Multiple phase-variable mechanisms, including capsular polysaccharides, modify bacteriophage susceptibility in *Bacteroides thetaiotaomicron*

*Nathan T. Porter[1], *Andrew J. Hryckowian[2], Bryan D. Merrill[2], Jackson O. Gardner[2], Shaleni Singh[1], Justin L. Sonnenburg[2], and Eric C. Martens[1]

[1]*Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI 48109, USA*

[2]*Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305*

*These authors contributed equally to this work

Correspondence to: emartens@umich.edu

Abstract

A variety of cell surface structures, including capsular polysaccharides (CPS), dictate the interactions between bacteria and elements of their environment including their viruses (bacteriophages). Members of the prominent human gut Bacteroidetes characteristically produce several phase-variable CPS, which have been demonstrated as key determinants in interacting with the host immune system. However, the contributions of Bacteroidetes CPS to bacteriophage interactions are unknown. We used engineered strains of the human symbiont *Bacteroides thetaiotaomicron*, which differ only in the CPS they express, to isolate bacteriophages from two locations in the United States. Testing each of 71 bacteriophages against a panel of strains that express wild-type phase-variable CPS, one of eight different single CPS, or no CPS at all, revealed that each infects only a subset of strains. Deletion of infection-permissive CPS from *B. thetaiotaomicron* was sufficient to abolish infection for several individual bacteriophages. Likewise, infection of wild-type *B. thetaiotaomicron* with one bacteriophage from this collection selected for a cell population expressing non-permissive CPS. Surprisingly, acapsular *B. thetaiotaomicron* also escapes complete killing by this bacteriophage, but surviving bacteria increased expression of a family of 9 phase-variable lipoproteins. When constitutively expressed, one of these lipoproteins enhances resistance to this bacteriophage. Our results reveal distinct roles for *Bacteroides* CPS in mediating bacteriophage susceptibility. Beneath this vanguard protective layer, additional mechanisms exist to protect these bacteria from being eradicated by bacteriophage. Given the diversity of CPS and other phase-variable surface structures encoded by gut-dwelling Bacteroidetes, our results provide important insight into the adaptations that allow these bacteria to persist despite bacteriophage predation and hold important implications for using bacteriophages therapeutically to target gut symbionts.

41

42 **Introduction**

43 The community of cellular microorganisms in the human intestinal tract is dominated by
44 a diverse population of bacteria, with hundreds of different species and even more strains
45 typically present within one person[1]. In the face of frequent perturbations from a variety of
46 environmental pressures, including diet changes, antibiotics and host immunity, the relative
47 abundances of different bacteria may vary greatly within an individual over short time periods.
48 However, the microbiota members within a host generally remain stable over long time
49 periods[2,3], suggesting that individual bacteria have evolved strategies that enable them to be
50 resilient in the face of challenges. One potential adaptive mechanism that has clearly been
51 diversified in gut resident Bacteroidetes is the ability of individual strains to alternately produce
52 several different capsular polysaccharides (CPS), which are extracellular matrix components
53 encoded by *cps* biosynthetic loci. Many *cps* loci in the Bacteroidetes are under the control of
54 phase variable promoters, allowing for substantial phenotypic heterogeneity within an isogenic
55 population. Furthermore, there is much broader diversification of *cps* loci among different strains
56 within a species[4-7]. While a few studies have shown that these CPS take part in evasion or
57 modulation of host immunity[4,8-11], the sheer diversity and number of CPS synthesis loci in the
58 Bacteroidetes suggests that they could also fill other fundamental roles[4,6,12,13].

59 Bacterial viruses or bacteriophages (herein, phages), like the bacteria on which they prey,
60 vary greatly across individual gut microbiomes and are even responsive to host dietary changes
61 or disease states[14-17]. Compared to its bacterial counterpart, far less is understood about the
62 human gut virome, especially the mechanisms governing phage-bacteria interactions.
63 Specifically, while phages that target several species of *Bacteroides* (the prominent genus of the

Bacteroidetes phylum in humans) have been shown to exhibit species- or strain-specificity[18-21], little is known about the molecular interactions that drive bacterial susceptibility[22] or the mechanisms by which these bacteria persist despite an abundance of phages in the gastrointestinal tract.

Given the observations that *Bacteroides* CPS are extremely variable, even within members of a single species[4,12], and employ complex regulatory mechanisms that diversify expression in members of a population[4,23], CPS are ideal candidates for modulating *Bacteroides*-phage interactions. Thus, we sought to test the hypothesis that CPS are direct positive or negative mediators of *Bacteroides*-phage interactions. To accomplish this, we employed a panel of engineered strains of the model symbiont *Bacteroides thetaiotaomicron* that each constitutively expresses a different single CPS or none at all. While our results clearly support the conclusion that specific CPS may either block or be required for phage infection, they also reveal that *B. thetaiotaomicron* possesses additional phage-evasion strategies that function in the absence of CPS. For one phage, CPS-independent survival involves altered expression of nutrient receptors and increased expression of phase-variable surface lipoproteins by the surviving bacteria. One hypothesis is that these latter functions also encode active resistance mechanisms. This idea is supported by locking on expression of one of these lipoproteins, which we show to be an additional determinant of phage tropism, conferring increased resistance to this phage. Our results provide a mechanistic glimpse into the intricacy of bacterial-phage interactions that exist in the human gut and provide a foundation for future work aimed at leveraging these interactions as a facet of targeted manipulation of the gut microbiome.

Results

Individual bacteriophages target *B. thetaiotaomicron* in a CPS-dependent fashion

To test the hypothesis that variable *Bacteroides* surface CPS mediate interactions with phages, we isolated phages that infect *B. thetaiotaomicron* VPI-5482 (ATCC 29148). To maximize our chances of collecting phages that could differ in their interactions with CPS, we used the wild-type strain of *B. thetaiotaomicron* that expresses 8 different, phase-variable CPS, each encoded by a different multi-gene *cps* locus; an isogenic panel of 8 single CPS-expressing strains (designated "*cps*1" through "*cps*8")[25]; and an acapsular strain in which all *cps* genes from all 8 loci are deleted[24] as independent hosts for phage isolation. Primary sewage effluent from two cities within the United States (Ann Arbor, Michigan and San Jose, California; separated by approximately 3,300 kilometers) was used as the source for these phages (for further details on phage isolation, see *Methods* and **Table S1**). All phages were plaque purified at least 3 times and high titer lysates were generated for each of the 71 phages. Plaque morphologies varied greatly among the strains, ranging in size from <1 mm to 3 mm or greater and in opacity from very turbid to clear (**Figures 1, S1**).

To determine if phages isolated on each individual host strain are restricted in their ability to infect other strains, we systematically tested each phage against each of the 10 *B. thetaiotaomicron* host strains that varied only in the CPS they are capable of expressing (n=3). Hierarchical clustering of the host infection profiles revealed a cladogram with 3 main branches that each encompasses phages from both collection sites, although substantial variation in host tropism exists for phages within each branch (**Figure 1**). Furthermore, individual phages within each branch displayed a range of plaque morphologies (**Figures 1, S1**), suggesting additional diversity in the collection that is not captured by this assay. Finally, host range assays were robust when performed by different experimenters at different research sites (**Figure S2**).

Phages in Branch 1 generally exhibited robust infection of the acapsular strain, although 3 of these phages did not form plaques on this host. Furthermore, phages in Branch 1 generally exhibited robust infection on strains expressing CPS7 or CPS8 alone, although a separate subset of 3 phages did not form plaques on the CPS8 expressing strain. Some Branch 1 phages also displayed less efficient infection of other strains with the exception of *cps*4, which was not infected by any phages in this group. Interestingly, ARB154 exclusively infected *cps*8, an uncommon CPS among *B. thetaiotaomicron* strains that appears to be contained in a mobile element[4]. Phages in Branch 2 generally exhibited robust infection of all strains except *cps*2, *cps*3 and *cps*4. However, subsets of this group were unable to infect *cps*1 or *cps*6. Finally, Branch 3 tended to exhibit strong infection of wild-type, *cps*1, *cps*2, and *cps*3, with some variations. Some Branch 3 phages also exhibited the ability to infect the *cps*7 and acapsular strains but were the only branch that poorly infected *cps*8. A subset of phages on Branches 1 and 3 failed to infect the acapsular strain, suggesting that they may require the presence of CPS for infection. Taken together, the observed variations in phage infection of *B. thetaiotaomicron* strains expressing different CPS, or none at all, provide initial support for our hypothesis that these surface structures are a key mediator of *B. thetaiotaomicron*-phage interactions.

Elimination of specific CPS subsets alters bacterial susceptibility to phages

One explanation for the differences in host infectivity described above is that there are distinct, CPS-dependent mechanisms of phage adsorption. For example, several phages robustly infect the acapsular strain, indicating that a capsule-independent cell surface receptor mediates infection for these examples. However, each of these phages also infect subsets of the single CPS-expressing strains, suggesting that some "non-permissive" CPS could block access to cell

surface receptors, while other "permissive" CPS fail to do so. For phages that do not efficiently infect the acapsular strain, one or more CPS may serve as a direct phage receptor(s) or as a required co-receptor.

To further define the roles of specific CPS during phage infection, we investigated a subset of 6 phages (ARB72, ARB78, ARB82, ARB101, ARB105, and ARB25). All 6 of these phages can infect the wild-type *B. thetaiotaomicron* strain that variably expresses its 8 different CPS and 5 of them infect the acapsular strain poorly or not at all (**Figure 1**). We tested our hypothesis that some CPS are required as receptors or co-receptors by deleting only the subsets of CPS biosynthetic genes encoding permissive capsules based on our prior experiments with single CPS-expressing strains. For ARB72, which most robustly infects the CPS1- and CPS3-expressing strains, simultaneous elimination of both of these capsules from wild-type *B. thetaiotaomicron* reduced infection below the limit of detection (**Figure 2A**). Likewise, elimination of the most permissive CPS for four other phages (ARB78, ARB82, ARB101 and ARB105) significantly reduced *B. thetaiotaomicron* infection by these phages, in some cases in the presence of permissive CPS (**Figure 2B-E**).

For ARB25, which infected 7 of the 10 strains tested in our initial plaque assays (**Figure 1**), some single and compounded *cps* gene deletions significantly reduced infection rates or reduced them below the limit of detection. The most noteworthy of these was deletion of the *cps*4 (initially determined in **Figure 1** to be non-permissive for ARB25) in combination with deleting the permissive *cps*1, which completely eliminated detectable infection (**Figure 2F**). While individual deletion of four other permissive CPS (*cps*1,6,7,8) led to partially reduced infection, so did single deletions of two CPS initially determined to be non-permissive (*cps*3 and *cps*4), suggesting the possibility of more complicated regulatory interactions, which are known

to occur with *Bacteroides* CPS[23,26]. Interestingly, strains lacking either *cps4* or *cps1/cps4* together compensated by significantly increased relative expression of the non-permissive *cps2* locus, which could contribute to ARB25 resistance (**Figure 2G**).

A strain expressing only two of the non-permissive CPS (CPS2 and CPS3) could not be detectably infected by ARB25 (**Figure 2F**, "2,3 only" condition). However, a strain expressing CPS2,3,4 regained some susceptibility (**Figure 2F**, "2,3,4 only" condition), indicating that CPS4 is capable of mediating some infection by this phage (addressed further below). In contrast to sole expression of *cps2* and *cps3* promoting resistance to ARB25, deletion of the *cps2* and *cps3* loci led to dominant expression of *cps1* and *cps4* genes, which increased infection efficiency and led to the production of clearer plaques (**Figure 2F,H**). Additional support for the idea that loss of *cps4* alone modifies ARB25 susceptibility comes from plaque morphologies arising from infection of the Δ*cps4* strain, which produced smaller and more turbid plaques, demonstrating that when infection does occur it is less productive (**Figure 2H**). Thus, CPS4-expressing cells are in some cases susceptible to ARB25, and loss of the genes encoding this capsule result in increased expression of CPS2, perhaps through alleviation of the UpxZ transcription termination mechanism, as described for capsule regulation in *B. fragilis*[23]. Because this programmed shift in CPS expression may increase resistance to phage, it is possible that phase-variation can equip *Bacteroides* populations to survive phage predation through selection of individual cell sub-populations that are expressing non-permissive CPS.

