**Title**

*Clostridium difficile* colonizes alternative nutrient niches during infection across susceptible murine gut environments

**Authors**

Jenior ML1, Leslie JL1, Young VB1,2, and Schloss PD1

**Affiliations**

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

2. Department of Internal Medicine; University of Michigan Health System; Ann Arbor, MI

**Abstract**

*Clostridium difficile* infection (CDI) as grown to be the greatest cause of hospital acquired infection in the United States. Susceptibility to CDI is induced by previous antibiotic exposure, which has been shown to alter the structure of the gut microbiome. These changes have been associated with changes in bacterial growth nutrient availability in the gut, often increasing concentrations of several useable by *C. difficile*. In this study, we orally challenged C57BL/6 mice with *C. difficile* str. 630 and demonstrated that it was able to colonize the ceca in three separate models of antibiotic induced susceptibility to the same high degree (~1×108 CFU/g content) within 18 hours of inoculation. However, despite equal vegetative cell load at this time point, the levels of both spore and toxin production vary between each antibiotic treatment group. The expression of both phenotypes have both been linked to environmental concentrations of certain substrates, and this indicated possible differences in the nutrient niche *C. difficile* inhabits between susceptible gut conditions. To more closely investigate the specific responses of *C. difficile* as it colonizes the GI tract of mice, we performed *in vivo* *C. difficile*-focused RNA-Seq analysis from cecal content of infected mice. This approach identified differences in expression for genes associated with life-cycle stages and pathogenicity between antibiotic pretreatments, in agreement with previous results. We then went on to observe numerous variations between condition in metabolic pathways associated with carbohydrate and amino acid catabolism, indicating that *C. difficile* likely colonizes alternative nutrient niches across the environments it colonizes. In order to assess which aspects of the gut environment *C. difficile* is exploiting during infection, we sought to identify the growth nutrients that are most likely being used the the pathogen in each colonized condition. To accomplish this we developed a novel substrate scoring algorithm within the genome-scale bipartite metabolic network of *C. difficile* str. 630. The calculation incorporates both network topology and transcript abundance for enzymes within the model to infer the likelihood that a given substrate in acquired from the environment as opposed to being produced by the organism itself. Using this technique, we first validated the approach by confirming the most central nodes reflect known components of *C. difficile* core metabolism. Applying the importance algorithm, we found that *C. difficile* indeed occupies alternative nutrient niches across each antibiotic class tested. Results from this analysis support specific sugar alcohols (D-sorbitol & mannitol) and host derived amino sugars (salicin & N-acetylneuraminate) as the most probable growth nutrients that *C. difficile* differentially utilizes across the susceptible gut environments it colonizes. We then confirmed the capacity of each implicated substrate to support substantial *C. difficile* growth *in vitro*. Furthermore, by employing liquid chromatography-mass spectrometry, we show that these substrates are significantly enriched in the conditions they were found to be most important by the network/transcript-based algorithm. This work has implications for elucidating specifics of the nutrient niche of *C. difficile* during infection, and may lead to the discovery of targeted measures for *C. difficile* colonization prevention.

**Introduction**

Infection by the Gram-positive, spore-forming bacterium Clostridium difficile has increased in both prevalence and severity across numerous countries over the last decade (Lessa 2012). In the United States in 2014, C. difficile was estimated to have caused >500,000 infections and resulted in ~$4.8 billion worth of acute care costs (Lessa, 2015). Clostridium difficile infection (CDI) causes an array of toxin-mediated symptoms ranging from abdominal pain and diarrhea, to the more life-threatening conditions pseudomembraneous colitis and toxin megacolon. Prior treatment with antibiotics is the most common risk factor associated with susceptibility to CDI (Leffler, 2015). It has been shown that antibiotic therapy alters the structure and function of the gut microbiota makes it susceptible to colonization by *C. difficile* (Young, 2004). This is referred to as colonization resistance, in which the gut microbiota inhibits the persistence or growth of a number of pathogenic bacteria. Colonization resistance can be achieved by multiple mechanisms including competition for physical space or growth nutrients (Lawley, 2013).

Mouse models have been an effective tool for studying the mechanisms of colonization resistance. Use of distinct antibiotic classes to vary the structure of the microbiota has been shown to result in varied levels of *C. difficile* colonization (Chen, 2008; Theriot, 2013; Schubert, 2015). Additionally, 18 hours after initially colonizing a treated mouse, *C. difficile* reaches its maximum density in the cecum before beginning to sporulate (Koenigsknecht, 2015). Building upon these results, others have shown that many of these antibiotic classes also alter the gut metabolome, increasing the concentrations of known *C. difficile* growth substrates (Antunes 2011; Theriot, 2014; Ferreyra, 2014; Jump, 2014). Taken together these results are a strong indication that the healthy gut microbiota inhibits the growth of *C. difficile* through limitation of substrates it needs to grow. The ability of a healthy gut community to exclude *C. difficile* colonization is suggestive of the nutrient-niche hypothesis (Freter, 1983; Wilson, 1988) in which an organism must be able to utilize a subset of available resources better than all competitors to colonize the intestine.

Based on its genome sequence and *in vitro* growth characteristics, *C. difficile* appears to be able to fill multiple nutrient niches. *C. difficile* has a relatively large and mosaic genome, it is amenable to a variety of growth substrates, and is able to colonize a diverse array of hosts suggesting that that it is a bacterial generalist (Sebaihia, 2006; Tracy, 2012; Songer, 2006). The ability to metabolize a variety of substrates is important since these substrates affect the regulation of genes involved in *C. difficile*’s pathogenesis. For example, *in vitro* transcriptomic analysis suggests that high concentrations of easily metabolized carbon sources, such as glucose or amino acids, inhibit toxin gene expression (Neumann-Schaal, 2015) and sporulation (Nawrocki, 2016). These genes are regulated by DNA-binding sigma factors, such as the pleiotropic regulator *ccpA,* which are under the control of environmental nutrient concentrations, especially carbohydrates (Bouillaut, 2015). Downstream effects of this regulation likely have enormous impact on the lifestyle and metabolic strategy of *C. difficile* when colonizing across sensitive hosts.

Previous transcriptomic studies of *C. difficile* have mainly focused on *in vitro* changes (Matamouros, 2007; Antunes, 2011), with a some work characterizing *in vivo* transcription during colonization of germ free mice (Janoir, 2013). It was also demonstrated that *C. difficile* up-regulates of several PTS and ABC transporters in germ free mice, alluding to metabolic adaptation to nutrient availability *in vivo* (Kansau, 2016). Although these analyses are informative, they are either primarily directed toward the expression of virulence factors or lack the context of the gut microbiota which *C. difficile* must compete against for the substrates it needs. Groups have also used metabolomic analysis to more directly assay changes in bacterial metabolism as they relate to CDI, but these methods cannot focus on *C. difficile* specifically and more closely exhibit echoes of metabolism, not currently active processes. With this in mind, *in vivo* *C. difficile* transcriptomic analysis from conventionally-raised animals can provide unique insight into its active metabolic pathways in a more realistic model of infection. Integrating these data with *in silico* metabolic modeling can also aid in identifying the most active aspects of *C. difficile* metabolism and infer which substrates the organism is differentially using across antibiotic models. This approach has been used to successfully \_\_\_in yeast

(Patio, 2005)

Based on the ability of *C. difficile* to grow on a diverse array of carbon sources and its ability to colonize a variety of communities, we hypothesized that it focuses its metabolism to fit the context of the community it is attempting colonize. To test this hypothesis, we employed a mouse model of infection to compare the response of *C. difficile* str. 630 the gut environment caused by different classes of antibiotics (Fig. 1A). The antibiotics used in this study included cefoperazone, clindamycin, and streptomycin (Table 1). These antibiotics differentially affect the structure of the gut microbiota (Schubert, 2015). Each has also been shown to alter the gut metabolome relative to untreated animals (Antunes, 2011; Theriot, 2014; Jump, 2014). As such, we predicted that *C. difficile* would encounter a unique subset of nutrients and competitors in each environment, which would necessitate distinct adaptive responses. To determine whether *C. difficile* is a generalist and differentially responds to each condition, we assayed for differences in the amount of sporulation and toxin activity phenotypes and used metabolic models built using *C. difficile* expression data. In each of the three antibiotic conditions we challenged with *C. difficile*, as well as in germ free mice, we observed that *C. difficile* altered its metabolic gene expression to colonize to high levels and produce varied amounts of toxin.

