Clostridioides difficile exploits toxin-mediated inflammation to alter the host nutritional landscape and exclude competitors from the gut microbiota

Authors: Joshua R. Fletcher[1], Colleen M. Pike[1], Ruth J. Parsons[1], Alissa J. Rivera[1], Matthew H. Foley[1], Michael R. McLaren[1], Stephanie A. Montgomery[2], Casey M. Theriot[1*]

Affiliations: [1]Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, 1060 William Moore Drive, Raleigh, NC 27607

[2]Department of Pathology and Laboratory Medicine, Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, NC

[*]**Corresponding author information:**

Casey M. Theriot

Department of Population Health and Pathobiology

College of Veterinary Medicine

Research Building 406

North Carolina State University

1060 William Moore Drive

Raleigh NC 27607

cmtherio@ncsu.edu

Keywords: *C. difficile*, toxin, inflammation, transcriptomics, microbiome, colonization, collagen, amino acids, carbohydrates

Introductory paragraph

Clostridioides difficile is a bacterial pathogen that causes a range of clinical disease from mild to moderate diarrhea, pseudomembranous colitis, and toxic megacolon. Typically, *C. difficile* infections (CDIs) occur after antibiotic treatment, which alters the gut microbiota, decreasing colonization resistance against *C. difficile.* Disease is mediated by two large toxins and the expression of their genes is induced upon nutrient depletion via the alternative sigma factor TcdR. Using *tcdR* mutants in two strains of *C. difficile*, we defined how toxin-induced inflammation alters *C. difficile* metabolism, tissue gene expression, and the gut microbiota to determine how inflammation by the host may be beneficial to *C. difficile*. Here we show that *C. difficile* metabolism is significantly different in the face of inflammation, with changes in many carbohydrate and amino acid uptake and utilization pathways. Host gene expression signatures suggest that degradation of collagen and other components of the extracellular matrix by matrix metalloproteinases is a major source of peptides and amino acids that supports *C. difficile* growth *in vivo.* Lastly, the inflammation induced by *C. difficile* toxin activity alters the gut microbiota, excluding members from the genus *Bacteroides* that are able to compete against *C. difficile* for the same essential nutrients released from collagen degradation.

Introduction

Clostridioides difficile (*C. difficile*) is a Gram-positive anaerobic gut pathogen that causes diarrhea, with severe cases resulting in significant morbidity and mortality[1]. *C. difficile* produces two large toxins, TcdA and TcdB, that glycosylate host Rho and Rac GTPases, leading to a disruption in the actin cytoskeleton and loss of epithelial barrier integrity; the subsequent apoptosis and tissue damage results in significant inflammation[2]. Previous work revealed inflammation can be beneficial for prominent enteric pathogens such as *Salmonella enterica* and

Vibrio cholerae by providing a metabolic niche for them in the gut. Whether inflammation can benefit *C. difficile* is not clear, yet patients with Inflammatory bowel disease are four times more likely to acquire *C. difficile* infection (CDI) compared to the general population, suggesting *C. difficile* may thrive in an inflamed environment[3–7]. CDI-mediated inflammation results in drastic shifts to the murine gut metabolome, with alterations in amino acid and peptide metabolite concentrations, indicating that toxin activity induces an altered gut metabolic profile[8]. Although a nutritional generalist, *C. difficile* is auxotrophic and requires multiple amino acids, including the branched chain amino acids and proline that are used in Stickland metabolism for ATP production and regeneration of NAD+; thus, *C. difficile* must acquire these nutrients from its environment[9–12]. We therefore hypothesized that *C. difficile* gains access to these nutrients by exploiting the host inflammatory response. We hypothesized that toxin-mediated inflammation alters the host gut environment to benefit *C. difficile* growth and persistence, either through nutrient availability and/or the composition of the gut microbiota, potentially excluding competitors or selecting for allies[13,14].

We addressed this hypothesis by taking a holistic approach to define the response of the pathogen, the host, and the gut microbiota in the face of inflammation induced by *C. difficile* toxins. Here we show that the toxin producing strain (wild type *C. difficile*) induces a unique *C. difficile* transcriptomic signature compared to the toxin null strain (isogenic *tcdR* mutant), indicating that inflammation shapes *C. difficile* metabolism *in vivo*. *C. difficile* transcripts for carbohydrate and branched chain amino acid metabolism genes were increased in response to toxin-induced inflammation, which is a reflection of the nutrients available in the inflamed gut. Host tissue extracellular matrix (ECM)-degrading matrix metalloproteinase (MMP) transcripts, encoding enzymes responsible for breaking down amino acid rich collagen, were also increased in expression during peak inflammation. Additionally, we show that toxin activity leads to a reduction and reorganization in collagen around cells *in vitro*, which provided *C. difficile* a

mechanism to acquire essential Stickland reaction substrates, supporting growth. Colonization with toxin producing *C. difficile* also led to alterations in the gut microbial community structure, with inflammation suppressing the return of members from the Bacteroidaceae Family. Our results were conserved across different strains, as toxin activity of the epidemic *C. difficile* R20291 strain elicits similar responses in a mouse model, suggesting that these effects may be conserved across toxigenic *C. difficile* strains from phylogenetically distinct backgrounds.

Materials and methods

Bacterial strains, growth conditions, and mutagenesis. *C. difficile* strains 630Δ*erm* and an isogenic *tcdR*::*ermB* ClosTron insertion mutant (both kindly gifted by Rita Tamayo), as well as the R20291 strain and its isogenic Δ*tcdR* mutant were routinely grown in and on Brain Heart Infusion (BHI) or Tryptone Yeast (TY) broth and agar; plates and cultures were grown at 37°C in an anaerobic chamber (Coy). For genetic manipulation of *C. difficile*, strains were grown on and in BHI agar and broth supplemented as necessary with 10 μg/ml of thiamphenicol, 50 μg/ml of kanamycin, and 16 μg/ml of cefoxitin to select for transconjugants or thiamphenicol alone for plasmid maintenance. Samples derived from *in vivo* studies were plated onto CCFA (cefoxitin, cycloserine, fructose agar) to select for and enumerate vegetative *C. difficile* CFUs, and TCCFA, containing the germinant taurocholate to enumerate spore CFUs.

The R20291 Δ*pyrE* strain was used to construct the Δ*tcdR* mutant[15]. Briefly, ~1.2 kb upstream and downstream of the *tcdR* gene were PCR amplified with Phusion High-Fidelity DNA polymerase (NEB M0530S). All primers used for construction of the mutagenesis vector and for PCR screening of transconjugants can be found in Supplemental Table 1. The homology arms were then combined into one single linear fragment via splice overlap extension PCR, then A-tailed with Taq polymerase (NEB M0267S). The A-tailed product was ligated into pCR2.1 (Thermo Fisher K202020); the resulting plasmid was digested with BamHI and KpnI and the

~2.4 kb homology arm fragment was ligated into the corresponding sites in pMTL_YN4 using T4 DNA Ligase (NEB M0202S). The final plasmid was conjugated into the R20291 Δ*pyrE* strain with *E. coli* SD46, and thiamphenicol resistant large colony variants were screened by PCR for plasmid integration into the chromosome. After confirming plasmid integration, colonies were grown in BHI broth overnight in the absence of selection to allow for plasmid excision, then plated onto minimal media agar supplemented with 5 μg/ml uracil and 2 mg/ml of 5-fluoroorotic acid to select for bacteria that had excised and lost the mutagenesis vector. Individual colonies were re-streaked twice on the same selective media, then PCR screened for loss of *tcdR*.

Spore preparation. Spores were prepared as in Edwards and McBride[16]. 500 μL of mid-log phase cultures was spread onto 70:30 agar plates and incubated at 37°C for 4 days, after which time the plates were removed from the anaerobic chamber. The bacterial lawns were scraped off and suspended in 10 mL sterile PBS, mixed 1:1 with 96% ethanol, vortexed vigorously for 30 sec, and allowed to sit at room temperature on the benchtop for 1 hr. The suspension was centrifuged at 3,000 rpm for 10 min. The pellet was suspended in 10 ml fresh sterile PBS, and centrifuged again; this was repeated twice. The final pellet was suspended in 1 ml PBS and serial dilutions were plated on BHI agar with 0.1% of the germinant taurocholate for spore CFU enumeration. Spore stocks were also enumerated one day prior to the day of challenge to confirm spore stock CFUs prior to making the inocula for *in vivo* studies; inocula were also diluted and plated the day of challenge.

Mouse model of *C. difficile* infection. C57BL/6J WT mice (5 to 8 weeks old; n=18 male and n=18 female) were purchased from Jackson Labs. The mice were given 0.5 mg/mL cefoperazone in their drinking water for five days to make them susceptible to *C. difficile* infection[17,18]. The mice were then given plain water for two days, after which time they (n=12, males and females) received 10^5 spores of either 630Δ*erm* or *tcdR*::*ermB* via oral gavage. One

group of mice (n=12, males and females) received no *C. difficile* spores. Mice were weighed daily and monitored for clinical signs of distress (ruffled fur, hunched posture, slow ambulation, etc.). Fecal pellets were collected 1 and 3 days after challenge and diluted 1:10 w/v in sterile PBS, then serially diluted in 96 well PCR plates and plated onto CCFA for vegetative *C. difficile* CFU enumeration. The serially diluted samples were then removed from the anaerobic chamber and heated to 65°C for 20 min to kill vegetative cells; the heat-treated dilution plate was passed back into the anaerobic chamber and the dilutions were plated onto TCCFA to enumerate spore CFUs.

At day 2 and 4 post challenge, mice were humanely sacrificed (n=6 per treatment), and necropsy was performed. Cecal content was harvested for enumeration of vegetative *C. difficile* and spore CFUs, as well as for RNA extraction and toxin activity. Cecal tissue was harvested for RNA extraction for gene expression analysis, 16S rRNA sequencing, and histopathology. Colon tissue was harvested for histopathology. Samples for sequencing and toxin activity were immediately flash frozen in liquid nitrogen and stored at -80°C until processing. Toxin activity in the cecal content was quantified using the Vero Cell cytotoxicity assay[18]. Briefly, the content was diluted 1:10 w/v in sterile PBS, and 10-fold dilutions were added to Vero cells in a 96-well dish for ~16 hr. The reciprocal of the lowest dilution in which ~80% of the cells have rounded was reported as the titer.

The R20291 study was conducted similarly to the one described above with some minor differences. C57BL/6J mice (5 to 8 weeks old, n=14 male and 14 female) were orally gavaged with 10^5 spores of R20291 or the Δ*tcdR* mutant (n=14 per strain). Weight and clinical signs of distress were monitored daily. Fecal pellets were collected at 1, 2, and 4 days after challenge and total *C. difficile* CFUs were enumerated on TCCFA agar; samples were then heat-treated to kill vegetative cells for spore CFU enumeration. Necropsy was performed 2 and 4 days after challenge, and cecal tissue was harvested for 16S rRNA sequencing (n=5 per group on day 4

post challenge); cecal tissue RNA was isolated from mice (n=3 per group) on day 2 for gene expression analysis via NanoString.

Histopathological examination of the mouse cecum and colon. At the time of necropsy, tissue from the cecum and colon were prepared for histology by placing the intact tissue into histology cassettes and stored in 10% buffered formalin for 48 hr at room temperature, then transferred to 70% ethyl alcohol for long term storage. Tissues were processed, paraffin embedded, sectioned at 4 μm thickness, and hematoxylin and eosin stained for histopathological examination (University of North Carolina Animal Histopathology & Lab Medicine core). Histological specimens were randomized and scored in a blinded manner by a board-certified veterinary pathologist (SM). Edema, inflammation (cellular infiltration), and epithelial damage for the cecum and colon were each scored 0-4 based on a previously published numerical scoring scheme[17]. Edema scores were as follows: 0, no edema; 1, mild edema with minimal (2x) multifocal submucosal expansion or a single focus of moderate (2-3x) sub-mucosal expansion; 2, moderate edema with moderate (2–3x) multifocal sub-mucosal expansion; 3, severe edema with severe (3x) multifocal sub-mucosal expansion; 4, same as score 3 with diffuse sub-mucosal expansion. Cellular infiltration scores were as follows: 0, no inflammation; 1, minimal multifocal neutrophilic inflammation of scattered cells that do not form clusters; 2, moderate multifocal neutrophilic inflammation (greater submucosal involvement); 3, severe multifocal to coalescing neutrophilic inflammation (greater submucosal ± mural involvement); 4, same as score 3 with abscesses or extensive mural involvement. Epithelial damage was scored as follows: 0, no epithelial changes; 1, minimal multifocal superficial epithelial damage (vacuolation, apoptotic figures, villus tip attenuation/necrosis); 2, moderate multifocal superficial epithelial damage (vacuolation, apoptotic figures, villus tip attenuation/necrosis); 3, severe multifocal epithelial damage (same as above) +/− pseudomembrane (intraluminal neutrophils, sloughed epithelium in a fibrinous matrix); 4, same

as score 3 with significant pseudomembrane or epithelial ulceration (focal complete loss of epithelium). Photomicrographs were captured on an Olympus BX43 light microscope with a DP27 camera using cellSens Dimension software.

RNA extraction from cecal tissue and cecal content. RNA was extracted from cecal tissue using the PureLink RNA Mini kit (Thermo Fisher, 12183025) following the manufacturer's protocol. The RNA was treated with Turbo DNase (Thermo Fisher, AM2239); the protocol was modified by increasing the amount of enzyme to 5 μl per sample. After 30 min of incubation in a water bath at 37°C, 2 μl of Turbo DNase enzyme was added to each sample for a further 30 min of incubation. The RNA was then column purified according to the manufacturer's instructions (Zymo, R1019). PCR with primers specific to intron 1 and exon 3 from the mouse β-actin gene were used to screen samples for genomic DNA after DNase treatment.

