#### Title:

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

#### Authors:

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

#### 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

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

### Abstract

Susceptibility to *Clostridium difficile* infection is primarily associated with previous exposure to antibiotics, which compromise the structure and function of the gut bacterial community. Specific antibiotic classes correlate more strongly with recurrent or persistent *C. difficile* infection. As such, we utilized a mouse model of infection to explore the effect of distinct antibiotic classes on the impact that infection has on community-level transcription and metabolic signatures shortly following pathogen colonization and how those changes may associate with persistence of *C. difficile*. Untargeted metabolomic analysis revealed that *C. difficile* infection has significantly larger impacts on the metabolic environment across cefoperazone and streptomycin-pretreated mice, which become persistently colonized compared to clindamycin-pretreated mice where infection quickly becomes undetectable. Through metagenome-enabled metatranscriptomics we observed that the infected microbial communities were greatly reduced in transcript for genes associated with carbon and energy acquisition, suggesting niches occupied instead by *C. difficile*. Furthermore, the largest degree of change in transcription were seen in the least abundant species within the respective groups indicating that *C. difficile* may “attack the loser” in gut environments where sustained infection occurs more readily. Overall, our results suggest that *C. difficile* is able to restructure the nutrient-niche landscape in the gut in order to promote persistent infection.

### Importance

*Clostridium difficile* has become the most common single cause of hospital-acquired infection over the last decade in the United States and colonization resistance to the nosocomial pathogen is primarily driven by the gut microbiota. This community is also involved in clearing the infection as the community recovers from perturbation. As distinct antibiotics are associated with different risk levels for CDI, we utilized a mouse model of infection with 3 separate antibiotic pretreatment regimes to generate alternative gut microbiomes that each allowed for *C. difficile* colonization but vary in clearance rate. To assess community-level dynamics, we implemented an integrative multi-omic approach that revealed infection significantly shifted many aspects of the gut ecosystem. This revealed a trend that the degree to which this change occurred inversely correlated with clearance during the first six days of infection, suggesting that *C. difficile* may differentially modify the gut environment to promote persistence. This is the first time metagenome-enabled metatranscriptomics have been employed to study the behavior of a host-associated microbiota in response to an infection. Our results allow for previously unseen understanding of the ecology associated with *C. difficile* infection and provides groundwork for identification of context-specific probiotic therapies.

### 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 transcription 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 eliminate *C. difficile* colonization.

### Results

#### Distinct antibiotic pretreatments are associated with alternative community structures that are equally susceptible to initial *C. difficile* colonization, but differ in patterns of clearance.

In previous work from our laboratory it was found that when conventionally-reared SPF 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 (19). Briefly, these drugs were chosen for not only the ability to to reduce *C. difficile* colonization resistance in a mouse model (18), but also for distinct and significant impacts on the structure and diversity of the cecal microbiota (Fig. 1A) (19). In each antibiotic pretreatment model, we observed equally high levels of *C. difficile* colonization on the day after infection, however,*C. difficile* fell below the limit of detection in only clindamycin-pretreated mice over the subsequent 9 days while the other two pretreatments remained highly colonized (*p* = 0.01; Fig. 1A). We hypothesized that this occurred in the clindamycin-pretreated mice because the resultant intestinal community occupied niche space with greater overlap to that of *C. difficile*.

Earlier publications have demonstrated that *C. difficile* virulence factor expression is regulated by availability of certain nutrients in the environment (20), so we first sought to measure spore and toxin production. We chose to focus all downstream experimentation on 18-hours post-infection to the assess behavior of *C. difficile* directly prior to the reduction in detectable *C. difficile*. This end 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, we found that communities indeed remained highly differentiated from untreated controls measured by 16S rRNA gene sequencing of the V4 region, classified to the family-level (Fig. 1B). The composition of streptomycin-pretreated communities was more variable between cages, but was generally enriched for members of phylum *Bacteroidetes*. Cefoperazone and clindamycin-pretreated cecal communities were consistently dominated by members of the *Lactobacillaceae* and *Enterobacteriaceae* families, respectively. Despite variation in the community structures, there were no significant differences in the number of vegetative cells between any antibiotic-pretreatment group (Fig. 1C). All susceptible mice were colonized with ~1x108 vegetative colony forming units (cfu) per gram of cecal content and untreated mice maintained *C. difficile* colonization resistance. We have also previously demonstrated that both *C. difficile* spore production and toxin activity differ between these pretreatment regimes (19). As both processes have been linked to environmental concentrations of specific growth nutrients (20), 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.

