Inflammation and bacteriophages affect DNA inversion states and functionality of the gut microbiota

Authors:

Shaqed Carasso[1*], Rawan Zaatry[1*], Haitham Hajjo[1*], Dana Kadosh-Kariti[1*], Nadav Ben-Assa[1], Rawi Naddaf[1], Noa Mandelbaum[1], Sigal Pressman[2,3], Yehuda Chowers[2,3,4], Tal Gefen[1], Kate L. Jeffrey[5,6], Juan Jofre[7], Michael J. Coyne[8], Laurie E. Comstock[8], Itai Sharon[9,10], Naama Geva-Zatorsky[1,11]

Affiliations:

[1] Department of Cell Biology and Cancer Science, Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Rappaport Technion Integrated Cancer Center (RTICC), Haifa, Israel

[2] Department of Gastroenterology, Rambam Health Care Campus, Haifa, Israel

[3] Clinical Research Institute, Rambam Health Care Campus, Haifa, Israel

[4] Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel

[5] Moderna, Inc., Cambridge, MA 02139, USA

[6] Center for the Study of Inflammatory Bowel Disease, Division of Gastroenterology, Department of Medicine, Massachusetts General Hospital Research Institute, Boston, MA 02114, USA; Harvard Medical School, Boston, MA 02115, USA; Program in Immunology, Harvard Medical School, Boston, MA 02115, USA.

[7] Department of Genetics, Microbiology and Statistics, School of Biology, University of Barcelona, Avda. Diagonal 643 08028 Barcelona Spain

[8] Duchossois Family Institute and Department of Microbiology, University of Chicago, Chicago, IL, USA

[9] Migal-Galilee Research Institute, P.O. Box 831, Kiryat Shmona 11016, Israel

[10] Faculty of Sciences and Technology, Tel-Hai Academic College, Upper Galilee 1220800, Israel

[11] CIFAR, MaRS Centre, West Tower 661 University Ave., Suite 505, Toronto, ON M5G 1M1, Canada

[*] These authors contributed equally

Summary

Reversible genomic DNA-inversions control expression of numerous bacterial molecules in the human gut, but how this relates to disease remains uncertain. By analyzing metagenomic samples from six human Inflammatory Bowel Disease (IBD) cohorts combined with mice experimentation, we identified

multiple invertible regions in which a particular orientation was correlated with disease. These include the promoter of the anti-inflammatory polysaccharide-A (PSA) of *Bacteroides fragilis*, which is mostly oriented 'OFF' during inflammation but switches to the 'ON' orientation when inflammation is resolved. We further detected increased abundances of *B. fragilis* associated bacteriophages in patients with the PSA 'OFF' orientation. Isolation and analysis of a *B. fragilis* associated bacteriophage revealed that it induced the PSA 'OFF' switch, thereby altering the bacterial induced immune modulation. Altogether, we reveal large-scale dynamic and reversible bacterial phase variations driven both by bacteriophages and the host inflammatory state signifying bacterial functional plasticity and suggesting potential clinical interventions.

Keywords

Phase variation, gut microbiome, bacteriophages, inflammatory bowel diseases, Crohn's disease, ulcerative colitis, *Bacteroides*, immunomodulation

Introduction

Phase variation is the process by which bacteria undergo frequent and reversible alterations in specific loci of their genome[1–4]. In Bacteroidales, the dominant order of bacteria in the human gut, phase variation is highly prevalent and is largely mediated by inversions of DNA segments between inverted repeats. These inversions often involve promoter regions dictating transcription initiation of genes or operons functioning as 'ON'\'OFF' switches[5]. In addition, DNA inversions can occur so that new genes are brought from an inactive to a transcriptionally active site by re-orientation or recombination of genomic shufflons and thus altering the expressed protein[6–8]. Analysis of the orientations of bacterial invertible regions in various host disease states can provide new insights into bacterial adaptation and functional contributions to the disease pathogenesis or its resolution. Phase variation in the gut Bacteroidales often modulates production of components presented on the bacterial surface[9] dictating which surface molecules interact, for example, with neighboring microbes or with the host. As such, phase variation confers altered functional phenotypes and bacterial functional plasticity.

Bacteroides fragilis, a common resident of the human gut, modulates its surface by the phase variable expression of its capsular polysaccharides (PS, denoted PSA-PSH). The biosynthesis loci of seven of its eight polysaccharides have invertible promoters that are orientated either 'ON' or 'OFF' in respect to the downstream PS biosynthesis operon[4]. Studies have shown that the *B. fragilis* polysaccharide A (PSA) modulates the host immune system by inducing regulatory T cells (Tregs) and secretion of the anti-inflammatory cytokine interleukin (IL)-10[10,11]. Moreover, PSA was shown to confer protection against experimental colitis[10,12–14], and thus is regarded as an anti-inflammatory polysaccharide.

Ulcerative colitis (UC) and Crohn's disease (CD) are multifactorial inflammatory bowel diseases (IBD), characterized by a compromised mucosal barrier, inappropriate immune activation, and mislocalization of the gut microbiota[15–18]. Genetic, immunological, and environmental factors all contribute to the development and progression of IBD. The gut microbiota is a major environmental factor in IBD, and as such, IBD has emerged as one of the most studied microbiota-linked diseases. A mechanistic analysis of bacterial functions and potential functional alterations will allow us to better understand the role of the gut microbiota in IBD.

Here we present an analysis of invertible DNA orientations in the gut microbiota of IBD patients from multiple cohorts across the globe, focusing on Bacteroidales species. Our combined analysis of IBD patient metagenomes and experimental mouse models reveals alterations in multiple DNA invertible sites during gut inflammation, with the potential to modulate the host immune system. Importantly, we show that filtered fecal extracts of IBD patients alter the orientation of the PSA promoter in *B. fragilis* populations. We show that bacteriophages are correlated with the PSA 'OFF' state and we show that a specific lytic bacteriophage of *B. fragilis* increases the PSA promoter 'OFF' population with concurrent decline of host colonic Treg cells. These findings reveal a dynamic interplay between gut inflammation, bacteriophage, and bacterial phase variation, with potential implications for the diagnosis and treatment of IBD.

Results

Multiple Bacteroidales invertible DNA regions have altered orientations in association with IBD

As a means to determine if Bacteroidales phase variable molecules may be differentially produced in IBD patients compared to controls, we used the PhaseFinder[9] software to identify and quantify DNA inversions in the gut metagenomes of cohorts of IBD patients and control subjects **[Table 1]**. To do this, we first analyzed 39 sequenced genomes representing different human associated strains (36 species) from the Bacteroidales order and identified 311 invertible regions in these genomes. Analysis of these regions in human gut metagenomes identified a total of 147 statistically significant DNA inversions associated with the disease from 25 of the reference bacterial genomes. These DNA inversions were located in both invertible promoters and intragenic regions of diverse genes, spanning from regulatory genes to outer surface molecules. **Table S1** details the locations of the inverted repeats (IR) sequences and genes flanking the phase variable regions **[Figure 1A]**. The invertible regions whose orientations differed most significantly between IBD patients and controls were within or in proximity to SusC/SusD-like outer membrane transport systems[19] and capsular polysaccharides (PS) promoters **[Figure 1B]**. Four of the five invertible capsular polysaccharide promoters of

Bacteroides thetaiotaomicron were differentially oriented between the groups, and three of the seven invertible PS promoters of *B. fragilis* were differentially oriented, with the anti-inflammatory PSA promoter being the most distinct between healthy and UC patients. The PSA promoter showed a higher percentage (71%) of reverse oriented reads (compared to its reference genome sequence) in IBD patients, compared to 56.2% in the healthy controls **[Figure 1C, 1D]**. To note, In the reference genome of *B. fragilis* NCTC9343 the promoter of PSA is in its 'ON' orientation, hence, the reverse orientation, found in IBD patients, represents the PSA promoter's 'OFF' orientation. We further identified differential promoter orientations in healthy and IBD cohorts for PULs and polysaccharides promoters of *Phocaeicola dorei*, a bacteria shown to be present in healthy subjects[20,21] and also correlated with disease activity in UC[22] **[Figure 1E]**.

Table 1: Cohorts of IBD patients and control subjects included in DNA inversion analysis.

	N (Samples included in the analysis)	CD	UC	Non-IBD	Country	Notes	
IBDMDB	1283	574	349	360	United States	Longitudinal	ref[23]
HMP (3 phases)	320			320	United States		ref[24]
MetaHit	122	4	21	97	Denmark and Spain		ref[25]
1000IBD	331	205	126		the Netherlands		ref[26]
GeversD_2014	50	36		14	United States and Canada		ref[27]
LewisJD_2015	303	303			United States and Canada	Children with Crohn's Disease Longitudinal, treated	ref[28]
SUM	2409	1122	496	791			

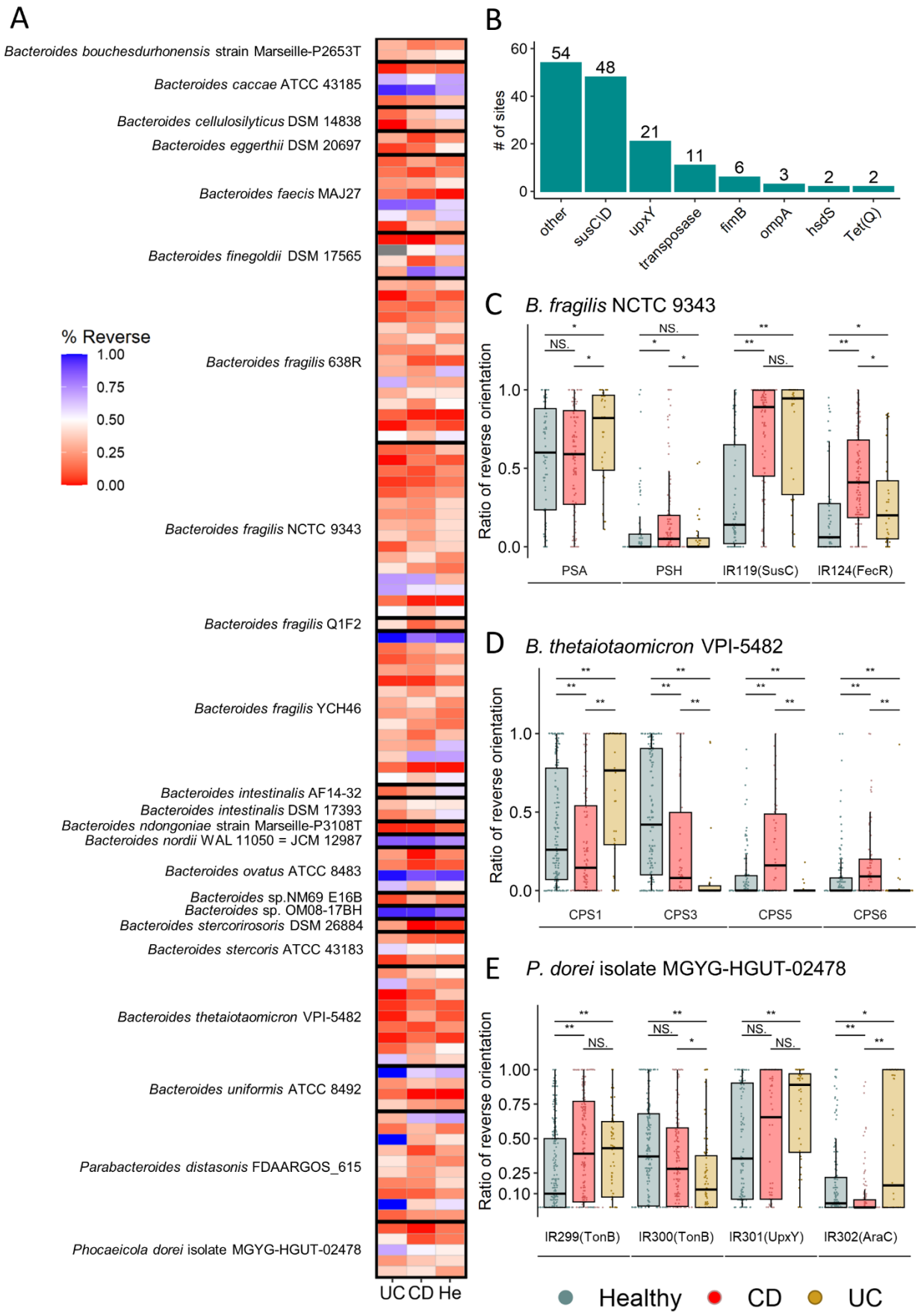


Figure 1: *Bacteroides* species exhibit phase variation in health and disease.
A. Selected significantly differentially oriented invertible DNA regions (inverted repeats (IR) segments, see **Table S1**) (Wilcoxon rank sum test, FDR p<0.05) in at least one comparison between Healthy, CD (Crohn's Disease) and UC (Ulcerative Colitis). Blue indicates the Forward orientation and red represents the Reverse orientation in comparison to the reference genome. **B.** Prevalence of functional genes in proximity to invertible DNA regions significantly different between healthy individuals and IBD patients. **C.** Differentially oriented invertible DNA regions, in *Bacteroides fragilis* NCTC 9343. PSA: Polysaccharide A, PSH: Polysaccharide H. **D.** Differentially oriented invertible DNA regions in *Bacteroides thetaiotaomicron* VPI-5482. CPS: Capsular Polysaccharide. **E.** Differentially oriented invertible DNA regions in *Phocaeicola dorei* MGYG-HGUT-02478. (Wilcoxon rank sum test, *$p < 0.05$; **$p < 0.01$)

Differential promoter orientation states are induced by inflammation in a dynamic and reversible manner

To assess the dynamics of the *B. fragilis* PSA promoter orientation under inflammatory conditions, we designed a longitudinal experimental mouse model during which we analyzed the PSA promoter orientation over time. Germ-free (GF) mice were colonized with a healthy human microbiota, which was spiked with *B. fragilis* NCTC 9343, termed "humanized" mice. Experimental colitis was induced by adding 3% Dextran sodium sulfate (DSS) to the drinking water[29] **[Figure 2A]** for 9 days after which DSS was replaced with water until day 14, and as a result the inflammation was reduced. Inflammation was confirmed by the levels of Calprotectin in the stool **[Figure 2C]**, a commonly used biomarker for inflammation, and with mouse weight loss **[Figure 2D]**. Stool was collected before, during and after DSS treatment to analyze orientations of specific invertible DNA regions, using quantitative PCR (qPCR). The *B. fragilis* PSA promoter orientation varied correlating with the inflammatory state **[Figure 2B]**. At the beginning of the experiment, ~45% of the population had the promoter in the 'ON' orientation **[Figure 2B]**. Nine days after DSS was introduced, these percentages declined to 24%. By day 14, the bacterial population returned to 40% 'ON', similar to the beginning of the experiment, and to the control group, which did not receive DSS (mean of 58% of the *B. fragilis* population 'ON').

