Title

Environment-directed activation of the Escherichia coliflhDC operon by transposons

Journal

Microbiology, 163(4)

ISSN

1350-0872

Authors

Zhang, Zhongge
Kukita, Chika
Humayun, M Zafri
et al.

Publication Date

2017-04-01

DOI

10.1099/mic.0.000426

Peer reviewed

Microbiology

Environment-directed Activation the Escherichia coli flhDC Operon by Transposons

--Manuscript Draft--

Manuscript Number:	MIC-D-16-00325R1
Full Title:	Environment-directed Activation the Escherichia coli flhDC Operon by Transposons
Article Type:	Standard
Section/Category:	Regulation
Corresponding Author:	Milton H. Saier, Jr. University of California, San Diego La Jolla, CA UNITED STATES
First Author:	Zhongge Zhang
Order of Authors:	Zhongge Zhang
	Chika Kukita
	Mir Zafri Humayun, PhD
	Milton H. Saier, Jr.
Abstract:	The flagellar system in E. coli K12 is expressed under the control of the flhDC-encoded master regulator, FlhDC. Transposition of insertion sequence (IS) elements to the upstream flhDC promoter region up-regulates transcription of this operon, resulting in more rapid motility. Wang and Wood (2011) provided evidence that insertion of IS5 into upstream activating sites occurs at higher rates in semisolid agar media in which swarming behavior is allowed as compared with liquid or solid media where swarming cannot occur. We confirm this conclusion, and show that three IS elements, IS1, IS3 and IS5, transpose to multiple upstream sites within a 370 bp region of the flhDC operon control region. Hot spots for IS insertion correlate with positions of stress-induced DNA duplex destabilization (SIDD). We show that IS insertion occurs at maximal rates in 0.24% agar with rates decreasing dramatically with increasing or decreasing agar concentrations. In mixed cultures, we show that these mutations preferentially arise from the wild type parent at frequencies of up to 3x10-3/cell/day when the inoculated parental and co-existing IS-activated mutant cells are entering the stationary growth phase. We rigorously show that the apparent increased mutation frequencies cannot be accounted for by increased swimming or by increased growth under the selective conditions used. Thus, our data are consistent with the possibility that appropriate environmental conditions, namely those that permit but hinder flagellar rotation, result in the activation of a mutational pathway that involves IS element insertion upstream of the flhDC operon.

Environment-directed Activation of the *Escherichia coli flhDC* Operon by Transposons

Zhongge Zhang[1,2], Chika Kukita[1,2], M. Zafri Humayun[3], and Milton H. Saier, Jr. [1]*

[1]Department of Molecular Biology, Division of Biological Sciences, University of California at San Diego, La Jolla, CA 92093-0116.
[3]Department of Microbiology, Biochemistry & Molecular Genetics, Rutgers - New Jersey Medical School, Newark, NJ 07101-1709

[2]These two authors contributed equally to the work reported.

*Corresponding Author:
Tel +1 858 534 4084
Fax +1 858 534 7108
E-mail: msaier@ucsd.edu

Key words: IS-mediated Directed Mutation, Flagellar Master Regulator, FlhDC, Transposon, Gene Activation, SIDD

SUMMARY

The flagellar system in *E. coli* K12 is expressed under the control of the *flhDC*-encoded master regulator, FlhDC. Transposition of insertion sequence (IS) elements to the upstream *flhDC* promoter region up-regulates transcription of this operon, resulting in more rapid motility. Wang and Wood (2011) provided evidence that insertion of IS*5* into upstream activating sites occurs at higher rates in semisolid agar media in which swarming behavior is allowed as compared with liquid or solid media where swarming cannot occur. We confirm this conclusion, and show that three IS elements, IS*1*, IS*3* and IS*5*, transpose to multiple upstream sites within a 370 bp region of the *flhDC* operon control region. Hot spots for IS insertion correlate with positions of stress-induced DNA duplex destabilization (SIDD). We show that IS insertion occurs at maximal rates in 0.24% agar with rates decreasing dramatically with increasing or decreasing agar concentrations. In mixed cultures, we show that these mutations preferentially arise from the wild type parent at frequencies of up to $3x10^{-3}$/cell/day when the inoculated parental and co-existing IS-activated mutant cells are entering the stationary growth phase. We rigorously show that the apparent increased mutation frequencies cannot be accounted for by increased swimming or by increased growth under the selective conditions used. Thus, our data are consistent with the possibility that appropriate environmental conditions, namely those that permit but hinder flagellar rotation, result in the activation of a mutational pathway that involves IS element insertion upstream of the *flhDC* operon.

INTRODUCTION

A transposon, a jumping gene, can hop from one location to another in the genome of an organism (1). Transposons thereby induce mutations, and the consequences can be beneficial or detrimental to the host cell (2-4). Transposition occurs by a conservative or replicative mechanism and can be regulated by host DNA binding proteins (5-8). Reversible transposon insertion/deletion can result in phase variation of phenotypic properties such as metabolism, virulence and biofilm formation (9-11).

The smallest bacterial transposons, about 1 kb long, are the so-called insertion sequences or "IS elements". Multiple IS elements found in *Escherichia coli* and other prokaryotes often play important roles in molecular and organismal evolution (4, 12). They sometimes modify the expression of global regulatory networks and even change mutation rates (13). Earlier studies had shown that IS insertions can activate genes adjacent to and usually downstream of the insertion sites by a number of distinct mechanisms (14-17). For example, IS*5* has been observed to be inserted in response to nutritional stress upstream of the *E. coli fuc* (18), *bgl* (19), *flhDC* (20), and *glpFK* operons (21-25).

Flagellar motility allows bacteria to reach more nutritive environments and escape from toxic or otherwise deleterious environments (26). Flagellar synthesis and function are energetically expensive; thus, the flagellar system, encoded by more than 50 genes, is highly regulated. Flagellar operons and their promoters in *E. coli* are categorized into three classes, Class I, Class II and Class III, according to their expression timing (27). In several bacteria including *E. coli*, the heterohexameric transcription factor, $FlhD_4C_2$ (hereafter referred to as FlhDC) is the only Class I protein, and is the master regulator of the entire flagellar cascade (28).

Transcription of *flhDC* is regulated by many transcription factors; in fact, no other bacterial operon has been shown to be subject to as complex a control network. Protein transcription factors reported to control *flhDC* expression include RssAB, H-NS, AtoSC, CsgD, GlgS, MatA, YdiV, FmrA, YiiQ, LrhA, HdfR, OmpR, PapX, RcsB, CysB, PhoB, NtrC and the cAMP-Crp complex (16, 17, 27-37). Small RNAs, ArcZ, OmrA, OmrB and OxyS, negatively regulate, and McaS positively regulates *flhDC* expression and cell motility in Hfq-dependent processes by binding to the 5' untranslated region of the *flhDC* mRNA (38, 39). *flhDC* operon expression is also autoregulated by FlhDC (40).

Fahrner and Berg (28) isolated several mutations that enhanced motility and *flhDC* expression, showing that both IS and a kanamycin resistance gene insertions into the upstream region can cause increased motility. None of these mutations eliminated the cAMP-Crp requirement, although several did eliminate the need for H-NS. These investigators concluded that the upstream region into which IS elements insert must represent a negative control region for *flhDC* expression.

Motility and flagellar synthesis are differentially regulated in a complex way during biofilm development (41). FlhDC influences biofilm formation and virulence (42), although the FlhDC regulon is surprisingly restricted with respect to the types of genes included (Fitzgerald et al., 2014). Thus, flagellar biosynthesis and function are sensitive to many external and internal conditions, serving several related functions.

As shown in Supplementary Figure S1, the chromosome of *E. coli* K12 strain BW25113 has ten identical copies of IS*5* (1195 bp long), six copies of IS*1* (768 bp) and five copies of IS*3* (1258 bp) (9). When any one of these three IS elements transposes into the upstream region of *flhDC*, transcription of the operon is up-regulated, resulting in increased transcription of Class II

and Class III genes, and consequent increased motility (17, 20, 26) and the present report). Usually IS elements closest to the target site on the chromosome are involved in the transposition event.

On solid agar (1.5%) plates, increased flagellar gene expression due to IS insertion upstream of the *flhDC* promoter does not allow *E. coli* to swarm, while in agitated liquid media, cells do not need flagella to move. In soft agar (<0.4%) plates, however, wild type cells migrate slowly through the gel (43), and they benefit from higher flagellar expression because the cells with more flagella have greater access to nutrients and can escape toxic end products of metabolism (17, 28, 44). In this article we shall refer to the rate of migration from the point of inoculation outward as the "colony expansion rate" or "swarming rate".

In their 2011 paper, Wang and Wood demonstrated the effects of different environmental conditions on the occurrence of IS*5* hopping into sites upstream of the *flhDC* operon. First, they measured *flhDC*-activating IS*5* hopping frequencies in soft-agar plates (0.3%). The *E. coli* K12 BW25113 strain, weakly motile without an IS element upstream of the *flhDC* promoter, was inoculated into soft-agar plates, and the appearance of swarming zones was recorded. The insertion frequencies on solid-agar plates (1.5% agar) as well as in LB and minimal nutrient liquid media remained low compared to the frequencies observed in soft-agar. Second, they tested the conditions essential for IS*5* hopping. A *flhD*-deleted strain as well as *motA* and *flgK* mutants showed no swarming halo on soft-agar, and no IS*5* insertions in the *flhD* upstream region were identified. Third, biofilm formation and growth rates for the strains with or without IS*5* upstream of *flhD* were measured. The IS*5*-activating insertion mutants gained a greater capacity for biofilm formation than the wild type strain. In fact, IS*5* hopped into the upstream region of *flhD* during biofilm formation, resulting in a subpopulation of more motile cells.

Finally, evidence was presented suggesting that IS*5* insertions upstream of *flhD* probably were not accompanied by a general increase in IS hopping into random insertion sites.

The observations reported by Wang and Wood (44) are consistent with the possibility that environmental conditions influence the rates of beneficial IS insertion mutations, and we sought to confirm and extend these results. We show here that in soft-agar plates, new insertions of IS elements preferentially arise as growth of the parental cells slows down and ceases. We further show that the IS insertional mutation frequencies are greatly enhanced in soft agar. The optimal agar concentration for promotion of IS insertional activation of *flhDC* proved to be 0.24% with decreasing mutational frequencies at higher and lower concentrations. In a carefully controlled study, we found that new IS*5* insertional mutants arise as a wild type culture simultaneously inoculated with an IS*5* insertional mutant, approached stationary phase. Moreover, using an assay that does not depend on motility, we showed that mutations that eliminate flagellar structure or motor function prevent IS insertion upstream of the *flhDC* operon, in agreement with the results of Wang & Wood (2011). These findings argue against the possibility that the increased frequency of IS-activated colony expansion mutations is an artifact due to an increase in growth and/or motility.

MATERIALS AND METHODS

Construction of bacterial strains

Strains and DNA oligonucleotides used in this study are described in Tables S1 and S2, respectively. All the strains were derived from *E. coli* K-12 strain BW25113 (45). To make the *flhDC* reporter strain, we first fused a promoter-less *cat* (encoding chloramphenicol acetyl transferase) structural gene (plus its own ribosomal binding site) to the 3' end of the *flhC* gene.

146 The resultant operon, *flhDC:cat*, is under the control of the *flhDC* promoter. To do this, the
147 region containing the *cat* gene and the downstream FRT-flanking kanamycin resistance gene
148 (*km*^{r}) in pKD-*cat* (Zhang and Saier, 2009a) was amplified using oligos, FlhC.cat-P1 and
149 FlhC.cat-P2 (Table S2). FlhC.cat-P1 is composed of a 20 bp region at the 3' end that is
150 complementary to the beginning of *cat*, and a 50 bp region at the 5' end that is homologous to
151 the 3' end of *flhC*. FlhC.cat-P2 is composed of a 20 bp region at the 3' end that is
152 complementary to the FRT-flanking *km*^{r} sequence, and a 50 bp region at the 5' end that is
153 homologous to the *flhC/motA* intergenic region. The PCR products (that is "*cat:km*^{r}") were gel
154 purified and electroporated into BW25113 cells expressing the lamada Red proteins encoded by
155 plasmid pKD46 (45). The cells were applied onto LB + Km plates, and Km^r colonies were
156 confirmed for the replacement of the *flhC/motA* intergenic region (-326 to -1 relative to the start
157 codon of *motA*) by the '*cat:km*^{r}' fragment by colony PCR, followed by DNA sequencing. The
158 FRT-flanking *km*^{r} gene was removed by first transforming the temperature-sensitive pCP20
159 plasmid and then growing a transformant at 40 °C overnight. The resultant strain was named
160 CAT_ΔAB (swarming negative).

