Title: *Clostridium difficile* differentially alters the structure and metabolism of distinct cecal microbiomes to promote persistent colonization during infection

Authors: Matthew L. Jenior[1], Jhansi L. Leslie[1], Vincent B. Young[1,2], and Patrick D. Schloss[1]

Corresponding Author: Patrick D. Schloss

Affiliations:

1.) Department of Microbiology & Immunology; University of Michigan, Ann Arbor MI,

2.) Department of Internal Medicine, Division of Infectious Diseases; University of Michigan, Ann Arbor MI

Contact Information:

Matthew L. Jenior – mljenior@umich.edu

Jhansi L. Leslie – jlleslie@umich.edu

Vincent B. Young – youngvi@umich.edu

Patrick D. Schloss – pschloss@umich.edu

Abstract

Clostridium difficile has become the most common single cause of hospital-acquired infection over the last decade in the United States. Susceptibility is primarily associated with previous exposure to antibiotics, which compromise the structure and function of the gut bacterial community. Furthermore, specific antibiotic classes correlate more strongly with recurrent or persistent *C. difficile* infection. We used a murine model of infection to explore the effect of distinct antibiotic classes on sustained *C. difficile* colonization, as well as the impact of infection on community-level gene expression and metabolism 18 hours post-infection. Utilizing untargeted metabolomic analysis, we found that *C. difficile* infection has larger impacts on the metabolic activity of the microbiota across cefoperazone and streptomycin-pretreated mice, which become persistently colonized compared to clindamycin-pretreated mice that cleared the infection within 8 days. Through metagenome-enabled metatranscriptomics, we observed that the infected microbial communities were enriched in pathways associated with amino acid metabolism and particularly in non-dominant species relative to mock-infected controls. Conversely, in mice pretreated with clindamycin, the effect of infection on the microbiota was only detectable in changes to the community structure but not in metabolic activity or gene expression. Our results suggest that *C. difficile* is able to restructure the nutrient-niche landscape in certain gut environments in order to promote persistent infection.

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 a person'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]. This is supported by the hypothesis that each differentially sensitized gut environment possesses an alternative nutrient niche landscape and *C. difficile* is able to adjust its metabolism accordingly. Although it is well established that *C. difficile* is able to colonize these communities, it is yet to be determined whether the differences in the metabolic capacity of communities following antibiotic treatment correlate with prolonged *C. difficile* colonization. Defining the functional status of the resident microbiota for any disease has been difficult and has led to a limited understanding of the specific species interactions that occur with *C. difficile* during infection. To address this fundamental question, we employed a murine model of CDI where the murine gut microbiota are disturbed with antibiotics and assessed for the ability of *C. difficile* to colonize and cause disease. We characterized the microbiome utilizing untargeted metabolomics and metagenome-enabled metatransciptomics. This approach allowed us to not only characterize the metabolic output of the community, but to also identify which subgroups of bacteria were differentially active under these conditions. Our data supports the hypothesis that *C. difficile* colonization indeed alters community-level gene expression and the metabolome of these communities. Furthermore, in mice that cleared the infection,

despite shifts in species abundance, the metatrascriptome and metabolome changed very little. This work highlights the need for a better appreciation of the effect CDI has on the gut microbiota is needed to develop more successful targeted therapies that eliminate *C. difficile* after persistent colonization.

Results

Distinct antibiotic pretreatments are associated with different patterns of clearance following primary infection. Conventionally-reared SPF mice were pretreated with one of three different antibiotics to sensitize the animals to *C. difficile* colonization. The selected antibiotics were streptomycin, cefoperazone, and clindamycin (Table S1). Each drug was chosen because they represented different spectrum of activity, had the ability to to reduce *C. difficile* colonization resistance in a mouse model [18], had distinct and significant impacts on the structure and diversity of the cecal microbiota (all $p < 0.001$; Fig. S1A), and the infected communities differed in their to recover and clear the infection (Fig. 1B). Selection of the toxigenic *C. difficile* strain 630 was based on its moderate clinical severity in mouse models [20], previous studies of *in vitro* metabolism [21], and well-annotated genome [22]. Briefly, mice were pretreated with the respective antibiotic and were subsequently challenged with ~$1x10^3$ *C. difficile* str. 630 spores via oral gavage (Fig. 1A). We then monitored for disease over the following 10 days and the carriage of *C. difficile* in feces over time. In each antibiotic pretreatment model, we observed equally high levels of *C. difficile* colonization on the day after infection; however, over the following 8 days, only clindamycin-pretreated mice cleared the infection while the other two pretreatments remained highly colonized (Fig. 1B).

***C. difficile* differentially expresses virulence factors across separate antibiotic pretreatments.** We hypothesized that the clindamycin-pretreated mice cleared the infection while animals in the other groups did not because the pretreated microbiota occupied niche space that with greater overlap with that of *C. difficile*. It had been previously demonstrated that *C. difficile* virulence factor expression is regulated by availability of certain nutrients in the environment [21], so we first sought to measure spore and toxin production. Continuing with the same pretreatment and primary infection protocols used in the 10-day colonization experiments, we chose to focus our analysis on 18-hours post-infection to the assess behavior of *C. difficile* directly prior to the beginning of clearance and utilized this time period for the remainder of experiments. This end point corresponded with a previous study where *C. difficile* reached maximum cecal vegetative cell load (i.e. number of colony forming units [cfu] per gram of cecal contents) with few detectable spores [23]. 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 were colonized with ~$1x10^8$ vegetative cfu per gram of cecal content and 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 [21]. Despite having similar vegetative *C. difficile* load, varying levels of spore load and toxin titer were observed across each of the antibiotic pretreatments. These results showed that *C. difficile*

colonized different antibiotic-pretreated mice to consistently high levels, and the distinct pretreatments corresponded with moderate differences in the level of *C. difficile* virulence factors.

Initial *C. difficile* colonization levels were consistent in spite of significantly different 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 models. 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 & S2B). We then confirmed that each antibiotic pretreatment induced distinct and significant shifts in the community structure (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). We then sought to identify specific OTUs that significantly differed in abundance between mock and *C. difficile*-infected communities within each pretreatment group. Utilizing LEfSe differential abundant feature detection [24], we identified 16 OTUs that discriminated between infected and uninfected communities in at least one antibiotic pretreatment (Fig. 2B-2D); 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 alternative markers of *C. difficile* colonization susceptibility. 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 as quantified by

Bray-Curtis dissimilarity ($p < 0.001$; Fig. 3A). As with the differences between resistant and susceptible states, the patterns of specific metabolite concentrations diverged depending on the antibiotic pretreatment (Fig. S3 & S5). These results demonstrated that each pretreatment lead to distinct susceptible metabolomic environments.

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 [25] 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 metabolite, we ranked metabolites in order of highest mean decrease accuracy (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 present in the intact microbiota [12]. N−methylpipecolate was the only metabolite we identified that consistently decreased during antibiotic pretreatment. This molecule is a by-product of amino acid catabolism and is typically oxidized under normal conditions back into glycine [26]. 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 [27]. We then applied this technique to identify metabolites that distinguish antibiotic pretreatment groups from each other (Fig. 3D). Interestingly, several additional carbohydrate or amino acid catabolism by-products were highlighted by these means including hydroxyisocaproate, methylvalerate, glutamylmethionine, and N−carbamoylaspartate [12]. Additionally, sucrose is a *C. difficile* growth substrate [28] which further promoted the hypothesis of differentially vacated nutrient niches due to antibiotic pretreatment that are now accessible to *C. difficile* (Fig. 2D). These data suggested that the populations responsible for carbohydrate fermentation that result in short-chain fatty acid (SCFA) production may be differentially affected by different classes of antibiotics. The distinct antibiotics allowed for different forms of nutrient catabolism to survive the pretreatment and potentially implies that there are alternative metabolite profiles that impact *C. difficile*'s ability to compete with the indigenous microbiota.

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 deeply connected to the life cycle of *C. difficile* [29] and whose bioconversion by the microbiota has been implicated as a driver of colonization resistance [30]. 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 alters each susceptible cecal metabolome distinctly.

