Parallel Genetic and Phenotypic Evolution of DNA Superhelicity in Experimental Populations of *Escherichia coli*

Estelle Crozat,[1,2] Cynthia Winkworth,†,[3] Joël Gaffé,[1,2] Peter F. Hallin,[4] Margaret A. Riley,[3] Richard E. Lenski,[5] and Dominique Schneider*,[1,2]

[1]Laboratoire Adaptation et Pathogénie des Micro-Organismes, Université Joseph Fourier Grenoble 1, Grenoble cedex 9, France

[2]CNRS UMR 5163, Laboratoire Adaptation et Pathogénie des Microorganismes, Grenoble, France

[3]Department of Biology, University of Massachusetts

[4]Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark

[5]Department of Microbiology and Molecular Genetics, Michigan State University

†Present address: Department of Zoology, University of Otago, Dunedin, New Zealand.

***Corresponding author:** E-mail: dominique.schneider@ujf-grenoble.fr.

Associate editor: Douglas Crawford

Abstract

DNA supercoiling is the master function that interconnects chromosome structure and global gene transcription. This function has recently been shown to be under strong selection in *Escherichia coli*. During the evolution of 12 initially identical populations propagated in a defined environment for 20,000 generations, parallel increases in DNA supercoiling were observed in ten populations. The genetic changes associated with the increased supercoiling were examined in one population, and beneficial mutations in the genes *topA* (encoding topoisomerase I) and *fis* (encoding a histone-like protein) were identified. To elucidate the molecular basis and impact of these changes, we quantified the level of genetic, phenotypic, and molecular parallelism linked to DNA supercoiling in all 12 evolving populations. First, sequence determination of DNA topology-related loci revealed strong genetic parallelism, with mutations concentrated in three genes (*topA*, *fis*, and *dusB*), although the populations had different alleles at each locus. Statistical analyses of these polymorphisms implied the action of positive selection and, moreover, suggested that *fis* and *dusB*, which belong to the same operon, have related functions. Indeed, we demonstrated that DusB regulates the expression of *fis* by both experimental and phylogenetic analyses. Second, molecular analyses of five mutations in *fis* and *dusB* affecting the transcription, translation, and protein activity of Fis also revealed strong parallelism in the resulting phenotypic effects. Third, artificially increasing DNA supercoiling in one of the two populations that lacked DNA topology changes led to a significant fitness increase. The high levels of molecular and genetic parallelism, targeting a small subset of the many genes involved in DNA supercoiling, indicate that changes in DNA superhelicity have been important in the evolution of these populations. Surprisingly, however, most of the evolved alleles we tested had either no detectable or slightly deleterious effects on fitness, despite these signatures of positive selection.

Key words: evolution, bacteria, DNA superhelicity, fitness, phenotypic and genetic parallelism.

Introduction

DNA superhelicity controls one of the most important regulatory networks in bacteria (Travers and Muskhelishvili 2005). It is involved in such vital processes as chromosome replication and segregation, transcription, translation, recombination, and cellular responses to environmental conditions. Many external stimuli produce transient modifications of DNA superhelicity, including changes in nutritional conditions (Balke and Gralla 1987; Reyes-Dominguez et al. 2003), oxidative stress (Weinstein-Fischer et al. 2000), osmotic stress (Higgins et al. 1988), thermal stress (Goldstein and Drlica 1984), pH stress (Karem and Foster 1993), presence or absence of oxygen (Bhriain et al. 1989), and intracellular growth of pathogens (O'Croinin et al. 2006). These transient alterations of superhelicity in turn modify the cell's global transcriptional patterns, thereby allowing phenotypic acclimation of bacteria to their environment (Higgins et al. 1988; Weinstein-Fischer et al. 2000; Peter et al. 2004).

DNA superhelicity impacts transcription through two mechanisms. The first acts locally, whereby the level of supercoiling alters the physical properties of promoter sequences, for example, by influencing the opening of the DNA duplex (Dai and Rothman-Denes 1999). The second mechanism acts more globally and involves the organization of the bacterial chromosome into topological domains (Postow et al. 2004; Valens et al. 2004; Deng et al. 2005). This domain organization is correlated with the existence of spatial patterns of transcription in the *Escherichia coli* genome (Jeong et al. 2004), which can be altered by modulating superhelicity. Structural organization of the chromosome, DNA superhelicity, and global transcription patterns are therefore highly interconnected processes that allow cells to cope with fluctuating environmental conditions.

The level of DNA superhelicity is tightly controlled by the combined activities of topoisomerases (Champoux 2001) and histone-like proteins (Dorman and Deighan

2003). In *E. coli*, topoisomerases I and IV, encoded by *topA* (Wang 1971) and *parC* and *parE* (Zechiedrich et al. 2000), respectively, act to relax DNA. In contrast, DNA gyrase, encoded by *gyrA* and *gyrB*, introduces negative supercoils into the chromosome (Gellert et al. 1976). Histone-like proteins can constrain negative supercoils (Luijsterburg et al. 2006) and produce diffusion barriers for the formation of topological domains (Hardy and Cozzarelli 2005). Most histone-like proteins, including Fis (Kelly et al. 2004), H-NS (Johansson et al. 2000; Hommais et al. 2001), HU (Kar et al. 2005), and IHF (Arfin et al. 2000), have also been shown to be global regulators of gene transcription in bacteria. These proteins are expressed at various levels during the cell cycle, generating yet another level of interplay between environmental signals, DNA superhelicity, and global transcription profiles.

Several lines of evidence suggest that Fis plays a crucial role in bacterial adaptation by modifying DNA superhelicity, thereby allowing cells to adjust their global gene expression profiles. Fis exerts a homeostatic control on DNA superhelicity, allowing a physiologically relevant supercoiling level in vivo. It also represses transcription of *gyrA* and *gyrB* (Schneider et al. 2000) and, in turn, its transcription is strongly dependent on the level of DNA supercoiling (Schneider et al. 2000). Nutrient upshifts lead to highly negative superhelicity, which triggers a high concentration of Fis (50,000 molecules per cell) (Talukder et al. 1999) that in turn activates transcription of ribosomal RNA (rRNA) operons and other genes necessary for cell growth. The transition to stationary phase is accompanied by DNA relaxation, leading to a dramatic decrease in Fis concentration (100 molecules per cell) and rRNA transcription. Besides responding to the nutritional state of the cell, Fis also helps bacteria cope with oxidative stress, again through its ability to adjust levels of DNA superhelicity (Weinstein-Fischer et al. 2000). After binding to many chromosomal sites, Fis acts as a global regulator of transcription (Kelly et al. 2004; Feldman-Cohen et al. 2006), regulating rRNA operons (Schneider et al. 2003), genes involved in metabolism (Gonzalez-Gil et al. 1996), and virulence genes (O'Croinin et al. 2006).

Evolution experiments, in which microbes are propagated for hundreds or thousands of generations in defined environments (Elena and Lenski 2003), provide a powerful approach to study the molecular basis of adaptation. In the longest running evolution experiment, Lenski and colleagues have propagated 12 populations of *E. coli* for more than 40,000 generations by daily serial transfer in a minimal medium supplemented with limiting glucose (Lenski et al. 1991; Cooper and Lenski 2000; Lenski 2004). All 12 populations achieved large gains in fitness. We showed previously that increases in negative DNA supercoiling evolved in parallel in most of these populations (Crozat et al. 2005). Parallel evolution, whereby multiple lineages evolve similar traits, is a strong indicator of adaptation by natural selection. For example, similar morphological changes have evolved repeatedly when populations of lizards, fishes, and flies have independently colonized similar environments in nature (Pigeon et al. 1997; Losos et al. 1998; Huey et al. 2000). Parallel changes sometimes extend even to the genome, as has been described for some viral and bacterial pathogens (Crandall et al. 1999; Reid et al. 2000). The parallel evolution of DNA superhelicity in the long-term experiment with *E. coli* provides therefore highly suggestive evidence that the resulting changes are adaptive. The change in DNA supercoiling was then further examined by genetic studies in one population, where two successive increases in DNA superhelicity were detected (Crozat et al. 2005). Mutations were found in *topA* and *fis* and, by using isogenic strains, it was demonstrated that these two mutations were responsible for the observed changes in DNA superhelicity in that population and, moreover, both were beneficial under the conditions of the evolution experiment. However, that previous study did not investigate whether parallel mutations had produced the parallel increases in DNA superhelicity observed in the other populations.

Several other phenotypic traits evolved in parallel in most or all populations, including cell size (Lenski et al. 1998; Philippe et al. 2009), growth parameters (Vasi et al. 1994), catabolic functions (Cooper and Lenski 2000; Pelosi et al. 2006), and global gene expression profiles. Genetic studies aimed at understanding these parallel phenotypic changes identified several genes that were also reproducibly altered in most or all populations, including *rbs*, the ribose utilization operon (Cooper et al. 2001); *spoT*, a gene involved in (p)ppGpp metabolism during the stringent response (Cooper et al. 2003; Pelosi et al. 2006); and *malT*, the transcriptional activator of the maltose utilization operons (Pelosi et al. 2006). Other parallel genetic changes were discovered in *nadR*, involved in NAD metabolism; *pykF*, which encodes pyruvate kinase I; *pbpA-rodA*, involved in cell wall biosynthesis; and *hokB-sokB*, a plasmid maintenance module (Schneider, Duperchy et al. 2000; Woods et al. 2006). Many of these genetic changes were demonstrated, by constructing and competing otherwise isogenic strains, to be beneficial in the environment of the evolution experiment. Complete genome sequences were recently obtained for evolved clones isolated at several generations from one population. A total of 45 mutations were found in the clone sampled after 20,000 generations (Barrick et al. 2009), of which about one-quarter had been identified by the previous genetic studies.

DNA supercoiling provides an interesting candidate function for further investigation into the extent as well as the pattern of evolutionary parallelism in these populations because more than 20 genes (encoding topoisomerases, histone-like proteins, and gene expression regulators) are involved in its regulation in *E. coli*. For the present study, we analyzed the extent of parallelism and its underlying molecular basis in three ways. First, we sequenced nine topology-related loci (including *topA* and *fis*) in the ancestor and evolved clones from all 12 populations. Under one hypothesis, mutations in many different genes that control DNA superhelicity may have caused the parallel increases in supercoiling. Alternatively, the same one or few genes may have been targeted by selection

Table 1. The 17 Topology-Related Genes and Upstream Open Reading Frames Sequenced in this Study.

Gene (upstream Open Reading Frames)	Number of Base Pairs Sequenced	Function	Reference
dps	802	Histone-like protein	Almiron et al. (1992)
fis (*dusB*)	1,695	Histone-like protein	Finkel and Johnson (1992)
gyrA	2,774	Gyrase subunit A	Champoux (2001)
gyrB	2,673	Gyrase subunit B	Champoux (2001)
hns	1,160	Histone-like protein	Yamada et al. (1991)
parC (*ygiS*)	2,599	Topoisomerase IV subunit A	Peng and Marians (1993)
parE (*yqiA*, *icc*, *yqiB*, *yqiE*)	4,652	Topoisomerase IV subunit B	Peng and Marians (1993)
topA	2,905	Topoisomerase I	Wang (1971)
topB (*selD*, *ydjA*)	3,842	Topoisomerase III	Champoux (2001)

in all of the independently evolving populations. Second, given that substantial genetic parallelism was observed, we analyzed how these changes related to gene regulation and to fitness. Third, two populations showed no increases in their DNA superhelicity (Crozat et al. 2005). We determined whether moving topology-related mutations from another evolved population affected superhelicity and fitness in one of these two populations that retained the ancestral superhelicity.

