

# 1 **Inflammation and bacteriophages affect DNA inversion states and functionality** 2 **of the gut microbiota**

3 Authors:

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

7 Affiliations:

8 <sup>1</sup> Department of Cell Biology and Cancer Science, Rappaport Faculty of Medicine, Technion-Israel  
9 Institute of Technology, Rappaport Technion Integrated Cancer Center (RTICC), Haifa, Israel

10 <sup>2</sup> Department of Gastroenterology, Rambam Health Care Campus, Haifa, Israel

11 <sup>3</sup> Clinical Research Institute, Rambam Health Care Campus, Haifa, Israel

12 <sup>4</sup> Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel

13 <sup>5</sup> Moderna, Inc., Cambridge, MA 02139, USA

14 <sup>6</sup> Center for the Study of Inflammatory Bowel Disease, Division of Gastroenterology, Department of  
15 Medicine, Massachusetts General Hospital Research Institute, Boston, MA 02114, USA; Harvard  
16 Medical School, Boston, MA 02115, USA; Program in Immunology, Harvard Medical School, Boston,  
17 MA 02115, USA.

18 <sup>7</sup> Department of Genetics, Microbiology and Statistics, School of Biology, University of Barcelona,  
19 Avda. Diagonal 643 08028 Barcelona Spain

20 <sup>8</sup> Duchossois Family Institute and Department of Microbiology, University of Chicago, Chicago, IL,  
21 USA

22 <sup>9</sup> Migal-Galilee Research Institute, P.O. Box 831, Kiryat Shmona 11016, Israel

23 <sup>10</sup> Faculty of Sciences and Technology, Tel-Hai Academic College, Upper Galilee 1220800, Israel

24 <sup>11</sup> CIFAR, MaRS Centre, West Tower 661 University Ave., Suite 505, Toronto, ON M5G 1M1, Canada

25 \* These authors contributed equally

26 <sup>§</sup> Corresponding author, [naamagvz@gmail.com](mailto:naamagvz@gmail.com)

## 27 Summary

28 Reversible genomic DNA-inversions control expression of numerous bacterial molecules in the  
29 human gut, but how this relates to disease remains uncertain. By analyzing metagenomic samples  
30 from six human Inflammatory Bowel Disease cohorts combined with mice experimentation, we  
31 identified multiple invertible regions where a particular orientation was correlated with disease.  
32 These include the promoter of the anti-inflammatory polysaccharide-A (PSA) of *Bacteroides fragilis*,  
33 which is mostly oriented 'OFF' during inflammation but is present in the 'ON' orientation when  
34 inflammation is resolved. We further detected increased abundances of *B. fragilis*-associated  
35 bacteriophages in patients with the PSA 'OFF' orientation, and a significant reduction in the  
36 frequency of the 'ON' orientation, in the presence of the *B. fragilis*-associated bacteriophage,  
37 thereby altering the bacterial induced immune modulation. Altogether, we reveal dynamic and  
38 reversible bacterial phase-variations driven both by bacteriophages and the host inflammatory state,  
39 signifying bacterial functional plasticity during inflammation and opening future research avenues.

## 40 Keywords

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

43

## 44 Introduction

45 Phase variation is the process by which bacteria undergo reversible alterations in specific loci of their  
46 genome, resulting in ON-OFF expression of genes<sup>1-4</sup>. In Bacteroidales, the dominant order of  
47 bacteria in the human gut, phase variation is highly prevalent and is largely mediated by inversions  
48 of DNA segments between inverted repeats. These inversions often involve promoter regions  
49 dictating transcription initiation of genes or operons functioning as 'ON'\OFF' switches<sup>5</sup>. In addition,  
50 DNA inversions can occur so that new genes are brought from an inactive to a transcriptionally  
51 active site by re-orientation or recombination of genomic "shufflons"— mobile genetic elements  
52 that facilitate rearrangements, serving as dynamic tools for altering the expressed gene<sup>6-8</sup>. Analysis  
53 of the orientations of bacterial invertible regions in various host disease states can provide new  
54 insights into bacterial adaptation and functional contributions to the disease pathogenesis or its  
55 resolution. Phase variation in the gut Bacteroidales often modulates production of components  
56 presented on the bacterial surface<sup>9</sup> dictating which surface molecules interact, for example, with  
57 neighboring microbes (e.g. bacteria, bacteriophages) or with the host.

58 *Bacteroides fragilis*, a common resident of the human gut, modulates its surface by the phase  
59 variable synthesis of its capsular polysaccharides (PS, denoted PSA-PSH). The biosynthesis loci of  
60 seven of its eight polysaccharides have invertible promoters that are orientated either 'ON' or 'OFF'  
61 in respect to the downstream PS biosynthesis operon<sup>5</sup>. Studies have shown that the *B. fragilis*  
62 polysaccharide A (PSA) modulates the host immune system by inducing regulatory T cells (Tregs) and  
63 secretion of the anti-inflammatory cytokine interleukin (IL)-10<sup>10,11</sup>. Moreover, PSA was shown to  
64 confer protection against experimental colitis<sup>10,12-14</sup>, and thus is regarded as an anti-inflammatory  
65 polysaccharide.

66 Bacteriophages, viruses that specifically target and infect bacteria, can influence the composition of  
67 bacterial populations and potentially affect their functionality. Several studies have highlighted the  
68 dynamic relationship between Bacteroidales and phages. Campbell et al.<sup>15</sup> found that infection of  
69 *Bacteroides vulgatus* with temperate bacteriophage BV01 results in a repression of bile salt  
70 hydrolase activity - changing the bacteria transcriptional profile depending on the integration of the  
71 phages into their genome. Porter et al.<sup>8</sup> and Hryckowian et al.<sup>16</sup> demonstrated that some capsular PS  
72 and other outer surface molecules can affect the susceptibility of *Bacteroides thetaiotaomicron*, to  
73 specific phages. Consistent with that, Shkoporov et al.<sup>17</sup> revealed that phase variation of capsular PS  
74 in *Bacteroides intestinalis* influences the infection of CrAss-like phages, establishing a dynamic  
75 equilibrium for their prolonged coexistence in the gut. These studies suggest that the interactions  
76 between phages and *Bacteroides* may affect the host through changes in bacterial functionality.

77 Ulcerative colitis (UC) and Crohn's disease (CD) are multifactorial inflammatory bowel diseases (IBD),  
78 characterized by a compromised mucosal barrier, inappropriate immune activation, and  
79 mislocalization of the gut microbiota<sup>18-21</sup>. As such, IBDs have emerged as some of the most studied  
80 microbiota-linked diseases<sup>22</sup> and an interesting setting for studying gut bacterial functions, their  
81 potential effects on the host, and bacteria-phage interactions.

82 Here we present an analysis of invertible DNA orientations in the gut microbiota of IBD patients from  
83 multiple cohorts and test the hypothesis that inflammation and bacteriophages mediate these  
84 differential orientations. Our combined analysis of IBD patient metagenomes and experimental  
85 mouse models, focusing on Bacteroidales species, reveals alterations in the orientations of multiple  
86 DNA invertible regions during gut inflammation with potential to modulate the host immune system.  
87 Importantly, we show that filtered fecal extracts of IBD patients alter the orientation of the PSA  
88 promoter of *B. fragilis*. We show that bacteriophages are correlated with the PSA 'OFF' state, and  
89 that a specific lytic bacteriophage of *B. fragilis* increases the PSA promoter 'OFF' population with

90 concurrent decline of host colonic Treg cells. These findings reveal a dynamic interplay between gut  
91 inflammation, bacteriophages, and bacterial phase variation, with potential implications for the  
92 diagnosis and treatment of IBD.

## 93 Results

### 94 IBD is correlated with quantitative differences in the orientations of multiple Bacteroidales 95 invertible DNA regions compared to healthy controls

96 As a means to determine if Bacteroidales phase variable molecules may be differentially produced in  
97 IBD patients compared to controls, we used the PhaseFinder<sup>9</sup> software to identify and analyze the  
98 relative orientations of DNA inversions in the gut metagenomes of cohorts of IBD patients and  
99 control subjects [Table 1]. In brief, we first detected putative invertible regions in Bacteroidales  
100 genomes by identifying inverted repeats, and then created a database containing their forward and  
101 reverse orientations (compared to the submitted genome sequences). Metagenomic sequences  
102 from publicly available datasets were then aligned to the database, resulting in the ratio of the  
103 orientations of each invertible region compared to their orientations in the published genome  
104 sequence. Our initial analysis of 39 sequenced genomes including 36 human gut Bacteroidales  
105 species identified 311 invertible regions. Of these 311 invertible regions, there were 147 that were  
106 statistically different in terms of their orientations in IBD patients compared to healthy controls.  
107 These included invertible regions from 25 of the reference bacterial genomes [Table S1]. These  
108 regions included both invertible promoters and intragenic regions of diverse genes, spanning from  
109 regulatory genes to those encoding genes for the production of outer surface molecules. Table  
110 S1 details the locations of the inverted repeats (IR) sequences and genes flanking the invertible  
111 regions [Figure 1A]. The regions in which the orientations differed most significantly between IBD  
112 patients and controls were within or in proximity to SusC/SusD-like outer membrane transport  
113 systems<sup>23</sup> and to *upxY* genes, which, in most cases, represent the first gene in the biosynthesis loci of  
114 capsular polysaccharides (PS) [Figure 1B]. Four of the five invertible capsular polysaccharide  
115 promoters of *Bacteroides thetaiotaomicron* were differentially oriented between the groups, and  
116 three of the seven invertible PS promoters of *B. fragilis* were differentially oriented, including the  
117 anti-inflammatory PSA promoter being differentially oriented between healthy and UC patients. The  
118 PSA promoter showed a higher percentage (71%) of reverse oriented reads in IBD patients,  
119 compared to 56.2% in the healthy controls [Figure 1C]. In the reference genome of *B. fragilis* NCTC  
120 9343, the PSA promoter is in its 'ON' orientation, hence, the reverse orientation, found in IBD  
121 patients, represents the PSA promoter's 'OFF' orientation. We further identified differential

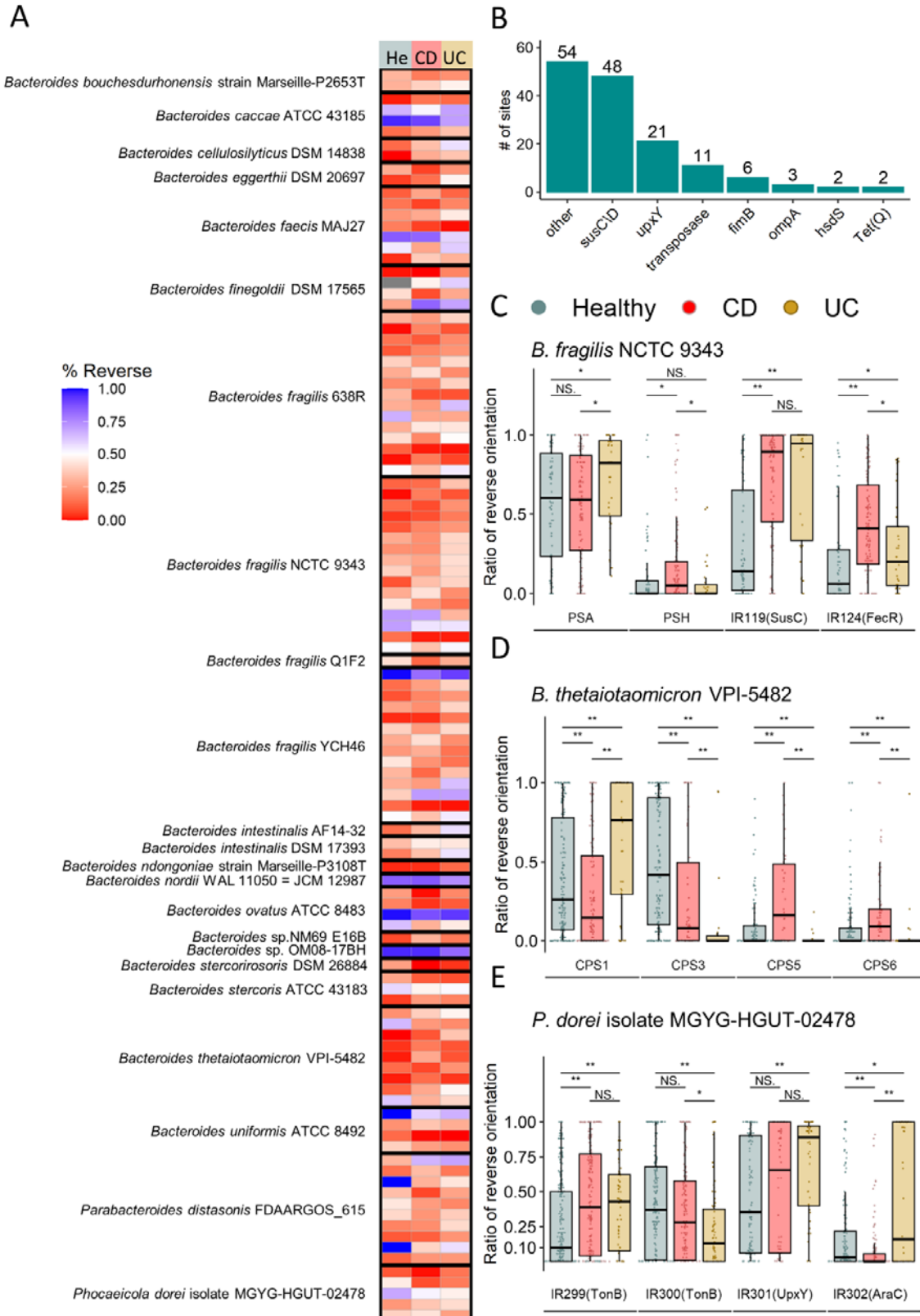
122 promoter orientations in healthy and IBD cohorts for PULs and polysaccharides promoters, such as  
 123 *Bacteroides thetaiotaomicron* VPI-5482 [Figure 1D], and *Phocaeicola dorei*, a bacteria shown to be  
 124 present in healthy subjects<sup>24,25</sup> and correlated with disease activity in UC<sup>26</sup> [Figure 1E].

125

126 **Table 1: Metagenomic datasets included in this study.**

Dataset	N (Samples)	CD	UC	Non-IBD	Country	Notes	
IBDMDB	1283	574	349	360	United States	Longitudinal	Ref <sup>27</sup>
HMP (3 phases)	320			320	United States		Ref <sup>28</sup>
MetaHit	122	4	21	97	Denmark and Spain		Ref <sup>29</sup>
1000IBD	331	205	126		the Netherlands		Ref <sup>30</sup>
GeversD_2014	50	36		14	United States and Canada		Ref <sup>31</sup>
LewisJD_2015	303	303			United States and Canada	Children with Crohn's Disease Longitudinal, treated	Ref <sup>32</sup>
Total	2409	1122	496	791			

127



128

129

130

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

## 139 **Differential promoter orientation states are induced by inflammation in a dynamic and reversible** 140 **manner**

141 To assess the dynamics of the *B. fragilis* PSA promoter orientation under inflammatory conditions,  
142 we designed a longitudinal experimental mouse model during which we analyzed the PSA promoter  
143 orientation over time. Germ-free (GF) mice were colonized with the gut microbiota from a healthy  
144 human microbiota, and subsequently spiked with *B. fragilis* NCTC 9343, termed "humanized" mice.  
145 Progeny of these mice were used for the experiments ( $n=8-12$  mice per experimental group, DSS vs.  
146 control). Experimental colitis was induced by adding 3% dextran sodium sulfate (DSS) to the drinking  
147 water<sup>33</sup> for 9 days after which DSS was replaced with water until day 14 **[Figure 2A]**. Inflammation  
148 was monitored by the levels of calprotectin in the stool **[Figure 2B]**, a commonly used biomarker for  
149 inflammation, and by mouse weight loss **[Figure 2C]**. Stool was collected before, during and after  
150 DSS treatment to analyze orientations of specific invertible DNA regions, using quantitative PCR  
151 (qPCR) to observe the state of the PSA promoter **[Figure S1A-C]**. The orientation of the *B. fragilis* PSA  
152 promoter is variable to some extent in both the control and DSS-treated mice and was significantly  
153 altered on day 9 in the DSS-treated mice along with the inflammatory state **[Figure 2B-D]**. At the  
154 beginning of the experiment, ~45% of the population had the promoter in the 'ON' orientation  
155 **[Figure 2D]**. A significant decline in these percentages was observed six days after DSS was  
156 introduced (down to an average of ~26%) and remained at around ~24% on average on day nine. By  
157 day 14, when the mice were starting to gain weight and the calprotectin levels were decreasing  
158 **[Figure 2B,2C]**, the bacterial population returned to an average of 40% 'ON', similar to the beginning  
159 of the experiment, and to the control group, which did not receive DSS (mean of 58% of the *B.*  
160 *fragilis* population 'ON') **[Figure 2D]**.

161 Since the frequency of *B. fragilis* with the PSA promoter oriented 'ON' declined in most of the  
162 inflamed mice **[Figure S1D]**, we performed metagenomic sequencing from stool collected on days  
163 zero and six for analysis of orientations of invertible regions in other Bacteroidales species and  
164 genomic regions **[Table S2]**.

165 This analysis validated the decline in the frequency of the ‘ON’ orientation of the *B. fragilis* PSA  
166 promoter [Figure 2E], aligned with the qPCR results [Figure S1E], and agreeing with the human  
167 metagenomic analysis [Figure 1C]. In addition, we identified a decline in the frequency of the ‘ON’  
168 orientation of the CPS3 promoter of *B. thetaiotaomicron* during inflammation in mice [Figure 2F],  
169 consistent with the results of the human metagenomic analysis where it also showed lower ratios of  
170 the ‘ON’ orientation, equivalent to reverse oriented reads in comparison to the reference genome  
171 [Figure 1D].

