



Parallel Evolution towards Increased Motility in Long-Term Cultures of *Escherichia coli*, Even Though Motility was Not Required for Long-Term Survival

Autumn L. Henderson,^{a*} Angie Moreno,^a  Karin E. Kram^a

^aDepartment of Biology, California State University, Dominguez Hills, Carson, California, USA

ABSTRACT *Escherichia coli* can survive for long periods in batch culture in the laboratory, where they experience a stressful and heterogeneous environment. During this incubation, *E. coli* acquires mutations that are selected in response to this environment, ultimately leading to evolved populations that are better adapted to these complex conditions, which can lead to a better understanding of evolutionary mechanisms. Mutations in regulatory genes often play a role in adapting to heterogeneous environments. To identify such mutations, we examined transcriptional differences during log phase growth in unaged cells compared to those that had been aged for 10 days and regrown. We identified expression changes in genes involved in motility and chemotaxis after adaptation to long-term cultures. We hypothesized that aged populations would also have phenotypic changes in motility and that motility may play a role in survival and adaptation to long-term cultures. While aged populations did show an increase in motility, this increase was not essential for survival in long-term cultures. We identified mutations in the regulatory gene *sspA* and other genes that may contribute to the observed differences in motility. Taken together, these data provide an overall picture of the role of mutations in regulatory genes for adaptation while underscoring that all changes that occur during evolution in stressful environments are not necessarily adaptive.

IMPORTANCE Understanding how bacteria adapt in long-term cultures aids in both better treatment options for bacterial infections and gives insight into the mechanisms involved in bacterial evolution. In the past, it has been difficult to study these organisms in their natural environments. By using experimental evolution in heterogeneous and stressful laboratory conditions, we can more closely mimic natural environments and examine evolutionary mechanisms. One way to observe these mechanisms is to look at transcriptomic and genomic data from cells adapted to these complex conditions. Here, we found that although aged cells increase motility, this increase is not essential for survival in these conditions. These data emphasize that not all changes that occur due to evolutionary processes are adaptive, but these observations could still lead to hypotheses about the causative mutations. The information gained here allow us to make inferences about general mechanisms underlying phenotypic changes due to evolution.

KEYWORDS motility, evolution, RNA-seq, transcriptome, bacterial genetics

In the natural environment, organisms regularly encounter heterogeneous and stressful environments. Studying how a species evolves can provide insight into why species have certain traits and help explain different adaptive strategies that lead to survival and increased fitness in stressful environments (1). Changes that occur during growth in new environments can either be sustained, increased, or removed by natural selection or genetic drift (1). Changes can also be driven by correlated evolution, where two or more traits are driven together by the evolutionary process (2).

Editor Beile Gao, South China Sea Institute of Oceanology

Copyright © 2022 Henderson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Karin E. Kram, kkram@csudh.edu.

*Present address: Autumn L. Henderson, University of California, Irvine, Biological Sciences III, Irvine, California, USA.

The authors declare no conflict of interest.

Received 19 November 2021

Accepted 31 May 2022

Escherichia coli is an ideal model organism for studying evolution and adaptation, not only because it grows rapidly but also because it is quite versatile and genetically diverse (3). Because there are infectious strains of *E. coli*, we must understand how it evolves in different environments to produce more effective antibiotics. Studying *E. coli* is beneficial for not only the development of more targeted treatments for infection but also for insight into basic biology, such as adaptive mechanisms.

E. coli, like many microbial organisms, has a five-phase life cycle in the laboratory: lag phase, log or exponential-phase, stationary-phase, death phase, and long-term stationary-phase (LTSP) (4). LTSP is a complex environment, where the total number of cells remains relatively constant because the replication and death rates of cells in this phase are balanced. During LTSP, cells undergo multiple stresses such as starvation, oxidative stress, and changes in pH (5). During this phase, different populations that have acclimated or adapted to the environment grow while other populations died, creating new conditions that allow new populations to grow (4). The continuous cycling of nutrients and waste products creates a heterogeneous, complex, and stressful environment in which cells need to adapt to survive. As the *E. coli* population ages, mutations that allow for better survival in stressful environments such as LTSP are selected for and can increase in frequency (4). While we and others have begun to explore this environment, the exact changes and stresses cells experience during LTSP are mostly unknown (6–10).

Cells can respond to their environment by altering gene expression in an impermanent way (physiological acclimation), but they can also adapt by accumulating mutations that permanently change expression for one or multiple genes. Mutations that affect gene expression have frequently been identified as one of the actors of evolutionary change, indicating that mutations in *trans* regulatory factors may play a role in adaptive evolution in multiple organisms and environments (11–13). One way to explore the conditions cells experience during this stressful period and determine what regulatory factors may help cells adapt to these stresses, is to examine transcriptional changes that occur after cells have been incubated into LTSP.

To identify transcriptional differences between unaged *E. coli* and cells aged into LTSP, we performed RNA-seq (14) on log-phase cells of both aged and unaged populations. Because we previously identified regulatory changes in LTSP (15), we specifically aimed to identify transcriptional changes caused by mutations in the genome accumulating during long-term incubation. Comparing aged and unaged populations when grown into log-phase allowed us to differentiate permanent transcriptional changes caused by mutations in the genome and those caused by plastic changes due to normal regulatory systems in the cell. We identified multiple genes involved in motility and chemotaxis that were upregulated in cells that had adapted to LTSP. Based on these data, we hypothesized that motility would be increased in cells that were adapted to LTSP, and that motility would be important for cells to survive LTSP. Our data supported our first hypothesis that aged cells were more motile but not the second because motility was not essential for survival in LTSP. We also identified mutations that may cause changes in motility and other expression differences. These data help to understand the types of changes that occur during survival in heterogeneous and stressful environments.

