Figure S1

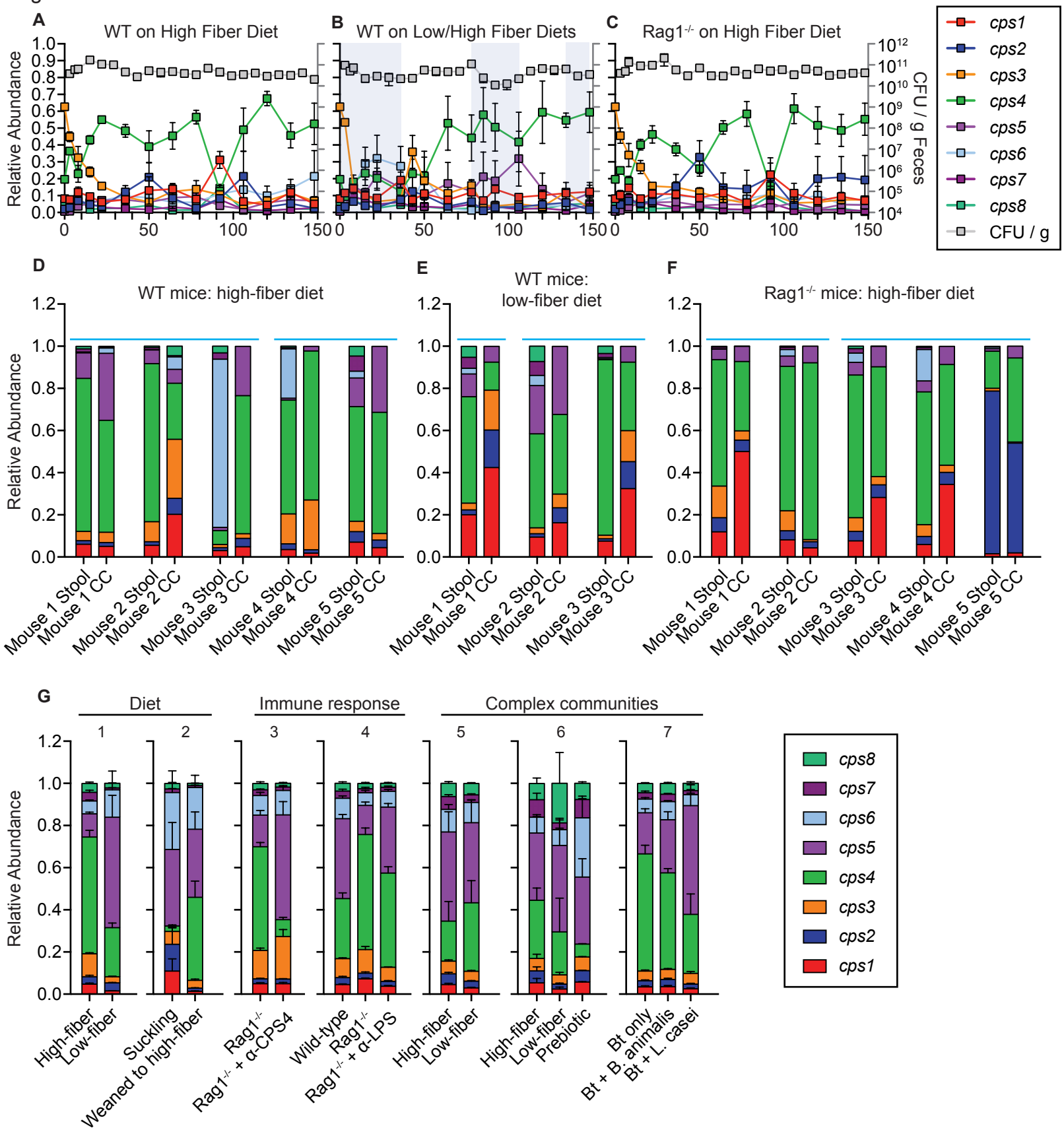


Figure S1. Expression of *cps* loci by wild-type *B. theta* in vivo. Related to Figure 1. A-C) Pooled relative *cps* locus expression data for mice shown in Figure 1 (n = 5 mice/group). D-F) Relative abundance of *cps* locus expression in stool and cecal contents, both on day of sacrifice, of mice shown in Figure 1. The stool sample for each mouse is the same as the last time point (Day 148) shown in Figure 1. Blue lines indicate co-housed mice. A, D) Wild-type mice fed a high-fiber diet. B, E) Wild-type mice oscillated between long-term feeding of a low-fiber and high-fiber diet. Mice had been fed the low-fiber diet for 2 weeks prior to sacrifice. Two of these mice died early, likely due to infighting between cage mates; thus no cecal samples are available for these two mice. C, F) Rag1$^{-/-}$ mice fed a high-fiber diet. G) Expression of *cps* loci by wild-type *B. theta* in various conditions, extracted from previously published datasets (n = 3-9 mice/group; see STAR Methods). Each panel represents a different study. Data are represented as mean ± SEM. WT: Wild-type (mice); CC: Cecal contents; Bt: *B. theta*; B. animalis: *Bifidobacterium animalis*; L. casei: *Lactobacillus casei*.

