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

Authors: Matthew L. Jenior, Jhansi L. Leslie, Vincent B. Young, and Patrick D. Schloss*

Abstract

Clostridium difficile infection (CDI) has grown to be the most prevalent cause of hospital acquired infection in the United States. Susceptibility to CDI is induced by recent antibiotic exposure, which is known to alter the structure of the gut microbiome and to affect the availability of growth nutrients in the gut. We hypothesized that *C. difficile* is a generalist that adapts its physiology to the nutrients available within the gut. We orally challenged C57BL/6 mice that previously received one of three antibiotics with *C. difficile* and demonstrated that it was able to colonize the cecum within 18 hours of infection. However, levels of both spore and toxin production, which are known to be affected by nutrient availability, varied between each antibiotic treatment group. To more closely investigate the specific responses of *C. difficile* as it colonized the cecum, we performed *in vivo* transcriptional analysis of *C. difficile* from cecal content of infected mice. This approach revealed variation in expression of genes that drive life-cycle switches as well as metabolic pathways associated with catabolizing a variety of carbon sources such as carbohydrates, amino acids, and amino sugars. To assess which substrates *C. difficile* was most likely exploiting in each antibiotic-perturbed microbiome,

we developed a novel metabolite scoring algorithm within the genome-scale bipartite metabolic network of *C. difficile* that incorporated both network topology and transcript abundance to infer the likelihood that a given metabolite was acquired from the environment. Applying this approach, we found that *C. difficile* indeed occupies alternative nutrient niches across each antibiotic-perturbed microbiome and that the highlighted metabolites support significant growth, *in vitro*. *In silico* findings were further validated using *in vivo* metabolomic analysis to compare resitant and susceptible cecal conditions. These results support the hypothesis that consumption of N-acetyl-D-glucosamine and Stickland fermentation substrates are central components of *C. difficile*'s metabolic strategy and pathogenesis. 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 to prevent *C. difficile* colonization including potential pre- or probiotic therapies.

Significance

Clostridium difficile infection has grown to be the largest single cause of hospital-acquired infection in the United States. A compromised gut microbiota, typically through previous antibiotic treatment, is a prerequisite feature of *C. difficile* colonization susceptibility. This loss of resistance has also been described across classes of antibiotics, in which many result in distinct gut communities and each presumably presents individual metabolic challenges to *C. difficile* as it colonizes. In this study we demonstrate that not only does *C. difficile* alter its pathogenesis between differentially sensitized hosts, but also exploits separate available nutrient niches across these environments. We accomplish this through an *in vivo* murine model of infection,

characterized using transcriptomic-enabled genome-scale metabolic modeling in combination with untargeted metabolomic analysis. Our results support that *C. difficile* possesses a highly plastic nutrient niche space, allowing it to successfully infect numerous different hosts and ultimately cause disease. This work also provides evidence that *C. difficile* virulence may be driven by the concentrations of specific carbohydrate sources that are accessible for growth during each unique infection.

Introduction

Infection by the Gram-positive, spore-forming bacterium *Clostridium difficile* has increased in both prevalence and severity across numerous countries during the last decade (1). In the United States, *C. difficile* was estimated to have caused >500,000 infections and resulted in ~$4.8 billion worth of acute care costs in 2014 (2). *C. 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 (3). It has been shown that antibiotic therapy alters the structure and function of the gut microbiota making it susceptible to colonization by *C. difficile* (4). 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 (5).

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 which have been shown to result in susceptibility *C. difficile* colonization (6–8). In this case,

the antibiotics chosen significantly impact the structure and diversity of the cecal microbiome (Fig. S1a & S1b). It has been further demonstrated that at 18 hours after being introduced to a cefoperazone treated mouse, *C. difficile* reached its maximum vegetative cell density in the cecum (9). This provided a single timepoint to measure the largest population of metabolically active *C. difficile*. 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 (7, 10–12). 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 an intact gut community to exclude *C. difficile* colonization is suggestive of the nutrient-niche hypothesis in which an organism must be able to utilize a subset of available resources better than all competitors to colonize the intestine (13, 14).

Based on its genome sequence and *in vitro* growth characteristics, *C. difficile* appears 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 (15–17). 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 and sporulation (18, 19). 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 (20, 21). 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 transcription of virulence factors, *in vitro* (22, 23), with some work characterizing transcription during colonization of germfree mice (24, 25). More relevant to nutrient acquisition, *C. difficile* up-regulated several phosphotransferase systems (PTS) and ABC transporters in germfree mice, alluding to metabolic adaptation to nutrient availability *in vivo* (25). 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 substrates. Metabolomic analyses have also been used to more directly assay changes in bacterial metabolism as they relate to CDI (7, 12); however, these methods cannot focus on *C. difficile*-specific metabolites and more closely resemble echoes of metabolism, not currently active processes. In contrast to these approaches, *in vivo C. difficile* transcriptomic analysis from specific pathogen free (SPF) animals may provide unique insight into its active metabolic pathways in a more realistic model of infection. Integrating transcriptomic data with genome-scale metabolic modeling has previously aided in identifying the most active aspects of an organism's metabolism and which substrates are preferred by the organism (26–28). Applying these methods to study *C. difficile* colonization would allow us to directly test the nutrient-niche hypothesis.

Founded 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* to the gut environment caused by different classes of antibiotics. The antibiotics used in this study included streptomycin (Fig. 1a), cefoperazone (Fig. 1b), and clindamycin (Fig. 1c). These antibiotics differentially affect the structure of the gut microbiota (8). Each has also been shown to alter the gut metabolome relative to untreated animals (7, 10, 12). 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 monoassociated germfree mice, we observed that *C. difficile* adapted its nutrient utilization profile to colonize to high levels and express its virulence factors.

Results

Levels of *C. difficile* sporulation and toxin activity vary between antibiotic-treated specific pathogen free and germfree mice.

Due to the connection between metabolism, sporulation, and toxin production in *C. difficile*, we measured sporulation and toxin production at 18 hours post infection in each group. There was not a significant difference in the number of vegetative cells between any susceptible condition tested (Fig. 2a). All antibiotic treated (Table 1) specific pathogen free (SPF) and germfree (GF) animals were colonized to ~1×10^8 colony forming units (c.f.u.) per gram of content, while untreated SPF 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 (P = 0.005, 0.008, 0.003) were detected in ex-GF mice than in the antibiotic treated mice (Fig. 2b). The spore densities in both streptomycin and clindamycin-treated mice were also generally higher than that in cefoperazone-treated mice. There was significantly more toxin activity in ex-GF animals than any other colonized group (all P <= 0.001), but toxin titer also varied between antibiotic treatment groups (Fig. 2c). Although similar toxin activity was found in both the cefoperazone and clindamycin-treated groups, toxin titer was below the limit of detection in most 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.

***C. difficile* adapts the expression of genes for virulence and key sigma factors that are under the control of environmental nutrient concentrations.** To more closely investigate the responses of *C. difficile* to colonizing distinct susceptible gut environments, we performed whole transcriptome analysis of *C. difficile* during infection of the antibiotic treatment models. We then narrowed our analysis to focus on genes that control or code for functions that have been linked to nutrient concentrations in the intestines during CDI. After observing differences in spore load, we first examined transcription of the most highly expressed genes in the *C. difficile* sporulation pathway (29–32) (Fig. 3a). Across the four conditions where *C. difficile* colonized, we observed transcriptional profiles consistent with observed spore levels (Fig. 2b). The mice treated with cefoperazone had the lowest spore density and had the highest level of expression

for the anti-sigma factors *spoVG* and *spoVS*. The produts of these genes are involved in suppressing expression of genes found later in the sporulation pathway (33). Streptomycin-treated mice had the next highest density of spores and the highest expression of genes associated with sporulation activation (*spoIIAB*/*spoIIE*), but they also had relatively high levels of expression of *sspA* and *sspB*, which are genes that code for effectors that protect DNA from damage during dormancy. Next, in mice treated with clindamycin, *C. difficile* expressed genes associated with late stages of sporulation, including those for spore coat components (*cdeC*, *cotD*, and *cotJB2*), spore formation (*spoIVA*, *spoVB*, and *spoVFB*), and *sspA* and *sspB*. Finally, GF mice harbored the highest density of spores and those *C. difficile* primarily expressed the dormancy genes linked with the latest stages sporulation. Together these data demonstrate that *C. difficile* differentially expressed genes associated with sporulation that corresponded to the presence of spores in the cecum.

Expression of genes for quorum sensing and pathogenicity have been linked to changes in the nutrients that can be found in the environment of *C. difficile*. Both the *agr* locus and *luxS* gene are thought to be associated with inducing the expression of *C. difficile* virulence in several strains (34, 35). Considering the link between quorum sensing genes and toxin production, we expected the expression of genes for quorum sensing and toxin production and toxin titer to be concordant. Based on this model, we expected GF mice to have the highest levels of expression of genes for toxin production (Fig. 3b) and quorum sensing (Fig. 3c); however, these transcripts were not found in the GF mice. We also observed the highest level of expression for quorum sensing genes in cefoperazone-treated mice, but *tcdA* expression in these animals was not the highest

among the different treatment groups. Interestingly, the levels of expression for genes associated with toxin production did not match the toxin titers observed in the animals. These results suggest that the relationship between toxin titer and the expression of genes for toxin production is even more complex than current models indicate.

We next focused on the regulators of metabolic pathways. Sigma factors are master regulators and a subset have been shown to integrate signals from intra- and extracellular nutrient concentrations (20, 21, 31, 36). The transcription of the global repressor *codY* is responsive to intracellular concentrations of *C. difficile* energy sources (37). Highest transcription for this gene was found in cefoperazone-treated and GF mice (Fig. 3d). The regulation networks of CodY and CcpA are highly interconnected, with the expression of *ccpA* specifically linked to local concentration of rapidly metabolizable carbon sources (38). Cefoperazone-treated mice also exhibited increased transcription of *ccpA*, but the GF condition did not follow the same pattern. CcpA acts directly on *spo0A* (Fig. 3d), which positively regulates initiation of the sporulation pathway in *C. difficile*. Transcripts for *spo0A* were highly abundant in all conditions tested except for clindamycin-treated mice, where it was still moderately detectable. The sig-family of sigma factors is under the control of spo0A and regulate different stages of sporulation. The genes from this family with the highest total transcription (*sigA1*, *sigF*, *sigG*, *sigH*, and *sigK*) each demonstrated a unique pattern of expression between conditions. These results indicate that complete expression of sporulation likely integrates multiple levels of signaling and is more complex than a single metabolic switch. Both CcpA and Spo0A also control pathogenicity by regulating toxin production (Fig. 3d). We found expression of the toxin negative regulator *tcdC* in

all of the antibiotic-treated groups, but no detectable transcripts for the positive toxin A/B regulator *tcdR* were seen in any treatment. In addition to its effects on sporulation and virulence, CcpA also regulates the expression of other sigma factors that generally mediate distinct forms of *C. difficile* metabolism as needed. These targets include *rex* (general fermentation regulator) and *prdR* (Stickland fermentation regulator) (Fig. 3d). Although the expression of both has been shown to be linked to environmental proline concentrations, *rex* integrates additional signals from the intracellular NADH/NAD+ ratio to also control carbohydrate fermentation. Low-level transcription of *prdR* was found across all conditions, however *C. difficile* expression the *rex* gene highly in both cefoperazone-treated and GF mice. Combined, the variable expression of these sigma factors support the hypothesis that *C. difficile* adapts expression metabolism to fit its needs between colonized environments.

Gene sets from multiple *C. difficile* metabolic pathways are differentially expressed between colonized environments. In the context of similar colonization between antibiotic-treated animals, differential expression of global metabolic control mechanisms that are under the control of specific nutrient concentrations suggests that *C. difficile* adapts to each environment when in competition with the resident microbiota. To test this further, we quantified the total expression for all KEGG annotated genes in the *C. difficile* genome (Fig. S2A). We then focused on general differences in carbohydrate (Fig. S2b) and amino acid (Fig. S2C) metabolism in order to asses for apparent differences in the utilization of carbon sources by *C. difficile* across environments. However aside from overall lower expression of most gene families in GF mice, no other clear trends were evident at this broad level of analysis so we moved

toward a more fine-scale resolution of annotation and focused on specific gene sets known to contribute to certain forms of *C. difficile* metabolism (Table S1). Also, to more effectively compare between colonized states, we calculated the percentage of total expression between antibiotic-treated conditions for each gene (Fig. 4). We then identified the condition in which each gene was most highly transcribed and adjusted the size of the corresponding point relative to the largest transcipt abundance. This demonstrated that genes involved in amino acid catabolism had the greatest amount of expression overall, relative to other gene sets. This category includes those enzymes involved in Stickland fermentation (*arg*, *fdh*, *grd*, and *prd* loci) as well as several general peptidases (*pep* family). These results indicated that catabolizing environmental amino acids may be important for the growth of *C. difficile* during infection.

