Landes Bioscience www.landesbioscience.com

Supplemental Material to:

Maria Elisa Perez-Muñoz, Kirk Bergstrom, Vincent Peng, Robert Schmaltz, Roberto Jimenez-Cardona, Nathan Marsteller, Sam McGee, Thomas Clavel, Ruth Ley, Lijun Xia, and Daniel A Peterson

Discordance between changes in the gut microbiota and pathogenicity in a mouse model of spontaneous colitis

Gut Microbes 2014; 5(3)
http://dx.doi.org/10.4161/gmic.28622

http://www.landesbioscience.com/journals/gutmicrobes/article/28622/

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Selected Bacterial Species

All the bacterial species used in monoassociation experiments were grown in an anaerobic chamber with environmental conditions set as follows: 37°C, 5% CO_2, 5% H and 90% N_2. *Bacteroides* species included *Bacteriodes thetaiotaomicron* VPI 5482, *and Bacteroides sartorii* A-C2-0. Both *Bacteroides* were grown in TYG (per liter: 10g Bacto™ Tryptone (BD #211705), 5g Difco™ Yeast Extract (BD #212750), and 11 mM (2g) Glucose (Acros Organics #170080025), supplemented with 8.2mM (1g) L-cysteine (Sigma #C7352), 1ml hematin-histidine solution (12mg hematin (Sigma #H3281) dissolved in 10ml of 0.2M histidine (Sigma #H7875) pH8), 1mg vitamin K3 (Sigma #M5625), 1ml of 0.8% calcium chloride (MP Biomedicals #153502) solution, 2.6μM (0.4mg) ferrous sulfate (MP Biomedicals #0219466383), 100ml of 1M potassium phosphate buffer (500ml of 1M KH2PO4 (Sigma #P0662) with 1M K2HPO4 (Sigma #P3786) added until pH is 7.2), and 40ml of TYG salts solution (per liter: 2mM (0.5g) MgSO4 heptahydrate (Amresco #0662), 119mM (10g) NaHCO3 (Sigma #54019), 34.2mM (2g) NaCl (Fisher #BP358)).

Other bacterial species tested included: *Akkermansia muciniphila* BAA-835, grown in Difco™ Brain Heart Infusion Broth (BD #237500, prepared according to instructions); and *Lactobacillus johnsonii,* isolated from mouse fecal pellets and identified by 16s rRNA sequencing using the universal primers 8F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1391R (5'-GACGGGCGGTGWGTRCA-3'), and grown in modified Difco™ MRS (BD #288210, per liter: 55g MRS, 29.2mM (10g) of maltose (Acros Organics #329911000), and 28mM (5g) of fructose (Alfa Aesar #A17718)).

618 *Bacteroides*-colonized mice were euthanized at 30 days. Mice colonized with other
619 selected bacterial species were sacrificed at 20 days post-TM treatment.
620
621 **Sample preparation for Immunohistochemical and Immunofluorescence staining**
622 Samples were embedded in paraffin after fixing them in 10% formalin (Azer Scientific
623 #CUNBF-5-G) or Carnoy's fixative (60% methanol (Fisher #A412), 30% chloroform (Sigma
624 #C-2432), 10% glacial acetic acid (Fisher #A38-212)). Once paraffin embedded, they were left
625 to dry at room temperature (RT) overnight. Paraffin sections of 5µm thickness were
626 deparaffinized and rehydrated as follows: 3 washes of xylene for 3 minutes; 1 wash with
627 xylene/ethanol (50/50) for 3 minutes; 3 washes of 100% ethanol for 3 minutes each; 3 washes of
628 95% EtOH (3 mins/each); and one wash with dH_2O for 2 minutes.
629
630 **Fluorescence in situ hybridization (FISH) labeling**
631 Bacterial probes used for FISH are listed in Table S3. FISH probes (MWG/Operon
632 Technologies) were diluted with FISH Hybridization buffer (0.9M NaCl, 0.1M TRIS pH 7.2,
633 30% Formamide, 0.1%SDS) to a final concentration of 2.5ng/µl, and aliquoted onto each tissue
634 section (100µL per slide). Slides were incubated at 37°C overnight in a humidified chamber in
635 the dark, and then washed with pre-warmed (37°C) FISH hybridization buffer for 15min. Excess
636 hybridization buffer was aspirated, and tissue samples were washed with FISH wash buffer
637 (0.9M NaCl, 0.1M TRIS pH 7.2) for 15 minutes. Afterwards, slides were rinsed twice for 2
638 minutes with distilled water, and 4′-6′-diamidino-2-phenylindole (DAPI, 25 ng/mL) was applied
639 for 3 minutes as a nuclear counterstain. Two additional washes with distilled water were done,
640 and slides were mounted using PermaFluor (ThermoScientific #TA-030-FM).

Hematoxylin and Eosin staining

Briefly, starting with paraffin embedded section, the sections were washed 3 times with 100% xylene, once with 50:50 xylene and 100% ethanol, twice with 100% ethanol and twice with 95% ethanol. All washes were 3 minutes long. The sections were rinsed with distilled water for 2-3 minutes after the washes. Harleco® Gill II Modified Hematoxylin (EMD/Millipore, cat# 65066-85) was added and incubated for 5 minutes. Sections underwent 4 washes with tap water and immersed for 1 second in acid alcohol (1% HCl in 70% ethanol) and washed again 4 times with tap water. Next, sections were slowly immersed in ammonia water (1ml NH_4OH in 1 liter of water) 5 to 6 times until sections darken. The wash step was repeated and slides were immersed in Eosin (VWR, cat # 95057-848) for 1 minute. After a final rinse with tap water, 2 minute sequential rinses were done as follows: 3 times with 95% ethanol, 3 times with 100% ethanol, once with 50:50 xylene/100% ethanol, and 3 times with 100% xylene. Slides were then mounted with Permount (Fisher #SP15-500)). For Alcian blue staining, Carnoy's-fixed paraffin-embedded sections were deparffinized and rehydrated as above, then stained with 1% Alcian Blue (pH 2.5, Newcomer Supply #1003) for 20 minutes, rinsed with 5 washes of distilled water, and then counterstained with nuclear fast red (Newcomer Supply #1255) for 5 minutes. Sections were then washed 3 times and dehydrated and mounted as above. (Images were obtained using a Nikon Eclipse E600 microscope equipped with a Nikon DS-2Mv camera, using the software NIS Freeware 2.10.

Tn Antigen staining

The protocol for Tn staining was done as described by Fu et al.[1] Briefly, deparaffinized sections underwent biotin blocking using the Streptavidin-Biotin blocking kit (Vector Laboratories #SP2002) according to manufacturer's instructions. Sections were then incubated overnight (4 °C) with biotinylated anti-Tn antibody (mouse IgM, provided by Tongzhong Ju and Richard Cummings, Emory University, Atlanta, Georgia, USA) or with biotinylated isotype-matched control anti mouse IgM (BD cat# 553472), each at 2 µg/mL. Antibodies were biotinylated using the ProtON™ Biotin Labeling Kit (Vector Laboratories #PLK-1202). Bound antibodies were detected with horseradish peroxidase-conjugated streptavidin (Jackson Laboratories #016-030-084).