**Results**

**Levels of *C. difficile* sporulation and toxin activity vary between germ free and conventionally-raised mice.**

Due to the connection between metabolism and sporulation and toxin production in *C. difficile*, we measured both phenotypes at 18 hours post infection. This time period was selected based on previous work indicating this was a time that would yield the highest number of *C. difficile* vegetative cells in the cecum relative to the number of spores (Koenigsknecht, 2015). There was not a significant difference in the number of vegetative cells between any susceptible condition tested (all *p*-values < 0.05; Fig. 1B). All antibiotic treated and gnotobiotic animals were colonized to ~1×108 CFU/g content, while untreated conventionally-raised mice maintained colonization resistance to *C. difficile*. Despite having the same number of vegetative *C. difficile* cells, large differences were detected in the density of *C. difficile* spores. Significantly more spores (all *p*-values < 0.01) were detected in gnotobiotic mice than in the antibiotic treated mice (Fig. 1C). The spore densities in both streptomycin and clindamycin-treated mice were significantly higher than that in cefoperazone-treated mice. There was significantly more toxin titer in gnotobiotic animals than any other colonized group (all *p*-values < 0.01), but toxin titer also varied between antibiotic treatment groups (Fig. 1D). Although similar toxin titers were found in both the cefoperazone and clindamycin-treated groups, toxin titer was not detectable in the streptomycin-treated animals. These results indicate that *C. difficile* was able to colonize different communities to a consistently high level, but that the density of spores and toxin titer varied by treatment. High levels of both spores and toxin titer in gnotobiotic mice is most likely due to temporal changes in infection dynamics when no other bacteria are present.

***C. difficile* displays flexible expression of key metabolic regulation factors as well as genes for pathogenesis across antibiotic pretreatments.**

To more closely investigate the responses of *C. difficile* to colonizing distinct susceptible gut environments, we performed stranded whole transcriptome analysis of *C. difficile* during infection of the antibiotic pretreatment models. In order to generate enough mRNA biomass contributed by *C. difficile*, we needed to pool cecal content from all mouse replicates into a single large isolation for each treatment group. This was beneficial because it not only normalized subtle variation in expression between samples, but also simplified the technical process of isolating high-quality RNA from multiple treatment groups at once. Sequencing, read curation, and normalization steps are fully described in Materials & Methods. Fully curated reads were the mapped to C*lostridium difficile* str. 630 genes. This resulted in ~150,000 curated reads per group left for downstream analysis. Each collection of reads was then 1000-fold iteratively subsampled to 97,930 reads to generate bootstrapped confidence intervals for each gene (Fig. 2). After observing very large temporal differences in *C. difficile* phenotypes between gnotobiotic and conventionally-raised mice, we chose to omit gnotobiotic animals from transcriptomic analysis and focus analysis on those groups containing potential competitor species.

To first observe markers of global changes in metabolism, we looked for differences in expression in key sigma factors that were previously found to be influenced by the nutrient environment of *C. difficile* (Fig. 2A). The global nutritional regulator *codY* transcript was highly detectable in each condition, despite apparent differences in toxin and spore production. Canonically, *codY* is a strong negative regulator of both phenotypes (Bouillaut, 2015) and could demonstrate a more complex regulatory network as *C. difficile* integrates multiple environment nutrient signals. Save for a few instances, it appeared that expression the sigma factors of interest were generally higher in cefoperazone-treated animals. The largest contrasts were seen in expression of *sigE*, *sigG*, and *sigK* which were approximately 10-fold less expressed in cefoperazone than the others. These sigma factors are involved in expressing genes for proper spore coat assembly and asymmetric division (Donnelly, 2016; Fimlaid, 2013). This agrees with the measured differences in CFU due to spores (Fig. 1C) where cefoperazone pretreatment resulted in delayed sporulation.

To further investigate the discrepancies in sporulation, we chose to quantify specific changes in expression within the *C. difficile* sporulation pathway (Fig. 2B). The gene list was derived from Saujet et. al., 2014 where the authors reviewed the complex cascade of regulation that occurs as *C. difficile* completes the steps of sporulation. The resulting profiles further support regulation differences across environments in which the pathogen has colonized. A striking difference was that in Cefoperzone treated animals, *C. difficile* highly expresses a number of putative anti-sigma factors including *spoIIAA,* *spoIIAB,* and *spoVG* involved in suppressing other sporulation-related gene expression (Matsuno, 1999). This pattern was not seen in either Streptomycin or Clindamycin pretreatments, possibly explaining the observed differences in spore CFU. Instead, these conditions induced expression of genes for spore coat components, coat assembly, and spore morphogenesis (including *cotD*, *cotJB2*, *dpaA*, *gpr*, *sodA*, *spoIVA*, *sspB*), all associated with later sporulation stages (Permpoonpattan, 2013). These results further reinforce the connections previously made to the differences seen in spore-specific CFU quantification.

Along with sporulation, expression of PaLoc genes has been linked to changes in the nutrients that can be found in the environment of *C. difficil*e (Fig. 2C). We found approximately 3-fold higher expression of the toxin A gene, *tcdA*, than the toxin B gene, *tcdB*, across all pretreatments. This relationship has long been known to exist at least *in vitro* with another strain of *C. difficile* (Hundsberger, 1997). This could have possible implications in the relative importance of toxin A for causing virulence, or temporal differences that toxin B may be more important later during infection. Additionally, we found large differences in the regulation factors most associated with toxin production, *tcdC* (negatively regulating anti-sigma factor) and *tcdR* (positive regulating sigma factor) (Carter, 2012). Between both genes, only *tcdC* transcript was detectable at 10-fold greater abundance than that of *tcdR* across all groups. This was interesting as toxin activity was detectable in each condition (Fig. 1A), but only the negative regulator of toxin expression was seen implying a more complex regulation of toxin production than this mechanism alone. Finally *tcdE*, the gene for a holin-like protein associated with *C. difficile* toxin secretion, was only detectable in cefoperazone pretreatment. Mutants for this gene have been associated with impaired toxin secretion and low *in vivo* expression of *tcdE* may partially explain the historically lowered virulence of *C. difficile* str. 630 (Theriot, 2011).

Given that many of the differences in expression dynamics seen may also be due to variations in bacterial quorum sensing, we went on to measure transcription of known quorum sensing genes in *C. difficile* (Fig. 2D). These systems have been linked to the lifestyle switch from growth/division to virulence/transmission of *C. difficile* (Darkoh, 2015; Carter, 2005). Transcripts for *agrB*, *agrD*, and *luxS* were found at nearly equivalent levels across all antibiotic pretreatments. This illustrated that the observed changes in phenotype and expression may not be directly due to changes in the levels of *C. difficile* quorum sensing occurring.

**Genes sets from multiple *C. difficile* energy metabolism pathways are differentially expressed between colonized environments.**

Differential expression of sigma factors indicated that environmental nutrient concentrations may be playing a role in the observed transcriptional patterns, we then decided to focus the analysis on specific genes that are integral to known *C. difficile* energy metabolism pathways. Mapped transcript read counts again were rarefied iteratively and the means of each gene were plotted as relative abundances along the three axes for the antibiotic pretreatment groups (Fig. 3). Each point represents a unique gene from the annotated genome of *C. difficile* str. 630 with position reflecting the ratio of transcription for that gene in all three antibiotic pretreatments. Points near a corner indicates approximately 100% of the total transcription for that gene across all pretreatments is contributed by *C. difficile* that has colonized mice given the antibiotic associated with given corner. By extension, points that are closer to the center of the plot area demonstrate nearly equal expression across each condition and imply that these functions may be constitutively expressed. Furthermore, a point directly along an axis conveys that 0% of its measured expression was found in the condition associated with the opposite corner. Genes included in each group and subsampled abundances can be found in Table S2.

