**Supplemental Information for "Quantitative Imaging of Gut Microbiota Spatial Organization"**

Kristen A. Earle[§*], Gabriel Billings[#*], Michael Sigal[§], Joshua S. Lichtman[‡], Gunnar C. Hansson[∞], Joshua E. Elias[‡], Manuel R. Amieva[§], Kerwyn Casey Huang[π,§,†], Justin L. Sonnenburg[§,†]

[§]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA

[#]Department of Physics, Stanford University, Stanford, CA 94305, USA

[∞]Department of Medical Biochemistry, University of Gothenburg, 40530, Gothenburg, Sweden

[π]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

[‡]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

[*]Contributed equally

## Supplemental Experimental Procedures

### *Gnotobiotic mouse experiments*

Germ-free Swiss-Webster mice were maintained in gnotobiotic isolators and fed either a standard polysaccharide-rich diet (Purina LabDiet 5K67) or a custom MAC-deficient diet composed of 68% glucose w/v, 18% protein w/v, and 7% fat w/v (Bio-Serv) (Sonnenburg et al., 2010). Mice were of mixed gender and approximately 6-8 weeks of age at the start of experiments. Sterile food and water were provided *ad libitum*. All mouse experiments were completed in accordance with the guidelines of the Institutional Animal Care and Use Committee of Stanford University.

Mice were monocolonized by gavage with $10^8$ CFUs in 200 µL TYG of *B. thetaiotaomicron* VPI-5482 (ATCC 29148, which we refer to as *Bt*). *Bt* was cultured anaerobically in TYG at 37 °C (Sonnenburg et al., 2005). For *Bt* monocolonization diet experiments, mice were switched to the MD diet the day of colonization, and maintained on this diet for 10 days before sacrifice.

For *Bt* and *Salmonella typhimurium*$^{avir}$ (*St*$^{avir}$; SL1344 ΔSPI-1ΔSPI-2 (*orgA::Tet*, *ssaV::Kan*) (Broz et al., 2010)) bicolonized mice, 200 µL of overnight cultures were pelleted and washed three times in sterile phosphate-buffered saline (PBS). Cultures were combined, pelleted, and resuspended in 200 µL PBS for oral gavage. Overnight cultures of *St*$^{avir}$ were grown aerobically in LB + 50 µg/mL kanamycin and 2 µg/mL tetracycline on a shaker at 37 °C. Bicolonized mice were maintained on a standard diet, and sacrificed 5 days post-gavage.

For humanized mouse experiments, germ-free mice were colonized with a complete human microbiota by gavage of a fecal sample from a healthy anonymous donor. All mice were gavaged with the same inoculum, which was prepared from frozen stock mixed with filter-sterilized pre-reduced PBS, and given 40 days for the microbiota to equilibrate prior to diet shift

(Marcobal et al., 2013). After community equilibration, mice were either maintained on the standard diet or shifted to the MD diet for short-term (10 days) or long-term (14 weeks) diet experiments.

### *Tissue collection and processing*

Mice were sacrificed by $CO_2$ asphyxiation followed by cervical dislocation, and sections of the ileum, proximal colon, and distal colon were collected and placed in cassettes. Tissues were fixed and processed as previously described (Johansson and Hansson, 2012). Briefly, tissues were fixed by immersion in methacarn solution (60% methanol, 30% chloroform, 10% acetic acid) for 48 h, followed by two successive washes each in methanol for 35 min, ethanol for 30 min, and xylene for 25 min. Tissue samples within cassettes were then submerged in melted paraffin at 68 °C for 1 h, removed, and kept at room temperature until sectioning. Paraffin blocks were cut into 4 µm-thick sections and deparaffinized for immunofluorescence and FISH (Johansson and Hansson, 2012).

### *Immunostaining*

After deparaffinization and rehydration, slides were incubated in antigen retrieval solution (10 mM sodium citrate, pH 6.0) at 90 °C for 30 min. Slides were blocked with blocking buffer (5% bovine serum albumin in PBS) for 30 min at room temperature in a humid chamber. For mucus visualization, a polyclonal rabbit anti-mouse Muc2-specific antibody (Santa Cruz Biotechnology) was diluted 1:100 in blocking buffer, applied to the slide, and incubated at 4 °C for 4-18 h. Slides were washed three times in PBS. The secondary antibody (Alexa Fluor 488 Affini-pure donkey anti-rabbit IgG, Jackson ImmunoResearch) was diluted 1:100 in blocking buffer with 10 µg/mL

DAPI (Sigma), added to the sample, and incubated in the dark at 4 °C for 1 h. REG3β staining was performed similarly with a polyclonal sheep anti-mouse REG3β antibody (R&D Systems) diluted 1:100 in blocking buffer and donkey anti-sheep secondary antibody diluted 1:250 in blocking buffer (Alexa Fluor 594 Affini-pure, Jackson ImmunoResearch). After incubation with the secondary antibody and DAPI, slides were washed three times in PBS, dried, and mounted with Vectashield mounting medium (Vector Labs). Slides were stored at 4°C in the dark until imaging.

### *FISH preparation*

FISH was performed as previously described (Johansson and Hansson, 2012). Briefly, deparaffinized sections were incubated with FISH probes (Stanford Protein and Nucleic Acids facility) diluted to 10 ng/μL in hybridization buffer (0.9 M NaCl, 20 mM Tris-Hcl (pH 7.4), 0.01% sodium dodecyl sulfate, 10% formamide) at 47 °C for 3-18 h. Slides were incubated with FISH washing buffer (0.9 M NaCl, 20 mM Tris-HCl (pH 7.4)) preheated to 47 °C for 10 min, and washed three times in PBS. Samples were then incubated with 10 μg/mL DAPI in PBS in the dark for 1 h at 4 °C, washed three times in PBS, allowed to dry, and mounted in Vectashield mounting medium. The fluorescently labeled DNA probes used in this study are listed in **Table S2**.

### **Bifidobacterium *fluorescence in situ hybridization***

*Bifidobobacterium bifidum*, *Bifidobacterium longum*, and *Bifodobacterium breve* were grown anaerobically overnight in MRS media at 37 °C. Ten μL of pure culture of each bacterium were smeared on a lysine-coated slide, allowed to dry, and fixed in fresh methacarn solution for 15 min. Slides were then rinsed in PBS and hybridized and DAPI stained as described above.

### *Mass spectrometry*

Protein was extracted from fecal pellets, digested with trypsin, and analyzed by liquid chromatography tandem mass spectrometry as previously described (Lichtman et al., 2013). Two fractions from the C4 chromatography (40% and 60% acetonitrile) were analyzed and no technical replicates were collected. Data were filtered to a 1% peptide false discovery rate and 5% protein false discovery rate using in-house software (Huttlin et al., 2010). Fold changes between groups were calculated with QSpec v. 1.2.2 (Choi et al., 2008) and are reported as natural logarithms.

### **H. pylori *infection and imaging***

For liquid cultures, *H. pylori* strain PMSS1 was grown shaking at 120 rpm in Brucella broth (Difco) containing 10% fetal bovine serum (Gibco) in a microaerobic atmosphere at 37 °C. Mice were maintained in autoclaved microisolator cages and provided with sterile drinking water and chow *ad libitum*. Six-week-old male C57BL/6 mice were infected with a single oral dose of $10^8$ *H. pylori* and sacrificed at 2 weeks post infection. At the time of harvest, the forestomach was removed, and the glandular stomach was opened along the lesser curvature, laid flat, and divided along the greater curvature. For microscopy analysis, a longitudinal section at the midline along the greater curvature was used. Tissue samples were processed for confocal immunofluorescence as previously described with minor modifications (Howitt et al., 2011). Tissue samples were fixed in 2% paraformaldehyde in 100 mM phosphate buffer (pH 7.4) for 1 h. Tissue was embedded in 4% agarose and 100-200-μm sections were generated using a vibratome (Leica). Tissue sections were permeabilized in PBS with 3% bovine serum albumin,

1% saponin, and 1% Triton X-100 prior to staining. Rabbit anti-*H. pylori* antibodies were generated against fixed PMSS1 bacteria (Covance) and anti-rabbit secondary antibodies (AlexaFluor-conjugated, Invitrogen) were used to visualize bacteria in gastric tissue. After staining for *H. pylori*, the tissue was washed and stained with the fluorescein isothiocyanate-conjugated Rabbit-anti-phospho-Histone-H3 (Ser10) antibody (CST), to visualize mitotic cells. DAPI and AlexaFluor 594 phalloidin (Invitrogen) were used to visualize nuclei and the actin cytoskeleton, respectively. After labeling, overlapping images from longitudinal sections of the entire stomach antrum were acquired.