161 As shown in Figure 1A, the promoter region of the *motAB* operon is located within the 3'
162 end of *flhC*. The presence of *cat* in the *flhC/motA* intergenic region appears to abolish expression
163 of the downstream *motAB* operon since its promoter is destroyed. This strain CAT_ΔAB would
164 be expected to be non-motile. To restore motility, the *motAB* promoter region was added back at
165 its proper location in front of *motA*. To do this, a DNA fragment (P*motAB*), containing the 3' end
166 of *flhC* and the *flhC/motA* intergenic region, was cloned into the *Xho*I/*BamH*I sites of pKDT,
167 yielding pKDT-P*motAB*. Present in this recombinant plasmid, the region (*km*^{r}:*rrnBT*:P*motAB*),
168 carrying *km*^{r}, the *rrnB* terminator (*rrnBT*) and P*motAB*, was amplified using the primers

PmotAB-P1 and PmotAB-P2 (Table S2). The amplified fragment, *km*^{r}:*rrnBT*:P*motAB*, was electroporated into CAT_ΔAB cells to replace the region between the 3' end of *cat* and the 5' end of *motA*. The transformants were selected for Km resistance. The Km^r colonies were confirmed for such replacement by colony PCR and subsequently by DNA sequencing. The *km*^{r} gene was removed as above. This yielded the reporter strain, CAT, in which *flhD*, *flhC* and *cat* form a single operon that is under the control of the *flhDC* promoter, with *motAB* expression under the control of its own promoter (Figure 1B). To create a second nonmotile strain, the *fliC* mutation in strain JW 1908 (CGSC #9586) was transferred into CAT by P_1 transduction, yielding strain CAT_Δ*fliC*.

In some cases, we needed to distinguish IS insertional mutants that arose after plating from the insertional mutants initially added. For this purpose, to facilitate isolation of IS insertional mutants, we inserted a *km*^{r} marker into the *ycaC/ycaD* intergenic region. The *km*^{r} gene was amplified from pKD4 using primers ycaD-P1 and ycaD-P2 (Table S2). The PCR products were purified and inserted into the *ycaC/ycaD* intergenic region of CAT. The resultant strain that is kanamycin resistant was named CAT_Km^R.

As an alternative, we made a second *flhDC* reporter strain in which the *flhD/flhC/cat* operon, under the control of the native *flhDC* promoter (P*flhDC-flhDC:cat*), was moved to another chromosomal location. To do this, the region containing the native *flhDC* promoter, *flhD* and the 5' end of *flhC* (but not the 3' end of *flhC* containing the *motAB* promoter region) was deleted. *flhDC:cat* was amplified from the genomic DNA of the first reporter strain CAT using primers PflhDC-Xho-L and Cat-Sal-R (Table S2), digested with *Xho*I and *Sal*I, then ligated into the same sites of pKDT, yielding pKDT-P*flhDC-flhDC:cat.* Using this plasmid as template, the region carrying *km*^{r}*:rrnBT*:P*flhDC-flhDC:cat* was amplified using primers IntC-P1 and IntC-P2.

The PCR products were purified and then electroporated into BW25113 cells. The integration of "*km*^{r}*:rrnB*T:P*flhDC-flhDC:cat*" into the *intS/yfdG* intergenic region was confirmed by colony PCR and DNA sequencing. The resultant strain was named CAT2, in which the native *flhDC* promoter, *flhD* and the 5' part of *flhC* were deleted, and a new copy of the same promoter driving the *flhDC:cat* operon was located in the *intS/yfdG* intergenic region (Figure 1C).

Swarming (colony expansion) rate measurements

Two wild type strains and 13 IS insertional mutant strains were each grown in 3 ml of liquid medium (0.5% tryptone, 0.5% NaCl) at 30°C overnight and diluted with an M9 salts solution to an OD of 1.0. Then 1.5 µl of each culture was used to inoculate the center of a soft agar plate (0.5% tryptone, 0.5% NaCl, 0.3% agar). The plates were incubated at room temperature, and the diameters of the swarming zones were measured from 8 to 17 hours after inoculation. Graphs of the diameters were plotted as a function of time, and the colony expansion (swarming) rates were calculated from the slopes.

The experiment was repeated with glucose (0.5%) present in the medium. Before inoculating the soft-agar plates, the strains were grown in liquid media (0.5% tryptone, 0.5% NaCl, 0.5% glucose). The soft-agar plates used were made with 0.5% tryptone, 0.5% NaCl, 0.5% glucose, and 0.3% agar. These plates were treated in the same way as for the experiments without glucose. The slopes were used to determine the rates of colony expansion for each strain, with and without glucose.

Growth rates

Growth rates of the *E. coli* BW25113 *flhDC:cat:'flhC:motAB* (CAT) strain as well as CAT_IS*1* and CAT_IS*5* were measured in liquid media (0.5% tryptone, 0.5% NaCl, 0% agar), on solid-agar plates (0.5% tryptone, 0.5% NaCl, 1.5% agar) and on soft-agar plates (0.5% tryptone, 0.5% NaCl, 0.3% agar). Each strain was grown in 3 ml of liquid medium (0.5% tryptone, 0.5% NaCl) at 30°C overnight, and the cultures were used for growth rate assays.

For liquid media, 5-10 µl of the overnight culture was added to 5 ml of medium. The tubes were incubated in a 30 °C shaking water bath for up to 24 hours. Every three hours, 100 µl of the cultures were removed, and the OD_{600} was measured to plot the growth curves.

For solid agar plate assays, 4 µl of the overnight culture (OD of 1.0) was inoculated in a single streak across the plate and incubated in a 30 °C incubator. At every sample point, the solid agar plates were washed with 3 ml of 1x M9 buffer, and the cell slurry was collected in a microcentrifuge tube. Then the cell culture was diluted with the same buffer to spread on LB plates. These plates were incubated at 37 °C before the colonies were counted.

The soft-agar plates were inoculated in the same way as the solid-agar plates. Every three hours, the gel was scraped off into a large test tube, each plate was rinsed with 3 ml of M9 buffer, and the wash buffer was added to the same tube. The gel was vortexed for 2 minutes to mix the preparation, and the resultant suspension was centrifuged at 800 rpm for 30 sec to separate the gel from the buffer. Such a centrifugation procedure was shown to remove only a small fraction (<10%) of the bacterial cells. The cells were collected in a microcentrifuge tube and diluted to plate onto LB plates. The plates were incubated at 37 °C to measure the cell populations.

Using soft-agar plates, the growth rates of CAT_Km^R (wild type) and CAT_ IS*5* were measured when these two types of cells were mixed before inoculation. The overnight cultures

were diluted with M9 buffer, and CAT_Km^R (2 x 10^9 cfu/ml) and CAT_ IS*5* (2 x 10^6 cfu/ml) were mixed in equal volumes. Then, the mixture (4 µl) was inoculated in a single streak across the center of the soft-agar plates and incubated at 30 °C. Every three hours, the cell samples were prepared as described above for the single strain soft agar assay. They were appropriately diluted before being applied unto LB plates, LB + Cm plates (LB with 6 µg/ml chloramphenicol), and LB + Cm + Km plates (LB with 6 µg/ml chloramphenicol and 25 µg/ml kanamycin). These plates were incubated in a 37 °C incubator before counting colonies. The numbers of colonies on LB plates gave the total population; that on LB + Cm plates gave the total population of the inoculated insertion mutant, and that on LB + Cm + Km gave the population of the newly emerging mutants. The growth rate of the IS*5* insertional mutant that was added at the beginning was calculated by subtracting the newly emerging population from the total insertion mutant population.

Mutant ratios

Mutant ratios were measured for three media, the same as for the growth rate experiments. The CAT strain was grown in 3 ml of liquid medium in a 30 °C water shaker overnight and diluted with M9 buffer to an OD of 1.0.

For liquid media, 5-10 µl of the overnight culture was inoculated into 5 ml of medium and incubated in a 30 °C shaker. At four different time points, the culture was plated onto LB plates (diluted with M9 buffer) and LB + Cm plates (undiluted). The mutant ratio was calculated by dividing the mutant population by the wild type population.

For solid-agar plates, 4 µl of an overnight wild type culture (OD of 1.0) was plated in a single streak across the plate and incubated in a 30 °C incubator. At four different time points,

each plate was washed with 3 ml of M9 buffer and plated onto LB plates (diluted with 1x M9 buffer) and LB + Cm plates and incubated in a 37 °C incubator. The colonies were counted, and the ratios of the total population of wild type to mutant cells were determined.

Soft-agar plates were inoculated and incubated in the same way as solid agar plates. Every three hours, the agar was collected in a test tube, and the plates were washed with 3 ml of M9 buffer. The collected agar was mixed by vortexing and centrifuged as described above, and the cultures were diluted to plate onto LB and LB + Cm plates.

Mutant ratios were determined for another wild type strain, CAT2, using these three media at two temperatures, 30 °C and 37 °C. The strain was grown overnight in 3 ml of liquid medium, diluted with M9 buffer to an OD of 1, and used to inoculate fresh media and incubate as described above. For this strain, a higher concentration of chloramphenicol (LB with 14 µg/ml chloramphenicol) was used for mutant isolation.

Mutant ratios were determined for another two non-motile strains, CAT_ Δ AB and CAT_ Δ fliC, using soft agar plates [0.4% nutrient broth (NB) + 0.3% agar]. The strains were cultured in 0.4% NB overnight, diluted to OD600 of 1.0, and 1.4 μ l of the diluted cells were streaked onto agar plates. The plates were incubated at 30 °C. After 18 h, the total and mutant populations were determined as described above.

To determine if agar concentrations affected the appearance of swarming mutants, the cells of strain CAT were inoculated onto NB agar plates that contained 0.4% NB and 0% to 1.2% agar. After the plates were incubated at 30 °C for 18 h, the total and mutant populations were determined. For the plates with 0% to 0.4% agar, the populations were determined as described for soft agar plate assays as above. For the plates with 0.6% agar and above, the cells were

282 washed off with 3 ml M9 buffer without breaking the agar. The total populations and mutant
283 populations were determined using LB and LB + Cm plates as above. Three plates were used for
284 each concentration of agar.
285
286 **Original locations of IS elements in the chromosome**
287 NCBI-BLAST was used to identify the locations and directions of original IS elements
288 on the chromosome of *E. coli* K-12 strain BW25113. Then, the relative locations were plotted on
289 the circular diagram of the chromosome. In Figure S1, the blue arrows indicate the IS*1*s; purple,
290 the IS*3*s, and green, the IS*5*s. The arrowheads indicate the direction of transcription of the
291 transposase gene.
292
293
294
295 **Determining DNA destabilization energy plots for the *flhDC* upstream region**
296 Destabilization energy G(x) is the incremental energy (kcal/mol) needed to guarantee
297 separation of base pair x under defined superhelical stress. Highly destabilized DNA regions
298 (referred to as SIDD for superhelicity-induced DNA duplex destabilization) have lower values of
299 G(x), and more stable regions have higher G(x) values (46). Destabilization energy calculations
300 were made by using the SIST programs provided by Prof. C. J. Benham (47) using default
301 settings to be described elsewhere [Humayun, Zhang & Saier, unpublished] for a 4.4 kb DNA
302 segment containing *flhD* flanked by 2000 bp upstream of the start codon for *flhD* and 2000 bp
303 downstream of the stop codon for the *flhD* gene (base pairs 1,970,104 to 1,974,454 from the *E.*

304 *coli* BW25513 genomic sequence, NCBI Reference NZ_CP009273.1). For clarity, Figure 5
305 shows the plot for a 2.4 kb portion of the sequence.
306
307 **RESULTS**
308 **Original locations of IS elements on the chromosome**
309 IS elements are present in most *E. coli* K12 chromosomes, but the numbers of copies and
310 their locations differ from strain to strain. In this study, *E. coli* K-12 BW25113 was used as the
311 parental wild type, and the numbers and locations of three IS elements, IS*1*, IS*3*, and IS*5*, were
312 determined using NCBI-BLAST. There were six copies of IS*1* in the chromosome, five copies of
313 IS*3*, and 10 copies of IS*5* (Figure S1). Some of them were in the direct orientation while others
314 were in the reverse orientation as shown. One IS*1* and three IS*5*s were located near the target
315 *flhDC* promoter, but which IS elements serve as the sources for insertion upstream of the *flhDC*
316 promoter are unknown.
317
318 **Key strains used in this study**
319 BW25113 is the parental strain used to make other strains. This strain has no IS element
320 present in the *flhDC* regulatory region, and is therefore a poor swarmer. Strain CAT carries a
321 promotor-less *cat* gene that is fused immediately downstream of *flhC* so that *flhD*, *flhC*, and *cat*
322 form a single operon under the control of the *flhDC* promoter. Strain CAT2 is the same strain as
323 CAT, except that the *flhDC* regulatory region and its downstream *flhD*:*flhC*:*cat* operon were
324 moved to the *intS/yfdG* intergenic region while the original promoter, *flhD* and the first half of
325 *flhC* are deleted. As alternative parental strains, CAT and CAT2 are sensitive to Cm at 3 μg/ml
326 while all IS mutants are resistant to Cm at 6.5 μg/ml.