A number of notable within-group differences stood out including one in relative expression abundance for genes associated with sugar alcohol catabolism (Fig. 3E). Transcripts from this gene set are almost undetectable in the clindamycin pretreated group, however a large difference was found between those genes more associated with cefoperazone than streptomycin pretreatment. Those genes closest to the cefoperazone edge are important for the utilization of mannitol (*mtl* operon) while those proximal to streptomycin are involved in sorbitol metabolism (*srl* operon), indicating potential niche segregation or changes in sorbitol availability between groups. Additionally, there was another within-group difference for polysaccharide catabolism between genes associated more with clindamycin pretreatment versus those associated with cefoperazone (Fig. 3D). Those genes most associated with clindamycin are largely members of the beta-glucoside utilization locus (*bgl* operon), canonically linked with salicin or arbutin metabolism (Schnetz, 1987). Genes more associated with cefoperazone pretreatment are primarily related to import and utilization of maltose or maltodextrin (*maa* & *map* operons) (Boos, 1998). Finally, amino sugar catabolism also displayed some degree of within-group trends (Fig. 3B). Streptomycin pretreatment was enriched for expression of *nanA* and an N-acylglucosamine-6-phosphate 2-epimerase. These gene products are important for the metabolism of N-acetylneuraminic acid and N-acetyleglucosamine respectively, which are both obtained from the cleavage of host derived amino sugars in the mucus layer by other bacterial species (Ng, 2013). Streptomycin also appears to be associated with high expression of *fruABC*, a fructose specific PTS transport system (Fig 3H), and *fruK*, which phosphorylates fructose upon import (Fig 3C), supporting that fructose utilization is occurring in this condition as well.

Addition analysis revealed several between-group significant differences in similarity. Genes in the sugar transport systems, PTS and sugar-specific ABC transporters (Fig. 3H & 3I), are among the groups that segregate relative to each other. Both systems are known carbohydrate importers, however activation of some PTS transporters in response to environmental concentrations of certain carbohydrates has been shown to negatively regulate the expression and function of many ABC transporters in some bacterial species (Deutscher, 2006). Further differences in carbohydrate metabolism were also visible when comparing the polysaccharide and monosaccharide catabolism expression sets (Fig. 3D & 3E). Clindamycin and streptomycin pretreatments had an enrichment of expression of glucosidase genes within polysaccharide catabolism, while cefoperazone pretreatment had higher expression of genes associated with fructose and mannose metabolism within monosaccharide catabolism. These differences could highlight the ability of *C. difficile* to adapt to environments with unique concentrations of resources and colonize a variety of nutrient niches during infection.

Conversely, the expression of genes associated with amino acid catabolism were consistently expressed at relatively equal levels across the conditions (Fig. 3F). These points all trend toward the center of the plotting area and show no apparent differences between any treatment group. This gene category includes those enzymes involved in Stickland fermentation (*arg*, *fdh*, *grd*, *had*,and *prd* operons) as well as several general peptidases (*pep* family). The known Stickland fermentation pathway in *C. difficile* couples the oxidation and reduction of proline an glycine to generate ATP (Neumann-Schaal, 2015). Transcription of specifically the *grd* and *prd* operons indicates that both proline and glycine are present and the associated metabolic pathways have most likely been activated (Bouillaut, 2013). This suggests that the catabolism of environmental amino acids may be a central part of the metabolism of *C. difficile* str. 630 during infection.

**Bipartite metabolic models underscore known bacterial metabolism and reveal adaptive nutritional strategies of *C. difficile* during infection.**

The enzymes encoded in an organism's genome, in many instances, mediate known biochemical reactions with input and output substrates inside the cell. These reactions can be organized into a web of substrates flowing between enzymes, modeled as a directed network of metabolism. In this instance we represented the metabolic network of *C. difficile* str. 630 as bipartite graph, where enzyme nodes only connect to substrate nodes and vice versa (Fig. 4A), using publicly-available gene content information from the Kyoto Encyclopedia of Genes and Genomes (KEGG). Examining the topology of a genome-scale metabolic graph can provide information about how well the architecture reflects known biology for the organism of interest. This is especially important when assembly of the graph is unsupervised and depends on the strength of the gene annotation used. This could potentially lead to incomplete connectedness of the network and result in a poor reflection of an organism's actual biology. Therefore, we calculated Betweenness-centrality for all enzyme and substrate nodes in the bipartite metabolic graph *C. difficile* str. 630 generated for this study (Table S3). This metric is the amount of influence a given node has on transfer through the directed network, which in biological terms reflects the substrates and functionalities that are most critical to the overall metabolism of the organism. Enzymes scoring highly include pyruvate kinase, 1-deoxy-D-xylulose-5-phosphate synthase, and transketolase while substrates with the highest scores included pyruvate, acetyl-CoA, and D-glyceraldehyde 3-phosphate. For both enzymes and substrates, these nodes are known to be involved in numerous central pathways including glycolysis, the citric acid cycle, and amino acid biosynthesis which are all critical to life for most free-living bacteria.

While traditional transcriptomic approaches are informative for identifying active pathways to highlight potential catabolism for families of substrates, it does not directly highlight specific nutrients that *C. difficile* may be using at a given time. With this in mind, we developed a novel approach employing the previously generated metabolic network and transcript abundances to infer which substrates are the mostly likely to be used by the bacterium at a given time. This is done in order to understand the relative importance of substrate nodes based on local enzyme transcription. To calculate substrate importance (Fig. 4A), we used rarefied transcriptomic abundances mapped to their respective enzymes (represented by **to** and **ti**), representing the reactions they mediate that either create or utilize the substrate (**s**). The sum is then taken of transcript abundances for enzyme nodes that connect with a given substrate node with the same direction edge (**eo** or **ei**). These values are then divided by the total number of edges that contributed to the previous calculation, creating an average transcription value as either input or output of reactions for the substrate node. The intent is to normalize for highly central substrate nodes that have many connecting enzymes and would score highly regardless of differential expression. The normalized transcription of output reactions is then subtracted from that of input reactions to yield a value that reflects whether synthesis or utilization of the substrate might be favored at the time of measurement. This value is then base-2 log transformed for ease of comparability between high and low transcript abundance associated with a substrate. Zeroes were left unchanged and negative values are accounted for prior to transformation using absolute value with the sign reapplied afterward following transformation. The presumption of our approach is that enzymes that are more highly transcribed are more likely to utilize the substrates they act on due to coupled bacterial transcription and translation. Furthermore, we applied a Monte-Carlo simulation style transcript assignment randomization in create a comparator for each substrate's computed importance score. By comparing the calculated scores to their simulated mean and standard deviation from iterative randomization, we were able to detect which substrate scores were most likely not from consistent high scoring due to a large degree of centrality. The entire process has been implemented in a generalized, automated workflow presented as BiGSMAll, available in a public Github repository located at <https://github.com/mjenior/bigsmall>. Applying this method, we extracted the relative importance values of the 758 unique compounds in the metabolic network of *C. difficile* str. 630, after 18 hours of infection in each treatment group. We included the transcriptome of *C. difficile* in gnotobiotic mice in the network-based analysis because although the dynamics of the infection may be altered, there is still active metabolism occurring and therefore provided a useful comparator through this approach.

A possible shortcoming of this technique is that some reactions being annotated as reversible universally across bacteria, but are not within certain species of interest. This may result in some metabolic byproducts appearing as important inputs, when in reality they are exclusively outputs of this network. To address this, we ranked scores for each treatment separately to identify the most important compounds for all individual groups (Table S5). In both cefoperazone and streptomycin pretreatments, several Stickland fermentation intermediates, including propanoate and phosphonoacetate, were scored as highly important, much greater than their simulated means. Alternatively in the clindamycin and gnotobiotic conditions, the mucus derived amino sugars N-acetylglucosamine and N-acetylneurminic acid as well as known glycolysis intermediates scored highly in instead of the substrates involved in fermentation. Both forms of carbohydrate metabolism are prominent in *C. difficile* and demonstrated that our importance-defining algorithm was robust could detect biologically relevant patterns of expression.

Although an analysis at the individual condition level can have some implications about nutrient utilization in each condition, a unified comparison of all conditions is more informative for understanding how *C. difficile* is most likely to behave when it colonizes a susceptible host. With this in mind, we then sought to identify core metabolites that are essential to *C. difficile* str. 630 in any condition. To accomplish this goal, we compared the top 100 scoring substrates from each group to find those that are shared across all conditions (Table S4). The host derived amino sugars N-acetyl-D-glucosamine was found to be consistently important, indicating that it may be an integral component of the nutrient niche for *C. difficile* under any condition. Other conserved features include several central fructose derivatives as well as a likely byproduct, lactate, from the anaerobic glycolytic pathway. This is not surprising as this is known to be a key component of *C. difficile* metabolism, and is likely active to some capacity across all conditions. The lactate produced in this manner is also possibly further fermented to short-chain fatty acids, primarily butyrate (Janoir, 2013). Additionally, acetate was found to be important in all conditions and is a common byproduct of Stickland fermentation and has been shown to be further metabolized to butyrate by other closely-related species of Clostridia (Li, 2008). We directly tested the relative concentration of acetate in cefoperazone treated, *C. difficile* str. 630 infected mice versus antibiotics alone and found that colonization led to a significant decrease in the levels of acetate (Fig. S7A). While this may be due to other factors, this could show that *C. difficile* similarly utilizes acetate for energy. Many of the top scoring compounds appear to be associated with both fermentation and glycolysis, suggesting that these forms of *C. difficile* metabolism are always active to some degree across conditions and are central to the organism’s nutritional strategy. Other substrates substrates such Fe2+ was consistently important, as well several forms of certain amino acids (proline, threonine, cysteine, leucine, isoleucine, and valine), many of which *C. difficile* is a known or suspected auxotrophe (Ho, 2015; Neumann-Schaal, 2015). This analysis further supported that our substrate importance algorithm was well-suited to find which substrates an organism is using at a given time.