For extraction of RNA from cecal content for bacterial RNAseq, the content was thawed on ice, then added to 10 mL of TRIzol Reagent (Thermo Fisher, 15596018) in a 15 mL conical. The content was dispersed by vortexing for 10 sec, and then was given 20 min on the benchtop to settle. The TRIzol-cecal content mix was then transferred in 1.2 mL aliquots to eight 1.7 mL centrifuge tubes. 350 μl of chloroform was added, and the tubes were vigorously inverted for 15 sec each, after which they were incubated at room temperature for 20 min. The samples were centrifuged at 14,000 rpm at 4°C for 20 min. The aqueous phase (~650 μl) was then added to 650 μl of isopropanol that had been supplemented with 5 μg/mL glycogen. Samples were vortexed and incubated on ice for 20 min, and then were centrifuged at 4°C for 30 min. Pellets were washed three times with 70% ethanol, and then dissolved in sterile deionized water. The samples were then treated with Turbo DNase, with the same augmentation of the protocol that was done for the cecal tissue RNA. For reasons that are unclear, the *in vivo* samples required multiple rounds of Turbo DNase treatment to remove contaminating genomic DNA, resulting in

197 degradation of the RNA in some samples. After column purification, PCR was performed with
198 primers specific to *tcdA* and *rpoC* to confirm removal of genomic DNA.

199 **RNA extraction from *C. difficile* cultures *in vitro.*** Three independent colonies of 630Δ*erm*
200 and *tcdR*::*ermB* each were inoculated into 5 ml of TY broth and grown overnight. These were
201 subcultured 1:500 in 5 ml fresh TY and allowed to grow for 18 hr at 37°C in the anaerobic
202 chamber. The cultures were centrifuged and supernatants were decanted. Pellets were
203 dissolved in 1 ml TRIzol Reagent for 20 min on the bench top, after which time 200 μl of
204 chloroform was added and the cultures were vigorously inverted for 15 sec and incubated on the
205 benchtop for a further 20 min. The RNA was precipitated as described above, with ~500 μl of
206 the aqueous phase added to 500 μl of isopropanol with glycogen supplementation. The Turbo
207 DNase treatment as described above was performed once, and RNA was confirmed to be free
208 of genomic DNA with the aforementioned primer sets.

209 **RNA sequencing and transcriptome analysis.** Sequencing of RNA derived from cecal content
210 and *in vitro* cultures was performed at the Roy J. Carver Biotechnology Center at the University
211 of Illinois at Urbana-Champaign. Ribosomal RNA was removed from the samples using the
212 RiboZero Epidemiology Kit (Illumina). RNAseq libraries were prepped with the TruSeq Stranded
213 mRNA Sample Prep Kit (Illumina). Library quantification was done via qPCR, and the samples
214 were sequenced on one lane for 151 cycles from each end of the fragments on a NovaSeq 6000
215 using a NovaSeq S4 reagent kit. The FASTQ files were generated and demultiplexed using the
216 bcl2fastq v2.20 Conversion Software (Illumina). Raw paired Illumina reads were imported into
217 Geneious 10.2.6, where adapters and low-quality reads were removed using BBDuk with a Kmer
218 length of 27, a minimum base quality score of 30, and a minimum average read quality of 30[19].
219 Reads less than 30 bases in length (and their paired read) were also discarded. The filtered
220 reads were mapped to the *C. difficile* 630Δ*erm* genome (NCBI accession no. NC_009089.1)

using BBMap with a Kmer length of 10 and no other changes to the default settings. Visual inspection of the data indicated that the majority of reads from three wild type and three *tcdR* samples, each from day 4, mapped to ribosomal RNA genes. These samples were excluded from the analysis, but are included in the SRA submission. The average number of reads that mapped from wild type (n=5) and the *tcdR* mutant (n=6) from day 2 were 7,395,921 and 22,768,296, respectively; average reads mapped from day 4 for wild type (n=3) and the *tcdR* mutant (n=3) were 12,959,255 and 4,319,864, respectively. Differential expression analysis was performed using DESeq2 with no changes to the default settings, and genes were considered differentially expression if they had ±1 log2 fold change and an adjusted p value of <0.05.[20]. Gene set enrichment analysis of differentially expressed genes was performed using the GSEA-Pro v3 program (http://gseapro.molgenrug.nl) with user defined cutoff values of -1 and 1. Bar plots of enriched Gene Ontology (GO) terms and log2 fold change values of individual transcripts were generated in GraphPad Prism 8.

Quantitative reverse transcription PCR. RNA from cecal tissue, cecal content, and *in vitro* bacterial cultures was used as template in reverse transcription reactions using the Murine Moloney Leukemia Virus Reverse Transcriptase (NEB M0253S) following the manufacturer's protocol. The resulting cDNA was diluted 1:5 in deionized water and used as template for quantitative PCR with the SsoAdvanced Universal SYBR Green Supermix (Bio Rad). Quantification of each gene assayed (Supplemental Table 1 for primers) was performed via standard curve and copy number was determined by comparison to the housekeeping genes *tbp* (TATA Binding Protein) for host genes and *rpoC* (RNA polymerase subunit beta) for *C. difficile* genes.

NanoString analysis. RNA from cecal tissue was submitted to the Lineberger Comprehensive Cancer Center Pre-Clinical Genomic Pathology Core at the University of North Carolina at Chapel Hill for quantification of transcripts via NanoString technology[21]. The RNA was hybridized

to probes on the Mouse nCounter Immunology Panel, plus custom probes targeting mouse *Mmps* and *Timps*. Raw data was imported into the nSolver Advanced Analysis software for data normalization and differential expression analysis. One mouse (challenged with wild type, day 4 post challenge) was excluded after principal components analysis and hierarchical clustering of the data identified it as an outlier with respect to all other samples. The data were normalized and differential expression analysis was performed within the nSolver Advanced Analysis software. Correction for multiple comparisons was performed using the method of Benjamini-Hochberg. Heatmaps of the data were constructed in R using the 'pheatmaps' package (https://cran.r-project.org/web/packages/pheatmap/index.html) and volcano plots were constructed in R with the 'EnhancedVolcano' package[22]. Gene set enrichment analysis was performed using the WebGestalt server (http://www.webgestalt.org) with the following changes to the default parameters: the minimum number of genes required for a pathway was lowered to 5 and the False Discovery Rate was adjusted to 0.1. Enriched pathways were visualized in Prism 8. Cecal tissue RNA from mice infected with R20291 and the Δ*tcdR* mutant was also used in a separate run with the NanoString Mouse nCounter Inflammation panel customized to include code sets for the *Mmp*s and *Timp*s. The data was analyzed identically in the nSolver Advanced program as described above.

Confocal microscopy. IMR90 human fibroblasts were cultured in Eagle's Minimum Essential Medium (EMEM) (ATCC, USA) supplemented with 10% fetal bovine serum at 37 °C with 5% CO_2. Cells were seeded on glass coverslips in 24-well plates for 3 days, followed by incubation with 0.5 pM of TcdA and TcdB. After 12 hr or 15 hr, cells were fixed in PBS with 4% paraformaldehyde for 20 min at room temperature and blocked in 10% normal goat serum (Sigma). Collagen was detected using a mix of antibodies against collagen types I, III, and V (Santa Cruz) in a 1:1:1 ratio, and Alexa Fluor 568-conjugated goat anti-mouse secondary antibody (ThermoFisher). Glass coverslips were mounted using VECTASHEIELD mounting

media with DAPI (Vector Laboratories). Confocal imaging was performed on Zeiss LSM 880 confocal microscope using a 40× Plan-Apochromat objective lens (numerical aperature of 1.4) and operated with ZEN software (Carl Zeiss, Inc).

***C. difficile* growth in defined minimal media supplemented with heat-degraded collagen.** Heat-degraded collagen was generated by heating type I collagen (Advanced BioMatrix) at 100°C for 4 hr. Degraded collagen was then concentrated and the pH was adjusted to 7.0. *C. difficile* was grown in a well-established defined minimal media (CDMM), from which proline was omitted and 0.5 mg of heat-degraded collagen was substituted. The media was passed into the anaerobic chamber and allowed to reduce for 24 hr before inoculation. At the same time, individual colonies of *C. difficile* were inoculated into TY media for overnight growth, after which they were subcultured 1:100 into fresh TY. After 4 hr of growth, the cultures were centrifuged in a microcentrifuge in the anaerobic chamber and washed three times in 1 ml sterile PBS, then inoculated 1:500 into the CDMM. Immediately after inoculation, the cultures were serially diluted and plated onto BHI agar for enumeration of *C. difficile* at 0 hr and 24 hr time points. Growth was calculated from 4 independent experiments.

***C. difficile* and *Bacteroides* growth in defined minimal media supplemented with Pro-Gly or Gly-Pro dipeptides, and other collagen degradation substrates.** To assess the ability of *C. difficile* to acquire the essential amino acid proline from dipeptides, we utilized a well-established defined minimal media (CDMM), from which proline was omitted and Proline-Glycine or Glycine-Proline dipeptides were substituted. The media was passed into the anaerobic chamber and allowed to reduce for 24 hr before inoculation. At the same time, individual colonies of *C. difficile* were inoculated into TY media for overnight growth, after which they were subcultured 1:100 into fresh TY for 4 hr. After the 4 hr growth, the cultures were centrifuged in a microcentrifuge in the anaerobic chamber and washed three times in 1 ml sterile PBS, then inoculated 1:500 into the CDMM. Immediately after inoculation, the cultures were serially diluted

and plated onto BHI agar for enumeration of *C. difficile* at the 0 hr time point. 24 hr later, the cultures were serially diluted and plated for CFU enumeration again. Growth was calculated from 3 independent experiments.

Bacteroides thetaiotaomicron VPI-5482 and *Bacteroides fragilis* Bf NCTC 9343 were cultured anaerobically at 37°C from glycerol stocks into tryptone-yeast extract-glucose (TYG) medium and grown overnight[23]. Cultures were back diluted to an OD 600 nm of ~0.1 the next day into minimal media (MM) containing 0.25% glucose, proline, or hydroxyproline[24]. CFUs were enumerated on BHI-blood agar plates at 0 hr and after 16 hr of growth. Fold change in growth was calculated from 3 independent experiments.

16S rRNA Illumina sequencing and microbiome analysis. DNA was isolated from cecal snips at the University of Michigan Microbial Systems Molecular Biology Laboratory. The Mag Attract Power Microbiome kit (Mo Bio Laboratories, Inc.) was used to isolate DNA from cecal snips. Dual-indexing sequencing approach was used to amplify the V4 region of the 16S rRNA gene. Each PCR mixture contained 2 µl of 10X Accuprime PCR buffer II (Life Technologies, CA, USA), 0.15 µl of Accuprime high-fidelity *polymerase* (Life Technologies, CA, USA), 5 µl of a 4.0 µM primer set, 1 µl DNA, and 11.85 µl sterile nuclease free water. The template DNA concentration was 1 to 10 ng/µl for a high bacterial DNA/host DNA ratio. The PCR conditions were as follows: 2 min at 95°C, followed by 30 cycles of 95°C for 20 sec, 55°C for 15 sec, and 72°C for 5 min, followed by 72°C for 10 min. Libraries were normalized using a Life Technologies SequalPrep normalization plate kit as per manufacturer's instructions for sequential elution. The concentration of the pooled samples was determined using the Kapa Biosystems library quantification kit for Illumina platforms (Kapa Biosystems, MA, USA). Agilent Bioanalyzer high-sensitivity DNA analysis kit (Agilent CA, USA) was used to determine the sizes of the amplicons in the library. The final library consisted of equal molar amounts from each of the plates, normalized to the pooled plate at the lowest concentration. Sequencing was done on the Illumina

MiSeq platform, using a MiSeq reagent kit V2 (Ilumina, CA, USA) with 500 cycles according to the manufacturer's instructions, with modifications[25]. Sequencing libraries were prepared according to Illumina's protocol for preparing libraries for sequencing on the MiSeq (Ilumina, CA, USA) for 2 or 4 nM libraries. PhiX and genomes were added in 16S amplicon sequencing to add diversity. Sequencing reagents were prepared according to the Schloss SOP (https://www.mothur.org/wiki/MiSeq_SOP#Getting_started), and custom read 1, read 2 and index primers were added to the reagent cartridge. FASTQ files were generated for paired end reads.

Raw reads were processed in QIIME2, with DADA2 used for de-noising and generating amplicon sequence variants (ASVs)[26,27]. Taxonomic assignment of the ASVs was done using the Silva reference database (silva-132-99-nb-classifier.qza)[28]. The code used to process the reads can be found in Supplemental File 7 for 630Δ*erm* and Supplemental File 8 for R20291. Percent relative abundances of Family level ASVs were calculated for each sample in Excel, averaged across treatment groups, and visualized in GraphPad Prism 8.

Principal components analysis (PCA) (Fig. 5b) was performed in the R statistical programming environment (https://www.r-project.org). ASVs from the V4 region that differ by just one base pair may come from the different 16S copies in the same genome[29]. We therefore aggregated ASVs into 99%-identity OTUs using complete-linkage clustering on the Levenshtein edit distances between ASV sequences. Principal components analysis was then performed on Hellinger-transformed OTU abundances[30] and the first two principal components plotted with sample scaling (scaling 1). The R code can be found in Supplemental File 9. The essential R packages used were biomformat, Biostrings, phyloseq, vegan, ggplot2, and data importing and manipulation packages from the tidyverse package collection[31,32].

