Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling

Bilal El Houdaigui[1], Raphaël Forquet[1], Thomas Hindré[2], Dominique Schneider[2], William Nasser[1], Sylvie Reverchon [1] **and Sam Meyer** [1,*]

[1]Université de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, CNRS UMR5240, Laboratoire de Microbiologie, Adaptation et Pathogénie, 69621 Villeurbanne, France and [2]Université Grenoble-Alpes, CNRS, Grenoble INP, TIMC-IMAG, 38000 Grenoble, France

ABSTRACT

DNA supercoiling acts as a global transcriptional regulator in bacteria, that plays an important role in adapting their expression programme to environmental changes, but for which no quantitative or even qualitative regulatory model is available. Here, we focus on spatial supercoiling heterogeneities caused by the transcription process itself, which strongly contribute to this regulation mode. We propose a new mechanistic modeling of the transcription-supercoiling dynamical coupling along a genome, which allows simulating and quantitatively reproducing *in vitro* and *in vivo* transcription assays, and highlights the role of genes' local orientation in their supercoiling sensitivity. Consistently with predictions, we show that chromosomal relaxation artificially induced by gyrase inhibitors selectively activates convergent genes in several enterobacteria, while conversely, an increase in DNA supercoiling naturally selected in a long-term evolution experiment with *Escherichia coli* favours divergent genes. Simulations show that these global expression responses to changes in DNA supercoiling result from fundamental mechanical constraints imposed by transcription, independently from more specific regulation of each promoter. These constraints underpin a significant and predictable contribution to the complex rules by which bacteria use DNA supercoiling as a global but fine-tuned transcriptional regulator.

INTRODUCTION

The role of DNA supercoiling (SC) in transcriptional regulation has attracted considerable attention in recent years. Due to the helical nature of DNA, mechanical torsion affects transcription at both initiation and elongation steps, and can thereby be considered as a non-conventional transcriptional regulator in eukaryotes as well as bacteria (1–5). In the latter, fast changes in DNA topology play a central role in the global transcriptional response to environmental stress (4,6). Inheritable changes in DNA topology are also under positive selection during evolution experiments with bacteria, in which SC-modifying mutations can provide a substantial fitness gain (7).

The regulatory action of SC is usually analysed from transcriptomes obtained after treatment by DNA gyrase inhibitors, causing global relaxation of the chromosome and changes in the transcription level of hundreds of genes (8–11). Since topoisomerases are found in all bacterial species, including those almost devoid of transcription factors such as *Mycoplasma* or *Buchnera* (11,12), they might underpin an ancestral and widespread global regulation mode. However, our understanding of the underlying mechanisms remains limited, as there is currently no quantitative or even qualitative systematic model of the features defining a supercoiling-sensitive gene, and how it responds to SC changes. Here, we hypothesise that the local genomic architecture (*i.e.*, relative orientations and distances between adjacent genes) is a primary factor in this response, by dictating the distribution of SC at the gene's promoter. Indeed, as recognized >30 years ago (13), transcription also generates significant torsional stress during the elongation step (14). This stress is positive in front of the elongating RNA polymerase (RNAP) and negative behind it, and is able to diffuse along the double-helix at distances of several kilobasepairs and reach nearby promoters (13,15,16). The relationship between transcription and SC is thus double-sided, and constitutes a dynamic and spatially organised coupling (17), hereafter quoted Transcription-Supercoiling Coupling (TSC). TSC has been proposed to underpin a complex and nonlinear interaction between adjacent genes depending on their relative orientations (17–20), which could play a fundamental yet unexplored role in the supercoiling regulation mode of gene expression (6).

