HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2019 July 11.

Published in final edited form as:

Science. 2019 January 11; 363(6423): 181–187. doi:10.1126/science.aau5238.

Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut

Xiaofang Jiang[1,2,*], **A. Brantley Hall**[2,3,*], **Timothy D. Arthur**[2], **Damian R. Plichta**[2,3], **Christian T. Covington**[2,3], **Mathilde Poyet**[1,2,4], **Jessica Crothers**[5], **Peter L. Moses**[6], **Andrew C. Tolonen**[2,3], **Hera Vlamakis**[2,3], **Eric J. Alm**[1,2,4,†], and **Ramnik J. Xavier**[1,2,3,7,†]

[1]Center for Microbiome Informatics and Therapeutics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[3]Center for Computational and Integrative Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

[4]MIT Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

[5]University of Vermont Medical Center Department of Pathology and Laboratory Medicine, Burlington, VT 05401, USA

[6]Division of Gastroenterology and Hepatology, University of Vermont, Burlington, VT 05401, USA

[7]Gastrointestinal Unit and Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

Abstract

Phase variation, the reversible alternation between genetic states, enables infection by pathogens and colonization by commensals. However, the diversity of phase variation remains underexplored. Here, we developed the PhaseFinder algorithm to quantify DNA inversion-mediated phase variation. A systematic search of 54,875 bacterial genomes identified 4,686 intergenic invertible DNA regions (invertons) revealing an enrichment in host-associated bacteria. Invertons containing promoters often regulate extracellular products, underscoring the importance

[†]Corresponding author - xavier@molbio.mgh.harvard.edu, ejalm@mit.edu.

Author Contributions: Conceptualization, X.J., A.B.H., T.D.A., H.V., A.C.T., E.J.A., R.J.X., Methodology X.J., A.B.H., D.R.P., C.T.C., T.D.A., Software X.J., A.B.H., Validation X.J., A.B.H., Formal Analysis, X.J., A.B.H., D.R.P., C.T.C., T.D.A., Investigation, X.J., A.B.H., T.D.A., Resources, M.P., P.L.M., J.C., Data Curation, X.J., A.B.H., T.D.A., Writing – Original Draft Preparation X.J., A.B.H., Writing – Review & Editing X.J., A.B.H., H.V., A.C.T., E.J.A., R.J.X., Visualization, X.J., A.B.H. Supervision, H.V., A.C.T., E.J.A., R.J.X., Project Administration, E.J.A., R.J.X., Funding Acquisition E.J.A., R.J.X.

[*]Contributed equally

Competing Interest: Eric J. Alm is a co-founder and shareholder of Finch Therapeutics.

Data and materials availability: Data used in the study is available from the NCBI. Isolate genomes: PRJNA496358. Dense longitudinal metagenomic data: PRJNA503484, FMT metagenomic data: PRJNA474024. PhaseFinder is available from GitHub: https://github.com/XiaofangJ/PhaseFinder.

Supplementary Materials:
Materials and Methods
Figures S1–S12
Tables S1–S16
References (28–49)

of surface diversity for gut colonization. We found invertons containing promoters regulating antibiotic resistance genes that shift to the ON orientation following antibiotic treatment in human metagenomic data and *in vitro*, thereby mitigating the cost of antibiotic resistance. We observed that the orientations of some invertons diverge after fecal microbiota transplant, potentially due to individual-specific selective forces.

One Sentence Summary:

Invertons are enriched in host-associated bacterial species and mediate antibiotic resistance and interhost adaptation.

Phase variation is a process employed by bacteria to generate frequent and reversible changes within specific hypermutable loci, introducing phenotypic diversity into clonal populations. Such phenotypic diversity plays an important role in mediating preemptive adaptation to abrupt and severe selective events and is often crucial for infection by pathogens and colonization by commensals(1–5). In bacteria, phase variation often manifests through regions of DNA that invert between two states in a predictable, reversible manner(6). The mechanism of inversion involves enzymes called invertases, which recognize a set of inverted repeats flanking the invertible DNA region and catalyze its inversion in a reversible manner(7). Invertible regions commonly contain promoters oriented such that in the ON orientation, the promoter is poised to activate transcription of an operon(7). In the opposite OFF orientation, the promoter is oriented away from the operon, which is therefore not transcribed(7). Additional types of regulatory elements, such as terminators, may also be contained within these invertible DNA regions(8). Invertases catalyze frequent inversions, for example, one inversion in every 100–1,000 *E. coli* cells, a rate at least three orders of magnitude higher than the rate of point mutations(7, 9, 10). Thus, invertible promoters generate genetic diversity in populations, enabling rapid and reversible adaptation. Isolated studies in specific pathogens and commensals have reported only a few invertible promoters that regulate genes involved in virulence or colonization such as fimbriae, flagella, and capsular polysaccharides(1, 2, 4, 7, 11–16).

Phase variation mediated by DNA inversion is an underexplored mechanism with broad consequences for adaptation to abrupt and severe selective events. In this study, we sought to systematically identify invertons, which we define as single intergenic invertible DNA regions flanked by inverted repeats likely recognized and inverted by invertase proteins in a reversible manner. The term inverton encompasses invertible promoters and intergenic invertible DNA regions containing alternate types of regulatory regions. Through our systematic search for invertons, we aimed to address three long-standing questions regarding this mechanism of regulation: first, how prevalent are invertons; second, what are the functions of genes regulated by invertons; and third, in the context of a host, do individual-specific selective pressures modulate inverton orientation? We show that invertons are widely distributed across bacteria, yielding fundamental insights into bacterial infection and colonization. We confirmed and expanded upon previous observations that the orientations of some invertons regulating capsular polysaccharide biosynthesis operons of human gut bacteria are stable within individuals and divergent between individuals(16). Using a fecal microbiota transplant study to observe the orientation of invertons from the same strain in

multiple individuals, we observed divergences in orientation between donor and patient. We also identified invertible promoters regulating antibiotic resistance genes. We observed that antibiotic treatment results in a shift from the OFF to ON orientation of these invertons in humans, and confirmed that antibiotics cause the orientation shift *in vitro*, which could mitigate the fitness cost of maintaining antibiotic resistance genes in the absence of antibiotics.