RESULTS

Multiple genes were differentially expressed in aged populations, including many involved in motility. To determine if incubation into LTSP led to differences in gene expression, we compared whole transcriptomes of log-phase cells (4 h) that were previously aged for 10 days (aged populations) to their parent strain which had not been previously aged (unaged populations). To ensure that these strains were in the same phase, we compared timing in lag, log, and entrance to stationary-phase and found no differences (Fig. S1). We identified 64 total genes that were differentially expressed (absolute log₂ change of at least 2 and q-value < 0.05) between the two populations (Table 1). Thirty-five of these genes had higher expression in aged populations, whereas 29 genes had higher expression in unaged populations. Fourteen of the genes more expressed in unaged cells were related to resistance to acid stress (out of 38 genes categorized with the response to elevated pH GO

TABLE 1 Differentially expressed genes between unaged and aged populations of *E. coli*

Gene	Log ₂ fold change aged/unaged	q-value	Description
Upregulated in aged cells			
Genes involved in chemotaxis and motility			
<i>flgB</i>	4.0	1.4×10^{-9}	flagellar basal-body rod protein
<i>flgC</i>	3.9	6.5×10^{-10}	flagellar basal-body rod protein
<i>fliC</i>	3.7	3.3×10^{-15}	flagellar filament structural protein
<i>flgE</i>	3.5	2.0×10^{-15}	flagellar hook protein
<i>flgF</i>	3.5	1.8×10^{-10}	flagellar basal-body rod protein
<i>tar</i>	3.5	1.6×10^{-9}	methyl-accepting chemotaxis protein
<i>flgD</i>	3.4	1.3×10^{-12}	flagellar biosynthesis, initiation of the hook assembly
<i>tap</i>	3.3	1.0×10^{-9}	methyl-accepting chemotaxis protein - dipeptide-sensing
<i>flgG</i>	3.2	1.2×10^{-11}	flagellar basal-body rod protein
<i>pdhE</i>	3.1	6.4×10^{-6}	c-di-GMP phosphodiesterase involved in the regulation of the switch from flagellar motility to sessile behavior
<i>cheA</i>	3.0	4.1×10^{-11}	chemotactic and aerotactic two-component signal transduction system histidine protein kinase
<i>fliF</i>	2.9	7.3×10^{-6}	flagellar basal-body MS-ring and collar protein
<i>fliN</i>	2.8	4.5×10^{-4}	flagellar motor switch protein
<i>fliD</i>	2.7	1.6×10^{-6}	flagellar filament capping protein
<i>fliA</i>	2.7	4.2×10^{-5}	sigma 28 (sigma F) factor responsible for genes involved in motility and flagellar synthesis
<i>flgH</i>	2.6	6.1×10^{-6}	flagellar L-ring protein
<i>cheW</i>	2.4	6.0×10^{-7}	chemotactic and aerotactic two-component signal transduction system coupling protein
<i>fliG</i>	2.3	7.2×10^{-5}	flagellar motor switch protein
<i>fliH</i>	2.3	1.1×10^{-2}	flagellar biosynthesis protein
<i>cheB</i>	2.3	1.8×10^{-4}	chemotactic and aerotactic two-component signal transduction system response regulator
<i>flgK</i>	2.3	1.7×10^{-5}	flagellar hook-filament junction protein 1
<i>fliS</i>	2.1	1.1×10^{-2}	flagellar biosynthesis protein
<i>motB</i>	2.1	1.6×10^{-3}	motility protein B
<i>flgI</i>	2.0	3.6×10^{-3}	flagellar P-ring protein
<i>fliL</i>	2.0	9.0×10^{-3}	flagellar protein
<i>ycgR</i>	2.0	1.3×10^{-3}	flagellar brake protein
<i>flgL</i>	2.0	3.4×10^{-5}	flagellar hook-filament junction protein 2
Genes involved in transport			
<i>ompF</i>	3.1	7.6×10^{-14}	outer membrane porin F
<i>nupG</i>	2.8	1.6×10^{-14}	nucleoside:H ⁺ symporter
<i>tsx</i>	2.5	3.5×10^{-10}	nucleoside-specific channel-forming protein
<i>gntP</i>	2.3	1.8×10^{-8}	fructuronate transporter
Genes involved in nucleoside metabolism			
<i>cdd</i>	2.7	3.2×10^{-5}	cytidine/deoxycytidine deaminase
<i>udp</i>	2.5	1.7×10^{-8}	uridine phosphorylase
Unknown function			
<i>ydiL</i>	2.8	5.0×10^{-9}	unknown
<i>ycdZ</i>	2.1	7.3×10^{-3}	unknown
Downregulated in aged cells			
Genes involved in acid resistance			
<i>gadB</i>	-3.8	5.8×10^{-19}	glutamate decarboxylase B
<i>dctR</i>	-3.7	8.4×10^{-11}	putative DNA-binding transcriptional regulator
<i>gadC</i>	-3.7	1.3×10^{-15}	L-glutamate:4-aminobutyrate antiporter
<i>yhiM</i>	-3.6	3.7×10^{-15}	unknown
<i>gadA</i>	-3.4	1.6×10^{-14}	glutamate decarboxylase A
<i>hdeD</i>	-3.3	7.8×10^{-9}	acid-resistance membrane protein
<i>gadY</i>	-3.2	2.6×10^{-9}	small regulatory RNA
<i>hdeB</i>	-3.1	8.4×10^{-11}	periplasmic acid stress chaperone
<i>gadE</i>	-3.1	4.1×10^{-12}	DNA-binding transcriptional activator
<i>hdeA</i>	-3.0	1.3×10^{-9}	periplasmic acid stress chaperone
<i>slp</i>	-3.0	1.2×10^{-13}	starvation lipoprotein

(Continued on next page)

TABLE 1 (Continued)