Statistical analysis. With the exception of the RNAseq analysis, all statistical tests were performed in GraphPad Prism 8. Kruskal-Wallis One-Way ANOVA with Dunn's correction for

multiple comparisons was used to test for significance when comparing *C. difficile* CFUs, mouse weights, and cecal toxin activity. A mixed effects model with Tukey's multiple comparison's test was used to test for significance on qRT-PCR data. Histopathology summary scores were tested for significance using a Geissner-Greenhouse corrected ordinary Two-Way ANOVA with Tukey's multiple comparisons test. A Mann-Whitney test was used to compare the fluorescence quantification of IMR90 cells treated with vehicle or toxins. DESeq2 identified statistically significant differentially expressed genes in the RNAseq study. A p-value of < 0.05 was considered statistically significant, with *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Results

Wild type *C. difficile* induces significantly more inflammation and tissue damage than a *tcdR* mutant. Antibiotic treated mice were challenged with 10^5 spores of *C. difficile* 630Δ*erm* (wild type, or wild type mice hereafter) or an isogenic *tcdR*::*ermB* (*tcdR,* or *tcdR* mice) mutant on day 0 and clinical signs of disease were monitored for four days post challenge (Fig. 1a). Mutation of *tcdR* has been reported to significantly reduce levels of both toxin gene expression and toxin protein production *in vitro*[33,34]. Nearly five-fold fewer *tcdR* vegetative cells were recovered in the feces relative to wild type at day 1 ($p=0.0314$, Kruskal-Wallis with Dunn's correction for multiple comparisons); however, there was no difference by day 3, nor were significant differences detected in fecal spores between the two strains at either day (Fig. 1b and 1c). Cecal content from day 2 and 4 did not harbor significantly different vegetative cells or spores, however nearly 10^5-fold more toxin activity was detected in wild type mice compared to *tcdR* mice at both days, indicating that the *tcdR* mutant behaves similarly *in vivo* as it does *in vitro* with respect to toxin production (Fig. S1a-c, Fig. 1d). Accordingly, histopathological analysis of cecal tissue found significantly increased inflammation in wild type mice when compared to uninfected controls (no *C. difficile*) at day 2 ($p=0.006$; Two-Way ANOVA with Tukey's multiple comparisons test), as well as significantly more epithelial damage when compared to both

uninfected controls and *tcdR* mice at day 4 (p=0.024 for both; Two-Way ANOVA with Tukey's multiple comparisons test) (Fig. 1e-f). While cecal inflammation, epithelial damage, and edema were lower in *tcdR* mice at day 2 relative to wild type mice, it was not statistically significant. Tissue damage in wild type mice was even more pronounced in colonic tissue (Fig. S1d). Together, these data show that the *tcdR* mutant fails to produce much detectable toxin activity *in vivo*, and consequently does not elicit significant inflammatory damage to host gut tissue.

Toxin-mediated inflammation significantly alters the *C. difficile* transcriptome *in vivo*. As colonization with wild type *C. difficile* leads to significant increases in inflammation and damage to the cecal epithelium, we hypothesized that *C. difficile* would shift its transcriptome to reflect such dramatic differences in the inflammatory environment[35]. However, a *tcdR* mutant in the R20291 strain has numerous differentially expressed genes *in vitro*[34]. To assess whether the *tcdR* mutation is pleiotropic in 630Δ*erm*, we performed RNAseq on wild type and the *tcdR* mutant grown for 18 hours in TY media as an *in vitro* control. We found that other than the genes of the Pathogenicity Locus (PaLoc), only two genes were differentially expressed (log2 fold change ±1 and adjusted p value<0.05) between the strains *in vitro* (*CD1917*, encoding *eutE*, and *CD3087*, encoding a transcription factor of the RpiR family) (Supplemental File 3). When comparing *in vivo* expression profiles between the two strains, the majority of differentially expressed genes were detected at day 2, with 86 transcripts increased and 82 decreased in wild type relative to *tcdR* (Supplemental File 3). After four days, wild type had 15 transcripts increased and 4 decreased. Consistent with the cecal content toxin activity assay, among the most significantly increased transcripts in wild type relative to the *tcdR* mutant were of the PaLoc. Interestingly, time was a more important variable with respect to the number of differentially expressed genes *in vivo* in both wild type and the *tcdR* mutant. When comparing wild type at day 4 to wild type at day 2, 249 genes had increased transcript levels and 155 decreased; the *tcdR* mutant at day 4 relative to day 2 had even more dramatic changes in gene expression, with 380 transcripts

increased and 338 decreased. A subset of genes identified as differentially expressed between wild type and the *tcdR* mutant were selected for quantitative real-time PCR (qRT-PCR) validation, which confirmed trends in expression from the RNAseq (Fig. S3a-d). Together, these data show that inflammation is an important environmental determinant of the *C. difficile* transcriptome *in vivo*.

Toxin-induced inflammation alters *C. difficile* metabolism. We used prokaryotic gene set enrichment analysis to summarize the main patterns in differential gene expression (Fig. 2a, Fig. S2, and Supplemental File 4). Gene Ontology (GO) terms related to carbohydrate metabolism were enriched in wild type relative to *tcdR* at both days post challenge, suggesting that wild type *C. difficile* may have had access to different carbohydrate nutrient sources. An operon encoding a phosphoenolpyruvate:carbohydrate phosphotransferase (PTS) system annotated to be specific to mannose/fructose/sorbose was significantly increased in wild type *C. difficile* at day 2 (Fig. 2b). PTS systems are typically induced by the presence of the carbohydrate that they import (and absence of a repressing carbohydrate) via a transcriptional antitermination mechanism[36]. Expression of this operon normalized between wild type and the *tcdR* mutant by day 4, though at this time point the wild type had increased expression of genes predicted to be involved in fructose/mannitol and tagatose metabolism, as well as genes involved in extracellular polysaccharide production (Fig. 2b).

Among the most abundant GO terms in the set of transcripts that were decreased in wild type were those for oxidation-reduction and catalytic processes, as well as those for leucine biosynthesis. The transcript levels for *ilvD*, involved in isoleucine and valine biosynthesis, were also decreased in wild type relative to *tcdR* (Fig. 2b). Both the *leu* operon and *ilvD* in *C. difficile* are transcriptionally repressed by CodY, whose repressive activity is high when bound by branched chain amino acids and/or GTP[37,38]. Similar patterns of decreased expression in wild type were seen in the CodY-regulated operon encoding the machinery for metabolism of

succinate to butyrate (Fig. 2b). CodY has also been shown to positively regulate the expression of some genes *in vitro*, including *pflB*[38]. We found that *pflB* expression *in vivo* was increased at day 2 in wild type *C. difficile* (Fig. 2b). These data, when combined with previous metabolomic studies, suggest that the metabolomic environment of the inflamed gut may be enriched for specific carbohydrates and oxidative Stickland reaction substrates in the form of branched chain amino acids.

Wild type *C. difficile* also had increased transcript levels of a number of genes involved in amino acid acquisition and metabolism compared to *tcdR in vivo*. One such gene, *CD3442*, encodes a putative M24 family Xaa-Pro prolidase. Eukaryotic prolidases are intimately linked to collagen metabolism, while prokaryotic prolidases are often involved in protein turnover and proline recycling[39,40]. In contrast, the gene encoding 4-hydroxyphenylacetate decarboxylase (*hpdB*) was decreased in wild type at day 4; the transcript from the *hpdC* gene immediately downstream was also decreased (p=0.08) (Fig. 2b). HpdB is involved in the fermentation of tyrosine to p-cresol, which has been shown to affect fitness *in vivo* in a murine relapse model and to modulate gut microbial community structure; the decreased *hpdB* transcript levels we observed may be consistent with lower levels of its substrate in the ceca of wild type mice[41–43].

Gene expression for multiple aspects of *C. difficile* physiology is altered in the presence of inflammation. Multiple genes encoding structural components of the flagella, including some that have been shown to induce inflammatory responses from host cells *in vitro* and *in vivo*, were decreased in expression in wild type (Fig. 2b). In contrast to flagellar genes, wild type had increased expression of the *cdd* operon, comprised of three genes (*cdd4*, *cdd3*, and *cdd2*) that are divergently transcribed from a two-component system (TCS) response regulator and histidine kinase (Fig. 2b). The *cdd* genes are annotated to encode the components of a multidrug/antibiotic ABC transport system, and given their genomic association with a TCS, it is tempting to speculate that it may encode an undescribed defense mechanism against

antimicrobial peptides that could function semi-analogously to the CprK-CprR/CprABC system that has been described for *C. difficile*[44,45].

Wild type *C. difficile* induces a robust inflammatory and proteolytic gene expression profile in host gut tissue. To determine how the host responds to *C. difficile* toxin-induced inflammation we compared the cecal tissue gene expression between three groups of mice (uninfected or no *C. diff*, wild type, and *tcdR*) with the NanoString Mouse Immunology panel modified to include probes targeting transcripts encoding matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) (Fig. 3, Fig. S4-S6, and Supplemental File 1). The top 50 differentially expressed transcripts (in terms of significance) from the wild type relative to *tcdR* comparisons from both days were combined and plotted in a heatmap with hierarchical clustering of samples, and two distinct clusters were observed (Fig. 3a). All samples from *tcdR* mice and uninfected controls formed one large cluster, while all wild type mice formed their own distinct cluster; neither cluster showed sub-clustering based on time points. Only 42% of the differentially expressed genes in wild type mice relative to *tcdR* mice were significant at both day 2 and 4, suggesting that the nature of the immune response to wild type *C. difficile* changed over the course of infection (Fig. S5). In contrast to wild type mice, *tcdR* mice had no significant differentially expressed genes in their cecal tissue when compared to uninfected controls (data not shown). A number of transcripts were selected for further analysis via qRT-PCR, which confirmed expression patterns observed via the NanoString approach (Fig.S6a-e). These data are consistent with the histopathological analysis, and show that in the absence of toxin activity, the *tcdR* mutant is relatively inert *in vivo* with respect to stimulating a host immune response.

We next performed gene set enrichment analysis for each group of differentially expressed genes, using GO biological process terms and considering the direction of expression for each transcript (Fig. 3b and Supplemental File 2). The GO term for regulation of inflammatory processes was enriched in transcripts with increased abundance in cecal tissue from wild type

mice at both days. The second most enriched at day 2 was regulation of peptide secretion, consistent with the role of host-derived antimicrobial peptides being produced as an arm of the innate immune response. Of particular interest was the enrichment of genes involved in the positive regulation of proteolysis, as peptide fragments derived from these processes may serve as nutrient sources for *C. difficile*, which is well-known for using amino acid fermentation as an energy source *in vitro* and *in vivo*[8,9,35,46–50]. In addition to the upstream regulators of various proteolytic processes, cecal tissue from wild type mice had significant increases in transcripts from genes encoding multiple MMPs including *Mmp3*, *Mmp10*, *Mmp12,* and *Mmp13*; probes targeting *Mmp3* were not included in the NanoString custom panel, so fold change for this transcript is reported based on qRT-PCR results (Fig. 3c, Fig S6c). Taken together, these data show that *C. difficile* toxin activity induces a highly inflammatory gut environment, and implicate MMP substrates, such as collagen and other ECM components, as reservoirs of Stickland substrate amino acids *in vivo*.

Toxin activity induces collagen degradation which supports *C. difficile* growth *in vitro*. The increase in numerous *Mmp* transcripts during toxin-induced inflammation suggested the ECM may be altered during CDI. We chose to examine whether toxin activity on cells *in vitro* affected collagen integrity, as it is highly abundant and an excellent source of proline, hydroxyproline, glycine, and alanine, all of which are amino acids that *C. difficile* can ferment via the Stickland reaction[11]. Confluent ECM-producing IMR90 human fibroblast monolayers were cultured on 24-well plates for three days prior to toxin treatment. Collagen remodeling was then visualized using immunofluorescence (Fig. 4a). By confocal microscopy, we observed toxins caused a notable disruption of the collagen network over a 12 hour period. Extensive networks of collagen fibrils were apparent in untreated cells, whereas collagen in toxin-treated cells appeared fragmented and condensed into globular structures. Moreover, we observed a

significant decrease in collagen fluorescence over a 15 hour period, validating that collagen was being degraded in the presence of toxins (Mann-Whitney test $p<0.0001$) (Fig. 4b).

Given that toxins induced the degradation of collagen in IMR90 cells, we next speculated that *C. difficile* can acquire nutrients from degraded collagen. To test this, *C. difficile* was grown in a minimal media with proline, without proline, or without proline and supplemented with heat-degraded collagen for 24 hours (Fig. 4c). Although not statistically significant ($p=0.233$, Kruskal Wallis one-way ANOVA), *C. difficile* grew approximately ten-fold higher in media supplemented with degraded collagen compared to media with no proline, indicating collagen can provide a source of proline for *C. difficile* growth. Additionally, when host collagen rich dipeptides Pro-Gly and Gly-Pro, were substituted for collagen degradation products, *C. difficile* grew to levels comparable to the standard minimal media control after 24 hours ($p=0.0015$ for both conditions compared to their respective 0 hour time point CFUs, two-way ANOVA with Sidak's multiple comparisons test) (Fig. 4d). This suggests that *C. difficile* can exploit host collagen degrading activity as a means to acquire nutrients.