To then look at the converse of the unified analysis, we then identified those compounds that are uniquely important to each condition in which *C. difficile* is growing. We pursued this by cross-referencing the top 25 positively scoring substrates from each treatment group to uncover patterns of unique nutrient utilization by *C. difficile* between environments. Results from this comparison left only a few substrates in each group (Fig. 4B). All substrates identified this way scored the comparator mean generated through Monte Carlo simulation (black vertical lines). Each group of substrates contained at least one known carbohydrate growth substrate of *C. difficile*; including D-fructose, mannitol, N-acetylneuraminic acid, salicin, and D-sorbitol (Ng, 2013; Theriot, 2014). These data support that *C. difficile* may exploit alternative carbon sources between the susceptible environments in colonizes.

**Carbon sources identified as important by metabolic networks support *C. difficile* growth *in vitro*.**

In order to validate the biological relevance of substrates identified in the previous network-based analysis, we went on to observe whether *C. difficile* str. 630 was able to utilize each substrate for *in vitro* growth. This was performed using the modified defined *C. difficile* minimal media, supplemented individually with the selected carbohydrates (Theriot, 2014) implicated by high importance scores. As *C. difficile* can use amino acids (AA) for growth through Stickland fermentation, but is auxotrophic for several, the best comparator are -glucose +AA control wells. This informed whether the presence of the selected carbohydrate effects has positive or negative effects on the growth rate of *C. difficile,* however this analysis does not preclude the possibility that *C. difficile* simply grows equally as well as it does using strictly AA for carbon.

At least one carbohydrate highlighted as important in each of the antibiotic treatments groups provided high levels of *C. difficile* growth (Fig. 4C). This included D-fructose (OD600 = 0.671), mannitol (OD600 = 0.464), and salicin (OD600 = 0.888). Each resulted in significantly more growth (corrected *p*-values << 0.001) than AA-only control (OD600 = 0.212). The growth curve for D-sorbitol showed no improvement over the AA-only controls. Additionally, N-acetylneuraminate was found to be highly important when colonizing Gnotobiotic mice, and it also allowed for significantly more growth (OD600 = 0.556) than AA-only control wells (corrected *p*-value << 0.001). These results indicate potential efficiency or preferences by *C. difficile* for the each carbohydrate source tested.

Calculation of maximal growth rates also revealed possible differences in utilization efficiency between the 4 carbon sources achieving the most growth over the 24 hour period (Table S5). The D-fructose growth curve reached its maximum positive growth rate of the earliest of all substrates at 6 hours (*m* = 0.089). Mannitol was second at 7 hours (*m* = 0.044), and salicin was third at 13.5 hours (*m* = 0.076). Interestingly, the salicin growth curve was still near its maximum growth rate at the 12 hour time point, while the D-fructose (*m* = 0.006) and mannitol (*m* = 0.004) curves had both reached stationary phase. When following the growth curves over the entire 24 hour period, it reveals that salicin provides for even greater OD600  (OD600  = 0.888) than BHI rich media control (OD600  = 0.663) (Fig. S6). This delay and subsequent progression into normal exponential phase with large outgrowth could indicate a difference in programming for uptake at the time of inoculation, or even differences in utilization efficiency. The stark drops in OD600 of some curves is most likely due to *C. difficile* ceasing division and switching to primarily sporulation, which has been associated with other bacterial species (Van Beilen, 2013).

Additionally, we tested an equal combination (5 mg/mL) of both carbohyrdates (Fig. S6) found to be especially important in the Streptomycin-associated metabolic network calculations (D-fructose and D-sorbitol). The resulting growth curve did not closely resemble either of those associated with its component carbohydrates. The maximum hour growth rate (*m* = 0.095 at 6.5 hours) is the highest reached by any media tested, yet the maximum OD600 of 0.510 was lower than both D-fructose and +glucose +AA control individually. In contrast, the carbohydrate combination resulted in significantly more (corrected *p*-value << 0.001) than D-sorbitol alone. This could suggest a combinatorial effect of D-fructose and D-sorbitol together on the growth of *C. difficile*.

***In vivo* concentrations of predicted nutrients are …**

In order to test if the identified carbohydrates are present or enriched in their respective antibiotic treatments, we performed metabolomic analysis from cecal content from the same animals. (Fig. 5A) (Fig. 5B)

**Discussion**

[from Introduction]

Our results indicate that *C. difficile* occupies separate nutrient niches across susceptible gut environments based on which substrates are available. Both *in vivo* and *in vitro* data support that *C. difficile* shifts its metabolic focus to alternative sources of carbon across susceptible gut environments. This phenotype could be due to a variety of elements including increased concentration of a given set of nutrients, or competition with other microorganisms for those substrates. Sporulation and toxin production have been linked to environmental concentrations of growth substrates for *C. difficile* (Antunes, 2012). In these experiments, the level of both phenotypes varied between antibiotic pretreatment groups and initially suggested differences in metabolic cues received by *C. difficile* in each of condition. By directly examining the transcription of specific metabolic pathways, we were able to gain a greater understanding of the plasticity that *C. difficile* possesses for nutrient utilization.

Initial transcriptional analysis revealed that there was a large amount of variability in the expression of sporulation genes between pretreatments. Aside from the regulating sigma factor *spo0A*, which was consistently highly expressed, *spo* family genes were only expressed highly in cefoperazone treated mice. These genes are integral to Stage V sporulation and contribute to the assembly of the spore coat (Pettit, 2014) whose expression illustrates that this population of *C. difficile* has begun its final phase before transmission either other antibiotic group. In addition to lifestyle associated genes, *C. difficile* str. 630 also demonstrated clear trends of expression for genes associated with separate aspects of its carbon metabolism between pretreatment groups. In both cefoperazone and streptomycin pretreatments, *C. difficile* displayed increased relative abundance in gene expression associated with catabolizing monosaccharides and simple sugar alcohols. However, *C. difficile* expression in the clindamycin treated animals seemed to favor fermentation of polysaccharides. These differences could signify a focus in nutrient catabolism for *C. difficile* in some treatments and virulence/transmission in others.

There is also evidence for other bacterial species contributing to pools of resources that are usable by *C. difficile* for growth (Ng, 2013; Ferreyra, 2014). Schubert et. al. 2015 demonstrated that the same Clindamycin model of infection used here, results in a bloom in relative abundance of the commensal species *Akkermansia muciniphila* at the time of infection. It has been previously shown that this bacterium is capable of cleaving host associated glycans and liberating short amino sugars, including salicin (Ganesh, 2013), which is a functionality *C. difficile* lacks (Sebaihia, 2006). Within the scope of this study, these findings are especially intriguing as salicin was one of the most important substrates to *C. difficile* when colonizing Clindamycin treated animals. Additionally, modeling analysis of *C. difficile* colonizing germfree mice revealed N-acetylneuraminic acid to be what was the most important carbon source in this condition. The capacity to catabolize this amino sugar and others like it is highly distributed among bacterial species that inhabit the mammalian gut (Vimr, 2004), meaning that it is an edge of strong competition with the normal microbiota and could demonstrate a preference of *C. difficile* for N-acetylneuraminic acid as a growth substrate. In light of all of these studies our method may also strongly support that *C. difficile* metabolizes alternative carbon sources across susceptible gut environments for growth. Together, they demonstrate that the novel metabolic modeling approach presented here highlights which substrates an organism is likely using at the time of transcriptomic sequencing. Ultimately, these genome-scale metabolic models provide a framework to predict phenotype changes of an organism across environments.