We further performed metagenomics analysis of these mice to study alterations in bacterial composition and to apply the PhaseFinder[9] algorithm for extensive phase variation analysis. This analysis revealed phase variation in multiple bacteria and genomic regions, including the PSA promoter of *B. fragilis* and the CPS3 promoter of *B. thetaiotaomicron*, overlapping with the results of the human databases analysis **[Figure 2E, 2F, Supp. Table 2]**.

Next, we identified alterations in bacterial composition of the DSS treated mice, whereas the control mice were stable throughout. Bacterial richness (alpha-diversity, observed species) declined on day 6 under inflammatory conditions [**Figure 2H]**, while evenness remained constant **[Figure 2I]**. Beta-diversity (Bray-Curtis distance) was altered upon DSS treatment and each mouse had a different bacterial composition five days after removal of the DSS **[Figure 2J]**. Inflamed mice showed a decrease in bacterial abundances of *Blautia* species genera and *Eubacterium ventriosum*, while showing an increase in *Barnesiella intestinihominis* and *Alistipes dispar* **[Figure 2G, Supp. table 2 and 3]**. The

150 relative abundances of *B. fragilis* and *B. thetaiotaomicron*, which exhibited differential orientations of
151 invertible DNA regions, increased in inflamed mice, although to a lesser extent **[Figure 2G]**.

152 To assess the role of inflammation on the orientation of the *B. fragilis* PSA promoter in the absence of
153 a complex microbiota, we repeated the experiment using gnotobiotic mice mono-colonized with *B.*
154 *fragilis* NCTC 9343. The PSA promoter orientation remained stable in the monocolonized mice over
155 the course of the experiment, suggesting a role of the microbiota in the relative 'OFF' orientation of
156 this promoter during inflammation. **[Figure 2K]**.

157

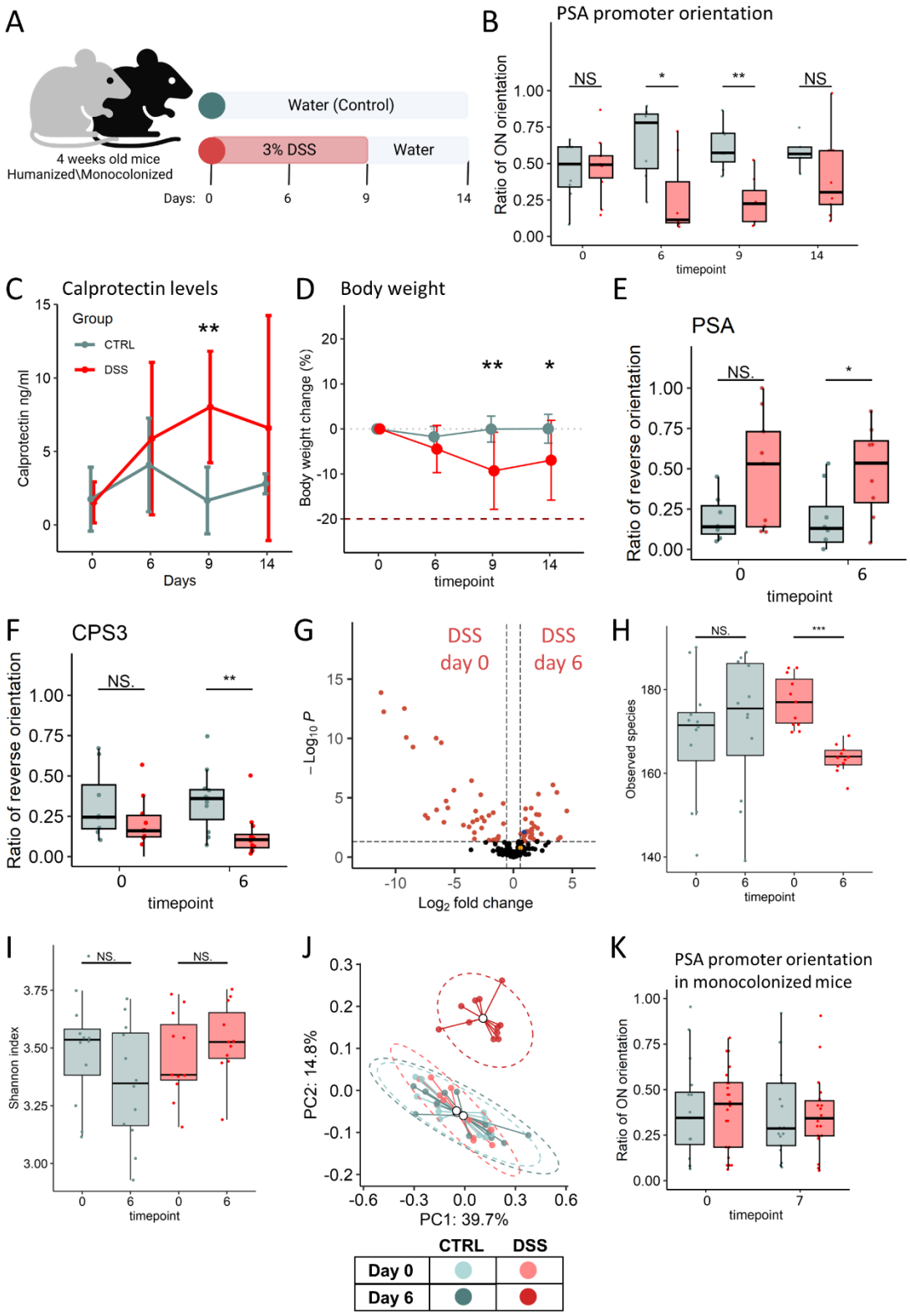


Figure 2: Relative orientation of the PSA promoter of *B. fragilis* from the population is affected by inflamed state of the gut.
A. Illustration of murine model of inflammation. Mice monocolonized with *B. fragilis* NCTC 9343 and mice colonized with a human microbiota spiked with *B. fragilis* NCTC 9343 - "Humanized", were exposed to 3% Dextran sodium sulfate (DSS) (Day=0 in the illustration) at four weeks of age; B. Ratio of *B. fragilis* PSA's promoter ON orientation measured by qPCR in different days of the experiment. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test, *$p < 0.05$; **$p < 0.01$). Blue: Control group, Red: DSS treated mice. **C.** Calprotectin level (ng/ml) measured in different days of the experiment. Lines represent the standard deviations. Blue: Control group, Red: DSS treated mice. **D.** The body weight change of mice measured in different days of the experiment. Lines represent the standard deviations. Blue: Control group, Red: DSS treated mice. **E.** Ratio of *B. fragilis* PSA's promoter reverse orientation measured by the PhaseFinder tool, in different days of the experiment. Data represent the median (line in box), IQR (box), and minimum/maximum

(whiskers). (Wilcoxon rank sum test, *p <0 .05; **p <0 .01). Blue: Control group, Red: DSS treated mice. **F.** Ratio of *B. thetaiotaomicron* CPS3's promoter reverse orientation measured by the PhaseFinder tool in different days of the experiment. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test, *p < 0.05; **p < 0.01). Blue: Control group, Red: DSS treated mice. **G.** Differential bacterial abundance between DSS treated mice in timepoint 0 and timepoint 6 detected by the Maaslin2 algorithm. Red dots indicate differentially abundant bacteria that were determined by adjusted P value < 0.05 and $\log_2$ fold change >1 and <-1, respectively. Yellow dot indicates *B. fragilis*, and blue dot indicates *Bacteroides thetaiotaomicron*. **H.** Alpha diversity (observed species) between groups and timepoints. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (t test, *p < 0.05; **p < 0.01, ***p<0.001). Blue: Control group, Red: DSS treated mice. **I.** Alpha diversity (Shannon index) between groups and timepoints. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (t test, *p <0 .05; **p < 0.01, ***p<0.001). Blue: Control group, Red: DSS treated mice. **J.** PCoA on Bray–Curtis dissimilarity distances between groups and timepoints. Each point represents a single sample, colored according to group and timepoints: light blue: Control at timepoint 0, blue: control at timepoint 6, light red: DSS treated at timepoint 0, red: DSS treated at timepoint 6. The mean (centroid) of samples in each group is indicated with a blank circle. Ellipses represent 0.95 confidence interval of each group. **K.** Ratio of *B. fragilis* PSA's promoter ON orientation measured by qPCR in different days in monocolonized mice. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test, *p <0 .05; **p < 0.01). Blue: Control group, Red: DSS treated mice.

The inflammatory milieu triggers DNA inversions in *B. fragilis*

We next sought to examine environmental factors of inflammation that result in a decreased percentage of the *B. fragilis* population with the PSA promoter orientated ON. To do so, we exposed *B. fragilis* NCTC 9343 to fecal filtrates from IBD patients. CD and UC patients were recruited from the Rambam Health Care Campus (RHCC). Fecal samples were collected before and after Infliximab (HR) or Humira (HuR) therapy, both are antibodies targeted against tumor necrosis factor-α (TNF-α), an inflammatory cytokine increased in IBD patients. *B. fragilis* NCTC 9343 was cultured in fecal filtrates until reaching mid-log phase, and subsequently DNA was extracted for qPCR analysis **[Figure 3A]**. *B. fragilis* exposed to fecal filtrates of patients before treatment, during inflammation, showed higher ratios of the population with the PSA promoter oriented 'OFF', while *B. fragilis* exposed to fecal filtrates after treatment showed higher ratios of the PSA promoter 'ON' orientation **[Figure 3B]**. This observation was in line with the fecal Calprotectin concentrations, which were reversely correlated with PSA promoter 'ON' orientation **[Figure 3C]**. These results demonstrated that the population of bacteria with the PSA promoter in each orientation vary in the inflamed and non-inflamed gut, with a higher percentage of the population in the 'OFF' orientation under inflammatory conditions, and a shift towards the 'ON' orientation following reduction in inflammation **[Figure 3B, 3C]**. This change could be due to alterations of the inversion under these conditions or due to selection of a preferred promoter state.

Investigating viral associations to orientation of the polysaccharide A promoter of *B. fragilis*

The fecal filtrates contain a mixture of bacterial and host metabolites, cytokines, antibodies, viruses, bacteriophages, and other components. Bacteriophages have been observed to induce phase variation in bacterial cells, leading to the switching 'OFF' of specific polysaccharides loci or increased

expression of alternative polysaccharides with possible protective effects[30]. Moreover, they have been shown to be associated with intestinal inflammation and IBD in several studies[31,32] , including Nishiyama et al.[33] who characterized temperate bacteriophages and their bacterial hosts in the IBDMDB metagenomics database. Here, using phage sequences from IBDMDB database, and compared the relative abundances of these phages between samples that displayed lower (<40%) or higher (>60%) ratios of the PSA promoter 'ON' orientation. Differential abundance analysis revealed that more bacteriophages were correlated with the PSA 'OFF' orientation **(Figure 3D, Supp. Table 3)**, three of which were predicted to infect *B. fragilis*[33]. To assess whether the relative abundances of the bacteriophages were accompanied with lower relative abundances of *B. fragilis*, we compared phage to host for each *B. fragilis* associated bacteriophage. The two viral OTUs with the highest phage to host ratios in patients, OTUs 0791 and 0820, were also correlated with low ratios of the PSA promoter 'ON' orientation (i.e., with the PSA promoter 'OFF' orientation) **[Figure 3E, 3F]**. Intriguingly, viral OTUs 0791 and 0820 were found to be more abundant in active CD and UC[33].

We conducted the same analysis comparing samples with high to low 'ON' orientation ratios of the CPS3 promoter of *B. thetaiotaomicron*. Although there were differences in the virome compositions between these groups, no *B. thetaiotaomicron* associated bacteriophages were correlated with the 'OFF' orientation of the CPS3 promoter (i.e., the orientation, which we identified as associated with the disease) **[Figure 3G, Supp. Table 3]**. We found three *B. thetaiotaomicron* associated viral OTUs - 0165, 1130, and 1490, slightly associated with the 'ON' orientation of the CPS3 promoter (the orientation that we identified as associated with healthy controls) **[Figure 3G]**. In addition, these bacteriophages displayed low phage to host ratios **[Figure 3E]**. To note, these OTUs were not previously associated with disease[33].