### *Gnotobiotic and humanized mice imaging*

Images were acquired on a Zeiss LSM 700 confocal microscope with the ZEN 2012 (**Figs. 5** and **S6**) or Zen 2009 (all other figures) software (Zeiss). With the exception of the images in **Figs. 2** and **7**, which were acquired at 20x, samples were imaged with a 63x oil-immersion objective with optical sections taken at 0.29-μm resolution. Images were acquired at a frame size of 2048x2048 with 16-bit depth. For **Figs. 5** and **S6**, tile scans were acquired on a microscope equipped with a mechanical stage. For all other figures, we acquired a collection of overlapping images by manual translation of the microscope stage, with several optical sections for each field of view to ensure that focal drift during stage translation would not preclude stitching of the images.

### *Automated stitching of large image collections*

To optimize stitching of large, tiled image sets, BacSpace first performs contrast matching for all pairs of overlapping images, which greatly improves precision. We generally imaged

regions that included a broad spectrum of features, for which stitching was challenging due to the large differences in contrast characteristics between the lumen and the epithelium. To address this obstacle, BacSpace performs several image transformations to improve stitching accuracy. First, BacSpace denoises through a relative noise transform (http://www.mathworks.com/matlabcentral/fileexchange/35556), and then equalizes contrast, both globally (via histogram equalization) and locally (via contrast-limited adaptive histogram equalization) using built-in MATLAB functions. Next, we improved the computational efficiency of stitching by down-sampling to lower image resolution to compute the approximate overlap positions, which are then used as seeds for stitching the full-resolution images. The computation of optimal overlaps, as well as the stitching itself, is performed by a freely available ImageJ (Schneider et al., 2012) plugin (Preibisch et al., 2009) that determines pairwise optimal image displacements via cross-correlation, and performs a global optimization of the displacements. The stitching plugin is integrated into BacSpace via the Java package MIJ. For the **Fig. 5B**, an additional normalization step was used to correct for a vertical gradient in illumination in each tile, which caused an artificial striping pattern in the stitched image. After stitching, the columns of the image were averaged to create a vertical intensity profile; each column of the image was then divided by this profile.

***Identification of host tissue***

To detect the epithelial boundary with high accuracy, BacSpace employs a mostly automated, iterative, multi-scale algorithm. Image intensities are initially re-mapped in order to saturate the signal of the host nuclei, reducing their prominence relative to other objects and thereby enhancing contrast. This contrast-enhanced image is then down-sampled and smoothed,

and an initial estimate for the epithelial contour is determined via thresholding to identify the bright epithelium, with user input to distinguish the true epithelial region from any bright debris. After the user defines the starting contour point, each point along the rest of the contour is refined sequentially, using the refined location of the previous contour point as a guide. To refine each contour point, BacSpace computes the magnitude of the image gradient and interpolates it along a local normal to the contour within a cutoff distance, typically 4 pixels. BacSpace then searches for peaks in this interpolated profile. The peak lying closest to the location of the previous contour point is selected as the new contour point. If no peaks are found, the location maximum value is used instead. After refinement, the contour is re-sampled so that the points are evenly spaced, and the contour is smoothed with a Gaussian filter.

The refined contour from the low-resolution version of the image is used as a seed to obtain a refined contour at higher resolution image in a similar manner. To span the scales from lowest resolution to the full resolution of the original image, BacSpace employs a geometric series of resolutions. Finally, the user can interactively curate the contour in the event of inaccuracies (for instance, a contour segment that followed the bright nuclear signal rather than dimmer autofluorescence). Any such corrections are then automatically refined using the algorithm outlined above.

***Identification of bacterial cells from DAPI fluorescence***

Since autofluorescent debris in the lumen of the gut is common, BacSpace removes debris based on differences in size and brightness from bacterial cells. First, large objects are subtracted using a grayscale morphological opening operation. The signal removed in this process is manually thresholded to generate an initial binary debris mask. BacSpace then uses a

secondary thresholding operation to identify pixels corresponding to debris that were contiguous with previously identified debris and were initially missed. The remaining luminal DAPI signal after debris subtraction is then identified as cells. For other channels (e.g., FISH images), debris is identified similarly. Any debris object identified in one channel is removed from the rest of the channels.

### *Image straightening*

To associate pixels in the image to particular contour positions, BacSpace computes a Voronoi diagram to find the closest contour point to each pixel. Each point on the contour corresponds to a row in the straightened image. To compute the intensity of the pixel (*i*,*j*) in the straightened image, which is the mean image intensity a distance *j* away from contour point *i*, we collect all points in the original image that lie in the Voronoi cell of the *i*th contour point. The pixel intensities of these points are then averaged with a Gaussian weight $\exp(-(d-j)^2/2\sigma^2)$, where *d* is the actual distance from the contour, and σ, the kernel width, is typically 5 pixels. When there are no contour points within the bin corresponding to a distance *j* in the Voronoi cell for a contour point *i*, a null value is assigned to the straightened image at point (*i*,*j*).

### *Considerations for measurements of methacarn-fixed colon mucus*

Previous studies of mucus structure in fixed sections have either employed widefield fluorescence microscopy or generated maximum intensity projections of *z*-stacks. To maximize resolution of bacterial cells, we analyzed single optical slices of laser scanning confocal microscopy *z*-stack images. In most instances, a single optical slice rendered a less dense, but faithful representation of the mucus (**Fig. S1C**), while in some images (**Fig. S1D**) a single optical

slice exaggerated the appearance of fixation and sectioning artifacts that were not present in maximum intensity projections. Regardless, because our mucus thickness measurements are based on gradients in fluorescence values, measurements are unlikely to be affected by any gradual variations in intensity resulting from the use of single optical slices as compared with maximum intensity projections.

Fixation and tissue processing induces the formation of certain artifacts: shrinkage of the inner mucus layer and sectioning cause the formation of unstained black spaces between the mucus and the epithelium not present *in vivo* (Johansson et al., 2008), while the less organized inner face of the attached mucus layer where the MUC2 mucin is expanding may expand during fixation and appear larger than the native state (Ambort et al., 2012). Because we cannot distinguish between actual artifacts of sample treatment and artifacts of a shallow focal plane intrinsic to confocal microscopy, and because all samples were treated equally, we did not remove any black space that appears between the epithelium and mucus from measurements of mucus thickness. However, a script to enable users to subtract these unstained regions within measurements has been included with BacSpace.

### *Measurement of bacterial density and layer thickness*

To measure the density of bacteria as a function of distance from the epithelium, BacSpace averages the non-null values in each column of the straightened image. To measure the variation in bacteria proximity to host tissue along the length of the epithelium, the same procedure is applied to the rows of the straightened image; the user can determine a maximum distance from the epithelium. To measure the thickness of the mucus layer, or the closeness of bacteria to the epithelium, BacSpace looks for edges in the straightened image. BacSpace allows

the user to input a rough location of the edge and then refines it as follows. First, the null values in the straightened image are interpolated, and the image is smoothed. BacSpace then seeks the maximum gradient in the direction perpendicular to the epithelium that lies within a given range (typically 100 pixels) of the user-provided estimate. Edge locations corresponding to interpolated regions of the gradient are excluded from average widths.