Several TSC models have been recently proposed (17–19,21) in a biophysical and essentially theoretical perspec-

tive, e.g., aiming at reproducing so-called 'transcription bursts' (22). But strikingly, no attempt has been made so far to simulate the expression of any specific promoter, or to reproduce the results of a specific experimental assay. These models indeed lack important components for such a purpose, such as explicit topoisomerase enzymes, which play a crucial role in the biological process. The new TSC model presented here is inspired by previous ones in that it describes the 1D stochastic binding and elongation of RNAPs along genes and the SC distribution resulting thereof; however, in contrast to these previous models, it was specifically developed and applied in a regulatory perspective, i.e., with the aim of simulating actual biological systems and quantifying the effect of TSC-mediated regulation *in vivo*. It includes new realistic modeling ingredients that allow (i) precisely mimicking specific experimental assays through a small number of biologically relevant input parameters (promoter initiation rates, topoisomerase concentrations, position of topological barriers), (ii) directly comparing the results of the simulations to observations and (iii) inferring the underlying dynamics of transcription and propagation of supercoils. In this paper, we use this model to develop the first systematic comparison between TSC simulations and gene expression data.

After presenting the model, we first simulate a series of *in vitro* or *in vivo* experimental systems on SC-sensitive model promoters. We show that our simplified description is able to reproduce the quantitative effect of TSC on gene expression on the chromosome, and demonstrate that it is largely dictated by local gene orientations. We then propose that the genomic context may be a strong determinant of the 'supercoiling-sensitivity' of many bacterial genes, independently from any sequence specificity of their promoter. We analyse existing and new transcriptomic data obtained in conditions of gyrase inhibition by antibiotics causing chromosomal relaxation, and show that convergent genes are significantly more activated than divergent ones in several bacterial species. We then demonstrate that this behavior results from the basic mechanical constraints imposed by transcription, independently from species- or gene-specific properties. These constraints define how DNA topology, globally controlled by the cell physiology, affects the expression of genes according to their local orientation, promoter strength and distance. Finally, we ask if this form of genome-printed regulation can contribute to bacterial evolvability; we analyse global transcription profiles obtained from the longest-running evolution experiment, in which SC-modifying modifications have been selected. As predicted by our TSC modeling, we demonstrate that genes' expression changes in the evolved strains with modified SC are related to their local orientation. This analysis suggests that the regulatory rules dictated by neighbor genes' topological interactions likely constitute a robust and fundamental constraint governing the evolution and regulation of bacterial genomes.

MATERIALS AND METHODS

Model equations

Our model describes the dynamic transcription-supercoiling coupling. Most hypotheses and components

of the model are described in Results and Discussion; here, we provide equations and parameter values. The promoter response curve (Figure 1C) is computed from a thermodynamic model of transcription, $f(\sigma) = \exp(m\,U(\sigma))$ where $U(\sigma)$ is the SC-dependent promoter opening free energy. It follows a sigmoidal curve (17): $U(\sigma) = 1/(1 + \exp((\sigma - \sigma_t)/\epsilon))$, where σ_t is the threshold of promoter opening, ϵ sets the width of the crossover, and $1/m$ is an effective thermal energy that sets the SC activation factor. Standard values shown on Figure 1 C are $\sigma_t = -0.042$, $\epsilon = 0.005$, $m = 2.5$ (calibrated on the *pelE* promoter, see below).

Topoisomerase activity curves $t(\sigma)$ (Figure 1D) follow sigmoidal curves (same equation as above) parameterised from experimental assays (23,24): bacterial topoisomerase I acts on negatively SC DNA only, whereas the DNA gyrase exhibits moderate ($\sim$30%) activity on relaxed DNA, and is fully active at $\sigma \simeq 0.1$. Accordingly, thresholds of the sigmoids were set at values -0.04 and 0.01, and crossover widths at 0.012 and 0.025 respectively (Figure 1 D). Basal activity constants were calibrated on *in vitro* assays of transcription-induced SC accumulation (Figure 2B, see Results): $k_{\text{topo}} = \pm 0.001$ s^{-1} (for topoisomerase I and gyrase respectively).

