SUPPLEMENTAL METHODS

Mouse model

Bone marrow was harvested from C57BL/6 or Luciferase$^+$ C57BL/6 (see '*In vivo* T cell imaging' below) donor mice following published protocols.[1] Bone marrow was depleted of T cells using anti-Thy1.2 monoclonal antibody (Pharmingen, San Diego CA) to strictly control T cell dose. Splenic T cells were purified using T cell enrichment columns (R&D Systems, Minneapolis MN) and anti-major histocompatibility complex (MHC) class II (Ia), CD11b, and DX5 microbeads (Miltenyi Biotec, Auburn, CA), following published protocols.[2]

1e7 T cell-depleted bone marrow cells and 1e6 purified splenic T cells were transplanted into BALB/c recipient mice on Day 0 (D0). Subsequent times are expressed relative to transplant day (D0) and refer to recipient mice; for example, D-2 occurs two days prior to transplant for recipient mice. Recipient mice were received from vendors on or before D-14 and allowed a minimum of one week to acclimate to facilities. Mice were maintained on PicoLab Rodent Diet 20 (LabDiet 5053) which did not contain any GOS, fed ad libitum. On D-7, recipient mice began receiving prebiotic GOS in drinking water at 4% estimated total caloric intake. This was achieved with a 4% w/v solution of GOS dissolved in drinking water. Based on patterns of chow and water consumption, mice are estimated to consume approximately 20 kcal/day and drink 4.5mL of water per day. Gavage was attempted but resulted in increased mortality in both cohorts; thus, consumption was likely variable across the population. GOS supplementation was maintained throughout the study. On D-3, D-2, and D-1, recipient mice received Imipenem-Cilastatin 500mg/500mg (NDC: 63323-322-25) at 50mg/kg body weight by gavage once daily. On D0, mice were subjected to 8.5Gy radiation (Cs-137) and then transplanted with a

suspension of donor T cell-depleted bone marrow and T cells. To supplement diet and reduce weight loss, mice received moistened food and gel water supplement (ClearH2O 70-01-5022) from D0 through D7. Mice were surveyed for GVHD score and body weight was measured approximately every 3 days following transplant. Stool samples were collected on D-7, D-3, D-1, D0, D2 or D3, D7, and D14. Endpoints were death, severe morbidity, or a loss of body weight in excess of 30% relative to D0, the latter two of which resulted in humane euthanasia. To ensure reproducibility, experiments were repeated in triplicate, with three mice per treatment group in experiments one and two, and six mice per treatment group in experiment three.

Assessment of GVHD

Presence of GVHD and its severity was calculated out of six categories: body weight, skin, fur, posture, and activity as previously published[3], as well as diarrhea. GVHD scores were assigned as detailed in Supplementary Table 1.

Histology

Tissue samples of skin from the back of the neck, large intestine, small intestine, and liver were collected in biopsy cassettes (Simport: M491) and stored in formalin (VWR: 89370-094) for a minimum of 24 hours before being embedded in paraffin. Samples were cut into 5 μm thick sections and stained with hematoxylin and eosin. Slides were analyzed by a single pathologist blinded to animal group or condition. For GI samples (large intestine and small intestine), tissues were scored with a semiquantitative system used to document the presence and severity of GVHD-associated features, as previously published.[5,6] The following features

were assessed for small and large intestine: villus blunting, crypt regeneration, ulceration, lamina propria inflammation, crypt loss, apoptosis, sloughing into the lumen, lymphocytic infiltrate, neutrophilic infiltrate, and edema. Each was graded as follows: 0 as normal, 0.5 as focal and rare, 1 as focal and mild, 2 as diffuse and mild, 3 as diffuse and moderate, and 4 as diffuse and severe. For skin samples, tissues were scored with a semiquantitative system for documenting GVHD damage with criteria as outlined in Supplementary Table 2.[7] For liver samples, tissues were graded using a semiquantitative scoring system as previously published[8], but no damage of relevance was observed so tissues were not fully evaluated (data not shown).