#### Multiple biological signatures in the bacterial community and metabolome differentiate cecal microbiomes that remained colonized by *C. difficile* from those that do not.

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 understand the ramifications each antibiotic had on the cecal metabolomic environment, 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 through these means. In combination with our 16S results, we first characterized the differences between the microbiomes (microbiota, plus associated metabolome) of the mock-infected animals to quantify possible drivers of communities that go on to clear infection. To focus our analysis on ascertaining changes in discrete populations within the microbiota, we generated operational taxonomic units (otu) clustered at 97% similarity which historically corresponds with species-level classification. We also removed all *C. difficile* 16S rRNA gene sequences, which represented an average of 2.113% sequencing reads across infection groups to eliminate its direct impact in downstream calculation. Using these methods we discovered that the Bray-Curtis dissimialrity of both the community structure (*p* < 0.001) and metabolome (*p* < 0.001) significantly differ between cleared and colonized groups early during infection (Fig. 2A & 2C). These results partially supported the hypothesis that the cecal environment created by clindamycin pretreatment was highly divergent from the other groups, and likely contributed to the clearance seen in the subsequent days.

To uncover the specific drivers at each level of biology, we then utilized Random Forest machine learning to identify the smallest optimal subset of features that successfully differentiate microbiomes that clear infection and and those that remain colonized after blinded cross-validation (AUCRF) (22). At the level of community membership (16S rRNA sequencing), it resulted it 5 otus which correctly classified all samples to their corresponding groups (Fig. 2B; Out-of-bag error=0%). Interestingly, these otus had very low consistency in differential abundance between groups and are often minority members of their respective communities, outside of a handful of outliers. Despite this these abundances provided a model that clearly distinguished communities that remained colonized by *C. difficile*. Similarly, machine learning with metabolomic results also revealed 5 metabolites which give a model that correctly differentiates the groups (Fig. 2D; Out-of-bag error=0%). Together these results further support that the environment of the cecum, even early during infection, is distinct between groups that clear the infection and those that maintain *C. difficile* at high levels. Furthermore, results from machine learning analysis indicate that it may be members in the minority of relative abundance that have disproportionate influence on the clearance patterns observed between pretreatment regimes and that changes in community structure may be less consistent than at other levels of biology.

#### Amino-acid metabolism by *C. difficile* appears important for sustained colonization across susceptible environments.

It has been recently established that the catabolism of amino acids (Stickland fermentation) by *C. difficile* may be a critical nutrient niche occupied by the pathogen early during infection in one antibiotic pretreament model of CDI (23). We were interested to observe if this behavior was conserved across multiple distinct gut environments where *C. difficile* was permitted to colonize. Utilizing the previously described approaches, we assessed the the distinction between any antibiotic-pretreated, mock-infected microbiomes and those of untreated, *C. difficile*-resistant animals. Not only was a Stickland substrate increased across susceptible environments but several secondary bile acids, which have been shown to be negatively correlated with *C. difficile* susceptibility were significantly decreased (Fig. S1D; *p* < 0.001). Additionally, multiple members of Clostridia which may also capable of metabolizing amino acids for growth (24) were selected by AUCRF and were significantly decreased during susceptibility (Fig. S1B; *p* < 0.001). Conversely, we also performed a similar analysis to investigate changes induced by *C. difficile* colonization itself in these susceptible conditions. This revealed that although CDI alone did not induce significant shifts in the global community structure or metabolome (Fig. S2 A & C; *p* = 0.185, 0.065), several features were able to discriminate infected and uninfected microbiomes with relatively OOB error. This analysis highlighted numerous known *C. difficile* growth substrates as enriched in all pretreatments including 6 Stickland substrates, 4 of which are proline conjugates, along with another carbon source arabonate/xlyonate (Fig. S2D). Furthermore 5-aminovalerate, the most common endproduct of Stickland fermentation, was significantly increased during infection in almost all metabolomes measured. More discrete inspection of these specific metabolites revealed that clindamycin pretreatment was only condition where both the inputs and outputs of Stickland fermentation were less abundant compared to controls (Fig. S3). These results strongly support Stickland fermentation as a primary nutritional strategy of *C. difficile* early in infection. Moreover, these data further suggest the degree to which the environment of the intestine is altered by infection may be linked to the ability of the pathogen to persist within those gut environments.

#### Infection corresponded with larger shifts in the metatranscriptomes of communities which allowed sustained *C. difficile* colonization.

