Table of contents

Supplemental Figures

Figure S1

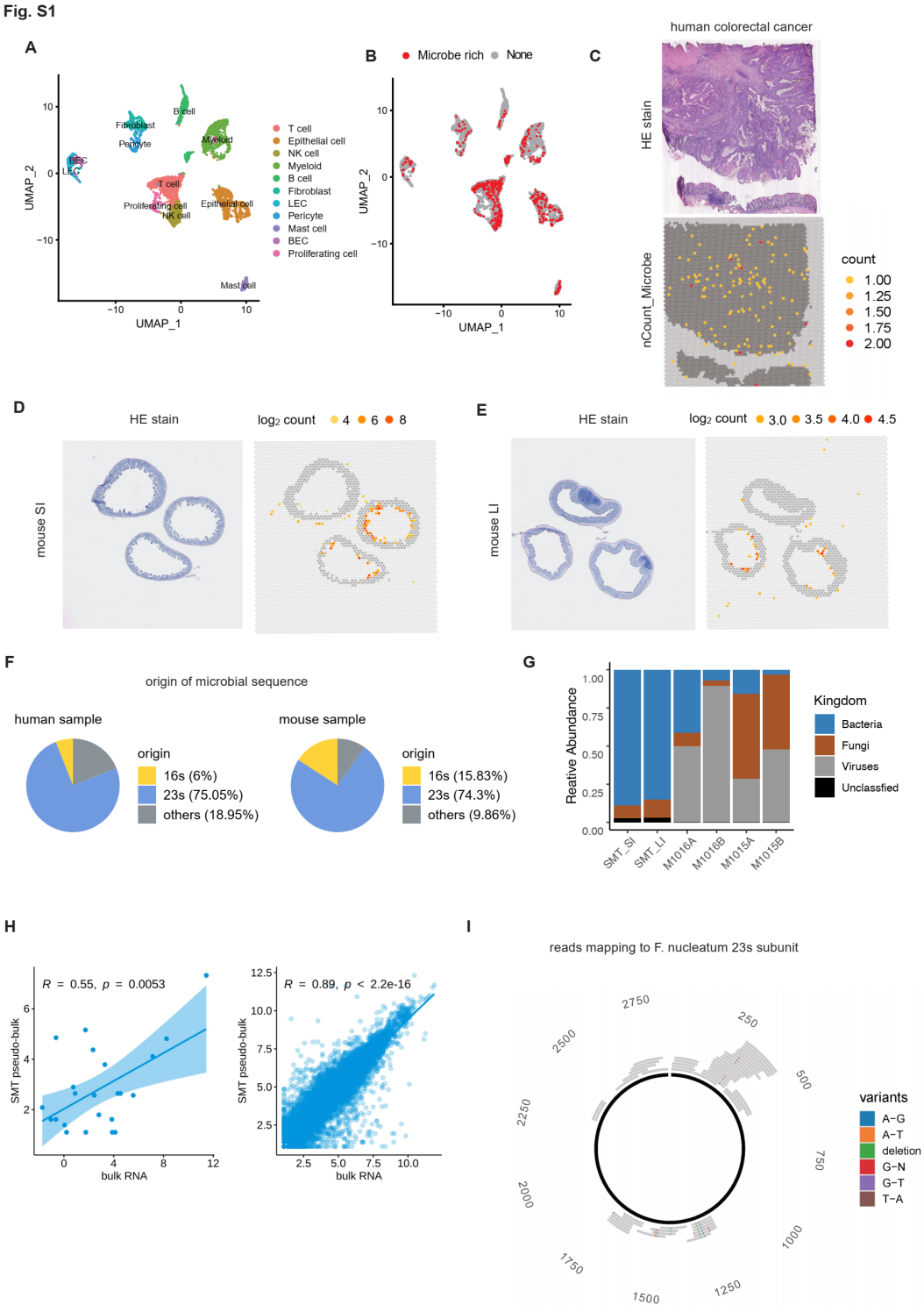

Fig. S1. Related to Fig 1. (A) TSNE plot showing single cell sequencing of an unsorted CRC sample, clustered into 11 major cell types. **(B)** Cells in (A) were highlighted in red if they were enriched in microbial sequences (>0.02%). **(C)** H&E stain (top) and microbial UMI count (bottom) for patient M1015. **(D-E)** H&E staining of the tissue section (left) and corresponding log scaled microbial sequence count per spot captured by SMT (right) for mouse small intestine (D) and large intestine (E). Note that 3 samples were placed in each window. **(F)** Pie chart showing origin of microbial sequence captured, for human samples (left) and murine samples (right). **(G)** Microbial composition cross human and murine samples at kingdom level. **(H)** correlation between SMT generated pseudo-bulk measurements (x-axis) and bulk RNA-seg measurements (y-axis), for microbial abundance at genus level (left) and host gene expression (right). Spearman's correlation coefficients and *p*-values were shown. **(I)** Representative read alignment to the *Fusobacterium* 23S gene, for M1016A.

