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

**Running title:** Colonization resistance against *C. difficile*

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

Perturbations to the gut microbiota can result in a loss of colonization resistance against gastrointestinal pathogens such as *Clostridium difficile*. Although *C. difficile* infection is commonly associated with antibiotic use, the precise alterations to the microbiota associated with this loss in function are unknown. We used a variety of antibiotic perturbations to generate a diverse array of gut microbiota structures, which were then challenged with *C. difficile* spores. Across these treatments we observed that *C. difficile* resistance was never attributable to a single organism, but rather it was the result of multiple microbiota members interacting in a context-dependent manner. Using relative abundance data, we built a machine learning regression model to predict the levels of *C. difficile* that were found 24 hours after challenging the perturbed communities. This model was able to explain 77.2% of the variation in the observed number of *C. difficile* per gram of feces. This model revealed important bacterial populations within the microbiota, which correlation analysis alone did not detect. Specifically, we observed that populations associated with the *Porphyromonadaceae*, *Lachnospiraceae*, *Lactobacillus*, and *Alistipes* were protective and populations associated with *Escherichia* and *Streptococcus* were associated with high levels of colonization. In addition, a population affiliated with *Akkermansia* indicated a strong context dependency on other members of the microbiota. Together, these results indicate that individual bacterial populations do not drive colonization resistance to *C. difficile*. Rather, multiple diverse assemblages act in concert to mediate colonization resistance.

### Importance

The gastrointestinal tract harbors a complex community of bacteria, known as the microbiota, which plays an integral role preventing its colonization by gut pathogens. This resistance has been shown to be crucial for protection against *C. difficile* infections (CDI), which are the leading source of hospital-acquired infections in the United States. Antibiotics are a major risk factor for acquiring CDI due to their effect on the normal structure of the indigenous gut microbiota. We found that diverse antibiotic perturbations gave rise to altered communities that varied in their susceptibility to *C. difficile* colonization. Instead of being dependent on one specific population of bacteria, we found that multiple co-existing populations conferred resistance. By understanding the relationships between *C. difficile* and members of the microbiota it will be possible to better manage this important infection.

### 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 {Vollaard, 1994 #1}. This native protection by the microbiota from invasive pathogenic species is termed colonization resistance. It is well established that the gut bacterial microbiota is critical in the host's defense against the pathogen *Clostridium difficile* {Chang, 2008 #3;Reeves, 2011 #4;Wilson, 1985 #2}. Perturbations to this indigenous community often lead 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 {Bauer, 2015 #5}. *C. difficile* infections (CDI) are the most reported hospital-acquired infection in the United States and are responsible for 14,000 deaths a year {Lessa, 2015 #6}.

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 unperturbed, untreated murine microbiome is completely resistant to *C. difficile* colonization. It was previously shown that C57Bl/6 mice treated with cefoperazone {Reeves, 2011 #4;Theriot, 2011 #7}, tigecycline {Bassis, 2014 #8}, clindamycin {Buffie, 2012 #9}, or clindamycin in combination with a five antibiotic cocktail {Chang, 2008 #3} had decreased colonization resistance. These studies suggest that a loss of *Lachnospiraceae* and *Barnesiella* and a bloom of *Lactobacillaceae* and *Enterobacteriaceae* are responsible for the loss of colonization resistance. These results are largely supported by human association studies {Schubert, 2014 #11;Vincent, 2013 #10}. We previously observed significant differences between the gut microbiota of hospitalized individuals with and without *C. difficile* and non-hospitalized controls {Schubert, 2014 #11}. In addition, fecal microbiota transplants have been shown to increase the relative abundance of *Bacteroidetes* and decrease the relative abundance of *Proteobacteria* and result in a successful restoration of colonization resistance in patients {Seekatz, 2014 #12}. The mechanisms involved in restoring colonization resistance is not fully understood, but it emphasizes the importance of the gut microbiota in protecting against *C. difficile*.