Toxin-mediated inflammation suppresses the return of the Bacteroidaceae in the gut microbiota. Given the importance of the microbiota in rendering the gut an inhospitable environment for *C. difficile*, we hypothesized that toxin-mediated inflammation may exclude or suppress members of the gut microbiota that contribute to colonization resistance, and/or select for microbes that may benefit *C. difficile* through further niche modification/preservation, or other mechanisms like cross-feeding. We performed 16S rRNA amplicon sequencing on cecal DNA from mice challenged with wild type, the *tcdR* mutant, and uninfected controls or no *C. diff* (Supplemental File 5 for ASV relative abundances and taxonomy). As expected, a significant driver of community similarity was *C. difficile* colonization status, however a number of mice from the uninfected controls (no *C. diff*) and the *tcdR* mice had community structures with an increased abundance of ASVs from the Akkermansaceae (at day 2) and Bacteroidaceae (both

days) Families, which were low or undetected in the cecal microbiota of wild type mice (Fig. 5a). Additionally, a number of low abundance Family members, including the Coriobacterales, Paenibacillaceae, and Burkholderiaceae, were present in either uninfected controls, *tcdR* mice, or both, but undetected in the cecal microbiota of wild type mice (Supplemental File 5). Amplicon sequence variants (ASVs) from the Staphylococcaceae Family were detected at relatively high abundance in uninfected controls at both days, but were in low abundance or undetected in mice challenged with either strain of *C. difficile* (Fig. 5a). ASVs generated by DADA2 were clustered into operational taxonomic units (OTUs) at 99% sequence identity, and Hellinger-transformed OTU abundances were analyzed by principal components analysis (PCA) to determine the similarity of each cecal community (Fig. 5b). At day 2, both wild type mice and all but one of the uninfected controls formed tight, distinct clusters, while *tcdR* mice showed no specific clustering. By day 4, no group clustered very closely. Some wild type and *tcdR* mice community structures were driven by ASVs from various Lachnospiraceae and Erysipelotrichaceae Families, while community structures of other *tcdR* and uninfected mice were driven by ASVs from the Bacteroidaceae, Akkermansaceae, and Staphylococcaceae Families. These data show that colonization with *C. difficile* and toxin-mediated inflammation can significantly impact the return of the gut microbial community structure.

The ASV from the Bacteroidaceae Family identified in this study classified as genus *Bacteroides*. Since the Bacteroidaceae were more abundant in some mice in the absence of inflammation, and hydroxyproline utilization genes are enriched in this Family, we sought to characterize the growth of two representative members, *Bacteroides thetaiotaomicron* and *Bacteroides fragilis*, in a minimal media with and without supplementation of collagen rich substrates, proline or hydroxyproline (Fig. 5c-d)[51]. Supplementation of the minimal media with either amino acid led to approximately ten-fold increased growth for both species over 16 hours relative to un-supplemented media. Glucose was required for robust growth, however when

minimal media with glucose was supplemented with either amino acid there was almost two-fold higher growth of *B. thetaiotaomicron* and *B. fragilis* relative to glucose alone, though this was not statistically significant. This trend of increased growth in supplemented media indicates that both proline and hydroxyproline can be utilized by members of the genus *Bacteroides*, and that these amino acids may be valuable nutrients to compete for in the gut.

Toxin-mediated alterations of host *Mmp* expression and the gut microbiota are conserved in mice challenged with epidemic R20291 *C. difficile* strain. As the wild type strain used in this study (630Δ*erm*) is a multi-passaged, erythromycin sensitive lab strain, we sought to replicate our findings in the clinically relevant R20291 strain. A Δ*tcdR* mutant was constructed via allelic replacement, and the parent strain was used to challenge antibiotic treated mice (Fig. 6). No differences in total fecal *C. difficile* load (vegetative cells + spore) were observed between the strains at any day post challenge, however, significantly fewer Δ*tcdR* spores were recovered from feces at day 2 ($p<0.0001$, Kruskal-Wallis with Dunn's correction for multiple comparisons), consistent with a previous report on a *tcdR* ClosTron mutant in R20291 showing decreased sporulation efficiency *in vitro* (Fig. 6a-b)[34]. Similar to what was observed for 630Δ*erm* and its *tcdR*::*ermB* derivative, cecal content from mice challenged with wild type R20291 had significantly higher toxin titers than that from Δ*tcdR* at both days (Fig. 6c). Importantly, comparable patterns of increased expression of the same *Mmp*s was observed in cecal tissue from wild type R20291 mice relative to the Δ*tcdR* mice (Fig. 6d). Given the pleiotropic nature of the *tcdR* mutation in the R20291 strain, we elected not to perform *C. difficile* gene expression studies *in vivo*. 16S rRNA amplicon sequencing on cecal tissue isolated at day 4 showed similar patterns in the microbial community structures as observed in mice challenged with the 630Δ*erm* strains. Wild type R20291 mice had very low or undetectable Bacteroidaceae ASVs, while Δ*tcdR* mice had considerably higher levels (2% average relative abundance vs. 32%, respectively) (Fig.

6e and Supplemental File 6). The ASV designated as *Bacteroides* in this study is identical to the *Bacteroides* ASV that was identified in the cecal communities of uninfected controls and *tcdR* mice from the 630Δ*erm* study in Fig. 5a-b. Collectively, these data show that induction of host *Mmp* gene expression is a conserved component of the immune response to *C. difficile* toxin activity, and that the decreased levels of Bacteroidaceae in the inflamed gut may be biologically significant and an additional mechanism where *C. difficile* is able to exploit a niche and thrive due to host inflammation.

Discussion

C. difficile is a major nosocomial pathogen and cases of CDI are beginning to be diagnosed in individuals who lack the classic predisposing traits of recent antibiotic use or compromised immune status[52]. While effective treatments exist for *C. difficile*, some patients require fecal microbiota transplants (FMTs) to resolve their infections, highlighting the need to better understand how this pathogen creates a niche for persistence in a host. CDI is highly inflammatory; therefore, we leveraged the power of bacterial genetics with *tcdR* mutants in two strains of *C. difficile* to define how toxin-mediated inflammation alters the gut environment *in vivo*, with a focus on bacterial metabolism and the gut microbiome. We observed changes in numerous *C. difficile* metabolic genes *in vivo*, consistent with the hypothesis that toxin-induced inflammation alters the nutrient landscape in the host gut environment. In particular, we found time-dependent increases and decreases in the expression of multiple PTS carbohydrate import genes, and decreased expression of CodY-regulated genes for branched chain amino acid biosynthesis and butyrate production in wild type *C. difficile*. On its face, this represents a paradox: CodY de-repression of *tcdR*, *tcdA*, and *tcdB* leads to extremely high levels of the toxins, yet in our study we found that a number of CodY regulated genes were repressed *in vivo* in the presence of inflammation. Two possible mechanisms may explain this discrepancy: phase variation in the expression of *sigD*, a positive regulator of *tcdR*, and bimodal expression of *tcdR*

and the toxin genes[33,53]. Our transcriptomic approach captured average expression of genes within large populations of bacteria in each sample, but cannot detect differences in expression between individual cells. Future studies examining the per-cell expression levels of the CodY-regulated genes identified in our RNAseq studies are warranted. Regardless, our RNAseq approach has identified numerous targets for further study via mutagenesis with respect to *C. difficile* metabolic requirements *in vivo*, or how it resists the deleterious effects of host inflammation. Further, it highlights the need to fully understand the spectrum of behavior of individual bacterial cells across a population during the infection process. This may open novel avenues for therapeutic targeting of specific subsets of pathogens within a metabolically heterogeneous population.

Our data support a model where the activity of the toxins stimulate an inflammatory host response that includes gene expression signatures consistent with degradation of collagen and other components of the ECM. Collagen is rich in Stickland reaction substrates like proline (and hydroxyproline, which *C. difficile* can dehydrate to proline) and glycine, amongst other amino acids that *C. difficile* can ferment[10,51,54]. Hence, the ECM and collagen may serve as a reservoir of preferred amino acid nutrients that sustain the metabolic burden of large bacterial populations and production of the toxins over time within the host gut environment. Aberrant MMP activity has been reported as a factor in the pathogenesis of Inflammatory bowel disease (IBD), and IBD patients may be more likely to contract CDI than the non-IBD population, suggesting that ECM remodeling could contribute to creating a niche in humans that *C. difficile* can more readily colonize and thrive[3–7]. The gastrointestinal pathogens, *Salmonella enterica* and *Vibrio cholerae,* have also been shown to benefit from inflamed host tissue, gaining access to nutrients that increase their fitness as pathogens[13,14,55–57]. Future studies are needed to determine the extent to which MMP activity contributes to the peptides and amino acids that *C. difficile* has access to in the inflamed gut.

Lastly, we show that the community structure of the gut microbiota is altered by the presence of *C. difficile* and the activity of its toxins, supporting the hypothesis that inflammation can benefit *C. difficile* by selecting against competitors and for potential allies. In particular, we found that the Bacteroidaceae tended to bloom in the uninflamed ceca of antibiotic treated uninfected controls, or mice challenged with two *tcdR* mutants from phylogenetically divergent *C. difficile* strains[58]. In the latter case, the experiments were conducted over a year apart, suggesting that the inflammation induced by wild type *C. difficile* in our mouse model selects against a Bacteroidaceae population expansion. In support of this, negative associations between *C. difficile* and members of the Bacteroidaceae have been reported in human studies, as well as *in vitro*, and in mouse models of CDI[59–63]. Further work is necessary to identify which mediator(s) of host inflammation are responsible for the restriction of the Bacteroidaceae in our model of CDI. *B. thetaiotaomicron* has been shown to cross-feed *C. difficile* with succinate in a co-colonization model and our RNAseq studies support this, in that *tcdR* mice colonized with the Bacteroidaceae had increased expression of the succinate to butyrate operon. Interestingly, homologs of the *C. difficile* hydroxyproline dehydratase gene, *hypD,* are enriched in the Bacteroidaceae, and *C. difficile* can satisfy its proline requirements through utilization of hydroxyproline *in vitro*[51]. Given the high levels of hydroxyproline in collagen, it is possible that toxin-mediated exclusion or suppression of the Bacteroidaceae removes a competitor for a vital nutrient. Additional studies are needed to determine if members of this Family produce antagonistic effects against *C. difficile in vitro* and *in vivo*.

Use of a toxin null or *tcdR* mutant combined with omic technologies and *in vivo* models represents a powerful approach for asking how *C. difficile* toxin-induced inflammation alters the host gastrointestinal environment in ways that may create or preserve a niche during colonization and disease. These results provide multiple avenues for future study of the basic biology of CDI at the level of host response, pathogen response to inflammation, and

manipulation of the host gut microbiota. While it was outside the scope of this work, we think the approach of querying the gut microbiota of mice colonized with wild type and mutant strains of *C. difficile*, in particular mutants in key metabolic pathways, may be fruitful for identifying bacterial taxa that bloom in the presence of a nutrient(s) that a mutant *C. difficile* population can no longer use. This approach may, with enough mutants in important metabolic pathways, contribute to a rationally designed consortium of bacteria that could compete with *C. difficile* for essential nutrients in models of colonization and disease.

Data availability.

Raw sequences have been deposited in the Sequence Read Archive (SRA) with SRA accession number SUB6663505 and BioProject ID PRJNA612095.

Biological materials availability.

Available upon request.

Acknowledgements.

We thank Dr. Rita Tamayo for the isogenic *tcdR*::*ermB* mutant. The Microscopy Services Laboratory, Department of Pathology and Laboratory Medicine, is supported in part by P30 CA016086 Cancer Center Core Support Grant to the UNC Lineberger Comprehensive Cancer Center. J.R.F. and M.H.F. are supported by the University of North Carolina Center for Gastrointestinal Biology and Disease T32DK07737 postdoctoral fellowship and a North Carolina State University College of Veterinary Medicine intramural award. C.M.T. and C.M.P. are funded by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM119438.

Author contributions.

J.R.F., C.M.P., R.J.P., A.J.R., and M.H.F. performed the experiments. S.A.M. conducted blinded histology scoring, imaging, and analysis of murine cecal and colonic tissue. J.R.F, C.M.P., M.H.F., and C.M.T. designed the experiments and J.R.F., C.M.P., M.H.F., and M.R.M. analyzed

669 and interpreted the data. J.R.F., C.M.P., M.H.F., M.R.M., and C.M.T. wrote the paper. All authors
670 edited the manuscript.

671 **Competing interests.**

672 C.M.T. is an advisor of Locus Biosciences and also consults for Vedanta Biosciences, Inc. and
673 Summit Therapeutics.

674 **Figure legends.**

675 **Fig. 1. Inflammation is attenuated in *tcdR* mice in a mouse model of *C. difficile* infection.**
676 A) Schematic depicting experimental design. All mice (n=36) received the antibiotic
677 cefoperazone in their drinking water. Subsets were orally gavaged with wild type (n=12) or *tcdR*
678 (n=12) after antibiotic treatment. B) *C. difficile* vegetative cell CFUs in feces (n=6-8 per strain).
679 C) *C. difficile* spore CFUs in the feces (n=6-8 per strain). D) Toxin activity in the cecal content of
680 mice (n=4-6). E) Histopathological summary scores of the cecum. F) Representative images of
681 H&E stained ceca; scale bar, 500 μm. Kruskal-Wallis test with Dunn's correction for multiple
682 comparisons was used to test for statistical significance in B, C, and D. Geissner-Greenhouse
683 corrected ordinary Two-Way ANOVA with Tukey's multiple comparisons test was used in E. *p
684 < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

685 **Fig. 2. Metabolic gene expression in *C. difficile* is significantly altered by toxin-mediated**
686 **inflammation.** A) Gene set enrichment analysis of the differentially expressed genes *in vivo*
687 from wild type *C. difficile* relative to the *tcdR* mutant from both days 2 and 4. GO terms that had
688 transcripts with decreased levels are depicted in black bars and GO terms containing transcripts
689 with increased levels are shown as red bars. B) Heatmap of the log2 fold change of key operons
690 and transcripts that were differentially expressed in wild type *C. difficile* (n=5 on day 2, n=3 on
691 day 4) relative to *tcdR* (n=6 on day 2, n=3 on day 4). The labels of known CodY regulated
692 transcripts are color-coded in red if they increased in expression in a *codY* mutant and green if
693 they decreased.

