*Clostridium difficile* alters the structure and metabolism of distinct cecal microbiomes during initial infection to promote sustained colonization

Matthew L. Jenior1, Jhansi L. Leslie1, Vincent B. Young1,2, and Patrick D. Schloss1\*

**1. Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan.**

**2. Department of Internal Medicine/Infectious Diseases Division, University of Michigan Medical Center, Ann Arbor, Michigan.**

**Corresponding author.** Correspondence to Patrick D. Schloss ([pschloss@umich.edu](mailto:pschloss@umich.edu))

### Abstract

Susceptibility to *Clostridium difficile* infection and disease severity are associated with different antibiotic therapies, which are known to alter the structure and function of the gut microbiota. We hypothesized that the varied pathogenicity of the same strain of *C. difficile* is due to a metabolic interaction between the gut microbiota and *C. difficile*. To test this hypothesis, we applied a multi-omic approach to a mouse model of infection where the murine microbiota were perturbed by pretreating the mice with either cefoperazone, streptomycin, or clindamycin. Within 18 hours of challenging the three groups of mice with *C. difficile*, the pathogen was able to colonize to high levels. Although *C. difficile* persisted for the next 10 days at high levels in cefoperazone and streptomycin-pretreated mice, it did not persist in clindamycin-pretreated mice. Untargeted metabolomic analysis of the cecal community at 18 hours revealed that *C. difficile* infection had significantly larger impacts on the metabolic environment across cefoperazone and streptomycin-pretreated mice. In cefoperazone and streptomycin-pretreated mice, Stickland fermentation byproducts were most affected. In clindamycin-pretreated mice, however, no metabolites were affected. In support of these results, we observed that in the infected communities genes associated with carbon and energy acquisition had altered expression in the presence of *C. difficile*. Furthermore, the largest changes in gene expression occurred among the least abundant species, suggesting that *C. difficile* “attacks the loser”. Overall, our results suggest that *C. difficile* is able to restructure the nutrient-niche landscape in the gut to promote persistent infection.

**Importance**

Over the past decade, *Clostridium difficile* has become the most common cause of hospital-acquired infection in the United States. The gut microbiota of healthy individuals have previously been shown to provide resistance against colonization by the nosocomial pathogen, but perturbations to the microbiota render individuals susceptible. The gut microbiota are also involved in clearing the infection as the community recovers from the antibiotic-induced perturbation. We utilized a mouse model of infection with three separate antibiotic pretreatment regimes to generate perturbed gut microbiomes that allowed for initial *C. difficile* colonization but different levels of long-term persistence. Using an integrative multi-omic approach we found that in conditions where the microbiota cleared *C. difficile*, *C. difficile* did not meaningfully impact the community during the early stages of infection. When the microbiota could not clear the pathogen, *C. difficile* colonization significantly shifted many aspects of the gut ecosystem to promote its own persistence.

### Introduction

One of the many beneficial functions provided by the indigenous gut bacterial community is its ability to protect the host from infection by pathogens (1). This attribute, termed colonization resistance, is one of the main mechanisms that protect healthy individuals from the gastrointestinal pathogen *Clostridium difficile* (2–4). *C. difficile* infection (CDI) is responsible for most cases of antibiotic-associated colitis, a toxin-mediated diarrheal disease that has dramatically increased in prevalence over the last 10 years. There are an estimated 453,000 cases of CDI resulting in 29,000 deaths in the US annually (5). Antibiotics are a major risk factor for CDI and are thought to increase susceptibility to CDI as they disrupt the gut bacterial community structure; however, it is still unclear what specific changes to the microbiota contribute to this susceptibility (6, 7). While most classes of antibiotics have been associated with initial susceptibility to CDI, fluoroquinolones, clindamycin, and cephalosporins are linked to increased risk of recurrent or persistent infection (8–10). This raises questions about the groups of bacteria that are differentially impacted by certain therapies and how these changes effect duration or severity of the infection.

Associations between the membership and functional capacity of the microbiota as measured by the metabolic output suggest that antibiotics increase susceptibility by altering the nutrient milieu in the gut to one that favors *C. difficile* metabolism (11–13). One hypothesis is that *C. difficile* colonization resistance is driven by competition for growth substrates by an intact community of metabolic specialists. This has been supported by animal model experiments over the past several decades (14–16). This line of reasoning has been carried through to the downstream restoration of colonization resistance with the application of fecal microbiota transplant (FMT). Although an individual’s microbiota may not return to its precise original state following FMT, it is hypothesized that the functional capacity of the new microbiota is able to outcompete *C. difficile* for resources and clear the infection (13, 17).

Leveraging distinct antibiotic treatment regimens in a murine model of CDI (18), we and others have shown that *C. difficile* adapts its physiology to the distinct cecal microbiomes that resulted from exposure to antibiotics (18, 19). We went on to show that *C. difficile* appears to adapts portions of its metabolism to fit alternative nutrient niche landscapes presented. As the diet of the mice remained unchanged, changes in the cecal metabolome are likely driven by the intestinal microbiota. Although it has been established that *C. difficile* colonizes these communities effectively, it is unknown whether the differences in the metabolic activity of communities following antibiotic treatment are impacted by *C. difficile* colonization or if they correlate with prolonged infection. Historically, it has been difficult to ascribe specific metabolic contributions to individual taxa within the microbiota during perturbations, especially within the context of a host. To address this limited understanding, we employed an integrative untargeted metabolomic and metagenome-enabled metatransciptomic approach to investigate specific responses to infection of the gut microbiota in a murine model of CDI. This high-dimensional analysis allowed us to not only characterize the metabolic output of the community, but to also identify which subgroups of bacteria were differentially active during mock infection and CDI. Our results supported the hypothesis that CDI was indeed associated with altered community-level gene expression and metabolomic profile of susceptible environments. This effect was significantly more pronounced in communities where *C. difficile* was able to maintain colonization. This work highlights the need for increased appreciation differential, combined effects of antibiotics and CDI on the gut microbiota in order to develop more successful targeted therapies that prevent *C. difficile* colonization and promote its clearance.

### Results

**Distinct antibiotic pretreatments are associated with different patterns of clearance and virulence factor expression.** We previously found that when conventionally-reared mice were pretreated with one of three different antibiotics (streptomycin, cefoperazone, and clindamycin; Table S1) to sensitize the animals to *C. difficile* colonization, each pretreatment was associated with altered patterns of *C. difficile* virulence factor expression (Fig. 1A) (19). Briefly, these drugs were chosen not only for the ability to to reduce *C. difficile* colonization resistance (18), but also for distinct and significant impacts on the structure and diversity of the cecal microbiota (all *p* < 0.001; Fig. S1A), as well as differential patterns of *C. difficile* clearance following initial infection as quantified from stool (Fig. 1B). In each antibiotic pretreatment model, we observed equally high levels of *C. difficile* colonization on the day after infection. Over the subsequent 8 days, *C. difficile* only fell below the limit of detection in the clindamycin-pretreated mice. In contrast, streptomycin and cefopherazone-pretreated mice remained highly colonized over the next 10 days of the experiment. We hypothesized that this difference in persistence occurred because the mice harbored microbiota that varied in the degree to which their niche space overlapped with that of *C. difficile*.

Because *C. difficile* virulence factor expression is regulated by availability of certain nutrients in the environment (20), we compared spore and toxin production in the different pretreatment groups. We chose to focus our analyses at 18-hours post-infection to assess the behavior of *C. difficile* directly prior to the reduction in detectable *C. difficile*. This time point corresponded with a previous study where *C. difficile* reached maximum cecal vegetative cell load with few detectable spores (21). We also elected to examine cecal content because it was more likely to be a site of active bacterial metabolism compared to stool and would allow for an assessment of functional differences in the microbiota. At 18 hours after infection, there were no significant differences in the number of vegetative cells between any antibiotic-pretreatment group. All susceptible mice harbored ~1x108 vegetative colony forming units (cfu) per gram of cecal content. Untreated mice maintained *C. difficile* colonization resistance (Fig. 1C). We also measured both spore production and toxin activity as activation of both processes has been linked to environmental concentrations of specific growth nutrients (20). Despite having similar vegetative *C. difficile* loads, varying levels of both cfu due to spores and toxin titer were observed across each of the antibiotic pretreatments prior to the onset of detectable inflammation in the epithelium (21). These results suggested that despite high initial *C. difficile* colonization, the microbiomes across pretreatments may vary in available nutrients or profiles of competitors for those niches.