Strain CAT_Km^r is the same as CAT except that it carries a Km resistance gene in the chromosome. This strain is resistant to Km at 25 μg/ml. All IS insertional mutants derived from this strain are resistant to both Cm and Km. Strains CAT_IS1 and CAT_IS5 are two insertional mutants that contain IS1 (at -107) and IS5 (at -99), respectively, in the *flhDC* regulatory region. Both strains are resistant to Cm at 6.5 μg/ml. Strains CAT_ΔAB is same as CAT except that the *motAB* operon is not expressed due to the lack of its promoter. CAT_ΔfliC is the same as CAT, except that the *fliC* gene is deleted. Both CAT_ΔAB and CAT_Δ*fliC* are non motile and sensitive to Cm at 3 μg/ml.

Swarm colony expansion

When an IS element was inserted upstream of the *flhDC* promoter, operon expression was increased, causing the cells to display increased motility (44). Including two parental "wild type" *E. coli* K-12 strains (BW25113 and CAT), 15 strains were tested for colony expansion (swarming) rates when glucose was absent from the media (Figure 2A). All the IS mutants migrated at higher rates than the wild type cells, while both types of wild type colonies migrated at similar rates (1 to 1.14 mm/h). Thus, the slowest mutant showed a 1.9 fold increase while the fastest mutant showed a 2.7-fold increase over wild type cells.

The colony expansion rates for these strains were measured again with glucose (0.5%) present in the media. In the presence of this nutrient, all of the mutant strains migrated at 40% to 60% faster rates relative to the rates in the absence of glucose, whereas the wild type strain virtually lost its swarming ability in the presence of glucose (Figure 2B and Table 1).

Growth rates for wild type and mutants when separately inoculated

Three strains, CAT (wild type, in which *flhDC* and *cat* form a single operon under the control of the *flhDC* promoter), CAT_IS*1* (an IS*1* insertion mutant), and CAT_IS*5* (an IS*5* insertion mutant) (Table S1) were inoculated into three different types of media. The media differed only by the concentration of agar, which was 0%, 0.3%, and 1.5% for liquid media, soft agar plates, and solid agar plates, respectively. For liquid media, growth rates were determined by measuring the optical densities (OD_{600}) and plotting OD_{600} values in a semi-log format. For 0.3% and 1.5% agar media, growth rates were determined by measuring absolute viable cell numbers and plotting the data in linear formats. The growth curves for these three strains overlapped in all three media up to 24 hours, and no significant difference was identified (Figures 3A, B and C). It should be noted, however, that these experiments would not allow detection of small growth rate differences (see below). For example, if the growth data shown in Figure 3C are plotted in semi-log format, CAT_IS*1* and CAT_IS*5* had slightly greater populations than wild type cells after 15 hours of growth.

Growth rates for wild type and mutant strains when inoculated together

When the mixture of CAT_Km^R (wild type) and CAT_ IS*5* cells grew in a soft agar medium (0.5% tryptone, 0.5% NaCl, 0.3% agar) together, the IS*5* strain showed a slight increase in growth rate where the doubling time of wild type was 62 min and that of the IS*5* mutant was ~49 min (Table 2). This difference was nearly the same when the tryptone concentration in the medium was increased 5-fold. The wild type had a doubling time of 57 min, and the mutant doubled in 47 min (Table 2). The "wild type" strain (CAT_Km^R) is resistant only to Km. Cells only become resistant to Cm when the *flhDC* operon is activated by IS insertion. This allowed us to distinguish between the originally inoculated IS5 mutant, CAT_ IS*5*, resistant only to

chloramphenicol (Cm^R Km^S), and the newly derived mutants from CAT_Km^R, which were resistant to both chloramphenicol and kanamycin (Cm^R Km^R). This gave us more accurate growth curves for the wild type and IS*5* strains. In both 0.5% tryptone and 2.5% tryptone soft-agar plates, all strains reached stationary phase around 15 hours after inoculation (Figures 4A and B). For these experiments, growth rates were determined by measuring absolute viable cell numbers over time and plotting the data in linear format.

To determine if the newly derived mutants were IS insertion mutants, 30 Cm^RKm^R mutants isolated from 0.5% tryptone soft agar plates (incubated for 15 h) were purified, and their *flhDC* regulatory regions were amplified by PCR. All mutants were found to carry an IS element. DNA sequencing showed that these mutants remained as the IS5, IS1, or IS3 insertional mutants as present originally (data not shown).

Insertion Sites

IS insertion sites were identified by PCR-amplification of the intergenic region between *flhD* and the upstream gene, *yecG* (*uspC*), and subsequent DNA sequencing of the upstream region of the *flhDC* promoter. IS*1* was found at six different locations, five in the reverse orientation between nucleotides -469 and -470, -214 and -215, -180 and -181, -120 and -121 and -107 and -108, and one in the direct orientation between -415 and -416 (Figure 5A). A single IS*3* insertion site was identified between -199 and -200 in the reverse orientation. IS*5* was found to insert at three sites, between -318 and -319, -169 and -170, and -99 and -100. At all three of these insertion sites, IS*5* was inserted in both orientations. Thus, while IS*1* was never observed in both orientations at a particular site, IS*5* was found in both orientations at all three sites. All insertion sites were upstream of the promoter, with none found downstream of the transcriptional start site

of the *flhDC* promoter. Further analyses showed that all identified IS insertion sites are located between, not within, known operator sites in the *flhDC* upstream control region.

We observed that all insertions were in a 370 bp region (-100 to -470) that constitutes a strong SIDD (superhelicity-induced DNA duplex destabilization) region (Figure 5B). SIDD sequences are susceptible to strand separation under negative superhelical stress (48, 49) (50). As considered in greater detail elsewhere, IS elements preferentially insert at SIDD sequences (M. Z. Humayun, Z. Zhang and M.H. Saier, unpublished results).

Mutant ratios obtained in liquid media and on solid-agar plates

CAT was used as the wild type strain in this experiment. A chloramphenicol resistance gene was fused at the end of the *flhC* gene so that it was under the control of the *flhDC* promoter and was up-regulated whenever the *flhDC* promoter was activated by an IS insertion upstream of *flhDC*. This allowed us to measure the populations of wild type and IS mutant strains on LB plates and LB + Cm plates, respectively.

On solid-agar plates (0.5% tryptone, 0.5% NaCl, 1.5% agar), the mutant ratio was low (<5 mutants/10^8 cells 25 hours after plating) (Figure 3A). In liquid media (0.5% tryptone, 0.5% NaCl, 0% agar), the insertion frequencies also remained low (<5 mutants/10^8 cells) (Figure 5B). CAT, CAT_IS*1*, and CAT_IS*5* had the same growth rates in liquid media within experimental error when these strains were cultured separately (Figure 3B).

Both in liquid media and on solid-agar plates, up-regulating *flhDC* expression is not expected to benefit cells. On the contrary, the biosynthesis and function of the flagellar system would be expected to be detrimental because of the high energy consumption resulting from flagellar biosynthesis and function. Wang and Wood (2011), in fact, reported slower growth

rates for an IS*5* insertional mutant compared to its wild type strain when cells were grown in liquid LB medium, presumably reflecting the increased energetic burden in mutant cells. However, under these experimental conditions, there was little difference in growth rates of wild type cells and insertion mutants in liquid media and on solid agar (but see next section).

IS insertion frequencies obtained in soft agar

Soft agar plates were tested with two different concentrations of tryptone in order to evaluate the effect of nutrient deficiency on mutation frequency. While wild type (Cm^S Km^R) and mutant (Cm^R Km^S) cells entered the stationary growth phase about 15 hours after inoculation, the emerging insertion mutants (Cm^R Km^R) began to appear after about 5 hours and continued to increase exponentially when both inoculated cell types (wild type and IS*5* insertional mutant cell) were barely growing (Figures 6A, B).

Table 2 compares the doubling times for the CAT_Km^R (wild type), CAT_IS5, and emerging insertion mutants in 0.5% and 2.5% tryptone soft-agar plates. For 0.5% tryptone soft-agar plates, between 1 and 15 hours these three populations showed *apparent* doubling times of 62 min, 49 min, and 65 min, respectively. Thus, the IS mutants had a faster doubling time than the wild type. Between 15 and 25 hours, growth of both the initially inoculated wild type and IS mutant cells approached the stationary phase when the doubling times were roughly 288 min and 866 min, respectively (Figure 4A and Table 2). The new mutant population, however, kept increasing as it maintained its *apparent* doubling time of 60 min, presumably reflecting the appearance of new insertional mutants and not rapid growth of this population. The average rate of appearance of these mutants, calculated from the slope of the curve, was about $3x10^{-3}$/cell/day.

This trend was the same in 2.5% tryptone soft-agar media (Figure 4B and Table 2). For the first 15 hours after inoculation, the doubling times for CAT_Km^R, CAT_ IS*5*, and the emerging insertion mutants were 57 min, 47 min, and 60 min, respectively. When both the wild type and IS*5* strains initially plated approached the stationary phase, the emerging mutant population kept increasing with an *apparent* doubling time of 56 min. It should be noted that the *apparent* doubling times for the newly appearing mutants represent a sum of the insertion rate and the doubling rate for these cells alone. Since both the wild type and initially inoculated IS*5* mutant were barely growing, we assume the increased numbers of newly emerging cells is primarily due to the appearance of new mutants, not to growth. This experiment, with a 5-fold increase in all nutrients, suggests that nutrient limitation was not the cause of growth cessation observed for the parent and co-inoculated IS5 insertion mutant.

The frequencies of mutant appearance were also determined with another wild type strain, CAT2, in which the *flhDC*:*cat* operon plus its upstream regulatory region was moved to another (ectopic) chromosomal location. The experiments were conducted at both 30 °C and 37 °C, and the trends were the same for these two conditions (Figures 6A, B). The mutant population remained low for solid-agar and liquid media during the entire 25 hour period. By contrast, emerging mutant cells kept increasing exponentially throughout this time period in soft-agar, even after the parental wild type cells had reached the stationary phase.

IS insertion frequencies of non-motile cells in soft agar

To see if IS insertional mutations occur in non-motile cells, we measured the IS insertion frequencies using strains CAT_ΔAB and CAT-ΔfliC that lack motility. As seen in Figure 7, almost no IS mutants arose in these non-motile strains. As expected, insertional

mutations arose at a high frequency from wild type cells under the same soft agar conditions. These observations are consistent with the results of Wang and Wood (2011), who deleted the *flhD, motA or flgK* gene and did not observe the appearance of IS5 insertional mutations.

Effects of agar concentration on the appearance of IS insertional mutants

The pore size of agar is inversely proportional to the agar concentration. To determine how the agar concentration affects the rates of IS insertion, we measured the frequencies of IS insertional mutations by incubating strain CAT in media containing various concentrations of agar, ranging from 0 to 1.2%. As shown in Figure 8, 0.2% agar gave rise to IS insertion mutants with the highest frequency. When the agar concentration was 0.1%, the mutation frequency was far less than when it was at 0.2%. When agar was absent, almost no mutants arose. Similarly, the mutation frequencies decreased greatly as agar concentrations increased from 0.3% to 0.5%. When the agar concentration was 0.6% or above, the mutational frequencies were negligible. The pore diameters of agar are roughly 1.6 μm, 0.7 μm, and 0.2 μm in 0.25% agar, 0.5% agar, and 1.0% agar, respectively (51; Zhang & Saier, unpublished data). When the pore diameters are smaller than 1 μm, cells can hardly migrate inside the agar. Our data with different concentrations of agar provide the first quantitative evidence showing the effect of agar concentration (and thus, pore size) on mutation rate. It should be noted, however, that increased agar concentration also results in an increase of viscosity that could play a role in determining the rate of

mutation. We propose that inhibition of flagellar rotation is the signal promoting IS insertion.

DISCUSSION

Activation of the *flhDC* operon by insertional mutations

Barker, Pruss (20) were the first to report a connection between IS element insertion upstream of the *flhDC* promoter and increased motility of *E.coli* K12, an observation that has been confirmed and extended by several investigators (17, 27, 28). Many other types of mutations can also activate *flhDC* expression as expected since expression is regulated by over two dozen transcription factors (see the Introduction). Though the mechanism(s) of activation of the operon downstream of the insertion sites remain(s) unknown, transposon insertion increased motility. Three IS elements (IS*1*, IS*3* and IS*5*) hop into the regulatory region of the operon, increasing expression of the downstream genes (20, 27, 52). Increased flagellar gene expression following IS insertion allows *E. coli* K12 to migrate more rapidly through soft-agar (0.2 - 0.3%) but not through solid-agar (1.5%), and flagella are not useful in agitated liquid media where the medium is of a uniform consistency. Wang and Wood (44) reported that IS5 insertional mutations upstream of the *flhDC* operon arose at high rates only in soft agar, precisely the condition where the consequences of the mutations were beneficial.