Fig. 3. *C. difficile* induces expression of numerous transcripts associated with inflammation and ECM degradation. A) Heatmap of the top 50 differentially regulated transcripts (by adj. p value) in the ceca of uninfected controls, wild type mice, and *tcdR* mice (n=5-6 per treatment and time point). B) Gene set enrichment analysis of the differentially expressed genes in wild type mice relative to *tcdR* mice. C) Log2 fold changes of various *Mmp*s and associated transcripts from wild type vs. *tcdR* mouse ceca. Significance for all transcripts except *Mmp3* was determined using differential expression analysis within the NanoString nSolver Advanced analysis software. *Mmp3* expression levels were determined via qRT-PCR on cDNA generated from the same RNA used in the NanoString analysis.

Fig. 4. Toxin-mediated degradation of collagen supports *C. difficile* growth *in vitro*. A) Representative images of collagen (red) produced by IMR90 cells. Confluent cell monolayers were treated with 0.5 pM TcdA and TcdB and images were collected 12 hours later. Collagen was stained with a mix of antibodies against collagen types I, III and V in a 1:1:1 ratio; scale bar, 10 μm. B) Mean fluorescent intensity of Alexa Fluor 568 stained collagen produced by IMR90 cells cultured in the presence or absence of 0.5 pM TcdA and TcdB for 15 hours calculated using ImageJ software. Statistical significance was determined by Mann-Whitney rank-sum test. C) *C. difficile* was grown in complete CDMM, CDMM lacking proline, or CDMM lacking proline and supplemented with heat-degraded collagen. CFUs/ml were enumerated at 0 and 24 hours. D) Similar to the scheme in C, except that purified Pro-Gly or Gly-Pro dipeptides were added to CDMM lacking proline. Statistical significance was determined by two-way ANOVA with Sidak's multiple comparisons test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Fig. 5. *C. difficile* toxin activity suppresses the Bacteroidaceae that are able to compete with *C. difficile* for amino acids. A) Averaged percent relative abundance of Family-level ASVs in each treatment group per timepoint. ASVs with less than 1% relative abundance in all samples were not included. B) PCA biplot of 16S rRNA amplicon sequences derived from cecal tissue

719 from uninfected or no *C. diff* controls (n=5 on day 2, n=6 on day 4), and wild type mice (n=4 on
720 day 2, n=6 on day 4) or *tcdR* mice (n=5 on day 2, n=6 on day 4). Each colored symbol represents
721 an individual mouse's cecal microbiome, with circles being those from day 2 and triangles from
722 day 4. 99% of OTUs are shown as gray crosses; the 10 OTUs furthest from the origin are labeled
723 by the finest taxonomic rank identified (family, genus, or species). C) 16 hour fold change in
724 CFUs of *B. thetaiotaomicron* in minimal media with or without glucose, supplemented with either
725 proline or hydroxyproline. D) 16 hour fold change growth of *B. fragilis* in identical media
726 conditions as in C.

727 **Fig. 6. *C. difficile* R20291 toxin activity similarly shapes the host gut transcriptome and**
728 **microbiota community structure in mice.** A) Total *C. difficile* CFUs (vegetative and spores) in
729 feces over time (n=6-8 per strain). B) Fecal spore CFUs over time (n=6-8 per strain). C) Toxin
730 activity in the cecal content of R20291 or Δ*tcdR* mice, as assessed by the Vero cell cytotoxicity
731 assay (n=5 per strain on day 2, n=4 per strain on day 4). D) Log2 fold change of *Mmp* and *Timp*
732 transcripts (n=3 per strain). E) Averaged percent relative abundances of 16S rRNA amplicon
733 sequences from cecal tissue isolated at day 4 (n=5 per strain). Kruskal-Wallis test with Dunn's
734 correction for multiple comparisons was used to test for statistical significance. *p < 0.05, **p <
735 0.01, ****p < 0.0001.

736 **Fig. S1. Extended data from the mouse model of *C. difficile* infection.** A) Mouse weights
737 over the course of the experiments (n=12 on day 2, n=6 on day 4). B) Vegetative cell CFUs in
738 the cecal content on day 2 and 4 (n=6 per strain). C) Spore CFUs in cecal content from B (n=6
739 per strain). D) Histopathological summary scores of the colon. Geissner-Greenhouse corrected
740 ordinary Two-Way ANOVA with Tukey's multiple comparisons test was used to test for
741 significance. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

742 **Fig. S2. *C. difficile* transcriptome shifts over the course of infection.** A) Enriched GO terms
743 in the differentially expressed genes in wild type *C. difficile* and B) *tcdR* at day 4 (n=3 per strain)

744 relative to day 2 (n=5-6 per strain). GO terms for transcripts with decreased or increased levels
745 are shown in black and red bars, respectively.

746 **Fig. S3. Validation of differentially expressed *C. difficile* transcripts by qRT-PCR.** A-D)
747 Expression was quantified from cDNA generated from the same RNA that was used in the
748 RNAseq experiment, including samples from *in vitro C. difficile* cultures.

749 **Fig. S4. Volcano plots depicting log2 fold change in expression and the adjusted p-value**
750 **for mouse cecal tissue transcriptomes.** A and B) Gene expression in wild type ceca relative
751 to *tcdR* ceca at days 2 and 4, respectively. C and D) Gene expression in wild type ceca relative
752 to ceca from uninfected, cefoperazone treated controls at 2 and 4 days, respectively. E and F)
753 Gene expression at day 4 relative to day 2 in wild type ceca and *tcdR* ceca, respectively. The
754 number of genes increased in expression are in red font and decreased are in green font.

755 **Fig. S5. Shared and unique differentially expressed genes in ceca of wild type relative to**
756 ***tcdR* mice over time.** A) Transcripts that increased in expression. B) Transcripts that
757 decreased in expression. Venn diagrams were created using Venny 2.1
758 (https://bioinfogp.cnb.csic.es/tools/venny/).

759 **Fig. S6. Validation of transcripts from NanoString cecal tissue analysis by qRT-PCR.** A-E)
760 Expression was quantified in cDNA generated from the same cecal RNA that was used in the
761 NanoString analysis. F) Unscaled heatmap depicting the log2 fold change of each *Mmp* and
762 *Timp* in every differential expression analysis. The heatmap was created using the R package
763 pheatmap (https://cran.r-project.org/web/packages/pheatmap/index.html). A mixed effects
764 model with Tukey's multiple comparison's test was used to test for significance. *$p < 0.05$, **$p <$
765 0.01.

766

767

768

References

1. Lessa, F. C., Winston, L. G., McDonald, L. C. & Emerging Infections Program C. difficile Surveillance Team. Burden of Clostridium difficile infection in the United States. *N. Engl. J. Med.* **372**, 2369–2370 (2015).
2. Aktories, K., Schwan, C. & Jank, T. Clostridium difficile Toxin Biology. *Annu. Rev. Microbiol.* **71**, 281–307 (2017).
3. de Bruyn, M. *et al.* The molecular biology of matrix metalloproteinases and tissue inhibitors of metalloproteinases in inflammatory bowel diseases. *Crit. Rev. Biochem. Mol. Biol.* **51**, 295–358 (2016).
4. Jen, M.-H., Saxena, S., Bottle, A., Aylin, P. & Pollok, R. C. G. Increased health burden associated with Clostridium difficile diarrhoea in patients with inflammatory bowel disease. *Aliment. Pharmacol. Ther.* **33**, 1322–1331 (2011).
5. Binion, D. G. Clostridium difficile Infection in Patients with Inflammatory Bowel Disease. *Gastroenterol. Hepatol.* **8**, 615–617 (2012).
6. Murthy, S. K. *et al.* Impact of Clostridium difficile colitis on 5-year health outcomes in patients with ulcerative colitis. *Aliment. Pharmacol. Ther.* **36**, 1032–1039 (2012).
7. Rodemann, J. F., Dubberke, E. R., Reske, K. A., Seo, D. H. & Stone, C. D. Incidence of Clostridium difficile infection in inflammatory bowel disease. *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.* **5**, 339–344 (2007).
8. Fletcher, J. R., Erwin, S., Lanzas, C. & Theriot, C. M. Shifts in the Gut Metabolome and Clostridium difficile Transcriptome throughout Colonization and Infection in a Mouse Model. *mSphere* **3**, (2018).
9. Karasawa, T., Ikoma, S., Yamakawa, K. & Nakamura, S. A defined growth medium for Clostridium difficile. *Microbiol. Read. Engl.* **141 (Pt 2)**, 371–375 (1995).

10. Jackson, S., Calos, M., Myers, A. & Self, W. T. Analysis of proline reduction in the nosocomial pathogen Clostridium difficile. *J. Bacteriol.* **188**, 8487–8495 (2006).
11. Bouillaut, L., Self, W. T. & Sonenshein, A. L. Proline-dependent regulation of Clostridium difficile Stickland metabolism. *J. Bacteriol.* **195**, 844–854 (2013).
12. Hofmann, J. D. *et al.* Metabolic Reprogramming of Clostridioides difficile During the Stationary Phase With the Induction of Toxin Production. *Front. Microbiol.* **9**, 1970 (2018).
13. Winter, S. E. *et al.* Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* **467**, 426–429 (2010).
14. Rivera-Chávez, F. & Mekalanos, J. J. Cholera toxin promotes pathogen acquisition of host-derived nutrients. *Nature* **572**, 244–248 (2019).
15. Ng, Y. K. *et al.* Expanding the repertoire of gene tools for precise manipulation of the Clostridium difficile genome: allelic exchange using pyrE alleles. *PloS One* **8**, e56051 (2013).
16. Edwards, A. N. & McBride, S. M. Isolating and Purifying Clostridium difficile Spores. *Methods Mol. Biol. Clifton NJ* **1476**, 117–128 (2016).
17. Theriot, C. M. *et al.* Cefoperazone-treated mice as an experimental platform to assess differential virulence of Clostridium difficile strains. *Gut Microbes* **2**, 326–334 (2011).
18. Winston, J. A., Thanissery, R., Montgomery, S. A. & Theriot, C. M. Cefoperazone-treated Mouse Model of Clinically-relevant Clostridium difficile Strain R20291. *J. Vis. Exp. JoVE* (2016) doi:10.3791/54850.
19. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinforma. Oxf. Engl.* **28**, 1647–1649 (2012).
20. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

818 21. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded
819 probe pairs. *Nat. Biotechnol.* **26**, 317–325 (2008).
820 22. Blighe, K., Rana, S & Lewis, M. EnhancedVolcano: Publication-ready volcano plots with
821 enhanced colouring and labeling. https://github.com/kevinblighe/EnhancedVolcano.
822 23. L.V. Holdeman, Cato, E. D. & Moore, W. E. C. . *Anaerobe Laboratory Manual*. (Virginia
823 Polytechnic Institute and State University Anaerobe Laboratory, 1977).
824 24. Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness
825 and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* **4**,
826 447–457 (2008).
827 25. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development
828 of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence
829 data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120
830 (2013).
831 26. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data
832 science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
833 27. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon
834 data. *Nat. Methods* **13**, 581–583 (2016).
835 28. Yilmaz, P. *et al.* The SILVA and 'All-species Living Tree Project (LTP)' taxonomic
836 frameworks. *Nucleic Acids Res.* **42**, D643-648 (2014).
837 29. Thanissery, R. *et al.* Characterization of C. difficile strains isolated from companion animals
838 and the associated changes in the host fecal microbiota. *bioRxiv* (2019)
839 doi:https://doi.org/10.1101/822577.
840 30. Legendre, P. & Gallagher, E. Ecologically meaningful transformations for ordination of
841 species data. *Oecologia* 271–280 (2001) doi:https://doi.org/10.1007/s004420100716.

842 31. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis
843 and graphics of microbiome census data. *PloS One* **8**, e61217 (2013).
844 32. McMurdie, P. J. *biomformat: An interface package for the BIOM file format*.
845 33. Ransom, E. M., Kaus, G. M., Tran, P. M., Ellermeier, C. D. & Weiss, D. S. Multiple factors
846 contribute to bimodal toxin gene expression in Clostridioides (Clostridium) difficile. *Mol.*
847 *Microbiol.* **110**, 533–549 (2018).
848 34. Girinathan, B. P. *et al.* Effect of tcdR Mutation on Sporulation in the Epidemic Clostridium
849 difficile Strain R20291. *mSphere* **2**, (2017).
850 35. Jenior, M. L., Leslie, J. L., Young, V. B. & Schloss, P. D. Clostridium difficile Colonizes
851 Alternative Nutrient Niches during Infection across Distinct Murine Gut Microbiomes.
852 *mSystems* **2**, (2017).
853 36. Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related
854 protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol.*
855 *Biol. Rev. MMBR* **70**, 939–1031 (2006).
856 37. Brinsmade, S. R., Kleijn, R. J., Sauer, U. & Sonenshein, A. L. Regulation of CodY activity
857 through modulation of intracellular branched-chain amino acid pools. *J. Bacteriol.* **192**,
858 6357–6368 (2010).
859 38. Dineen, S. S., McBride, S. M. & Sonenshein, A. L. Integration of metabolism and virulence
860 by Clostridium difficile CodY. *J. Bacteriol.* **192**, 5350–5362 (2010).
861 39. Lowther, W. T. & Matthews, B. W. Metalloaminopeptidases: common functional themes in
862 disparate structural surroundings. *Chem. Rev.* **102**, 4581–4608 (2002).
863 40. Kitchener, R. L. & Grunden, A. M. Prolidase function in proline metabolism and its medical
864 and biotechnological applications. *J. Appl. Microbiol.* **113**, 233–247 (2012).

