Supplementary Materials for

Gut microbiota utilize immunoglobulin A for mucosal colonization

G.P. Donaldson, M.S. Ladinsky, K.B. Yu, J.G. Sanders, B.B. Yoo, W.C. Chou, M.E. Conner, A.M. Earl, R. Knight, P.J. Bjorkman, S.K. Mazmanian
correspondence to: gdonalds@caltech.edu, sarkis@caltech.edu

This PDF file includes:

Materials and Methods

Bacterial strains, media, and plasmids

Bacteroides fragilis NCTC9343 was grown in Brain Heart Infusion (BD) supplemented with 5 μg/ml hemin (Frontier Scientific) and 5 μg/ml vitamin K1 (Sigma) in an anaerobic atmosphere of 80% nitrogen, 10% carbon dioxide, and 10% hydrogen. For selection, 200 μg/ml gentamicin, 10 μg/ml erythromycin, 10 μg/ml chloramphenicol, or 2 μg/ml tetracycline was added. For *in vivo* studies, *B. fragilis* carried a marker plasmid pFD340-Chlor or pFD340-Tet which confer resistance to chloramphenicol and tetracycline, respectively, allowing identification of multiple strains within an animal without affecting their fitness (*26*). PSC was deleted by allelic exchange with a suicide plasmid (pNJR6) as previously described (*26*). Details for all strains and plasmids are listed in Table S4 and primers for cloning in Table S5.

Mice

All mouse experiments were performed in accordance with the NIH Guide for the Care and Use of Laboratory Animals using protocols approved by the Institutional Animal Care and Use Committee at the California Institute of Technology. Swiss Webster, C57BL/6, *Rag1*$^{-/-}$, and BALB/c mice from Taconic Farms and BALB/c IgA$^{-/-}$ mice from the Baylor College of Medicine were C-section rederived germ-free and bred in flexible film isolators. Unless specifically stated, Swiss Webster mice were used for all experiments due to ease of breeding in gnotobiotic conditions. Eight-week-old germ-free mice were transferred to autoclaved microisolator cages supplied with autoclaved chow (LabDiet 5010) and autoclaved water containing 10μg/ml erythromycin (to select for marker plasmids) and 100μg/ml gentamicin (to which *Bacteroides* are naturally resistant). Mice were mono-colonized by a single oral gavage of 10^8 CFU in 100μl of HBSS with 1.5% sodium bicarbonate and maintained for at least 4 weeks prior to any imaging, mucosal plating, gene expression, or colonization experiments. Colonization was confirmed for all mice and monitored over time using freshly collected fecal samples. Feces were weighed, mashed and vortexed in 1 ml BHI, spun at 400 g for 1 minute to pellet debris, and then diluted for CFU plating.

For depletion of B cells, 8-week-old germ-free Swiss Webster mice were injected intraperitoneally with 250 μg of monoclonal anti-mouse CD20 antibody (clone 5D2 from Genentech) or 250 μg of an isotype control (IgG2a from Bio X Cell). One week later, mice were mono-colonized. One week after that, the mice were given a second identical antibody injection. Four weeks after mono-colonization, the mice were used for horizontal transmission experiments.

Horizontal transmission and sequential gavage colonization assays

For horizontal transmission experiments, pairs of mono-colonized female mice were cohoused in a new autoclaved cage for 4 hours during the day (16 hours overnight for BALB/c mice, which exhibited a higher threshold for invasion), then individually housed in new autoclaved cages. For each experiment, the antibiotic resistance markers (Tet and Chlor) were swapped for half the animals (for example: WT-Tet vs. Δ*ccf*-Chlor, and then WT-Chlor vs. Δ*ccf*-Tet) to ensure that the difference in horizontal transmission was due to the bacterial or mouse genotype and not due to the different antibiotic resistance markers. The same assay was performed in specific pathogen free (SPF) BALBc mice by first treating with ciprofloxacin and metronidazole as described

previously (*26*) to allow colonization with *B. fragilis*. For sequential gavage experiments, mice mono-colonized for three weeks were gavaged with 10^8 CFU of the challenge strain in HBSS with 1.5% sodium bicarbonate and subsequently single-housed. The community-level comparison of BALB/c mice (Fig. 4H and I) was made by gavaging 8-10 week-old mice with 100 µl cecal contents (donors: 8 week-old Swiss Webster mice delivered from Taconic, 2 male and 2 female cecums dissected and pooled) homogenized to saturation in HBSS with 1.5% sodium bicarbonate. Mucus and lumen samples were collected 5 weeks after colonization.

Sample preparation for electron microscopy

A 1 cm portion of ascending colon from mono-colonized mice was excised and immediately fixed with an ice-cold solution of 3% glutaraldehyde, 1% paraformaldehyde, 5% sucrose in 0.1 M sodium cacodylate trihydrate, without flushing of the intestinal contents. Tissues were prefixed for 1 hour at 4°C, then transferred to a petri dish containing 5% sucrose in 0.1 M cacodylate buffer. Tissues were cut into ~1-2 mm^3 blocks with a #11 scalpel and placed into brass planchettes (Type A; Ted Pella, Inc., Redding, CA), prefilled with cacodylate buffer containing 10% Ficoll (70 kD, Sigma-Aldrich) which serves as an extracellular cryoprotectant. Excess buffer was removed with Whatman filter paper and the sample covered with a Type B brass planchette. Samples were ultra-rapidly frozen with a HPM-010 high-pressure freezing machine (Leica Microsystems, Vienna Austria), then transferred immediately under liquid nitrogen to cryotubes (Nunc) containing a frozen solution of 2.5% osmium tetroxide, 0.05% uranyl acetate in acetone. The tubes were loaded into a AFS-2 freeze-substitution machine (Leica Microsystems) precooled to -100°C. Samples were processed at -90°C for 72 hours, warmed over 12 hours to -20°C, held at that temperature for 6-10 hours, then warmed to 4°C for 1 hour. The fixative was removed and the samples rinsed 4 x with cold acetone, infiltrated into Epon-Araldite resin (Electron Microscopy Sciences, Port Washington PA) and flat-embedded between two Teflon-coated glass slides. Resin was polymerized at 60°C for 48 hours.

Electron microscopy and dual-axis tomography

Flat-embedded colon samples were observed with a stereo dissecting microscope to ascertain preservation quality and select appropriate regions for EM study. These were extracted with a microsurgical scalpel and glued to the tips of plastic sectioning stubs. Semi-thick (300 or 400 nm) serial sections were cut with a UC6 ultramicrotome (Leica Microsystems) using a diamond knife (Diatome, Ltd. Switzerland). Sections were placed onto Formvar-coated copper-rhodium 1 mm slot grids (Electron Microscopy Sciences) and stained with 3% uranyl acetate and lead citrate. Gold beads (10 nm) were placed on both surfaces of the grid to serve as fiducial markers for subsequent image alignment. Grids were placed in a dual-axis tomography holder (Model 2040, E.A. Fischione Instruments, Export PA) and imaged with a Tecnai TF30ST-FEG transmission electron microscope (300 KeV; FEI Company/ThermoFisher Scientific, Hillsboro OR) equipped with an XP1000 2k x 2k CCD camera (Gatan, Inc. Pleasanton CA). Tomographic tilt-series and large-area montaged overviews were acquired automatically using the SerialEM software package (*49*). For tomography, samples were tilted +/- 64° and images collected at 1° intervals. The grid was then rotated 90° and a similar series taken about the orthogonal axis. Tomographic data was calculated, analyzed and modeled using the IMOD software package (*50*, *51*). For quantification of epithelial associated bacteria (Fig. 1F and 4J), sections were surveyed until a bacterium on the apical epithelial surface was identified, and then a ~20 x 20 µm montaged projection image was acquired. Bacterial cells were identified by morphology and

counted. About 1 mm length of epithelium was scanned per mouse, including an area of around 20 μm from the apical epithelial surface into the mucus. For measurement of the capsule thickness (Fig. 1I), the distance between the outer membrane and edge of the electron-dense capsule layer was measured in tomograms. Ten different measurement sites on the same cell were averaged for each cell.

