Comprehensive mutagenesis of the fimS promoter regulatory switch reveals novel regulation of type 1 pili in uropathogenic Escherichia coli

Huibin Zhang⁎, Teodorus T. Susanto⁎, Yue Wan⁎, and Swaine L. Chen⁎,†,1

⁎Infectious Diseases Group, Genome Institute of Singapore, Singapore 138672; ⁎Stem Cell and Development, Genome Institute of Singapore, Singapore 138672; and ⁎Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119074

Edited by Roy Curtis III, University of Florida, Gainesville, FL, and approved March 7, 2016 (received for review December 6, 2015)

Type 1 pili (T1P) are major virulence factors for uropathogenic Escherichia coli (UPEC), which cause both acute and recurrent urinary tract infections. T1P expression therefore is of direct relevance for disease. T1P are phase variable (both piliated and nonpiliated bacteria exist in a clonal population) and are controlled by an invertible DNA switch (fimS), which contains the promoter for the fim operon encoding T1P. Inversion of fimS is stochastic but may be biased by environmental conditions and other signals that ultimately converge at fimS itself. Previous studies of fimS sequences important for T1P phase variation have focused on laboratory-adapted E. coli strains and have been limited in the number of mutations or by alteration of the fimS genomic context. We surmounted these limitations by using saturating genomic mutagenesis of fimS coupled with accurate sequencing to detect both mutations and phase status simultaneously. In addition to the sequences known to be important for biasing fimS inversion, our method also identifies a previously unknown pair of 5' UTR inverted repeats that act by altering the relative fimA levels to control phase variation. Thus, we have uncovered an additional layer of T1P regulation potentially impacting virulence and the coordinate expression of multiple pilius systems.

Urinary tract infection | fimS | type 1 pili | phase variation | saturating chromosomal mutagenesis

Urinary tract infections (UTIs) are common infections that mostly affect women with an annual cost of billions of dollars (1). They are caused mainly by uropathogenic Escherichia coli (UPEC), which rely on several virulence factors to cause disease (2); type 1 pili (T1P) are perhaps the most important of these (3). T1P present a tip adhesin, FimH, that binds specifically to mannose, which is found on several glycosylated surface proteins, such as uroplakins and α1- and β3-integrins (4, 5), in the mammalian bladder. Binding to mannosylated proteins initiates a cascade of events including UPEC invasion of host tissues (2), formation of intracellular structures (6), and activation of the host immune response (7). The interplay among these events determines the outcome of infection (8), ranging from full clearance to chronic active cystitis. T1P are important for most of these UTI stages and demonstrate dynamic regulation of expression that can be seen directly in human infections and urine (9, 10).

Thus, understanding T1P regulation is fundamental to understanding UTI. T1P have been well studied (11), mostly in laboratory-adapted E. coli strains. T1P are encoded by the fim operon; their expression is phase variable via inversion of fimS, a chromosomal DNA segment containing the fim promoter (12). This binary epigenetic switch results in bacteria switching between fimbriated and nonfimbriated states. Inversion is mediated by the FimB and FimE recombinases, whose genes are located immediately upstream of fimS (13) and are found in most E. coli strains (14). Inversion is influenced by numerous signals, such as temperature, pH, osmolality, and the stress and stringent responses (15–19). Several global gene regulators also affect phase variation and pilus expression by either direct association with fimS or regulation of fim recombinase expression (11). These regulators act through a variety of mechanisms, such as DNA structure and supercoiling (20–22), transcription termination and RNA stability (23, 24), or dual effects on both phase variation and fimA transcription [such as through cAMP receptor protein (CRP), integration host factor (IHF), or growth phase] (18, 20, 25).

In vitro, the recombinases bind to sites flanking and overlapping fimS inverted repeats (IRs) (26). On plasmid substrates, FimB mediates switching in both directions, whereas FimE performs only ON-to-OFF switching (27). However, FimE can mediate OFF-to-ON inversion in vivo (28). In addition to fimB and fimE, UPEC strains often carry other unlinked recombinases (such as fimX and ipuA) whose products also mediate fimS inversion (14, 29, 30), and even the presence of the “core” fimB gene can vary (31). Thus, both the presence and function of recombinase vary among E. coli, highlighting the importance of studying T1P directly in pathogenic E. coli strains.

In general, previous studies of T1P phase variation used in vitro assays, candidate gene approaches, and reporter fusions in nonpathogenic E. coli strains (11). As noted above for FimE (28), in vitro studies do not capture the full complexity of in vivo regulation, and an in vitro array-based screen for fimS-binding protein factors did not identify FimB (32). Directed genetic approaches have been invaluable for connecting T1P regulation to global regulatory networks (11, 33, 34) but are limited in scale. Reporter assays enable unbiased screens but modify the genomic context and/or copy number of fimS, potentially affecting regulation. Finally, laboratory-adapted strains provide convenience in

Significance

Urinary tract infections are common infections caused mostly by uropathogenic Escherichia coli (UPEC). Type 1 pili (T1P) are important UPEC virulence factors and thus are attractive therapeutic targets. T1P expression is highly regulated and phase variable, resulting in piliated and nonpiliated bacteria in an otherwise clonal population. Regulation converges at the invertible fimS (fim switch) regulatory sequence. We performed the first (to our knowledge) comprehensive analysis of fimS mutations affecting phase variation in vivo in a clinical UPEC strain. We identified a previously unknown regulatory element that affects T1P expression and coordinates the expression of other pili, processes that facilitate infection by UPEC. This method may enable comprehensive studies of phase variation in other pathogens and other regulated DNA dynamics in general.

Author contributions: H.Z., Y.W., and S.L.C. designed research; H.Z., T.T.S., and Y.W. performed research; H.Z. and S.L.C. analyzed data; and H.Z. and S.L.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

1To whom correspondence should be addressed. Email: slchen@gis.a-star.edu.sg.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1522958113/-/DCSupplemental.
genetic manipulation, but, as mentioned above, regulation relevant to infection may differ in clinical isolates (14, 29–31).

We have leveraged accurate PacBio sequencing and negative selection cassettes (35, 36) to overcome these limitations, enabling a high-throughput assay to quantify the effects of all fimS mutations on TIP phase variation in vivo. Because inversion is a genomic change, sequencing allows the direct association of phase status with mutations in fimS (35). A negative selection system usable in unmodified E. coli clinical isolates enables the creation of markerless and scarless chromosomal mutant libraries, similar to genome editing mediated by CRISPR (36, 37). Inclusion of sequence context flanking fimS enabled us to discover a previously unknown pair of IRs (termed “UIRs” for 5′ UTR inverted repeats) that ensure high expression of fimA compared with other fim operon genes and may coordinate TIP expression with other pilus systems. Our results advance the understanding of TIP regulation and provide a general method for studying other phase-variable and epigenetically regulated systems.