865 41. Selmer, T. & Andrei, P. I. p-Hydroxyphenylacetate decarboxylase from Clostridium difficile.
866 A novel glycyl radical enzyme catalysing the formation of p-cresol. *Eur. J. Biochem.* **268**,
867 1363–1372 (2001).
868 42. Dawson, L. F. *et al.* The analysis of para-cresol production and tolerance in Clostridium
869 difficile 027 and 012 strains. *BMC Microbiol.* **11**, 86 (2011).
870 43. Passmore, I. J. *et al.* Para-cresol production by Clostridium difficile affects microbial
871 diversity and membrane integrity of Gram-negative bacteria. *PLoS Pathog.* **14**, e1007191
872 (2018).
873 44. McBride, S. M. & Sonenshein, A. L. Identification of a genetic locus responsible for
874 antimicrobial peptide resistance in Clostridium difficile. *Infect. Immun.* **79**, 167–176 (2011).
875 45. Suárez, J. M., Edwards, A. N. & McBride, S. M. The Clostridium difficile cpr locus is
876 regulated by a noncontiguous two-component system in response to type A and B
877 lantibiotics. *J. Bacteriol.* **195**, 2621–2631 (2013).
878 46. Jenior, M. L., Leslie, J. L., Young, V. B. & Schloss, P. D. Clostridium difficile Alters the
879 Structure and Metabolism of Distinct Cecal Microbiomes during Initial Infection To Promote
880 Sustained Colonization. *mSphere* **3**, (2018).
881 47. Yamakawa, K., Kamiya, S., Meng, X. Q., Karasawa, T. & Nakamura, S. Toxin production
882 by Clostridium difficile in a defined medium with limited amino acids. *J. Med. Microbiol.* **41**,
883 319–323 (1994).
884 48. Neumann-Schaal, M., Hofmann, J. D., Will, S. E. & Schomburg, D. Time-resolved amino
885 acid uptake of Clostridium difficile 630Δerm and concomitant fermentation product and
886 toxin formation. *BMC Microbiol.* **15**, 281 (2015).
887 49. Riedel, T. *et al.* High metabolic versatility of different toxigenic and non-toxigenic
888 Clostridioides difficile isolates. *Int. J. Med. Microbiol. IJMM* **307**, 311–320 (2017).

889 50. Battaglioli, E. J. *et al.* Clostridioides difficile uses amino acids associated with gut microbial
890 dysbiosis in a subset of patients with diarrhea. *Sci. Transl. Med.* **10**, (2018).
891 51. Huang, Y. Y., Martínez-Del Campo, A. & Balskus, E. P. Anaerobic 4-hydroxyproline
892 utilization: Discovery of a new glycyl radical enzyme in the human gut microbiome
893 uncovers a widespread microbial metabolic activity. *Gut Microbes* **9**, 437–451 (2018).
894 52. McDonald, L. C. *et al.* Clinical Practice Guidelines for Clostridium difficile Infection in Adults
895 and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and
896 Society for Healthcare Epidemiology of America (SHEA). *Clin. Infect. Dis. Off. Publ. Infect.*
897 *Dis. Soc. Am.* **66**, 987–994 (2018).
898 53. Anjuwon-Foster, B. R. & Tamayo, R. A genetic switch controls the production of flagella
899 and toxins in Clostridium difficile. *PLoS Genet.* **13**, e1006701 (2017).
900 54. Levin, B. J. *et al.* A prominent glycyl radical enzyme in human gut microbiomes
901 metabolizes trans-4-hydroxy-l-proline. *Science* **355**, (2017).
902 55. Barman, M. *et al.* Enteric salmonellosis disrupts the microbial ecology of the murine
903 gastrointestinal tract. *Infect. Immun.* **76**, 907–915 (2008).
904 56. Sekirov, I. *et al.* Salmonella SPI-1-mediated neutrophil recruitment during enteric colitis is
905 associated with reduction and alteration in intestinal microbiota. *Gut Microbes* **1**, 30–41
906 (2010).
907 57. Gillis, C. C. *et al.* Dysbiosis-Associated Change in Host Metabolism Generates Lactate to
908 Support Salmonella Growth. *Cell Host Microbe* **23**, 54-64.e6 (2018).
909 58. He, M. *et al.* Evolutionary dynamics of Clostridium difficile over short and long time scales.
910 *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7527–7532 (2010).
911 59. Tvede, M. & Rask-Madsen, J. Bacteriotherapy for chronic relapsing Clostridium difficile
912 diarrhoea in six patients. *Lancet Lond. Engl.* **1**, 1156–1160 (1989).

913 60. Goldberg, E. *et al.* The correlation between Clostridium-difficile infection and human gut
914 concentrations of Bacteroidetes phylum and clostridial species. *Eur. J. Clin. Microbiol.*
915 *Infect. Dis. Off. Publ. Eur. Soc. Clin. Microbiol.* **33**, 377–383 (2014).
916 61. Anonye, B. O. *et al.* Probing Clostridium difficile Infection in Complex Human Gut Cellular
917 Models. *Front. Microbiol.* **10**, 879 (2019).
918 62. Deng, H. *et al.* Bacteroides fragilis Prevents Clostridium difficile Infection in a Mouse Model
919 by Restoring Gut Barrier and Microbiome Regulation. *Front. Microbiol.* **9**, 2976 (2018).
920 63. Ghimire, S. *et al.* Identification of Clostridioides difficile-Inhibiting Gut Commensals Using
921 Culturomics, Phenotyping, and Combinatorial Community Assembly. *mSystems* **5**, (2020).
922
923

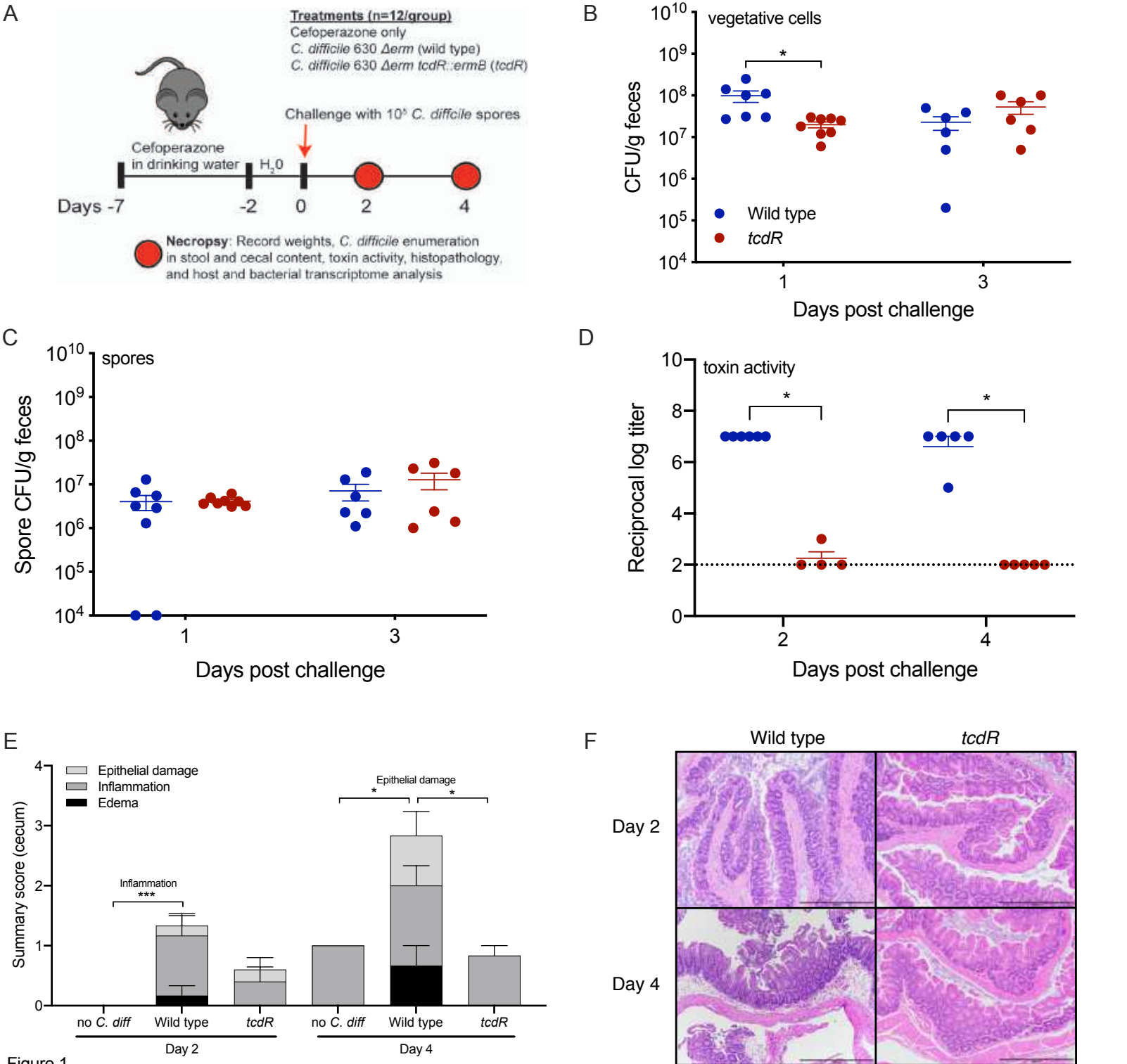
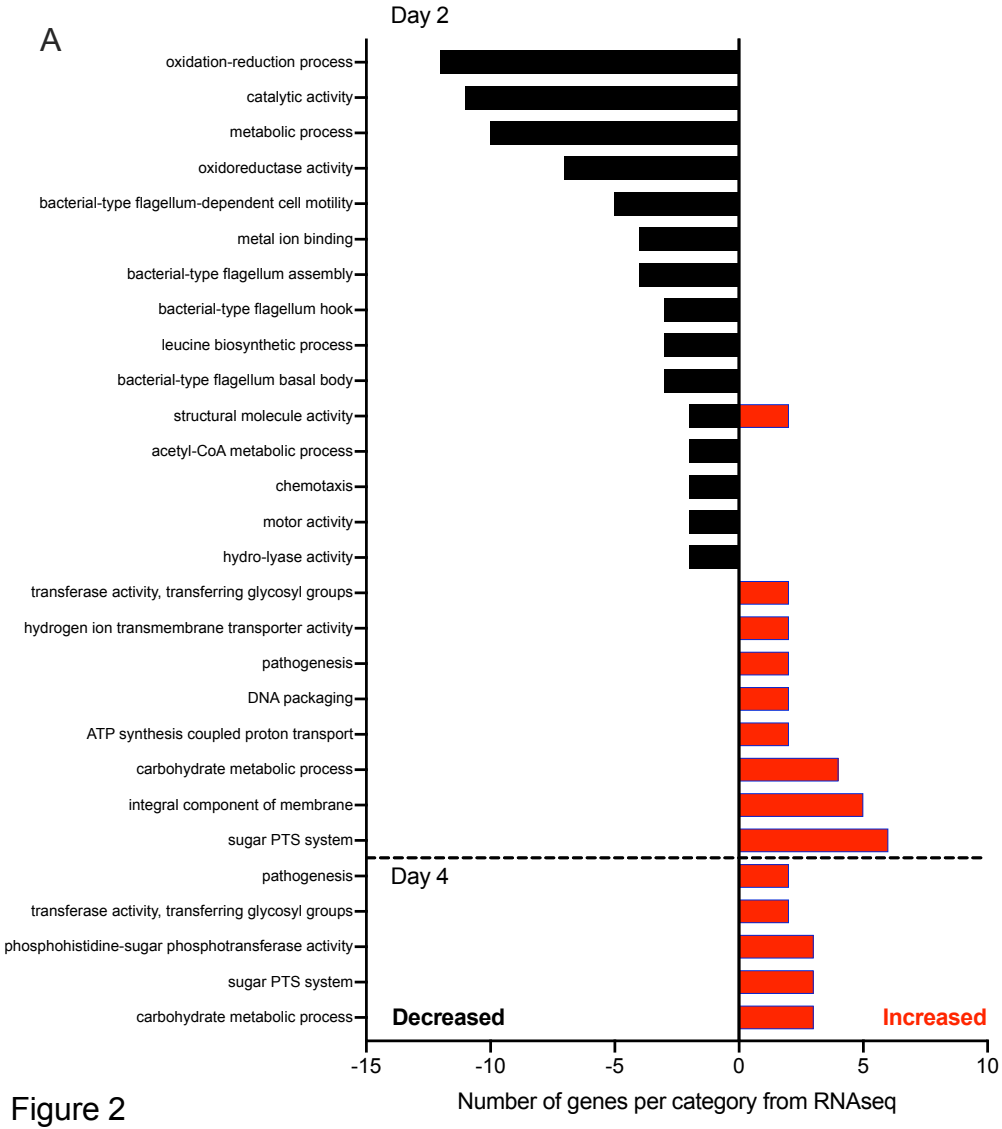