Despite shared signal between community structure and metabolome in the cecal environments of interest, it is difficult to connect specific microbes with changes in concentration of particular metabolites that may ultimately influence the duration of infection. To gain a more specific understanding of how the microbiota or *C. difficile* shaped the metabolic environment, we employed parallel metagenomic and metatranscriptomic shotgun sequencing following paired DNA and RNA samples collected from the cecal content of the mice used in the previous analyses. To achieve usable concentrations of bacterial mRNA after rRNA depletion, we were required to performed a combined sequencing effort for each treatment and infection group. To establish confidence in continuing pooled analysis, we calculated within-group sample variance replicates from cfu, otu relative abundance, and untargeted metabolomic (Table S3). With only limited outliers, this revealed low levels of variance within control and experimental groups at multiple levels of biology. Following sequencing, metagenomic reads from mock-infected cecal communities were assembled *de novo* into contigs and putative genes were identified resulting in 234,868 (streptomycin), 83,534 (cefoperazone), and 35,681 (clindamycin) likely open reading frames in each metagenome. Of these putative genes, 28.5% were able to be annotated with known function KEGG, and many of these annotations were homologs among species within the dataset. Streptomycin pretreatment resulted in a significantly more diverse community (Table S1) than other groups, so a more diverse metagenome was expected. 2408 unique functionally annotated genes were detected in the metagenome resulting from strptomycin pretreatment, at least 1163 more genes than either cefoperazone or clindamycin (Fig. S4A-D). Metagenome-enabled mapping of both metagenomic and metatranscriptomic reads revealed that we were able to obtain informative depths of sequencing that adequately sampled the landscape of shotgun libraries (Fig. S4E-F). As expected, genes with any detectable transcript in any metatranscriptome were a subset of their corresponding metagenome. Metatranscriptomic read abundances were normalized to corresponding metagenomic coverage per gene in order to normalize for abundance of the contributing bacterial taxa. This step was followed by a final even subsampling of reads between conditions in order to create an equal standard of comparison across groups and identify genes with the largest increases in transcription above baseline.

We hypothesized that degree of change in the metatranscriptome corresponding with *C. difficile* colonization would reflect the concurrent shifts seen at the metabolomic-level. As disparate bacterial taxa possess vastly different metabolic capabilities, and antibiotic induced distinct species profiles of each community, we addressed this question by delineating the transcriptomic contributions of separate bacterial taxa within each metatranscriptomic dataset. Additionally since many genes lack a specific functional annotation in KEGG but retain general taxonomic information, we continued the analysis at the genus level of classification for all genes contributed to each metagenome. Using this approach, we 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 transcription (Fig. 3). 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 transcription outliers between uninfected and infection conditions compared with those of the other antibiotic groups. This suggested that the degree to which the metatranscriptome was shifted by infection corresponded to prolonged colonization.

Perhaps even more prominently, this analysis revealed that outlier genes were largely contributed by underrepresented genera across infection models. In streptomycin-pretreated mice, the largest differences were in 937 genes belonging to *Lactobacillus* with increased transcription during *C. difficile* infection (Fig. 3A) which represented a median proportion of ~1.25% of the overall community (Fig. 1B). In cefoperazone-pretreated mice 2290 genes belonging to *Bacteroides* had higher transcription in mock infected mice (Fig. 3B) and composed a median of ~1.35% of the cecal communities in these mice (Fig. 1B). A consistent trend in streptomycin and cefoperazone pretreatments was an overrepresentation of highly transcribed 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* = 0.334 & *R* = 0.031). In clindamycin-pretreated mice the largest difference in transcription was for 510 *Lactobacillus* genes with increased transcription during CDI (Fig. 3C). Infected and uninfected metatranscriptomes associated with this antibiotic more strongly correlated than either other pretreatment (*R* = 0.864). These data also supported that the transcriptional activity of lowly abundant bacterial taxa experienced larger degrees of change during infection by *C. difficile*.

#### Largest amount of metatranscriptomic changes in response to infection are concentrated in the minority taxa of each pretreatment group.