Simulation methods

The genome was discretised with a 60-nt unit length, and simulations were performed with the Euler algorithm using a constant timestep $dt = 2$ s. At each timestep, free RNAPs are stochastically assigned to the available promoters according to their instantaneous initiation rate, or stay unbound; elongating RNAPs are moved by one unit length (speed 30 nt/s), and unbind when they reach a terminator. The SC levels of all topological regions (separated by fixed proteins or RNAPs) are then updated. Each elongating RNAP adds a constant number $\pm\Gamma$ of supercoils (in the regions ahead or behind it, respectively), which is normalised by the region length to compute the increment in superhelical density. Γ represents the fraction of DNA coils that are converted into torsional supercoils by the elongating RNAP, and might depend on the environment (molecular crowding, viscosity); a value of 0.2 per unit length was used in all simulations presented here. Topoisomerases then increment the SC level of each domain by $k_{\text{topo}}\,t(\sigma)\,dt$. In absence of topoisomerases, we checked that the genome-averaged SC level is constant throughout the simulation. Simulations, as well as all data analyses, were carried in Python with the NumPy numerical package. Note that the effective initiation rate of a gene in a simulation depends not only on its basal rate, but also on the number of RNAPs in the simulation, SC levels, and initiation rates of the other promoters. In all simulations of *in vivo* systems, we used a concentration of 0.25 μM for gyrase, and 0.025 μM for topoisomerase I (22).

Our unidimensional description of the chromosome disregards the 3D organisation of topological domains (25), especially plectonemes, loops, etc. This organisation is strongly influenced by the specific DNA sequence of the domain, as well as the action of nucleoid-associated proteins which may play a crucial role in organising the different types of spatial deformations, e.g., by modifying the
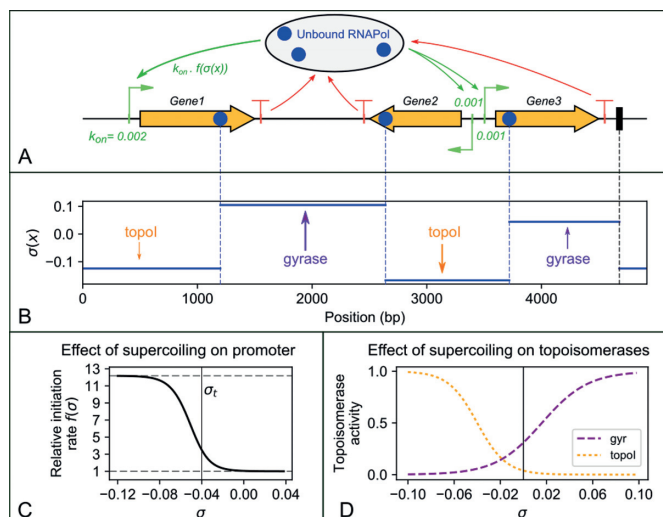
Figure 1. Illustration and main components of the transcription-supercoiling coupling model. (**A**) Snapshot of the simulation of the stochastic binding (green arrows; the basal initiation rate k_{on} of each promoter is shown), elongation, and dissociation (red arrows) of a set of RNAPs along a 1D genome (here a 5-kb plasmid). (**B**) The SC profile is updated at each timestep, and is affected by elongating RNAPs as well as by topoisomerase activity. This level is constant between topological barriers, i.e., either elongating RNAPs (blue) or fixed proteic barriers (black). (**C**) The local SC level affects each promoter through an activation curve derived from thermodynamics of open complex formation, which modulates its specific strength (basal initiation rate). (**D**) Topoisomerases bind in a deterministic but heterogeneous way, according to the local SC level (see text).

twist/writhe equilibrium of supercoiled DNA, or by promoting the localisation of gene promoters at apical loops located at the tip of plectonemic DNA, where they remain accessible to regulatory proteins (26). Our neglecting the topological action of nucleoid-associated proteins within domains is therefore consistent with a unidimensional description of the chromosome. These proteins are still taken into account either as topological barriers or by modulating promoter strengths (i.e. in the role of global transcription factors). Similarly, we disregard the presence of specific sequences favoring structural transitions of DNA (2,27) or DNA gyrase binding.