Mucosal scraping for CFU quantification or sequencing analysis

The ascending colon was dissected and cut open longitudinally. Lumen content was removed, and the flat tissue was washed with HBSS until no feces could be observed. A 1 cm length of the tissue was cut, making an approximately 1 cm x 1 cm square of flat tissue. Mucus was removed from the epithelial surface by holding one corner of the tissue and using light pressure with a sterile plastic 1.8 cm cell scraper (BD Falcon). For sequencing analysis, mucus was frozen at this point. For CFU quantification, the mucus was moved into a 2 ml screw cap tube with 1.4 mm ceramic beads (Lysing Matrix D, MP Biomedicals) and 1 ml of HBSS and bead beat on medium for 1 minute to homogenize the mucus. This homogenate was diluted for CFU plating.

RNA isolation

Pelleted mid-log bacterial cultures, freshly harvested colon lumen contents, or whole colon tissue were immediately lysed by bead-beating in a mixture of 500 μl buffer (0.2 M NaCl and 20 mM EDTA), 210 μl 20% SDS, and 500 μl phenol, chloroform, and isoamyl alcohol (Ambion AM9720). The aqueous phase was separated by centrifugation and moved to a new tube for a second extraction with phenol, chloroform, and isoamyl alcohol. Then, 50 μl of 3 M sodium acetate and 500 μl of cold ethanol were mixed in to the aqueous fraction and placed on ice for 20 minutes. RNA was pelleted and washed once with cold 70% ethanol then resuspended in 100 μl water. RNA was further purified using the Qiagen RNeasy mini kit according to the manufacturer instructions. DNA was removed using Turbo DNase (Ambion AM2238) for one hour at 37 C before applying to a second Qiagen RNeasy column, including an on-column Qiagen RNase-free DNase digest.

qRT-PCR

First-strand cDNA synthesis from total RNA was achieved using the manufacturer's instructions for the iScript cDNA synthesis kit (Bio-Rad). Reactions were run on an ABI PRISM 7900HT Fast Real-Time PCR System (Applied Biosystems) with Power SYBR Green PCR Master Mix (Applied Biosystems). Relative quantification was performed using the ΔΔCt method with *gyrB* (DNA gyrase) as the housekeeping control gene. The mean Ct value from 3 technical replicates was used for each biological replicate. Primers are listed in Table S5.

RNAseq

For bacterial RNAseq (Fig. 2A), 10 ml liquid cultures were harvest at mid-log phase and total RNA was isolated as described above (RNA isolation). Libraries were prepared using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB E7530) according to the manufacturer's instructions and sequenced on an Illumina HiSeq2500 in single read rapid run mode with single-end 50 bp reads. For each sample, at least 2.5 million reads mapped to protein-coding genes. The bacterial RNAseq analysis package Rockhopper (*52*) was used for quality filtering, mapping, and differential expression analysis.

For mouse RNAseq (Fig. 3A), a 1 cm length of opened and washed ascending colon tissue was prepped for total RNA as described above (RNA isolation). Ribosomal RNA was depleted using a Ribo-Zero Magnetic Gold Epidemiology Kit (Epicentre/Illumina). Sequencing was done on an Illumina HiSeq2500 to produce 101-bp paired-end reads. Reads were trimmed with a Phred quality score cut-off of 20 using fastq_quality_trimmer from the FASTX toolkit, version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/). Reads shorter than 20 bp after adaptor and poly(A)-trimming were discarded. Trimmed reads were aligned to the mouse genome (build GRCm38.p4) by STAR (*53*), read counts were quantified using HTseq (*54*), and differentially expressed genes were identified using edgeR (*55*).

Bacterial flow cytometry

Fresh fecal pellets were mashed and vortexed in 1 ml of HBSS per 100 mg feces. Large debris was removed by centrifugation at 400 x g for 5 minutes. Supernatants, containing fecal bacteria, were diluted to 10^6 CFU per well in 96-well v-bottom plates, washed with HBSS, and stained with a monoclonal rat anti-mouse IgA conjugated to PE (eBioscience clone mA-6E1) or monoclonal rat anti-mouse IgM conjugated to PE (Biolegend clone RMM-1) at 1:250 and SYTO 9 nuclear stain (Molecular Probes) at 1:1000. After washing twice with HBSS, stained bacteria were analyzed on a MACSQuant VYB (Miltenyi). Data were analyzed using FlowJo.

For IgA affinity studies, IgA was extracted from fresh fecal samples by mashing and vortexing in 1 ml HBSS per 100 mg feces, then removing the insoluble fraction by centrifuging at 16,000 x g for 10 minutes at 4 C, and collecting the supernatant containing soluble IgA. The pellet was resuspended in the same volume for a second extraction. The combined supernatant was then sterile filtered, heat inactivated at 56 C for 30 minutes, and then stored at 4 C for up to 1 month (degradation was not observed by ELISA in this timeframe). IgA concentrations were measured using a Ready-SET-Go mouse IgA ELISA kit (eBioscience) according to the manufacturers protocol. Binding assays were performed on fecal-derived bacteria from mono-colonized *Rag1*$^{-/-}$ mice so that the bacteria were *in vivo*-adapted (in which *ccf* is induced and the thick capsule is formed), but IgA-free (because *Rag1*$^{-/-}$ mice do not make antibody). 10^6 CFU per well in 96-well plates was incubated on a shaker with 200 µl IgA at a concentration of 1 µg / ml for 30 minutes and then stained for flow cytometry.

Western blot

Fresh fecal pellets from *Rag1*$^{-/-}$ mice were mashed and vortexed in 1 ml of BHI per 100 mg feces. Large debris was removed by centrifugation at 400 x g for 5 minutes. This supernatant or late-log cultures of bacteria were pelleted at 8,000 g and resuspended at 10X concentration in BHI. The concentrate was mixed 1:1 with 2X Laemmli sample buffer and boiled for 10 minutes to lyse bacteria. 20 µl of lysate (10^8 CFU) was loaded per well in tris-glycine gels with a 4-20% polyacrylamide gradient (Novex WedgeWell), ran in a tris-glycine buffer with 0.1% SDS (Bio-Rad). Gels were transferred at 4 C overnight to a PVDF membrane in the same buffer supplemented with 20% methanol. PVDF blots were blocked for 1 hour in PBS-T (PBS with 0.1% Tween20) with 5% powdered milk, washed with PBS-T, and then stained with IgA, extracted from feces as described above, at a final concentration of 0.1 µg / ml in PBS-T with 2.5% powdered milk for 2 hours. Blots were washed four times with PBS-T, stained with an HRP goat anti-mouse IgA (Southern Biotech 1040-05) diluted 1:1000 in PBS-T with 2.5%

powdered milk for 1 hour, washed four times again, and then developed with a Clarity western ECL substrate (Bio-Rad).

Tissue-cultured epithelial cell adherence assay

Intestinal epithelial cell lines HT29, HT29-MTX, and Caco-2 were maintained in high glucose DMEM with 4 mM L-glutamine, 4.5 g / L glucose (HyClone), supplemented with 10% fetal bovine serum (FBS) (Gibco), and penicillin-streptomycin solution (Corning). Cells were seeded in flat 96-well plates at a density of 25,000 cells per well and then grown for seven days past confluence to allow production of mucus. IgA was extracted and bound to fecal-derived bacteria from mono-colonized *Rag1*$^{-/-}$ mice so that the bacteria were *in vivo*-adapted but IgA-free as described above (Bacterial flow cytometry for IgA coating and affinity). IgA-coated bacteria were pelleted and resuspended in high-glucose DMEM without antibiotics or FBS. Epithelial cells were washed once with HBSS and then incubated with 10^6 coated bacteria for 2 hours at 37 C in an anaerobic atmosphere of 80% nitrogen, 10% carbon dioxide, and 10% hydrogen to allow adherence of bacteria. Bacterial growth was not observed within this timeframe either on epithelial cells or in media alone. Wells were washed twice with 200 µl HBSS to remove unbound bacteria, and then trypsinized with 50 µl of 0.05% trypsin in HBSS (Corning) for 20 minutes at 37 C to disassociate epithelial cells. Then 50 µl of BHI was added and cells were vigorously resuspended before dilution plating for CFU. Fraction bound was calculated as the output CFU / input CFU. For all epithelial cell adherence assays, the average of 4 technical replicates (4 separate binding reactions with the same IgA and bacteria on 4 wells of cells) is reported for each biological replicate.