In this study, we confirm and extend the provocative results of Wang and Wood (44), showing that the frequency of IS element insertional events in the *flhDC* promoter increases during exposure to soft agar, but not solid agar or liquid medium. We showed that the concentration that optimally promotes IS insertion is about 0.24%, a concentration correlating with a pore size in the agar equivalent to that just barely allowing entry of the cells into the agar

channels, presumably inhibiting motility by reducing the flagellar rotation rate. Further, we showed that this increase cannot simply be an artifact of a post-insertional growth advantage. We constructed two reporter systems, based on transcriptional fusion of a *cat* gene downstream of the *flhC* gene, such that IS-activation of the *flhDC* operon leads to chloramphenicol resistance. This system provided a way to track both the emergence of novel mutants, and expansion of the newly emerging mutant population in the presence of the inoculated wild type and an innoculated IS-*flhDC*-activated mutant. While the assay system does not allow for discrimination between mutation and selection, the collection of experiments performed does yield compelling evidence that the growth and motility advantages are not sufficient to account for the accumulation of the newly arising more motile/chloramphenicol-resistant mutants.

The role of environment on insertional activation of the *flhDC* operon

The finding that IS insertions only occurred after growth in soft agar raises the possibility that an environmental trigger activated (or facilitated) insertion into *flhDC*-activating sites. Presumably, cells in the wild with more flagella swarm faster through soft-agar and are capable of reaching nutrients ahead of cells with fewer flagella. Both Wang and Wood (44) and we noted that the increase in the insertion-mutant population could not be accounted for simply by faster growth. Wang and Wood (44) reported slower growth rates for an insertion mutant when the cells were grown in liquid LB medium, but they did not measure these rates in soft agar. The results from our mixing experiment provide the most conclusive demonstration that growth rates cannot explain the apparent rapid increase in the emerging insertion-mutant population on soft-agar. Our results showed that the growth rate in soft agar differed by ~20% when cells were in the log phase of growth, whereas the newly emerging mutant population increased dramatically

only after most cells (both wild type and the co-inoculated IS5 insertional mutant tested) had reached stationary phase. Between 15 and 25 hours, the ratio of the wild type and the originally inoculated IS insertion strains did not differ appreciably, but the numbers of emerging IS mutants increased dramatically when the cells were approaching the stationary growth phase (Figures 4A and B). This result suggests that the rates of insertional mutant appearance are upregulated when the inoculated cells are entering the stationary phase in soft-agar. This trend was the same for the other parental wild type strain examined (CAT2, data not shown).

Our observations, together with those of Wang and Wood (44) indicate that higher expression of the flagellar genes may confer an advantage to cells in soft-agar plates, but not in liquid media. The advantage is presumed to be the ability to reach a new nutrient-rich niche, or the ability to escape growth-inhibitory metabolic waste products, both of which are only feasible in soft agar. This appears to be an example of a mechanism whereby gene expression regulation provides adaptive evolutionary benefit in overcoming stress (53).

Effect of environment on insertional mutagenesis in other experimental systems

At present, molecular mechanisms by which the environment is sensed to enhance insertion frequency upstream of the *flhDC* operon are unknown. However, this behavior is reminiscent of, and could be mechanistically related to, the swarming behavior of *Vibrio parahaemolyticus* which induces a second, lateral, swarming flagellar system in response to rotational inhibition of the first polar one (54-56). In the case of *E. coli* IS-mediated *flhDC* activation studied here, weak interference with flagellar rotation could provide the signal for IS insertion. One possibility is that flagellar rotation senses the viscosity of the medium.

In a different respect, this work is also reminiscent of the work of Vandecraen, Monsieurs (57) who reported that the presence of toxic levels of zinc in the growth medium induced transposition and insertion of several IS elements in the bacterium *Cupriavidus metallidurans*, increasing its adaptation to growth in the presence of zinc. This situation resembles the situation described here in that multiple IS sequences insert at multiple sites to activate gene expression.

In its responsiveness to environmental conditions, the *flhDC* insertional mutagenesis system is also reminiscent of our previous extensive work on IS*5* insertional activation of the glycerol-utilization *glpFK* operon (21-25). The *glpFK* system, which is mechanistically much better described than any other insertional mutagenesis system, is based on the inability of Crp^- or $CyaA^-$ cells to utilize glycerol as a carbon source (i.e., Crp^- and $CyaA^-$cells have a Glp^- phenotype). Prolonged incubation of these cells on minimal agar plates with glycerol as the sole carbon source, however, leads to an accumulation of promoter-activating IS*5* insertional mutations upstream of the *glpFK* promoter that render the cells Glp^+. The insertional mutagenesis appears to be specific to the *glpFK* region under these conditions, as no other examined sites showed elevated insertional mutagenesis (24). We have recently proposed a plausible mechanism that connects glycerol-starvation stress to localized mutagenesis of the *glpFK* promoter region, mediated by two DNA-binding proteins (25).

Effect of IS insertion events on *flhDC* expression

IS insertion into the *flhDC* operon upstream region of *E. coli* always leads to faster migration rates relative to parental cells (Figure 2A, B). The mechanisms by which transposition increased expression of the downstream *flhDC* operon in *E. coli* is however, not yet clear [but see (17, 23, 28, 44)]. Activation may be accomplished by one of the following four mechanisms;

1) the IS element could directly activate an existing chromosomal promoter, 2) it could use a promoter already on the transposon to activate expression, 3) it could create a new hybrid-promoter (e.g., the -35 region on the transposon with the -10 region on the chromosome), and 4) it could inactivate a repressor or inhibitory region of the upstream DNA while still using the native *flhDC* promoter. All of these mechanisms have been documented in previous studies with various operons (23, 57-59). We favor the last of these four mechanisms for *flhDC* activation in agreement with previously published work (17, 28). The fact that activation of the *flhDC* promoter occurs to a similar degree regardless of IS element direction (orientation) is in agreement with this conclusion. In this regard, it would be interesting to investigate if the presence of these IS elements in the *flhDC* upstream region affects the binding and action of any of the many transcription factors that influence *flhDC* expression.

No apparent correlation between the insertion sites and swarming rates was seen for IS5, but a possible correlation could be seen for IS*1* (Figure 9). Although the detailed mechanisms by which IS elements influence motility, are unknown, insertions at sites immediately upstream of the *flhDC* promoter yielded strains with higher expression based on increased motility in soft-agar (see also 28).

Glucose effect on motility

The IS insertional mutants showed greater rates of colony expansion (swarming) in the presence of glucose than in the absence of glucose, although glucose strongly inhibited motility of the wild type strains, presumably by repressing flagellar synthesis (Table 1). The native *flhDC* operon is positively activated by the cyclic AMP receptor protein (Crp) which forms a complex with cAMP to bind to a specific site in the promoter (60). Wild type parental *E. coli* K12 cells

lose their motility when glucose is present and cAMP levels are low (61), presumably because *flhDC* expression is dependent on the Crp-cAMP complex (60, 62). By contrast, the increased rates of colony expansion of the mutants in the presence of glucose is paradoxical because the mutants also proved to be Crp-cAMP-dependent. Whether a lower cAMP concentration is sufficient for motility in the mutants, or whether the glucose repressive effect observed in wild type cells is partially or fully independent of Crp, remains to be investigated. The increased swarming of the mutants in the presence of glucose might be due to an increased cellular energy level driving motility rather than an effect on gene expression. This suggestion will be examined in future experiments.

Role of DNA structural features in IS insertion and gene activation

Different transposable genetic elements display different target-site preferences. Even though there is a preference for short sequences (3-10 bp), the preference is not absolute for many IS elements, and genomic IS element distributions have been interpreted to suggest that transposition is more-or-less random, although with striking preferences for intergenic regions (63). The question of what drives insertions to specific targets in experimental systems such as those discussed here is largely unexplored. Interestingly, the 370 bp region into which IS elements are inserted in the *flhDC* upstream region has an embedded SIDD element (for superhelicity-induced DNA duplex destabilization), and every inserted IS element occurs within this region (Figure 5B). Examination of other hot spots of insertion elsewhere on the *E. coli* chromosome have revealed a strong preference for SIDD-overlapping regions. This last mentioned work will be reported in a subsequent communication (M. Z. Humayun, Z. Zhang and M.H. Saier, manuscript in preparation).

In future studies, there are many additional aspects of IS-mediated *flhDC* mutational activation that need to be examined. For example, how does each IS element activate *flhDC* transcription? Which IS elements are transposed to which locations and why? How do the environmental cues trigger transposition to the upstream sites? These exciting questions will be of interest to many microbial physiologists.

Funding

This work was supported by the National Institutes of Health [GM077402 and GM109895].

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Professor James Golden for valuable discussions, Professor Thomas Wood and Dr. Xiaoxue Wang for critically reading the manuscript prior to submission for publication and Anne Chu, Harry Zhou and Sabrina Phan for assistance with the preparation of this manuscript.

References

1. Munoz-Lopez M, Vilar-Astasio R, Tristan-Ramos P, Lopez-Ruiz C, Garcia-Perez JL. Study of Transposable Elements and Their Genomic Impact. Methods Mol Biol. 2016;1400:1-19.
2. Blot M. Transposable elements and adaptation of host bacteria. Genetica. 1994;93(1-3):5-12.
3. Bonchev G, Parisod C. Transposable elements and microevolutionary changes in natural populations. Mol Ecol Resour. 2013;13(5):765-75.
4. Zhang Z, Saier MH, Jr. Transposon-mediated adaptive and directed mutations and their potential evolutionary benefits. J Mol Microbiol Biotechnol. 2011;21(1-2):59-70.
5. Derbyshire KM, Grindley ND. Replicative and conservative transposition in bacteria. Cell. 1986;47(3):325-7.
6. Haniford DB, Ellis MJ. Transposons Tn10 and Tn5. Microbiol Spectr. 2015;3(1):MDNA3-0002-2014.
7. Skipper KA, Andersen PR, Sharma N, Mikkelsen JG. DNA transposon-based gene vehicles - scenes from an evolutionary drive. J Biomed Sci. 2013;20:92.
8. Chandler M, Fayet O, Rousseau P, Ton Hoang B, Duval-Valentin G. Copy-out-Paste-in Transposition of IS911: A Major Transposition Pathway. Microbiol Spectr. 2015;3(4).
9. Sousa A, Bourgard C, Wahl LM, Gordo I. Rates of transposition in Escherichia coli. Biol Lett. 2013;9(6):20130838.
10. Zhang Z, Yen MR, Saier MH, Jr. Precise excision of IS5 from the intergenic region between the fucPIK and the fucAO operons and mutational control of fucPIK operon expression in Escherichia coli. J Bacteriol. 2010;192(7):2013-9.
11. Ziebuhr W, Krimmer V, Rachid S, Lossner I, Gotz F, Hacker J. A novel mechanism of phase variation of virulence in Staphylococcus epidermidis: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS256. Mol Microbiol. 1999;32(2):345-56.
12. Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol Rev. 2014;38(5):865-91.
13. Gaffe J, McKenzie C, Maharjan RP, Coursange E, Ferenci T, Schneider D. Insertion sequence-driven evolution of Escherichia coli in chemostats. J Mol Evol. 2011;72(4):398-412.
14. Bonnefoy V, Fons M, Ratouchniak J, Pascal MC, Chippaux M. Aerobic expression of the nar operon of Escherichia coli in a fnr mutant. Mol Microbiol. 1988;2(3):419-25.
15. Sawers RG. Expression of fnr is constrained by an upstream IS5 insertion in certain Escherichia coli K-12 strains. J Bacteriol. 2005;187(8):2609-17.
16. Li B, Li N, Wang F, Guo L, Huang Y, Liu X, et al. Structural insight of a concentration-dependent mechanism by which YdiV inhibits Escherichia coli flagellum biogenesis and motility. Nucleic Acids Res. 2012;40(21):11073-85.
17. Lee C, Park C. Mutations upregulating the flhDC operon of Escherichia coli K-12. J Microbiol. 2013;51(1):140-4.