Transcriptomics data

Transcriptomes of *Dickeya dadantii* were analysed as previously described (10). *D. dadantii* 3937 cells were cultivated in M63 minimal medium supplemented with 0.2% (wt/vol) sucrose as carbon source, with or without 0.2% (wt/vol) polygalacturonate (PGA, a pectin derivative). Cells were harvested in early exponential phase ($OD_{600} = 0.2$, sucrose) or transition to stationary phase ($OD_{600} = 1.5$ in sucrose and $OD_{600} = 1.8$ in sucrose + PGA). For each condition, total RNAs were extracted using the frozen-phenol procedure (28), either from untreated cells or from cells treated with 100 μg/ml novobiocin for 15 min, with two biological replicates. At this concentration, novobiocin has no impact on *D. dadantii* growth (4). Control experiments for RNA extraction quality, absence of DNA contamina-

tion, and qRT-PCR validation of selected genes were conducted as previously (10). Further steps were carried out by Vertis Biotechnologie AG (http://www.vertis-biotech.com): rRNA depletion using the Illumina Ribo-Zero kit, RNA fragmentation, strand-specific cDNA library preparation, and Illumina NextSeq500 paired-end sequencing (∼15 million paired reads per sample).

Transcriptomes of *Escherichia coli* clones isolated from the Long-term Evolution Experiment (LTEE) (29,30) were obtained as previously described (31). Briefly, bacterial strains were grown in LTEE conditions (Davis minimal medium containing 25 μg/mL of glucose) but in 200 ml cultures to obtain sufficient amounts of RNA. Bacteria were harvested at mid-exponential phase and total RNAs were extracted from cells pellets using Qiagen's RNeasy Cell Tissue Kit, following the manufacturer's protocol. RNaseOUT (Invitrogen) was added to the RNA extracts that were subsequently handled and analysed by Beckman Coulter Genomics according to the standard Affymetrix GeneChip protocol for bacteria, and using the Affymetrix *E. coli* Genome 2.0 Microarray. Quality checks and normalisation were performed using the Affymetrix GeneChip Command Console and Expression Console according to standard Affymetrix protocols. The three strains used in this study were the ancestral strain of the LTEE, REL606, and two clones previously isolated from the Ara-1 population at 2000 (REL1164A, quoted '2K' below) and 20 000 (REL8593A, quoted '20K' below) generations (32). For each strain, 6 independent cultures were grown and 6 RNA extracts were analysed (18 microarrays in total).

Statistics and data analysis

Sequenced reads from *D. dadantii* were mapped on the reference genome (NCBI NC_014500.1) with Bowtie2, counted with htseq-count (33) against the reference annotation, and gene differential expression was analysed with DESeq2 (34). A threshold of 0.05 on the adjusted *P*-value was chosen to define differentially expressed genes (between 1250 and 1750 for novobiocin treatment).

The normalised dataset from *E. coli* (10 208 probes per microarray) was cured to only retain data from Affymetrix probes that entirely match within one coding sequence of the REL606 reference genome (NCBI NC_012967.1) and with the highest score of sequence homology in case several probes target the same gene. The cured dataset (3839 probes per assay) was obtained for all three strains (six replicates each) and $\log_2$ transformed. Differential expression was analysed with the optimal discovery procedure (35) implemented in the EDGE v1.1.290 software (36), and based on the false-discovery rate. Differentially expressed genes were first defined using a q-value threshold of 0.1; for the orientation analysis which required larger datasets (see below), a looser threshold of 0.25 on the *P*-value was used instead.