Following the changes 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 surrounding community. First, in a similar approach to identifying susceptibility markers, we observed the global Bray-Curtis dissimilarities of mock-infected and *C. difficile*-infected metabolomes within each antibiotic pretreatment group individually. We compared the cecal metabolomes from all mice across pretreatment groups, and were unable to detect a consistent difference between groups ($p = 0.075$; Fig. S3A). We moved on to individual antibiotic groups and found that both streptomycin ($p = 0.039$) and cefoperazone ($p = 0.016$) pretreated metabolomes deviated significantly from that of mock infection (Fig. 4A-4B). However, clindamycin-pretreated cecal metabolomes were not significantly altered by the presence of the pathogen ($p = 0.127$; Fig. 4C). These results diverged from what was seen in the paired OTU relative abundance results where instead the cecal

community structure of mice pretreated with cefoperazone was unchanged and clindamycin-pretreated communities were significantly different (p = 0.003; Fig. 4B, 4C). Interestingly, streptomycin-pretreated microbiomes were significantly altered by *C. difficile* infection at both the OTU and metabolomic levels (Fig. 4A & S1D). These data indicated that large shifts of populations in the cecal microbiota were not implicitly associated with 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 pooled analysis across all models tested. To accomplish this, we again applied the Random Forest machine-learning approach. We were able to distinguish those microbiomes infected with *C. difficile* from those that were not and identified the 10 metabolites with the greatest 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-pretreatment, the only group to reduce *C. difficile* colonization over 10 days, 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* [31]. *C. difficile* is able to catabolize proline along with glycine through a set of paired biochemical reactions known as Stickland fermentation

[32]. Along these lines, when we measured for unified metabolite signals across all antibiotic pretreatment models, proline and 3 proline-conjugated amino acids (i.e. pro-hydroxy-pro, thioproline, and trans-4-hydroxyproline) were significantly decreased during infection (Fig. S3B; all $p < 0.001$). These combined results demonstrated that while distinct metabolic challenges may exist in each susceptible metabolome, Stickland fermentation could be a preferred energy acquisition pathway for *C. difficile in vivo*. Random Forest machine-learning models were able to reliably classify infected and uninfected mice in both streptomycin and cefoperazone pretreatments (Fig. 4A-4B, Out-of-bag error = 0%). However, infection did not provide for the same level of discrimination in clindamycin-pretreated mice (Fig. 4C, Out-of-bag error = 44.44%). None of the top 5 MDA metabolites in each antibiotic pretreatment were shared across all groups during individual analysis. 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. However, it was not possible to isolate where changes to the metabolome were a result of altered community metabolism, altered host metabolism, or from *C. difficile*.

***C. difficile* colonization induced shifts in the expression of several catabolic pathways and nutrient acquisition systems across susceptible communities.** 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. Metagenomic reads from mock-infected cecal communities were assembled *de novo* into contigs and putative genes were identified resulting in 234,124 (streptomycin), 83,096 (cefoperazone), and 35,977 (clindamycin) potential genes in each metagenome. Streptomycin pretreatment resulted in a significantly more diverse community (Fig. S1A) than other groups, so a more diverse metagenome was expected. Utilizing metatranscriptomic read abundances normalized to corresponding metagenomic coverage per gene, we first focused on differences in gene expression in broadly defined pathways and gene categories for each antibiotic pretreatment with and without *C. difficile*-infection. Based on the metabolomic results, we hypothesized that pathways with the greatest differences would include those involved in the metabolism of carbohydrates and amino acids. We then calculated the difference in cDNA abundance for each pathway between infected and uninfected conditions, represented as delta-cDNA abundance. To highlight the largest differences, we limited the analysis to the top 5 KEGG pathways with the most change between mock and *C. difficile*-infected conditions within each pretreatment (Fig. 5A). In streptomycin and clindamycin-pretreated mice, increased expression of KEGG pathways was observed in the *C. difficile*-infected metatranscriptomes. Both groups displayed large changes in amino sugar metabolism and ABC transporters, however other distinct carbon metabolism pathways were upregulated. Genes for glycolysis and oxidative phosphorylation were overrepresented in streptomycin pretreated mice while

starch/sucrose metabolism and PTS systems were more abundantly associated with clindamycin-pretreated mice. Together these shifts suggested that communities differentially adapt carbon metabolism pathways in response to colonization of *C. difficile*. Conversely in cefoperazone-pretreated mice, a large number of genes were instead over-expressed in the absence of *C. difficile*. These pathways included the replication or manipulation of genetic material (i.e. RNA Processing, tRNA Synthesis, & Homologous Recombination). Among the genes that were upregulated in the presence of *C. difficile*, there were many genes involved in anaerobic glucose metabolism and select ABC-transporters (Table S3). These results indicate that the cecal microbiota of infected mice shifted its metabolism toward catabolizing simple carbohydrates.

We then performed a more fine-scale analysis of changes at the gene level, by selecting the genes in each antibiotic pretreatment group with the largest disparity in normalized cDNA abundance between mock and *C. difficile*-infected groups (Fig. 5B-5D). In agreement with pathway-level differences in expression, the majority of genes with the greatest difference between mock and infected mice belonged to the pathways highlighted in Fig. 5A with three additional pathways relating to amino acid metabolism (I: Glutathione metabolism, J: Valine/Leucine/Isoleucine metabolism, & K: Glycine/Serine/Threonine metabolism). Numerous genes for transport of simple carbohydrates and glycolysis were also differentially overrepresented in infected conditions across pretreatment groups when compared to mock infection of untreated, resistant communities (Fig. S4). These data expand our interpretation of pathway-level analysis and support our hypothesis that *C. difficile* colonization leads to changes in community-level expression of genes for nutrient acquisition and catabolism.

***C. difficile* colonization corresponds with large-scale changes in expression of genes from specific 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. We then directly compared the normalized cDNA abundances for each gene between infected and uninfected states for each antibiotic pretreatment and applied linear correlation to identify distinct patterns of expression (Fig. 3A-3C). This resulted in 2473 outliers at an average distance of 2.545 arbitrary units (AU) associated with streptomycin, 2930 outliers at an average distance 3.854 AU with cefoperazone, and only 727 outliers at an average distance of 2.414 AU 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 community gene expression was not significantly altered in the presence of *C. difficile*.

As 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* that were upregulated with *C. difficile* infection (Fig. 6A). In cefoperazone-pretreated mice 2290 genes belonging to *Bacteroides* were upregulated 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 (r = 0.03 & r = 0.03). In clindamycin-pretreated mice the largest difference in transcription was for 510 *Lactobacillus* genes that were upregulated during CDI (Fig. 6C). Infected and uninfected metatranscriptomes associated with this antibiotic more strongly correlated than either other pretreatment (r = 0.86), further supporting that *C. difficile* colonization had a low impact on transcription of the cecal microbiota. These data indicate that *C. difficile* may differentially modify the transcriptional activity of separate microbial populations based on the context of the community in which it colonizes.

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 to OTU abundances, transcriptional activity of the microbiota, and cecal metabolome shortly after infection suggest that *C. difficile* manipulates the niche landscape of the intestinal tract in order to better maintain long-term colonization. This hypothesis is best supported by the disparity of community-level phenotypes between the clindamycin and cefoperazone/ or streptomycin pretreatment models. Only in the clindamyin pretreatment model 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 previously for prominent symbiotic bacterial species in gnotobiotic mice [33], 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 metabolic environment of the community at large [34]. 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 [35], and this concept would likely apply to bacterial groups recovering population density following cessation of antibiotic pretreatment. As such, *C. difficile* may preferentially seek to compete with 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 *in vivo* observations that a medically relevant bacterial pathogen may alter the metabolic activity of a host-associated community to promote its own colonization. Others have identified potential metabolite markers of *C. difficile* infection in patient feces [36], 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 [37]. 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 [38]. In previous work from our group, we 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. While 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 mRNA detected for a given gene does not necessarily reflect the amount of functional protein made by a cell. Furthermore, interpretation of timing may also be an issue since a large influx of transcript for a specific product may signal an initial upregulation before subsequent translation has been able to occur. This also omits consideration of any post-translational modifications that are required for ultimately functional enzymes. Additionally, due to the low relative abundance of *C. difficile* in these communities as well as the high density of other surrounding bacterial species required focusing sequencing efforts on depth rather than multiple replicates within each group. Greater transcript read abundance per gene allows for improved survey for the activity of lowly abundant species as well as greater confidence in genes found to be highly transcribed. 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 [19,39–41]. 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 replicate this effect, but each has been met with incomplete restoration of colonization resistance or function [30,42–44]. 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 [45]. Moreover, the current 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 require specific microbes or functionalities to restore colonization resistance to *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. For a more detailed description of the animal model, refer to Jenior et al. [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. They were administered one of three antibiotics; cefoperazone, streptomycin, or clindamycin before oral *C. difficile* infection (Table S1). 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.