***In vitro* fermentations**

In vitro fermentations seeded with stool microbiota were carried out following previously published protocols for human stool fermentations [9], adapted for use with mouse stool. Fresh, never-frozen stool was obtained directly from the anus of mice at Day 0 prior to transplant, and immediately transferred to an anaerobic environment (COY anaerobic chamber). The stool was mixed with anaerobic 1X PBS (pH 7.0 +- 0.1) at a 10% weight to volume ratio. This mixture was allowed to rest at room temperature for 10 minutes to soften stool pellets, and was then gently vortexed for 60 seconds to dissociate the stool to form a slurry. The stool slurry was centrifuged at 1000g for 60 seconds to pellet large particles, then was mixed 1:1 with either a 1% w/v solution of galactooligosaccharides (Bimuno Powder) or a control solution of carbon-free PBS and incubated in duplicate in a 96-well cell culture plate for 24 hours in anaerobic conditions. The resulting fermentation conditions where therefore 5% fecal slurry with 0.5% prebiotic (w/v). Following fermentation, the 96-well plate was placed on

ice for five minutes to halt fermentation reactions, and then a 100μL aliquot was taken from each well for SCFA quantification. For further info on the development of these methods, see our previous publication [9].

DNA extraction, PCR amplification, and sequencing

We performed 16S rRNA gene amplicon sequencing on human stool samples to determine microbiota community composition following protocols previously published [9]. DNA was extracted from frozen fecal samples with the Qiagen DNeasy PowerSoil DNA extraction kit (ID 12888-100). Amplicon sequencing was performed using custom barcoded primers targeting the V4 region of the 16S gene[10], using published protocols.[10-12] The sequencing library was diluted to a 5pM concentration and sequenced using an Illumina MiniSeq and a MiniSeq Mid Output Kit (FC420-1004) with paired-end 150bp reads.

Identifying sequence variants and taxonomy assignment

We used an analysis pipeline with DADA2[13] to identify and quantify sequence variants, as previously published.[9,14] In brief, the following steps were performed: 1) 16S rRNA primer sequences were trimmed from paired sequencing reads using Trimmomatic v0.36 without quality filtering[15]; 2) reads were demultiplexed without quality filtering using python scripts provided with Qiime v1.9[16]; 3) reads were trimmed and quality filtered using fastqPairedFilter provided with the DADA2 R package (v1.8.0); 4) sequence variants were inferred by DADA2 independently for the forward and reverse reads using error profiles learned from ~1 million reads; 5) forward and reverse reads were merged and bimeras were removed using the

function removeBimeraDenovo with default settings; 6) taxonomy was assigned using the function assignTaxonomy from DADA2, trained using version 123 of the Silva database.

Statistical methods

Statistical analyses on parametric data were carried out using Student's t-tests or analysis of variance (ANOVA) followed by Tukey's Honestly Significant Difference post-hoc tests. Nonparametric murine survival data were analyzed using Kaplan-Meier and log-rank statistics. Microbiome sequencing data were compared with permutational multivariate analysis of variance (PERMANOVA) using weighted UniFrac distance, and Wilcoxon rank-sum tests adjusted for false discovery using Benjamini-Hochberg correction. Statistical significance was defined as $p < 0.05$. All statistical tests were conducted in R.

Random Forest Analysis

Taxonomic data was agglomerated to the family and genus level. Random forest models were trained using LOOCV with default hyperparameters using the caret package in R (v6.0-88).[17] To capture variation in the training process, 100 iterations of model training and evaluation were performed. Accuracy and Cohen's kappa metrics were compiled across training iterations to assess model performance. Feature importance scores across model iterations were averaged and ranked to determine the most important taxonomic features for model classifications.