Lamina propria CD45$^+$ cells labeling

Lamina propria CD45$^+$ cell staining was performed as previously described.[2] Briefly, five µm-thick formalin or Carnoy's-fixed, paraffin-embedded tissue sections were deparaffinized with xylene and rehydrated through an ethanol gradient to water. Sections were blocked for 15 minutes using serum-free protein block (Dakom #X0909); incubated overnight at 4°C with rat-anti mouse CD45 (Abcam #ab25386, 5ug/mL) in primary antibody dilution buffer (1°ADB) (1% BSA, 0.1% Triton-x 100, 0.05% Tween-20 in PBS); and washed twice with 1°ADB, and twice with wash buffer (WB) (0.05% Tween-20 in PBS), 5 minutes/wash. Epifluorescent labeling for all stains was carried out using AlexaFluor 488-conjugated donkey anti-rat IgG (Invitrogen #A21208) in wash buffer, for 1 hr at RT in the dark, followed by rinsing in WB (3 x 5 mins). Samples were then fixed in 4% paraformaldehyde (PFA) in PBS for 5–10 mins, rinsed in WB (3 x 5 mins), counterstained with DAPI, washed 3 x 5 min in dH_2O, air-dried in the dark for 20 mins, and finally mounted with PermaFluor. Sections were viewed on a Nikon Eclipse 80*i*

686 microscope equipped with a Nikon DS-Qi1MC camera operating through NIS Elements AR
687 software (v3.0).
688
689 **DNA isolation for Pyrosequencing of 16S rRNA PCR**
690 Samples (FP, CC and LCC) from conventional mice and cecal contents from
691 conventionalized mice were weighted and diluted in phosphate buffer (pH 7) in a 1:10 ratio,
692 followed by centrifuging at 10,000 rpm (Eppendorf 5424) for 5 minutes. This washing step was
693 repeated two more times to remove inhibitors. Bacterial cell pellets were resuspended in 750µl of
694 lysis buffer (200mM NaCl, 20mM EDTA, 100mM Tris [pH 8.0], 20 mg/ml lysozyme (Fisher
695 #BP535)), transferred to a sterile microcentrifuge tube containing 300mg of 0.1mm zirconium
696 beads (BioSpec Products #11079101z), and incubated at 37°C for 30min. Afterwards, 85µl of
697 10% sodium dodecyl sulfate solution (Accugene #51213) and 40µl of proteinase K (Amresco
698 #0706, 15mg/ml) were added, and tubes were incubated at 60°C for 30min. After incubation,
699 500µl of phenol-chloroform-isoamyl alcohol (Acros Organics #327115000, 25:24:1) were added,
700 and samples were homogenized in a MiniBeadbeater-8 (BioSpec Products) at maximum speed
701 for 2 minutes, followed by cooling in ice. Layers were separated by centrifugation at 14,000rpm
702 for 5min before extraction of the top layer. This step was repeated three times with
703 phenol:chlorofom:isoamyl, and twice with chloroform-isoamyl alcohol (Acros Organics
704 #327155000, 24:1). DNA was recovered by inducing precipitation through the addition of 100%
705 ethanol and 3M sodium acetate. The amount of eluting reagents used depended on the final
706 volume of sample obtained after the last extraction step (2.5 times of ethanol per volume of
707 sample and 1/10 volume of sample of 3M sodium acetate). Samples were stored at -20°F
708 overnight followed by centrifugation at 14,000rpm (Eppendorf 5424) for 30min. Ethanol was

discarded, DNA pellets were allowed to dry at room temperature, and resuspended in 100µl of 10mM Tris-HCL buffer (pH8).

Primers used for Pyrosequencing of 16S rRNA

The primer sequences used for the analysis of conventional mice samples were as follows: 454_27F 5'-*GCCTTGCCAGCCCGCTCAGTC***AGTCAGAGTTTGATCCTGGCTCAG**-3' and 454_338R 5'-*GCCTCCCTCGCGCCATCAG*NNNNNNNNNNNNCA**TGCTGCCTCCCGTAGGAGT**-3'; where the italicized sequences are the 454 Life Sciences primers B and A, and the bold sequences are the broadly conserved bacterial primer 27F and the broad-range bacterial primer 338R, respectively. The NNNNNNNNNNNN region in the reverse primer represents the unique 12-base barcode used to tag each PCR product with a "CA" inserted as a linker between the barcode and the reverse primer. PCR conditions were: 5mins at 95°C; followed by 30 cycles of 45s at 95°C, 45s at 57°C, and 2min at 72°C; and lastly 10min at 72°C in an Eppendorf Mastercycler® thermocycler. PCR products were visualized by gel electrophoresis. Amplicons from individuals samples were mixed in equal amounts based on concentration, and purified using the QIAGEN MiniElute Gel extraction kit (QIAGEN #28604). The final purified product, with a concentration of 29.99ng/ml was submitted to emulsion PCR and sequenced using a Roche 454 FLX pyrosequencer at the Cornell University Life Sciences Core Laboratories Center. Sequences obtained from pyrosequencing were filtered and submitted to quality control parameters.

DNA extracted from the conventionalized mice was amplified and processed for pyrosequencing by the Johns Hopkins University Sequencing Core Facility using a Roche GS Junior System. The primer sequences used for these set of samples were as follows: 357F 5'-

*CCTATCCCCTGTGTGCCTTGGCAGTCTCAG***CCTACGGGAGGCAGCAG**-3' and 926R 5'-*CCATCTCATCCCTGCGTGTCTCCGAC*TCAGNNNNNNNNNN**CCGTCAATTCMTTTRAGT**-3'; where the italicized sequences are the 454 Life Sciences primers B and A, the bold sequences are the bacterial primer 357F and 926R, respectively, and the NNNNNNNNNN region, with a "TCAG" inserted as a linker between the barcode and the primer, represents the unique 10-base barcode used to tag each sample.

Selection and Analysis of OTUs

Taxonomic-independent methods were used to estimate species diversity in both datasets. Sequences were aligned and clustered using the *Aligner* and *Complete Linkage Clustering* algorithm available through the *Pyrosequencing pipeline* of the RDP. Clustering was done with a threshold of 97% pairwise identity for assignment into operational taxonomic units (OTUs). Initial clustering of the sequences in the *conventional* dataset by *Complete Linkage Clustering* produced a total of 3,641 OTUs. After selecting representative sequences for each OTU, a single best BLAST hit was identified for all sequences in the dataset using BioEdit Blastall function by blasting them against the dataset of non-chimeric sequences. OTUs over 300 bases long, over 97% identity, and containing more than 100 hits in total were kept for analysis. This filtering step produced a total of 891 OTUs. A new set of representative sequences was selected from the group of 891 OTUs, and the blasting step was repeated. Sequences that overlapped between OTUs were determined by comparing the sequences that comprise each OTU, and were merged together, resulting in 206 OTUs, from which representative sequences were selected and blasted against the dataset. Hits were counted using Excel. The filtering step was repeated as before producing a total of 184 unique OTUs. Proportions were calculated and significance of the OTUs

was determined by calculating the average mean of proportions and Z score.[3] A critical value of 3.29 (α=0.001) was selected to determine statistical significance using the following formula:

$$Z = \frac{(p_1 - p_2)}{\left\{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}^{1/2}}$$

$$\text{where } P = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$$

being n_1 and n_2 the random sizes of samples 1 and 2, and p_1 and p_2 their respective proportions.