***C. difficile* infection and necropsy.** On the day of challenge, $1x10^3$ *C. difficile* 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 CO_2 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 [46]. In parallel, undiluted samples were heated at 60°C for 30 minutes to eliminate vegetative cells and enrich for spores [29]. 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* load 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 rounding assay was performed as described elsewhere [19,47]. 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 the PowerSoil-htp 96 Well Soil DNA isolation kit (MO BIO Laboratories) and an epMotion 5075 automated pipetting system

(Eppendorf). The V4 region of the bacterial 16S rRNA gene was amplified using custom barcoded primers [48]. 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. For RNA extraction, a more detailed description of the procedure can be found in [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 [49]. 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 [48]. 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 settings and was repeated across 4 lanes for normalization and to obtain necessary coverage [40]. This gave 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 [50] on a per library basis. Reads were quality trimmed using Sickle [51] 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. Actual read abundances for individual metagenomic and metatranscriptomic sequencing efforts can be found in Table S2.

Metagenomic contig assembly and gene annotation. Metagenomic contigs were assembled using Megahit [52] 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 [53] using Diamond implementation of BLASTp [54]. 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 [55] mapping databases from downstream analyses.

DNA/cDNA read mapping and normalization. Mapping was accomplished using Bowtie2 [55] and the default stringent settings. Optical and PCR duplicates were then removed using Picard MarkDuplicates (http://roadinstitute.github.io/picard/). The remaining mappings were converted to idxstats format using Samtools [56] 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 upregulation 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.

Quantification of *in vivo* metabolite relative concentrations. Metabolomic analysis was performed by Metabolon (Durham, NC), or 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. Peaks were quantified using area under the curve.

Statistical methods. All statistical 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, vegetative and spore load, 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 *C. difficile* load 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 [57], with the informative threshold of MDAs greater than the absolute value of the lowest MDA defined by [58]. Distances of outlier points from center line during metatranscriptomic comparisons was accomplished using 2-dimensional linear geometry.

Funding Information

This work was supported by funding from the National Institutes of Health to PDS (R01GM099514, P30DK034933, U19AI09087, and U01AI124255), VBY (P30DK034933, U19AI09087, and U01AI124255), a Translational Research Education Certificate grant to JLL (MICHR; UL1TR000433), and was partially supported by a predoctoral fellowship from the Cellular Biotechnology Training Program to MLJ (T32GM008353).

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. 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_eLife_2017.

Author Contributions M.L.J. conceived, designed and performed experiments, analyzed data, and drafted the manuscript. J.L.L. performed experiments and contributed to the manuscript. V.B.Y conceived of experiments and contributed to the manuscript. P.D.S. interpreted data and contributed to the manuscript.

Competing Interests The authors declare no conflicts of interest.

References

1. Vollaard EJ, Clasener HAL. Colonization resistance. 1994. pp. 409–414. doi:10.1128/AAC.38.3.409

2. Freter R. The Fatal Enteric Cholera Infection in the Guinea Pig, Achieved by Inhibition of Normal Enteric Flora. The Journal of Infectious Diseases. 1955;97: 57–65. Available: www.jstor.org/stable/30092353

3. Fekety R, Silva J, Toshniwal R, Allo M, Armstrong J, Browne R, et al. Antibiotic-associated colitis: Effects of antibiotics on clostridium difficile and the disease in hamsters. Reviews of Infectious Diseases. 1979;1: 386–397. doi:10.1093/clinids/1.2.386

4. Britton RA, Young VB. Interaction between the intestinal microbiota and host in Clostridium difficile colonization resistance. Trends in microbiology. 2012;20: 313–9. doi:10.1016/j.tim.2012.04.001

5. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden of Clostridium difficile Infection in the United States. New England Journal of Medicine. 2015;372: 825–834. doi:10.1056/NEJMoa1408913

6. Antonopoulos DA, Huse SM, Morrison HG, Schmidt TM, Sogin ML, Young VB. Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. Infection and Immunity. 2009;77: 2367–2375. doi:10.1128/IAI.01520-08

7. Buffie CG, Jarchum I, Equinda M, Lipuma L, Gobourne A, Viale A, et al. Profound alterations of intestinal microbiota following a single dose of clindamycin results in

sustained susceptibility to Clostridium difficile-induced colitis. Infection and Immunity. 2012;80: 62–73. doi:10.1128/IAI.05496-11

8. Thomas C, Stevenson M, Riley TV. Antibiotics and hospital-acquired Clostridium difficile-associated diarrhoea: A systematic review. 2003;51: 1339–1350. doi:10.1093/jac/dkg254

9. Brown KA, Khanafer N, Daneman N, Fisman DN. Meta-analysis of antibiotics and the risk of community-associated Clostridium difficile infection. Antimicrobial Agents and Chemotherapy. 2013;57: 2326–2332. doi:10.1128/AAC.02176-12

10. Bignardi G. Risk factors for Clostridium difficile infection. Journal of Hospital Infection. 1998;40: 1–15. doi:10.1016/S0195-6701(98)90019-6

11. Antunes LCM, Han J, Ferreira RBR, Loli P, Borchers CH, Finlay BB. Effect of antibiotic treatment on the intestinal metabolome. Antimicrobial Agents and Chemotherapy. 2011;55: 1494–1503. doi:10.1128/AAC.01664-10

12. Jump RLP, Polinkovsky A, Hurless K, Sitzlar B, Eckart K, Tomas M, et al. Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. PLoS ONE. 2014;9. doi:10.1371/journal.pone.0101267

13. Theriot CM, Koenigsknecht MJ, Carlson PE, Hatton GE, Nelson AM, Li B, et al. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. Nature communications. 2014;5: 3114. doi:10.1038/ncomms4114

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

15. Sambol SP, Merrigan MM, Tang JK, Johnson S, Gerding DN. Colonization for the Prevention of Clostridium difficile Disease in Hamsters. The Journal of infectious diseases. 2002;186: 14–16. doi:10.1086/345676

16. P??rez-Cobas AE, Artacho A, Ott SJ, Moya A, Gosalbes MJ, Latorre A. Structural and functional changes in the gut microbiota associated to Clostridium difficile infection. Frontiers in Microbiology. 2014;5. doi:10.3389/fmicb.2014.00335

17. Zaura E, Brandt BW, Mattos MJT de, Buijs MJ, Caspers MPM, Rashid MU, et al. Same Exposure but two radically different responses to antibiotics: Resilience of the salivary microbiome versus long-term microbial shifts in feces. mBio. 2015;6. doi:10.1128/mBio.01693-15

18. Schubert AM, Sinani H, Schloss PD. Antibiotic-induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against Clostridium difficile. mBio. 2015;6. doi:10.1128/mBio.00974-15

19. Jenior ML, Leslie JL, Young VB, Schloss PD. Clostridium difficile colonizes alternative nutrient niches during infection across distinct murine gut microbiomes. Turnbaugh PJ, editor. mSystems. American Society for Microbiology Journals; 2017;2. doi:10.1128/mSystems.00063-17

20. Theriot CM, Koumpouras CC, Carlson PE, Bergin II, Aronoff DM, Young VB. Cefoperazone-treated mice as an experimental platform to assess differential virulence

of Clostridium difficile strains. Gut microbes. 2011;2: 326–334. doi:10.4161/gmic.2.6.19142

21. Bouillaut L, Dubois T, Sonenshein AL, Dupuy B. Integration of metabolism and virulence in Clostridium difficile. Research in Microbiology. 2015;166: 375–383. doi:10.1016/j.resmic.2014.10.002

22. Eijk E van, Anvar S, Browne HP, Leung W, Frank J, Schmitz AM, et al. Complete genome sequence of the Clostridium difficile laboratory strain 630∆erm reveals differences from strain 630, including translocation of the mobile element CTn5. BMC Genomics. 2015;16: 31. doi:10.1186/s12864-015-1252-7