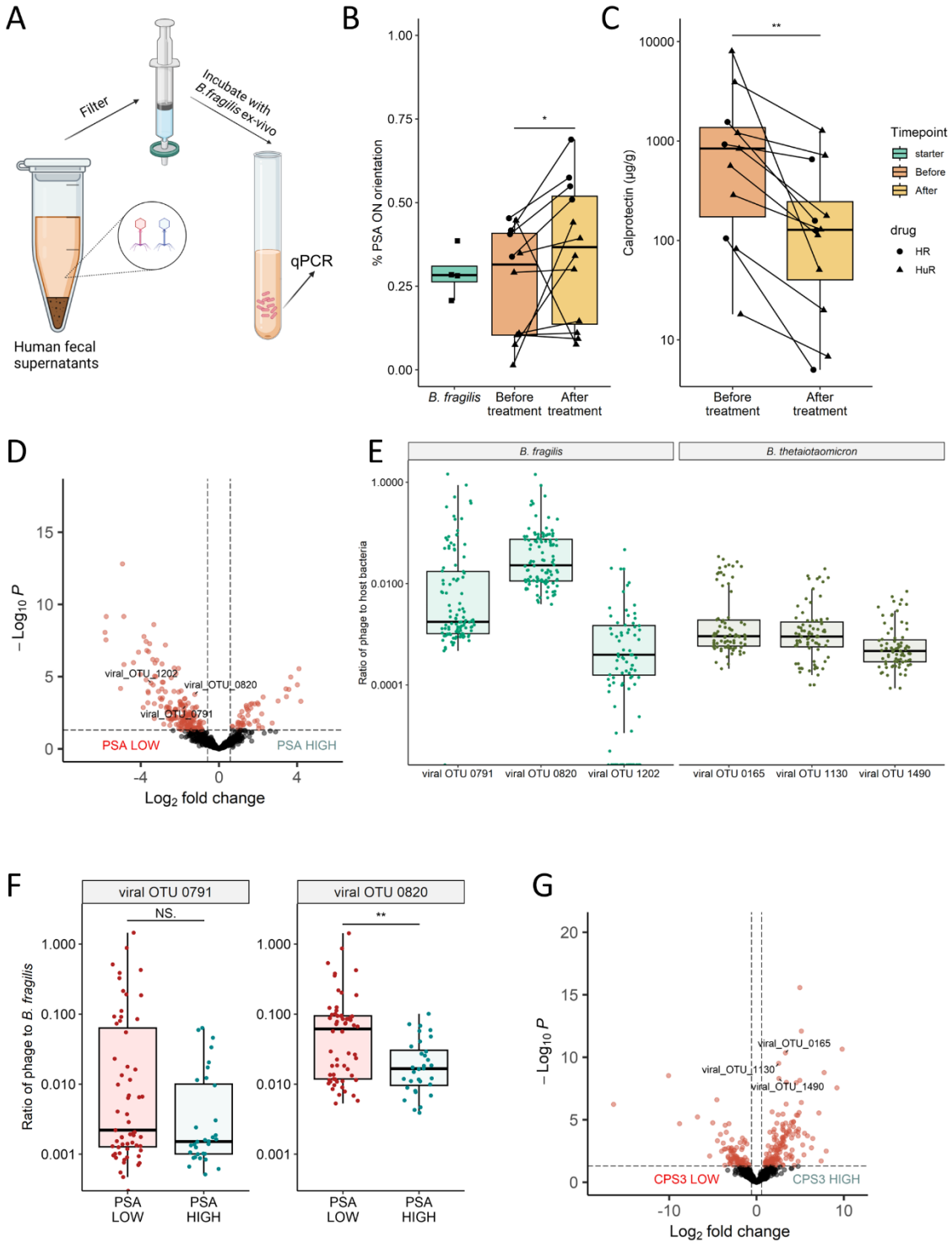


Figure 3: Evidence of viral association to *B. fragilis* Polysaccharide A's promoter genomic orientation.
A. Experimental design of culturing *B. fragilis* in patient's fecal filtrates. **B.** Ratio of the ON orientation of the PSA promoter of *B. fragilis,* measured by qPCR, after ex-vivo exposure to fecal filtrates of CD patients before and after treatment with anti-TNF (either Infliximab (HR) or Humira (HuR)). Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (One-sided Wilcoxon rank sum test, *p <0 .05). green: *B. fragilis* not exposed to fecal filtrates, orange: *B. fragilis* exposed to fecal filtrates of patients before anti-TNF treatment. yellow: *B. fragilis* exposed to fecal filtrates of patients after treatment. Dots represent individual experiments; lines connect experiments from the same patient; Shapes are determined by the patients' treatments, circle: HR, triangle: HuR. **C.** Calprotectin levels (µg/g) measured in patients' feces. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (One-sided Wilcoxon rank sum test, *p <0 .05). Dots represent samples; lines connect samples from the same patient; Shapes are determined by the patients' treatments, circle: HR, triangle: HuR. **D.** Differential viral taxonomic units' abundance, (from the IBDMDB cohort, count table from Nishiyama et al. (2020).), between samples with low ON orientation of the PSA promoter (<40%) and high ON orientation (>60%), in the IBDMDB cohort count table. Differentially abundant viral taxonomic units were detected by the DeSeq2 algorithm (Wald test, p < 0.01). Red dots indicate differentially abundant bacteria that were determined by P value < 0.01 and fold change >1.5 and <-1.5, respectively. **E.** Phage to host ratios of viral OTUs detected in Figure 3D and 3F. **F.** Phage to host ratios of viral OTUs detected in Figure 3D against samples' levels

of PSA's promoter ON orientation analyzed from the IBDMDB cohort, based on the count table from Nishiyama et al. (2020). **G.** Differential viral taxonomic units' abundance, (from the IBDMDB cohort, count table from Nishiyama et al. (2020).), between samples with low ON orientation of the CPS3 promoter (<40%) and high ON orientation (>60%), in the IBDMDB cohort. Differentially abundant viral taxonomic units were detected by the DeSeq2 algorithm (Wald test, $p < 0.01$). Red dots indicate differentially abundant bacteria that were determined by P value < 0.01 and fold change >1.5 and <-1.5, respectively.

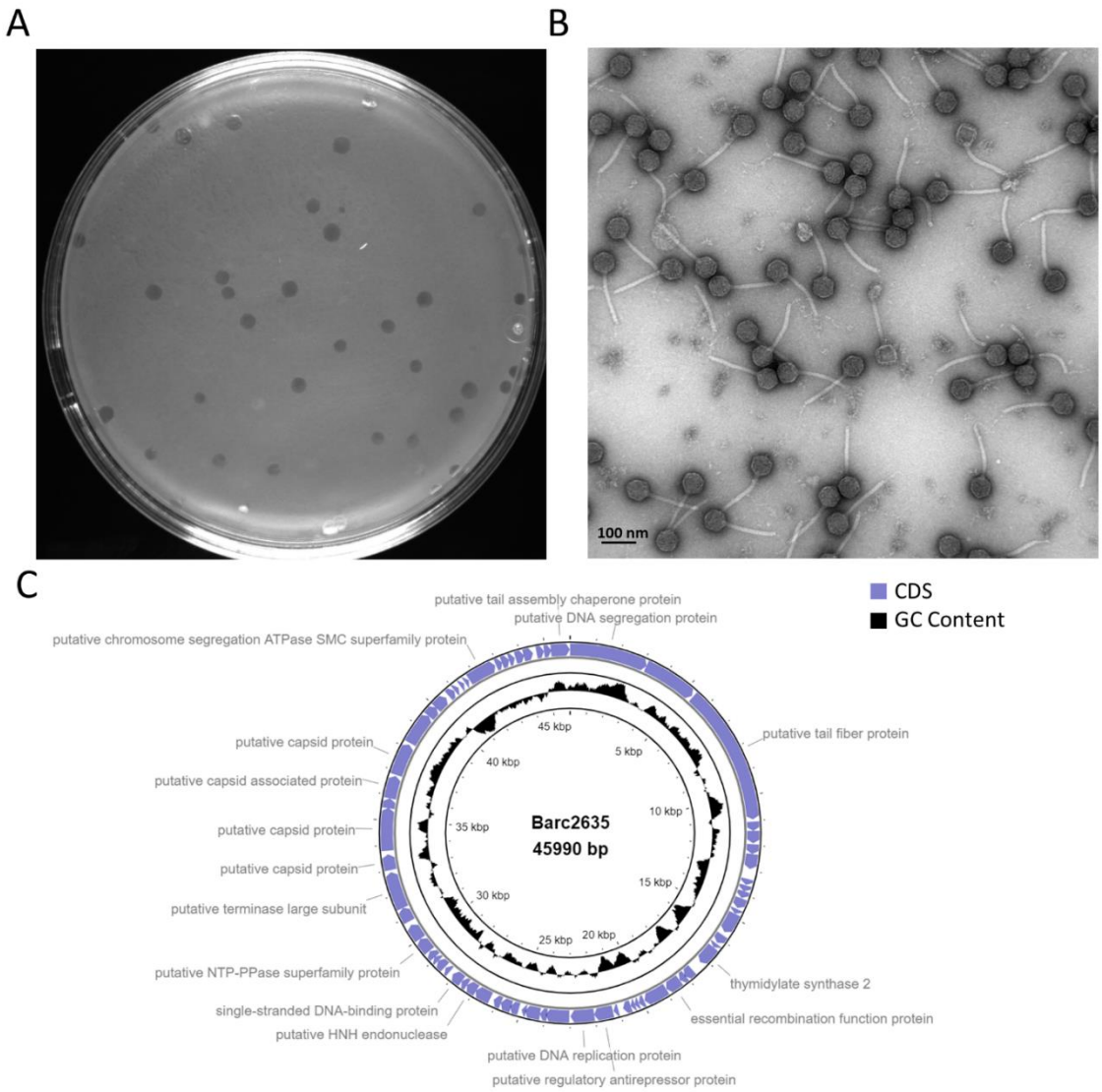
Bacteriophage exposure drives phase variation and reduction in Treg cells population

Since a higher abundance of bacteriophages were correlated with the PSA promoter in the 'OFF' orientation, we sought to study whether encounter with bacteriophage can alter the orientation of this promoter in *B. fragilis*. We started by isolating a bacteriophage from sewage that we designated bacteriophage Barc2635 that makes plaques on a lawn of *B. fragilis* NCTC934 **[Supp Figure 1A]**. We sequenced the phage genome (GenBank accession: MN078104) and characterized its morphology by electron microscopy **[Supp. Figure 1B].** Barc2635 is a double-stranded DNA lytic bacteriophage of 45990 bp with a GC content of 38.9%, containing 67 putative CDS belonging to the *Siphoviridae* family and the *Caudovirales* order **[Supp. Figure 1C]**. Interestingly, increased abundance of bacteriophages from the *Caudovirales* order were found to be correlated with IBD patients (UC and CD)[34]. We analyzed the sequence similarities of Barc2635, with the *B. fragilis* associated bacteriophages identified in the IBDMDB[33] and found that Barc2635 is most similar to a cluster of bacteriophages, including viral OTUs 0791 and 0820 which were associated with active disease[33] **[Figure 4A]**.

To test whether this bacteriophage causes changes in relative orientations of the PSA promoter in populations of *B. fragilis*, we monocolonized GF mice with *B. fragilis* NCTC 9343 in the presence or absence of bacteriophage Barc2635 and analyzed genomic-wide DNA inversion states **[Figure 4B]**. In the presence of Barc2635, *B. fragilis* exhibited prolonged alterations in the orientations of invertible regions affecting outer surface components such as polysaccharide utilization loci (PULs), ABC transporters and capsular polysaccharides, including PSA **[Supp. Table 4]**. In response to bacteriophage, a higher percentage of the population had the PSA promoter in the 'OFF' state **[Supp. Figure 2A]**, with a concomitant increase in the percentage of the population with the PSF promoter orientated 'ON' **[Supp. Figure 2B]**. Throughout the experiment (42 days), Barc2635 and *B. fragilis* co-existed in gnotobiotic mice, with mild fluctuations in their abundances **[Supp. Figure 2C].** Ten days after the initial colonization, we quantified expression of the first gene of the PSA locus, *upaY*[35]. *upaY* expression requires that the PSA promoter be oriented 'ON', and that no inhibitory UpxZ products from other loci are produced[36]. **Figure 4C** shows a significant reduction in *upaY* expression in mice containing Barc2635 in comparison to mice with bacteria alone. To directly measure the levels of PSA on the surface of the *B. fragilis* population, we used specific antibodies and monitored surface PSA by flow cytometry. The percentage of bacterial cells expressing PSA from mice with bacteriophage Barc2635 was significantly lower in comparison to bacterial cells from mice without the bacteriophage

[Figure 4D], in agreement with the *upaY* expression levels. In both groups of mice (i.e., *B. fragilis* alone and *B. fragilis* with Barc2635), the CFU levels of *B. fragilis* were comparable, with a $10^{0.5}$ difference **[Supp. Figure 2D]**, and no phages were detected in the *B. fragilis* monocolonized group (i.e., without bacteriophage inoculation) **[Supp. Figure 2E]**.

In a previous study, we demonstrated that phase variation of specificity proteins for a Type I restriction-modification system (R-M) of *B. fragilis* results in altered transcription of some capsular polysaccharides with potential altered immunomodulatory functionality of these bacteria[7]. Since the Type I R-M systems are known to provide protection against bacteriophage[37–40], we used long-reads sequencing to quantify the orientations of the genes encoding the Type I R-M specificity proteins upon exposure to Barc2635. Our results demonstrate different patterns of specificity gene combination in the expression locus in *B. fragilis* isolated from mice inoculated with the phage compared to those not exposed to phage **[Figure 4E]**.



Supplemental Figure 1: Characterization of Bacteriophage Barc2635
A. A representative transmission electron microscopy of Barc2635. **B.** Barc2635 plaques on a lawn *B. fragilis* NCTC 9343. **C.** Genome structure, GC content, and putative annotations of Barc2635. CDS: coding sequence.

The PSA of B*. fragilis* induces regulatory T cells (Tregs, $CD4^+Foxp3^+Rorg^+$) in the colonic lamina propria of mice[41–44]. To monitor the effect of intestinal Barc2635 during *B. fragilis* colonization on the $CD4^+Foxp3^+Rorg^+$ Tregs population, we extracted cells from the colonic lamina propria and immunophenotyped them by flow cytometry. We found that in the presence of Barc2635, the $CD4^+Foxp3^+Rorg^+$ Tregs population is decreased in concert with the decrease in the *upaY* transcription, coinciding with less surface production of the immunomodulatory PSA **[Figure 4F, Figure 4G]**. These data link phage to alterations in bacterial functionality with concomitant effects on host physiology.