These results may also indicate that further considerations are needed when attempting to design more targeted probiotics for the prevention or elimination of *C. difficile* from the human gut. Different classes of antibiotics may each result in a unique gut environment which *C. difficile* can exploit. Therefore, the best approach may be to consider each of these as a separate problem which different collections of bacteria can restore to a resistant state. Supporting this hypothesis is that the current, most effective microbiome-focused treatment for *C. difficile* infection is fecal microbiota transplant. By even transiently implanting the entire metabolic functionality of a resistant gut community, one could imagine a rapid out-competition for resources by a healthy community of specialists and leading to *C. difficile* clearance. This could be supported by combining the metabolic modeling technique described here using metatranscriptomic analysis of susceptible versus resistance gut bacterial communities.

In conclusion, our findings indicate that *C. difficile* is able to cope with unique nutrient availabilities across vulnerable gut environments which it colonizes. The data also support that separate classes of antibiotics may cause susceptibility to CDI through slightly different means by allowing for increases in relative concentrations of different carbon sources. Limitation of specific subsets of *C. difficile* growth nutrients by specific of the gut microbiota to may represent the key to targeted colonization prevention. If we begin to understand the limited number of dimensions that *C. difficile* can inhabit within its niche space, we may be able to design targeted probiotics to fill those niches in susceptible individuals. There could also be many not yet appreciated environmental factors that influence the metabolic program of *C. difficile* during infection. In the future, this may need to a consideration during treatment as the correct course of action could be situational based on the factors are initially causing patients to become susceptible to infection.

**Materials and Methods**

**Animal care and antibiotic administration.** Six-to-eight week-old specific pathogen free (SPF) C57BL/6 mice were obtained from a single breeding colony maintained at the University of Michigan for all experiments. Six-to-eight week-old germfree C57BL/6 mice were obtained from a single breeding colony maintained at the University of Michigan for all experiments. 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. Specified SPF animals were administered one of three antibiotics; cefoperazone, streptomycin, or clindamycin (Table 1). Cefoperazone (0.5 mg/ml) and streptomycin (5.0 mg/ml) were administered in the drinking water for 5 days with 2 days recovery with untreated drinking water prior to infection. Clindamycin (10 mg/kg) was given via intraperitoneal injection 24 hours before time of infection.

***C. difficile* infection and necropsy.** *C. difficile* strain 630 spores were prepared from a single large batch whose concentration was determined a week prior to challenge for all experiments. On the day of challenge, 1×103 *C. difficile* spores were administered to mice via oral gavage in ×1 phosphate-buffered saline (PBS) vehicle. Subsequent quantitative plating for CFU was performed to ensure correct dosage. Infection negative control animals were given an oral garage of 100 μl ×1 PBS at the same time as those mice administered *C. difficile* spores.

18 hours following infection, mice were euthanized by carbon dioxide asphyxiation. Necropsy was then performed and cecal content was split into three small aliquots (~100 μl). Two were flash frozen immediately for later DNA extraction and toxin titer analysis respectively. The third aliquot was quickly moved to an anaerobic chamber for CFU quantification. The remaining content in the ceca (~1 ml) was emptied into a stainless steel mortar in a dry ice/ethanol bath using 1 ml of sterile PBS. This process was repeated for each mouse within a treatment group to pool content into a single large sample (9 mice across 3 cages) to compensate for cage effects as much as possible while maximizing sequencing depth. The content was then finely ground and stored at -80 C for subsequent RNA extraction.

***C. difficile* cultivation and quantification.** Differential plating was performed to quantify both C. difficile spores and vegetative cells respectively. Cecal samples were weighed and serially diluted under anaerobic conditions (6% H, 20% CO2, 74% N2) with anaerobic PBS. Samples were plated on CCFAE plates (fructose agar plus cycloserine, cefoxitin, and erythromycin) and counted for CFU after 24 hours of anaerobic growth at 37° C (Buggy, 1997). It is important to note that the germination agent taurocholate was omitted from these plates in order to only quantify vegetative cells. Undiluted samples with then heated at 60° C for 30 minutes in order to eliminate vegetative cells and leave only spores (Sorg, 2001). These samples were then serially diluted under anaerobic conditions in anaerobic PBS and plated on TCCFAE (CCFAE plus taurocholate). Plating was simultaneously done for heated samples on CCFAE to ensure all vegetative cells had been eliminated. After 24 hours of anaerobic incubation at 37° C, CFUs from spores were quantified. Significant differences were determined by Holm-Bonferroni corrected Wilcoxon tests.

***C. difficile* toxin titer assay** Vero cell rounding assay was performed on mouse cecal content as previously described (Leslie, 2014). Cells were grown to a confluent monolayer in DMEM, supplemented with 10% heat-inactivated fetal bovine serum and 1% penicillin-streptomycin . The cells then were transferred to a conical tube and spun at 1,000 rpm for 5 minutes to pellet the cells. 1×105 cells were seeded in each well of a 96-well plate and incubated for 4 hours. Filtered cecal content was serially diluted 1:10 in ×1 PBS. Control wells were given a volume of either a 1:25 dilution of anti-toxin serum. Following the incubation, samples were added to the Vero cells and the plate was incubated overnight at 37° C. Plates were viewed after 24 hours at ×10 magnification for cell rounding. The cytotoxic titer was defined as the reciprocal of the highest dilution that produced rounding in 80% of the cells. Significant differences were determined by Holm-Bonferroni corrected Wilcoxon tests. A detailed protocol with product information can be found at:

<https://github.com/jlleslie/Intraspecific_Competition/blob/master/methods/Verocell_ToxinActivity_Assay.Rmd>

**RNA extraction and library preparation.** This procedure was adapted from the RNA isolation protocol defined by Lopez-Medina et. al., 2011. 3 ml of lysis buffer (2% SDS, 16 mM EDTA and 200 mM NaCl) contained in a 50 ml polypropylene conical tube was first heated for 5 minutes in a boiling water bath. 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 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 g at 4° 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 g at 4° C for 45 minutes. The pellet was washed with 0° C 100% ethanol and resuspended in 200 μl of RNase-free water.

Following the manufacterer's protocol, samples were then treated with 2 μl of Turbo DNase for 30 minutes at 37° C. RNA samples were retrieved using the Zymo Quick-RNA MiniPrep according the manufacterer's protocol. Completion of the reaction was assessed using PCR for the V6 region of the 16S rRNA gene (Kozich, 2013). Quality and integrity of RNA was assessed using the Agilent RNA 6000 Nano kit for Total Prokaryotic RNA. The Ribo-Zero Gold rRNA Removal Kit Epidemiology was then used to deplete Prokaryotic and Eukaryotic rRNA from the samples according the manufacterer's protocol. Also, prior to library construction, quality and integrity as measured again using the Agilent RNA 6000 Pico Kit and stranded RNA-Seq libraries were made constructed with the TruSeq Total RNA Library Preparation Kit v2, both using the manufacturer's protocol. 2x50 bp sequencing was finally performed using an Illumina HiSeq 2500. A detailed protocol for this entire procedure containing product numbers and quality-control primer sequences can be found in the Github repository associated with this project.

<https://github.com/SchlossLab/Jenior_Transcriptomics_2015/blob/master/doc/wetlab_protocol.txt>

**Sequence curation, read mapping, and normalization.** Raw transcript sequencing read curation was performed in a two step process. Residual 5’ and 3’ Illumina adapter sequences were trimmed using CutAdapt (Martin, 2011) on a per library basis. Reads were quality trimmed using Sickle (Joshi, 2011) on the default settings. An average of ~300,000,000 total reads (both paired and orphaned) remained after quality trimming. Mapping was accomplished using Bowtie2 (Langmead, 2009) and the default stringent settings. ~1,600,000 reads in sample each mapped to the annotated nucleotide gene sequences of *Peptoclostridium difficile* str. 630 from the Kyoto Encyclopedia of Genes and Genomes (KEGG). Optical and PCR duplicates were then removed using Picard MarkDuplicates (<http://broadinstitute.github.io/picard/>), leaving ~150,000 reads per sample for final analysis. The remaining mappings were converted to idxstats format using Samtools (Li, 2009) and the read counts per gene were tabulated. Discordant pair mappings were discarded and counts were then normalized to read length and gene length. Generally, transcriptomic datasets were subsampled to the same number of reads as 90% of reads in the sample with the fewest mapped.

**Bipartite Metabolic Modeling.** The metabolism of *C. difficile* str. 630 was represented as a directed bipartite graph with both enzymes and substrates as nodes. Briefly, models were constructed using KEGG ortholog gene annotations to which transcripts had been mapped. Using the reaction\_mapformula.lst file located in /kegg/ligand/reaction/, KEGG orthologs were converted to lists of input and output compounds associated with each enzymatic reaction respectively. As compounds are associated with the enzyme, it is labeled as either an input or output of the reaction that enzyme mediates. This process is repeated for all enzymes in the given bacterial genome. The resulting data structure is an associative array of enzymes associated with lists of both categories of substrates (input and output), which serves as the architecture of the network. The final metabolic network of *C. difficile* str. 630 contains a total of 1205 individual nodes (447 enzymes and 758 substrates) with 2135 edges. Transcriptomic mapping data is then re-associated with the respective enzyme nodes prior to substrate importance calculations.