All processed data are available in a Supplementary File. The relation between gene response and local orientation was carried with a homemade Python code. All error bars shown are 95% confidence intervals. Proportions were compared with the χ^2 test.
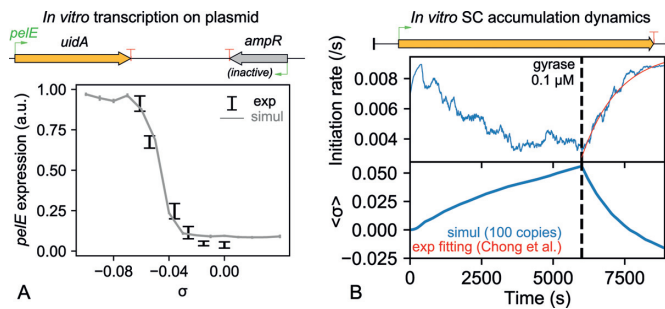
Figure 2. Calibration of the model on *in vitro* transcription experiments with plasmids. (**A**) The promoter activation curve (Figure 1C) is calibrated from *pelE* expression levels measured on purified plasmids prepared at different SC levels (4). Due to the absence of topological barriers in the plasmid, transcription-induced supercoils do not accumulate and SC levels remain constant. In this assay, this promoter from *Dickeya dadantii* is activated around 20-fold by negative SC. The weakly expressed promoter *ampR* was kept inactive in the simulations. (**B**) Topoisomerase activity constants are calibrated from a SC accumulation assay (22). The instantaneous initiation rate (top) was measured over time, and is reproduced by our simulations where the average plasmid SC level can be inferred (bottom). Positive supercoils first accumulate in the absence of DNA gyrase, resulting in a progressive repression of the promoter, and are then released in an exponential timecurve (red curve) reflecting gyrase activity.

Note that when we discuss the chromosomal SC level, the expression 'increased SC' refers to the absolute level (although negative), as often in microbiological literature.

RESULTS AND DISCUSSION

Stochastic modeling of the transcription-supercoiling dynamical coupling

The transcription-supercoiling coupling (TSC) is complex and dynamic, and must therefore be simulated with a suitable biophysical model. As previously proposed (17–19), we simulated the transcription dynamics of a circular genome (which can be a plasmid, but also an entire chromosome) as a unidimensional system, as illustrated on Figure 1 in a linearised depiction. At each timestep, (A) RNA Polymerases (RNAPs) bind stochastically at promoters (with different basal initiation rates k_{on}), transcribe the genes, and eventually unbind at terminators; and (B) the local distribution of SC is affected by transcription and by the action of the two major topoisomerases, DNA gyrase and topoisomerase I. These two aspects of the genome (gene expression and physical state) affect each other in a reciprocal way: (i) elongating RNAPs act as topological motors that pump positive supercoils from back to front, resulting in heterogeneous SC distributions; (ii) the initiation rate of each promoter is modulated by the local level of SC, following a single response function (Figure 1C). Qualitatively, these two effects are sufficient to generate a local form of transcriptional regulation mediated by SC (17,18). For a quantitative modeling of the process, a series of simplifications, hypotheses and additional components are required, as follows.

Within our unidimensional description, only the *twist* (torsional) contribution to SC is considered, rather than the complete (*twist* + *writhe*) level that involves 3D deforma-

tions of the double-helix (25). Torsional deformations are the main modulators of the free energy of DNA melting required for transcription initiation, which may constitute a major mechanism of regulation by SC *in vivo* (3,17), and was used to calibrate the regulatory part of the model. We assume that this regulation follows a single response curve (Figure 1C), obtained from the thermodynamic opening curve of a bacterial promoter (4,17).

The elongation speed is assumed to be constant, and to give rise to a constant rate of torsional SC generation (see Materials and Methods). This approximation signifies that the generated writhe can be effectively dissipated by topoisomerases (and especially the DNA gyrase) so that the elongating RNAP is not stalled by positive supercoils. Since the latter effect occurs at highly positive SC levels, this hypothesis differs from that of a previous model based on such a stalling effect (19), the latter being probably relevant for highly transcribed genes. In contrast, we note that elongation speeds of the majority of *E. coli* genes were found to be remarkably similar, even when their initiation rates differ by orders of magnitude (37). Based on the latter data, our modeling may thus be applicable to most moderately transcribed genes in the globally underwound bacterial genome, where the key regulatory step occurs at the initiation step and elongation does not generate comparably drastic mechanical constraints. This model might thus be inaccurate for very strong promoters like those of ribosomal RNAs (19), where the latter constraints may be handled by specific mechanisms disregarded here, such as a particular 3D organisation of the domain and high-affinity gyrase downstream binding sites (38).