23. Koenigsknecht MJ, Theriot CM, Bergin IL, Schumacher CA, Schloss PD, Young VB. Dynamics and establishment of Clostridium difficile infection in the murine gastrointestinal tract. Infection and Immunity. 2015;83: 934–941. doi:10.1128/IAI.02768-14

24. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biology. 2011;12: R60. doi:10.1186/gb-2011-12-6-r60

25. Liaw a, Wiener M. Classification and Regression by randomForest. R news. 2002;2: 18–22. doi:10.1177/154405910408300516

26. Goyer A, Johnson TL, Olsen LJ, Collakova E, Shachar-Hill Y, Rhodes D, et al. Characterization and Metabolic Function of a Peroxisomal Sarcosine and Pipecolate Oxidase from Arabidopsis. Journal of Biological Chemistry. 2004;279: 16947–16953. doi:10.1074/jbc.M400071200

27. Fonknechten N, Chaussonnerie S, Tricot S, Lajus A, Andreesen JR, Perchat N, et al. Clostridium sticklandii, a specialist in amino acid degradation:revisiting its metabolism through its genome sequence. BMC genomics. 2010;11: 555. doi:10.1186/1471-2164-11-555

28. Nakamura S, Nakashio S, Yamakawa K, Tanabe N, Nishida S. Carbohydrate Fermentation by Clostridium difficile. Microbiology and Immunology. 1982;26: 107–111. doi:10.1111/j.1348-0421.1982.tb00159.x

29. Sorg JA, Sonenshein AL. Inhibiting the initiation of Clostridium difficile spore germination using analogs of chenodeoxycholic acid, a bile acid. Journal of Bacteriology. 2010;192: 4983–4990. doi:10.1128/JB.00610-10

30. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, et al. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. Nature. 2014;517: 205–208. doi:10.1038/nature13828

31. Neumann-Schaal M, Hofmann JD, Will SE, Schomburg D. Time-resolved amino acid uptake of Clostridium difficile 630Δerm and concomitant fermentation product and toxin formation. BMC Microbiology. 2015; 281. doi:10.1186/s12866-015-0614-2

32. Bouillaut L, Self WT, Sonenshein AL. Proline-dependent regulation of Clostridium difficile stickland metabolism. Journal of Bacteriology. 2013;195: 844–854. doi:10.1128/JB.01492-12

33. Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, Wollam A, et al. Characterizing a model human gut microbiota composed of members of its two

dominant bacterial phyla. Proceedings of the National Academy of Sciences. 2009;106: 5859–5864. doi:10.1073/pnas.0901529106

34. Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, Kurm V, et al. Where less may be more: how the rare biosphere pulls ecosystems strings. The ISME Journal. 2017; doi:10.1038/ismej.2016.174

35. Shade A, Jones SE, Gregory Caporaso J, Handelsman J, Knight R, Fierer N, et al. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. mBio. 2014;5. doi:10.1128/mBio.01371-14

36. Rojo D, Gosalbes MJ, Ferrari R, Pérez-Cobas AE, Hernández E, Oltra R, et al. Clostridium difficile heterogeneously impacts intestinal community architecture but drives stable metabolome responses. The ISME Journal. 2015;9: 2206–2220. doi:10.1038/ismej.2015.32

37. Abraham NM, Liu L, Jutras BL, Yadav AK, Narasimhan S, Gopalakrishnan V, et al. Pathogen-mediated manipulation of arthropod microbiota to promote infection. Proceedings of the National Academy of Sciences. 2017; 201613422. doi:10.1073/pnas.1613422114

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

39. Sheik CS, Jain S, Dick GJ. Metabolic flexibility of enigmatic SAR324 revealed through metagenomics and metatranscriptomics. Environmental Microbiology. 2014;16: 304–317. doi:10.1111/1462-2920.12165

40. Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, et al. Relating the metatranscriptome and metagenome of the human gut. Proceedings of the National Academy of Sciences. 2014;111: E2329–E2338. doi:10.1073/pnas.1319284111

41. Jorth P, Turner KH, Gumus P, Nizam N, Buduneli N, Whiteley M. Metatranscriptomics of the human oral microbiome during health and disease. mBio. 2014;5. doi:10.1128/mBio.01012-14

42. Reeves AE, Koenigsknecht MJ, Bergin IL, Young VB. Suppression of Clostridium difficile in the gastrointestinal tracts of germfree mice inoculated with a murine isolate from the family Lachnospiraceae. Infection and Immunity. 2012;80: 3786–3794. doi:10.1128/IAI.00647-12

43. Lawley TD, Clare S, Walker AW, Stares MD, Connor TR, Raisen C, et al. Targeted Restoration of the Intestinal Microbiota with a Simple, Defined Bacteriotherapy Resolves Relapsing Clostridium difficile Disease in Mice. PLoS Pathogens. 2012;8. doi:10.1371/journal.ppat.1002995

44. Petrof EO, Gloor GB, Vanner SJ, Weese SJ, Carter D, Daigneault MC, et al. Stool substitute transplant therapy for the eradication of Clostridium difficile infection: ‘RePOOPulating’ the gut. Microbiome. 2013;1: 3. doi:10.1186/2049-2618-1-3

45. Adlerberth I, Huang H, Lindberg E, Åberg N, Hesselmar B, Saalman R, et al. Toxin-Producing clostridium difficile strains as long-term gut colonizers in healthy infants. Journal of Clinical Microbiology. 2014;52: 173–179. doi:10.1128/JCM.01701-13

46. Wilson KH, Kennedy MJ, Fekety FR. Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. Journal of Clinical Microbiology. 1982;15: 443–446.

47. Leslie JL, Huang S, Opp JS, Nagy MS, Kobayashi M, Young VB, et al. Persistence and toxin production by Clostridium difficile within human intestinal organoids result in disruption of epithelial paracellular barrier function. Infection and Immunity. 2015;83: 138–145. doi:10.1128/IAI.02561-14

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

49. Lopez-Medina E, Neubauer MM, Pier GB, Koh AY. RNA isolation of Pseudomonas aeruginosa colonizing the murine gastrointestinal tract. Journal of visualized experiments : JoVE. 2011; 6–9. doi:10.3791/3293

50. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. 2011;17: 10. doi:10.14806/ej.17.1.200

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

52. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2014;31: 1674–1676. doi:10.1093/bioinformatics/btv033

53. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. 1999. pp. 29–34. doi:10.1093/nar/27.1.29

54. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nature methods. 2015;12: 59–60. doi:10.1038/nmeth.3176

55. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods. 2012;9: 357–9. doi:10.1038/nmeth.1923

56. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078–2079. doi:10.1093/bioinformatics/btp352

57. Breiman L. Random forests. Machine Learning. 2001;45: 5–32. doi:10.1023/A:1010933404324

58. Segal MR. Machine Learning Benchmarks and Random Forest Regression. Biostatistics. 2004; 1–14. Available: http://escholarship.org/uc/item/35x3v9t4.pdf

Figure Legends

Figure 1 | Experimental models of *C. difficile* infection and distinct virulence patterns. (A) Experimental timelines of antibiotic pretreatment and infection mouse models. **(B)** *C. difficile* 630 load 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* load 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 $p < 0.001$).

Figure 2 | Impact of *C. difficile* colonization on other bacterial populations abundances in the gut microbiota. (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 | Metabolite markers of *C. difficile* 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 and out-of-bag error from internal cross-validation is labeling along the bottom axis. 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 | *C. difficile* infection has differential effects 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 | *C. difficile* infection alters community-level select pathway and gene expression of the gut microbiota across perturbed communities

Metatranscriptomic (cDNA) read abundances associated with each gene were normalized to their associated metaganomic (DNA) coverage, resulting in values that reflect upregulation. **(A)** Differences in read abundance for top 5 KEGG pathways with the largest amount of change between Mock and *C. difficile*-infected states within each antibiotic pretreatment. Values above the center line were expressed more during *C. difficile* infection, and those below the line were expressed more in Mock-infected animals. Dotted lines indicate average expression of pathways associated with each condition. **(B-D)** cDNA read abundances of the top 10 genes with the largest differences in expression within each indicated antibiotic pretreatment group. Shown are the expression levels for the genes displayed on the left during Mock (white) or *C. difficile* (black) infection. Gene names and member pathways indicated by letter codes along the left axis (pathway legend below).