Figure S2

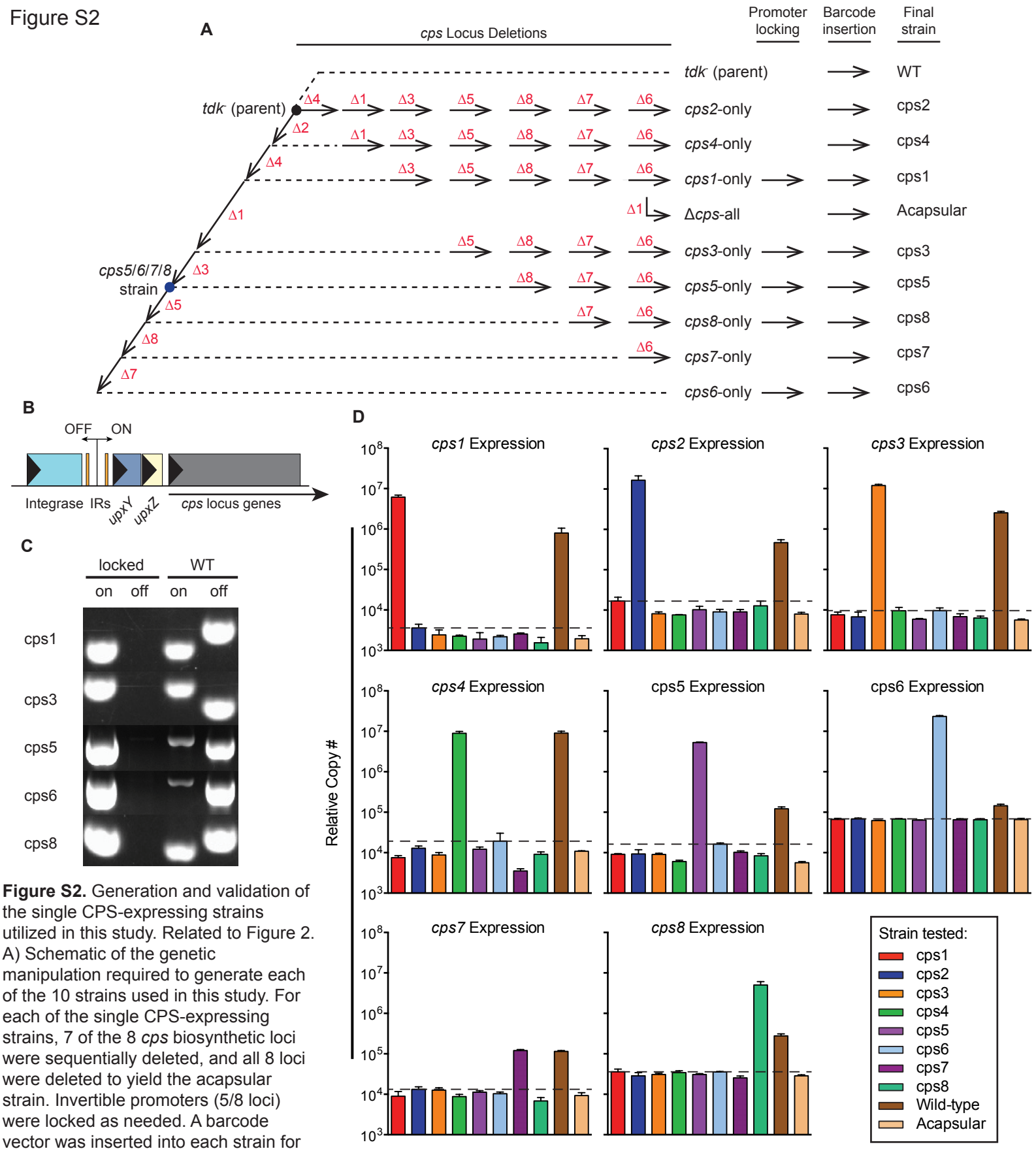
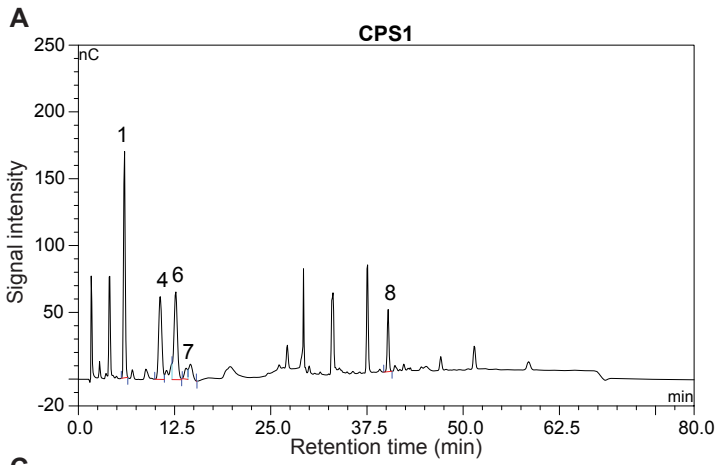
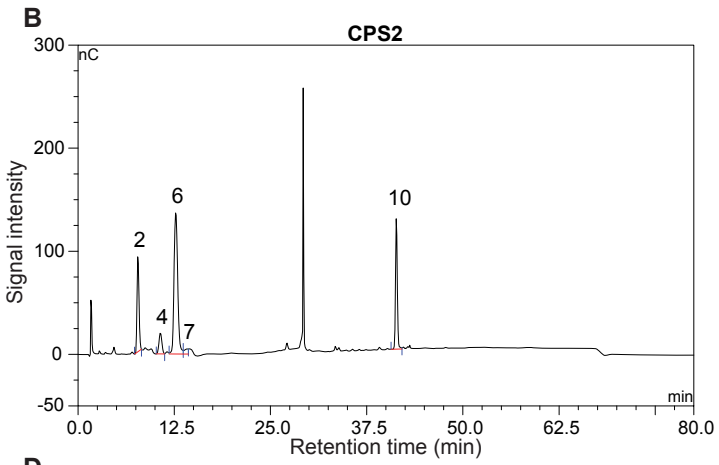
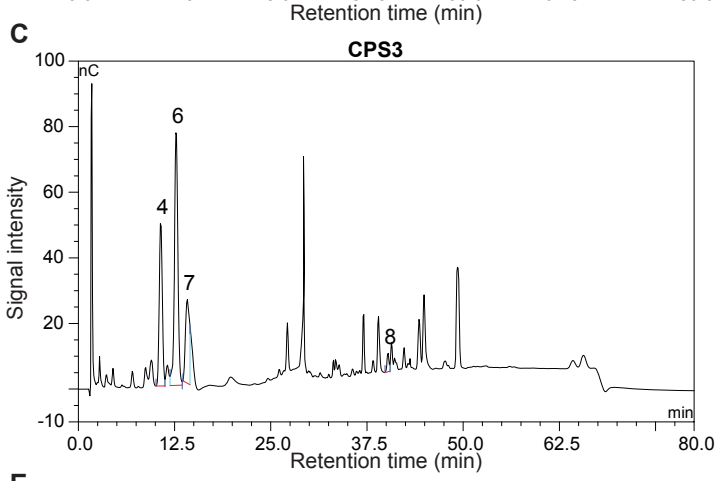