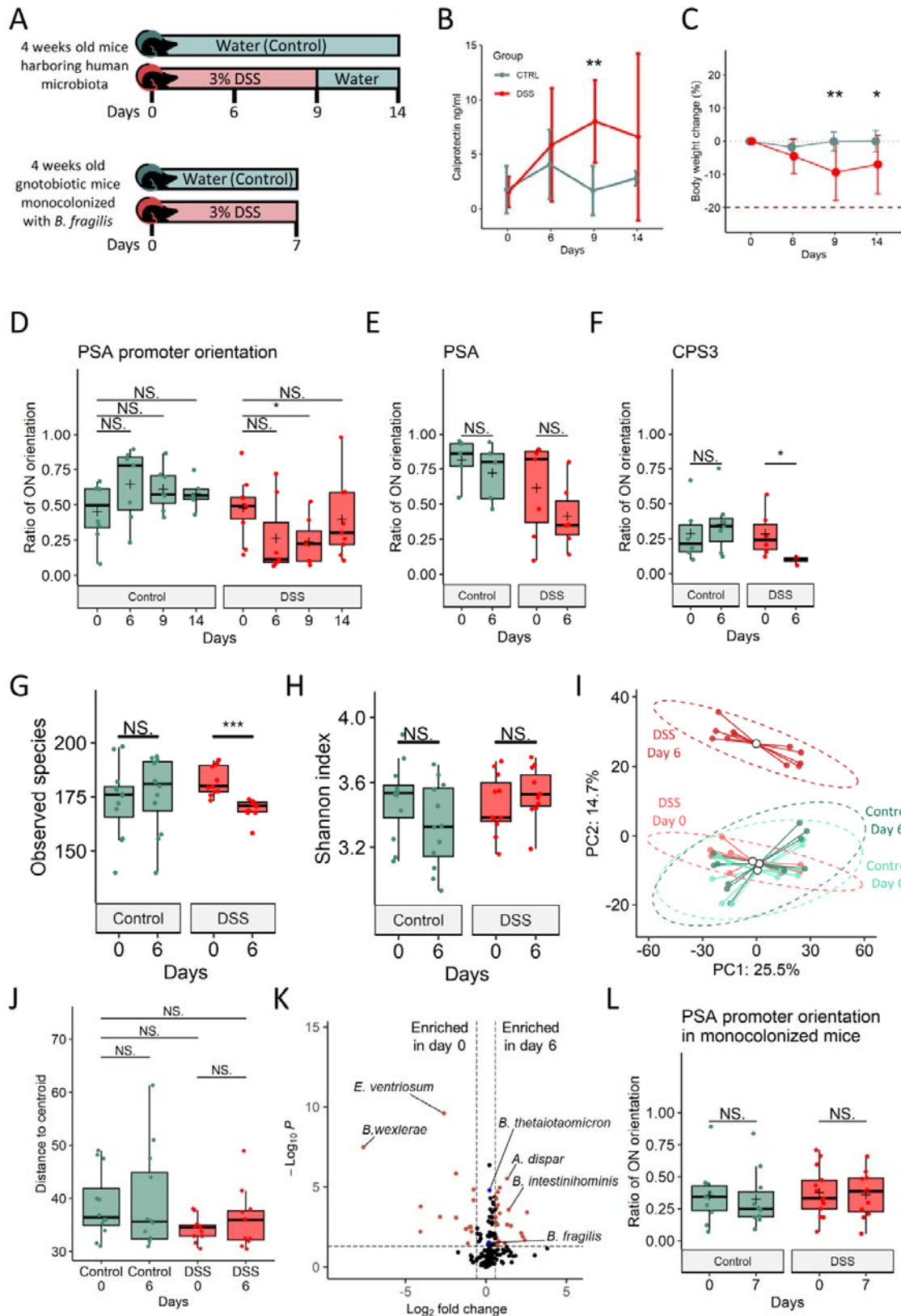
172 The bacterial composition of the gut microbiota in control mice was stable throughout the  
173 experiment but varied in the DSS-treated mice. Bacterial richness (alpha-diversity, observed species)  
174 declined during inflammatory conditions [Figure 2G], while evenness (alpha-diversity, Shannon  
175 index) remained constant [Figure 2H]. The beta-diversity (Aitchison distance) was significantly  
176 altered [Figure 2I], while the beta-dispersion (Aitchison distance) was not significantly altered  
177 [Figure 2J].

178 Inflamed mice showed a decrease in relative abundances of *Blautia* species and *Eubacterium*  
179 *ventriosum*, with a concomitant increase in *Barnesiella intestinihominis* and *Alistipes dispar* [Figure  
180 2K, Table S2]. The relative abundance of *B. fragilis*, which exhibited differential orientations of  
181 invertible DNA regions, was not significantly altered during inflammation. [Figure 2K].

182 To better assess the role of inflammation on the orientation of the *B. fragilis* PSA promoter we  
183 repeated the experiment using gnotobiotic mice mono-colonized with *B. fragilis* NCTC 9343 (n=8-12  
184 mice per experimental group, DSS vs. control), excluding influences of altered microbiota [Figure 2A,  
185 S1F]. The PSA promoter orientation was not altered in the monocolonized mice [Figure 2L], along  
186 with the bacterial abundances [FIGURE S1G] over the course of the experiment (with or without  
187 DSS), suggesting a role of the microbiota in driving the ‘OFF’ orientation of the PSA promoter during  
188 inflammation.



189

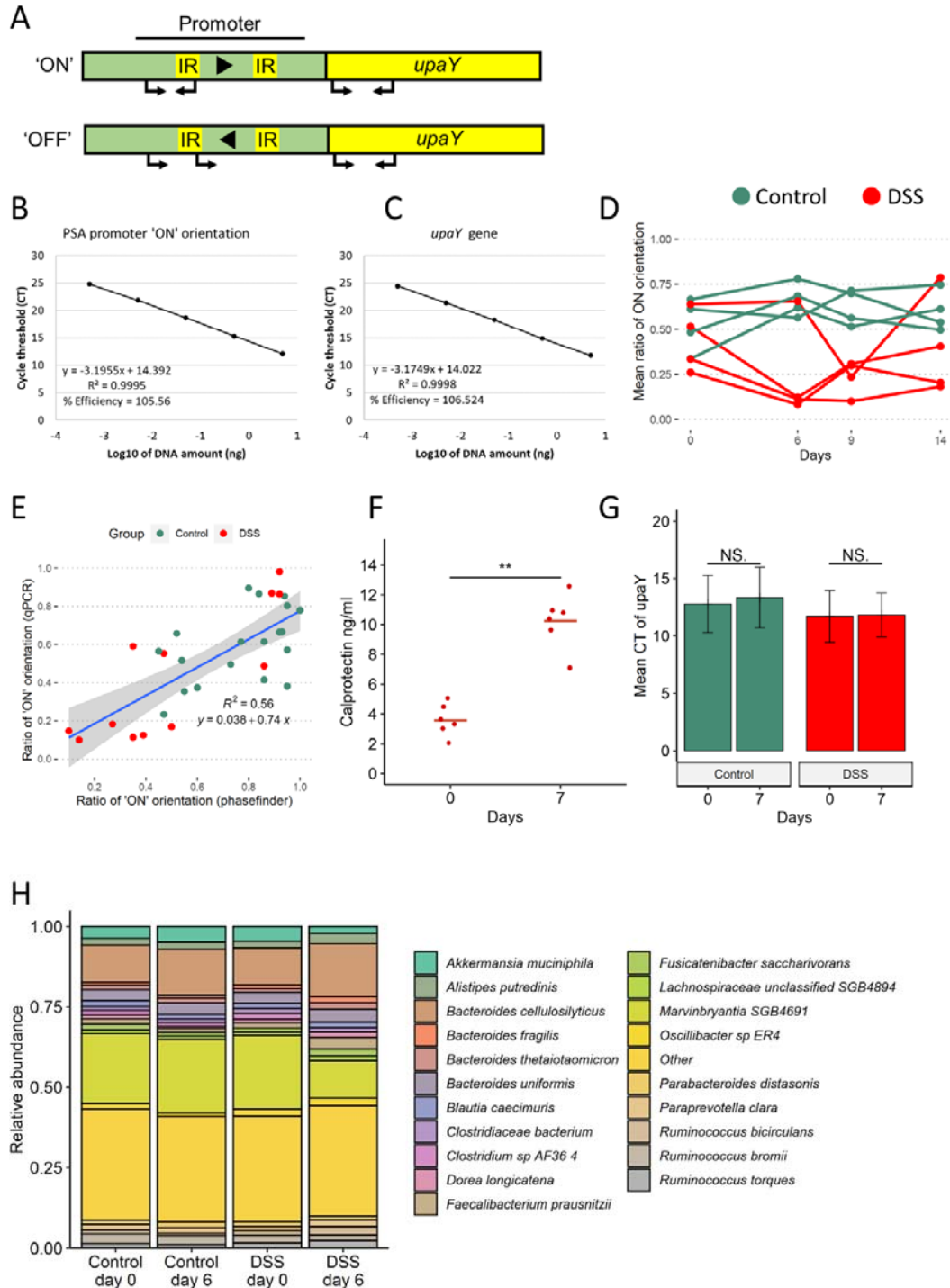


190

191 **Figure 2: Relative orientation of the PSA promoter of *B. fragilis* is affected by inflammation.**  
192 **A.** Illustration of murine model of inflammation. "Humanized" mice harboring human microbiota spiked with *B. fragilis* NCTC 9343, and  
193 monocolonized with *B. fragilis* NCTC 9343 were exposed to 3% DSS (Day=0 in the illustration) at four weeks of age; **B.** Calprotectin levels  
194 (ng/ml) measured in different days of the experiment. Lines represent the standard deviations. Green: Control group, Red: DSS treated  
195 mice. **C.** The body weight change of mice measured in different days of the experiment. Lines represent the standard deviations. Green:  
196 Control group, Red: DSS treated mice. **D.** Ratio of *B. fragilis* PSA's promoter 'ON' orientation measured by qPCR on different days of the  
197 experiment (n=8 in each time point). Data represents the mean (cross), median (line in box), IQR (box), and minimum/maximum  
198 (whiskers). (Wilcoxon rank sum test, \* $p < 0.05$ ; \*\* $p < 0.01$ ). Green: Control group, Red: DSS treated mice. **E.** Ratio of *B. fragilis* PSA's  
199 promoter reverse orientation measured by the PhaseFinder tool, in different days of the experiment. Data represent the median (line in  
200 box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test, \* $p < 0.05$ ; \*\* $p < 0.01$ ). Green: Control group, Red: DSS  
201 treated mice. **F.** Ratio of *B. thetaiotaomicron* CPS3's promoter reverse orientation measured by the PhaseFinder tool on different days of  
202 the experiment. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test,  
203 \* $p < 0.05$ ; \*\* $p < 0.01$ ). Green: Control group, Red: DSS treated mice. **G.** Alpha diversity (observed species) between groups and timepoints.  
204 Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test, \* $p < 0.05$ ; \*\* $p < 0.01$ ,  
205 \*\*\* $p < 0.001$ ). Green: Control group, Red: DSS treated mice. **H.** Alpha diversity (Shannon index) between groups and timepoints. Data  
206 represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (Wilcoxon rank sum test, \* $p < 0.05$ ; \*\* $p < 0.01$ ,  
207 \*\*\* $p < 0.001$ ). Green: Control group, Red: DSS treated mice. **I.** Beta diversity (Aitchison distance) between groups and timepoints. Principal  
208 Coordinates Analysis (PCoA) of Aitchison distances between bacterial communities of different groups and timepoints. Each point  
209 represents a single sample, colored according to group and timepoints: light Green: Control on day 0, Green: control on day 6, light red:  
210 DSS treated on day 0, red: DSS treated on day 6. The mean (centroid) of samples in each group is indicated with a blank circle. Ellipses  
211 represent 0.95 confidence intervals of each group. **J.** Beta-dispersion values (Aitchison distances from the centroid) between groups and  
212 timepoints. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers). (betadisper,  $p > 0.05$ ). **K.** Differential  
213 bacterial abundance between DSS treated mice in day 0 and day 6 detected by the Maaslin2 algorithm. Red dots indicate differentially  
214 abundant bacteria that were determined by adjusted  $P$  value  $< 0.05$  and  $\log_2$  fold change  $> 1$  and  $< -1$ , respectively. Blue dots indicate *B.*  
215 *fragilis* and *B. thetaiotaomicron*. **L.** Ratio of *B. fragilis* PSA's promoter 'ON' orientation measured by qPCR in different days in gnotobiotic  
216 mice monocolonized with *B. fragilis* NCTC 9343. Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers).  
217 (Wilcoxon rank sum test, \* $p < 0.05$ ; \*\* $p < 0.01$ ). Green: Control group, Red: DSS treated mice.

218

219



220

221 **FigureS1: Relative orientation of the PSA promoter of *B. fragilis* is affected by inflammation.**

222 A. Schematics of the qPCR assay used to assess the ratio of PSA 'ON' promoter orientation. Primers (arrows) were directed upstream and  
 223 within the PSA promoter invertible region (IR: inverted repeats) targeting only the 'ON' oriented promoters. Another set of primers was  
 224 directed at the *upaY* and used to normalize the results to the number of genomes in a sample. **B. & C.** Primer efficiencies for the two sets  
 225 of primers. Standard curves were calculated for each primer set from the Cycle thresholds (CTs) and the log of the initial DNA  
 226 concentrations. Each sample represents the mean CT of 3 technical repeats. **D.** Mean ratio of *B. fragilis* PSA's promoter 'ON' orientation

227 measured by qPCR for different experiments on different days (n=8-12 in each time point). Green: Control group, Red: DSS treated mice. E.  
228 Scatter plot of the ratio of *B. fragilis* PSA's promoter 'ON' orientation measured by qPCR and by the PhaseFinder tool. Each point  
229 represents a single sample, colored according to group and timepoints: green: Control group, red: DSS treated mice. Blue line represents  
230 linear regression (gray area represents 95% confidence intervals). F. Calprotectin levels (ng/ml) measured in gnotobiotic mice colonized  
231 with *B. fragilis* and treated with DSS on different days of the experiment. Line represents the mean. (Wilcoxon rank sum test,  $**p < 0.01$ )  
232 G. Mean qPCR cycle threshold (CTs) of *upaY* gene DNA, representing the abundance of *B. fragilis* in gnotobiotic mice monocolonized with  
233 *B. fragilis* of different groups and timepoints. Bars represent the means of the CTs in each group and timepoint. Error bars represent the  
234 standard error. Green: Control group, Red: DSS treated mice. (Wilcoxon rank sum test,  $p > 0.05$ ). H. Stacked bar plots representing the  
235 relative abundance of the 20 most abundant bacterial species in the humanized mice in each experimental group in days 0 and 6.

## 236 **The inflammatory milieu of a complex human microbiota results in the PSA promoter "OFF"** 237 **orientation**

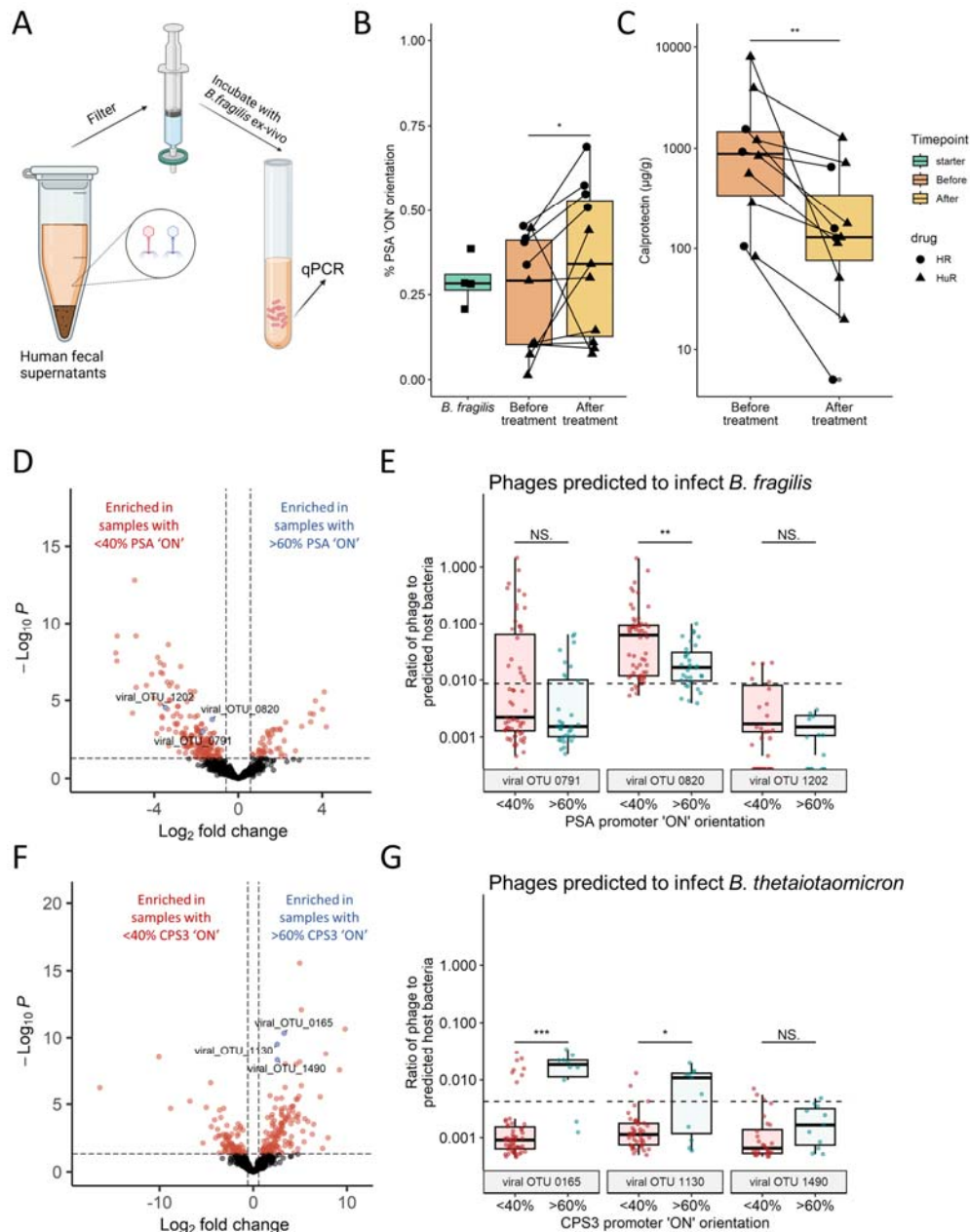
238 We next sought to examine factors of the complex microbiota that result in a decreased percentage  
239 of the *B. fragilis* population with the PSA promoter orientated 'ON' during inflammation. To do so,  
240 we exposed *B. fragilis* NCTC 9343 to fecal filtrates from IBD patients. CD and UC patients were  
241 recruited from the Rambam Health Care Campus (RHCC). Fecal samples were collected before and  
242 after treatment (Infliximab (HR) or Humira (HuR) therapy, 4 and 7 patients respectively). Both  
243 treatments are antibodies targeted against tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), an inflammatory  
244 cytokine increased in IBD patients. *B. fragilis* NCTC 9343 was cultured in fecal filtrates to mid-log  
245 phase, at which time, DNA was extracted for qPCR analysis of the PSA promoter orientation [Figure  
246 3A]. *B. fragilis* exposed to fecal filtrates of patients before treatment showed higher ratios of the  
247 population with the PSA promoter oriented 'OFF', while *B. fragilis* exposed to fecal filtrates after  
248 treatment showed higher ratios of the PSA promoter oriented 'ON' [Figure 3B]. This observation was  
249 in line with the fecal calprotectin concentrations which were high before treatment and decreased  
250 after treatment, inversely correlating with PSA promoter 'ON' orientation [Figure 3C]. These results  
251 demonstrate that the population of bacteria with the PSA promoter in each orientation varies in the  
252 inflamed and non-inflamed gut, with a higher percentage of the population in the 'OFF' orientation  
253 under inflammatory conditions, and a shift towards the 'ON' orientation following reduction in  
254 inflammation [Figure 3B, 3C]. Particularly, treatment responders had a significant correlation of PSA  
255 'ON' with lower calprotectin levels, compared to the non-responders [Figure S2A]. This change could  
256 be due to induction of inversion under these conditions or to selection of a preferred promoter  
257 state.