### *Segmentation of FISH images*

After BacSpace processing to identify the epithelium and remove non-cellular debris, each channel was independently segmented by a two-step thresholding method. An initial threshold was used to identify bright objects, and a less restrictive, secondary threshold was applied to the regions around those objects to ensure that complete cells were identified rather than bright regions at the center of cells with a dimmer periphery. We identified *St* cells as objects that appeared in both the DAPI and Cy5 channels; to reduce false-negative rejection of *St* cells with dim DAPI labeling, we applied a more permissive threshold to the DAPI masks used to identify *St* cells, but used more accurate, restrictive thresholds for identifying *Bt* cells. Some cells appeared larger in the DAPI masks than in the *St* masks, resulting in a halo of *Bt* surrounding some *St* cells; these regions were identified based on their proximity to *St* cells, and the masks re-assigned so that they were considered part of *St* cells. Finally, small objects were removed through morphological opening, and any remaining debris objects were manually removed.

### *Identification of Bacteroidales and* Akkermansia muciniphila *clusters*

The Bac303 FISH probe, in addition to labeling Bacteroidales cells, formed puncta that were both brighter and smaller than cells (see **Fig. 5D**). Thus, we first masked these puncta by thresholding. Segmentation of images was performed as described above. Bacteroidales cells

were identified as the overlap of the masks of the DAPI and Bac303 channels. *A. muciniphila* cells were identified as the overlap of the DAPI and Muc1437 channels. A morphological closing operation was used on the resulting mask to ensure that cells in close proximity to each other (<0.4 μm) were counted as a single cluster.

### *Cluster-size analysis*

After processing through BacSpace, FISH images were contrast-adjusted and scaled to values between 0 and 1. For each pixel, we considered circular neighborhoods of varying radii; For each neighborhood, we computed 1) the fraction of total area (weighted by fluorescence intensity) occupied by each taxon, and 2) the fraction of all bacteria (also weighted by fluorescence intensity) that each taxon constituted. The effective cluster size was the smallest neighborhood radius for which either the fraction of total area was less than 0.3, or the fraction of total bacteria was less than 0.7. Total bacteria were computed, for each pixel, as the maximum of the signal from both taxa. This algorithm was implemented using image convolution, and for speed of computation, Gaussian rather than disk-shaped kernels were used.

### *Measurement of* Hp *and mitotic cell location*

After background subtraction via morphological opening, we applied a threshold to each channel of straightened image output by BacSpace using Otsu's method to determine a threshold value. Due to cross-talk between fluorescence channels, we removed all *Hp* pixels that overlapped with pixels identified as mitotic cells. Each row in the straightened image corresponds to a point on the mucosal boundary, and we measured the median position of bright pixels for

each row. For computing statistics, we sampled every 30 μm, approximately the width of a gastric gland, to ensure that we did not double-count glands.

## Supplemental Figures

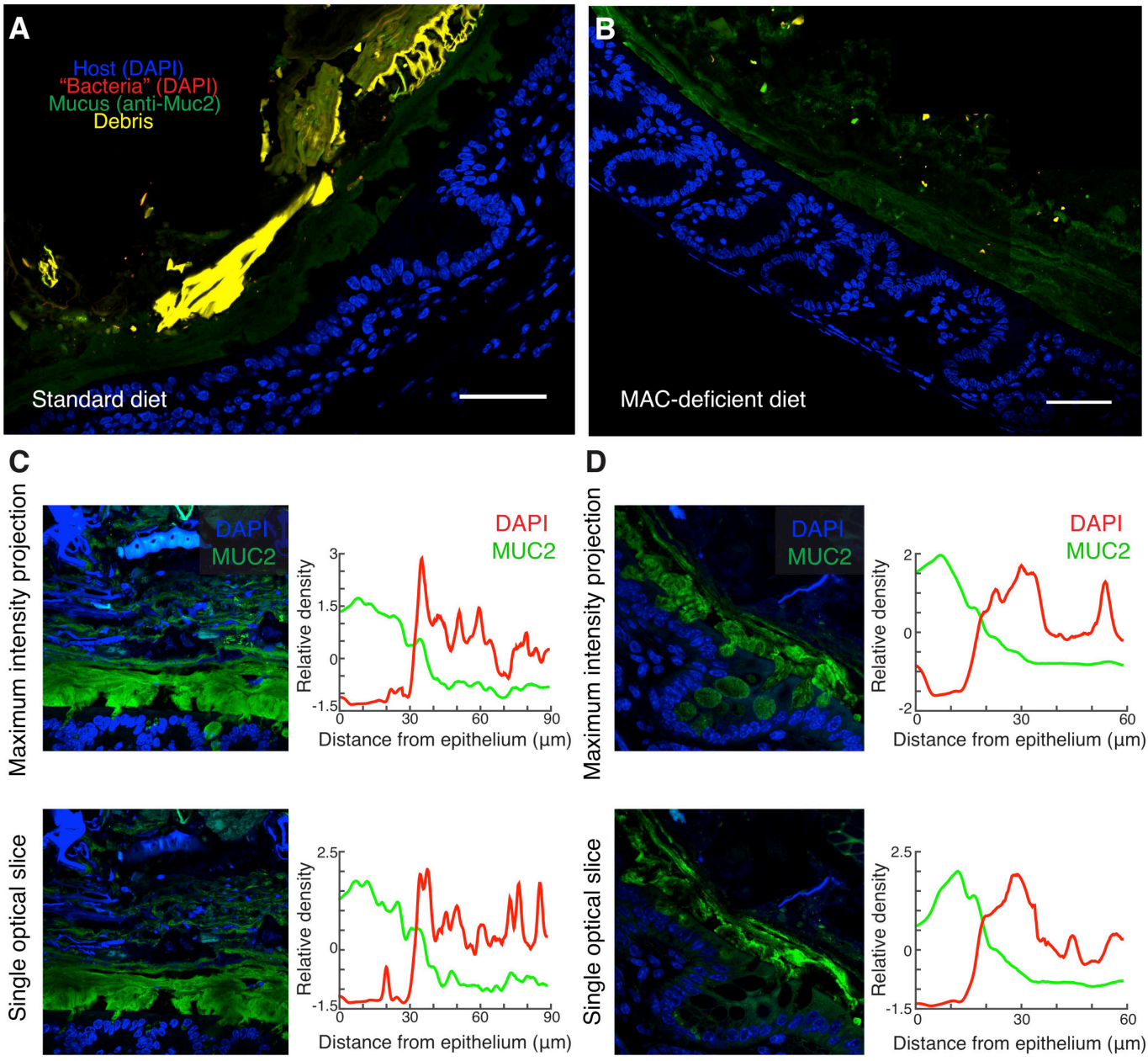


**Figure S1, Related to Figure 1. Validation of imaging and processing methods.**
(A, B) MUC2- and DAPI-labeled images from germ-free mice processed with BacSpace. BacSpace removes non-specific luminal fluorescence by segmenting out large, bright objects (yellow). Images from germ-free mice suggest a low false positive rate, wherein small, dim autofluorescent material will be segmented as bacteria (red). Consistent with previous reports (Petersson et al., 2011), we found mucus quality in germ-free mice to be diffuse and poorly organized, apparent from the dim Muc2 signal (green). Host nuclei are stained with DAPI (blue). Scale bar = 50 μm.
(C, D) MUC2- and DAPI-labeled images from *Bt*-monocolonized mice fed standard chow.

