Bacterial Evolution Through the Selective Loss of Beneficial Genes: Trade-Offs in Expression Involving Two Loci

Erik R. Zinser,*[,1] Dominique Schneider,[†] Michel Blot[†,2] and Roberto Kolter*[,3]

**Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115 and [†]Laboratoire Plasticité et Expression de Génomes Microbiens CNRS FRE2383, Université Joseph Fourier, 38041 Grenoble Cedex 9, France*

Manuscript received December 14, 2002
Accepted for publication March 26, 2003

ABSTRACT

The loss of preexisting genes or gene activities during evolution is a major mechanism of ecological specialization. Evolutionary processes that can account for gene loss or inactivation have so far been restricted to one of two mechanisms: direct selection for the loss of gene activities that are disadvantageous under the conditions of selection (*i.e.*, antagonistic pleiotropy) and selection-independent genetic drift of neutral (or nearly neutral) mutations (*i.e.*, mutation accumulation). In this study we demonstrate with an evolved strain of *Escherichia coli* that a third, distinct mechanism exists by which gene activities can be lost. This selection-dependent mechanism involves the expropriation of one gene's upstream regulatory element by a second gene via a homologous recombination event. Resulting from this genetic exchange is the activation of the second gene and a concomitant inactivation of the first gene. This gene-for-gene expression tradeoff provides a net fitness gain, even if the forfeited activity of the first gene can play a positive role in fitness under the conditions of selection.

MICROBIAL genomes are in a constant state of flux, whereby new genes are acquired by horizontal transfer and preexisting genes are lost by mutation (Lawrence 1999; Lawrence and Roth 1999). It is these changes in the repertoire of genes that are thought to be the primary mechanisms of prokaryotic adaptation that lead to speciation. Extensive evidence for the continual expansion and retraction of genomes has been discovered among the species whose whole genomes have been sequenced (Andersson *et al.* 1998; Alm *et al.* 1999; Cole *et al.* 2001; Perna *et al.* 2001). Horizontally acquired genetic units range from regions within a single gene to entire gene "islands," often of coordinated function (Lawrence and Roth 1999). Loss of genetic material over the course of the natural history of microbes is also readily apparent. Both large-scale losses via deletion and gene inactivations by revertable point mutations (Morishita *et al.* 1981; Andersson *et al.* 1998; Maurelli *et al.* 1998; Cole *et al.* 2001) have been observed in modern microbes.

Genes and gene activities are generally thought to be lost from populations over the course of evolution by one of two mechanisms: (i) selection against deleterious genes and (ii) selection-independent genetic drift of neutral or nearly neutral genes (Lawrence and Roth 1999; Cooper and Lenski 2000). Adaptive losses of gene activities have been observed in recent studies on the virulence of Shigella. Whereas nonpathogenic *Escherichia coli* has both *ompT* and *cadA*, Shigella has lost both of these genes, whose presence attenuates the virulence of the bacterium (Nakata *et al.* 1993; Maurelli *et al.* 1998). Adaptive mutation in one environment can often result in lower fitness values for other environmental conditions. This is known as antagonistic pleiotropy and has been observed in a number of experimentally studied processes, including senescence in fruit flies and catabolic decay in serially transferred glucose-grown *E. coli* (Rose and Charlesworth 1980; Cooper and Lenski 2000).

Independent of adaptive mutations, genes can also be lost from a population by the fixation of neutral or nearly neutral mutations via selection-independent genetic drift. This mechanism, termed mutation accumulation (Hughes and Charlesworth 1994; Cooper and Lenski 2000), is thought to account for much of the loss of genes in prokaryotes: it is impossible to protect every gene against mutation and loss by genetic drift if their disappearance does not result in a severe loss in fitness (Lawrence and Roth 1999).

We have investigated the processes of microbial evolution under conditions of prolonged starvation. While the majority of the population quickly dies out during starvation, in bacteria such as *E. coli* a significant minority remain viable for months and even years of prolonged starvation (Finkel *et al.* 1997; Finkel and Kolter 1999). Prolonged starvation in these cultures is marked by rapid, continuous evolution among the surviving minor-

[1]*Present address:* Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

[2]Deceased September 2002.

[3]*Corresponding author:* Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Ave., Boston, MA 02115. E-mail: rkolter@hms.harvard.edu

ity (Zambrano *et al.* 1993; Finkel and Kolter 1999). One adapted mutant of *E. coli* isolated from a starved culture was found to have acquired four *g*rowth *a*dvantage in *s*tationary *p*hase (GASP) mutations (Zambrano and Kolter 1993; Zambrano *et al.* 1993; Zinser and Kolter 1999). These GASP mutations confer the ability of a minority population to grow and invade a wild-type majority population during the course of culture starvation. Two of these GASP mutations were identified as mutant alleles of the global gene regulators, *rpoS* and *lrp* (Zambrano *et al.* 1993; Zinser and Kolter 2000).