Pyrosequencing of the *conventionalized* samples produced a total of 61,384 sequences. Eighteen percent were removed as chimeras, leaving a total of 56,527 sequences for analysis. Taxonomic and non-taxonomic based analyses were performed following the procedure described above (as for the conventional dataset). For OTUs, clustering of sequences for the conventionalized mice dataset produced a total of 709 OTUs. After filtering and removal or overlapping OTUs, OTUs were reduced to 340 accounting for 99% of the sequences.

Taxonomy was assigned for each OTU representative sequence by selecting the Best Blast result from a local type strain database using BioEdit Blast. Additionally, representative sequences from each OTU were aligned using PyNast, and used to construct a reference phylogenetic tree through QIIME. This phylogenetic tree was used for Unweighted UniFrac analysis.

SUPPLEMENTAL REFERENCES

1. Fu, J., Wei, B., Wen, T., Johansson, M.E.V., Liu, X., Bradford, E., Thomsson, K.A., McGee, S., Manssur, L., Tong, M. et al. (2011). Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. J Clin Invest 121:4:1657-66
2. Tobisawa, Y., Imai, Y., Fukuda, M., and Kawashima, H. Sulfation of colonic mucins by N-acetylglucosamine 6-O-sulfotransferase-2 and its protective function in experimental colitis in mice. J. Biol Chem (2010) 285:6750-60
3. Kanji, G.K. (2006) Z test for the equality of two proportions. In 100 Statistical Tests. SAGE Publications Ltd, London. 3rd edition, p27

Table S1. Mice average daily weight gain (ADG) and standard deviations (SD) per treatment and per genotype to assess development of colitis

Treatment	ADG WT ± SD (g)	ADG TM-IEC C1*galt*1$^{-/-}$ ± SD (g)	P-value
Germ-free 20d	0.14±0.07	0.12±0.03	0.4769
Germ-free 30d	0.11±0.03	0.10±0.03	0.5993
B. thetaiotaomicron	0.10±0.04	0.07±0.04	0.2474
B. sartorii	0.08±0.06	0.06±0.02	0.6371
A. muciniphila	0.11±0.02	0.13±0.06	0.5528
L. johnsonii	0.14±0.06	0.16±0.08	0.6850
Conventionalized	0.07±0.06	0.10±0.04	0.3511

Table S2. Statistically Significant Operational Taxonomic Units per anatomical site

OTU number	Representative sequence	cecal TM-IEC C1*galt*1-/-	cecal WT	cecal Z value	fecal TM-IEC C1*galt*1-/-	fecal WT	fecal Z value	LCC TM-IEC C1*galt*1-/-	LCC WT	LCC Z value	Percent (%) identity	SpeciesBest Blast Results Type strain database
1744	fecalwt4_100687	0.0307	0.0161	3.3937	0.0181	0.0210	-0.7301	0.0257	0.0233	0.6134	91	Algoriphagus halophilus
15434	LCCwt2_14989	0.0032	0.0136	-4.2406	0.0008	0.0038	-2.3511	0.0028	0.0298	-8.1531	91	Allobaculum stercoricanis
19122	cecalwt4_156667	0.0036	0.0056	-1.0974	0.0013	0.0050	-2.3880	0.0009	0.0073	-3.8169	96	Anaerotruncus colihominis
79	cecalwt4_233197	0.0526	0.0463	1.0294	0.0348	0.0755	-6.3494	0.0724	0.0769	-0.6776	92	Bacteroides fragilis
17587	cecalko1_43511	0.0076	0.0080	-0.1612	0.0069	0.0172	-3.3789	0.0068	0.0062	0.3055	91	Bacteroides tectus
889	LCCko4_107859	0.0066	0.0180	-3.8147	0.0108	0.0263	-4.1128	0.0128	0.0119	0.3148	92	Barnesiella viscericola
3	cecalwt3_8802	0.0037	0.0180	-5.1414	0.0063	0.0040	1.0980	0.0025	0.0031	-0.4158	91	Blautia hydrogenotrophica
7142	cecalwt4_225040	0.0004	0.0077	-4.4140	0.0005	0.0020	-1.5281	0.0002	0.0031	-2.6719	91	Clostridium aldenense
165	cecalko4_30502	0.0077	0.0001	4.1634	0.0147	0.0000	5.5299	0.0108	0.0002	6.0149	97	Clostridium disporicum
9022	cecalwt1_107537	0.0059	0.0181	-4.1329	0.0060	0.0046	0.6411	0.0041	0.0094	-2.5225	90	Clostridium jejuense
6028	fecalko1_63023	0.0678	0.0736	-0.8158	0.1059	0.1756	-7.0355	0.0935	0.1266	-4.1777	97	Eubacterium biforme
770	LCCwt1_102773	0.0004	0.0038	-2.8031	0.0011	0.0108	-4.6699	0.0006	0.0061	-3.6155	94	Eubacterium biforme
871	fecalko3_101523	0.0332	0.0072	6.3751	0.0476	0.0168	5.8349	0.0377	0.0129	6.4325	99	Lactobacillus intestinalis
10920	LCCko1_17826	0.1689	0.0879	8.5389	0.3361	0.1399	15.6044	0.2579	0.1360	12.3351	99	Lactobacillus johnsonii
4363	cecalko3_75532	0.0250	0.0201	1.1914	0.0093	0.0251	-4.3586	0.0201	0.0296	-2.3904	94	Parabacteroides distasonis
11	cecalwt3_15009	0.0036	0.0094	-2.6655	0.0020	0.0090	-3.4410	0.0045	0.0134	-3.6872	94	Parabacteroides distasonis
8594	fecalwt2_204110	0.0144	0.0255	-2.8946	0.0052	0.0196	-4.7119	0.0168	0.0284	-3.0631	96	Parabacteroides merdae
161	LCCko3_63397	0.0058	0.0019	2.1982	0.0085	0.0011	3.4978	0.0073	0.0032	2.3174	91	Parasutterella excrementihominis
12	fecalko2_41438	0.0436	0.1098	-9.1405	0.0627	0.0709	-1.1368	0.0516	0.0536	-0.3486	86	Prevotella loescheii

Values represent Weighted Means of Proportions and Z values. Critical value = 3.29 (α=0.001)

Table S3. Bacterial probes used for FISH analysis

Probe	Target	Sequence	5' conjugation
EUB338	Universal bacterial probe	5'-GCT GCC TCC CGT AGG AGT-3'	Texas Red
CFB286	Chlorobi/ Fibrobacteres Bacteroidetes Superphylum	5'-TCC TCT CAG AAC CCC TAC-3'	Alexa Fluor 488
MUC1437	*Akkermansia muciniphila*	5'-CCT TGC GGT TGG CTT CAG AT -3'	Alexa Fluor 488
NON338	Nonsense/non-specific probe (negative control)	5'- ACT CCT ACG GGA GGC AGC-3'	Texas Red
GAM42a	Gamma Proteobacter Phylum	5'- GCC TTC CCA CAT CGT TT-3'	Alexa Fluor 488