***C. difficile*** **colonization corresponded with significantly altered cecal community structures.** It has not yet been established whether *C. difficile* colonization impacts the structure of the gut microbiota during infection in susceptible mice. To evaluate changes in bacterial population abundances in response to antibiotic perturbation and *C. difficile* colonization, we sequenced the V4 region of the 16S rRNA gene from the cecal content of both mock and *C. difficile*-infected mice across antibiotic pretreatment groups. To focus our analysis specifically on the surrounding bacterial communities we removed all *C. difficile* 16S rRNA gene sequences, which represented less than 3.54% of sequencing reads in any infection group. We confirmed that each antibiotic pretreatment significantly impacted both cecal community structure (Bray-Curtis distances) and diversity (inverse-Simpson) compared to untreated control mice (all *p* < 0.001, Fig. S1A & S1B). We then confirmed that each antibiotic pretreatment induced distinct and significant shifts in the community structure prior to being challenged with *C. difficile* (all *p* < 0.001, Fig. S1C). The composition of streptomycin-pretreated communities was more variable between cages, but was generally enriched for members of phylum *Bacteroidetes* (Fig. 2A). Cefoperazone and clindamycin-pretreated cecal communities were consistently dominated by members of the *Lactobacillaceae* and *Enterobacteriaceae* families, respectively (Fig. 2A). Despite variation in the community structures generated by streptomycin pretreatment, those communities were colonized by *C. difficile* to the same level (Fig. 1B, 1C, & S1).

Next, we measured the effect that *C. difficile* colonization had on overall community structure and composition within each of the antibiotic groups individually. We found that the structure of both streptomycin and clindamycin-pretreated infected communities were significantly different from their corresponding uninfected controls (streptomycin *p* = 0.014, Fig. S1D; clindamycin *p* = 0.003, Fig. S1F). Cefoperazone-pretreated communities did not significantly differ from their control group (Fig. S1E). Utilizing LEfSe differential abundant feature detection (22), we identified 16 OTUs that discriminated between infected and uninfected communities in at least one antibiotic pretreatment (Fig. 2BCD); however, these OTUs were generally near the limit of detection and very few were present across multiple groups tested. This suggested that CDI susceptible communities were only modestly impacted by the presence of *C. difficile*.

**Distinct antibiotic classes lead to enrichment of alternative *C. difficile* nutrient niches.** Pretreatment with antibiotics not only alters the structure of the resident microbiota, but also has a dramatic impact on the intestinal metabolome (11–13). To test the metabolic features of the altered communities, we performed untargeted metabolomic analysis on the cecal contents that were also utilized in the 16S rRNA gene sequencing. A total of 727 distinct metabolites were identified. First, we characterized the differences between the metabolomes of the mock-infected communities to measure the impact of antibiotic pretreatment that generated *C. difficile*-susceptible conditions (Fig. 3). All of the antibiotic pretreatments significantly altered the cecal metabolome compared to untreated, *C. difficile*-resistant mice (*p* < 0.001; Fig. 3A). As with the differences between resistant and susceptible states, the patterns of specific metabolite concentrations diverged depending on the antibiotic pretreatment (Fig. 3). These results demonstrated that each pretreatment lead to susceptible communities with unique metabolomes.

We were interested in identifying those metabolites with the highest degree of change between resistant and susceptible groups. Due to the large number of individual metabolites with significant differences in relative concentration between groups, we employed the Random Forest machine-learning algorithm (23) to highlight those metabolites that most easily differentiated the antibiotic groups in the absence of infection. To further limit the analysis to only the most informative metabolites, we ranked metabolites in order of highest mean decrease accuracy (MDA; i.e. the amount in which their removal negatively impacts correct sample classification) and reported the top 7 metabolites for each resistant/susceptible comparison (Fig. 3C). These lists primarily included precursors or intermediates of polysaccharide fermentation such as chiro−inositol, malonate, erythritol, 4−guanidinobutanoate, and soyasaponin II, the majority of which were increased in susceptible conditions. Decreases in these metabolites, *in vivo*, have previously been associated with a reduction in the normal levels of polysaccharide fermentation metabolites that were present in the intact microbiota (12). N−methylpipecolate was the only metabolite we identified that was consistently lower following antibiotic pretreatment. This molecule is a by-product of amino acid catabolism and is typically oxidized under normal conditions into glycine (24). These data support the hypothesis that amino acid catabolism is disrupted at a community-level following antibiotic pretreatment and may suggest open nutrient niches for *C. difficile*, a known fermenter of peptides (25). We then applied this approach to identify metabolites that differed between antibiotic pretreatment groups (Fig. 3D). Interestingly, several additional carbohydrate or amino acid catabolism by-products were identified including hydroxyisocaproate, methylvalerate, glutamylmethionine, and N−carbamoylaspartate (12). Additionally, sucrose is a *C. difficile* growth substrate (26), which provided further support for the hypothesis that nutrient niches were vacated due to antibiotic pretreatment that were then accessible to *C. difficile* (Fig. 2D). These data suggested that the populations responsible for fermentation of carbohydrates into short-chain fatty acids (SCFAs) may be differentially affected by different classes of antibiotics. The distinct antibiotics permitted survival of separate bacterial groups, likely with varied nutrient catabolism strategies, that each may compete with *C. difficile* along different axes.

Following this unsupervised approach for identification of susceptibility markers, we also measured the differences in relative concentration of metabolites previously connected to potential *C. difficile* colonization susceptibility (13). We assessed differences in bile acids, which are connected to the germination-sporulation life cycle of *C. difficile* (27) and whose bioconversion by the microbiota has been implicated as a driver of colonization resistance (28). We found that there were no persistently increased bile acids across the antibiotic pretreatments (Fig. S2A). There were also no amino acids that were always highly available (Fig. S2B). These results suggested that despite varying efficiencies, *C. difficile* can recognize a subset of these molecules in various contexts leading to germination and outgrowth. In contrast, we found that several carbohydrates were significantly increased across all antibiotic groups including arabitol/xylitol, ribitol, and sucrose (Fig. S2B). Together, our results supported that each susceptible environment was distinguishable from the others, with its own subset of enriched *C. difficile* growth substrates. This could be an indication that particular competitors were eliminated during antibiotic pretreatment or those community members normally responsible for the consumption of these metabolites have altered their metabolic program to exploit alternative nutrient sources.

***C. difficile*** **colonization induced distinct metabolomic changes of each pretreatment group.** Given the changes we observed to the cecal metabolome in response to antibiotic pretreatment, we assessed the degree to which *C. difficile* colonization altered the cecal metabolome of susceptible animals. We hypothesized that the introduction of a new competitor, *C. difficile*, would impact the metabolome either through signatures of its own metabolism or by causing a shift in the metabolism of other members in the community. First, in a similar approach to identifying susceptibility markers, we calculated the Bray-Curtis dissimilarities between the mock- and *C. difficile*-infected metabolomes within each antibiotic pretreatment group. We compared the differences in the structures of the metabolomes within each individual antibiotic group and found that the *C. difficile*-infected metabolomes of both streptomycin- (*p* = 0.039) and cefoperazone- (*p* = 0.016) pretreated mice deviated significantly from that of mock-infected mice (Fig. 4A-4B). However, the metabolomes of clindamycin-pretreated mice were not significantly altered by the presence of the pathogen (*p* = 0.127; Fig. 4C). These results differed from what we observed when analyzing the OTU relative abundance data. Based on the 16S rRNA gene sequence data, the cecal community structure of mice pretreated with cefoperazone was unchanged and clindamycin- and streptomycin-pretreated mice were significantly different when challenged by *C. difficile* (*p* = 0.003; Fig. 4, S1D). These data indicated that large shifts of populations in the cecal microbiota did not result in concordant shifts in the metabolome. This supported the hypothesis that divergent community structures can ultimately share a convergent metabolic output despite changes to community structure and membership.

We then sought to identify changes in those metabolites that were potential markers of *C. difficile* infection through a unified analysis across all pretreatment groups. To accomplish this, we again applied the Random Forest machine-learning approach to distinguish mock- and *C. difficile*-infected mice from each other. Through this modeling approach, we identified the 10 metabolites with the largest MDA (Fig. S3B). The strongest single predictor and only metabolite among the top 10 that was increased during infection was 5-aminovalerate. Presence of 5-aminovalerate appears to be dependent on bacterial processes as relative concentrations of the molecule were significantly lower in ceca of antibiotic-pretreated, mock-infected animals (Fig. 4G). Conversely, it was found to be highly abundant in most infected animals compared to both susceptible and resistant controls. Only in clindamycin-pretreated mice did the concentration of 5-aminovalerate remain near undetectable in a subset of samples. This amino acid analog is a known byproduct of D-proline fermentation in *C. difficile* (29). *C. difficile* is able to catabolize proline along with glycine through a set of paired biochemical reactions known as Stickland fermentation (30). Along these lines, proline and 3 proline-conjugated amino acids (i.e. pro-hydroxy-pro, thioproline, and trans-4-hydroxyproline) were significantly decreased during infection in all pretreatment groups (Fig. S3B; all *p* < 0.001). Combined, these results demonstrated that while each of the pretreatment groups had unique metabolomic profiles, Stickland fermentation products were shared across each group indicating that it is a preferred energy acquisition pathway for *C. difficile*. Random Forest machine-learning models were able to reliably classify infected and uninfected mice in both streptomycin and cefoperazone pretreatments (Fig. 4AB, Out-of-bag error < 10%). However, among the clindamycin-pretreated mice, it was not possible to reliably discriminate between mock- and *C. difficile*-infected mice (Fig. 4C, Out-of-bag error = 44.44%). Among the top 5 MDA metabolites in each of the models discriminating between mock- and *C. difficile*-infected mice, none were shared across all antibiotic-pretreatment groups. These data supported our hypothesis that not only does each antibiotic pretreatment create different community structures, but that the metabolism of these distinct communities responds to *C. difficile* colonization in a unique manner.