Because the gut microbiota is a complex community we need tools that enable us to dissect the interactions within the community and with *C. difficile*. One approach is the use of mathematical models to identify associations between members of the microbiota and *C. difficile*. Mathematical models have been used to predict *C. difficile* {Schubert, 2014 #11;Stein, 2013 #13} and *Citrobacter* infection {Belzer, 2014 #14}, colon cancer {Zackular, 2013 #15}, and psoriasis {Statnikov, 2013 #16} based on the composition of the gut microbiota. We similarly sought to identify the subset of the normal murine microbiota that are responsible for colonization resistance by using mathematical models to explain the relationship between members of the gut microbiota.

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, and recovery times we altered the murine gut microbiota and challenged the communities with *C. difficile* spores to quantify 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 regression 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 giving rise to colonization resistance were non-linear and context dependent. These findings show we can successfully apply modeling techniques to accurately predict the colonization resistance 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**). Following the cessation of antibiotics, each treatment group was given one day of recovery prior to challenge with *C. difficile* spores. One day post challenge 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 (**Figures 1 and S1**). We observed the most variance in levels of *C. difficile* colonization in mice treated with ampicillin. None of the mice that received ciprofloxacin were colonized. In addition to administering ciprofloxacin by oral gavage, we provided ciprofloxacin by intraperitoneal injection (10 mg/mL). For both approaches we provided one or two days of recovery from the antibiotic treatment. Regardless of the method, the resulting communities were resistant to *C. difficile* colonization. Only one of six mice receiving vancomycin was colonized with *C. difficile.* We suspected that this was due to residual vancomycin repressing *C. difficile* growth. In fact, two days post *C. difficile* challenge, *C. difficile* bloomed in this treatment group to a median of 9.1x107 (interquartile range 7.6x107–1.1x108) CFU/g feces. Furthermore, given two days of post-vancomycin recovery, there was no delay in *C. difficile* colonization to high levels, and on day one post challenge we observed a median of 3.0x107 (interquartile range 2.6x107–3.6x107, N=4) CFU/g feces. These results suggest that although vancomycin is not absorbed by the gut tissue, the absence of *C. difficile* in the remaining five vancomycin-treated mice may have been due to residual antibiotics lingering in the environment. Overall, the various antibiotic perturbations provided varying levels of colonization by *C. difficile,* which suggested that the resulting communities varied in their composition.

To test this hypothesis, 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. All of the antibiotic treatments, except for the ciprofloxacin-treated mice (AMOVA, P=0.09), resulted in distinct and reproducible changes to the structure of the microbiota relative to the untreated animals (AMOVA, P<0.001). The similarity in the structure of the microbiota in ciprofloxacin-treated and untreated mice suggests that a higher dose of ciprofloxacin may have been necessary to significantly perturb the microbiota to allow *C. difficile* to overcome colonization resistance. Comparisons of the microbiota between antibiotic classes indicated that their structures were significantly different from each other (AMOVA, P<0.03). The community structures of mice receiving beta-lactams (i.e. cefoperazone and ampicillin) were not significantly different from each other (AMOVA, P=0.37). 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. Consequently we expected a greater maintenance of resistance against *C. difficile* colonization in these titrated treatment groups. In addition to the previous treatments, we treated mice with lower concentrations of cefoperazone, streptomycin, and vancomycin (*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 in all mice receiving titrated doses of cefoperazone, *C. difficile* colonization levels decreased significantly (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 3) and a number of rare OTUs. We also observed increases in the relative abundances of OTUs associated with the family *Porphyromonadaceae* (OTU 5, 10, 11, 13, and 21; **Figure 2**). 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, 5, 6, 10, and 11), *Alistipes* (OTU 12), 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 *C. difficile* 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 vancomycin resulted in similarly high levels of *C. difficile* colonization (P=0.96). Seven OTUs were differentially represented across the three doses of vancomycin. Surprisingly, even though the colonization levels of *C. difficile* 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, 5, and 6) 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 in which specific members of the microbiota are found is important in determining the overall resistance to *C. difficile*. For example, the relationship between the *Bacteroides*-affiliated OTU (OTU 1) and *C. difficile* colonization was positive in streptomycin-treated mice and it was negative in cefoperazone-treated mice. In addition, cefoperazone and streptomycin-treated mice had high levels of *C. difficile* although the former had significantly higher levels of an *Escherichia*-affiliated OTU (OTU 3), which were 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 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 {Reeves, 2011 #118; Theriot, 2014 #956}. 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 -fold reduction in colonization (P<0.001; **Figure 3**). In addition, 7 of the 14 mice given the longer recovery period 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 were two OTUs that affiliated with a member of the *Barnesiella* (OTU 2) and the *Escherichia* (OTU 3). The relative abundance of this *Barnesiella*-affiliated OTU increased with the delay, and the relative abundance of this *Escherichia*-affiliated OTU 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 *C. difficile* colonization (P=0.03). As before, we observed a similar increase and decrease in relative abundances for *Barnesiella* (OTU 2) and *Escherichia* (OTU 3)-affiliated OTUs. However untreated, fully resistant mice harbored significantly lower levels of the *Barnesiella*-affiliated OTU (OTU 2). Rather, untreated mice had high levels of various *Porphyromonadaceae*-affiliated OTUs (**Figure 1**). These findings further confirm the context-dependency of colonization resistance suggested by the results of our titration experiments.