To more clearly indentify associations of gene sets with each condition, we also analyzed each set separately. First, we found that the expression of genes associated with amino acid catabolism were expressed at nearly consistent levels across the conditions. This was in agreement with the high level of overall expression associated with these genes. Additionally, genes for the metabolism of the host-derived amino sugars N-acetylglucosamine and N-acetylmannosamine were also expressed at consistent levels across each treatment group (*glm*, *nan*, *mur*, and *acd* loci). Along similar lines with related molecules, a number of genes for certain monosaccharides entering (*gal*, *man*, *pmi*, and *tag* loci) and leading through glycolysis (*fba*, *fbp*, *gap*, and *pfk*), as well as catabolism of the polysaccharides trehalose and cellibiose (*treA* and *celG*) were expressed relatively evenly between each condition. Combined, these

findings suggest that catabolism of amino acids and specific carbohydrates are likely core components of the *C. difficile* nutritional strategy during infection.

Aside from those gene sets that were equally expressed across conditions, there were also large scale differences in expression of certain pathways between groups of mice. We chose to assess sugar transport systems have been associated with adaptive expression of phosphotransferase systems (PTS) and ABC transporters with many known differences in substrate specificities (25). Among the genes classified as PTS transporters were overrepresented in both clindamycin and streptomycin-treated mice, while ABC sugar transporters were overrepresented in the cefoparazone-treated mice. The most stark differences were seen in transcription for genes involved in sugar alcohol catabolism. Expression of these genes was entirely absent from clindamycin-treated mice and expression of genes for mannitol utilization (*mtl* operon) were overrepresented in cefoparazone-treated mice and expression of genes for sorbitol utilization (*srl* operon) were overrepresented in streptomycin-treated mice. Concordant patterns also emerged in genes associated with fermentation end steps and polysaccharide degradation. Short chain fatty acids (SCFAs) and alcohols are the end products of both carbohydrate and amino acid fermentation in *C. difficile* through separate pathways with shared terminal steps. Transcripts for genes involved in *C. difficile* butyrate/butanol metabolism (*ptb*, *buk1*, *cat2*, and *adhE*) were more abundant in clindamycin-treated mice. Additionally, alpha/beta-galactosidase genes (*aglB* and *bglA*) were also overrepresented in clindamycin-treated mice. Together these patterns suggested that polysaccharide fermentation occurred this condition. More subtle differences were seen in those gene associated with glycolysis. This category includes

genes for not only the steps of glycolysis, but also several genes that mediate entry points of monosaccharides to glycolysis. Transcripts for several genes in this group (*eno*, *gapA*, *gpmI*, *tpi*, and *pyk*) were overrepresented in cefoparazone-treated mice, however *fruK* was overrepresented in streptomycin-treated mice which catalyzes the committed step of glycolysis. Overall, these results support the hypothesis that *C. difficile* is able to adapt its metabolism to fit the nutrient availability across different susceptible environments.

Structure of genome-scale bipartite metabolic model underscores known bacterial metabolism. To further investigate which metabolites were differentially utilized between conditions, we represented the metabolic network of *C. difficile* as a directed bipartite graph using the genome annotation. Enzymes and metabolites were represented by nodes and their interaction by the edge between the nodes (Fig. 5a). To validate our metabolic network, we calculated betweenness centrality (BC) and overall closeness centralization index (OCCI) for all enzyme and metabolite nodes in the bipartite metabolic network of *C. difficile* generated for this study (Table S2). In biological terms, BC reflects the amount of influence a given hub has on the overall flow of metabolism through the network (39) and OCCI indicates those enzymes and substrates that are the most central components of the organism's metabolism (40). For both enzymes and substrates, the 18 of top 20 nodes with the highest BC values were involved in glycolysis, fermentation, and amino acid synthesis. In agreement, almost all nodes with the largest OCCI values were involved in glycolysis and amino acid synthesis as well. Enzymes that scored highly in both metrics included pyruvate kinase, aspartate aminotransferase, and formate C-acetyltransferase while substrates

consistently scoring most highly were pyruvate, acetyl-CoA, D-glyceraldehyde 3-phosphate. This indicated to us that the topology of the network reflects established bacterial physiology.

Metabolite importance algorithm reveals adaptive nutritional strategies of *C. difficile* during infection across distinct environments. Moving beyond a strictly topological analysis of the *C. difficile* metabolic network, we sought to utilize transcriptomic data to infer which metabolites *C. difficile* is most likely to obtain from its environment in each condition. To accomplish this we mapped normalized transcript abundances to the enzyme nodes in the network. Due to the coupling of transcription and translation in bacteria, we were able to use this information as a proxy for enzyme levels. The importance of each metabolite was measured as the log-transformed difference between the average transcript levels of enzymes that use the metabolite as a substrate and those that generate it as a product (Fig. 5b). A metabolite with a high importance score is most likely obtained from the environment because the expression of genes for enzymes that produce the metabolite are low. Then, using a Monte Carlo-style simulation, we generated a random transcript abundance distribution for each enzyme node to then calculate new metabolite importance scores for each iteration. We then created a confidence interval of scores for each metabolite that would likely result from random noise (41). This provided a standard of comparison for actual importance scores from single timepoint measurments, and ultimately allow for computing the significance level that a given score has a high probability of being excluded from its associated null hypothesis score distribution.

Applying these methods to the *C. difficile* transcriptomic data collected from the *in vivo* CDI models, we sought to assess differential patterns of metabolite importance. We first ranked the importance scores to identify the most important metabolites for each treatment group (Table S3). To identify the core metabolites that are essential to *C. difficile* in any condition, we compared the highest 50 scoring, significant metabolites from each treatment group ($P < 0.05$) (Fig. 6a). The host derived amino sugar N-acetyl-D-glucosamine was found to be consistently important. Components of the Stickland fermentation pathway were also found to be important to *C. difficile* in all conditions tested including proline, 3−hydroxybutanoyl−CoA, and formate (42–44). This indicated that these metabolites may be an integral component of the nutrient niche for *C. difficile* in any infection condition. Additionally, acetate was found to be important in all conditions, but was just below the significance cutoff in GF mice (Table S3). It has been shown that *C. difficile* metabolizes acetate for use in glycolysis (45). We directly tested the relative concentration of acetate in cefoperazone-treated *C. difficile*-infected mice versus mock-infected mice. We found that *C. difficile* colonization led to a significant decrease in the levels of acetate (Fig. S4) suggesting that *C. difficile* was utilizing acetate in the cecum. These findings provided validation for our metabolite importance algorithm as well as supporting known elements of *C. difficile* metabolism.

Returning to our hypothesis that *C. difficile* adapts its metabolism to fit the surrounding community, we identified those metabolites that were uniquely important to each condition in which *C. difficile* colonized. We cross-referenced the top 25 positively scoring, significant substrates ($P < 0.05$) between treatment groups to uncover the most important patterns of nutrient utilization by *C. difficile* in each (Fig. 6b). Each group of

metabolites contained at least one known carbohydrate growth substrate of *C. difficile* (7, 46). This included sorbitol, galacititol, mannitol, N-acetylneuraminic acid, and salicin. Furthermore, in GF mice where no other competitors are present, our model indicated that *C. difficile* was more likely to acquire numerous amino acids from the environment instead of expending energy to produce them itself. These data support the hypothesis that *C. difficile* may exploit alternative nutrient sources between the susceptible environments it colonizes.

Carbohydrates predicted to be important using network-based approach differentially supported *C. difficile* growth *in vitro*. To validate the biological relevance of substrates identified as uniquely important to *C. difficile* metabolism through our network-based analysis, we tested whether *C. difficile* was able to utilize each substrate for *in vitro* growth (Fig. 6c). This was performed using a modified defined *C. difficile* minimal media (7), supplemented individually with the selected carbohydrates implicated by high importance scores. Because *C. difficile* is auxotrophic for several amino acids, it was necessary to include amino acids in the minimal media despite the capability to achieve modest growth through Stickland fermentation of these substrates. This focused our analysis on carbohydtrates and made the most effective negative control growth in media lacking carbohydrates but containing amino acids.

Relying on network-based analysis a single carbohydrate source, N-acetyl-D-glucosamine, was found to be important to *C. difficile* in each condition tested (Fig. 6a). When tested for improved growth, significantly more growth (Max OD_{600} = 0.774) was observed compared to negative controls ($P < 0.001$). This provided evidence that N-acetyl-D-glucosamine, derived from the host mucus layer, may be a central component

364 of the *C. difficile* nutritional niche during infection. Upon analysis of carbon sources
365 identified as distinctly important across treatment groups, we found that at least one
366 substrate from cefoperazone-treated, clindamycin-treated, and GF groups provided high
367 levels of *C. difficile* growth relative to negative controls ($P < 0.001$). This included
368 mannitol (cefoperazone; Max $OD_{600} = 0.464$), salicin (clindamycin; Max $OD_{600} = 0.869$),
369 and N-acetylneuraminate (GF; Max $OD_{600} = 0.439$). Dinstinctly important from
370 streptomycin-treated mice, D-sorbitol (streptomycin; Max $OD_{600} = 0.202$) and galactitol
371 (streptomycin; Max $OD_{600} = 0.211$), significantly altered growth compared to negative
372 controls ($P < 0.001$), however did not overall growth rate (Table S4). Maximum growth
373 rate analysis for each carbohydrate also indicated potential hierarchy in carbohydrate
374 preference (Table S4) which is as follows: N-acetyl-D-glucosamine ($k = 0.085$ hr^{-1}),
375 salicin ($k = 0.077$ hr^{-1}), mannitol ($k = 0.044$ hr^{-1}), N-acetylneuriminate ($k = 0.024$ hr^{-1}),
376 sorbitol ($k = 0.022$ hr^{-1}), and finally galactitol ($k = 0.02$ hr^{-1}). Although not a
377 carbohydrate, we also tested acetate for the ability to support *C. difficile* growth *in vitro*,
378 but showed no improvement over negative controls (Fig. S5). These data suggested
379 that *C. difficile* was well-suited to adapt its metabolism toward nutrient sources that were
380 differentially available across susceptible mouse ceca.

381 **Metabolomic analysis supported network-based results for metabolic adaptation**
382 **and uncovered additional aspects of *C. difficile*'s nutrient niche.** In order to validate
383 results from transcriptomic and metabolic modeling analyses, we utilized untargeted
384 mass-specrometry (Table S5) to measure the relative *in vivo* concentrations of
385 metabolites highlighted by network-based importance scores. We compared cecal
386 content from mock *C. difficile*-infected susceptible and resistant mice at the same 18

hour time point as the previous analyses. Initially N-acetylglucosamine + N-acetylgalactosamine (Fig. 7a) and salicylate (Fig. 7b) were found to be significantly decreased in all susceptible treatment groups compared to resistant controls ($P < 0.05$). Additionally, we found that galactitol (Fig. 7c), mannitol + sorbitol (Fig. 7d), and N-acetylneuraminate (Fig. 7e) were significantly increased in both cefoperazone-treated SPF and GF mice ($P < 0.05$). Finally, these assays also revealed that proline (Fig. 7f) was significantly increased in all *C. difficile* susceptible conditions ($P < 0.05$), while glycine (fig. 7g) was found to also be concordantly increased. These amino acids are required together for *C. difficile* to perform Stickland fermentation. Together metabolomic results further supported network-derived importance scores and our hypothesis that *C. difficile* adapts its metabolic strategy in distinct susceptible environments.