We developed the PhaseFinder algorithm to computationally identify invertons and quantify their orientations with genomic or metagenomic sequencing reads by identifying regions flanked by inverted repeats, mimicking their inversion *in silico*, and identifying regions where sequencing reads support both orientations (Fig. S1 and S2). Simulations to benchmark the performance of PhaseFinder reveal that given enough coverage, PhaseFinder can identify most invertons without a substantial rate of false positives. We reasoned that if inversion rates were high, both orientations of invertons would coexist allowing for the identification of invertons in bacterial populations used for genome sequencing. Therefore, we used PhaseFinder to search for invertons in all available NCBI genomes from RefSeq that were sequenced using Illumina paired-end sequencing with data deposited in the NCBI Short Read Archive (SRA). In total, 54,875 bacterial genomes spanning the breadth of cultured bacterial diversity were searched, leading to the discovery of 4,686 putative invertons in 2,414 genomes (Table S1 and S2). Invertons were found in 10 of 19 bacterial phyla. Five phyla harbored invertons in at least 20% of their genomes (Table 1). The lack of a systematic method to identify invertons was the impetus for our study; however, the limited scope of known inverton examples may lead to biases in the PhaseFinder algorithm against invertons with features divergent from known invertible promoters. Additionally, the identification of invertons with PhaseFinder relies on the presence of both orientations of the inverton in sequenced samples. Therefore, applying the PhaseFinder algorithm to additional bacterial genomes derived from diverse conditions and sequenced at higher coverage or with longer reads will likely lead to the discovery of many more invertons.

To explore how invertons are distributed across environmental niches, we used information from ProGenomes and the Joint Genome Institute to categorize species into aquatic, terrestrial, and host-associated habitats(17, 18). The prevalence of invertons was higher in host-associated species (Fisher’s Exact Test, host vs. aquatic FDR p=3.5e-5, odds ratio=6.4; host vs. terrestrial FDR p=5.3e-3, odds ratio=4.8) (Fig. 1A; Table S3). This overall enrichment in the prevalence of invertons is due to phylum-level enrichment in Bacteroidetes (FDR p=2.35e-15) and Proteobacteria (FDR p=8.51e-5) and the fact that all Spirochaetes and Verrucomicrobia found with invertons were associated with vertebrate hosts. Additionally, we observed an increase in the number of invertons per genome in host-associated species (Wilcoxon rank sum test, host vs. aquatic FDR p=2.3e-4, W=361190; host vs. terrestrial FDR p=7.0e-3, W=180680) (Fig. 1B). The enrichment of invertons in host-associated species is not due to habitat-specific differences in coverage (Fig. S3). Overall, our results suggest that diverse species likely use invertons to increase their fitness in host-associated niches.

We acquired detailed information on the niches inhabited by species from the phylum Bacteroidetes (Table S4). In Bacteroidetes, the prevalence of invertons was higher in host-

gut-associated species (Fisher's Exact Test, host-gut vs. aquatic FDR p=3.3e-35, odds ratio = 328.4; host-gut vs. terrestrial FDR p=2.3e-34, odds ratio=220.8; host-gut vs. host-other, FDR p=3.9e-17, odds ratio=35.2) (Fig. 1C). The number of invertons per genome was also higher in host-gut-associated isolates (Wilcoxon rank sum test, host-gut vs. aquatic FDR p=1.8e-26, W=6911; host-gut vs. terrestrial FDR p=2.3e-26, W=6967; host-gut vs. host-other FDR p=4.9e-13, W=3956) (Fig. 1D).

Because of the observed enrichment for invertons in gut species, we performed an in-depth analysis and curation of invertons in a non-redundant selection of 49 representative species from human stool using longitudinal metagenomic data instead of the reads used to assemble reference genomes (Table S5). We identified 459 putative invertons (Table S6), 87.6% of which were from species in the phylum Bacteroidetes, which had an average of 19 invertons per genome. We also identified invertons in additional phyla. We found 53 invertons from two *Akkermansia* species (phylum Verrucomicrobia), two invertons from a *Eubacterium* species (phylum Firmicutes), and one inverton from a *Bifidobacterium* species (phylum Actinobacteria) (Fig. S4).

We categorized the invertons based on their flanking inverted repeats (IR) and identified four canonical motifs in Bacteroidetes: three corresponding to known IR motifs in *B. fragilis* and one uncharacterized motif (Fig. 2A; Table S7 and S8)(13). We also identified a distinctive motif class with tandem repeats within each inverted repeat, which we call motif 0 (Fig. S5). We found conserved promoter consensus motifs in 98% (231/235) of invertons with IR motifs 1–4 (Fig. 2B)(19). In contrast, promoter motifs were not observed in any of the invertible regions of IR motif 0-containing sequences (Table S9). Due to the lack of promoter motifs combined with their location downstream of operons, invertons containing motif 0 may function as another type of regulatory element, such as a phase-variable terminator(8). In gut Bacteroidetes, we identified a total of 255 invertible promoters. Based on the orientation of the promoter consensus motif in relation to surrounding genes, we determined which genes/operons were regulated by invertible promoters and whether the promoter was in the ON or OFF orientation with respect to the downstream gene (Table S7 and S8). In both species of *Akkermansia*, all invertons are flanked by inverted repeats with the same motif (Fig. 2C), contain a promoter motif (Fig. 2D), and lack upstream invertases, suggesting they are all invertible promoters co-regulated by a single, master invertase (EAJ16_05345 in *Akkermansia muciniphila*; MK095134 in *Akkermansia sp. aa_0143*).