Materials and Methods

Long-Term Evolution Experiment

The long-term evolution experiment is described in detail elsewhere (Lenski 2004). In brief, 12 populations were founded using two *E. coli* B clones as ancestors (Jeong et al. 2009) and then propagated by serial daily transfers at 37 °C for 20,000 generations (3,000 days) in Davis minimal medium supplemented with glucose at 25 mg/ml (DM25) (Lenski et al. 1991). Each day, after 24-h incubation, cells were diluted 100-fold into fresh medium. One ancestral strain, REL606, cannot grow on arabinose (Ara^-), and it was used to initiate six populations, Ara−1 to Ara−6. The other six populations, Ara+1 to Ara+6, were started with REL607, a spontaneous Ara^+ mutant of the source strain. The Ara phenotype serves as a neutral marker in competition experiments for measuring relative fitness (Lenski et al. 1991; Lenski 2004).

Strains, Plasmids, and Culture Conditions

Evolved clones used in this study were isolated either at generation 20,000 (for sequence analyses and Western blots) or at generations 2,000 and 5,000 (for allele exchange experiments) from the 12 populations (Cooper and Lenski 2000); strains were stored as 15% glycerol suspensions at −80 °C. Strains that are isogenic except for *topA* and *fis* alleles were previously constructed in the ancestral background (Crozat et al. 2005) using the suicide plasmid pKO3 (Link et al. 1997), and they were used again in this study. These strains are 606^{topA-1} and 606^{fis-1}, which bear evolved *topA* and *fis* alleles, respectively, from population Ara−1 (Crozat et al. 2005), and $606^{\Delta fis}$, which has an in-frame *fis* deletion. Strains TOP10 (Invitrogen, San Diego, CA) and JM109 (Yanisch-Perron et al. 1985) were used for cloning experiments.

Plasmids pUC18 (Yanisch-Perron et al. 1985) were used for DNA supercoiling measurements, pKO3 (Link et al. 1997) for allele replacements, and pCRII-Topo (Invitrogen) for cloning experiments. Plasmids pKG13 and pKK223-3 (Gosink et al. 1996) were used to evaluate the DNA-binding activity of ancestral and evolved Fis proteins.

All experiments were performed by growing strains either in the DM25 minimal medium that was used in the evolution experiment (Lenski et al. 1991) or in Luria Broth (LB) (Sambrook et al. 1989). Ampicillin (100 μg/ml), chloramphenicol (30 μg/ml), or kanamycin (50 μg/ml) was added as needed.

DNA Sequencing

Nine loci encoding topoisomerases or histone-like proteins were sequenced in an evolved clone isolated at 20,000 generations from each of the 12 populations. That generation was chosen because most of the previous genetic studies, including the one on DNA topology (Crozat et al. 2005), focused on this time point. Moreover, the focal population Ara−1 evolved a mutator phenotype after 26,000 generations (Barrick et al. 2009), which complicates the analyses. Four of the sequenced genes are located within operons and the whole operons were then sequenced, such that 17 genes in all were sequenced in each clone. Table 1 details the genes sequenced in this study, as well as function of the gene products and number of base pairs sequenced. All 17 genes (totaling 23,102 bp) were sequenced in 14 clones, including the 12 evolved clones as well as the 2 ancestral variants, REL606 and REL607; as expected, the ancestors were identical. Genomic DNA preparations, polymerase chain reaction (PCR) amplifications, sequencing reactions, and sequence analyses were performed as previously described (Lenski et al. 2003). All primer pairs used for the genes in table 1 are available upon request to the corresponding author. All polymorphic sites were confirmed by resequencing the relevant regions twice with independent DNA extractions and reanalyzing the sequence chromatogram using Sequencher 3.1.1. Polymorphisms were analyzed using the Molecular Evolutionary Analysis Package MEA (Etsuko Moriyama, Yale University version 6/22/2000), based on the algorithm of Nei and Gojobori (1986).

Construction of Isogenic Strains

Evolved *fis* alleles from two populations (Ara+1 and Ara−3) and an evolved *dusB* allele from another (Ara+5) were moved into the ancestral chromosome by homologous recombination using the suicide plasmid pKO3 (Link et al.

1997), as described previously (Crozat et al. 2005). The constructed strains were checked for the presence of the evolved allele. Both evolved *fis* alleles generated new restriction sites, and screening employed a PCR–restriction fragment length polymorphism (PCR-RFLP) approach using *Alu*I and *Dde*I (Euromedex) to distinguish between the ancestral and evolved alleles. To distinguish between ancestral and evolved alleles of *dusB*, the region of *dusB* bearing the evolved mutation was sequenced. All constructed strains were also deconstructed, such that the evolved alleles were replaced by their ancestral counterparts, to confirm the absence of secondary mutations during the construction process. These experiments generated the following isogenic strains: 606^{fis+1}, 606^{fis-3}, and 606^{dusB+5}. All have the ancestral REL606 genome except they carry the indicated evolved alleles. We could not move the evolved *fis* allele from Ara−4 into the ancestral strain, despite numerous attempts, for reasons unknown. The same strategy was also used to replace the Ara+5 evolved *dusB* allele by its ancestral allele in a 5,000-generation clone, with the resulting strain called $5K^{dusBanc}$.

The same strategy was used to move the evolved *topA* allele from population Ara−1 (Crozat et al. 2005) into the chromosome of an evolved clone, called 1158C, isolated from population Ara+1 at 2,000 generations. The 1158C clone contains the ancestral *topA* allele, which was thus replaced by the evolved Ara−1 *topA* allele to produce strain $1158C^{topA-1}$.

Fitness Assays

The relative fitness of two strains was estimated by competing them, as described elsewhere (Pelosi et al. 2006). Briefly, isogenic strains were separately competed against the ancestor carrying the opposite arabinose marker. Each pairwise competition was replicated at least 6-fold. Competitions were performed in DM25, the same medium used for the evolution experiment, for either 1 or 6 days, with the longer assays used to measure smaller fitness differences. Samples were taken immediately after mixing at day 0 and again after 1 or 6 days of competition to measure the abundance of both competitors. Using the initial and final cell counts and dilution factors allows one to calculate the realized (net) population growth of each competitor. Fitness was then calculated as the ratio of their realized growth rates, and *t*-tests were used to evaluate whether the values differed significantly from the null hypothetical value of one. To determine the fitness effect of the evolved Ara+5 *dusB* allele in the genetic background of a 5,000-generation clone, the 5K and $5K^{dusBanc}$ clones competed separately against the REL606 ancestor and their fitness values were compared. We also tried to compete the 5K and $5K^{dusBanc}$ clones directly against each other. To distinguish them, we introduced the *araA* mutation responsible for the Ara^- phenotype of REL606 (Jeong et al. 2009) into each clone. However, we obtained inconsistent results with these marked pairs of clones, suggesting that a secondary mutation occurred during one of the constructions. In order to achieve better resolution of their relative fitness, we also competed the 5K and $5K^{dusBanc}$ clones separately against a 5,000-generation clone sampled from population Ara−1.

DNA Superhelicity Measurements

Topological changes were measured by the relative abundance of different topoisomers of a reporter plasmid, as described previously (Crozat et al. 2005). Briefly, plasmid pUC18 was introduced by electrotransformation into the strains of interest. After growing cells in LB-ampicillin media to an $OD_{600\ nm}$ of 2, plasmids were extracted (Qiagen Qiaprep Spin Miniprep Kit) and quantified, with 200 ng loaded onto 1% agarose gels containing 1.5 μg/ml chloroquine. Levels of DNA supercoiling were estimated by mean σ values. The limitations and validity of this plasmid-based assay are discussed in Crozat et al. (2005).

Electrophoresis and Immunoblot Analysis of Proteins

Triplicate cultures of the ancestral strains bearing the different *fis* and *dusB* alleles, and of evolved clones sampled at 20,000 generations from the 12 populations, were grown in LB medium at 37 °C. Aliquots were sampled during exponential phase after 1.5 h of incubation. After centrifugation, cells were resuspended in lysis buffer (70 mM Tris pH 7.4, 1 mM ethylenediaminetetraacetic acid [EDTA], 1 mM dithiothreitol [DTT], and 10% glycerol) before sonication. Lysates were centrifuged at 10,000 × g for 30 min at 4 °C, and total protein concentrations were determined using the Bradford Protein Assay Kit (Bio-Rad) with bovine serum albumin as a standard. Equal amounts of protein were loaded onto 16% sodium dodecyl sulfate–polyacrylamide gels, separated by electrophoresis and electrotransferred (Bio-Rad) onto nitrocellulose membranes (Amersham Pharmacia). Prestained protein standards (Euromedex) were used to estimate molecular weights. Immunoblot analyses were performed with antibodies against Fis (courtesy of A. Ishihama, Nippon Institute for Biological Sciences, Japan) and RpoA (courtesy of M. Cashel, National Institutes of Health, USA) and developed using the ECL plus Western Blotting Detection System Kit (Amersham Pharmacia).

Cell cultures for these assays were grown in LB rather than DM25 for three reasons. First, LB gave higher cell densities than DM25, especially during early exponential phase. Second, we wanted to be consistent with previous analyses (Crozat et al. 2005). The previous work, moreover, had demonstrated that the observed changes in superhelicity in the evolved clones were consistent between LB and DM25 media, although it was much more difficult to make the measurements in DM25 (Crozat et al. 2005). Third, it was previously shown that the expression profile of Fis, with a peak in exponential phase, was similar in LB and minimal media (Ball et al. 1992; Mallik et al. 2004), although the maximum level is lower in minimal medium.

Transcriptional Fusions and β-Galactosidase Assays

Ancestral and evolved Ara+1 *fis* promoter alleles were PCR amplified from positions −241 to +84 relative to the +1 *dusB* transcription start site (Mallik et al. 2004) using

primers (fisEcoRI 5′-cacgtaatttgccgaattcttacgcaaatt-3′ and fisBamHI 5′-ccatgggcgctgggatccaggcgatttctg-3′). These primers were designed to contain *Eco*RI and *Bam*HI restriction sites (underlined in the sequences). Each 325-bp product was cloned into the pRS551 plasmid, upstream of the *lacZ* reporter gene encoding β-galactosidase (Simons et al. 1987). These transcriptional fusions were transferred onto λ bacteriophage DNA by recombination, using λRS45 as described elsewhere (Simons et al. 1987). Both λRS45 derivatives were then introduced into the λ *att* site of the chromosome of ancestral strain REL606. The λRS45 derivative carrying the ancestral *fis* promoter was also introduced into the λ *att* site in strains 606^{dusB+5} and $606^{\Delta dusB}$. For both fusions, the sequences of the *fis* promoters were verified, and the single lysogenization of fusion-carrying λ phages was checked by using a previously described PCR method (Powell et al. 1994).

For β-galactosidase assays, the cultures were grown in kanamycin-containing LB media, and samples were taken at different times during growth. The β-galactosidase specific activities were assayed by using *o*-nitrophenyl β-D-galactopyranoside (ONPG) as a substrate, and they are expressed in nanokatals per milligram (nanomoles of ONPG transformed per second per milligram of cellular proteins, nkat/mg). The amount of total cellular protein was quantified using the Bradford Protein Assay Kit (Bio-Rad). All reported activity values are the average of three independent experiments.