1. Chen BJ, Cui X, Sempowski GD, et al. A comparison of murine T-cell-depleted adult bone marrow and full-term fetal blood cells in hematopoietic engraftment and immune reconstitution. *Blood.* 2002;99(1):364-371.
2. Chen BJ, Deoliveira D, Cui X, et al. Inability of memory T cells to induce graft-versus-host disease is a result of an abortive alloresponse. *Blood.* 2007;109(7):3115-3123.
3. Cooke KR, Kobzik L, Martin TR, et al. An experimental model of idiopathic pneumonia syndrome after bone marrow transplantation: I. The roles of minor H antigens and endotoxin. *Blood.* 1996;88(8):3230-3239.
4. Cao YA, Wagers AJ, Beilhack A, et al. Shifting foci of hematopoiesis during reconstitution from single stem cells. *Proc Natl Acad Sci U S A.* 2004;101(1):221-226.
5. Cooke KR, Hill GR, Crawford JM, et al. Tumor necrosis factor- alpha production to lipopolysaccharide stimulation by donor cells predicts the severity of experimental acute graft-versus-host disease. *The Journal of Clinical Investigation.* 1998;102(10):1882-1891.
6. Hill GR, Crawford JM, Cooke KR, Brinson YS, Pan L, Ferrara JLM. Total Body Irradiation and Acute Graft-Versus-Host Disease: The Role of Gastrointestinal Damage and Inflammatory Cytokines. *Blood.* 1997;90(8):3204-3213.
7. Kaplan DH, Anderson BE, McNiff JM, Jain D, Shlomchik MJ, Shlomchik WD. Target antigens determine graft-versus-host disease phenotype. *J Immunol.* 2004;173(9):5467-5475.
8. Gowdy KM, Cardona DM, Nugent JL, et al. Novel role for surfactant protein A in gastrointestinal graft-versus-host disease. *Journal of immunology (Baltimore, Md : 1950).* 2012;188(10):4897-4905.
9. Holmes ZC, Silverman JD, Dressman HK, et al. Short-Chain Fatty Acid Production by Gut Microbiota from Children with Obesity Differs According to Prebiotic Choice and Bacterial Community Composition. *mBio.* 2020;11(4).
10. Caporaso JG, Lauber CL, Walters WA, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences.* 2011;108(Supplement 1):4516.
11. Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The Isme Journal.* 2012;6:1621.
12. Maurice Corinne F, Haiser Henry J, Turnbaugh Peter J. Xenobiotics Shape the Physiology and Gene Expression of the Active Human Gut Microbiome. *Cell.* 2013;152(1):39-50.
13. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods.* 2016;13(7):581-583.
14. Silverman JD, Durand HK, Bloom RJ, Mukherjee S, David LA. Dynamic linear models guide design and analysis of microbiota studies within artificial human guts. *Microbiome.* 2018;6(1):202.
15. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 2010;7(5):335-336.
16. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *Plos One.* 2013;8(4):e61217.
17. *caret: Classification and Regression Training.* [computer program]. 2021.

SUPPLEMENTARY FIGURES

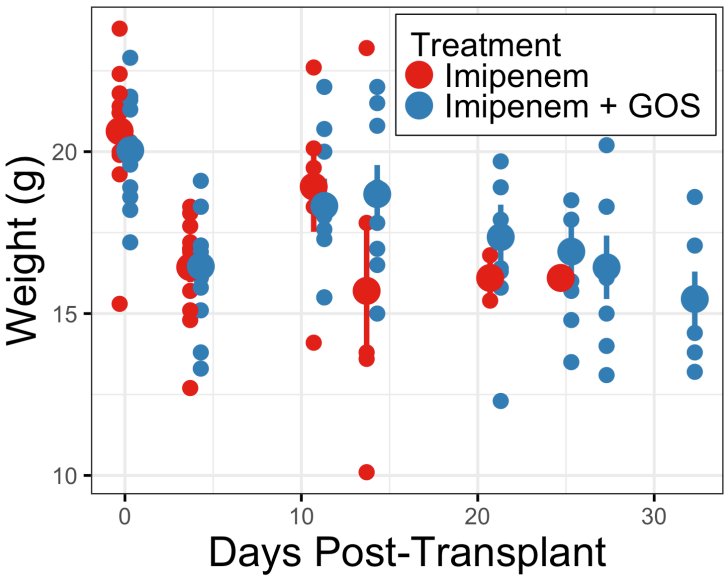

Figure S1: Body weights of GOS-supplemented and un-supplemented Imipenem-treated mice after undergoing BMT. Imipenem group is not represented in later timepoints because all mice failed to survive (Figure 1A). There were no differences between groups (n = 12 per group, repeated-measures ANOVA).