Through functional annotation, we found that 73% (228/312) of invertible promoters regulate genes involved in the biosynthesis of polysaccharides, fimbriae, outer membrane proteins, and autotransporters, genes involved in the utilization of polysaccharides, or PEP-CTERM domain-containing proteins (Fig. 2A; Table S7 and S8; Fig. S4). All of these functional categories except polysaccharide utilization are enriched in genes regulated by invertible promoters (Table S10 and S11). Genes regulated by invertible promoters are enriched for cell surface products (e.g. GO:0016020, membrane, p=1.93e-14) (Table S11). The most enriched functional class is capsular polysaccharide biosynthesis loci, for which we found at least one example regulated by an inverton in four of the five major gut phyla: Bacteroidetes, Actinobacteria, Verrucomicrobia, and Firmicutes (Table S11). Previous studies show that a repertoire of phase-variable capsular polysaccharides is necessary for

competitive gut colonization by *Bacteroides species*(15, 20, 21). Our data show that phase variable capsular polysaccharide biosynthesis loci are not just a peculiarity of *Bacteroides* species but are likely a convergent response to a strong selective mechanism in the vertebrate gut that possibly originates from the immune response or phages(20, 21).

We found that invertible promoters could regulate antibiotic resistance genes, such as IBP132, which is upstream of the macrolide resistance gene *ermG* in *Bacteroides stercoris*. To investigate this mechanism *in vivo*, we searched specifically for antibiotic resistance genes regulated by invertible promoters in a cohort of 39 Finnish children, 19 of whom had never been exposed to antibiotics and 20 of whom had been administered 9–15 antibiotic treatments over a three-year period(22). By coupling PhaseFinder with metagenomic assembly analysis, we found three antibiotic resistance genes regulated by invertible promoters including the same *ermG* macrolide resistance gene as IBP132, a *cmeABC* multidrug resistance cassette conferring resistance primarily to macrolides and cephalosporins, and *pmrEL* genes conferring resistance to cationic antimicrobial peptides such as polymixin B (Fig. 3A). At least one antibiotic resistance gene regulated by an invertible promoter was found in 38% (15/39) of individuals from this Finnish cohort: 40% (8/20) of individuals were administered antibiotics and 37% (7/19) of individuals were untreated. Invertons regulating antibiotic resistance genes were also detected in metagenomic data from healthy adults in the USA. All examples of invertons regulating antibiotic resistance were found in *Bacteroides* species, which are increasingly associated with multi-drug resistant infections(23). Surprisingly, all *cmeABC* and *ermG* antibiotic resistance genes were regulated by an identical invertible promoter. Based on genomic context, the *cmeABC/ermG* invertible promoter is likely located on an integrative conjugative element homologous to CTNhyb, an antibiotic resistance transmitting mobile element (Fig. S6)(24).

We examined the orientation of the invertible promoters regulating antibiotic resistance genes in longitudinal metagenomic data from the Finnish children. The mean orientation of the *cmeABC/ermG* invertible promoter was 94% OFF in untreated individuals and 84% OFF in individuals administered antibiotics. We observed an individual in which the *cmeABC/ermG* invertible promoter was 99% OFF 7 days before the macrolide azithromycin was administered and 74% ON 27 days after treatment (Fig. 3B). The *cmeABC/ermG* invertible promoter reverted to 99% OFF within 5 months after azithromycin administration (Fig. 3B). A similar phenomenon was observed in a second individual (Fig. S7). A permutation test revealed that macrolide treatment was positively associated with the ON orientation of the *cmeABC/ermG* invertible promoter (qPCR $p=0.0005$, metagenomic $p=0.0465$). Thus, it appears that macrolides may select for the ON orientation of the *cmeABC/ermG* invertible promoter, and the orientation of the invertible promoter drifts towards OFF after cessation of antibiotic treatment.

To test whether antibiotics select for resistance genes with invertible promoters in the ON orientation, we first verified that the genes regulated by the *cmeABC/ermG* invertible promoter confer macrolide resistance. We cultivated 13 *B. stercoris* isolates with and 1 isolate without the *cmeABC/ermG* invertible promoter upstream of the macrolide resistance gene *ermG* (Table S12) (25). The invertible promoter was primarily ON (>75%) in 10

isolates derived from erythromycin-containing media and primarily OFF (>97%) in three isolates derived from media without erythromycin (Fig. 3C). We established that all *B. stercoris* isolates with *ermG* regulated by the *cmeABC/ermG* invertible promoter were resistant to erythromycin, while the isolate without the *ermG* gene was susceptible (Fig. S8).

Next, we showed that erythromycin treatment selects for the ON orientation of the *cmeABC/ermG* invertible promoter. To quantify changes in the relative abundances of cells with the *cmeABC/ermG* invertible promoter in ON and OFF orientations, we performed qPCR comparing relative amplification using a static primer downstream of the invertible promoter paired with either a primer that amplifies the ON orientation or one that amplifies the OFF orientation (Fig. S9). We transferred *B. stercoris* isolates with the *cmeABC/ermG* invertible promoter oriented either primarily ON or OFF into media with (+Erm) or without (-Erm) erythromycin and quantified the percentage of cells in the ON or OFF orientation after 24 hours (Fig. 3D). The OFF cultures grew in +Erm medium only after an extended lag phase relative to growth in -Erm medium, whereas ON cultures grew similarly in +Erm and -Erm media (Fig. 3E). OFF isolates grown in +Erm medium became predominantly ON, while OFF isolates grown in -Erm medium remained OFF. ON isolates remained predominantly ON in both +Erm and -Erm media (Fig. 3D).

Finally, during serial transfers in both +Erm and -Erm media, we monitored the ON:OFF ratio of the *cmeABC/ermG* invertible promoter in a *B. stercoris* OFF isolate that had previously been switched to ON in +Erm medium. Over the course of 24 transfers at 1:1,000 dilution, the orientation of the *cmeABC/ermG* invertible promoter remained ON in +Erm medium but gradually drifted away from 100% ON in -Erm medium, recapitulating the *in vivo* observations from metagenomic data (Fig. 3F).

Although strongly favored in the presence of antibiotics, high expression of antibiotic resistance genes likely incurs a significant fitness cost, which could explain the reversion to the OFF orientation after antibiotic treatment. Many compensatory mechanisms to maintain antibiotic resistance in the absence of antibiotics have been noted(26), but invertons are akin to catastrophe insurance; a certain percentage of the population is always prepared to resist future antibiotic treatment and reintroduce heterogeneity after antibiotic selection.