Finally, we used ARB25 to test if purified, exogenous CPS could modify phage infection as has been shown for other phage-CPS interactions[27]. Given that strains expressing CPS1 and CPS2 show different susceptibility to infection by ARB25, we tested purified preparations of both CPS types, expecting that exogenous CPS2 might inhibit ARB25 infection of the acapsular

strain if it is capable of directly interfering with recognition of a surface receptor. Arguing against this hypothesis, pre-incubation with purified preparations of CPS2 or CPS1 (a permissive capsule control) both resulted in no significant difference in plaquing efficiency on the acapsular strain (**Figure S3A**), suggesting that exogenous CPS cannot block infection by phage ARB25.

B. thetaiotaomicron **acquires transient resistance to phage infection**

Interestingly, we observed that liquid cultures of the various *B. thetaiotaomicron* strains infected with ARB25 did not show evidence of complete lysis after 36 hours of growth, as determined by optical density at 600 nm (OD_{600}) (**Figure 3**). Previous reports demonstrated that *B. fragilis*[20] and *B. intestinalis*[21] exhibited transient resistance to phage infection that could be "reset" through removal of the phage from the culture, although the underlying mechanism of this transient resistance was not determined. Based on these observations, we questioned whether similar transient resistance occurs with *B. thetaiotaomicron* and whether this resistance could be dependent on CPS expression. Growth curves of each of the CPS-expressing strains inoculated with active or heat-killed ARB25 confirmed our initial host range assays, except that cultures containing the CPS4-expressing strain were sensitive to killing by this phage in liquid culture, while cultures of the CPS6-expressing strain had no significant decrease in OD_{600} (**Figure 3**). In these experiments, most strains deemed to be susceptible via plaque assay (**Figure 1**) exhibited an initial lag in growth or a drop in OD_{600}. As expected, the *cps*2 and *cps*3 strains remained resistant to ARB25 infection for the duration of the experiments. However, after initial growth inhibition, the susceptible strains displayed either growth stagnation without complete loss of culture density (*cps*4, *cps*5, *cps*7, *cps*8) or resumption of growth (wild-type, acapsular, *cps*1) to near uninfected levels, the latter suggesting outgrowth of a resistant subpopulation of bacteria. Culture supernatants taken from ARB25 post-infected, wild-type *B. thetaiotaomicron* still

contained high phage titers when exposed to naïve bacteria that had not been exposed to phage, excluding the possibility that the phages were inactivated (**Figure S3B**).

We next determined whether strains that had survived or proliferated after exposure to phage retained resistance after removal of phage. In order to isolate phage-free bacterial clones, we isolated individual colonies by sequentially streaking each twice from a subset of the cultures that gained resistance to ARB25 (WT, acapsular, *cps*1, and *cps*4) as well as the inherently ARB25-resistant *cps*2 strain. The majority of clones isolated using this process were free from detectable phage (see *Methods*). We then re-infected each clone with live ARB25 and monitored susceptibility by delayed growth or drop in the culture density as compared to infection with heat-killed phage. As expected, the *cps*2 strain remained resistant. On the other hand, the majority of clones (42/61 total, ~69%) of the other four strains regained susceptibility (**Table 1**), suggesting that resistance to this phage is not caused by a permanent genetic alteration in most cases.

Phage resistant wild-type *B. thetaiotaomicron* populations exhibit altered *cps* locus expression

Given that CPS type is correlated with resistance to phage infection (e.g., ARB25 fails to infect strains expressing CPS2 or CPS3 under all conditions tested), we hypothesized that wild-type *B. thetaiotaomicron* cells inherently expressing resistant capsules would be positively selected in the presence of phage. To test this, we infected wild-type *B. thetaiotaomicron* with ARB25 and monitored bacterial growth. For cells treated with a high MOI (MOI ≈ 1), culture turbidity increased very slightly, declined before 3 hours after infection, and finally increased again to ultimately achieve a high growth level as previously observed (**Figure 4A**).

Interestingly, bacterial cultures originating from different single colonies displayed variable growth kinetics and possibly resistance frequency to ARB25, with the growth of one clone barely delayed by treatment with live ARB25. Next, we measured if infection with ARB25 resulted in a change in CPS expression by the phage-resistant *B. thetaiotaomicron* population. In support of our hypothesis, *B. thetaiotaomicron* exposed to heat-killed phage predominantly expressed genes encoding CPS3 and CPS4, which we typically observe in *in vitro* culture. Treatment with live ARB25 resulted in a dramatic loss of *cps*1 and *cps*4 expression (capsules for which combined loss eliminates ARB25 infection, **Figure 2F**), with a concomitant increase of expression of the non-permissive *cps*3 (**Figure 4B**). Similar growth and expression phenotypes occurred in cultures treated with a low MOI (MOI $\approx 10^{-4}$), albeit with higher culture turbidity before a decline and subsequent resumption of growth (**Figure S4**). Dirichlet regression (see *Methods*) supported significant *cps* expression changes for *cps1*, *cps3*, and *cps4* in response to ARB25 ($p < 0.01$ for experiments with both low and high MOI). Notably, the most resistant of the three bacterial clones (as evidenced by faster outgrowth post-infection) in each of the two experiments (low and high MOI) exhibited similar *cps* expression to the other clones after treatment with live phage, but expressed lower levels of permissive *cps*1 and *cps4* and higher levels of non-permissive *cps*3 in heat-killed phage treatment groups (**Figure S5)**. This alteration in expression of permissive and non-permissive CPS may contribute to the ability of these clones to resume growth more rapidly after phage challenge because it was already skewed towards expression of non-permissive CPS.

Multiple layers of phase-variable resistance functions equip *B. thetaiotaomicron* to survive phage predation

The results described above support a model in which some individual cells within a *B. thetaiotaomicron* population are pre-adapted to resist eradication by a single phage like ARB25 through expression of different CPS. Complex phase-variation mechanisms have already been described in controlling CPS expression in *B. thetaiotaomicron* and *B. fragilis*[4,23,28] and we demonstrate the ability of CPS to dictate interactions between phages and *B. thetaiotaomicron* (**Figures 1, 2, 3**). However, in acapsular *B. thetaiotaomicron* infected with ARB25, we observed significant growth after initial lysis that frequently regained susceptibility when isolated from and subsequently re-challenged with ARB25 (**Figure 3, Table 1**), suggesting the emergence of a transiently phage-resistant subpopulation in the absence of CPS. To determine if additional phage resistance mechanisms are involved, we performed whole genome transcriptional profiling by RNA-sequencing (RNA-seq) to measure transcriptional differences between ARB25 post-infected and mock-infected *B. thetaiotaomicron*. In a wild-type *B. thetaiotaomicron* population, in which cells retain the ability to phase-vary expression of their eight different CPS, the transcriptional profiles of bacterial populations surviving after ARB25 infection (n=3) were largely characterized by decreased gene expression: among a total of 56 genes that exhibited significant expression differences $\geq$3-fold between *B. thetaiotaomicron* exposed to live and heat-killed ARB25, 51 genes were decreased in post-infected bacteria (**Figure 5A, Table S2**). Most of these genes with decreased transcription (44/51) belong to the loci encoding CPS1 and CPS4, consistent with our findings by qPCR that lower expression of these CPS occurs after ARB25 challenge (**Figure 4B**). Correspondingly, increased expression of genes encoding CPS2 and CPS3 was also apparent by RNA-seq, but did not reach statistical significance. Interestingly, two additional gene clusters encoding different outer-membrane "Sus-like systems", which are well-described Bacteroidetes mechanisms for import and degradation of carbohydrates and other

nutrients[2,29], were also decreased in post-infected bacteria (6/7 remaining non-*cps* genes). The central features of these systems are outer membrane TonB-dependent transporters (similar to *E. coli* TonA, or **T** **on**e phage receptor **A**; the first described phage receptor[30]), suggesting the possibility that the proteins encoded by these genes are part of the receptor for ARB25.

To further investigate CPS-independent mechanisms that allow *B. thetaiotaomicron* to avoid eradication by phage, we performed identical RNA-seq experiments on acapsular *B. thetaiotaomicron* populations exposed to heat-killed or live ARB25. In this case, we hypothesized that eliminating the ability to phase vary capsules would better reveal the mechanism(s) that allows the acapsular strain to survive ARB25 predation, possibly by further reducing expression of putative phage receptors like the Sus-like systems identified above. Expression of the same two Sus-like systems (*BT2170-73, BT2365*) that were decreased in ARB25-exposed wild-type were also decreased to similar levels in acapsular *B. thetaiotaomicron* (**Figure 5B**). Otherwise, the transcriptional profiles of acapsular *B. thetaiotaomicron* surviving after ARB25 infection (n=3) were largely characterized by increased gene expression: 71 of the 81 genes differentially regulated ≥3-fold between ARB25-infected and the heat-killed reference showed increased expression in ARB25 post-infected cells. All but 3 of these genes were unique to the post-infection transcriptome of acapsular *B. thetaiotaomicron* compared to wild-type *B. thetaiotaomicron* (**Figure 5B**, **Table S2**).

Among the 71 genes with increased expression in post-infected acapsular *B. thetaiotaomicron*, 24 genes were part of 8 different gene clusters that encode putative tyrosine recombinases along with pairs of outer membrane lipoproteins and OmpA β-barrel proteins (**Figure 5C**). One of these genes (*BT1927*) was previously characterized as encoding a phase-variable, S-layer protein, which organizes into a tessellated structure on the cell surface and

294 when locked into the "on" orientation promoted increased *B. thetaiotaomicron* resistance to
295 complement-mediated killing[31]. The remaining 7 gene clusters share both syntenic organization
296 and homology to this original S-layer gene cluster. Closer scrutiny of the promoter regions
297 upstream of the 7 newly identified gene clusters encoding putative lipoproteins revealed that
298 each is also flanked by a pair of imperfect, 17 nucleotide palindromic repeats (**Figure 5C**). Three
299 of these newly identified repeats are identical to the repeats known to mediate recombination at
300 the *BT1927* promoter[31]. The remaining 4 sequences only varied by the sequence of a
301 trinucleotide located in the middle of each imperfect palindrome (colored blocks in **Figure 5C**
302 middle). Finally, amplicon sequencing of each promoter supported the existence of the proposed
303 recombination events in 5 of the 7 newly identified loci (**Figure S6)**.
304 Among the remaining genes that were significantly up- or down-regulated in post
305 ARB25-infected acapsular *B. thetaiotaomicron*, there was an additional signature of genes for
306 which DNA recombination may be involved in re-organizing expression of cell surface proteins.
307 Specifically, the expression of 3 of 4 genes in an operon (*BT1042-45*) previously implicated in
308 utilization of host glycans[32] were expressed an average of 6.9-fold less in ARB25-infected
309 acapsular cells compared to heat-killed controls. Correspondingly, 5 genes in an adjacent operon
310 (*BT1046-51*) with similar arrangement and predicted functions exhibited an average of 8.3-fold
311 increased expression in acapsular *B. thetaiotaomicron* exposed to live ARB25. Both of these
312 operons have been previously linked to transcriptional regulation by a nearby extra-cytoplasmic
313 function sigma (ECF-σ), anti-σ factor pair, such that when the single ECF-σ coding gene
314 (*BT1053*) is deleted, the ability to activate the adjacent operons is eliminated[33]. Based on 1) the
315 ARB25-dependent shift in gene expression described above; 2) previously established common
316 ECF-σ regulation of the *BT1042-45* and *BT1046-51* operons; 3) the observation that two genes

317 encoding TonB-dependent transporters (*BT1040, BT1046*) appear to be truncated at their 5' ends
318 compared to *BT1042* (**Figure 4**, **S7A**) and only the full-length BT1042 sequence harbors a
319 required anti-σ contact domain[33]; and 4) annotation of a gene encoding a putative tyrosine
320 recombinase (*BT1041*) located in the middle of this locus, we hypothesized that this gene cluster
321 possesses the ability to undergo recombination and that specific combinatorial variants are
322 selected under phage pressure.