## 258 **Investigating associations of phage with PSA promoter orientation.**

259 Fecal filtrates contain a mixture of bacterial and host metabolites, cytokines, antibodies, viruses,  
260 bacteriophages, and other components. Bacteriophages were shown to be associated with intestinal

261 inflammation and IBD in several studies<sup>34–38</sup>, including Nishiyama *et al.*<sup>39</sup> who characterized  
262 temperate bacteriophages and their bacterial hosts in the IBDMDB metagenomics database. Here,  
263 using phage sequences from the IBDMDB database, we compared the relative abundances of these  
264 phages between samples that displayed lower (<40%) or higher (>60%) ratios of the PSA promoter  
265 'ON' orientation. Differential relative abundance analysis revealed 108 viral OTUs enriched in  
266 patients with lower ratios of the PSA promoter 'ON' orientation after FDR correction, compared to  
267 24 viral OTUs enriched in patients with the higher ratios of the 'ON' orientation] **Figure 3D, Table**  
268 **S3**]. Three of the bacteriophages enriched in the lower ratios of the PSA promoter 'ON' orientation  
269 was predicted to infect *B. fragilis*<sup>39</sup> - viral OTUs 0791, 0820, and 1202. To assess whether the relative  
270 abundances of the bacteriophages were accompanied with lower relative abundances of *B. fragilis*,  
271 the phage-to-host abundance ratio was calculated for each phage-*B. fragilis* pair. Two of the OTUs  
272 whose relative abundances were correlated to the PSA promoter 'OFF' orientation, viral OTUs 0791  
273 and 0820, also showed higher phage-to-host abundance ratios [**Figure 3E, Figure S2**]. Intriguingly,  
274 these two viral OTUs were found to be more abundant in active CD and UC<sup>39</sup>.

275 We conducted the same analysis comparing samples with high to low 'ON' orientation ratios of the  
276 CPS3 promoter of *B. thetaiotaomicron*. Although there were differences in the virome compositions  
277 between these groups, no *B. thetaiotaomicron* associated bacteriophages were correlated with the  
278 'OFF' orientation of the CPS3 promoter (i.e., the orientation, which we identified as associated with  
279 the disease) [**Figure 3F, Table S3**]. We found three *B. thetaiotaomicron* associated with viral OTUs -  
280 0165, 1130, and 1490, slightly associated with the 'ON' orientation of the CPS3 promoter (the  
281 orientation that we identified as associated with healthy controls) [**Figure 3G, Figure S2**]. In addition,  
282 these bacteriophages displayed low phage to host ratios [**Figure 3G**]. To note, these OTUs were not  
283 previously associated with disease<sup>39</sup>.

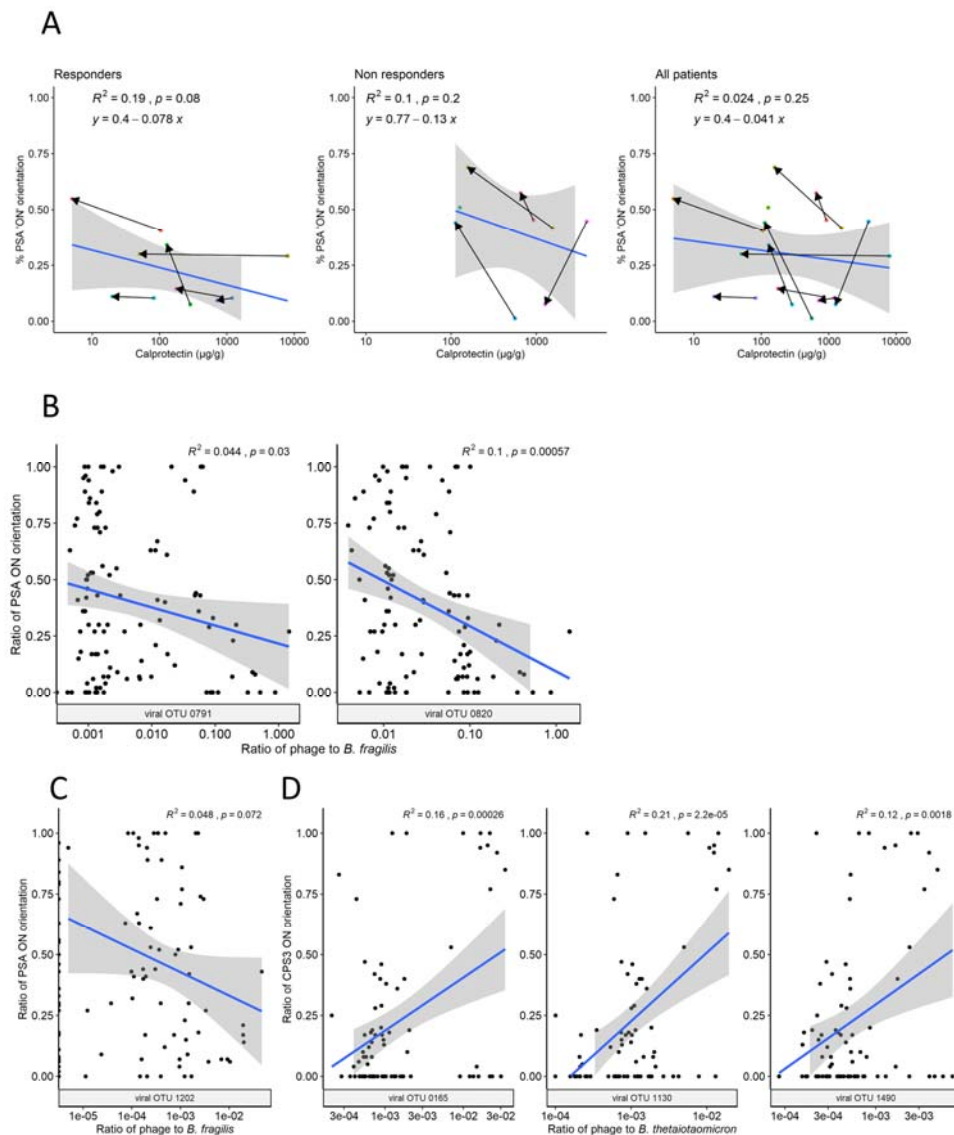


284

285 **Figure 3: Evidence of phage correlation with *B. fragilis* polysaccharide A promoter orientation.**

286 **A.** Experimental design of culturing *B. fragilis* in patients' fecal filtrates. **B.** Ratio of the 'ON' orientation of the PSA promoter of *B. fragilis*,  
 287 measured by qPCR, after ex-vivo exposure to fecal filtrates of IBD patients before and after treatment with anti-TNF (4 patients on  
 288 Infliximab (HR) and 7 patients on Humira (HuR)). Data represent the median (line in box), IQR (box), and minimum/maximum (whiskers).  
 289 (One-sided Wilcoxon rank sum test, \* $p < .05$ ). green: *B. fragilis* grown in a 1:1 ratio of M9 and PBS not exposed to fecal filtrates, orange:  
 290 *B. fragilis* exposed to fecal filtrates of patients before anti-TNF treatment. yellow: *B. fragilis* exposed to fecal filtrates of patients after  
 291 treatment. Each dot represents 3 individual experiments; lines connect experiments from the same patient; Shapes are determined by the  
 292 patients' treatments, circle: HR, triangle: HuR. **C.** Calprotectin levels ( $\mu\text{g/g}$ ) measured in patients' feces. Data represent the median (line in  
 293 box), IQR (box), and minimum/maximum (whiskers). (One-sided Wilcoxon rank sum test, \* $p < .05$ ). Dots represent samples; lines connect  
 294 samples from the same patient; Shapes are determined by the patients' treatments, circle: HR, triangle: HuR. **D.** Differential viral  
 295 taxonomic units' abundances, (from the IBDMDB cohort, count table from Nishiyama *et al.* (2020).), between samples with low 'ON'

296 orientation of the PSA promoter (<40%) and high 'ON' orientation (>60%). Differentially abundant viral taxonomic units were detected by  
 297 the DeSeq2 algorithm (Wald test,  $p < 0.01$ ). Red dots indicate differentially abundant bacteria that were determined by P value < 0.01 and  
 298 fold change >1.5 and <-1.5, respectively. E. Phage to host ratios of viral OTUs predicted to infect *B. fragilis* and detected in Figure 3D. Data  
 299 represent the median (line in box), IQR (box), minimum/maximum (whiskers), dots represent individual samples. Dashed line represents  
 300 the mean phage to host ratio of all phages predicted to infect *B. fragilis* in all samples (mean = 0.009). F. Phage to host abundances ratios  
 301 of viral OTUs 0791 and 0820 against samples' levels of PSA's promoter 'ON' orientation (red: <40%, blue: >60%) analyzed from the  
 302 IBDMDB cohort, based on the count table from Nishiyama *et al.* (2020). (Wilcoxon rank sum test, \* $p < 0.05$ ; \*\* $p < 0.01$ ) G. Differential viral  
 303 taxonomic units' abundances, (from the IBDMDB cohort, count table from Nishiyama *et al.* (2020).), between samples with low 'ON'  
 304 orientation of the CPS3 promoter (<40%) and high 'ON' orientation (>60%). Differentially abundant viral taxonomic units were detected by  
 305 the DeSeq2 algorithm (Wald test,  $p < 0.01$ ). Red dots indicate differentially abundant bacteria that were determined by P value < 0.01 and  
 306 fold change >1.5 and <-1.5, respectively. H. Phage to host abundances ratios of viral OTUs predicted to infect *B. thetaiotaomicron* and  
 307 detected in Figure 3G. Data represent the median (line in box), IQR (box), minimum/maximum (whiskers), dots represent individual  
 308 samples. Dashed line represents the mean phage to host ratio of all phages predicted to infect *B. thetaiotaomicron* in all samples (mean =  
 309 0.03).  
 310



312

313 **Figure S2: Phage correlation with *B. fragilis* polysaccharide A promoter orientation.**

314 **A.** Scatter plots of the ratio of *B. fragilis* PSA's promoter 'ON' orientation measured by qPCR and fecal calprotectin levels ( $\mu\text{g/g}$ ). Each point  
315 represents a single sample, colors indicate the same patient, arrows denote the change from samples before anti-TNF treatment to after.  
316 Blue line represents linear regression (gray area represents 95% confidence intervals). **B. & C.** Scatter plots of the ratio of *B. fragilis* PSA's  
317 promoter 'ON' orientation measured by qPCR and phage to host ratios of viral OTUs predicted to infect *B. fragilis* in Nishiyama *et al.*  
318 (2020). Blue line represents linear regression (gray area represents 95% confidence intervals). **D.** Scatter plots of the ratio of *B. fragilis*  
319 PSA's promoter 'ON' orientation measured by qPCR and phage to host ratios of viral OTUs predicted to infect *B. thetaiotaomicron* in  
320 Nishiyama *et al.* (2020). Blue line represents linear regression (gray area represents 95% confidence intervals).

321

### 322 **Bacteriophage exposure correlates with *B. fragilis* PSA promoter 'OFF' state**

323 Since a higher abundance of bacteriophages were correlated with the PSA promoter in the 'OFF'  
324 orientation, we sought to study whether encounter with bacteriophage can be associated with  
325 altered orientation of the PSA promoter in *B. fragilis*. To this end, we isolated from sewage Barc2635  
326 - a *B. fragilis* NCTC 9343 specific lytic bacteriophage [**Figure S3A**], sequenced its genome, (deposited  
327 to GenBank, accession: MN078104), and characterized its morphology by electron microscopy  
328 [**Figure S3B**].

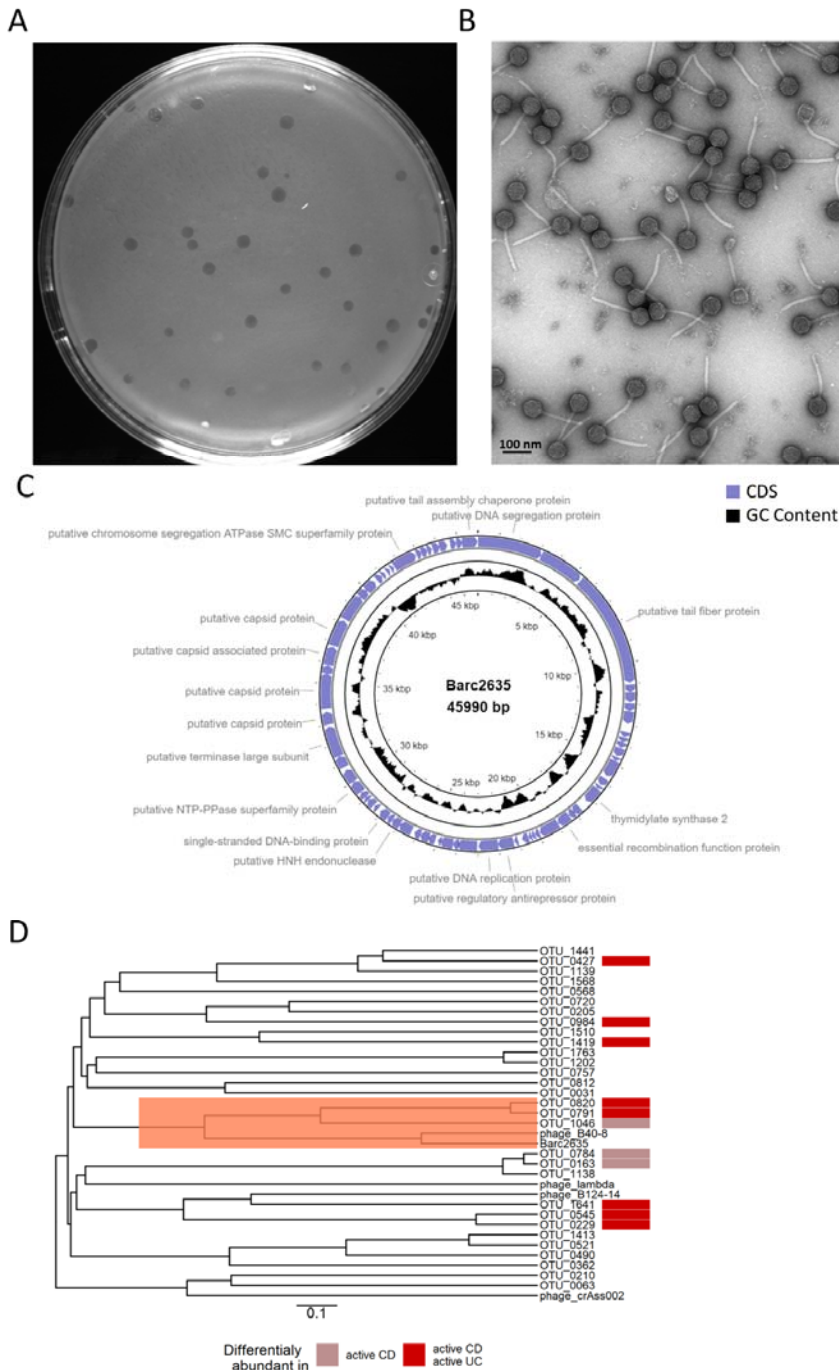
329 Barc2635 is a lytic double-stranded DNA bacteriophage of 45,990 bp with a GC content of 38.9%,  
330 containing 67 putative CDS belonging to the tailed phages unified in the class Caudoviricete,  
331 formerly *Caudovirales*, [**Figure S3C**]. Interestingly, increased abundance of bacteriophages from this  
332 class were found to be correlated with IBD patients (UC and CD)<sup>35</sup>. We analyzed the sequence  
333 similarities of Barc2635, with the *B. fragilis* associated bacteriophages identified in the IBDMD<sup>39</sup> as  
334 well as from other studies<sup>40-42</sup>, and found that Barc2635 is most similar to a cluster of  
335 bacteriophages, including viral OTUs 0791 and 0820 which were associated with active disease<sup>39</sup>  
336 [**Figure 4A, S3D**].

337

338 To test whether exposure to this bacteriophage leads to changes in relative orientations of the PSA  
339 promoter of the *B. fragilis* population, we monocolonized GF mice with *B. fragilis* NCTC 9343 in the  
340 presence or absence of Barc2635 and analyzed the PSA promoter DNA inversion state [**Figure 4B**]. In  
341 response to bacteriophage, a higher percentage of the population had the PSA promoter in the 'OFF'  
342 state [**Figure 4C**]. To further analyze the consequence of PSA promoter inversion we measured the  
343 mRNA expression of the first gene in the PSA biosynthesis locus - *upaY*<sup>43</sup>, which requires the PSA  
344 promoter to be oriented 'ON'. **Figure 4D** shows a significant reduction in *upaY* expression in mice  
345 containing Barc2635 in comparison to mice with bacteria alone. To directly measure the levels of  
346 PSA on the surface of the *B. fragilis* population, we used specific antibodies and monitored surface  
347 PSA by flow cytometry. The percentage of bacterial cells synthesizing PSA from mice with



348 bacteriophage Barc2635 was significantly lower in comparison to bacterial cells from mice without  
 349 the bacteriophage [Figure 4E], in agreement with the *upaY* expression levels and the PSA promoter  
 350 orientation. In the two groups of mice (i.e., *B. fragilis* alone and *B. fragilis* with Barc2635), the CFU  
 351 levels of *B. fragilis* were different by 0.5 a log [Supp. Figure S4A], and no phages were detected in  
 352 the *B. fragilis* monocolonized group (i.e., without bacteriophage) [Supp. Figure S4B].



353

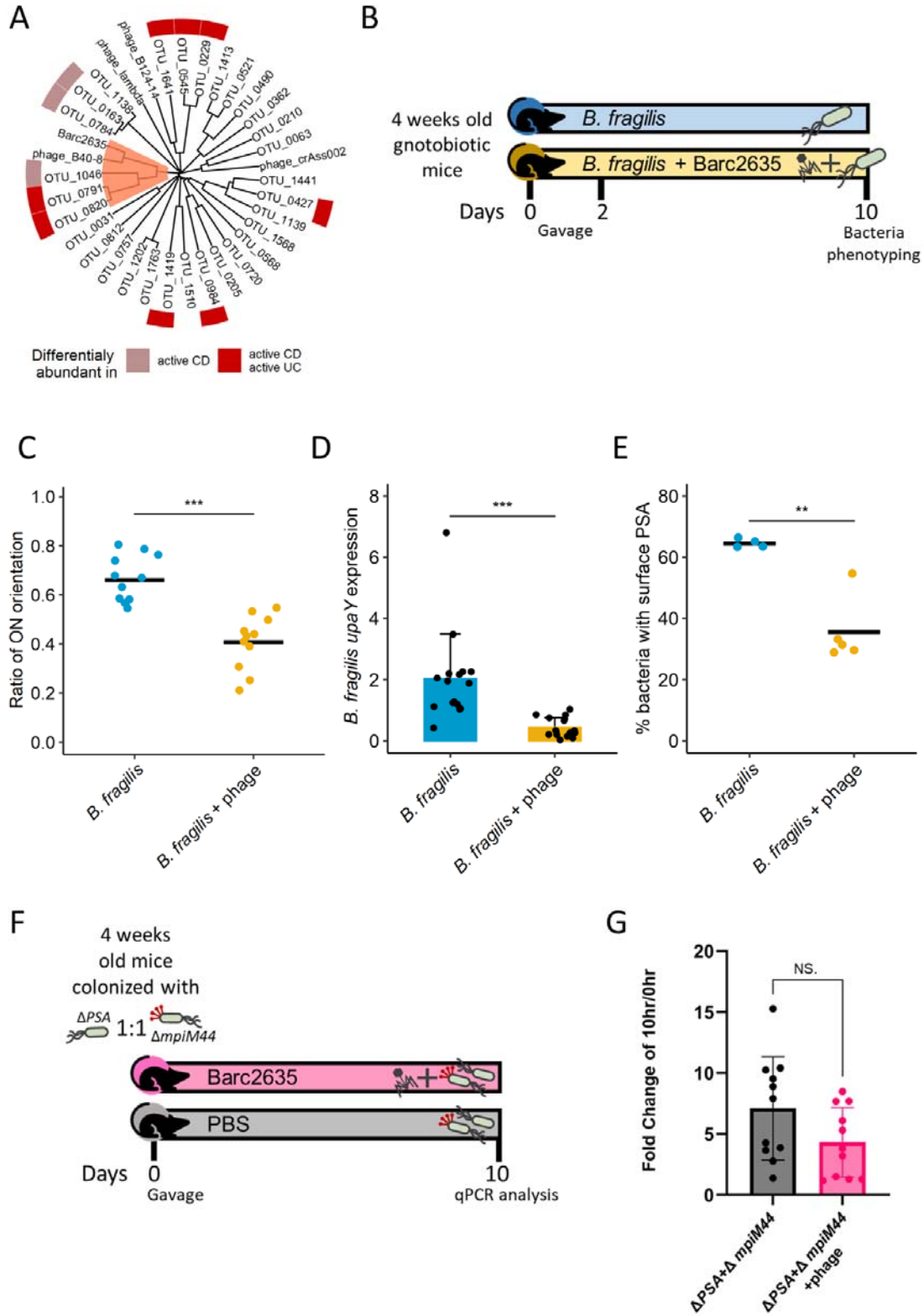
354 Figure S3: Characterization of Bacteriophage Barc2635

355 A. Barc2635 plaques on a lawn *B. fragilis* NCTC 9343. B. A representative transmission electron microscopy of Barc2635. C. Genome  
356 structure, GC content, and putative annotations of Barc2635. CDS: coding sequence. D. Phylogenetic tree based on the whole genome of  
357 viral OTUs identified as bacteriophages against *B. fragilis* as well as *Bacteroides* bacteriophages Barc2635, B40-8, B124-14, crAss002, and  
358 *Enterobacteria* phage lambda. MAFFT was used to perform multiple sequence alignment and the average-linkage method was used to  
359 construct the phylogenetic tree. Colors denote association between the viral OTUs abundances in Nishiyama *et al.* (2020), IBDMDB cohort,  
360 with active Crohn's disease (light red) or with both active Ulcerative colitis and active Crohn's disease (red).