Dense longitudinal metagenomic data allow for a detailed view of the dynamics of invertons over time in the human gut. We analyzed a dataset of samples from 54 individuals, four of whom (subject designations: ae, am, an, ao) were sampled densely over 5–18 months, and tracked the orientations of invertons. The F orientation is the same orientation of the inverton in the reference genome while the R orientation is the opposite orientation of the inverton in the reference genome. We identified 423 invertons with sufficient coverage to track their temporal dynamics in these individuals. Of these, 322 were predominantly found in one orientation within an individual with little or no fluctuation (mean > 95% and min >75% for either the F or R orientation) (Fig. 4A); the orientations of 59 were relatively stable (max-min %R <= 50%) within an individual (Fig. 4B), while the orientations of 42 were unstable (max-min %R > 50%) within an individual (Fig. 4C; Table S13).

Although the orientations of 90% of invertons in the same individual were relatively stable over time, the orientations between individuals varied extensively (Fig. 4D; Fig. S10). The mean %R orientation of 214 out of 423 of the invertons varied by more than 50% between individuals. In 122 examples, averaging across time, the inverton was predominantly (>95%) in the F orientation in at least one individual and predominantly (>95%) in the R orientation in another. Additionally, 119 out of 238 invertons were significantly different (Kruskal-Wallis H test, FDR $p < 0.05$) between the four individuals for whom we had dense longitudinal metagenomic data. The differences in the orientations of invertons between individuals could be explained by divergent selective forces between individuals, different optimal orientations between strains, or stochastic variation in orientation.

Sculpted by the individual's diet, lifestyle, immune response and genetics, the gut of every individual is a distinctive environment for resident bacteria. We observed the influence of the individual on the orientations of invertons in a cohort of patients with ulcerative colitis who were the recipients of fecal microbiota from healthy donors. The source of the fecal microbiota transplant (FMT) was donor "am", whose longitudinal metagenomic data was analyzed above. Therefore, we could monitor the orientations of invertons from the same strain for 18 months in a healthy donor and up to 5 months in the patients.

First, we identified strains from the donor that engrafted into the patient's microbiome. To identify engraftment, we found cases where the same strain was present in the donor and in at least one patient after FMT, but absent in the same patient(s) before FMT (Fig. 4E, G, and I; Fig. S11). We found three high-abundance species with invertons from the donor that engrafted in patients: *Bacteroides fragilis*, *Bacteroides vulgatus*, and *Bacteroides ovatus*.

Then, we compared the orientations of invertons from the donor and patient after engraftment and found the orientations of 42.8% (48/112) invertons diverged from the donor orientation (Wilcoxon rank sum test, FDR $p<0.05$) (Fig. 4F, H and J; Fig. S12). In *B. fragilis*, two invertons (IBP183 and IBP198) engrafted in the opposite orientation of the donor strain and remain predominantly (87.2% and 87.1%) in the R orientation, but drifted towards F near the end of the sampling (Fig. 4F). For *B. ovatus*, a strain existed in patient 014 before FMT but was replaced by the donor *B. ovatus* strain (Fig. 4G; compare orange to purple). IBP155 from both the donor strain and pre-FMT strain were predominantly (90.1% and 100%) in the R orientation, while the newly-engrafted strain was oriented entirely in the F orientation and remained in the F orientation over the course of sampling (Fig. 4H). In *B. vulgatus*, an inverton was initially present completely in the F orientation but over the course of 145 days completely reversed to the R orientation (Fig. 4J, IBP121). In addition to the invertons that reversed their orientations, we also identified examples of invertons that maintained their orientations (Fig. 4F, IBP189; 4H, IBP166; 4J, IBP125).

Our findings highlight the role invertons play in host-microbe co-existence. Genes regulated by invertons were highly enriched for products located on the exterior of the cell where they are exposed to the host immune system and phages indicating they may be primary targets for selection that are beneficial for gut commensals to diversify their surface architectures or essential processes, such as antibiotic resistance, whose expression has a high fitness cost. The high prevalence of antibiotic resistance genes regulated by invertons containing

promoters suggests this is an example of bet-hedging(27). This could lead to longer persistence of antibiotic resistance genes in a microbial community, further increasing the burden to combat antibiotic resistance. Our results indicate that in the human gut, invertons help bacterial populations regain heterogeneity after bottlenecks encountered during colonization of a new host or severe perturbations. Overall, our study provides insights into a mechanism allowing adaptive tradeoffs in bacteria that have evolved to successfully colonize the gastrointestinal tract of vertebrate hosts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Tiffany Poon, Tommi Vatanen and Mathieu Groussin for their assistance in obtaining samples and data. We thank Theresa Reimels for comments on the manuscript. We thank the Microbial Omics Core of the Broad Institute and the Broad Technology Labs for their assistance with DNA extraction, library preparation, and sequencing. The FMT research was approved by the Committee on Human Research in the Medical Sciences (CHRMS) (CHRMS 15–373).

Funding: This work was funded by grants from the National Institutes of Health (DK043351 and AT009708), the Crohn's and Colitis Foundation of America, the Juvenile Diabetes Research Foundation, and the Center for Microbiome Informatics and Therapeutics at MIT to R.J.X. A.B.H. is a Merck Fellow of the Helen Hay Whitney Foundation.