Results

Mapping the Effects of fimS Mutations on Phase Variation. To study native fimS regulation, we created a complete library of mutations in fimS and its flanking sequences directly in the chromosome of the prototypical cystitis strain UTB89 (38). A two-step knockout/knockin strategy with a dual positive/negative selection cassette was used to recombine an error-prone PCR product (covering fimS and some flanking DNA) into the chromosomal fimS locus in UTB89 (36) without residual markers or scars and preserving native fimS context and copy number. Sanger sequencing of seven individual clones verified that the correct recombinants containing fimS mutations (average 2.1 mutations per 480 bp; 0.44%) were made. These mutant libraries were grown under phase ON- or OFF-inducing conditions, and the fimS locus was amplified by PCR, attached to unique molecular identifiers (UMIs) to facilitate quantitation (39), and sequenced (Fig. L4).

To cover the 480-bp mutated region, we used the Illumina MiSeq (2 x 250-bp reads) and PacBio (average 6-kb reads) sequencing platforms. Illumina sequencing showed high mutation percentages for both unmutated control and mutated libraries (0.23% and 0.85% of all positions, respectively, vs. expected 0% and 0.44%) (Fig. S1A). These mutations were likely systematic errors that might be mitigated with variable-length adapters (40, 41), but overall sequencing quality also declined at the 3′ ends of reads. In contrast, circular consensus sequences from PacBio sequencing resulted in mutation percentages of 0.0075% for an unmutated control and 0.41% for the mutated library, in outstanding agreement with Sanger sequencing (Fig. S1A) and with uniform base accuracy throughout the entire fimS locus. We therefore performed the rest of our studies with the PacBio platform.

At every fimS position, we sequenced at least 22 and 6 mutant amplicons with UMIs (combining across all mutations at that position) for ON and OFF induction, respectively. Mutations were distributed relatively uniformly among reads (average 2.5 and 3.0 mutations per read for ON and OFF induction, respectively) (Fig. S1B) and across fimS (although some peaks were detected) (Fig. S1C), indicating the absence of large jackpot effects through the entire protocol. The percent of phase OFF reads was calculated for each mutation at each position, yielding a comprehensive map of changes in phase variation bias associated with mutations throughout fimS (Fig. S2). Regions of fimS known to have large effects on phase variation when mutated were clearly identified (Fig. 1B and Figs. S2 and S3A), including the fim promoter transcriptional start site at position 349 (42) and binding sites for fim recombinases at positions 36–72 and 343–376 (26, 43).

Sequencing-Based Analysis of fimS Phase Identifies Known and Previously Unknown Phase-Regulation Loci. To validate our sequencing-based approach, we compared our results with data from a qualitative study of plasmid-based IR mutations on phase variation in a K12 strain (43). The recombinases bind the fimS IRs denoted “IR left” (IRL) and “IR right” (IRR) during inversion (26, 43), but not all IR positions are equally important. We indeed saw relatively large effects for IR mutations (Fig. 2A and Fig. S4A). Quantitative measurements of the percent of phase OFF cells calculated from sequencing reads corresponded well with the qualitative categories previously reported (Fig. 2B and Fig. S4B) (43), particularly for the “large effect” category. We found substantial variation in “partial effect” mutations; this variation may reflect differences in strains, context, copy number, or growth conditions.

Therefore, as further validation, we created single point mutations in the chromosomal fimS in UTB89. Both G50A (IRL) and C358T (IRR) mutations yielded results quantitatively similar to the PacBio sequencing data by a PCR/restriction digest assay under phase ON-inducing conditions (92–98% phase OFF by digest assay vs. 96–98% by PacBio for mutants and 17% phase OFF by digest assay vs. 28% by PacBio for WT) (Fig. 2A and C). Furthermore, the mutants had lower HA titers, a functional assay for TIP expression and assembly (Fig. S4C).

Among sequences not previously implicated in regulating phase variation, one region beyond IRR, in the 5′-UTR of fimA, stood out (Figs. 1B, purple boxes, and 2D) because it contained...
In addition, as predicted, a single C440T mutation that would allow G-U pairing had little effect on phase variation or HA titers.

To address whether the UIRs might act at the RNA level, we used RNA structural probing on the fimA 5′ UTR. This assay uses 2-methylthio nicotinic acid imidazolide (NAI) to modify single-stranded or flexible RNA selectively; modified positions can be identified by primer extension (44). Changes in primer extension patterns are interpreted as changes in RNA structure. We saw prominent changes at three residues: C441 (within the UIR), A443, and A449 (Fig. 3A). At these positions, the WT and G398C/C440G mutants (both with preserved G-C pairing) had similar patterns. In contrast, all mutants with disrupted base-pairing or weaker A-U pairing showed altered patterns at C441 and A449. Interestingly, changes at position A443 (2 bp from the edge of the UIR) were detected only when base-pairing was fully disrupted (in G398A and G398C mutants). Thus, UIR mutations can affect fimA RNA structure, and these changes correlate well with in vivo phase and HA titers of fimA mutants (Fig. S5 A and B).

Because the UIRs affect both phase regulation and fimA 5′ UTR structure, we measured fim transcription using quantitative real-time PCR (qRT-PCR) (Fig. 3B). G398 single mutants had reduced expression of all fim genes compared with WT. Again, restoration of base-pairing in the G398C/C440G mutant also restored expression to near WT transcript levels, whereas C440T and G398A/C440T mutants showed a partial rescue. Interestingly, fimA transcript levels in all mutants (except G398C/C440G)

Fig. 2. Identification of loci involved in phase regulation. (A) Detailed phase variation map for the fimS IRs (purple arrows) under OFF induction. Numbers correspond to nucleotide positions within fimS as in Fig. 1; the WT IR sequence is indicated below. Each position is represented by four rectangles (representing all four possible nucleotides) whose brightness and color indicate the number of mutant sequencing reads observed and the percentage of reads that are phase OFF, respectively (based on the scale at the right). (B) Comparison of phase variation effects with those found by McCusker et al. (43) under OFF induction. Each point represents the expected phase OFF for a single IR mutation, with the x-axis categories taken from ref. 43. The solid black lines indicate the median. Only IR mutations with at least 10 sequencing reads were plotted (n = 34). (C) TIP phase assay of IR mutants. Mutations are indicated at the top. Bands corresponding to phase ON and OFF orientation are indicated on the left. The % phase OFF based on band intensities is shown below. (D) Phase variation map for UIRs (purple solid arrows) under ON induction. Notation is as in A. (E) HA assay for TIP function. Average logy, HA titer), with SEs (y-axis) for each mutant (x-axis) are plotted using data from at least three independent experiments. Colors represent different mutations; solid bars represent HA titers with mannose added. Only one transversion per position is shown; other mutations are shown in Fig. S4 D. Purple solid lines denote mutations within the UIR.