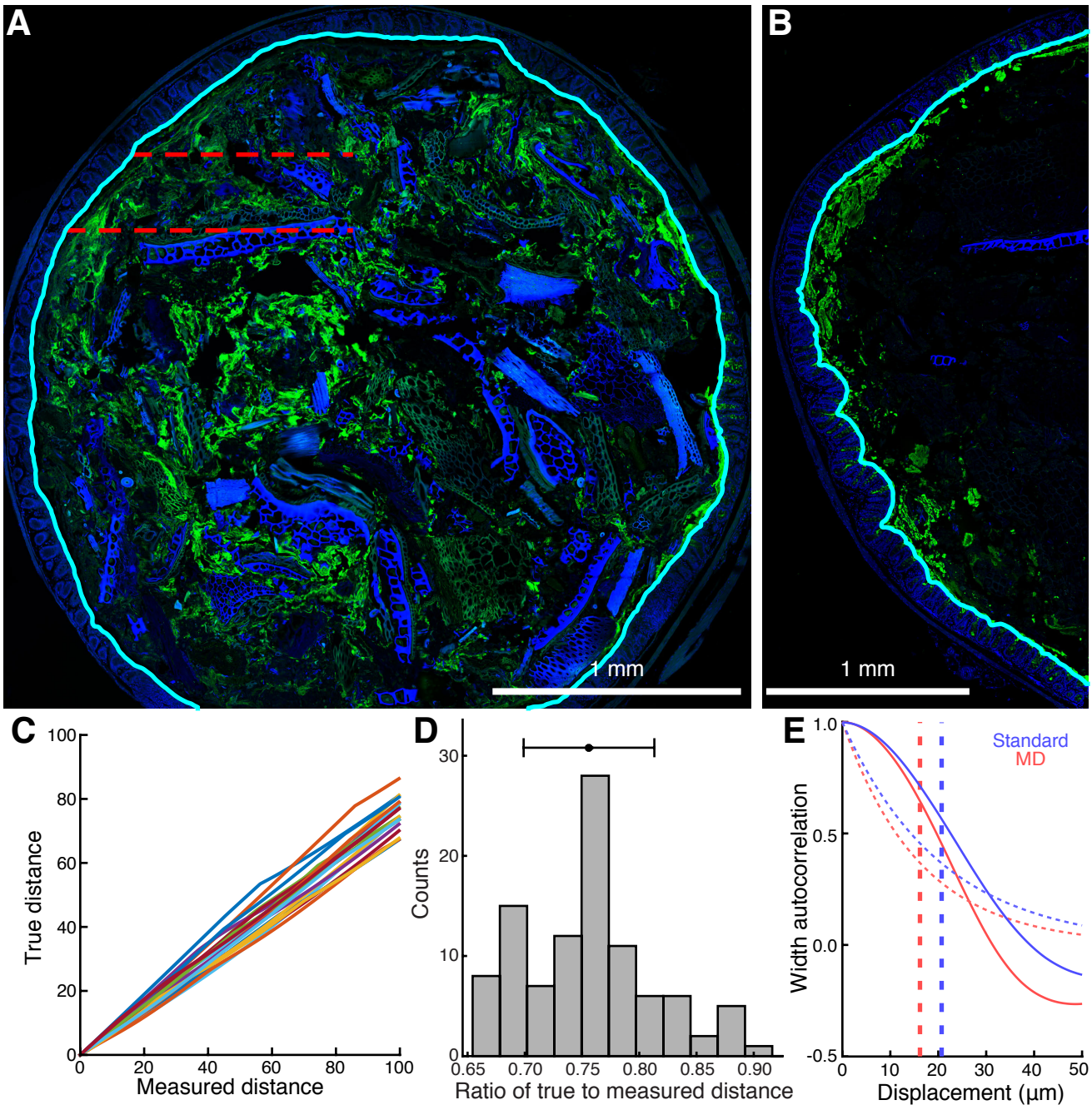


**Figure S2, Related to Figure 2. Effect of 3D geometry on measurements of mucus thickness.**

(A,B) Stitched images of transverse sections of the distal colon from mice fed a standard (A) or MD (B) diet. Samples were labeled with DAPI (blue) and a MUC2 antibody (green), and the

computationally identified apical boundary is overlaid in cyan. In (A), the range over which simulated longitudinal sections were computed in (C) is indicated by red lines.
(C) True distance to epithelial boundary vs distance measured in a longitudinal section, for simulated longitudinal sections taken at positions indicated in (A) by red dashed lines.
(D) Histogram of the ratio of true to measured distance, computed by fitting curves in (C) to a line passing through the origin.
(E) Autocorrelation of the width of the inner mucus layer along the length of the epithelium for sections in **Fig. 2A** (blue) and (**2B**) (red), along with fits to an exponential function. Dashed vertical lines indicate the autocorrelation length scales, determined by fitting to a single exponential.

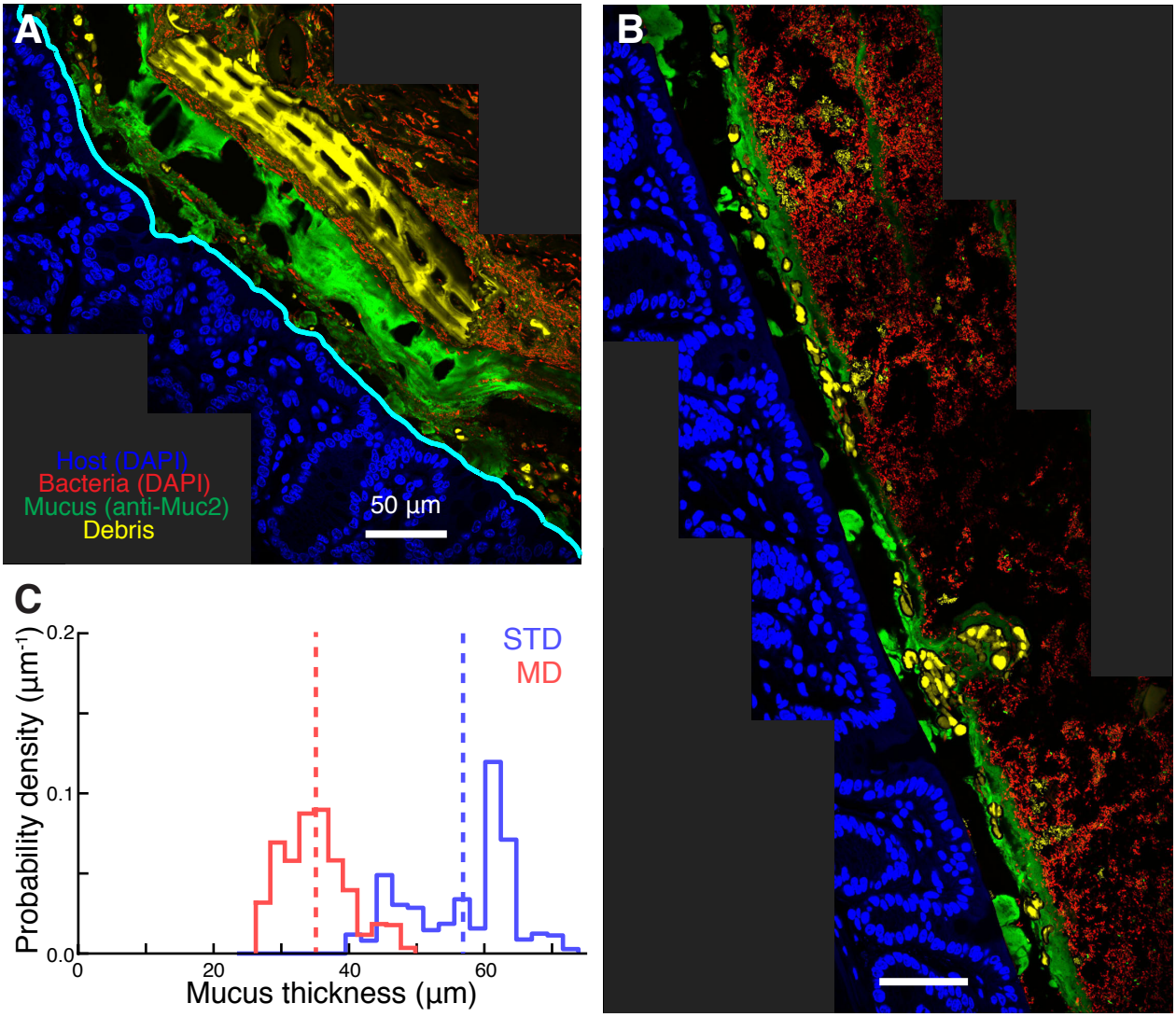


**Figure S3, Related to Figure 3. Dietary effect on humanized mouse colonic mucus thickness.**

(A and B) MUC2- and DAPI-labeled distal colon sections from humanized mice fed a standard (A) or MD (B) diet, distinct from the mice shown in **Fig. 4**. Images were processed in BacSpace to remove debris (yellow) from measurements. Luminal DAPI signal that is not subtracted as debris is false-colored red. Scale bar = 50 µm.