682 18. Chen YM, Lu Z, Lin EC. Constitutive activation of the fucAO operon and silencing of
683 the divergently transcribed fucPIK operon by an IS5 element in Escherichia coli mutants
684 selected for growth on L-1,2-propanediol. J Bacteriol. 1989;171(11):6097-105.
685 19. Hall BG. Activation of the bgl operon by adaptive mutation. Mol Biol Evol.
686 1998;15(1):1-5.
687 20. Barker CS, Pruss BM, Matsumura P. Increased motility of Escherichia coli by
688 insertion sequence element integration into the regulatory region of the flhD operon. J
689 Bacteriol. 2004;186(22):7529-37.
690 21. Saier MH, Jr., Zhang Z. Transposon-mediated directed mutation controlled by DNA
691 binding proteins in Escherichia coli. Front Microbiol. 2014;5:390.
692 22. Saier MH, Jr., Zhang Z. Control of Transposon-Mediated Directed Mutation by the
693 Escherichia coli Phosphoenolpyruvate:Sugar Phosphotransferase System. J Mol Microbiol
694 Biotechnol. 2015;25(2-3):226-33.
695 23. Zhang Z, Saier MH, Jr. A novel mechanism of transposon-mediated gene activation.
696 PLoS Genet. 2009;5(10):e1000689.
697 24. Zhang Z, Saier MH, Jr. A mechanism of transposon-mediated directed mutation. Mol
698 Microbiol. 2009;74(1):29-43.
699 25. Zhang Z, Saier MH, Jr. Transposon-mediated activation of the Escherichia coli glpFK
700 operon is inhibited by specific DNA-binding proteins: Implications for stress-induced
701 transposition events. Mutat Res. 2016;793-794:22-31.
702 26. Soutourina OA, Bertin PN. Regulation cascade of flagellar expression in Gram-
703 negative bacteria. FEMS Microbiol Rev. 2003;27(4):505-23.
704 27. Fitzgerald DM, Bonocora RP, Wade JT. Comprehensive mapping of the Escherichia
705 coli flagellar regulatory network. PLoS Genet. 2014;10(10):e1004649.
706 28. Fahrner KA, Berg HC. Mutations That Stimulate flhDC Expression in Escherichia coli
707 K-12. J Bacteriol. 2015;197(19):3087-96.
708 29. Rahimpour M, Montero M, Almagro G, Viale AM, Sevilla A, Canovas M, et al. GlgS,
709 described previously as a glycogen synthesis control protein, negatively regulates motility
710 and biofilm formation in Escherichia coli. Biochem J. 2013;452(3):559-73.
711 30. Vikram A, Jayaprakasha GK, Uckoo RM, Patil BS. Inhibition of Escherichia coli
712 O157:H7 motility and biofilm by beta-sitosterol glucoside. Biochim Biophys Acta.
713 2013;1830(11):5219-28.
714 31. Allison SE, Silphaduang U, Mascarenhas M, Konczy P, Quan Q, Karmali M, et al. Novel
715 repressor of Escherichia coli O157:H7 motility encoded in the putative fimbrial cluster OI-
716 1. J Bacteriol. 2012;194(19):5343-52.
717 32. Kitagawa R, Takaya A, Yamamoto T. Dual regulatory pathways of flagellar gene
718 expression by ClpXP protease in enterohaemorrhagic Escherichia coli. Microbiology.
719 2011;157(Pt 11):3094-103.
720 33. Lehti TA, Bauchart P, Dobrindt U, Korhonen TK, Westerlund-Wikstrom B. The
721 fimbriae activator MatA switches off motility in Escherichia coli by repression of the
722 flagellar master operon flhDC. Microbiology. 2012;158(Pt 6):1444-55.
723 34. Reiss DJ, Mobley HL. Determination of target sequence bound by PapX, repressor of
724 bacterial motility, in flhD promoter using systematic evolution of ligands by exponential

725 enrichment (SELEX) and high throughput sequencing. J Biol Chem. 2011;286(52):44726-
726 38.
727 35. Theodorou MC, Theodorou EC, Kyriakidis DA. Involvement of AtoSC two-component
728 system in Escherichia coli flagellar regulon. Amino Acids. 2012;43(2):833-44.
729 36. Wiebe H, Gurlebeck D, Gross J, Dreck K, Pannen D, Ewers C, et al. YjjQ Represses
730 Transcription of flhDC and Additional Loci in Escherichia coli. J Bacteriol.
731 2015;197(16):2713-20.
732 37. Wada T, Hatamoto Y, Kutsukake K. Functional and expressional analyses of the anti-
733 FlhD4C2 factor gene ydiV in Escherichia coli. Microbiology. 2012;158(Pt 6):1533-42.
734 38. De Lay N, Gottesman S. A complex network of small non-coding RNAs regulate
735 motility in Escherichia coli. Mol Microbiol. 2012;86(3):524-38.
736 39. Thomason MK, Fontaine F, De Lay N, Storz G. A small RNA that regulates motility
737 and biofilm formation in response to changes in nutrient availability in Escherichia coli.
738 Mol Microbiol. 2012;84(1):17-35.
739 40. Chilcott GS, Hughes KT. Coupling of flagellar gene expression to flagellar assembly in
740 Salmonella enterica serovar typhimurium and Escherichia coli. Microbiol Mol Biol Rev.
741 2000;64(4):694-708.
742 41. Guttenplan SB, Kearns DB. Regulation of flagellar motility during biofilm formation.
743 FEMS Microbiol Rev. 2013;37(6):849-71.
744 42. Sule P, Horne SM, Logue CM, Pruss BM. Regulation of cell division, biofilm formation,
745 and virulence by FlhC in Escherichia coli O157:H7 grown on meat. Appl Environ Microbiol.
746 2011;77(11):3653-62.
747 43. Copeland MF, Weibel DB. Bacterial Swarming: A Model System for Studying
748 Dynamic Self-assembly. Soft Matter. 2009;5(6):1174-87.
749 44. Wang X, Wood TK. IS5 inserts upstream of the master motility operon flhDC in a
750 quasi-Lamarckian way. ISME J. 2011;5(9):1517-25.
751 45. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in
752 Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A. 2000;97(12):6640-5.
753 46. Wang H, Noordewier M, Benham CJ. Stress-induced DNA duplex destabilization
754 (SIDD) in the E. coli genome: SIDD sites are closely associated with promoters. Genome
755 research. 2004;14(8):1575-84.
756 47. Zhabinskaya D, Madden S, Benham CJ. SIST: stress-induced structural transitions in
757 superhelical DNA. Bioinformatics (Oxford, England). 2015;31(3):421-2.
758 48. Benham CJ. Duplex destabilization in superhelical DNA is predicted to occur at
759 specific transcriptional regulatory regions. J Mol Biol. 1996;255(3):425-34.
760 49. Bi C, Benham CJ. WebSIDD: server for predicting stress-induced duplex destabilized
761 (SIDD) sites in superhelical DNA. Bioinformatics. 2004;20(9):1477-9.
762 50. Wang H, Benham CJ. Superhelical destabilization in regulatory regions of stress
763 response genes. PLoS Comput Biol. 2008;4(1):e17.
764 51. Zimm BH, Levene SD. Problems and prospects in the theory of gel electrophoresis of
765 DNA. Q Rev Biophys. 1992;25(2):171-204.
766 52. Martinez-Vaz BM, Xie Y, Pan W, Khodursky AB. Genome-wide localization of mobile
767 elements: experimental, statistical and biological considerations. BMC Genomics.
768 2005;6:81.

769 53. Stoebel DM, Hokamp K, Last MS, Dorman CJ. Compensatory evolution of gene
770 regulation in response to stress by Escherichia coli lacking RpoS. PLoS Genet.
771 2009;5(10):e1000671.
772 54. Gode-Potratz CJ, Kustusch RJ, Breheny PJ, Weiss DS, McCarter LL. Surface sensing in
773 Vibrio parahaemolyticus triggers a programme of gene expression that promotes
774 colonization and virulence. Mol Microbiol. 2011;79(1):240-63.
775 55. Kim YK, McCarter LL. Cross-regulation in Vibrio parahaemolyticus: compensatory
776 activation of polar flagellar genes by the lateral flagellar regulator LafK. J Bacteriol.
777 2004;186(12):4014-8.
778 56. McCarter LL. Dual flagellar systems enable motility under different circumstances. J
779 Mol Microbiol Biotechnol. 2004;7(1-2):18-29.
780 57. Vandecraen J, Monsieurs P, Mergeay M, Leys N, Aertsen A, Van Houdt R. Zinc-
781 Induced Transposition of Insertion Sequence Elements Contributes to Increased
782 Adaptability of Cupriavidus metallidurans. Front Microbiol. 2016;7:359.
783 58. Al-Aboud KM, Al-Natour S. Lipoid proteinosis: a report of 2 siblings and a brief
784 review of the literature. Saudi Med J. 2008;29(12):1835; author reply
785 59. Teras R, Horak R, Kivisaar M. Transcription from fusion promoters generated during
786 transposition of transposon Tn4652 is positively affected by integration host factor in
787 Pseudomonas putida. J Bacteriol. 2000;182(3):589-98.
788 60. Soutourina O, Kolb A, Krin E, Laurent-Winter C, Rimsky S, Danchin A, et al. Multiple
789 control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic
790 AMP-catabolite activator protein complex in transcription of the flhDC master operon. J
791 Bacteriol. 1999;181(24):7500-8.
792 61. Saier MH, Jr. Protein phosphorylation and allosteric control of inducer exclusion and
793 catabolite repression by the bacterial phosphoenolpyruvate: sugar phosphotransferase
794 system. Microbiol Rev. 1989;53(1):109-20.
795 62. Stella NA, Kalivoda EJ, O'Dee DM, Nau GJ, Shanks RM. Catabolite repression control
796 of flagellum production by Serratia marcescens. Res Microbiol. 2008;159(7-8):562-8.
797 63. Lee H, Doak TG, Popodi E, Foster PL, Tang H. Insertion sequence-caused large-scale
798 rearrangements in the genome of Escherichia coli. Nucleic Acids Res. 2016;44(15):7109-19.
799

800

Figure Legends

Figure 1. The schematic diagram of the *flhDC* and *motAB/cheAW* operons and constructs used in the present studies.

A). Schematic diagram of the wild type *flhDC* and *motAB/cheAW* operons. The *flhD* and *flhC* genes are transcribed from one promoter while the *motA, motB, cheA* and *cheW* genes are transcribed from a second promoter located in the 3' region of *flhC*. Red rectangles indicate binding sites for some negative regulators, and blue rectangles indicate the same for certain positive regulatory proteins. The lines below the gene depictions are binding sites for small RNAs. The two arrows indicate the promoters.

B). Schematic diagram of the constructs used to understand the regulated expression of the *flhDC* operon in the CAT strain. The *cat* gene is fused immediately after the *flhC* gene, yielding the *flhDC*:*cat* operon that is under the control of the native *flhDC* promoter. The 3' region of the *flhC* gene bearing the *motA* promoter was fused to the *motA* gene. This promoter is therefore between the *cat* and *motA* genes, so that the downstream genes are expressed under the control of the native promoter.

C). Schematic diagram of the CAT2 strain. The region containing the *flhDC* promoter and the *flhDC*:*cat* operon was moved into the *intS/yfdG* intergenic region (bottom). The original *flhDC* promoter, *flhD* and the 5' part of *flhC* were deleted (top).

Figure 2. Swarming rates of wild type (CAT) and IS mutants without (A) or with (B) glucose. Diameters of swarming zones of IS mutants and the wild type strain when plated on soft agar without (A) or with (B) glucose (0.5%) in the medium as a function of time. All plates were incubated at room temperature. The experiments were conducted in triplicate, and the error bars (not always visible) represent the standard deviations (SD) of the three measurements. The error

was always less than 10% of the measured values. The strains examined are those listed in Table 1, and the raw numbers plotted here are those presented in that table.

Figure 3. Growth and mutant appearance frequencies on solid agar plates (A), in liquid medium (B) and in soft (0.3%) agar (C).

A. CAT (orange rhombuses), CAT_IS*1* (blue squares), and CAT_IS*5* (green triangles) were inoculated separately on solid agar plates (0.5% tryptone, 0.5% NaCl, 1.5% agar) and incubated at 30 °C for 25 hours. Cells were washed off the plates, diluted, plated and grown for colony counting. The growth curves were plotted for each strain as a function of time. These three curves use the left y-axis. Insertion mutant populations were determined by plating onto LB + Cm plates. The insertion mutant ratios were determined at four time points throughout the 25 hour period. The ratio is recorded on the right y-axis. Error, expressed in S.D. values, was less than 5% of the measured values in all Figures, A, B and C.

B. CAT, CAT_IS*1* and CAT_IS*5* strains were inoculated separately into liquid media (0.5% tryptone, 0.5% NaCl, 0% agar) at 30 °C for 24 hours, and cultures were rotated at 250 rpm. The growth curves were plotted for each strain. Insertion frequencies in liquid media from the CAT strain were measured, and the insertion mutant ratios were measured at four time points through the 25 hour period. Format of presentation is as for Figure A.

C. CAT, CAT_IS*1* and CAT_IS*5* strains were inoculated into soft-agar plates (0.5% tryptone, 0.5% NaCl, 0.3% agar) separately and incubated at 30 °C for 24 hours. The populations were measured every three hours. Format of presentation is as for Figure A. Insertion frequencies are presented in Figures 4.

Figure 4. Growth curves and mutant appearance frequencies in soft agar plates containing 0.5% tryptone (A) or (B) 2.5% tryptone (B).

For Figure A, overnight cultures of CAT_Km^R (wild type, Km^R, $2x10^9$ cfu/ml; squares) and a CAT_ IS*5* mutant (Cm^R, $2x10^6$ cfu/ml; circles) were mixed and inoculated into soft-agar plates (0.5% tryptone, 0.5% NaCl, 0.3% agar) at 30 °C. For Figure B, the same strains were inoculated into soft-agar with 2.5% tryptone plus NaCl and 0.3% agar. All soft-agar plates were incubated at 30 °C. Every three hours, the cells were collected and plated on LB and LB + Cm plates to measure the two inoculated populations. The emerging insertion mutant populations (Cm^R Km^R, triangles) were determined on LB + Cm + Km plates. All measurements were conducted in triplicate, and the error, expressed in standard deviations (SDs), was less than 10% of the measured values. Error bars are often obscured by the symbols representing the averages of the three determinations.