323 To test this hypothesis, we designed PCR primer pairs (**Figure 5D**, green dumbbells) to
324 detect both the originally annotated sequence orientation and 3 potential alternative
325 recombination states derived from either moving the full-length 5' end of *BT1042* to one of two
326 alternative *susC*-like genes or an internal rearrangement derived from recombination of two
327 incomplete *susC*-like genes (**Figure 5D**, variants 1-3). In support of our hypothesis, we were
328 able to detect by both PCR (**Figure 5E**) and amplicon sequencing (**Figure S7B**) the presence of
329 all 5 predicted alternative recombination states (**Figure 5D,E** amplicons 2, 3, 4, 6, 7), plus the 3
330 expected from the originally published genome assembly[34]. In further support of our hypothesis,
331 an insertion mutation in the associated tyrosine recombinase-coding gene (*BT1041*) locked the
332 corresponding mutant into the native genomic architecture as determined by the presence of
333 amplicon 1, but the absence of amplicons 2 and 3 (**Figure 5E**). Further sequence analysis and
334 tracking of single nucleotide polymorphisms in the 5' ends of the three recombinationally active
335 *susC*-like genes narrowed the recombination site down to a 7 bp sequence that is flanked by an
336 imperfect direct repeat encompassing over 132 additional downstream bp that may also influence
337 recombination specificity (**Figure S7B**). Thus, three separate operons that are under the
338 transcriptional control of a single ECF-σ regulator and are involved in utilization of host glycans,
339 are also able to undergo recombinational shuffling via a tyrosine recombinase/direct repeat

mediated mechanism to vary which of the three operons is expressed to produce its corresponding surface proteins. This strategy is similar to recombinational shufflons involving nutrient utilization functions that have been characterized in *B. fragilis*[35,36], with the exception that in the example described here, recombination occurs between direct repeats instead of palindromes. One explanation for this phenomenon is that shufflons evolved to subvert phage infection by expressing alternate cell surface receptors for important nutrients, which are also targeted by phages. Contrary to this hypothesis that the proteins encoded by BT1042-45 might serve as a receptor for ARB25, elimination of the genes spanning BT1033-52 did not eliminate ARB25 infection in the acapsular strain, suggesting that an additional or different receptor(s) exists. Interestingly, the BT1033-52 mutant exhibited variable plaquing efficiency compared to the acapsular parent (**Figure S8**), suggesting that loss of these genes cold exert global effects that mediate susceptibility to ARB25.

Since the gene encoding the original outer membrane S-layer protein (*BT1927*), and its downstream gene (*BT1926*), were among the most highly activated (147- and 114-fold, respectively) in post-infected, acapsular *B. thetaiotaomicron,* and can be alternatively locked into the "on" or "off" orientations by mutating the recombination site upstream of the phase-variable promoter[31], we chose to focus on the role of this single function in resisting phage infection. We re-engineered acapsular *B. thetaiotaomicron* to contain locked "on" and locked "off" versions of this promoter and evaluated sensitivity to ARB25 and another phage, SJC01, which has a similar infection profile on *B. thetaiotaomicron* (**Figure 1**). Consistent with the observation that S-layer is highly activated in acapsular *B. thetaiotaomicron* infected with ARB25, acapsular S-layer "off" cells were more effectively killed in the presence of live phage relative to acapsular S-layer "on" cells (**Figure 6A**). Interestingly, SJC01 showed the opposite effect, as it more effectively

363 killed cells with S-layer locked "on" versus cells with S-layer locked "off" (**Figure 6B**). These
364 data indicate that in addition to the 8 surface-exposed capsular polysaccharide types, S-layer
365 lipoproteins can function as positive or negative determinants of phage tropism in *B.*
366 *thetaiotaomicron*. Finally, the observation that CPS[5,6] and homologous S-layer like functions are
367 broadly represented in gut Bacteroidetes[31], suggest that these two mechanisms help to diversify
368 members of this phylum under phage-mediated selection.
369
370 **Discussion**
371 Production of multiple phase-variable CPS is a hallmark of human gut Bacteroidetes.
372 Previous work has revealed the importance of *Bacteroides* CPS in interactions with the host
373 immune system[4,8,37,38]. However, other biological roles for *Bacteroides* CPS remain relatively
374 unexplored. Using a panel of *B. thetaiotaomicron* strains that express individual CPS, we tested a
375 previously inaccessible hypothesis: that *Bacteroides*-targeting phage can be both inhibited and
376 assisted by the repertoire of capsules expressed by their host bacteria. Our data clearly indicate
377 that production of specific CPS is associated with alterations in phage susceptibility, which is
378 underscored by the observation that none of the 71 phages characterized here infect every CPS-
379 variable strain that we tested (**Figure 1**). Phage-mediated selection and interactions with the host
380 immune system help to explain both the extensive diversification of CPS structures in gut-
381 resident Bacteroidetes[4,12] and their complex phase-variable regulation mechanisms within a
382 given strain or species[23,25]. Surprisingly, our results also reveal that additional phase variable
383 markers are expressed by *B. thetaiotaomicron* under phage-mediated selection, highlighting that
384 other strategies exist in *Bacteroides* for surviving in the face of phage predation.

There are several mechanisms through which CPS could promote or prevent phage infection. First, CPS may sterically mask surface receptors to block phage binding, although additional specificity determinants must be involved because no individual phage that infects the acapsular strain is blocked by all single *B. thetaiotaomicron* CPS. These specificity determinants could be driven by CPS structure (physical depth on the cell surface, polysaccharide charge, permeability) or be actively circumvented by the presence of polysaccharide depolymerases on the phage particles, as has been described in other phage-bacterium systems (e.g., *E. coli* K1 and phiK1-5[39]). Alternatively, certain permissive CPS could serve as obligate receptors[40] (i.e., phage that do not infect acapsular) or more generally increase the affinity of a phage for the bacterial cell surface – similar to what has been proposed in the "bacteriophage adhering to mucus" model, whereby hypervariable domains on phage capsids facilitate adherence to mucus and increase the frequency of bacteria-phage interaction[41]. This latter type of adherence to CPS might increase the likelihood that a phage would contact its receptor by sustained interaction with the extracellular matrix. Some combination of these possibilities is likely to explain the host range infection profile for the majority of the phages in our collection. Collectively, our observations provide the foundation for future mechanistic work, beginning with phage genome sequencing, aimed at understanding the physical and chemical interactions that mediate infection of *B. thetaiotaomicron* and other *Bacteroides* by their phages.

Using ARB25 as a representative from our larger collection, we demonstrate that infection with this single phage does not fully eradicate presumably susceptible *B. thetaiotaomicron* populations. Rather, resistant cells grow, often quickly, after what we interpret to be an initial lytic event (**Figure 3**). Similar observations were previously made with ΦCrass001, a phage that infects *B. intestinalis* [21]. Specifically, though ΦCrass001 robustly

formed plaques on lawns of *B. intestinalis*, it failed to eradicate this bacterium in liquid culture. We hypothesize that ΦCrass001 undergoes lytic growth on a subset of *B. intestinalis* cells expressing permissive CPS. Subsequently, clones expressing resistant CPS would grow to dominate a culture after phage challenge, as we observed with ARB25.

Given the roles of CPS in mediating *B. thetaiotaomicron*-phage interactions, the outgrowth of a phage resistant sub-population was especially surprising in the context of acapsular *B. thetaiotaomicron*. While wild-type *B. thetaiotaomicron* primarily appears to survive continued ARB25 predation by pre-adaptive CPS variation or by shifting the population to express non-permissive CPS, the acapsular strain instead shifts to increased expression of phase-variable sets of surface proteins, at least one of which (*BT1927-26*) confers increased resistance to ARB25 when locked "on" in acapsular *B. thetaiotaomicron*. A previous study measured that only 1:1000 *B. thetaiotaomicron* cells in an unchallenged population express the S-layer encoded by *BT1927* [31]. Given that ARB25 non-permissive CPS can comprise up to 40% of the expressed capsule population (e.g., CPS3 in **Figure 4B**), the rapid emergence of cells expressing alternative CPS could be explained by this frequency difference, in which case S-layer expressing wild-type bacteria might also emerge given longer phage exposure times. The original *B. thetaiotaomicron* S-layer study also demonstrated that locking the invertible promoter for the *BT1927* S-layer into the "on" orientation facilitated survival against complement-mediated killing[31], suggesting that orthogonal roles for this and related proteins exist in *B. thetaiotaomicron* and likely facilitate survival in the face of diverse environmental disturbances. Combined with our data on CPS-mediated phage tropism, our observations that the *BT1927*-encoded S-layer confers resistance to some phages, that 7 other homologous systems are also upregulated after exposure to ARB25, and that a shufflon harboring three recombinationally-variable nutrient acquisition operons exists

in *B. thetaiotaomicron*, together reveal that there are at least 17 independent cell surface structures in *B. thetaiotaomicron* that could be altered in cells exposed to phages. The fact that almost all of these surface structures (14/17) are under the control of independent phase-variable promoters with associated tyrosine recombinases[4,31,36] and that their products show altered expression after phage exposure speaks to the effectiveness of this strategy in pre-adapting some *B. thetaiotaomicron* cells within a population to thrive in the face of phage predation.

Phages are the most abundant biological entities in the gut microbiome[41] and interest in the roles and identities of these gut-resident viruses is increasing as metagenomic sequencing approaches are unveiling a more comprehensive view of their dynamics during health and disease[16,17,42]. Although sequence-based approaches are powerful for describing the phages that are present, they do not generate information on the definitive hosts or the mechanisms of individual bacteria-phage interactions, which are likely to be elaborate. These limitations will prohibit full dissection of the ecological interactions that phage exert on bacterial populations in the gut. The approach taken here of isolating phages for a particular host of interest, with added layers of detail like systematic variation of surface CPS when possible, will be an essential complement to high throughput sequencing studies and will help build a foundation of mechanistic gut bacterium-phage interactions.

Our results, with a single strain of bacteria commonly found in human gut microbiomes, point to the existence of a very complex relationship between bacteria and phage in the gut microbiome. Considering the possibilities that these interactions could vary over time, differ by host species, and evolve differently within individuals or regionally distinct global populations, the landscape becomes even more complex. Given the diverse adaptive and counter-adaptive strategies that have apparently evolved in the successful gut symbiont *B. thetaiotaomicron* and

its relatives, our findings hold important implications for the use of phages to intentionally alter the composition or function of the gut microbiota. While a cocktail of multiple phages could theoretically be harnessed together to elicit more robust alteration of target populations within a microbiome, the complexity of host tropisms and bacterial countermeasures that exist for *B. thetaiotaomicron* imply that a deliberate selection of complementary phage would be needed. If selections of effective phage cocktails need to be further tailored to individual microbiomes, or elicit resistance within individuals or populations the way antibiotics do, the prospects for effective gut microbiome-targeting phage therapy could indeed become very complicated. Given these considerations, our findings here contribute an important early step towards building a deep functional understanding of the bacterium-virus interactions that occur in the human gut microbiome. As such, this work contributes to the overall goal of understanding the ecology of this important microbial community and developing rational approaches to shape its physiology.

Methods

Bacterial strains and culture conditions.

The bacterial strains used in this study are listed in **Table S3.** Frozen stocks of these strains were maintained in 25% glycerol at -80˚C and were routinely cultured in an anaerobic chamber or in anaerobic jars (using GasPak EZ anaerobe container system sachets w/indicator, BD) at 37˚C in *Bacteroides* Phage Recovery Medium (BPRM), as described previously[43]: per 1 liter of broth, 10 g meat peptone, 10 g casein peptone, 2 g yeast extract, 5 g NaCl, 0.5 g L-cysteine monohydrate, 1.8 g glucose, and 0.12 g $MgSO_4$ heptahydrate were added; after autoclaving and cooling to approximately 55 °C, 10 ml of 0.22 µm-filtered hemin solution (0.1% w/v in 0.02% NaOH), 1 ml of 0.22 µm-filtered 0.05 g/ml $CaCl_2$ solution, and 25 ml of 0.22µm-

filtered 1 M Na_2CO_3 solution were added. For BPRM agar plates, 15 g/L agar was added prior to autoclaving and hemin and Na_2CO_3 were added as above prior to pouring the plates. For BPRM top agar used in soft agar overlays, 3.5 g/L agar was added prior to autoclaving. Hemin, $CaCl_2$, and Na_2CO_3 were added to the top agar as above immediately before conducting experiments. Bacterial strains were routinely struck from the freezer stocks onto agar plates of Brain Heart Infusion supplemented with 10% horse blood (Quad Five, Rygate, Montana) (BHI-blood agar; or for the SJC phages used in **Figure 1**, on BPRM agar) and grown anaerobically for up to 3 days. A single colony was picked for each bacterial strain, inoculated into 5 mL BPRM, and grown anaerobically overnight to provide the starting culture for experiments.