The substrate importance algorithm favors metabolites that are more likely acquired from the environment (not produced within the network), and will award them a higher score. This is accomplished by first calculating the relative importance of compounds as both inputs and outputs of reactions in the network and then subtracting the output score from the input score:

insert the equation here again

If a compound is more likely to be produced, the more negative the resulting score will be. The general concept for the bipartite metabolic networks and substrate importance algorithm were partially based on methods created to infer metabolite preference of *Saccharomyces cerevisiae* grown *in vitro* and using repeated transcriptomic sequencing efforts (Patil, 2005). These methods have been written into a single python script, along with supporting reference files, and is presented as BiGSMAll (BacterIal Genome-Scale Metabolic models for AppLied reverse ecoLogy), available in a public Github repository at <https://github.com/mjenior/bigsmall>.

**Mathematical validation of metabolite importance scores.** As sequencing replicates of *in vivo* transcriptomes were not feasible, we sought to apply a modified Monte-Carlo simulation to distinguish calculated metabolite importances due to transcriptional patterns, from consistently important compounds or those scores derived from random chance. With this in mind we took an iterative approach of randomizing assignment of transcript abundance to enzyme nodes and recalculating metabolite importances. Each substrate node accumulated a distribution of importance scores that were then used to calculate a confidence interval. Ultimately, scores from measured values were then compared against simulated medians and confidence intervals to assess for those outside the expectation by random chance for a network with the same topology. This is a superior approach to topology randomization as the landscape of the network itself was created through natural selection and any large-scale alterations to that would yield biologically uninformative comparators. These simulations and comparisons are available within the BigSmall workflow presented above.

**Anaerobic *in vitro C. difficile* growth curves.** The carbon-free variation of *C. difficile* Basal Defined Medium (NCMM) was prepared as previously described in Theriot et. al., 2014. Respective carbohydrate sources were added at a final concentration of 5000 mg/L, also as outlined in the aforementioned study. A solution of the required amino acids was made separately and added when noted at identical concentrations to the same study. 245 μl of final media mixes were added to a 96-well sterile plate arranged as shown in Table S4. A rich media growth control was also included, consisting of liquid Brain-Heart Infusion + 0.5% cysteine. Edge wells were skipped in order to avoid false changes to optical density due to dessication.

All culturing and growth measurement were performed anaerobically in a Coy Type B Vinyl Anaerobic Chamber (3.0% H, 5.0% CO2, 92.0% N, 0.0% O). *C. difficile* str. 630 was grown for 14 hours at 37° C in 3 ml BHI + 0.5% cysteine. Cultures were then centrifuged at 2000 rpm for 5 minutes and resulting pellets were then washed twice with sterile, anaerobic 1x phosphate-buffered saline (PBS). Washed pellets were resuspended in 3 ml more PBS and 5 μl of prepped culture was added the each growth well of the plate containing aliquoted media. The plate was then placed in a Tecan Sunrise plate reader, heated to 37° C. Plates were incubated for 24 hours with automatic optical density readings at 600 nm taken every 30 minutes. OD600 values were normalized to readings from wells containing sterile media of the same type at equal time of incubation. Growth rates and other curve metrics were determined by differentiation analysis of the measured OD600 over time in R to obtain the slope at each time point. Significant differences were determined by Holm-Bonferroni corrected 2-way ANOVA.

**Untargeted liquid chromatography-mass spectrometry for metabolite content of mouse cecal content.** Flash frozen aliquot of cecal content was

**Figure Legends**

**Figure 1 | Experimental timelines and *C. difficile* phenotype quantification 18 hours after infection.** **A)** Experimental timelines for the antibiotic treatments used to sensitize animals to *C. difficile* colonization and infection (n = 9 per treatment group). **B)** Toxin titer measured by activity in Vero cell rounding assay. *C. difficile* monoassociated with germfree mice displayed significantly more toxin activity than any antibiotic treatment (corrected *p*-values < 0.05). **C)** Vegetative *C. difficile* CFU assessed through differential plating (plated without germinant) and **D)** *C. difficile* spore CFU measured with an alternative plating strategy (heating to 65° C for 30 minutes, plated with germinant). No statistical differences were observed in vegetative cell quantification, however germfree mice had significantly more spores than antibiotic treatment groups (corrected *p*-values < 0.05). Differences in spore quantification were noted between antibiotic pretreatment groups, however they were not significant. Conventionally raised mice that received no antibiotics remained resistant to *C. difficile* colonization.

**Figure 2 | Select *C. difficile* gene set expression compared between treatment group.** The median of subsampled gene abundances, shown with 95% confidence interval. **A)** Select sigma factors which their expression has been linked to environment cues. **B)** Majority of known genes from the *C. difficile* sporulation pathway. Many key genes for sporulation display differential expression between treatment groups. **C)** Transcript recruitment to pathogenicity loci, including both the standard PaLoc genes as well as CdtR but not the genes for binary toxin as they are absent in str. 630. **D)** Genes associated with quorum sensing in *C. difficile* are detected at approximately even levels in each antibiotic pretreatment.

**Figure 3 | *C. difficile* expression of gene sets for alternative carbon metabolism across antibiotic pretreatments. A)** Ternary plot of all genes with with any transcripts mapping to them are included. Coordinates for each gene are determined based on their relative abundance across each of the antibiotic treatment. The more over-represented transcription of a gene is in a single condition, the more it is placed near the corner corresponding to that treatment. Points placed near the center have equal expression in each of the conditions measured, conversely those genes toward the sides and corners are overrepresented in the groups associated with those edges. Colored points within each gene set are sized relative to the greatest transcript abundance among the three conditions. Groups from (A) are also plotted separately in **B)** through **I)** without size differences so all points are visible.

**Figure 4 | Genome-scale bipartite metabolic modeling results using the transcriptome of *C. difficile* str. 630 in each colonized environment. A)** Largest component from the bipartite genome-scale metabolic model of *C. difficile* str. 630. This portion of the graph includes 1070 nodes (404 enymes and 666 substrates), with 2000 edges connecting them. The size of enzyme nodes is relative to the number of normalized reads mapping to nucleotide gene sequence.Transcript abundances are from *C. difficile* str. 630 during infection of cefoperazone treated mice. The expanded window displays metabolite importance algorithm and an example of a single importance calculation based on local gene transcription. **B)** Top scoring unique important substrates in each antibiotic treatment, top 25 from all were cross-referenced to each other. Monte Carlo simulated means are indicated by the lines, all substrates are calculated as more important than if transcripts abundances were randomly assigned throughout the entire metabolic network. **C)**  *In vitro* growth curves validating identified growth nutrients from (B). In black are CDMM controls variations with the indicated constituents. At least one experimental growth substrate indicated as uniquely important from each group supports *C. difficile* growth significantly more than – Glucose + Amino Acids control wells (corrected *p*-values << 0.001).

**Figure 5 | *in vivo* concentrations from targeted LC-MS analysis of predicted *C. difficile* nutrient sources from metabolic models. A)** Heatmap, group by treatment group on x-axis and hierarchical clustering on y-axis. Comparing No Antibiotics to each treatment group. **B)** Bar charts of the actual concentrations for the substrates identified through metabolic modeling, shown in Figure 4.

**Supplementary Figure 1 | Spore % of total CFU and toxin to spore correlations.** Continued analysis of wetlab phenotypic data. **A)** Spore %. **B)** Spore toxin correlation w/o germfree. **C)** Spore toxin correlation w/ germfree.

**Supplementary Figure 2 | *C. difficile* str. 630 exome coverage for each condition.** Normalized read abundances mapped to the full length genome of *C. difficile* str. 630 for each of the antibiotic pretreatments.

**Supplementary Figure 3 | Expression of specific KEGG gene families.** Description. A). B). C).

**Supplementary Figure 4 | Substrate scores outside of at least two standard deviations from simulated means.** Description. A). B). C).

**Supplementary Figure 5 | Additional growth curves.** Description.

**Supplementary Figure 6 | Acetate data supporting network output.** Description.

**Acknowledgments**

The authors would like to acknowledge members of the Young laboratory, Germfree Center, UM Sequencing Core, and Metabolon for their assistance in experimental design, execution, and data collection.