Figure 5. The upstream *flhDC* promoter region showing IS insertion sites and the orientations of the elements characterized in this study (A), and depiction of the SIDDs that overlap the 370 bp region into which IS elements insert in the *flhDC* upstream region (B).

A. The positions of the insertion sites and their orientations for the newly inserted IS*1*, IS*3*, and IS*5* elements in the upstream region of the *flhD* promoter reported in this study are presented. The transcriptional start site is indicated as +1, and other locations, in parentheses, were assigned relative to this start site. The directions of the arrows correspond to the orientations of the insertions. IS*1* has six different locations, IS*3*, one location, and IS*5*, three different locations. IS*1*, IS*3* and IS*5* insertion sites are gray highlighted, bolded, and underlined, respectively. Single headed errors indicate the direction of transcription of the transposase gene. Double headed arrows for IS*5* indicate that this IS element was inserted in both orientations at each of its 3 sites. B. Destabilization energy plot for a DNA segment that includes the *flhD/uspC* intergenic region (see *Materials and Methods*). Destabilization energy G(x) is the incremental energy (kcal/mol)

needed to guarantee separation of base pair x under defined superhelical stress. Highly destabilized (SIDD) DNA regions have lower values of G(x), and stable regions have higher G(x) values (46). The green arrow shows the start codon of the *flhD* gene; the open arrow shows the transcriptional start site (+1) of the *flhDC* operon; the black arrows correspond to the IS*5* insertion sites; blue arrows show the IS*1* insertion sites, and the red arrow reveals the IS*3* insertion site.

Figure 6. Mutant appearances from CAT2 [(A, B and C, 37 °C), (D, E and F, 30 °C)].

An overnight culture of CAT2 ($1x10^9$cfu/ml) was inoculated onto solid-agar plates (0.5% tryptone, 0.5% NaCl, 1.5% agar) (A and D), into liquid medium (0% agar) (B and E), and into soft agar plates (0.3% agar) (C and F). The solid agar and soft agar plates were incubated at 37 °C (A and C, repectively) or 30 °C (D and F, respectively) as specified, while the liquid media were incubated with rotation (250 rpm) at 37 °C (B) or 30 °C (E). The total populations (diamonds) and the emerging insertion mutant populations (squares) were determined in LB and LB + Cm media, respectively. Error for triplicate determinations was within 10% of the measured values.

Figure 7. Frequencies of appearance of insertion mutants from wild type and non-motile cells in soft agar

Overnight LB cultures of strains CAT, CAT_ΔAB and CAT_ΔfliC were washed once using 1x M9 salts and then diluted to an OD_{600} of 1. Four µl was streaked onto soft NB agar plates (0.4% NB + 0.3% agar), and the plates were incubated at 30 °C. After 18 hours, the total populations and IS insertional mutant populations were determined as described in the legend to Figure 4. Three soft agar plates were used for each strain. CAT_ΔAB and CAT_ΔfliC are nonmotile.

Figure 8. Effects of agar concentration on the appearance of insertional mutants.

893 An overnight LB culture of CAT was washed once and diluted to OD_{600} of 1. Four µl of diluted
894 culture was streaked onto NB agar plates (0.4% NB + 0 to 1.2% agar), and the plates were
895 incubated at 30 °C. After 18 h, the total populations and mutant populations in the plates with
896 0% to 0.4% agar were determined as described in the legend to Figure 4. For the plates with
897 0.6% agar or above, the cells were washed off the plates without breaking the agar, and the total
898 populations and mutant populations were determined as described in the legend to Figure 4.
899 These experiments were conducted in triplicate, and error bars represent standard deviations.
900 **Figure 9. Insertion sites vs. swarming rates.**
901 Colony expansion swarming rates in soft agar for the different insertion mutants were plotted as
902 a function of distance from the *flhD* transcriptional start site. The data are from Table 1 (no
903 glucose). The IS element type and its insertional orientation are indicated (direct (dir), or reverse
904 (rev)). Distance from the transcriptional start site is indicated on the x-axis. Errors of triplicate
905 determinations were usually within 5% of the measured values.
906

907 **Table 1. Colony expansion swarming rates of wild type and IS mutants with and without**

908 **glucose***

Strains	Swarming Rate (mm/h) without glucose	with glucose	ratio
IS*1*(-469 rev)	2.20	3.60	1.6
IS*1*(-415 dir)	2.50	3.90	1.6
IS*1*(-214 rev)	2.75	3.83	1.4
IS*1*(-180 rev)	2.80	3.92	1.4
IS*1*(-120 rev)	2.90	4.35	1.5
IS*1*(-107 rev)	3.05	4.30	1.4
IS*3*(-199 rev)	2.70	4.20	1.5
IS*5*(-318)	2.40	3.80	1.6
IS*5*(-318 rev)	2.30	3.90	1.7
IS*5*(-169)	2.35	3.80	1.7
IS*5*(-169 rev)	2.50	3.90	1.6
IS*5*(-99)	2.70	4.20	1.5
IS*5*(-99 rev)	2.60	4.10	1.5
BW25113 (wild-type)**	1.14	0.50	0.4
CAT (wild-type)**	1.00	0.40	0.4

909

910 * The diameters of the swarming zones were measured in soft-agar plates (0.5% tryptone, 0.5%

911 NaCl, 0.3% agar) with or without 0.5% glucose. Each strain was grown in a liquid medium for

912 10-12 hours before inoculation. 1.5 µl of diluted culture was inserted into the soft-agar in the

913 middle of the plates, and the plates were incubated at room temperature. The diameters of the

914 swarms were measured every two hours to calculate the swarming rates (mm/h). n=3. The three

915 determinations were averaged. The error, expressed in standard deviations of these

916 determinations, was <0.4 mm/h.

917 ** The sizes of the colonies were measured for wild-type and CAT (wild-type) instead of the

918 swarming zone because of their minimal motility. Thus, the ratio ± glucose, represents a

919 maximal value.

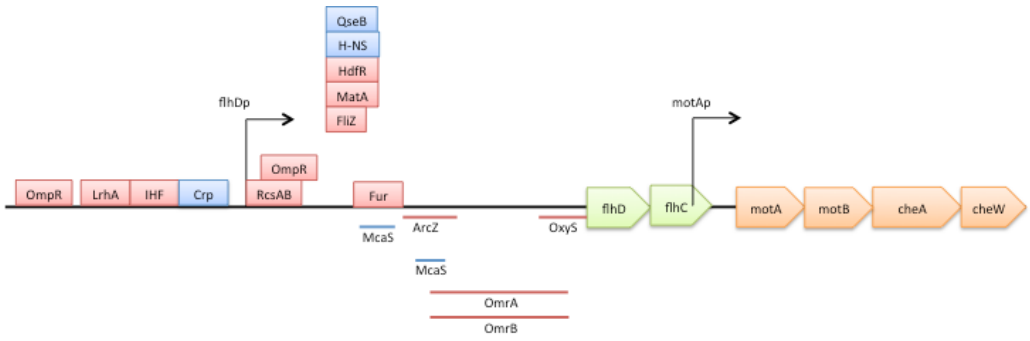
920

Table 2. Doubling time comparisons in 0.5% (top) and 2.5% (bottom) tryptone soft-agar plates.

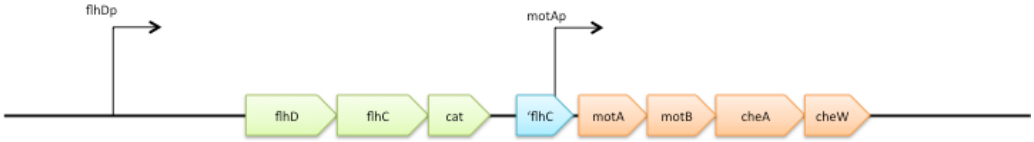
Apparent growth rate*	CAT_Km^R	CAT_IS*5*	Emerging mutants
Maximal (0.5% Tryptone; 1-15h)	62	49	65
Average (0.5% Tryptone; 15-25h)	288	866	60
Maximal (2.5% Tryptone; 1-15h)	57	47	60
Average (2.5% Tryptone; 15-25h)	238	492	56

*Doubling times for CAT_Km^R (wild type), CAT_IS*5*, and emerging insertion mutants in 0.5% tryptone soft-agar plates (top) and 2.5% tryptone soft agar plates (bottom). The maximal growth rate for the 1 to 15 h interval was the steepest slope before the 15 hour time point, and the apparent average growth rate for the subsequent time period was the slope between 15 and 25 hours. Values are expressed in minutes. Note that the values for the emerging mutants represent a sum of their appearance rates and their growth rates, with the former presumably predominating.

(a)



(b)



(c)