For the experiment described in **Figure 2G**, liquid cultures of *B. thetaiotaomicron* were grown in BPRM using the pyrogallol method as described previously[4]. Briefly, a sterile cotton ball was burned and then pushed midway into the tube, after which 200 µl of saturated $NaHCO_3$ and 200 µl of 35% pyrogallol solution were added to the cotton ball. A rubber stopper was used to seal the tubes, and tubes were incubated at 37 °C.

Bacteriophage isolation from primary wastewater effluent.

The bacteriophages described in this study were isolated from primary wastewater effluent from two locations at the Ann Arbor, Michigan Wastewater Treatment Plant and from the San Jose-Santa Clara Regional Wastewater Treatment Facility. After collection, the primary effluent was centrifuged at 5,500 rcf for 10 minutes at room temperature to remove any remaining solids. The supernatant was then sequentially filtered through 0.45 µm and 0.22 µm polyvinylidine fluoride (PVDF) filters to yield "processed primary effluent." Initial screening for plaques was done using a soft agar overlay method[44] where processed primary effluent was

combined with 1 part overnight culture to 9 parts BPRM top agar and poured onto a BPRM agar plate (e.g. 0.5 mL overnight culture and 4.5 mL BPRM top agar was used for standard circular petri dishes [100 mm x 15 mm]). Soft agar overlays were incubated anaerobically at 37 ˚C overnight. Phages were successfully isolated using three permutations of this assay: (1) Direct plating, where processed primary effluent was directly added to overnight culture prior to plating. (2) Enrichment, where 10 mL processed primary effluent was mixed with 10 mL 2XBPRM and 3 mL exponential phase *B. thetaiotaomicron* culture and grown overnight. The culture was centrifuged at 5500 rcf for 10 minutes and filtered through a 0.22 µm PVDF filter. (3) Size exclusion, where processed primary effluent was concentrated up to 500-fold via 30 or 100 kDa PVDF or polyethersulfone size exclusion columns. Up to 1 mL of processed primary effluent, enrichment, or concentrated processed primary effluent was added to the culture prior to adding BPRM top agar, as described above. To promote a diverse collection of phages, no more than 5 plaques from the same plate were plaque purified and a diversity of plaque morphologies were selected as applicable. When using individual enrichment cultures, only a single plaque was purified.

Single, isolated plaques were picked into 100 µL phage buffer (prepared as an autoclaved solution of 5 ml of 1 M Tris pH 7.5, 5 ml of 1 M MgSO4, 2 g NaCl in 500 ml with ddH_2O). Phages were successfully plaque purified using one of two methods: (1) a standard full plate method, where the diluted phage samples were combined with *B. thetaiotaomicron* overnight culture and top agar and plated via soft agar overlay as described above or (2) a higher throughput 96-well plate-based method, where serial dilutions were prepared in 96-well plates and 1 µL of each dilution was spotted onto a solidified top agar overlay. This procedure was

repeated at least 3 times to plaque purify each phage. For more details on the phages isolated in this work, see **Figure S1**.

High titer phage stocks were generated by flooding a soft agar overlay that plate yielded a "lacey" pattern of bacterial growth (near confluent lysis). Following overnight incubation of each plate, 5 ml of sterile phage buffer was added to the plate to resuspend the phage. After at least 2 hours of incubation at room temperature, the lysate was spun at 5,500 rcf for 10 minutes to clear debris and then filter sterilized through a 0.22 μm PVDF filter.

Quantitative host range assays.

To accommodate the large number of phage isolates in our collection, we employed a spot titer assay for semi-quantitative comparisons of infectivity on each bacterial strain. High titer phage stocks were prepared on their "preferred host strain," which was the strain that yielded the highest titer of phages in a pre-screen of phage host range. Lysates were then diluted to approximately 10^6 PFU/mL, were added to the wells of a 96-well plate, and further diluted to 10^5, 10^4, and 10^3 PFU/mL using a multichannel pipettor. One microliter of each of these dilutions was plated onto solidified top agar overlays containing single bacterial strains indicated in each figure. After spots dried, plates were incubated anaerobically for 15-24 hours prior to counting plaques. Phage titers were normalized to the bacterial strain that typically exhibited the highest phage titer, which was designated as the "preferred host strain".

Images of phage plaques.

To document the morphologies of plaques formed by the purified phages, two sets of plaque pictures were taken: the first set were taken with a Color QCount Model 530 (Advanced

Instruments) with a 0.01 second exposure. Images were cropped to 7.5 mm^2 but were otherwise unaltered. The second set of images were taken on a ChemiDoc Touch instrument (BioRad) with a 0.5 second exposure. Images were cropped to 7.5 mm^2 and background unnaturally high pixels were removed (Image Lab, BioRad) to facilitate viewing of the plaques.

Incubation of phage with extracted CPS.

Approximately 50-100 PFU of ARB25 in 50 µl phage buffer were mixed with an equal volume of H_2O or capsule (2 mg/ml) extracted by the hot water-phenol method (as described in Reference 4) and incubated at 37 °C for 30 minutes. Samples were then plated on the acapsular strain, and plaques were counted after 15-24 hours anaerobic incubation at 37 °C. Counts from two replicates on the same day were then averaged, and the experiment was performed three times.

Growth curves.

For growth curve experiments, 3 individual clones of each indicated strain were picked from agar plates and grown overnight in BPRM. Then, for experiments in **Figures 3** and **6**, each clone was diluted 1:100 in fresh BPRM and 100 µl was added to a microtiter plate. 10 µl of approximately $5*10^6$ PFU/ml live or heat-killed phage were added to each well, plates were covered with an optically clear gas-permeable membrane (Diversified Biotech, Boston, MA) and optical density at 600 nm (OD_{600}) values were measured using an automated plate reading device (BioTek Instruments). Phages were heat killed by heating to 95 °C for 30 minutes, and heat-killed phage had no detectable PFU/ml with a limit of detection of 100 PFU/ml.

In **Figure S3B**, wild-type *B. thetaiotaomicron* was infected with live or heat-killed ARB25, and bacterial growth was monitored via optical density at 600 nm (OD_{600}) on an automated plate reader for 12 hours. At 0, 6.02, 8.36, and 11.7 hours post inoculation, replicate cultures were vortexed in 1:5 volume chloroform, centrifuged at 5,500 rcf at 4 °C for 10 minutes, and the aqueous phase was titered on the acapsular strain. No phages were detected in heat-killed controls.

Generation of phage-free bacterial isolates and determination of their phage susceptibility.

To isolate phage-free bacterial clones from ARB25-infected cultures (Table 1), each culture was streaked on a BHI-blood agar plate. Eighteen individual colonies were picked from each plate, and each of these clones was re-streaked onto a new BHI-blood agar plate. One colony was picked from each of these secondary plates and was inoculated into 150 µl BPRM broth and incubated anaerobically at 37 °C for 2 days. Only one of the clones (a *cps*4 isolate) failed to grow in liquid media. To determine whether cultures still contained viable phage, 50 µl of each culture was vortexed with 20 µl chloroform, then centrifuged at 5,500 rcf for 10 minutes. 10 µl of the lysate was spotted on BPRM top agar containing naïve acapsular bacteria and was incubated anaerobically overnight at 37 °C. Loss of detectable phage in the twice passaged clones was confirmed for most of the clones (79/89, 89%) by the absence of plaques on the naïve acapsular strain.

To determine whether the resultant phage-free plaques were resistant to ARB25 infection, each culture was diluted 1:100 in fresh BPRM, 100 µl was added to a microtiter plate, and 10 µl of either live or heat-killed ARB25 (approximately $5*10^6$ PFU/ml) was added. Plates were incubated anaerobically at 37 °C for 48 hours, and OD_{600} was measured as described above.

590 Cultures were determined to be susceptible to ARB25 by demonstration of delayed growth or
591 drop in OD_{600}, as compared to heat-killed controls.
592
593 ***Measurement of* cps *gene expression***.
594 For **Figures 2G**, **4**, and **S4**, overnight cultures were subcultured into fresh BPRM to an
595 OD_{600} of 0.01. For Figure 4B, 200 µL of approximately 2 x 10^8 PFU/mL live phage or heat killed
596 phage were added to 5 mL of the diluted cultures. For **Figure S4**, 200 µL of approximately 2 x
597 10^5 PFU/mL live phage or heat killed phage were added to 5 mL of the diluted cultures. Bacterial
598 growth was monitored by measuring OD_{600} every 15-30 minutes using a GENESYS 20
599 spectrophotometer (Thermo Scientific). Cultures were briefly mixed by hand before each
600 measurement. For determination of relative *cps* gene expression, cultures were grown to OD_{600}
601 0.6-0.8, were centrifuged at 7700 rcf for 2.5 minutes, the supernatant was decanted, and the
602 pellet was immediately resuspended in 1 ml RNA-Protect (Qiagen). RNA-stabilized cell pellets
603 were stored at -80 °C.
604 Total RNA was isolated using the RNeasy Mini Kit (Qiagen) then treated with the
605 TURBO DNA-free Kit (Ambion) followed by an additional isolation using the RNeasy Mini Kit.
606 cDNA was then synthesized using SuperScript III Reverse Transcriptase (Invitrogen) according
607 to the manufacturer's instructions using random oligonucleotide primers (Invitrogen). qPCR
608 analyses for *cps* locus expression were performed on a Mastercycler ep realplex instrument
609 (Eppendorf). Expression of each of the 8 *cps* synthesis loci was quantified using primers to a
610 single gene in each locus (primers are listed in **Table S4**) and normalized to a standard curve of
611 DNA from wild-type *B. thetaiotaomicron*. Relative abundance of expression for each locus was
612 then calculated. A custom-made SYBR-based master mix was used for qPCR: 20 µl reactions

were made with ThermoPol buffer (New England Biolabs), and containing 2.5 mM $MgSO_4$, 0.125 mM dNTPs, 0.25 µM each primer, 0.1 µl of a 100 X stock of SYBR Green I (Lonza), and 500 U Hot Start *Taq* DNA Polymerase (New England Biolabs). 10 ng of cDNA was used for each sample, and samples were run in duplicate. A touchdown protocol with the following cycling conditions was used for all assays: 95 °C for 3 minutes, followed by 40 cycles of 3 seconds at 95 °C, 20 seconds of annealing at a variable temperature, and 20 seconds at 68 °C. The annealing temperature for the first cycle was 58 °C, then dropped one degree each cycle for the subsequent 5 cycles. The annealing temperature for the last 34 cycles was 52 °C. These cycling conditions were followed by a melting curve analysis to determine amplicon purity.

Transcriptomic analysis of* B. thetaiotaomicron *after phage infection.