### 361 **Barc2635 infects both $\Delta$ PSA and constitutively expressing PSA mutants**

362 To distinguish whether the PSA promoter 'OFF' state that resulted from exposure to bacteriophages  
363 was due to induction or selection, we studied the interaction of Barc2635 with two mutants of *B.*  
364 *fragilis*:  $\Delta$ *psa*<sup>44</sup> and  $\Delta$ *mpiM44*<sup>43</sup>. The former mutant lacks the PSA biosynthesis locus, and the latter  
365 has the PSA promoter locked in the 'ON' state, and therefore, constitutively synthesizes PSA. *In vitro*  
366 phage infection assays revealed comparable infection efficacy [FIGURE S4C] and no competitive  
367 advantage of either of the mutants during exposure to the bacteriophage [FIGURE S4D-F]. In  
368 agreement to the *in vitro* results, there was no selectivity of the bacteriophage towards either of the  
369 mutants in mice that were initially colonized with equal ratios of the two mutants [FIGURE 4F, G]. To  
370 note, the  $\Delta$ *psa* mutant exhibited higher fitness in colonizing GF mice, however, this advantage  
371 remained unchanged regardless of exposure to the Barc2635 [FIGURE S4G-H]. This was further  
372 confirmed by comparing the ratio of  $\Delta$ *psa*/ $\Delta$ *mpiM44* on Day 10 to the ratio on Day 0 between the  
373 treated group and the control group [FIGURE 4G]. These data show that mutants lacking PSA do not  
374 have a fitness advantage during exposure to Barc2635.

375

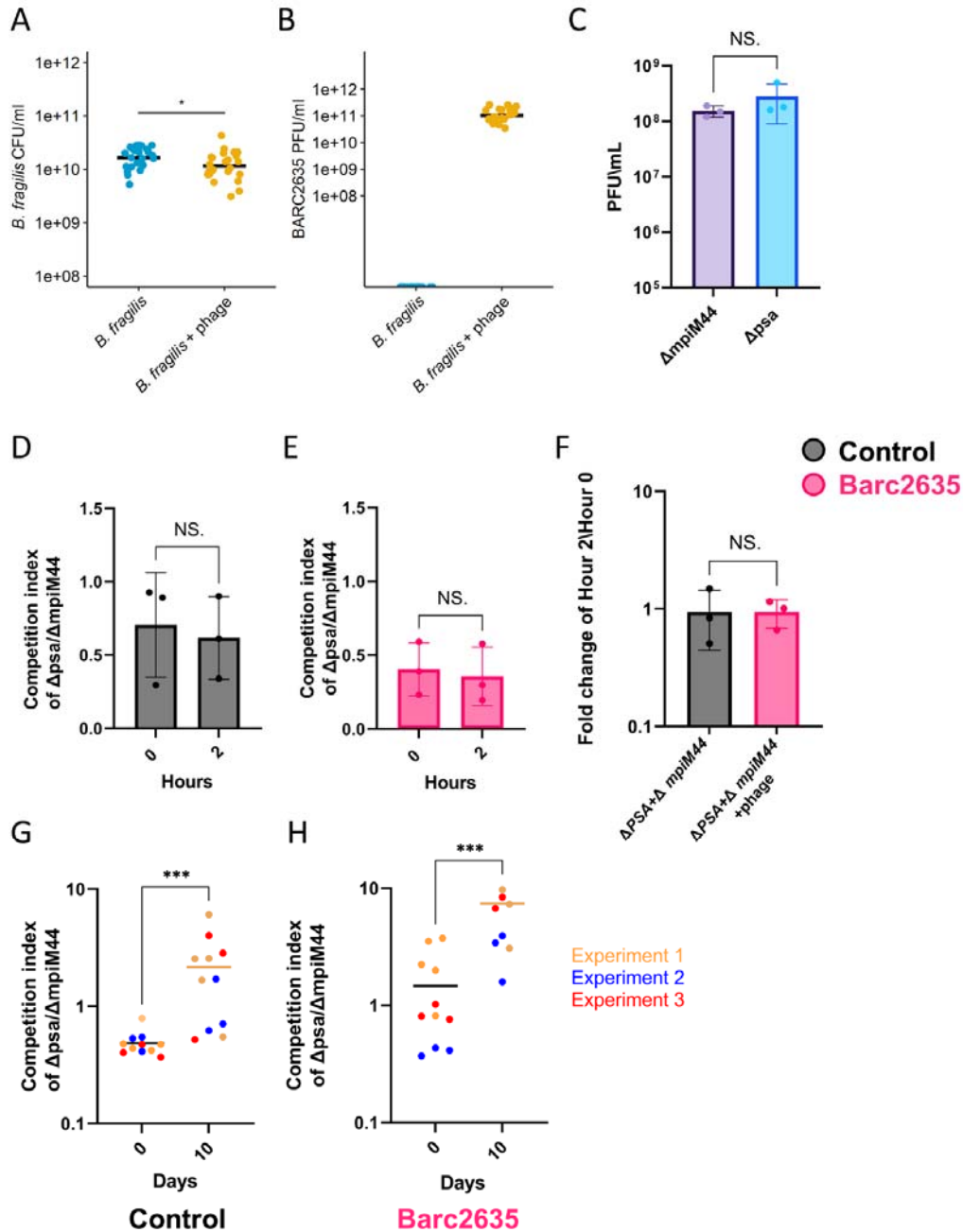


376  
 377

378

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

380 A. Phylogenetic tree based on the whole genome of viral OTUs identified as bacteriophages against *B. fragilis* as well as *Bacteroides*  
381 bacteriophages Barc2635, B40-8, B124-14, crAss002, and *Enterobacteria* phage lambda. MAFFT was used to perform multiple sequence  
382 alignment and the average-linkage method was used to construct the phylogenetic tree. Colors denote association between the viral OTUs  
383 abundances in Nishiyama *et al.* (2020), IBDMDB cohort, with active Crohn's disease (light red) or with both active Ulcerative colitis and  
384 active Crohn's disease (red). B. Experimental design of *in vivo* experiments. C. Ratio of *B. fragilis* PSA's promoter 'ON' orientation  
385 measured on day 10 by qPCR in fecal samples of gnotobiotic mice monocolonized with *B. fragilis* with and without the Barc2635  
386 bacteriophage. Horizontal lines represent the mean. Blue: Control group, Red: Barc2635 treated mice. Each dot represents a mouse.  
387 (Mann-Whitney test, \*\*\* $p < 0.001$ ). D. Expression levels of *upaY* in ceca of gnotobiotic mice monocolonized with *B. fragilis* with and  
388 without the Barc2635 bacteriophage on day 10. Levels are shown as  $2^{-(\Delta CT)}$  with *rpsL* as a reference gene. Each dot represents a mouse.  
389 (Mann-Whitney test, \*\*\*\* $P < 0.0001$ ). E. PSA presence on the surface of *B. fragilis* exposed to the Barc2635 bacteriophage, detected by  
390 anti-PSA antibodies on day 10. Bacteria were analyzed by flow cytometry for the expression of PSA, using Rabbit anti PSA antibodies. Each  
391 dot represents a mouse. F. Experimental design of *in vivo* competition experiments. G. Fold change of  $\Delta ps a \Delta mp i M 44$  at Tp10 and Tp0 in  
392 the Barc2635 bacteriophage treated group (pink) compared to  $\Delta ps a \Delta mp i M 44$  at Tp10 and Tp0 in the control group (black). Each dot  
393 represents a mouse. (Mann-Whitney test,  $p > 0.05$ ).



394

395 **Figure 54: Barc2635 infects both  $\Delta$ mpi and  $\Delta$ psa in the same manner in vivo.**

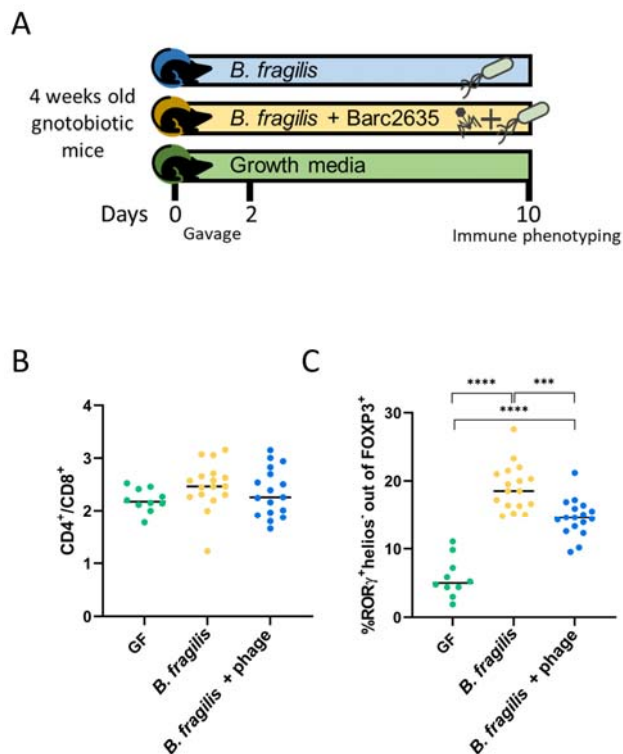
396 **A.** CFUs of *B. fragilis* in fecal samples of mice monocolonized with *B. fragilis* (blue) or colonized with *B. fragilis* + phage (yellow) measured  
 397 on day 10 (Wilcoxon rank sum test, \*p < 0.05). **B.** PFUs of Barc2635 in fecal samples of mice monocolonized with *B. fragilis* (blue) or  
 398 colonized with *B. fragilis* + phage (yellow) measured on day 10. **C.** Barc2635 plaque forming units of  $\Delta$ mpiM44 (purple) compared to  $\Delta$ psa  
 399 ( $\Delta$ psa). Each dot represents a different experiment. (Mann-Whitney test, p > 0.05). **D.** Competition index of  $\Delta$ psa/ $\Delta$ mpiM44 in the control  
 400 group at the beginning of the experiment (0 horse) compared to 2 hours, measured by qPCR. Each dot represents a different experiment  
 401 (Mann-Whitney test, p > 0.05). **E.** Competition index of  $\Delta$ psa/ $\Delta$ mpiM44 in Barc2635 treated group at the beginning of the experiment (0  
 402 horse) compared to 2 hours, measured by qPCR. (Mann-Whitney test, p > 0.05) **F.** Fold change of  $\Delta$ psa/ $\Delta$ mpiM44 at the beginning of the  
 403 experiment (0 horse) compared to 2 hours in the Barc2635 bacteriophage treated group (pink) compared to  $\Delta$ psa/ $\Delta$ mpiM44 at Tp2 and

404 Tp0 in the control group (black). Each dot represents a different experiment. (Mann-Whitney test,  $p > 0.05$ ). G. Competition index of  
405  $\Delta$ psa $\Delta$ mpiM44 in the control group on day 0 compared to day 10, measured by qPCR. Each dot represents a mouse (Mann-Whitney test,  
406  $***p < 0.001$ ). H. Competition index of  $\Delta$ psa $\Delta$ mpiM44 in the treated group on day 0 compared to day 10, measured by qPCR. Each dot  
407 represents a mouse (Mann-Whitney test,  $***p < 0.001$ )

#### 408 Bacteriophage-driven phase variation in *B. fragilis* results in reduction of Treg cells

409 It is well-established that the PSA of *B. fragilis* NCTC 9343 induces regulatory T cells (Tregs,  $CD4^+Foxp3^+ROR\gamma^+$ ) in the colonic lamina propria of mice<sup>13,45-47</sup>. To monitor the effect of Barc2635  
410 during *B. fragilis* colonization to the  $CD4^+Foxp3^+ROR\gamma^+$  Tregs population, we extracted cells from the  
411 colonic lamina propria and immunophenotyped them by flow cytometry [Figure 5A]. We found that  
412 in the presence of Barc2635, the  $CD4^+Foxp3^+ROR\gamma^+$  Tregs population is decreased concomitant with  
413 the decrease in *upaY* transcription, coinciding with less surface production of the  
414 immunomodulatory PSA [Figure 5B, Figure 5C].  
415

416

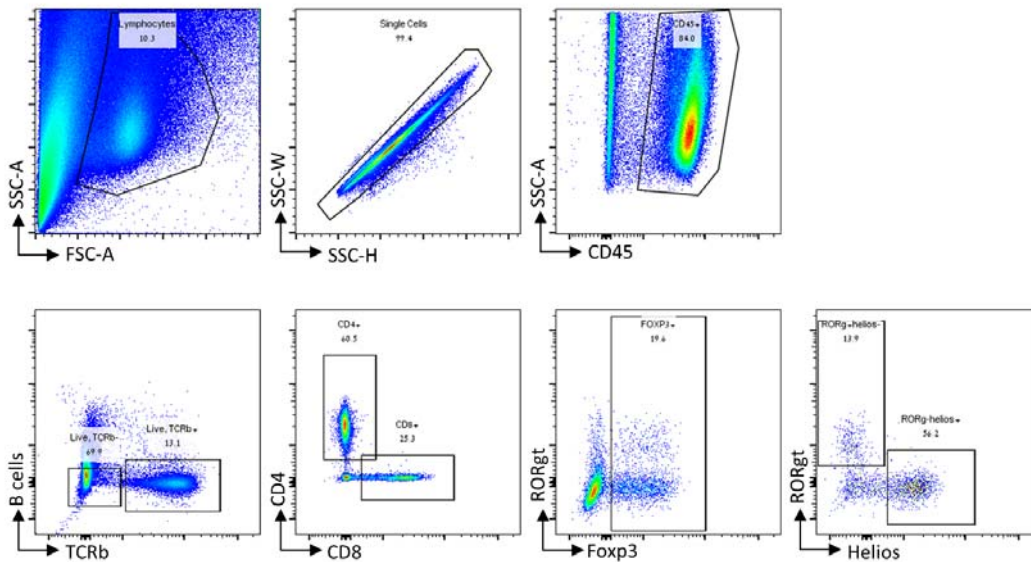


417

418

419 **Figure 5: Bacteriophage-driven phase variation in *B. fragilis* results in reduction of Treg cells.**

420 A. Experimental design of *in vivo* experiments. B. CD4<sup>+</sup> to CD8<sup>+</sup> ratio out of CD45<sup>+</sup>TCR<sup>+</sup> live cells in colon. Single cells were isolated from  
421 colon lamina propria. Immune cells were analyzed by flow cytometry. Each dot represents a mouse. C. RORγ<sup>+</sup> helios<sup>+</sup> percentages out of  
422 FOXP3<sup>+</sup> cells in colon. Single cells were isolated from colon lamina propria. Immune cells were analyzed by flow cytometry. Each dot  
423 represents a mouse.  $***P < 0.001$ ,  $****P < 0.0001$ , one-way analysis of variance (ANOVA).



424

425 **Figure S5: Gating strategy.**

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

427

428 **Discussion**

429 Phase variations, prevalent in host-associated species, especially the gut Bacteroidales<sup>9</sup>, contribute  
430 to bacterial fitness in changing ecosystems. Reversible DNA inversions lead to phase variable  
431 synthesis of numerous molecules (e.g. surface, regulatory, and other molecules), and as such, confer  
432 functional plasticity. Our study reveals that the phase variable states of certain molecules correlate  
433 with gut inflammation, with potential implications on host physiology. By analyzing six different  
434 human gut metagenomic datasets with samples from IBD patients and healthy individuals, and using  
435 controlled mouse experiments, we identified multiple invertible genomic regions that map to 25  
436 different Bacteroidales strains that are differentially oriented in disease and health. Notably, we find  
437 that both intergenic regions (e.g. promoters) and intragenic regions (e.g. genomic “shufflons”)  
438 exhibit altered orientations in gut inflammatory conditions - affecting gene expression.

439 The most prevalent genomic regions with differential orientations during inflammation and health  
440 were in polysaccharide (PS) promoters and near *susC/susD* homologs - both with  
441 immunomodulatory potential. SusC-like proteins are abundant  $\beta$ -barrel outer membrane proteins  
442 involved in nutrient acquisition in Bacteroidetes. Recently, an epitope of SusC proteins was shown to  
443 elicit T cell responses in IBD patients and healthy controls<sup>48</sup>, suggesting that operons that include  
444 SusC homologs might confer immunomodulatory properties to the bacteria. Among the phase-

445 variable PSs, the anti-inflammatory PSA promoter of *B. fragilis* showed a higher percentage of  
446 reverse oriented reads in IBD patients compared to healthy controls, indicating that the 'OFF'  
447 orientation was more prevalent in IBD patients, potentially limiting the protective, anti-  
448 inflammatory, effects of PSA. These results align with a previous study<sup>49</sup> that focused on the PSA  
449 promoter of *B. fragilis* in IBD patients using PCR digestion of biopsy samples. We further find that  
450 some DNA inversion states, like the PSA promoter of *B. fragilis*, are dynamic and can respond to  
451 changes in their local environment, specifically during inflammation.

452 Intriguingly, the preferential 'OFF' orientation of the PSA promoter of *B. fragilis* did not occur to the  
453 same extent in inflamed monocolonized mice as it did in inflamed mice colonized with a human  
454 microbiota and when bacteria were exposed to patients' fecal filtrates. These results suggest that  
455 environmental factors in inflamed humanized mice and IBD patients are necessary for this effect.  
456 Our analysis of phage OTUs in the IBDMDB cohorts revealed enrichment of *B. fragilis* associated  
457 phage OTUs in samples where the PSA promoter was present in the 'OFF' orientation. In addition,  
458 exposure of *B. fragilis* to an isolated bacteriophage, Barc2635, resulted in higher frequency of the  
459 PSA promoter 'OFF' orientation, suggesting a role for gut bacteriophages in bacterial phase variation.