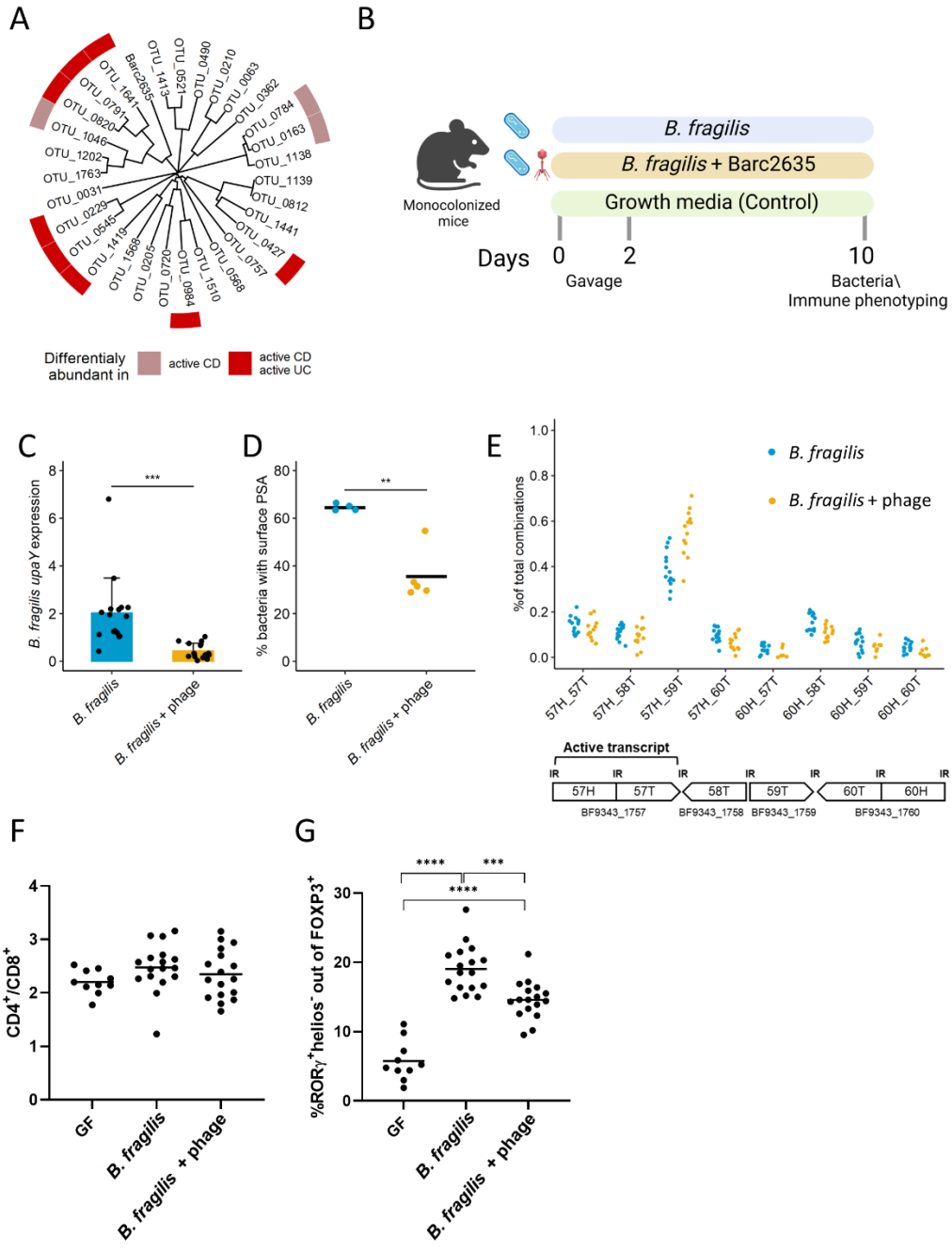
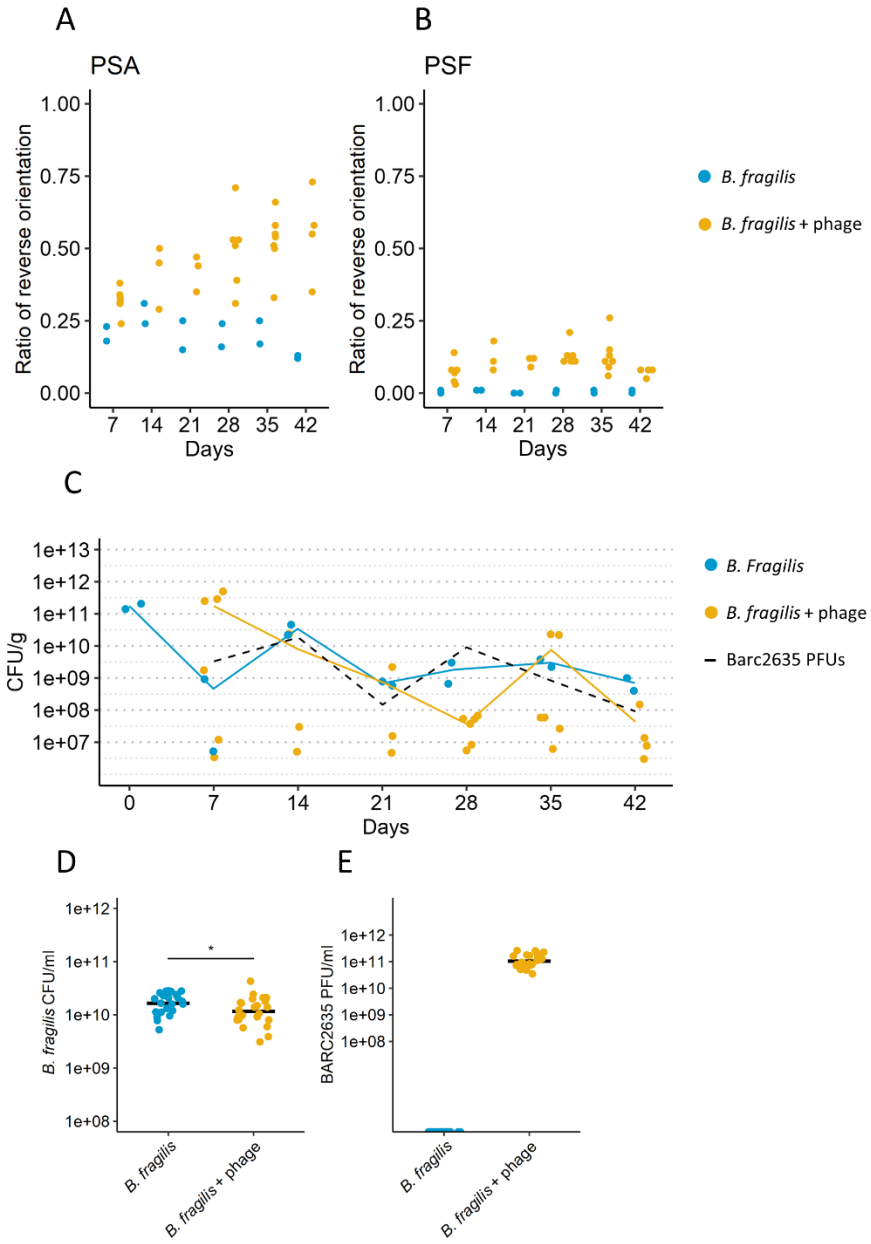


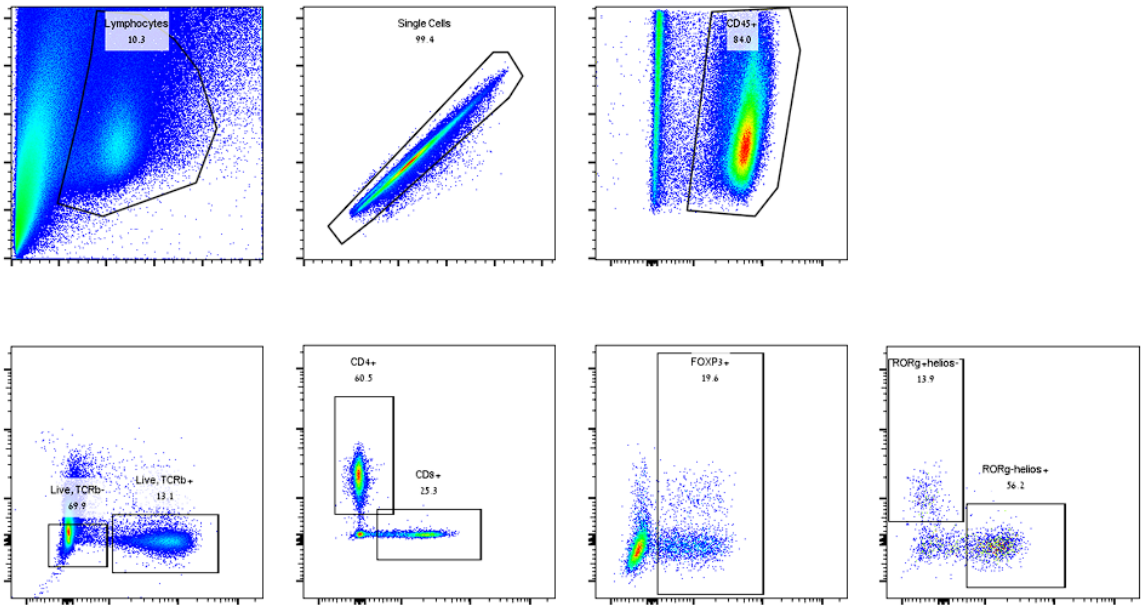
Figure 4: Phage exposure alters expression of the PSA locus and PSA surface

A. Phylogenetic tree based on the whole genome of viral OTUs identified as bacteriophages against *B. fragilis* as well as bacteriophage Barc2635. MAFFT was used to perform multiple sequence alignment. The average-linkage method was used to construct the phylogenetic tree with 1000 bootstrap replicates. Colors denote association between the viral OTUs abundances in Nishiyama et al. (2020), IBDMDB cohort, with active Crohn's disease (light red) or with both active Ulcerative colitis and active Crohn's disease (red). **B.** Experimental design of *in vivo* experiments. **C.** Expression levels of *upaY* in ceca of mice monocolonized with *B. fragilis* vs. mice monocolonized and infected with the Barc2635 bacteriophage. Levels are shown as 2^(-**Δ**CT) with *rpsL* as a reference gene. Each dot represents a mouse. ****P<0.0001, Mann-Whitney test. **D.** PSA presence on the surface of *B. fragilis* exposed to the Barc2635 bacteriophage, detected by anti-PSA antibodies. Bacteria were analyzed by flow cytometry for the expression of PSA, using Rabbit anti PSA antibodies. Each dot represents a mouse. **E.** Percent of the bacterial population with each of the phase-variable Type-I R-M specificity gene combinations at the expression locus. Each dot represents a mouse. IR: inverted repeat. **F.** CD4$^+$ to CD8$^+$ ratio out of CD45$^+$TCRβ^+ live cells in colon. Single cells were isolated from colon lamina propria. Immune cells were analyzed by flow cytometry. Each dot represents a mouse. **G.** RORγ^+ helios$^-$ percentages out of FOXP3$^+$ cells in colon. Single cells were isolated from colon lamina propria. Immune cells were analyzed by flow cytometry. Each dot represents a mouse. ***P<0.001, ****P<0.0001, one-way analysis of variance (ANOVA).



Supplemental Figure 2

A. Ratio of *B. fragilis* PSA promoter reverse orientation in mice fecal samples measured by the PhaseFinder tool, on different days of the experiment. Blue: *B. fragilis*, yellow: *B. fragilis* + Barc2635. **B.** Ratio of *B. fragilis* PSF promoter reverse orientation in mice fecal samples measured by the PhaseFinder tool, on different days of the experiment. Blue: *B. fragilis*, yellow: *B. fragilis* + Barc2635. **C.** CFUs of *B. fragilis* in fecal samples of mice monocolonized with *B. fragilis* (blue) or colonized with *B. fragilis* + phage (yellow); PFUs of Barc2635 in fecal samples of mice colonized with *B. fragilis* + phage (dashed black line). **D.** CFUs of *B. fragilis* in fecal samples of mice monocolonized with *B. fragilis* (blue) or colonized with *B. fragilis* + phage (yellow) (Wilcoxon rank sum test, $*p < 0.05$). **E.** PFUs of Barc2635 in fecal samples of mice monocolonized with *B. fragilis* (blue) or colonized with *B. fragilis* + phage (yellow).



Supplemental Figure 3 – gating strategy

Representative flow cytometry plots demonstrating the gating strategy for the staining panel.

Discussion

Phase variations, prevalent in host-associated species, especially in the abundant gut Bacteroidales order[9], contribute to bacterial fitness in changing ecosystems, such as the human gut. Reversible DNA inversions lead to phase variable synthesis of numerous molecules (e.g. surface, regulatory, and other molecules), and as such, confer functional plasticity. Our study reveals that phase variable states of certain molecules correlate with gut inflammation, with potential implications on host physiology. By analyzing six different public databases of IBD patients (CD and UC) and controlled mice experiments, we identify multiple invertible genomic regions that map to 25 different Bacteroidales strains. Notably, we find that both intergenic regions (e.g. promoters) and intragenic regions (e.g. genomic shufflons) exhibit altered genomic phase variable states in gut inflammatory conditions - both affecting gene expression.

The most prevalent genomic regions that demonstrated differential orientations of invertible regions were in susC/susD homologs and polysaccharide (PS) promoters - both with immunomodulatory potential. An epitope of SusC-like proteins, part of the starch-utilization system in *Bacteroidetes*, was recently shown to elicit T cell responses in IBD patients and healthy controls[45], suggesting that operons that include SusC homologs might confer immunomodulatory properties to the bacteria, and phase variation in such genes might alter bacteria-host interactions. Among the phase-variable PSs, the anti-

358 inflammatory polysaccharide A (PSA) promoter of *B. fragilis* showed a higher percentage of reverse
359 oriented reads in IBD patients compared to healthy controls, indicating that the 'OFF' orientation was
360 more prevalent in IBD patients, potentially limiting the protective, anti-inflammatory, effects of PSA.
361 These results align with a previous study[46] that focused on the PSA promoter of *B. fragilis* in IBD
362 patients using PCR digestion of biopsy samples.

363 We further find that some DNA inversion states, like the PSA promoter of *B. fragilis*, are dynamic and
364 can respond to changes in their local environment, specifically during inflammation. Our analysis
365 focused on the bacterial population level, and therefore the DNA inversions could either occur in
366 individual bacterial cells in response to inflammation, or as a selection process on the bacterial
367 population level, where those with the promoter OFF orientation are enriched under inflammatory
368 conditions.

369 Intriguingly, the preferential 'OFF' orientation of the PSA promoter of *B. fragilis* did not occur to the
370 same extent in inflamed monocolonized mice as it did in both the inflamed mice colonized with a
371 human microbiota and when bacteria are grown in patients' fecal filtrates. These results suggest that
372 environmental factors in humanized mice and patients are necessary for this effect. Our analysis of
373 phage OTUs in the IBDMDB cohorts revealed enrichment of *B. fragilis* associated viral OTUs in samples
374 where the PSA promoter was present in the 'OFF' orientation. In addition, exposure of *B. fragilis* to
375 Barc2635 resulted in the PSA promoter 'OFF' orientation, suggesting a role for gut bacteriophages in
376 bacterial phase variation.

377 Type I restriction-modification (R-M) systems provide protection against bacteriophages by identifying
378 and restricting foreign DNA. We previously characterized a phase variable Type I R-M system of *B.*
379 *fragilis* that regulates distinct transcriptional programs affecting polysaccharide synthesis. Along with
380 the effects on the PSA promoter orientation, infection with Barc2635 also triggered genomic inversion
381 of this Type I R-M system towards inhibition of PSA synthesis[7,36].

382 The anti-inflammatory effects of PSA are mediated by upregulation of Tregs and subsequent induction
383 of IL-10 secretion[12,13]. We found that the Treg cells populations decreased in *B. fragilis* monocolonized
384 mice infected with Barc2635, suggesting that the PSA 'OFF' trigger of Barc2635 has implications on
385 the host immune system[11]. Applying the same analysis to *B. thetaiotaomicron*'s CPS3, we observed
386 bacteriophages that are more abundant in samples with a higher ratio of CPS3 'ON' oriented
387 promoters, however, the bacteriophage to host ratios per sample were too low to conclude that *B.*
388 *thetaiotaomicron* bacteriophages lead to populations with altered CPS3 synthesis. Altogether, we
389 demonstrate differential orientations of invertible regions in IBD and elucidate the bacteriophages as

a potential trigger. These phase variations result in bacterial functional plasticity affecting the host immune system. Bacteriophages are an example of an environmental factor that can alter phase variable states. However, within the inflamed gut there are numerous factors that can potentially alter phase variable states. These include the microbiome[33,47–50] (e.g. neighboring bacteria and viruses), the gut metabolome[47], an altered immune system[51] and physical alterations in the intestine such as abnormal pH concentrations[52], osmotic[53,54] and oxidative stress[55,56]. For example, Tropini et al.[57] showed that PEG induced osmotic perturbation resulted in an increased ratio of CPS4 to CPS5 in *B. thetaiotaomicron*.