Discussion

Collectively, our results support the hypothesis that *C. difficile* can adapt its metabolism to the available niche landscape across susceptible gut environments and give insight to the adaptive strategies that *C. difficile* can use to colonize diverse human microbiota. All of our *in vivo*, *in vitro*, and *in silico* data demonstrate the plasticity of *C. difficile* to effectively change its metabolism to utilize alternative resources for growth. This may be the result of increased concentration of particular metabolites as a consequence of concordant decreases in the population of one or more competitors for those resources. In addition to uncovering adaptive strategies of *C. difficile*, our method is also able to identify consistent trends in metabolism across environments. The findings that N-acetyl-D-glucosamine and Stickland fermentations substrates were consistently among

the highest scoring shared metabolite among all tested conditions strongly indicates that these metabolites are central to the nutritional strategy of *C. difficile* and may be utilized in and condition in which they are available. The metabolism of both substrate types provides not only carbon and energy to *C. difficile*, but are also a source for nitrogen which is a limited resource in the mammalian lower GI tract (47).

Apart from exploring differential patterns in known metabolism, our modeling approach also allowed for the identification of emergent properties for the metabolic strategy of *C. difficile* during infection. One interesting result is the appearance of CO_2, an apparent metabolic end product, in the list of shared important metabolites (Fig. 6a). While this may be a shortcoming of the annotation, one group has posited that *C. difficile* may actually be autotrophic under certain conditions and could explain our results (48). This highlights that our method does not only identify growth substrates, but also reports any metabolites that are being removed from the environment. Although our modeling results are consistent with previously published work on the metabolism of *C. difficile*, there are potential limitations of this approach. Ultimately, the metabolite importance calculation is dependent on correct and existing gene annotation. In this regard it has been shown that the pathway annotations in KEGG are robust to missing elements (49), however this does not completely eliminate the possibility for this type of error. Due to the topology of the metabolic network, we were also unable to integrate stoichiometry for each reaction which may effect rates of consumption or production. In addition to computational limitations, our network-based approach simplifies several aspects of bacterial physiology including the assumptions that all detectable transcript is translated to effector protein, that each metabolite is available, and that only the neighboring

pathway is important in determining a metabolite's importance. In spite of these assumptions, the method outlined here supports known elements of *C. difficile* biology and our subsequent *in vitro* and *in vivo* metabolomic analysis confirmed the predictions made here.

Based on the evidence presented, our results support the hypothesis that *C. difficile* is a metabolic generalist and is able to catabolize alternative carbon sources across susceptible gut environments. This may be due to an inability to outcompete a collection of metabolic specialists in an intact community, and separate classes on antibiotics differentially eliminate these populations and allow for *C. difficile* colonization. This may be particularly true with competitors for proline and glycine as our data also suggests that overall *C. difficile* colonization resistance may be driven by the availability of fermentable amino acids. *C. difficile* may also have evolved to utilize the host-derived aminoglycan N-acetyl-D-glucosamine in all infection conditions as it is a consistently available nutrient source across susceptible hosts. Additionally, differential importance scores for carbohydrates in agreement with metabolomic results across distinct environments, and when considered with variable sporulation and toxin production, may indicate that *C. difficile* pathogenesis is directed by which carbohydrates are available for use. Furthermore , competition for nutrients may also potentially explain the success rate of fecal microbial transplant (FMT), in that wholesale installation of a diverse range of specialized metabolic strategies is enough to outcompete *C. difficile* from the majority of perturbed gut environments. In light of these findings, our metabolic network platform may also prove informative for generating hypotheses through reverse ecology that could ultimately lead to uncovering new interaction between species that ultimately

impact host health (50). In conclusion, *C. difficile* is is able to optimize its nutritional strategy for each colonized gut environment. Our results implicate that further considerations are needed when attempting to design targeted prebiotic and probiotic therapies for the prevention or elimination of *C. difficile* from the human gut.

Methods

Animal care and antibiotic administration Adapted from the previously described model (51), six-to-eight week-old 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 GF C57BL/6 mice were obtained from a single breeding colony maintained at the University of Michigan and fed Laboratory Rodent Diet 5001 from LabDiet 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 distilled drinking water *ad libitum* for 5 days with 2 days recovery with untreated distilled 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, 1×10^3 *C. difficile* spores were administered to mice via oral gavage in phosphate-buffered saline (PBS) vehicle. Subsequent quantitative plating for c.f.u. was performed to ensure correct dosage. Infection negative

control animals were given an oral garage of 100 µl 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 c.f.u. 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** Cecal samples were weighed and serially diluted under anaerobic conditions (6% H, 20% CO_2, 74% N_2) with anaerobic PBS. Differential plating was performed to quantify both *C. difficile* spores and vegetative cells by plating diluted samles on CCFAE plates (fructose agar plus cycloserine, cefoxitin, and erythromycin) at 37° C for 24 hours under anaerobic conditions (52). It is important to note that the germination agent taurocholate was omitted from these plates in order to only quantify vegetative cells. In parallel, undiluted samples were heated at 60° C for 30 minutes to eliminate vegetative cells and leave only spores (53). These samples were serially diluted under anaerobic conditions in anaerobic PBS and plated on CCFAE with taurocholate at 37° C for 24 hours. Plating was simultaneously done for heated samples on CCFAE to ensure all vegetative cells had been eliminated.

***C. difficile* toxin titer assay** To quantify the titer of toxin in the cecum, a Vero cell rounding assay was performed (54). Briefly, filtered-sterilized cecal content was serially diluted 1:5 in PBS. As a control for toxin-mediated cell rounding the cecal content was diluted a further 1:2 by the addition of an equal volume of goat anti-toxin serum (T5000; TechLab). Vero cells were grown to a confluent monolayer in DMEM (Dulbecco's Modified Eagle's medium), supplemented with 10% heat-inactivated fetal bovine serum and 1% penicillin-streptomycin. The cells then were transferred to a conical tube and centrifuged at 1,000 rpm for 5 minutes to pellet the cells. The old media was removed and the cells were re-suspended in fresh media to a final concentration of 1×10^5 cells per 90µL. 90µL of the cell suspension were seeded in each well of a 96-well plate and incubated at 37° C in a 5% CO_2 humidified incubator for 4 hours. Following the incubation, cecal samples were added to the Vero cells and the plate was incubated overnight at 37° C. Plates were viewed after 24 hours at 10x magnification for cell rounding. The cytotoxic titer was defined as the $\log_{10}$ transformed reciprocal of the highest dilution that produced rounding in 80% of the cells. A more detailed protocol with product information can be found at: https://github.com/jlleslie/Intraspecific_Competition/blob/master/methods/Verocell_ToxinActivity_Assay.Rmd

16S rRNA gene sequencing DNA 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

523 and sequenced as described previously using an Illumina MiSeq sequencer (55). All 63
524 samples were sequenced on a single sequencing run.

525 **Sequence curation** The 16S rRNA gene sequences were curated using the mothur
526 software package (v1.36), as described previously (55). In short, paired-end reads were
527 merged into contigs, screened for quality, aligned to SILVA 16S rRNA sequence
528 database, and screened for chimeras. Sequences were classified using a naive
529 Bayesian classifier trained against a 16S rRNA gene training set provided by the
530 Ribosomal Database Project (RDP) (56). Curated sequences were clustered into
531 operational taxonomic units (OTUs) using a 97% similarity cutoff with the average
532 neighbor clustering algorithm. The number of sequences in each sample was rarefied to
533 2,500 per sample to minimize the effects of uneven sampling.

534 **RNA extraction, shotgun library preparation, and sequencing** To generate enough
535 mRNA biomass contributed by *C. difficile*, we pooled cecal content from all mouse
536 replicates into a single large isolation for each treatment group. Pooling was performed
537 in a sterile stainless steel mortar resting in dry ice and a small amount of 100% ethanol.
538 After all content for the given group was added, the sample was ground with a sterile
539 pestle to a fine powder and scraped into a sterile 50 ml polypropylene conical tube.
540 Samples were stored at -80° C until the time of extraction. Immediately before RNA
541 extraction, 3 ml of lysis buffer (2% SDS, 16 mM EDTA and 200 mM NaCl) contained in
542 a 50 ml polypropylene conical tube was first heated for 5 minutes in a boiling water bath
543 (57). The hot lysis buffer was added to the frozen and ground cecal content. The
544 mixture was boiled with periodic vortexing for another 5 minutes. After boiling, an equal
545 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° 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 µl of RNase-free water. Following the manufacturer'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 manufacturer's protocol. Completion of the reaction was assessed using PCR for the V4 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 manufacturer's protocol. Prior to library construction, quality and integrity as measured again using the Agilent RNA 6000 Pico Kit. Stranded RNA-Seq libraries were made constructed with the TruSeq Total RNA Library Preparation Kit v2, both using the manufacturer's protocol. The Agilent DNA High Sensitivity Kit was used to measure concentration and fragment size distribution before sequencing. High-throughput sequencing was performed by the University of Michigan Sequencing Core in Ann

Arbor, MI. For all groups, sequencing was repeated across 4 lanes of an Illumina HiSeq 2500 using the 2x50 bp chemistry.

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 (35) 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 (58) 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 KEGG: Kyoto Encyclopedia of Genes and Genomes (59). 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 (60) 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. Unless indicated otherwise, each collection of reads was then 1000-fold iteratively subsampled to 90% of the lowest sequence total within each analysis, and a median expression value for each gene was calculated.

Reaction Annotation & Bipartite Network Construction. The metabolism of *C. difficile* str. 630 was represented as a directed bipartite graph with both enzymes and metabolites as nodes. Briefly, models were semi-automatically constructed using KEGG ortholog (KO) gene annotations to which transcripts had been mapped. Reactions that each KEGG ortholog mediate were extracted from ko_reaction.list located in

/kegg/genes/ko/. KOs that do not mediate simple biochemical reactions (ex. mediate interactions of macromolecules) were omitted. Metabolites linked to each reaction were retrieved from reaction_mapformula.lst file located in /kegg/ligand/reaction/ from the KEGG release. Those reactions that did not have annotations for the chemical compounds the interact with are discarded. Metabolites were then associated with each enzyme and the directionality and reversibility of each biochemical conversion was also saved. This process was repeated for all enzymes in the given bacterial genome, with each enzyme and metabolite node only appearing once. The resulting data structure was an associative array of enzymes associated with lists of both categories of substrates (input and output), which could then be represented as a bipartite network. The final metabolic network of C. difficile str. 630 contained a total of 1205 individual nodes (447 enzymes and 758 substrates) with 2135 directed edges. Transcriptomic mapping data was then re-associated with the respective enzyme nodes prior to substrate importance calculations. Betweenness-centrality and overall closeness centralization indices were calculated using the igraph R package found at http://igraph.org/r/.

Metabolite Importance Calculation. The substrate importance algorithm (Fig. 5a) favors metabolites that are more likely acquired from the environment (not produced within the network), and will award them a higher score (Fig. 6b & 6c). The presumption of our approach was that enzymes that were more highly transcribed were more likely to utilize the substrates they act on due to coupled bacterial transcription and translation. If a compound was more likely to be produced, the more negative the resulting score would be. To calculate the importance of a given metabolite (m), we used rarefied

transcript abundances mapped to respective enzyme nodes. This was represented by t_o and t_i to designate if an enzyme created or utilized m. The first step was to calculate the average expression of enzymes for reactions that either created a given metabolite (i) or consumed that metabolite (ii). For each direction, the sum of transcripts for enzymes connecting to a metabolite were divided by the number of contributing edges (e_o or e_i) to normalize for highly connected metabolite nodes. Next the raw metabolite importance score was calculated by subtracting the creation value from the consumption value to weight for metabolites that are likely acquired exogenously. The difference was $\log_2$ transformed for comparability between scores of individual metabolites. This resulted in a final value that reflected the likelihood a metabolite was acquired from the environment.Untransformed scores that already equaled to 0 were ignored and negative values were accounted for by transformation of the absolute value then multiplied by -1. These methods have been written into a single python workflow, 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.

Transcriptome Bootstrapping and Probability Distribution Comparison. As sequencing replicates of *in vivo* transcriptomes was not feasible, we applied a Monte Carlo style simulation to distinguish calculated metabolite importances due to distinct transcriptional patterns for the environment measured from those metabolites that were constitutively important. We employed a 10,000-fold bootstrapping approach of randomly reassigning transcript abundance for enzyme nodes and recalculating metabolite importances. This approach was chosen over fitting a simulated

transcriptome to a negative binomial distribution because it created a more relevant standard of comparison for lower coverage sequencing efforts. Using this method, each substrate node accumulated a random probability distribution of importance scores which were then used to calculate the median and confidence interval in order to ultimately generate a p-value for each metabolite. This was a superior approach to switch randomization since the connections of the network itself was created through natural selection and any large-scale alterations would yield biologically uninformative comparisons(61). These calculations are also included within the standard bigSMALL workflow presented above and examples can be found in Fig. S6.