Whole genome transcriptional profiling of wild-type and acapsular *B.thetaiotaomicron* infected with live or heat-killed ARB25 was conducted using total bacterial RNA that was extracted the same as described above (Qiagen RNAEasy, Turbo DNA-free kit) and then treated with Ribo-Zero rRNA Removal Kit (Illumina Inc.) and concentrated using RNA Clean and Concentrator -5 kit (Zymo Research Corp, Irvine, CA). Sequencing libraries were prepared using TruSeq barcoding adaptors (Illumina Inc.), and 24 samples were multiplexed and sequenced with 50 base pair single end reads in one lane of an Illumina HiSeq instrument at the University of Michigan Sequencing Core. Demultiplexed samples were analyzed via Arraystar software (DNASTAR, Inc.) using RPKM normalization and the default parameters. Changes in gene expression in response to live ARB25 infection were determined by comparison to the heat-killed reference: retained were genes with $\geq$ 3-fold expression changes up or down, Benjamini-Hochberg corrected *P* value $\leq$ 0.05, and genes for which the lower value in the fold-change

636 calculation was at least 1% of the mean RPKM-normalized value for all genes in the
637 transcriptome. The latter cutoff was implemented to reduce the noise effects of changes in genes
638 with very low expression values.
639
640 ***PCR and sequencing of phase variable* B. thetaiotaomicron *chromosomal loci.***
641 We found that each of the 8 chromosomal loci shown in **Figure 5C** had nearly identical
642 301 bp promoter sequences, including both of the imperfect palindromes that we predict to
643 mediate recombination and the intervening sequence at each locus. While the 8 S-layer genes
644 and the 7/8 of the upstream regions encoding putative tyrosine recombinases (all but the BT1927
645 region) shared significant nucleotide identity, we were able to design primers that were specific
646 to regions upstream and downstream of each invertible promoter and used these to generate an
647 amplicon for each locus that spanned the predicted recombination sites. After gel extracting a
648 PCR product of the expected size for each locus, which should contain promoter orientations in
649 both the "on" and "off" orientations, we performed a second PCR using a universal primer that
650 lies within the 301 bp sequence of each phase-variable promoter and extended to unique primers
651 that anneal within each S layer protein encoding gene. Bands of the expected size were excised
652 from agarose gels, purified and sequenced using the primer that anneals within each S layer
653 encoding gene to determine if the predicted recombined "on" promoter orientation can be
654 detected. (Note that the assembled *B. thetaiotaomicron* genome architecture places all of these
655 promoters in the proposed "off" orientation. We were able to detect 6/8 of these loci in the "on"
656 orientation in ARB25-treated cells by this method, **Figure S6**.) A similar approach was used to
657 determine the re-orientation of DNA fragments in the *B. thetaiotaomicron* PUL shufflon shown
658 in **Figure 5D**, using PCR primer amplicons positioned according to the schematic followed by

659 sequencing with the primer on the "downstream" end of each amplicon according to its position
660 relative to the shuffled promoter sequence. For a list of primers used see **Table S4**.
661
662 ***Construction of acapsular* B. thetaiotaomicron *S-layer 'ON' and S-layer 'OFF' mutants.***
663 Acapsular *B. thetaiotaomicron* S-layer 'ON' and 'OFF' mutants (Δ*cps* BT1927-ON and
664 Δ*cps* BT1927-OFF, respectively) were created using the Δ*tdk* allelic exchange system[45]. To
665 generate homologous regions for allelic exchange, the primers BT_1927_XbaI-DR and
666 BT_1927_SalI-UF were used to amplify the BT1927-ON and BT1927-OFF promoters from the
667 previously-constructed BT1927-ON and BT1927-OFF strains[31] via colony PCR using Q5 High
668 Fidelity DNA polymerase (New England Biolabs). Candidate Δ*cps* BT1927-ON and Δ*cps*
669 BT1927-OFF mutants were screened by PCR using the primer pair BT1927_Diagnostic_R and
670 BT1927_Diagnostic_F and confirmed by Sanger sequencing using these diagnostic primers. All
671 plasmids and primers are listed in **Supplementary Tables S3** and **S4**, respectively.
672
673 ***Data representation and statistical analysis***.
674 The heatmaps for **Figures 1** and **S2** and the dendrogram for **Figure 1** were generated in R
675 using the "heatmap" function. Other graphs were created in Prism software (GraphPad Software,
676 Inc., La Jolla, CA). Statistical significance in this work is denoted as follows unless otherwise
677 indicated: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$. Statistical analyses other than
678 Dirichlet regression were performed in Prism. Dirichlet regression was performed in R using the
679 package "DirichletReg" (version 0.6-3), employing the alternative parameterization as used
680 previously[4,46]. Briefly, the parameters in this distribution are the proportions of relative *cps* gene
681 expression and the total *cps* expression, with *cps7* expression used as a reference. The variable of

interest used in **Figure 2G** is bacterial strain, whereas the variable of interest used in **Figure 4B** is phage viability (live versus heat-killed phage). Precision was allowed to vary by group given this model was superior to a model with constant precision, as determined by a likelihood ratio test at significance level $p < 0.05$.

Acknowledgements

We thank Rey Honrada at the San Jose-Santa Clara Wastewater Treatment Plant and the staff at the Ann Arbor Wastewater treatment plant for assistance in collecting primary sewage effluent. This work was funded by NIH grants (GM099513 and DK096023 to E.C.M), an NIH postdoctoral NRSA (5T32AI007328 to A.J.H.), a Stanford University School of Medicine Dean's Postdoctoral Fellowship (A.J.H.), and the NIH Cellular Biotechnology Training Program (N.T.P., T32GM008353).

Author Contributions

NTP, AJH, BDM, JOG, and SS performed the experiments. NTP, AJH, and ECM designed the experiments, and analyzed and interpreted the data. JLS and ECM provided tools and reagents. NTP, AJH and ECM prepared the display items. NTP, AJH and ECM wrote the paper. All authors edited the manuscript prior to submission.

References

1 Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science (New York, N.Y.)* **308**, 1635-1638, doi:10.1126/science.1110591 (2005).

2 Cockburn, D. W. & Koropatkin, N. M. Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease. *Journal of molecular biology* **428**, 3230-3252, doi:10.1016/j.jmb.2016.06.021 (2016).

3 Hooper, L. V., Littman, D. R. & Macpherson, A. J. Interactions between the microbiota and the immune system. *Science (New York, N.Y.)* **336**, 1268-1273, doi:10.1126/science.1223490 (2012).

4 Porter, N. T., Canales, P., Peterson, D. A. & Martens, E. C. A Subset of Polysaccharide Capsules in the Human Symbiont Bacteroides thetaiotaomicron Promote Increased Competitive Fitness in the Mouse Gut. *Cell host & microbe* **22**, 494-506.e498, doi:10.1016/j.chom.2017.08.020 (2017).

5 Donia, M. S. *et al.* A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **158**, 1402-1414, doi:10.1016/j.cell.2014.08.032 (2014).

6 Coyne, M. J. & Comstock, L. E. Niche-specific features of the intestinal bacteroidales. *Journal of bacteriology* **190**, 736-742, doi:10.1128/jb.01559-07 (2008).

7 Xu, J. *et al.* Evolution of symbiotic bacteria in the distal human intestine. *PLoS biology* **5**, e156, doi:10.1371/journal.pbio.0050156 (2007).

8 Peterson, D. A., McNulty, N. P., Guruge, J. L. & Gordon, J. I. IgA response to symbiotic bacteria as a mediator of gut homeostasis. *Cell host & microbe* **2**, 328-339, doi:10.1016/j.chom.2007.09.013 (2007).

9 Round, J. L. *et al.* The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science (New York, N.Y.)* **332**, 974-977, doi:10.1126/science.1206095 (2011).

10 Shen, Y. *et al.* Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell host & microbe* **12**, 509-520, doi:10.1016/j.chom.2012.08.004 (2012).

11 Neff, C. P. *et al.* Diverse Intestinal Bacteria Contain Putative Zwitterionic Capsular Polysaccharides with Anti-inflammatory Properties. *Cell host & microbe* **20**, 535-547, doi:10.1016/j.chom.2016.09.002 (2016).

12 Patrick, S. *et al.* Twenty-eight divergent polysaccharide loci specifying within- and amongst-strain capsule diversity in three strains of Bacteroides fragilis. *Microbiology (Reading, England)* **156**, 3255-3269, doi:10.1099/mic.0.042978-0 (2010).

13 Porter, N. T. & Martens, E. C. The Critical Roles of Polysaccharides in Gut Microbial Ecology and Physiology. *Annual review of microbiology* **71**, 349-369, doi:10.1146/annurev-micro-102215-095316 (2017).

14 Reyes, A. *et al.* Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **466**, 334-338, doi:10.1038/nature09199 (2010).

15 Minot, S. *et al.* The human gut virome: inter-individual variation and dynamic response to diet. *Genome research* **21**, 1616-1625, doi:10.1101/gr.122705.111 (2011).

16 Duerkop, B. A. *et al.* Murine colitis reveals a disease-associated bacteriophage community. *Nature microbiology*, doi:10.1038/s41564-018-0210-y (2018).

17 Manrique, P. *et al.* Healthy human gut phageome. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 10400-10405, doi:10.1073/pnas.1601060113 (2016).

18 Booth, S. J., Van Tassell, R. L., Johnson, J. L. & Wilkins, T. D. Bacteriophages of Bacteroides. *Reviews of infectious diseases* **1**, 325-336 (1979).

19 Cooper, S. W., Szymczak, E. G., Jacobus, N. V. & Tally, F. P. Differentiation of Bacteroides ovatus and Bacteroides thetaiotaomicron by means of bacteriophage. *Journal of clinical microbiology* **20**, 1122-1125 (1984).

20 Keller, R. & Traub, N. The characterization of Bacteroides fragilis bacteriophage recovered from animal sera: observations on the nature of bacteroides phage carrier cultures. *The Journal of general virology* **24**, 179-189, doi:10.1099/0022-1317-24-1-179 (1974).

21 Shkoporov, A. N. *et al.* PhiCrAss001 represents the most abundant bacteriophage family in the human gut and infects Bacteroides intestinalis. *Nature communications* **9**, 4781, doi:10.1038/s41467-018-07225-7 (2018).

22 Puig, A., Araujo, R., Jofre, J. & Frias-Lopez, J. Identification of cell wall proteins of Bacteroides fragilis to which bacteriophage B40-8 binds specifically. *Microbiology (Reading, England)* **147**, 281-288, doi:10.1099/00221287-147-2-281 (2001).

23 Chatzidaki-Livanis, M., Weinacht, K. G. & Comstock, L. E. Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont Bacteroides fragilis. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 11976-11980, doi:10.1073/pnas.1005039107 (2010).

24 Rogers, T. E. *et al.* Dynamic responses of Bacteroides thetaiotaomicron during growth on glycan mixtures. *Molecular microbiology* **88**, 876-890, doi:10.1111/mmi.12228 (2013).

25 Hickey, C. A. *et al.* Colitogenic Bacteroides thetaiotaomicron Antigens Access Host Immune Cells in a Sulfatase-Dependent Manner via Outer Membrane Vesicles. *Cell host & microbe* **17**, 672-680, doi:10.1016/j.chom.2015.04.002 (2015).

26 Chatzidaki-Livanis, M., Coyne, M. J. & Comstock, L. E. A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of Bacteroides fragilis. *Journal of bacteriology* **191**, 7288-7295, doi:10.1128/jb.00500-09 (2009).

27 Pelkonen, S., Aalto, J. & Finne, J. Differential activities of bacteriophage depolymerase on bacterial polysaccharide: binding is essential but degradation is inhibitory in phage infection of K1-defective Escherichia coli. *Journal of bacteriology* **174**, 7757-7761 (1992).

28 Krinos, C. M. *et al.* Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature* **414**, 555-558, doi:10.1038/35107092 (2001).

29 Martens, E. C., Koropatkin, N. M., Smith, T. J. & Gordon, J. I. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *The Journal of biological chemistry* **284**, 24673-24677, doi:10.1074/jbc.R109.022848 (2009).

30 Braun, V. FhuA (TonA), the career of a protein. *Journal of bacteriology* **191**, 3431-3436, doi:10.1128/jb.00106-09 (2009).

31 Taketani, M., Donia, M. S., Jacobson, A. N., Lambris, J. D. & Fischbach, M. A. A Phase-Variable Surface Layer from the Gut Symbiont Bacteroides thetaiotaomicron. *mBio* **6**, e01339-01315, doi:10.1128/mBio.01339-15 (2015).

32 Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell host & microbe* **4**, 447-457, doi:10.1016/j.chom.2008.09.007 (2008).

