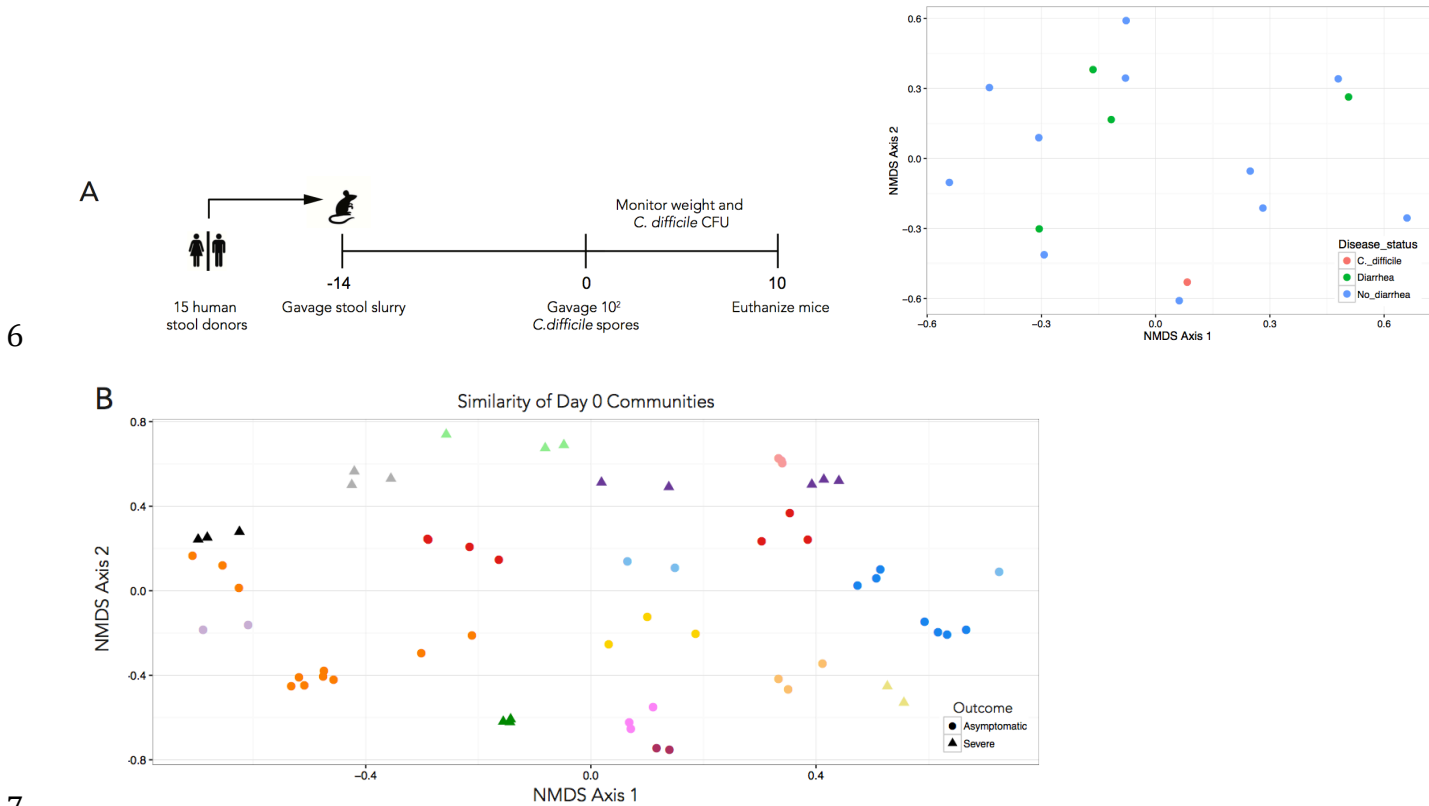


1 **Microbiota predict *Clostridium difficile* severity in germ-free mice colonized with human feces**  
2 Kaitlin J Flynn\*, Nicholas Lesniak\*, Alyxandria M. Schubert, Hamide Sinani, anyone-else-ERIN-people?,  
3 maybe-Ingrid Bergin?, Patrick D. Schloss