**Correlation analysis reveals potentially protective bacteria.** To identify bacterial taxa that could be associated with resistance or susceptibility to *C. difficile* across the three sets of experiments, we measured the correlation between the relative abundance of each OTU on the day of inoculation with the level of *C. difficile* colonization 24 hours later (**Figure 4**). OTUs associated with providing resistance against *C. difficile* (N=22) outnumbered those with associated with susceptibility (N=9). The *Porphyromonadaceae*-affiliated OTUs (ρaverage=-0.52, N=11 OTUs) were consistently associated with low levels of *C. difficile* colonization. Among the three *Proteobacteria*-affiliated OTUs with a significant positive association with *C. difficile* colonization, the strongest was affiliated with the *Escherichia* (OTU 3; ρ=0.54). By performing an OTU-based analysis we were able to observe intra-family and genus differences in association with *C. difficile* colonization. For example, the *Lachnospiraceae* have been associated with protection against *C. difficile*. Although within the *Lachnospiraceae* familiy we observed 3 OTUs that were associated with low levels of *C. difficile* colonization, 1 OTU was associated with high levels of *C. difficile*. In addition we observed 3 significantly correlated *Lactobacillus*-affiliated OTUs (family *Lactobacillaceae*, 2) of which were associated with low levels of *C. difficile* and 1 was associated with high levels. The broad taxonomic representation of OTUs associated with low levels of *C. difficile* again suggests that a diverse community may be advantageous in preventing *C. difficile* colonization.

**The composition of the disturbed gut microbiota is predictive of** ***C. difficile*** **colonization levels.** These three sets of experiments demonstrated that in certain contexts individual OTUs could be associated with *C. difficile* colonization, but in other contexts those OTUs had the opposite or no association. This suggests that colonization is a phenotype that is driven by multiple populations that act independently and possibly in concert to resist colonization. Correlation-based analyses cannot predict these types of context dependencies because they do not take into account the non-linearity and statistical interactions between populations. 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 using the composition of the microbiota at the time of challenge as predictor variables. The model explained 77.2% of the variation in the observed *C. difficile* colonization levels (**Figures 5**). 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 77.1% of the variation in the observed *C. difficile* colonization levels. The OTUs that were ranked as being the most important in defining the random forest model further validated the observations from the correlation-based analysis (**Figures S4**). According to the random forest model, colonization resistance was associated with OTUs that affiliated with the *Porphyromonadaceae* (OTU 15, 10, 6, 18, and 11), *Lachnospiraceae* (OTU 25), *Lactobacillus* (OTU 23), *Alistipes* (OTU 12), and *Turicibacter* (OTU 9; **Figure 6**). A loss in these populations, concurrently with a gain in OTUs affiliated with the *Escherichia* (OTU 3) or *Streptococcus* (OTU 90), was associated with an increased susceptibility to infection (**Figure 6**). As we observed in the titration experiments, the relationship between an *Akkermansia*-affiliated OTU (OTU 4) and *C. difficile* indicated that wide variation in the relative abundance of *Akkermansia* was associated with varying levels of *C. difficile*. There were varying abundances of the *Akkermansia*-affiliated OTU in mice regardless of the level of *C. difficile* colonization. Finally, as indicated by the number of OTUs with relative abundances below the limit of detection, those mice could harbor varying levels of *C. difficile*. These observations bolster the hypothesis that colonization resistance is context dependent.