33 Martens, E. C., Roth, R., Heuser, J. E. & Gordon, J. I. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. *The Journal of biological chemistry* **284**, 18445-18457, doi:10.1074/jbc.M109.008094 (2009).

34 Xu, J. *et al.* A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science (New York, N.Y.)* **299**, 2074-2076, doi:10.1126/science.1080029 (2003).

35 Nakayama-Imaohji, H. *et al.* Identification of the site-specific DNA invertase responsible for the phase variation of SusC/SusD family outer membrane proteins in Bacteroides fragilis. *Journal of bacteriology* **191**, 6003-6011, doi:10.1128/jb.00687-09 (2009).

36 Kuwahara, T. *et al.* Genomic analysis of Bacteroides fragilis reveals extensive DNA inversions regulating cell surface adaptation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14919-14924, doi:10.1073/pnas.0404172101 (2004).

37 Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107-118, doi:10.1016/j.cell.2005.05.007 (2005).

38 Donaldson, G. P. *et al.* Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science (New York, N.Y.)* **360**, 795-800, doi:10.1126/science.aaq0926 (2018).

814 39 Tomlinson, S. & Taylor, P. W. Neuraminidase associated with coliphage E that
815 specifically depolymerizes the Escherichia coli K1 capsular polysaccharide. *Journal of*
816 *virology* **55**, 374-378 (1985).
817 40 Gupta, D. S. *et al.*
818 41 Barr, J. J. *et al.* Bacteriophage adhering to mucus provide a non-host-derived immunity.
819 *Proceedings of the National Academy of Sciences of the United States of America* **110**,
820 10771-10776, doi:10.1073/pnas.1305923110 (2013).
821 42 Norman, J. M. *et al.* Disease-specific alterations in the enteric virome in inflammatory
822 bowel disease. *Cell* **160**, 447-460, doi:10.1016/j.cell.2015.01.002 (2015).
823 43 Tartera, C., Araujo, R., Michel, T. & Jofre, J. Culture and decontamination methods
824 affecting enumeration of phages infecting Bacteroides fragilis in sewage. *Applied and*
825 *environmental microbiology* **58**, 2670-2673 (1992).
826 44 Araujo, R. *et al.* Optimisation and standardisation of a method for detecting and
827 enumerating bacteriophages infecting Bacteroides fragilis. *Journal of virological methods*
828 **93**, 127-136 (2001).
829 45 Koropatkin, N., Martens, E. C., Gordon, J. I. & Smith, T. J. Structure of a SusD
830 homologue, BT1043, involved in mucin O-glycan utilization in a prominent human gut
831 symbiont. *Biochemistry* **48**, 1532-1542, doi:10.1021/bi801942a (2009).
832 46 MJ, M. DirichletReg: Dirichlet Regression in R. (2015).
833

834 **Figure Legends**

835 **Figure 1**. Host range of *B. thetaiotaomicron* phages on strains expressing different CPS.
836 Seventy-one bacteriophages were isolated and purified on the wild-type, Δ*cps* (acapsular), or the
837 8 single CPS-expressing *B. thetaiotaomicron* strains. High titer phage stocks were prepared on
838 their "preferred host strain," which was the strain that yielded the highest titer of phages in a pre-
839 screen of phage host range; phages were then tested in a quantitative host range assay. Tenfold
840 serial dilutions of each lysate ranging from approximately 10^6 to 10^3 plaque-forming units (PFU)
841 / ml were spotted onto top agar plates containing each of the 10 bacterial strains. Plates were then
842 incubated overnight, and plaques on each host were counted. Titers were calculated for each host
843 and normalized to the titer on the "preferred host strain" for each replicate and replicates (n = 3)
844 were averaged. The phages were then clustered based on their plaquing efficiencies on the
845 different strains (see *Methods*). Each row in the heatmap corresponds to an individual phage,
846 whereas each column corresponds to one of the 10 host strains. Images at the far right of the

847 figure illustrate the range of plaque morphologies of select phages from the collection (see
848 **Figure S1** for images of plaques for all phages). ARB25, a phage that is the focus of deeper
849 investigation is highlighted in red text. Scale bar = 2mm.
850
851 **Figure 2**. Infection by the Branch 3 phages ARB72 (A), ARB78 (B), ARB82 (C), ARB101 (D)
852 and ARB105 (E) is inhibited by eliminating most or all of the permissive CPS from wild-type *B.*
853 *thetaiotaomicron*. Each phage was titered on the wild-type strain, the acapsular strain, their
854 respective preferred host strain (blue bars), and a set of bacterial strains harboring limited CPS
855 locus deletions that correspond to their predetermined host range (n = 6 replicates/phage). (F)
856 Elimination of permissive CPS from Branch 2 phage ARB25 reduces infection, but complete
857 reduction of infection only occurs in the context of deleting more than one permissive CPS. The
858 number of replicates conducted on each strain is annotated in parentheses next to the strain name.
859 (G) Relative *cps* locus expression of the 8 *cps* loci in the indicated strains. (H) Representative
860 pictures of phage plaques on the indicated host strains. The top row of images for each phage is
861 unaltered; background and unnaturally saturated pixels were removed from images in in the
862 bottom row to facilitate plaque visualization. Scale bar = 2mm. For panels A-F, significant
863 differences in phage titers on the preferred host strain were calculated via Kruskal Wallis test
864 followed by Dunn's multiple comparisons test. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, **** $p <$
865 0.0001. For panel G, deletion of *cps* loci resulted in significant changes in *cps* gene expression
866 for the missing loci; additionally, significant changes in *cps2* expression were observed in Δ4
867 and Δ1,4 strains ($p < 0.05$ for each change, determined by Dirichlet regression).
868

869 **Figure 3.** Effect of CPS and phage infection on bacterial growth. Ten strains: the wild-type
870 (WT), the acapsular strain (Δcps), or the eight single CPS-expressing strains were infected with
871 either live or heat-killed ARB25. Growth was monitored via optical density at 600 nm (OD_{600})
872 on an automated plate reading instrument as described in *Methods*.

873

874 **Figure 4.** ARB25 infection of wild-type *B. thetaiotaomicron* causes altered *cps* gene expression.
875 Wild-type *B. thetaiotaomicron* was infected with live or heat-killed ARB25 at an MOI of ~1. (A)
876 Growth was monitored by measuring OD_{600} every 15-30 minutes. (B) *cps* gene transcript
877 analysis was carried out by qPCR. The end of the growth curve in panel A represents the point at
878 which cultures were harvested for qPCR analysis (e.g. the first observed time point where culture
879 surpassed OD_{600} of 0.6). Significant changes in *cps1*, *cps3*, and *cps4* expression were observed
880 between groups treated with live or heat-killed ARB25 ($p < 0.01$ for each, determined by
881 Dirichlet regression).

882

883 **Figure 5.** Infection of acapsular *B. thetaiotaomicron* selects for the expression of multiple phase-
884 variable loci. (A) Wild-type *B. thetaiotaomicron* was infected with ARB25 or was alternatively
885 exposed to heat-killed ARB25 and cultures were grown to OD_{600}=0.6-0.7. Cells were harvested
886 and RNA-seq analysis was carried out as described in *Methods* (n=3 independent experiments
887 for each treatment group). Transcript abundance was compared between treatment groups and
888 fold change between ARB25-treated cells versus cells treated with heat killed ARB25 was
889 calculated and is plotted against the P value (T test with Benjamini-Hochberg correction) for
890 each fold change calculation. (B) Acapsular *B. thetaiotaomicron* was treated with ARB25 or
891 heat-killed ARB25 and fold change in transcript abundance was calculated, as described in panel

892 A. (C) Among the genes with increased expression in post-infected acapsular *B.*
893 *thetaiotaomicron*, 24 genes were part of 8 different gene clusters predicted to encode tyrosine
894 recombinases along with outer membrane lipoproteins and OmpA-like proteins. These gene
895 clusters are shown. The numbers inside the gene schematics represent the fold change in
896 expression in ARB25-treated cells relative to those treated with heat-killed ARB25. Flanking the
897 promoters of each of these loci are pairs of imperfect, 17 nucleotide palindromic repeats. qPCR
898 analysis and amplicon sequencing of each orientation of these 8 promoters revealed expected
899 increases in orientation to the "ON" position in ARB25-exposed acapsular *B. thetaiotaomicron*.
900 (D) Another chromosomal locus with signatures of phage-selected recombination was identified
901 by RNA-seq. Specifically, 3 of 4 genes in an operon (BT1042-BT1045) were significantly down-
902 regulated after exposure to phage and 5 genes in an adjacent operon (BT1046-BT1051) were up-
903 regulated an average of 8.3-fold. (E) PCR using oligonucleotides flanking direct repeats within
904 the BT1032-BT1053 locus (green dumbbells, panel D) were used to demonstrate locus
905 architecture in wild type *B. thetaiotaomicron* and in a mutant lacking the tyrosine recombinase
906 within this locus (*B. thetaiotaomicron* ΔBT1041).

907

908 **Figure 6**. S-layer differentially affects *B. thetaiotaomicron* susceptibility to ARB25 and the
909 related phage SJC01. Acapsular *B. thetaiotaomicron* S-layer 'ON' and 'OFF' mutants (Δ*cps*
910 BT1927-ON and Δ*cps* BT1927-OFF, respectively) were infected with (A) ARB25 or (B) SJC01
911 in liquid culture. Growth was monitored via optical density at 600 nm (OD_{600}) on an automated
912 plate reading instrument as described in *Methods*.

913

914 **Table 1.** Susceptibility of strains to infection by ARB25 after infection and passaging.

915

916 **Figure S1.** Representative pictures of phage plaques for all phages from this study: A) phages
917 from Ann Arbor (ARB); B) phages from San Jose (SJC). The top row of images for each phage
918 are unaltered; background and saturated pixels were removed from images in the bottom row to
919 facilitate viewing of the plaques. Scale bar = 2 mm

920

921 **Figure S2.** Replication of subset of host range assays of *B. thetaiotaomicron*-targeting phages on
922 strains expressing different CPS types. Ten bacteriophages isolated and purified on the wild-
923 type, acapsular, or the 8 single CPS-expressing strains were re-tested in a spot titer assay to
924 determine phage host range. 10-fold serial dilutions of each phage ranging from approximately
925 10^6 to 10^3 plaque-forming units (PFU) / ml were spotted onto top agar plates containing the 10
926 bacterial strains. Plates were then grown overnight, and phage titers were calculated. Titers are
927 normalized to the titer on the preferred host strain for each replicate. Each row in the heatmap
928 corresponds to a replicate for an individual phage, whereas each column corresponds to one of
929 the 10 host strains. One to three replicates of the assay were conducted for each phage by the two
930 lead authors (AJH and NTP). Assays were carried out at the same time, and each author used the
931 same set of cultures and phage stocks. For comparison, individual replicates from **Figure 1** are
932 included (marked with *)

933

934 **Figure S3.** A) Phage titers in infected cultures incubated with purified capsules. ARB25 was
935 incubated with purified CPS1 or CPS2 (1 mg/ml) before plating on the acapsular strain, and
936 plaques were counted after overnight incubation. Titers are normalized to mock (H_2O) treatment.

937 No significant differences in titers were found compared to mock treatment, as determined by
938 Welch's t test. B) Wild-type *B. thetaiotaomicron* was infected with live or heat-killed ARB25,
939 and bacterial growth was monitored via optical density at 600 nm (OD600) on an automated
940 plate reader for 12 hours. At 0, 6.02, 8.36, and 11.7 hours post inoculation, replicate cultures
941 were removed and phage levels were titered. No phage were detected in heat-killed controls.

942

943 **Figure S4.** Infection of wild-type *B. thetaiotaomicron* at a low multiplicity of infection and
944 subsequent effects on *cps* ene expression. (A) The wild-type (WT) strain was infected at a low
945 multiplicity of infection (MOI = 1 x 10^{-4}) of live or heat-killed ARB25, and bacterial growth was
946 monitored via OD_{600}. (B) RNA was harvested from cultures after reaching an OD of 0.6-0.7,
947 cDNA was generated, and relative expression of the 8 *cps* loci was determined by qPCR.