**Infection corresponded with large shifts in metatranscriptomic expression that strongly differ across antibiotic pretreatments.** Combined, 16S rRNA gene sequencing and metabolomics results demonstrated that antibiotic pretreatments resulted in distinct bacterial communities that led to altered metabolite profiles. It is also possible that metabolism of the host or *C. difficile* contributed to the differences in the metabolomes during infection. To gain a more specific understanding of how the microbiota or *C. difficile* shaped the metabolic environment, we employed a metagenome-enabled metatranscriptomic shotgun sequencing approach with paired DNA and RNA samples collected from the cecal content of the mice used in the previous analyses. To achieve usable concentrations of bacterial RNA and obtain informative depths of sequencing that adequately sampled the diversity in metagenomic and metatranscriptomic libraries (Fig. S6EF), we were required to pool the limited amount of RNA from each treatment and infection group. We calculated the within-group sample variance replicates from cfu, otu relative abundance, and untargeted metabolomic to establish confidence in continuing pooled analysis (Table S5). This analysis revealed limited amounts of variation at multiple levels of biology.

Following sequencing, metagenomic reads from mock-infected cecal communities were assembled *de novo* into contigs and we identified 234,868 (streptomycin), 83,534 (cefoperazone), and 35,681 (clindamycin) genes within each metagenome. Of these genes, 28.5% were annotated to a known KEGG function and many of these annotations were homologous among species across the dataset. Streptomycin pretreatment resulted in a community that had significantly more taxonomic diversity than the other pretreatment groups (Fig. S1A). Therefore, we expected a more diverse metagenome. In fact, 2408 unique functionally annotated genes were detected in the metagenome resulting from streptomycin pretreatment, at least 1163 more genes than either cefoperazone or clindamycin (Fig. S5). Also, as expected genes with any detectable transcript in any metatranscriptome were a subset of their corresponding metagenome. Metatranscriptomic read abundances were normalized to the corresponding coverage per gene in the metagenome after normalizing the number of reads across metagenomic sequence collections. To highlight the largest magnitude of difference in transcription, we identified the top 15% genes (Fig. 5A) or pathways (Fig. 5B) in each antibiotic pretreatment group with the highest disparity in normalized cDNA abundance between mock and *C. difficile*-infected groups and cross-referenced the pretreatments. Based on the metabolomic results we hypothesized that pathways with the greatest differences would be related to catabolism of metabolites in established *C. difficile* nutrient niches.

<<<<<<< HEAD This analysis revealed modest conservation of transcriptional changes induced by *C. difficile* colonization across the antibiotic pretreatments. This trend was partially driven by significant differences in community membership (Fig. 2A), and consequently the metagenomic repertoire available for transcription. The most stark finding was that only 8 genes were among those shared across pretreatment groups (Fig. 5A). Although mice receiving the streptomycin and clindamycin pretreatments shared a large number of altered genes, cefoperazone-pretreated mice shared relatively few with the metatranscriptome of the other pretreament groups. This result was likely due to large differences in community compositions. Upon inspection of 8 genes that were shared across the three pretreatment groups, a Dps family protein gene was the most varied in transcription. The protein product of this gene is associated with DNA protection and iron binding under periods of environmental stress (31). Furthermore, formate acetyltransferase and phosphopyruvate hydratase are highly varied in their transcription and are central nodes of metabolism, which lend themselves to the hypothesis that the microbiota must process energy sources differently in the presence of *C. difficile*. Pathway-level analysis revealed a larger proportion of overlap between pretreatments than at the gene-level (Fig. 5B). Among these differences, the greatest degree of change was in ABC transporter transcription further supported large-scale shifts in nutrient import across the individual communities. More to this point, numerous genes for transport of simple carbohydrates and glycolysis were also differentially overrepresented in infected conditions across pretreatment groups when compared to mock-infection mice (Fig. S4). Among other pathways in the shared differences grouping, the microbial metabolism in diverse environments KEGG category is a broad classification that includes genes for not only carbohydrate and amino acid catabolism, but also the degradation and metabolism of xenobiotics. Inclusion of this category supports the hypothesis that the microbiota may be restructuring energy acquisition strategies or dealing with additional metabolites introduced due to pathogen colonization. Parallel to these results, large changes in oxidative phosphorylation as well as starch/sucrose metabolism also indicate modification of nutrient niche segregation between mock and infected states that may drive the ability of *C. difficile* to persist. Together, these shifts demonstrated that *C. difficile* colonization is associated with substantial changes in the transcription of nutrient acquisition pathways, which vary greatly between antibiotic-induced susceptible states. ======= This analysis revealed low levels of conservation of transcriptional changes between the antibiotic pretreatments induced by *C. difficile* colonization. This trend is at least partially driven by significant differences in community membership (Fig. 2A), and consequently metagenomic content repertoire available for transcription. The most stark finding was how few genes were among those shared as highly differentially expressed between all pretreatments (Fig. 5A). Interestingly, streptomycin and clindamycin pretreatments shared a large number of altered genes, while cefoperazone-pretreatment shared relatively few with either other susceptible condition. These were again likely due to large differences in community compositions. Upon inspection of those 8 shared genes, a Dps family protein gene is the most varied in transcription among them. The protein product of this gene is associated with DNA protection and iron binding under periods of environmental stress (31). Furthermore, format acetyltransferase and phosphopyruvate hydratase are highly varied in their transcription and are central nodes of metabolism, which lend themselves to the hypothesis that the microbiota must process energy sources differently in the presence of *C. difficile*. Pathway-level analysis revealed a larger proportion of overlap between pretreatments than at the gene-level (Fig. 5B). Among these differences, the greatest degree of change was in ABC transporter transcription further supported large-scale shifts in nutrient import across the individual communities. More to this point, numerous genes for nutrient transport or utilization were differentially overrepresented between infected and mock-infected conditions across pretreatment groups when compared to untreated/resistant communities (Fig. S4). Following up on other pathways in the shared differences grouping, the KEGG category microbial metabolism in diverse environments is a broad classification that includes genes for not only carbohydrate and amino acid catabolism, but also the degradation and metabolism of xenobiotics. Inclusion of this category supports that the microbiota may be restructuring energy acquisition strategies or dealing with additional metabolites introduced due to pathogen colonization. Parallel to these results, large changes in oxidative phosphorylation as well as starch/sucrose metabolism also indicate modification of nutrient niche segregation between mock and infected states that may drive the ability of *C. difficile* to colonize long-term. Altogether, these shifts supported that *C. difficile* colonization is associated with substantial changes in the transcription of nutrient acquisition pathways in the microbiota, which vary greatly between antibiotic-induced susceptible states. >>>>>>> upstream/lab\_computer

***C. difficile*** **colonization disproportionately impacted the gene transcription of low-abundance bacterial taxa.** As not all bacterial taxa share identical metabolic capabilities, we hypothesized that expression in specific subsets of bacteria were differentially affected by the presence or metabolic activity of *C. difficile*. We sought to delineate the transcriptomic contributions of separate bacterial taxa within each metatranscriptomic dataset. To accomplish this we utilized the genus level taxonomic information associated with each KEGG annotation to identify which group likely contributed a given gene to the metagenome. Many genes in the KEGG database are annotated as hypothetical or uncharacterized but still possess a taxonomic annotation. This resulted in substantially more genes from the total being conserved for analysis in each group. With these data, we narrowed the focus to transcription for genera that represented >0.01% of genes receiving taxonomic annotations in any of the metagenomic assemblies. Final metatranscriptomic abundances were then obtained by dividing the cDNA totals of each gene by its metagenomic coverage in order to normalize for abundance of the contributing bacterial taxa. We then directly compared the normalized cDNA abundances for each gene between infected and uninfected states for each antibiotic pretreatment and applied Spearman correlation to identify distinct patterns of expression (Fig. 6A-C). This resulted in 2473 genes that had an average distance of 2.545 units of deviation (UD) associated with streptomycin-pretreatment, 2930 genes at an average distance 3.854 UD with cefoperazone-pretreatment, and only 727 outliers at an average distance of 2.414 UD in clindamycin-pretreatment. Overall, the clindamycin pretreatment was associated with the fewest expression outliers between uninfected and infection conditions compared with those of the other antibiotic groups. This indicated that the gene expression in the clindamycin-pretreated community was not significantly altered in the presence of *C. difficile*.