4 **Figures**

5 **Figure 1: Germ-free mice inoculated with human feces as a model for *C. difficile* infection**



7  
8 Figure 1. Germ-free mice inoculated with human feces as a model for *C. difficile* infection. A) Stool was  
9 collected from 16 healthy, diarrheal and CDI patients and inoculated into 3-4 germ-free mice per donor  
0 by oral gavage. After allowing the community to stabilize for 14 days, mice were orally gavaged with 100  
1 spores of *C. difficile* strain 431. Weight and stool CFU was monitored for up to 10 days post infection. B)  
2 NMDS ordination of donor stool communities prior to inoculating mice. Each point represents one donor  
3 and donors are colored by clinical diagnosis. C) NDMS ordination of the stool communities on day 0. Each  
4 symbol represents one mouse and is colored by donor. Circles represent mice that survived the 10 days  
5 of infection and triangles represent those who suffered severe disease.

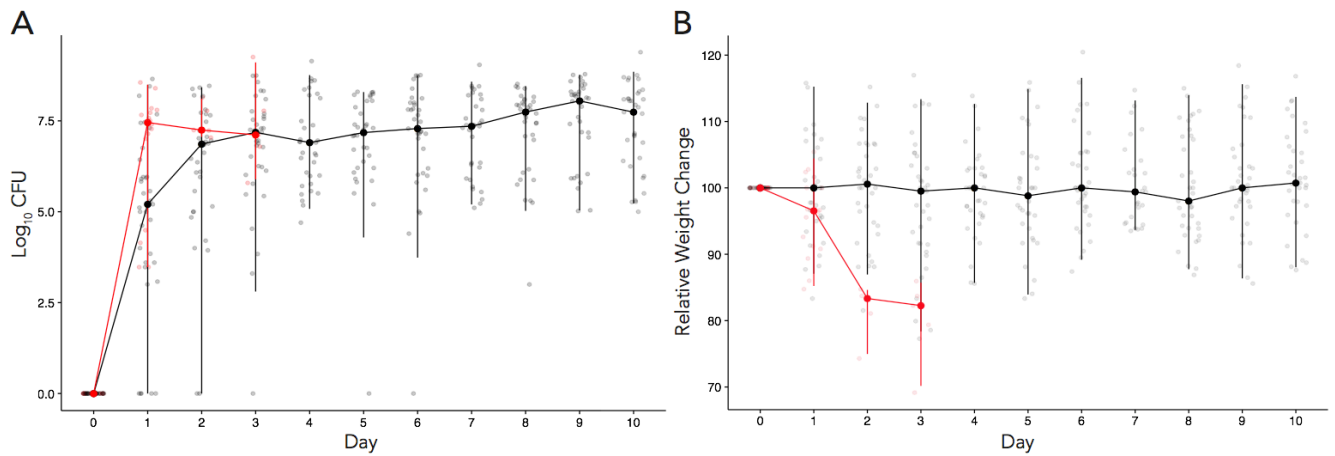
6  
7 Figure to dos:

8 A) Consider making this timeline a bit more streamlined/using R like Jenior's timelines look

9 B) Do ADONIS to show no differences or mantel test, report value on figure or in text?

0 C) Do mantel test for correlation of distances between severe/mild. Change legend to be mild/severe.  
1 Decide if colors are distinguishable enough (I think they are)