Fis Purifications

Two of the evolved mutations in *fis* affect its coding sequence. We purified the ancestral and both evolved Fis proteins using a previously described protocol (Pan et al. 1996). The three *fis* alleles were first cloned into the pET41 expression vector and overexpressed in BL21(DE3) cells (Novagen). Plasmid-carrying cells were grown in 250-ml cultures of LB containing kanamycin at 37 °C to an $OD_{600\ nm}$ of 0.6–0.8, and Fis expression was induced with 0.8 mM of isopropyl-beta-D-thiogalactopyranoside (IPTG) for 4 h. After centrifugation, pellets were washed and resuspended in extraction buffer (70 mM Tris pH 7.4, 1 M NaCl, and 10% glycerol). After lysis with a French Press (1,000 psi; Thermo Spectronic, Rochester, NY), the lysates were centrifugated for 15 min at 5,000 × g and sonicated for 1 min on ice. The subsequent purification steps were performed as described elsewhere (Pan et al. 1996). Fis concentrations were determined with the Bradford Protein Assay reagent using bovine serum albumin as a standard.

DNA Labeling

Three promoters known to contain Fis-binding sites were analyzed: 1) a fragment with the *fis* promoter extending from positions −215 to +58 was PCR amplified using primers ODS234 (5′-cgcaaatttgcgtgccaaaatt-3′) and ODS435 (5′-tgagctgatattgctcgatg-3′); 2) a fragment with the *rrnB*P1 promoter, extending from position −210 to +56 (Appleman et al. 1998), was amplified using primers ODS436 (5′-aacatgaagccccggatgca-3′) and ODS437 (5′-tctcaggagaaccccgct-3′); and 3) a fragment with the *proP* promoters, extending from positions −114 to +197 (Xu and Johnson 1997), was amplified using primers ODS438 (5′-ctcgaggaaatcttctctgt-3′) and ODS439 (5′-tttccttttcagcatagcttt-3′). All positions are given relative to corresponding +1 transcription start sites. In each case, prior to the PCRs, the forward primer was labeled with γ-^{32}P-dATP (6,000 Ci/mmol; ICN) using T4 polynucleotide kinase according to the manufacturer's recommendations (Euromedex). Labeled fragments were purified using the column-based PCR Purification Kit (Qiagen).

Electromobility Shift Assays

Binding reactions were performed by incubating ≤5 fmol (≤0.6 nM) of labeled DNA fragments with various amounts (0–40 nM) of purified Fis protein in a 15 μl solution containing 20 mM Tris pH 7.5, 1 mM EDTA pH 8, 80 mM NaCl, 4% glycerol, 0.5 mM DTT, 0.5 mg/ml bovine serum albumin, and 10 μg/ml of sonicated herring sperm DNA. After 20 min of incubation at 37 °C, 5 μl of loading buffer (0.08% bromophenol blue in 50% sucrose) was added and samples were electrophoresed on 5% polyacrylamide gels (19:1) in 0.5× Tris-Borate EDTA buffer. The gels were then dried, exposed to Fujifilm Imaging Plates, and revealed on an FLA8000 scanner (Fujifilm). The percentage of bound DNA was quantified with the Image Gauge software (Fujifilm). Several types of controls were performed. From 5- to 1,000-fold excesses of unlabeled DNA fragments were added to the binding mixtures. The fragments consisted of a 30-bp oligonucleotide containing either a consensus Fis-binding site or the same oligonucleotide except with the Fis site mutated. A supershift assay was also performed by adding the antibody raised against Fis to the binding reactions.

Results

Evidence for Selection on Topology-Related Genes

DNA sequences were determined for 17 candidate genes, distributed over nine operons, for the ancestor and for clones sampled from all 12 evolving populations at generation 20,000 (table 2). These operons are involved in DNA supercoiling regulation in *E. coli*; six encode DNA topoisomerases and three major histone-like proteins. The 12 evolved clones have a total of 27 mutations.

To evaluate whether any of the candidate genes known to be involved in DNA topology, or any other genes in the same operons, were targets of selection during the long-term evolution experiment, we applied four statistical tests (Woods et al. 2006). The first test compares the overall rate of substitution in these 17 genes versus 36 previously sequenced randomly chosen gene regions (Lenski et al. 2003). A total of 23,102 bp were sequenced for the topology genes, whereas 18,374 bp were sequenced for the random genes. Table 3 shows that 11 of the 12 populations had higher rates of change in the topology-related genes than in the random genes. This difference is highly significant (one-tailed sign test, $P = 0.0032$) in the direction predicted

Table 2. Mutations Found in 12 Evolving Populations after 20,000 Generations of Evolution among Topology-Related Genes and Upstream Open Reading Frames.

Population[a]	Gene	Nucleotide Position[b]	Mutation	Amino Acid Change
Ara−1	*fis*	−4	A to C	NA (RBS)[c]
	topA	+97	C to T	H33 to Y
Ara−2 (M)	*topB*	+628	C to T	D210 to N
	ydjA (*topB* locus)	+270	G to A	No change
	ydjA (*topB* locus)	+15	T to C	No change
	topA	+1012	T to C	Y338 to H
	topA	+1526	A to G	Y509 to C
Ara−3	*fis*	+152	A to C	Y51 to S
	topA	+2315	C to A	T792 to K
Ara−4 (M)	*fis*	+177	+ A	CTD change
Ara−5	*topA*	−112	G to T	NA (promoter)[d]
	topA	+599	C to A	A200 to E
Ara−6	*dusB* (*fis* locus)	+904	G to A	A302 to T
Ara+1	*fis*	−63[e]	T to G	NA (promoter)[d]
Ara+2	*fis*	+329	C to A	NA (downstream)
Ara+3 (M)	*dusB* (*fis* locus)	+936	G to A	No change
	parC	+460	G to A	P154 to S
Ara+4	*dusB* (*fis* locus)	+942	+ 9 bp	+ ALE after E314
	dps	+143	IS*1* insertion	NA
Ara+5	*dusB* (*fis* locus)	+943	G to A	A315 to T
Ara+6 (M)	*dusB* (*fis* locus)	+899	T to G	F300 to C
	selD (*topB* locus)	+960	T to G	No change
	yqiB (*parE* locus)	+134	T to G	Y45 to S
	yqiA (*parE* locus)	+251	A to C	L84 to R
	hns	−89	A to C	NA (promoter)[d]
	gyrA	+1925	T to G	Q657 to H
	gyrB	+1971	T to G	E662 to A

NOTE.—NA, not applicable.

[a] Populations showing a mutator phenotype (Sniegowski et al. 1997) are indicated by (M).

[b] Nucleotide positions are given relative to the first position of the translational start codon of the corresponding gene unless indicated otherwise.

[c] Mutation within the ribosome binding site.

[d] Mutation within the promoter region.

[e] The nucleotide position is given relative to the first position of the translational start codon of *dusB*.

Table 3. Number of Mutations in Random and Topology-Related Genes in 12 *Escherichia coli* Populations that Evolved for 20,000 Generations.

	Random Genes[b] (18,374 bp total)		Topology-Related Genes[c] (23,102 bp total)	
Population[a]	No. of Mutations	Rate per 1,000 bp	No. of Mutations	Rate per 1,000 bp
Ara−1	0	0.000	2	0.087
Ara−2 (M)	0	0.000	5	0.216
Ara−3	0	0.000	2	0.087
Ara−4 (M)	3	0.163	1	0.043
Ara−5	0	0.000	2	0.087
Ara−6	0	0.000	1	0.043
Ara+1	0	0.000	1	0.043
Ara+2	0	0.000	1	0.043
Ara+3 (M)	1	0.054	2	0.087
Ara+4	0	0.000	2	0.087
Ara+5	0	0.000	1	0.043
Ara+6 (M)	2	0.109	7	0.303

[a] Populations Ara−2, Ara−4, Ara+3, and Ara+6 became mutators (M), whereas the others retained the low ancestral mutation rate.

[b] Including those mutations at or near fixation (Lenski et al. 2003).

[c] Topology-related genes include regulatory regions as well as other ORFs found in the same loci.

if one or more of the topology-related genes were undergoing accelerated evolution, consistent with being targets of selection. The difference is significant even if we exclude population Ara−1 ($P = 0.0059$), which had previously been shown to harbor beneficial mutations in both *topA* and *fis* (Crozat et al. 2005).

Of the 27 mutations identified in the topology-related genes, 11 (40.7%) were in *topA*, *fis*, or their regulatory regions, although together these two genes constituted only 3,634 bp (15.7%). Another gene, *dusB*, had five mutations (18.5%) in only 966 bp of coding sequence (4.2%). The *dusB* gene encodes one of three dihydrouridine synthases in *E. coli* (Bishop et al. 2002). Dihydrouridine is the most abundant modified base found at conserved positions in the D-loop of transfer RNA (tRNA) in prokaryotes and eukaryotes, but its physiological role is unknown (Bishop et al. 2002). The other 11 mutations (40.7%) were dispersed across ten different genes that comprised 80.1% of the sequencing effort, including 2 in *ydjA* and 1 each in *topB*, *parC*, *dps*, *selD*, *yqiB*, *yqiA*, *hns*, *gyrB*, and *gyrA*. We conclude that mutations in *dusB*, as well as those in *topA* and *fis*, provide evidence of positive selection based on this first statistical test. Moreover, the fact that *dusB* mutations were found in five populations (Ara−6, Ara+3, Ara+4, Ara+5, and Ara+6) that evolved changes in supercoiling (Crozat et al. 2005), while lacking any mutations in *topA* or *fis*, suggests that *dusB* mutations were responsible for the changes in DNA topology. The remaining statistical tests therefore focused on *topA*, *fis*, and *dusB*.

The second test compares the pattern of synonymous and nonsynonymous substitutions in these 3 topology-related genes with the pattern observed in the 36 randomly chosen genes. The random genes had equal numbers of synonymous and nonsynonymous mutations (three each) while nonsynonymous changes predominated in the topology-related genes (11/12 nonsynonymous including insertions in the coding sequences). This difference is, again, in the direction expected if the topology genes experienced selection for changes in the encoded proteins, although the difference is only marginally significant (one-tailed Fisher's exact test, $P = 0.0833$).

The third test compares the distribution of point mutations between four populations that became mutators, owing to evolved defects in DNA repair, and eight populations that retained the low ancestral mutation rate (table 3). Mutations in topology genes were more common in the nonmutator (10 mutations) than in the mutator (4 mutations) populations. In contrast, all six mutations in the randomly chosen genes were found in the four mutator populations. This difference is in the expected direction under the hypothesis that both mutator and nonmutator populations accumulate mutations in target genes, whereas mutator populations differentially accumulate mutations in other loci that are not under positive selection (one-tailed

FIG. 1. Substitutions in *fis* and *dusB* genes and structure of the *dusB-fis* operon. The −35 and −10 promoter elements (Mallik et al. 2004) are shown by black boxes and coding sequences by white rectangles. The broken arrow marks the initiation site and direction of transcription. The locations of ten mutations substituted in the long-term evolution experiment are indicated by arrows and labeled according to the population in which they arose. In this study, the three evolved *fis* alleles from populations Ara−1, Ara+1, and Ara−3 and the evolved *dusB* allele from population Ara+5 were genetically manipulated to construct otherwise isogenic strains. We were unable to move the *fis* allele from population Ara−4 to the ancestral background despite multiple attempts. However, we purified the evolved Fis protein from Ara−4, together with the corresponding protein from Ara−3, for biochemical analyses. We did not analyze the *fis* allele from population Ara+2 because it lies 3′ to the *dusB-fis* operon. All five evolved *dusB* alleles have mutations within a 45-bp region that encodes the C-terminal region of DusB, which is known not to be involved in the dihydrouridine synthase activity (Savage et al. 2006). We therefore genetically manipulated and studied only one *dusB* allele from population Ara+5.