Gene	Log ₂ fold change aged/unaged	q-value	Description
<i>gadX</i>	−3.0	1.2×10^{-10}	DNA-binding transcriptional dual regulator
<i>gadW</i>	−2.6	5.9×10^{-10}	DNA-binding transcriptional dual regulator
<i>glsA</i>	−2.1	2.1×10^{-8}	glutaminase 1
Genes involved in other stress responses			
<i>yhiD</i>	−4.4	1.3×10^{-15}	unknown
<i>ybaT</i>	−2.3	3.1×10^{-8}	putative transporter
<i>aidB</i>	−2.2	1.7×10^{-6}	isovaleryl-CoA dehydrogenase and DNA-binding transcriptional repressor
<i>cbpA</i>	−2.0	3.5×10^{-7}	curved DNA-binding protein
<i>appY</i>	−2.0	5.1×10^{-4}	DNA-binding transcriptional activator
Genes involved in metabolism			
<i>hyaA</i>	−2.3	1.7×10^{-5}	hydrogenase 1 small subunit
<i>hyaF</i>	−2.1	2.1×10^{-8}	unknown
<i>sufC</i>	−2.1	9.2×10^{-5}	Fe-S cluster scaffold complex subunit
Genes involved in transport			
<i>mdtE</i>	−3.1	3.2×10^{-6}	multidrug efflux pump membrane fusion protein
<i>mdtF</i>	−2.9	5.4×10^{-6}	multidrug efflux pump RND permease
<i>lsrC</i>	−2.2	4.9×10^{-5}	autoinducer-2 ABC transporter membrane subunit
Genes involved in motility			
<i>csgD</i>	−2.0	9.7×10^{-5}	DNA-binding transcriptional dual regulator
Genes involved in other functions			
<i>lysQ</i>	−2.8	1.5×10^{-2}	tRNA-Lys(UUU)
<i>cnu</i>	−2.7	6.6×10^{-7}	H-NS- and StpA-binding protein
<i>ryeA</i>	−2.2	6.9×10^{-3}	small antisense RNA

terms in *E. coli*), and 27 of the more expressed genes in aged cells were related to motility and chemotaxis (out of 50 genes involved in these processes).

One of the motility-related genes, *flhA*, encodes a sigma factor (sigma 28/sigma F) that is responsible for the initiation of transcription of many genes involved in motility and flagellar synthesis (16). *flhA* was expressed 6.7 times more highly in aged populations than in unaged populations and was responsible for the transcription of 19 out of the 27 motility-related genes that were upregulated in aged populations (Table 1); (17). Further, *csgD*, which was expressed at 4-fold lower levels in aged populations, encodes a repressor of *flhA* (18).

Other categories of genes that had greater expression in aged populations included multiple genes encoding proteins involved in nucleoside transport or metabolism (*nupG*, *tsx*, *cdd*, and *udp*) (19), as well as other transport proteins (*ompF* and *gntP*) (20, 21). Genes that were expressed at a lower level in aged populations included those encoding proteins involved in stress responses, anaerobic respiration (*hyaA* and *hyaF* 22); the efflux pump MdtEF (23), and a quorum-sensing transporter LsrC (24).

Cells aged to long-term stationary-phase were more motile than unaged cells. We hypothesized that the higher expression of motility-related genes described above would lead to more motile cells phenotypically. To test this hypothesis, we performed motility assays on both unaged and aged populations. To have more biological replicates and determine if this phenomenon would be re-evolved in independent cultures, we aged 6 new cultures for 10 days into LTSP (see Materials and Methods). We then saved these strains and regrew them for motility assays so that any phenotypic change observed would be due to genotypic changes from incubation into LTSP, not physiological acclimation. The average zone of motility for the unaged populations ($n = 20$) strain was 3.8 cm and it was 4.9 cm for the new aged strains ($n = 60$) (Fig. 1), which represents a significant increase in motility of 29% (Welch's *t* test; $t[38] = -2.7$; $P = 0.01$) and recapitulation of the increased motility phenotype.

Motility was not necessary for survival in LTSP. We hypothesized based on the transcriptomic and motility data that the increase in motility may be advantageous to cells in LTSP. To test this hypothesis, we chose four of the genes upregulated in long-term cultures which contribute to motility (*cheA*, *flgB*, *flhA*, and *tar*) (17), and competed strains missing each of these genes with wild-type cultures. These four genes were each in different operons and

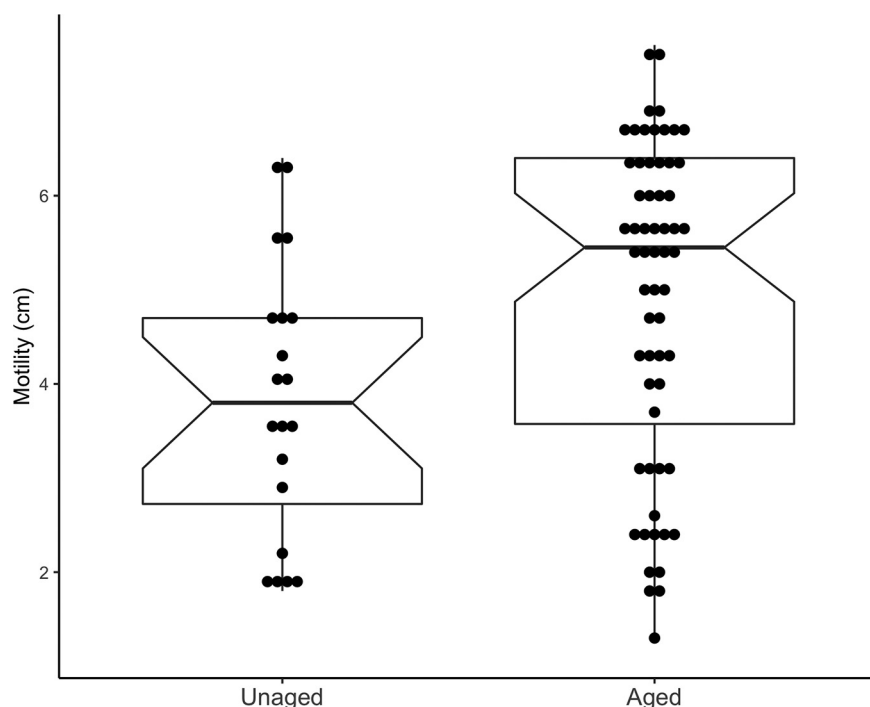


FIG 1 Motility is increased in aged populations of cells. Motility assays were performed on cells that were first aged for 10 days into LTSP as well as the parental strain of those populations that were unaged. Motility was measured in centimeters (cm) for 20 replicates of unaged populations and 60 replicates of aged populations. A Welch's *t* test showed a significant difference between the groups: *t* (38) = -2.7, *P* = 0.01.

code for proteins that were in different parts of the flagellar biosynthesis or chemotaxis pathways. To confirm that the loss of these genes affected motility in our background strain in a way we expected, we first confirmed that mutant strain cell densities were similar to the cell density seen in WT cultures during lag, log, and stationary-phase (Fig. S1) and then performed motility assays for these strains ($\Delta cheA$, $\Delta flgB$, $\Delta fliA$, and Δtar) (Fig. 2I). The $\Delta fliA$ and $\Delta flgB$ mutants showed significantly less movement than the wild-type after 12 h, which was expected as neither of these strains should be able to form flagella. The $\Delta cheA$ mutant also had significantly less movement than the wild-type strain after 12 h (Fig. 2I). However, the Δtar mutant had a zone of motility of 3.6 cm after 12 h, significantly more than the wild-type strain, which may mean that loss of this specific chemotactic response leads to increased motility overall (Fig. 2I).