Measuring *in vivo* concentrations of acetate. Cecal contents were flash frozen in liquid nitrogen at the time of necropsy and subjected to short chain fatty acid quantification analysis using GC-MS (gas chromatography–mass spectrometry) as described in the targeted metabolomics section of Theriot et al., 2014. All assays were performed at the Michigan Regional Comprehensive Metabolomics Resource Core in Ann Arbor, MI.

Anaerobic *in vitro C. difficile* growth curves. The carbon-free variation of *C. difficile* Basal Defined Medium (NCMM) was prepared as previously described (7). Individual carbohydrate sources were added at a final concentration of 5 mg/mL and pair-wise carbohydrate combinations were added at 2.5 mg/mL each (5 mg/mL total). 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 clear-bottom plate. A rich media growth control was also included, consisting of liquid Brain-Heart Infusion + 0.5% cysteine. All culturing and growth measurement were

performed anaerobically in a Coy Type B Vinyl Anaerobic Chamber (3.0% H, 5.0% CO_2, 92.0% N, 0.0% O_2). *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 1 × 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. OD_{600} 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 OD_{600} over time in R to obtain the slope at each time point.

Quantification of *in vivo* metabolite relative concentrations. Metabolomic analysis performed by Metabolon (Durham, NC), a brief description of their methods is as follows. 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 possitive conditions using a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 µm) 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. Finally, samples were analyzed via negative ionization following elution from a hydrophilic interaction chromatography column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS n scans using dynamic exclusion. The scan range varied slighted between methods but covered 70-1000 m/z. Library matches for each compound were checked for each sample and corrected if necessary. Peaks were quantified using area under the curve.

Statistical methods. All statistical analyses were performed using R (v.3.2.0). Significant differences between community structure of treatment groups from 16S rRNA gene sequencing were determined with AMOVA in the mothur software package. Significant differences of Inv. Simpson diversity, CFU, toxin titer, and acetate concentration were determined by Wilcoxon rank-abundance test with Holm-Bonferroni correction. Significant differences for growth curves compared to no carbohydrate control (+ amino acids) were calculated using 2-way ANOVA with Holm-Bonferroni correction. Significance for metabolite importance scores was determined as described above. Wilcoxon rank-abundance test with Holm-Bonferroni correction was also used to calculate significant differences is relative metabolite concentrations from mass-spectrometry data.

Acknowledgements

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. This work was supported by funding from the National Institutes of Health to PDS (R01GM099514, P30DK034933, U19AI09087, and U01AI124255), VBY (P30DK034933, U19AI09087, and U01AI124255), a Translational Research Education Certificate grant to JLL (MICHR; UL1TR000433), and was partially supported by a predoctoral fellowship from the Cellular Biotechnology Training Program to MLJ (T32GM008353). Pooled and quality trimmed transcriptomic read data and experiment metadata are available through the NCBI Sequence Read Archive (SRA; PRJNA354635). Data processing steps for beginning from raw sequence data to the final manuscript are hosted at http://www.github.com/SchlossLab/Jenior_Modeling_PNAS_2016.

Author Information

Affiliations Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan. Matthew L. Jenior, Jhansi L. Leslie, & Patrick D. Schloss Ph.D.

Department of Internal Medicine & Infectious Diseases, University of Michigan Medical Center, Ann Arbor, Michigan. Vincent B. Young M.D. Ph.D.

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

Competing interests The authors declare no competing interest.

Corresponding author Correspondence to Patrick D. Schloss

729 **Table 1. Antibiotics used during *C. difficile* infection models.**
730

Antibiotic	Class	Target	Activity	Administration	Dosage
Cefoperazone	Cephalosporin (3rd generation)	Primarily Gram-positive bacteria, with increased activity against Gram-negative bacteria	Irreversibly crosslink bacterial transpeptidases to peptidoglycan and prevents cell wall synthesis	Drinking water Ad libitum	0.5 mg/ml
Streptomycin	Aminoglycoside	Active against most Gram-negative aerobic and facultative anaerobic bacilli	Protein synthesis inhibitor through binding the 30S portion of the 70S ribosomal subunit	Drinking water Ad libitum	5.0 mg/ml
Clindamycin	Lincosamide	Primarily active against Gram-positive bacteria, most anaerobic bacteria, and some mycoplasma	Protein synthesis inhibition through binding to the 23s portion of the 50S ribosomal subunit	Intraperitoneal injection	10 mg/kg body weight

Figure Legends

Figure 1 | Experimental time lines for mouse model pretreatments and *C. difficile* infection. 9 wild-type C57BL/6 mice across 3 cages were included in each treatment group. **(A)** Streptomycin or **(B)** cefoperazone administered *ad libitum* in drinking water for 5 days with 2 days recovery with untreated drinking water before infection, **(C)** a single clindamycin intraperitoneal injection one day prior to infection, or **(D)** no antibiotic pretreatment (for both SPF control and GF mice). If no antibiotics were administered in the drinking water, mice were given untreated drinking water for the duration of the experiment beginning 7 days prior to infection. At the time of infection, mice were challenged with 1×10^3 *C. difficile* str. 630 spores at the time of infection. Sacrifice and necropsy was done 18 hours post-challenge and cecal content was then collected.

Figure 2 | *C. difficile* sporulation and toxin activity quantification after 18 hours of infection. (A) Vegetative *C. difficile* c.f.u. per gram of cecal content. No significant differences were observed in between any group colonized by *C. difficile*. **(B)** *C. difficile* spore c.f.u. per gram of cecal content. Significantly more spores were detectable in GF mice compared to any of the antibiotic-treated SPF groups ($P < 0.05$). **(C)** Toxin titer from cecal content measured by activity in Vero cell rounding assay. GF mice also displayed significantly more toxin activity than all other groups ($P < 0.05$). Untreated mice in a,b,c had no detectable *C. difficile* or toxin activity and were significantly different from all other groups in each assay ($P < 0.01$). Median values are shown for each group with significant differences calculated using Wilcoxon rank-sum test with the Holm-Bonferroni correction. Dotted lines denote the limit of detection for both assays, and undetectable points are shown just below the limit of detection for clarity.

Figure 3 | Select *C. difficile* gene set expression compared between treatment group. Relative abundances of *C. difficile* transcript for specific genes of interest, rarefied to 90% of the total number of reads within each colonized condition shown. **(A)** Transcription for select genes from the *C. difficile* sporulation pathway with the greatest variation in expression between the conditions tested. **(B)** Relative abundances of transcript for genes that encode effector proteins from the *C. difficile* pathogenicity locus. **(C)** Transcript abundances for genes associated with quorum sensing in *C. difficile*. **(D)** Transcript relative abundance of select sigma factors which expression or activity is influenced by environmental metabolite concentrations. Asterisks (*) indicate genes from which transcript was undetectable.

Figure 4 | *C. difficile* expression of gene sets for carbon metabolism pathways across antibiotic pretreatments. Ternary plot indicating the relative abundance of transcripts for all *C. difficile* str. 630 genes across the three colonized antibiotic-treated conditions (gray points). Raw transcript abundances were iteratively rarefied and the median expression of each gene was calculated (~24x coverage). 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. Transcripts for genes that are over-represented in a single condition are placed more proximal to the corner associated with that treatment group. Points placed near the center are equally expressed across all of the conditions measured at 18 hours post-infection. Points are colored based on inclusion in specific carbon metabolic pathways, and point sizes within groups of interest were determined based on the highest expression value for each gene from a single condition. Gene groups from are

shown individually below without abundance information for ease of comparison. Genes included in each group with normalized transcript abundances can be found in Table S1, and refer to Fig. S3 for additional explanation of figure interpretation.

Figure 5 | 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. The complete network contains 447 enzymes and 758 metabolites, with 2135 directed edges. Size of enzyme nodes is relative to the number of normalized reads mapped to the corresponding gene. The sizes shown reflect the transcriptome of *C. difficile* str. 630 during infection of cefoperazone-treated mice after 18 hours of infection. Below the representative network is the metabolite importance algorithm separated into 3 components; (i) relative transcription of reactions consuming a metabolite, (ii) relative transcription of reactions consuming a metabolite, and (iii) difference of consumption and creation of the given metabolite. **(B)** The expanded window displays an example of a single metabolite importance calculation based on local enzyme gene transcription. White values in the red nodes represent the number of normalized transcript reads mapping to the gene sequence for each enzyme node. Average expression of input and output reactions surrounding metabolite **m** are calculated at then the difference of these values found to get the relative importance of **m**. Log_2 transformation is then performed for uniform comparison between metabolites.

Figure 6 | Results from network-based metabolite importance calculation and *in vitro* growth with important carbohydrates. Prior to importance calculation, transcript abundances for each condition were evenly rarefied for even comparison across

colonized environments (~18x coverage). **(A)** Median shared significant metabolites among the 50 highest scoring metabolites from each condition ($P < 0.05$). Median importance scores and pooled random distribution were recalculated per metabolite using the scores from each condition tested. **(B)** Distinctly important significant metabolites from each treatment group ($P < 0.05$). The top 25 scoring metabolites from each group was cross-referenced against each other group resulting in metabolites that are differentially important between environments. **(C)** *in vitro* growth curves validating identified growth nutrients from network analysis. One metabolite that is consistently important to *C. difficile* and at least one metabolite indicated as distinctly important from each group supported growth significantly more ($P < 0.001$) than no carbohydrate control (+ amino acids, gray line). Only those carbon sources that significantly improved *C. difficile* growth over control are displayed (remainder are located in Table S4). Significant differences were calculated using 2-Way ANOVA with Holm-Bonferroni correction.

Figure 7 | Relative *in vivo* concentrations of metabolites highlighted through network-based analysis Select metabolite concentrations using normalized untargeted UPLC-MS quantification from cecal content of *C. difficile* susceptible mice compared to resistant controls. Samples were collected at the same 18 hour timepoint as the analyses performed in infection experiments (n = 9 mice per group). Highlighted carbohydrates indlude: **(A)** N−Acetylglucosamine + N−Acetylgalactosamine, **(B)** Galactitol. **(C)** Mannitol + Sorbitol (These metabolites are quantified simultaneously as they only differ by chirality), **(D)** Salicylate, and **(E)** N−Acetylneuriminate. **(F)** Proline and **(G)** Glycine were also quantified. Paired metabolites were quantified simultaneously as

the only differ by chirality making differentiation impossible. Significant differences between experimental, susceptible groups and resistant SPF mice ($P < 0.05$) are denoted by a single asterisk (*) and were determined using Wilcoxon rank-abundance test with Holm-Bonferroni correction.

Supplementary Figure 1 | Analysis of bacterial community structure resulting from antibiotic treatment. Results from 16S rRNA gene amplicon sequencing from bacterial communities of cecal content in both mock-infected and *C. difficile* 630-infected animals 18 hours post-infection across pretreatment models. **(A)** Non-metric multidimensional scaling (NMDS) ordination based on $Theta_{YC}$ distances for the gut microbiome of all conventionally-raised mice used in these experiments (n = 63). All treatment groups are significantly different from all other groups by AMOVA ($P < 0.001$). **(B)** Inverse Simpson diversity for each cecal community from the mice in (A). Cecal communities from mice not treated with any antibiotics are significantly more diverse than any antibiotic-treated condition by Wilcoxon test ($P < 0.001$).

Supplementary Figure 2 | Expression of specific KEGG gene families. Abundances of normalized transcriptomic reads from *C. difficile* str. 630 in each tested condition. **(A)** All KEGG families. **(B)** Those sub-families within Carbohydrate metabolism. **(C)** Sub-families within Amino acid metabolism.

Supplementary Figure 3 | Additional explanation for Figure 4 interpretation. Relative abundance of transcription for *C. difficile* 630 genes during infection across the 3 antibiotic pretreatment models used during this study. Points that are located closer to a corner are more highly transcribed in the condition associated with that corner compared to the others. As this shows a 3-dimensional data set in 2 dimensions, there

is an amount of distortion proximal to each corner. Simply put for points that are nearer to an edge, a greater percentage of their total transcription was contributed by *C. difficile* colonizing those mice. **(A)** This point represents the transcription for a gene that is overrepresented in cefoperazone-treated mice. **(B)** This point represents a gene in which transcripts are equally detectable in all 3 conditions. **(C)** Transcripts for this gene are only underrepresented in only cefoperazone-treated mice, and are equally detectable in clindamycin and streptomycin-treated animals.