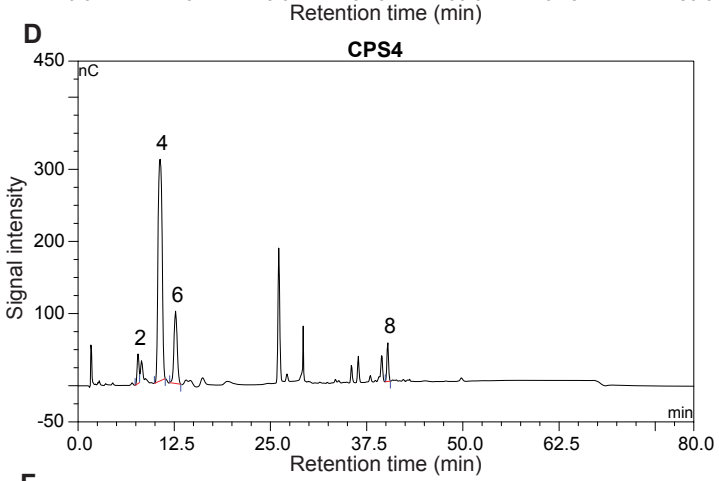
Figure S2. Generation and validation of the single CPS-expressing strains utilized in this study. Related to Figure 2. A) Schematic of the genetic manipulation required to generate each of the 10 strains used in this study. For each of the single CPS-expressing strains, 7 of the 8 *cps* biosynthetic loci were sequentially deleted, and all 8 loci were deleted to yield the acapsular strain. Invertible promoters (5/8 loci) were locked as needed. A barcode vector was inserted into each strain for identification in a mixed population. The CPS5-expressing strain was rederived from the *cps5/6/7/8*-encoding strain (blue dot) and this rederived strain was tested in Figure S4E only. B) General schematic of the promoter region of phase-variable *cps* loci (*cps1*, *cps3*, *cps5*, *cps6*, and *cps8*). Inverted terminal repeats (IRs) flanking the promoter region facilitate promoter inversion between "on" and "off" states. C) Phase-variable promoters are locked into the "on" orientation. Genomic DNA from the 5 single CPS-expressing strains containing invertible promoters was used in a PCR to determine promoter orientation after locking the promoter as described above, with the parent strain (*tdk*-, WT) employed as a promoter-inverting control. The forward primer for each strain is found within the phase variable region and a reverse primer is found downstream and within the locus ("on"), or upstream of the locus ("off"). *cps3* locus data is from Hickey et al., 2015. D) Each of the barcoded single CPS-expressing strains expresses a single *cps* locus. Each strain was grown in triplicate to mid-log, and qRT-PCR for a single gene in each locus was performed. Data are represented as mean ± SEM.