This study demonstrates that alterations in invertible genomic regions, and consequently molecular phase variations in gut bacteria, during IBD, can potentially influence the host immune system and inflammatory state. These phase variations can occur in response to changes in gut environmental factors during inflammation, including, but not limited to bacteriophages. Future studies integrating bacterial DNA inversions and phase variation analyses may illuminate the role of bacterial functional plasticity in additional physiological states, possibly resulting from different environmental factors, which may drive alterations in microbe-host interactions.

Limitations of the Study:

The six databases analyzed in this study included adults and children, healthy controls and inflamed patients (before and after treatment) and longitudinal samples. The impact of each of these factors, and the impact of diet and lifestyle, merit further research. The PhaseFinder[9] algorithm that we applied identifies DNA inversions in short reads. As such, our analysis does not include other phase variation mechanisms like duplications or insertions. Also, analysis of short reads of metagenomics sequencing data might overlook inversions resulting from long structural genomic alterations[3,7,58]. Furthermore, metagenomic sequencing data averages the whole bacterial population of samples, implying a dilution factor for low abundant bacteria. Moreover, genomic inversion sites might interfere with genome sequence assemblies, leading to incomplete genomes, and thus could be overlooked when applying tools that search inversions in reference genomes. This study used an identified and isolated lytic bacteriophage for *B. fragilis*, further studies are required to elucidate the functional effects of IBD associated lytic and temperate bacteriophages. With these limitations notwithstanding, our study highlights the importance of considering bacterial phase variation in the context of IBD and its potential impact on inflammation and on altered microbe-host interactions.

Methods

Mice

All mouse work was in accordance with protocols approved by the local IACUC committee under approval numbers: IL-151-10-21 and IL-105-06-21.

Males, 4-5 weeks-old Germ-Free (GF) C57BL/6 or Swiss Webster mice from the Technion colony were used. Mice were housed and maintained in a GF care facility and were provided with food and water *ad libitum*; they were exposed to a 12:12 h light-dark cycle at room temperature.

Humanized mice were created by oral gavage of GF C57BL/6 mice with human feces from a single healthy human donor, spiked with *B. fragilis* NCTC 9343.
Monocolonized mice were created by oral gavage of GF C57BL/6 mice with *B. fragilis* NCTC 9343.

For the prolonged experiments of Barc2635 and *B. fragilis* co-existence in gnotobiotic mice, female Swiss Webster GF mice were administered with *B. fragilis* NCTC 9343 and bacteriophage Barc2635, or *B. fragilis* only. All administrations were performed by oral gavage. On day 49 mice were sacrificed for bacteria and immune phenotyping.

For phage immunomodulation studies, GF mice were gavaged twice (on day 0 and on day 2) with *B. fragilis* NCTC 9343 and bacteriophage Barc2635, *B. fragilis* only, or growth media as control. On day 10 mice were sacrificed for bacteria and immune phenotyping.

DSS model

For acute DSS-induced colitis, humanized mice were treated with 3% Dextran sodium sulfate in their drinking water for 9 days followed by 5 days recovery before sacrificing. The control animals were administered distilled water.
Fecal samples were obtained on days 0, 6, 9 and 14 (before DSS treatment, during and after discontinuing DSS treatments), for further analyses.

Mice Inflammation measurement

Fecal supernatants were prepared after suspension in 1:10 sterile PBS and centrifugation at 4,500xg for 15 minutes. Enzyme-linked immunosorbent assay (ELISA) to measure calprotectin concentrations was performed using Mouse S100A8/S100A9 Heterodimer kit according to the manufacturer protocol [R&D systems].
Mice weights were assessed using an electronic scale at the same daytime.

453

454 qPCR and primers

455 DNA was extracted from fecal samples using ZymoBIOMICS DNA Miniprep Kit [Zymo research]. The
456 'ON'/'OFF' orientation of the PSA promoter in the extracted DNA samples was determined by
457 quantitative polymerase chain reaction (qPCR) using SYBR® Green mix [Thermo Fisher Scientific]. Two
458 sets of primers were designed to target the PSA locus. One set was used as a proxy to the number of
459 bacteria in the samples and targeted *UpaY*, the first gene immediately downstream to the promoter
460 region. The second set of primers targeted the promoter region and would only produce a product
461 when the orientation is 'ON'. The ratio of 'ON'/'OFF' PSA orientation in the samples was calculated
462 using the 2−ΔΔCT method calculated against the PSA locked 'ON' results (100% 'ON' orientation).

463

464 *upaY* gene expression was determined by RT-qPCR. RNA was extracted from fecal samples using
465 zymoBIOMICS RNA miniprep kit [Zymo research]. Reverse transcription of RNA to cDNA was
466 performed using the qScript cDNA Synthesis Kit [Quantabio]. *upaY* mRNA levels were determined by
467 qPCR using SYBR® Green mix [Thermo Fisher Scientific] with primers against *rpsL* as a reference gene.
468 The 2−ΔΔCT method was employed for the specificity fold change tests.
469 Primers:

B_Frag_upaY_F	CGCTCGGACAAAGAAGGACC
B_Frag_upaY_R	ACTTCTACCCTACGACGACGA
B_Frag_PSA_M	GGTGTTCCAAAAGACGAACGT
B_Frag_PSA_F	TGTGTAAATGATAGGAGGCTAGGG
rpsL_F	CCGAACTCTGCAATGCGTAA
rpsL_R	CGCGAACCAGTACGATTGAG

470

471 Mice feces sequencing

472 DNA samples underwent quality control by Qubit fluorescence analysis to determine concentration of
473 DNA for downstream analysis (ThermoFisher, Cat. Q32850). Libraries were prepared using the Illumina
474 Tagmentation DNA prep streamlined library preparation protocol according to manufacturer's
475 instructions with a minimum of 50 ng of DNA starting mass and 8 cycles of PCR enrichment, ending
476 with a fragment size of 550 bp. IDT for Illumina DNA/RNA UD indexes and Nextera DNA CD indexes
477 were used (Illumina IDT, Cat. 20027213; Illumina Nextera, Cat. 20018708).

All libraries were diluted to 15 pM in 96-plex pools and validated on 100-cycle paired-ends read Miseq V2 runs (Illumina, Cat. MS-102-2002), before shipping to the US at 4 nM for sequencing on the Novaseq 6000 in S4 mode at 96-plex in a 300-cycle paired-end reads run, with an estimated read depth of 30 Gbp per sample (Illumina, Cat. 20028312). Final loading concentration of 600 pM. All sequencing runs were performed with a spike-in of 1% PhiX control library V3 (Illumina, Cat. FC-110-3001).

Taxonomic profiling

Community profiling was performed using metaphlan4 v4.0.0[59] with mpa database vJan21. For each sample, the forward reads were first aligned against the mpa database using bowtie2 v2.3.5.1[60] (flags **`--sam-no-hd --sam-no-sq --no-unal --very-sensitive`**). Next, the resulting sam file was analyzed by metaphlan4 with default parameters. In each analysis, species at abundance <=0.1% were ignored.

Microbiome analysis

Statistical analysis of sequenced data was initially performed using MicrobiomeAnalyst[61] followed by comprehensive analysis using R packages: Phyloseq[62], Vegan[63] and DESeq2[64]. Differences in microbial taxa and functional modules were assessed by differential abundance analyses using DESeq2[64]. Alpha diversity indices (observed species and Shannon) were compared using the Kruskal–Wallis rank sum test. Beta diversity distance matrices (Bray-Curtis) were compared using the vegan package's function ADONIS, a multivariate ANOVA based on dissimilarity tests and visualized using PCoA. Results were visualized using the ggplot2 R package[65].

Identification of DNA inversion sites

Representative *Bacteroides* species were selected from an in-house database of human-associated microbial species. The database was constructed from 118K metagenomics assembled genomes (MAGs) recovered from human-associated metagenomics samples. Taxonomy assignment was performed using gtdbtk v2.0.0 and GTDB release 207 with the classify_wf program and default parameters[66]. Overall, 36 *Bacteroides* species were identified of which 35 had species-level GTDB taxonomy assignments. For each of these species, the representative genome from GTDB was used for the PhaseFinder analysis. We have also included the genomes of two additional *B. fragilis* strains, *Phocaeicola dorei*, and *Parabacteroides distasonis*. Overall, 40 genomes were included. Information about the genomes is provided in Table S1.

PhaseFinder[9] (v1.0) was used to identify invertible sites in the metagenomics samples. The default parameters of PhaseFinder were used. We filtered the results by removing identified sites with <20

reads supporting either the forward or reverse orientations combined from the paired-end method, mean Pe_ratio <1% across all samples, and sites within coding regions of rRNA products. To compare DNA orientation patterns between CD and UC patients with those of healthy controls, we focused on sites that displayed a difference of over 10% between at least one of the groups. The Wilcoxon rank sum test was employed to conduct these comparisons, and the Benjamini-Hochberg method was utilized to correct for multiple comparisons, with a false discovery rate (FDR) set at less than 0.1. Each invertible region was manually curated to assess its coding regions, gene annotations, and their putative functions. Briefly, genomic regions were visualized online using the NCBI Graphical Sequence Viewer (version 3.47.0). Invertible regions with coding sequences (CDS) within them, were annotated according to the CDS name(s). Invertible regions lacking CDS within them were searched for CDSs that start in proximity to the invertible DNA sites (<200~bp). CDSs in the region (four upstream and four downstream) were used to assess the functionality of the region. Regions containing or in proximity to rRNA and tRNA genes were filtered from the comparisons as well as invertible regions with no CDSs start in proximity to the inverted repeats.

Bacteriophage isolation

For isolation of bacteriophages, inflowing raw sewage for a waste-water treatment plant (WWTP) from Barcelona (Spain) was filtered through low protein binding 0.22 µm pore size polyethersulfone (PES) membrane filters (Millex- GP, Millipore, Bedford, Massachusetts) to remove bacteria. Isolated plaques were obtained by the double-agar layer technique[67]. Briefly, tubes containing 2.5 ml of soft BPRM–agar kept at 45°C were inoculated with 1 ml of an exponential growth phase culture (OD600=0.3, corresponding to ca $2x10^8$ CFU / ml) of the host bacteria grown in BPRM broth and 1 ml of the filtered sewage sample. After gently mixing, the contents of each tube were poured onto a plate of BRPM-agar and incubated inside GasPak (BBL) jars at 37°C. Plaques were clearly spotted after 18 h of incubation.

For phage isolation, discrete well-isolated plaques were stabbed with a sterile needle and inoculated in a tube containing 5 ml of BRPM broth. Then 1-ml of a culture of *B. fragilis* NCTC 9343in exponential growth was inoculated into the tube, which was then incubated for 18h at 37 °C. After incubation, an aliquot of the culture was treated with chloroform (1:10 (v;v), vigorously mixed for 5 minutes and centrifuged at 16,000x*g* for 5 minutes[68]. The supernatant containing the phage suspensions were further filtered through low protein binding 0.22 µm pore size polyethersulfone (PES) membrane filters (Millex-GP, Millipore, Bedford, Massachusetts), diluted and plated as indicated in the previous paragraph to verify the uniformity of the plaques. Then, one well differentiated plaque was stabbed

and the whole operation repeated to obtain a high titer, over $1x10^9$ plaque forming units (PFU), phage suspensions.

Phage genome sequencing

5hrs induction of phage DNA, phage particles were PEG-precipitated (6,000-12,000 MW, 8%), and isolated by a CsCl gradient; 33 g, 41 g, 55 g in 50 ml TM buffer (50mM Tris-Cl pH8.0, 10mM MgCl2), ultracentrifuged at 40,000 rpm for 1.5hrs, and dialyzed overnight in 100mM Tris ph7.5, 1M NaCl, 1mM EDTA 56. Genomic DNA was extracted using phenol-chloroform as described 57.

Illumina sequencing of Bacteroides phage Barc2635 was performed at the Biopolymers Facility, Harvard Medical School, Department of Genetics, producing paired-end reads of 150 bp. Adapter sequence removal and quality trimming was performed using BBDuk, part of the BBTools (v 37.50) suite of programs. The reads were further screened against NCBI's UniVec_Core database (build 10.0) and the B. fragilis NCTC 9343 genome sequence using blastn and reads that returned a significant hit to either were removed. The phage genome was assembled de novo using Velvet 1.2.10 under a k-value determined by Velvet Optimizer (v. 2.2.5). The genome was annotated using an in-house customized version of Prokka v1.12, submitted to NCBI, and assigned GenBank accession MN078104. Phage Genome map **[Supp. Figure 1C]** was visualized using the online tool Proksee (https://proksee.ca/, accessed on 21 January 2023).

Viral OTUs multiple alignment

Alignments and phylogenetic tree of whole genomes of viral OTUs identified as bacteriophages against *B. fragilis*[33] as well as bacteriophage Barc2635 were generated using MAFFT online service (version 7). Multiple sequence alignment was done using the default settings of the site. Phylogenetic tree was constructed using the average-linkage method with 1000 bootstrap replicates and was visualized with R package ggtree[69] (version 3.4.4).

MinION library preparation and sequencing

The specificity region of BF9343_1757–1760 was amplified by PCR from a population of bacteria grown in vitro and in vivo from fecal content as described above. The primers annealed outside the invertible region. The amplicons were purified by Wizard SV Gel and PCR Clean-Up System (Promega) and measured by nanodrop.

Primers:

Type1RM_hsdS_F	GACAATCGAGATGAAGAACAAC
Type1RM_hsdS_R	CCATAGGCGTATGATTTCCTG

575

576 DNA quantity was measured again using Qubit fluorometry (Thermo Fisher Scientific, Waltham, MA,
577 USA). Nanopore sequencing libraries were prepared from 200 fmol purified amplicons using Ligation
578 Sequencing Kit 1D (SQK-LSK109) and PCR-free Native Barcoding Expansion Kit (EXP-NBD104) (Oxford
579 Nanopore Technologies, Oxford, England). The barcoded libraries were loaded and sequenced on the
580 MinION device controlled by MinKNOW software (v.19.12.5) using MinION flow cells (FLO-MIN106D
581 R9.4.1, Oxford Nanopore Technologies, Oxford, England) after quality control runs. The raw data were
582 base called and demultiplexed by Guppy Basecalling Software (v. 3.3.3+fa743a6).

583 MinION data analysis

584 Adapters and barcodes sequences were removed from the reads using Porechop (v0.2.4, available
585 from https://github.com/rrwick/Porechop). Reads were oriented using the 'Preparing Reads for
586 Stranded Mapping' protocol (Eccles, D.A. (2019). Protocols.io[70]). We aligned the reads to the PCR
587 forward primer using LASTAL (v.1060)[71], and then reverse-complemented the reverse-oriented reads.
588 The reads were combined to an all forward oriented file and cropped to the first 1300 bases using
589 Trimmomatic (v.0.39)[72]. The reads were then split according to their alignment to the 1757-57 or 1757-
590 60 5" half sequences using LASTAL. Reads were mapped to the full sequences with Minimap2 (v.2.17-
591 r941)[73] using the –for-only and asm20 options. Mapped read counts were extracted from the
592 Minimap2 SAM output using SAMtools (v.1.7)[74].