To explore the concept of increased transcriptional shift in the lowest abundance taxa responding to infection we tabulated the absolute difference between mock-infection and CDI in all transcriptomic abundances per genus in each antibiotic pretreatment. We further normalized these values for the number of genes detected per each genera to adjust for species that were more successfully assembled or annotated and we eliminated genera where less than 50 genes were detected per group. Taxa were then stratified into categories based on their relative abundance in each community from 16S rRNA gene sequencing (Fig. 4). This revealed that indeed most change occurred within relatively rare bacterial taxa, and their degree of change inversely correlated with sustained colonization. To this point, metatranscriptomic changes in lowly abundant taxonomic groups were significantly reduced in clindamycin pretreatment (*p* < 0.001). Additionally, the proportion of taxa in the lowest abundance bracket was largely similar across pretreatment groups (~88.9%). As a corollary, it stood to reason that the majority of unique genes or metabolic potential was possessed by lowly abundant species, and subsequent quantification supported this (Table S4). As a consequence, the downstream impacts on functionality may affect a disproportionately large effect in the overall environment of the intestine as a function of their collective metabolism. We then hypothesized that genes of pathways with the largest degree of change in communities where *C. difficile* maintains colonization may be closely linked to the acquisition of resources that the pathogen prefers for growth.

#### Altered transcription within minority taxa favors reduced nutrient competition with *C. difficile* in communities that permit sustained colonization.

Based on the previous combine metabolomic and metatranscriptomic results, we hypothesized that pathways with the greatest differences would be related to catabolism of metabolites usable by *C. difficile* for growth. In order to assess these changes, we cross referenced genus-level classification of 16S results (Materials & Methods) below the 0.1% relative abundance threshold in each community with the genes attributed to those genera in the metagenome-enabled metatranscriptome mapping. We further focused this analysis on genes that had also received functional KEGG annotation in addition to taxonomic classification (Fig. 5). This resulted in 585 genes differentially transcribed between clindamycin pretreatment and both other groups that allow for sustained *C. difficile* colonization. We then subset each metatranscriptome to genes unique to either the pretreatment group which cleared *C. difficile* or to those that did not. Lastly, we limited these groups to those genes that had any degree of change between mock and *C. difficile*-infected groups in each antibiotic pretreatment respectively leaving 29 genes in 11 total pathways. Included at the end was also a comparison of transcription for each gene in untreated mice to provide a reference to the degree that each perturbation and subsequent pathogen colonization impacts transcription in the gut microbiota.

Results from this analysis ultimately revealed that primarily genes and pathways involved in simple carbohydrate-containing molecule acquisition/utilization were differentially overrepresented in infected conditions across pretreatment groups when compared to mock-infection mice (Fig. 5). Interestingly, a much larger proportion of these genes have decreased transcription during infection compared to mock-infected controls. At the pathway-level, many genes associated with galactose and amino sugar acquisition (both *C. difficile* growth substrates) are reduced during infection in both streptomycin and cefoperazone pretreatments. Conversely, pathways uniquely associated with clindamycin-pretreated cecal communities were related to a diverse array of carbon sources, and could indicate ineffective competition by *C. difficile* with this community for any particular subset of resources. Parallel to these results from differential analysis, consistent changes to genes in the starch/sucrose metabolism pathway across all conditions tested may indicate some level of preference and competition for these carbohydrates the pathogen during colonization of each tested conditions. These data agree with concordant in the metabolome where sucrose is significantly associated with discriminating metabolomes that clear and those that remain colonized (Fig. 2D). Our results indeed suggest that *C. difficile* colonization induces a shift transcriptional activity for a minority subset of species, possibly in an effort to segregate a desired nutrient niche, prior to the introduction of the hallmark disease phenotypes associated with CDI.

### 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, metatranscriptomic activity, and metabolome composition 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. Instances of active nutrient niche restructuring in the gut have been documented previously for prominent symbiotic bacterial species in gnotobiotic mice (25), 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. 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 (26). 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 (27), 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. 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 (28), 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 (29). 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 (30). 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 the metatranscriptome. 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 transcription 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, 31–33). 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 (34–37). 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 (38). 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 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.

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 (genders were housed separately). 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). A detailed description of these animal models was outlined previously (19). A similar experimental design was implemented for gnotobiotic mice and was performed with the University of Michigan Germfree Mouse Center as described previously (19). 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* 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 *C. difficile* 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 (39). Quantification of total *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.

#### 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 (40). 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 (41). 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 rRNA Removal Kit (Epidemiology) was then used to deplete prokaryotic and eukaryotic rRNA from the samples according the manufacturer’s protocol (Illumina). 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 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 (40). The 16S rRNA gene sequences were curated using the mothur software package (v1.36) and otu-based analysis was performed as described in (19). Genus-level classification-based analysis of 16S data was accomplished using the phylotype workflow in mothur and the full Silva bacterial taxonomy (release 132). 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 280,000,000 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 normalize for technical variation between lans and to obtain necessary coverage (32). This gave an average of 380,000,000 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 (42) on a per library basis. Reads were quality trimmed using Sickle (43) 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 S4.