UIRs Likely Base Pair as RNA to Affect fimA Transcript Levels. Based on their position, the UIRs are likely transcribed by the fimA promoter. In addition, two mutations within the UIRs, A400G and C403T, had intermediate effects on phase variation and HA titers compared with other UIR mutations (Fig. S4 D and E). We hypothesized that the intermediate phenotypes resulted from the preservation of RNA base-pairing by suboptimal G-U base pairs (for example, A400G would convert an A-U pair to G-U). This hypothesis predicted that compensatory mutations to restore base pairing would rescue UIR mutant phenotypes, which we tested at the predicted G398/C440 base pair (Fig. 2D). G398 mutants with disrupted base pairing had high levels of phase OFF cells under phase ON growth conditions and indeed were rescued by compensatory mutations (G398A/C440T and G398C/C440G) (Fig. S5 A and B), suggesting that the UIRs do base pair.

another pair of IRs, which we denote as “UIRs,” as explained above. Validation with single point mutations in the UIRs resulted in high percentages of phase OFF bacteria despite growth in phase ON-inducing conditions, matching the PacBio data well (Fig. S4D). These mutants had only slightly reduced average mannose-sensitive HA (MSHA) titers (Fig. 2E and Fig. S4 E; mannose sensitivity is indicative of T1P-mediated HA) but, in contrast, had markedly increased mannose-resistant HA (MRHA) titers. The MRHA suggests the expression of other pili, possibly co-ordinated with the lower expression of T1P. Importantly, mutations immediately flanking the UIR had little effect in these assays (Fig. 2 D and E and Fig. S4 D and E).

UIRs Likely Base Pair as RNA to Affect fimA Transcript Levels. Based on their position, the UIRs are likely transcribed by the fimA promoter. In addition, two mutations within the UIRs, A400G and C403T, had intermediate effects on phase variation and HA titers compared with other UIR mutations (Fig. S4 D and E). We hypothesized that the intermediate phenotypes resulted from the preservation of RNA base-pairing by suboptimal G-U base pairs (for example, A400G would convert an A-U pair to G-U). This hypothesis predicted that compensatory mutations to restore base pairing would rescue UIR mutant phenotypes, which we tested at the predicted G398/C440 base pair (Fig. 2D). G398 mutants with disrupted base pairing had high levels of phase OFF cells under phase ON growth conditions and indeed were rescued by compensatory mutations (G398A/C440T and G398C/C440G) (Fig. S5 A and B), suggesting that the UIRs do base pair.
were specifically reduced (~2× lower; note log scale in Fig. 3B) compared with other fim transcripts. To verify that the low fimA transcript level was the primary defect, we overexpressed FimA using the fimS promoter (preserving the WT UIRs in trans) or the lacZ promoter (completely replacing the native fimA 5’ UTR, including the UIRs, in the plasmid-encoded fimA) in the G398A mutant. In both cases, supplying additional fimA transcript restored phase variation, MSHA, and MRHA titers to WT levels (Fig. 3C and Fig. S5C), implying that the UIRs act through fimA levels.

UIR Mutants Regulate Phase Variation Through a Previously Unknown Mechanism. Transcript levels of fimA are known to be higher than those of other fim genes (45), but misregulation of this ratio has not been observed. We tested whether misregulation of relative fimA transcript levels was a common feature of TIP regulation by comparing fimA with fimH expression in a mutant with altered phase variation, FimH-Q133K. This mutation abrogates mannose binding by FimH but does not affect TIP assembly, leading to phase variation, FimH-Q133K. This mutation abrogates mannose binding by FimH but does not affect TIP assembly, leading to phase variation, FimH-Q133K. This mutation abrogates mannose binding by FimH but does not affect TIP assembly, leading to phase variation. FimH-Q133K mutants showed no HA with desialylated erythrocytes (Fig. S6C). In contrast, UIR mutants still had MSHA on desialylated erythrocytes, indicating that they have HA from both TIP and S pil. As further evidence, we saw variable increases in SfaA protein expression (SfaA is known to cross-react with the FimA antibodies used) in strains with sialylation-dependent MRHA (Fig. S6G).

Discussion
Phase variation is a common mechanism for regulating virulence factor expression, enabling immune evasion and bet-hedging at a population level (48). Phase variation often relies on epigenetic changes: recombination/gene conversion, short nucleotide repeat length polymorphism, DNA inversion, and DNA methylation (49). TIP phase variation is one of the best-studied examples of DNA inversion (12). Interestingly, although TIP are a primary virulence factor for UPEC causing UTI (3), the TIP fim operon is found in most E. coli strains (50). Therefore, most studies of TIP phase variation have used commensal, laboratory-adapted strains (11). Laboratory strains are convenient for genetic manipulation and have contributed greatly to detailed knowledge about fimS switching, but sequence context, copy number, and transcriptional and posttranscriptional regulation have frequently been altered, leading to a formal gap with native TIP regulation. Potential regulatory variation among different strains, particularly in UPEC (9, 51), presents a further complication. We surmounted these complications by performing a comprehensive screen of cis-acting sites at fimS, associating phase variation directly with mutations, in the native chromosomal context of a clinical UPEC strain. Our discovery of a previously unknown regulatory input into TIP expression and phase variation outside the invertible fimS region highlights the value of preserving context and copy number and emphasizes the complexity of TIP regulation.

This work relied on two recent advances: an efficient negative selection system usable in unmodified clinical isolates of E. coli (39), and accurate, long read PacBio sequencing (35). We killed read coverage for high accuracy; however, we still achieved essentially complete coverage of single mutations in the 480-bp target region. We thus analyzed mutations individually, without considering that some individual clones had multiple mutations. Hence, we would miss combinations of mutations with synthetic phenotypes. Higher read coverage could overcome this problem but would require a corresponding increase in mutant library complexity. Despite the high efficiency of the negative selection cassette (36), such libraries (~10^3 to 10^6 chromosomal mutants) are still difficult to achieve. Therefore, the fimS locus falls into a size range that is well matched to the genetic and genomics tools currently available for studying single mutations.

Our results identified known loci important for phase variation and correspond well with a recent complete mutagenesis study of the IIRs (Fig. 2B and Fig. S4B) (43). However, we saw marginal effects at binding sites for leucine responsive protein (LRP) and IHF that also affect phase variation (52, 53), likely because single mutations have minor influences on binding affinities. Indeed, large effects on fimS recombination rates were observed only with multiple mutations in these binding sites that disrupted in vitro binding (52, 53). Furthermore, we have studied fimS under classical phase ON and phase OFF induction conditions. Some fimS regulatory mechanisms integrate with growth phase and other global regulatory networks, such as CRP and DNA supercoiling (18,22); thus expanding studies to additional conditions may be warranted.
The UIRs do not appear to be directly involved in the fimS inversion event itself but instead are involved through feedback from low FimA levels or disrupted TIP assembly. The UIRs help maintain high levels of fimA, likely through a base-paired RNA structure. Increased fimA transcript levels have been attributed to transcriptional attenuation between fimA and fimT (45), and the structure of the UIRs raises the possibility of regulation by a small RNA. Notably, another IR in the 5’ UTR of fimA, upstream of the UIRs, has been identified by sequence analysis (42); however, neither that report nor our data show any function for that sequence. An RNA structure at the 3’ end of fimE has been noted to influence phase switching as well (54).