We competed strains missing motility genes with WT strains in two ways: first, we combined cultures in a 1:1 ratio at lag phase and incubated them through LTSP, so that the strains experienced all phases together (Fig. 2A, C, E, and G). Second, we competed for strains with each other only in LTSP. We did this by growing both WT and mutant strains on their own for 4 days, into the beginning of LTSP, and then combining the cultures in a 1:1 ratio and incubating for another 11 days (Fig. 2B, D, F, and H). The most surprising result was observed in competition between the $\Delta cheA$ strain and the wild-type strain (Fig. 2A and B). *cheA* encodes the histidine kinase protein of the two-component signaling pathway, CheA, which signals the flagellar motor switch to change directions as it senses chemoattractants and oxygen (25). When competitions began during lag phase, the $\Delta cheA$ mutant outcompeted the wild-type cells after 24 h of incubation by a 10,000-fold difference, with wild-type cells decreasing from their inoculated cell density ($\sim 10^7$ CFU/mL at inoculation to between 10^5 and 10^6 CFU/mL at 24 h) (Fig. 2A). A drastic decline in wild-type cell densities continued and by day 11 there was no detectable wild-type strain present. However, after normal cell growth for the $\Delta cheA$ mutant strain, the mutant strain maintains a cell density of $\sim 10^8$ CFU/mL after 48 h of incubation (Fig. 2A). For competitions beginning in LTSP, we found that this effect was not as clear (Fig. 2B). In only one out of three competitions, the WT

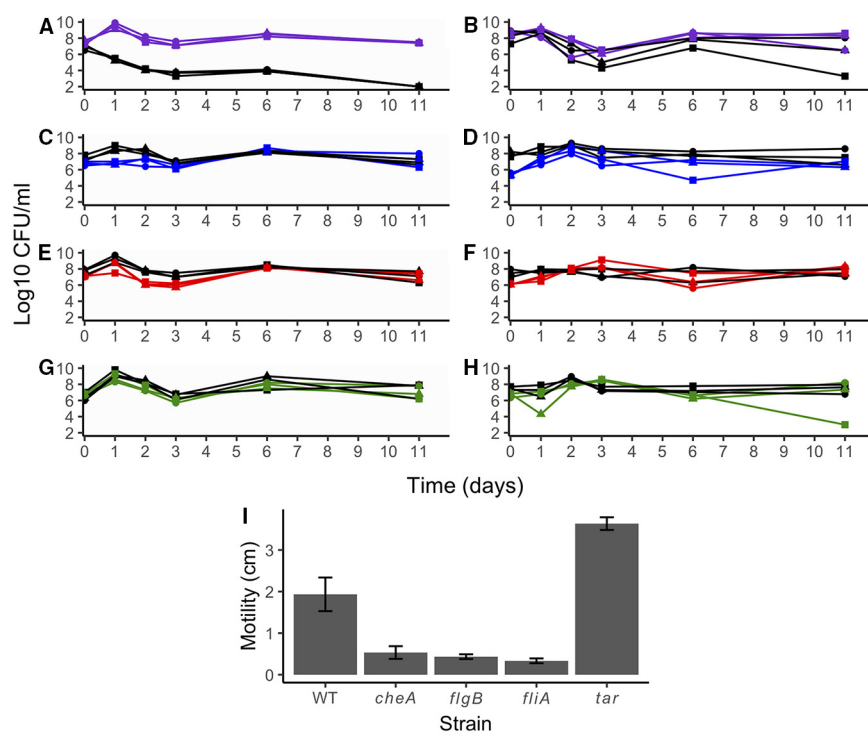


FIG 2 Motility is not correlated with survival or advantage in LTSP. (A to H) Wild type (WT) cells (black lines) competed with mutant strains ($\Delta cheA$, purple; $\Delta flgB$, blue; $\Delta fliA$, red; Δtar , green) by inoculating cultures in a 1:1 ratio and incubating them for 10 days in LTSP. Cultures were either coinoculated in a 1:1000 ratio with fresh LB after overnight growth (A, C, E, and G), and, therefore, competed during outgrowth or after 4 days of incubation in monoculture (B, D, F, and H) and only during LTSP. If the loss of the gene in the mutant strain did not affect survival in LTSP, we would have expected to see the WT strain and mutant strain at similar levels throughout the growth curve. Symbols represent the same culture for each of the three competitions. (I) Motility assays were performed with the WT strain and each mutant strain with the zone of motility measured at 12 h. The bar represents the average of three replicates of each strain, and the error bars represent the standard deviation. Each mutant strain's motility measurement was significantly different than the wild-type strain (Student's *t* test; *P* < 0.005).

strain decreased in CFU/mL after 11 days of incubation, and in the other two competitions, mutant and WT strains were similar in CFU/mL throughout. It was possible that $\Delta cheA$ strains could grow so quickly when competing with WT cells that the WT cells could not recover, although the growth rates were similar to those when they were grown on their own (Fig. S1). Comparatively, there was little outgrowth of $\Delta cheA$ cells in competition only in LTSP.

flgB encodes a protein in the flagellar basal body, and unlike *cheA*, is not transcribed by sigma F (26). Cells missing *flgB* were not able to form flagella and, therefore, were nonmotile, as observed in our motility assays (Fig. 2I). After being inoculated at approximately the same cell density (10^7 CFU/mL) as wild-type cells, the $\Delta flgB$ strain did not grow whereas wild-type cells did, which led to the wild-type strain being present in the culture at ~10- to 100-fold higher after 24 to 48 h of incubation (Fig. 2C). However, after 72 h of incubation, the wild-type cell counts decreased, and the wild-type and mutant cells had a similar cell density of between 10^6 to 10^8 CFU/mL for the rest of the incubation period (Fig. 2B). In competition during LTSP only, the $\Delta flgB$ strain was present at a similar cell density to the WT strain (Fig. 2D).