Supplementary Figure 4 | *in vivo* acetate concentrations with GC-MS analysis from cefoperazone-treated mouse cecal content. 2 groups of 5 mice each were pretreated with cefoperazone as outlined Fig. 1b. A single cage was infected with *C. difficile* in the same fashion as other experiments described here. Animals were necropsied at 18 hours post-infection and cecal content was flash frozen for later GC-MS analysis. Significance was determined using Wilcoxon signed-rank test.

Supplementary Figure 5 | Additional growth curves for additional carbon sources and controls. Significant metabolites from network analysis that did not provide improved growth over no carbohydrate (+ amino acids) control. Also included is the negative control of minimal media with no amino acids as well as *C. difficile* growth in standard Brain-Heart Infusion broth.

Supplementary Figure 6 | Example Mont-Carlo simulation results for a single metabolite. Shown here are the distributions resulting from 10000-fold Monte-Carlo simulation of transcript randomization for enzymes interacting with select metabolites. These scores were derived from transcriptome of *C. difficile* colonizing streptomycin-treated mice. The black solid line represents the median of the distribution, the red

dotted lines denote the 95% confidence interval, and the blue arrow indicates the actual measured score for each metabolite. **(A)** Starch (Importance score = 5.492, $P < 0.01$). **(B)** D-Glyceraldehyde 3-phosphate (Importance score = -4.139, P = n.s.).

Supplementary Table 1 | Sets of genes included in Figure 4 with normalized abundances and citations.

Supplementary Table 2 | Topology metrics for enzyme and metabolite nodes in the *C. difficile* str. 630 metabolic network.

Supplementary Table 3 | All metabolites with significant important scores for *C. difficile* in each colonized condition.

Supplementary Table 4 | Growth curve analysis for each tested carbon source.

Supplementary Table 5 | Normalized untargeted metabolomics results for molecules highlighted by network analysis.

References

1. Lessa FC, Gould CV, McDonald LC (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.

2. Lessa FC, et al. (2015) Burden of *Clostridium difficile* Infection in the United States. *The New England journal of medicine* 372(9):825–834.

3. Leffler DA, Lamont JT (2015) *Clostridium difficile* Infection. *New England Journal of Medicine* 372(16):1539–1548.

4. Young VB, Schmidt TM (2004) Antibiotic-Associated Diarrhea Accompanied by Large-Scale Alterations in the Composition of the Fecal Microbiota. *Journal of Clinical Microbiology* 42(3):1203–1206.

5. Lawley TD, Walker AW (2013) Intestinal colonization resistance. *Immunology* 138(1):1–11.

6. Chen X, et al. (2008) A Mouse Model of *Clostridium difficile*-Associated Disease. *Gastroenterology* 135(6):1984–1992.

7. Theriot CM, et al. (2014) Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. doi:10.1038/ncomms4114.Antibiotic-induced.

8. Schubert AM, Sinani H, Schloss PD (2015) Antibiotic-induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against *Clostridium difficile*. *mBio* 6(4). doi:10.1128/mBio.00974-15.

9. Koenigsknecht MJ, et al. (2015) Dynamics and establishment of *Clostridium difficile* infection in the murine gastrointestinal tract. *Infection and Immunity* 83(3):934–941.

10. Antunes LCM, et al. (2011) Effect of antibiotic treatment on the intestinal metabolome. *Antimicrobial Agents and Chemotherapy* 55(4):1494–1503.

11. Ferreyra JA, et al. (2014) Gut microbiota-produced succinate promotes *Clostridium difficile* infection after antibiotic treatment or motility disturbance. *Cell Host and Microbe* 16(6):770–777.

12. Jump RLP, et al. (2014) Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. *PLoS ONE* 9(7). doi:10.1371/journal.pone.0101267.

13. Freter R, Brickner H, Botney M, Cleven D, Aranki A (1983) Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora. *Infection and Immunity* 39(2):676–685.

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

15. Sebaihia M, et al. (2006) The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature genetics* 38(7):779–786.

16. Tracy BP, Jones SW, Fast AG, Indurthi DC, Papoutsakis ET (2012) Clostridia: The importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. 23:364–381.

17. Songer JG, Anderson MA (2006) *Clostridium difficile*: An important pathogen of food animals. 12:1–4.

18. Neumann-Schaal M, Hofmann JD, Will SE, Schomburg D (2015) Time-resolved amino acid uptake of *Clostridium difficile* 630 Delta-erm and concomitant fermentation product and toxin formation. *BMC Microbiology*:281.

19. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM (2016) CodY-dependent regulation of sporulation in *Clostridium difficile*. *Journal of Bacteriology* 198(15):2113–2130.

20. Antunes A, et al. (2012) Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. 40:10701–10718.

21. Bouillaut L, Dubois T, Sonenshein AL, Dupuy B (2015) Integration of metabolism and virulence in *Clostridium difficile*. *Research in Microbiology* 166(4):375–383.

22. Matamouros S, England P, Dupuy B (2007) *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Molecular Microbiology* 64(5):1274–1288.

23. Antunes A, Martin-Verstraete I, Dupuy B (2011) CcpA-mediated repression of *Clostridium difficile* toxin gene expression. *Molecular Microbiology* 79(4):882–899.

24. Janoir C, et al. (2013) Adaptive strategies and pathogenesis of *Clostridium difficile* from *in vivo* transcriptomics. *Infection and Immunity* 81(10):3757–3769.

25. Kansau I, et al. (2016) Deciphering adaptation strategies of the epidemic *Clostridium difficile* 027 strain during infection through in vivo transcriptional analysis. *PLoS ONE* 11(6). doi:10.1371/journal.pone.0158204.

26. Patil KR, 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.

27. Borenstein E, Kupiec M, Feldman MW, Ruppin E (2008) Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proceedings of the National Academy of Sciences* 105(38):14482–14487.

28. Suthers PF, et al. (2009) Genome-scale metabolic reconstruction Of mycoplasma genitalium, iPS189. *PLoS Computational Biology* 5(2). doi:10.1371/journal.pcbi.1000285.

29. Putnam EE, Nock AM, Lawley TD, Shen A (2013) SpoIVA and sipl are *Clostridium difficile* spore morphogenetic proteins. *Journal of Bacteriology* 195(6):1214–1225.

30. Underwood S, et al. (2009) Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *Journal of Bacteriology* 191(23):7296–7305.

31. Fimlaid KA, et al. (2013) Global Analysis of the Sporulation Pathway of *Clostridium difficile*. *PLoS Genetics* 9(8). doi:10.1371/journal.pgen.1003660.

32. Saujet L, Pereira FC, Henriques AO, Martin-Verstraete I (2014) The regulatory network controlling spore formation in *Clostridium difficile*. *FEMS Microbiology Letters* 358(1):1–10.

33. Matsuno K, Sonenshein AL (1999) Role of SpoVG in asymmetric septation in *Bacillus subtilis*. *Journal of Bacteriology* 181(11):3392–3401.

34. Lee ASY, Song KP (2005) LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. *Biochemical and Biophysical Research Communications* 335(3):659–666.

35. Martin MJ, et al. (2013) The *agr* locus regulates virulence and colonization genes in *Clostridium difficile* 027. *Journal of Bacteriology* 195(16):3672–3681.

36. Donnelly ML, Fimlaid KA, Shen A (2016) Characterization of *Clostridium difficile* spores lacking either SpoVAC or DPA Synthetase. *Journal of Bacteriology* (April):JB.00986–15.

37. Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL (2007) Repression of *Clostridium difficile* toxin gene expression by CodY. *Molecular Microbiology* 66(1):206–219.

38. Martin-Verstraete I, Peltier J, Dupuy B (2016) The regulatory networks that control *Clostridium difficile* toxin synthesis. 8. doi:10.3390/toxins8050153.

39. Potapov AP, Voss N, Sasse N, Wingender E (2005) Topology of mammalian transcription networks. *Genome informatics International Conference on Genome Informatics* 16(2):270–278.

40. Ma HW, Zeng AP (2003) The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* 19(11):1423–1430.

41. McGill R, Tukey JW, Larsen Wa (1978) Variations of Box Plots. *The American Statistician* 32(1):12–16.

42. Aboulnaga EH, et al. (2013) Effect of an oxygen-tolerant bifurcating butyryl coenzyme a dehydrogenase/electron-transferring flavoprotein complex from *Clostridium difficile* on butyrate production in *Escherichia coli*. *Journal of Bacteriology* 195(16):3704–3713.

43. Fonknechten N, et al. (2010) *Clostridium sticklandii*, a specialist in amino acid degradation:revisiting its metabolism through its genome sequence. *BMC genomics* 11:555.

44. Jackson S, Calos M, Myers A, Self WT (2006) Analysis of proline reduction in the nosocomial pathogen *Clostridium difficile*. *Journal of Bacteriology* 188(24):8487–8495.

45. Karlsson S, Burman LG, Åkerlund T (2008) Induction of toxins in *Clostridium difficile* is associated with dramatic changes of its metabolism. *Microbiology* 154(11):3430–3436.

46. Ng KM, et al. (2013) Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* 502(7469):96–9.

47. Fuller MF, Reeds PJ (1998) Nitrogen cycling in the gut. *Annual review of nutrition* 18:385–411.

48. Köpke M, Straub M, Dürre P (2013) *Clostridium difficile* Is an Autotrophic Bacterial Pathogen. *PLoS ONE* 8(4). doi:10.1371/journal.pone.0062157.

49. Green ML, Karp PD (2006) The outcomes of pathway database computations depend on pathway ontology. *Nucleic Acids Research* 34(13):3687–3697.

50. Levy R, Borenstein E (2012) Reverse Ecology: from systems to environments and back. *Advances in experimental medicine and biology* 751:329–345.

51. Theriot CM, et al. (2011) Cefoperazone-treated mice as an experimental platform to assess differential virulence of *Clostridium difficile* strains. *Gut microbes* 2(6):326–334.

52. Wilson KH, Kennedy MJ, Fekety FR (1982) Use of sodium taurocholate to enhance spore recovery on a medium selective for *Clostridium difficile*. *Journal of Clinical Microbiology* 15(3):443–446.

53. Sorg Ja, Sonenshein AL (2010) Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. *Journal of Bacteriology* 192(19):4983–4990.

54. Leslie JL, et al. (2015) Persistence and toxin production by *Clostridium difficile* within human intestinal organoids result in disruption of epithelial paracellular barrier function. *Infection and Immunity* 83(1):138–145.

55. Kozich J(of M, Schloss P (2013) 16S Sequencing with the Illumina MiSeq Personal Sequencer. *University of Michigan Health System SOP* 3.1:1–16.

56. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* 73(16):5261–5267.