There is unexpectedly large regulatory variation among strains of the same bacterial species, most prominently in Salmonella (55). Regarding TIP, UPEC isolates carry a variable complement of alternative recombinases (ipuA, ipuB, fimX) that themselves have varying effects on fimS inversion; ipuA inverts bidirectionally, whereas fimX favors OFF-to-ON switching (29, 30, 56). FimX also inverts another genomic locus, hyaR, using different IRs (14). Even with a similar complement of recombinases, further regulatory variation may exist. ST131 strain EC958 carries an insertion sequence that inactivates fimB but retains OFF-to-ON switching ability, presumably through FimX or a noncanonical activity of FimE (31). Finally, there are hints of evolutionary interactions between the expression of TIP and other virulence factors. In many UPEC strains, such as 536, the leuX-associated pathogenicity island (PAI) can be excised spontaneously (57). In UTI89, however, this excision results in the loss of leuX-encoded tRNALeu1 causing inefficient translation of FimB and reduced TIP expression (56). Thus, TIP expression is evolutionarily connected to the maintenance of this PAI and its virulence factors, including hemolysin and CNF1 (56). These examples of regulatory variation highlight the importance of working directly with pathogenic clinical isolates.

In contrast to this regulatory variation, the sequence of fimS is relatively well conserved across E. coli strains. Across 302 E. coli strains, including both pathogens and nonpathogens, fimS is on average 96.5% identical (using UTI89 as the reference sequence) (Fig. S3A). In general, less conserved fimS sequences are less likely to affect phase variation when mutated (Fig. S3B), although the overall high level of conservation makes such an analysis less useful for directly identifying functional sites.

Other chaperone-usher pilus systems in E. coli undergo phase variation via distinct mechanisms. Phase variation of P pili is mediated by the methylation of dam sites near its operon promoter; a similar mechanism controls S pili (49). Although phase variation appears to be stochastic, it also is subject to directed regulatory controls. For instance, disruption of P pilus assembly leads to periplasmic accumulation of pilus subunits, activating an unfolded protein response through the Cpx two-component system and feeding back to turn the P pilus promoter phase OFF and inhibit transcription (58). Additional mechanisms may coordinate expression of different pili within a bacterium, so that only one pilus is expressed at a time (33). Because some E. coli carry more than 10 distinct pilus systems, these mechanisms may be extremely complex. The UIRs also may act in such pilus cross-regulation. The sensing of TIP misassembly is not as well characterized as that for P pili, although it does not seem to proceed through the Cpx system (45). Interestingly, bacteria can detect nonfunctional TIP subunits; disrupting the FimH mannose-binding pocket by mutation (Q133K) or small molecule inhibition turns fimS phase OFF (10). In UTI mutants, the loss of appropriate pilus subunit stoichiometry (possibly affecting pilus assembly) also leads to negative feedback regulation of fimS and up-regulation of S pili independent of the fim recombinases.

As an epigenetic phenomenon, phase variation is often difficult to study, particularly when it is mediated by DNA methylation. Although we have focused on inversion, improvements in targeted methylation analysis (59) could enable the study of cis-acting sites at the P- and S-pilus promoters (49). Furthermore, PacBio sequencing can detect transposition, short sequence repeats, and slipped-strand mispairing; these capabilities may be useful for studying phase variation of other nonplus structures, such as flagella and lipoproteins (49), or restriction-modification systems, which impact phage defense and global gene expression (60–62). Additionally, with appropriate screen design, unbiased discovery of trans-acting factors for phase variation may be feasible. Finally, similar methods could be extended to perform traditional “promoter bashing” with native copy number and context (63) or studies of untranslated and genic regions (37).

In this paper we describe a high-throughput approach, coupling technological advances in next-generation sequencing and recombineering, that allowed us to interrogate the effects of more than 1,000 cis-acting mutations on fimS switching. This method provides further opportunities for phase variation studies in other systems and also in studies involving any genomic changes that can be detected by sequencing. The resultant discovery of previously unknown regulatory elements in fimS has provided additional insights into phase variation of type 1 pili, cross-talk between pilus systems, and regulatory complexities of operons.

Materials and Methods

Detailed materials and methods describing the mutagenesis approach, computational and statistical analyses, mutant generation, and in vitro and in vivo assays can be found in SI Materials and Methods. The strains, plasmids, and primers used are given in Tables S1–S3, respectively. Phase assay primers and expected band sizes are given in Table S4.

ACKNOWLEDGMENTS. We thank members of the S.L.C. laboratory for technical assistance and useful discussions, especially Liow Luting (who made the strains SLC-490, -532, and -533), Majid Eshaghi, and Jade Chung; Scott Hultgren for the generous gift of FimA antibodies; and William Burkholder and Lewis Hong for assistance with the protocol for adding unique molecular identifiers. This work was supported by National Research Foundation, Singapore Grant NRF-RF2010-10 (to S.L.C.), and the Genome Institute of Singapore/Agency for Science, Technology, and Research.


Supporting Information

Zhang et al. 10.1073/pnas.1522958113

SI Materials and Methods

Media and Culture Conditions. For ON induction of type 1 pili, a single bacterial colony on an LB agar plate was used to inoculate 5 mL of LB medium in a 25-mL flask and was statically incubated at 37 °C for 18–24 h. Ten microliters of this culture were used to inoculate a fresh 10 mL of LB medium in a 50-mL flask that was incubated for 24 h under the same conditions. The procedure for OFF induction of type 1 pili is similar, except that cultures were incubated at 25 °C under shaking conditions.

For plasmid maintenance, transformed bacterial strains were grown in LB medium supplemented with kanamycin (50 μg/mL).

Generation of fimS Mutagenesis Libraries. The fimS sequence with the flanking 50 nt was amplified with PCR from fimS LON (SLC-532) or locked-OFF strains (SLC-533) using HZpri030 and HZpri037; these locked strains have a WT fimS but lack all three fim recombinases required for phase switching. This PCR product was used as the template for PCR mutagenesis of fimS with the GeneMorph II Random Mutagenesis Kit (Agilent Technologies). The mutated fimS product was used as substrate for homologous recombination and negative selection (36) in UTI89 to create a library of fimS mutants. Several libraries originating from either fimS (ON) or fimS (OFF) templates were built, with the number of clones ranging from 1.3 × 10^3 to 2.1 × 10^4. For ON induction, libraries made from fimS (OFF) were used, and vice versa for OFF induction. This method ensured that the cells started at the same phase before induction.

Illumina Sequencing Library Preparation and Analyses. gDNA was purified from cultures using phenol-chloroform extraction. fimS was amplified using fimS-specific primers that incorporate the necessary sequencing adaptor sequences (HZpri082, HZpri085, HZpri086, and HZpri089–091). Sequencing was performed on an Illumina MiSeq Sequencer, and the reads were aligned to fimS from UTI89 (38) using the Burrows–Wheeler alignment (BWA) tool (65) with default parameters. Alignments were analyzed for mutations and corresponding phase variation using SAMtools (66).