B cell flow cytometry

Cells were isolated from the peritoneum, mesenteric lymph nodes, and colon tissues. The peritoneal space was injected with PBS and massaged for thirty seconds to suspend cells. Mesenteric lymph nodes were dissected and single cell suspensions were obtained by grinding tissues through a 100 µm cell strainer. Colonic lamina propria lymphocytes were isolated by first flushing lumen contents away with PBS, and removing the longitudinal and circular muscle layers of the colon by micro-dissection. The remaining tissue was incubated in 1 mg / mL Collagenase II (Dibco) for 1 hour and then cells were filtered through a 100 µm cell strainer. Cells were spun down, resuspended and separated by a 40%/80% (v/v) Percoll (GE Healthcare) density gradient. All cells were washed in complete RPMI buffer, allowed to rest at 37 C for at least 1 hour in a 5 % CO_2 tissue culture incubator, and incubated in 5% mouse serum for 15 minutes before proceeding to staining for flow cytometry. Cells were stained with a LIVE/DEAD fixable Violet dye (Life Technologies), FITC-conjugated B220 antibody (eBioscience clone RA3-6B2), and PerCP-Cy5.5-conjugated CD3e antibody (eBioscience clone 145-2C11) for 20 minutes at 4 C. Cells were washed in HBSS with 1% BSA and fixed in 2% PFA prior to acquisition on a MACSQuant Analyzer (Miltenyi). Data were analyzed using FlowJo.

Sorting IgA-bound bacteria for sequencing

This method was carried out as previously described (*35*). Two fecal pellets were mashed and resuspended in 1 ml PBS with 0.5% BSA (this buffer used throughout the protocol). The homogenate was centrifuged at 400 g for 5 minutes to remove debris, and the supernatant was passed through a 30 µm filter. Bacteria were pelleted by centrifugation at 8000 g for 5 minutes

and resuspended in 1 ml buffer plus 35 µl of anti-IgA PE (eBioscience clone mA-6E1) and stained for 20 minutes at 4 C. Stained bacteria were pelleted, washed once, then resuspended in 1 ml buffer plus 100 µl ultrapure anti-PE magnetic beads (Miltenyi) and incubated for 15 minutes at 4 C. Bead-bound, stained bacteria were pelleted and resuspended in 500 µl buffer and then run over an MS column (Miltenyi) on an OctoMACS separator (Miltenyi). Bound bacteria were eluted and re-loaded on the column three times to further purify.

16S amplicon sequencing

DNA was extracted from tissue scrapings, fecal samples, and IgA-sorted bacterial pellets using the Qiagen PowerMag Soil extraction kit and a ThermoFisher KingFisher magnetic bead purification robot. Extracted DNA was amplified and sequenced as in Caporaso et al (*56*). Briefly, the V4 region of the 16S ribosomal RNA gene was amplified in triplicate using the 515F/805R primers from the Earth Microbiome Project (*57*) and amplicons pooled and sequenced on an Illumina MiSeq instrument. Sequences were uploaded to the Qiita analysis platform for demultiplexing and primary data processing (https://qiita.ucsd.edu/). Exact sequence variants (ESVs) were identified and chimeric sequences removed using the Deblur plugin in Qiita (*58*). For phylogenetic analyses, ESVs were inserted into the GreenGenes reference phylogeny (release 13_8, (*59*))using SEPP (*60*). Taxonomy was assigned per ESV according to its placement in the GreenGenes phylogeny using the fragment-insertion pluging in Qiime 2 v2017-12 (https://qiime2.org). Samples yielding less than 1000 Deblurred sequences were excluded from subsequent analysis.

16S diversity analyses

We used the Galaxy implementation LEfSe (*61*) to identify taxonomic groups that were significantly enriched (LDA effect size ≥ 2) in bacterial pellets purified using anti-IgA antibodies, coding pellet sort status as a class and mouse sex as subclass and normalizing to 10^6 counts per sample. To identify specific ESVs differentially represented in tissues or feces of BALB/c and BALB/c $IgA^{-/-}$ mice, we used the DESeq2 (*62*) in R (*63*), adding a pseudocount of 1 to all ESVs and fitting models using a negative binomial distribution. Analyses of microbial community beta-diversity were performed using the Weighted UniFrac metric (*64*) calculated from count tables rarefied to 1000 sequences in Qiime2. Principle Coordinate Analyses and Adonis tests of group differences in central tendency (*65*) were performed in the Vegan package in R (*66*).

Statistical methods

All bar graphs show mean values with standard error of measurement. Log-scale line graphs for horizontal transmission and sequential gavage assays show the geometric mean and 95% confidence interval. Post-hoc corrections for multiple comparisons were the Tukey method when comparing all groups with each other and Sidak for other types of comparisons.

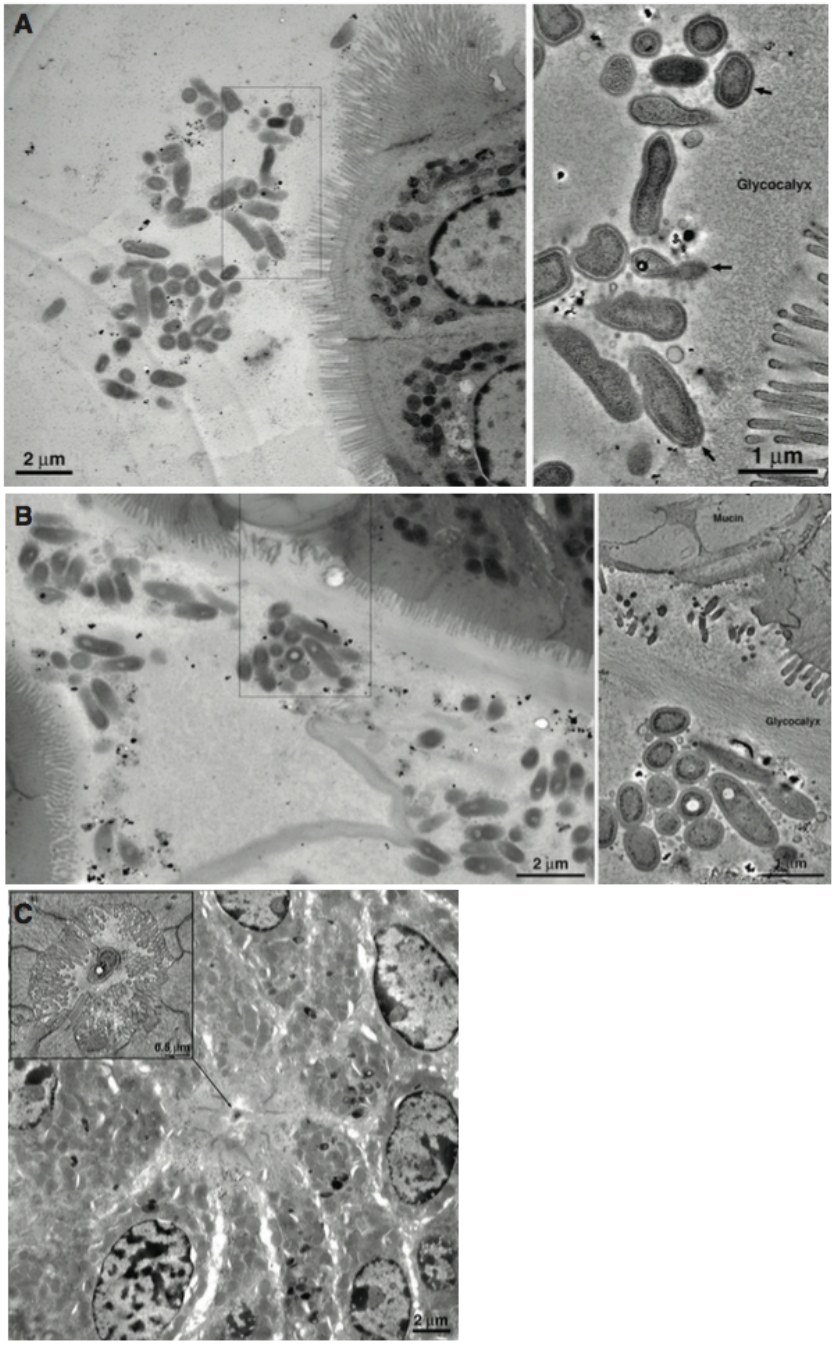

Fig. S1. Additional EM images from mice mono-colonized with wild-type *B. fragilis*. (**A** and **B**) Example aggregates of *B. fragilis* on the epithelial surface in projection images (left) with high resolution tomograms of the marked region (right). Arrows indicate examples of bacteria penetrating the glycocalyx. (**C**) Projection and inset tomogram of an additional example of *B. fragilis* in a crypt of Lieberkühn.