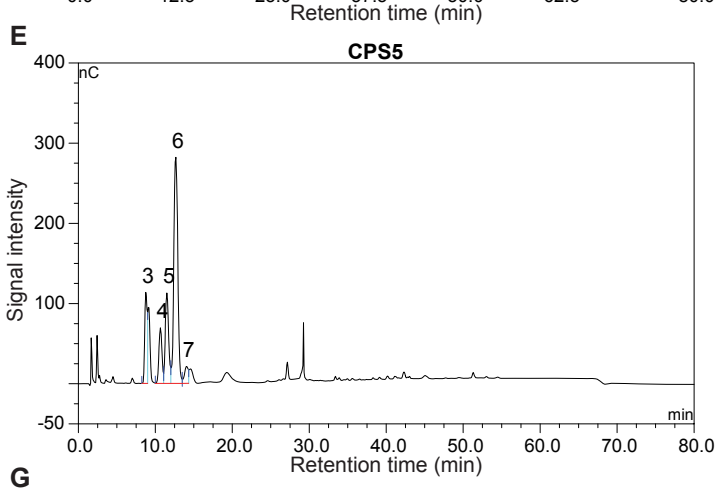
Figure S3

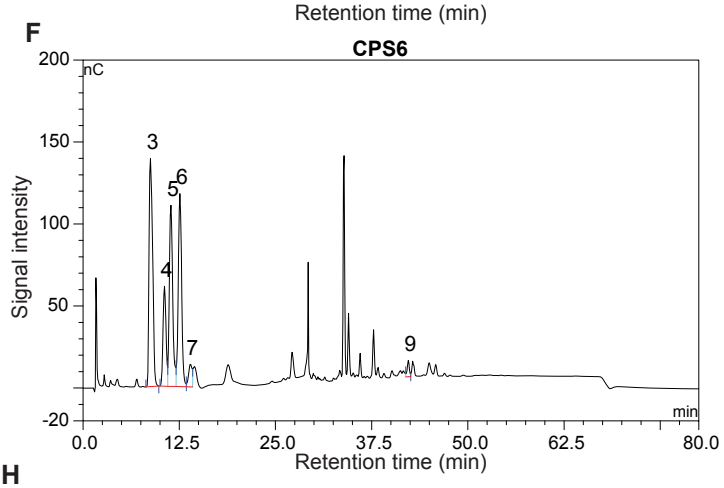


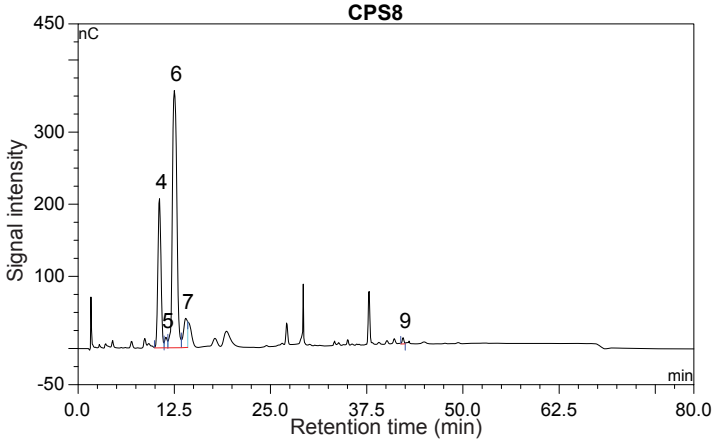












H

	3	4	5	6	7	8	9	Gly
CPS1		22		33	9	36		5
CPS2		8		85	7			0
CPS3		18		42	34	6		0
CPS4		61		21		18		2
CPS5	10	7	27	49	8			24
CPS6	23	8	37	23	6		3	7
CPS8		19	2	63	14		1	21

Figure S3. Monosaccharide composition of purified CPS. Related to STAR Methods. HPAEC-PAD traces for purified preparations of A) CPS1, B) CPS2, C) CPS3, D) CPS4, E) CPS5, F) CPS6, and G) CPS8. Due to low yield, CPS7 was not analyzed. H) Summary of relative percentage of quantified monosaccharides, and percentage of glycogen (”Gly”) contamination in each sample. Numbers on peaks (panels A-G) and numbers on column headings (panel H) correspond with the following detected monosaccharides: 1) N-acetyl-fucosamine, 2) Rhamnose, 3) N-acetyl-galactosamine, 4) N-acetyl-glucosamine, 5) Galactose, 6) Glucose, 7) Mannose, 8) Galacturonic acid, 9) Glucuronic acid, 10) unknown.