Since many genes lack a specific functional annotation in KEGG but do map to individual bacterial species, we continued the analysis by measuring change in transcript abundances contributed by the more prevalent genera. Incorporation of the genus-level taxonomic information for each transcript revealed that outlier genes were contributed by underrepresented genera. In streptomycin-pretreated mice, the most prominent differences were in 937 genes belonging to *Lactobacillus* with increased expression during *C. difficile* infection (Fig. 6A). In cefoperazone-pretreated mice, 2290 genes belonging to *Bacteroides* had higher transcription in mock infected mice (Fig. 6B). A consistent trend in streptomycin and cefoperazone pretreatments was an overrepresentation of highly-expressed genes from genera belonging to *Bacteroidetes* during mock infection. The metatransciptomes within both of these pretreatment conditions poorly correlated between mock and infected conditions, indicating a high degree of change induced by *C. difficile* colonization (*ρ* = 0.03 & *ρ* = 0.03). In clindamycin-pretreated mice, the largest difference in transcription was for 510 *Lactobacillus* genes with increased transcription during colonization (Fig. 6C). Mock- and *C. difficile*-infected metatranscriptomes in clindamycin-pretreated mice were strongly correlated with each other (*ρ* = 0.86). This suggested that the degree to which the metatranscriptome was shifted by infection corresponded to prolonged colonization. We further characterized this by tabulating the absolute difference between mock- and *C. difficile*-infected communities transcriptomic abundances per genus in each antibiotic pretreatment, normalizing for the number of genes detected in each genera. Taxa were then separated into groups based on their relative abundance in each community based on 16S rRNA gene sequencing (Fig. 6D). Most changes occurred among the relatively rare taxa. As the number and diversity of low abundance taxa was much larger than that of high abundance taxa, the majority of unique genes and metabolic potential was expressed from the rare taxa.

### Discussion

Our results demonstrate that distinct intestinal ecosystems are differently impacted by *C. difficile* colonization and that these changes to community metabolism could have implications on the ability of the pathogen to persist in those environments. Furthermore, discordant modifications shortly after infection to OTU abundances, metabolome composition, and metatranscriptomic activity suggest that *C. difficile* manipulates the niche landscape of the intestinal tract in order to better maintain long-term colonization even prior to the onset of toxin-mediated inflammation. This hypothesis was supported by the disparity of community-level phenotypes between the clindamycin and cefoperazone or streptomycin pretreatment groups. Only in the clindamycin-pretreatment group were animals able to go on to clear *C. difficile*. Unlike the other conditions, clindamycin-pretreated communities were significantly altered in community structure following *C. difficile* colonization, but not at the metabolomic level (Fig. S1F & 4A). This disconnect between community structure and metabolic output was explained by their similarity in gene expression (Fig. 4D & 5C). Collectively, these results advance the idea that for *C. difficile* to maintain colonization it must exploit and restructure the available niche space.

Instances of active nutrient niche restructuring in the gut have been documented for prominent symbiotic bacterial species in gnotobiotic mice (32), but never before in a conventionally-reared animal model of infection following antibiotic pretreatment. Interestingly, the taxonomic groups that were represented as outliers in the normalized metatranscriptomes of each antibiotic pretreatments were rare in their cecal community (on average less than 5% of the community; Fig. 2A & 6). Previous studies have found that rare taxonomic groups, even those at a low abundance as a result of a spontaneous perturbation, may have disproportion effects on the metabolome (33). For example, this strategy has been observed in temperate lakes where conditionally rare microbes were found to be far more metabolically active than highly abundant taxa (34). This would also be likely to apply to bacterial groups recovering population density following cessation of antibiotic pretreatment. As such, *C. difficile* may outcompete these organisms to ultimately affect greater change to the entire ecosystem and open a long-lasting nutrient niche. While this hypothesis requires further exploration, it provides an ecological framework to study the interactions between *C. difficile* and members of susceptible communities.

This study is one of the first to provide *in vivo* evidence that a medically relevant bacterial pathogen may alter the metabolic activity of a host-associated community to promote its own colonization. This is also the first application of metatranscriptomic analysis of the gut microbiota *in vivo* and in response to a pathogen. Other groups have identified potential metabolite markers of *C. difficile* infection in patient feces (35), but they were not able to make correlations with changes in community metabolism that were afforded to us by paired untargeted metabolomic analysis and metatranscriptomic sequencing. In a recent study, a tick-vectored bacterial pathogen altered the ability of the resident microbiota of the tick by interrupting proper biofilm formation and allowing lasting colonization (36). Furthermore, it was recently found that bacterial metabolic generalists may be more likely to actively antagonize the growth of other species in an environment which they are colonizing (37). We previously found that *C. difficile* has a wide nutrient niche-space *in vivo* and is most likely utilizes its role as a metabolic generalist to colonize diverse gut microbiomes (19). The ability to simultaneously antagonize the metabolism of surrounding species in cecal environments that support persistence would explain the more significant shifts in metatranscriptomic expression. Although we acknowledge that this study may not elucidate the specific mechanism by which this interaction occurs, the combined systems analysis strengthens each individual level of observation and only when employed together does a clearer definition of *C. difficile*-related microbial ecology in the gut emerge. This research lays the groundwork for a more rationale consideration of the metabolic functionalities of bacterial taxa to consider when attempting to rebuild *C. difficile* colonization resistance across differentially perturbed gut environments.

In spite of consistent signals across multiple levels of -omics datasets, possible shortcomings to our interpretation of the presented data do exist. First, as with all transcriptomic studies, the relative level of RNA detected for a given gene does not necessarily reflect the amount of functional protein made by a cell. Furthermore, the timing of translation after transcription may also confound our results and our analysis clearly does not capture the possibility that post-translational modifications may be required to activate some enzymes. Additionally, the low relative abundance of *C. difficile* required us to focus our efforts on greater sequencing depth rather than breadth across multiple replicates within each pretreatment group. Greater transcript abundance would improve our ability to quantify the activity of low abundance species. Although the lack of animal-based replication for the metatranscriptomic data does potentially limit the ability to generalize our results, this approach has been successfully utilized by numerous groups in the past to accurately characterize transcriptionally activity across communities of bacteria (19, 38–40). In terms of metabolomics, alternative possible interpretations of the data also exist. For example, we considered metabolites that did not change in concentration between uninfected and infected conditions as not impacted by changes in bacterial metabolism induced by *C. difficile* colonization. However, this may instead indicate that the metabolism of *C. difficile* itself may simply replace the level present in the uninfected community. Such instances may not be detectable through untargeted mass spectrometry alone, however the combination of methods utilized here present a much more unified description of the system than any of the component techniques alone.

Several groups have attempted to identify single bacterial species or limited strain consortia that are able to provide colonization resistance, but each has been met with incomplete restoration of colonization resistance or function (28, 41–43). The effect of *C. difficile* colonization on community metabolic activity could also be linked to pathogen strain and may offer an explanation to the propensity of some strains to persist over others (44). Moreover, our work contributes to the existing concept that the healthy gut microbiota maintains colonization resistance to *C. difficile* by outcompeting the pathogen for preferred nutrient niche space. Ultimately, our results suggest that each susceptible and subsequently infected microbiome may be unique and will require specific microbes or functionalities to restore colonization resistance against *C. difficile* in that specific context. Conversely, colonization resistance against *C. difficile* may be the result of contributions by distinct subcommunities of bacteria across each unique resistant gut community. As the microbiome is so intimately connected to colonization resistance against the bacterium, it has become imperative to understand what factors allow some gut environments to be persistently colonized while others are not. This research lays the groundwork for future studies to assess context dependent restoration of *C. difficile* colonization resistance and what factors are able to interfere with the ability of *C. difficile* to modify gut ecology in order to promote clearance.

### Materials and Methods

**Animal care and antibiotic administration.** Our murine model of antibiotic-enabled *C. difficile* colonization has been described elsewhere (19). Briefly, approximately equal numbers of male and female conventionally-reared six-to-eight week-old C57BL/6 mice were randomly assigned to each experimental group. Nine mice were used in each experimental and control group. They were administered one of three antibiotics: cefoperazone, streptomycin, or clindamycin before oral *C. difficile* infection (Table S1, Fig. 1A). All animal protocols 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 from the Office of Laboratory Animal Welfare (OLAW), United States Department of Agriculture (USDA) registration, and the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). The protocol license Institutional Animal Care and Use Committee (IACUC) number for all described experiments is PRO00006983.

***C. difficile*** **infection and necropsy.** On the day of challenge, 1x103 *C. difficile* str. 630 spores were administered to mice via oral gavage in phosphate-buffered saline (PBS) vehicle. Mock-infected animals were given an oral gavage of 100 ul PBS at the same time as those mice administered *C. difficile* spores. 18 hours following infection, mice were euthanized by CO2 asphyxiation and necropsied to obtain the cecal contents. Aliquots were immediately flash frozen for later DNA extraction and toxin titer analysis. A third aliquot was transferred to an anaerobic chamber for quantification of *C. difficile* abundance. The remaining content in the ceca was mixed in a stainless steel mortar housed in a dry ice and ethanol bath. Cecal contents from all mice within each pretreatment group were pooled into the mortar prior to grinding to a fine powder. The ground content was then stored at -80°C for subsequent RNA extraction. For 10-day colonization studies, fresh stool was collected from infected mice each day beginning on the day of infection. Mice were monitored for overt signs of disease and were euthanized after the final stool collection.