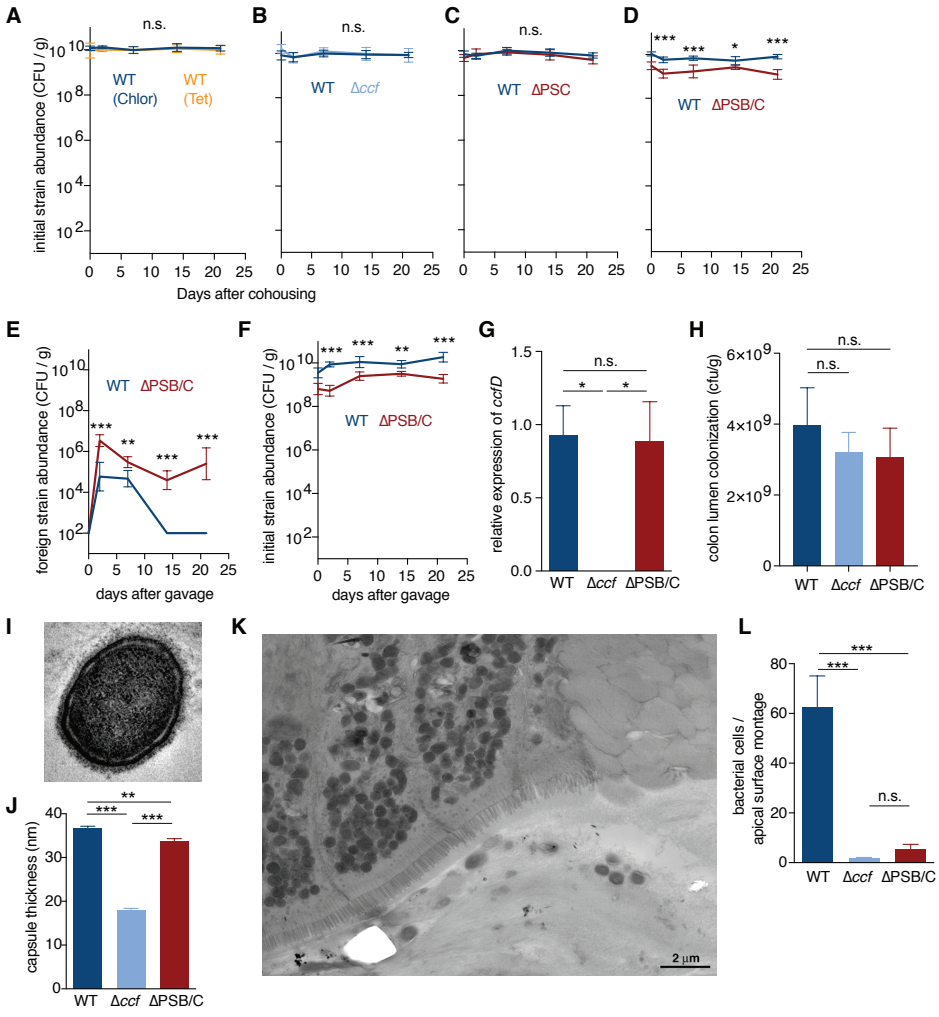

Fig. S2. Additional *in vivo* phenotypes for various *B. fragilis* strains. (**A**) Abundance of the initial strains corresponding to Fig. 2D, (**B**) Fig. 2E, (**C**) Fig. 2F, and (**D**) 2G (n = 9-12). (**E**) Abundance of foreign strains in mice originally mono-colonized with indicated strain (n = 6) and challenged by gavage. (**F**) Abundance of the original strains corresponding to E (geometric mean and 95% confidence interval plotted for all line graphs with Sidak repeated measure 2-way ANOVA on log-transformed data). (**G**) qRT-PCR analysis for *ccf* expression of RNA extracted from colon lumen contents of mono-colonized mice, indicating that PSB/C mutation does not affect *ccf* expression (Tukey ANOVA, n = 3). (**H**) Quantification of bacteria in the colon lumen of mono-colonized mice, corresponding to Fig. 2I (Tukey ANOVA, n = 8). (**I**) High resolution tomogram (image width = 600 nm) of *B. fragilis* ΔPSB/C in the colon mucosa of a mono-colonized mouse and (**J**) quantification of the capsule thickness including data from Fig. 1I for wild-type and Δ*ccf* (Tukey ANOVA, n = 10). (**K**) Example projection montage of the apical surface of the epithelium of mice mono-colonized with *B. fragilis* ΔPSB/C shows aggregates of only small numbers of bacteria. (**L**) Quantification of number of epithelial associated bacteria including data from Fig. 1F for wild-type and Δ*ccf* (Tukey ANOVA, n = 7, 8, 8 images from 4 mice per group) (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

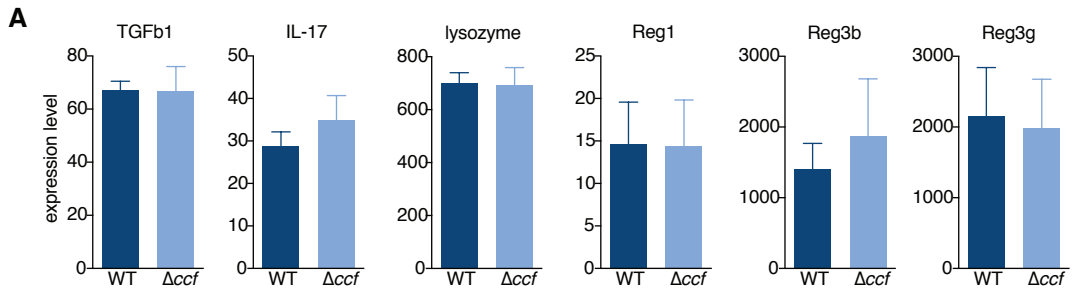

Fig. S3. Host inflammatory profiles of *B. fragilis* and *B. fragilis* Δ*ccf* show no differences. (**A**) Expression levels of inflammatory cytokines, lysozyme, and antimicrobial peptides from ascending colon RNAseq of mice mono-colonized with *B. fragilis* or *B. fragilis* Δ*ccf* reveal an indistinguishable inflammatory response (n = 3). Beta-defensins, TNFα, IL-1β, and IL-6 were not expressed at detectable levels.