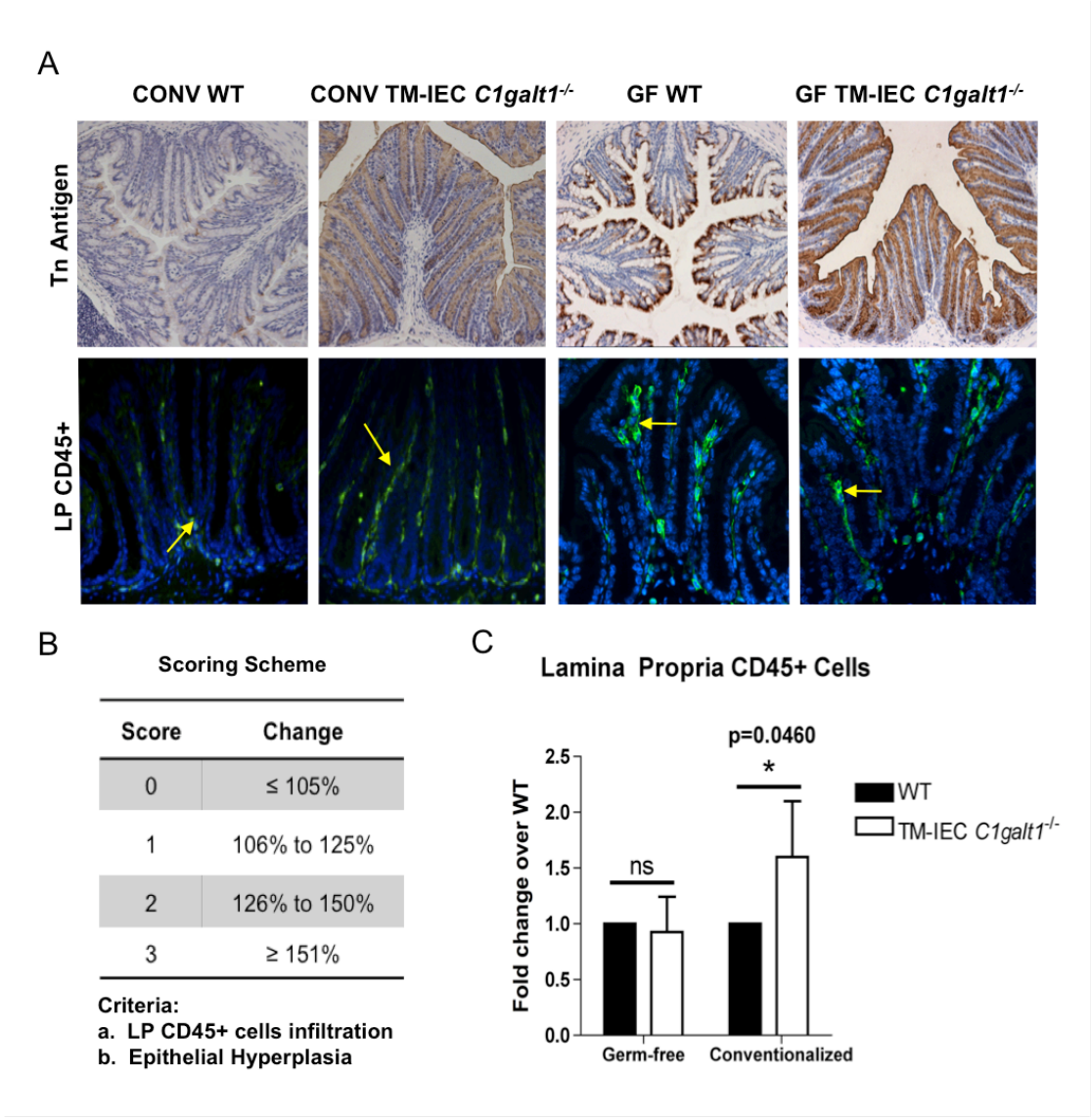

Figure S1. **Increases in Tn antigen expression and infiltration of lamina propria CD45+ cells were observed in the conventionalized TM IEC *C1galt1*$^{-/-}$ mice.** [A] Tn antigen staining and CD45^{+} cells staining showed increases in both Tn antigen (stained brown) and infiltration of LP CD45^{+} cells (green dots, pointed by yellow arrows) in the colon of TM IEC *C1galt1*$^{-/-}$ mice relative to the WT. [B] Parameters used to determine histological scoring for assessment of inflammation in the colon (infiltration of CD45^{+} cells in the tissue and epithelial hyperplasia). [C] The number of CD45^{+} cells in the conventionalized TM-IEC *C1galt1*$^{-/-}$ mice doubled, possibly due to an increased exposure of the microbiota to the intestinal epithelia caused by a breach in the mucus layer. Bars represent fold change over WT; error bars represent standard deviations. P-values were estimated for changes between genotypes (WT *versus* TM-IEC *C1galt1*$^{-/-}$).