Figure 6 | *C. difficile* colonization alters gene expression of taxonomic groups differentially between antibiotic pretreatments. Each point represents a unique gene from the respective metagenomic assembly. Coordinates were determined by the $\log_2$-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 arbitrary units or AU. 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.

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 | Gene-level expression for each community compared to the level expressed in resistant mice for each gene Murine cecal metatransctipromic sequencing results. Both reads and genes atributed to *C. difficile* were removed prior to analysis to focus on the changes in the community caused by infection. Shown are genes in each groups with the largest disparity from a metatranscriptome from a community resistant to *C. difficile* colonization. Panels are displayed as follows; Streptomycin pretreatment: **(A)** mock infection, **(B)** *C. difficile* infection. Cefoperazone pretreatment: **(C)** mock infection, **(D)** *C. difficile* infection. Clindamycin pretreatment: **(E)** mock infection, **(F)** *C. difficile* infection.

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

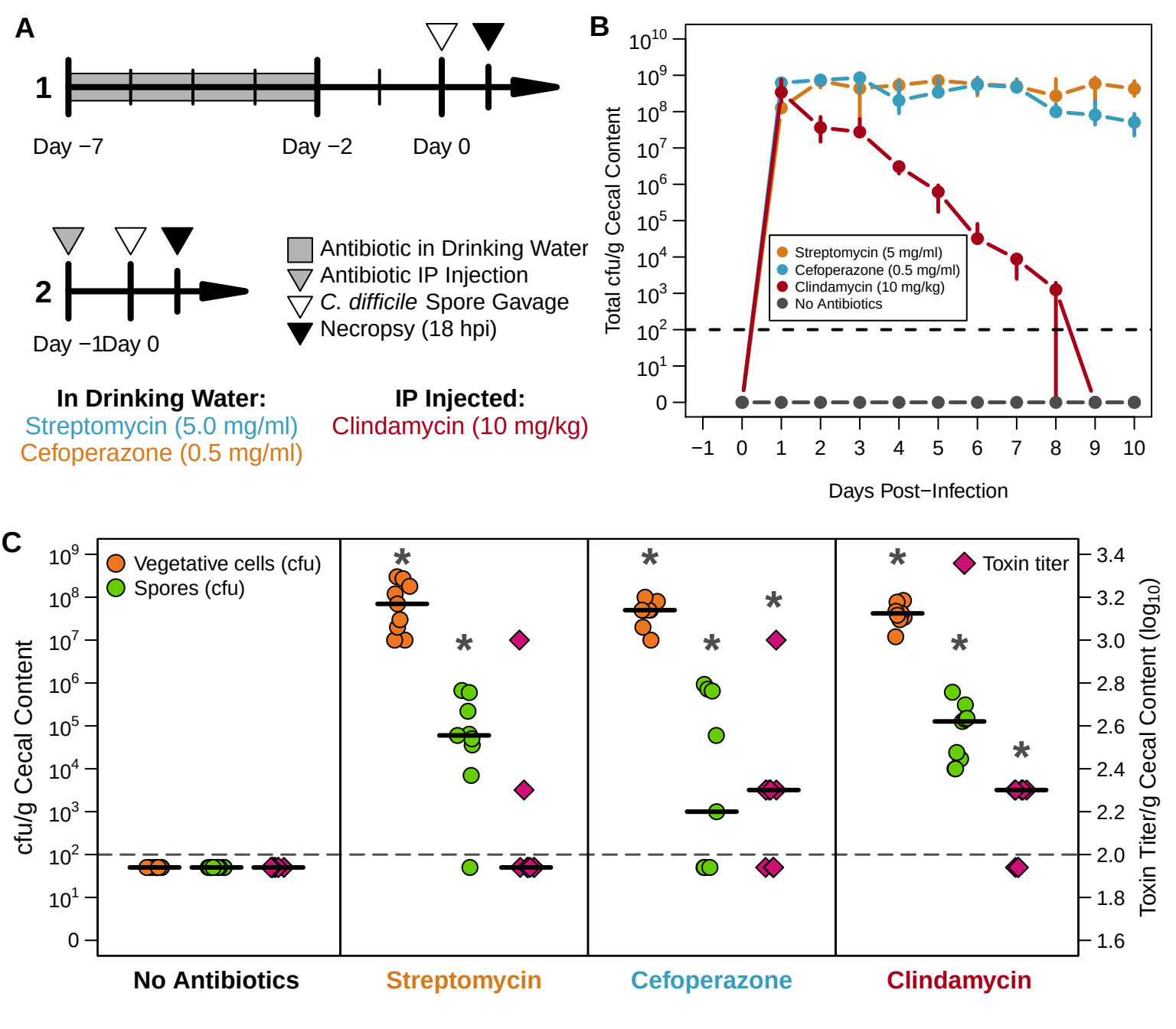
Supplementary Table 2 | High-throughput sequencing read counts and metagenomic assembly quality Raw and curated read abundances for both metagenomic and metatranscriptomic sequencing efforts. Raw read curation steps are

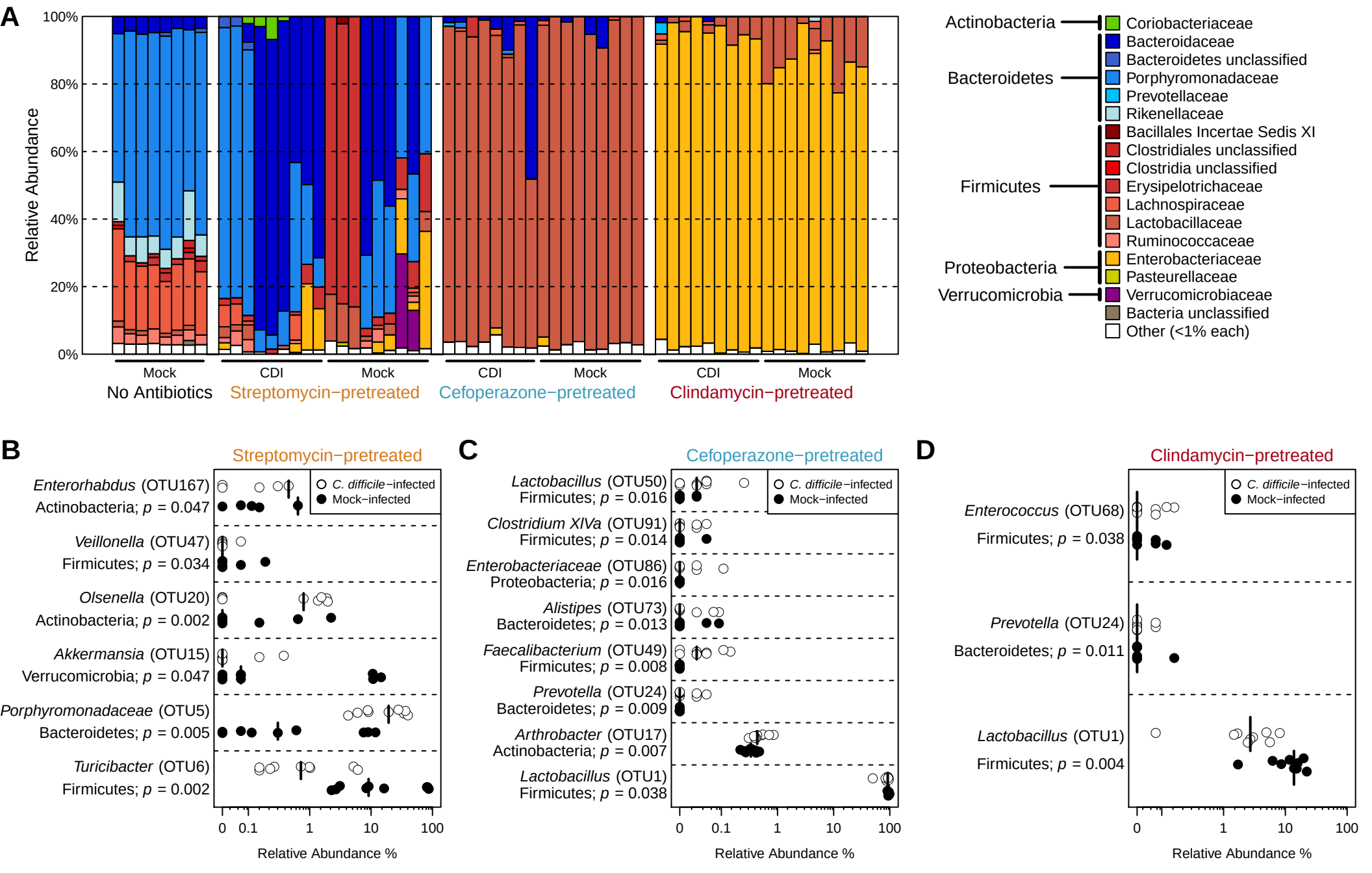
outlined in Materials & Methods. Metagenomic contig summary statistics reflect the quality of assembly for each group.