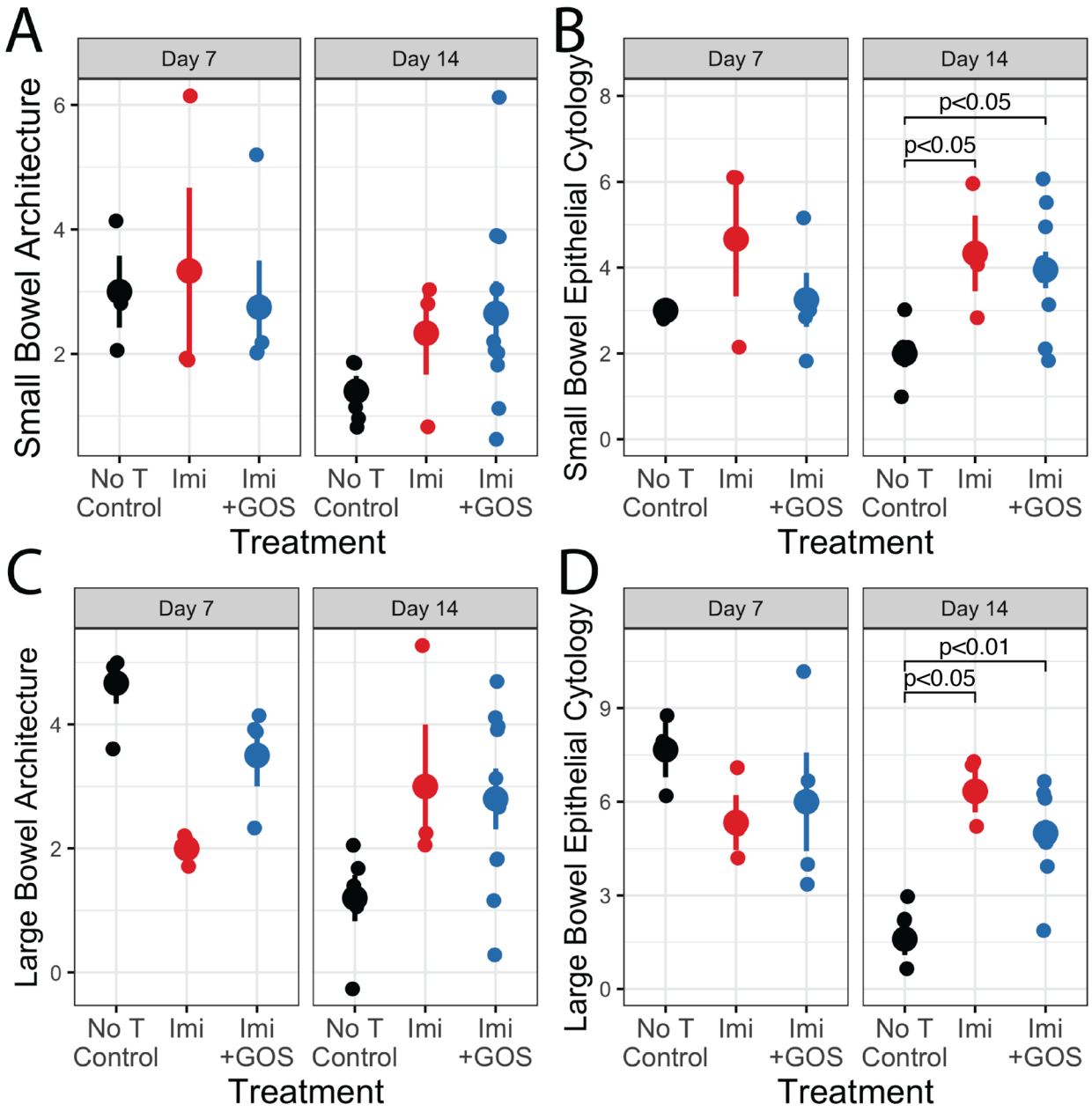

Figure S2: Histopathological analysis of GI tissue showing (A) small bowel architecture, (B) small bowel epithelial cytology, (C) large bowel architecture, and (D) large bowel epithelial cytology at day 7 (n = 3 in control group, 3 in Imipenem group, and 4 in Imipenem + GOS group) and day 14 (n = 5 in control group, 