In this report we identify a third adaptive GASP mutation (*sgaA*) in this mutant as a genomic rearrangement, which concomitantly activates one locus while inactivating another. Through an insertion of an IS*5* element followed by an inversion between it and a preexisting IS*5* element, the initially inactive locus effectively expropriated the regulatory element of the initially active locus. This gene-for-gene expression trade-off is particularly interesting because the inactivated locus plays a positive role in fitness under the conditions in which it was inactivated. The concept of such an expropriation event as a novel evolutionary mechanism of gene loss is then discussed.

MATERIALS AND METHODS

Fitness assays: All experiments were performed at 37°. Stationary-phase competitions were performed as described (Zinser and Kolter 1999). Briefly, competing strains were marked with either of two selectively neutral chromosomal markers: growth on salicin as a sole carbon source (Bgl^+) or valine-resistant growth on glucose (Val^r). To rule out marker effects on fitness, the Bgl^+ and Val^r markers were switched between competing strains in half of the pairwise competitions. Strains were subcultured 1:100 into fresh Luria broth (LB) and incubated for 24 hr prior to mixing to facilitate entry into stationary phase. Cultures were mixed 1:1000 (v:v), and the two populations in the mixed culture were monitored by serial dilution in M63 medium and plating on minimal salicin (0.2%) and minimal glucose (0.2%) plus valine (0.02%) plates. The invasion index, as a measure of relative fitness, was defined as the change in the minority to total population ratio between days 1 and 3 of the stationary-phase competition.

Genetic techniques: Insertional mutageneses with the mini-Tn*10* Cm^r and Tn*phoA′-1* (*lacZ*) transposons were performed as described (Kleckner *et al.* 1991; Wilmes-Riesenberg and Wanner 1992). Phage P1*vir* transduction was performed as described elsewhere (Miller 1992). The Tn*phoA′*-1 insertion alleles were moved into the insertion-/inversion-less parental background by P1 transduction, selecting for kanamycin resistance and then confirming the absence of the insertion/inversion in the strain by PCR.

Restriction fragment length polymorphism analysis of the GASP mutants: Restriction fragment length polymorphism (RFLP) analyses (Papadopoulos *et al.* 1999) using eight known *E. coli* K-12 insertion sequence (IS) elements as probes were performed with *Eco*RV-digested DNA from ZK819 (wild type) and ZK1141 (*sgaA*; Zambrano and Kolter 1993). Among several IS-associated mutations detected in ZK1141, the loss of one IS*5*-containing fragment (a 3.9-kb fragment, called A) and the gain of two other IS*5*-containing fragments (3- and 6-kb fragments, called B and C, respectively) were shown to involve a genomic region (14′ on the chromosome) known to include the *sgaA* GASP mutation (Zinser and Kolter 1999). Sequences adjacent to IS*5* fragments in A (using genomic DNA from ZK819), B, and C (using genomic DNA from ZK1141) were cloned after inverse PCR (Schneider *et al.* 2000). Sequence comparisons of the IS*5* junctions with the *E. coli* genome sequences suggested a complex rearrangement—a transposition event with an associated inversion. A new IS*5* copy inserted upstream of the *cstA* gene at 13.6′, between its promoter and a CRP-binding site. A recombination event between this new IS*5* copy and IS*5*D (located at 14.8′ between *ybeJ* and *cute*; Blattner *et al.* 1997) led to the inversion of the intervening sequence. This inversion was confirmed by hybridization experiments, using all adjacent sequences of IS*5* in fragments A, B, and C as probes (data not shown). This rearrangement explains the loss of fragment A and the gain of fragments B and C in ZK1141.

Construction of the *cstA*::*lacZ* fusions: We analyzed transcriptional activities of plasmid-borne *cstA*::*lacZ* constructs, as similar qualitative results with plasmid and chromosomal *cstA* transcriptional fusions have been reported (Schultz and Matin 1991). We constructed plasmid fusions by cloning the promoter regions of the wild-type parent and the inversion mutant onto the pRS550 plasmid (Simons *et al.* 1987). The promoter of *cstA* in ZK819 together with the CRP-binding site (prior to the inversion) was cloned into pRS550 by PCR, leading to plasmid pGBE153. The promoter of *cstA* in ZK1141 (after the inversion) was cloned into pRS550 by PCR. The resulting plasmid, pGBE155, carries the *cstA* promoter as well as the upstream IS*5* element. The sequence of the primers was designed to introduce suitable *Bam*HI and *Eco*RI restriction sites, allowing the easy cloning of the PCR products into pRS550. The two plasmid constructs were introduced into ZK126 by standard techniques.