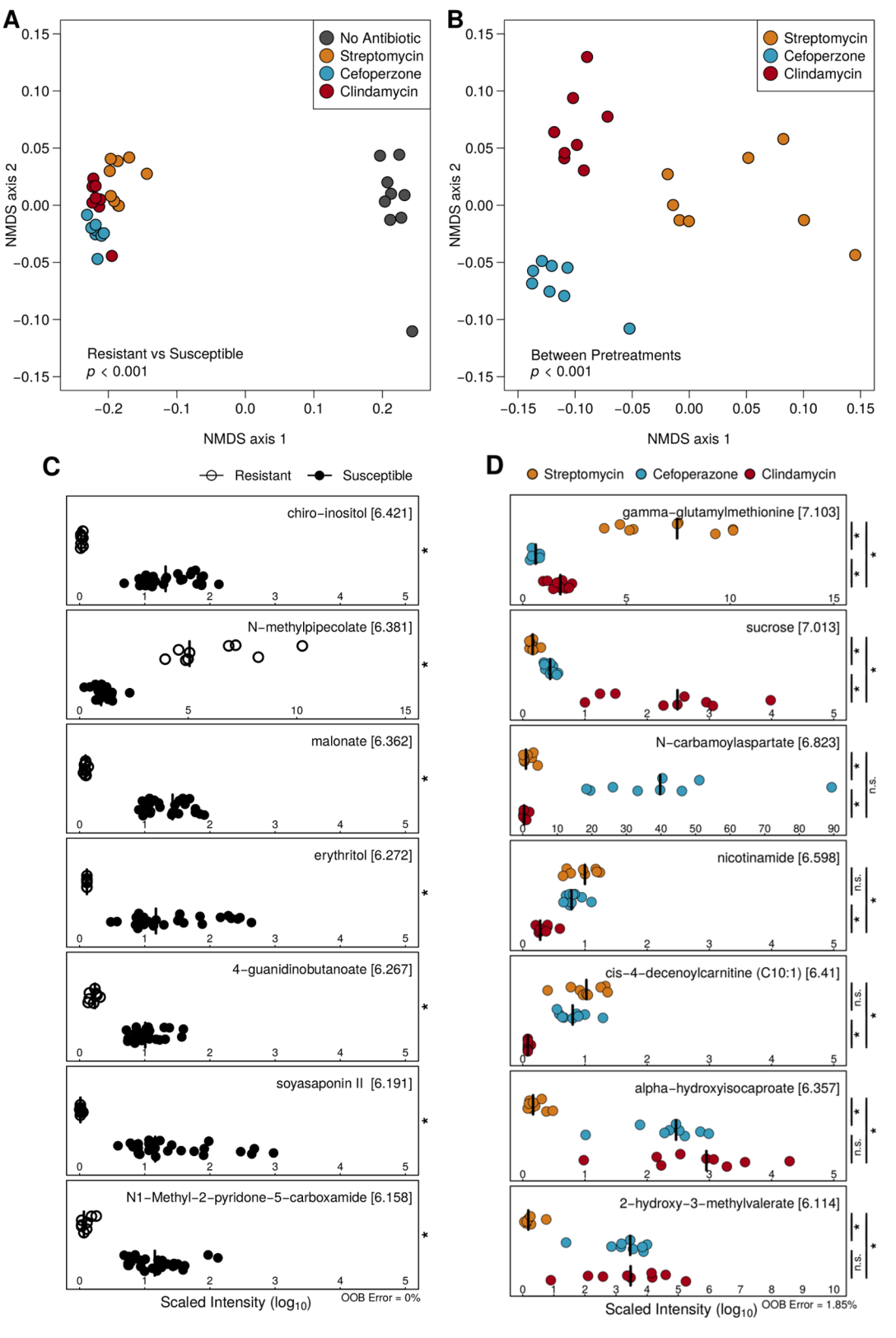
Supplementary Table 3 | Normalized cDNA abundances for each antibiotic pretreated group Final cDNA read abundances for each gene from the corresponding metagenomic assembly. Normalization reflects standardizing cDNA read abundances to both read length and target gene length as well as metagenomic coverage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

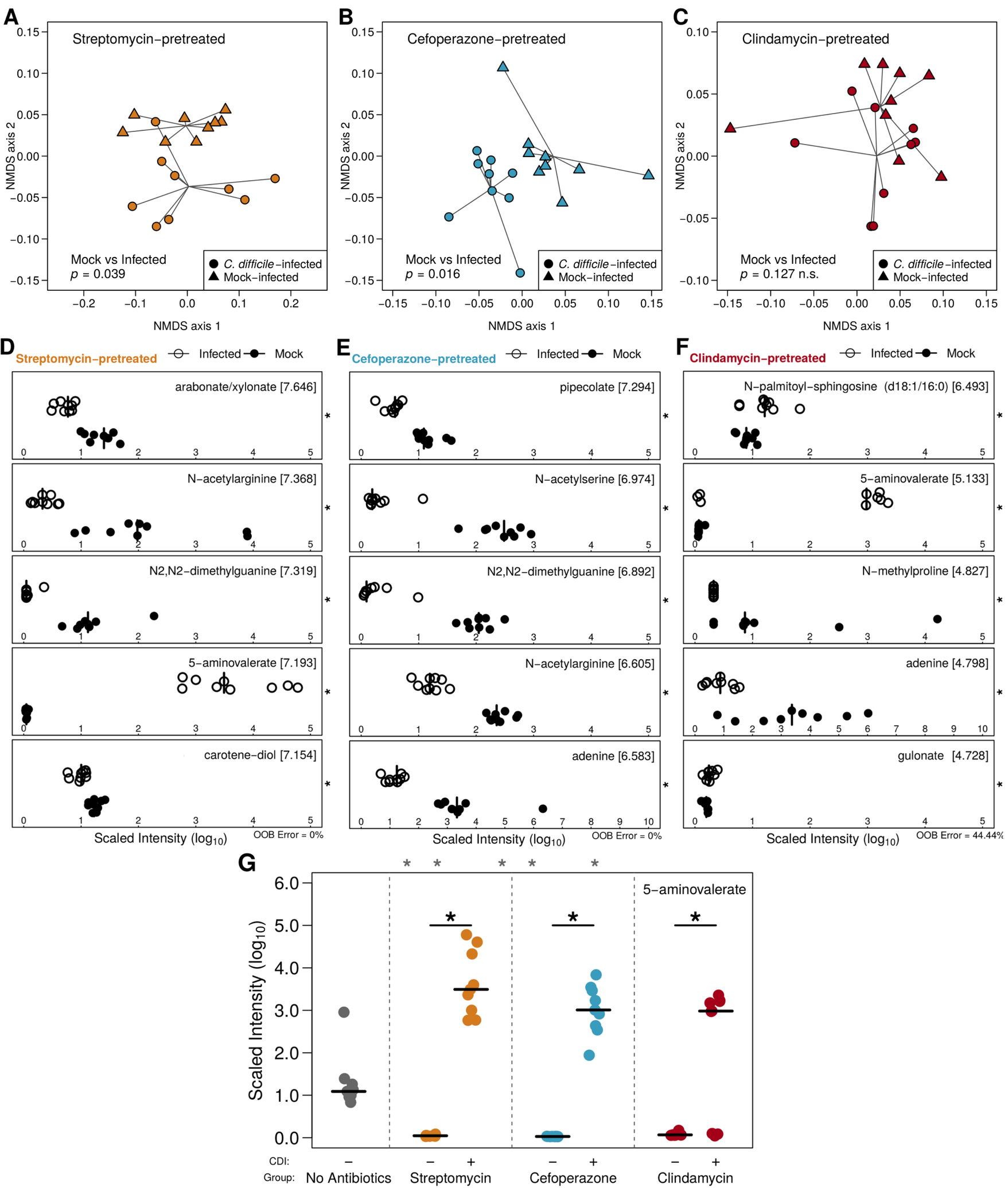
Supplementary Table 4 | Median scaled intensity of metabolites for each antibiotic pretreatment and infection group