**Conflict of Interest**

The authors declare no conflict of interest.

**References**

Antunes, A., Camiade, E., Monot, M., Courtois, E., Barbut, F., Sernova, N. V., & Dupuy, B. (2012). Global transcriptional control by glucose and carbon regulator CcpA in Clostridium difficile. *Nucleic Acids Research*. http://doi.org/10.1093/nar/gks864

Antunes, A., Martin-Verstraete, I., & Dupuy, B. (2011). CcpA-mediated repression of Clostridium difficile toxin gene expression. *Molecular Microbiology*, *79*(4), 882–899. http://doi.org/10.1111/j.1365-2958.2010.07495.x

Antunes, L. C. M., Han, J., Ferreira, R. B. R., Loli, P., Borchers, C. H., & Finlay, B. B. (2011). Effect of antibiotic treatment on the intestinal metabolome. *Antimicrobial Agents and Chemotherapy*, *55*(4), 1494–1503. http://doi.org/10.1128/AAC.01664-10

Antunes, L. C. M., Han, J., Ferreira, R. B. R., Loli, P., Borchers, C. H., & Finlay, B. B. (2011). Effect of antibiotic treatment on the intestinal metabolome. *Antimicrobial Agents and Chemotherapy*, *55*(4), 1494–1503. http://doi.org/10.1128/AAC.01664-10

Boos, W., & Shuman, H. (1998). Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation. *Microbiology and Molecular Biology Reviews* *: MMBR*, *62*(1), 204–29. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=98911&tool=pmcentrez&rendertype=abstract

Bouillaut, L., Dubois, T., Sonenshein, A. L., & Dupuy, B. (2015). Integration of metabolism and virulence in Clostridium difficile. *Research in Microbiology*, *166*(4), 375–383. http://doi.org/10.1016/j.resmic.2014.10.002

Bouillaut, L., Dubois, T., Sonenshein, A. L., & Dupuy, B. (2015). Integration of metabolism and virulence in Clostridium difficile. *Research in Microbiology*, *166*(4), 375–383. http://doi.org/10.1016/j.resmic.2014.10.002

Bouillaut, L., Dubois, T., Sonenshein, A. L., & Dupuy, B. (2015). Integration of metabolism and virulence in Clostridium difficile. *Research in Microbiology*, *166*(4), 375–383. http://doi.org/10.1016/j.resmic.2014.10.002

Bouillaut, L., Self, W. T., & Sonenshein, A. L. (2013). Proline-dependent regulation of Clostridium difficile stickland metabolism. *Journal of Bacteriology*, *195*(4), 844–854. http://doi.org/10.1128/JB.01492-12

Carter, G. P., Purdy, D., Williams, P., & Minton, N. P. (2005). Quorum sensing in Clostridium difficile: Analysis of a luxS-type signalling system. In *Journal of Medical Microbiology* (Vol. 54, pp. 119–127). http://doi.org/10.1099/jmm.0.45817-0

Carter, G. P., Rood, J. I., & Lyras, D. (2012). The role of toxin A and toxin B in the virulence of Clostridium difficile. *Trends in Microbiology*. http://doi.org/10.1016/j.tim.2011.11.003

Chen, X., Katchar, K., Goldsmith, J. D., Nanthakumar, N., Cheknis, A., Gerding, D. N., & Kelly, C. P. (2008). A Mouse Model of Clostridium difficile-Associated Disease. Gastroenterology, 135(6), 1984–1992. http://doi.org/10.1053/j.gastro.2008.09.002

Darkoh, C., Dupont, H. L., Norris, S. J., & Kaplan, H. B. (2015). Toxin synthesis by Clostridium difficile is regulated through quorum signaling. *mBio*, *6*(2). http://doi.org/10.1128/mBio.02569-14

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., & Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infection and Immunity*, *80*(8), 2704–2711. http://doi.org/10.1128/IAI.00147-12

Deutscher, J., Francke, C., & Postma, P. W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and Molecular Biology Reviews* *: MMBR*, *70*(4), 939–1031. http://doi.org/10.1128/MMBR.00024-06

Donnelly, M. L., Fimlaid, K. A., & Shen, A. (2016). Characterization of *Clostridium difficile* spores lacking either SpoVAC or DPA Synthetase. *Journal of Bacteriology*, (April), JB.00986–15. http://doi.org/10.1128/JB.00986-15

Elsden, S. R., & Hilton, M. G. (1978). Volatile acid production from threonine, valine, leucine and isoleucine by clostridia. *Archives of Microbiology*, *117*(2), 165–172. http://doi.org/10.1007/BF00402304

Fabich, A. J., Jones, S. A., Chowdhury, F. Z., Cernosek, A., Anderson, A., Smalley, D., & Conway, T. (2008). Comparison of carbon nutrition for pathogenic and commensal Escherichia coli strains in the mouse intestine. *Infection and Immunity*, *76*(3), 1143–1152. http://doi.org/10.1128/IAI.01386-07

Ferreyra, J. A., Wu, K. J., Hryckowian, A. J., Bouley, D. M., Weimer, B. C., & Sonnenburg, J. L. (2014). Gut microbiota-produced succinate promotes C. Difficile infection after antibiotic treatment or motility disturbance. *Cell Host and Microbe*, *16*(6), 770–777. http://doi.org/10.1016/j.chom.2014.11.003

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D., & Shen, A. (2013). Global Analysis of the Sporulation Pathway of Clostridium difficile. *PLoS Genetics*, *9*(8). http://doi.org/10.1371/journal.pgen.1003660

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D., & Shen, A. (2013). Global Analysis of the Sporulation Pathway of Clostridium difficile. *PLoS Genetics*, *9*(8). http://doi.org/10.1371/journal.pgen.1003660

Fimlaid, K. A., & Shen, A. (2015). Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Current Opinion in Microbiology*. http://doi.org/10.1016/j.mib.2015.01.006

Freter, R. 1983. Mechanisms that control the microflora in the large intestine, p. 33-54. In D. J. Hentges (ed.), *Human*

*intestinal microflora in health and disease*. Academic Press, Inc., New York, NY.

Ho, T. D., & Ellermeier, C. D. (2015). Ferric uptake regulator fur control of putative iron acquisition systems in Clostridium difficile. *Journal of Bacteriology*, *197*(18), 2930–2940. <http://doi.org/10.1128/JB.00098-15>

Hundsberger, T., Braun, V., Weidmann, M., Leukel, P., Sauerborn, M., & von Eichel-Streiber, C. (1997). Transcription analysis of the genes tcdA-E of the pathogenicity locus of Clostridium difficile. *European Journal of Biochemistry / FEBS*, *244*(3), 735–742. <http://doi.org/10.1111/j.1432-1033.1997.t01-1-00735.x>

Janoir, C., Denève, C., Bouttier, S., Barbut, F., Hoys, S., Caleechum, L., & Dupuy, B. (2013). Adaptive strategies and patho genesis of clostridium difficile from In vivo transcriptomics. *Infection and Immunity*, *81*(10), 3757–3769.

<http://doi.org/10.1128/IAI.00515-13>

Joshi, N., & Fass, J. (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.

Jump, R. L. P., Polinkovsky, A., Hurless, K., Sitzlar, B., Eckart, K., Tomas, M., & Donskey, C. J. (2014). Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. *PLoS ONE*, *9*(7). http://doi.org/10.1371/journal.pone.0101267

Jump, R. L. P., Polinkovsky, A., Hurless, K., Sitzlar, B., Eckart, K., Tomas, M., & Donskey, C. J. (2014). Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. *PLoS ONE*, *9*(7). http://doi.org/10.1371/journal.pone.0101267

Kansau, I., Barketi-Klai, A., Monot, M., Hoy, S., Dupuy, B., & Janoir, C. (2016). Deciphering Adaptation Strategies of the Epi demic Clostridium difficile 027 Strain during Infection through In Vivo Transcriptional Analysis. *PLoS ONE*, 11(6). http://dx.doi.org/10.1371/journal.pone.0158204

Karasawa, T., Maegawa, T., Nojiri, T., Yamakawa, K., & Nakamura, S. (1997). Effect of arginine on toxin production by Clostridium difficile in defined medium. *Microbiology and Immunology*, *41*(8), 581–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9310936

Karlsson, S., Dupuy, B., Mukherjee, K., Norin, E., Burman, L. G., & Åkerlund, T. (2003). Expression of Clostridium difficile toxins A and B and their sigma factor TcdD is controlled by temperature. *Infection and Immunity*, *71*(4), 1784–1793. http://doi.org/10.1128/IAI.71.4.1784-1793.2003

