# Microbiota predict *Clostridium difficile* severity in germ-free mice colonized with human feces

Running title: Microbiota predict *C. difficile* severity in humanized mice

Kaitlin J. Flynn[1+], Nicholas Lesniak[1+], Alyxandria M. Schubert[2], Hamide Sinani[?], and Patrick D. Schloss[1†]

† Corresponding author: pschloss@umich.edu

+ These authors contributed equally to this work

1. Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan 48109

2. Food and Drug Administration?

3. Department for Hamide?

## Abstract

*Clostridium difficile* causes diarrheal disease when it successfully colonizes a disrupted gut microbial community. Current mouse models to study *C. difficile* infection (CDI) rely on pre-treatment with antibiotics to alter the mouse microbiome prior to inoculation. This model does not allow for analysis of human-associated microbial community members that modulate *C. difficile* colonization and expansion. To study human-associated microbes in the context of CDI, we inoculated germ-free C57BL/6 mice with one of 16 human fecal samples from diarrheal or healthy donors and challenged with *C. difficile* 14 days later. Five unique donor-mice combinations resulted in severe CDI while the remaining 11 only experienced mild disease. Both healthy and diarrheal donors were susceptible to colonization and severe symptoms of CDI. To determine if specific microbes were associated with disease severity outcomes, we built a classification Random Forest machine learning model based on relative abundance data of the communities prior to infection. The model identified a number of bacterial populations associated with the development of severe CDI, including *Bacilliales, Ruminococcaceae, Ruminococcus, Staphylococcus, Streptococcus and Bacteriodetes*. Additionally, a regression model accurately predicted colonization levels of *C. difficile* at one to ten days post-infection. This model explained 99% of the variance in the number of CFU isolated from mouse stool. Members of *Lachnospiraceae, Parabacteroides, Bacteroidales, Bacteroidetes, Porphyromonadaceae* and unclassified *Bacteria* families were predictive of future *C. difficile* colonization levels. Finally, challenging these mice with different strains of *C. difficile* revealed that susceptible human-associated microbial communities were prone to severe disease independent of strain type. Taken together these results suggest that human-associated microbial communities can be recapitulated in germ-free mice and used to characterize dynamics of CDI. Because both healthy and diarrheal patients were susceptible to severe CDI, machine-learning models are useful to identify bacterial populations that allow colonization and contribute to the development of *C. difficile* associated disease in humans.

## Introduction

*Clostridium difficile* is an opportunistic pathogen of the human lower gastrointestinal tract. *C. difficile* forms spores that can persist on abiotic surfaces and are not readily killed by ethanol-based hand-sanitizers, putting hospital patients particularly at risk. Indeed, ~12% of hospital acquired infections in the United States are due to *C. difficile* and result in up to 15,000 deaths annually [1]. Disruption of the native microbial community is the most common risk factor for development of *C. difficile* infection (CDI) [2]. Antibiotic use and inflammatory bowel diseases are associated with loss of colonization resistance to *C. difficile* through the loss of potentially protective bacterial families such as *Barnesiella* and *Lachnospiraceae* in both mouse models and human association experiments [3,4]. The composition of the community is clearly important for the acquisition, resistance and treatment of *C. difficile* infection, as giving patients a healthy fecal microbiome transplant is the most effective treatment for this disease (cite Anna 2014). The precise mechanisms of colonization resistance and restoration of a healthy community are yet to be discovered.

Murine models to study CDI typically rely on treating conventionally-raised mice with antibiotics either in drinking water or by injection to induce susceptibility (3, 4). This model provides a convenient way to study *C. difficile* pathogenesis and virulence factors. Numerous microbiome studies have been performed using this model to determine the antibiotic classes (5), starting microbial community (6) and metabolites (7) that impact development and severity of *C. difficile* infection. While informative, these studies are somewhat removed from human disease because they only examine mouse-associated microbial communities.

Gnotobiotic or germ-free mouse models have been used for a range of studies of CDI, including assessment of species-specific interactions between *C. difficile* and competing microbial community members (8), analysis of nutrient restriction (9), *in vivo* transcriptomics of *C. difficile* and examination of host immune response to CDI (10). Further, CDI therapeutics such as antibiotics and fecal microbiota transplants have been tested extensively in a gnotobiotic-piglet or piglet-to-gnotobiotic-mouse model of disease (11), (12). Pigs have a longer digestive tract with components more similar to humans than mice and are typically infected by strains typical in human infection (13). However, the murine and porcine microbiomes typically do not resemble

those of the human gut.