### Discussion

Previous attempts to study the role of the gut microbiota in colonization resistance against *C. difficile* infection have utilized a single perturbation to the community. Here, we used seven antibiotics from six classes that were given to mice in varying doses and with varying post-antibiotic recovery periods. The result was a combination of 15 different perturbations and the non-perturbed microbiota, which allowed us to generate distinct community profiles that displayed varying susceptibility to *C. difficile* colonization. Our findings demonstrated that colonization resistance was not driven by individual populations, but by a consortium of organisms. Others have demonstrated that *Barnesiella* or *Lachnospiraceae* are partially protective against *C. difficile* {Buffie, 2012 #9;Reeves, 2012 #18}. Although we observed similar results in a subset of our perturbations, by using a large number of perturbations, we were able to demonstrate that a varied collection of populations was important for complete colonization resistance. Thus colonization resistance can be described as an emergent property of the microbiome, in which individual bacterial populations integrated in a community contribute to the overall resistance to *C. difficile* {Novikoff, 1945 #34}.

There is clear need for more efficient therapies for treatment of *C. difficile* infections in humans aimed at restoration of the microbiota. Current first line treatments of CDI include regimens of either metronidazole or vancomycin, which further perturb the microbiota. As such, relapse rates for CDI are typically around 25-30% {Wilcox, 1998 #19}. Interestingly, we observed that treatment with either antibiotic induced susceptibility to *C. difficile* in mice. This result has implications for understanding the causes of recurrent infections. Previous efforts to restore the microbiota and reestablish colonization resistance also support our findings. For instance, association of germ-free mice with a *Lachnospiraceae* isolate only reduced the level of *C. difficile* colonization by 10 to 100-fold {Reeves, 2012 #18}. Using conventional mice, mixtures of bacteria rather than individual populations have been shown to restore colonization resistance and mediate clearance of *C. difficile* {Buffie, 2015 #21;Lawley, 2012 #20}. Fecal transplants, which represent a diverse collection of bacterial populations, have been highly effective in treating humans with recurrent *C. difficile* {Kassam, 2013 #22;Seekatz, 2014 #12;Weingarden, 2015 #23}. By generating a diverse collection of communities that were challenged with *C. difficile*, we have identified a subset of populations using random forest modeling that could be used as a probiotic cocktail to provide colonization resistance. These would include members of the *Porphyromonadaceae*, *Lachnospiraceae*, *Lactobacillus*, and *Alistipes*. Several of these populations have been examined for their potential as a probiotic for preventing *C. difficile* infection. A 6-species cocktail, including isolates of *Porphyromonadaceae*, *Lachnospiraceae*, *Lactobacillus*, *Coriobacteriaceae*, *Staphylococcus*, and *Enterococcus*, successfully resolved CDI in mice {Lawley, 2012 #20}. In humans, *Lactobacillus*-based probiotics have been co-administered with antibiotics to deter the onset of antibiotic-associated diarrhea (AAD) and *C. difficile* infection {Lawley, 2012 #20}. A more diverse probiotic, which contained 33 bacterial species including *Porphyromonadaceae*, *Lachnospiraceae*, *Ruminococcaceae*, *Eubacteriaceae*, and *Lactobacillus* isolates, successfully restored colonization resistance in recurrent *C. difficile* infection and eliminated diarrhea up to 6 months post treatment {Petrof, 2013 #25}. Given this evidence, we feel confident that an effective probiotic mixture could be designed based on our findings to recover colonization resistance against *C. difficile*. Moreover, this line of study will be useful towards the development of personalized treatments based on an individual's specific gut microbiota, which may be a more efficient strategy for preventing and treating CDI. Further examination of the bacterial populations identified in this study is necessary to identify causal relationships and assess the specific mechanisms of colonization resistance. Such research will further advance the development of protocols to prevent and treat CDI.