Figure S2

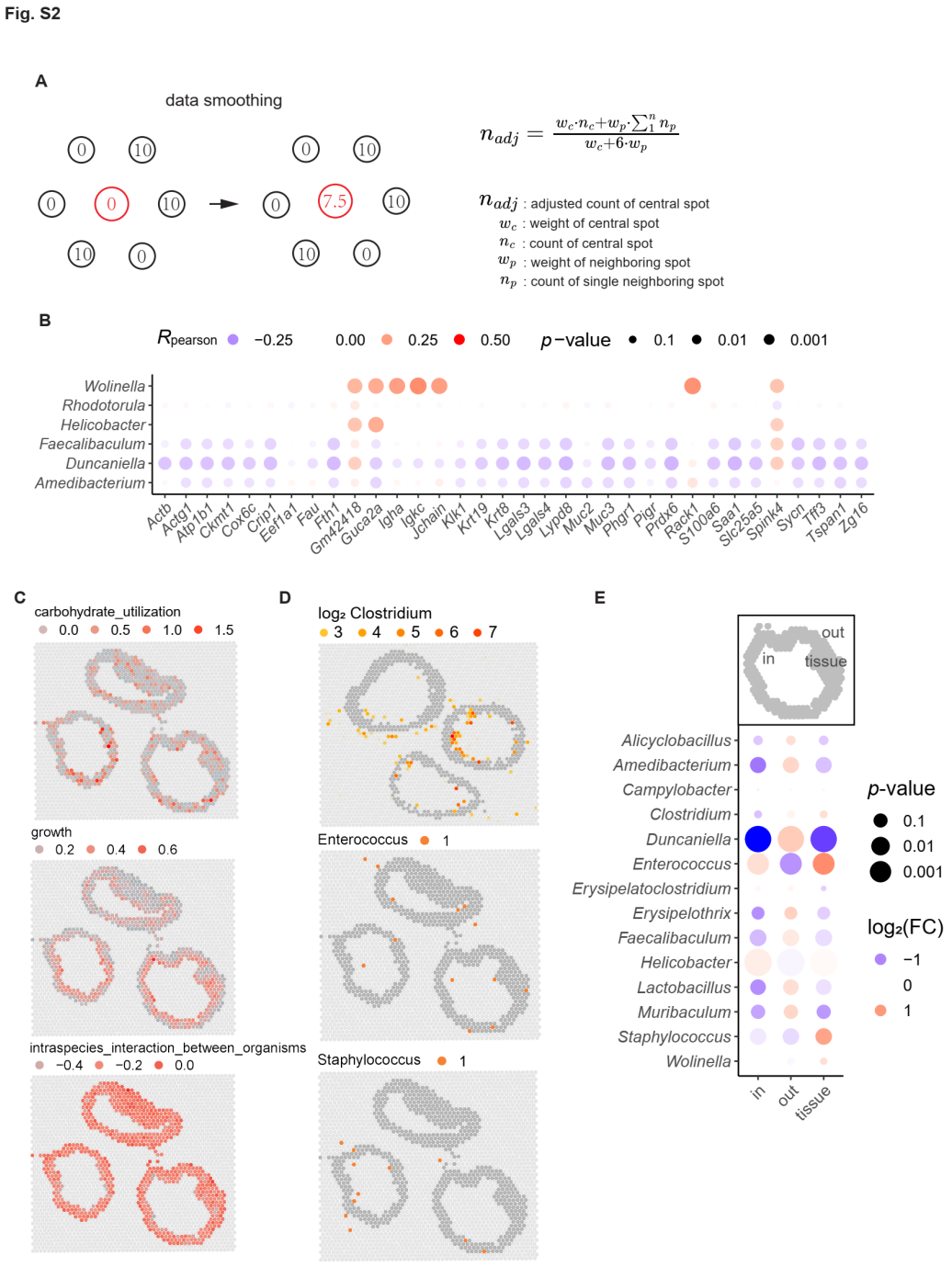

Fig. S2. Related to Fig 2. (A) Schema of kernel smoothing performed before microbe-host gene interaction detection. **(B)** Representative dot plot showing the spatial correlation between microbe and host genes in mouse large intestine, *p*-values shown were before multi-test correction. **(C)** Representative spatial feature plot showing host pathways that did

not show inter sample difference, in mouse large intestine. **(D)** Spatial feature plot for *Clostridium* in the small intestine, *Enterococcus* and *Staphylococcus* in the large intestine, showing their elevated concentration on tissue and further out, suggesting their trans-intestinal wall mobility. **(E)** Dot plot showing fold change and *p*-value for microbial genera that were differentially distributed in the three regions designated 'in', 'tissue' and 'out', by Chi-square test, for mouse large intestine.