Construction of the Δ*cstA* mutants: A *cstA* deletion was constructed in the ZK819 background. The *cstA* gene was cloned by PCR using the PCR-Script Cam cloning kit (Stratagene, La Jolla, CA). A deletion within *cstA* was created by cutting the resulting plasmid with *Mlu*I, which cuts twice within *cstA*, and religating the large fragment. The new plasmid contains an in-frame 1275-bp deletion within *cstA*. The insert was isolated as a *Sac*I-*Sal*I fragment and cloned into the suicide-plasmid pCVD442 cut with *Sac*I-*Sal*I (Donnenberg and Kaper 1991). The in-frame deletion mutant allele of *cstA* was introduced, as described (Donnenberg and Kaper 1991), in the chromosome of ZK819, generating the Δ*cstA* derivative, GBE111. The presence of the deletion was confirmed by PCR and hybridization experiments (data not shown). The Δ*cstA* derivatives of the ZK2552 and ZK2553 strains were constructed by P1 transduction. First, the *lipA150*::Tn*1000*d allele was transduced with P1 phage from KER176 into ZK2552 or ZK2553 (Zinser and Kolter 1999), selecting for kanamycin resistance. Then the Δ*cstA* allele was transduced from GBE111 into the *lipA150*::Tn*1000*d transductants, selecting for lipoic acid prototrophs. Transductants carrying the Δ*cstA* allele were identified by PCR using the primers that flanked the deleted region.

RESULTS

The *sgaA* GASP mutation is a genomic rearrangement involving insertion sequence elements: Previous work identified two of the four adaptive mutations of a survivor of an aged culture of *E. coli* as loss-of-function alleles of the global regulators, *rpoS* and *lrp* (Zambrano *et al.*

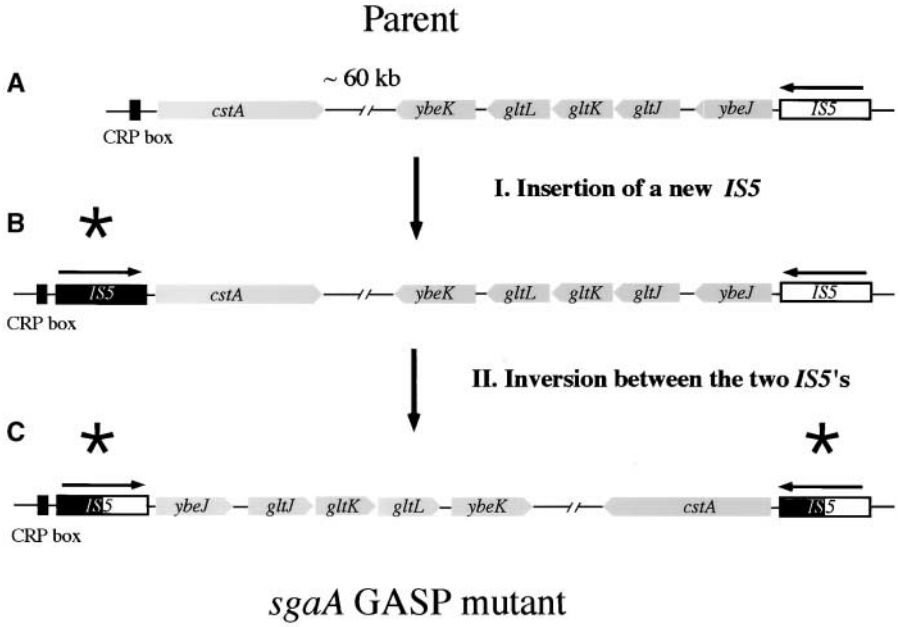

FIGURE 1.—The *sgaA* GASP mutation is a two-step genomic rearrangement involving two IS*5* elements. Asterisks denote the sites of these genomic rearrangements. (A) In the wild-type parent, the putative *ybeJ-gltJKL-ybeK* operon is ~60 kb away from the *cstA* locus. The two loci are transcribed in opposite orientation. (B) The first rearrangement was an insertion of a new IS*5* between the transcriptional promoter and the upstream CRP box of *cstA*. (C) The second rearrangement was an inversion between this new IS*5* and the IS*5* (IS*5*D) upstream of the putative *ybeJ-gltJKL-ybeK* operon. The exact point of crossover between the IS*5* elements is unknown. Arrows above the IS*5* elements indicate the direction of the *ins5A* transcript.

1993; ZINSER and KOLTER 2000). All four of the GASP mutations in this mutant confer growth phenotypes in addition to GASP. One of the two unidentified GASP mutations, *sgaA*, confers faster growth on glutamate, asparagine, and proline as carbon sources and confers the new ability to grow on aspartate as a sole carbon source (ZINSER and KOLTER 1999). It was hypothesized that these growth phenotypes allow the *sgaA* mutant to outcompete the wild-type parent for amino acids as nutrients released by the dying majority population during prolonged starvation (ZINSER and KOLTER 1999). To identify the *sgaA* GASP locus, an insertional mutagenesis strategy was employed. Mutants of the *sgaA* strain with random mini-Tn*10* insertions were screened initially for a loss of growth on aspartate, followed by a secondary screen for loss of the GASP phenotype relative to its *sgaA*$^+$ parent. An insertion in the *gltJ* locus eliminated both the aspartate growth and GASP phenotypes of the *sgaA* strain (data not shown).