As mentioned above, *fliA* encodes sigma F, and so cells missing *fliA* should not be able to form flagella or perform chemotaxis (18), as confirmed in our motility assays (Fig. 2I). For competition between the $\Delta fliA$ strain and the wild-type strain starting in lag phase, both strains showed similar densities after 24-h of incubation (Fig. 2E). The $\Delta fliA$ mutant showed a sharp decrease after 48 to 72 h of ~100 to 1000-fold, whereas the wild-type strain only decreased by ~10-fold after 48 h and ~100-fold after 72 h. By day 6 of incubation, the wild-type and mutant cells both recovered to a cell density of $\sim 10^8$ CFU/mL and remained at similar cell

TABLE 2 Single nucleotide polymorphisms (SNPs) identified in aged populations via RNA-seq

Gene	SNP identified	% of population
Population A		
<i>hokB</i> and <i>mokB</i>	Y29C (TAC→TGC) and L48L (CTA→CTG)	68.6%
<i>argV</i>	noncoding (35/77 nt)	67.2%
<i>sspA</i>	R65C (CGT→TGT)	66.9%
<i>leuQ</i>	noncoding (48/87 nt)	55.5%
Population B		
<i>argV</i>	noncoding (35/77 nt)	65.3%
<i>sspA</i>	V25F (GTC→TTC)	79.5%
<i>cytR</i>	A257T (GCG→ACG)	85.9%

densities for the rest of the incubation period (Fig. 2E). Similarly, in competition during LTSP only, the $\Delta fliA$ strain was present at a similar cell density to the WT strain throughout (Fig. 2F).

tar encodes a methyl-accepting chemotaxis protein that senses the chemoattractant aspartate and other amino acids (25). Cells missing *tar* should still be able to move but may not be able to sense some of the amino acids that *tar* generally senses. There is some overlap with other methyl-accepting chemotaxis proteins (27). In our motility assays, strains missing the *tar* gene were more motile than WT cells (Fig. 2I). In the competition of the Δtar strain and wild-type cells starting in lag phase, both strains grew similarly throughout long-term incubation (Fig. 2G). In competition during LTSP only, in one competition the Δtar strain was at a lower cell density by 11 days of incubation, whereas in the other two competitions the Δtar and WT strains had similar cell densities throughout incubation (Fig. 2H). None of the mutant strains grew differently than wild-type strains when grown without competition from wild-type cells in monoculture (unpublished data).

Given these observations, in comparing growth phenotypes to motility phenotypes, these results suggest that our hypothesis that increased motility would allow cells to have an advantage in survival in LTSP was not supported because a gain of motility (from loss of *tar*) did not give cells an advantage, and the loss of motility (from loss of *cheA*, *fliA*, or *flgB*) did not give cells a disadvantage (Fig. 2).

Mutation analysis of RNA-seq data identified mutations potentially involved in motility. Because motility did not seem to affect the cells' ability to survive in LTSP, we hypothesized that we could identify mutations in aged populations that were isolated to one gene that had pleiotropic effects, and the increase in motility was a "hitchhiker" to other effects which were advantageous to the cell. To identify mutations that we knew were associated with the increase in motility-related genes, we identified mutations in our RNA-seq data using the program *breseq* in population mode (28).

We identified 5 unique mutations in the two aged populations, and one mutation present in both populations (Table 2). Because we used the "population" setting in *breseq*, we were able to determine what percentage of the reads, and therefore likely the population, contained each mutation. Two of the unique mutations were both mapped to the *sspA* gene, encoding the stringent starvation protein (29), which had previously been noted as a gene that is frequently mutated in long-term cultures that often confer an advantage to cells (6, 7). One *sspA* mutation (R65C) was present in 66.9% of the reads in aged population A, and the other mutation (V25F) was present in 79.5% of the reads in aged population B. SspA regulates transcription through both regulation of H-NS (30) and directly via RNA polymerase (31), and R65 is an essential amino acid for its interactions with RNAP (31). Loss of function of SspA may be responsible for regulating motility via the DNA-binding regulatory protein H-NS. SspA negatively regulates H-NS, which in turn negatively regulates the expression of *csgD* (29). Loss of function of SspA may lead to the higher activity of H-NS, which would then lead to lower expression of *csgD*, as we saw in our transcriptome data (Table 1). CsgD is a negative regulator of *fliA* (18). Thus, in these cells, the repression of *csgD* might lead to a derepression of *fliA*, and, therefore, other motility- and chemotaxis-related genes. Further, H-NS is a positive regulator of *fliA* (along with

other flagellar-related genes) (32), so higher activity of H-NS could also lead to a direct increase in *flaA* expression.

There was also a mutation mapped to *argV* in both populations, encoding an arginine tRNA (ACG) (33). Because this exact mutation was present in both aged populations, there was a chance it occurred during the initial overnight growth, or it could be due to molecular convergence leading to selection for this allele.

In population A, we also identified mutations that mapped to the *hokB* and *mokB* genes (68.6% of reads), which overlap in the genome. HokB is part of the HokB-SokB type 1 toxin-antitoxin system (34). MokB is a putative regulator of HokB (34). However, the base change which leads to a missense mutation in HokB leads to a silent mutation in MokB. We also identified a mutation in *leuQ* (55.5% of reads), another gene encoding a tRNA, this one for leucine (CAG) (35).

In population B, the only other mutation identified was mapped to *cytR* (85.9% of reads), a regulator of nucleoside metabolism (19). Like *sspA*, *cytR* frequently accumulates mutations in cells incubated to LTSP, and *sspA* and *cytR* mutations often occur in the same populations (6, 7).