Fisher's exact test, $P = 0.0054$). Importantly, and consistent with our separation of *topA*, *fis*, and *dusB* from the other genes in topology-related loci, all ten point mutations in those other genes were in the four mutator lines (one mutation in these other genes was in a nonmutator population, but it was an IS*1* insertion in *dps*). This distributional difference between *topA*, *fis*, and *dusB* and the other topology-related genes is also highly significant (one-tailed Fisher's exact test, $P = 0.0005$). The other topology-related genes thus accumulate mutations in a manner similar to the set of randomly chosen genes.

The fourth sequence-based statistical test evaluates the distribution of mutations across the 12 independently evolved populations. If mutations accumulated randomly by drift, and if the populations have the same mutation rate, then one would expect a Poisson distribution of these mutations. Given the variation in mutation rates caused by the defects in DNA repair in four populations, one would expect a highly clustered distribution under the drift hypothesis. By contrast, if selection drives sequence evolution and if the different mutations at a given locus, or set of loci, produce the same or similar benefit (such that multiple mutations exhibit diminishing returns), then one would expect a more uniform distribution of mutations across the replicate lines. Two of the candidate genes, *fis* and *dusB*, show maximally uniform distributions, each gene harboring single mutations in five different populations, although the number of mutations for either gene alone is too small for statistical significance. However, if we combine these two genes, which together comprise the full extent of an operon (Mallik et al. 2004), we see that ten populations each have exactly one mutation in either *fis* or *dusB*. That distribution is maximally uniform and significantly different from a random Poisson distribution ($P = 0.0039$). This outcome is driven not only by the uniform distribution of mutations in each separate gene but also by a significant negative association between the two genes (two-tailed Fisher's exact test, $P = 0.0079$). This pattern therefore led us to analyze the molecular and phenotypic interactions of *fis* and *dusB* and their associated gene products. In the next section, we confirm the sequence-based hypothesis that mutations in the *fis* and *dusB* genes can produce similar phenotypic effects.

The substitution of the evolved *topA* and *fis* alleles was previously demonstrated in the Ara−1 population by sequencing those genes in multiple clones sampled at several generations (Crozat et al. 2005). We similarly confirmed fixation of the evolved *topA*, *fis*, and *dusB* alleles in the other populations (table 2). In population Ara−2, the two evolved *topA* alleles had fixed by generations 10,000 (mutation at position +1526, 3/3 clones tested) and 20,000 (position +1012, 3/3 clones). In Ara−3, the *topA* allele fixed by generation 4,000 (3/3 clones tested). The two *topA* mutations in Ara−5 fixed by generations 2,000 (position +599, 3/3 clones tested) and 20,000 (position −112, 3/3 clones). The evolved *fis* alleles were substituted by generation 7,000 in population Ara−3 (3/3 clones), generation 20,000 in Ara−4 (3/3 clones), generation 20,000 in Ara+1 (3/3 clones), and generation 10,000 in Ara+2 (2/2 clones). The *dusB* alleles in populations Ara+5 and Ara+6 were substituted by 5,000 generations (13/13 clones) and 10,000 generations (2/2 clones), respectively.

Phenotypic Effects of *fis* and *dusB* Mutations

Features of Several Evolved Alleles

We observed striking genetic parallelism in terms of the concentration of mutations in three topology-related genes. However, there was also considerable divergence, with different alleles at each locus as well as different combinations of genes with mutations. Therefore, we analyzed the functional effects of several mutations to assess whether they confer similar phenotypes. We focused on four evolved alleles, three with mutations affecting *fis* and one affecting *dusB* (table 2). The *fis* mutations in populations Ara+1, Ara−1, and Ara−3 impact different aspects of *fis* gene expression (fig. 1). Specifically, the *fis* allele in Ara−1 has an A-to-C transversion located in the RBS that was previously described (Crozat et al. 2005), which caused a 3-fold decrease in the level of Fis, presumably by affecting translation initiation. The *fis* allele in Ara+1 presumably affects transcription because it altered the sixth position of the *fis* promoter −35 element (Walker et al. 1999), a T-to-G transversion at position −31 relative to the +1 transcription start site. This change does not bring the −35 element any closer to the consensus

5′-TTGACA-3′ sequence. The Ara−3 *fis* allele has a Y-to-S amino acid replacement at residue 51, which may affect the activity of the Fis protein. The ancestral Y51 residue is highly conserved, is located at the start of the B helix of a Fis monomer, and is involved in the formation of the active Fis dimers (Yuan et al. 1991). Finally, the evolved *dusB* allele from population Ara+5 involves an A-to-T amino acid replacement at residue 315. All five evolved *dusB* alleles have mutations within a 45-bp region encoding the C-terminal region of DusB, and none of them appear to impact the dihydrouridine synthase activity (Savage et al. 2006).

In order to analyze the phenotypic effects of these alleles, we examined a set of isogenic strains made by allelic replacements in the ancestral background of strain REL606: 606^{fis-1}, 606^{fis+1}, 606^{fis-3}, and $606^{\Delta fis}$, bearing the evolved *fis* alleles from Ara−1, Ara+1, Ara−3, and an in-frame *fis* deletion, respectively, and 606^{dusB+5} and $606^{\Delta dusB}$, which have the Ara+5 *dusB* evolved allele and an in-frame deletion of *dusB*, respectively. We also analyzed $5K^{dusBanc}$, in which the ancestral *dusB* allele replaced its evolved counterpart in a 5,000-generation clone sampled from the Ara+5 population.

The fis^{-1}, *fis*$^{+1}$, and *dusB*$^{+5}$ Evolved Alleles Affect the Level of Fis Protein

We measured the effects of the various *fis* and *dusB* alleles on the level of Fis protein. Soluble proteins were extracted from early exponential-phase cultures, when the level of Fis is maximal, and immunodetection was performed using an antibody raised against Fis (fig. 2A). An anti-RpoA (α subunit of RNA polymerase) antibody served as a control. The isogenic mutant strains (606^{fis-1}, 606^{fis+1}, 606^{fis-3}, $606^{\Delta fis}$, 606^{dusB+5}, and $606^{\Delta dusB}$), along with clones from all 12 populations at 20,000 generations, were compared with the ancestral strain REL606. Ancestral and evolved backgrounds gave qualitatively similar results for each *fis* and *dusB* allele when isogenic constructs were available (data not shown). Consistent with previous results (Crozat et al. 2005), the Ara−1 allele caused a roughly 3-fold reduction in the level of Fis, and a similar decline was observed for the Ara+1 allele (fig. 2A). In contrast, there were no comparable reductions for the Ara−3 and Ara−4 alleles, consistent with their mutations being located in the Fis coding sequence. Introduction of the evolved Ara+5 *dusB* allele into the ancestor also reduced the amount of Fis to about the same level as seen in the 20,000-generation Ara+5 clone. Moreover, deletion of *dusB* led to the absence of any detectable Fis protein, suggesting a positive regulatory role of DusB on *fis* expression. Complementation of the $606^{\Delta dusB}$ strain with a plasmid overexpressing *dusB* restored the expression of *fis* (fig. 2A). This restoration, together with the fact that the *dusB* deletion was in-frame, eliminates potential polar effects that might have resulted from the *dusB-fis* operon structure.

A comparison of *fis* transcription levels in *fis*::*lacZ* chromosomal transcriptional fusions constructed in strains REL606 and 606^{dusB+5} revealed no obvious differences,

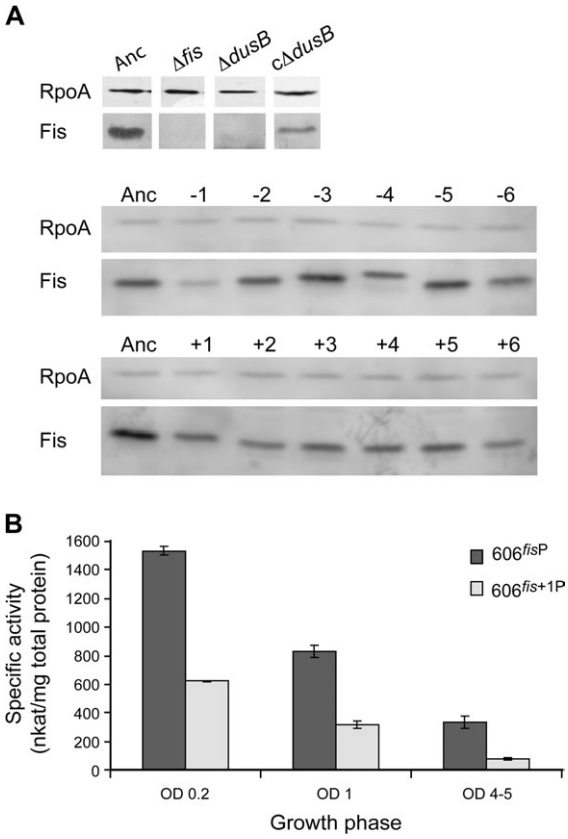


FIG. 2. Expression levels of *fis*. (*A*) Immunodetection of the Fis protein in a set of isogenic strains in the ancestral background and in evolved clones. Proteins were extracted in early exponential phase after 1.5 h in LB medium from the ancestral strain REL606 (labeled Anc), ancestral strains with in-frame deletions of *fis* (*Δfis*) and *dusB* (*ΔdusB*), ancestral strain with plasmid-encoded complementation of the deleted *dusB* gene (c*ΔdusB*), and 20,000-generation clones from each of the 12 populations (A−1 to −6 and A+1 to +6). Complementation of *ΔdusB* was achieved by transforming $606^{\Delta dusB}$ with the pCRII-Topo vector (Invitrogen) carrying the ancestral *dusB* gene under the control of its own promoter. Anti-RpoA served as an internal control to check for equivalent amounts of protein loaded into each well. The images show typical data from one of three replicate experiments. (*B*) Single-copy *lacZ* fusions were constructed in the REL606 ancestral chromosome with either the ancestral *fis* promoter (606^{fisP}) or the evolved Ara+1 *fis* promoter (606^{fis+1P}). Cultures were grown in LB medium with kanamycin, and β-galactosidase specific activities were measured from samples taken during exponential phase ($OD_{600\ nm}$ of 0.2), transition to stationary phase ($OD_{600\ nm}$ of 1), and stationary phase ($OD_{600\ nm}$ of 4–5). The data are averages of three independent assays, with error bars representing standard errors.

indicating that the regulation of *fis* by DusB might occur at a posttranscriptional level (data not shown). However, deleting either *fis* ($606^{\Delta fis}$) or *dusB* ($606^{\Delta dusB}$) resulted in an approximately 5-fold increase in transcription of *fis*. This increase was probably a direct consequence of the absence of Fis protein in both strains, relieving the autorepression of Fis on its own transcription and also confirming the regulation of *fis* by DusB. The absence of an effect on transcription in the 606^{dusB+5} strain suggests that enough Fis

protein is still present to bind to the high-affinity sites in the *fis* promoter region.