Figure S4

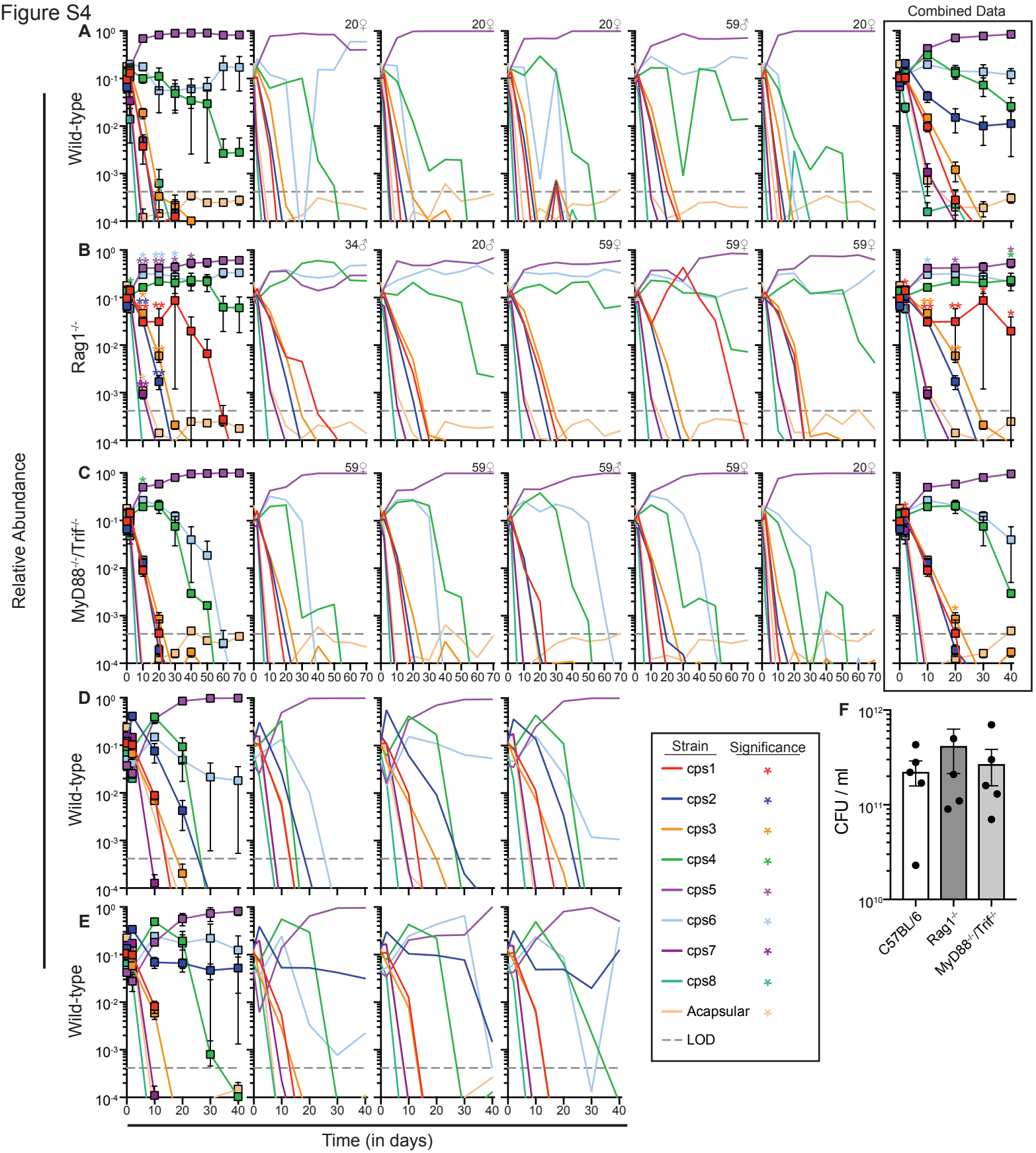


Figure S4. Competition of single CPS-expressing strains in vivo, data from individual mice. Related to Figure 2. A-C) 9 strains (8 single CPS-expressing strains and an acapsular strain) were pooled and inoculated into germ-free mice of 3 genotypes: A) C57BL/6 (wild-type, WT); B) Rag1$^{-/-}$; C) MyD88$^{-/-}$/Trif$^{-/-}$ (n = 5 mice/group). All mice were fed a high-fiber diet. Relative abundance of each strain in stool was determined at regular intervals. Each mouse in panels A-C was individually housed. Age at inoculation (in days) and gender are indicated at the top right of each panel. D-E) The experiment shown in Figure 2 was repeated in wild-type mice (n = 3 mice/group) with the same community of strains as before, including either: D) the original cps5 strain (mice co-housed); or E) a rederived cps5 strain (mice co-housed). The first panel in A-E represents pooled data, with subsequent panels representing individual mice. The last panel in A represents pooled data from all WT mice in Figures 2A, S4D-E, and S6A (n = 14 mice). Asterisks in panels B and C indicate significant differences (* $p < .05$, ** $p < .01$) in relative strain abundance between WT (panel A) and the indicated group, using a Kruskall-Wallis test with Benjamini-Hochberg correction. F) Colony forming unit (CFU) counts in cecal contents for mice in panels A-C. Data are represented as mean ± SEM as applicable. LOD: Limit of detection.