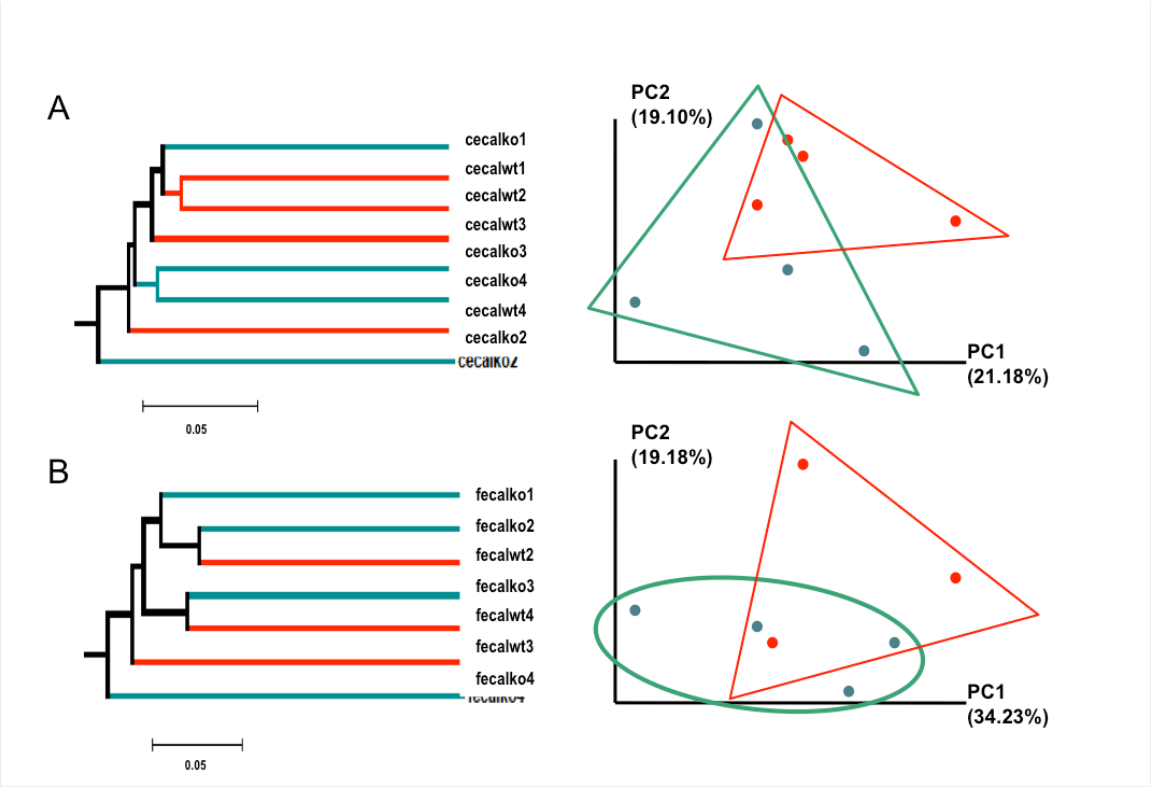

Figure S2. The microbial communities from cecal contents and fecal pellets of WT and TM-IEC *C1galt1*$^{-/-}$ mice share similarities in composition. Unlike LCC samples, where two distinct microbial communities were found between the genotypes, the microbial communities of the cecal contents and the fecal pellets between the WT and the TM-IEC *C1galt1*$^{-/-}$ mice share similarities in composition. KO: TM-IEC *C1galt1*$^{-/-}$ mice. Red dots and lines represent the WT mice. Green dots and lines represent the TM-IEC *C1galt1*$^{-/-}$ mice.

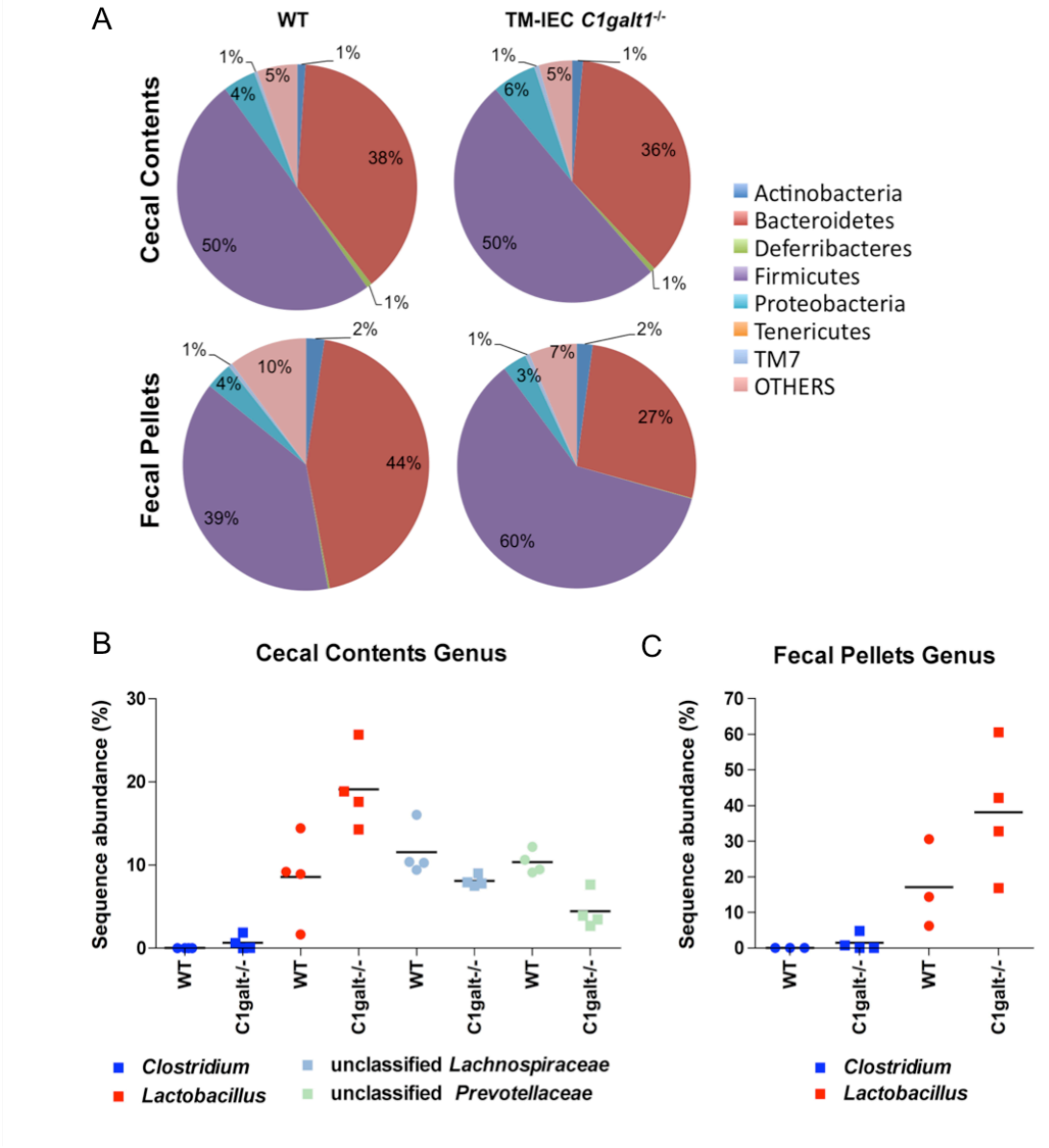

Figure S3. Lack of core 1-derived *O*-glycans caused increases in Firmicutes and decreases in Bacteroidetes. [A] Pie charts represent the abundance (in percentage) of bacterial phyla present per genotype. An increase of 21% in Firmicutes and a relative decrease of 17% in *Bacteroidetes* were found in the fecal population of the TM-IEC *C1galt1*$^{-/-}$ mice relative to the WT. However, these changes were not observed in the contents of the cecum. [B-C] Anatomical site-specific microbial decreases at the genus level were found in the TM-IEC *C1galt1*$^{-/-}$ mice. Statistically significant decreases in unclassified *Prevotellaceae* were found exclusively in the cecal contents. Horizontal bars represent means.