(C) Distribution of mucus layer thickness in mice fed a standard (blue) or MD (red) diet. Dashed line represents the mean.

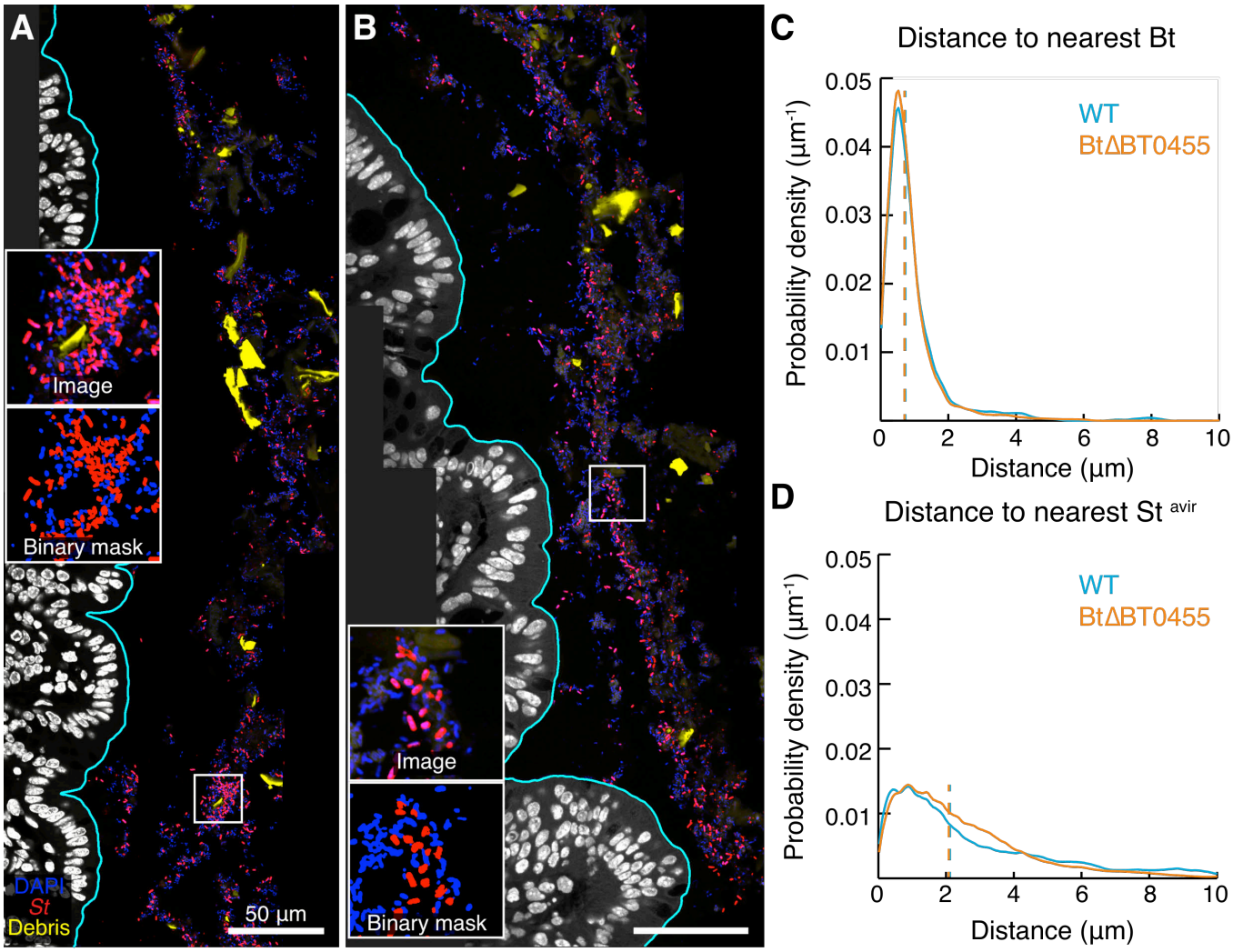


**Figure S4, Related to Figures 4 and 5. Spatial organization in a cross-feeding intestinal community.**

(A,B) Proximal colon of gnotobiotic mice bicolonized with *St*$^{avir}$ and either wildtype *Bt* (A) or *Bt*ΔBT0455, a sialidase deficient mutant (Ng et al., 2013). (B), labeled with DAPI (blue) and *Salmonella*-specific Cy5 FISH probes (red). Insets: comparison of fluorescence images (top) with computed binary masks used for species identification (below).

(C,D) Distribution of interspecies distances in mice colonized with *St* and wild-type (WT) *Bt* (cyan) and *Bt*ΔBT0455 (orange), measured from binary masks in (A) and (B). Dashed lines indicate median distances. Distances from *St* to nearest *Bt* (C) and from *Bt* to nearest *St* (D) were not affected by *Bt*'s ability to liberate sialic acid.