Figure S5

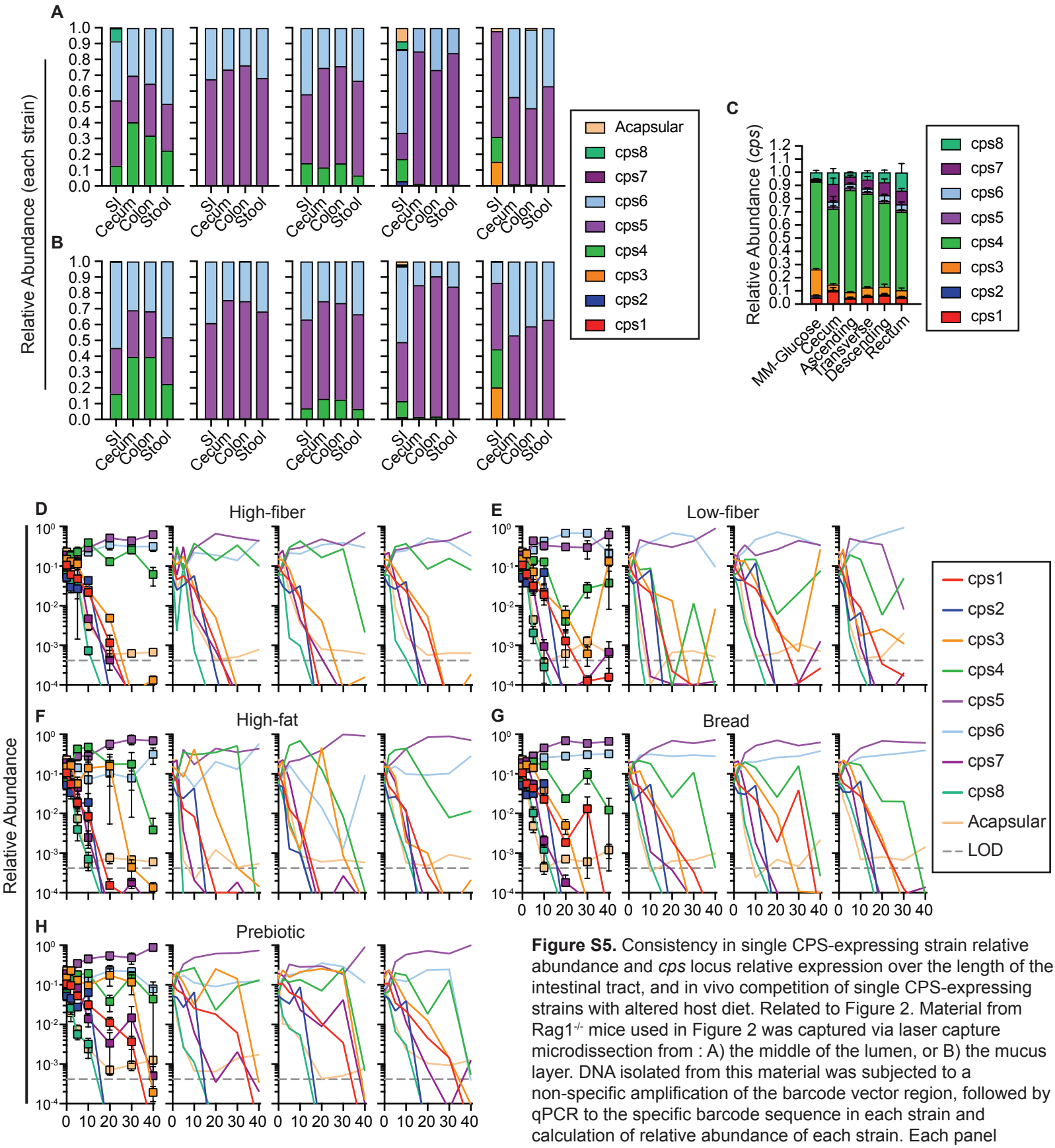


Figure S5. Consistency in single CPS-expressing strain relative abundance and *cps* locus relative expression over the length of the intestinal tract, and in vivo competition of single CPS-expressing strains with altered host diet. Related to Figure 2. Material from Rag1$^{-/-}$ mice used in Figure 2 was captured via laser capture microdissection from : A) the middle of the lumen, or B) the mucus layer. DNA isolated from this material was subjected to a non-specific amplification of the barcode vector region, followed by qPCR to the specific barcode sequence in each strain and calculation of relative abundance of each strain. Each panel represents an individual mouse, and upper and lower panels correspond to the same mouse. C) Mice (n = 3 mice) were monoassociated with *B. theta* for 10 days and RNA was extracted from whole intestinal segments as used in a previous study (Hickey et al., 2015). Relative *cps* locus expression was determined from each intestinal segment (small intestine and cecum, as well as regions of the distal colon). D-H) 9 strains (8 single CPS-expressing strains and an acapsular strain) were pooled and inoculated into wild-type germ-free mice fed one of 5 diets: D) high-fiber diet, E) low-fiber diet, F) high-fat diet, G) bread diet, or H) prebiotic diet. At regular intervals, DNA was extracted from stool for determining relative abundance of each strain via qPCR. The first panel in D-H represents pooled data, with each subsequent panel representing individual, co-housed mice (n = 3 mice/group). Pooled data are represented as mean ± SEM. SI, small intestine; MM-Glucose: minimal medium containing glucose; LOD: limit of detection.