PacBio Sequencing Library Preparation. gDNA was purified from cultures using phenol-chloroform extraction. Barcoding PCR to introduce unique molecular identifiers was performed on gDNA with HZpri197 and HZpri198 (39, 67). Exonuclease I (New England Biolabs) was added to degrade the first set of primers before HZpri199 and HZpri200 were used to amplify the barcoded fimS sequences. The resultant PCR product was purified with AMPure beads. PacBio sequencing libraries were prepared from the PCR products using the SMRTbell Template Prep Kit 1.0 (Pacific Biosciences).

PCR Sequencing and Analyses. Sequencing was performed on a PacBio RS II (Pacific Biosciences), and raw sequencing reads were processed with the RS_ReadsOfInsert.1 protocol on SMRT Portal versions 2.2.0 or 2.3.0 (Pacific Biosciences) to obtain circular consensus sequences. The reads were aligned to fimS using the BWA tool (65), and alignments were analyzed for mutations and corresponding phase variation using SAMtools (66). P values were calculated using a binomial test with Bonferroni correction in R version 3.2.0 (68).

For comparison with the study by McCusker et al. (43), P values were calculated using the Spearman correlation test based on Monte Carlo simulation in R (68–70). Assignment of IR mutations to qualitative categories was taken from the FimB-mediated ON-to-OFF inversion experiments (43).

fimS Conservation Analyses. The fimS sequences from 302 complete and draft E. coli strains (available on GenBank as of June 2013) were extracted using BLASTN (71) with the UTI89 fimS sequence as a query and were aligned using MUSCLE (72). Conservation at each position was calculated as the percentage of fimS sequences that had the same nucleotide as the UTI89 fimS sequence.

Generation of fimS Mutant Strains. Red recombinase-mediated homologous recombination (73) was used to replace genomic fimS in UTI89 with a positive–negative selection cassette. This cassette contains a toxin gene (relE) under rhamnose induction and confers kanamycin resistance (36). The resultant fimS-knockout strain, HZ142, was used to generate other fimS mutant strains by a second homologous recombination with the desired fimS mutant sequences.

The phase LON strains with the fimS (G398A) mutation were created by replacing genomic fimS in phase LON strains with the positive–negative selection cassette (36). The phase LON strains used were HZ215 and SLC-490. HZ215 harbored deletions in fimB, fimE, and fimX, whereas SLC-490 contains an inversion of IRL; both strains are unable to recombine at the fim switch. The selection cassette was replaced by fimS (G398A) in a second homologous recombination.

T1P Phase Assay. The T1P phase assay was modified from Chen et al. (74). To determine the phase of fimS, the genomic fimS region was amplified with PCR. The resultant PCR product was digested with HinII, and the restriction fragments were run on an agarose gel to determine phase (Table S4). To calculate % phase ON/OFF, we used ImageJ (75) for analyses of band intensities.

HA Assay. The HA assays were performed as previously described with minor modifications (76). One milliliter of bacterial cells (OD_600 = 1.0) was gently pelleted and resuspended in 100 μL PBS with or without 4% mannose. Twenty-five microliters of bacterial suspension were serially diluted (twofold) in a row of 96-well V-bottomed plates (#3897; Corning); each well contained 25 μL of PBS with or without 4% mannose. Twenty-five microliters of guinea pig erythrocytes were added to each well. The plate was gently mixed and incubated overnight at 4 °C. The HA titer is the greatest dilution of bacteria that resulted in visible clumping of erythrocytes.

To assess agglutination by S pili, guinea pig erythrocytes were desialylated before being added to diluted bacteria (77). Briefly, erythrocytes were resuspended in 900 μL of 1x PBS and 1x GlycoBuffer 1. Two microliters of α2-3,6,8 neuraminidase (New England Biolabs) was added to the erythrocytes and incubated at 37 °C for 1 h with gentle rotation. The desialylated erythrocytes were washed once with 1x PBS before use.

qRT-PCR. RNA was extracted from bacterial cultures after ON induction of type 1 pili using hot phenol. Contaminating gDNA was removed using the TURBO DNA-free Kit (Life Technologies), and cDNA was synthesized using SuperScript II (Life Technologies). The qRT-PCR primers used were HZpri227 and HZpri228 for fimA; HZpri229 and HZpri230 for fimI; HZpri231 and HZpri232 for fimC; HZpri233 and HZpri234 for fimD; HZpri235 and HZpri236 for fimF; HZpri237 and HZpri238 for fimG; HZpri239 and HZpri240 for fimH; and HZpri243 and
HZpri244 for rrsA. qRT-PCR was performed using SYBR Fast Universal qRT-PCR Kit (KAPA Biosystems) on either a LightCycler 480 system (Roche) or an Applied Biosystems 7500 Fast Real-time PCR system (Life Technologies) with technical triplicates. Fold change was calculated from cycle threshold values using rrsA (housekeeping gene) for normalization, unless stated otherwise.

**Construction of fimA-Expressing Plasmid.** To construct a plasmid expressing fimA, genomic fimS-fimA and fimA were amplified by PCR using HZpri260 and HZpri261 and HZpri064 and HZpri279, respectively. The PCR products were cloned into the pCR-Blunt II-TOPO vector (Life Technologies), and the correct plasmids were identified by Sanger sequencing (pTOPO::fimS-fimA and pTOPO::fimA). pTOPO::fimA drives the expression of fimA through a lacZ promoter, which is upstream of the cloning site in the pCR-Blunt II-TOPO vector. The control TOPO vector (pTOPO) was obtained from a false-positive clone and was verified by Sanger sequencing not to contain fimA. Transformation was achieved by electroporation.

**Structure Probing.** RNA was transcribed in vitro from a DNA template using the MEGAscript T7 Transcription Kit (Thermo Fisher Scientific) according to the manufacturer’s instructions, and structure probing was performed using a SHAPE molecule, NAI (78). Briefly, RNA was heated to 90 °C to denature the RNA, was cooled on ice quickly, and then was heated to 37 °C in the presence of 10× RNA structure buffer [150 mM NaCl, 10 mM MgCl2, Tris (pH 7.4)]. Next, the RNA was incubated with NAI at 37 °C for 20 min, followed by inactivation using phenol-chloroform extraction. WT RNA that was not incubated with NAI was used as an untreated control.