The power of the gnotobiotic models to study CDI has been further realized by first inoculating germ-free mice and piglets with human stool microbes. In one study, germ-free piglets were acutely colonized with human feces for one week and then treated with tigecycline. After challenge with *C. difficile* none of the antibiotic-treated piglets succumbed to infection, while some of the untreated human-colonized pigs did (11). In another study, germ-free mice colonized with human feces were bred over several generations to create a cohort of mice with identical human-derived microbiomes (14). These mice were subsequently treated with a five-antibiotic cocktail to induce dysbiosis and then were successfully colonized by *C. difficile* (14). While informative, these studies were limited in their use of only one human donor as input inoculum.

Here we designed a laboratory system that allows for studying and modeling microbial coumminity interactions with *C. difficile* infection in mice with human derived microbiomes. To test the impact of individual human microbiomes on CDI, we colonized germ-free mice with 16 different human stool donors. We then characterized human-associated microbiome response to *C. difficile* challenge. Further, we built a predictive model based on bacterial composition that classified "at-risk" microbiomes prior to infection with *C. difficile*. These findings show that human-associated microbiomes can be at risk for CDI even in the absence of antibiotic-induced dysbiosis. Finally, our results indicate that study of mice colonized with human feces can replicate a range of clinical outcomes.

## Results

**Germ-free mice inoculated with human feces as model for C. difficile infection.** To generate mice with human-derived microbiomes, we inoculated one cage each of gnotobiotic C57/BL6 mice with one of 16 different human fecal donors. Five donors were patients that had diarrhea that was not attributable to *C. difficile* infection while 11 donors were healthy at time of donation. Stool from a patient that was colonized with virulent *C. difficile* was used as a positive control. After inoculation with human stool, mice microbiomes were allowed to equilibrate for 14 days. Prior to

infection, stool samples were taken from each mouse to establish baseline microbiome composition. Then, 100 spores of *C. difficile* strain 430 (isolated from the positive control patient) were used to orally infect each mouse. Mice were monitored for weight loss and clinical signs of disease. Fecal samples were taken to enumerate *C. difficile* CFU and for microbiome analysis each day up to 10 days post-infection (Fig 1A). To ensure that the donors we selected represented a diverse array of human microbiomes, we sequenced the 16S rRNA genes from donor fecal inocula. Ordination of the distances between donor communities showed that the donors each had distinctly different communities, independent of whether the sample came from a sick or healthy person (Fig 1B). Likewise, the microbial communities of the mice on day 0 were characterized by sequencing of fecal pellet DNA prior to infection (Table S1). Ordination of day 0 communities shows that mice were similar to each other within each cage and donor, but distinct from other donors (Fig 1C). This result confirmed that human-associated microbes were able to colonize gnotobiotic mice and provide various initial communities to test *C. difficile* dynamics.

**C. difficile infection in mice with human-derived microbiota cause a range of outcomes.** *C. difficile* colonization was monitored by daily plating of stool pellets for *C. difficile* CFU. Nearly all of the mice were colonized to $10^5 - 10^7$ CFU by one day post-infection and remained colonized at that level until the end of the experiment (Fig 2A). As one indicator of disease, mouse weights were taken each day post-infection and weight-loss was monitored alongside clinical signs. When mice were judged to be too ill to continue they were humanely euthanized. Overall, disease phenotypes fell into two classes. Mice that became severely ill and lost 20% or more of their starting body weight within one to two days post-infection were classified as “severe” whereas mice that were colonized with *C. difficile* but did not show signs of disease or severe weight loss were considered to have “mild” disease (Fig 2A, 2B). Interestingly, *C. difficile* was able to cause severe disease in both mice that had been colonized with healthy stool and those colonized with diarrheal stool, suggesting susceptibility to CDI is dependent on the composition of the starting microbiome and not associated with donor clinical metadata.