***C. difficile*** **cultivation and quantification.** Cecal samples were weighed and serially diluted under anaerobic conditions with anaerobic PBS. Differential plating was performed to quantify both *C. difficile* spores and vegetative cells by plating diluted samples on CCFAE plates (fructose agar plus cycloserine, cefoxitin, and erythromycin) at 37°C for 24 hours under anaerobic conditions (45). In parallel, undiluted samples were heated at 60°C for 30 minutes to eliminate vegetative cells and enrich for spores (27). These samples were serially diluted under anaerobic conditions in anaerobic PBS and plated on CCFAE with taurocholate at 37°C for 24 hours. Plating was simultaneously done for heated samples on CCFAE to ensure all vegetative cells had been eliminated. Quantification of *C. difficile* cfu for the 10-day colonization experiments was performed from stool using TCCFAE to measure total *C. difficile* load in these animals over time.

***C. difficile*** **toxin titer assay.** To quantify the titer of toxin in the cecum a Vero cell (ATCC CCL-81) rounding assay was performed (19, 46). Briefly, filtered-sterilized cecal content was serially diluted in PBS and added to Vero cells in a 96-well plate where it was then incubated at 37°C for 24 hours. Plates were then inspected for evidence of cell rounding.

**DNA/RNA extraction and sequencing library preparation.** DNA for shotgun metagenomic and 16S rRNA gene sequencing was extracted from approximately 50 mg of cecal content from each mouse using an epMotion 5075 automated pipetting system (Eppendorf) with the the PowerSoil-htp 96 Well Soil DNA isolation kit (MO BIO Laboratories). The V4 region of the bacterial 16S rRNA gene was amplified using custom barcoded primers (47). Equal molar ratios of raw isolated DNA within each treatment group were then pooled and ~2.5 ng of material was used to generate shotgun libraries with a modified 10-cycle Nextera XT genomic library construction protocol (Illumina). This was done to mimic the pooling strategy necessary for metatranscriptomic library preparation. Final libraries were pooled at equal molar ratios and stored at -20°C. Previously described methods were used to extract RNA from cecal contents (19). Briefly, immediately before RNA extraction, 3 mL of lysis buffer (2% SDS, 16 mM EDTA and 200 mM NaCl) contained in a 50-mL polypropylene conical tube was heated for 5 minutes in a boiling water bath (48). The hot lysis buffer was added to the frozen and ground cecal content. The mixture was boiled with periodic vortexing for another 5 minutes. After boiling, an equal volume of 37°C acid phenol/chloroform was added to the cecal content lysate and incubated at 37°C for 10 minutes with periodic vortexing. The mixture was the centrifuged at 2,500 x g at 4°C for 15 minutes. The aqueous phase was then transferred to a sterile tube and an equal volume of acid phenol/chloroform was added. This mixture was vortexed and centrifuged at 2,500 x g at 4°C for 5 minutes. The process was repeated until aqueous phase was clear. The last extraction was performed with chloroform/isoamyl alcohol to remove acid phenol. An equal volume of isopropanol was added and the extracted nucleic acid was incubated overnight at -20°C. The following day the sample was centrifuged at 12000 x g at 4°C for 45 minutes. The pellet was washed with 0°C 100% ethanol and resuspended in 200 ul of RNase-free water. Following the manufacturer’s protocol, samples were then treated with 2 ul of Turbo DNase for 30 minutes at 37°C. RNA samples were retrieved using the Zymo Quick-RNA MiniPrep according the manufacturer’s protocol. The Ribo-Zero Gold, immediately before RNA extraction, 3 mL of lysis buffer (2% SDS, 16 mM EDTA and 200 mM NaCl) contained in a 50-mL polypropylene conical tube was heated for 5 minutes. RNA Removal Kit Epidemiology was then used to deplete prokaryotic and eukaryotic rRNA from the samples according the manufacturer’s protocol. Stranded RNA-Seq libraries were made constructed with the TruSeq Total RNA Library Preparation Kit v2, both using the manufacturer’s protocol. Completed libraries were pooled in equal molar ratios within their respective groups and stored at -20°C until time of sequencing.

**High-throughput sequencing and raw read curation.** Sequencing of 16S rRNA gene amplicon libraries was performed using an Illumina MiSeq sequencer as described previously (47). The 16S rRNA gene sequences were curated using the mothur software package (v1.36) as described in (19). Shotgun metagenomic sequencing was performed in 2 phases. Libraries from mock-infected communities, that were also to be utilized for *de novo* contig assembly, were sequenced using an Illumina HiSeq 2500 on 2x250 paired-end settings and was repeated across 2 lanes to normalize for inter-run variation. *C. difficile*-infected metagenomic libraries were sequenced with an Illumina NextSeq 300 with 2x150 settings across 2 runs to also normalize for inter-run variation. These efforts resulted in an average of 280 million paired raw reads per sample. Metatranscriptomic sequencing was performed on an Illumina HiSeq 2500 with 2x50 reagents and was repeated across 4 lanes to normalize technical variation between lanes and to obtain the necessary level of coverage (39). This resulted in an average of 380 million raw cDNA reads per library. Both metagenomic and metatranscriptomic sequencing was performed at the University of Michigan Sequencing Core. Raw sequence read curation for both metagenomic and metatranscriptomic datasets was performed in a two step process. Residual 5-prime and 3-prime Illumina adapter sequences were trimmed using CutAdapt (49) on a per library basis. Reads were quality trimmed using Sickle (50) with a quality cutoff of Q30. This resulted in approximately 270 million reads per library (both paired and orphaned) for both metagenomic and metatranscriptomic sequencing (Table S2).

**Metagenomic contig assembly and gene annotation.** Metagenomic contigs were assembled using Megahit (51) with the following settings: minimum kmer size of 87, maximum kmer size of 127, and a kmer step size of 10. Prodigal was utilized to to identify putative gene sequences, and were screened for a minimum length of 250 nucleotides. These sequences were translated to amino acids and peptides were annotated against the KEGG protein database (52) using Diamond implementation of BLASTp (53). Peptide-level gene annotations were assigned to the corresponding nucleotide sequence, and genes failing to find a match in KEGG were preserved as unannotated genes. Final nucleotide fasta files with KEGG annotations were then utilized in the construction of Bowtie2 (54) mapping databases from downstream analyses.

**DNA/cDNA read mapping and normalization.** Mapping was accomplished using Bowtie2 (54) and the default stringent settings. Optical and PCR duplicates were then removed using Picard MarkDuplicates (<http://broadinstitute.github.io/picard/>). The remaining mappings were converted to idxstats format using Samtools (55) and the read counts per gene were tabulated. Discordant pair mappings were discarded and counts were then normalized to read length and gene length to give a per base report of gene coverage. Transcript abundance was then normalized to gene abundance to yield overall level of transcription for each gene. Unless indicated otherwise, each collection of reads was then 1000-fold iteratively subsampled to 90% of the lowest sequence total within each analysis, and a median expression value for each gene was calculated. Reads contributed by *C. difficile* were removed from analysis using Bowtie2 against the *C. difficile* str. 630 genome with settings allowing up to 2 mismatches.

**Quantification of *in vivo* metabolite relative concentrations.** Metabolomic analysis was performed by Metabolon (Durham, NC), for a detailed description of the procedure refer to (19). Briefly, all methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer at 35,000 mass resolution. Samples were dried then reconstituted in solvents compatible to each of the four methods. The first, in acidic positive conditions using a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 um) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). The second method was identical to the first but was chromatographically optimized for more hydrophobic compounds. The third approach utilized a basic negative ion optimized conditions using a separate dedicated C18 column. Basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM Ammonium Bicarbonate at pH 8. Samples were then analyzed via negative ionization following elution from a hydrophilic interaction chromatography column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 um) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS n scans using dynamic exclusion. The scan range varied slighted between methods but covered 70-1000 m/z. Library matches for each compound were checked for each sample and corrected if necessary.

**Statistical methods.** Analyses were performed using R (v.3.2.0). Significant differences between community structure of treatment groups from 16S rRNA gene sequencing were determined with AMOVA in the mothur software package. Significant differences of inverse Simpson diversity, cfu, toxin titer, and metabolite concentrations were determined by Wilcoxon signed-rank test with Benjamini-Hochberg correction using a study-wide Type I error rate of 0.05. Undetectable points used half the limit of detection for cfu and toxin statistical calculations. LEfSe analysis with OTU data was performed with a cutoff that each OTU must appear in all samples from their respective groups to be considered true signal (n=9). Random forest was performed using the implementation in R (56), with the informative threshold of MDAs greater than the absolute value of the lowest MDA defined by (57). Distances of outlier points from center line during metatranscriptomic comparisons was accomplished using 2-dimensional linear geometry.