460 Our analyses were not able to distinguish the cause of the change in DNA orientation during  
461 inflammation and phage exposure. It is possible that these conditions induce inversions to the 'OFF'  
462 orientation but may also result from selection of bacteria that are not expressing PSA. Comparison  
463 of mutants that are unable or that constitutively synthesize PSA revealed similar relative PSA  
464 promoter orientations after Barc2635 exposure, suggesting that the observed decrease in PSA  
465 expression may be the result of induction of inversion rather than selection pressures. If so, this area  
466 is ripe for future mechanistic understanding.

467 The anti-inflammatory effects of PSA are mediated by upregulation of induced Tregs  
468 (Foxp3+ROrgt+) and subsequent induction of IL-10 secretion<sup>12,13</sup>. We found that the Treg cells  
469 populations decreased in *B. fragilis* monocolonized mice infected with Barc2635, suggesting that the  
470 increase in bacteria with the PSA promoter in its 'OFF' orientation has implications on the host  
471 immune system, linking bacteriophage exposure to alterations in bacterial functionality with  
472 concomitant effects on host physiology. Applying the same analysis to *B. thetaiotaomicron*'s CPS3,  
473 we observed bacteriophages that are more abundant in samples with a higher ratio of CPS3 'ON'  
474 oriented promoters, however, the bacteriophage to host abundances ratios per sample were too  
475 low to conclude that *B. thetaiotaomicron* bacteriophages lead to populations with altered CPS3  
476 synthesis.



477 Altogether, we demonstrate differential orientations of Bacteroidales invertible regions in human  
478 IBD and identify bacteriophages as a potential trigger. These phase variations result in bacterial  
479 functional plasticity affecting the host immune system, correlating with inflammation.  
480 Bacteriophages are one factor that alter bacterial phase variable states. However, within the  
481 inflamed gut, there are numerous factors that may potentially alter the invertible states of  
482 Bacteroidales populations. These include the microbiome<sup>35,39,50-52</sup> (e.g. neighboring microbes), the  
483 gut metabolome<sup>50</sup>, an altered immune system<sup>53</sup> and physical alterations in the intestine such as  
484 abnormal pH concentrations<sup>54</sup>, osmotic<sup>55,56</sup> and oxidative stress<sup>57,58</sup>. For example, Tropini et al.<sup>59</sup>  
485 showed that PEG induced osmotic perturbation resulted in an increased ratio of CPS4 to CPS5 in *B.*  
486 *thetaiotaomicron*.

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

#### 494 **Limitations of the Study:**

495 The six datasets analyzed in this study included adults and children, healthy controls and inflamed  
496 patients (before and after treatment) and longitudinal samples. The impact of each of these factors,  
497 and the impact of diet and lifestyle, merit further research. The PhaseFinder<sup>9</sup> algorithm that we  
498 applied identifies DNA inversions in short reads. As such, our analysis does not include other phase  
499 variation mechanisms like duplications or insertions. Also, analysis of short reads of metagenomics  
500 sequencing data may overlook inversions resulting from long structural genomic alterations<sup>3,7,60</sup>.  
501 Furthermore, metagenomic sequencing data averages the whole bacterial population of samples,  
502 implying a dilution factor for low abundant bacteria. Moreover, genomic inversion sites might  
503 interfere with genome sequence assemblies, leading to incomplete genomes, and thus could be  
504 overlooked when applying tools that search inversions in reference genomes. Also, our  
505 bacteriophage computational analysis, which is based on Nishiyama *et al.*<sup>39</sup>, might have overlooked  
506 some of the bacteriophages in the patients, since it relies on prior analysis of MAGs that were  
507 predicted as temperate phages. This study used an identified and isolated lytic bacteriophage for *B.*  
508 *fragilis*, further studies are required to elucidate the functional effects of IBD associated lytic and

509 temperate bacteriophages. With these limitations notwithstanding, our study highlights the  
510 importance of considering bacterial phase variation in the context of IBD and its potential impact on  
511 inflammation and on altered microbe-host interactions.

## 512 [Methods](#)

### 513 [Bacterial strains and growth conditions](#)

514 *B. fragilis* NCTC 9343 was grown in Brain heart infusion (BHIS) supplemented with 50mg/L hemin  
515 (Alfa Aesar) in 1N NaOH, and 2.5µg/L vitamin K under strictly anaerobic conditions (80% N<sub>2</sub>, 10%  
516 H<sub>2</sub>, 10% CO<sub>2</sub>) at 37 °C in an anaerobic chamber.

### 517 [Mice](#)

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

520 4-5 weeks-old Germ-Free (GF) C57BL/6 mice (males and females) from the Technion colony were  
521 used. Mice were housed and maintained in a GF care facility and were provided with food and water  
522 *ad libitum*; they were exposed to a 12:12 h light-dark cycle at room temperature.

523 Humanized mice were generated by oral gavage of GF C57BL/6 mice with 200 µl mix of human  
524 fecal microbiota and cultured *B. fragilis* NCTC 9343. To prepare this mix, human feces from a single  
525 healthy human donor were suspended in sterile PBS 1:10 w/v (1 gram feces was suspended in 10  
526 ml of sterile PBS). Next, 10<sup>8</sup> CFUs of *B. fragilis* NCTC 9343 were resuspended in 1 ml of sterile PBS  
527 and added to 1 ml of fecal supernatant. *B. fragilis* NCTC 9343 were grown on Brain Heart Infusion  
528 agar plates (BHI, BD BBLTM) supplemented with 50µg/ml hemin (Alfa Aesar) in 1N NaOH and  
529 2.50µg/ml vitamin K (Thermo Fisher Scientific) in 100% EtOH, at 37°C in an anaerobic chamber,  
530 85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% H<sub>2</sub> (COY). The progeny of these mice were co-housed, randomized, and  
531 grouped before treatment.

532 Gnotobiotic mice monocolonized with *B. fragilis* were created by oral gavage of GF C57BL/6 mice  
533 with 200 µl of 5x10<sup>8</sup> CFU/ml of *B. fragilis* NCTC 9343, grown in the same conditions as above. The  
534 mice were co-housed, randomized, and grouped before treatment. Both gavages, in humanized and  
535 monocolonized mice, were performed once.

536 For phage immunomodulation studies, GF mice were gavaged twice (on day 0 and on day 2) with *B.*  
537 *fragilis* NCTC 9343 (10<sup>8</sup> CFU in 200 µl Bacteroides phage recovery medium (BPRM) media) and  
538 bacteriophage Barc2635 (10<sup>9</sup> PFU in 200 µl 0.22-µm-filtered BPRM), *B. fragilis* only (10<sup>8</sup> CFU in 200 µl  
539 BPRM media), or growth media as control. On day 10 mice were sacrificed for bacteria and immune  
540 phenotyping.

541

542 DSS model

543 For DSS experiments, male and female mice were co-housed separately. Mice were randomized and  
544 grouped (water control and DSS separately) before treatment. Throughout the experiments, close  
545 attention was paid to monitor any distress responses exhibited by the mice: mice with a 20%  
546 decrease in body weight, immobility, or an extreme reaction to touch, were excluded from the  
547 experiments.

548 For acute DSS-induced colitis in humanized mice, progeny of humanized mice were treated with  
549 3% DSS in their drinking water for 9 days followed by 5 days recovery before sacrificing for the  
550 experiment. The control animals were administered distilled water.

551 Fecal samples were obtained on days 0, 6, 9 and 14 (before DSS treatment, during and after  
552 discontinuing DSS treatments), for further analyses.

553 For acute DSS-induced colitis in gnotobiotic mice colonized with *B. fragilis*, monoclonized mice  
554 were treated with 3% DSS in their drinking water for 7 days. The control animals were administered  
555 distilled water.

556 Fecal samples were obtained on days 0 and 7 (before and after DSS treatment), for further analyses.

557 Mice Inflammation measurement

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

562 Mice weights were assessed using an electronic scale at the same day-time.

563 qPCR and primers

564 DNA was extracted from fecal samples (30-50 mg) using ZymoBIOMICS DNA Miniprep Kit [Zymo  
565 research]. The 'ON'/'OFF' orientation of the PSA promoter in the extracted DNA samples was  
566 determined by quantitative polymerase chain reaction (qPCR) using SYBR® Green mix [Thermo Fisher  
567 Scientific]. Two sets of primers were designed to target the PSA locus [Figure S1A-C]. One set  
568 targeted the reference gene *upaY*, the first gene immediately downstream to the promoter region,

569 and used as a proxy to the number of bacteria in the samples. The second set of primers targeted  
570 the promoter region and would only produce a product when the orientation is 'ON'. The ratio of  
571 'ON'/'OFF' PSA orientation in the samples was calculated using the  $2^{-\Delta\Delta CT}$  method<sup>61</sup> calculated  
572 against the PSA locked 'ON' results (100% 'ON' orientation).

573 *upaY* gene expression was determined by RT-qPCR with the *upaY* primers. RNA was extracted from  
574 fecal samples using zymoBIOMICS RNA miniprep kit [Zymo research]. Following DNase treatment  
575 (NEB), reverse transcription of RNA to cDNA was performed using the qScript cDNA Synthesis Kit  
576 [Quantabio]. *upaY* mRNA levels were determined by qPCR using SYBR® Green mix [Thermo Fisher  
577 Scientific] with primers against *rpsL* as a reference gene. Controls for non-template and non-reverse  
578 transcription were included. Primer efficiency was determined for each set of primers<sup>62</sup>. The  $2^{-\Delta\Delta CT}$   
579 method was employed for the specificity fold change tests.

580 Primers:

581

B_Frag_upaY_F	CGCTCGGACAAAGAAGGACC
B_Frag_upaY_R	ACTTCTACCCTACGACGACGA
B_Frag_PSA_M	GGTGTTC AAAAGACGAACGT
B_Frag_PSA_F	TGTGTAATGATAGGAGGCTAGGG
rpsL_F	CCGAACTCTGCAATGCGTAA
rpsL_R	CGCGAACCAAGTACGATTGAG

582

### 583 Mice feces sequencing

584 DNA was extracted from fecal samples (30-50 mg) using ZymoBIOMICS DNA Miniprep Kit [Zymo  
585 research]. Samples underwent quality control by Qubit fluorescence analysis to determine the  
586 concentration of DNA for downstream analysis (ThermoFisher, Cat. Q32850). Libraries were

587 prepared using the Illumina Tagmentation DNA prep streamlined library preparation protocol  
588 according to the manufacturer's instructions with a minimum of 50 ng of DNA starting mass and 8  
589 cycles of PCR enrichment, ending with a fragment size of 550 bp. IDT for Illumina DNA/RNA UD  
590 indexes and Nextera DNA CD indexes were used (Illumina IDT, Cat. 20027213; Illumina Nextera, Cat.  
591 20018708).

592 All libraries were diluted to 15 pM in 96-plex pools and validated on 100-cycle paired-ends read  
593 Miseq V2 runs (Illumina, Cat. MS-102-2002), before shipping to the US at 4 nM for sequencing on the  
594 Novaseq 6000 in S4 mode at 96-plex in a 300-cycle paired-end reads run, with an estimated read  
595 depth of 30 Gbp per sample (Illumina, Cat. 20028312). Final loading concentration of 600 pM. All  
596 sequencing runs were performed with a spike-in of 1% PhiX control library V3 (Illumina, Cat. FC-110-  
597 3001). The sequencing mean library size was 134,629,540.5 reads [range: 10,107,679 - 396,239,822].

#### 598 Taxonomic profiling

599 Community profiling was performed using metaphlan4<sup>63</sup> v4.0.059 with mpa database vJan21. For  
600 each sample, the forward reads were first aligned against the mpa database using bowtie2  
601 v2.3.5.160 (flags `--sam-no-hd --sam-no-sq --no-unal --very-sensitive`). Next,  
602 the resulting sam file was analyzed by metaphlan4 with default parameters, resulting in a merged  
603 relative abundances table.

#### 604 Microbiome analysis

605 Initial visual exploration of sequenced data was conducted using the MicrobiomeAnalyst<sup>64</sup> web  
606 platform and followed by a comprehensive statistical analysis of using the Phyloseq<sup>65</sup> version 1.44.0  
607 and vegan<sup>66</sup> version 2.6-4 packages in R v4.3.1. Read counts were transformed into relative  
608 abundances by normalization to the total number of reads per sample. Low-abundance filters were  
609 applied to discard species whose relative abundance did not reach 0.1% and did not appear in at  
610 least 20% of the samples. Alpha diversity (observed species and Shannon diversity) was calculated at  
611 the species level using the 'plot\_richness' function and compared with the Kruskal–Wallis rank sum  
612 test. Beta diversity distance matrices (Aitchison distance) were ordinated using the 'ordinate'  
613 function, compared using the vegan<sup>66</sup> package's 'adonis' function (Permutational Multivariate  
614 Analysis of Variance), and visualized using PCoA. Beta dispersion (Aitchison distance) was calculated  
615 using the vegan<sup>66</sup> package's 'betadisper' function. Differential abundance of species between mice  
616 groups and timepoints were computed with R package Maaslin 2<sup>67</sup>, version 1.14.1 with CLR

617 transformation of the data. Plotting was performed using R packages tidyr version 1.3.0, reshape2  
618 version 1.4.4, ggpubr version 0.6.0, ggplot2<sup>68</sup> version 3.4.3, and EnhancedVolcano version 1.18.0.

619

620 Analysis of differential abundances of phages from Nishiyama *et al.*<sup>39</sup>

621 Count tables, metadata, and the sequences of MAGs and viral regions from Nishiyama *et al.*<sup>39</sup>  
622 analysis of the IBDMDB dataest were obtained from the following site:

623 <ftp://ftp.genome.jp/pub/db/community/ibd-phage/>.

624 Statistical analysis was done using the Phyloseq<sup>65</sup> version 1.44.0 and DESeq2<sup>69</sup> version 1.40.2 version  
625 2.6-4 packages in R v4.3.1. The data was filtered to include only viral OTUs, and Low-abundance  
626 filters were applied to discard OTUs whose had less than 10 read counts and did not appear in at  
627 least 20% of the samples. Read counts were normalized to relative abundances. Samples were split  
628 to low (<40%) or high (>60%) ratios of the PSA promoter 'ON' orientation, according to the  
629 PhaseFinder results. Differential abundance of viral OTUs between the groups were computed with  
630 R package DESeq2<sup>69</sup> with default settings. Plotting was performed using R packages tidyr version  
631 1.3.0, reshape2 version 1.4.4, ggpubr version 0.6.0, ggplot2 version 3.4.3, and EnhancedVolcano  
632 version 1.18.0. Phage-to-host abundance ratio was calculated for each phage-*B. fragilis* and phage-*B.*  
633 *thetaiotaomicron* pairs according to the bacterial host predictions done by Nishiyama *et al.*<sup>39</sup>.

634 Identification of DNA inversion regions

635 Representative Bacteroides species were selected from using the following approach: database of  
636 human-associated microbial species was constructed from 118K metagenomics assembled genomes  
637 (MAGs) recovered from human-associated metagenomics samples acquired from Pasolli *et al.*<sup>70</sup>,  
638 Almeida *et al.*<sup>71</sup>, and Forster *et al.*<sup>72</sup>. Taxonomy assignment was performed using gtdbtk v2.0.0 and  
639 GTDB release 207 with the classify\_wf program and default parameters<sup>73</sup>. Overall, 34 *Bacteroides*  
640 species were identified of which 33 had species-level GTDB taxonomy assignments. For each of these  
641 species, the representative genome from GTDB was used for the PhaseFinder analysis. We have also  
642 included the genomes of three additional *B. fragilis* strains, *Phocaeicola dorei*, and *Parabacteroides*  
643 *distasonis*. Overall, 39 genomes were included. Information about the genomes is provided in Table  
644 S1.

645 PhaseFinder<sup>9</sup> (v1.0) was used to identify invertible regions in metagenomics samples. The default  
646 parameters of PhaseFinder were used. Putative inversion regions were detected by identifying  
647 inverted repeats in bacteroidales genomes using the 'locate' function. A database containing the  
648 inversion regions forward and reverse orientations was created using the tool's 'create' function.  
649 Using the tool's 'ratio' function, metagenomic samples from the publicly available cohorts were then  
650 aligned to the database, resulting in the ratio of reverse oriented reads out of all reads assigned to  
651 each region. We filtered the results by removing identified regions with <20 reads supporting either  
652 the forward or reverse orientations combined from the paired-end method, mean Pe\_ratio <1%  
653 across all samples, and sites within coding regions of rRNA products. To compare DNA orientation  
654 patterns between CD and UC patients with those of healthy controls, we focused on sites that  
655 displayed a difference of over 10% between at least one of the groups. The Wilcoxon rank sum test  
656 was employed to conduct these comparisons, and the Benjamini-Hochberg method was utilized to  
657 correct for multiple comparisons, with a false discovery rate (FDR) set at less than 0.1. Each  
658 invertible region was manually curated to assess its coding regions, gene annotations, and their  
659 putative functions. Briefly, genomic regions were visualized online using the NCBI Graphical  
660 Sequence Viewer (version 3.47.0). Invertible regions with coding sequences (CDS) within them, were  
661 annotated according to the CDS name(s). Invertible regions lacking CDS within them were searched  
662 for CDSs that start in proximity to the invertible DNA sites (<200~bp). CDSs in the region (four  
663 upstream and four downstream) were used to assess the functionality of the region. Regions  
664 containing or in proximity to rRNA and tRNA genes were filtered from the comparisons as well as  
665 invertible regions with no CDSs start in proximity to the inverted repeats.

#### 666 Bacteriophage isolation

667 Bacteriophage BARC2635 was isolated from raw sewage. Briefly, inflowing raw sewage for a waste-  
668 water treatment plant (WWTP) from Barcelona (Spain) was filtered through low protein binding 0.22  
669  $\mu\text{m}$  pore size polyethersulfone (PES) membrane filters (Millex- GP, Millipore, Bedford,  
670 Massachusetts) to remove bacteria. Isolated plaques were obtained by the double-agar layer  
671 technique<sup>74</sup>. Briefly, tubes containing 2.5 ml of soft BPRM-agar kept at 45°C were inoculated with 1  
672 ml of an exponential growth phase culture (OD<sub>600</sub>=0.3, corresponding to ca  $2 \times 10^8$  CFU / ml) of the  
673 host bacteria grown in BPRM broth and 1 ml of the filtered sewage sample. After gently mixing, the  
674 contents of each tube were poured onto a plate of BRPM-agar and incubated inside GasPak (BBL)  
675 jars at 37°C. Plaques were clearly spotted after 18 h of incubation.



676 For phage isolation, discrete well-isolated plaques were stabbed with a sterile needle and inoculated  
677 in a tube containing 5 ml of BRPM broth. Then 1 ml of a culture of *B. fragilis* NCTC 9343 in  
678 exponential growth was inoculated into the tube, which was then incubated for 18h at 37 °C. After  
679 incubation, an aliquot of the culture was treated with chloroform (1:10 (v:v), vigorously mixed for 5  
680 minutes and centrifuged at 16,000xg for 5 minutes<sup>75</sup>. The supernatant containing the phage  
681 suspensions were further filtered through low protein binding 0.22 µm pore size polyethersulfone  
682 (PES) membrane filters (Millex-GP, Millipore, Bedford, Massachusetts), diluted and plated as  
683 indicated in the previous paragraph to verify the uniformity of the plaques. Then, one well  
684 differentiated plaque was stabbed and the whole operation was repeated to obtain a high titer, over  
685 1x10<sup>9</sup> plaque forming units (PFU), phage suspensions.