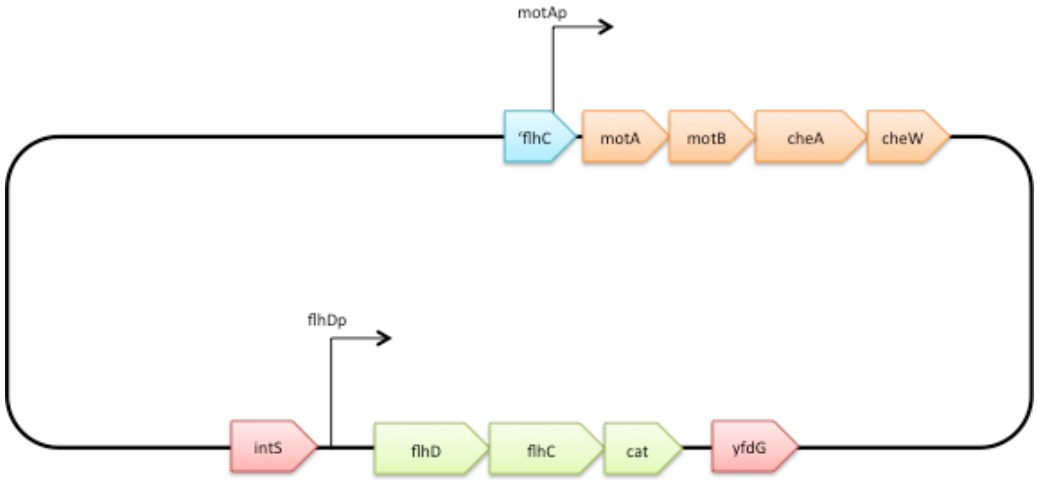


Figure 1

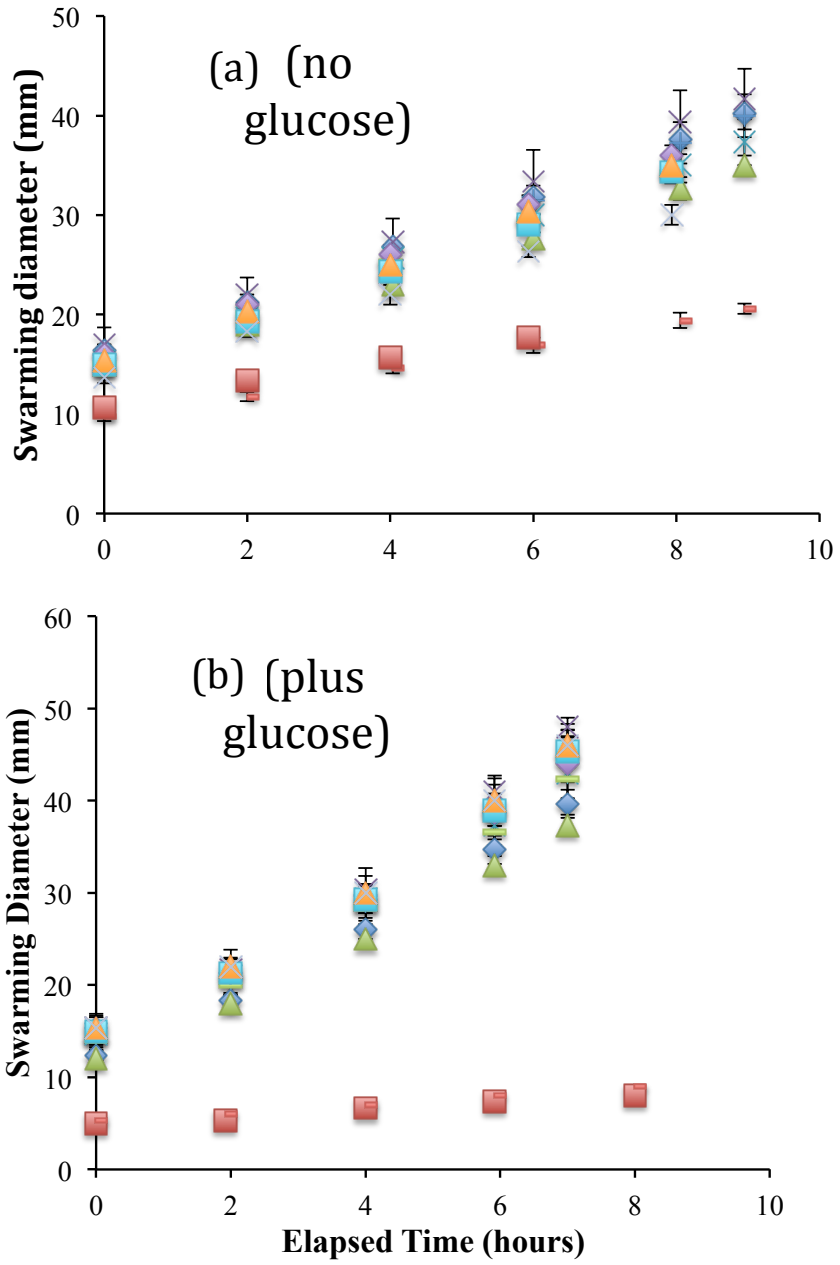


Figure 2

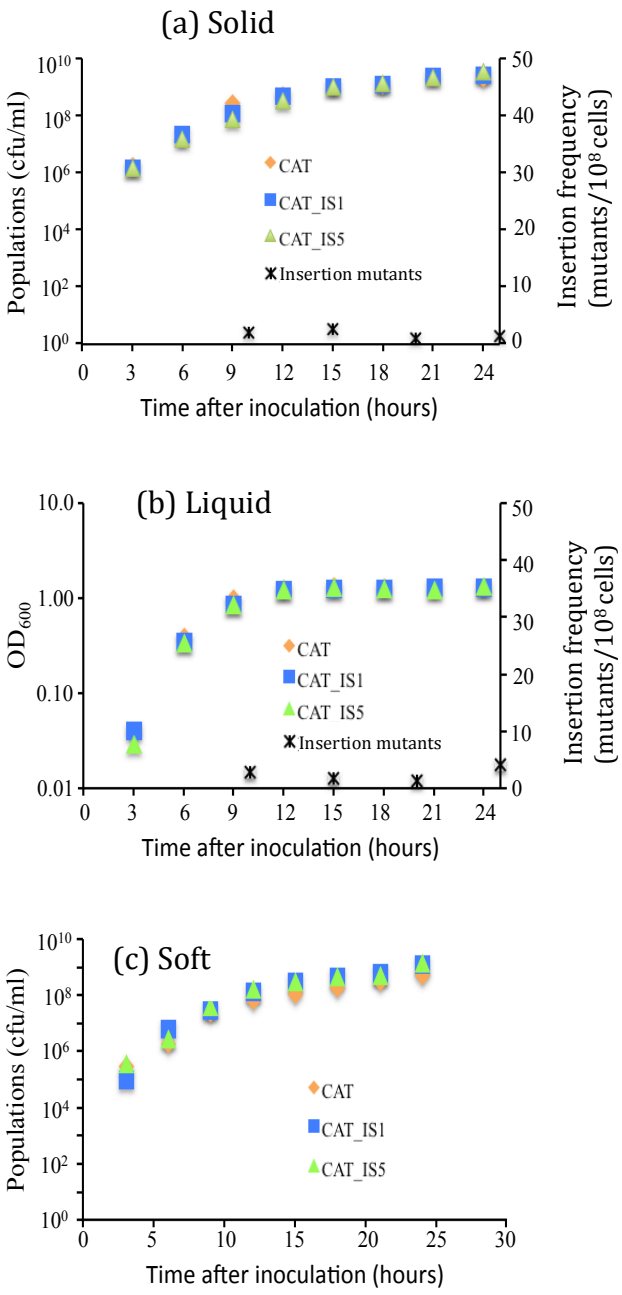


Figure 3

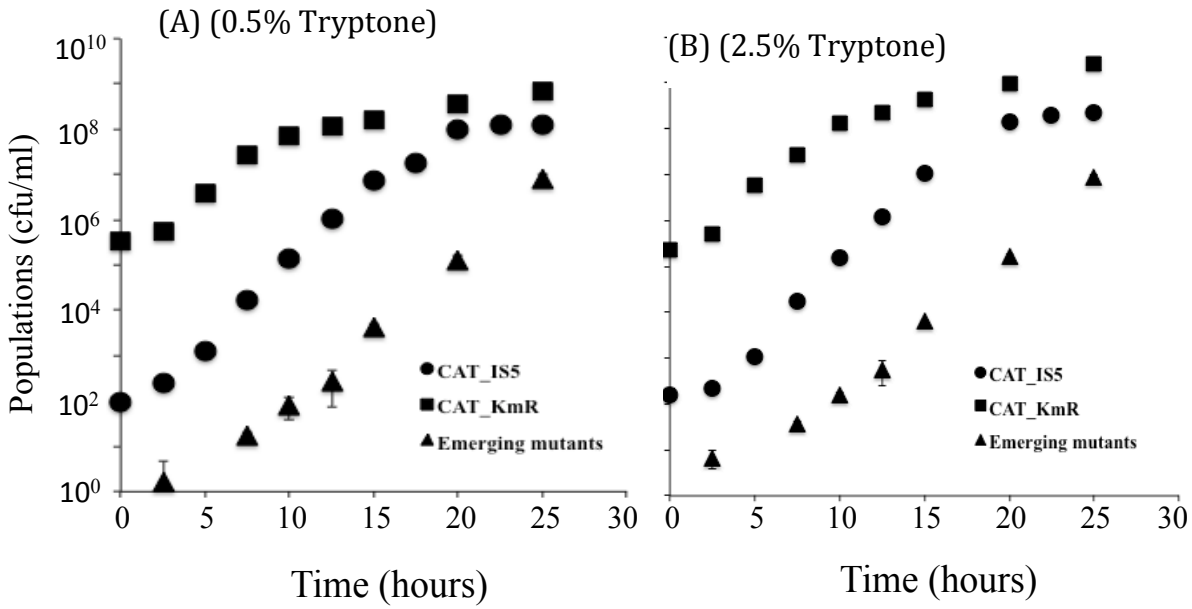
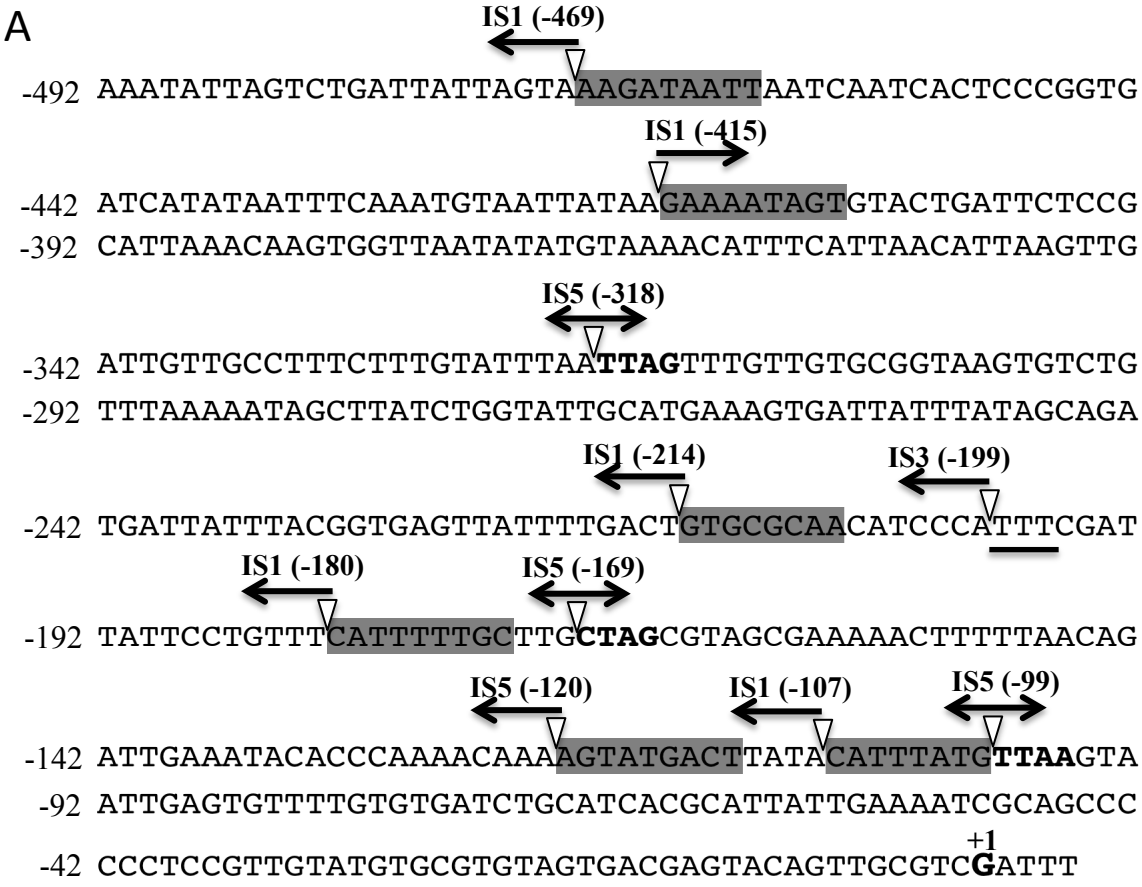
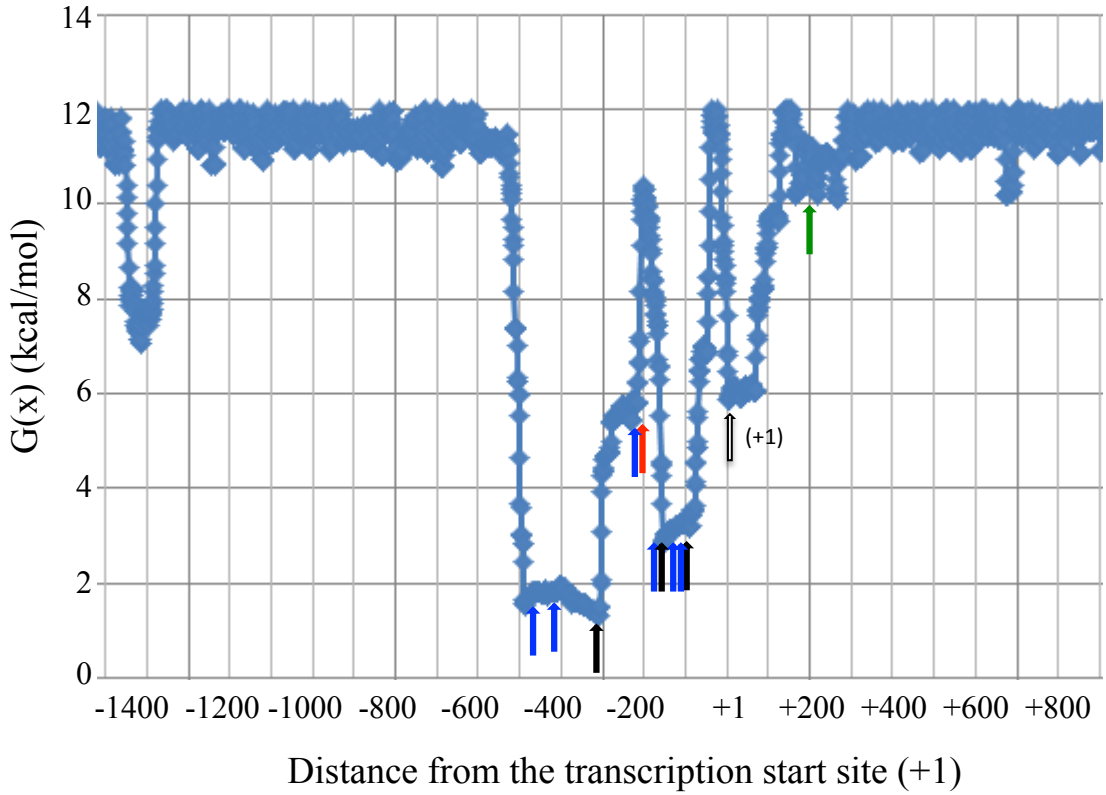