3 in Imipenem group, and 10 in Imipenem + GOS group) post-transplant.

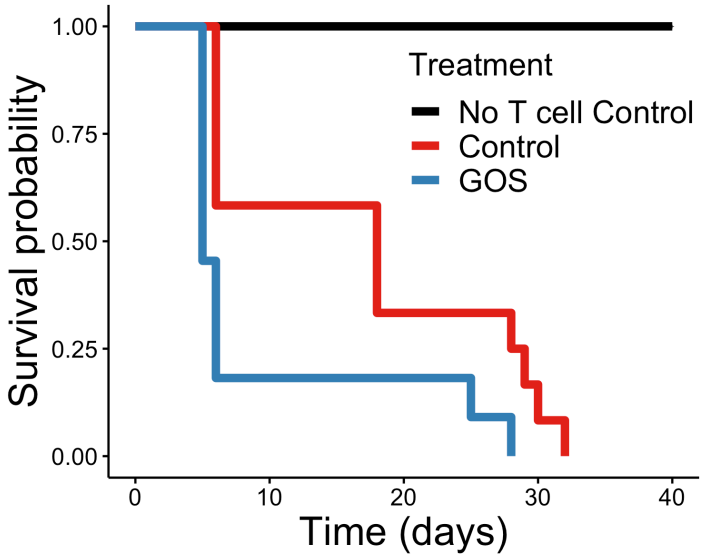

Figure S3: Survival of Taconic mice with and without GOS supplementation during transplant in a model that lacked pre-treatment with antibiotics. No significant difference was observed between groups (n = 12 per group, p = 0.26, log-rank test).