Figure S3

Fig. S3

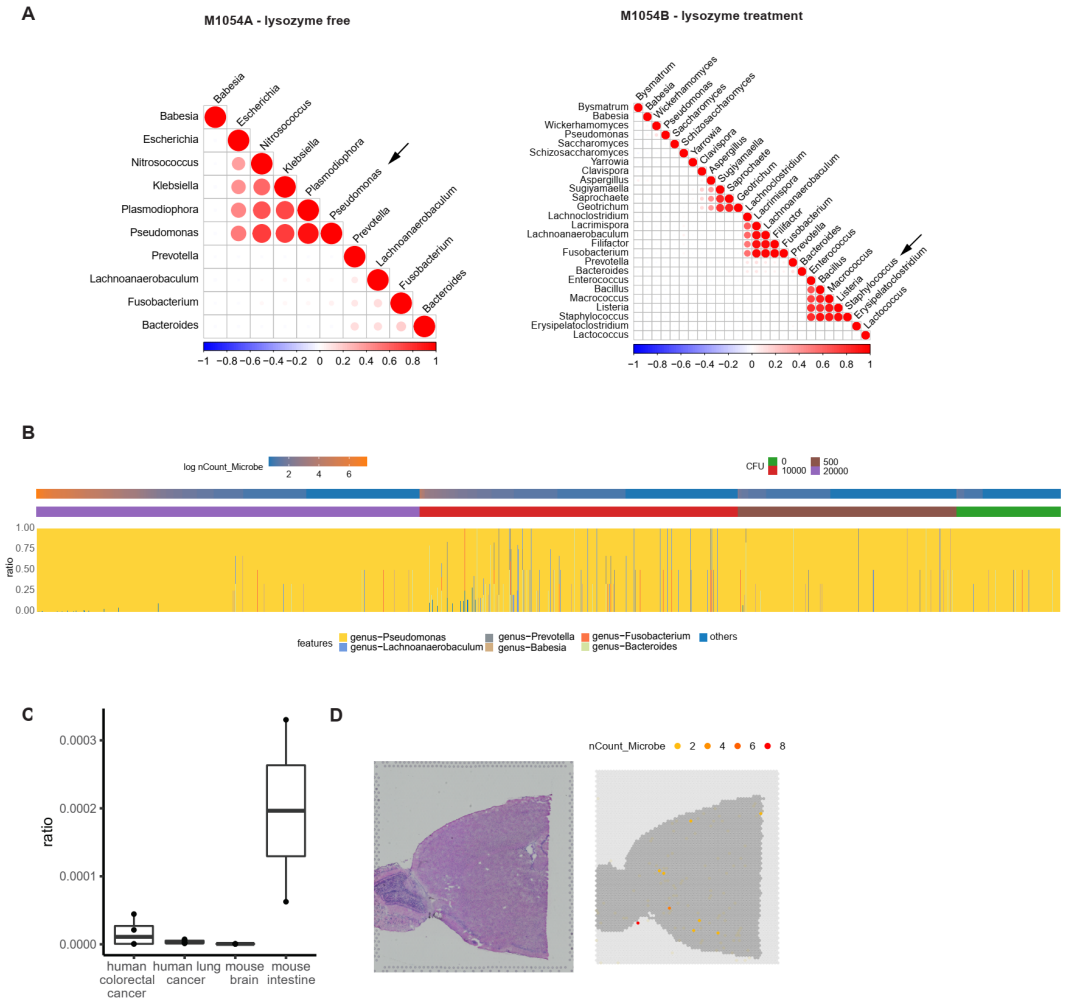

Fig. S3. Related to Fig 3. (A) Spatial correlation among bacteria genera as called by SMT, some spillover of the highly abundant control bacteria, *Pseudomonas* and *Staphylococcus*, did occur, manifesting as sequences assigned to closely related taxa. **(B)** Representative stacked bar plot showing microbial composition for individual spot in M1054A, their overall microbial sequence count and amount of control applied were also shown. **(C)** Box plot showing ratio of microbial reads in samples from human colorectal cancer and mouse intestine in this study, as well as a dataset of mouse brain obtained from in 10× Genomics' official website, and a human lung cancer dataset downloaded from SRA. **(D)** Representative spatial feature plot showing distribution of microbe sequence from mouse brain.