DISCUSSION

We identified multiple genes that were differentially expressed during log phase in populations of *E. coli* that have been incubated into LTSP compared to their unaged parental strain. Several genes involved in motility and chemotaxis had higher expression in cells that had been aged into LTSP (Table 1), leading to increased motility. The increase in motility-related genes or the motility phenotype occurred in multiple aged cultures, suggesting that the phenomenon of increased motility may be common in cells adapted to LTSP. Because the RNA-seq data showed overexpression in motility-related genes and aged populations were more motile, we hypothesized that increased motility would be advantageous in LTSP. We would not have hypothesized this *a priori*. Because these cultures were shaken in tubes, we originally thought cells would not need their motile or chemotactic abilities to access nutrients or oxygen. Contrary to our predictions, based on single-gene deletion mutations (Fig. 2) we determined that motility does not correlate with survival in LTSP.

However, we did observe a surprising result when *cheA* was deleted. These cells could outcompete wild-type cells beginning in log phase and throughout the *E. coli* life cycle (Fig. 2A). When CheA was inhibited, the cells performed little to no clockwise movement, which meant that the cells would mostly be engaging in counterclockwise movement and long smooth runs (25). Two hypotheses could explain this result. First, the long smooth runs could be giving an advantage to the mutant cells over the wild-type cells, which would be performing a normal run and tumble pattern. Long smooth runs would allow cells to steadily move toward chemoattractants. However, because the cells were grown and competed in test tubes in a shaking incubator, it seemed unlikely that the long runs would give an advantage in that environment because cells were being moved around the tubes likely without needing to use their flagella. A second hypothesis is that the $\Delta cheA$ mutant was conserving energy by not having to switch between the run and tumble movement. The conservation of energy could provide an advantage to the cells.

Because loss of motility did not lead to a loss of ability to survive in LTSP, we then hypothesized that the increase in motility may be a pleiotropic effect of a mutation in a regulatory gene. We thus identified multiple mutations in the aged populations using the RNA-seq data. Unique mutations in the gene *sspA* were identified in both aged populations (Table 2), which may lead to an increase in motility via increased activity of H-NS. Along with explaining the increase in motility-related genes, the increased H-NS activity could also explain the decreased expression of acid-resistance genes in aged populations (Table 1) because it is a negative regulator of many of those genes (36). We hypothesize that this loss or decrease in function of SspA leads to the expression differences and phenotypes we observed in the aged populations, and ongoing testing will determine how this change allows cells to survive in LTSP.

We also identified mutations in *cytR* (Table 2), which is a negative regulator of other genes that were more highly expressed in aged cells: *nupG*, *tsx*, *cdd*, and *udp* (Table 1) (19).

The mutations in *cytR* and other identified mutations likely did not affect the motility phenotype we observed. However, these data indicate that identifying mutations in RNA-seq data can be useful in explaining expression changes in aged populations. Understanding which mutations occur in long-term cultures, and if and why they were advantageous, gives insight into how cells adapt to this stressful condition.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The *E. coli* K-12 lineage MG1655-derivative PFM2 (37) and a chloramphenicol resistant (PFM2-CM) derivative of PFM2 (7) were used in this study. To start overnight cultures of cells, frozen stocks of cells stored in a -80°C freezer in 5 mL of Difco Luria-Bertani Broth (LB; BD Difco) were inoculated and incubated overnight (16 to 18 h) at 37°C in a 150 mm $\times$ 18 mm test tube in a shaking incubator at 200 RPM. These frozen stocks consist of a population of cells originally started from a single clone. A 1:1000 dilution was then inoculated into fresh 5 mL of LB and incubated as described above. The long-term cultures were incubated under these conditions for 10 days with no addition of nutrients. After 10 days, they were stored in a 10% glycerol solution in LB at -80°C .

Mutant strains were created for competition assays against wild-type nonmutant (PFM2) cells using bacteriophage P1 transduction. Strains in the KEIO collection, which is a collection of single-gene knock-out mutants with a kanamycin resistance cassette in the open reading frame coding regions (38), were used as donors to generate the desired P1 lysates to transfer mutant genotypes. P1 transduction protocol as described previously (15) was used.

RNA preparation, sequencing, and analysis. Cultures of wild-type and aged cells were started as described above, incubated overnight, and diluted at 1:1000 the next day. Two cultures of the wild-type strain and the aged strain were grown for 4 h into the log phase. RNeasy Protect Bacteria Minikit (Qiagen) was added per the manufacturer's instructions. Cells were pelleted and frozen at -80°C until total cellular RNA extraction was performed using the RNeasy Protect Bacteria Minikit (Qiagen). Total RNA was processed, assessed for quality, and sequenced by the University of Southern California Genomics Core, where rRNA was removed using the Ribo-zero kit (Illumina, Inc.). Seventy-five base pair single-end reads were generated from each sample on the Illumina HiSeq 2500 platform. We received an average of 8.8 million reads per sample, with a range of ~ 7.5 million to 10.4 million reads per sample. Using a custom Python pipeline, we used TopHat2 to align the raw reads to the *E. coli* MG1655 genome (39), SAMtools to generate binary alignment files (BAM) (40), and HTSeq to calculate read counts per gene and per sample (41). In all samples, 96.5% of reads mapped to our reference genome (*Escherichia coli* K-12 MG1655). All rRNA gene reads and any gene that had no expression were removed from the data set. We used the HTSeq output files as input to analyze gene expression levels using DESeq2 (42), with comparisons between unaged and aged populations. We considered genes to be differentially expressed between these populations if the $\log_2$ fold change was greater than or equal to 2, and if the *q*-value was less than or equal to 0.05.

Motility assays. Motility assays were used to assess the movement differences between the samples. Swimming motility assays were performed using 0.18% semisolid agar (0.54g Bacto Agar [BD Difco™ Bacto™], 6.0g LB broth powder [BD Difco™], and 300 mL MilliQ water). The semi solid agar was autoclaved for 15 min at 121°C . 17.6 mL of the semisolid agar was poured into 100 mm $\times$ 15 mm plastic Petri plates creating 2.2 mm thick semisolid agar in each plate. Overnight cultures of the wild-type (unaged) cells along with overnight cultures of the aged or mutant populations were inoculated in liquid media as described above. The cells were then further diluted in a 1:1000 ratio in 5 mL of LB and placed in the 37°C shaking incubator for 4 h. After the incubation, the center of the semisolid agar plates was inoculated with cells from each culture using an inoculation loop without touching the bottom of the plate. The plates were then placed in a 37°C incubator for 12 h. The zone of motility was recorded by measuring the diameter of the halo or the distance the cells traveled from the center inoculation point. A Welch's or Student's *t* test was used to determine the significance of the difference between the aged and unaged populations and wild-type and mutant populations using the ggpubr package in R.