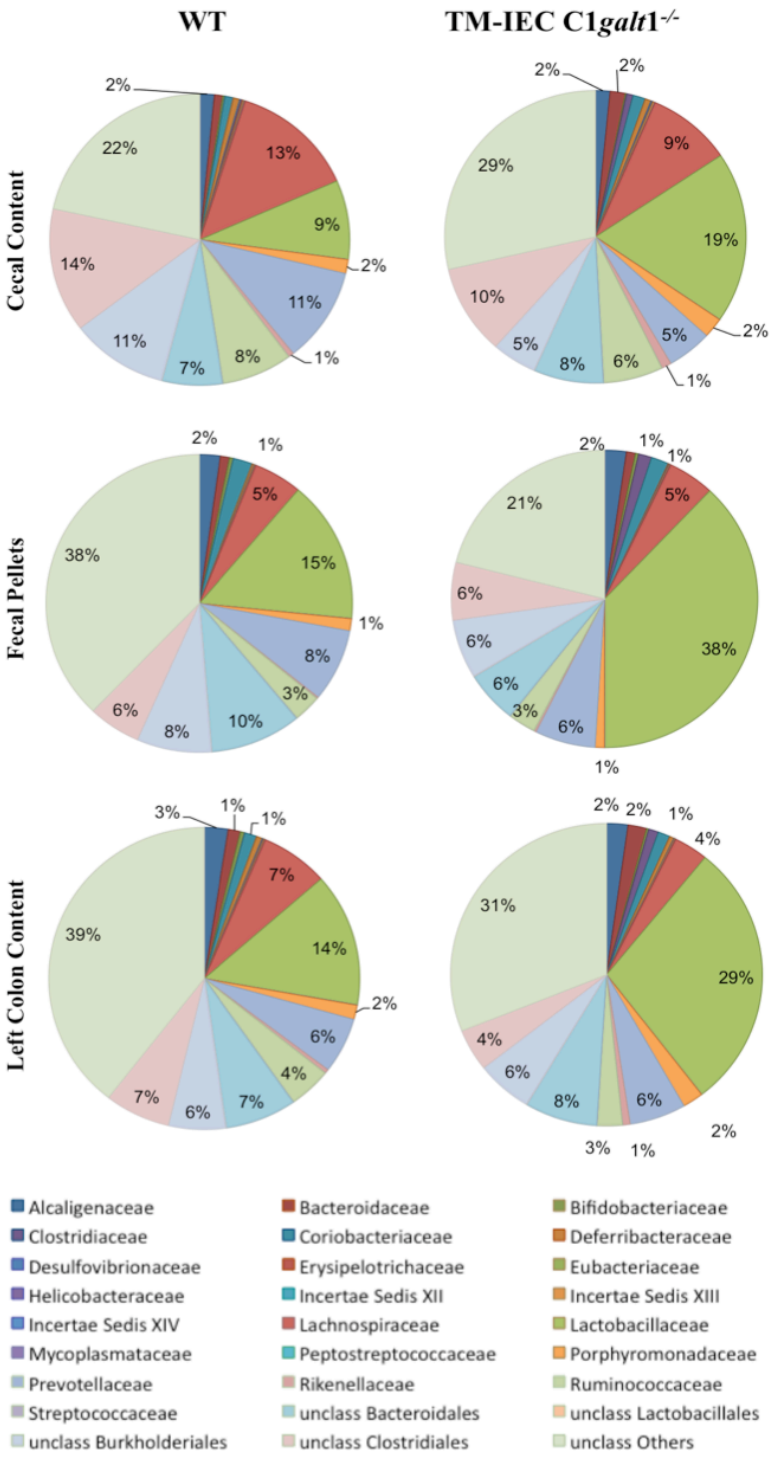

Figure S4. The net change in the population of Firmicutes at the Family level is the result of increases in *Clostridiaceae* and *Lactobacillaceae,* and decreases in *Ruminococcaceae* and *Lachnospiraceae,* *Clostridiaceae* and *Lactobacillaceae* consistently increased in all anatomical sites in the TM-IEC *C1galt1*$^{-/-}$ mice. Pie charts represent the abundance (in percentage) of bacterial families present in each anatomical site per genotype.

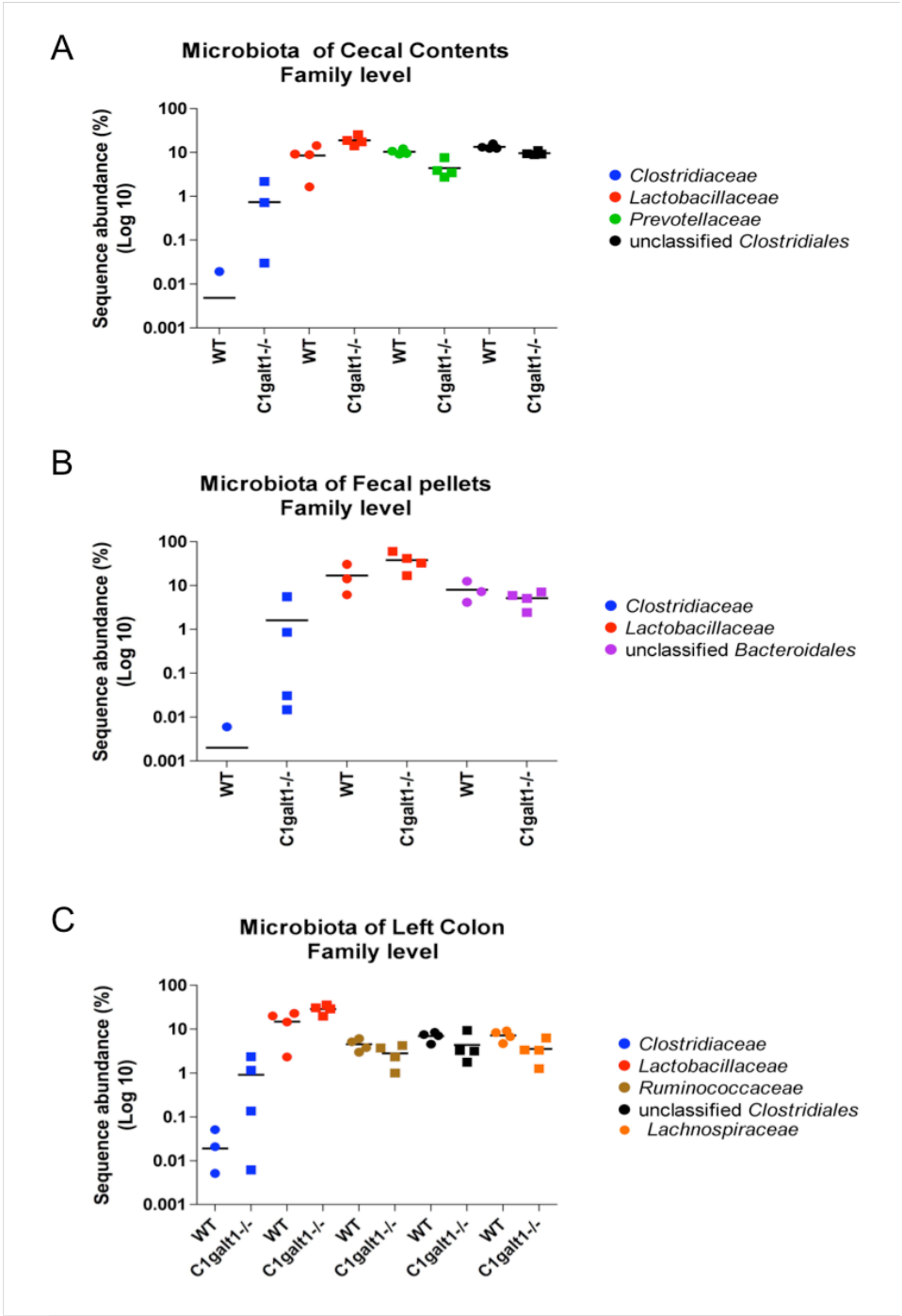

Figure S5. Anatomical site-specific decreases were found in the TM-IEC *C1galt1*$^{-/-}$ mice at the Family level. Decreases in *Prevotellaceae* were found exclusively in the cecal contents [A], while decreases in unclassified *Bacteroidales* were found exclusively in the fecal pellets [B], and decreases in *Ruminocccaceae* were found exclusively in the LCC [C] of the TM-IEC *C1galt1*$^{-/-}$ mice. Horizontal bars represent means. N=3-4 for each anatomical site tested, data points equal to zeros are not shown as the y-axis is in log10 scale.