The *gltJ* gene is a member of a putative five-gene operon (Figure 1A; BLATTNER *et al.* 1997), the first four members of which are thought to encode the components of the known aspartate/glutamate binding protein-dependent ABC-type transporter of *E. coli* (WILLIS and FURLONG 1975; D. LUM and B. J. WALLACE, unpublished data). The predicted YbeJ product is most similar to known periplasmic binding protein components, the GltJ and GltK products are most similar to known integral membrane protein components, and the GltL product is most similar to known ATPases. The predicted product of *ybeK*, the final gene of the putative operon, shares closest homology to known nucleoside hydrolases (alignments not shown).

RFLP and sequence analyses identified the *sgaA* GASP mutation as a complex two-step recombination event involving an IS*5* element (designated IS*5*D in the published *E. coli* genome; BLATTNER *et al.* 1997) present in the wild-type parent 103 bp upstream of the start of the *ybeJ* open reading frame (Figure 1A). The first event was an insertion of a new IS*5* element at a position 60 kb away from IS*5D*, between the transcriptional promoter and the upstream CRP-binding site (CRP box) of the *cstA* gene, which is thought to encode an oligopeptide permease (SCHULTZ and MATIN 1991; Figure 1B). The second event was a homologous recombination event between the new IS*5* and IS*5*D, resulting in a chromosomal inversion (Figure 1C). This rearrangement effectively placed the upstream region, including the CRP box, of *cstA* upstream of the *ybeJ-gltJKL-ybeK* operon. This two-step recombination event, termed IN(*cstA*::IS*5*-IS*5*D), was the only mutation detected between *ybeJ* and the CRP box in the *sgaA* mutant (data not shown).

The *sgaA* insertion/inversion event activates one locus while inactivating another: Given that the *ybeJ-gltJKL-ybeK* operon encodes a putative amino acid transporter, it was hypothesized that the *sgaA* genomic rearrangement enhanced the growth rate on amino acids by increasing expression of this operon, thereby increasing the transport capacity of the cell. Analysis of chromosomal *lacZ* transcriptional fusions within the *ybeJ-gltJKL-ybeK* operon, constructed by transposon mutagenesis (WILMES-RIESENBERG and WANNER 1992), confirmed this prediction. Expression of the *ybeJ*::*lacZ* fusion was significantly higher at all cell densities in the *sgaA* strain compared to the *sgaA*$^+$ parent (Figure 2A). Fusions to *gltJ* and *gltK* behaved similarly to the *ybeJ* fusion (data not shown), suggesting that all three genes are members of an operon.

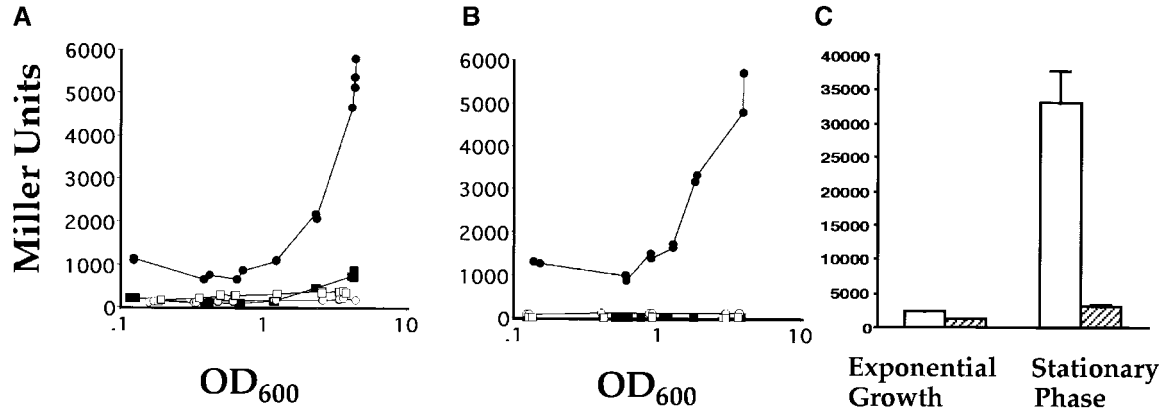

FIGURE 2.—Transcriptional activities of genes on either side of the *sgaA* insertion/inversion. (A and B) β-Galactosidase activities (MILLER 1992) of mutants with chromosomal *lacZ* insertions in the *ybeJ* (A) or IS*5* (B) loci. The genetic backgrounds were *sgaA*$^+$ (■), *sgaA* (●), *sgaA*$^+$ Δ*crp-45* (□), and *sgaA* Δ*crp-45* (○). Each point is the average of duplicate assays performed on an LB-grown culture, inoculated 1/100 from a stationary-phase culture. (C) β-Galactosidase activities of the ZK126 (wild-type) strain with plasmid-borne *lacZ* fusions to the upstream regulatory elements of the *cstA* allele of the *sgaA*$^+$ (pGBE153, open bars) or *sgaA* (pGBE155, hatched bars) strains. Cultures were assayed during midexponential growth and stationary phase in LB broth. The values are means ±95% confidence intervals of three independent experiments.