948

949 **Figure S5.** Single replicates of *cps* expression in heat-killed versus live phage-treated *B.*
950 *thetaiotaomicron*. Relative *cps* CPS transcript abundance in ARB25 infection experiments at
951 high MOI (A) and low MOI (B). In the high MOI experiment, replicate 2 showed higher starting
952 expression of the non-permissive CPS3 compared to others. In the low MOI experiment,
953 replicate 3 showed higher starting expression of the non-permissive CPS3. In both experiments,
954 post phage-exposed replicates displayed nearly identical CPS expression profiles characterized
955 by high expression of CPS3.

956

957 **Figure S6.** Determination of phase-variable promoter switching for six loc encoding putative S-
958 layer proteins. The hypothesis that the promoters associated with seven newly identified *B.*

959 *thetaiotaomicron* S-layer like lipoproteins was validated using a PCR amplicon sequencing
960 strategy. Because of high nucleotide identity in both the regions flanking the 7 new loci, a nested
961 PCR approach was required to specifically amplify and sequence each site. In the first step, a
962 primer lying in each S-layer gene (**Table S3** "S-layer gene" primers) was oriented towards the
963 promoter and used in a PCR extension to a primer in the upstream recombinase gene (**Table S3**
964 "recombinase gene 3" primer). The products of this PCR were purified without gel extraction
965 and used in a second reaction with a nested primer that lies internal to the previous recombinase
966 gene primer (**Table S3** "recombinase 2" primer). The expected PCR products from this reaction,
967 which are ~1 kb and span promoter sequences in both the ON and OFF orientations, were
968 excised and used for an orientation-specific PCR using the original S-layer gene primer for each
969 site and a universal primer (green schematic) that was designed for each promoter and is oriented
970 to extend upstream of the S-layer gene (e.g., OFF orientation). Resulting products from this third
971 reaction, which should correspond to the ON orientation if a promoter inversion has occurred in
972 some cells, were obtained for 5/7 of the newly identified loci and the BT1927 S-layer locus as a
973 control. In all cases in which an amplicon and sequence were obtained, the expected
974 recombination occurred between the inverted repeat site proximal to the S-layer gene start (new
975 DNA junction), which would orient the promoter to enable expression of the downstream S-layer
976 gene. The sequences shown are the consensus between forward and reverse reads for each
977 amplicon. The putative core promoter -7 sequence is shown in bold/red text, the coding region of
978 each S-layer gene is shown in bold/blue text and the S-layer gene proximal recombination site is
979 noted and highlighted in bold/gold text. Note that the 5'-end of the sequenced amplicon was not
980 resolved for the BT2486 locus.

981

Figure S7. Recombination between the genes BT1040, BT1042, and BT1046. (A) Pfam domain schematics of the amino acid sequences of these three genes highlighting that BT1040 and BT1046, as originally assembled in the *B. thetaiotaomicron* genome sequence, lack additional N-terminal sequences that are present on BT1042. (B) Sequencing of the 8 PCR amplicons schematized in **Figure 5C**. Amplicons 1, 5 and 8 represent the original genome architecture, while the others represent inferred recombination events that are validated here by sequencing. The 5' and 3' ends of the BT1042, BT1040 and BT1046 genes are color-coded to assist in following their connectivity changes after recombination. A series of single-nucleotide polymorphisms (SNPs) present in BT1042, downstream of the proposed recombination site, are highlighted in yellow. The transfer of these SNPs to a fragment containing the 5' end of BT1040 (Amplicon 4) was used to narrow the recombination region to the 7 nucleotide sequence highlighted in red. Additional SNPs that are specific to the regions upstream of this recombination site are shown in white text for each sequence.

Figure S8. The BT1033-52 locus does not affect susceptibility of acapsular *B. thetaiotaomicron* to ARB25. Ten-fold serial dilutions of ARB25 were spotted onto lawns of *B. thetaiotaomicron* Δ*cps* (n=5) and *B. thetaiotaomicron* Δ*cps* ΔBT1033-52 (n=5, n=3 independent clones each). Plaquing efficiency was determined by normalizing plaque counts on *B. thetaiotaomicron* Δ*cps* ΔBT1033-52 relative to plaque counts on *B. thetaiotaomicron* Δcps for each replicate. Statistical significance was determined using the Mann-Whitney test.

Table S1. Phages used in this study and details on their isolation.

Table S2. Genes that are differentially regulated in post ARB25-infected wild-type and acapsular *B. thetaiotaomicron*.

Table S3. Bacterial strains and plasmids used in this study.

Table S4. Primers used in this study.