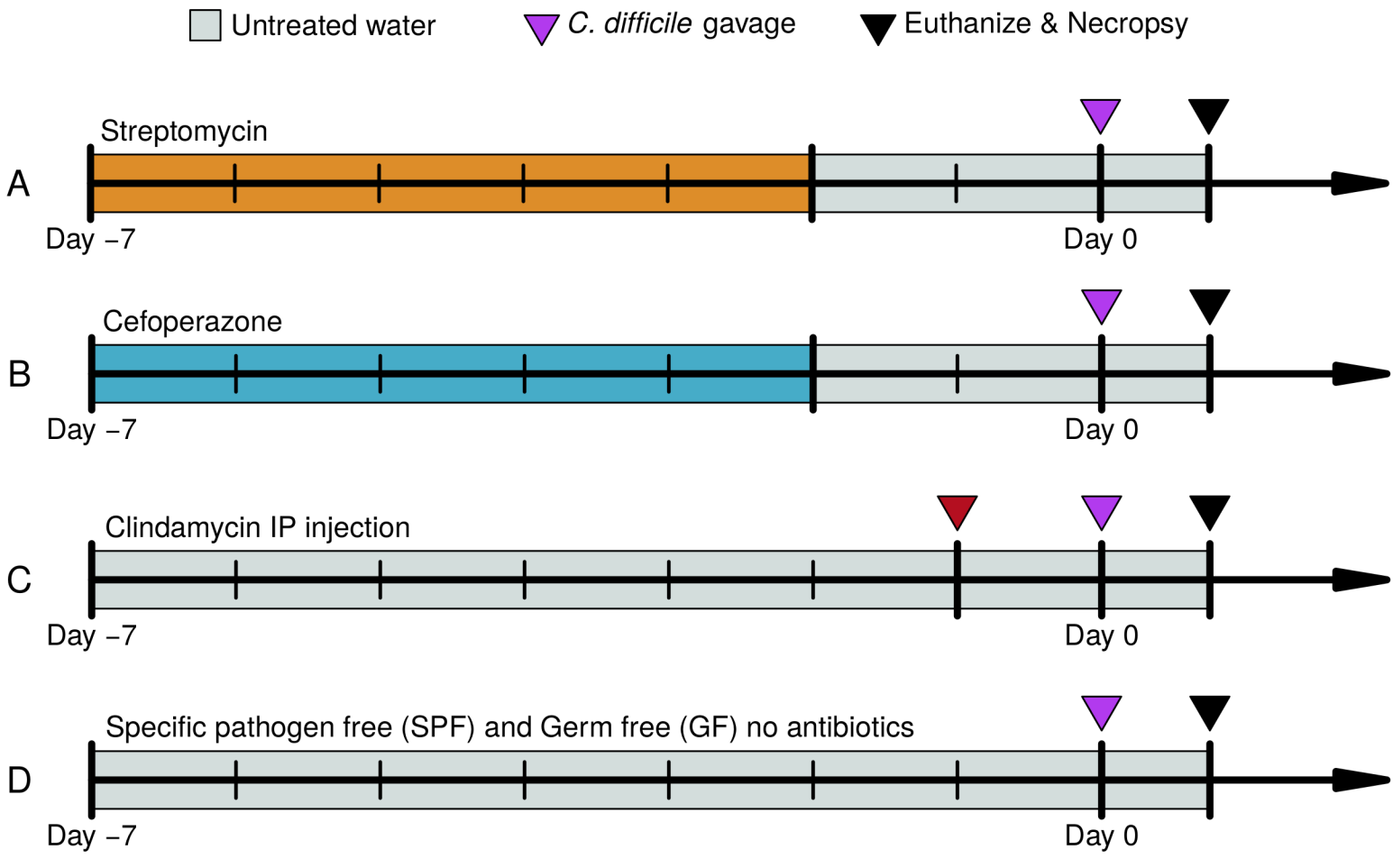
57. Lopez-Medina E, Neubauer MM, Pier GB, Koh AY (2011) RNA isolation of *Pseudomonas aeruginosa* colonizing the murine gastrointestinal tract. *Journal of visualized experiments : JoVE* (55):6–9.

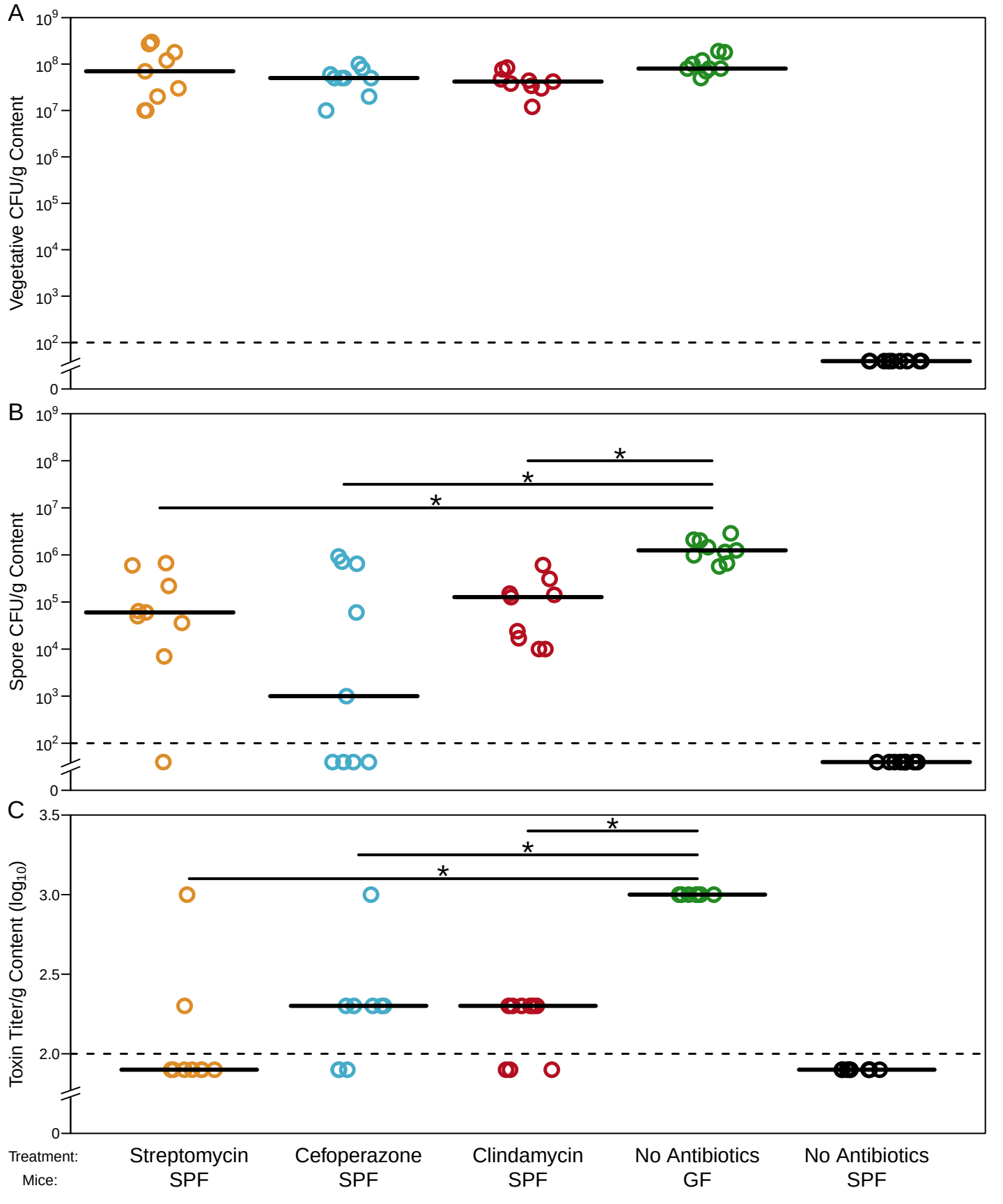
58. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*:1–10.

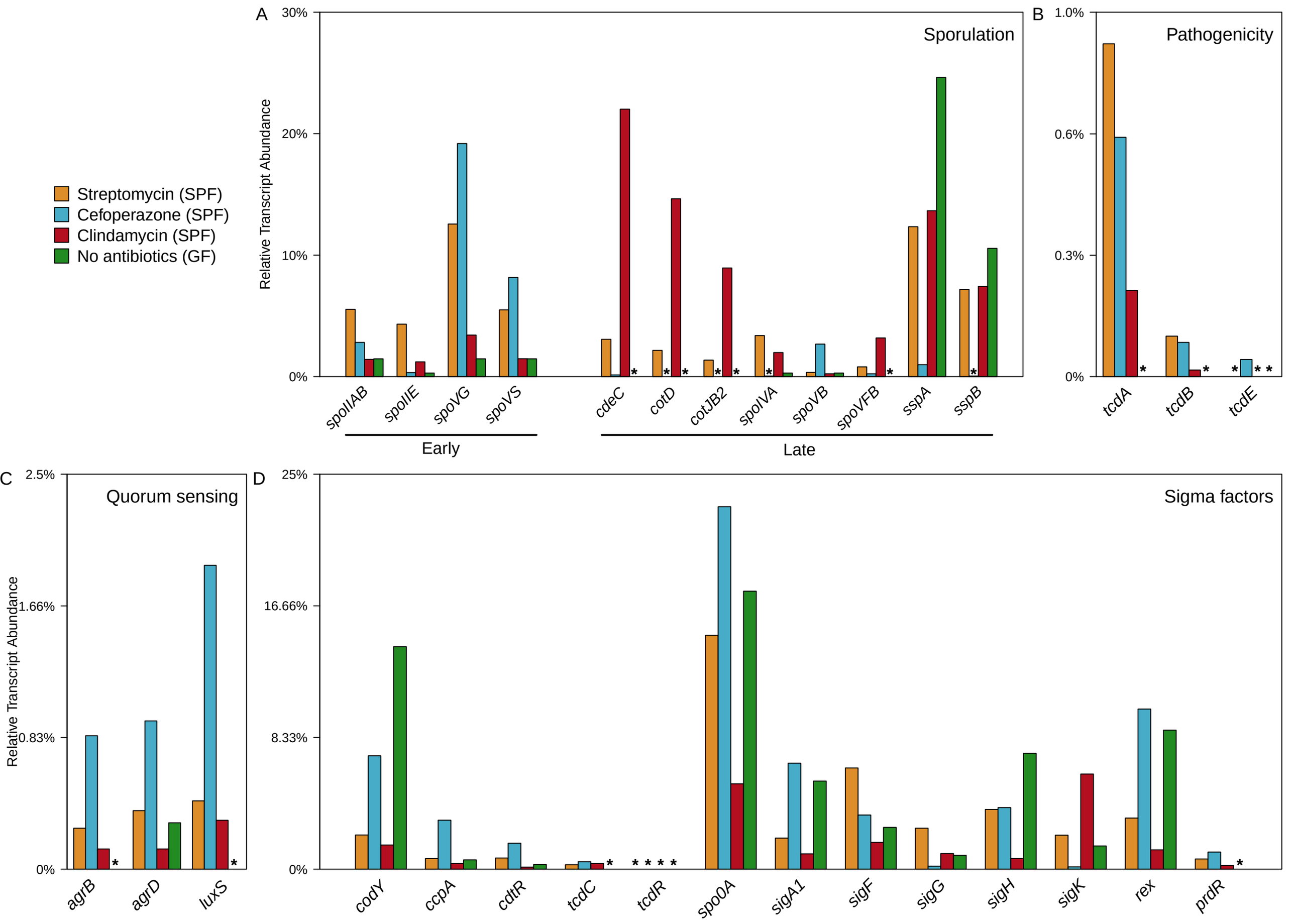
59. Ogata H, et al. (1999) KEGG: Kyoto encyclopedia of genes and genomes. 27:29–34.

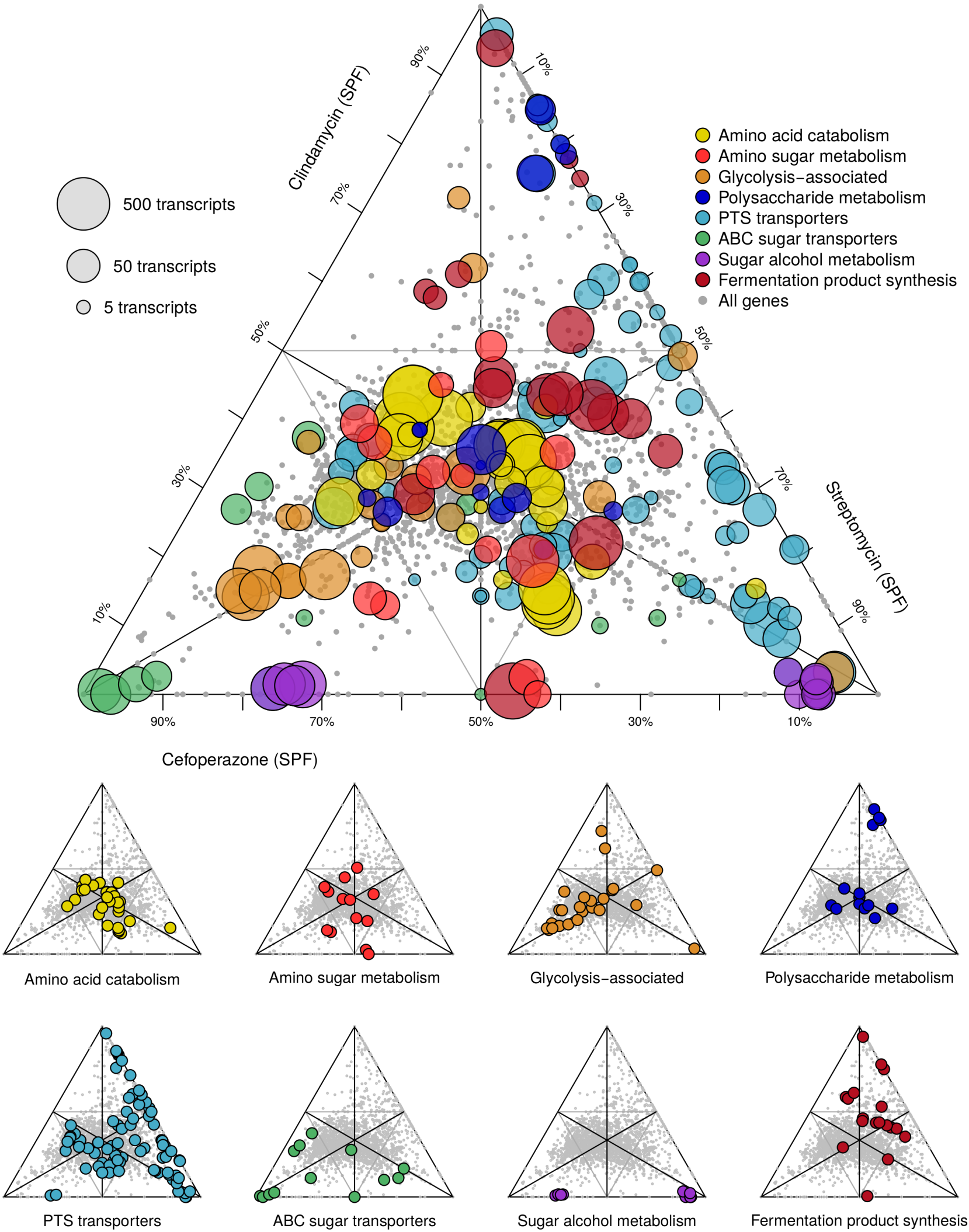
60. Li H, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.