Modification sites along the RNA were identified using primer extension. Briefly, a primer (HZpri280) located ~30–50 bases downstream of the region of interest was labeled with γ-P32-ATP using T4 PNK kinase and was purified using a 15% Tris/borate/EDTA (TBE)-urea PAGE gel. The labeled primer was incubated with the RNA at 65 °C for 5 min and then at 35 °C for 5 min and then was cooled to 4 °C. To detect structure-probed sites, 3 μL of enzyme mix (4:1:1 of First-strand buffer:DTT:NTP) was added to the reaction and incubated at 52 °C for 1 min before SuperScript III was added to the reaction for another 10 min of incubation at 52 °C. A sequencing ladder for the RNA was generated by adding 1 μL of ddNTP (5 mM) to the reaction between the addition of the enzyme mix and SuperScript III. The RNA samples were denatured with 4 M sodium hydroxide before being loaded onto a 7 M TBE-Urea PAGE sequencing gel for visualization of modification sites.

**Western Blots.** After ON induction, 1 mL of bacterial culture at OD600 = 1.0 was pelleted and resuspended in 200 μL of Tricine sample buffer (Bio-Rad). To disassemble type 1 pili, 4 μL of 1 M HCl was added, and the suspension was boiled at 100 °C for 5 min. Four microliters of 1 M NaOH was added for neutralization. As controls, the samples were boiled without the addition of acid or were left at room temperature for 5 min. After centrifugation (21,000 × g for 5 min), the supernatant was analyzed by SDS/PAGE (15% polyacrylamide gel). Proteins were transferred to a nitrocellulose membrane using a Trans-Blot SD Semi-Dry transfer cell (Bio-Rad). Blots were probed with anti-type 1 pili rabbit antibody (1:3,000) for 3 h. The secondary antibody used was Amersham ECL anti-rabbit IgG HRP-linked whole antibody (1:5,000). Blots were visualized using Amersham HRP substrate in a ChemiDoc machine (Bio-Rad).
Fig. S1. Mapping the effect of fimS mutations on phase variation. (A) Comparison of mutation distributions of sequencing libraries sequenced with different platforms. The percentage of sequenced reads mutated (y axis) at each fimS position (x axis) is plotted. (Top) Unmutagenized library (PacBio versus Illumina). Note the different scale for PacBio. (Middle) PacBio sequencing (unmutagenized versus mutagenized libraries). (Bottom) Illumina sequencing (unmutagenized versus mutagenized libraries). (B) Frequency graphs of mutations per sequenced read. The number of reads $\times 10^3$ (y axis) is plotted against the number of mutations per read (x axis) under ON induction (Upper) and under OFF induction (Lower). (C) Mutation distributions across fimS. The percentage of sequenced reads mutated (y axis) at each position (x axis) is plotted under ON induction (Upper) and OFF induction (Lower).
Fig. S2. Detailed phase variation map with significance. Each position in fimS is represented by four rectangles (representing all four possible nucleotides) whose brightness and color indicate the number of mutant sequencing reads observed and the percentage of reads that are phase OFF (based on the scale at bottom right), respectively. WT levels of OFF frequencies are shown in white, lower values are shown in green, and higher values are shown in red. The nucleotide position within fimS is shown below the map (position 1 maps to genomic coordinate 4907377 in UTI89), and the WT fimS sequence is shown above the map. Selected features are indicated by labeled colored bars (12, 26, 42, 52, 64). “FimBE”, “LRP”, and “IHF” indicate the binding sites the respective proteins; “−10” and “−35” indicate the relative positions of the transcriptional start site, and “IRL” and “IRR” indicate the pair of IRs. Yellow asterisks denote P values less than 0.01. (A) Under ON induction. (B) Under OFF induction.
Fig. S3. Analysis of fimS conservation and correlation with phase variation. (A) Phase variation map (under OFF induction) of fimS mutations with conservation. Numbers correspond to nucleotide positions within fimS (position 1 maps to genomic coordinate 4907377 in UTI89). A genetic map of fimS is depicted as a gray line with selected features indicated by labeled colored bars and triangles (12, 26, 42, 52, 64). “FimBE,” “LRP,” and “IHF” indicate the binding sites for the respective proteins; “−10” and “−35” indicate the relative positions of the transcriptional start site, and “IRL” and “IRR” indicate the pair of IRs. Below this genetic map, each position in fimS is represented by a rectangle whose brightness and color indicate the number of mutant sequencing reads observed and the percentage of reads that are phase OFF (based on the scale at top right), respectively; data are averaged for all mutations at that position. WT levels of OFF frequencies are shown in white, lower values are shown in green, and higher values are shown in red. The black bars below the phase variation map represent percentage similarity (y axis) of fimS sequences from 302 E. coli strains to UTI89. (Upper) Conservation (identity to the UTI89 fimS sequence) is scaled from 90 to 100%. (Lower) Conservation is scaled from 0 to 100%. (B) Percentage of conservation with UTI89 (y axis) is plotted against percentage of phase OFF reads (x axis). Each point represents each fimS position (480 total), with its corresponding percentage similarity and phase OFF percentage when mutated. The vertical green line denotes the phase OFF percentage observed for WT fimS, and the red box denotes data points that are either <0.5-fold (OFF induction) or more than twofold (ON induction) of that value. These two sets of data points were used to determine significance using the Mann-Whitney-Wilcoxon test in R (68). (Left) Under OFF induction. (Right) Under ON induction.
Fig. S4. Accurate sequencing-based analysis of fimS phase identifies known and previously unknown loci involved in phase regulation. (A) Detailed phase variation map for the fimS IRs (purple arrows) under ON induction. Numbers correspond to nucleotide positions within fimS. Each position in the IRs is represented by four rectangles (representing all four possible nucleotides) whose brightness and color indicate the number of mutant sequencing reads observed and the percentage of reads that are phase OFF (based on the scale below the maps), respectively. WT levels of OFF frequencies are shown in white, lower values are shown in green, and higher values are shown in red. The WT IR sequence is depicted below the map. (B) Comparison of phase variation effects with those reported by McCusker et al. (43) under ON induction. Each point represents the % phase OFF for a single IR mutation, with the x-axis categories taken from ref. 43. The median for each x-axis category is indicated by a black solid line. Only IR mutations with at least 10 sequencing reads were plotted (n = 41). (C) HA assay of the IR mutants G50A and C358T under ON induction. Average log$_{2}$(HA titers), with SEs (y-axis), for each mutant (x-axis) are plotted using data from at least three independent experiments. Blue and orange bars represent HA titers performed without mannose and with 4% mannose added, respectively. (D) Type 1 pilus phase assay of UIR mutants under ON induction. Mutations are shown above the blots; capitalized letters represent mutations in UIRs. "WT" refers to the parental UTI89 strain band; sizes corresponding to the phase ON or OFF orientation are indicated on the left. (E) HA assay. Average log$_{2}$(HA titer), with SEs (y-axis), for each mutant (x-axis) are plotted using data from at least three independent experiments. Colors represent different mutations, as indicated in the keys. The purple solid lines denote mutations within UIRs. (Upper) Without mannose. (Lower) With mannose.