Figure 1

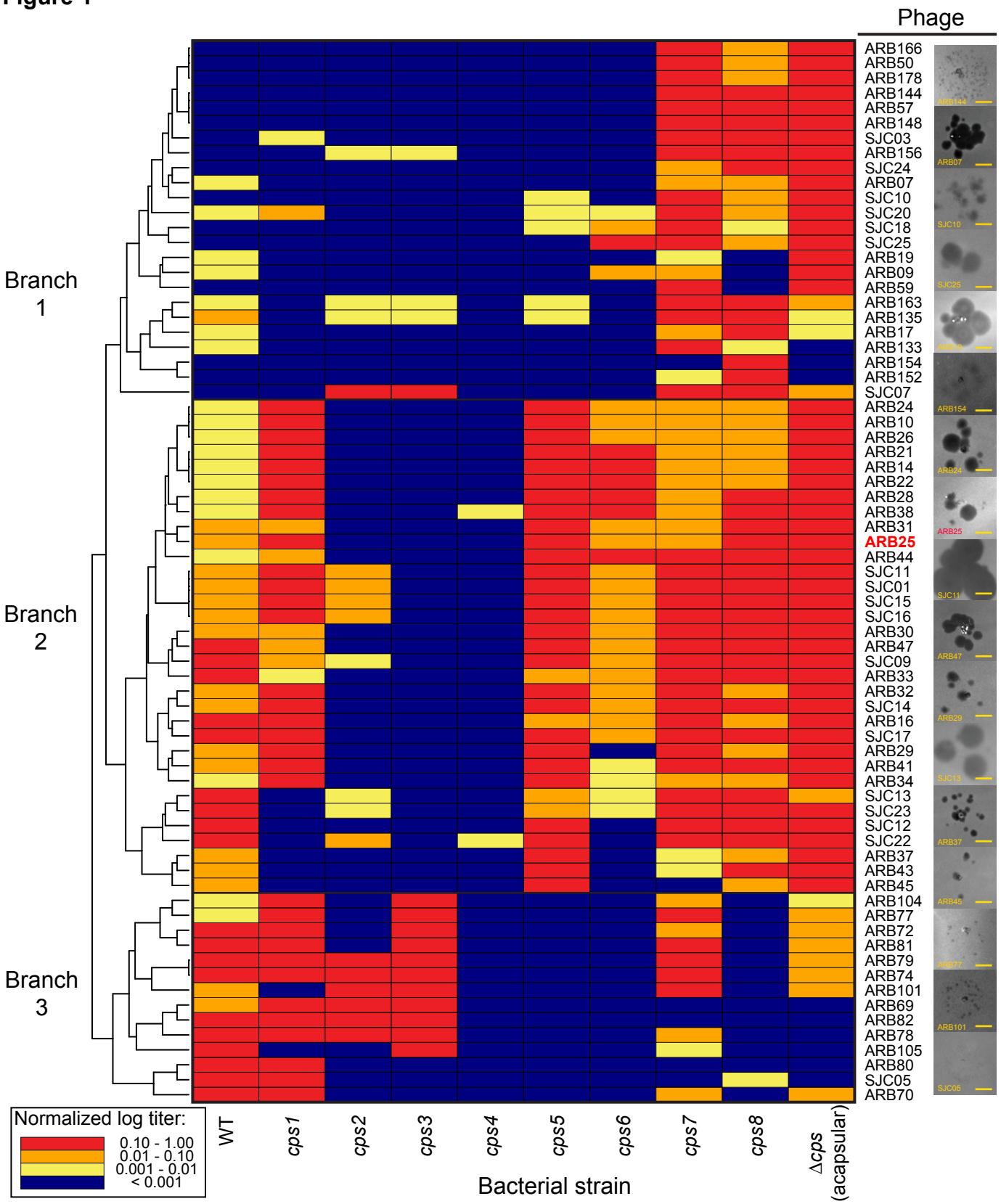

Figure 2

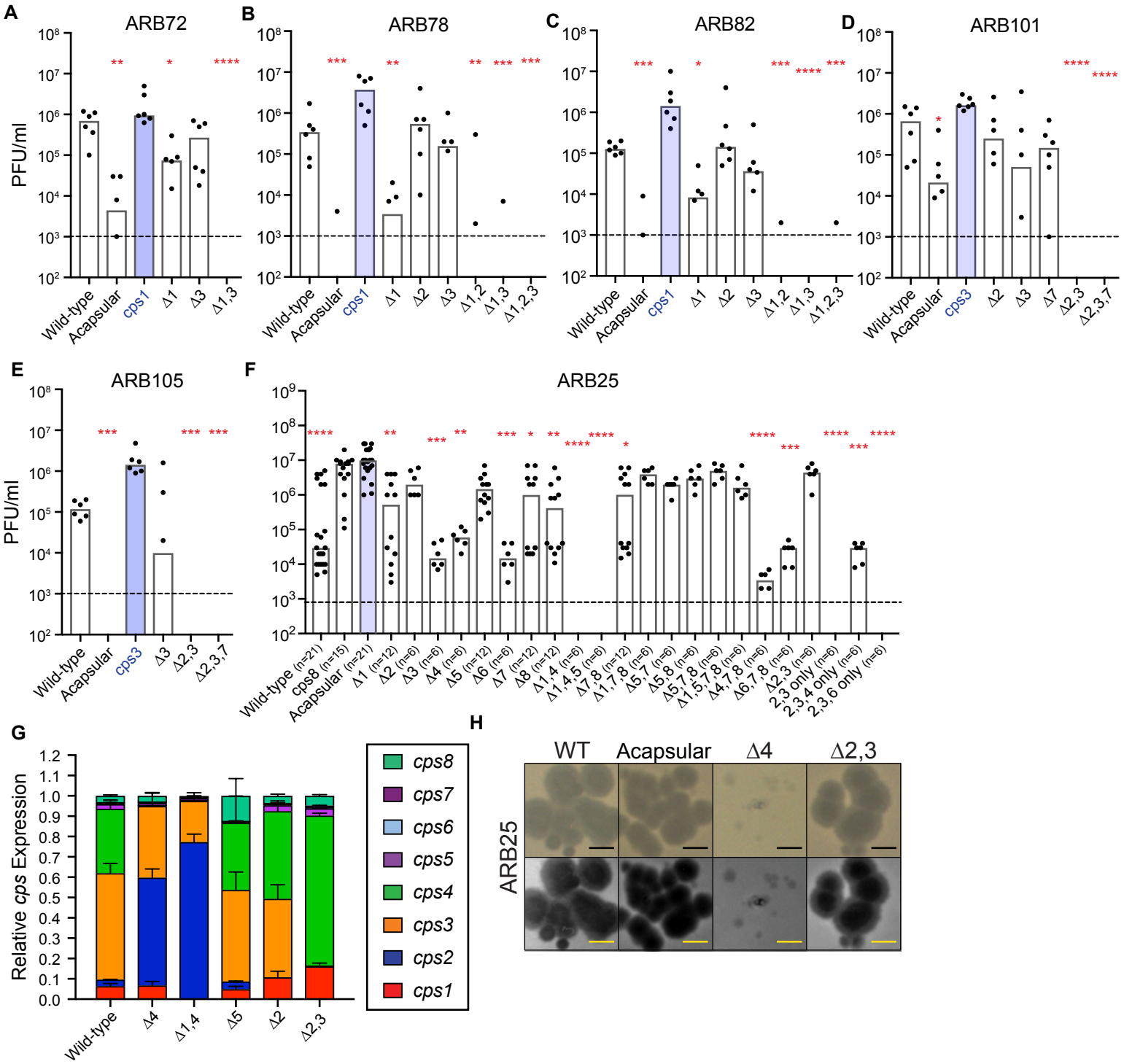

Figure 3

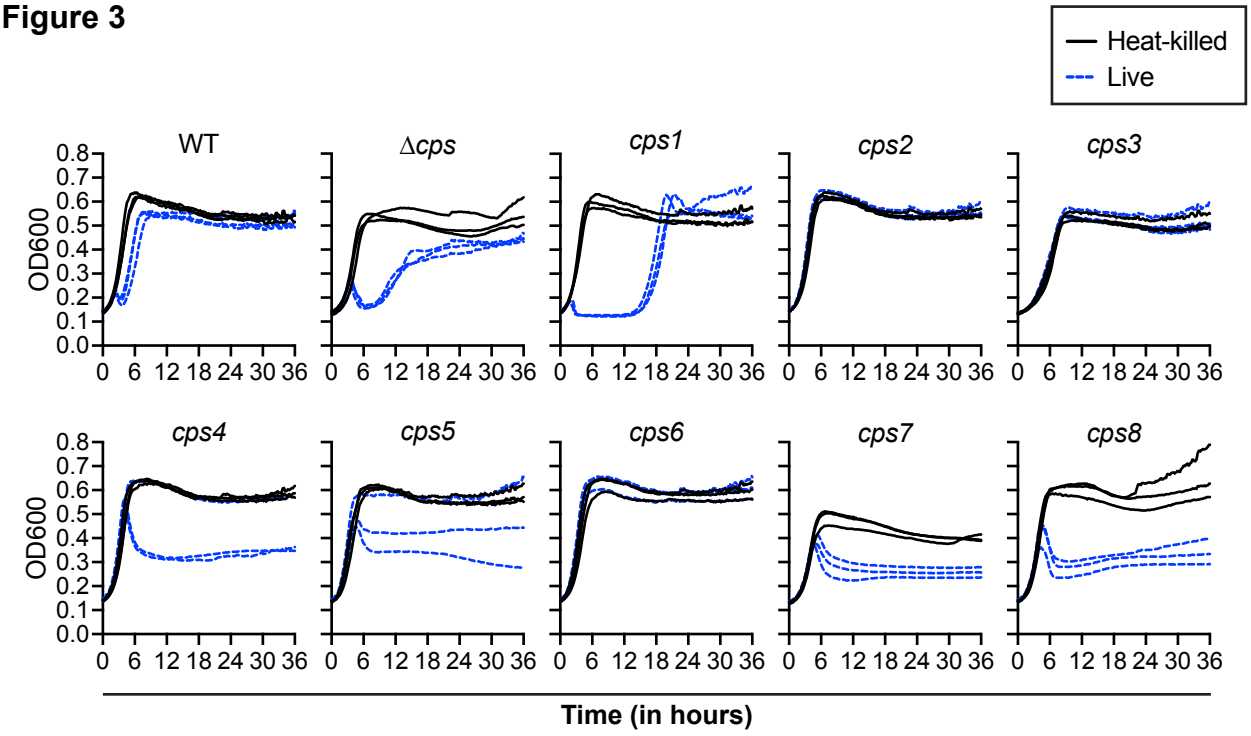

Figure 4

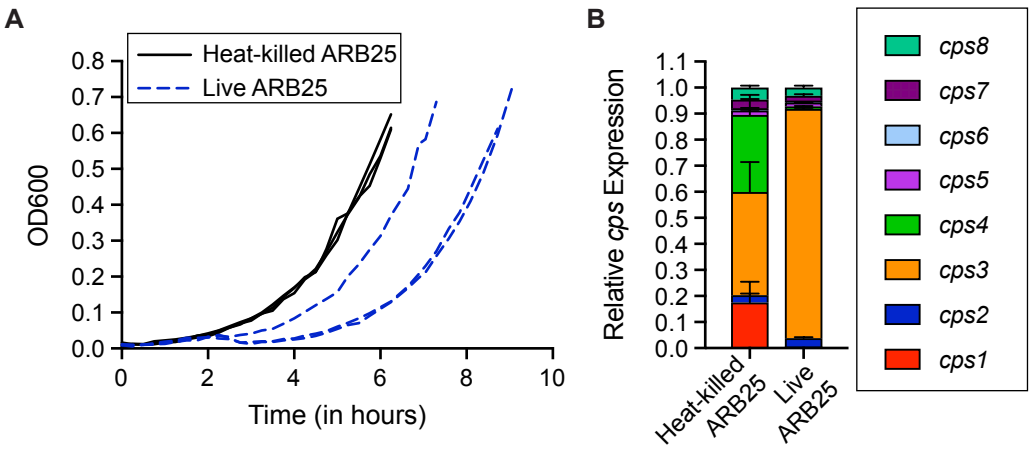

Figure 5

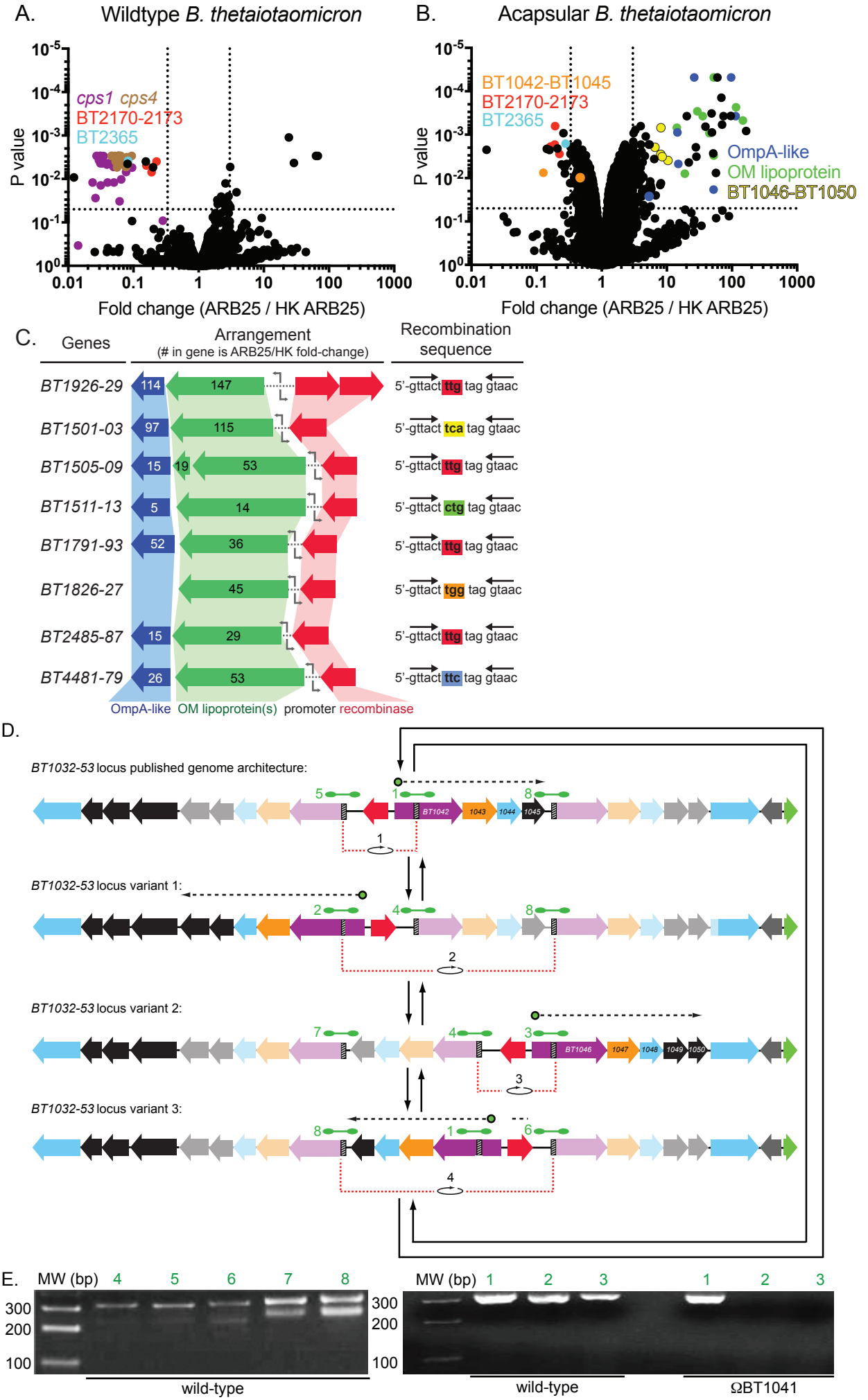

Figure 6

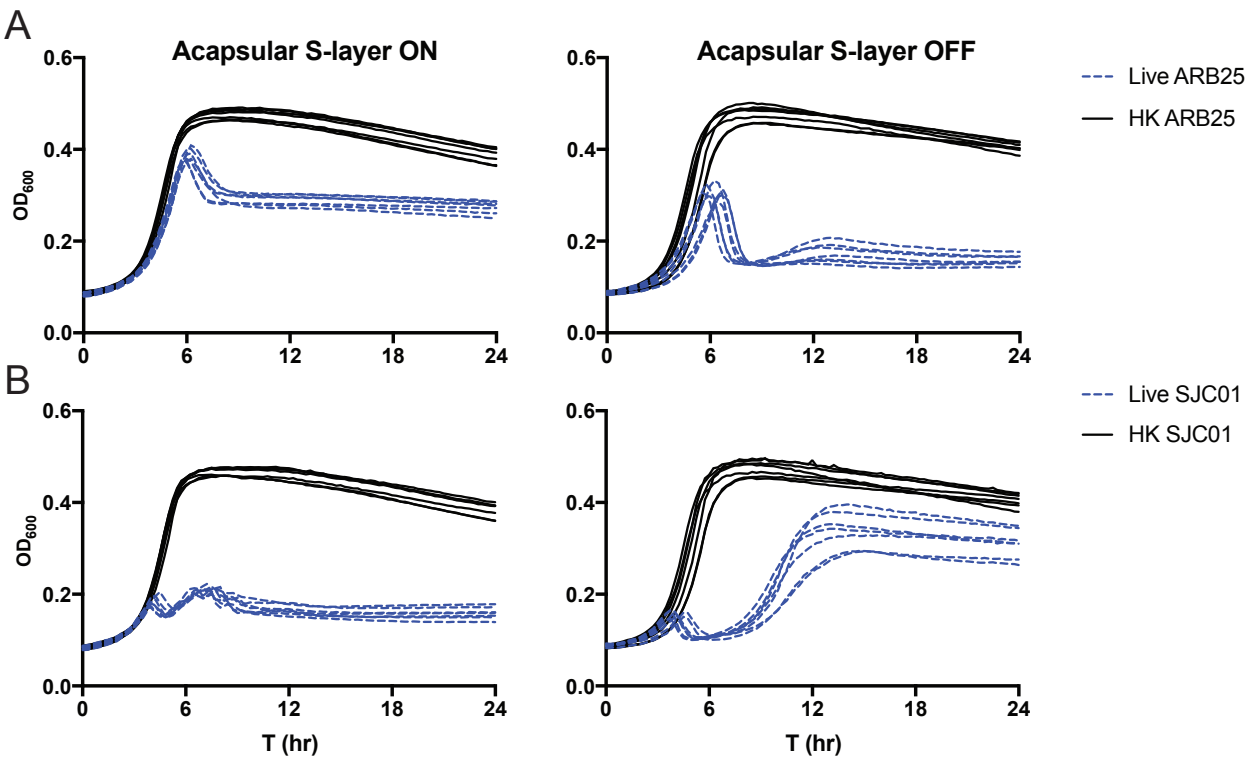

Table 1. Bacterial susceptibility to secondary ARB25 infection

Strain	# susceptible isolates after first infection	# susceptible isolates after second infection
Wild-type	9/9 (100%)	9/17 (53%)
cps1	9/9 (100%)	15/17 (88%)
cps2	0/9 (0%)	0/18 (0%)
cps4	2/9 (22%)	8/11 (73%)
Acapsular	9/9 (100%)	10/16 (63%)
Total of phage-susceptible strains[a]	**29/36 (81%)**	**42/61 (69%)**

[a]Wild-type, cps1, cps4, and acapsular