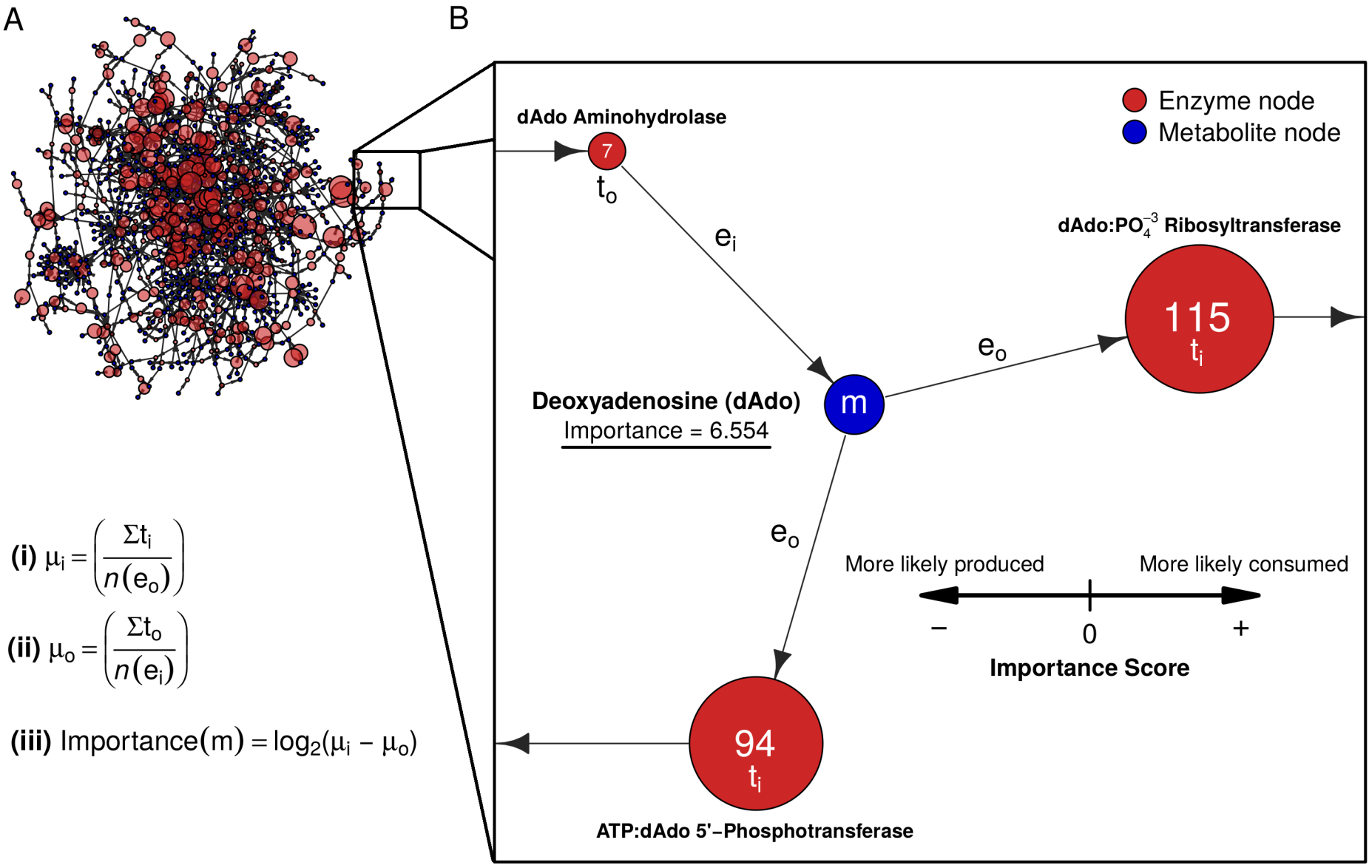
61. Basler G, Ebenhöh O, Selbig J, Nikoloski Z (2011) Mass-balanced randomization of metabolic networks. *Bioinformatics* 27(10):1397–1403.


Untreated water
C. difficile gavage
Euthanize & Necropsy
A
Streptomycin
Day −7
Day 0
B
Cefoperazone
Day −7
Day 0
C
Clindamycin IP injection
Day −7
Day 0
D
Specific pathogen free (SPF) and Germ free (GF) no antibiotics
Day −7
Day 0


A
Vegetative CFU/g Content
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
0
B
Spore CFU/g Content
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
0
*
*
*
C
Toxin Titer/g Content ($\log_{10}$)
3.5
3.0
2.5
2.0
0
*
*
*
Treatment:
Mice:
Streptomycin
SPF
Cefoperazone
SPF
Clindamycin
SPF
No Antibiotics
GF
No Antibiotics
SPF


Streptomycin (SPF)
Cefoperazone (SPF)
Clindamycin (SPF)
No antibiotics (GF)
A
Sporulation
Relative Transcript Abundance
30%
20%
10%
0%
spoIIAB
spoIIE
spoVG
spoVS
cdeC
cotD
cotJB2
spoIVA
spoVB
spoVFB
sspA
sspB
Early
Late
B
Pathogenicity
1.0%
0.6%
0.3%
0%
tcdA
tcdB
tcdE
C
Quorum sensing
2.5%
1.66%
0.83%
0%
agrB
agrD
luxS
D
Sigma factors
25%
16.66%
8.33%
0%
codY
ccpA
cdtR
tcdC
tcdR
spo0A
sigA1
sigF
sigG
sigH
sigK
rex
prdR

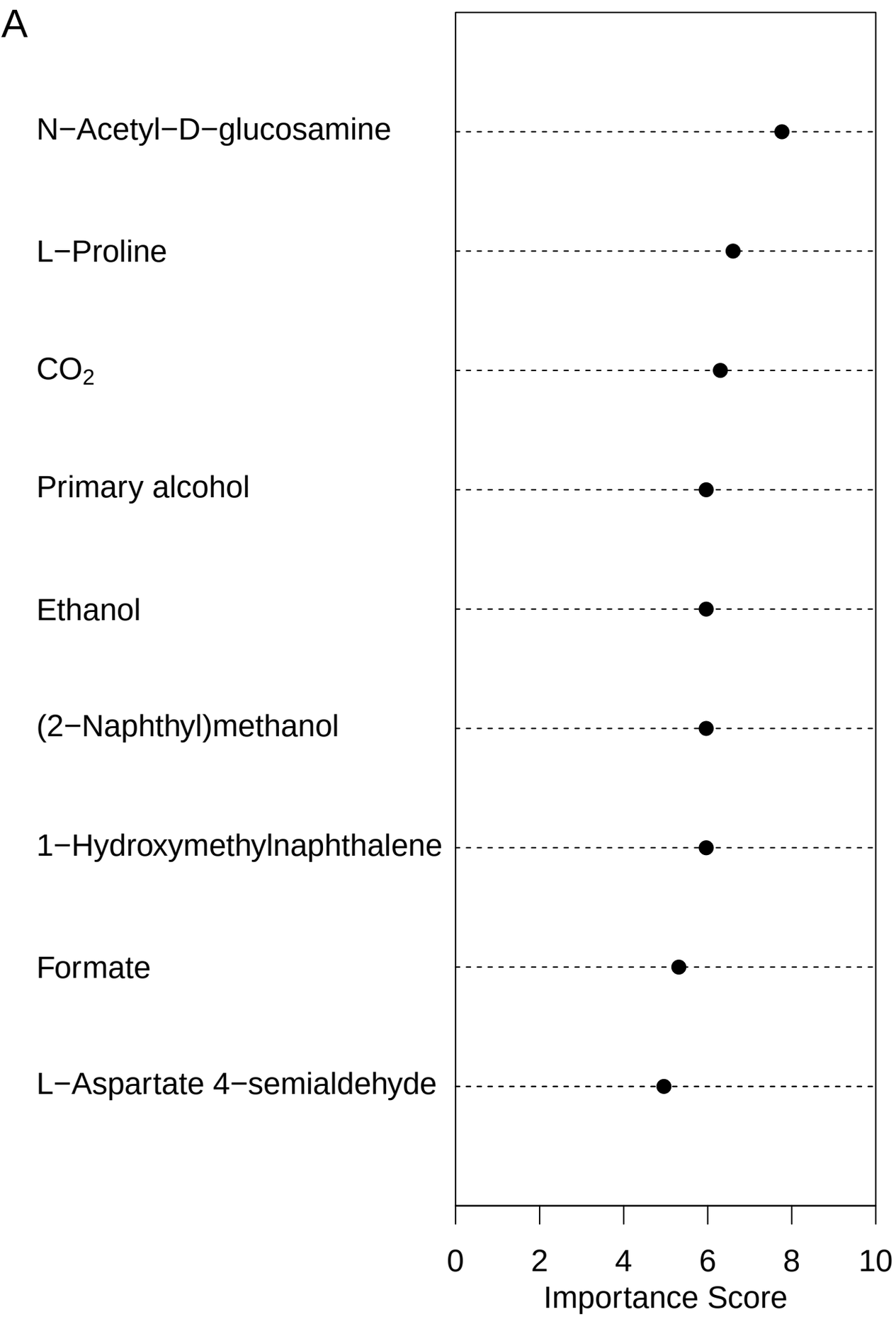

Clindamycin (SPF)
Streptomycin (SPF)
Cefoperazone (SPF)
90%
70%
50%
30%
10%
500 transcripts
50 transcripts
5 transcripts
Amino acid catabolism
Amino sugar metabolism
Glycolysis−associated
Polysaccharide metabolism
PTS transporters
ABC sugar transporters
Sugar alcohol metabolism
Fermentation product synthesis
All genes
Amino acid catabolism
Amino sugar metabolism
Glycolysis−associated
Polysaccharide metabolism
PTS transporters
ABC sugar transporters
Sugar alcohol metabolism
Fermentation product synthesis