References and Notes:

1. Coyne MJ, Chatzidaki-Livanis M, Paoletti LC, Comstock LE, Role of glycan synthesis in colonization of the mammalian gut by the bacterial symbiont Bacteroides fragilis. Proc. Natl. Acad. Sci. U. S. A 105, 13099–13104 (2008). [PubMed: 18723678]
2. Porter NT, Canales P, Peterson DA, Martens EC, 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.e8 (2017). [PubMed: 28966055]
3. Gauntlett JC, Nilsson H-O, Fulurija A, Marshall BJ, Benghezal M, Phase-variable restriction/modification systems are required for Helicobacter pylori colonization. Gut Pathog. 6, 35 (2014). [PubMed: 25349630]
4. Taketani M, Donia MS, Jacobson AN, Lambris JD, Fischbach MA, A Phase-Variable Surface Layer from the Gut Symbiont Bacteroides thetaiotaomicron. MBio. 6, e01339–15 (2015). [PubMed: 26419879]
5. Lim JK et al., In vivo phase variation of Escherichia coli type 1 fimbrial genes in women with urinary tract infection. Infect. Immun 66, 3303–3310 (1998). [PubMed: 9632599]
6. Henderson IR, Owen P, Nataro JP, Molecular switches--the ON and OFF of bacterial phase variation. Mol. Microbiol 33, 919–932 (1999). [PubMed: 10476027]
7. Coyne MJ, Weinacht KG, Krinos CM, Comstock LE, Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. Proc. Natl. Acad. Sci. U. S. A 100, 10446–10451 (2003). [PubMed: 12915735]
8. Hinde P, Deighan P, Dorman CJ, Characterization of the detachable Rho-dependent transcription terminator of the fimE gene in Escherichia coli K-12. J. Bacteriol 187, 8256–8266 (2005). [PubMed: 16321930]
9. Kawula TH, Orndorff PE, Rapid site-specific DNA inversion in Escherichia coli mutants lacking the histone-like protein H-NS. J. Bacteriol 173, 4116–4123 (1991). [PubMed: 1648076]
10. Bayliss CD, Determinants of phase variation rate and the fitness implications of differing rates for bacterial pathogens and commensals. FEMS Microbiol. Rev 33, 504–520 (2009). [PubMed: 19222587]

11. Silverman M, Simon M, Phase variation: genetic analysis of switching mutants. Cell. 19, 845–854 (1980). [PubMed: 6247071]
12. Abraham JM, Freitag CS, Clements JR, Eisenstein BI, An invertible element of DNA controls phase variation of type 1 fimbriae of Escherichia coli. Proc. Natl. Acad. Sci. U. S. A 82, 5724–5727 (1985). [PubMed: 2863818]
13. 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. J. Bacteriol 191, 6003–6011 (2009). [PubMed: 19648246]
14. Nakayama-Imaohji H et al., DNA Inversion Regulates Outer Membrane Vesicle Production in Bacteroides fragilis. PLoS One. 11, e0148887 (2016). [PubMed: 26859882]
15. Troy EB, Carey VJ, Kasper DL, Comstock LE, Orientations of the Bacteroides fragilis capsular polysaccharide biosynthesis locus promoters during symbiosis and infection. J. Bacteriol 192, 5832–5836 (2010). [PubMed: 20729352]
16. Zitomersky NL, Coyne MJ, Comstock LE, Longitudinal analysis of the prevalence, maintenance, and IgA response to species of the order Bacteroidales in the human gut. Infect. Immun 79, 2012–2020 (2011). [PubMed: 21402766]
17. Mende DR et al., proGenomes: a resource for consistent functional and taxonomic annotations of prokaryotic genomes. Nucleic Acids Res. 45, D529–D534 (2017). [PubMed: 28053165]
18. Grigoriev IV et al., The genome portal of the Department of Energy Joint Genome Institute. Nucleic Acids Res. 40, D26–32 (2012). [PubMed: 22110030]
19. Bayley DP, Rocha ER, Smith CJ, Analysis of cepA and other Bacteroides fragilis genes reveals a unique promoter structure. FEMS Microbiol. Lett 193, 149–154 (2000). [PubMed: 11094294]
20. Liu CH, Lee SM, Vanlare JM, Kasper DL, Mazmanian SK, Regulation of surface architecture by symbiotic bacteria mediates host colonization. Proc. Natl. Acad. Sci. U. S. A 105, 3951–3956 (2008). [PubMed: 18319345]
21. Porter NT, Canales P, Peterson DA, Martens EC, 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.e8 (2017). [PubMed: 28966055]
22. Yassour M et al., Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. Sci. Transl. Med 8, 343ra81 (2016).
23. Hartmeyer GN, Sóki J, Nagy E, Justesen US, Multidrug-resistant Bacteroides fragilis group on the rise in Europe? J. Med. Microbiol 61, 1784–1788 (2012). [PubMed: 22956754]
24. Husain F et al., The Ellis Island effect: a novel mobile element in a multi-drug resistant Bacteroides fragilis clinical isolate includes a mosaic of resistance genes from Gram-positive bacteria. Mob. Genet. Elements 4, e29801 (2014). [PubMed: 25165618]
25. Monod M, Mohan S, Dubnau D, Cloning and analysis of ermG, a new macrolide-lincosamide-streptogramin B resistance element from Bacillus sphaericus. J. Bacteriol 169, 340–350 (1987). [PubMed: 3025178]
26. Andersson DI, Hughes D, Persistence of antibiotic resistance in bacterial populations. FEMS Microbiol. Rev 35, 901–911 (2011). [PubMed: 21707669]
27. Philippi T, Seger J, Hedging one's evolutionary bets, revisited. Trends Ecol. Evol 4, 41–44 (1989). [PubMed: 21227310]
28. Coyne MJ, Comstock LE, Niche-specific features of the intestinal Bacteroidales. J. Bacteriol 190, 736–742 (2008). [PubMed: 17993536]
29. Holden N et al., Comparative analysis of FimB and FimE recombinase activity. Microbiology. 153, 4138–4149 (2007). [PubMed: 18048927]
30. Somvanshi VS et al., A single promoter inversion switches Photorhabdus between pathogenic and mutualistic states. Science. 337, 88–93 (2012). [PubMed: 22767929]
31. Horino A, Sasaki Y, Sasaki T, Kenri T, Multiple promoter inversions generate surface antigenic variation in Mycoplasma penetrans. J. Bacteriol 185, 231–242 (2003). [PubMed: 12486060]
32. Czurda S, Jechlinger W, Rosengarten R, Chopra-Dewasthaly R, Xer1-mediated site-specific DNA inversions and excisions in Mycoplasma agalactiae. J. Bacteriol 192, 4462–4473 (2010). [PubMed: 20562305]