593 MinION sequence data has been deposited in the NCBI sequence read archive (SRA) under the
594 BioProject accession number: PRJNA948162

595 Fecal filtrates of patients

596 Recruitment of IBD patients for this study was conducted at the Rambam Health Care Campus (RHCC).
597 The study was approved by the local institutional review boards with study numbers 0052-17 and
598 0075-09 in which all patients consented to be included in it. During therapy, patients were treated
599 either with Infliximab or Humira for at least two weeks.

600 Fecal samples were collected by patients at home prior to their clinic visit and collected from each at
601 the hospital. The samples were then stored at -80°C until they were shipped to the laboratory for
602 analysis.

Calprotectin levels were measured in each fecal sample using LIAISON Calprotectin (catalogue No. 318960) according to the manufacturer's instructions. The levels of calprotectin in the fecal samples were used as a measure of disease activity in IBD patients.

In vitro assays with fecal filtrates from IBD patients

Bacteroides fragilis NCTC 9343 was grown in Brain heart infusion (BHIS) supplemented with 5 mg/l hemin and 2.5 µg/l vitamin K to OD_{600}~0.6 and centrifuged at 4,500xg for 5 minutes. Bacterial pellets were washed twice with sterile PBS to remove BHIS components and then suspended with 1mL supplemented M9 minimal media. Patients' fecal samples were suspended in sterile PBS (1:5), centrifuged at 4,500xg for 15 minutes, and supernatants were collected and filtered using the Medical Millex-VV Syringe Filter Unit, 0.22 µm, PVDF membrane. For the *in vitro* assay, *B. fragilis* was cultured in patients' supernatants and diluted in M9 minimal media in a 1:25:25 ratio (bacteria: fecal supernatants: media). Subsequently, 200µl of each culture was collected for DNA extraction at OD_{600}~0.6.

M9 Medium was prepared as follows:
Per L:
1g NH_4Cl
6g $Na2HPO_4$
3g KH2PO4
0.5g NaCl

Supplements:
14.7mg $CaCl_2.2H_2O$ (0.1ml 1M stock)
246mg $MgSO_4.7H_2O$ (1ml 1M stock)
0.5% glucose (20ml 25%)
0.05% L-cysteine (10ml 5% stock)
5mg/ml Hemin (1ml 0.5% stock)
2.5mg/ml $VitK_1$ (0.5ml 0.5% stock)
2mg/ml $FeSO_4.7H_2O$ (0.1ml 2% stock)
5ng/ml VitB12 (0.05ml 0.01% stock)

Gut lamina propria preparation

For lamina propria immunophenotyping, mice colons were removed by cutting the colon from the cecum-colon junction to the anus. Fat tissue was carefully removed from colon tissue and further proceeded for single cell suspension preparation using lamina propria dissociation kit (Miltenyi), according to the manufacturer's protocol.

640

641 Flow cytometry

642 Cell preparations for flow cytometry analysis were performed in 5 ml tubes or U shape 96 wells plates.
643 Single cells were washed with PBS and stained for live/dead staining using 1:1,000 in PBS, Zombie
644 fixable viability dye (Biolegend) for 10 minutes, at room temperature, and washed once with FACS
645 buffer, by centrifuge at 300xg for 5 minutes. For FcR blocking, cells were incubated with 0.5 µg
646 CD16/CD32 antibody for 10 minutes on ice and proceeded to further staining without a washing step.
647 Extracellular markers were stained with the relevant antibody panels for 30 minutes on ice and
648 washed twice with FACS buffer, by centrifuge at 300xg for 5 minutes. After the last wash, cells were
649 fixed with Foxp3 Fixation/Permeabilization working solution (Thermo) for 16 hours at 4 °C in the dark.
650 For Intracellular staining, cells were permeabilized using 1X Foxp3 permeabilization buffer (Thermo)
651 according to the manufacturer protocol. For intracellular blocking, 2 µl of 2% rat serum (Stemcell
652 technologies) was added to each well for 15 minutes at room temperature and proceeded to further
653 staining without a washing step.
654 To quantify the percentage of *B. fragilis* from monocolonized mice with PSA on their surface, feces
655 from monocolonized mice with *B. fragilis* with or without phage were suspended 1:10 in ice cold PBS
656 (mg/µl) and centrifuged at 300xg for 5 minutes, 4°C. Supernatants were separated from pellets and
657 further centrifuged at 4,500xg for 5 minutes, 4°C. Bacterial pellets were resuspended in an ice cold
658 FACS buffer, 1:10 from initial PBS suspension. 100 µl of resuspended bacteria were incubated with
659 1:1,000 Rabbit anti *B. fragilis* PSA for 30 minutes at 4°C. Bacteria were washed twice using an ice cold
660 FACS buffer by centrifuge at 4,500xg for 5 minutes and then incubated with a donkey anti rabbit
661 fluorophore conjugated secondary antibody. After staining steps, the bacteria were washed twice with
662 an ice cold FACS buffer and finally resuspended in 500 µl ice cold PBS plus 1:1,000 Hoechst dye and
663 analyzed by flow cytometry using FSC and SSC thresholds of 1,000, and logarithmic scale. Gating
664 strategy is detailed in **Supplemental Figure 3**.

665

666 Funding

667 This work was supported by the Technion Institute of Technology, 'Keren Hanasi', Cathedra, the
668 Rappaport Technion Integrated Cancer Center, the Alon Fellowship for Outstanding Young
669 Researchers, the Israeli Science Foundation (grant 1571/17 and 3165/20), Israel Cancer Research Fund
670 Research Career Development Award, the Seerave Foundation, the Canadian Institute for Advanced
671 Research (Azrieli Global Scholars; grant FL-000969), Human Frontier Science Program Career
672 Development Award (grant CDA00025/2019-C), the Gutwirth foundation award and the European

673 Union (ERC, ExtractABact, 101078712). Views and opinions expressed are, however, those of the
674 author(s) only and do not necessarily reflect those of the European Union or the European Research
675 Council Executive Agency. Neither the European Union nor the granting authority can be held
676 responsible for them. NGZ is an Azrieli Global Scholar at the Canadian Institute for Advanced Research,
677 and a Horev Fellow (Taub Foundation). MJC and LC are supported by the Duchossois Family Institute.
678 SC is supported by the Gutwirth Excellence Scholarship and by Teva Pharmaceutical Industries as part
679 of the Israeli National Forum for BioInnovators (NFBI). HH is supported by Leonard and Diane Sherman
680 Interdisciplinary Graduate School fellowship, Wjuniski Fellowship Fund for the MD/PhD Medical
681 Scientist Program, and by VATAT fellowship for outstanding doctoral students from the Arab
682 community fellowship.

683 ## Acknowledgements

684 We would like to thank the Geva-Zatorsky lab members for fruitful discussions and contributions.
685 Thanking Zachary Merenstein and Neekoo Farahmandpour for helping in establishing the qPCR
686 method; Dr. Svetlana Friedman, Ran Tahan and Prof. Debbie Lindell for advising us on the
687 bacteriophage characterization and Dr. Carolina Tropini for insightful discussions. We thank Drs. Amir
688 Grau, Ofer Shenker, and the biomedical core facility at the Technion Rappaport faculty of Medicine
689 for their help with flow cytometry; BiotaX for metagenomics sequencing, and Prof. Ogata and Dr.
690 Nishiyama from the Institute for Chemical Research, Kyoto University, for helping in accessing their
691 data. Figures 2a, 3a and 4b were created with BioRender.com.
692

693 ## Author contributions

694 SC HH RZ DKK TG IS and NGZ conceived and designed the project. SC HH RZ DKK and TG performed
695 and analyzed the experiments. SC, IS and NGZ planned the computational and analytical aspects; SC
696 and IS performed the computational analyses. MJC and LC contributed to the computational analysis
697 and interpretation of the results. SP and YC designed, lead, and executed the clinical study. KJ
698 contributed to the phage-bacteria-host study and to the bacteriophage morphological
699 characterization. JJ isolated the bacteriophage. NM RN, and NBA helped with the experiments. SC HH
700 RZ TG IS LC and NGZ wrote the manuscript. SC RZ HH and DKK contributed equally to the study. All
701 authors read and approved the manuscript. NGZ conceived and planned the study, supervised it,
702 interpreted the experiments, and wrote the manuscript.
703
704
705

706

References

1. Moxon R, Bayliss C, Hood D. Bacterial contingency loci: The role of simple sequence DNA repeats in bacterial adaptation. *Annu Rev Genet*. 2006;40:307-333. doi:10.1146/annurev.genet.40.110405.090442

2. Phillips ZN, Tram G, Seib KL, Atack JM. Phase-variable bacterial loci: how bacteria gamble to maximise fitness in changing environments. *Biochem Soc Trans*. 2019;47(4):1131-1141. doi:10.1042/BST20180633

3. West PT, Chanin RB, Bhatt AS. From genome structure to function: insights into structural variation in microbiology. *Curr Opin Microbiol*. 2022;69:102192. doi:10.1016/J.MIB.2022.102192

4. Goldberg A, Fridman O, Ronin I, Balaban NQ. Systematic identification and quantification of phase variation in commensal and pathogenic Escherichia coli. *Genome Med*. 2014;6(11):112. doi:10.1186/s13073-014-0112-4

5. Krinos CM, Coyne MJ, Weinacht KG, Tzianabos AO, Kasper DL, Comstock LE. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature*. 2001;414(6863):555-558. doi:10.1038/35107092

6. Van Der Woude MW, Bäumler AJ. Phase and Antigenic Variation in Bacteria. *Clin Microbiol Rev*. 2004;17(3):581. doi:10.1128/CMR.17.3.581-611.2004

7. Ben-Assa N, Coyne MJ, Fomenkov A, et al. Analysis of a phase-variable restriction modification system of the human gut symbiont Bacteroides fragilis. *Nucleic Acids Res*. 2020;48(19):11040-11053. doi:10.1093/nar/gkaa824

8. Porter NT, Hryckowian AJ, Merrill BD, et al. Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. *Nat Microbiol*. 2020;5(9):1170-1181. doi:10.1038/s41564-020-0746-5

9. Jiang X, Brantley Hall A, Arthur TD, et al. Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. *Science (1979)*. 2019;363(6423):181-187. doi:10.1126/science.aau5238

10. Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature*. 2008;453(7195):620-625. doi:10.1038/nature07008

11. Surana NK, Kasper DL. The yin yang of bacterial polysaccharides: lessons learned from B. fragilis PSA. *Immunol Rev*. 2012;245(1):13-26. doi:10.1111/j.1600-065X.2011.01075.x

12. Dasgupta S, Erturk-Hasdemir D, Ochoa-Reparaz J, Reinecker HC, Kasper DL. Plasmacytoid dendritic cells mediate anti-inflammatory responses to a gut commensal molecule via both innate and adaptive mechanisms. *Cell Host Microbe*. 2014;15(4):413-423. doi:10.1016/j.chom.2014.03.006

13. Round JL, Mazmanian SK. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc Natl Acad Sci U S A*. 2010;107(27):12204-12209. doi:10.1073/PNAS.0909122107

14. Chang YC, Ching YH, Chiu CC, et al. TLR2 and interleukin-10 are involved in Bacteroides fragilis-mediated prevention of DSS-induced colitis in gnotobiotic mice. *PLoS One*. 2017;12(7). doi:10.1371/JOURNAL.PONE.0180025

15. Ramos GP, Papadakis KA. Mechanisms of Disease: Inflammatory Bowel Diseases. *Mayo Clin Proc*. 2019;94(1):155-165. doi:10.1016/J.MAYOCP.2018.09.013

16. Guan Q. A Comprehensive Review and Update on the Pathogenesis of Inflammatory Bowel Disease. *J Immunol Res*. 2019;2019. doi:10.1155/2019/7247238

17. Corridoni D, Arseneau KO, Cominelli F. Inflammatory bowel disease. *Immunol Lett*. 2014;161(2):231-235. doi:10.1016/J.IMLET.2014.04.004

18. Ha CWY, Martin A, Sepich-Poore GD, et al. Translocation of Viable Gut Microbiota to Mesenteric Adipose Drives Formation of Creeping Fat in Humans. *Cell*. 2020;183(3):666-683.e17. doi:10.1016/J.CELL.2020.09.009

19. Bolam DN, Koropatkin NM. Glycan recognition by the Bacteroidetes Sus-like systems. *Curr Opin Struct Biol*. 2012;22(5):563-569. doi:10.1016/J.SBI.2012.06.006

20. Kulagina E v., Efimov BA, Maximov PY, Kafarskaia LI, Chaplin A v., Shkoporov AN. Species composition of Bacteroidales order bacteria in the feces of healthy people of various ages. *Biosci Biotechnol Biochem*. 2012;76(1):169-171. doi:10.1271/BBB.110434

21. Kraal L, Abubucker S, Kota K, Fischbach MA, Mitreva M. The Prevalence of Species and Strains in the Human Microbiome: A Resource for Experimental Efforts. *PLoS One*. 2014;9(5):e97279. doi:10.1371/JOURNAL.PONE.0097279

22. Nomura K, Ishikawa D, Okahara K, et al. Bacteroidetes species are correlated with disease activity in ulcerative colitis. *J Clin Med*. 2021;10(8):1749. doi:10.3390/JCM10081749/S1

23. Proctor LM, Creasy HH, Fettweis JM, et al. The Integrative Human Microbiome Project. *Nature 2019 569:7758*. 2019;569(7758):641-648. doi:10.1038/s41586-019-1238-8

24. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The Human Microbiome Project. *Nature 2007 449:7164*. 2007;449(7164):804-810. doi:10.1038/nature06244

25. Qin J, Li R, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature 2010 464:7285*. 2010;464(7285):59-65. doi:10.1038/nature08821

26. Imhann F, van der Velde KJ, Barbieri R, et al. The 1000IBD project: Multi-omics data of 1000 inflammatory bowel disease patients; Data release 1. *BMC Gastroenterol*. 2019;19(1):1-10. doi:10.1186/S12876-018-0917-5/TABLES/2

27. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe*. 2014;15(3):382-392. doi:10.1016/J.CHOM.2014.02.005

28. Lewis JD, Chen EZ, Baldassano RN, et al. Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease. *Cell Host Microbe*. 2015;18(4):489-500. doi:10.1016/J.CHOM.2015.09.008

29. Chassaing B, Aitken JD, Malleshappa M, Vijay-Kumar M. Dextran Sulfate Sodium (DSS)-Induced Colitis in Mice. *Current protocols in immunology / edited by John E Coligan . [et al]*. 2014;104(SUPPL.104):Unit. doi:10.1002/0471142735.IM1525S104