Figure S4

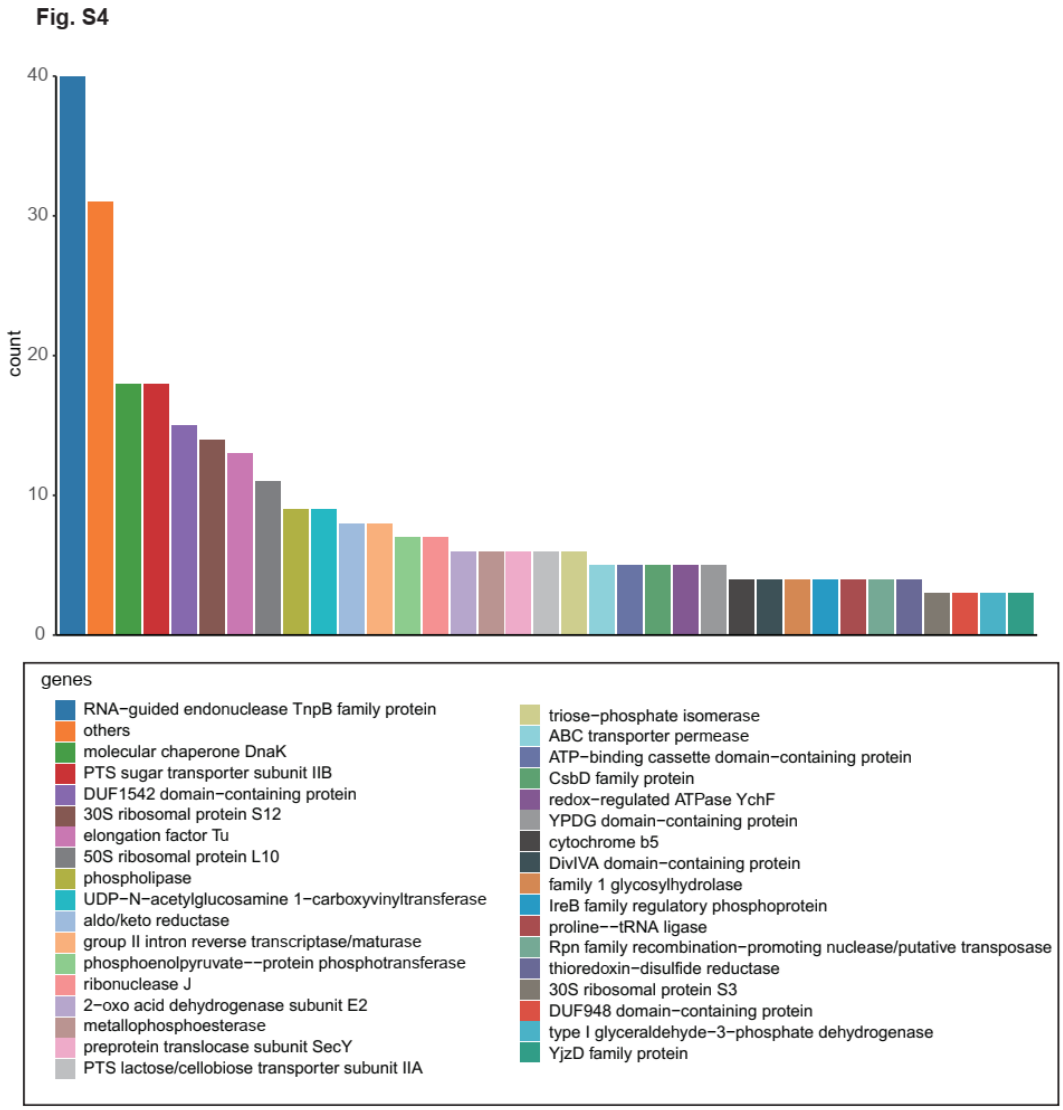

Fig. S4. The genic annotation for a bacteria recovered by SMT. For *Lactobacillus*, the most abundant genus in mouse small intestine, a total of 304 mRNA UMI were recovered and their annotations shown.

Supplemental Table

Table S1

SMT & kraken2 comparison

	Kraken2	SMT
Microbial reads called[a]	63493	38977
Microbial UMI called	8593	1453
Read/UMI[b]	7.3	26.8

a. Kraken2 and SMT used same dataset, and the total number of input read is 3233863.

b. Used as a measure of quality, overall read/UMI is 717337698/27998444 = 25.6.