33. Boot HJ, Kolen CP, Pouwels PH, Interchange of the active and silent S-layer protein genes of Lactobacillus acidophilus by inversion of the chromosomal slp segment. Mol. Microbiol 21, 799–809 (1996). [PubMed: 8878042]
34. Rice P, Longden I, Bleasby A, EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16, 276–277 (2000). [PubMed: 10827456]
35. Schwan WR, Regulation of fim genes in uropathogenic Escherichia coli. World J. Clin. Infect. Dis 1, 17–25 (2011). [PubMed: 23638406]
36. Molina N, van Nimwegen E, Universal patterns of purifying selection at noncoding positions in bacteria. Genome Res. 18, 148–160 (2008). [PubMed: 18032729]
37. Wagner A, Lewis C, Bichsel M, A survey of bacterial insertion sequences using IScan. Nucleic Acids Res. 35, 5284–5293 (2007). [PubMed: 17686783]
38. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M, ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 34, D32–6 (2006). [PubMed: 16381877]
39. Holtgrewe M, Mason--a read simulator for second generation sequencing data. Technical Report FU Berlin (2010) (available at http://publications.mi.fu-berlin.de/962/).
40. Tatusova T et al., NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44, 6614–6624 (2016). [PubMed: 27342282]
41. Bailey TL et al., MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37, W202–8 (2009). [PubMed: 19458158]
42. Grant CE, Bailey TL, Noble WS, FIMO: scanning for occurrences of a given motif. Bioinformatics. 27, 1017–1018 (2011). [PubMed: 21330290]
43. Andreani J, La Scola B, Metagenomic binning reconstruction coupled with automatic pipeline annotation and giant viruses: A potential source of mistake in annotations. Virus Res. 255, 36–38 (2018). [PubMed: 29981362]
44. Ermolaeva MD, White O, Salzberg SL, Prediction of operons in microbial genomes. Nucleic Acids Res. 29, 1216–1221 (2001). [PubMed: 11222772]
45. Jones P et al., InterProScan 5: genome-scale protein function classification. Bioinformatics. 30, 1236–1240 (2014). [PubMed: 24451626]
46. Yu G, Wang L-G, Han Y, He Q-Y, clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 16, 284–287 (2012). [PubMed: 22455463]
47. Beissbarth T, Speed TP, GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics. 20, 1464–1465 (2004). [PubMed: 14962934]
48. McArthur AG et al., The comprehensive antibiotic resistance database. Antimicrob. Agents Chemother 57, 3348–3357 (2013). [PubMed: 23650175]
49. Truong DT et al., MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods 12, 902–903 (2015). [PubMed: 26418763]

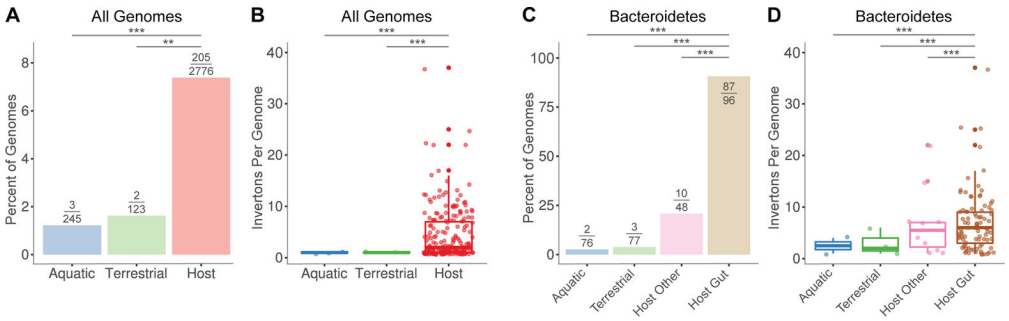

Fig. 1. The prevalence and number of invertons per genome are enriched in host-associated species.

(A) The percentage of genomes identified with invertons and **(B)** the number of invertons per genome from aquatic, terrestrial, and host-associated isolates. **(C)** In the phylum Bacteroidetes, the percentage of genomes identified with invertons and **(D)** the number of invertons per genome from aquatic, terrestrial, host sites other than gut, and host gut-associated isolates.

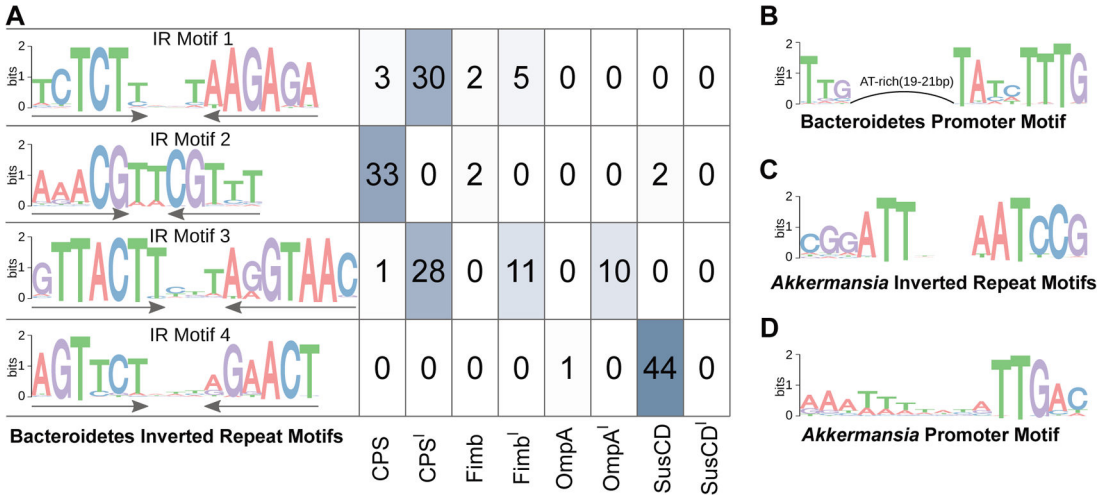

Fig. 2. Motifs found in the inverted repeats of Bacteroidetes and *Akkermansia* invertons consist of 5–7 base pair palindromes with 2–3 intervening base pairs.