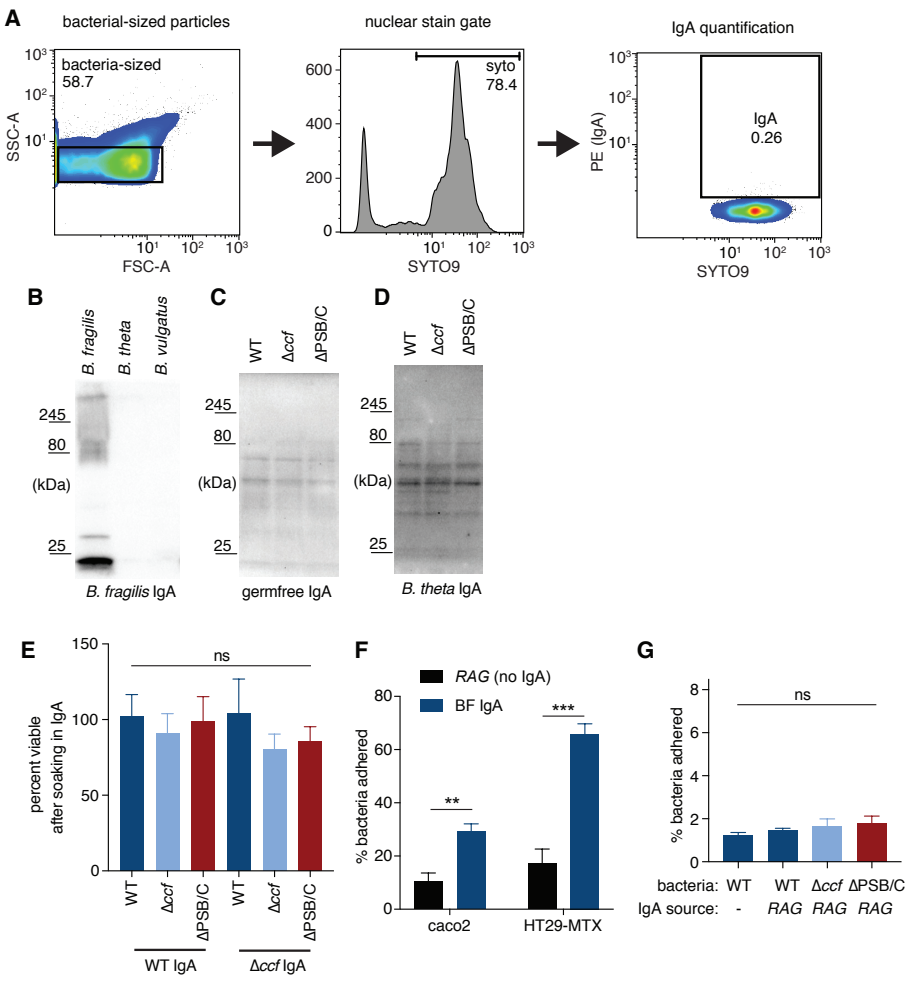

Fig. S4. Interactions between IgA and *B. fragilis*. (**A**) Gating strategy for assessing IgA coating of bacteria. Gates were defined using single-stained samples. (**B**) Lysates of culture-grown *Bacteroides* species (10^8 CFU loaded per well for all blots) probed in western blots with fecal IgA from *B. fragilis* mono-colonized mice, indicating minimal cross-reactivity to other species. (**C**) Lysates of *B. fragilis* strains from the feces of mono-colonized *Rag1*$^{-/-}$ mice probed in western blots with fecal IgA from germ-free and (**D**) *B. thetaiotaomicron* mono-colonized mice, with images overexposed to show minimal binding (blots B-D repeated at least 3 times). (**E**) Plating CFU before and after coating with IgA *in vitro* suggests no impact of IgA on the viability of *B. fragilis* (Tukey ANOVA, n = 8). (**F**) Tissue-cultured epithelial cell adherence assay with various cell lines (in addition to those used in the main text), using wild-type *B. fragilis* from feces of mono-colonized *Rag1*$^{-/-}$ mice and IgA from mice mono-colonized with *B. fragilis* (Sidak 2-way ANOVA, n = 5 mice as the source of bacteria) (**G**) Additional controls for Fig. 3H (first two bars are the same data as in Fig. 3H) indicate no baseline difference in epithelial adherence between strains of *B. fragilis* from feces of mono-colonized *Rag1*$^{-/-}$ mice, without the addition of IgA (Tukey ANOVA, n = 4) (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

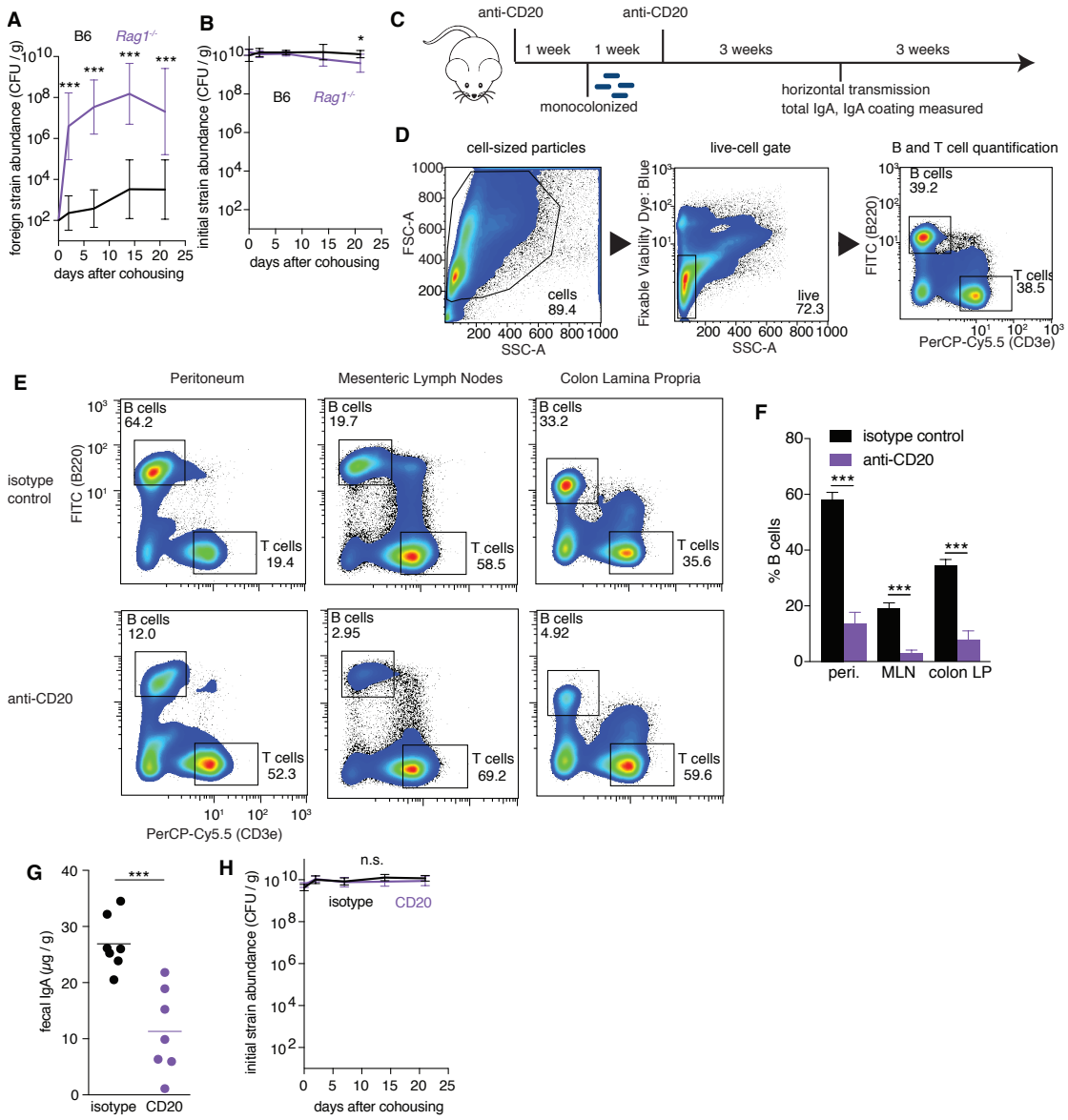

Fig. S5. *B. fragilis* colonization phenotypes in *RAG1*$^{-/-}$ and B cell depleted mice. (**A**) Abundance of foreign strains exchanged between pairs of wild-type *B. fragilis* mono-colonized mice, either C57BL/6 (B6) and *Rag1*$^{-/-}$, in colony forming units (CFU) per gram of feces (geometric mean and 95% confidence interval plotted for all line graphs with Sidak repeated measure 2-way ANOVA on log-transformed data, n = 9). (**B**) Abundance of initial strains corresponding to **A**. (**C**) Anti-CD20 experimental timeline with germ-free Swiss Webster mice. (**D**) Gating strategy for quantification of B cells to confirm depletion by anti-CD20 antibody and (**E**) example plots. (**F**) Quantification of B cells as a proportion of cell populations in the peritoneum, mesenteric lymph nodes, and lymphocyte fraction of the colon lamina propria (LP) following injection of anti-CD20 or isotype control antibody (Sidak 2-way ANOVA, n = 4, 5). (**G**) IgA concentration in feces as assessed by ELISA 3 weeks after second injection of anti-CD20 or isotype control antibody (unpaired t test, n = 7). (**H**) Abundance of initial strains for Fig. 4C (n = 10) (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