As previously proposed (19,39), RNAP-generated supercoils are assumed to diffuse instantaneously (at the timescale of the simulation, *i.e.*, seconds) along kilobase distances, until reaching another RNAP or a fixed topological barrier, which can represent H-NS or other DNA-bound proteins with the same ability to isolate supercoils (3,6). The resulting instantaneous SC profile (Figure 1 B) appears as a succession of flat regions, in contrast to the more continuous and heterogeneous profiles expected if the diffusion was slower (18). Finally, the activities of topoisomerases are considered as deterministic and continuous, whereby DNA gyrase introduces negative supercoils and topoisomerase I relaxes DNA with nonlinear activation curves calibrated from experimental knowledge (see Figure 1D and Materials and Methods), without any sequence preference. As a result of these simplifications and hypotheses, the proposed model involves only a small number of mechanistic parameters, all of which are either obtained from experiments or calibrated on *in vitro* transcription assays, as follows (equations in Materials and Methods).

The promoter response curve was calibrated from the *pelE* promoter (Figure 2 A), which encodes a major virulence factor of the phytopathogenic enterobacterium *Dickeya dadantii*, and is strongly SC-sensitive (4). In this assay, plasmids carrying a *pelE* promoter were prepared at different SC levels; since they were free to rotate in solution and carried no topological barrier, the supercoils generated by

transcription merged instantaneously, and the SC level thus remained constant with time. The measured expression levels thus directly reflected the promoter response curve (Figure 1C) and match the dependency expected from promoter opening thermodynamics (17). Note that this curve is not specific to *pelE*, but is compatible with expression levels measured on various promoters of several species (Supplementary Figure S1).

Conversely, in order to calibrate the dynamics of SC accumulation resulting from topoisomerase activity, we simulated an experiment involving a plasmid with a long (12 kb) gene, anchored to a surface that hampers rotation (Figure 2B and (22)). Topoisomerase I was first introduced at a well-controlled concentration (41 nM): as a result, during each elongation event, only the negative twin-domain was relaxed by topoisomerase I, whereas the positive supercoils were not relaxed; this unbalance resulted in a slow accumulation of positive supercoils (in a timescale of hours), and promoter repression (by a factor 3). The sudden addition of DNA gyrase (at 0.1 μM) then restored the initiation rate within <1 h. After calibrating the topoisomerase activity constants, simulations closely reproduce the observed behaviour, and allow inferring the underlying SC level of the plasmid. Note that in this artificial construct involving strongly positive SC levels, the latter probably hampers transcription not only at the initiation step, but also during elongation (22); accordingly, the standard initiation curve (Figure 1C) had to be replaced by an effective curve repressed at positive SC levels (Supplementary Figure S2A). We now show that in spite of its simplicity, our model is able to reproduce quantitative measurements of TSC-induced transcriptional interaction between adjacent genes in well-controlled *in vivo* experimental setups.