(A) Functional profiling of operons regulated by invertons reveals specializations of each inverted repeat (IR) motif. The heat map represents the number of operons per functional class regulated by invertons with either global or local invertases. CPS: capsular polysaccharide; Fimb: fimbriae; OmpA: outer membrane protein A. Superscript l on the gene abbreviation indicates the presence of a local invertase. Based on the absence of local invertases directly upstream, IR motifs 2 and 4 are likely regulated by global invertases. **(B)** Promoter motif identified from invertons from Bacteroidetes spp. **(C)** The inverted repeat motif found in all identified invertons from *Akkermansia* spp. **(D)** Promoter motif identified from invertons from *Akkermansia* spp.

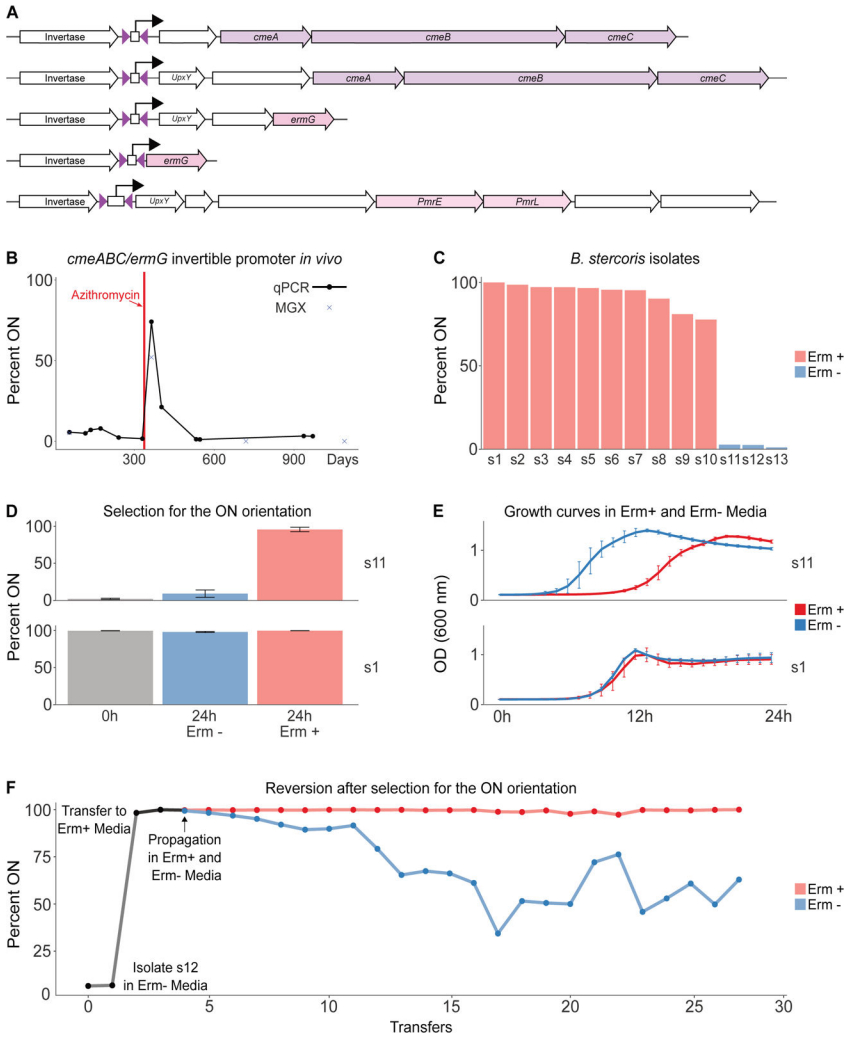

Fig. 3. Invertible promoters regulate antibiotic resistance genes.

(A) Three classes of genes conferring antibiotic resistance were regulated by invertible promoters. The genomic context of the loci are shown with antibiotic resistance genes colored. Promoters are designated by hooked arrows, and purple triangles represent inverted repeats. **(B)** An invertible promoter regulating the *cmeABC* multidrug efflux cassette in individual E011878 was 99% OFF before antibiotic treatment, shifted to 74% ON 7 days afterward, and then drifted back to 99% OFF over 5 months as measured by both qPCR and metagenomic (MGX) data. **(C)** The *cmeABC*/*ermG* invertible promoter is oriented ON in *B. stercoris* isolated in medium containing erythromycin (+Erm) and OFF in *B. stercoris* isolated in medium without erythromycin (-Erm). **(D)** The orientation of the *cmeABC*/*ermG* invertible promoter under antibiotic selection. Isolate s11 grown in -Erm medium remained OFF, while the same isolate transferred to +Erm medium shifted to ON. Isolate s1 remained ON in -Erm and +Erm media. **(E)** The growth of isolates under antibiotic selection. Isolate s1 grows in +Erm medium without a lag phase relative to -Erm medium, whereas isolate s11 grows in +Erm medium after an extended lag phase, consistent with a lower number of initially erythromycin-resistant cells. **(F)** Kinetics of the *cmeABC*/*ermG* invertible promoter. An isolate with the *cmeABC*/*ermG* invertible promoter initially in the OFF orientation, s12,

was exposed to erythromycin. Half of the culture was then propagated in the presence of erythromycin and the other half in the absence of erythromycin for 24 1:1,000 dilution transfers. In +Erm medium, the promoter remained ON, while in -Erm medium, it drifted towards OFF.