The UIRs likely form a base-paired RNA structure to affect *fimA* transcript levels. "WT" refers to WT UTI89 strain. Assays were performed after phase ON induction. (A) Type 1 pilus phase assay. Mutations are shown above the blots. WT band sizes corresponding to the phase ON or OFF orientation are indicated on the left. (B) HA assay. Average log₂ (HA titers) with SEs (y axis) for each mutant (x axis) are plotted using data from at least three independent experiments. Blue and orange bars represent HA titers performed without mannose and with 4% mannose added, respectively. (C) HA of strains over-expressing *fimA in trans*. Average log₂ (HA titers) with SEs (y axis) for each strain (x axis) are plotted using data from at least three independent experiments. Blue and orange bars represent HA titers performed without mannose and with 4% mannose added, respectively.
UIR mutants regulate phase variation through a previously unknown mechanism. "WT" refers to the WT UTI89 strain. All assays were performed after growth in ON-inducing conditions. (A) Type 1 pilus phase assay of T1P-deficient strains. Mutations are shown above the blots. Band sizes corresponding to the phase ON or OFF orientation are indicated on the right. (B) qRT-PCR of fim transcripts of T1P-deficient strains. Colors represent different mutations as indicated in the key. Average fold change (normalized to fimH) and SE are plotted (y-axis) using data from at least three independent experiments. (C) HA assay of T1P-deficient strains. Average log2(HA titers), with SEs (y-axis), for each strain (x-axis) are plotted using data from at least three independent experiments. Blue and orange bars represent HA titers performed without mannose and with 4% mannose added, respectively. Data for desialylated erythrocytes with 4% mannose were 20 ± 0 for all samples and were not plotted. (D) Type 1 pilus phase assay of phase LON strains with fimS (G398) mutation. Mutations are shown above the blots. Band sizes corresponding to the phase ON or OFF orientation are indicated on the right. (Left) Mutant IR LON. (Right) Δfim(BEX) LON. (E) qRT-PCR of fim transcripts for Δfim(BEX) LON strains. Colors represent different mutations as indicated in the key. Average fold change and SE (y-axis, normalized to WT) for each fim transcript (x-axis) are plotted using data from at least three independent experiments. (F) HA assay of phase LON strains with the fimS (G398) mutation. Average log2(HA titers), with SEs (y-axis), for each mutant (x-axis) are plotted using data from at least three independent experiments. Blue and orange bars represent HA titers performed without mannose and with 4% mannose added, respectively. (Left) Mutant IR LON. (Right) Δfim(BEX) LON. (G) Western blot using an antibody that detects both FimA and SfaA. The different mutant strains are shown above the blots. Protein ladder sizes are on the left, and positions of bands corresponding to FimA and SfaA are identified on the right.
### Table S1. Strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>The following strains were created from UTI89 and carry chromosomal alterations as described without any need for selection during growth</td>
<td></td>
</tr>
<tr>
<td>HZ086, HZ087</td>
<td>fimS (G50A)</td>
<td>G to A mutation at position 50 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ089, HZ090</td>
<td>fimS (C358T)</td>
<td>C to T mutation at position 358 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ164, HZ165</td>
<td>fimS (G398A)</td>
<td>G to A mutation at position 398 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ142</td>
<td>∆fimS::kan-prhaB-reIE</td>
<td>Replacement of fimS with positive-negative selection cassette</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ183–185</td>
<td>fimS_ON (INV48-56, G398A)</td>
<td>Inversion of left inverted repeat, G to A mutation at position 398 in fimS, phase locked ON</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ200–202</td>
<td>∆fimX, ∆(fimB-fimE), fimS_ON (G398A)</td>
<td>Deletion of fim recombinases, phase locked ON, G to A mutation at position 398 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ215–217</td>
<td>∆fimX, ∆(fimB-fimE), fimS_ON</td>
<td>Deletion of fim recombinases, phase locked ON</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ222, HZ293</td>
<td>fimS (C403G)</td>
<td>C to G mutation at position 403 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ224, HZ225</td>
<td>fimS (C403A)</td>
<td>C to A mutation at position 403 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ230, HZ231, HZ237</td>
<td>fimS (T396G)</td>
<td>T to G mutation at position 396 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ236, HZ275, HZ276</td>
<td>fimS (T396A)</td>
<td>T to A mutation at position 396 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ238, HZ277</td>
<td>fimS (A400T)</td>
<td>A to T mutation at position 400 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ240, HZ234–239</td>
<td>fimS (A400G)</td>
<td>A to G mutation at position 400 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ241, 268, 279</td>
<td>fimS (A400C)</td>
<td>A to C mutation at position 400 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ242, HZ247</td>
<td>fimS (A404G)</td>
<td>A to G mutation at position 404 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ244, HZ245</td>
<td>fimS (A404T)</td>
<td>A to T mutation at position 404 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ246</td>
<td>fimS (A404C)</td>
<td>A to C mutation at position 404 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ252, HZ253</td>
<td>fimS (C440T)</td>
<td>C to T mutation at position 440 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ255, HZ256</td>
<td>fimS (G398C, C440T)</td>
<td>G to A mutation at position 398, C to T mutation at position 440 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ272–274</td>
<td>fimS (T396C)</td>
<td>T to C mutation at position 396 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ295, HZ297, HZ299</td>
<td>fimS (G397A)</td>
<td>G to A mutation at position 397 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ296, HZ298, HZ301</td>
<td>fimS (G397T)</td>
<td>G to T mutation at position 397 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ300, HZ302, HZ303</td>
<td>fimS (G397C)</td>
<td>G to C mutation at position 397 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ325, HZ326, HZ400, HZ402–205</td>
<td>fimS (C403T)</td>
<td>C to T mutation at position 403 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ376–378</td>
<td>fimS (G398C, C440G)</td>
<td>G to C mutation at position 398, C to G mutation at position 440 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ379–383</td>
<td>fimS (G398C)</td>
<td>G to C mutation at position 398 of fimS</td>
<td>This paper</td>
</tr>
<tr>
<td>SLC282</td>
<td>∆fimB-fimH</td>
<td>Deletion of fim operon</td>
<td>This paper</td>
</tr>
<tr>
<td>SLC478</td>
<td>fimH (Q133K)</td>
<td>fimH with Q to K substitution at amino acid 133</td>
<td>(1)</td>
</tr>
<tr>
<td>SLC490</td>
<td>fimS_ON (INV48-56)</td>
<td>Inversion of left inverted repeat in fimS, phase locked ON</td>
<td>This paper</td>
</tr>
<tr>
<td>SLC532</td>
<td>∆fimX, ∆(fimB-fimE), fimS_On</td>
<td>Deletion of fim recombinases, phase locked ON</td>
<td>This paper</td>
</tr>
<tr>
<td>SLC533</td>
<td>∆fimX, ∆(fimB-fimE), fimS_Off</td>
<td>Deletion of fim recombinases, phase locked OFF</td>
<td>This paper</td>
</tr>
<tr>
<td>UTI89</td>
<td>WE</td>
<td>Clinical cystitis isolate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The following strains carry plasmids that require kanamycin selection</td>
<td></td>
</tr>
<tr>
<td>HZ342–344, HZ347–349, HZ352–354</td>
<td>fimS (G398A)/ pTOPO::fimS-fimA</td>
<td>G to A mutation at position 398 of fimS; Carries TOPO vector expressing fimS-fimA</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ357–358</td>
<td>UTI89 with pTOPO</td>
<td>Wildtype carrying control TOPO vector</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ362–363, HZ367–368</td>
<td>fimS (G398A)/pTOPO</td>
<td>G to A mutation at position 398 of fimS; Carries control TOPO vector</td>
<td>This paper</td>
</tr>
<tr>
<td>HZ387–389, HZ390–392</td>
<td>fimS (G398A)/ pTOPO::fimA</td>
<td>G to A mutation at position 398 of fimS; Carries TOPO vector expressing fimS-fimA</td>
<td>This paper</td>
</tr>
</tbody>
</table>

