102 to 106 CFU/g feces. These included several streptomycin (0.1 mg/ml) treated, ampicillin treated (both 1 and 6 days recovery), and cefoperazone treated (both 0.1 mg/ml and 0.3 mg/ml) microbiotas. Similarly, having low amounts of one "protective" bacteria such as *Lachnospriaceae* OTU39 does not automatically render the host susceptible to colonization. For instance streptomycin treated (0.1 mg/ml) mice did not harbor *Lachnospiraceae* OTU39, yet had undetectable levels of *C. difficile*. Thus the loss or absence of one protective bacterial species can be compensated by others, either related species or others with protective capacities.

Furthermore, the dynamics of *C. difficile* colonization does not appear to be influenced solely by the loss of protective species. Both *Escherichia* OTU11 and *Streptococcus* OTU78 were positively correlated with *C. difficile* colonization and important in the random forest model. Notably, *Escherichia* OTU11 was among the strongest positively correlated OTUs with *C. difficile* across the majority of antibiotic treatments. Though *Escherichia* may aid in colonization of *C. difficile*, our data suggests it is not essential. *Escherichia* OTU11 was near absent in streptomycin treated (5 mg/ml) mice, but *C. difficile* colonized to high levels (*Figure 1*). Together These observations lead to a model of colonization by *C. difficile* in which antibiotic perturbation leads to a significant loss of protective species. This in turn leads to a loss of immune homeostasis [Smith et al. 2013 Science], an increase in available resources [Theriot 2014, Ferreyra et al 2014, Ng KM et al 2013, Koenigsknecht et al 2015] and eliminates direct or indirect inhibitory effects of bacterial products made by these protective bacteria (e.g. bacteriocins or short chain fatty acids) [Buffie CG et al 2015, Theriot 2014]. While this may be suitable enough for *C. difficile* to colonize, other opportunistic species such as *Escherichia coli* and other *Proteobacteria* may also take advantage of these microbiota perturbations. The bloom in these bacteria may further change the gut milieu in favor of *C. difficile* colonization. Increased growth by these bacteria can trigger inflammation, which has been shown to be advantageous for *Salmonella tyhpumurium* to further disrupt and suppress the indigenous microbiota [Stecher et al 2007 Plos Biol]. Increased *Proteobacteria* abundances were also found among mice with increased mucus layer penetrability [Jakobsson et al. 2015 EMBO reports].

We have shown across several different types of perturbations that we can accurately predict the resistance levels of these communities to *C. difficile*. This has great applications for predicting patient risk of CDI in the hospital. Furthermore, this analysis can be used to inform the design of personalized medicine, whether through tailoring probiotic or antibiotic regimens. Future studies are needed to examine of the human-specific bacteria in colonization resistance and also to examine the changes in models necessary to account for physiological differences between *C. difficile* strains.

### Materials and Methods

***Animal care.*** We used 5-8 week old C57Bl/6 mice for all of our experiments. These mice were reared under SPF conditions within the animal facility at the University of Michigan. All animal-related protocols and experiments were approved by the University Committee on Use and Care of Animals at the University of Michigan and carried out in accordance with the approved guidelines.

***Antibiotic administration.*** Mice were administered one of seven different antibiotics, including cefoperazone, vancomycin, metronidazole, streptomycin, ciprofloxacin, ampicillin, and clindamycin. The route of administration depended on the specific antibiotic. Cefoperazone (either 0.5 mg/ml, 0.3mg/ml, or 0.1 mg/ml), vancomycin (0.625 mg/ml, 0.3 mg/ml, or 0.1 mg/ml), streptomycin (5 mg/ml, 0.5 mg/ml, or 0.1 mg/ml), metronidazole (0.5 mg/ml), and ampicillin (0.5 mg/ml) were all administered in the mouse drinking water for 5 days. Ciprofloxacin (10 mg/kg) was administered via oral gavage and clindamycin (10 mg/kg) was administered via IP injection. Mice that did not receive antibiotics were used as negative controls for these experiments.

***C. difficile preparation and challenge*** All antibiotic-treated mice were given 24 hours to recover with untreated drinking water prior to *C. difficile* challenge. *C. difficile* strain 630Δerm spores were used in all experiments [cite dena lyras]. Spores were prepared from a single large batch whose concentration was determined within the week prior to each *C. difficile* challenge [cite paul?]. On the day of challenge 10^3 *C. difficile* spores were administered to mice via oral gavage. Immediately following this challenge, the remaining *C. difficile* inoculum was diluted in a series and plated to confirm the correct dosage.

***Sample collection and plating*** Fecal samples were freshly collected for each mouse on the day of *C. difficile* challenge (day 0). On the day after the challenge (day 1), another fecal sample was weighed and diluted under anaerobic conditions with anaerobic PBS. The number of colony forming units (CFU) were counted following 24 hrs growth on TCCFA plates at 37°C under anaerobic conditions [cite plates].

***DNA extraction and sequencing*** Total bacterial DNA was extracted from each day 0 stool sample using the MOBIO PowerSoil®-htp 96 Well Soil DNA Isolation Kit. Sequencing preparation and sequencing was carried out as previously described (Kozich JJ et al. 2013). Briefly, the V4 region of the 16S rRNA gene was amplified, purified, and pooled in equimolar concentrations. These libraries were then mixed with Illumina-prepared PhiX control libraries.