#### 686 Phage genome sequencing

687 Phage particles were PEG-precipitated (6,000-12,000 MW, 8%), and isolated by a CsCl gradient; 33 g,  
688 41 g, 55 g in 50 ml TM buffer (50mM Tris-Cl pH8.0, 10mM MgCl<sub>2</sub>), ultracentrifuged at average  
689 152,000xg for 1.5hrs, and dialyzed overnight in 100mM Tris pH7.5, 1M NaCl, 1mM EDTA<sup>58</sup>. Genomic  
690 DNA was extracted using phenol-chloroform as described<sup>76</sup>.

691 Illumina sequencing of Bacteroides phage Barc2635 was performed at the Biopolymers Facility,  
692 Harvard Medical School, Department of Genetics, producing paired-end reads of 150 bp. Adapter  
693 sequence removal and quality trimming was performed using BBDuk, part of the BBTools (v 37.50)  
694 suite of programs. The reads were further screened against NCBI's UniVec\_Core database (build  
695 10.0) and the *B. fragilis* NCTC 9343 genome sequence using blastn and reads that returned a  
696 significant hit to either were removed. The phage genome was assembled de novo using Velvet  
697 1.2.10 under a k-value determined by Velvet Optimizer (v. 2.2.5). The genome was annotated using  
698 a customized version of Prokka<sup>77</sup> v1.12, altered to additionally utilize the profile Hidden Markov  
699 Model (HMM) libraries of Pfam<sup>78</sup> version 35, TIGRFAM<sup>79</sup> version 15, and the Clusters of Orthologous  
700 Genes<sup>80</sup> (COGs, 2020 update, and PRotein K(c)usters (PRK) portions of NCBI's Conserved Domain  
701 Database<sup>81</sup> during annotation., submitted to NCBI, and assigned GenBank accession MN078104.  
702 Phage Genome map [Figure S3C] was visualized using the online tool Proksee (<https://proksee.ca/>,  
703 accessed on 21 January 2023).

704

705 Barc2635 susceptibility and competition assays:

706 Frozen stocks of *B. fragilis* NCTC 9343  $\Delta$ *mpiM44*<sup>43</sup> (genetically engineered bacteria that constitutively  
707 express PSA) and *B. fragilis* NCTC 9343  $\Delta$ *PSA*<sup>44</sup> (lacks PSA biosynthesis locus) maintained in 25%  
708 glycerol at  $-80^{\circ}\text{C}$  were thawed on Brain Heart Infusion agar plates (BHI, BD BBLTM) supplemented  
709 with  $5\mu\text{g/ml}$  hemin (Alfa Aesar) in  $1\text{N}$  NaOH and  $2.5\mu\text{g/ml}$  vitamin K (Thermo Fisher Scientific) in  
710 100% EtOH, at  $37^{\circ}\text{C}$  in an anaerobic chamber, 85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% H<sub>2</sub> (COY). Then strains were  
711 grown anaerobically for up to 3 days, and a single colony was picked for each bacterial strain,  
712 inoculated into  $5\text{ml}$  BPRM and grown anaerobically overnight to provide the starting culture for  
713 experiments. A dilution of 1:10 was done the following day and the diluted culture was incubated at  
714 the same conditions until mid-logarithmic phase of growth.

715 At the beginning of the experiment, comparable CFU amounts of each strain were verified at OD 0.5  
716 ( $\Delta$ *mpiM44*  $2 \times 10^8$  and  $\Delta$ *PSA*  $2 \times 10^8$ ).

717 For susceptibility assays:  $300\mu\text{l}$  of each strain was infected with Barc2635 in 1:100 ratio, and then  
718 incubated for 5 minutes in  $37^{\circ}\text{C}$ , followed by a centrifuge of  $4,500\text{g} \times 5$  minutes and re-suspension  
719 with clean  $300\mu\text{l}$  BPRM to get rid of free phages that did not infect the cells. The infected bacterial  
720 cells were added to 3 ml of the molten soft top agar and mixed well before being poured onto the  
721 bottom agar. The following day plaques were counted according to the formula:  
722  $\text{PFU/ml} = \# \text{ of plaques} / (\text{dilution} * \text{infection volume in ml})$ .

723 For competition assays *in vitro*: Both bacterial strains ( $\Delta$ *mpiM44* and  $\Delta$ *PSA*) were grown at the same  
724 anaerobic conditions until mid-logarithmic phase of growth. Equal CFU's were verified.  
725 Next, bacterial cells were mixed in 1:1 ratio and then Barc2635 was added to the mix at 1:100 ratio.  
726  $200\mu\text{l}$  of Time point (Tp) 0 (starting point of the experiment) and Tp2 (2 hours after Barc2635  
727 infection) were taken to qPCR analysis.

728 *In vivo* competition was done by gavaging C57BL/6 mice with equal ratios of  $\Delta$ *mpiM44* and  
729  $\Delta$ *psa*, wait two days for the bacteria to settle in and then Barc2635 ( $10^9$  PFU in  $200\mu\text{l}$   $0.22\text{-}\mu\text{m}$ -  
730 filtered BPRM) was added on Tp0. 10 days after the inoculation mice fecal samples were collected  
731 for qPCR analysis.

732 Primers used:

733 For  $\Delta$ *mpiM44*

734 wcf\_F: GGC CTC CTT CAT CTC AGG TTT ATC C

735 wcf\_R: GAT AAT CGC GGC ACC CTA TGG G

736 For  $\Delta$ PSA  
737 mpi\_F: AAG AGG GCT ATG TGT TTC AGG ACG  
738 mpi\_R: CTG CGT GCG AGA GCT TCT TTG

739

#### 740 Viral OTUs multiple alignment

741 Alignments and phylogenetic tree of whole genomes of viral OTUs identified as bacteriophages  
742 against *B. fragilis*<sup>39</sup> as well as bacteriophage Barc2635 were generated using MAFFT<sup>82</sup> online service  
743 (version 7). Multiple sequence alignment was done using the default settings of the site. An average  
744 linkage UPGMA guide tree was constructed using average-linkage UPGMA with the online MAFFT<sup>82</sup>  
745 service and was visualized with R package ggtree<sup>83</sup> (version 3.4.4).

#### 746 Fecal filtrates of patients

747 Recruitment of IBD patients for this study was conducted at the Rambam Health Care Campus  
748 (RHCC). The study was approved by the local institutional review boards with study numbers 0052-  
749 17 and 0075-09 in which all patients consented to be included in it. Fecal samples were collected by  
750 the patients, prior to their clinic visit. The samples were then stored at -80°C until they were shipped  
751 to the laboratory for analysis.

752 Calprotectin levels were measured in each fecal sample using LIAISON Calprotectin (catalog No.  
753 318960) according to the manufacturer's instructions. The levels of calprotectin in the fecal samples  
754 were used as a measure of disease activity in IBD patients.

#### 755 In vitro assays with fecal filtrates from IBD patients

756 *B. fragilis* NCTC 9343 was grown in Brain heart infusion (BHIS) supplemented with 5 $\mu$ g/L hemin  
757 (Alfa Aesar) in 1 $\mu$ N NaOH, and 2.5 $\mu$ g/L vitamin K to OD<sub>600</sub>~0.6 and centrifuged at 4,500xg for 5  
758 minutes. Bacterial pellets were washed twice with sterile PBS to remove BHIS components and then  
759 suspended with 1 mL supplemented M9 minimal media.

760 Patients' fecal samples were suspended in sterile PBS (1:5), centrifuged at 4,500xg for 15 minutes,  
761 and supernatants were collected and filtered using the Medical Millex-VV Syringe Filter Unit, 0.22  
762  $\mu$ m, PVDF membrane.

763 For the *in vitro* assay, *B. fragilis* was cultured in a mixture of M9 and patients' fecal supernatants in a  
764 1:25:25 ratio (bacteria suspended in M9: M9: fecal supernatants) and grown at 37°C in an anaerobic  
765 chamber, 85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% H<sub>2</sub> (COY). As a control, *B. fragilis* was grown in M9 and mixed with  
766 the same ratios but with sterile PBS instead of fecal filtrates. Subsequently, 200µl of each culture  
767 was collected for DNA extraction at OD<sub>600</sub>~0.6 using ZymoBIOMICS DNA Miniprep Kit [Zymo  
768 research]. qPCR analysis was done using the primers used in mice experiments, with 3 technical  
769 replicates from each individual patient:

770 B\_Frag\_upaY\_F CGCTCGGACAAAGAAGGACC

771 B\_Frag\_upaY\_R ACTTCTACCCTACGACGACGA

772 B\_Frag\_PSA\_M GGTGTTCCAAAAGACGAACGT

773 B\_Frag\_PSA\_F TGTGTAATGATAGGAGGCTAGGG

774

775 Each sample was technically repeated 3 times.

776

777 M9 Medium was prepared as follows:

778 18.7 mM NH<sub>4</sub>Cl

779 42.2 mM Na<sub>2</sub>HPO<sub>4</sub>

780 22 mM KH<sub>2</sub>PO<sub>4</sub>

781 8.5 mM NaCl

782 Supplements:

783 0.1 mM CaCl<sub>2</sub>·2H<sub>2</sub>O

784 1 mM MgSO<sub>4</sub>·7H<sub>2</sub>O

785 0.5% glucose

786 0.05% L-cysteine

787 5 g/L Hemin

788 2.5 mg/ml VitK<sub>1</sub>

789 2 mg/ml FeSO<sub>4</sub>.7H<sub>2</sub>O

790 5 ng/ml VitB12

791

792 Gut lamina propria preparation

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

797

798 Flow cytometry

799 Cell preparations for flow cytometry analysis were performed in 5 ml tubes or U shape 96 wells  
800 plates. Single cells were washed with PBS and stained for live/dead staining using 1:1,000 in PBS,  
801 Zombie fixable viability dye (Biolegend) for 10 minutes, at room temperature, and washed once with  
802 FACS buffer, by centrifuge at 300xg for 5 minutes. For FcR blocking, cells were incubated with 0.5 µg  
803 CD16/CD32 antibody for 10 minutes on ice and proceeded to further staining without a washing  
804 step. Extracellular markers were stained with the relevant antibody panels for 30 minutes on ice and  
805 washed twice with FACS buffer, by centrifuge at 300xg for 5 minutes. After the last wash, cells were  
806 fixed with Foxp3 Fixation/Permeabilization working solution (Thermo) for 16 hours at 4 °C in the  
807 dark. For Intracellular staining, cells were permeabilized using 1X Foxp3 permeabilization buffer  
808 (Thermo) according to the manufacturer protocol. For intracellular blocking, 2 µl of 2% rat serum  
809 (Stemcell technologies) was added to each well for 15 minutes at room temperature and proceeded  
810 to further staining without a washing step.

811 To quantify the percentage of *B. fragilis* from monocolonized mice with PSA on their surface, feces  
812 from monocolonized mice with *B. fragilis* with or without phage were suspended 1:10 in ice cold  
813 PBS (mg/µl) and centrifuged at 300xg for 5 minutes, 4°C. Supernatants were separated from pellets  
814 and further centrifuged at 4,500xg for 5 minutes, 4°C. Bacterial pellets were resuspended in an ice  
815 cold FACS buffer, 1:10 from initial PBS suspension. 100 µl of resuspended bacteria were incubated  
816 with 1:1,000 rabbit antibodies (Antibodies preparation previously described<sup>1</sup>) against *B. fragilis* PSA

817 for 30 minutes at 4°C. Bacteria were washed twice using an ice cold FACS buffer by centrifuge at  
818 4,500xg for 5 minutes and then incubated with a donkey anti rabbit fluorophore conjugated  
819 secondary antibody. After staining steps, the bacteria were washed twice with an ice cold FACS  
820 buffer and finally resuspended in 500 µl ice cold PBS plus 1:1,000 Hoechst dye and analyzed by flow  
821 cytometry using FSC and SSC thresholds of 1,000, and logarithmic scale. Gating strategy is detailed in  
822 **Figure S5**. Antibody specificity for PSA was verified using PSA mutant *B. fragilis*. Non-specific binding  
823 controls were included.

824

## 825 Funding

826 This work was supported by the Technion Institute of Technology, ‘Keren Hanasi’, Cathedra, the  
827 Rappaport Technion Integrated Cancer Center, the Alon Fellowship for Outstanding Young  
828 Researchers, the Israeli Science Foundation (grant 1571/17 and 3165/20), Israel Cancer Research  
829 Fund Research Career Development Award, the Seerave Foundation, the Canadian Institute for  
830 Advanced Research (Azrieli Global Scholars; grant FL-000969), Human Frontier Science Program  
831 Career Development Award (grant CDA00025/2019-C), the Gutwirth foundation award and the  
832 European Union (ERC, ExtractABact, 101078712). Views and opinions expressed are, however, those  
833 of the author(s) only and do not necessarily reflect those of the European Union or the European  
834 Research Council Executive Agency. Neither the European Union nor the granting authority can be  
835 held responsible for them. NGZ is an Azrieli Global Scholar at the Canadian Institute for Advanced  
836 Research, and a Horev Fellow (Taub Foundation). MJC and LC are supported by the Duchossois  
837 Family Institute. SC is supported by the Gutwirth Excellence Scholarship and by Teva Pharmaceutical  
838 Industries as part of the Israeli National Forum for BioInnovators (NFBI). HH is supported by Leonard  
839 and Diane Sherman Interdisciplinary Graduate School fellowship, Wjuniski Fellowship Fund for the  
840 MD/PhD Medical Scientist Program, and by VATAT fellowship for outstanding doctoral students from  
841 the Arab community fellowship.

842

## 843 Data availability

844 Sequencing data has been deposited in the NCBI sequence read archive (SRA) under the BioProject  
845 accession number: PRJNA916364.

846

847 **Acknowledgements**

848 We would like to thank the Geva-Zatorsky lab members for fruitful discussions and contributions.  
849 Thanking Zachary Merenstein and Neekoo Farahmandpour for helping in establishing the qPCR  
850 method; Dr. Svetlana Friedman, Ran Tahan and Prof. Debbie Lindell for advising us on the  
851 bacteriophage characterization and Dr. Carolina Tropini for insightful discussions. We thank Drs.  
852 Amir Grau, Ofer Shenker, and the biomedical core facility at the Technion Rappaport faculty of  
853 Medicine for their help with flow cytometry; Biotax labs for metagenomics sequencing, and Prof.  
854 Ogata and Dr. Nishiyama from the Institute for Chemical Research, Kyoto University, for helping in  
855 accessing their data. Figure 3A was created with BioRender.com.

856

857 **Author contributions**

858 SC HH RZ DKK TG IS and NGZ conceived and designed the project. SC HH RZ DKK and TG performed  
859 and analyzed the experiments. SC, IS and NGZ planned the computational and analytical aspects; SC  
860 and IS performed the computational analyses. MJC and LC contributed to the computational  
861 analysis, phage sequencing and analysis, and interpretation of the results. SP and YC designed, led,  
862 and executed the clinical study. KJ contributed to the phage-bacteria-host study and to the  
863 bacteriophage morphological characterization. JJ isolated the bacteriophage. NM RN, and NBA  
864 helped with the experiments. SC HH RZ TG IS LC and NGZ wrote the manuscript. SC RZ HH and DKK  
865 contributed equally to the study. All authors read and approved the manuscript. NGZ conceived and  
866 planned the study, supervised it, interpreted the experiments, and wrote the manuscript.

867

868

869

870

## 871 References

- 872 1. Moxon R, Bayliss C, Hood D. Bacterial contingency loci: The role of simple sequence DNA  
873 repeats in bacterial adaptation. *Annu Rev Genet.* 2006;40:307-333.  
874 doi:10.1146/annurev.genet.40.110405.090442
- 875 2. Phillips ZN, Tram G, Seib KL, Attack JM. Phase-variable bacterial loci: how bacteria gamble to  
876 maximise fitness in changing environments. *Biochem Soc Trans.* 2019;47(4):1131-1141.  
877 doi:10.1042/BST20180633
- 878 3. West PT, Chanin RB, Bhatt AS. From genome structure to function: insights into structural  
879 variation in microbiology. *Curr Opin Microbiol.* 2022;69:102192.  
880 doi:10.1016/J.MIB.2022.102192
- 881 4. Goldberg A, Fridman O, Ronin I, Balaban NQ. Systematic identification and quantification of  
882 phase variation in commensal and pathogenic *Escherichia coli*. *Genome Med.* 2014;6(11):112.  
883 doi:10.1186/s13073-014-0112-4
- 884 5. Krinos CM, Coyne MJ, Weinacht KG, Tzianabos AO, Kasper DL, Comstock LE. Extensive surface  
885 diversity of a commensal microorganism by multiple DNA inversions. *Nature.*  
886 2001;414(6863):555-558. doi:10.1038/35107092
- 887 6. Van Der Woude MW, Bäumlér AJ. Phase and Antigenic Variation in Bacteria. *Clin Microbiol*  
888 *Rev.* 2004;17(3):581. doi:10.1128/CMR.17.3.581-611.2004
- 889 7. Ben-Assa N, Coyne MJ, Fomenkov A, et al. Analysis of a phase-variable restriction  
890 modification system of the human gut symbiont *Bacteroides fragilis*. *Nucleic Acids Res.*  
891 2020;48(19):11040-11053. doi:10.1093/nar/gkaa824
- 892 8. Porter NT, Hryckowian AJ, Merrill BD, et al. Phase-variable capsular polysaccharides and  
893 lipoproteins modify bacteriophage susceptibility in *Bacteroides thetaiotaomicron*. *Nat*  
894 *Microbiol.* 2020;5(9):1170-1181. doi:10.1038/s41564-020-0746-5
- 895 9. Jiang X, Brantley Hall A, Arthur TD, et al. Invertible promoters mediate bacterial phase  
896 variation, antibiotic resistance, and host adaptation in the gut. *Science (1979).*  
897 2019;363(6423):181-187. doi:10.1126/science.aau5238
- 898 10. Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal  
899 inflammatory disease. *Nature.* 2008;453(7195):620-625. doi:10.1038/nature07008
- 900 11. Surana NK, Kasper DL. The yin yang of bacterial polysaccharides: lessons learned from *B.*  
901 *fragilis* PSA. *Immunol Rev.* 2012;245(1):13-26. doi:10.1111/j.1600-065X.2011.01075.x
- 902 12. Dasgupta S, Erturk-Hasdemir D, Ochoa-Reparaz J, Reinecker HC, Kasper DL. Plasmacytoid  
903 dendritic cells mediate anti-inflammatory responses to a gut commensal molecule via both  
904 innate and adaptive mechanisms. *Cell Host Microbe.* 2014;15(4):413-423.  
905 doi:10.1016/J.CHOM.2014.03.006
- 906 13. Round JL, Mazmanian SK. Inducible Foxp3+ regulatory T-cell development by a commensal  
907 bacterium of the intestinal microbiota. *Proc Natl Acad Sci U S A.* 2010;107(27):12204-12209.  
908 doi:10.1073/PNAS.0909122107