### Table S2. Plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Genotype</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH082-084</td>
<td>pTOPO::fimS-fimA</td>
<td>fimS-fimA in TOPO vector</td>
<td>This paper</td>
</tr>
<tr>
<td>pH085-087</td>
<td>pTOPO</td>
<td>Control TOPO vector</td>
<td>This paper</td>
</tr>
<tr>
<td>pH092</td>
<td>pTOPO::fimA</td>
<td>fimA in TOPO vector</td>
<td>This paper</td>
</tr>
</tbody>
</table>

Zhang et al. www.pnas.org/cgi/content/short/1522958113
**Table S3. Primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>HZpri030</td>
<td>GGGAAAGAAATAATCTCATAAAC</td>
</tr>
<tr>
<td>HZpri037</td>
<td>CTCAGGGAGCAAGGGCCACAG</td>
</tr>
<tr>
<td>HZpri044</td>
<td>CGGTATTTGAGCTGGAAACCCG</td>
</tr>
<tr>
<td>HZpri045</td>
<td>GGTTTGATCAACAGAGCCCTG</td>
</tr>
<tr>
<td>HZpri063</td>
<td>GCTGCTTTCCTTTTTAAAAACACTATTTTC</td>
</tr>
<tr>
<td>HZpri064</td>
<td>ATGAAAAATTAAACCTCTCGCAATTG</td>
</tr>
<tr>
<td>HZpri082</td>
<td>GACGCTCTTCCGATCTAGTTTTAATATTGATATTAAAGTTAAAAC</td>
</tr>
<tr>
<td>HZpri085</td>
<td>GACGCTCTTCCGATCTTGTTTTAATATTGATATTAAAGTTAAAAC</td>
</tr>
<tr>
<td>HZpri086</td>
<td>AATGATACGGCGACCACCGAGATCTACACTCTTTCTCTACAGCTACTGCTTCTGATCTACGTA</td>
</tr>
<tr>
<td>HZpri089</td>
<td>AATGATACGGCGACCACCGAGATCTACACTCTTTCTCTACAGCTACTGCTTCTGATCTACGTA</td>
</tr>
<tr>
<td>HZpri090</td>
<td>GGGTTGATCAACAGAGCCCTG</td>
</tr>
<tr>
<td>HZpri091</td>
<td>CAAGCGGAGACGGATCAAGATGTTTTAATTTTATCAT</td>
</tr>
<tr>
<td>HZpri151</td>
<td>TTTACACTAATTTGATAAAGTTTTAAAAAAACAAAAATA</td>
</tr>
<tr>
<td>HZpri197</td>
<td>CGACGTTAAACAGCGCTTCGAGTTTTAATTTTATCAT</td>
</tr>
<tr>
<td>HZpri198</td>
<td>CAAGCGGAGACGGATCAAGATGTTTTAATTTTATCAT</td>
</tr>
<tr>
<td>HZpri199</td>
<td>CGACGTTAAACAGCGCTTCGAGTTTTAATTTTATCAT</td>
</tr>
<tr>
<td>HZpri200</td>
<td>CAAGCGGAGACGGATCAAGATGTTTTAATTTTATCAT</td>
</tr>
<tr>
<td>HZpri223</td>
<td>GAACCGCTTCTCTGGCAAC</td>
</tr>
<tr>
<td>HZpri224</td>
<td>AATGATACGGCGACCACCGAGATCTACACTCTTTCTCTACAGCTACTGCTTCTGATCTACGTA</td>
</tr>
<tr>
<td>HZpri225</td>
<td>GGGTTGATCAACAGAGCCCTG</td>
</tr>
<tr>
<td>HZpri226</td>
<td>GGGTTGATCAACAGAGCCCTG</td>
</tr>
<tr>
<td>HZpri227</td>
<td>GGTTGATCAACAGAGCCCTG</td>
</tr>
<tr>
<td>HZpri228</td>
<td>GAACCGCTTCTCTGGCAAC</td>
</tr>
<tr>
<td>HZpri229</td>
<td>TTTACACTAATTTGATAAAGTTTTAAAAAAACAAAAATA</td>
</tr>
<tr>
<td>HZpri230</td>
<td>TTTACACTAATTTGATAAAGTTTTAAAAAAACAAAAATA</td>
</tr>
<tr>
<td>HZpri231</td>
<td>TTTACACTAATTTGATAAAGTTTTAAAAAAACAAAAATA</td>
</tr>
<tr>
<td>HZpri232</td>
<td>CAGCTCTCTTCTCTGGCAAC</td>
</tr>
<tr>
<td>HZpri233</td>
<td>CATTTATACGGCGAGCCAC</td>
</tr>
<tr>
<td>HZpri234</td>
<td>ATCCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri235</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri236</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri237</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri238</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri239</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri240</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri241</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri242</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri243</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri244</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri245</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri246</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri247</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri248</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri249</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri250</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri251</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri252</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri253</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri254</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri255</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri256</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri257</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri258</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri259</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri260</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri261</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri262</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri263</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri264</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri265</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri266</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri267</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri268</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri269</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri270</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri271</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri272</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri273</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri274</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri275</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri276</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri277</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri278</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri279</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
<tr>
<td>HZpri280</td>
<td>GGTCTCCGAGCTTTCTCAGAGAATCGCTTCTGATCTAC</td>
</tr>
</tbody>
</table>

*With phosphorothioates at the 3′ end to prevent exonuclease degradation.

**Table S4. Phase assay primers and expected band sizes**

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Primers</th>
<th>Phase ON fimS fragment sizes, nt</th>
<th>Phase OFF fimS fragment sizes, nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δfim(BEX) strains</td>
<td>HZpri063 and HZpri151</td>
<td>420, 60</td>
<td>343, 137</td>
</tr>
<tr>
<td>Strains with fimA-expressing plasmid</td>
<td>HZpri044 and HZpri045</td>
<td>588, 262, 27</td>
<td>545, 305, 27</td>
</tr>
<tr>
<td>All others</td>
<td>HZpri241 and HZpri242</td>
<td>485, 74</td>
<td>357, 202</td>
</tr>
</tbody>
</table>