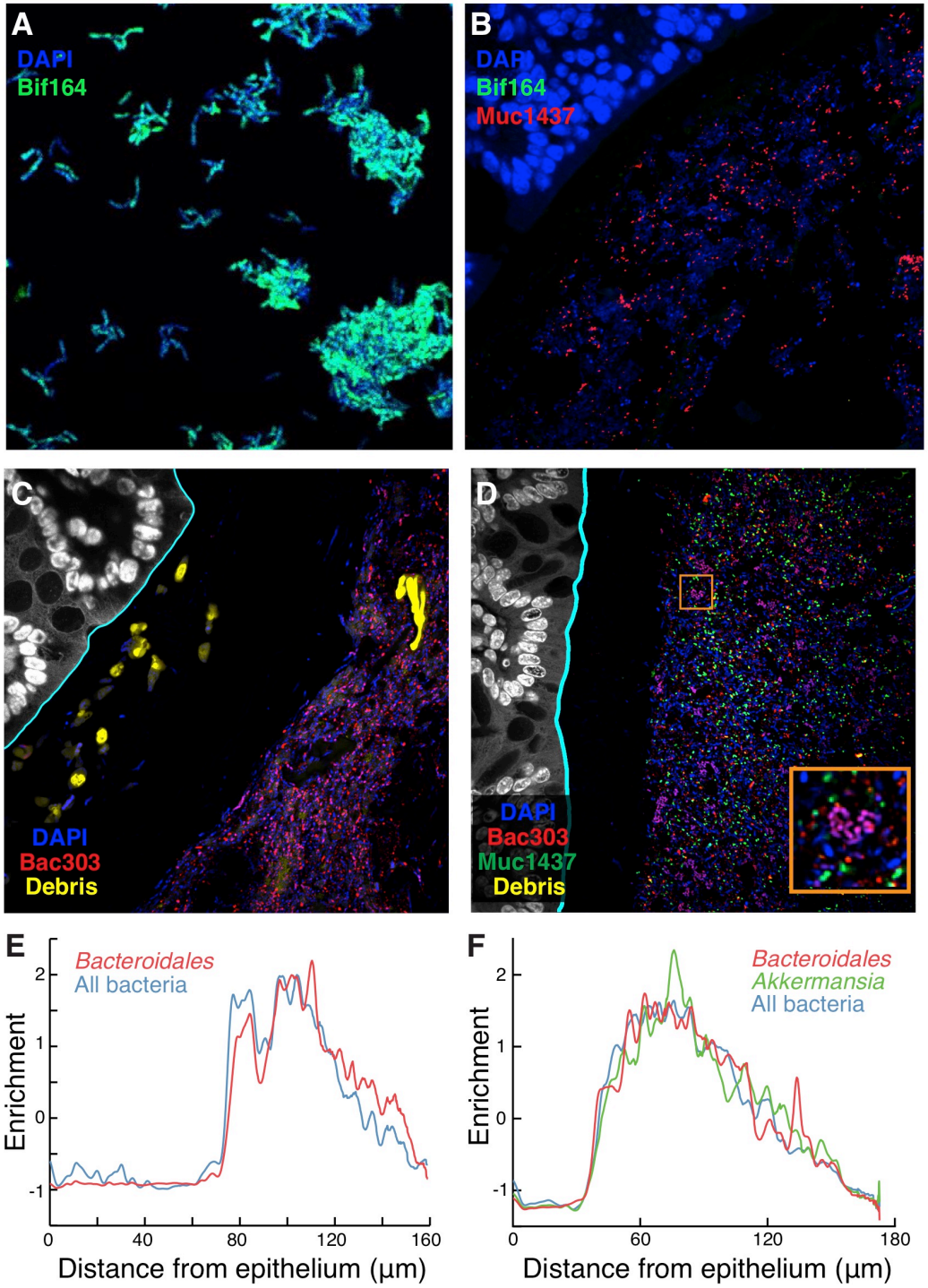


**Figure S5, Related to Figure 5. Fluorescence *in situ* hybridization of known mucin utilizers in humanized mice.**

(A) The Bif164 FISH probe robustly labeled methacarn-fixed pure cultures of *Bifidobacterium bifidum* (A), *Bi. breve* (data not shown) and *Bi. longum* (data not shown).

(B) Hybridization with Bif164 was undetectable in MD-fed (B) or standard diet-fed mice (data not shown). Hybridization with *Akkermansia* probe Muc1437 is shown for FISH labeling control.

(C,D) Representative single images of FISH probe-hybridized distal colon sections from humanized mice fed a standard (C) or MD (D) diet. All bacteria are labeled with luminal DAPI signal (blue), Bacteroidales are labeled with the Bac303 probe (red), and *Akkermansia* are labeled with the Muc1437 probe (green).
(E,F) Enrichment of Bacteroidales and *Akkermansia* FISH probe signals and total bacterial signal, as a function of distance from the epithelium, for the images in panels (C) and (D), respectively. *Akkermansia* enrichment is not plotted in (E), as Muc1437 signal was undetectable in standard diet-fed mice. Curves are mean-subtracted and divided by the standard deviation for normalization.

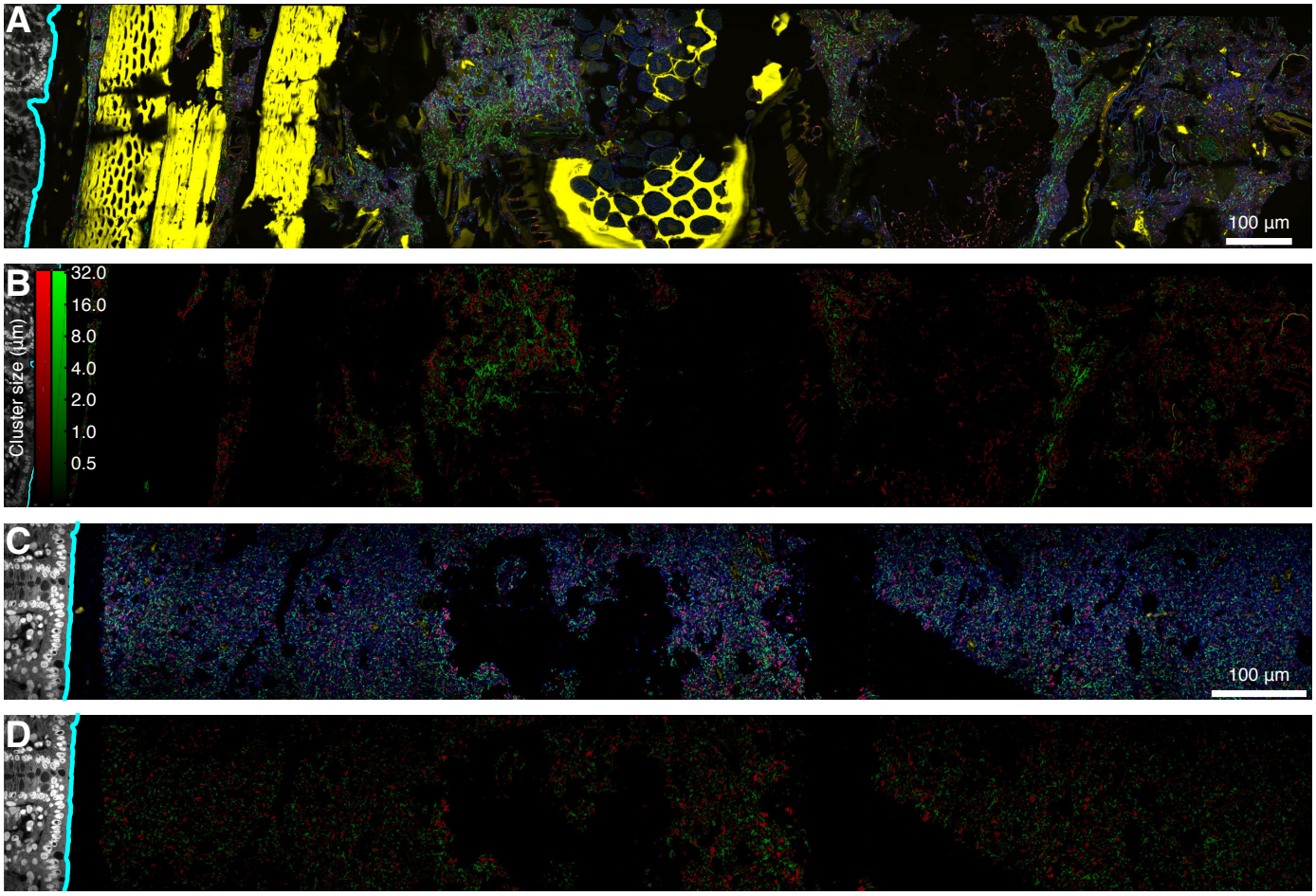


**Figure S6, Related to Figure 5. Clustering in a complex bacterial community.**
(A,C) Distal colon sections from humanized mice, distinct from those in **Fig. 5**, fed a standard (A) or MD (C) diet. Sections were stained with DAPI (epithelial signal, white; luminal signal, blue), and FISH probes Bac303 (Bacteroidales, red) and LGC354A-C and Erec482 (Firmicutes, green). Computed epithelial boundary is outlined in cyan, and debris appears in yellow.
(B,D) Pixel intensity corresponds to measured cluster size in images (A) and (C), respectively, for Bacteroidales (red) and Firmicutes (green). Intensities are on a log scale for clarity, and the same color scale is used for both images.