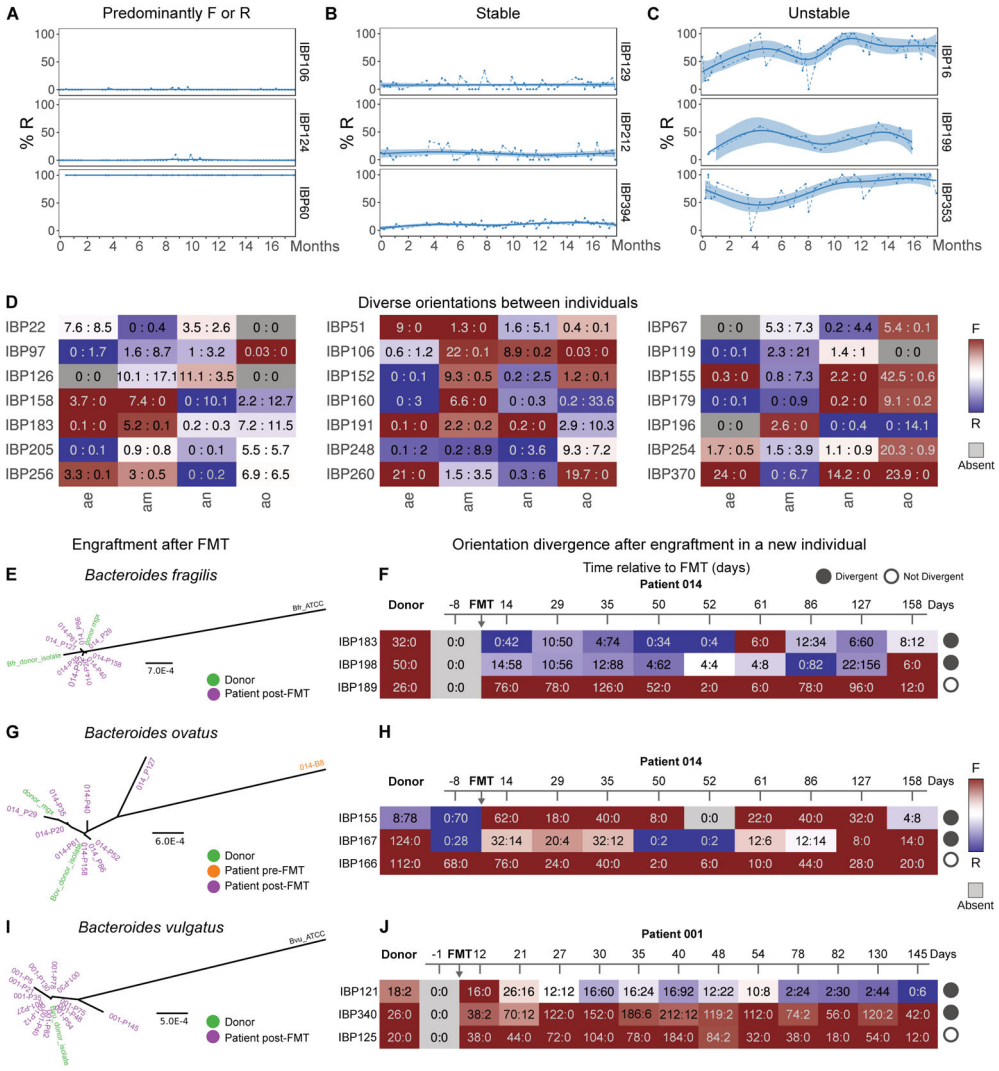

Fig. 4. The orientations of invertons are generally stable within individuals but divergent between individuals.

(A) Examples of invertons from donor “am” found in the forward (F: IBP106, IBP124) or reverse (R: IBP60) orientation. **(B)** Examples of invertons from donor am (IBP129, IBP212, and IBP394) that are longitudinally stable. **(C)** Examples of invertons from donor am that are longitudinally unstable (IBP16, IBP199, and IBP353). **(D)** Selected examples of invertible promoters with different orientations between four individuals (ae, am, an, and ao). Numbers in each heatmap cell represent the average counts in the F and R orientations (F:R) while red indicates the F orientation and blue represents the R orientation. **(E,G,I)** Phylogenetic trees produced by StrainPhlAn from metagenomic data from FMT donor (green) and patient (orange pre-FMT, purple post-FMT) as well as isolate genomes from the donor (green) and unrelated reference genomes from the same species (black). The phylogenetic trees demonstrate engraftment and persistence of (E) *B. fragilis*, (G) *B. ovatus*, and (I) *B. vulgatus* strains from the donor to the patient. Phylogenetic tree legends are the number of nucleotide substitutions per site. **(F, H, J)** Examples of divergence of invertible promoter orientations after engraftment in a patient. Black circles denote invertible

promoters whose orientation was significantly different between donor and patient after FMT. White circles denote invertible promoters whose orientation is not significantly different (Wilcoxon rank sum test, FDR $p<0.05$). Numbers in each heatmap cell represent the counts in the F and R orientations (F:R). (F) After engraftment, the orientations of IBP183 and IBP198 were reversed compared to the donor. (H) A strain of *B. ovatus* was present in Patient 14 before FMT but was outcompeted by the donor strain. After engraftment, the orientation of IBP155 and IBP167 was reversed compared to the donor. (J) After *B. vulgatus* engraftment, the orientation of IBP121 was only found in the F orientation, but over the course of 135 days, the orientation reverses.

Table 1.

Invertons per phylum identified in a systematic search of bacterial genomes with PhaseFinder.

Phylum	Genomes Searched	Genomes with Invertons	Percentage of Genomes with Invertons	Total Invertons
Acidobacteria	6	0	0	0
Actinobacteria	5262	67	1.3	200
Armatimonadetes	2	0	0	0
Bacteroidetes	491	160	32.6	1254
Chlamydiae	45	3	6.7	5
Chloroflexi	1	0	0	0
Cyanobacteria	20	4	20	7
Deinococcus -Thermus	14	1	7.1	1
Fibrobacteres	16	0	0	0
Firmicutes	17010	986	5.8	1422
Fusobacteria	10	0	0	0
Nitrospirae	1	0	0	0
Proteobacteria	14872	1133	7.6	1628
Spirochaetes	138	57	41.3	140
Synergistetes	8	2	25	2
Tenericutes	21	0	0	0
Thermodesulfobacteria	3	0	0	0
Thermotogae	7	0	0	0
Verrucomicrobia	5	1	20	27