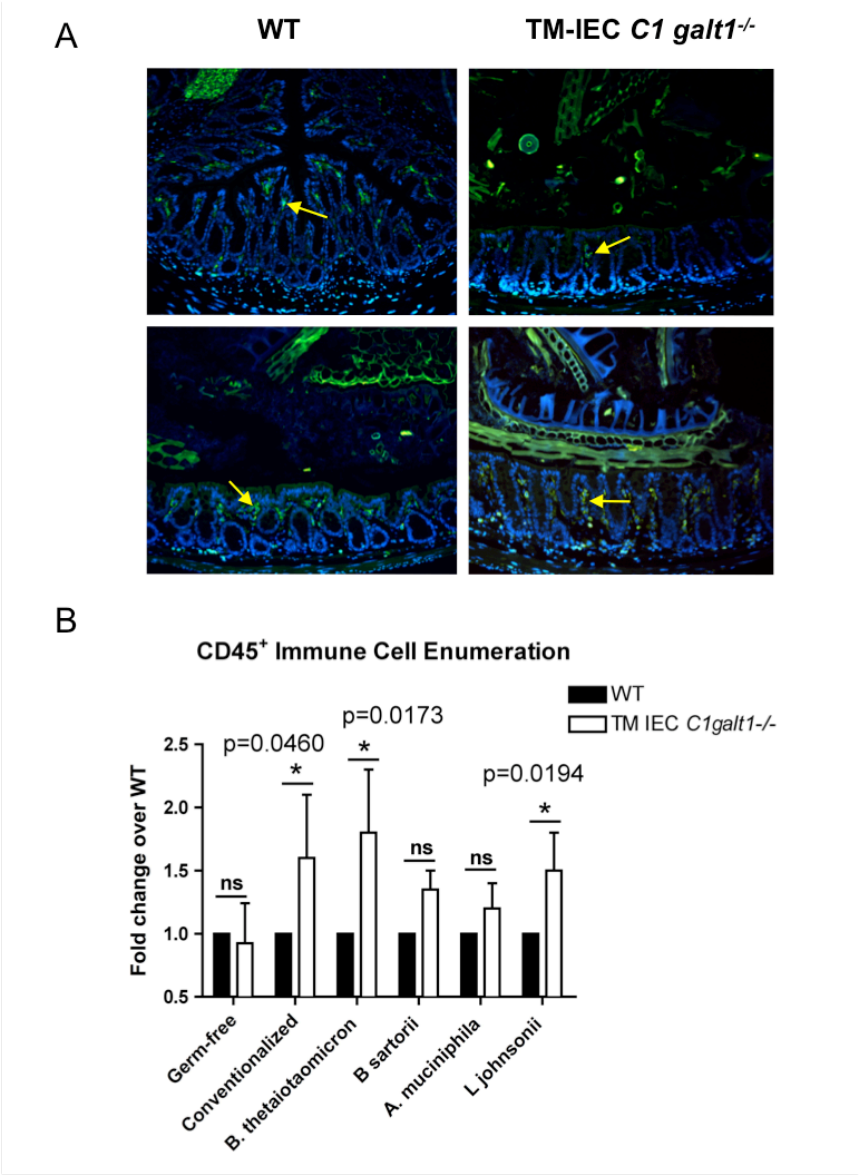

Figure S6. Intestinal symbiont *Bacteroides thetaiotaomicron* induced the highest degree of infiltrated lamina propria CD45+ cells into the colon tissue in the absence of core 1-derived *O*-glycans. [A] Changes in number of $CD45^+$ cells (pointed by yellow arrows) were estimated to evaluate the immune response caused by the bacteria in the case of a breach in the mucus layer, relative to bacterial association in an intact mucosal barrier. [B] Lack of core 1 *O*-glycans and single association of GF mice with *B. thetaiotaomicron* caused the strongest immune response than mice associated with a complete microbiota, as measured by changes in infiltrated lamina propria $CD45^+$ cells. *B. thetaiotaomicron* caused a 1.8 fold-change increase in the number of $CD45^+$ cells. Vertical bars represent fold change over WT. Error bars represent standard deviations. P-values were estimated for changes between genotypes (WT *versus* TM-IEC *C1galt1*$^{-/-}$).