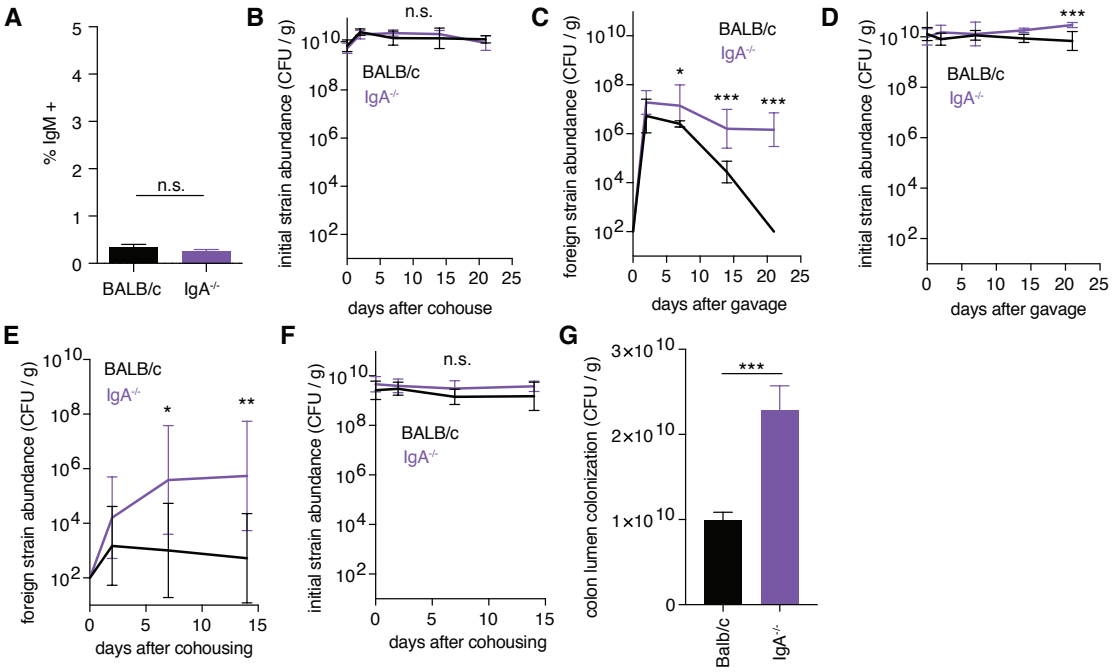

Fig. S6. *In vivo* phenotypes for *B. fragilis* in BALB/C and BALB/c IgA$^{-/-}$ mice. (**A**) Percent of *B. fragilis* coated in IgM in mono-colonized BALB/c and BALB/c IgA$^{-/-}$ mice (unpaired t test, n = 4). (**B**) Abundance of initial strains corresponding to Fig 4D (geometric mean and 95% confidence interval plotted for all line graphs with Sidak repeated measure 2-way ANOVA on log-transformed data, n = 9). (**C**) Abundance of foreign strains in mice mono-colonized with *B. fragilis* and challenged by gavage with *B. fragilis* (n = 5). (**D**) Abundance of initial strains corresponding to **C**. (**E**) Abundance of foreign strains of *B. fragilis* in SPF mice colonized with *B. fragilis* and co-housed (n = 10). (**F**) Abundance of initial strains corresponding to **E**. (**G**) Quantification of colon lumen colonization corresponding to Fig 4E (unpaired t test, n = 9). (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

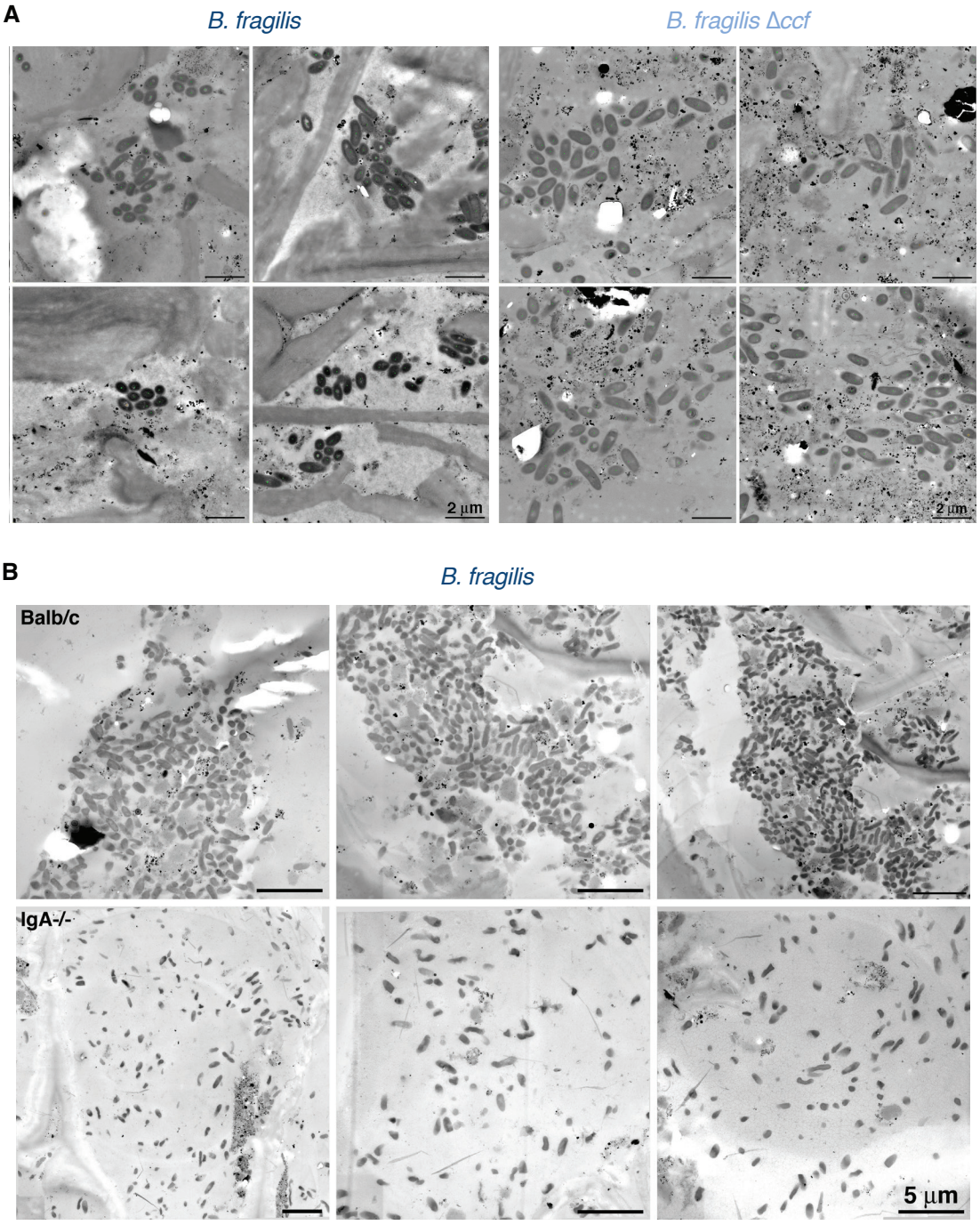

Fig. S7. Aggregation of *B. fragilis in vivo* is affected by IgA and *ccf*. (**A**) Transmission electron micrographs of colon lumen contents of Swiss Webster mice mono-colonized with *B. fragilis* or *B. fragilis* Δ*ccf* (4 representative images for each). (**B**) Transmission electron micrographs of feces of BALB/c and BALB/c IgA$^{-/-}$ mice mono-colonized with *B. fragilis* (3 representative images for each).