Significantly, the expression pattern of *ybeJ* in the *sgaA* strain was essentially the same as *cstA* in the wild type (SCHULTZ and MATIN 1991). Expression greatly increased when the culture entered stationary phase (Figure 2A). Expression of *ybeJ* was also found to be CRP dependent, as it was eliminated when the Δ*crp-45* allele was introduced (Figure 2A). The loss of *ybeJ* expression in the Δ*crp-45* strains was due only to the absence of CRP protein, as normal induction was observed after introduction of a wild-type *crp* allele in a pBR322 plasmid (data not shown). These results indicate that the regulatory components of the region upstream of *cstA* were effectively transferred and are indeed functional upstream of the *ybeJ-gltJKL-ybeK* operon in the *sgaA* mutant.

Interestingly, a transposon insertion within the IS*5* element upstream of the *ybeJ-gltJKL-ybeK* operon inactivated the operon, as seen by a concomitant loss of the GASP phenotype and the amino acid growth phenotypes in the *sgaA* background. Expression from the IS*5*::*lacZ* fusion, oriented in the same direction as the *ybeJ-gltJKL-ybeK* operon, demonstrated that expression of the *ins5A* transposase gene within the element (KROGER and HOBOM 1982) is also dependent on the insertion/inversion mutation, growth phase, and CRP (Figure 2B). CRP therefore appears to activate the *ins5A* promoter, p_{5I} (KROGER and HOBOM 1982), by binding at the CRP box, located 73.5 bp upstream, which is near the optimal distance of 61.5 bp for CRP activation (BUSBY and EBRIGHT 1999).

Concomitant to the activation of the *ybeJ-gltJKL-ybeK* operon by the expropriation of the regulatory region of *cstA* was the inactivation of the *cstA* locus. We analyzed transcriptional activities of plasmid-borne *cstA*::*lacZ* constructs. Expression from the *cstA* promoter of the *sgaA* mutant was significantly lower than that of the wild-type parent, in both exponential and stationary-phase cultures (Figure 2C), indicating a severe decrease in *cstA* activity for the *sgaA* strain. The overall consequence of the insertion/inversion event was therefore to activate the *ybeJ-gltJKL-ybeK* operon while simultaneously inactivating the *cstA* locus.

The *cstA* locus plays a beneficial role during stationary phase: The *cstA* gene encodes a starvation-inducible oligopeptide permease (SCHULTZ and MATIN 1991) and may therefore play a role in stationary-phase fitness by allowing the surviving cells to scavenge from the dying majority oligopeptides as nutrient resources (ZINSER and KOLTER 1999). If so, the decrease in *cstA* activity of the *sgaA* mutant suggested that the insertion/inversion mutation may have compromised its ability to compete for oligopeptides during stationary phase. We therefore examined the role *cstA* plays in establishing stationary-phase fitness and the consequences of the loss of this gene's activity on the fitness of the *sgaA* mutant.

Competition experiments with a constructed in-frame Δ*cstA* mutant strain confirmed a positive role for *cstA* activity in establishing stationary-phase fitness. We defined relative fitness as the ability of one population, when inoculated as a thousandfold minority, to invade a differentially marked majority population over the course of the stationary-phase incubation, a period in which the total population decreases only gradually (ZAMBRANO *et al.* 1993). A wild-type minority population that expressed the *cstA* gene (WT) invaded a majority that could not (Δ*cstA*; Figure 3A). However, this same minority could not invade a majority that could express *cstA* (Figure 3A). Hence, there is a clear competitive advantage for the ability to express *cstA* during stationary phase.

The *sgaA* mutant has lost the fitness benefit of the *cstA* locus: *sgaA* was identified as a GASP mutation, and *sgaA* mutants can invade an *sgaA*$^+$ (WT) majority popu-