- 909 14. Chang YC, Ching YH, Chiu CC, et al. TLR2 and interleukin-10 are involved in *Bacteroides*  
910 *fragilis*-mediated prevention of DSS-induced colitis in gnotobiotic mice. *PLoS One*. 2017;12(7).  
911 doi:10.1371/JOURNAL.PONE.0180025
- 912 15. Campbell DE, Ly LK, Ridlon JM, Hsiao A, Whitaker RJ, Degnan PH. Infection with *Bacteroides*  
913 Phage BV01 Alters the Host Transcriptome and Bile Acid Metabolism in a Common Human  
914 Gut Microbe. *Cell Rep*. 2020;32(11):108142.  
915 doi:10.1016/J.CELREP.2020.108142/ATTACHMENT/3C255998-9456-4BE2-B212-  
916 43C137BAD94A/MMC7.XLSX
- 917 16. Hryckowian AJ, Merrill BD, Porter NT, et al. *Bacteroides thetaiotaomicron*-Infecting  
918 Bacteriophage Isolates Inform Sequence-Based Host Range Predictions. *Cell Host Microbe*.  
919 2020;28(3):371-379.e5. doi:10.1016/J.CHOM.2020.06.011
- 920 17. Shkoporov AN, Khokhlova E V., Stephens N, et al. Long-term persistence of crAss-like phage  
921 crAss001 is associated with phase variation in *Bacteroides intestinalis*. *BMC Biol*. 2021;19(1).  
922 doi:10.1186/S12915-021-01084-3
- 923 18. Ramos GP, Papadakis KA. Mechanisms of Disease: Inflammatory Bowel Diseases. *Mayo Clin*  
924 *Proc*. 2019;94(1):155-165. doi:10.1016/J.MAYOCP.2018.09.013
- 925 19. Guan Q. A Comprehensive Review and Update on the Pathogenesis of Inflammatory Bowel  
926 Disease. *J Immunol Res*. 2019;2019. doi:10.1155/2019/7247238
- 927 20. Corridoni D, Arseneau KO, Cominelli F. Inflammatory bowel disease. *Immunol Lett*.  
928 2014;161(2):231-235. doi:10.1016/J.IMLET.2014.04.004
- 929 21. Ha CWY, Martin A, Sepich-Poore GD, et al. Translocation of Viable Gut Microbiota to  
930 Mesenteric Adipose Drives Formation of Creeping Fat in Humans. *Cell*. 2020;183(3):666-  
931 683.e17. doi:10.1016/J.CELL.2020.09.009
- 932 22. Lee M, Chang EB. Inflammatory Bowel Diseases (IBD) and the Microbiome—Searching the  
933 Crime Scene for Clues. *Gastroenterology*. 2021;160(2):524-537.  
934 doi:10.1053/J.GASTRO.2020.09.056
- 935 23. Bolam DN, Koropatkin NM. Glycan recognition by the *Bacteroidetes* Sus-like systems. *Curr*  
936 *Opin Struct Biol*. 2012;22(5):563-569. doi:10.1016/J.SBI.2012.06.006
- 937 24. Kulagina E V., Efimov BA, Maximov PY, Kafarskaia LI, Chaplin A V., Shkoporov AN. Species  
938 composition of *Bacteroidales* order bacteria in the feces of healthy people of various ages.  
939 *Biosci Biotechnol Biochem*. 2012;76(1):169-171. doi:10.1271/BBB.110434
- 940 25. Kraal L, Abubucker S, Kota K, Fischbach MA, Mitreva M. The Prevalence of Species and Strains  
941 in the Human Microbiome: A Resource for Experimental Efforts. *PLoS One*. 2014;9(5):e97279.  
942 doi:10.1371/JOURNAL.PONE.0097279
- 943 26. Nomura K, Ishikawa D, Okahara K, et al. *Bacteroidetes* species are correlated with disease  
944 activity in ulcerative colitis. *J Clin Med*. 2021;10(8):1749. doi:10.3390/JCM10081749/S1
- 945 27. Proctor LM, Creasy HH, Fettweis JM, et al. The Integrative Human Microbiome Project.  
946 *Nature* 2019 569:7758. 2019;569(7758):641-648. doi:10.1038/s41586-019-1238-8

- 947 28. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The Human  
948 Microbiome Project. *Nature* 2007 449:7164. 2007;449(7164):804-810.  
949 doi:10.1038/nature06244
- 950 29. Qin J, Li R, Raes J, et al. A human gut microbial gene catalogue established by metagenomic  
951 sequencing. *Nature* 2010 464:7285. 2010;464(7285):59-65. doi:10.1038/nature08821
- 952 30. Imhann F, Van Der Velde KJ, Barbieri R, et al. The 1000IBD project: Multi-omics data of 1000  
953 inflammatory bowel disease patients; Data release 1. *BMC Gastroenterol.* 2019;19(1):1-10.  
954 doi:10.1186/s12876-018-0917-5/TABLES/2
- 955 31. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naïve microbiome in new-onset  
956 Crohn's disease. *Cell Host Microbe.* 2014;15(3):382-392. doi:10.1016/j.CHOM.2014.02.005
- 957 32. Lewis JD, Chen EZ, Baldassano RN, et al. Inflammation, Antibiotics, and Diet as Environmental  
958 Stressors of the Gut Microbiome in Pediatric Crohn's Disease. *Cell Host Microbe.*  
959 2015;18(4):489-500. doi:10.1016/j.CHOM.2015.09.008
- 960 33. Chassaing B, Aitken JD, Malleshappa M, Vijay-Kumar M. Dextran Sulfate Sodium (DSS)-  
961 Induced Colitis in Mice. *Current protocols in immunology / edited by John E Coligan . [et al].*  
962 2014;104(SUPPL.104):Unit. doi:10.1002/0471142735.IM1525S104
- 963 34. Lepage P, Colombet J, Marteau P, Sime-Ngando T, Doré J, Leclerc M. Dysbiosis in  
964 inflammatory bowel disease: a role for bacteriophages? *Gut.* 2008;57(3):424-425.  
965 doi:10.1136/GUT.2007.134668
- 966 35. Norman JM, Handley SA, Baldrige MT, et al. Disease-specific alterations in the enteric  
967 virome in inflammatory bowel disease. *Cell.* 2015;160(3):447-460.  
968 doi:10.1016/j.cell.2015.01.002
- 969 36. Wang W, Jovel J, Halloran B, et al. Metagenomic analysis of microbiome in colon tissue from  
970 subjects with inflammatory bowel diseases reveals interplay of viruses and bacteria. *Inflamm*  
971 *Bowel Dis.* 2015;21(6):1419-1427. doi:10.1097/MIB.0000000000000344
- 972 37. Adiliaghdam F, Amatullah H, Digumarthi S, et al. Human enteric viruses autonomously shape  
973 inflammatory bowel disease phenotype through divergent innate immunomodulation. *Sci*  
974 *Immunol.* 2022;7(70):eabn6660. doi:10.1126/SCIIMMUNOL.ABN6660
- 975 38. Qv L, Mao S, Li Y, Zhang J, Li L. Roles of Gut Bacteriophages in the Pathogenesis and  
976 Treatment of Inflammatory Bowel Disease. *Front Cell Infect Microbiol.* 2021;11.  
977 doi:10.3389/FCIMB.2021.755650
- 978 39. Nishiyama H, Endo H, Blanc-Mathieu R, Ogata H. Ecological Structuring of Temperate  
979 Bacteriophages in the Inflammatory Bowel Disease-Affected Gut. *Microorganisms.*  
980 2020;8(11):1-15. doi:10.3390/MICROORGANISMS8111663
- 981 40. Ogilvie LA, Caplin J, Dedi C, et al. Comparative (meta)genomic analysis and ecological profiling  
982 of human gut-specific bacteriophage  $\phi$ B124-14. *PLoS One.* 2012;7(4).  
983 doi:10.1371/JOURNAL.PONE.0035053
- 984 41. Puig M, Gironés R. Genomic structure of phage B40-8 of *Bacteroides fragilis*. *Microbiology*  
985 *(Reading).* 1999;145 ( Pt 7)(7):1661-1670. doi:10.1099/13500872-145-7-1661

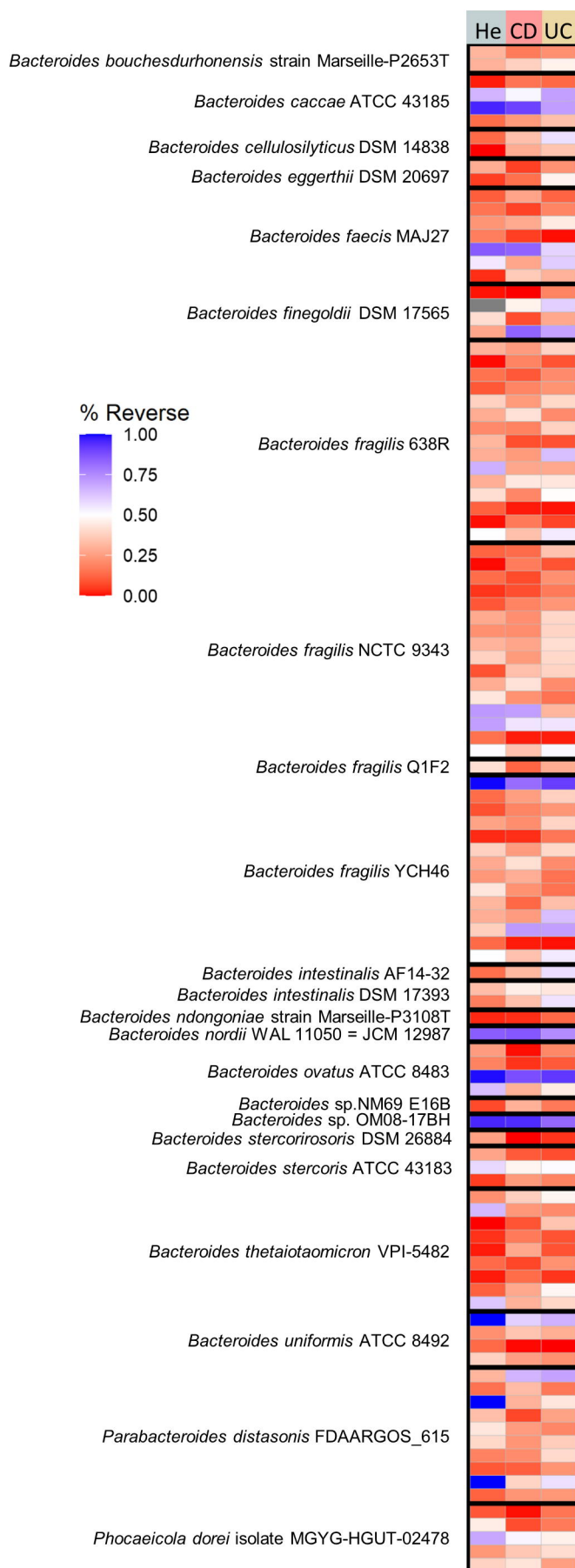
- 986 42. Guerin E, Shkoporov AN, Stockdale SR, et al. Isolation and characterisation of  $\Phi$ crAss002, a  
987 crAss-like phage from the human gut that infects *Bacteroides xyloxylophilus*. *Microbiome*.  
988 2021;9(1):1-21. doi:10.1186/S40168-021-01036-7/FIGURES/7
- 989 43. Chatzidaki-Livanis M, Weinacht KG, Comstock LE. Trans locus inhibitors limit concomitant  
990 polysaccharide synthesis in the human gut symbiont *Bacteroides fragilis*. *Proc Natl Acad Sci U*  
991 *S A*. 2010;107(26):11976-11980. doi:10.1073/PNAS.1005039107
- 992 44. Coyne MJ, Tzianabos AO, Mallory BC, Carey VJ, Kasper DL, Comstock LE. Polysaccharide  
993 biosynthesis locus required for virulence of *Bacteroides fragilis*. *Infect Immun*.  
994 2001;69(7):4342-4350. doi:10.1128/IAI.69.7.4342-4350.2001
- 995 45. Erturk-Hasdemir D, Oh SF, Okan NA, et al. Symbionts exploit complex signaling to educate the  
996 immune system. *Proc Natl Acad Sci U S A*. 2019;116(52):26157-26166.  
997 doi:10.1073/PNAS.1915978116
- 998 46. Johnson JL, Jones MB, Cobb BA. Polysaccharide-experienced effector T cells induce IL-10 in  
999 FoxP3+ regulatory T cells to prevent pulmonary inflammation. *Glycobiology*. 2018;28(1):50-  
1000 58. doi:10.1093/GLYCOB/CWX093
- 1001 47. Telesford KM, Yan W, Ochoa-Reparaz J, et al. A commensal symbiotic factor derived from  
1002 *Bacteroides fragilis* promotes human CD39(+)Foxp3(+) T cells and Treg function. *Gut*  
1003 *Microbes*. 2015;6(4):234-242. doi:10.1080/19490976.2015.1056973
- 1004 48. Pedersen TK, Brown EM, Plichta DR, et al. The CD4+ T cell response to a commensal-derived  
1005 epitope transitions from a tolerant to an inflammatory state in Crohn's disease. *Immunity*.  
1006 2022;55(10):1909-1923.e6. doi:10.1016/J.IMMUNI.2022.08.016
- 1007 49. Blandford LE, Johnston EL, Sanderson JD, Wade WG, Lax AJ. Promoter orientation of the  
1008 immunomodulatory *Bacteroides fragilis* capsular polysaccharide A (PSA) is off in individuals  
1009 with inflammatory bowel disease (IBD). *Gut Microbes*. Published online February 7, 2019:1-9.  
1010 doi:10.1080/19490976.2018.1560755
- 1011 50. Lloyd-Price J, Arze, C Ananthakrishnan, AN Schirmer M, Avila-Pacheco J, Poon T, Andrews, E  
1012 Ajami N, Bonham, KS Brislawn C. Multi-omics of the gut microbial ecosystem in inflammatory  
1013 bowel diseases. *Nature*. 2019;569(7758):655-662. Accessed March 18, 2021.  
1014 <https://www.nature.com/articles/s41586-019-1237-9>
- 1015 51. Federici S, Kredo-Russo S, Valdés-Mas R, et al. Targeted suppression of human IBD-associated  
1016 gut microbiota commensals by phage consortia for treatment of intestinal inflammation. *Cell*.  
1017 2022;185(16):2879-2898.e24. doi:10.1016/J.CELL.2022.07.003
- 1018 52. Duerkop BA, Kleiner M, Paez-Espino D, et al. Murine colitis reveals a disease-associated  
1019 bacteriophage community. *Nat Microbiol*. 2018;3(9):1023. doi:10.1038/S41564-018-0210-Y
- 1020 53. Wallace KL, Zheng LB, Kanazawa Y, Shih DQ. Immunopathology of inflammatory bowel  
1021 disease. *World Journal of Gastroenterology*: *WJG*. 2014;20(1):6. doi:10.3748/WJG.V20.I1.6
- 1022 54. Nugent SG, Kumar D, Rampton DS, Evans DF. Intestinal luminal pH in inflammatory bowel  
1023 disease: possible determinants and implications for therapy with aminosaliculates and other  
1024 drugs. *Gut*. 2001;48(4):571-577. doi:10.1136/GUT.48.4.571

- 1025 55. Vertzoni M, Koulouri C, Poulou A, Goumas K, Reppas C. Exploring the impact of Crohn's  
1026 disease on the intragastric environment of fasted adults. *ADMET DMPK*. 2020;8(2):122.  
1027 doi:10.5599/ADMET.846
- 1028 56. Barkas F, Liberopoulos E, Kei A, Elisaf M. Electrolyte and acid-base disorders in inflammatory  
1029 bowel disease. *Annals of Gastroenterology: Quarterly Publication of the Hellenic Society of*  
1030 *Gastroenterology*. 2013;26(1):23. Accessed February 11, 2023. /pmc/articles/PMC3959504/
- 1031 57. Tian T, Wang Z, Zhang J. Pathomechanisms of Oxidative Stress in Inflammatory Bowel Disease  
1032 and Potential Antioxidant Therapies. *Oxid Med Cell Longev*. 2017;2017.  
1033 doi:10.1155/2017/4535194
- 1034 58. Bourgonje AR, Feelisch M, Faber KN, Pasch A, Dijkstra G, van Goor H. Oxidative Stress and  
1035 Redox-Modulating Therapeutics in Inflammatory Bowel Disease. *Trends Mol Med*.  
1036 2020;26(11):1034-1046. doi:10.1016/J.MOLMED.2020.06.006
- 1037 59. Tropini C, Moss EL, Merrill BD, et al. Transient osmotic perturbation causes long-term  
1038 alteration to the gut microbiota. *Cell*. 2018;173(7):1742. doi:10.1016/J.CELL.2018.05.008
- 1039 60. Milman O, Yelin I, Kishony R. Systematic identification of gene-altering programmed  
1040 inversions across the bacterial domain. *Nucleic Acids Res*. 2023;51(2):553-573.  
1041 doi:10.1093/NAR/GKAC1166
- 1042 61. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative  
1043 PCR and the 2<sup>-Delta Delta C(T)</sup> Method. *Methods*. 2001;25(4):402-408.  
1044 doi:10.1006/METH.2001.1262
- 1045 62. Bustin SA, Benes V, Garson JA, et al. The MIQE Guidelines: Minimum Information for  
1046 Publication of Quantitative Real-Time PCR Experiments. *Clin Chem*. 2009;55(4):611-622.  
1047 doi:10.1373/CLINCHEM.2008.112797
- 1048 63. Blanco-Míguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic  
1049 taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nature Biotechnology*  
1050 2023. Published online February 23, 2023:1-12. doi:10.1038/s41587-023-01688-w
- 1051 64. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: A web-based tool  
1052 for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res*.  
1053 2017;45(W1):W180-W188. doi:10.1093/nar/gkx295
- 1054 65. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and  
1055 Graphics of Microbiome Census Data. *PLoS One*. 2013;8(4):e61217.  
1056 doi:10.1371/JOURNAL.PONE.0061217
- 1057 66. Oksanen J, Blanchet FG, Friendly M, et al. vegan: Community Ecology Package, vegan:  
1058 Community Ecology Package. R package version 2.5-6. Published 2020. Accessed December  
1059 29, 2020. <https://CRAN.R-project.org/package=vegan>
- 1060 67. Himmel M, Rahnavard A, McIver L. Maaslin2: Maaslin2. R package version 1.0.0. Published  
1061 online 2019.
- 1062 68. Wickham H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*. Published  
1063 online 2016.