Random forest regression models allowed us to describe community resistance as a byproduct of an assemblage of bacterial populations rather than as individual populations. A correlation-based analysis was unable to identify populations that had a context dependent or non-linear association with *C. difficile* colonization. Although the murine and human microbiota do not fully overlap, our previous analysis of humans infected with *C. difficile* supports the populations that we associated with colonization {Schubert, 2014 #11}. For instance, *Escherichia* was previously associated with infected individuals and *Lachnospiraceae*, *Ruminococcaceae*, and *Alistipes* were absent from infected individuals. The overlap between the results from the current study and past human studies along with the power of random forest models suggest that it should be possible to model a patient’s risk of developing a *C. difficile* infection based on their gut microbiota composition at admission. As a demonstration of this, we generated a random forest model to predict the binary outcome of whether a mouse would become colonized, regardless of *C. difficile* abundance. Using the same OTUs, we observed an error rate of 10.7%. This suggests that such an approach would be valuable and could perhaps be improved by incorporating other clinical data {Schubert, 2014 #11}. Overall these findings demonstrate the significance of the microbiota as an interconnected bacterial community in assessing resistance to pathogen colonization.

### Materials and Methods

**Animal care.** We used 5-8 week old C57Bl/6 mice obtained from a single breeding colony maintained at the University of Michigan 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 (**Table 1**). The route of administration depended on the specific antibiotic. Cefoperazone (0.5, 0.3, or 0.1 mg/ml), vancomycin (0.625, 0.3, or 0.1 mg/ml), streptomycin (5, 0.5, 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 intraperitoneal injection. Mice that had not received antibiotics were used as negative controls for these experiments. *C. difficile* is unable to colonize mice that are not perturbed by antibiotics.

***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. Spores were prepared from a single large batch whose concentration was determined within the week prior to each *C. difficile* challenge {Sorg, 2009 #26}. Spores were stored long term at 4°C. On the day of challenge 103 *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. On the day after the challenge another fecal sample was weighed and diluted under anaerobic conditions with anaerobic PBS. The number of colony forming units (CFU) were counted following 24 hours growth on TCCFA plates at 37°C under anaerobic conditions {Buggy, 1983 #27}.

**DNA extraction and sequencing.** Total bacterial DNA was extracted from each stool sample collected prior to challenge using the MOBIO PowerSoil®-htp 96 Well Soil DNA Isolation Kit. We generated amplicons of the V4 region within the 16S rRNA gene and sequenced the fragments using an Illumina MiSeq as previously described {Kozich, 2013 #28}.

**Sequence curation.** These sequences were curated using mothur (v.1.35) as previously described {Kozich, 2013 #28;Schloss, 2009 #29}. Briefly, 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 10) requiring an 80% bootstrap confidence score {Wang, 2007 #1809}. In parallel to the fecal samples, we also sequenced 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 curation procedure was 0.02%. All 16S rRNA gene sequence data and metadata are available through the Sequence Read Archive under accession SRP057386.

**Statistical analysis and modeling.** Complete scripts for regenerating our analysis and this paper are available at the online repository for this study (https://github.com/SchlossLab/Schubert\_AbxD01\_mBio\_2015). Comparisons between the antibiotic-treated communities were made by calculating dissimilarity matrices based on the metric of Yue and Clayton {Yue, 2005 #1061}. To avoid biases due to uneven sampling, the dissimilarity matrices were calculated by rarefying the samples to 1,625 sequences per sample. We then used analysis of molecular variance (AMOVA) to test for differences in community structure using 10,000 permutations {Anderson, 2001 #804}. OTU-based analyses were performed using R (v.3.1.2). After subsampling the OTU frequency data to 1,625 sequences per sample, OTUs were considered for analysis if their average relative abundance within any treatment group was at least 1% (N=38 OTUs). All OTU-by-OTU comparisons were performed using the Kruskal-Wallis rank sum test followed by pairwise Wilcoxon rank sum tests. Comparison of log (base 10) transformed *C. difficile* CFU/g feces between experimental groups was calculated using the Kruskal-Wallis rank sum test followed by pairwise Wilcoxon rank sum tests. Spearman rank correlation analysis was performed between OTU counts and *C. difficile* CFU/g feces. All P-values were corrected using a Benjamini and Hochberg adjustment with an experiment-wide Type I error rate of 0.05 {Benjamini, 1995 #604}. Random forest regression models were constructed using the randomForest R package using 10,000 trees {Cutler, 2007 #31}. To construct each tree, two-thirds of the samples are used to train the model and the other one-third is used to test the model. The regression was performed using the log (base 10) transformation of the number of CFU/g fecal material as the dependent variable and the 38 OTUs as predictor variables.