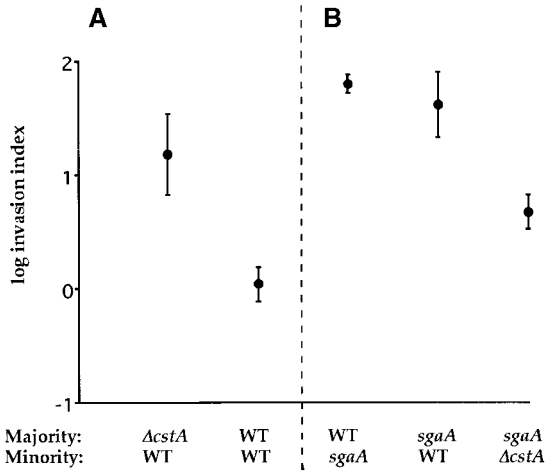

FIGURE 3.—Relative fitness of strains with different *cstA* and *sgaA* alleles. Differentially marked (Bgl^+ and Val^r) LB-grown stationary-phase cultures (24 hr) were mixed 1:1000 and assayed for viable counts on selection plates. Invasion index represents the mean change in the minority-to-total population ratio between days 1 and 3 of the stationary-phase competition ($\pm$SD, $n \geq 6$). For each pairwise competition, two to four replicate experiments were performed for each neutral marker configuration. To rule out marker effects on fitness, the Bgl^+ and Val^r markers were switched between competing strains in half of the pairwise competitions. (A) The wild type (WT) showed a positive invasion index when competed as a minority *vs.* the *cstA* null mutant (Δ*cstA*), but did not when competed *vs.* a differentially marked wild type. (B) The wild type and the *sgaA* mutant each showed a positive invasion index when competed as a minority *vs.* each other. The *cstA* null mutant showed a lower invasion index compared to the wild type when competed as a minority *vs. sgaA*. Data for the *sgaA* minority *vs.* the wild-type majority competition are from ZINSER and KOLTER (1999).

lation (ZINSER and KOLTER 1999; Figure 3B). However, because the *sgaA* mutant has lost transcriptional activity of the *cstA* locus, it should have also lost the fitness benefit of *cstA*. Indeed, the wild type was clearly able to invade the *sgaA* majority population when placed as a minority (Figure 3B). This advantage was due primarily to the *cstA* products, as the Δ*cstA* strain could not invade nearly to the same level as the wild type (Figure 3B). Hence, both the *sgaA* mutant and the wild type demonstrate a fitness advantage when competed *vs.* the other strain as a minority population. Such dependence of relative fitness on relative cell density indicates that the two strains compete for different nutrient resources and occupy different niches in stationary-phase cultures.

DISCUSSION

We have characterized a two-step genomic rearrangement that occurred during prolonged starvation and have demonstrated that this mutation provides a fitness gain in stationary phase. The IN(*cstA*::IS*5*-IS*5*D) mutation activated the *ybeJ-gltJKL-ybeK* operon by placing a CRP-binding element upstream of the operon. Concomitant with the activation of the *ybeJ-gltJKL-ybeK* operon was the inactivation of the *cstA* locus. Expression of either locus provided a clear ability to invade a majority that was incapable of expressing that locus, indicating that both loci contribute toward fitness in the same selection environment. From these data we propose that the insertion/inversion mutation *sgaA* resulted in an obligate resource trade-off, in which the initial ability to exploit one resource at equal fitness with the parent (oligopeptides), is forfeited for the new ability to outcompete the parent for another resource (monomeric amino acids). Even with the loss in fitness for the first resource, the insertion/inversion mutation resulted in a net fitness gain, likely because it acted to create a new niche for the mutant during starvation conditions.

We have shown that transposable elements were responsible for the activation of the *ybeJ-gltJKL-ybeK* operon (and inactivation of the *cstA* gene). Transposable elements have been found to activate transcription of adjacent genes by introducing complete or partial promoters located within the element itself or by disrupting or displacing a negative element that normally shuts down transcription (SYVANEN 1984). The mechanism of gene activation by the IN(*cstA*::IS*5*-IS*5*D) mutation is distinct from the ones described above. In this case, a transcription-activating DNA fragment is brought into proximity of the adjacent gene, but it is not carried by the transposable element (otherwise, expression of the *ybeJ-gltJKL-ybeK* operon would not have been dependent upon the inversion event).

Recombination-mediated regulatory element expropriation as a mechanism of gene activation/inactivation has been well established as a type of phase variation in bacteria (HENDERSON *et al.* 1999) and as a causative agent in the human cancer, Burkitt lymphoma (DALLA-FAVERA *et al.* 1982; SHEN-ONG *et al.* 1982; TAUB *et al.* 1982). Directed evolution studies in *Salmonella enterica* have also demonstrated occurrence of expropriation events in the restoration of inactive genes under intense selection pressures (SCHMID and ROTH 1983).

Here we describe regulatory element expropriation as a selection-dependent mechanism for the loss of genes or gene activities during evolution. Gene inactivation is a byproduct of the regulatory element expropriation event and, as long as the net fitness change is positive, selection will favor the loss of the gene's activity. These expropriation events can effectively select against gene activities that are deleterious, neutral, or beneficial to the organism under the conditions of selection. We note a distinction between the expropriation and the classical mutation accumulation mechanisms of gene activity loss. In expropriation, the loss of gene activity is selection dependent, because it is a necessary consequence of the overall fitness gain. By contrast, under mutation accumulation, loss of gene activity is a neutral or nearly neutral event and results from selection-independent genetic drift (ROSE and CHARLESWORTH 1980; HUGHES

and Charlesworth 1994; Cooper and Lenski 2000). Expropriation may also differ from antagonistic pleiotropy, in which there is a direct selection against a deleterious gene activity (Rose and Charlesworth 1980; Hughes and Charlesworth 1994; Cooper and Lenski 2000). Expropriation may therefore represent a third mechanism that can account for the loss of gene activity.