A
Treatments (n=12/group)
Cefoperazone only
C. difficile 630 Δerm (wild type)
C. difficile 630 Δerm tcdR::ermB (tcdR)
Challenge with 10⁵ C. diffcile spores
Cefoperazone in drinking water
H2O
Days -7
-2
0
2
4
Necropsy: Record weights, C. difficile enumeration in stool and cecal content, toxin activity, histopathology, and host and bacterial transcriptome analysis
B
vegetative cells
CFU/g feces
Wild type
tcdR
Days post challenge
C
spores
Spore CFU/g feces
Days post challenge
D
toxin activity
Reciprocal log titer
Days post challenge
E
Epithelial damage
Inflammation
Edema
Summary score (cecum)
Inflammation
***
Epithelial damage
*
*
no C. diff
Wild type
tcdR
Day 2
Day 4
F
Wild type
tcdR
Day 2
Day 4

Figure 1


A
Day 2
oxidation-reduction process
catalytic activity
metabolic process
oxidoreductase activity
bacterial-type flagellum-dependent cell motility
metal ion binding
bacterial-type flagellum assembly
bacterial-type flagellum hook
leucine biosynthetic process
bacterial-type flagellum basal body
structural molecule activity
acetyl-CoA metabolic process
chemotaxis
motor activity
hydro-lyase activity
transferase activity, transferring glycosyl groups
hydrogen ion transmembrane transporter activity
pathogenesis
DNA packaging
ATP synthesis coupled proton transport
carbohydrate metabolic process
integral component of membrane
sugar PTS system
Day 4
pathogenesis
transferase activity, transferring glycosyl groups
phosphohistidine-sugar phosphotransferase activity
sugar PTS system
carbohydrate metabolic process
Decreased
Increased
-15
-10
-5
0
5
10
Number of genes per category from RNAseq

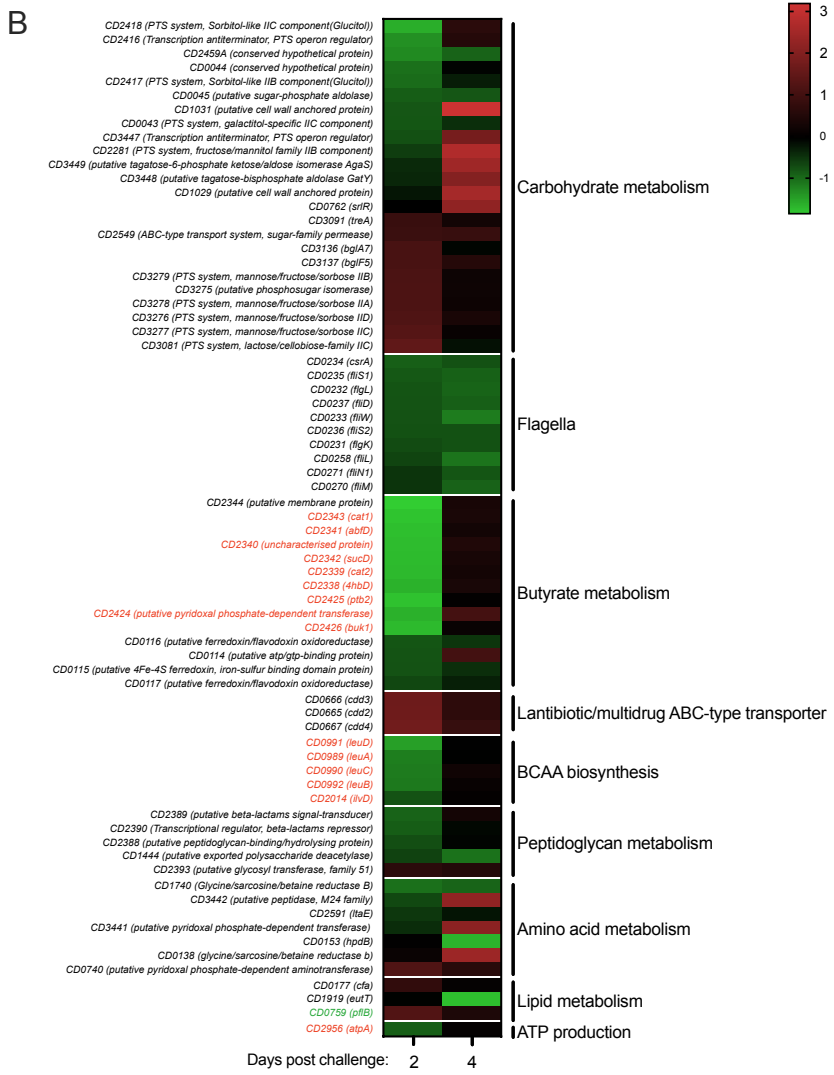

B
CD2418 (PTS system, Sorbitol-like IIC component(Glucitol))
CD2416 (Transcription antiterminator, PTS operon regulator)
CD2459A (conserved hypothetical protein)
CD0044 (conserved hypothetical protein)
CD2417 (PTS system, Sorbitol-like IIB component(Glucitol))
CD0045 (putative sugar-phosphate aldolase)
CD1031 (putative cell wall anchored protein)
CD0043 (PTS system, galactitol-specific IIC component)
CD3447 (Transcription antiterminator, PTS operon regulator)
CD2281 (PTS system, fructose/mannitol family IIB component)
CD3449 (putative tagatose-6-phosphate ketose/aldose isomerase AgaS)
CD3448 (putative tagatose-bisphosphate aldolase GatY)
CD1029 (putative cell wall anchored protein)
CD0762 (srlR)
CD3091 (treA)
CD2549 (ABC-type transport system, sugar-family permease)
CD3136 (bglA7)
CD3137 (bglF5)
CD3279 (PTS system, mannose/fructose/sorbose IIB)
CD3275 (putative phosphosugar isomerase)
CD3278 (PTS system, mannose/fructose/sorbose IIA)
CD3276 (PTS system, mannose/fructose/sorbose IID)
CD3277 (PTS system, mannose/fructose/sorbose IIC)
CD3081 (PTS system, lactose/cellobiose-family IIC)
Carbohydrate metabolism
CD0234 (csrA)
CD0235 (fliS1)
CD0232 (flgL)
CD0237 (fliD)
CD0233 (fliW)
CD0236 (fliS2)
CD0231 (flgK)
CD0258 (fliL)
CD0271 (fliN1)
CD0270 (fliM)
Flagella
CD2344 (putative membrane protein)
CD2343 (cat1)
CD2341 (abfD)
CD2340 (uncharacterised protein)
CD2342 (sucD)
CD2339 (cat2)
CD2338 (4hbD)
CD2425 (ptb2)
CD2424 (putative pyridoxal phosphate-dependent transferase)
CD2426 (buk1)
CD0116 (putative ferredoxin/flavodoxin oxidoreductase)
CD0114 (putative atp/gtp-binding protein)
CD0115 (putative 4Fe-4S ferredoxin, iron-sulfur binding domain protein)
CD0117 (putative ferredoxin/flavodoxin oxidoreductase)
Butyrate metabolism
CD0666 (cdd3)
CD0665 (cdd2)
CD0667 (cdd4)
Lantibiotic/multidrug ABC-type transporter
CD0991 (leuD)
CD0989 (leuA)
CD0990 (leuC)
CD0992 (leuB)
CD2014 (ilvD)
BCAA biosynthesis
CD2389 (putative beta-lactams signal-transducer)
CD2390 (Transcriptional regulator, beta-lactams repressor)
CD2388 (putative peptidoglycan-binding/hydrolysing protein)
CD1444 (putative exported polysaccharide deacetylase)
CD2393 (putative glycosyl transferase, family 51)
Peptidoglycan metabolism
CD1740 (Glycine/sarcosine/betaine reductase B)
CD3442 (putative peptidase, M24 family)
CD2591 (ltaE)
CD3441 (putative pyridoxal phosphate-dependent transferase)
CD0153 (hpdB)
CD0138 (glycine/sarcosine/betaine reductase b)
CD0740 (putative pyridoxal phosphate-dependent aminotransferase)
Amino acid metabolism
CD0177 (cfa)
CD1919 (eutT)
CD0759 (pflB)
Lipid metabolism
CD2956 (atpA)
ATP production
3
2
1
0
-1
Days post challenge:
2
4

Figure 2

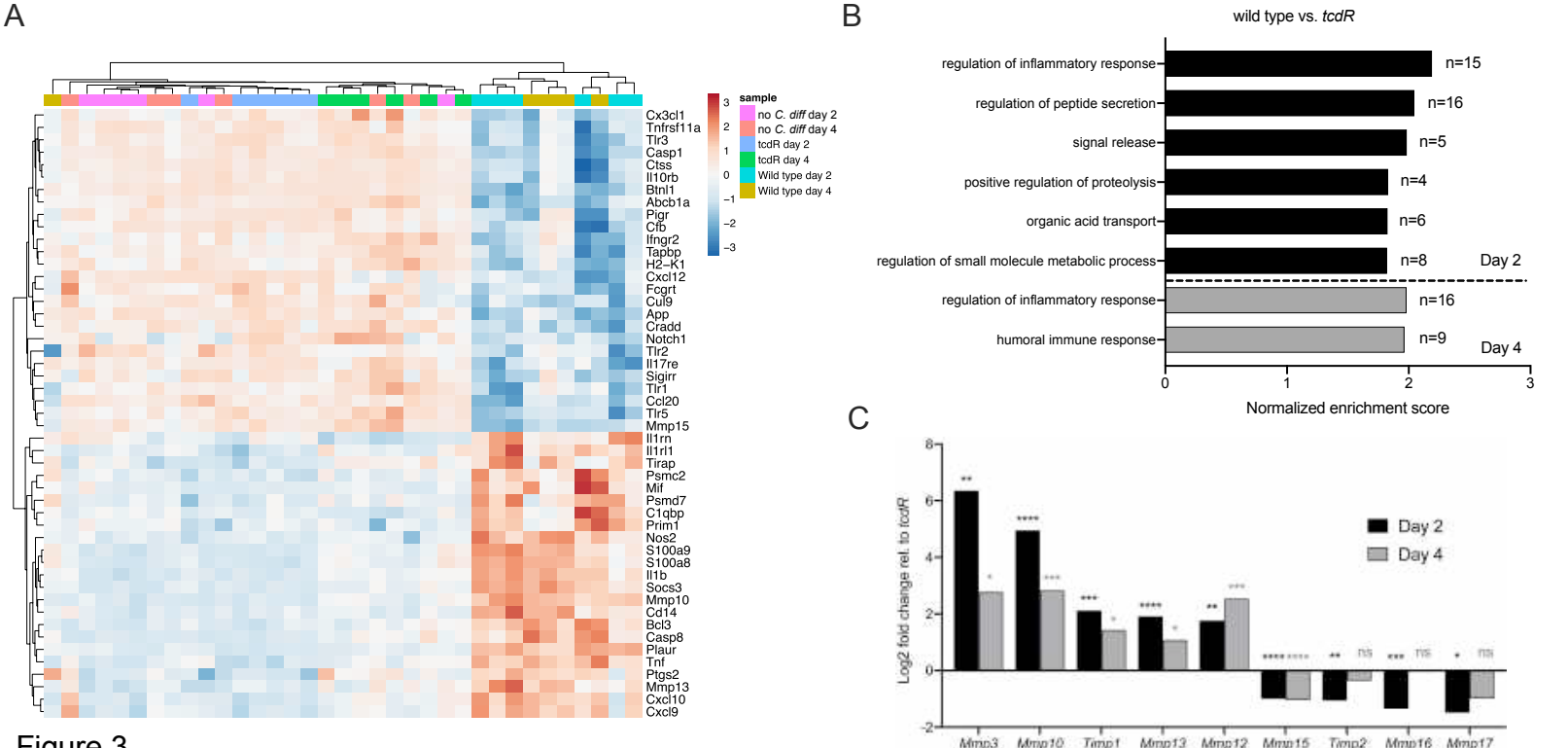

A
sample
no C. diff day 2
no C. diff day 4
tcdR day 2
tcdR day 4
Wild type day 2
Wild type day 4
Cx3cl1
Tnfrsf11a
Tlr3
Casp1
Ctss
Il10rb
Btnl1
Abcb1a
Pigr
Cfb
Ifngr2
Tapbp
H2−K1
Cxcl12
Fcgrt
Cul9
App
Cradd
Notch1
Tlr2
Il17re
Sigirr
Tlr1
Ccl20
Tlr5
Mmp15
Il1rn
Il1rl1
Tirap
Psmc2
Mif
Psmd7
C1qbp
Prim1
Nos2
S100a9
S100a8
Il1b
Socs3
Mmp10
Cd14
Bcl3
Casp8
Plaur
Tnf
Ptgs2
Mmp13
Cxcl10
Cxcl9
B
wild type vs. tcdR
regulation of inflammatory response
regulation of peptide secretion
signal release
positive regulation of proteolysis
organic acid transport
regulation of small molecule metabolic process
regulation of inflammatory response
humoral immune response
n=15
n=16
n=5
n=4
n=6
n=8
n=16
n=9
Day 2
Day 4
Normalized enrichment score
C
Log2 fold change rel. to tcdR
Day 2
Day 4
Mmp3
Mmp10
Timp1
Mmp13
Mmp12
Mmp15
Timp2
Mmp16
Mmp17