In the light of these regulatory effects, the Fis levels measured in the 12 evolved clones agree well with the types of alleles they possess (fig. 2*A*). Five of them have mutations in *dusB* (Ara−6, Ara+3, Ara+4, Ara+5, and Ara+6), and all show reductions in Fis levels, as do two or three other clones with mutations in regulatory elements of the *fis* gene (Ara−1, Ara+1, and perhaps Ara+2). Two evolved clones with mutations in the *fis* coding sequence (Ara−3 and Ara−4), as well as two with no mutations in either *fis* or *dusB* (Ara−2 and Ara−5), express levels of Fis that are similar to the ancestral strain REL606.

Because the fis^{+1} allele affects the promoter region (fig. 1), we also examined its effect on transcription. Single-copy chromosomal *lacZ* transcriptional fusions were created with either the ancestral or the Ara+1 evolved *fis* promoter using bacteriophage λ derivatives (Simons et al. 1987). Each fusion was introduced into the λ *att* site of the ancestral REL606 chromosome. The β-galactosidase activities were then measured in exponential phase, during the transition into stationary phase, and in stationary phase (fig. 2*B*). The ancestral and evolved promoters exhibit similar dependence of *fis* transcription on growth rate, with maximum transcription during exponential growth followed by declining activity as cells transition into stationary phase. However, the activity of the evolved promoter was substantially reduced in all growth phases, consistent with the protein data (fig. 2*A* and *B*).

The fis^{-3} and fis^{-4} Evolved Alleles Affect the Activity of Fis Protein

We examined the effects of the Ara−3 and Ara−4 *fis* alleles by measuring in vivo the DNA-binding activities of Fis proteins using two compatible plasmids (Gosink et al. 1996). The expression plasmid contains either the *fis* gene or a variant under the control of the *tac* promoter, whereas the reporter plasmid carries the *cat* gene whose expression is repressed by Fis binding (Gosink et al. 1996). The *cat* gene encodes chloramphenicol acetyltransferase, which confers resistance to chloramphenicol, and its expression depends on modified *lac* −35 and −10 promoter elements. In particular, the *lac* promoter has been modified by introducing the UP elements of the *rrnB*P1 promoter immediately upstream of the −35 box to strongly activate transcription and a strong Fis-binding site immediately downstream of the −10 box. Binding by Fis protein will repress *cat* transcription and thereby reduce the level of resistance to chloramphenicol of a strain carrying both plasmids. Moreover, the decrease in resistance should be proportional to the strength of Fis DNA binding.

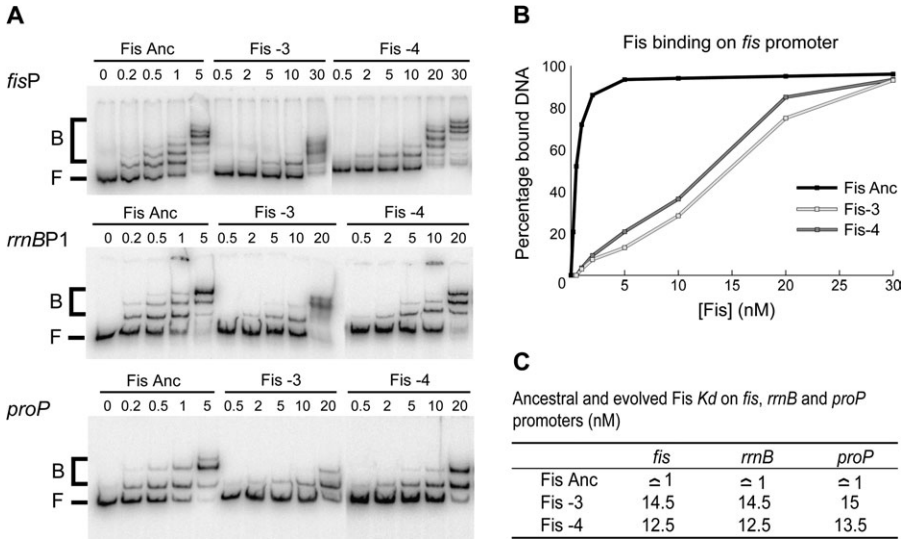
The ancestral, Ara−3, and Ara−4 *fis* genes were each cloned into the expression plasmid under the control of P*tac*. Each expression plasmid was introduced with the reporter plasmid into the ancestral strain carrying the in-frame *fis* deletion. All three strains were then plated onto two types of LB plates, one containing 100 μg/ml of chloramphenicol and the other containing 600 μg/ml of chloramphenicol; both types of plates also contained ampicillin to ensure retention of the expression plasmid. Relative survival was calculated as the ratio of colony number on plates with high and low chloramphenicol concentrations and used as an indirect measure of Fis binding, as previously described (Gosink et al. 1996). Survival was almost 9-fold higher when the expression plasmid carried the evolved *fis* allele from Ara−3 or Ara−4 (3.5% and 3.4%, respectively) compared with the ancestral gene (0.4%), indicating reduced DNA binding by the evolved Fis proteins.

To test this inference further, we purified the ancestral and evolved Fis proteins (Pan et al. 1996) and assessed their DNA-binding abilities by performing gel mobility shift assays (fig. 3). Three representative promoter sequences, all known to contain Fis-binding sites, were used as binding targets: the *fis* promoter (Mallik et al. 2004), the *rrnB*P1 promoter (Ross et al. 1990), and the *proP* promoters (McLeod et al. 2002). For the ancestral protein, we observed a shift in electrophoretic mobility even at the lowest Fis concentration (0.2 nM), whereas retardation was maximal at the highest concentration used (fig. 3*A*). The multiple retarded bands indicate the presence of several Fis-binding sites (Xu and Johnson 1995). In contrast, both evolved Fis proteins exhibit reduced DNA-binding abilities, with higher concentrations (2–5 nM) necessary to produce a discernible shift (fig. 3*A*). Much higher concentrations (at least 20–30 nM) of the evolved proteins were also required for maximal retardation.

Binding specificity was checked by using 1) a 5- to 1,000-fold excess of an unlabeled 30-bp oligonucleotide containing a consensus Fis-binding site, 2) the same DNA fragment but with the Fis site mutated, and 3) an antibody against Fis in the DNA-binding reactions. The first assay resulted in a reversal of the binding observed in figure 3*A*; the second assay had no detectable effect on binding, and the third assay resulted in a supershift (data not shown). To compare more precisely the behavior of the ancestral and evolved Fis proteins, we also quantified the amount of DNA bound by Fis based on the gel mobility shift assays. Figure 3*B* shows a typical quantitation curve, whereas figure 3*C* provides relative dissociation constants calculated for each of the three Fis proteins and each of the three promoter region. These values are 12–15 times higher for both evolved proteins compared with the ancestral protein, confirming their decreased DNA-binding ability.

Fitness Effects of the Evolved Alleles

We performed competition assays to estimate the fitness effects of two evolved *fis* alleles and one evolved *dusB* allele under the same conditions as the evolution experiment. All evolved alleles were introduced into the chromosome of the Ara^- ancestral strain REL606, and these strains competed against the Ara^+ ancestral strain REL607; the Ara marker is itself selectively neutral under these conditions (Lenski et al. 1991; Lenski 2004). Surprisingly, each evolved allele was either deleterious or neutral in the ancestral background (606^{fis+1}: relative fitness 0.9482, $t_s = 2.9232$, 5 df, $P = 0.0329$; 606^{fis-3}: 0.9535, $t_s = 5.5278$, 5 df, $P =$



	fis	*rrnB*	*proP*
Fis Anc	≏1	≏1	≏1
Fis -3	14.5	14.5	15
Fis -4	12.5	12.5	13.5

FIG. 3. DNA binding of ancestral and evolved Fis proteins. (*A*) Binding of the ancestral (Fis Anc), Ara−3 evolved (Fis−3), and Ara−4 evolved (Fis−4) Fis proteins to the *fis*P (top), *rrnB*P1 (middle), and *proP* P1 and P2 (bottom) promoter regions in vitro. The interaction of Fis proteins with these transcription regulatory regions was assessed by performing electrophoretic mobility shift assays. The regulatory sequences were amplified by PCR with radiolabeled primers and incubated with increasing concentrations of purified Fis proteins (shown in nM), and the complexes were electrophoresed. Free (F) and Fis-bound (B) DNA bands are indicated. Note that different concentrations of Fis protein were used for the different *fis* alleles. (*B*) Quantitation of bound DNA. The intensities of the free and bound DNA bands were measured using the Image Gauge software (Fujifilm), and the percentage of bound DNA is plotted against Fis concentration. All experiments were replicated three times, and typical curves are shown for each of the three Fis variants using the *fis*P promoter. (*C*) Dissociation constants. *Kd* values were calculated from quantitation curves, including those in (*B*), as the amount of Fis protein yielding one-half retardation in the shift assays in (*A*). Values are presented for each combination of the three Fis proteins and three promoters.

0.0027; and 606^{dusB+5}: 1.0002, $t_s = 0.1205$, 11 df, $P = 0.9063$; in each case, $H_0 = 1$ was tested using a two-tailed *t*-test). These outcomes contrast with the small, but significant, benefit measured for the evolved Ara−1 *fis* allele in the ancestral background (Crozat et al. 2005). We also examined the fitness effect of the *dusB* allele in the genetic background of a 5,000-generation clone (designated 5K for short) from population Ara+5, which was done by replacing the evolved allele with the ancestral one. A clone isolated at 5,000 generations was chosen because the evolved *dusB* allele had fixed by 5,000 generations but was rare or absent at 4,000 generations. As explained in the Materials and Methods, we could not compete 5K and $5K^{dusBanc}$ directly. Instead, 5K and $5K^{dusBanc}$ competed separately with the ancestor and with a 5,000-generation clone from population Ara−1. Against the ancestor, the fitness values for 5K and $5K^{dusBanc}$ were 1.5080 and 1.5507, respectively, and that difference is not significant ($t_s = 0.9707$, 10 df, $P = 0.3546$, based on a two-tailed *t*-test). Against the Ara−1 evolved clone, the fitness values for 5K and $5K^{dusBanc}$ were 1.0647 and 1.0801, respectively, which are also not significant ($t_s = 0.5714$, 12 df, $P = 0.5783$, based on a two-tailed *t*-test). Thus, the evolved *dusB* allele appears to be neutral or perhaps slightly deleterious even in its own background, although we cannot exclude a small benefit of, say, one or a few percent. We also measured the fitness effects, in the ancestral background, of in-frame deletions of *fis* and *dusB* by competing $606^{\Delta fis}$ and $606^{\Delta dusB}$ against REL607. Each deletion significantly reduced fitness ($606^{\Delta fis}$: 0.9123, $t_s = 6.9591$, 5 df, $P = 0.0009$, and $606^{\Delta dusB}$: 0.9275, $t_s = 5.9920$, 5 df, $P = 0.0019$; in each case, $H_0 = 1$ was tested using a two-tailed *t*-test).

To summarize our results to this point, we observed strong genetic parallelism during the evolution experiment, with *fis* and *dusB* repeatedly being targets of natural selection. This parallelism was achieved, however, through different mutations that affected all levels of regulation: transcription, translation, protein activity, and regulation via a newly discovered regulator of *fis*, the DusB protein. Despite these mutational differences, however, all of the evolved variants are phenotypically similar insofar as they cause a reduction in either the amount or the activity of Fis protein. However, we were unable to measure significant fitness advantages for the mutations we tested, perhaps because benefits are very small or depend subtly on context, although the genetic and phenotypic parallelisms strongly imply that they were beneficial during the long-term evolution experiment.