Regulatory element expropriation may play a significant role in the process of ecological specialization and evolution in microbes. Growing evidence suggests that the predominant mechanism of microbial evolution is lateral gene transfer (Guttman and Dykhuizen 1994; Lawrence and Roth 1996; Lawrence and Ochman 1998; Nelson *et al.* 1999). For lateral gene transfer to modify the functional activities of the recipient and therefore alter fitness, not only must genetic material be transferred to the recipient, but also the expression of the transferred genes must be properly regulated (Lawrence and Roth 1996). Expropriation may be an important method to provide immediate selective value to introduced genes, for it would avoid incompatibilities of the donor's promoters and regulatory elements with the recipient's transcriptional machinery and regulators (Lawrence and Roth 1996). IS element-mediated gene transfer events can insert foreign genes into positions where they become regulated by endogenous promoters (Kasak *et al.* 1993). Subsequent excision or decay of the IS element can remove all traces of its involvement in the transfer event. Such expropriation events may therefore be common in the evolutionary history of microbes, because they provide an overall fitness gain to the organism, notwithstanding fitness losses due to inactivation of the preexisting gene.

Initial loss of function by regulatory element expropriation leaves the structural gene intact. If the lost activity is beneficial under the selection conditions there will be considerable pressure to restore the activity, which can be accomplished by reversion (in this case, by "back" inversion). Such revertants would be at an advantage among a majority of nonrevertants (as seen in Figure 3B), and frequency-dependent selection could establish a form of phase variation (Henderson *et al.* 1999), keeping the gene activity within the population in a metastable state. However, if prior to a reversion event secondary mutations arise within the structural component of the inactivated gene (which would be neutral events), then genetic drift could lead to permanent loss of the gene.

We thank S. Dove for strains and valuable discussions. We also thank S. Finkel for valuable discussions and for critical reading of the manuscript. Research was supported by the National Institutes of Health and National Science Foundation (R.K.) and Centre National de la Recherche Scientifique and Commissariat à la Énergie Atomique (M.B.).

LITERATURE CITED

Alm, R. A., L. L. Ling, D. T. Moir, B. L. King, E. D. Brown *et al.*, 1999 Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:** 176–180.

Andersson, S. G. E., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. M. Alsmark *et al.*, 1998 The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396:** 133–143.

Blattner, F. R., G. R. Plunkett, C. A. Bloch, N. T. Perna, V. Burland *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. Science **277:** 1453–1462.

Busby, S., and R. H. Ebright, 1999 Transcription activation by catabolite activator protein (CAP). J. Mol. Biol. **293:** 199–213.

Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson *et al.*, 2001 Massive gene decay in the leprosy bacillus. Nature **409:** 1007–1011.

Cooper, V. S., and R. E. Lenski, 2000 The population genetics of ecological specialization in evolving *Escherichia coli* populations. Nature **407:** 736–739.

Dalla-Favera, R., M. Bregni, J. Erikson, D. Patterson, R. C. Gallo *et al.*, 1982 Human c-*myc onc* gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA **79:** 7824–7827.

Donnenberg, M. S., and J. B. Kaper, 1991 Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. Infect. Immun. **59:** 4310–4317.

Finkel, S. E., and R. Kolter, 1999 Evolution of microbial diversity during prolonged starvation. Proc. Natl. Acad. Sci. USA **96:** 4023–4027.

Finkel, S. E., E. Zinser, S. Gupta and R. Kolter, 1997 *Molecular Microbiology*, pp. 3–16. Springer-Verlag, Berlin.

Guttman, D. S., and D. E. Dykhuizen, 1994 Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. Science **266:** 1380–1383.

Henderson, I. R., P. Owen and J. P. Nataro, 1999 Molecular switches—the ON and OFF of bacterial phase variation. Mol. Microbiol. **33:** 919–932.

Hughes, K. A., and B. Charlesworth, 1994 A genetic analysis of senescence in *Drosophila*. Nature **367:** 64–66.

Kasak, L., R. Horak, A. Nurk, K. Talvik and M. Kivisaar, 1993 Regulation of the catechol 1,2-dioxygenase- and phenol monooxygenase-encoding *pheBA* operon in *Pseudomonas putida* PaW85. J. Bacteriol. **175:** 8038–8042.

Kleckner, N., J. Bender and S. Gottesman, 1991 Uses of transposons with emphasis on Tn*10*. Methods Enzymol. **204:** 139–180.