#### Metagenomic contig assembly and gene annotation.

Metagenomic contigs were assembled using Megahit (44) 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 (45) using Diamond implementation of BLASTp (46). 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 (47) mapping databases from downstream analyses.

#### DNA/cDNA read mapping and normalization.

Mapping was accomplished using Bowtie2 (47) 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 (48) 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 transcription 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 for up to 2 mismatches.X

#### 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.

All statistical analyses were performed using R (v.3.2.0) and the vegan package (49). 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. Dynamic time warping was performed with the dtw package in R (50). Random forest was performed using the AUCRF implementation (22) as well as the standard package (51) in R. Distances of outlier points from center line during metatranscriptomic comparisons was accomplished using 2-dimensional linear geometry.

#### 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.

#### 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.

#### Data Availability

Pooled and quality trimmed *C. difficile*-infected metatranscriptomes (SRA; PRJNA354635) and 16S rRNA gene amplicon read data (SRA; PRJNA383577) from infection experiments are available through the NCBI Sequence Read Archive. Metagenomic reads and 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>.

### Figure and Table Legends

#### Figure 1 | Distinct antibiotic pretreatments have differential impacts on *C. difficile* colonization and cecal microbiota community structure.

**(A)** *C. difficile* 630 cfu in stool of infected mice following each antibiotic-pretreated group over 10 days of infection. Median and interquartile range are shown for each time point. Both cefoperzone and streptomycin pretreatments had more significantly detectable cfu on the final day of detectable cfu associated with clindamycin-pretreatment (*p* < 0.001). **(B)** Relative abundance of family-level OTU taxonomic classification in each pretreatment group from 16S rRNA gene sequencing. **(C)** Quantification of terminal vegetative *C. difficile* cfu in cecal content across 18 hour colonization models. Black lines indicate median values and each pretreatment group had significantly greater detectable cfu than no antibiotic controls. Significant differences in A were determined through permANOVA wth dynamic time warping and in C were found by Wilcoxon rank-sum test with Benjamni-Hochberg correction when necessary. The limit of detection was used in place of undetectable values for statistical testing.

#### Figure 2 | Significant differences in cecal community structure and metabolomes track with downstream *C. difficile* clearance across antibiotic pretreatment regimes.

**(A)** NMDS ordination of Bray-Curtis distances of otu relative abundances between mouse cecal communities that remain colonized by *C. difficile* and those that eventually clear the infection. **(B)** Relative abundance of OTUs included the optimal model generated by AUCRF discriminating the same groups as in panel A. **(C)** NMDS ordination of Bray-Curtis distances using metabolite intensities between the aforementioned groups of animals. **(D)** Scaled intensity of metabolites included the optimal model generated by AUCRF discriminating colonized and clearing mouse cecal microbiomes. Differences for ordinations in A & C were calculated using permANOVA. Optimal AUCRF models demonstrated 0% out of bag error, and significant differences in B & D were determined by Wilcoxon rank-sum test with Benjamni-Hochberg correction.

#### Figure 3 | *C. difficile* colonization alters gene transcription of taxonomic groups differentially between antibiotic pretreatments.

Each point represents a unique gene from the respective metagenomic assembly. Coordinates were determined by the log2-transformed transcription 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 transcription between conditions (defined by outlier analysis). Antibiotic pretreatments; **(A)** Streptomycin-pretreated, **(B)** Cefoperazone-pretreated, and **(C)** Clindamycin-pretreated.

#### Figure 4 | Majority of metatranscriptomic changes are focused within minority members of each microbiota.

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 first panel represent the quantity of genera in each group as some points are obscured.

#### Figure 5 | Metatranscriptomic changes due to infection in certain metabolic pathways are overrepresented in the minority taxa.

Log2 metagenome-normalized cDNA abundances for genes with differential transcription during infection belonging to genera that compose >0.1% of their respective antibiotic pretreated communities. Levels of transcription for each in the resistant community is also included. Bar colors refer to antibiotic pretreatment group, while hashed lines represent *C. difficile* infected metatranscriptomes and solid colors are mock infected. Letters in brackets next to gene names indicate KEGG pathway membership labeled in the lower legend. Double asterisks denotes genes shared between pretreatment groups.