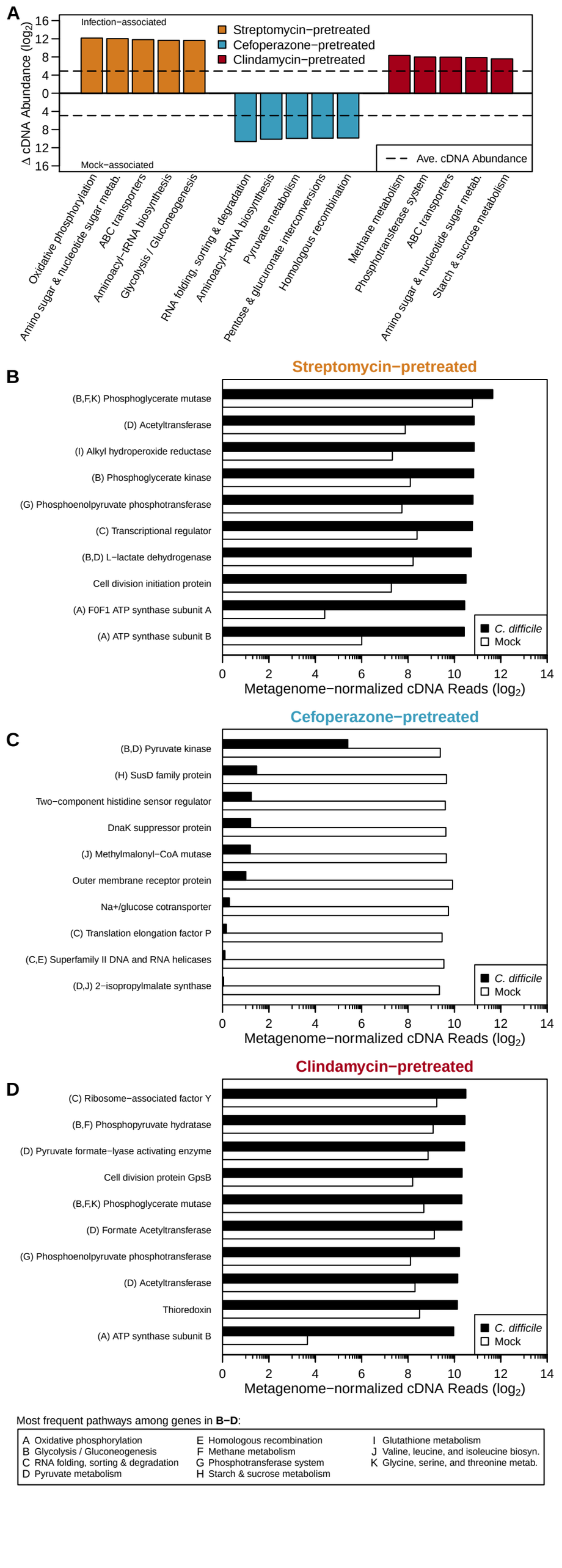
Supplementary Figure 5 | Heatmap comparing Resistant and Susceptible mice for all measured metabolites Relative concentrations of metabolites in each animal across all expremerimental groups. Sample names are listed along the right side, and *C. difficile* susceptibility is along the left. Hierarchical clustering was perormed for each KEGG compound category included in the untargeted metabolomic analysis separately and are listed in order under the figure panel.

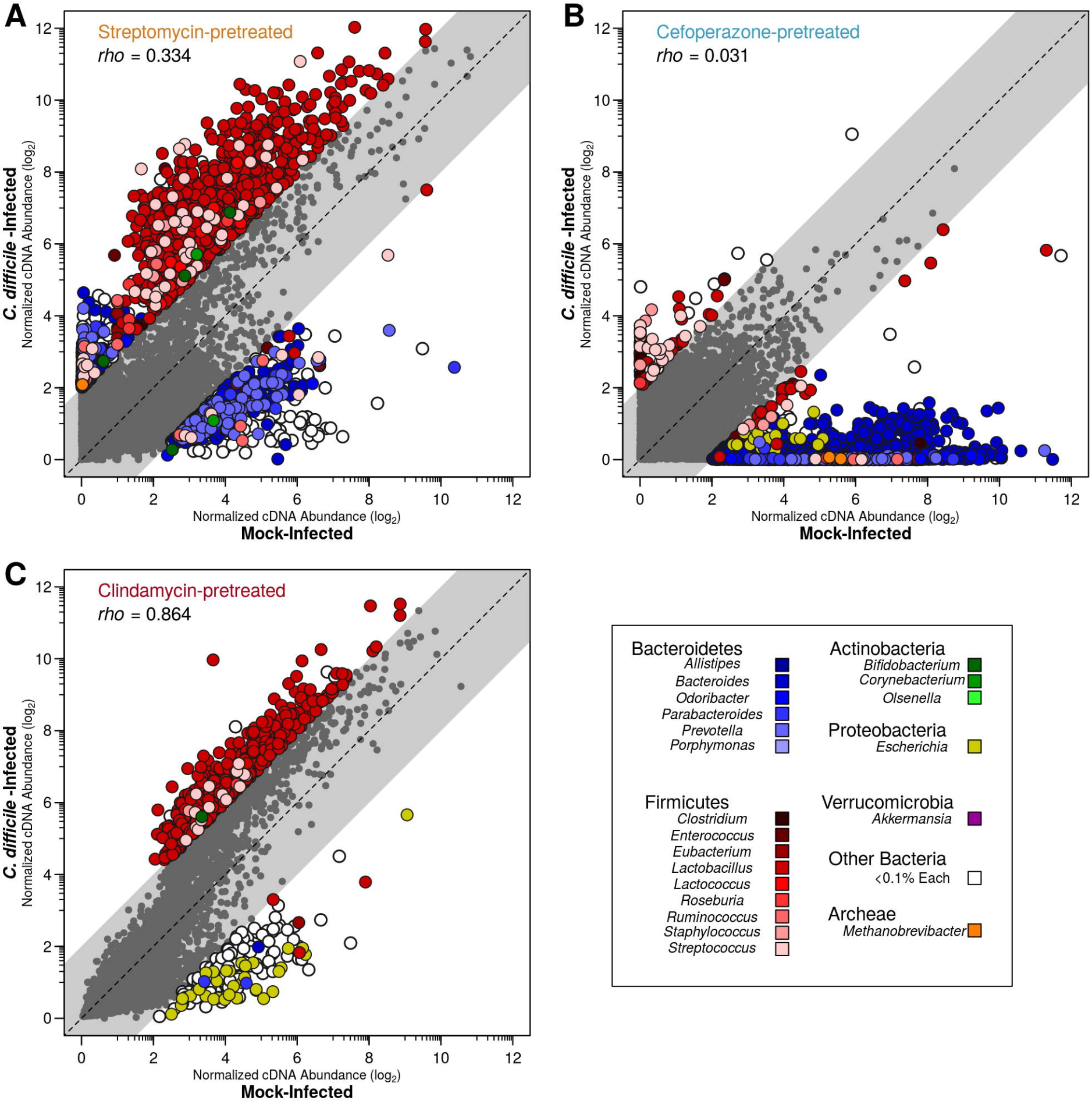

A
1
Day −7
Day −2
Day 0
2
Day −1
Day 0
Antibiotic in Drinking Water
Antibiotic IP Injection
C. difficile Spore Gavage
Necropsy (18 hpi)
In Drinking Water:
Streptomycin (5.0 mg/ml)
Cefoperazone (0.5 mg/ml)
IP Injected:
Clindamycin (10 mg/kg)
B
Total cfu/g Cecal Content
Days Post−Infection
Streptomycin (5 mg/ml)
Cefoperazone (0.5 mg/ml)
Clindamycin (10 mg/kg)
No Antibiotics
C
cfu/g Cecal Content
Toxin Titer/g Cecal Content (log10)
Vegetative cells (cfu)
Spores (cfu)
Toxin titer
No Antibiotics
Streptomycin
Cefoperazone
Clindamycin