Figure 4

A



B



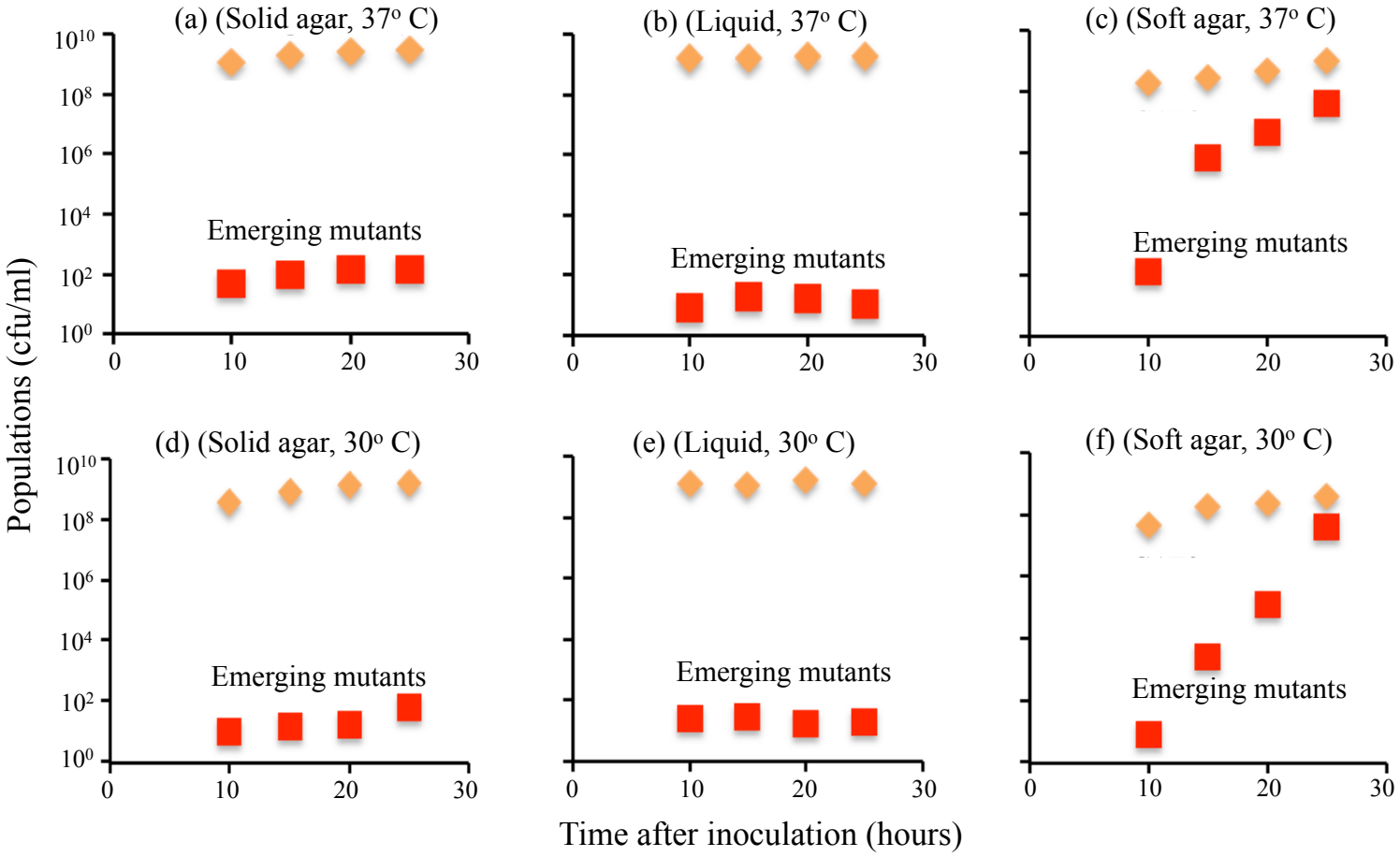


Figure 6

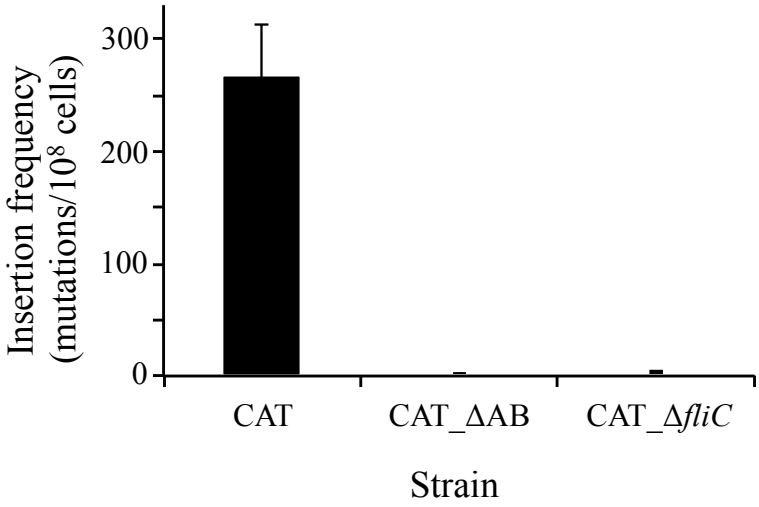


Figure 7

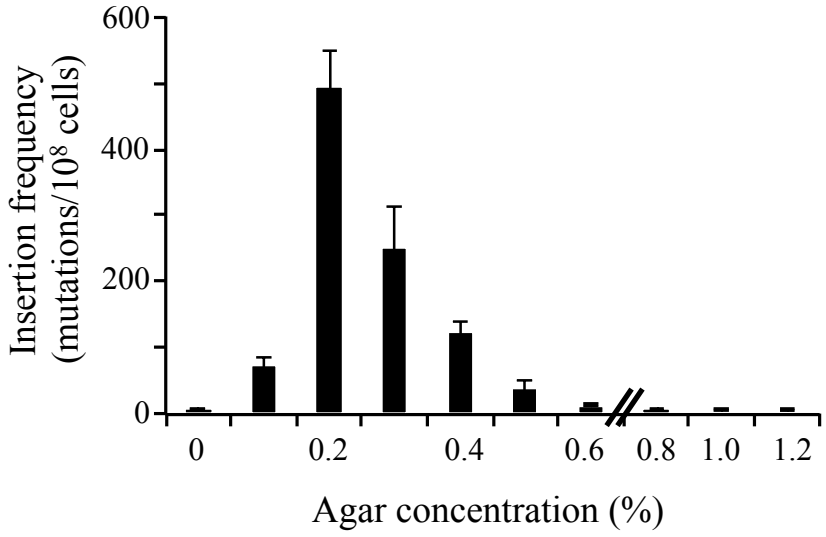


Figure 8

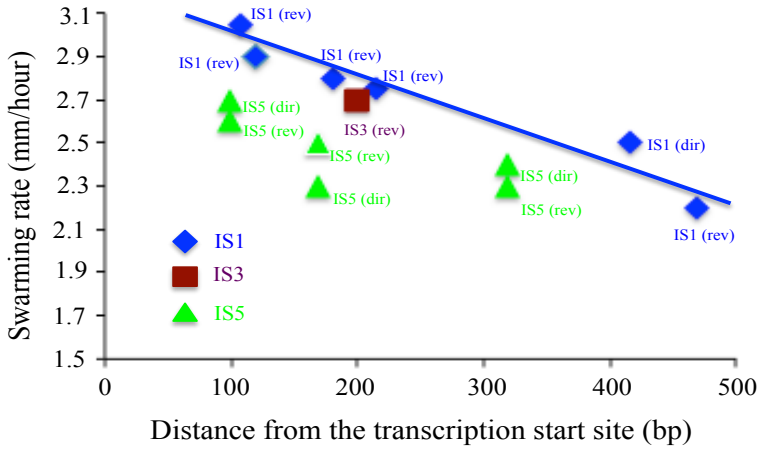


Figure 9

Table S1. *E. coli* strains used in this study

Strains	Genotype/relevant characteristics	Source
BW25113	wild type, weak swarmer	Datsenko and Wanner, 2005
CAT	*flhDC:cat:'flhC:motA*/Cm^R, used as a wild type	This study
CAT_Km^R	Km^R, in a wild type background	This study
CAT2	Δ*PflhDC-flhD*-5'-*flhC*; P*flhDC*:*flhD*:*flhC*:*cat* located in the *intS/yfdG* intergenic region, used as a wild type	This study
CAT_IS*1*	CAT strain with IS*1* upstream of the *flhD* promoter	This study
CAT_IS*5*	CAT strain with IS*5* upstream of the *flhD* promoter	This study
CAT_ΔAB	CAT, in which *motAB* is not expressed; this study nonmotile	
CAT_ΔfliC	CAT, in which *flic* is not deleted; this study nonmotile	
IS*1*(-107 rev)	IS*1* insertion into the -107 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*1*(-120 rev)	IS*1* insertion into the -120 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*1*(-180 rev)	IS*1* insertion into the -180 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*1*(-214 rev)	IS*1* insertion into the -214 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*1*(-415)	IS*1* insertion into the -415 bp position relative to the *flhD* start site (+1) in direct orientation	This study
IS*1*(-469 rev)	IS*1*insertion into the -469 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*3*(-199 rev)	IS*3* insertion into the -206 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*5*(-99)	IS*5* insertion into the -99 bp position relative to the *flhD* start site (+1) in direct orientation	This study
IS*5*(-99rev)	IS*5* insertion into the -99 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*5*(-169)	IS*5* insertion into the -169 bp position relative to the *flhD* start site (+1) in direct orientation	This study
IS*5*(-169 rev)	IS*5* insertion into the -169 bp position relative to the *flhD* start site (+1) in reverse orientation	This study
IS*5*(-318)	IS*5* insertion into the -318 bp position relative to the *flhD* start site (+1) in direct orientation	This study
IS*5*(-318rev)	IS*5* insertion into the -318 bp position relative to the *flhD* start site (+1) in reverse orientation	This study

Table S2. Oligonucleotides used in this study

Name	Sequence	Use
PflhDC-Xho-F	atactcgagcttattctgtgaacttcaggtgac	Amplification of the *flhDC* promoter
PflhDC-Rn	tgcagatcacacaaaacactcaattac	Amplification of the *flhDC* promoter
FlhC.cat-P1	gatattatcccacaactgctggatgaacagagagtacaggctgtttaactgaaggagctaaggaagctaaaatggag	Integration of *cat* downstream of *flhC*
FlhC.cat-P2	catcatccttccactgttgaccatgacaggatgttcagtcgtcaggcgttaactgtgtaggctggagctgcttc	Integration of *cat* downstream of *flhC*
motA-ver-R	gtagctggcgaagatctcgctctcac	Verification of *cat* integration
flhC-Xho-F	tatctcgagcggaatgattactgcaactttccagctgcaac	Cloning the 3' region of *flhC* and the 5' region of *motA* to pKDT
motA-Bam-R	aatggatccgtatttggagcgacgaaacagcaacggcagc	Cloning the 3' region of *flhC* and the 5' region of *motA* to pKDT
PmotAB-P1	gaattacaacagtactgcgatgagtggcagggcggggcgtaaggcgcgccattgtgtaggctggagctgcttc	Integration of the 3' region of *flhC* and the *flhC*/*motA* intergenic region upstream of *motA*
PmotAB-P2	gtatttggagcgacgaaacagcaacggcagc	Integration of the 3' region of *flhC* and the *flhC*/*motA* intergenic region upstream of *motA*
ycaD-P1	atcagacgcgatgcattgctctgaaagcatagacgggaaatatgagtttgctgtgtaggctggagctgcttc	Insertion of km^r gene into *ycaD/intS* intergenic region
ycaD-P2	ggtgaaaatacgcgatatcccagcggcggtattatcgatttatattaccatatgaatatcctccttag	Insertion of km^r gene into *ycaD/intS* intergenic region
ycaD-ver-R	gccagagtcaacaaaagcaggc	Verification of km^r gene insertion into *ycaD/intS* intergenic region
Cat-Sal-R	atagtcgacattacgccccgccctgccactcatc	Cloning P*flhDC-flhDC.cat* into pKDT
intC-P1	tgagaaggtggagtgagcgaccttaacaactatcgaatagcacaaagtcttgtgtaggctggagctgcttc	Integration of P*flhDC-flhDC.cat* into *intS/yfdG* intergenic region
intC-P2	ttcttctatcagctaataatcaaaggaatgaagtctatcatccaagtcttcatgagaattaattccgggatcc	Integration of P*flhDC-flhDC.cat* into *intS/yfdG* intergenic regions

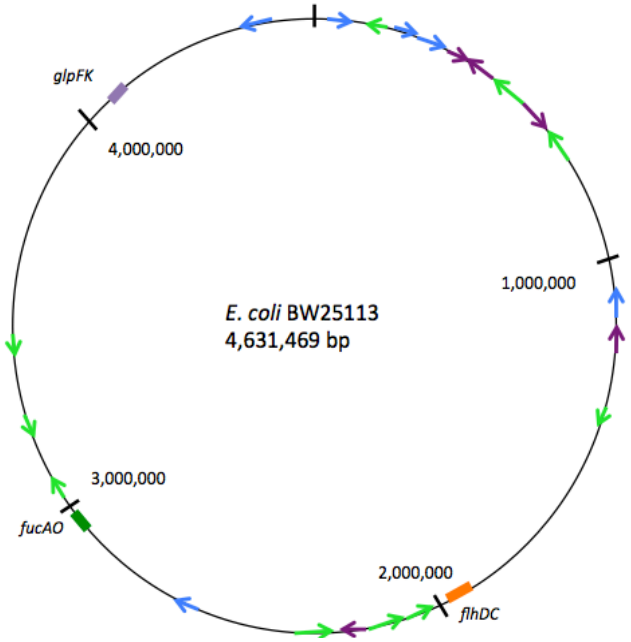


Figure S1. Original locations of relevant IS elements on the *E. coli* K12 chromosome. Figure S1

The original locations of IS elements and related genes at the beginning of this study are indicated on the *E. coli* K-12 BW25113 (GCA_000750555.1) chromosome. Green arrows correspond to IS*5*, purple to IS*3*, and blue to IS*1*. The heads of the arrows indicate the orientations of transcription of the transposase genes. The *flhDC* operon is located between 1972344 and 1973236 bp.