Phylogenetic Confirmation of the Functional Link between *dusB* and *fis*

An earlier study suggested a functional link between *dusB* and *fis* based on a phylogenetic analysis of the distribution of the DusB and Fis proteins (Morett and Bork 1998). Sequence alignments of diverse γ-proteobacterial Fis proteins and α- and γ-proteobacterial NtrC proteins indicate that Fis is probably derived from some ancestral version of the C-terminal DNA-binding domain of the α-proteobacterial NtrC. Moreover, DusB has significant similarity to NifR3, encoded by a gene located in the same operon as *ntrC*

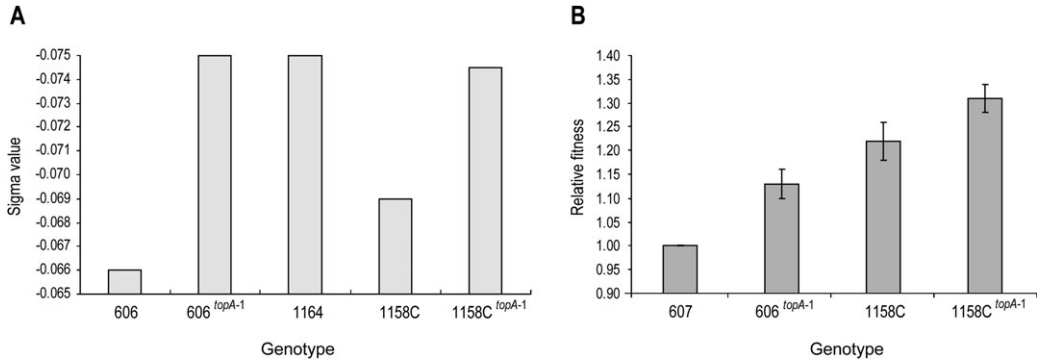


Fig. 4. Effects of an evolved *topA* allele on DNA supercoiling and fitness. (*A*) DNA supercoiling levels were measured using a reporter plasmid in five different genetic backgrounds. 606 is the REL606 ancestor of the evolution experiment (bearing the ancestral *topA* allele); 606^{topA-1} is REL606 except with the Ara−1 evolved *topA* allele introduced (Crozat et al. 2005); 1164 is a 2,000-generation clone from population Ara−1 with the evolved *topA* allele; 1158C is a 2,000-generation clone from population Ara+1 that still carries the ancestral *topA* allele, and $1158C^{topA-1}$ is the same Ara+1 clone except with the Ara−1 evolved *topA* allele introduced. (*B*) Fitness effects of the Ara−1 evolved *topA* allele when introduced into the ancestral and evolved Ara+1 genetic backgrounds. Competitions were performed in the same medium used in the long-term evolution experiment, with the reference competitor being the ancestral strain with the opposite Ara marker state. Error bars are 95% confidence intervals based on six replicate competitions for each genotype. The Ara−1 evolved *topA* allele increases fitness in both backgrounds, and the effects are significant based on two-tailed *t*-tests ($P = 0.0402$ in the Ara+1 background; $P < 0.0001$; [Crozat et al. 2005] in the ancestral background). The genotypes are the same as in (*A*), with 607 being the Ara^+ ancestral counterpart of REL606, which is Ara^-. The Ara marker itself has no discernible effect on fitness; competitions between 606 and 607 yield a mean relative fitness that does not differ significantly from a value of one.

in several α-proteobacteria. These earlier sequence analyses further suggested that the *dusB-fis* operon resulted from some rearrangement involving horizontal transfer into a lineage ancestral to the γ-proteobacteria (Morett and Bork 1998). These patterns were offered as evidence for some interaction between the proteins to "carry out an as yet unknown function."

This previous study on the association of Fis and DusB was performed in 1998 using fewer than 30 proteins in the alignments. We revisited the phylogenetic association by analyzing the *fis* and *dusB* sequences in 323 bacteria with available genome sequences. The *fis* gene was not found at all outside the proteobacteria. The association of *dusB* and *fis* in the same operon was widely observed in both the γ-proteobacteria (53 of 79 genomes examined) and the β-proteobacteria (22 of 23 genomes). By contrast, this same association was rare in the α-proteobacteria (1/46 genomes) and δ-proteobacteria (1/11 genomes). These observations reinforce the functional association between *fis* and *dusB*. They also suggest that the *dusB-fis* operon arose before the γ- and β-proteobacteria diverged, whereas the earlier phylogenetic analysis suggested a more recent origin within the γ-proteobacteria (Morett and Bork 1998).

Test of the Effect of Genetic Context on the Evolution of Increased Supercoiling

Unlike the other ten populations, two lines, Ara+1 and Ara+3, had not evolved increased supercoiling within 20,000 generations (Crozat et al. 2005), nor had either line substituted any mutations in the *topA* gene (table 2), and yet they had improved fitness similar to the other populations (Cooper and Lenski 2000). We asked, therefore, whether mutations elsewhere in the genome had eliminated the advantage of increased DNA supercoiling. We moved the *topA* allele that had been substituted in population Ara−1 (Crozat et al. 2005), and which strongly increases both fitness (+13%) and DNA superhelicity (from −0.066 to −0.075), into the genome of an evolved clone isolated at generation 2,000 from population Ara+1. This clone, called 1158C, still carries the ancestral *topA* allele, and it shows only a slight increase in DNA supercoiling (−0.069) relative to the ancestor (Crozat et al. 2005). We then compared the supercoiling and fitness levels of 1158C and its $1158C^{topA-1}$ counterpart. Replacing the ancestral *topA* with the evolved $topA^{-1}$ allele increased supercoiling in the 1158C clone to the same level as generated by the $topA^{-1}$ allele in the ancestral genetic background (fig. 4*A*). The evolved $topA^{-1}$ allele also significantly increased fitness in both the ancestral (Crozat et al. 2005) and the evolved Ara+1 backgrounds (fig. 4*B*).

Thus, in at least one lineage that evolved little or no increase in DNA supercoiling, the introduction of a beneficial supercoiling-related mutation from another line generated similar phenotypic effects, at least at the 2,000-generation time point that we tested. Therefore, the absence of an altered *topA* in that particular population does not appear to be the result of a very early substitution of some other mutation that eliminated the beneficial effect of *topA*.

Discussion

In a previous study, parallel phenotypic increases in DNA supercoiling were observed in 10 of 12 evolving *E. coli* populations (Crozat et al. 2005). In one population that evolved two successive increases in DNA supercoiling, two mutations were discovered, one in *topA* and the other in *fis*, and together, they were shown to be responsible for the observed changes in DNA topology in that population.

Both mutations were also demonstrated to increase fitness, indicating that supercoiling was a target of selection in the experimental environment (Crozat et al. 2005). More recently, whole-genome sequencing of several clones isolated from this same population through 20,000 generations revealed no other mutations in genes known to control DNA topology (Barrick et al. 2009).

Here, we extend the investigation of the DNA topology-associated regulatory network to all 12 populations. We discovered substantial parallelism at both the genetic and the molecular levels. Sequencing nine DNA topology-related genes encoding topoisomerases and histone-like proteins suggested that only two of them, *topA* and *fis*, had evolved in many or most of the populations. We also sequenced eight other genes in the same operons as the nine known topology-related genes and discovered that one of them, *dusB*, had also changed in several populations and, moreover, was involved in regulating DNA topology. All of the mutations in these three genes were fixed in their respective populations. In contrast to this parallelism at the level of the affected genes, however, there was some divergence in the mode of action of the substituted alleles. In particular, phenotypic analyses revealed that some of the mutations—those in the promoter of the *dusB-fis* operon, in the *fis* RBS, and in the coding sequence of *dusB*—reduced the level of Fis protein, whereas those in the coding region of *fis* decreased the protein's activity.

The three genes identified here as the targets of selection on DNA topology—*topA*, *fis*, and *dusB*—represented only 20% of our sequencing effort, yet they harbored 16 of the 27 mutations found in the 17 genes sequenced. Moreover, all but 1 of the other 11 mutations occurred in the four populations that had evolved mutator phenotypes, which are known to accumulate neutral and mildly deleterious mutations at much higher rates (Sniegowski et al. 1997; Cooper and Lenski 2000; Funchain et al. 2000; Giraud et al. 2001; Lenski et al. 2003; Barrick et al. 2009). In previous work on the genetic basis of adaptation in these same lines, but using different genes, mutations under positive selection were shown to be overdispersed (Woods et al. 2006); for example, each line substituted one nonsynonymous mutation in a particular gene, but none harbored two or more substitutions in the same gene. This nonrandom distribution implies that one mutation suffices to provide a benefit, whereas a second one in the same gene yields little or no further advantage. With respect to *fis*, we observed that 5 populations had single substitutions in this gene, which is consistent with the overdispersed distribution, although the pattern is not significant given only five mutations in 12 populations. However, we also noticed that five lines each had single mutations in *dusB*, which occurs in the same operon as *fis*. Moreover, no line had mutations in both *fis* and *dusB*, such that the ten mutations total in this operon were maximally, and significantly, overdispersed among the 12 populations. This pattern led us to hypothesize that these two genes interact in some way, such that a population could benefit similarly from a mutation in either *fis* or *dusB*. Our genetic experiments and molecular assays demonstrated that the DusB protein regulates expression of *fis* and confirmed that some mutations in *dusB* had similar functional effects to certain mutations in *fis* and its known regulatory elements. These experiments showed for the first time a regulatory function of DusB on the *fis* gene in addition to its previously described dihydrouridine synthase activity (Bishop et al. 2002). Our phylogenetic analysis of the linkage between *dusB* and *fis* also implies that DusB has a similar regulatory function in many other species besides *E. coli*.

Despite propagating the same ancestral strain in identical environments, all 12 lines have followed different evolutionary trajectories. In some cases, certain of the early steps have closed off particular avenues or opened new pathways for adaptation (Cooper et al. 2003; Blount et al. 2008); in other cases, the same opportunity may exist until each population has solved the challenge in a similar way. With respect to DNA topology, two lines had not yet changed even after 20,000 generations. When we moved a *topA* allele from a line that had evolved in this respect to a 2,000-generation clone from one of the populations that had not changed, that allele was also beneficial in this other genetic background, indicating that the potential for adaptive evolution of this trait had persisted at least to that point in time. More generally, despite parallelism at the level of genes, there was substantial allelic diversity, with no two evolved alleles identical at the nucleotide level. In particular, different mutations in *fis* and *dusB* affected different aspects of *fis* expression—transcription, translation, and protein activity—though they had similar effects because they all reduced either the amount or the activity of Fis. The mutation in population Ara+1 affects the promoter (Walker et al. 1999) and thereby reduces transcription. The mutation in Ara−1 affects the RBS element, reducing complementarity between the translational initiation region of *fis* messenger RNA (mRNA) and the 3′ end of 16S rRNA (Owens et al. 2004). The mutation in Ara−4 is a 1-bp insertion, leading to a change in the C-terminus of the protein such that each monomer is 103, rather than 98, residues long. This mutation modifies the D helix of Fis, which affects the DNA-binding helix-turn-helix (HTH) motif (Yuan et al. 1991); important residues are also absent from the evolved protein, which might explain its reduced DNA-binding activity. The evolved Ara−3 allele, by contrast, replaces the Y51 residue with an S residue, and the Y51 residue constitutes an important part of a hydrogen-bonded network that is involved in both the HTH DNA-binding motif and the Fis dimerization (Yuan et al. 1991). The *dusB* mutation in population Ara+5 disrupts positive regulation of Fis by DusB at the translational level. DusB is known to modify tRNAs (Bishop et al. 2002), and so we hypothesize that DusB might exert its regulatory effect by binding the *dusB-fis* mRNA, thereby facilitating the translation of *fis*. However, additional experiments are needed to test this hypothesis and further investigate the functional relationship between DusB and Fis. More generally, these evolved *fis* and *dusB* alleles, by reducing either the amount or the activity of Fis, may lead to an

increase in the level of DNA superhelicity by partially relieving the repression Fis has been shown to exert on the *gyrAB* genes that encode gyrase (Schneider et al. 1999).