**Nick will be updating the next two sections Microbes present in the gut prior to infection are predictive of *C. difficile* colonization levels** Previous work in our group has demonstrated that the microbes present in the mouse microbiome prior to *C. difficile* challenge can predict future

colonization levels (cite Alyx). To determine if the human-associated microbiome in the mouse could similarly predict *C. difficile* colonization we employed a Random Forest machine learning algorithm similar to the one used in previous studies. Relative abundance of OTUs with greater than 1% abundance at day 0 were used as input to the model. We found that the model explained X% of the variance in colonization levels of *C. difficile* (Fig 3A). By refining the model to include only the top X predictive OTUs, the model explained X% of the variance. Partial dependency plots of the top X OTUs predictive of *C. difficile* colonization reveal that colonization is positively associated with *Bacteriodetes*, *Parabacteriodes* and *Porphyromonadaceae* (really?) (Fig 3B). Colonization was negatively associated with the presence of *Lachnospiraceae* (Fig 3B) though none of the human-associated mouse communities were completely resistant to colonization by *C. difficile* (Fig 2A, 3B). These results confirm that as in mouse microbiome studies, the human-associated gut community prior to challenge is predictive of *C. difficile* colonization levels.

**The microbiome predicts *C. difficile* infection severity.** As *C. difficile* colonization levels can be predicted by the microbiome composition on day 0, we postulated that severity of disease could also be predicted from day 0. Using weight loss as a proxy for severity, we built a Random Forest classification model based on relative abundance to determine which community members were associated with either mild or severe disease. We optimized the model using the AUC-RF algorithm to generate a Reciever Operating Curve (ROC) with maximal Area Under the Curve (AUC) (cite Calle 2011), as used previously in our group to classify colon cancer outcomes (cite Niel 2015 gen med). The optimized model used X number of OTUs and identified several that predicted increased weightloss and subsequent disease severity (Fig 4A, 4B). The populations that were most commonly found in mice that succumbed to *C. difficile* infection were members of the *Clostridiales*, *Bacilliales*, *Streptococcus* (Lactobacilles?) and *Bacteriodes*. (need more OTUs here, or number that were used for model and then group by rank). The AUC was calculated using the Out-Of-Bag (OOB) error for each sample (Fig 4A). Cross-validation using 10-fold, leave-one-mouse-out and leave-one-cage-out resulted in AUCs not significantly different from the OOB (Fig 4A). This model reveals that the levels of X OTUs on day 0 can predict disease outcome in mice with human-associated microbiomes. (probably also going to want a supplemental table for predictive OTUs for this and Fig 3)

**Susceptibility to infection is community-dependent.** *C. difficile* strains are known to vary

in disease phenotypes based both on pathogenicity factors of the strain and host microbiome composition (cite someone). To test the susceptibility of human-derived microbiomes to *C. difficile* infection, two donor-mouse microbiomes that were previously shown to be resistant to *C. difficile* strain 430 induced weight loss (Fig 2B) (DA00396, DA00430-wait isnt this a control) and one that was susceptible to severe disease (DA00578) were independently challenged with one of three unique *C. difficile* strains (*table or reference for toxin prod*, etc). As done previously, *C. difficile* stool CFU was enumerated daily and mouse weight-loss tracked. Infection of resistant mice with virulent strain 458 or the weaker strain 299 resulted in colonization to levels near that of 431 (Fig 5A). Accordingly, these mice did not experience severe weight loss or show clinical signs of disease (Fig 5B). In contrast, infection of susceptible mice (DA00578, Fig 1B) with pathogenic strain 395 resulted in colonization and severe weight loss by two days post-infection (Fig 5A, 5B). Strain 395 is similarly virulent as strain 431, and shows a similar infection phenotype in these mice. These preliminary results suggest that human-associated microbiomes may be more predictive of disease severity than the *C. difficile* strain.

## Discussion

Study of human-associated microbiomes and their interactions with *C. difficile* has been limited by the lack of an appropriate model and available samples. Here we colonized 16 groups of mice with human donor stool from patients with different clinical profiles or healthy volunteers. This created a diverse set of human-associated mouse microbiomes that allowed for testing *C. difficile* infection dynamics and community interactions. Our results showed that both healthy and diarrheal-associated human communities were susceptible to severe disease. Additionally, our results confirmed previous findings that resistance to *C. difficile* disease is dependent on the entire microbial community and not just one or two taxa (cite alyx). Further, we were able to learn classification models that identified specific OTUs within susceptible communities that predispose mice to infection. By subsequently infecting susceptible or resistant communities with varying strains of *C. difficile* we confirmed that the initial state of the human-associated microbiome is more predictive of disease severity than the strain used.