Growth assays. Competition assays were performed in liquid culture using 5 mL of LB broth. Competitions were completed in two ways: first, cells were competed starting in lag phase using cultures that were inoculated with a 1:1000 dilution of wild-type cells with a chloramphenicol-resistance marker (PFM2-CM) that were incubated as described above, and a 1:1000 dilution of mutant cells ($\Delta fliA$, $\Delta flgB$, Δtar , or $\Delta cheA$) with a kanamycin-resistance marker that was incubated as described above. The cells were placed in the shaking incubator at 37°C and allowed to compete for resources for 11 days. Second, cells were competed in LTSP only by first incubating cells in monoculture for 4 days into LTSP, and then mixing them in a 1:1 ratio (2.5 mL each) into a new tube. Serial dilutions were performed to measure the cell densities of both strains on days 0, 1, 2, 3, 6, and 11 of incubation (6). The dilutions were plated on chloramphenicol LB agar plates and kanamycin LB agar plates to distinguish between the cell types, with wild-type cells growing on chloramphenicol plates and mutant cells growing on kanamycin plates.

Monocultures of the wild-type cells (PFM2) and the $\Delta fliA$, $\Delta flgB$, Δtar , and $\Delta cheA$ mutants were also incubated as described above, with each strain grown in its own culture. The cells were plated on LB agar plates with no antibiotic.

Mutation analysis. The RNA-seq Illumina reads were aligned to the *E. coli* K-12 (strain MG1655) genome (NCBI reference sequence [NC_000913.3](https://www.ncbi.nlm.nih.gov/assembly/GCF_000013325.1)) using the population analysis settings with the BreSeq package in Python (28).

Data availability. Both count files derived from HTseq and raw sequence reads have been deposited in NCBI's Gene Expression Omnibus (41) and are accessible through GEO Series accession number [GSE188790](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE188790).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

We thank Steven Finkel, Sonal Singhal, Kathryn Theiss, Alejandra Verdugo-Rodriguez, Lacey Westphal, and members of the Kram lab for helpful discussion of the project and manuscript.

This work was supported in part by the National Science Foundation RUI number 1715066 (to K.E.K.). K.E.K. was also partially supported by the CSUDH Research, Scholarship, and Creative Activity Award and the CSUDH Norris Foundation Award. The funders had no role in study design, data collection, interpretation, or the decision to submit the work for publication.

We declare no conflict of interest.