Two interesting comparisons can be made with other evolution experiments that have been performed with various organisms and under diverse conditions. First, two of the main bacterial global regulatory networks, DNA superhelicity (Crozat et al. 2005, this work) and the stringent response (Cooper et al. 2003; Pelosi et al. 2006), have changed in many of the replicate populations in this long-term experiment with *E. coli*, as have other genes that may have profound metabolic effects (Philippe et al. 2007). More generally, changes in regulatory and metabolic networks appear to be a common and an important feature of laboratory evolution experiments with microbes, including *E. coli* propagated in other environments (Herring et al. 2006; Ferenci 2008; Conrad et al. 2009), as well as many other bacteria (Velicer et al. 2006; Bantinaki et al. 2007; Gresham et al. 2008) and yeast (Rosenzweig et al. 1994; Ferea et al. 1999; Gresham et al. 2008; Lang et al. 2009). These results indicate the important role of regulatory networks in sustaining the evolvability of living organisms.

Second, parallel evolution across replicate populations has been reported in many of these experiments, although the functional level and quantitative extent of parallelism vary between different systems. At one extreme, replicate populations of bacteriophage ΦX174 exhibited striking parallelism even at the DNA sequence level, with about half of all the fixed mutations identical across two replicate populations (Wichman et al. 1999). By contrast, in several studies of evolving bacteria, including this one, replicate populations frequently show parallelism at the level of genes and cellular functions but rarely at the level of the mutations themselves (Herring et al. 2006; Woods et al. 2006; Bantinaki et al. 2007; Barrick et al. 2009; Conrad et al. 2009). Woods et al. (2006) have hypothesized that the smaller size and lower complexity of viral genomes relative to bacteria might explain their greater propensity for parallel sequence evolution.

The parallel evolution that we have observed at the level of DNA superhelicity and the underlying genes provides strong evidence that the corresponding mutations had beneficial fitness effects, especially in the absence of any specific sequence motifs that might suggest hypermutability. Indeed, this interpretation has been confirmed for several other genes in these same evolving population system by constructing and competing isogenic strains that differ only by the alleles in question (Cooper et al. 2001, 2003; Pelosi et al. 2006), including previous studies of *topA* and *fis* alleles that changed supercoiling in one of the populations (Crozat et al. 2005). Yet, despite the signature of positive selection in the genetic and functional parallelism, two other *fis* and one *dusB* evolved alleles that we tested here were neutral or even slightly deleterious when moved into the ancestral background.

We do not know why the evolved alleles tested in this study failed to show the expected benefits, but we can suggest three possible explanations. One possibility is a methodological problem during the construction of the isogenic strains, in particular the possible introduction of some secondary mutation that arose during the genetic manipulations. Most mutations are deleterious, and so any such random mutation would be far more likely to reduce fitness than to increase it (hence, such problems are unlikely to explain beneficial effects when those are observed). A second possibility involves subtle differences in the environmental conditions used to assess fitness effects and those that existed during the evolution experiment itself. In particular, fitness assays involve pairwise competitions in which the strains have equal initial frequency. By contrast, many genotypes were often present during the evolution experiment (Papadopoulos et al. 1999; Barrick and Lenski 2009), and some of them have subtle fitness advantages only when they are in the minority (Elena and Lenski 1997; Rozen et al. 2005).

A third possibility is that evolved alleles may be beneficial in some genetic contexts but not in others. In particular, an evolved allele might be beneficial only in association with one or more mutations that arose previously, in which case the allele in question would not be advantageous when tested in the ancestral strain. This possibility may be particularly relevant here because pervasive epistatic interactions involving global regulatory networks have been demonstrated in the long-term populations (Cooper et al. 2008). Moreover, it is known that the stringent response and DNA superhelicity regulatory networks are intertwined (Travers and Muskhelishvili 2005), and both have been evolving in these populations (Cooper et al. 2003, 2008; Crozat et al. 2005; Philippe et al. 2007). Also, different evolved alleles may have different pleiotropic effects, so that compensatory changes might be specific to a given allele, thereby giving rise to epistatic interactions with mutations in some other genes. Given these potential subtleties and complications, it would be interesting to reconstruct and analyze in detail the effects of all of the mutational steps in one or more of these populations, though that remains a daunting challenge in light of the many mutations and the difficulty of ensuring strict isogenic comparisons. For now, we sought to test this epistatic hypothesis by replacing the evolved *dusB* allele (which had appeared to be neutral in the ancestral background) in an evolved clone with the ancestral *dusB* allele. The result was inconclusive at best, however, with the evolved allele again having a negative, albeit not significant, effect in this context.

In conclusion, a central goal of evolutionary biology is to characterize the often complex relationships between genotype, phenotype, and fitness. We previously showed that mutations in *topA* and *fis* were responsible for changes in DNA superhelicity in 1 of 12 populations in a long-term experiment with *E. coli* (Crozat et al. 2005). Here, we extended these results to the other populations in that experiment by showing that mutations in *topA*, *fis*, and a third gene, *dusB*, caused parallel changes in superhelicity in most of them. The *dusB* gene was not previously known to have any role in controlling DNA topology prior to our analyses. We showed that DusB exerts its effect on

superhelicity by regulating the expression of the Fis protein. However, it remains unknown how a reduction in the amount or activity of Fis leads to the observed changes in supercoiling and if other yet unknown genetic changes may also be involved. Ongoing genome sequencing and global expression analyses of isolates from the long-term evolution experiment should provide new insights into these questions. It will be especially useful to obtain global transcription profiles of strains that are isogenic except for the various *fis* and *dusB* mutations to evaluate the impact of those mutations on other genes that may affect such phenotypes as DNA superhelicity and fitness. Comparisons of these profiles across the ancestral and evolved backgrounds should also give insights into other mutations that may interact with the topology-related alleles. Histone-like proteins such as Fis are involved in both global transcription control and global chromosomal organization, and thus, it would also be informative to obtain and compare the complete maps of Fis-binding sites in the ancestral and evolved genomes. These maps could be obtained by chromatin immunoprecipitation coupled with microarray technology or ChIP-chip (Grainger et al. 2006). Taken together, such analyses would provide valuable insights into the role of changing gene regulatory networks in the relationship between genotype, phenotype, and fitness.

Acknowledgments

The authors thank Mohamed-Ali Hakimi (Laboratoire Adaptation et Pathogénie des Micro-Organismes, France) for help in Fis purification, Akira Ishihama (Nippon Institute for Biological Sciences, Japan) and Michael Cashel (National Institutes of Health, USA) for the anti-Fis and anti-RpoA antibodies, respectively, and Tamas Gaal and Richard Gourse (University of Wisconsin, USA) for plasmids pKG13 and pKK223-3. This work was supported by grants from the French Centre National de la Recherche Scientifique (to D.S., BDI fellowship to E.C.), Université Joseph Fourier (to D.S.), Agence Nationale de la Recherche Program "Génomique Microbienne à Grande Echelle" (ANR-08-GENM-023 to D.S.), the United States National Science Foundation (DEB-9981397 and DEB-0515729 to R.E.L.), and the United States Defense Advanced Research Projects Agency "Fun Bio" Program (HR0011-09-1-0055 to R.E.L.).

References

Almiron M, Link AJ, Furlong D, Kolter R. 1992. A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev*. 6:2646–2654.

Appleman JA, Ross W, Salomon J, Gourse RL. 1998. Activation of *Escherichia coli* rRNA transcription by FIS during a growth cycle. *J Bacteriol*. 180:1525–1532.

Arfin SM, Long AD, Ito ET, Tolleri L, Riehle MM, Paegle ES, Hatfield GW. 2000. Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *J Biol Chem*. 275:29672–29684.

Balke VL, Gralla JD. 1987. Changes in the linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. *J Bacteriol*. 169:4499–4506.

Ball CA, Osuna R, Ferguson KC, Johnson RC. 1992. Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J Bacteriol*. 174:8043–8056.

Bantinaki E, Kassen R, Knight CG, Robinson Z, Spiers AJ, Rainey PB. 2007. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of Wrinkly Spreader diversity. *Genetics* 176:441–453.

Barrick JE, Lenski RE. 2009. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. *Cold Spring Harb Symp Quant Biol*. Advance Access published September 23, 2009, doi:10.1101/sqb.2009.74.018.

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.

Bhriain NN, Dorman CJ, Higgins CF. 1989. An overlap between osmotic and anaerobic stress responses: a potential role for DNA supercoiling in the coordinate regulation of gene expression. *Mol Microbiol*. 3:933–942.

Bishop AC, Xu J, Johnson RC, Schimmel P, Crecy-Lagard V. 2002. Identification of the tRNA-dihydrouridine synthase family. *J Biol Chem*. 277:25090–25095.

Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A*. 105:7899–7906.

Champoux JJ. 2001. DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem*. 70:369–413.

Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y, Palsson BO. 2009. Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol*. 10(10):R118.

Cooper TF, Remold SK, Lenski RE, Schneider D. 2008. Expression profiles reveal parallel evolution of epistatic interactions involving the CRP regulon in *Escherichia coli*. *PLoS Genet*. 4:e35.

Cooper TF, Rozen DE, Lenski RE. 2003. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc Natl Acad Sci U S A*. 100:1072–1077.

Cooper VS, Lenski RE. 2000. The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407:736–739.

Cooper VS, Schneider D, Blot M, Lenski RE. 2001. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J Bacteriol*. 183:2834–2841.

Crandall KA, Kelsey CR, Imamichi H, Lane HC, Salzman NP. 1999. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous substitution rate ratio to detect selection. *Mol Biol Evol*. 16:372–382.

Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D. 2005. Long-term experimental evolution in *Escherichia coli*. XII. DNA topology as a key target of selection. *Genetics* 169:523–532.

Dai X, Rothman-Denes LB. 1999. DNA structure and transcription. *Curr Opin Microbiol*. 2:126–130.

Deng S, Stein RA, Higgins NP. 2005. Organization of supercoil domains and their reorganization by transcription. *Mol Microbiol*. 57:1511–1521.

Dorman CJ, Deighan P. 2003. Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Genet Dev*. 13:179–184.

Elena SF, Lenski RE. 1997. Long-term experimental evolution in *Escherichia coli*. VII. Mechanisms maintaining genetic variability within populations. *Evolution* 51:1058–1067.

Elena SF, Lenski RE. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet*. 4:457–469.