**Data Availability.** Pooled and quality trimmed transcriptomic and 16S rRNA gene amplicon read data from infection experiments are available through the NCBI Sequence Read Archive (SRA; PRJNA354635). Metagenomeic reads, mock-infected metatranscriptomic reads can be found also on the SRA (PRJNA415307). Data processing steps for beginning from raw sequence data to the final manuscript are hosted at <https://github.com/mjenior/Jenior_Metatranscriptomics_mSphere_2018>.

### Acknowledgments

The authors would like to acknowledge Charles Koumpouras for assistance with DNA extractions and metabolomic sample preparation. We would also like to acknowledge members of the University of Michigan Germfree Mouse Center, University of Michigan Sequencing Core, and Metabolon for their assistance in experimental design, execution, and data collection. Thank you also to Greg Medlock and Maureen Carey for input on details of the analysis.

**Author Contributions.** M.L.J. conceived, designed and performed experiments, analyzed data, and drafted the manuscript. J.L.L. performed experiments, analyzed data, and contributed to the manuscript. V.B.Y. contributed to the manuscript. P.D.S. interpreted data and contributed the manuscript. The authors declare no conflicts of interest.

### Figure and Table Legends

**Figure 1 | Experimental models of *C. difficile* infection and distinct virulence patterns.** **(A)** Experimental time lines of antibiotic pretreatment and infection mouse models. **(B)** *C. difficile* 630 cfu in stool of infected mice following each antibiotic pretreatment regimen over 10 days of infection. Median and interquartile range are shown for each time point. **(C)** Quantification of *C. difficile* cfu and toxin titer in cecal content across antibiotic treatment models after 18 hours of infection. Black lines indicate median values. Gray asterisks indicate significant difference from no antibiotic controls by Wilcoxon rank-sum test with the Benjamini-Hochberg correction (all *p* < 0.001). Points below the limit of detection were excluded from statistical testing. Nine mice were used in each experimental and control group and samples from these animals were used for all downstream analyses.

**Figure 2 | Impact of *C. difficile* colonization on gut microbiota community composition.** **(A)** Relative abundance of family-level taxonomic classification for OTUs in each pretreatment group. **(B-D)** Discriminating OTUs with LEfSe analysis between Mock and *C. difficile*-infected communities within each pretreatment group, *p*-values from LEfSe are shown on the left with taxonomic information (all *p* < 0.05). Relative abundance of the respective OTUs from each mouse along with medians are shown for each pretreatment group.

**Figure 3 | Metabolomic markers of infection susceptibility.** Results from UPLC/MS metabolomic analysis of cecal content from mice also used in previous analyses. Only mock-infected metabolomic results were included this analysis to identify markers of susceptibility. **(A-B)** NMDS ordinations of Bray-Curtis distances comparing the cecal metabolome dissimilarity of mice receiving no pretreatment or one of the three distinct classes of antibiotics. (A) Metabolomes of resistant mice are significantly different from antibiotic pretreated animals (*p* < 0.001). (B) Antibiotic pretreated metabolomes and also significantly distinct from one another (*p* < 0.001). Significant differences for NMDS analyses were calculated with AMOVA. **(C-D)** Scaled intensities of metabolites with highest mean decrease in accuracy (MDA) from random forest feature selection discriminating groups from (A-B). MDA is labeled in brackets beside each metabolite name. Asterisks along the right axis indicate significant difference by Wilcoxon rank-sum test. (C) Relative concentrations of metabolites that distinguish cecal content from mice resistant to *C. difficile* colonization from susceptible animals. (D) Metabolites concentrations for those molecules that are able to differentiate antibiotic pretreatment groups. Multiple comparisons were accounted for using the Benjamini-Hochberg correction.

**Figure 4 | Differential effects of CDI on the cecal metabolome.** **(A-C)** NMDS ordinations of Bray-Curtis distances comparing infection groups within each antibiotic pretreatment: (A) Streptomycin, (B) Cefoperazone, and (C) Clindamycin. Significant differences were determined by AMOVA. **(D-F)** Relative concentrations of metabolites with highest Mean Decrease Accuracy (MDA) from Random Forest feature selection to discriminate between Mock and *C. difficile*-infected conditions within each antibiotic pretreatment. Respective MDA is labeled in brackets next to each metabolite name and out of bag error from internal cross-validation of each model is labeled under the plotting area. (D) Streptomycin pretreatment, (E) Cefoperazone pretreatment, and (F) Clindamycin pretreatment. **(G)** Relative concentrations of 5-aminovalerate across all experimental and control groups. Significant differences are labeled along the right axis and were calculated with Wilcoxon rank-sum test with Benjamini-Hochberg correction.

**Figure 5 | Genes and pathways with largest degree of change in metatranscriptome associated with CDI.** Venn diagrams showing the overlap in top 15% of differentially transcribed individual genes **(A)** or pathways **(B)** respectively in each antibiotic preatretment between CDI and mock infections. Below each venn diagram are the corresponding genes/pathways contained in the central shared regions for all three pretreatments. Solid colored bars indicated *C. difficile* infection and dashed bars are mock infected, separate colors denoted sntibiotic pretreatment. Genes/pathways are ranked in terms of decreasing sample variance between all analyzed groups. Genes are also labeled with numbers indicated membership in KEGG pathways (pathway legend below).

**Figure 6 | *C. difficile* colonization alters gene expression of taxonomic groups differentially between antibiotic pretreatments.** **(A-C)** Each point represents a unique gene from the respective metagenomic assembly. Coordinates were determined by the log2-transformed expression level of each gene between *C. difficile*-infected and mock-infected conditions. Metatranscriptomic read abundances were normalized to their associated metagenomic coverage. Outliers were defined using linear correlation and a squared residual cutoff of 2. Distance of outliers to the x=y line were also calculated and represented in unites of deviation or UD. Colored indicate genus of origin, and gray areas denote genes with consistent expression between conditions (defined by outlier analysis). Antibiotic pretreatments: (A) Streptomycin-pretreated, (B) Cefoperazone-pretreated, and (C) Clindamycin-pretreated. **(D)** Absolute difference in metatranscriptomic reads contributed by each genus in pretreatments between mock and *C. difficile*-infected conditions. Colored lines denoted antibiotic pretreatment. Each point represents all transcript contributed by that genus in each pretreatment group. Numbers at the base of pretreatment lines in the ≤0.1% panel represent the quantity of genera in each group as some points are obscured.

**Supplementary Figure 1 | Impact of antibiotic pretreatment and *C. difficile* infection on cecal bacterial community structure.** **(A)** Inverse-Simpson diversity of cecal communities from all pretreatment groups. Gray stars indicate significant difference from no antibiotic controls (all *p* < 0.001). Black stars denote within pretreatment group significant difference between mock and *C. difficile*-infected communities. Differences were calculated using Wilcoxon rank-sum test with Benjamini-Hochberg correction. NMDS ordinations of Bray-Curtis distances comparing the groups labeled in the bottom left of each plotting area. **(B)** Antibiotic-pretreated compared to Untreated controls and **(C)** comparison between only antibiotic pretreatment groups. **(D-F)** Within antibiotic pretreatment comparisons for the effect of *C. difficile* colonization on community structure. Significant differences and correlation coefficients for ordination analyses were found using ANOSIM.

**Supplementary Figure 2 | Relative concentrations of select metabolite groups in each group** Metabolites included in this analysis were chosen based on their previously published links to *C. difficile* physiology or susceptibility to infection. Groups are as follows: **(A)** Bile acids, **(B)** Carbohydrates, and **(C)** Amino Acids. Significant differences were determined by Wilcoxon rank-sum test with Benjamini-Hochberg correction.

**Supplementary Figure 3 | Effect of infection on the cecal metabolome across pretreatment groups** Pooled analysis of antibiotic pretreated animals only. **(A)** NMDS ordination of Bray-Curtis distances differentiating mock and *C. difficile*-infected metabolomes (*p* = 0.075). **(B)** Random forest classification results for metabolites that effectively distinguish infected and uninfected conditions. Shown are relative concentrations of the top 10 metabolites with the highest mean decrease in accuracy from internal cross-validation. Shown in the top right corner of each panel are the metabolite names and mean decrease accuracy of each. Significant differences in concentration between mock and *C. difficile*-infected groups were determined by Wilcoxon rank-sum test with Benjamini-Hochberg correction.

**Supplementary Figure 4 | Largest discrete gene and pathway, metatranscriptomic differences between antibiotic pretreatments and resistant microbiota** Included are the top 10 genes from unique subsections of venn diagrams from each pretreatment group in Figure 5A-B. Genes/Pathways with the largest disparity compared the metatranscriptome from the cecal community of untreated mice. Panels for gene-level analysis are as follows; **(A)** Streptomycin-pretreatment, **(B)** Cefoperazone-pretreatment, and **(C)** Clindamycin-pretreatment. **(D)** Pathway-level annotations for unique groupings in Figure 5B.