A
Relative Abundance
100%
80%
60%
40%
20%
0%
Mock
CDI
Mock
CDI
Mock
CDI
Mock
No Antibiotics
Streptomycin−pretreated
Cefoperazone−pretreated
Clindamycin−pretreated
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
Verrucomicrobia
Coriobacteriaceae
Bacteroidaceae
Bacteroidetes unclassified
Porphyromonadaceae
Prevotellaceae
Rikenellaceae
Bacillales Incertae Sedis XI
Clostridiales unclassified
Clostridia unclassified
Erysipelotrichaceae
Lachnospiraceae
Lactobacillaceae
Ruminococcaceae
Enterobacteriaceae
Pasteurellaceae
Verrucomicrobiaceae
Bacteria unclassified
Other (<1% each)
B
Streptomycin−pretreated
C. difficile−infected
Mock−infected
Enterorhabdus (OTU167)
Actinobacteria; p = 0.047
Veillonella (OTU47)
Firmicutes; p = 0.034
Olsenella (OTU20)
Actinobacteria; p = 0.002
Akkermansia (OTU15)
Verrucomicrobia; p = 0.047
Porphyromonadaceae (OTU5)
Bacteroidetes; p = 0.005
Turicibacter (OTU6)
Firmicutes; p = 0.002
0
0.1
1
10
100
Relative Abundance %
C
Cefoperazone−pretreated
C. difficile−infected
Mock−infected
Lactobacillus (OTU50)
Firmicutes; p = 0.016
Clostridium XIVa (OTU91)
Firmicutes; p = 0.014
Enterobacteriaceae (OTU86)
Proteobacteria; p = 0.016
Alistipes (OTU73)
Bacteroidetes; p = 0.013
Faecalibacterium (OTU49)
Firmicutes; p = 0.008
Prevotella (OTU24)
Bacteroidetes; p = 0.009
Arthrobacter (OTU17)
Actinobacteria; p = 0.007
Lactobacillus (OTU1)
Firmicutes; p = 0.038
0
0.1
1
10
100
Relative Abundance %
D
Clindamycin−pretreated
C. difficile−infected
Mock−infected
Enterococcus (OTU68)
Firmicutes; p = 0.038
Prevotella (OTU24)
Bacteroidetes; p = 0.011
Lactobacillus (OTU1)
Firmicutes; p = 0.004
0
1
10
100
Relative Abundance %


A
No Antibiotic
Streptomycin
Cefoperzone
Clindamycin
NMDS axis 2
NMDS axis 1
Resistant vs Susceptible
p < 0.001
B
Streptomycin
Cefoperzone
Clindamycin
NMDS axis 2
NMDS axis 1
Between Pretreatments
p < 0.001
C
Resistant
Susceptible
chiro−inositol [6.421]
N−methylpipecolate [6.381]
malonate [6.362]
erythritol [6.272]
4−guanidinobutanoate [6.267]
soyasaponin II [6.191]
N1−Methyl−2−pyridone−5−carboxamide [6.158]
Scaled Intensity (log10)
OOB Error = 0%
D
Streptomycin
Cefoperazone
Clindamycin
gamma−glutamylmethionine [7.103]
sucrose [7.013]
N−carbamoylaspartate [6.823]
nicotinamide [6.598]
cis−4−decenoylcarnitine (C10:1) [6.41]
alpha−hydroxyisocaproate [6.357]
2−hydroxy−3−methylvalerate [6.114]
n.s.
Scaled Intensity (log10)
OOB Error = 1.85%


A
Streptomycin−pretreated
Mock vs Infected
p = 0.039
C. difficile−infected
Mock−infected
NMDS axis 1
NMDS axis 2
B
Cefoperazone−pretreated
Mock vs Infected
p = 0.016
C
Clindamycin−pretreated
Mock vs Infected
p = 0.127 n.s.
D
Streptomycin−pretreated
Infected
Mock
arabonate/xylonate [7.646]
N−acetylarginine [7.368]
N2,N2−dimethylguanine [7.319]
5−aminovalerate [7.193]
carotene−diol [7.154]
Scaled Intensity (log10)
OOB Error = 0%
E
Cefoperazone−pretreated
pipecolate [7.294]
N−acetylserine [6.974]
N2,N2−dimethylguanine [6.892]
N−acetylarginine [6.605]
adenine [6.583]
OOB Error = 0%
F
Clindamycin−pretreated
N−palmitoyl−sphingosine (d18:1/16:0) [6.493]
5−aminovalerate [5.133]
N−methylproline [4.827]
adenine [4.798]
gulonate [4.728]
OOB Error = 44.44%
G
5−aminovalerate
Scaled Intensity (log10)
CDI:
Group:
No Antibiotics
Streptomycin
Cefoperazone
Clindamycin


A
Δ cDNA Abundance (log2)
Infection−associated
Mock−associated
Streptomycin−pretreated
Cefoperazone−pretreated
Clindamycin−pretreated
Ave. cDNA Abundance
Oxidative phosphorylation
Amino sugar & nucleotide sugar metab.
ABC transporters
Aminoacyl−tRNA biosynthesis
Glycolysis / Gluconeogenesis
RNA folding, sorting & degradation
Aminoacyl−tRNA biosynthesis
Pyruvate metabolism
Pentose & glucuronate interconversions
Homologous recombination
Methane metabolism
Phosphotransferase system
ABC transporters
Amino sugar & nucleotide sugar metab.
Starch & sucrose metabolism
B
Streptomycin−pretreated
(B,F,K) Phosphoglycerate mutase
(D) Acetyltransferase
(I) Alkyl hydroperoxide reductase
(B) Phosphoglycerate kinase
(G) Phosphoenolpyruvate phosphotransferase
(C) Transcriptional regulator
(B,D) L−lactate dehydrogenase
Cell division initiation protein
(A) F0F1 ATP synthase subunit A
(A) ATP synthase subunit B
C. difficile
Mock
Metagenome−normalized cDNA Reads (log2)
C
Cefoperazone−pretreated
(B,D) Pyruvate kinase
(H) SusD family protein
Two−component histidine sensor regulator
DnaK suppressor protein
(J) Methylmalonyl−CoA mutase
Outer membrane receptor protein
Na+/glucose cotransporter
(C) Translation elongation factor P
(C,E) Superfamily II DNA and RNA helicases
(D,J) 2−isopropylmalate synthase
C. difficile
Mock
Metagenome−normalized cDNA Reads (log2)
D
Clindamycin−pretreated
(C) Ribosome−associated factor Y
(B,F) Phosphopyruvate hydratase
(D) Pyruvate formate−lyase activating enzyme
Cell division protein GpsB
(B,F,K) Phosphoglycerate mutase
(D) Formate Acetyltransferase
(G) Phosphoenolpyruvate phosphotransferase
(D) Acetyltransferase
Thioredoxin
(A) ATP synthase subunit B
C. difficile
Mock
Metagenome−normalized cDNA Reads (log2)
Most frequent pathways among genes in B−D:
A Oxidative phosphorylation
B Glycolysis / Gluconeogenesis
C RNA folding, sorting & degradation
D Pyruvate metabolism
E Homologous recombination
F Methane metabolism
G Phosphotransferase system
H Starch & sucrose metabolism
I Glutathione metabolism
J Valine, leucine, and isoleucine biosyn.
K Glycine, serine, and threonine metab.


A
Streptomycin-pretreated
rho = 0.334
B
Cefoperazone-pretreated
rho = 0.031
C
Clindamycin-pretreated
rho = 0.864
C. difficile -Infected
Normalized cDNA Abundance (log2)
Mock-Infected
Normalized cDNA Abundance (log2)
Bacteroidetes
Allistipes
Bacteroides
Odoribacter
Parabacteroides
Prevotella
Porphymonas
Firmicutes
Clostridium
Enterococcus
Eubacterium
Lactobacillus
Lactococcus
Roseburia
Ruminococcus
Staphylococcus
Streptococcus
Actinobacteria
Bifidobacterium
Corynebacterium
Olsenella
Proteobacteria
Escherichia
Verrucomicrobia
Akkermansia
Other Bacteria
<0.1% Each
Archeae
Methanobrevibacter