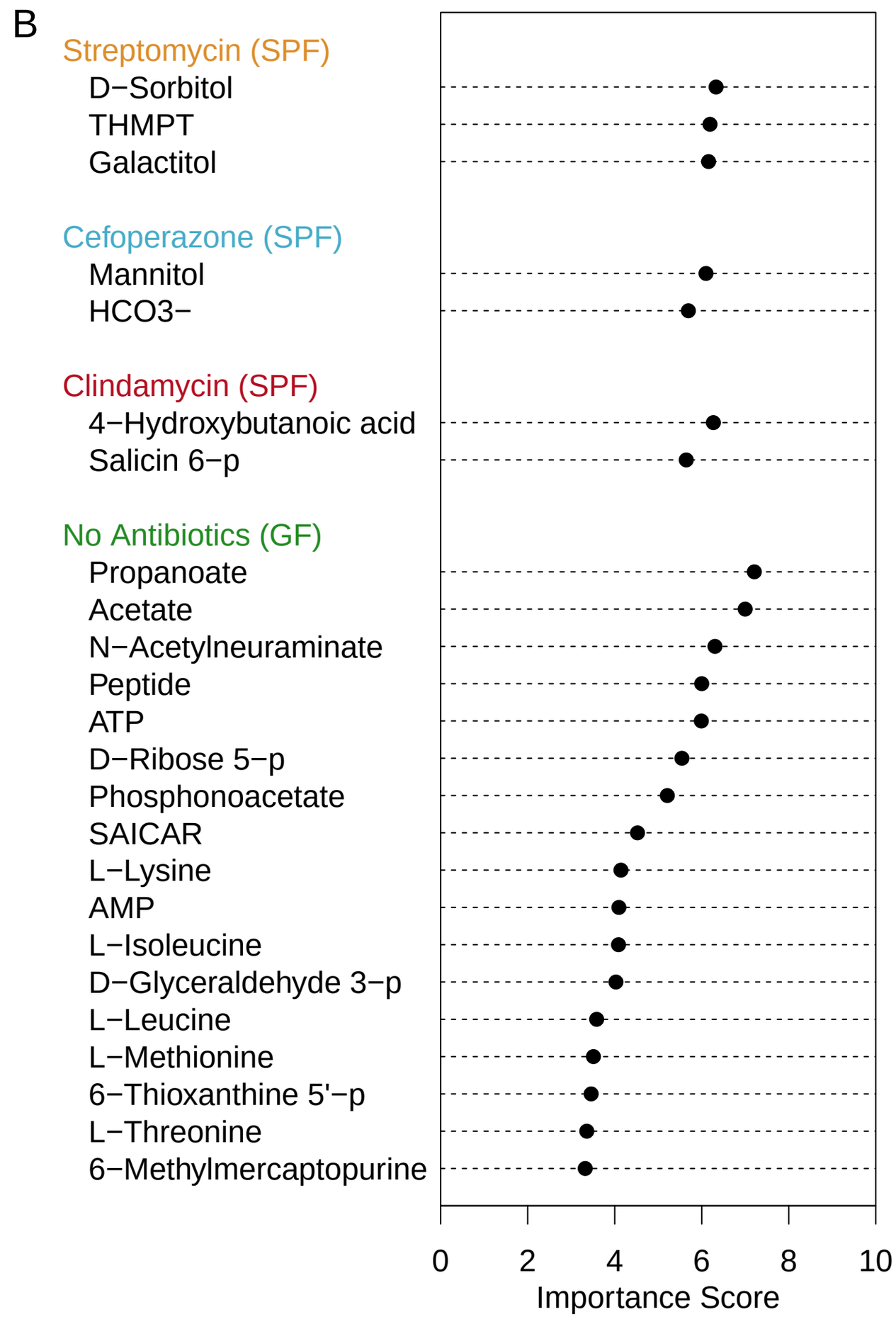


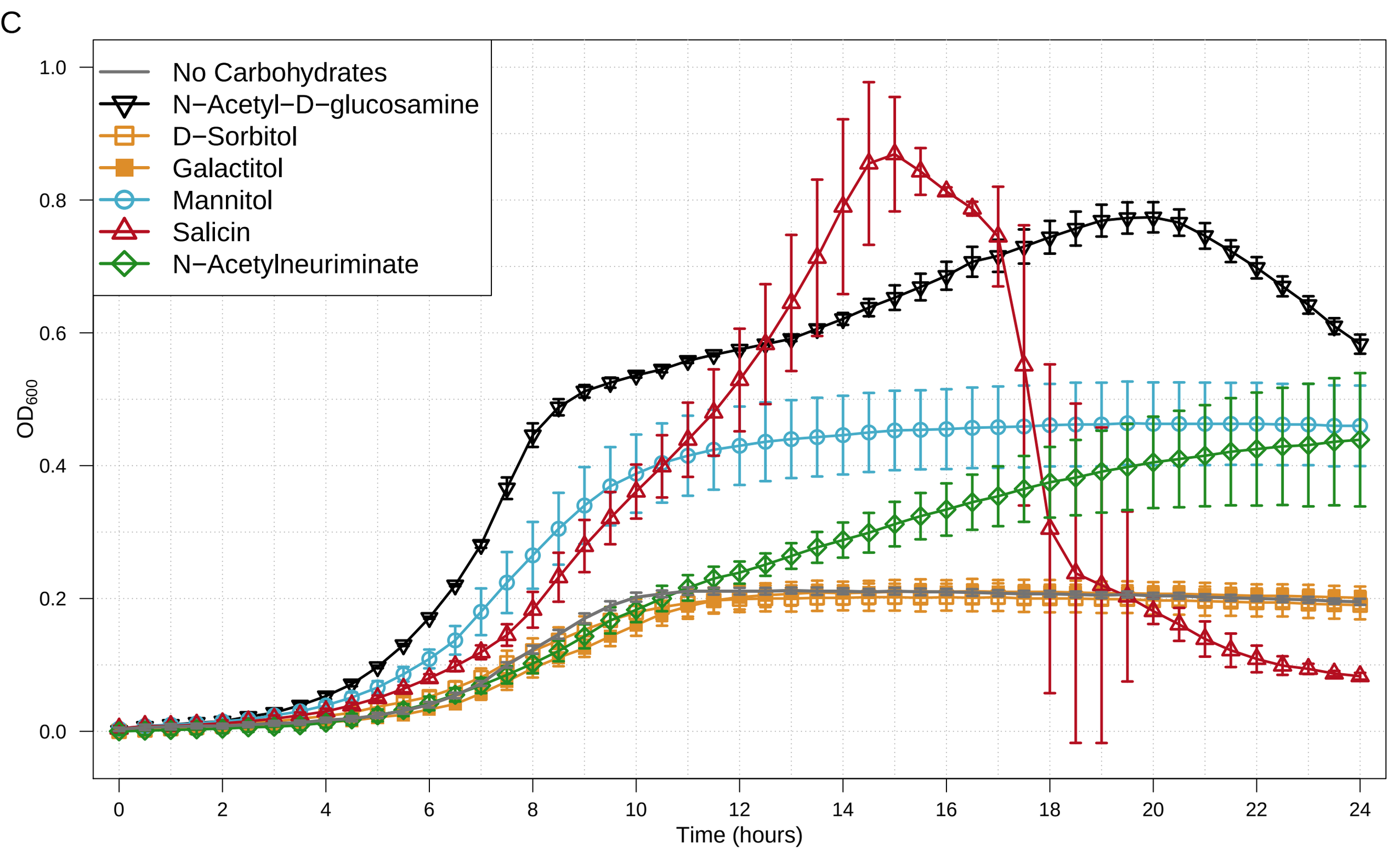
(i) $\mu_i = \left(\frac{\Sigma t_i}{n(e_o)}\right)$

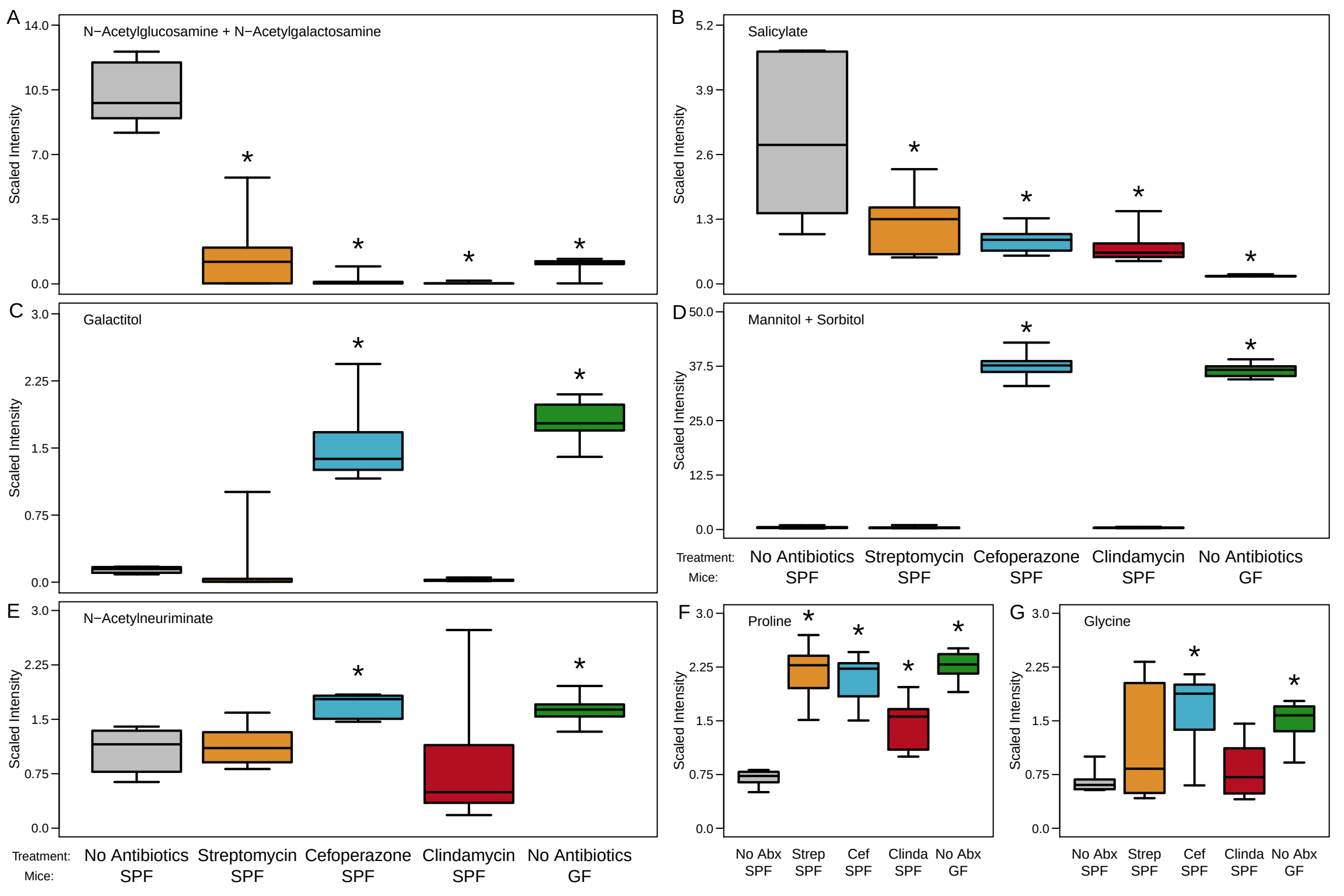
(ii) $\mu_o = \left(\frac{\Sigma t_o}{n(e_i)}\right)$

(iii) $\text{Importance}(m) = \log_2(\mu_i - \mu_o)$


A
N−Acetyl−D−glucosamine
L−Proline
CO_2
Primary alcohol
Ethanol
(2−Naphthyl)methanol
1−Hydroxymethylnaphthalene
Formate
L−Aspartate 4−semialdehyde
0
2
4
6
8
10
Importance Score


B
Streptomycin (SPF)
D−Sorbitol
THMPT
Galactitol
Cefoperazone (SPF)
Mannitol
HCO3−
Clindamycin (SPF)
4−Hydroxybutanoic acid
Salicin 6−p
No Antibiotics (GF)
Propanoate
Acetate
N−Acetylneuraminate
Peptide
ATP
D−Ribose 5−p
Phosphonoacetate
SAICAR
L−Lysine
AMP
L−Isoleucine
D−Glyceraldehyde 3−p
L−Leucine
L−Methionine
6−Thioxanthine 5'−p
L−Threonine
6−Methylmercaptopurine
0
2
4
6
8
10
Importance Score


C
No Carbohydrates
N−Acetyl−D−glucosamine
D−Sorbitol
Galactitol
Mannitol
Salicin
N−Acetylneuriminate
OD_{600}
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
14
16
18
20
22
24
Time (hours)


A
N−Acetylglucosamine + N−Acetylgalactosamine
Scaled Intensity
14.0
10.5
7.0
3.5
0.0
B
Salicylate
5.2
3.9
2.6
1.3
0.0
C
Galactitol
3.0
2.25
1.5
0.75
0.0
D
Mannitol + Sorbitol
50.0
37.5
25.0
12.5
0.0
E
N−Acetylneuriminate
F
Proline
G
Glycine
Treatment:
Mice:
No Antibiotics SPF
Streptomycin SPF
Cefoperazone SPF
Clindamycin SPF
No Antibiotics GF
No Abx SPF
Strep SPF
Cef SPF
Clinda SPF
No Abx GF