### Torsional interaction between adjacent genes *in vivo*

The asymmetric nature of TSC implies that adjacent genes interact in a strongly orientation-dependent manner: divergent genes are expected to activate each other, whereas convergent genes experience a mutually repressive interaction (17). To test this effect, Sobetzko (14) inserted a cassette containing two closely located divergent genes into the *Escherichia coli* chromosome (Figure 3 A), one of them being inducible (*tetP*). When the inducer concentration was varied, the expression of both genes changed in a coordinated manner, which was attributed to an activation of the second (non-inducible) gene by SC. We simulated this construct, using its exact architecture and physiologically relevant parameters (see Materials and Methods). Simulations reproduced a mutual activation when the basal level of *tetP* was progressively increased (Figure 3A, right panel). Interestingly, since both promoters are always located in the same topological domain and experience the same SC level in our simulations, the observed increase in the expression of *tetP* results not only from the inducer concentration, but also from the resulting increase in negative SC in the central region (that is responsible for *tufB* activation). As can be observed, the activation of *tetP* is only semi-quantitatively reproduced, as the simulated levels remain below experimental ones: this small discrepancy may be indicative of a differ-
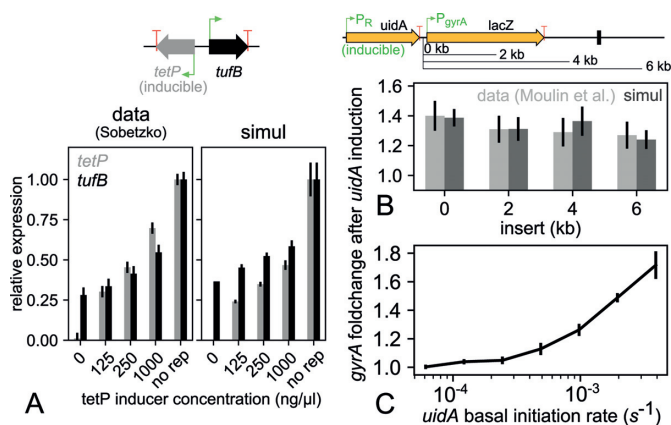
Figure 3. *In vivo* TSC-induced co-regulation of adjacent genes on the *Escherichia coli* chromosome. (**A**) The expression of the inducible *tetP* promoter (by increasing *tet* inducer concentration or absence of *tet* repressor) progressively activates a divergent promoter (*tufB*) (14). This effect is reproduced semi-quantitatively in our simulations by only changing the *tetP* basal initiation rate; small discrepancies with the data may reflect a different SC-sensitivity of the two promoters (see text). (**B**) The *gyrA* promoter is activated by the expression of the upstream isodirectional gene *uidA*, almost independently of the distance between the two genes (from 0 to 6 kb) (16). This observation is fully consistent with our hypothesis of fast diffusion of SC over kilobase distances. A reversed promoter activation curve was used here to account for the unusual properties of the *gyrA* promoter (see text and Supplementary Figure S2B). (**C**) Simulations (without insert) show that the *gyrA* activation factor is strongly dependent on the expression strength of the *uidA* gene.

ent SC-sensitivity for the two promoters (weaker for *tufB*), in contrast to our unique response curve (Figure 1C).