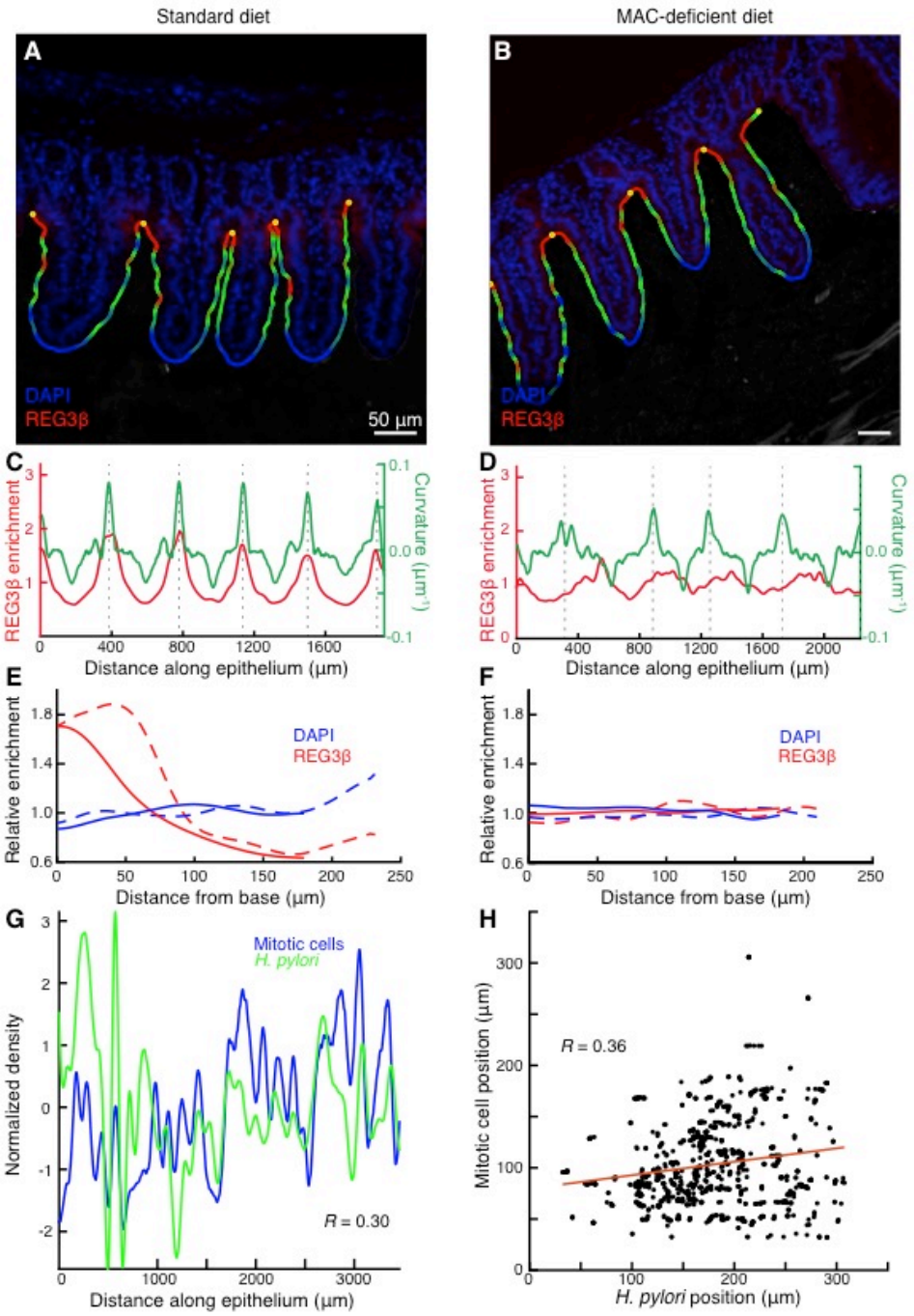


**Figure S7, Related to Figures 6 and 7. Repeatability of BacSpace quantification.**
(A, B) Ileum of mice, distinct from those in **Fig. 6**, fed a standard (A) or MD (B) diet. Samples were stained with DAPI (blue) and REG3β antibody (red). The computationally identified outline is colored by contour curvature, and the computationally identified base of the villi is shown with yellow dots. Villi that are not continuous with the imaged epithelium due to sectioning angle have been removed from the analysis.

(C, D) Variation in enrichment of REG3β normalized to the mean, as determined by fluorescence intensity (red) and curvature of epithelial contour (green) along the length of the epithelium. Dashed lines indicate the base of the villi.
(E, F) Distribution of REG3β signal (red) and cell nuclei (DAPI, blue) relative to the base of the villi in a standard (E) or MD (F) diet. The corresponding signals for the mice shown in **Fig. 6** are shown in dotted lines.
(G) Colocalization of mitotic cells and *Hp* 2 weeks post-infection, in a mouse distinct from that in **Fig. 7**, measured with BacSpace. Variation in density (normalized by z-score) of Hp (green) and mitotic cells (blue) along the length of the antrum indicates correlation (Pearson correlation coefficient 0.30) between the Hp and mitotic cell localization.
(H) Average distance from submucosal boundary of Hp and mitotic cells at positions along the antrumn for which both are present.

| Uniprot Identifier | Annotation | Protein length (aa) | Secreted | Hum Std 1 | Hum Std 2 | Hum MD 1 | Hum MD 2 | GF MD 1 | GF MD 2 | GF MD 3 | HumMD v GFMD log fold change | HumStd v. HumMD log fold change |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P35230 | Regenerating islet-derived protein 3-beta | 175 | Yes | 9 | 5 | 43 | 37 | 17 | 22 | 27 | -2.769 | 2.011 |
| P62821 | Ras-related protein Rab-1A | 205 | No | 0 | 0 | 2 | 8 | 0 | 2 | 0 | -2.058 | 2.058 |
| Q8BWF0 | Succinate-semialdehyde dehydrogenase | 523 | No | 0 | 0 | 0 | 9 | 0 | 0 | 0 | -2.355 | 2.099 |
| Q9CXY6 | Interleukin enhancer-binding factor 2 | 390 | No | 0 | 0 | 5 | 8 | 0 | 2 | 0 | -2.289 | 2.188 |
| P80313 | T-complex protein 1 subunit eta | 544 | No | 0 | 0 | 6 | 5 | 1 | 2 | 1 | -2.101 | 2.162 |
| Q9R0Q3 | Transmembrane emp24 domain-containing protein 2 | 201 | No | 0 | 0 | 8 | 3 | 0 | 1 | 0 | -2.438 | 2.117 |
| P49312 | Heterogeneous nuclear ribonucleoprotein A1 | 320 | No | 0 | 0 | 5 | 9 | 0 | 0 | 0 | -2.692 | 2.384 |
| A2NU19 | VH coding region (Fragment) | 119 | No | 0 | 0 | 0 | 12 | 0 | 0 | 0 | -2.785 | 2.357 |
| P10126 | Elongation factor 1-alpha 1 | 462 | No | 5 | 6 | 32 | 62 | 7 | 17 | 7 | -2.088 | 2.48 |
| Q9JHU4 | Cytoplasmic dynein 1 heavy chain 1 | 4644 | No | 0 | 0 | 7 | 7 | 0 | 0 | 0 | -2.68 | 2.302 |
| P59999 | Actin-related protein 2/3 complex subunit 4 | 168 | No | 0 | 0 | 10 | 5 | 0 | 0 | 0 | -2.683 | 2.416 |
| O88569 | Heterogeneous nuclear ribonucleoproteins A2/B1 | 353 | No | 0 | 0 | 5 | 9 | 1 | 2 | 0 | -2.363 | 2.393 |
| P61979 | Heterogeneous nuclear ribonucleoprotein K | 463 | No | 0 | 0 | 5 | 13 | 0 | 3 | 2 | -2.095 | 2.738 |
| P62267 | 40S ribosomal protein S23 | 143 | No | 0 | 0 | 8 | 10 | 0 | 2 | 0 | -2.532 | 2.643 |
| Q9CPQ3 | Mitochondrial import receptor subunit TOM22 homolog | 142 | No | 0 | 0 | 9 | 8 | 0 | 2 | 0 | -2.519 | 2.65 |
| Q68FD5 | Clathrin heavy chain 1 | 1675 | No | 0 | 1 | 11 | 11 | 0 | 3 | 0 | -2.556 | 2.84 |
| P35564 | Calnexin | 591 | No | 0 | 0 | 3 | 20 | 0 | 3 | 1 | -2.577 | 2.831 |
| Q78PY7 | Staphylococcal nuclease domain-containing protein 1 | 910 | No | 0 | 0 | 17 | 19 | 0 | 8 | 2 | -2.197 | 3.351 |
| P53026 | 60S ribosomal protein L10a | 217 | No | 0 | 0 | 6 | 27 | 1 | 4 | 1 | -2.766 | 3.325 |
| Q9CXW4 | 60S ribosomal protein L11 | 178 | No | 0 | 0 | 38 | 32 | 0 | 3 | 1 | -3.669 | 3.852 |
| D3YZH3 | MCG128013 | 192 | No | 0 | 0 | 23 | 40 | 0 | 3 | 0 | -3.581 | 3.89 |
| Q3TFD0 | Serine hydroxymethyltransferase | 501 | No | 0 | 0 | 49 | 81 | 1 | 12 | 2 | -3.148 | 4.477 |
| O09049 | Regenerating islet-derived protein 3-gamma | 174 | Yes | 12 | 5 | 7 | 25 | 6 | 17 | 14 | -2.646 | 0.915 |