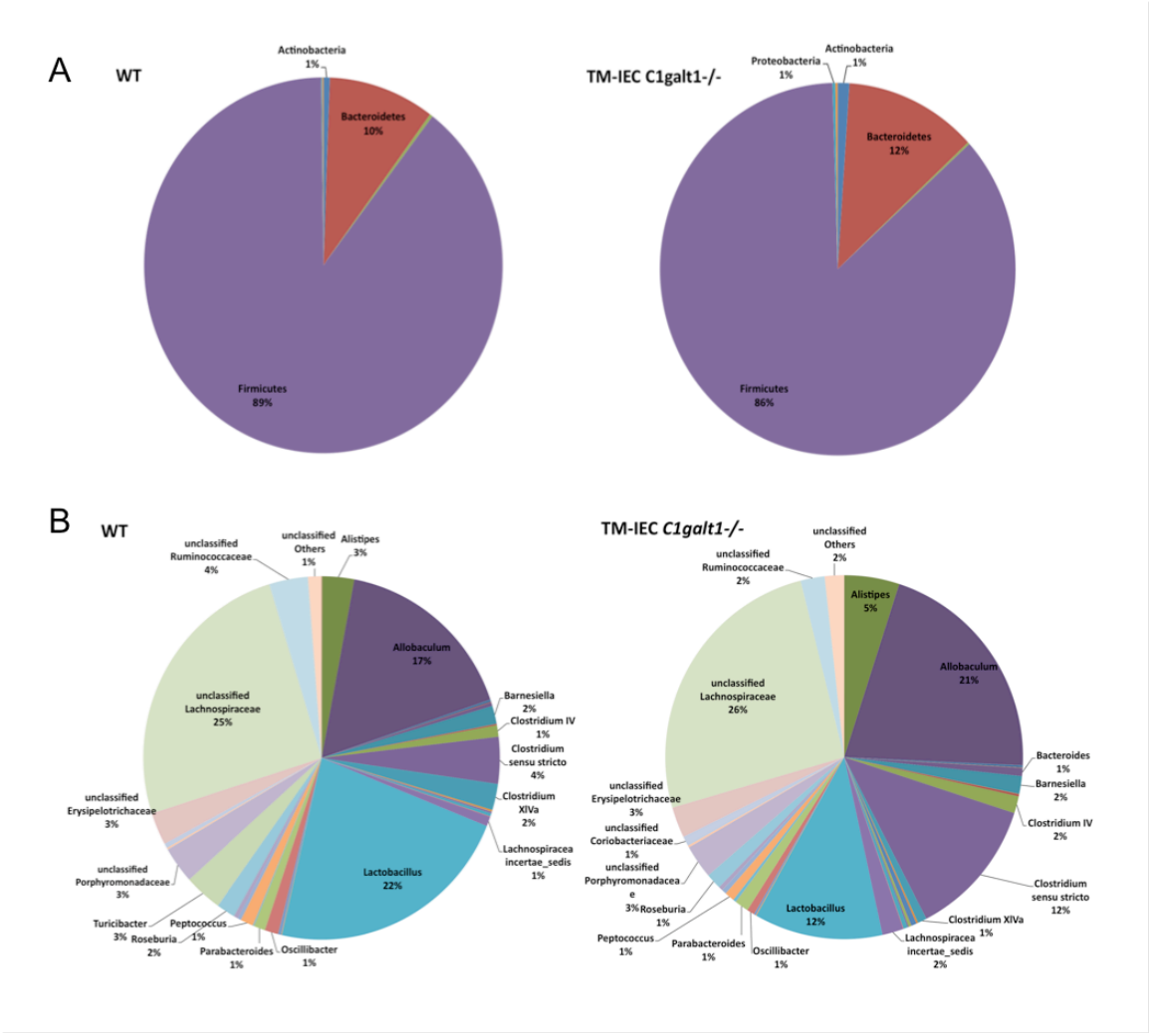

Figure S7: The microbiota of the conventionalized WT and TM-IEC *C1galt1*$^{-/-}$ differs in composition from the microbiota of the conventional WT and TM-IEC *C1galt1*$^{-/-}$ mice, although some similarites exist. [A] No significant changes in microbial taxa were found at the phylum level between the genotypes in the cecal contents of the conventionalized mice. [B] Significant changes at the genus level include increases in the *Clostridium senso strict*o group (which includes *Clostridium disporicum*), and decreases in *Clostridium* XIVa, *Lactobacillus* and *Turicibacte*r.

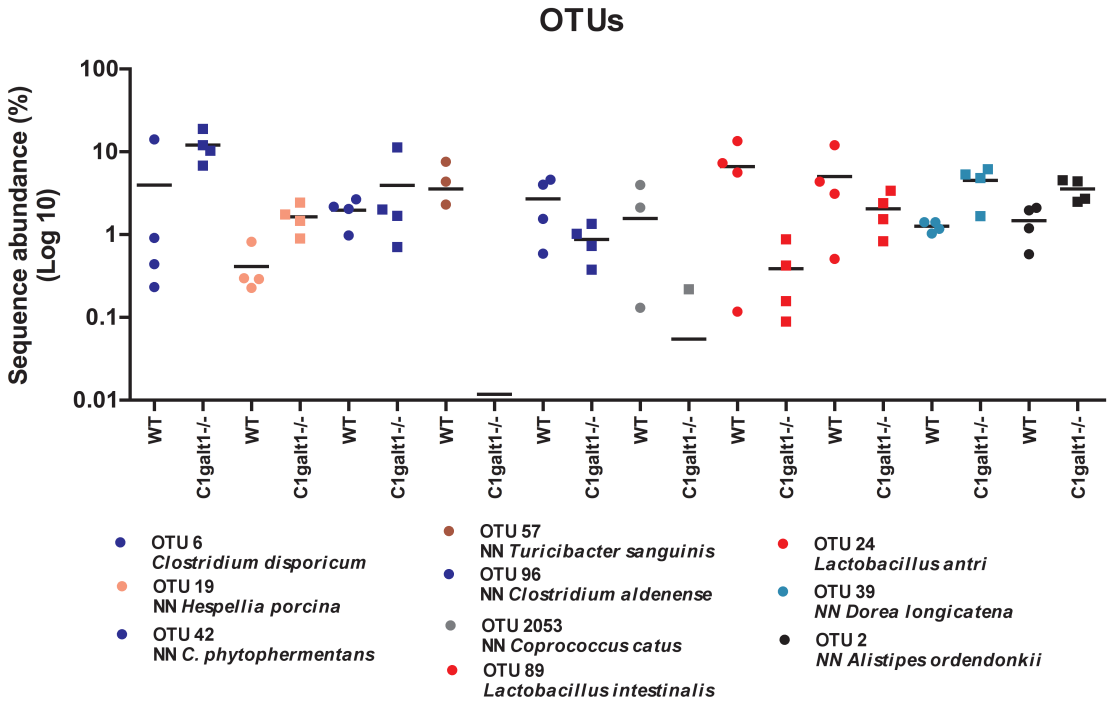

Figure S8. OTUs analysis for the *conventionalized* mice revealed *Clostridium disporicum* as the main similarity and significant change in the microbiota of the TM-IEC *C1galt1*$^{-/-}$ when compared to the conventional microbiota.

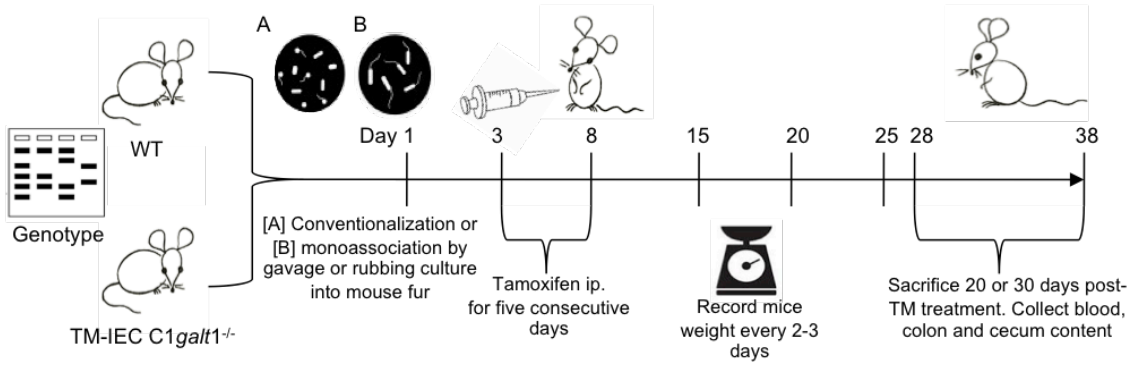

Figure S9. Schematic representation of the experimental protocol. GF mice were colonized with either 200µl of diluted cecal contents from conventional mice or with single bacterium pure culture. Subsequently (48 to 72 hours post-colonization), 1 mg of Tamoxifen (TM) diluted in an ethanol and sunflower oil mixture (1:9 v/v) was injected intraperitoneally for five consecutive days to both WT and TM-IEC *C1galt1*$^{-/-}$ mice. Mice were sacrificed 20 or 30 days after TM treatment, depending on bacterial species tested.