Here we colonized germ-free mice with a variety of human stool communities and allowed these communities to colonize the mice for 2 weeks, prior to *C. difficile* challenge. It is important to note that the mouse gut environment is different from that of the human, and as such exerts a selective pressure on the human stool microbial community that stabilizes over the 2 weeks. Thus these mice cannot be considered "humanized" but rather colonized by human-associated microbes. Interestingly, clinical status was not associated with proclivity to severe *C. difficile* infection as some healthy donor mice were susceptible. A surprising finding was that none of the human-associated mouse communities were completely resistant to colonization by *C. difficile*. This is in contrast to numerous reports of mouse microbiome studies (cite) and the knowledge that in general healthy humans without gut perturbations are resistant to *C. difficile* infection (cite). On the other hand, it is known that humans, especially health care employees, can be asymptomatic carriers of *C. difficile* (Furyura-Kanamari). This may explain the fact that the majority of our groups of mice were colonized by *C. difficile* but did not display outward signs of illness or infection-associated weight loss (Fig 2).

• Discuss prediction methods and outcomes - Nick - we have bugs that predict severity, but do we have bugs that predict mild/asymptomatic? probably, could discuss here • Discuss potential mechanisms for interesting OTUs - Nick

Finally, our pilot experiments testing the efficiency of infection with different *C. difficile* strains have allowed for testing community versus pathogen factors of infection. Our results are consistent with previous reports that the composition of the gut microbiome determines if an individual is successfully infected or not (cite Alyx and others). *C. difficile* strains are known to vary in pathogenicity, primarily via the amount of toxin that is produced during infection (cite). To confirm that disease risk is dependent on the microbiome, we challenged susceptibile mice with an additional pathogenic strain of *C. difficile* and found that the mice indeed fell ill. In contrast, we exposed an asymptomatic community to a relatively weak strain of *C. difficile* and found that the community could be colonized by this one as well– but not cause disease. This suggests that this particular colony is prone to colonization by any *C. difficile* strain. Further experiments will test if infecting susceptible communities with weak strains allows the mice to survive. It has been recently reported that carriage of a non-toxigenic strain of *C. difficile* can prevent additional colonization by

a pathogenic strain, suggesting a role of niche-filling or immunity this pre-colonization provides (cite cite). One interesting experiment would be to recapitulate this finding in human-associated mouse microbiomes by making a susceptible community resistant to infection. Further work with human-associated communities in mice and mathematical modeling will allow for uncovering of complex human-associated disease processes and dynamics as they relate to human *C. difficile* infection.

## Materials and Methods

**Mice** The experiments in this study used 6-8 week-old male and female gnotobiotic C57/BL6 mice. The mice were born and maintainted in a completely sterile environment in the Germ Free Mouse Facility at the University of Michigan. The University Committee on Use and Care of Animals approved all protocols and experiments conducted in this study.

**Human samples** Donor stool samples were chosen from a sample bank collected from patients from October 2010 to November 2012 at the University of Michigan. Sample collection methods and study was approved by the Institutional Review Board at the University of Michigan. Stool samples from University of Michigan Health System patients were submitted to the microbiology lab for *C. difficile* testing. *C. difficile* presence was measured using WHATEVER- Cdiff quik check? toxin AB?. Samples from these patients were stored in Cary-Blair medium at -80C. All patients consented to participation in the study. Donor stool for this study was selected from one patient that tested positive for *C. difficile* infection and five patients that presented with diarrhea and tested negative for *C. difficile*. Healthy donor stool was collected from adult volunteers living in the Ann Arbor, MI area. 100ul of stool slurry was gavaged into each mouse.

***Clostridium difficile* and sample processing** For all experiments, 100 spores of *Clostridium difficile* were used to orally infect mice as described previously (cite alyx). The *C. difficile* strains used were acquired from spore stocks of (WHERE) and comprised of strains 431, 395, 299 and 468 (cite?). After gavage, the remaining inoculum was diluted and plated to confirm infection dose. Bacteria were cultured anaerobically on TCCFA plates and incubated at 37C (cite plate recipe).

During infection, two fresh stool pellets were collected from each mouse daily. One pellet was resuspended in anaerobic phosphate-buffered saline and diluted and plated on TCCFA plates to enumerate *C. difficile* CFU. The other pellet was frozen at -80C until the end of the experiment.

**DNA isolation and sequencing** Bacterial DNA was isolated from banked donor stool and each mouse stool pellet using the MO-BIO PowerSoil DNA isolation kit. This DNA was used as template for amplification of the V4 region of the 16S rRNA gene and fragments were sequenced on an Illumina MiSeq as previously described (cite Jim paper).