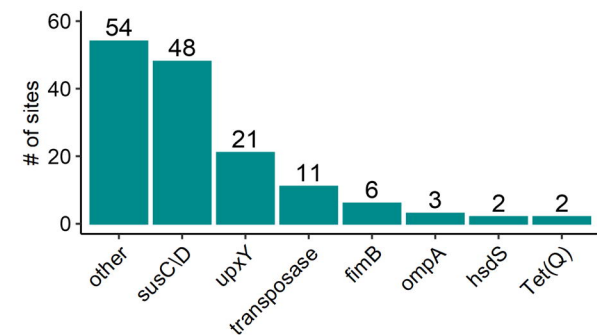
- 1064 69. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-  
1065 seq data with DESeq2. *Genome Biol.* 2014;15(12). doi:10.1186/s13059-014-0550-8
- 1066 70. Pasolli E, Asnicar F, Manara S, et al. Extensive Unexplored Human Microbiome Diversity  
1067 Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and  
1068 Lifestyle. *Cell.* 2019;176(3):649-662.e20.  
1069 doi:10.1016/J.CELL.2019.01.001/ATTACHMENT/31D7FB71-DCA8-4659-BC04-  
1070 FB25BCC69750/MMC7.XLSX
- 1071 71. Almeida A, Mitchell AL, Boland M, et al. A new genomic blueprint of the human gut  
1072 microbiota. *Nature* 2019 568:7753. 2019;568(7753):499-504. doi:10.1038/s41586-019-0965-  
1073 1
- 1074 72. Forster SC, Kumar N, Anonye BO, et al. A human gut bacterial genome and culture collection  
1075 for improved metagenomic analyses. *Nature Biotechnology* 2019 37:2. 2019;37(2):186-192.  
1076 doi:10.1038/s41587-018-0009-7
- 1077 73. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification  
1078 with the genome taxonomy database. *Bioinformatics.* 2022;38(23):5315-5316.  
1079 doi:10.1093/BIOINFORMATICS/BTAC672
- 1080 74. Adams MH. Bacteriophages. *Bacteriophages.* Published online 1959.
- 1081 75. García-Aljaro C, Muniesa M, Jofre J. Isolation of Bacteriophages of the Anaerobic Bacteria  
1082 Bacteroides. *Methods Mol Biol.* 2018;1693:11-22. doi:10.1007/978-1-4939-7395-8\_2
- 1083 76. *ISO 10705-4:2001, Water Quality - Detection and Enumeration of Bacteriophages - Part 4:*  
1084 *Enumeration of Bacteriophages Infecting Bacteroides Fragilis: ISO/TC 147/SC 4: Amazon.Com:*  
1085 *Books.* International Organisation for Standardisation; 2001.
- 1086 77. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068-  
1087 2069. doi:10.1093/BIOINFORMATICS/BTU153
- 1088 78. El-Gebali S, Mistry J, Bateman A, et al. The Pfam protein families database in 2019. *Nucleic*  
1089 *Acids Res.* 2019;47(D1):D427-D432. doi:10.1093/NAR/GKY995
- 1090 79. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. TIGRFAMs and Genome  
1091 Properties in 2013. *Nucleic Acids Res.* 2013;41(D1):D387-D395. doi:10.1093/NAR/GKS1234
- 1092 80. Galperin MY, Wolf YI, Makarova KS, Alvarez RV, Landsman D, Koonin E V. COG database  
1093 update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic*  
1094 *Acids Res.* 2021;49(D1):D274-D281. doi:10.1093/NAR/GKAA1018
- 1095 81. Tatusov RL, Koonin E V., Lipman DJ. A genomic perspective on protein families. *Science.*  
1096 1997;278(5338):631-637. doi:10.1126/SCIENCE.278.5338.631
- 1097 82. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:  
1098 improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772-780.  
1099 doi:10.1093/MOLBEV/MST010
- 1100 83. Yu G. Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinformatics.*  
1101 2020;69(1):e96. doi:10.1002/CPBI.96

1102

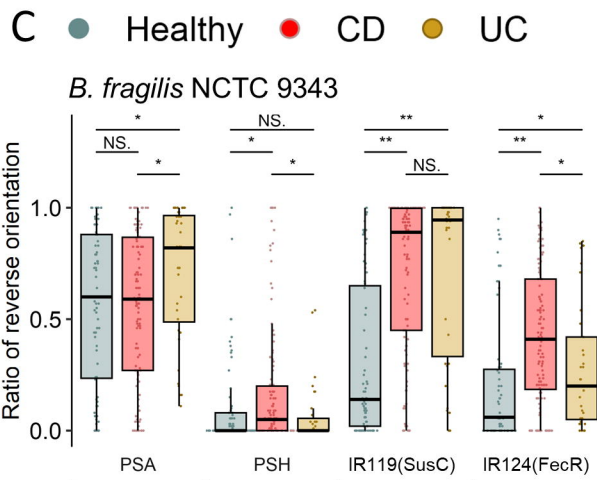
A



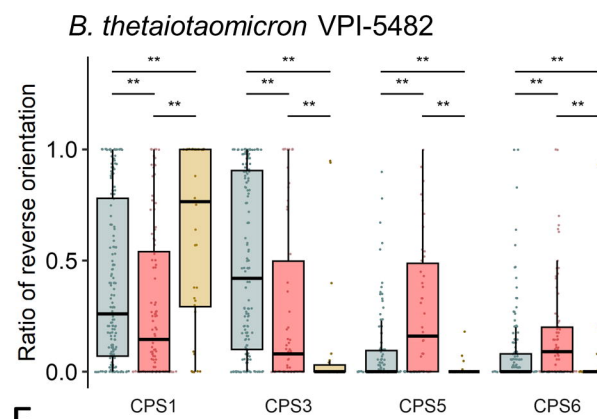
B



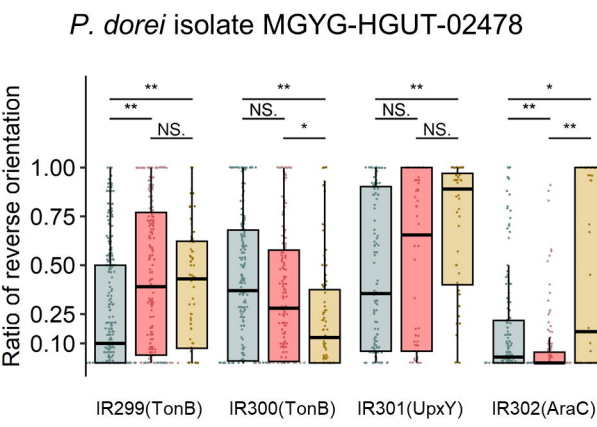
C

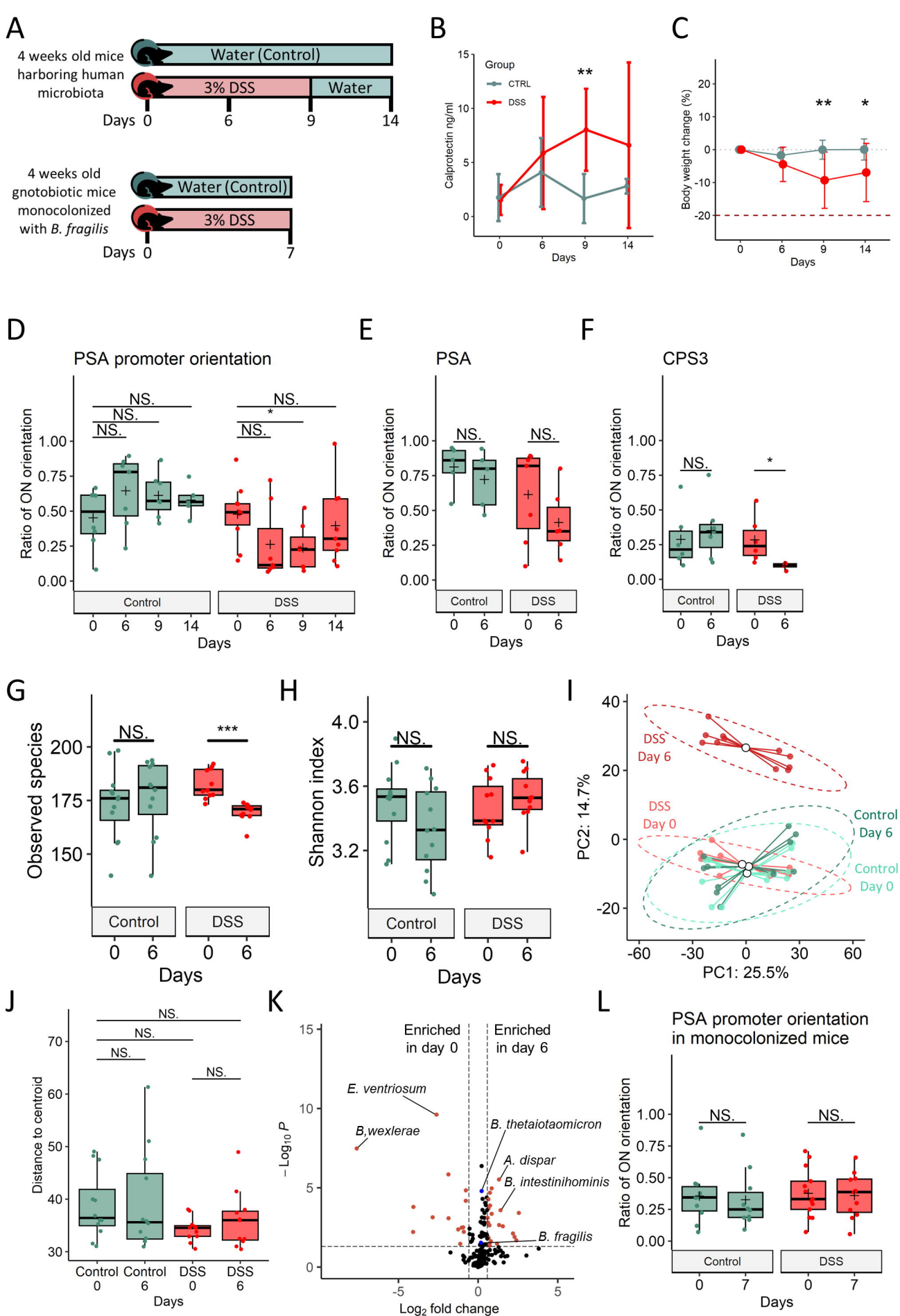


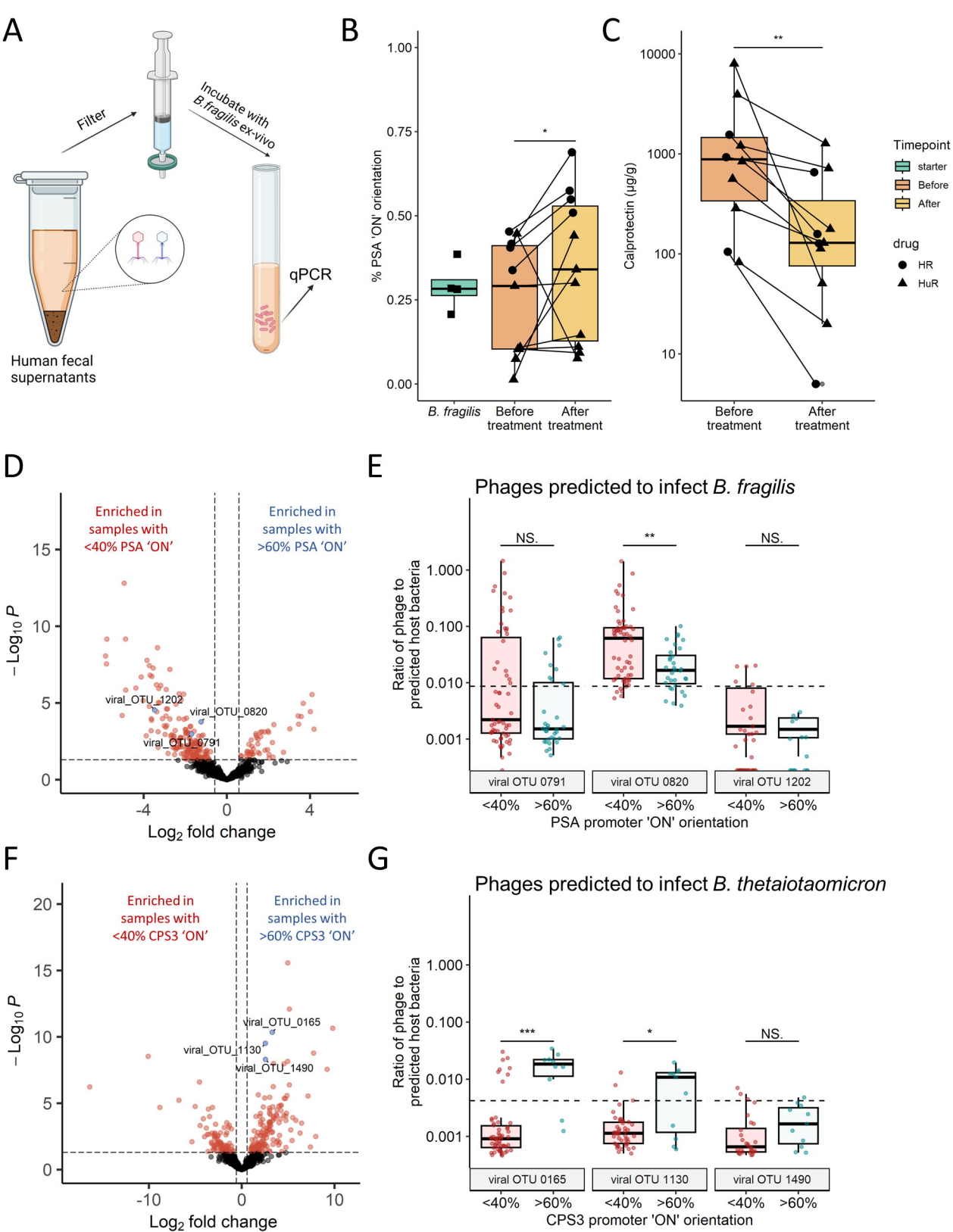
D



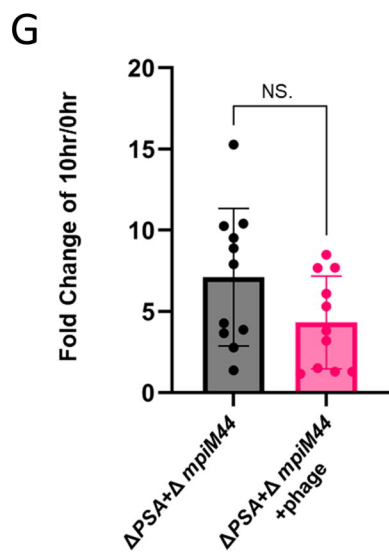
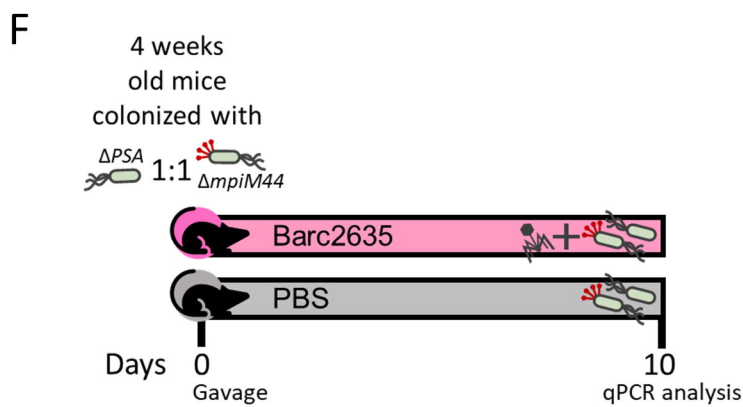
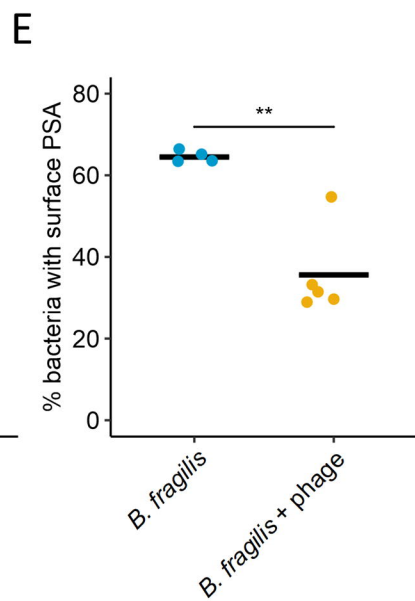
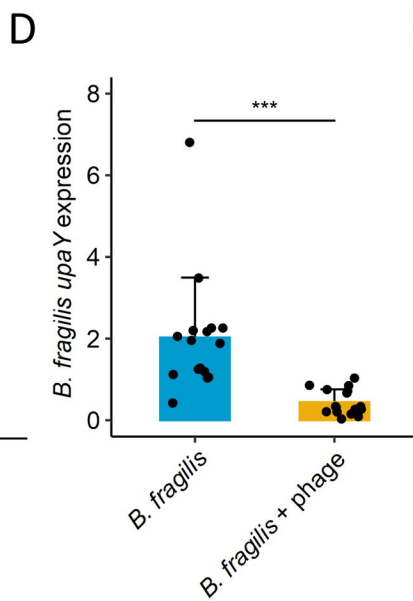
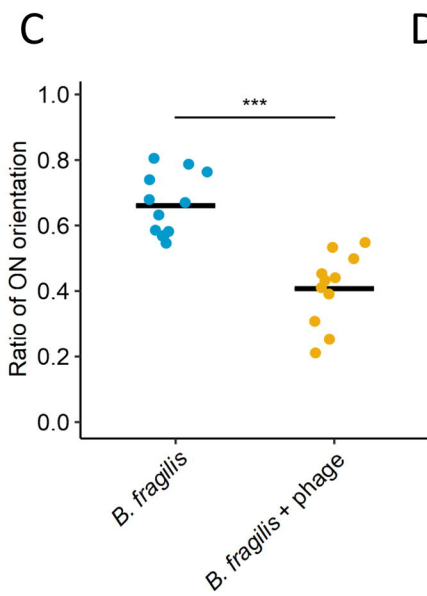
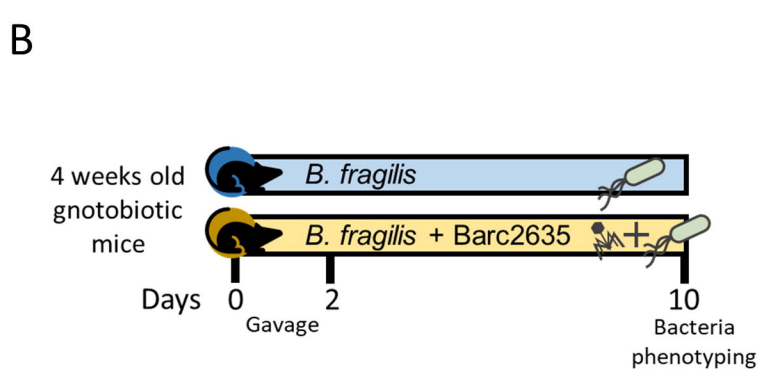
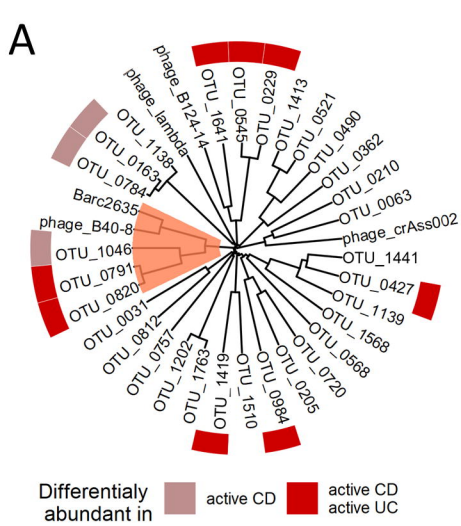
E



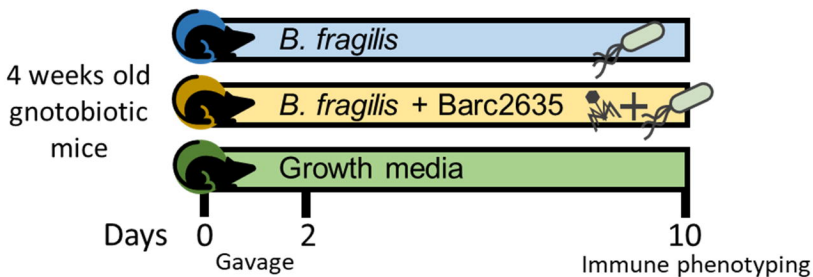




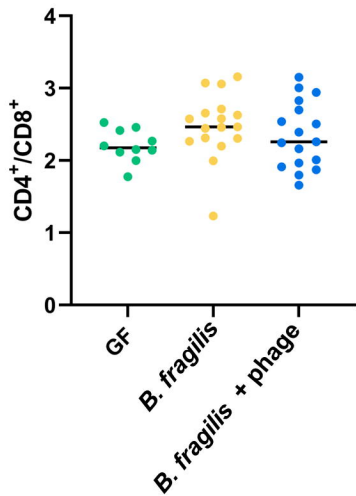




A



B



C