Figure S6

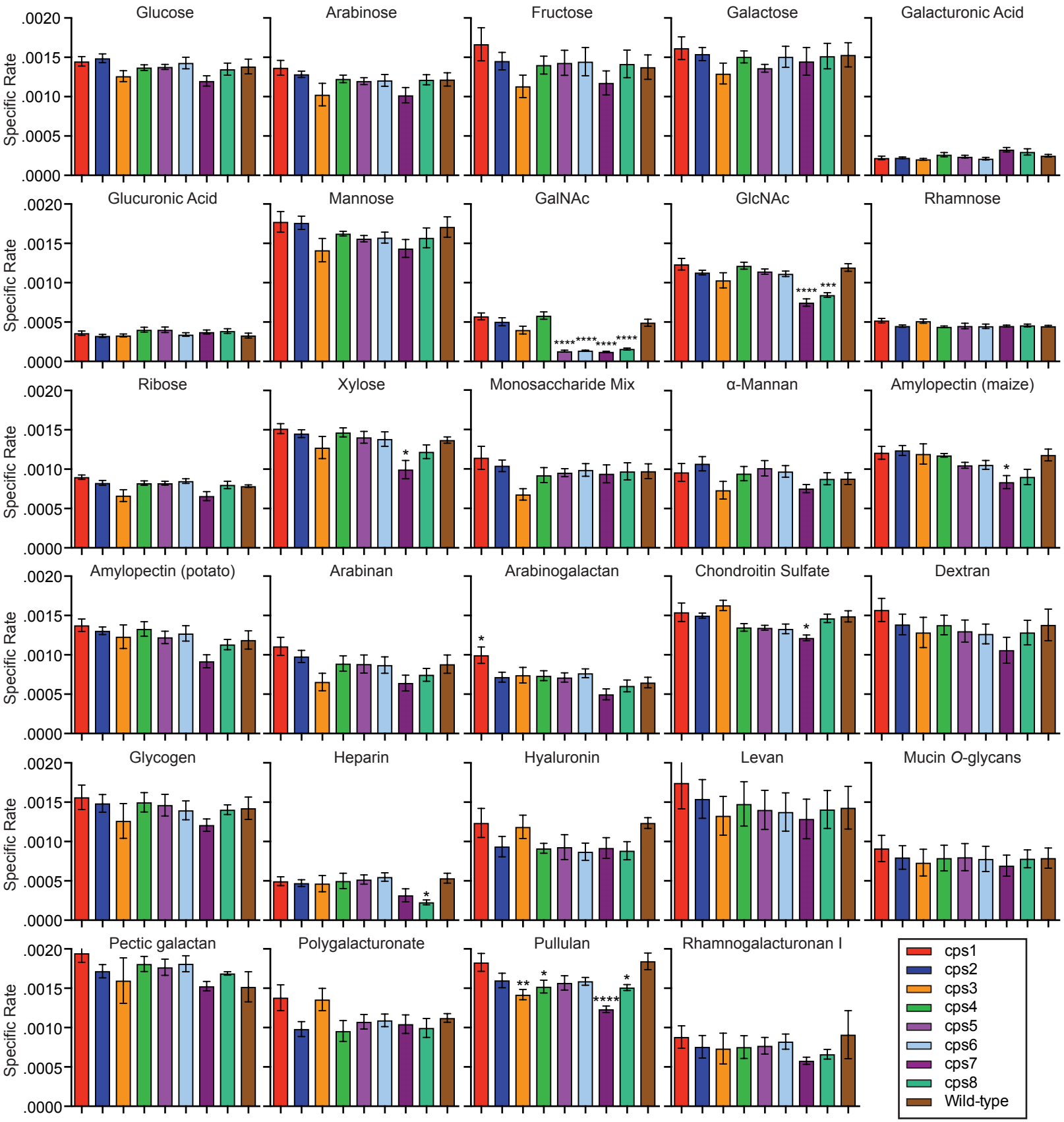


Figure S6. Specific growth rate of wild-type and single CPS-expressing strains in medium containing various monosaccharides or polysaccharides. Related to Figure 2. Strains were inoculated into fresh minimal medium containing 0.5% of the indicated carbohydrate (except 1.0% for O-glycans and rhamnogalacturonan I). The monosaccharide mix contains equal concentrations of each of the individual monosaccharides at a final concentration of 0.5%. Data are represented as mean ± SEM (n = 8 independent replicates for glucose, and 3-5 independent replicates for other sugars). Significant differences in growth rate were calculated via one-way ANOVA followed by Dunnett's multiple comparisons tests to the wild-type strain with Holm-Sidak correction. * ($p < 0.05$); ** ($p < 0.01$); **** ($p < 0.0001$).