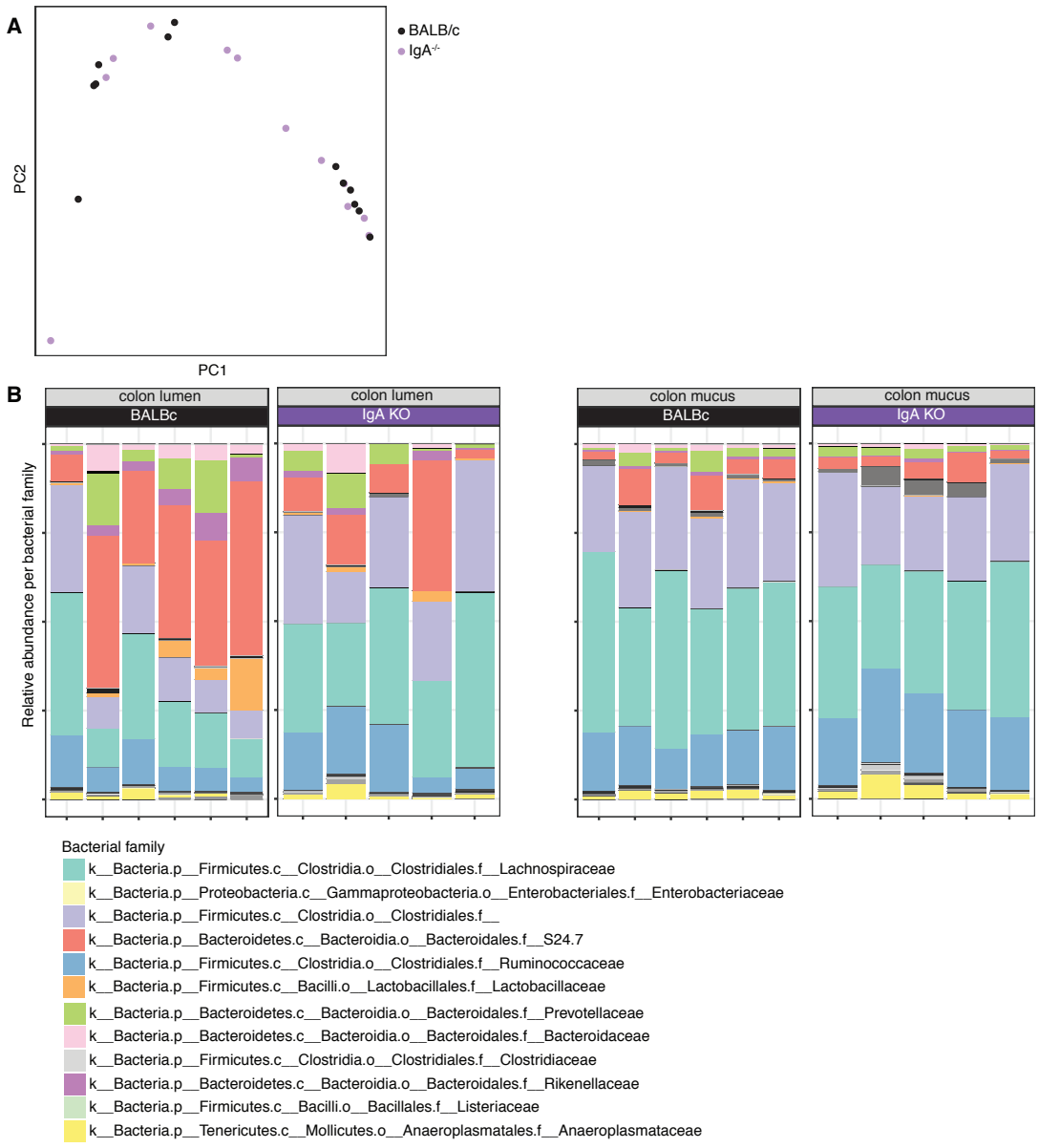

Fig. S8. 16S microbiome profiling of ex-germ-free mice. (**A**) Principle coordinate analysis of weighted UniFrac distances of fecal microbiomes between BALB/c and BALB/c IgA$^{-/-}$ mice. (**B**) Relative abundance of bacterial families in colon and ileal lumen and mucus in ex-germ-free BALB/c and IgA$^{-/-}$ mice. Most abundant 12 families assigned colors; remaining families assigned random shades of gray.

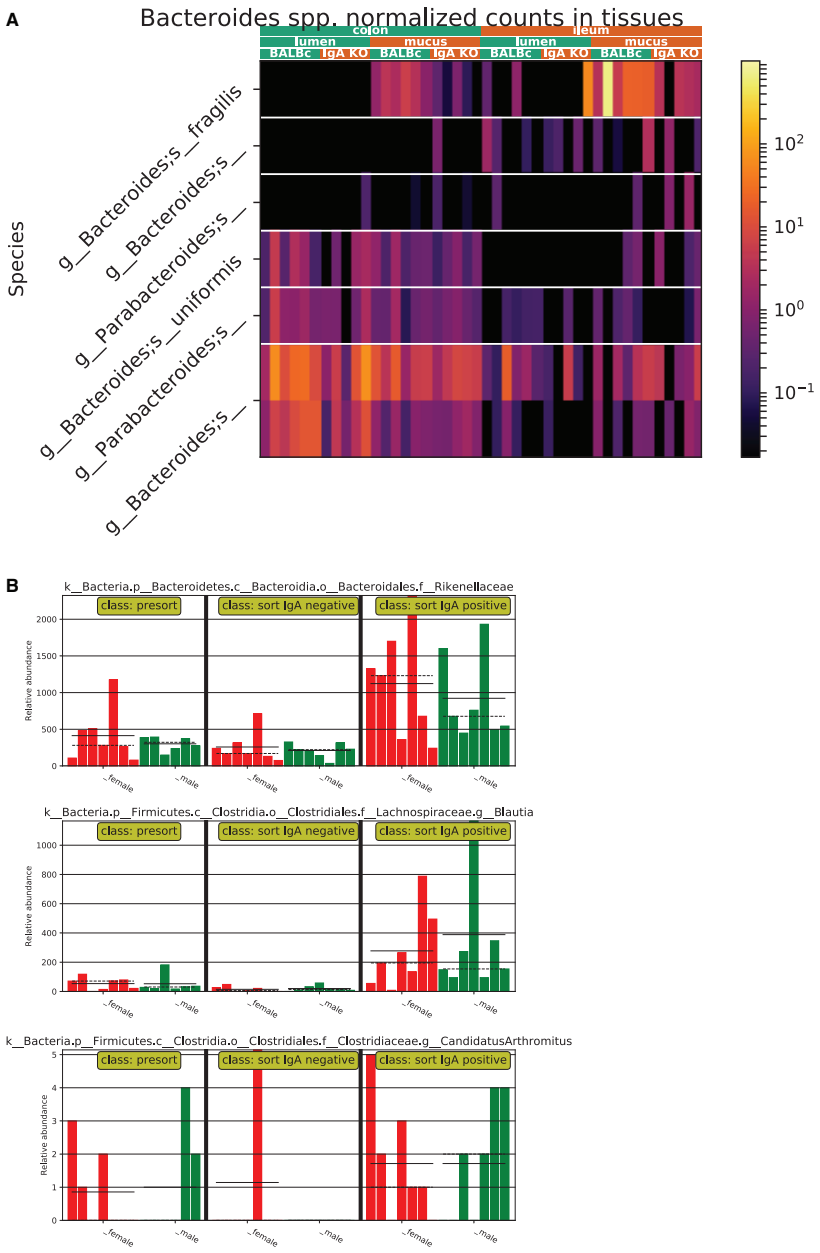

Fig. S9. Analyses of exact sequence variants (ESVs) of interest in ex-germ-free mice. (**A**) Heatmap of relative abundance of ESVs assigned to *Bacteroides* or *Parabacteroides* in the mucus and lumen of the colon and ileum of BALB/c and BALB/c IgA$^{-/-}$ mice. (**B**) Relative abundance of highly IgA-bound taxa (LEfSe LDA effect size > 2) before and after sorting (ESVs in Figure 4I are the most abundant representatives of these taxonomic groups; "Candidatus Arthromitus" = SFB).

Table S1.

Differentially expressed (adjusted p-value <0.05) genes in RNAseq of *B. fragilis* grown in culture with and without overexpression of *ccfA*. Genes are ordered by the fold-change.

Locus ID	Gene	Log2(fold-change) (ccfA overexpression / vector)	Adjusted P-value
BF3583	ccfA	4.08	2.139E-106
BF1021	wcfG (PSC)	2.93	1.29145E-15
BF1009	upcY (PSC)	2.91	9.25035E-13
BF1011	rmlA2 (PSC)	2.89	1.29145E-15
BF1023	wcfI (PSC)	2.85	6.7139E-14
BF1017	wcfD (PSC)	2.81	2.02687E-08
BF1022	wcfH (PSC)	2.66	4.63823E-10
BF1024	wcfJ (PSC)	2.44	1.07951E-07
BF1019	wcfE (PSC)	2.44	5.99448E-07
BF1012	rmlC1 (PSC)	2.42	4.56061E-06
BF1015	wcfB (PSC)	2.42	1.98807E-07
BF1025	wcfK (PSC)	2.42	1.65081E-07
BF1010	upcZ (PSC)	2.21	1.33346E-05
BF1020	wcfF (PSC)	2.19	0.000822972
BF1377	wcfS (PSA)	-2.07	0.005651271
BF1376	wcfR (PSA)	-2.15	0.001828779
BF1369	wzx3 (PSA)	-2.18	8.44926E-05
BF1373	wcfO (PSA)	-2.30	2.01292E-06
BF1374	wcfP (PSA)	-2.41	5.18014E-07
BF1372	wzy3 (PSA)	-2.44	9.8695E-08
BF1370	wcfM (PSA)	-2.45	7.08576E-08
BF1371	wcfN (PSA)	-2.45	3.10176E-08
BF0466	Putative protein	-2.48	1.65081E-07
BF1375	wcfQ (PSA)	-2.62	3.87573E-10
BF1368	upaZ (PSA)	-2.67	8.15223E-12
BF1367	upaY (PSA)	-2.75	6.50178E-13

Table S2.