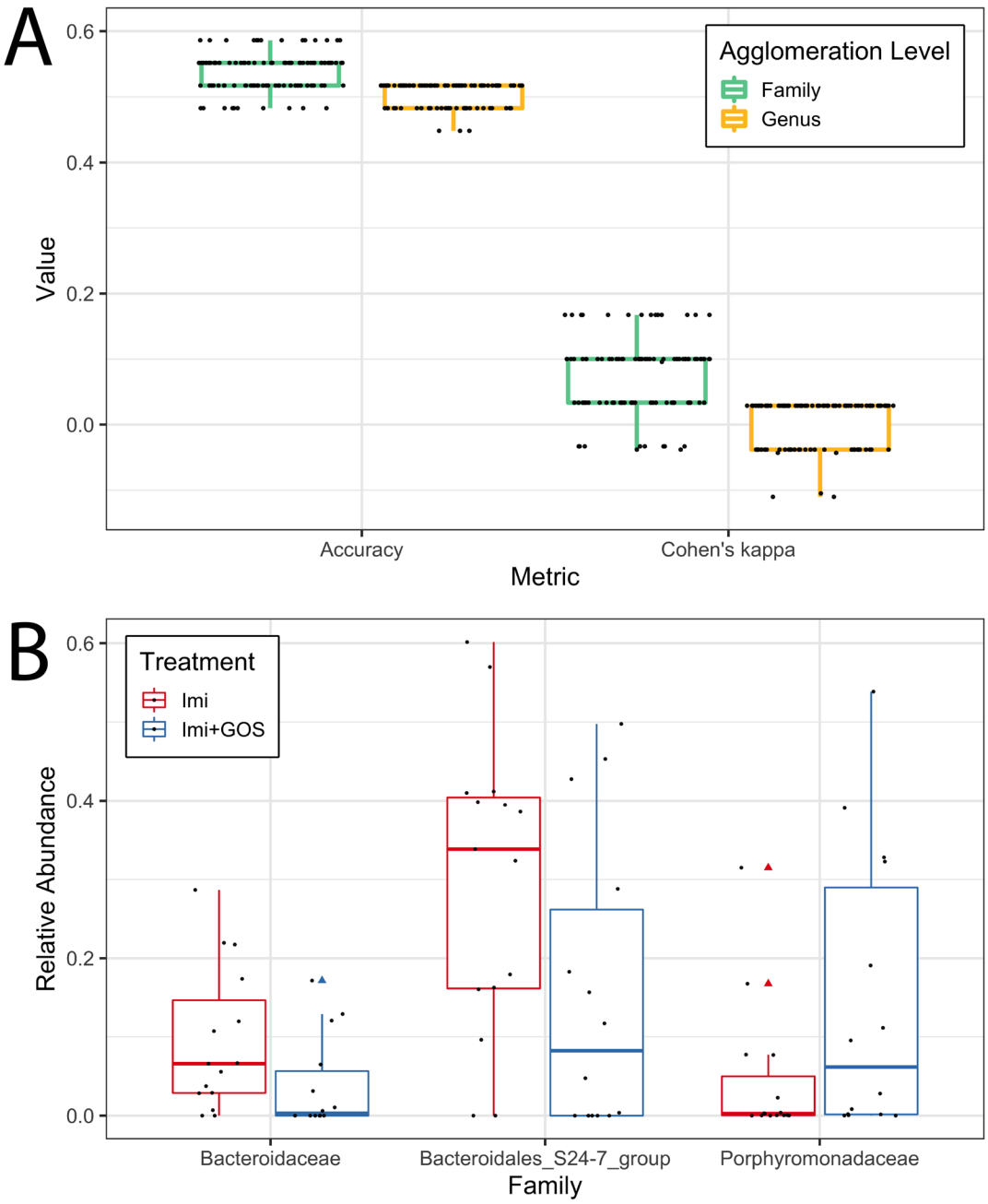

Figure S4: Random forest analysis predicting GOS-supplementation in Imipenem-treated mice following BMT. (A) Performance of random forest models classifying samples post-BMT as receiving GOS-supplementation or no supplementation. Models based on both family and genus level data exhibited median Cohen's kappa values greater than 0 (family: $p < 0.0001$, median $\kappa = 0.10$; genus: $p = 0.0008$, median $\kappa = 0.029$, indicating better agreement between

classifications and true classes than would be expected by random chance. Box plots display the accuracy and Cohen's kappa of model predictions across 100 iterations of leave-one-out cross-validation, with individual points representing values from single iterations. Data was agglomerated at the family and genus level. (B) The relative abundance of the top three most important features across random forest iterations using family-level agglomeration.