3 **Figure 2. *C. difficile* infection dynamics**

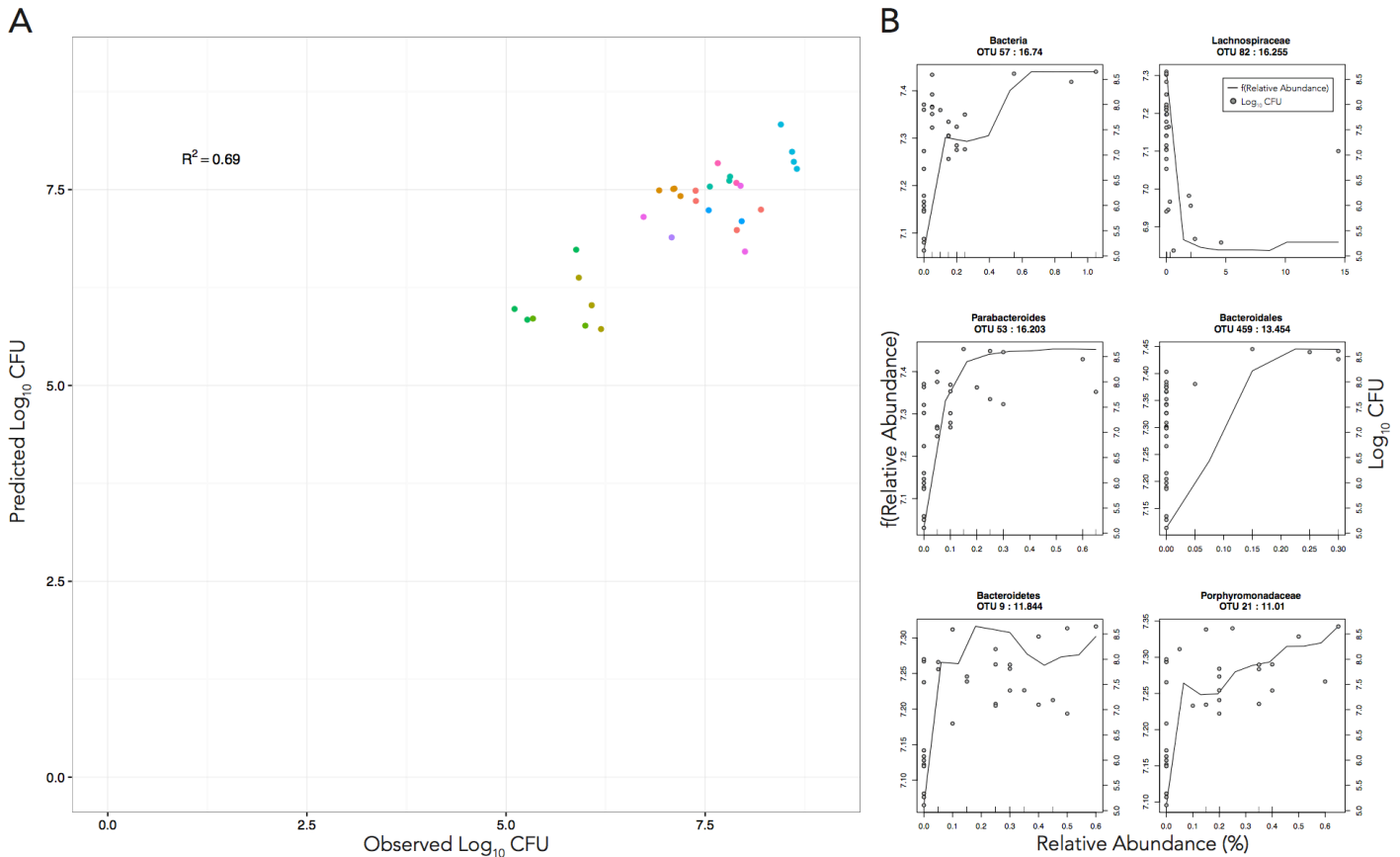


5 Figure 2. *C. difficile* infection dynamics. A) *C. difficile* CFU was enumerated by plating of mouse stool  
6 pellets daily. Each point represents a mouse and the lines represent the mean of mice in each cage and  
7 error bars are interquartile ranges. Red lines and points correspond to mice that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

2 **Figure to dos:**

3 **A + B) Decide if these are the final values/error we want to represent, add back mild/severe legend**

9 **Figure 3. Microbial community on day 0 predicts future *C. difficile* CFU**



1 Figure 3. Random Forest prediction of *C. difficile* colonization level. A) Day 0 microbial community  
2 members above 1% relative abundance were used to predict median log10 CFU of *C. difficile* after  
3 colonization. OTUs were chosen such that they were not predictive of cage or donor. Each point is a  
4 mouse colored by cage. B) Partial dependency plots of the top six predictive OTUs. Line displats the  
5 partial dependence of log10 CFU on the relative abundance of each predictive OUT. Each median log10  
6 CFU is plotted against its relative abundance for each predictive OTU.