**Table S1, Related to Fig. 6. Spectral counts of proteins upregulated by 2 log-fold in humanized long-term MD diet-fed mice, compared to germ-free mice on an MD diet and humanized mice on a standard diet**. REG3γ was not enriched, but is included in the table for comparison.

| Specificity | Probe | 5' fluorophore | Sequence (5'-3') | Source |
|---|---|---|---|---|
| *Akkermansia* | Muc1437 | Cy3 | CCTTGCGGTTGGCTTCAGAT | (Derrien et al., 2008) |
| Bacteroidales | Bac303 | Cy5 | CCAATGTGGGGGACCTT | (Manz et al., 1996) |
| *Bifidobacterium* | Bif164 | FITC | CATCCGGCATTACCACCC | (Langendijk et al., 1995) |
| Enterobacteriaceae | Ent1251 | Cy5 | TGCTCTCGCGAGGTCGCTTCTCTT | (Ootsubo et al., 2002) |
| Firmicutes | LGC354A<br>LGC354B<br>LGC354C<br>Erec482 | Cy3 | GGAAGATTCCCTACTGC<br>CGGAAGATTCCCTACTGC<br>CCGAAGATTCCCTACTGC<br>GCTTCTTAGTCAGGTACCG | (Meier et al., 1999)<br><br><br>(Franks et al., 1998) |
| *Salmonella* (23S rRNA) | Sal1713 | Cy5 | GTGCATCCACTTCACTAA | (Nordentoft et al., 1997) |

**Table S2, related to Figs. 4 and 5. Probes used in this study**. Previously validated probes were used in this study to identify subsets of bacteria in bicolonized or humanized mice. For identification of Salmonella in $St^{avir}$ and *Bt* bicolonized mice, the Ent1251 and Sal1713 probes were combined.

## Supplemental References


Ambort, D., Johansson, M.E., Gustafsson, J.K., Nilsson, H.E., Ermund, A., Johansson, B.R., Koeck, P.J., Hebert, H., and Hansson, G.C. (2012). Calcium and pH-dependent packing and release of the gel-forming MUC2 mucin. Proceedings of the National Academy of Sciences of the United States of America *109*, 5645-5650.

Broz, P., Newton, K., Lamkanfi, M., Mariathasan, S., Dixit, V.M., and Monack, D.M. (2010). Redundant roles for inflammasome receptors NLRP3 and NLRC4 in host defense against Salmonella. The Journal of experimental medicine *207*, 1745-1755.

Choi, H., Fermin, D., and Nesvizhskii, A.I. (2008). Significance analysis of spectral count data in label-free shotgun proteomics. Molecular & cellular proteomics : MCP *7*, 2373-2385.

Derrien, M., Collado, M.C., Ben-Amor, K., Salminen, S., and de Vos, W.M. (2008). The Mucin degrader Akkermansia muciniphila is an abundant resident of the human intestinal tract. Applied and environmental microbiology *74*, 1646-1648.

Franks, A.H., Harmsen, H.J., Raangs, G.C., Jansen, G.J., Schut, F., and Welling, G.W. (1998). Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Applied and environmental microbiology *64*, 3336-3345.

Howitt, M.R., Lee, J.Y., Lertsethtakarn, P., Vogelmann, R., Joubert, L.M., Ottemann, K.M., and Amieva, M.R. (2011). ChePep controls Helicobacter pylori Infection of the gastric glands and chemotaxis in the Epsilonproteobacteria. mBio *2*.

Huttlin, E.L., Jedrychowski, M.P., Elias, J.E., Goswami, T., Rad, R., Beausoleil, S.A., Villen, J., Haas, W., Sowa, M.E., and Gygi, S.P. (2010). A tissue-specific atlas of mouse protein phosphorylation and expression. Cell *143*, 1174-1189.

Johansson, M.E., and Hansson, G.C. (2012). Preservation of mucus in histological sections, immunostaining of mucins in fixed tissue, and localization of bacteria with FISH. Methods in molecular biology (Clifton, NJ) *842*, 229-235.

Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proceedings of the National Academy of Sciences of the United States of America *105*, 15064-15069.

Langendijk, P.S., Schut, F., Jansen, G.J., Raangs, G.C., Kamphuis, G.R., Wilkinson, M.H., and Welling, G.W. (1995). Quantitative fluorescence in situ hybridization of Bifidobacterium spp. with genus-specific 16S rRNA-targeted probes and its application in fecal samples. Applied and environmental microbiology *61*, 3069-3075.

Lichtman, J.S., Marcobal, A., Sonnenburg, J.L., and Elias, J.E. (2013). Host-centric proteomics of stool: a novel strategy focused on intestinal responses to the gut microbiota. Molecular & cellular proteomics : MCP *12*, 3310-3318.

Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K.H. (1996). Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. Microbiology (Reading, England) *142 ( Pt 5)*, 1097-1106.

Marcobal, A., Kashyap, P.C., Nelson, T.A., Aronov, P.A., Donia, M.S., Spormann, A., Fischbach, M.A., and Sonnenburg, J.L. (2013). A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. The ISME journal *7*, 1933-1943.

Meier, H., Amann, R., Ludwig, W., and Schleifer, K.H. (1999). Specific oligonucleotide probes for in situ detection of a major group of gram-positive bacteria with low DNA G + C content. Systematic and applied microbiology *22*, 186-196.

Ng, K.M., Ferreyra, J.A., Higginbottom, S.K., Lynch, J.B., Kashyap, P.C., Gopinath, S., Naidu, N., Choudhury, B., Weimer, B.C., Monack, D.M.*, et al.* (2013). Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. Nature *502*, 96-99.

Nordentoft, S., Christensen, H., and Wegener, H.C. (1997). Evaluation of a fluorescence-labelled oligonucleotide probe targeting 23S rRNA for in situ detection of Salmonella serovars in paraffin-embedded tissue sections and their rapid identification in bacterial smears. Journal of clinical microbiology *35*, 2642-2648.

Ootsubo, M., Shimizu, T., Tanaka, R., Sawabe, T., Tajima, K., Yoshimizu, M., Ezura, Y., Ezaki, T., and Oyaizu, H. (2002). Oligonucleotide probe for detecting Enterobacteriaceae by in situ hybridization. Journal of applied microbiology *93*, 60-68.

Petersson, J., Schreiber, O., Hansson, G.C., Gendler, S.J., Velcich, A., Lundberg, J.O., Roos, S., Holm, L., and Phillipson, M. (2011). Importance and regulation of the colonic mucus barrier in a mouse model of colitis. American journal of physiology Gastrointestinal and liver physiology *300*, G327-333.

Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics *25*, 1463-1465.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., and Harada, H. (1999). Fluorescence in situ hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Applied and environmental microbiology *65*, 1280-1288.

Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., and Sonnenburg, J.L. (2010). Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. Cell *141*, 1241-1252.

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science (New York, NY) *307*, 1955-1959.