Figure 3

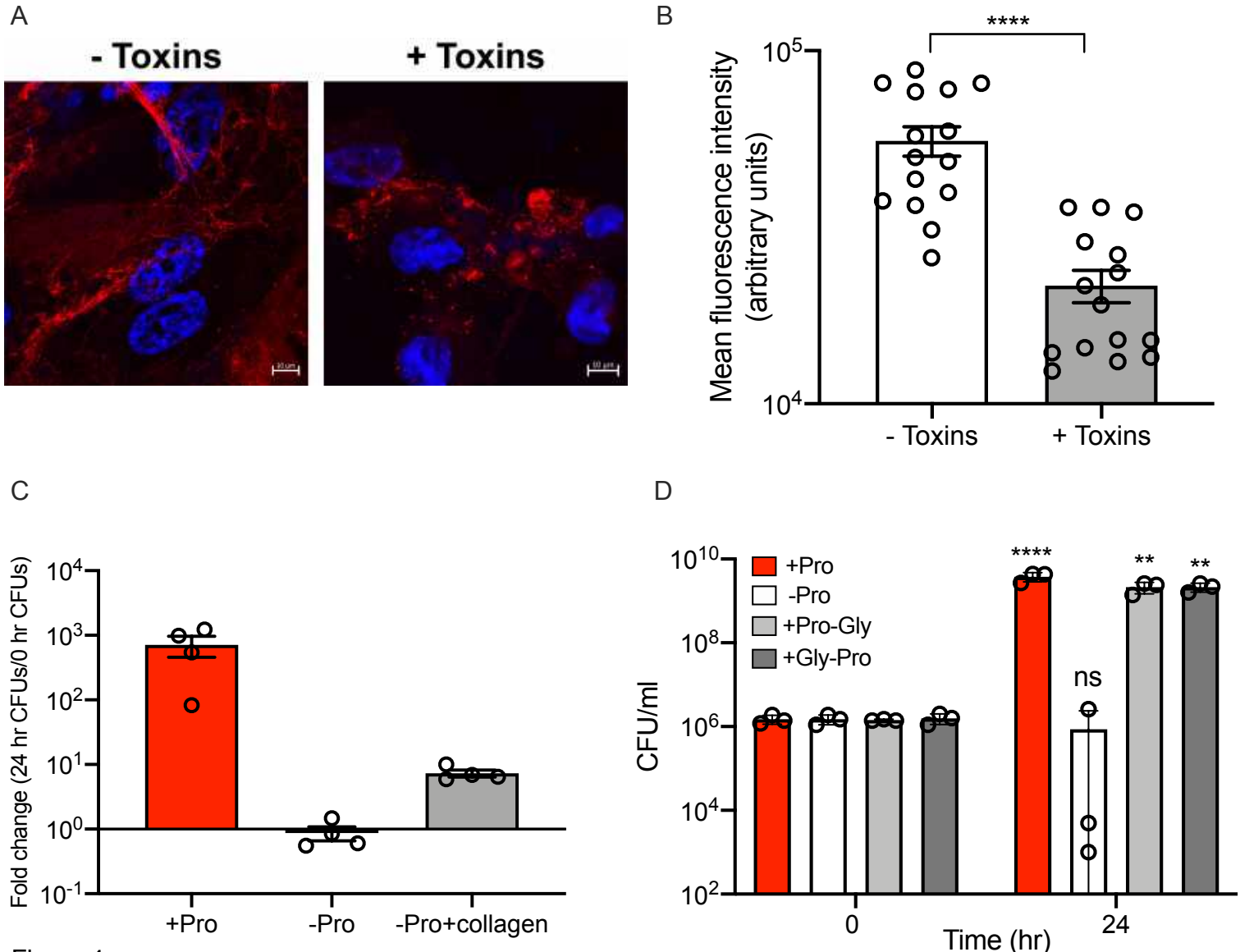

A
- Toxins
+ Toxins
B
****
Mean fluorescence intensity (arbitrary units)
10^5
10^4
- Toxins
+ Toxins
C
Fold change (24 hr CFUs/0 hr CFUs)
10^4
10^3
10^2
10^1
10^0
10^{-1}
+Pro
-Pro
-Pro+collagen
D
CFU/ml
10^{10}
10^8
10^6
10^4
10^2
+Pro
-Pro
+Pro-Gly
+Gly-Pro
****
ns
**
**
0
24
Time (hr)

Figure 4

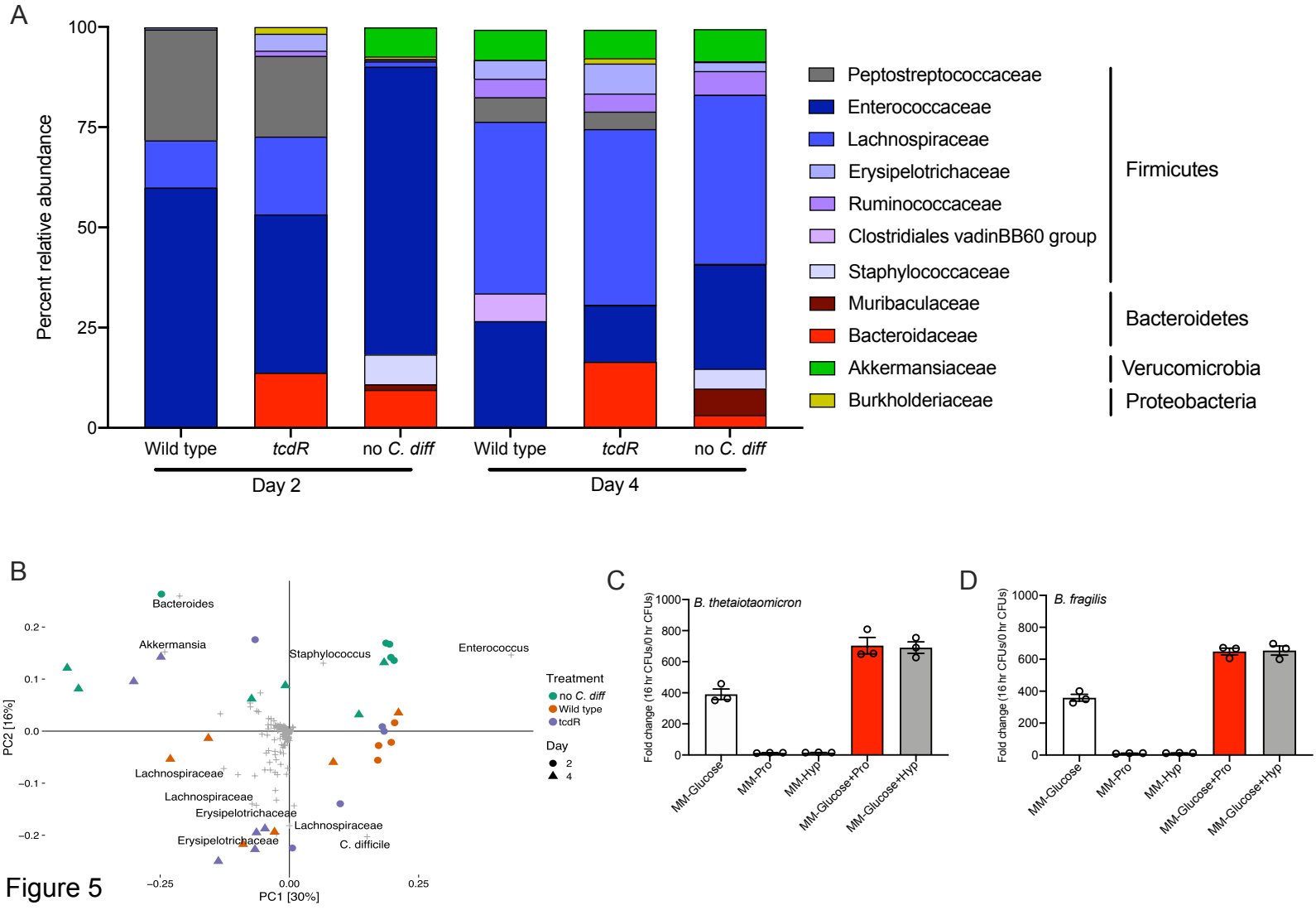

A
Percent relative abundance
100
75
50
25
0
Wild type
tcdR
no C. diff
Wild type
tcdR
no C. diff
Day 2
Day 4
Peptostreptococcaceae
Enterococcaceae
Lachnospiraceae
Erysipelotrichaceae
Ruminococcaceae
Clostridiales vadinBB60 group
Staphylococcaceae
Muribaculaceae
Bacteroidaceae
Akkermansiaceae
Burkholderiaceae
Firmicutes
Bacteroidetes
Verucomicrobia
Proteobacteria
B
Bacteroides
Akkermansia
Staphylococcus
Enterococcus
Lachnospiraceae
Lachnospiraceae
Erysipelotrichaceae
Lachnospiraceae
Erysipelotrichaceae
C. difficile
PC2 [16%]
PC1 [30%]
Treatment
no C. diff
Wild type
tcdR
Day
2
4
C
B. thetaiotaomicron
Fold change (16 hr CFUs/0 hr CFUs)
MM-Glucose
MM-Pro
MM-Hyp
MM-Glucose+Pro
MM-Glucose+Hyp
D
B. fragilis
Fold change (16 hr CFUs/0 hr CFUs)
MM-Glucose
MM-Pro
MM-Hyp
MM-Glucose+Pro
MM-Glucose+Hyp

Figure 5

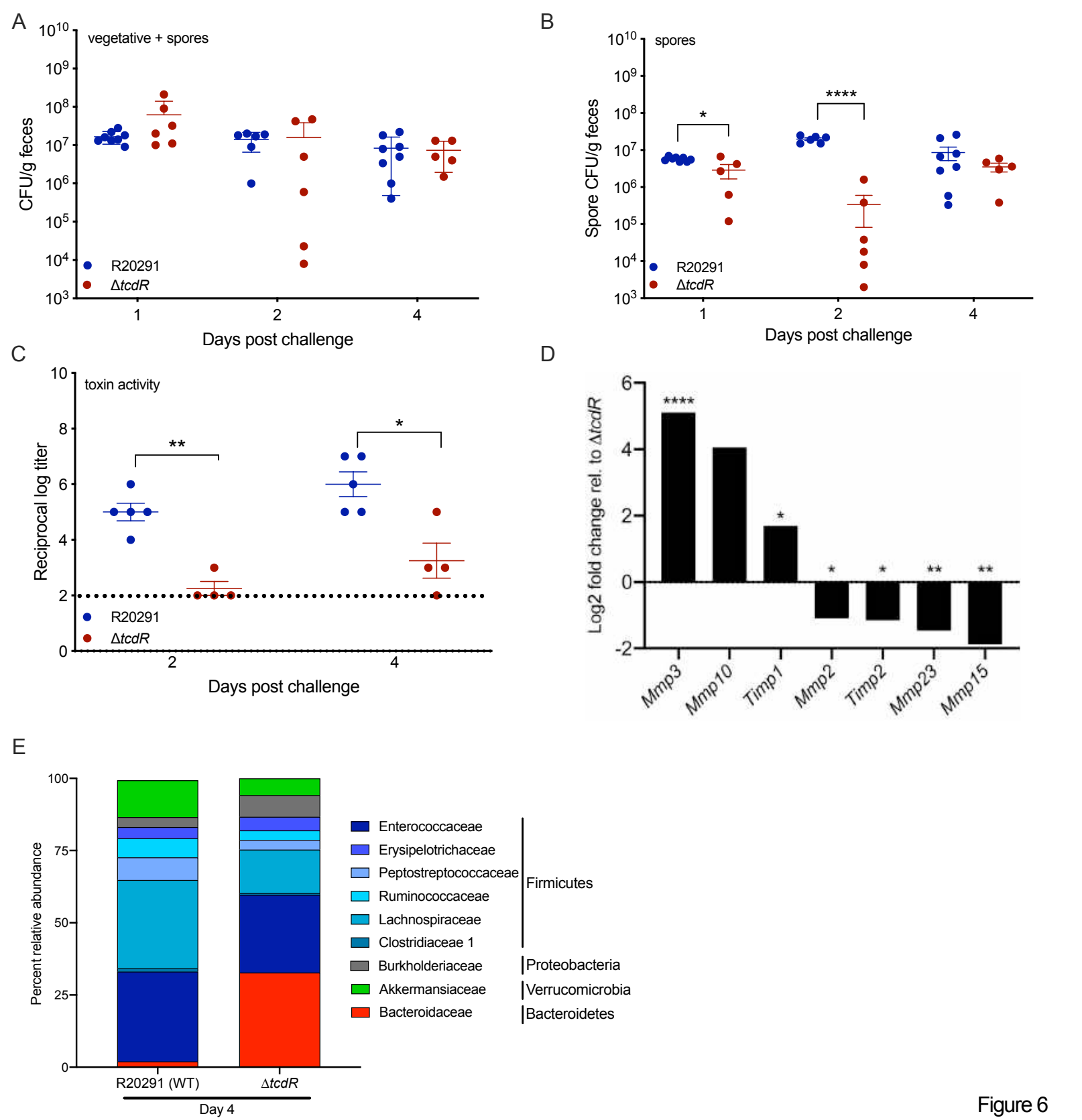

A
vegetative + spores
CFU/g feces
R20291
ΔtcdR
Days post challenge
B
spores
Spore CFU/g feces
*
****
R20291
ΔtcdR
Days post challenge
C
toxin activity
Reciprocal log titer
**
*
R20291
ΔtcdR
Days post challenge
D
Log2 fold change rel. to ΔtcdR
****
*
*
*
**
**
Mmp3
Mmp10
Timp1
Mmp2
Timp2
Mmp23
Mmp15
E
Percent relative abundance
R20291 (WT)
ΔtcdR
Day 4
Enterococcaceae
Erysipelotrichaceae
Peptostreptococcaceae
Ruminococcaceae
Lachnospiraceae
Clostridiaceae 1
Burkholderiaceae
Akkermansiaceae
Bacteroidaceae
Firmicutes
Proteobacteria
Verrucomicrobia
Bacteroidetes

Figure 6