**Supplementary Figure 5 | Unique genes with functional annotation detectable within each metagenome and metatranscriptome** Genes in each datasets were derived from respective metagenomic assemblies, with only those genes that mapped to a KEGG pathway-level annotation. Results are shown as follows: **(A)** Untreated, **(B)** Streptomycin-pretreated, **(C)** Cefoperazone-pretreated, and **(D)** Clindamycin-pretreated mice. Each panel includes that treatments’ unique genes from metagenomic assembly and genes that recruited at least one cDNA read from the corresponding metatranscriptomes. Collector’s curves from rarefaction analysis of reads mapped to genes from **(E)** metagenomes and **(F)** metatranscriptomes.

**Supplementary Table 1 | Chosen antibiotic pretreatment regimes** Antibiotic classes, mechanisms, and dosage information for each pretreatment.

**Supplementary Table 2 | High-throughput sequencing read counts and metagenomic assembly quality** Raw and curated read abundances for both metagenomic and metatranscriptomic sequencing efforts. Raw read curation steps are outlined in Materials & Methods. Metagenomic contig summary statistics reflect the quality of assembly for each group.

**Supplementary Table 3 | Normalized cDNA abundances of complete metatranscriptomes for each antibiotic pretreated group** Final cDNA read abundances for each gene from the corresponding metagenomic assembly. Normalization reflects standardizing cDNA read abundances to both read length and target gene length as well as metagenomic coverage for each gene.

**Supplementary Table 4 | Median scaled intensity of metabolites for each antibiotic pretreatment and infection group**

**Supplementary Table 5 | Interquartile range of sample variances across datasets with replication generated during this study**

### References

1. **Vollaard, E. J.**, and **H. A. L. Clasener**. 1994. Colonization resistance. U.S. Patent 3.

2. **Freter, R.** 1955. The Fatal Enteric Cholera Infection in the Guinea Pig, Achieved by Inhibition of Normal Enteric Flora. The Journal of Infectious Diseases **97**:57–65.

3. **Fekety, R.**, **J. Silva**, **R. Toshniwal**, **M. Allo**, **J. Armstrong**, **R. Browne**, **J. Ebright**, and **G. Rifkin**. 1979. Antibiotic-associated colitis: Effects of antibiotics on clostridium difficile and the disease in hamsters. Reviews of Infectious Diseases **1**:386–397.

4. **Britton, R. A.**, and **V. B. Young**. 2012. Interaction between the intestinal microbiota and host in Clostridium difficile colonization resistance. Trends in microbiology **20**:313–9.

5. **Lessa, F. C.**, **Y. Mu**, **W. M. Bamberg**, **Z. G. Beldavs**, **G. K. Dumyati**, **J. R. Dunn**, **M. M. Farley**, **S. M. Holzbauer**, **J. I. Meek**, **E. C. Phipps**, **L. E. Wilson**, **L. G. Winston**, **J. A. Cohen**, **B. M. Limbago**, **S. K. Fridkin**, **D. N. Gerding**, and **L. C. McDonald**. 2015. Burden of Clostridium difficile Infection in the United States. New England Journal of Medicine **372**:825–834.

6. **Antonopoulos, D. A.**, **S. M. Huse**, **H. G. Morrison**, **T. M. Schmidt**, **M. L. Sogin**, and **V. B. Young**. 2009. Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. Infection and Immunity **77**:2367–2375.

7. **Buffie, C. G.**, **I. Jarchum**, **M. Equinda**, **L. Lipuma**, **A. Gobourne**, **A. Viale**, **C. Ubeda**, **J. Xavier**, and **E. G. Pamer**. 2012. Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to Clostridium difficile-induced colitis. Infection and Immunity **80**:62–73.

8. **Thomas, C.**, **M. Stevenson**, and **T. V. Riley**. 2003. Antibiotics and hospital-acquired Clostridium difficile-associated diarrhoea: A systematic review **51**:1339–1350.

9. **Brown, K. A.**, **N. Khanafer**, **N. Daneman**, and **D. N. Fisman**. 2013. Meta-analysis of antibiotics and the risk of community-associated Clostridium difficile infection. Antimicrobial Agents and Chemotherapy **57**:2326–2332.

10. **Bignardi, G.** 1998. Risk factors for Clostridium difficile infection. Journal of Hospital Infection **40**:1–15.

11. **Antunes, L. C. M.**, **J. Han**, **R. B. R. Ferreira**, **P. Loli**, **C. H. Borchers**, and **B. B. Finlay**. 2011. Effect of antibiotic treatment on the intestinal metabolome. Antimicrobial Agents and Chemotherapy **55**:1494–1503.

12. **Jump, R. L. P.**, **A. Polinkovsky**, **K. Hurless**, **B. Sitzlar**, **K. Eckart**, **M. Tomas**, **A. Deshpande**, **M. M. Nerandzic**, and **C. J. Donskey**. 2014. Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. PLoS ONE **9**.

13. **Theriot, C. M.**, **M. J. Koenigsknecht**, **P. E. Carlson**, **G. E. Hatton**, **A. M. Nelson**, **B. Li**, **G. B. Huffnagle**, **J. Z Li**, and **V. B. Young**. 2014. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. Nature communications **5**:3114.

14. **Wilson, K. H.**, and **F. Perini**. 1988. Role of competition for nutrients in suppression of Clostridium difficile by the colonic microflora. Infection and Immunity **56**:2610–2614.

15. **Sambol, S. P.**, **M. M. Merrigan**, **J. K. Tang**, **S. Johnson**, and **D. N. Gerding**. 2002. Colonization for the Prevention of Clostridium difficile Disease in Hamsters. The Journal of infectious diseases **186**:14–16.

16. **P??rez-Cobas, A. E.**, **A. Artacho**, **S. J. Ott**, **A. Moya**, **M. J. Gosalbes**, and **A. Latorre**. 2014. Structural and functional changes in the gut microbiota associated to Clostridium difficile infection. Frontiers in Microbiology **5**.

17. **Zaura, E.**, **B. W. Brandt**, **M. J. T. de Mattos**, **M. J. Buijs**, **M. P. M. Caspers**, **M. U. Rashid**, **A. Weintraub**, **C. E. Nord**, **A. Savell**, **Y. Hu**, **A. R. Coates**, **M. Hubank**, **D. A. Spratt**, **M. Wilson**, **B. J. F. Keijser**, and **W. Crielaard**. 2015. Same Exposure but two radically different responses to antibiotics: Resilience of the salivary microbiome versus long-term microbial shifts in feces. mBio **6**.

18. **Schubert, A. M.**, **H. Sinani**, and **P. D. Schloss**. 2015. Antibiotic-induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against Clostridium difficile. mBio **6**.

19. **Jenior, M. L.**, **J. L. Leslie**, **V. B. Young**, and **P. D. Schloss**. 2017. Clostridium difficile colonizes alternative nutrient niches during infection across distinct murine gut microbiomes. mSystems. American Society for Microbiology Journals **2**.

20. **Bouillaut, L.**, **T. Dubois**, **A. L. Sonenshein**, and **B. Dupuy**. 2015. Integration of metabolism and virulence in Clostridium difficile. Research in Microbiology **166**:375–383.

21. **Koenigsknecht, M. J.**, **C. M. Theriot**, **I. L. Bergin**, **C. A. Schumacher**, **P. D. Schloss**, and **V. B. Young**. 2015. Dynamics and establishment of Clostridium difficile infection in the murine gastrointestinal tract. Infection and Immunity **83**:934–941.

22. **Segata, N.**, **J. Izard**, **L. Waldron**, **D. Gevers**, **L. Miropolsky**, **W. S. Garrett**, and **C. Huttenhower**. 2011. Metagenomic biomarker discovery and explanation. Genome Biology **12**:R60.

23. **Liaw, a**, and **M. Wiener**. 2002. Classification and Regression by randomForest. R news **2**:18–22.

24. **Goyer, A.**, **T. L. Johnson**, **L. J. Olsen**, **E. Collakova**, **Y. Shachar-Hill**, **D. Rhodes**, and **A. D. Hanson**. 2004. Characterization and Metabolic Function of a Peroxisomal Sarcosine and Pipecolate Oxidase from Arabidopsis. Journal of Biological Chemistry **279**:16947–16953.

25. **Fonknechten, N.**, **S. Chaussonnerie**, **S. Tricot**, **A. Lajus**, **J. R. Andreesen**, **N. Perchat**, **E. Pelletier**, **M. Gouyvenoux**, **V. Barbe**, **M. Salanoubat**, **D. Le Paslier**, **J. Weissenbach**, **G. N. Cohen**, and **A. Kreimeyer**. 2010. Clostridium sticklandii, a specialist in amino acid degradation:revisiting its metabolism through its genome sequence. BMC genomics **11**:555.

26. **Nakamura, S.**, **S. Nakashio**, **K. Yamakawa**, **N. Tanabe**, and **S. Nishida**. 1982. Carbohydrate Fermentation by Clostridium difficile. Microbiology and Immunology **26**:107–111.