Kim, J., Hetzel, M., Boiangiu, C. D., & Buckel, W. (2004). Dehydration of (R)-2-hydroxyacyl-CoA to enoyl-CoA in the fermentation of α-amino acids by anaerobic bacteria. *FEMS Microbiology Reviews*, *28*(4), 455–468. http://doi.org/10.1016/j.femsre.2004.03.001

Koenigsknecht, M. J., Theriot, C. M., Bergin, I. L., Schumacher, C. A., Schloss, P. D., & Young, V. B. (2015). Dynamics and establishment of Clostridium difficile infection in the murine gastrointestinal tract. *Infection and Immunity*, *83*(3), 934–941. http://doi.org/10.1128/IAI.02768-14

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, 1–10. http://doi.org/gb-2009-10-3-r25 [pii]\r10.1186/gb-2009-10-3-r25

Lawley, T. D., & Walker, A. W. (2013). Intestinal colonization resistance. *Immunology*, *138*(1), 1–11. http://doi.org/10.1111/j.1365-2567.2012.03616.x

Leffler, D. A., & Lamont, J. T. (2015). Clostridium difficile Infection. *New England Journal of Medicine*, *372*(16), 1539–1548. http://doi.org/10.1056/NEJMra1403772

Lessa, F. C., Gould, C. V, & McDonald, L. C. (2012). Current status of Clostridium difficile infection epidemiology. *Clinical Infectious Diseases* *: An Official Publication of the Infectious Diseases Society of America*, *55 Suppl 2*(Suppl 2), S65–70. http://doi.org/10.1093/cid/cis319

Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R., & McDonald, L. C. (2015). Burden of Clostridium difficile Infection in the United States. *The New England Journal of Medicine*, *372*(9), 825–834. <http://doi.org/10.1056/NEJMoa1408913>

Li, F., Hinderberger, J., Seedorf, H., Zhang, J., Buckel, W., & Thauer, R. K. (2008). Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from Clostridium kluyveri. *Journal of Bacteriology*, *190*(3), 843–850. http://doi.org/10.1128/JB.01417-07

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. http://doi.org/10.1093/bioinformatics/btp352

López-Maury, L., Marguerat, S., & Bähler, J. (2008). Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, *9*(8), 583–593. http://doi.org/10.1038/nrg2398

López-Maury, L., Marguerat, S., & Bähler, J. (2008). Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, *9*(8), 583–593. http://doi.org/10.1038/nrg2398

Lopez-Medina, E., Neubauer, M. M., Pier, G. B., & Koh, A. Y. (2011). RNA isolation of Pseudomonas aeruginosa colonizing the murine gastrointestinal tract. *Journal of Visualized Experiments* *: JoVE*, (55), 6–9. http://doi.org/10.3791/3293

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10. http://doi.org/10.14806/ej.17.1.200

M, N.-S., JD, H., SE, W., & D, S. (2015). Time-resolved amino acid uptake of Clostridium difficile 630Δerm and concomitant fermentation product and toxin formation. *BMC Microbiology*, 281. http://doi.org/10.1186/s12866-015-0614-2

M, N.-S., JD, H., SE, W., & D, S. (2015). Time-resolved amino acid uptake of Clostridium difficile 630Δerm and concomitant fermentation product and toxin formation. *BMC Microbiology*, 281. http://doi.org/10.1186/s12866-015-0614-2

Matamouros, S., England, P., & Dupuy, B. (2007). Clostridium difficile toxin expression is inhibited by the novel regulator TcdC. *Molecular Microbiology*, *64*(5), 1274–1288. http://doi.org/10.1111/j.1365-2958.2007.05739.x

Moncrief, J. S., Barroso, L. A., & Wilkins, T. D. (1997). Positive regulation of Clostridium difficile toxins. *Infection and Immunity*, *65*(3), 1105–1108.

Naaber, P., Smidt, I., Štšepetova, J., Brilene, T., Annuk, H., & Mikelsaar, M. (2004). Inhibition of Clostridium difficile strains by intestinal Lactobacillus species. *Journal of Medical Microbiology*, *53*(6), 551–554. http://doi.org/10.1099/jmm.0.45595-0

Ng, K. M., Ferreyra, J. a, Higginbottom, S. K., Lynch, J. B., Kashyap, P. C., Gopinath, S., & Sonnenburg, J. L. (2013). Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature*, *502*(7469), 96–9. http://doi.org/10.1038/nature12503

Patil, K. R., & Nielsen, J. (2005). Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the National Academy of the Sciences of the United States of America*, *102*(8), 2685–2689.

Patil, K. R., & Nielsen, J. (2005). Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the National Academy of the Sciences of the United States of America*, *102*(8), 2685–2689.

Permpoonpattana, P., Phetcharaburanin, J., Mikelsone, A., Dembek, M., Tan, S., Brisson, M. C., & Cutting, S. M. (2013). Functional characterization of Clostridium difficile spore coat proteins. *Journal of Bacteriology*, *195*(7), 1492–1503. http://doi.org/10.1128/JB.02104-12

Saujet, L., Pereira, F. C., Henriques, A. O., & Martin-Verstraete, I. (2014). The regulatory network controlling spore formation in Clostridium difficile. *FEMS Microbiology Letters*, *358*(1), 1–10. http://doi.org/10.1111/1574-6968.12540

Scaria, J., Suzuki, H., Ptak, C. P., Chen, J.-W., Zhu, Y., Guo, X.-K., & Chang, Y.-F. (2015). Comparative genomic and phenomic analysis of Clostridium difficile and *Clostridium sordellii*, two related pathogens with differing host tissue preference. *BMC Genomics*, *16*(1), 448. http://doi.org/10.1186/s12864-015-1663-5

Schnetz, K., Toloczyki, C., & Rak, B. (1987). B-glucoside (bgl) operon of Escherichia coli K-12: Nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two Bacillus subtilis genes. *Journal of Bacteriology*, *169*(6), 2579–2590.

Schubert, A. M., Sinani, H., & Schloss, P. D. (2015). Antibiotic-induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against Clostridium difficile. *mBio*, *6*(4). http://doi.org/10.1128/mBio.00974-15

Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., & Parkhill, J. (2006). The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nature Genetics*, *38*(7), 779–786. http://doi.org/10.1038/ng1830

Songer, J. G., & Anderson, M. A. (2006). Clostridium difficile: An important pathogen of food animals. *Anaerobe*. http://doi.org/10.1016/j.anaerobe.2005.09.001

Theriot, C. M., Koenigsknecht, M. J., Jr, P. E. C., Hatton, G. E., Nelson, A. M., Li, B., & Young, V. B. (2014). Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. http://doi.org/10.1038/ncomms4114.Antibiotic-induced

Theriot, C. M., Koumpouras, C. C., Carlson, P. E., Bergin, I. I., Aronoff, D. M., & Young, V. B. (2011). Cefoperazone-treated mice as an experimental platform to assess differential virulence of Clostridium difficile strains. *Gut Microbes*, *2*(6), 326–334. http://doi.org/10.4161/gmic.19142

Tracy, B. P., Jones, S. W., Fast, A. G., Indurthi, D. C., & Papoutsakis, E. T. (2012). Clostridia: The importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. *Current Opinion in Biotechnology*. http://doi.org/10.1016/j.copbio.2011.10.008

Van Beilen, J. W. A., & Brul, S. (2013). Compartment-specific pH monitoring in Bacillus subtilis using fluorescent sensor proteins: A tool to analyze the antibacterial effect of weak organic acids. *Frontiers in Microbiology*, *4*(JUN). http://doi.org/10.3389/fmicb.2013.00157

Wilson, K. H., & Perini, F. (1988). Role of competition for nutrients in suppression of Clostridium difficile by the colonic microflora. *Infection and Immunity*, *56*(10), 2610–2614.

Young, V. B., & Schmidt, T. M. (2004). Antibiotic-Associated Diarrhea Accompanied by Large-Scale Alterations in the Composition of the Fecal Microbiota. Journal of Clinical Microbiology, 42(3), 1203–1206. http://doi.org/10.1128/JCM.42.3.1203-1206.2004

Yuille, S., Mackay, W. G., Morrison, D. J., & Tedford, M. C. (2015). Optimising gut colonisation resistance against Clostridium difficile infection. *European Journal of Clinical Microbiology and Infectious Diseases*. <http://doi.org/10.1007/s10096-015-2479-6>