• Sequence curation/mothur, data analysis, code + availability of data(SRA) and code (Github) - Nick

## Acknowledgments

We would like to thank members of the Schloss lab and Jhansi Leslie for analysis advice. We would also like to thank the Microbial Systems Laboratory at the University of Michigan for sequencing assistance.

## Figure Legends

**Figure 1. Germ-free mice inoculated with human feces as a model for *C. difficile infection.*** A) Experimental design. Stool from 16 healthy, diarrheal and CDI patients were independently inoculated into 3-4 germ-free mice by oral gavage. 14 days later mice were orally infected with 100 spores of C. difficile strain 431. Weight and stool CFU were monitored for up to 10 days post infection. B) NMDS ordination of donor stool communities prior to inoculating mice. Each point represents one donor and are colored by clinical diagnosis. C) NDMS ordination of the stool communities on day 0. Each symbol represents one mouse and is colored by donor. Circles represent mice that experienced mild disease and triangles represent those that suffered severe disease.

**Figure 2. *C. difficile* infection results in mild or severe disease.** A) *C. difficile CFU* was enumerated by plating of mouse stool pellets daily. Each point represents a mouse and the lines represent the median CFU in each group. Error bars are interquartile ranges. Red lines and points correspond to mice that succumbed to severe disease whereas black lines and points correspond to mice that had mild or no disease. B) Mouse weights were recorded and daily percent weight loss calculated for each mouse. Data is presented as the median of each group and interquartile ranges. Mice that succumbed to severe infection typically lost a significant amount of weight by day 1 or 2 post infection. Red lines correspond to severely ill mice, black to mice with mild disease.

**Figure 3. Random Forest prediction of *C. difficile* colonization level.** A) OTUs above 1% relative abundance on day 0 were used to predict median $\log_{10}$ CFU of *C. difficile* after colonization. OTUs were chosen such that they were not predictive of cage or donor. Each point is a mouse colored by cage. B) Partial dependency plots of the top six predictive OTUs. Line displats the partial dependence of $\log_{10}$ CFU on the relative abundance of each predictive OUT. Each median $\log_{10}$ CFU is plotted against its relative abundance for each predictive OTU.

**Figure 4. Random Forest prediction of CDI severity.** OTUs above 1% relative abundance on day 0 were used to predict disease severity. OTUs were chosen such that they were not predictive of cage or donor. Predictive classification tested via 10-fold (gray), leave-one-cage-out (purple

dashed) or leave-one-mouse-out (blue dashed) models are displayed in (A). B) Partial dependency plots of most predictive OTUs. Line displays the partial dependence of $\log_{10}$ CFU on OTU relative abundance. Points are the OTU relative abundance of each mouse colored by outcome (red, severe, black, mild).

**Figure 5. Infection of mice with different *C. difficile strains*.** 3 strains of *C. difficile* were used to infect mice colonized with susceptible (DA00578) or resistant (DA00369, DA00430) human donor stool. A) *C. difficile* stool CFU was enumerated over 10 days. B) Percent weight loss was calculated each day for each mouse. In both plots, each mouse is a point and lines represent the mean of each cage.

**Supplement Table S1: Mouse day 0 communities by donor genera (avg + stdev of cage)**

## References


1. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden ofClostridium difficileInfection in the united states. New England Journal of Medicine. New England Journal of Medicine (NEJM/MMS); 2015;372: 825–834. doi:10.1056/nejmoa1408913

2. Britton RA, Young VB. Interaction between the intestinal microbiota and host in clostridium difficile colonization resistance. Trends in Microbiology. Elsevier BV; 2012;20: 313–319. doi:10.1016/j.tim.2012.04.001

3. Schubert AM, Rogers MAM, Ring C, Mogle J, Petrosino JP, Young VB, et al. Microbiome data distinguish patients with clostridium difficile infection and non-c. difficile-associated diarrhea from healthy controls. mBio. American Society for Microbiology; 2014;5: e01021–14–e01021–14. doi:10.1128/mbio.01021-14

4. Vincent C, Stephens DA, Loo VG, Edens TJ, Behr MA, Dewar K, et al. Reductions in intestinal clostridiales precede the development of nosocomial clostridium difficile infection. Microbiome. Springer Nature; 2013;1: 18. doi:10.1186/2049-2618-1-18