Differentially expressed (adjusted p-value < 0.05) genes in mice mono-colonized with *B. fragilis* or *B. fragilis* Δ*ccf*. Genes are ordered by the absolute value of the fold-change.

Gene	**Log2(fold-change) (WT/Δ*ccf*)**	**Adjusted P-value**
Igkv4-61	-4.29	0.003125755
Igkv8-19	4.18	1.51E-06
Igkv3-12	-3.99	0.049083696
Ighv1-70	3.74	2.28E-07
Ighv1-5	-3.35	0.024434919
Gpr63	-2.44	0.038541457
Igkv6-23	2.03	3.13E-05
Cyr61	-1.83	6.56E-07
Ighv8-5	1.82	0.001789717
Nr4a1	-1.74	2.69E-10
Dusp1	-1.39	1.27E-07
Fos	-1.36	1.02E-06
Egr1	-1.10	1.19E-06
Sik1	-0.97	3.69E-05
Klf2	-0.90	0.017155274

Table S3.

DEseq analysis of exact sequence variants (ESVs) differing in abundance according to mouse genotype in either feces or the lumen and mucus of the ileum and colon of ex-germ-free mice.

Enriched in:	Sample type	log2(fold-change)	Adjusted p-value	Best taxonomic assignment
BALB/c	mucus and lumen	-8.634412444	0.000312846	Mollicutes RF39
BALB/c	mucus and lumen	-3.905449189	0.000551256	*Bacteroides fragilis*
BALB/c	mucus and lumen	-3.519511063	0.001725039	Lachnospiraceae
IgA$^{-/-}$	mucus and lumen	2.821434619	0.001983777	*Ruminoccus*
IgA$^{-/-}$	mucus and lumen	5.7119259	0.001221637	Clostridiales
IgA$^{-/-}$	mucus and lumen	5.039308072	2.76E-05	Lachnospiraceae
IgA$^{-/-}$	mucus and lumen	3.467202384	2.76E-05	*Ruminoccus gnavus*
IgA$^{-/-}$	mucus and lumen	7.097420438	4.84E-08	Alphaproteobacteria BD7-3
BALB/c	feces	-24.05671228	3.37E-23	Mollicutes RF39
BALB/c	feces	-7.354148271	1.01E-06	Lachnospiraceae
BALB/c	feces	-5.70390907	0.00543052	*Coprococcus*
IgA$^{-/-}$	feces	4.450515436	0.008991039	Clostridiales
IgA$^{-/-}$	feces	2.283274756	0.001376256	*Adlercreutzia*
IgA$^{-/-}$	feces	8.12403428	2.37E-08	Alphaproteobacteria BD7-3
IgA$^{-/-}$	feces	20.67955327	3.84E-20	Clostridiales

Table S4.

Bacterial strains and plasmids used in this study.

Strain / Plasmid	Description	Source
Bacteroides fragilis NCTC9343	Type strain, parent for all mutants in this study	ATCC
Bacteroides thetaiotaomicron ATCC 29148	Type strain	ATCC
Bacteroides vulgatus ATCC8482	Type strain	ATCC
B. fragilis Δ*ccf*	An in-frame deletion within the operon containing *ccfC*, *ccfD*, and *ccfE* (BF3581-79)	(*26*)
B. fragilis ΔPSC	An in-frame deletion of non-regulatory genes in the PSC locus (BF1011-26)	This study
B. fragilis ΔPSB	An in-frame deletion of non-regulatory genes in the PSB locus (BF1895-1914)	(*67*)
B. fragilis ΔPSB/C	ΔPSC mutation made in a ΔPSB background	This study
pFD340	*Escherichia coli* (Carbenicillin) / *Bacteroides* (Erythromycin) shuttle vector	(*68*)
pFD340-*ccfA*	Overexpression of the sigma factor *ccfA* using the IS4351 promoter on pFD340	This study
pFD340-chlor	Marker plasmid for mouse colonization experiments (Erythromycin, Chloramphenicol)	(*26*)
pFD340-tet	Marker plasmid for mouse colonization experiments (Erythromycin, Tetracycline)	(*26*)
pNJR6	Suicide plasmid for allelic exchange (Kanamycin for *E. coli*, Erythromycin for *B. fragilis*)	(*69*)

Table S5.

Primers used for cloning, generating mutants, and qPCR.

Primer	Sequence	Purpose	Source
ΔPSC primer-1	**GTGGATCC**AAATGCGTTGCTTTTGCTTT	Left flank PSC 5’	This study
ΔPSC primer-2	**AAACCATGG**TTCGAAATCGTTTTGCTTCA	Left flank PSC 3’	This study
ΔPSC primer-3	**GATTTCGAA**CCATGGTTTATGCTGGCTTT	Right flank PSC 5’	This study
ΔPSC primer-4	**TTGGATCC**AACACTACGCCTACCCGATG	Right flank PSC 3’	This study
PSC wildtype 1	GGAGGATGTTTGAATTGGTGG	PSC WT check 5’	This study
PSC wildtype 2	CCCGCTTAATGCCCTAAAAT	PSC WT check 3’	This study
PSC KO 1	GGAGGATGTTTGAATTGGTGG	PSC KO check 5’	This study
PSC KO 2	TATCCTGATGTTCTGCTTTTCCG	PSC KO check 3’	This study
ccfA 1	**AAGGATCC**TGCGCAACTGATATTGTTAGAA	*ccfA* (BF3583) cloning 5’	This study
ccfA 2	**AAGAGCTCC**GAAATCTACTCAGTGTAAATGGA	*ccfA* (BF3583) cloning 3’	This study
q. gyrB 1	GTGAATGAGGACGGCAGTTT	qPCR gyrase	This study
q. gyrB 2	CTCGATGGGGATGTTTTGTT	qPCR gyrase	This study
q. PSA 1	TTGTATCCGCAAGGGAGAGA	qPCR PSA	This study
q. PSA 2	CGCTCCATACTGCCCATATT	qPCR PSA	This study
q. PSB 1	GCTTTTGGCTTAATGCTTGTTGG	qPCR PSB	This study
q. PSB 2	GCCTAGAAGTACAATTAGCCCGA	qPCR PSB	This study
q. PSC 1	TGTTTGGTGGCTGCTACTTG	qPCR PSC	This study
q. PSC 2	AGGTGAAGTTTGAAGCCAAGG	qPCR PSC	This study
q. PSD 1	CAATTTGGGAGGTGCGTTGT	qPCR PSD	This study
q. PSD 2	ACGACCAATCCAAAACCCCA	qPCR PSD	This study
q. PSE 1	TGCCTCCCTGTTGGTGAAAA	qPCR PSE	This study
q. PSE 2	AGCGTTAGCCAAACTCCGTA	qPCR PSE	This study
q. PSF 1	TTCTATCGTTCAGCGTGCGA	qPCR PSF	This study
q. PSF 2	TGCCCATACGCCAAATCCTT	qPCR PSF	This study
q. PSG 1	CAAGTACACCTGTCAGTAGTTTGC	qPCR PSG	This study
q. PSG 2	GCAACTTCCAATTCCTAACAAAAGA	qPCR PSG	This study
q. PSH 1	GGAAAACAGTCGGAATGGCTC	qPCR PSH	This study
q. PSH 2	TTCCACACACGCAGACACAA	qPCR PSH	This study
q. *ccfA* 1	GGAATTTGCATGACACTTAT	qPCR *ccfA*	(*26*)
q. *ccfA* 2	CTGAGAGGTTTCATCTTCTG	qPCR *ccfA*	(*26*)
q. *ccfB* 1	AGTGTCCCCACTTCATCGTC	qPCR *ccfB*	(*26*)
q. *ccfB* 2	TGAAACTTTTGCCGGAGAAT	qPCR *ccfB*	(*26*)
q. *ccfC* 1	GATGAACTGATAGCCCATTA	qPCR *ccfC*	(*26*)
q. *ccfC* 2	TAGCGATGACTAAAGGTGTT	qPCR *ccfC*	(*26*)
q. *ccfD* 1	CGGTTATATGCTTTTCAAAC	qPCR *ccfD*	(*26*)
q. *ccfD* 2	CAAATAGAAATCTGCCAAAC	qPCR *ccfD*	(*26*)