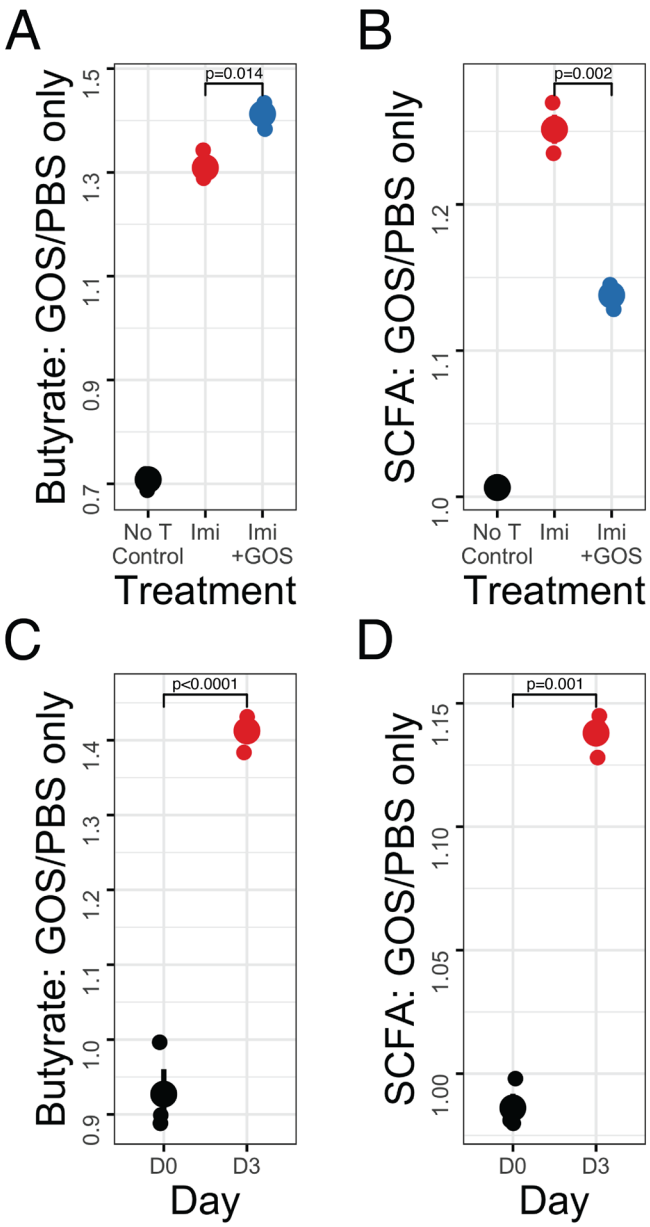

Figure S5: *In vitro* butyrate (A,C) and total SCFA (B,D) production capacity of mouse stool on Day 3 (A-B) post-transplant, and compared between Days 0 and 3 post-transplant (C-D). GOS supplementation resulted in a durable increase in butyrate production over un-supplemented mice at Day 3 ($p = 0.014$, t-test; A), but total SCFA production was lower in the GOS-supplemented group ($p = 0.002$, t-test; B). Both butyrate production and total SCFA production

in GOS-supplemented mice were higher on Day 3 compared to Day 0 ($p < 0.001$, $p = 0.001$, respectively; t-tests; C-D).

SUPPLEMENTARY TABLES

Attribute	Score	Criteria
Weight	0	less than 10% loss of body weight
	1	between 10% and 20% loss of body weight
	2	more than 20% loss of body weight
Skin	0	no scaling
	1	scaling of paws and tail (less than 25%)
	2	obvious areas of denuded skin (more than 25%)
Fur	0	no fur loss
	1	mild to moderate ruffling
	2	severe ruffling and poor grooming
Posture	0	normal
	1	hunching noted at rest
	2	severe hunching which impairs movement
Activity	0	normal
	1	mild to moderately decreased
	2	stationary unless stimulated
Diarrhea	0	none
	1	diarrhea present
	2	excessive diarrhea

Table 1: Scoring criteria for assessing GVHD clinical score.

Attribute	Score	Criteria
Dermis	0	normal
	1	mild increased collagen density
	2	marked increased collagen density
Hair folliccles	0	normal number of approximately 5 per linear millimeter
	1	between 1 and 5 follicles per linear millimeter
	2	<1 follicle per linear millimeter
Epidermis	0	normal
	1	interface damage in <20% of section with occasional necrotic keratinocytes
	2	widespread interface damage (>20% of section)
Inflammation	0	none
	1	focal infiltrates
	2	widespread infiltrates

Table 2: Scoring criteria for histological assessment of GVHD in skin.