30. Shkoporov AN, Khokhlova E V., Stephens N, et al. Long-term persistence of crAss-like phage crAss001 is associated with phase variation in Bacteroides intestinalis. *BMC Biol*. 2021;19(1). doi:10.1186/S12915-021-01084-3

31. Qv L, Mao S, Li Y, Zhang J, Li L. Roles of Gut Bacteriophages in the Pathogenesis and Treatment of Inflammatory Bowel Disease. *Front Cell Infect Microbiol*. 2021;11. doi:10.3389/FCIMB.2021.755650

32. Adiliaghdam F, Amatullah H, Digumarthi S, et al. Human enteric viruses autonomously shape inflammatory bowel disease phenotype through divergent innate immunomodulation. *Sci Immunol*. 2022;7(70):eabn6660. doi:10.1126/SCIIMMUNOL.ABN6660

33. Nishiyama H, Endo H, Blanc-Mathieu R, Ogata H. Ecological Structuring of Temperate Bacteriophages in the Inflammatory Bowel Disease-Affected Gut. *Microorganisms*. 2020;8(11):1-15. doi:10.3390/MICROORGANISMS8111663

34. Norman JM, Handley SA, Baldridge MT, et al. Disease-specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. *Cell*. 2015;160(3):447. doi:10.1016/J.CELL.2015.01.002

35. Chatzidaki-Livanis M, Coyne MJ, Comstock LE. A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of Bacteroides fragilis. *J Bacteriol*. 2009;191(23):7288-7295. doi:10.1128/JB.00500-09

36. Chatzidaki-Livanis M, Weinacht KG, Comstock LE. Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont Bacteroides fragilis. *Proc Natl Acad Sci U S A*. 2010;107(26):11976-11980. doi:10.1073/PNAS.1005039107

37. Noyer-Weidner M, Trautner TA. Methylation of DNA in prokaryotes. *EXS*. 1993;64:39-108. doi:10.1007/978-3-0348-9118-9_4/COVER

38. Bickle TA, Krüger DH. Biology of DNA restriction. *Microbiol Rev*. 1993;57(2):434-450. doi:10.1128/MR.57.2.434-450.1993

39. Murray NE, Daniel AS, Cowan GM, Sharp PM. Conservation of motifs within the unusually variable polypeptide sequences of type I restriction and modification enzymes. *Mol Microbiol*. 1993;9(1):133-143. doi:10.1111/J.1365-2958.1993.TB01675.X

40. Sistla S, Rao DN. S-Adenosyl-L-methionine-dependent restriction enzymes. *Crit Rev Biochem Mol Biol*. 2004;39(1):1-19. doi:10.1080/10409230490440532

41. Round JL, Mazmanian SK. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc Natl Acad Sci U S A*. 2010;107(27):12204-12209. doi:10.1073/PNAS.0909122107

42. Erturk-Hasdemir D, Oh SF, Okan NA, et al. Symbionts exploit complex signaling to educate the immune system. *Proc Natl Acad Sci U S A*. 2019;116(52):26157-26166. doi:10.1073/PNAS.1915978116

43. Johnson JL, Jones MB, Cobb BA. Polysaccharide-experienced effector T cells induce IL-10 in FoxP3+ regulatory T cells to prevent pulmonary inflammation. *Glycobiology*. 2018;28(1):50-58. doi:10.1093/GLYCOB/CWX093

44. Telesford KM, Yan W, Ochoa-Reparaz J, et al. A commensal symbiotic factor derived from Bacteroides fragilis promotes human CD39(+)Foxp3(+) T cells and Treg function. *Gut Microbes*. 2015;6(4):234-242. doi:10.1080/19490976.2015.1056973

45. Pedersen TK, Brown EM, Plichta DR, et al. The CD4+ T cell response to a commensal-derived epitope transitions from a tolerant to an inflammatory state in Crohn's disease. *Immunity*. 2022;55(10):1909-1923.e6. doi:10.1016/J.IMMUNI.2022.08.016

46. Blandford LE, Johnston EL, Sanderson JD, Wade WG, Lax AJ. Promoter orientation of the immunomodulatory *Bacteroides fragilis* capsular polysaccharide A (PSA) is off in individuals with inflammatory bowel disease (IBD). *Gut Microbes*. Published online February 7, 2019:1-9. doi:10.1080/19490976.2018.1560755

47. Lloyd-Price J, Arze, C Ananthakrishnan, AN Schirmer M, Avila-Pacheco J, Poon T, Andrews, E Ajami N, Bonham, KS Brislawn C. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*. 2019;569(7758):655–662. Accessed March 18, 2021. https://www.nature.com/articles/s41586-019-1237-9

48. Federici S, Kredo-Russo S, Valdés-Mas R, et al. Targeted suppression of human IBD-associated gut microbiota commensals by phage consortia for treatment of intestinal inflammation. *Cell*. 2022;185(16):2879-2898.e24. doi:10.1016/J.CELL.2022.07.003

49. Norman JM, Handley SA, Baldridge MT, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell*. 2015;160(3):447-460. doi:10.1016/J.CELL.2015.01.002

50. Duerkop BA, Kleiner M, Paez-Espino D, et al. Murine colitis reveals a disease-associated bacteriophage community. *Nat Microbiol*. 2018;3(9):1023-1031. doi:10.1038/s41564-018-0210-y

51. Wallace KL, Zheng LB, Kanazawa Y, Shih DQ. Immunopathology of inflammatory bowel disease. *World Journal of Gastroenterology : WJG*. 2014;20(1):6. doi:10.3748/WJG.V20.I1.6

52. Nugent SG, Kumar D, Rampton DS, Evans DF. Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. *Gut*. 2001;48(4):571-577. doi:10.1136/GUT.48.4.571

53. Vertzoni M, Koulouri C, Poulou A, Goumas K, Reppas C. Exploring the impact of Crohn's disease on the intragastric environment of fasted adults. *ADMET DMPK*. 2020;8(2):122. doi:10.5599/ADMET.846

54. Barkas F, Liberopoulos E, Kei A, Elisaf M. Electrolyte and acid-base disorders in inflammatory bowel disease. *Annals of Gastroenterology : Quarterly Publication of the Hellenic Society of Gastroenterology*. 2013;26(1):23. Accessed February 11, 2023. /pmc/articles/PMC3959504/

55. Tian T, Wang Z, Zhang J. Pathomechanisms of Oxidative Stress in Inflammatory Bowel Disease and Potential Antioxidant Therapies. *Oxid Med Cell Longev*. 2017;2017. doi:10.1155/2017/4535194

56. Bourgonje AR, Feelisch M, Faber KN, Pasch A, Dijkstra G, van Goor H. Oxidative Stress and Redox-Modulating Therapeutics in Inflammatory Bowel Disease. *Trends Mol Med*. 2020;26(11):1034-1046. doi:10.1016/J.MOLMED.2020.06.006

57. Tropini C, Moss EL, Merrill BD, et al. Transient osmotic perturbation causes long-term alteration to the gut microbiota. *Cell*. 2018;173(7):1742. doi:10.1016/J.CELL.2018.05.008

58. Milman O, Yelin I, Kishony R. Systematic identification of gene-altering programmed inversions across the bacterial domain. *Nucleic Acids Res*. 2023;51(2):553-573. doi:10.1093/NAR/GKAC1166

59. Blanco-Míguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nature Biotechnology 2023*. Published online February 23, 2023:1-12. doi:10.1038/s41587-023-01688-w

60. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357. doi:10.1038/NMETH.1923

61. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res*. 2017;45(W1):W180-W188. doi:10.1093/nar/gkx295

62. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One*. 2013;8(4):e61217. doi:10.1371/JOURNAL.PONE.0061217

63. Simpson GL. *The Vegan Package*.; 2009. Accessed December 29, 2020. http://cran.r-project.org/,

64. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12). doi:10.1186/s13059-014-0550-8

65. Wickham H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*. Published online 2016.

66. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics*. 2022;38(23):5315-5316. doi:10.1093/BIOINFORMATICS/BTAC672

67. Adams MH. Bacteriophages. *Bacteriophages*. Published online 1959.

68. García-Aljaro C, Muniesa M, Jofre J. Isolation of Bacteriophages of the Anaerobic Bacteria Bacteroides. *Methods Mol Biol*. 2018;1693:11-22. doi:10.1007/978-1-4939-7395-8_2

69. Yu G. Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinformatics*. 2020;69(1):e96. doi:10.1002/CPBI.96

70. Eccles DA. Preparing Reads for Stranded Mapping. Published 2022. Accessed February 11, 2023. https://www.protocols.io/view/preparing-reads-for-stranded-mapping-5qpvon2zzl4o/v7

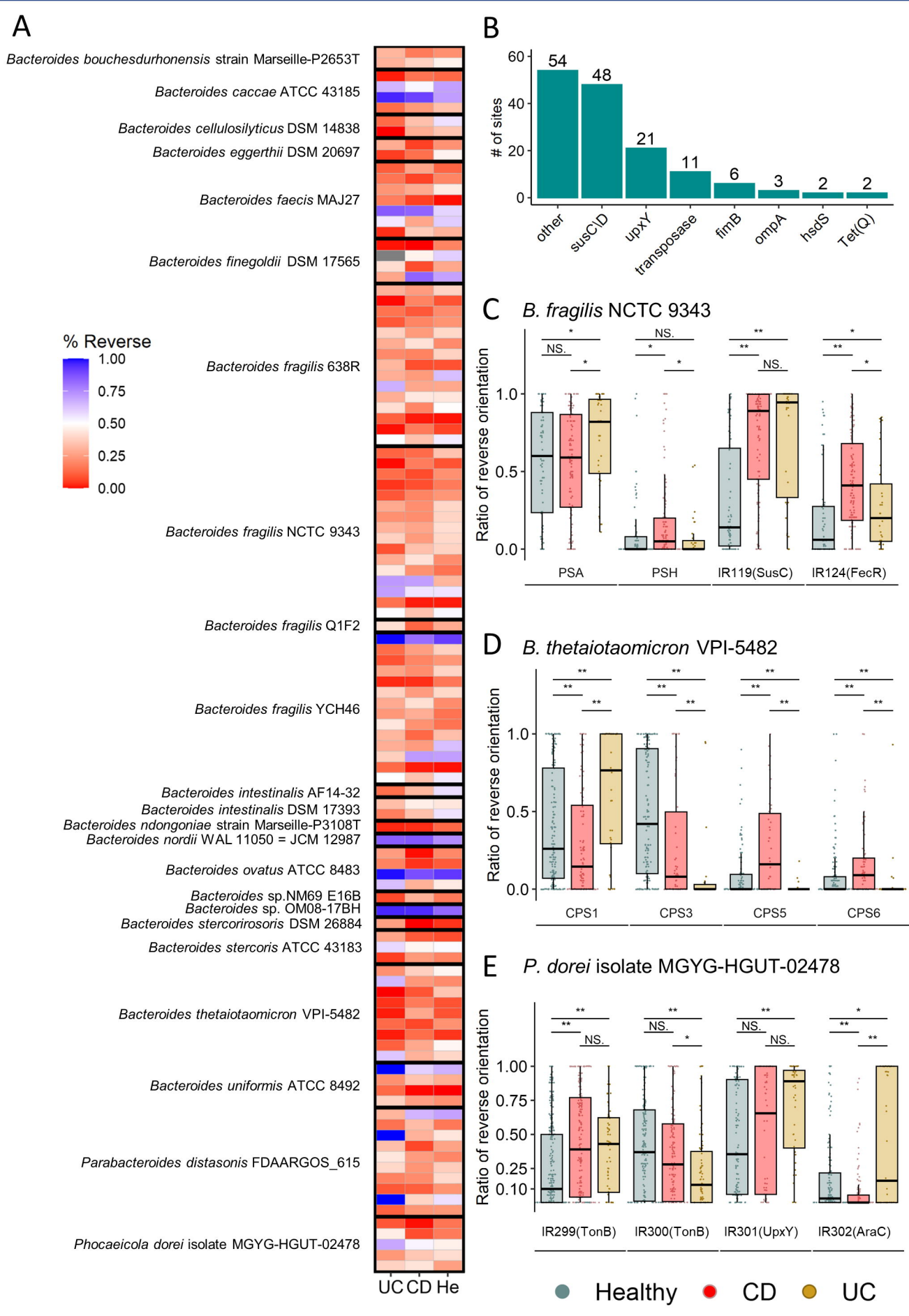
71. Frith MC, Hamada M, Horton P. Parameters for accurate genome alignment. *BMC Bioinformatics*. 2010;11(1):1-14. doi:10.1186/1471-2105-11-80/FIGURES/7

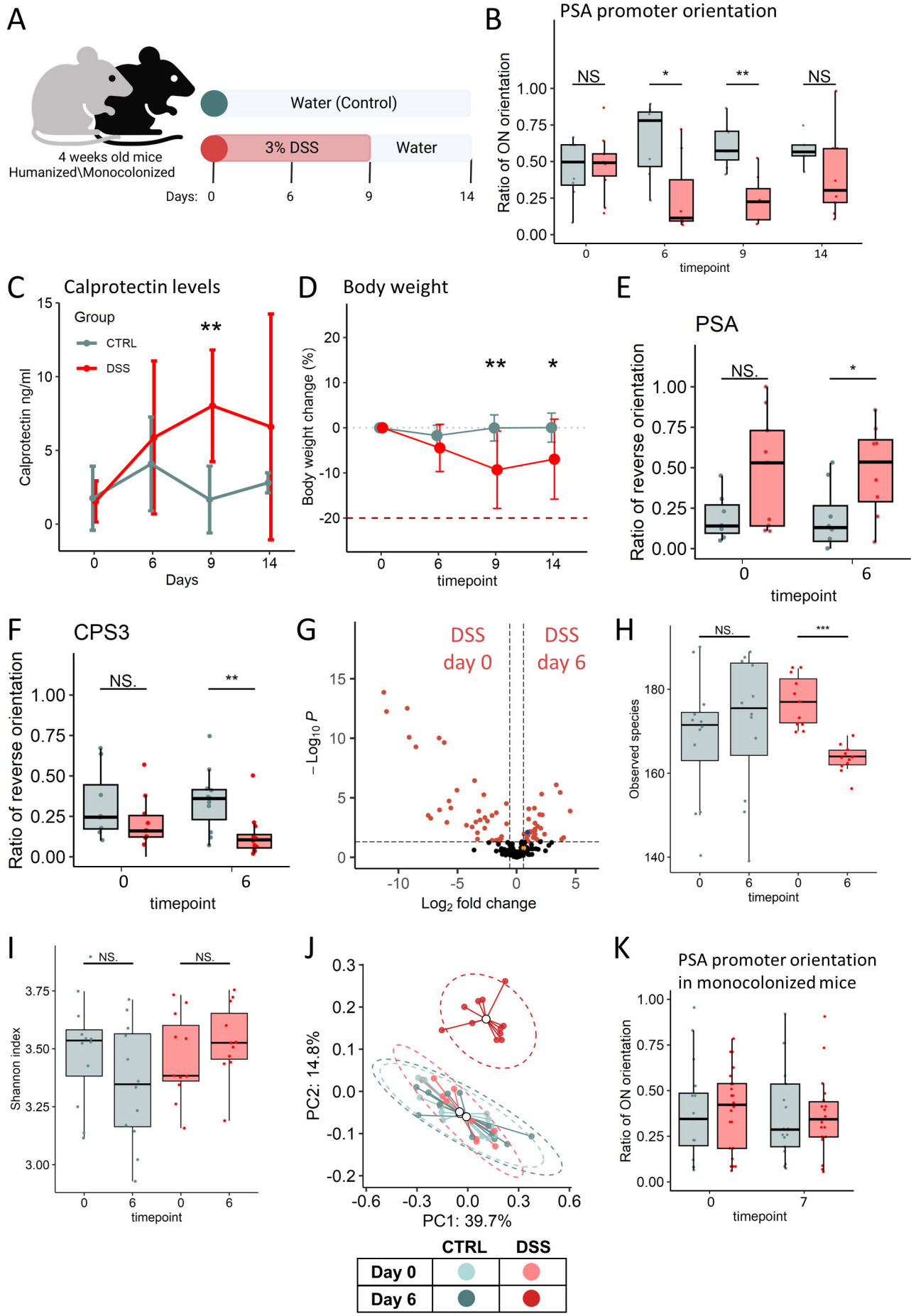
899 72. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
900 *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/BIOINFORMATICS/BTU170