REFERENCES

- Flatt T, Heyland A. 2011. Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs. Oxford University Press, Oxford, UK.
- Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B: Biological Sciences* 255:37–45. <https://doi.org/10.1098/rspb.1994.0006>.
- Blount ZD. 2015. The unexhausted potential of *E. coli*. *Elife* 4:e05826. <https://doi.org/10.7554/eLife.05826>.
- Finkel SE. 2006. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* 4:113–120. <https://doi.org/10.1038/nrmicro1340>.
- Farrell MJ, Finkel SE. 2003. The growth advantage in stationary-phase phenotype conferred by *rpoS* mutations is dependent on the pH and nutrient environment. *J Bacteriol* 185:7044–7052. <https://doi.org/10.1128/JB.185.24.7044-7052.2003>.
- Kram KE, Geiger C, Ismail WM, Lee H, Tang H, Foster PL, Finkel SE. 2017. Adaptation of *Escherichia coli* to long-term serial passage in complex medium: evidence of parallel evolution. *mSystems* 2:e00192-16. <https://doi.org/10.1128/mSystems.00192-16>.
- Westphal LL, Lau J, Negro Z, Moreno IJ, Ismail Mohammed W, Lee H, Tang H, Finkel SE, Kram KE. 2018. Adaptation of *Escherichia coli* to long-term batch culture in various rich media. *Res Microbiol* 169:145–156. <https://doi.org/10.1016/j.resmic.2018.01.003>.
- Chib S, Ali F, Seshasayee ASN. 2017. Genomewide mutational diversity in *Escherichia coli* population evolving in prolonged stationary phase. *mSphere* 2:e00059-17. <https://doi.org/10.1128/mSphere.00059-17>.
- Ali F, Seshasayee ASN. 2020. Dynamics of genetic variation in transcription factors and its implications for the evolution of regulatory networks in bacteria. *Nucleic Acids Res* 48:4100–4114. <https://doi.org/10.1093/nar/gkaa162>.
- Ratib NR, Seidl F, Ehrenreich IM, Finkel SE. 2021. Evolution in long-term stationary-phase batch culture: emergence of divergent *Escherichia coli* lineages over 1,200 Days. *mBio* 12:e03337-20. <https://doi.org/10.1128/mBio.03337-20>.
- Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS. 2012. The molecular diversity of adaptive convergence. *Science* 335:457–461. <https://doi.org/10.1126/science.1212986>.
- Conrad TM, Frazier M, Joyce AR, Cho B-K, Knight EM, Lewis NE, Landick R, Palsson BØ. 2010. RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc Natl Acad Sci U S A* 107:20500–20505. <https://doi.org/10.1073/pnas.0911253107>.
- Madan Babu M, Teichmann SA. 2003. Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucleic Acids Res* 31:1234–1244. <https://doi.org/10.1093/nar/gkg210>.
- Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63. <https://doi.org/10.1038/nrg2484>.
- Kram KE, Henderson AL, Finkel SE. 2020. *Escherichia coli* has a unique transcriptional program in long-term stationary phase allowing identification of genes important for survival. *mSystems* 5:e00364-20. <https://doi.org/10.1128/mSystems.00364-20>.
- Baremburch C, Hengge R. 2007. Cellular levels and activity of the flagellar sigma factor FlhA of *Escherichia coli* are controlled by FlgM-modulated proteolysis. *Mol Microbiol* 65:76–89. <https://doi.org/10.1111/j.1365-2958.2007.05770.x>.
- Clarke MB, Sperandio V. 2005. Transcriptional regulation of *flhDC* by QseBC and σ^{28} (FlhA) in enterohaemorrhagic *Escherichia coli*. *Mol Microbiol* 57:1734–1749. <https://doi.org/10.1111/j.1365-2958.2005.04792.x>.
- Dudin O, Geiselmann J, Ogasawara H, Ishihama A, Lacour S. 2014. Repression of flagellar genes in exponential phase by CsgD and CpxR, two crucial modulators of *Escherichia coli* biofilm formation. *J Bacteriol* 196:707–715. <https://doi.org/10.1128/JB.00938-13>.
- Sernova NV, Gelfand MS. 2012. Comparative genomics of CytR, an unusual member of the LacI family of transcription factors. *PLoS One* 7:e44194. <https://doi.org/10.1371/journal.pone.0044194>.
- Peekhaus N, Tong S, Reizer J, Saier MH, Murray E, Conway T. 1997. Characterization of a novel transporter family that includes multiple *Escherichia coli* gluconate transporters and their homologues. *FEMS Microbiol Lett* 147:233–238. <https://doi.org/10.1111/j.1574-6968.1997.tb10247.x>.
- Cowan S, Garavito R, Jansonius J, Jenkins J, Karlsson R, König N, Pai E, Paupit R, Rizkallah P, Rosenbusch J, Rummel G, Schirmer T. 1995. The structure of OmpF porin in a tetragonal crystal form. *Structure* 3:1041–1050. [https://doi.org/10.1016/S0969-2126\(01\)00240-4](https://doi.org/10.1016/S0969-2126(01)00240-4).
- Dubini A, Pye RL, Jack RL, Palmer T, Sargent F. 2002. How bacteria get energy from hydrogen: a genetic analysis of periplasmic hydrogen oxidation in *Escherichia coli*. *Int J Hydrogen Energy* 27:1413–1420. [https://doi.org/10.1016/S0360-3199\(02\)00112-X](https://doi.org/10.1016/S0360-3199(02)00112-X).
- Zhang Y, Xiao M, Horiyama T, Zhang Y, Li X, Nishino K, Yan A. 2011. The multidrug efflux pump MdtEF protects against nitrosative damage during the anaerobic respiration in *Escherichia coli*. *J Biol Chem* 286:26576–26584. <https://doi.org/10.1074/jbc.M111.243261>.
- Xavier KB, Bassler BL. 2005. Regulation of uptake and processing of the quorum-sensing autoinducer AI-2 in *Escherichia coli*. *J Bacteriol* 187:238–248. <https://doi.org/10.1128/JB.187.1.238-248.2005>.
- Hansen CH, Endres RG, Wingreen NS. 2008. Chemotaxis in *Escherichia coli*: a molecular model for robust precise adaptation. *PLoS Comput Biol* 4:e1. <https://doi.org/10.1371/journal.pcbi.0040001>.
- Homma M, Kutsukake K, Hasebe M, Iino T, Macnab RM. 1990. FlgB, FlgC, FlgF and FlgG: a family of structurally related proteins in the flagellar basal body of *Salmonella typhimurium*. *J Mol Biol* 211:465–477. [https://doi.org/10.1016/0022-2836\(90\)90365-S](https://doi.org/10.1016/0022-2836(90)90365-S).
- Wolff C, Parkinson JS. 1988. Aspartate taxis mutants of the *Escherichia coli* tar chemoreceptor. *J Bacteriol* 170:4509–4515. <https://doi.org/10.1128/jb.170.10.4509-4515.1988>.
- Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol* 1151:165–188. https://doi.org/10.1007/978-1-4939-0554-6_12.
- Hansen A-M, Jin DJ. 2012. SspA up-regulates gene expression of the LEE pathogenicity island by decreasing H-NS levels in enterohemorrhagic *Escherichia coli*. *BMC Microbiol* 12:231. <https://doi.org/10.1186/1471-2180-12-231>.
- Hansen A-M, Qiu Y, Yeh N, Blattner FR, Durfee T, Jin DJ. 2005. SspA is required for acid resistance in stationary phase by downregulation of H-NS in *Escherichia coli*: SspA and acid resistance in stationary phase. *Mol Microbiol* 56:719–734. <https://doi.org/10.1111/j.1365-2958.2005.04567.x>.

31. Wang F, Shi J, He D, Tong B, Zhang C, Wen A, Zhang Y, Feng Y, Lin W. 2020. Structural basis for transcription inhibition by *E. coli* SspA. *Nucleic Acids Res* 48:9931–9942. <https://doi.org/10.1093/nar/gkaa672>.
32. Bertin P, Terao E, Lee EH, Lejeune P, Colson C, Danchin A, Collatz E. 1994. The H-NS protein is involved in the biogenesis of flagella in *Escherichia coli*. *J Bacteriol* 176:5537–5540. <https://doi.org/10.1128/jb.176.17.5537-5540.1994>.
33. Neidhardt FC, Curtiss R, III, Ingraham JL, Lin ECC, Low KB, Jr, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umberger HE. 1996. *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, Second Edition. American Society for Microbiology.
34. Kawano M. 2012. Divergently overlapping *cis*-encoded antisense RNA regulating toxin-antitoxin systems from *E. coli*: *hok/sok*, *ldr/rdl*, *symE/symR*. *RNA Biol* 9:1520–1527. <https://doi.org/10.4161/ma.22757>.
35. Mohanty BK, Kushner SR. 2007. Ribonuclease P processes polycistronic tRNA transcripts in *Escherichia coli* independent of ribonuclease E. *Nucleic Acids Res* 35:7614–7625. <https://doi.org/10.1093/nar/gkm917>.
36. Giangrossi M, Zattoni S, Tramonti A, De Biase D, Falconi M. 2005. Antagonistic role of H-NS and GadX in the regulation of the glutamate decarboxylase-dependent acid resistance system in *Escherichia coli*. *J Biol Chem* 280: 21498–21505. <https://doi.org/10.1074/jbc.M413255200>.
37. Lee H, Popodi E, Tang H, Foster PL. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci U S A* 109:E2774–E2783.
38. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008. <https://doi.org/10.1038/msb4100050>.
39. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36. <https://doi.org/10.1186/gb-2013-14-4-r36>.
40. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
41. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169. <https://doi.org/10.1093/bioinformatics/btu638>.
42. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.