#### Supplementary Figure 1 | Conserved markers of *C. difficile* colonization susceptibility in mouse cecal microbiomes.

**(A)** NMDS ordination of Bray-Curtis distances of otu relative abundances between mouse cecal communities that are susceptible to colonization by *C. difficile* and those that are resistant. **(B)** Relative abundance of OTUs included the optimal model generated by AUCRF discriminating the same groups as in panel A, labeled with the most finest resolution RDP taxonomic classification. **(C)** NMDS ordination of Bray-Curtis distances between metabolite intensities between the aforementioned groups of animals. **(D)** Scaled intensity of metabolites included the optimal model generated by AUCRF discriminating resistant and susceptible cecal microbiomes. Significant differences in A & C were also calculated using permANOVA. The AUCRF models generated in this analysis also had 0% out of bag error and significant differences in B & D were calculated as in Figure 2.

#### Supplementary Figure 2 | Signatures of infection effect on the cecal microbiomes conserved across pretreatment groups.

**(A)** NMDS ordination of Bray-Curtis distances of otu relative abundances between antibiotic-pretreated mouse cecal communities that are either *C. difficile*-colonized or mock-infected. **(B)** Relative abundance of otus included the optimal model generated by AUCRF discriminating the same groups as in panel A. **(C)** NMDS ordination of Bray-Curtis distances using metabolite intensities between the same classes. **(D)** Scaled intensity of metabolites included the optimal model discriminating infected and uninfected cecal microbiomes. Statistical differences were performed as in Figure 2.

#### Supplementary Figure 3 | Relative concentrations of select *C. difficile* Stickland fermentation metabolites across infection models.

Metabolites included in this analysis were chosen based on their previously published interaction with *C. difficile* Stickland fermentation. Panels are as follows; **(A)** Proline, **(B)** 4-Hydroxyproline, **(C)** Glycine, **(D)** 5-Aminovalerate. Significant differences were determined by Wilcoxon rank-sum test with Benjamini-Hochberg correction when necessary. Black asterisks in the plotting area represent within group differences, while green asterisks along the top border denote significant differences compared to untreated control.

#### Supplementary Figure 4 | 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 | Antibiotic pretreatment regime summaries

Antibiotic classes, mechanisms, and dosage information for each pretreatment. Quantified effect on alpha- and beta-diversities of the cecal microbiota are also included.

#### Supplementary Table 2 | Summary of impact of infection on cecal community structure and metbaolome

Global effect as well as changes to specific metabolites are included.

#### Supplementary Table 3 | Summary statistics for datasets containing replicates generated during this study

#### Supplementary Table 4 | High-throughput sequencing read counts and metagenomic assembly quality

Raw and curated read abundances for both metagenomic and metatranscriptomic sequencing efforts. Read curation steps are outlined in Materials & Methods.

### 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. **Perez-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. **Calle, M. L.**, **V. Urrea**, **A. L. Boulesteix**, and **N. Malats**. 2011. AUC-RF: A new strategy for genomic profiling with random forest. Human Heredity **72**:121–132.

23. **Fletcher, J. R.**, **S. Erwin**, **C. Lanzas**, and **C. M. Theriot**. 2018. Shifts in the Gut Metabolome and Clostridium difficile Transcriptome throughout Colonization and Infection in a Mouse Model. mSphere **3(2)**.

24. **Dai, Z.**, **Z. Wu**, **S. Hang**, **W. Zhu**, and **G. Wu**. 2015. Amino acid metabolism in intestinal bacteria and its potential implications for mammalian reproduction. Molecular human reproduction **21**:389–409.

25. **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.

26. **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.

27. **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**.

28. **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.

29. **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.

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

31. **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.

32. **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.

33. **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**.

34. **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.

35. **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**.

36. **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.

37. **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.

38. **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.

39. **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.

40. **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.

41. **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.

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

43. **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.

44. **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.

45. **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.

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

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

48. **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.

49. **Oksanen, J.**, **F. G. Blanchet**, **M. Friendly**, **R. Kindt**, **P. Legendre**, **D. McGlinn**, **P. R. Minchin**, **R. B. O’Hara**, **G. L. Simpson**, **P. Solymos**, **M. H. H. Stevens**, **E. Szoecs**, and **H. Wagner**. 2018. vegan: Community Ecology Package.

50. **Giorgino, T.** 2009. Computing and Visualizing Dynamic Time Warping Alignments in R : The dtw Package. Journal of Statistical Software **31**:1–24.

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