Feldman-Cohen LS, Shao Y, Meinhold D, Miller C, Colon W, Osuna R. 2006. Common and variable contributions of Fis

residues to high-affinity binding at different DNA sequences. *J Bacteriol*. 188:2081–2095.
Ferea TL, Botstein D, Brown PO, Rosenzweig RF. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc Natl Acad Sci U S A*. 96:9721–9726.
Ferenci T. 2008. The spread of a beneficial mutation in experimental bacterial populations: the influence of the environment and genotype on the fixation of *rpoS* mutations. *Heredity* 100:446–452.
Finkel SE, Johnson RC. 1992. The Fis protein: it's not just for DNA inversion anymore. *Mol Microbiol*. 6:3257–3265.
Funchain P, Yeung A, Stewart JL, Lin R, Slupska MM, Miller JH. 2000. The consequences of growth of a mutator strain of *Escherichia coli* as measured by loss of function among multiple gene targets and loss of fitness. *Genetics* 154:959–970.
Gellert M, Mizuuchi K, O'Dea MH, Nash HA. 1976. DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc Natl Acad Sci U S A*. 73:3872–3876.
Giraud A, Matic I, Tenaillon O, Clara A, Radman M, Fons M, Taddei F. 2001. Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* 291:2606–2608.
Goldstein E, Drlica K. 1984. Regulation of bacterial DNA supercoiling: plasmid linking numbers vary with growth temperature. *Proc Natl Acad Sci U S A*. 81:4046–4050.
Gonzalez-Gil G, Bringmann P, Kahmann R. 1996. FIS is a regulator of metabolism in *Escherichia coli*. *Mol Microbiol*. 22:21–29.
Gosink KK, Gaal T, Bokal AJ, Gourse RL. 1996. A positive control mutant of the transcription activator protein FIS. *J Bacteriol*. 178:5182–5187.
Grainger DC, Hurd D, Goldberg MD, Busby SJ. 2006. Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res*. 34:4642–4652.
Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D, Dunham MJ. 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet*. 4:e1000303.
Hardy CD, Cozzarelli NR. 2005. A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol Microbiol*. 57:1636–1652.
Herring CD, Raghunathan A, Honisch C, et al. (11 co-authors). 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet*. 38:1406–1412.
Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, Bremer E. 1988. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *Salmonella typhimurium* and *Escherichia coli*. *Cell* 52:569–584.
Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer JP, Danchin A, Bertin P. 2001. Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol*. 40:20–36.
Huey RB, Gilchrist GW, Carlson ML, Berrigan D, Serra L. 2000. Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287:308–309.
Jeong H, Barbe V, Lee CH, et al. (19 co-authors). 2009. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J Mol Biol*. 394:644–652.
Jeong KS, Ahn J, Khodursky AB. 2004. Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol*. 5:R86.
Johansson J, Balsalobre C, Wang SY, Urbonaviciene J, Jin DJ, Sonden B, Uhlin BE. 2000. Nucleoid proteins stimulate stringently controlled bacterial promoters: a link between the cAMP-CRP and the (p)ppGpp regulons in *Escherichia coli*. *Cell* 102:475–485.
Kar S, Edgar R, Adhya S. 2005. Nucleoid remodeling by an altered HU protein: reorganization of the transcription program. *Proc Natl Acad Sci U S A*. 102:16397–16402.
Karem K, Foster JW. 1993. The influence of DNA topology on the environmental regulation of a pH-regulated locus in *Salmonella typhimurium*. *Mol Microbiol*. 10:75–86.
Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JC, Dorman CJ. 2004. A global role for Fis in the transcriptional control of metabolism and type III secretion in *Salmonella enterica* serovar *Typhimurium*. *Microbiology* 150:2037–2053.
Lang GI, Murray AW, Botstein D. 2009. The cost of gene expression underlies a fitness trade-off in yeast. *Proc Natl Acad Sci U S A*. 106:5755–5760.
Lenski RE. 2004. Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*. *Plant Breed Rev*. 24:225–265.
Lenski RE, Mongold JA, Sniegowski PD, Travisano M, Vasi F, Gerrish PJ, Schmidt TM. 1998. Evolution of competitive fitness in experimental populations of *Escherichia coli*: what makes one genotype a better competitor than another? *Antonie Van Leeuwenhoek*. 73:35–47.
Lenski RE, Rose MR, Simpson SC, Tadler SC. 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat*. 138:1315–1341.
Lenski RE, Winkworth CL, Riley MA. 2003. Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. *J Mol Evol*. 56:498–508.
Link AJ, Phillips D, Church GM. 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol*. 179:6228–6237.
Losos JB, Jackman TR, Larson A, Queiroz K, Rodriguez-Schettino L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279:2115–2118.
Luijsterburg MS, Noom MC, Wuite GJ, Dame RT. 2006. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol*. 156:262–272.
Mallik P, Pratt TS, Beach MB, Bradley MD, Undamatla J, Osuna R. 2004. Growth phase-dependent regulation and stringent control of fis are conserved processes in enteric bacteria and involve a single promoter (*fis* P) in *Escherichia coli*. *J Bacteriol*. 186:122–135.
McLeod SM, Aiyar SE, Gourse RL, Johnson RC. 2002. The C-terminal domains of the RNA polymerase alpha subunits: contact site with Fis and localization during co-activation with CRP at the *Escherichia coli proP* P2 promoter. *J Mol Biol*. 316:517–529.
Morett E, Bork P. 1998. Evolution of new protein function: recombinational enhancer Fis originated by horizontal gene transfer from the transcriptional regulator NtrC. *FEBS Lett*. 433:108–112.
Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 3:418–426.
O'Croinin T, Carroll RK, Kelly A, Dorman CJ. 2006. Roles for DNA supercoiling and the Fis protein in modulating expression of virulence genes during intracellular growth of *Salmonella enterica* serovar *Typhimurium*. *Mol Microbiol*. 62:869–882.
Owens RM, Pritchard G, Skipp P, Hodey M, Connell SR, Nierhaus KH, O'Connor CD. 2004. A dedicated translation factor controls the synthesis of the global regulator Fis. *EMBO J*. 23:3375–3385.
Pan CQ, Finkel SE, Cramton SE, Feng JA, Sigman DS, Johnson RC. 1996. Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J Mol Biol*. 264:675–695.

Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE, Blot M. 1999. Genomic evolution during a 10,000-generation experiment with bacteria. *Proc Natl Acad Sci U S A*. 96:3807–3812.

Pelosi L, Kuhn L, Guetta D, Garin J, Geiselmann J, Lenski RE, Schneider D. 2006. Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* 173:1851–1869.

Peng H, Marians KJ. 1993. *Escherichia coli* topoisomerase IV. Purification, characterization, subunit structure, and subunit interactions. *J Biol Chem*. 268:24481–24490.

Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR. 2004. Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol*. 5:R87.

Philippe N, Crozat E, Lenski RE, Schneider D. 2007. Evolution of global regulatory networks during a long-term experiment with *Escherichia coli*. *Bioessays* 29:846–860.

Philippe N, Pelosi L, Lenski RE, Schneider D. 2009. Evolution of penicillin-binding protein 2 concentration and cell shape during a long-term experiment with *Escherichia coli*. *J Bacteriol*. 191:909–921.

Pigeon D, Chouinard A, Bernatchez L. 1997. Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis* Salmonidae). *Evolution* 51:196–205.

Postow L, Hardy CD, Arsuaga J, Cozzarelli NR. 2004. Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev*. 18:1766–1779.

Powell BS, Rivas MP, Court DL, Nakamura Y, Turnbough CL Jr. 1994. Rapid confirmation of single copy lambda prophage integration by PCR. *Nucleic Acids Res*. 22:5765–5766.

Reid SD, Herbelin CJ, Bumbaugh AC, Selander RK, Whittam TS. 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* 406:64–67.

Reyes-Dominguez Y, Contreras-Ferrat G, Ramirez-Santos J, Membrillo-Hernandez J, Gomez-Eichelmann MC. 2003. Plasmid DNA supercoiling and gyrase activity in *Escherichia coli* wild-type and *rpoS* stationary-phase cells. *J Bacteriol*. 185:1097–1100.

Rosenzweig RF, Sharp RR, Treves DS, Adams J. 1994. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* 137:903–917.

Ross W, Thompson JF, Newlands JT, Gourse RL. 1990. *Escherichia coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J*. 9:3733–3742.

Rozen DE, Schneider D, Lenski RE. 2005. Long-term experimental evolution in *Escherichia coli*. XIII. Phylogenetic history of a balanced polymorphism. *J Mol Evol*. 61:171–180.

Sambrook J, Fritsch JE, Maniatis T. 1989. Molecular cloning: a laboratory manual. Plainview (NY): Cold Spring Harbor Laboratory Press.

Savage DF, de Crécy-Lagard V, Bishop AC. 2006. Molecular determinants of dihydrouridine synthase activity. *FEBS Lett*. 580:5198–5202.

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

Schneider DA, Ross W, Gourse RL. 2003. Control of rRNA expression in *Escherichia coli*. *Curr Opin Microbiol*. 6:151–156.

Schneider R, Travers A, Kutateladze T, Muskhelishvili G. 1999. A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol Microbiol*. 34:953–964.

Schneider R, Travers A, Muskhelishvili G. 2000. The expression of the *Escherichia coli fis* gene is strongly dependent on the superhelical density of DNA. *Mol Microbiol*. 38:167–175.

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

Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387:703–705.

Talukder AA, Iwata A, Nishimura A, Ueda S, Ishihama A. 1999. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol*. 181:6361–6370.

Travers A, Muskhelishvili G. 2005. DNA supercoiling—a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol*. 3:157–169.

Valens M, Penaud S, Rossignol M, Cornet F, Boccard F. 2004. Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J*. 23:4330–4341.

Vasi F, Travisano M, Lenski RE. 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. *Am Nat*. 144:432–456.

Velicer GJ, Raddatz G, Keller H, Deiss S, Lanz C, Dinkelacker I, Schuster SC. 2006. Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc Natl Acad Sci U S A*. 103:8107–8112.

Walker KA, Atkins CL, Osuna R. 1999. Functional determinants of the *Escherichia coli fis* promoter: roles of -35, -10, and transcription initiation regions in the response to stringent control and growth phase-dependent regulation. *J Bacteriol*. 181:1269–1280.

Wang JC. 1971. Interaction between DNA and an *Escherichia coli* protein omega. *J Mol Biol*. 55:523–533.

Weinstein-Fischer D, Elgrably-Weiss M, Altuvia S. 2000. *Escherichia coli* response to hydrogen peroxide: a role for DNA supercoiling, topoisomerase I and Fis. *Mol Microbiol*. 35:1413–1420.

Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ. 1999. Different trajectories of parallel evolution during viral adaptation. *Science* 285:422–424.

Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. 2006. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci U S A*. 103:9107–9112.

Xu J, Johnson RC. 1995. Fis activates the RpoS-dependent stationary-phase expression of *proP* in *Escherichia coli*. *J Bacteriol*. 177:5222–5231.

Xu J, Johnson RC. 1997. Activation of RpoS-dependent *proP* P2 transcription by the Fis protein in vitro. *J Mol Biol*. 270:346–359.

Yamada H, Yoshida T, Tanaka K, Sasakawa C, Mizuno T. 1991. Molecular analysis of the *Escherichia coli hns* gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Mol Gen Genet*. 230:332–336.

Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103–119.

Yuan HS, Finkel SE, Feng JA, Kaczor-Grzeskowiak M, Johnson RC, Dickerson RE. 1991. The molecular structure of wild-type and a mutant Fis protein: relationship between mutational changes and recombinational enhancer function or DNA binding. *Proc Natl Acad Sci U S A*. 88:9558–9562.

Zechiedrich EL, Khodursky AB, Bachellier S, Schneider R, Chen D, Lilley DM, Cozzarelli NR. 2000. Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *J Biol Chem*. 275:8103–8113.