27. **Sorg, J. A.**, and **A. L. Sonenshein**. 2010. Inhibiting the initiation of Clostridium difficile spore germination using analogs of chenodeoxycholic acid, a bile acid. Journal of Bacteriology **192**:4983–4990.

28. **Buffie, C. G.**, **V. Bucci**, **R. R. Stein**, **P. T. McKenney**, **L. Ling**, **A. Gobourne**, **D. No**, **H. Liu**, **M. Kinnebrew**, **A. Viale**, **E. Littmann**, **M. R. M. van den Brink**, **R. R. Jenq**, **Y. Taur**, **C. Sander**, **J. R. Cross**, **N. C. Toussaint**, **J. B. Xavier**, and **E. G. Pamer**. 2014. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. Nature **517**:205–208.

29. **Neumann-Schaal, M.**, **J. D. Hofmann**, **S. E. Will**, and **D. Schomburg**. 2015. Time-resolved amino acid uptake of Clostridium difficile 630erm and concomitant fermentation product and toxin formation. BMC Microbiology 281.

30. **Bouillaut, L.**, **W. T. Self**, and **A. L. Sonenshein**. 2013. Proline-dependent regulation of Clostridium difficile stickland metabolism. Journal of Bacteriology **195**:844–854.

31. **Ilari, A.**, **S. Stefanini**, **E. Chiancone**, and **D. Tsernoglou**. 2000. The dodecameric ferritin from Listeria innocua contains a novel intersubunit iron-binding site. Nature Structural Biology **7**:38–43.

32. **Mahowald, M. A.**, **F. E. Rey**, **H. Seedorf**, **P. J. Turnbaugh**, **R. S. Fulton**, **A. Wollam**, **N. Shah**, **C. Wang**, **V. Magrini**, **R. K. Wilson**, **B. L. Cantarel**, **P. M. Coutinho**, **B. Henrissat**, **L. W. Crock**, **A. Russell**, **N. C. Verberkmoes**, **R. L. Hettich**, and **J. I. Gordon**. 2009. Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proceedings of the National Academy of Sciences **106**:5859–5864.

33. **Jousset, A.**, **C. Bienhold**, **A. Chatzinotas**, **L. Gallien**, **A. Gobet**, **V. Kurm**, **K. Küsel**, **M. C. Rillig**, **D. W. Rivett**, **J. F. Salles**, **M. G. A. van der Heijden**, **N. H. Youssef**, **X. Zhang**, **Z. Wei**, and **W. H. G. Hol**. 2017. Where less may be more: how the rare biosphere pulls ecosystems strings. The ISME Journal.

34. **Shade, A.**, **S. E. Jones**, **J. Gregory Caporaso**, **J. Handelsman**, **R. Knight**, **N. Fierer**, and **J. A. Gilbert**. 2014. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. mBio **5**.

35. **Rojo, D.**, **M. J. Gosalbes**, **R. Ferrari**, **A. E. Pérez-Cobas**, **E. Hernández**, **R. Oltra**, **J. Buesa**, **A. Latorre**, **C. Barbas**, **M. Ferrer**, and **A. Moya**. 2015. Clostridium difficile heterogeneously impacts intestinal community architecture but drives stable metabolome responses. The ISME Journal **9**:2206–2220.

36. **Abraham, N. M.**, **L. Liu**, **B. L. Jutras**, **A. K. Yadav**, **S. Narasimhan**, **V. Gopalakrishnan**, **J. M. Ansari**, **K. K. Jefferson**, **F. Cava**, **C. Jacobs-Wagner**, and **E. Fikrig**. 2017. Pathogen-mediated manipulation of arthropod microbiota to promote infection. Proceedings of the National Academy of Sciences 201613422.

37. **Russel, J.**, **H. Roder**, **J. Madsen**, **M. Burmell**, and **S. Soresen**. 2017. Antagonism correlates with metabolic similarity in diverse bacteria. PNAS.

38. **Sheik, C. S.**, **S. Jain**, and **G. J. Dick**. 2014. Metabolic flexibility of enigmatic SAR324 revealed through metagenomics and metatranscriptomics. Environmental Microbiology **16**:304–317.

39. **Franzosa, E. A.**, **X. C. Morgan**, **N. Segata**, **L. Waldron**, **J. Reyes**, **A. M. Earl**, **G. Giannoukos**, **M. R. Boylan**, **D. Ciulla**, **D. Gevers**, **J. Izard**, **W. S. Garrett**, **A. T. Chan**, and **C. Huttenhower**. 2014. Relating the metatranscriptome and metagenome of the human gut. Proceedings of the National Academy of Sciences **111**:E2329–E2338.

40. **Jorth, P.**, **K. H. Turner**, **P. Gumus**, **N. Nizam**, **N. Buduneli**, and **M. Whiteley**. 2014. Metatranscriptomics of the human oral microbiome during health and disease. mBio **5**.

41. **Reeves, A. E.**, **M. J. Koenigsknecht**, **I. L. Bergin**, and **V. B. Young**. 2012. Suppression of Clostridium difficile in the gastrointestinal tracts of germfree mice inoculated with a murine isolate from the family Lachnospiraceae. Infection and Immunity **80**:3786–3794.

42. **Lawley, T. D.**, **S. Clare**, **A. W. Walker**, **M. D. Stares**, **T. R. Connor**, **C. Raisen**, **D. Goulding**, **R. Rad**, **F. Schreiber**, **C. Brandt**, **L. J. Deakin**, **D. J. Pickard**, **S. H. Duncan**, **H. J. Flint**, **T. G. Clark**, **J. Parkhill**, and **G. Dougan**. 2012. Targeted Restoration of the Intestinal Microbiota with a Simple, Defined Bacteriotherapy Resolves Relapsing Clostridium difficile Disease in Mice. PLoS Pathogens **8**.

43. **Petrof, E. O.**, **G. B. Gloor**, **S. J. Vanner**, **S. J. Weese**, **D. Carter**, **M. C. Daigneault**, **E. M. Brown**, **K. Schroeter**, and **E. Allen-Vercoe**. 2013. Stool substitute transplant therapy for the eradication of Clostridium difficile infection: ‘RePOOPulating’ the gut. Microbiome **1**:3.

44. **Adlerberth, I.**, **H. Huang**, **E. Lindberg**, **N. Åberg**, **B. Hesselmar**, **R. Saalman**, **C. E. Nord**, **A. E. Wold**, and **A. Weintraubb**. 2014. Toxin-Producing clostridium difficile strains as long-term gut colonizers in healthy infants. Journal of Clinical Microbiology **52**:173–179.

45. **Wilson, K. H.**, **M. J. Kennedy**, and **F. R. Fekety**. 1982. Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. Journal of Clinical Microbiology **15**:443–446.

46. **Leslie, J. L.**, **S. Huang**, **J. S. Opp**, **M. S. Nagy**, **M. Kobayashi**, **V. B. Young**, and **J. R. Spence**. 2015. Persistence and toxin production by Clostridium difficile within human intestinal organoids result in disruption of epithelial paracellular barrier function. Infection and Immunity **83**:138–145.

47. **Kozich, J.**, **S. Westcott**, **N. Baxter**, **S. Highlander**, and **P. Schloss**. 2013. 16S Sequencing with the Illumina MiSeq Personal Sequencer. University of Michigan Health System SOP **3.1**:1–16.

48. **Lopez-Medina, E.**, **M. M. Neubauer**, **G. B. Pier**, and **A. Y. Koh**. 2011. RNA isolation of Pseudomonas aeruginosa colonizing the murine gastrointestinal tract. Journal of visualized experiments : JoVE 6–9.

49. **Martin, M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet **17**:10.

50. **Joshi, N.**, and **J. Fass**. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at https://github.com/najoshi/sickle. 2011.

51. **Li, D.**, **C. M. Liu**, **R. Luo**, **K. Sadakane**, and **T. W. Lam**. 2014. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics **31**:1674–1676.

52. **Ogata, H.**, **S. Goto**, **K. Sato**, **W. Fujibuchi**, **H. Bono**, and **M. Kanehisa**. 1999. KEGG: Kyoto encyclopedia of genes and genomes. U.S. Patent 1.

53. **Buchfink, B.**, **C. Xie**, and **D. H. Huson**. 2015. Fast and sensitive protein alignment using DIAMOND. Nature methods **12**:59–60.

54. **Langmead, B.**, and **S. L. Salzberg**. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods **9**:357–9.

55. **Li, H.**, **B. Handsaker**, **A. Wysoker**, **T. Fennell**, **J. Ruan**, **N. Homer**, **G. Marth**, **G. Abecasis**, and **R. Durbin**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–2079.

56. **Breiman, L.** 2001. Random forests. Machine Learning **45**:5–32.

57. **Segal, M. R.** 2004. Machine Learning Benchmarks and Random Forest Regression. Biostatistics 1–14.