***Sequence curation*** These sequences were curated using mothur as previously described (Kozich JJ et al and Schloss PD et al). Sequences were binned into operational taxonomic units (OTUs) using a 3% dissimilarity cutoff. Taxonomic assignments were determined by using a naïve Bayesian classifier with the Ribosomal Database Project (RDP) training set (version 9) requiring an 80% bootstrap confidence score. In parallel to the fecal samples, we also sequeced a mock community where we knew the true sequence of the 16S rRNA gene sequences. Analysis of the mock community data indicated that the error rate following our cuation procedure was 0.07%. In order to avoid biases due to uneven sampling, samples were normalized to 1,625 sequences per samples [Schloss, reducing effects]. All 16S rRNA gene sequence data and metadata are available through the Sequence Read Archive under accession XXXXXX.

***Statistical analysis and modeling*** All analyses were conducted using R version 3.1.2. OTUs were considered for analysis if their average relative abundance within any treatment group was at least 1% (43 OTUs). Spearman rank correlation analysis was performed with a Benjamini and Hochberg adjustment between OTU counts or diversity levels and *C. difficile* CFU/g feces. Comparison of bacterial levels among titration groups or delayed groups of the same antibiotic was performed using the Kruskall-Wallis rank sum test followed by pairwise Wilcoxon rank sum tests with a Benjamini and Hochberg adjustment. Comparison of log (base 10) transformed *C. difficile* CFU/g feces between experimental groups was calculated using ANOVA and subsequently pairwise T tests with a Benjamini and Hochberg adjustment. Random forest regression models were constructed using the randomForest R package using 5,000 trees [ref]. The regression was performed using the log10 transformation of the number of CFU/g fecal material as the dependent variable and the 43 OTUs as predictor variables. Complete analysis scripts are available at the online repository for this study (https://github.com/SchlossLab/Schubert\_AbxD01\_mBio\_2015).

**Figure Legends**

**Figure 1. Antibiotic treatments result in significant alterations to the structure of the microbiota and variation in colonization resistance.** Bars indicate the median percent relative abundance of those selected OTUs from all treatment groups on the day of *C. difficile* challenge. Stars along the x-axis indicate those OTUs that were significantly different from untreated mice for that antibiotic treatment. The error bars indicate the interquartile range. The median level *C. difficile* colonization found 24 hours post microbiota sampling is indicated in the top right for each treatment with the interquartile range in parentheses.

**Figure 2. Titration of antibiotic perturbations results in altered community structures and *C. difficile* colonization resistance.** Bars indicate the median percent relative abundance of those selected OTUs from all treatment groups on the day of *C. difficile* challenge. Stars along the x-axis indicate those OTUs that varied significantly across doses of the same antibiotic. The error bars indicate the interquartile range. The median level *C. difficile* colonization found 24 hours post microbiota sampling is plotted on the right for each treatment with error bars indicating the interquartile range.

**Figure 3. Increasing the recovery time following antibiotic perturbation restores colonization resistance.** Bars indicate the median percent relative abundance of those selected OTUs from all treatment groups on the day of *C. difficile* challenge. Stars along the x-axis indicate those OTUs that varied significantly between those mice that were allowed 1 or 6 days of recovery. The error bars indicate the interquartile range. The median level *C. difficile* colonization found 24 hours post microbiota sampling is plotted on the right for each recovery period and antibiotic with error bars indicating the interquartile range.

**Figure 3. Broader taxonomic groups share similar relationships with *C. difficile* colonization.** The relative abundance of OTUs found on the day that mice were challenged with *C. difficile* spores were correlated with the level of C. difficile CFU/g feces 24 hours later. Only OTUs with an average abundance of at least 0.1% across all original treatments were considered in the correlation analysis. All correlations shown in this graph are significant after controlling for multiple comparisons using the Benjimani-Hochberg correction (P<0.05). OTUs are grouped by the taxonomic family that its sequences classified as or the next deeper classification that provided a classification. The letters in the parentheses correspond to the phylum that the taxa belongs to. B: Bacteroidetes, F: Firmicutes, P: Proteobacteria, V: Verrucomicrobia, A: Actinobacteria, T: Tenericutes.

**Figure 5. Random forest model to predict *C. difficile* colonization levels based on the microbiota.** (A) The fit of the random forest model by experimental group. The dashed line represents the y=x line. (B) The 15 OTUs with the highest percent increase in mean squared error when dropped from the model. (C) The percent relative abundance of the top 8 OTUs from (B) on day 0 are plotted against the log10 *C. difficile* CFU/g feces.

**Figure S1. Effect of antibiotic perturbations on phylum-level representation of communities on day of *C. difficile* challenge.** Bars depict the median relative abundance across mice within the treatment group and error bars indicate the interquartile range.

**Figure S2. Effect of titrated antibiotic treatments on phylum-level representation of communities on day of *C. difficile* challenge.** Bars depict the median relative abundance across mice within the treatment group and error bars indicate the interquartile range.

**Figure S3. Effect of recovery period following antibiotic treatments on phylum-level representation of communities on day of *C. difficile* challenge.** Bars depict the median relative abundance across mice within the treatment group and error bars indicate the interquartile range.

**Figure S2. Changes between untreated and recovery days.** A) The average thetaYC distance between pretreatment (day -11) and recovery days (days -5 through -1) is represented by circles connected by solid lines. Dashed lines with the same symbol indicate individual mice from the same cage. Antibiotic treatment is indicated by color. B) The average level of the inverse simpson diversity index is plotted by antibiotic treatment across time. Error bars show standard error of the mean. Mice were challenged with *C. difficile* on day 0.