Figure S7

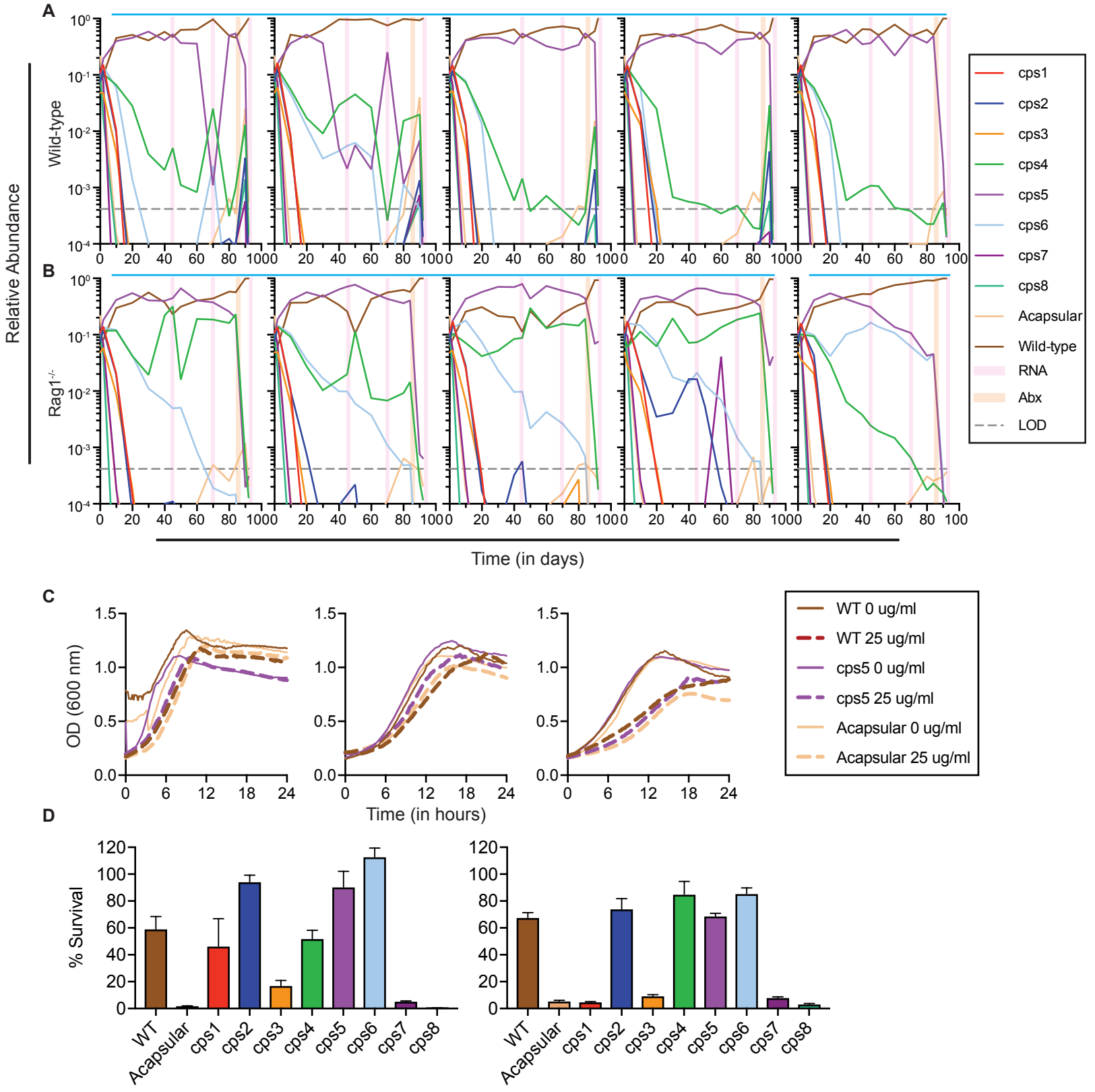


Figure S7. Individual mice from in vivo competition with wild-type (WT) B. theta and susceptibility to innate immune factors. Related to Figure 4. 10 strains (WT, the 8 single CPS-expressing strains, and the acapsular strain) were pooled and inoculated into germ-free mice fed a high-fiber diet. Mice were either A) wild-type, or B) Rag1$^{-/-}$. DNA was extracted from stool for determining relative abundance of each strain via qPCR. Ciprofloxacin was administered via drinking water for the indicated time, shown by the pink bars ("Abx"). Orange bars ("RNA") indicate times at which relative *cps* locus expression was also determined. Each panel represents an individual mouse in each group. Blue lines indicate co-housed mice. C) WT, cps5, and acapsular strains were grown in TYG medium in the presence (25 µg/ml) or absence of human beta-defensin 3. Each panel represents a separate experiment. D) 3 replicates of each strain (WT, acapsular, and the 8 single CPS-expressing strains) were incubated with normal human serum or freshly inactivated serum for 1 hour before plating on solid media. The percentage of serum-killed cells [(serum-treated - inactivated serum-treated) / inactivated serum-treated] was calculated. Two independent experiments are shown. Data for all experiments are represented as mean ± SEM, as applicable. LOD: Limit of detection.