7 **Figure to-dos**

8 **A) make points bigger, probably entire plot can be smaller**

9 **B)New partial dependency representation/graph style? Nick**

**Figure 4. Random forest predicts CDI severity from day 0 microbiome.**

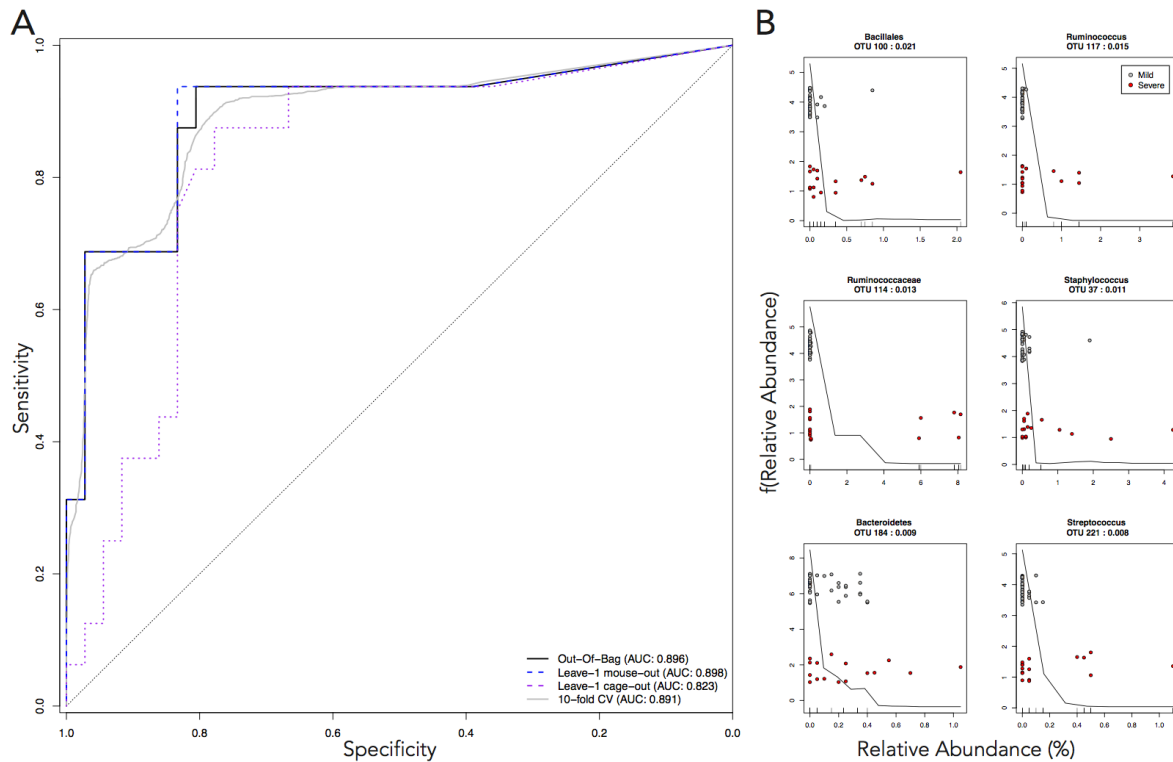


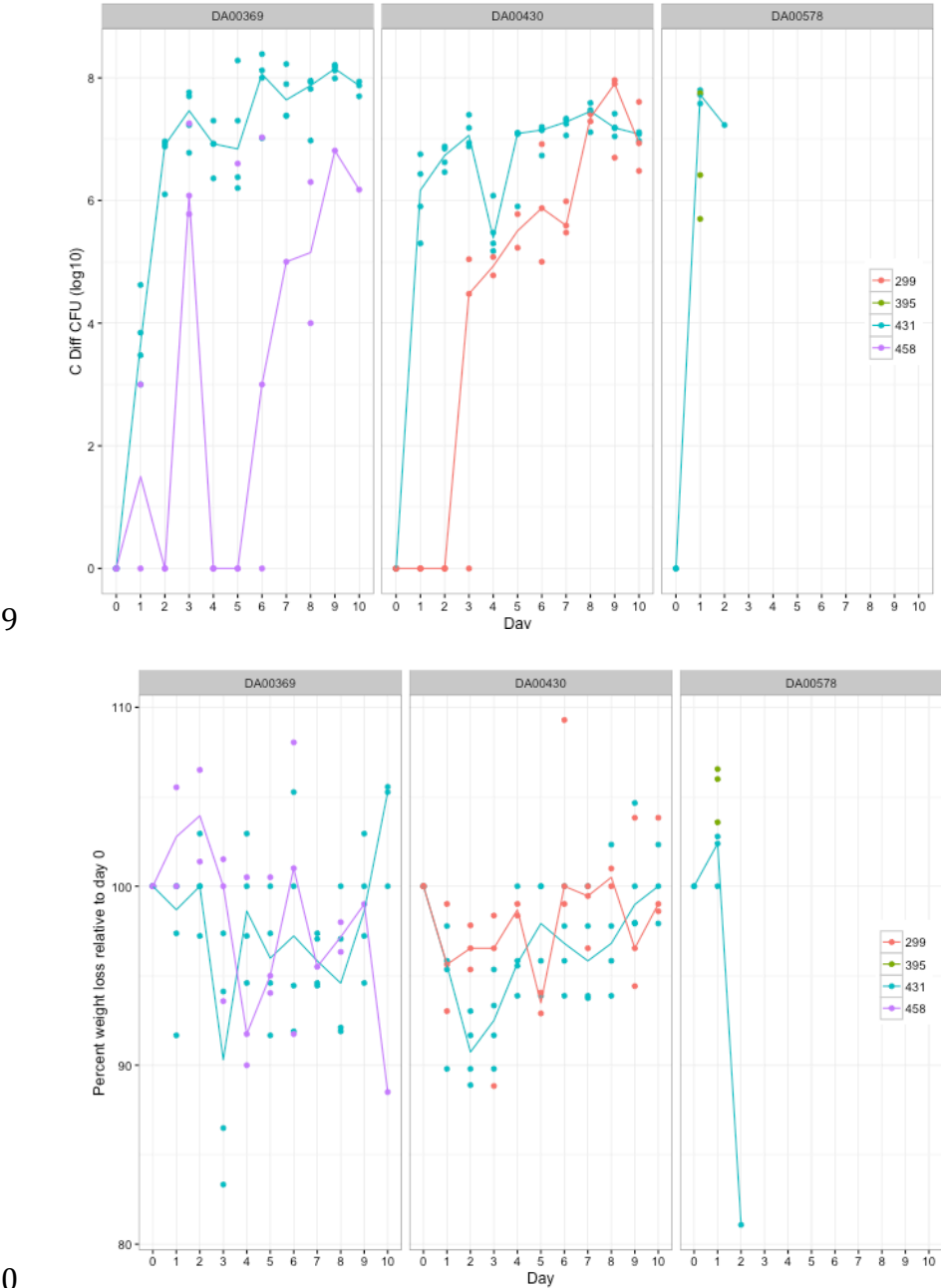
Figure 4. Random Forest prediction of CDI severity. Day 0 microbial community members above 1% relative abundance were used to predict disease severity. OTUs were chosen such that they were not predictive of cage or donor. Predictive classification tested via 10-fold (gray), leave-one-cage-out (purple 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 log10 CFU on OTU relative abundance. Points are the OTU relative abundance of each mouse colored by outcome (red, severe, black, mild).

Figure to -dos

A) decide on final models to present

B) new partial dependency plots/format, Nick?

8 **Figure 5. Propensity for severe CDI is community-dependent**



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

5 **Figure to-dos:**

- 6 A) decide how we want to label donors (numbers aren't really used elsewhere)  
7 B) Need to add lines for green points in DA00578 plots

8 **Supplement**

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

0

1     **Citations**

2