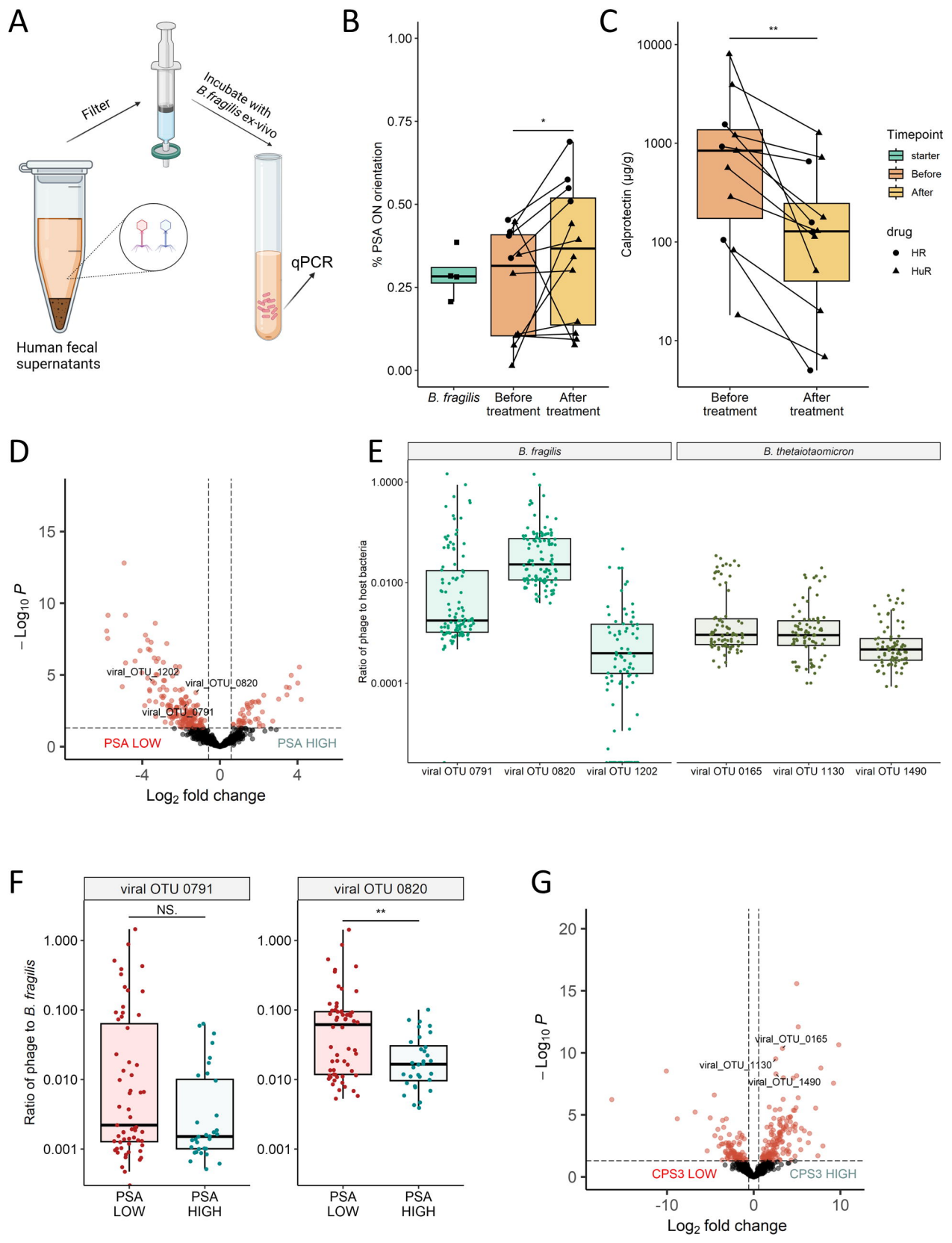
901 73. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*.
902 2018;34(18):3094-3100. doi:10.1093/BIOINFORMATICS/BTY191

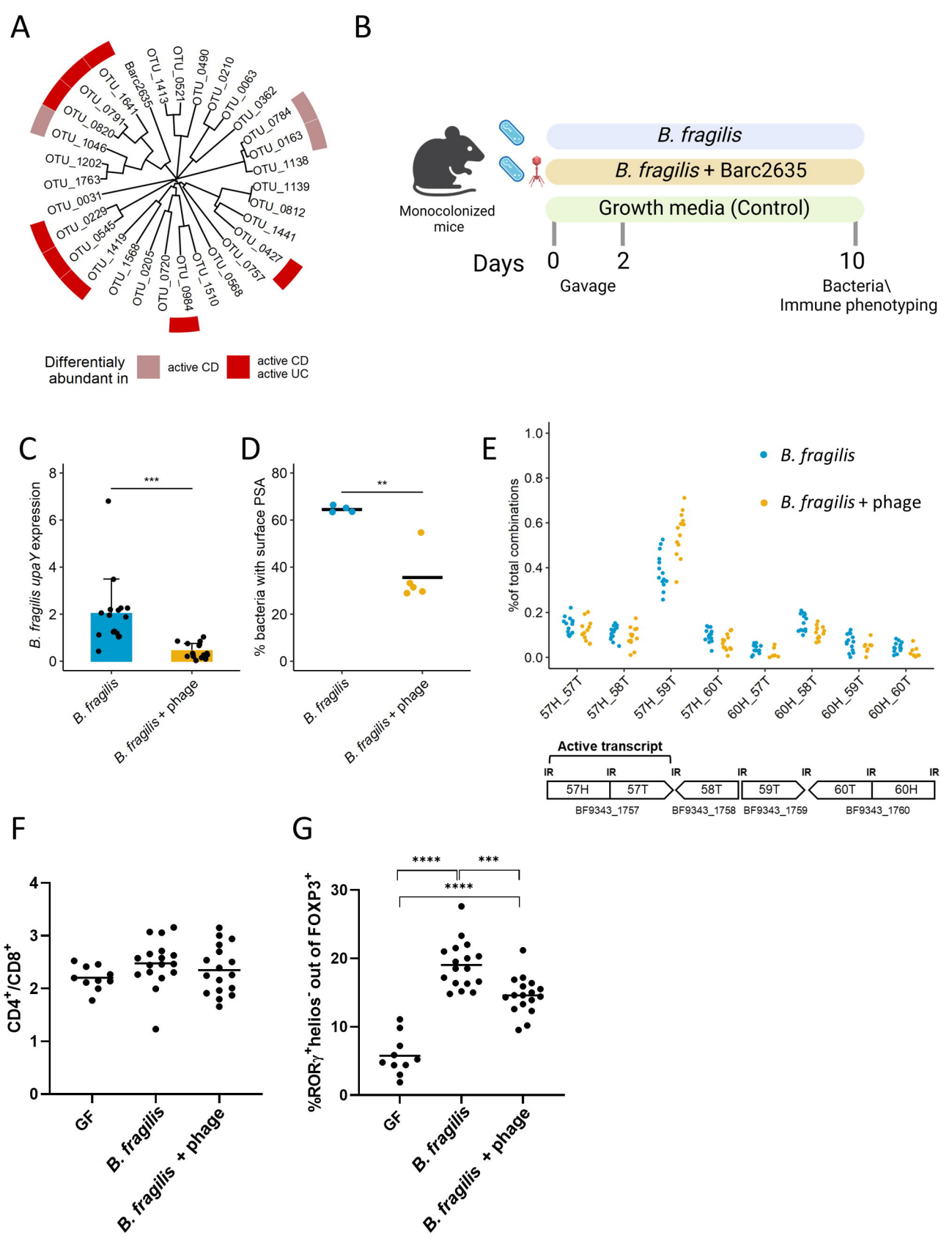
903 74. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools.
904 *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/BIOINFORMATICS/BTP352

905


A
Bacteroides bouchesdurhonensis strain Marseille-P2653T
Bacteroides caccae ATCC 43185
Bacteroides cellulosilyticus DSM 14838
Bacteroides eggerthii DSM 20697
Bacteroides faecis MAJ27
Bacteroides finegoldii DSM 17565
Bacteroides fragilis 638R
Bacteroides fragilis NCTC 9343
Bacteroides fragilis Q1F2
Bacteroides fragilis YCH46
Bacteroides intestinalis AF14-32
Bacteroides intestinalis DSM 17393
Bacteroides ndongoniae strain Marseille-P3108T
Bacteroides nordii WAL 11050 = JCM 12987
Bacteroides ovatus ATCC 8483
Bacteroides sp.NM69 E16B
Bacteroides sp. OM08-17BH
Bacteroides stercorirosoris DSM 26884
Bacteroides stercoris ATCC 43183
Bacteroides thetaiotaomicron VPI-5482
Bacteroides uniformis ATCC 8492
Parabacteroides distasonis FDAARGOS_615
Phocaeicola dorei isolate MGYG-HGUT-02478
% Reverse
1.00
0.75
0.50
0.25
0.00
UC CD He
B
# of sites
54
48
21
11
6
3
2
2
other
susCID
upxY
transposase
fimB
ompA
hsdS
Tet(Q)
C
B. fragilis NCTC 9343
Ratio of reverse orientation
PSA
PSH
IR119(SusC)
IR124(FecR)
D
B. thetaiotaomicron VPI-5482
CPS1
CPS3
CPS5
CPS6
E
P. dorei isolate MGYG-HGUT-02478
IR299(TonB)
IR300(TonB)
IR301(UpxY)
IR302(AraC)
Healthy
CD
UC


A
4 weeks old mice
Humanized\Monocolonized
Water (Control)
3% DSS
Water
Days: 0 6 9 14
B
PSA promoter orientation
Ratio of ON orientation
NS
*
**
NS
timepoint
C
Calprotectin levels
Group
CTRL
DSS
Calprotectin ng/ml
Days
D
Body weight
Body weight change (%)
timepoint
E
PSA
Ratio of reverse orientation
NS.
*
timepoint
F
CPS3
Ratio of reverse orientation
NS.
**
timepoint
G
DSS day 0
DSS day 6
$-\mathrm{Log}_{10} P$
Log_2 fold change
H
Observed species
NS.
***
timepoint
I
Shannon index
NS.
NS.
timepoint
J
PC2: 14.8%
PC1: 39.7%
CTRL
DSS
Day 0
Day 6
K
PSA promoter orientation
in monocolonized mice
Ratio of ON orientation
timepoint


A
Filter
Incubate with
B.fragilis ex-vivo
qPCR
Human fecal
supernatants
B
% PSA ON orientation
1.00
0.75
0.50
0.25
0.00
*
B. fragilis
Before
treatment
After
treatment
C
Calprotectin (µg/g)
10000
1000
100
10
**
Before
treatment
After
treatment
Timepoint
starter
Before
After
drug
HR
HuR
D
– Log10 P
15
10
5
0
viral_OTU_1202
viral_OTU_0820
viral_OTU_0791
PSA LOW
PSA HIGH
-4
0
4
Log2 fold change
E
B. fragilis
B. thetaiotaomicron
Ratio of phage to host bacteria
1.0000
0.0100
0.0001
viral OTU 0791
viral OTU 0820
viral OTU 1202
viral OTU 0165
viral OTU 1130
viral OTU 1490
F
viral OTU 0791
viral OTU 0820
NS.
**
Ratio of phage to B. fragilis
1.000
0.100
0.010
0.001
PSA
LOW
PSA
HIGH
G
– Log10 P
20
15
10
5
0
viral_OTU_0165
viral_OTU_1130
viral_OTU_1490
CPS3 LOW
CPS3 HIGH
-10
0
10
Log2 fold change


A
OTU_0521
OTU_0490
OTU_0210
OTU_0063
OTU_0362
OTU_0784
OTU_0163
OTU_1138
OTU_1139
OTU_0812
OTU_1441
OTU_0427
OTU_0757
OTU_0568
OTU_1510
OTU_0984
OTU_0720
OTU_0205
OTU_1568
OTU_1419
OTU_0545
OTU_0229
OTU_0031
OTU_1763
OTU_1202
OTU_1046
OTU_0820
OTU_0791
OTU_1641
Barc2635
OTU_1413
Differentialy abundant in
active CD
active CD
active UC
B
Monocolonized mice
B. fragilis
B. fragilis + Barc2635
Growth media (Control)
Days
0
2
10
Gavage
Bacteria\ Immune phenotyping
C
B. fragilis upaY expression
***
B. fragilis
B. fragilis + phage
D
% bacteria with surface PSA
**
B. fragilis
B. fragilis + phage
E
%of total combinations
B. fragilis
B. fragilis + phage
57H_57T
57H_58T
57H_59T
57H_60T
60H_57T
60H_58T
60H_59T
60H_60T
Active transcript
IR
57H
57T
58T
59T
60T
60H
BF9343_1757
BF9343_1758
BF9343_1759
BF9343_1760
F
CD4+/CD8+
GF
B. fragilis
B. fragilis + phage
G
****
***
****
%RORγ+helios- out of FOXP3+
GF
B. fragilis
B. fragilis + phage