Kroger, M., and G. Hobom, 1982 Structural analysis of insertion sequence IS*5*. Nature **297:** 159–162.

Lawrence, J. G., 1999 Gene transfer, speciation, and the evolution of bacterial genomes. Curr. Opin. Microbiol. **2:** 519–523.

Lawrence, J. G., and H. Ochman, 1998 Molecular archaeology of the *Escherichia coli* genome. Proc. Natl. Acad. Sci. USA **95:** 9413–9417.

Lawrence, J. G., and J. R. Roth, 1996 Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics **143:** 1843–1860.

Lawrence, J. G., and J. R. Roth, 1999 Genomic flux: genome evolution by gene loss and acquisition, pp. 263–289 in *Organization of the Prokaryotic Genome*, edited by R. L. Charlebois. ASM Press, Washington, DC.

Maurelli, A. T., R. E. Fernandez, C. A. Bloch, C. K. Rode and A. Fasano, 1998 "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. Proc. Natl. Acad. Sci. USA **95:** 3943–3948.

Miller, J. H., 1992 *A Short Course in Bacterial Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Morishita, T., Y. Deguchi, M. Yajima, T. Sakurai and T. Yura, 1981 Multiple nutritional requirements of Lactobacilli: genetic lesions affecting amino acid biosynthetic pathways. J. Bacteriol. **148:** 64–71.

Nakata, N., T. Tobe, I. Fukuda, T. Suzuki, K. Komatsu *et al.*, 1993 The absence of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the *ompT* and *kcpA* loci. Mol. Microbiol. **9:** 459–468.

Nelson, K. E., R. A. Clayton, S. Gill, M. L. Gwinn, R. J. Dodson *et al.*, 1999 Evidence for lateral gene transfer between Archaea and Bacteria from the genome sequence of *Thermotoga maritima*. Nature **399:** 323–329.

Papadopoulos, D., D. Schneider, J. Meier-Eiss, W. Arber, R. E.

Lenski *et al.*, 1999 Genomic evolution during a 10,000-generation experiment with bacteria. Proc. Natl. Acad. Sci. USA **96:** 3807–3812.

Perna, N. T., G. Plunkett, V. Burland, B. Mau, J. D. Glasner *et al.*, 2001 Genomic sequence of enterohaemorrhagic *Escherichia coli* 0157:H7. Nature **409:** 529–533.

Rose, M., and B. Charlesworth, 1980 A test of evolutionary theories of senescence. Nature **287:** 141–142.

Schmid, M. B., and J. R. Roth, 1983 Selection and endpoint distribution of bacterial inversion mutations. Genetics **105:** 539–557.

Schneider, D., E. Duperchy, E. Coursange, R. E. Lenski and M. Blot, 2000 Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. Genetics **156:** 477–488.

Schultz, J. E., and A. Matin, 1991 Molecular and functional characterization of a carbon starvation gene of *Escherichia coli*. J. Mol. Biol. **218:** 129–140.

Shen-Ong, G. L.C., E. J. Keath, S. P. Piccoli and M. D. Cole, 1982 Novel *myc* oncogene RNA from abortive immunoglobulin-gene recombination in mouse plasmacytomas. Cell **31:** 443–452.

Simons, R. W., F. Houman and N. Kleckner, 1987 Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:** 85–96.

Syvanen, M., 1984 The evolutionary implications of mobile genetic elements. Annu. Rev. Genet. **18:** 271–293.

Taub, R., I. Kirsch, C. Morton, G. Lenoir, D. Swan *et al.*, 1982 Translocation of the c-*myc* gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc. Natl. Acad. Sci. USA **79:** 7837–7841.

Willis, R. C., and C. E. Furlong, 1975 Interactions of a glutamate-aspartate binding protein with the glutamate transport system of *Escherichia coli*. J. Biol. Chem. **250:** 2581–2586.

Wilmes-Riesenberg, M. R., and B. L. Wanner, 1992 *TnphoA* and *TnphoA′* elements for making and switching fusions for study of transcription, translation, and cell surface localization. J. Bacteriol. **174:** 4558–4575.

Zambrano, M. M., and R. Kolter, 1993 *Escherichia coli* mutants lacking NADH dehydrogenase-I have a competitive disadvantage in stationary phase. J. Bacteriol. **175:** 5642–5647.

Zambrano, M. M., D. A. Siegele, M. Almiron, A. Tormo and R. Kolter, 1993 Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. Science **259:** 1757–1760.

Zinser, E. R., and R. Kolter, 1999 Mutations enhancing amino acid catabolism confer a growth advantage in stationary phase. J. Bacteriol. **181:** 5800–5807.

Zinser, E. R., and R. Kolter, 2000 Prolonged stationary-phase incubation selects for *lrp* mutations in *Escherichia coli* K-12. J. Bacteriol. **182:** 4361–4365.

Communicating editor: J. B. Walsh