### Acknowledgements

We thank Dr. Vincent Young for providing a critical review of an earlier version of this manuscript. This work was supported by several grants from the National Institutes for Health R01GM099514, R01HG005975, U19AI090871, and P30DK034933. The funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Table 1. Description of Antibiotics used in this study**

| Antibiotic | Administration | Class | Mechanism | Target | |:---------:|:--------------:|:-----:|:---------:|:---------:| | Ampicillin | *Ad libitum* in drinking water, five days | β-lactam: Penicillin | Inhibits peptidoglycan synthesis | Gram +/- | | Cefoperazone | *Ad libitum* in drinking water, five days | β-lactam: Cephalosporin | Inhibits peptidoglycan synthesis | Gram +/- | | Ciprofloxacin | Oral gavage, one time | Fluoroquinolone | Inhibits DNA gyrase | Gram +/- | | Clindamycin | Intraperitoneal injection, one time | Lincosamide | Inhibits protein synthesis | Anaerobes | | Metronidazole | *Ad libitum* in drinking water, five days | Nitromidazole | Destabilizes bacterial DNA | Anaerobes | | Streptomycin | *Ad libitum* in drinking water, five days | Aminoglycoside | Inhibits protein synthesis | Gram +/- | | Vancomycin | *Ad libitum* in drinking water, five days | Glycopeptide | Inhibits peptidoglycan synthesis | Gram + |

### 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 after correcting for multiple comparisons. 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. The dose of antibiotic and the number of animals used in each treatment group is indicated for antibiotic treatment group. The treatment groups are sorted according to level of *C. difficile* colonization.

**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 after correting for multiple comparisons. 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. The number of animals used in each treatment group is indicated in the legend, which depicts the doses of each antibiotic that were used.

**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 after correting for multiple comparisons. 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. The number of mice used in each treatment group is indicated above the *C. difficile* colonization data. The dose of each antibiotic is indicated next to the name of the antibiotic.

**Figure 4. Diverse taxonomic groups are associated with low levels of** ***C. difficile*** **colonization.** Spearman correlation coefficients were calculated using the relative abundance of OTUs found on the day that mice were challenged with *C. difficile* spores and the amount of *C. difficile* observed 24 hours later. Only significant correlations are presented after correting for multiple comparisons. OTUs are grouped by the taxonomic family and the letters in the parentheses correspond to the phylum that the taxa belong to. B: *Bacteroidetes*, F: *Firmicutes*, P: *Proteobacteria*, A: *Actinobacteria*, T: *Tenericutes*.

**Figure 5. Random forest regression model predicts** ***C. difficile*** **colonization levels based on the structure of the microbiota.** The overall model explained 77.2% of the variation in the data. Each pane shows antibiotic treatment groups in color and the other points as gray circles. The panels are sorted by the level of C. difficile colonization when mice were treated with the highest dose.

**Figure 6. Relationship between OTU relative abundance and** ***C. difficile*** **colonization levels indicates non-linearity and context-dependency.** The 12 OTUs that resulted in the greatest change in percent mean squared error when removed from the random forest regression model are shown in each pane and together explain 77.1% of the variation in the data. The Spearman correlation value between that OTUs abundance and *C. difficile* levels are shown for each pane when the corrected P-value was significant. The color and symbols represent the same antibiotic dose and recovery period as in Figure 5.

**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 S4. The change in percent mean squared error when each OTU was removed from the random forest regression model.** The top 12 OTUs with the highest percent increase in mean square error are denoted in red text and are depicted in Figure 6.