This experiment shows that 300-bp distant promoters can significantly influence each other through TSC; however, can the resulting supercoils then diffuse at kilobase distances, as relevant to most genomic loci? Moulin *et al.* (16) measured the induction of the *gyrA* promoter by an upstream inducible gene located at various distances up to more than 6 kb (Figure 3B). Note that *gyrA* is a very unusual promoter, as its SC-sensitivity goes opposite to most other promoters with an activation by DNA relaxation, which ensures homeostasis of the SC level in the cell. In the considered construct, this activation is produced by the positive supercoils resulting from the activity of the upstream *uidA* gene. Strikingly, the activation level is almost independent of the distance between the two genes, showing that TSC may be able to couple most genes located within topological domains of 10–20 kb (40), probably resulting in complex collective interactions. By construction (instantaneous diffusion), our model allows SC to diffuse at kilobase distances until it reaches a topological barrier; it is therefore no surprise that the experimental profile is easily reproduced under physiological topoisomerase concentrations (Figure 3B), using an unusually reversed promoter activation curve relevant to *gyrA* (Supplementary Figure S2B). The level of activation then depends on the expression strength of the upstream gene (Figure 3C); the experimentally observed value is obtained for a moderate rate of one transcription event every 5 min. Altogether, these results show that our modeling quantitatively reproduces the effect of TSC on well-defined constructions on the chromo-
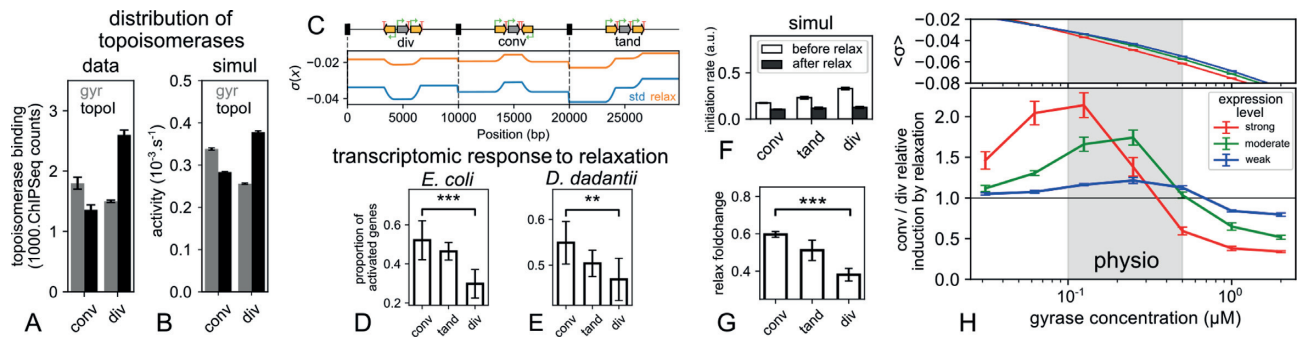
Figure 4. (**A**) ChIP-Seq data show that topoisomerase I and the DNA gyrase preferentially bind in divergent and convergent regions, respectively (43). (**B**) The observed distribution naturally emerges from our modeling of TSC in physiologically relevant conditions (moderate gene expression and 0.25 μM gyrase concentration, see Materials and Methods). (**C**) Simulations were carried on a model genome with three distinct topological domains, each carrying two active genes (yellow) in different orientations, and a central inactive gene (gray) as a regulatory probe. The heterogeneous SC levels are shown before (blue) and after (orange) relaxation by gyrase inhibition, and they give rise to the heterogeneous recruitment of topoisomerases observed in B. (**D**) Transcriptomics data in *Escherichia coli* from (9) show that genes' response to chromosomal relaxation is tightly related to their local orientation. (**E**) The same is observed in new RNA-Seq data from *Dickeya dadantii*, in exponential as well as at transition to stationary (Supplementary Figure S4B) phase, confirming that this feature is not organism- or condition-specific. (**F**) Simulating the action of the antibiotics exhibits a different effect on the promoters' initiation rates, depending on their local orientation. (**G**) As a result, relaxation favours convergent genes versus divergent ones, in agreement with genome-wide experimental data. (**H**) The average SC level (upper panel) and relative convergent/divergent foldchange due to relaxation (lower panel) was computed for a range of initial DNA gyrase concentrations and expression strengths (of all active genes), corresponding to effective waiting times of 1, 2.5 and 10 min between transcription events at each promoter, respectively. Data shown in B, C, F and G correspond to the central datapoint with moderate expression.

some; we can now turn our attention to its contribution to the global regulation by SC along entire genomes.

Topoisomerase distribution and regulatory activity on the chromosome is dictated by genomic architecture

As shown above, adjacent genes interact on the chromosome through TSC, according to complex yet predictable rules that already emerge from our simplified modeling. Previous studies already showed that, in contrast to the 'classical dogma' of transcriptional regulation, the expression of a promoter is affected by its position on the genome (41); our present results also emphasize an additional orientational effect. Considering the density of bacterial genomes, we may thus expect that the mechanical constraints induced by transcription of adjacent genes, and tightly related to their orientation, play an important role in the genome-wide coordination of gene expression. To directly test this effect, the local distribution of supercoils along the chromosome should be measured, but the only available data of this kind lack sufficient spatial resolution (42). However, an indirect measurement was provided by ChIP-Seq of topoisomerases. In *Mycobacterium tuberculosis*, the binding distributions of topoisomerase I and DNA gyrase (43) were found to be enriched in divergent and convergent regions respectively (Figure 4 A), in agreement with our expectations based on a simplified description of TSC (Figure 4B) where these preferences reflect the different SC levels induced by neighbouring transcription rather than sequence preferences (Figure 1D). This result is further confirmed by recent data in *E. coli* obtained with poisons that trap the DNA gyrase at its activity sites at high resolution (38). These data not only show that convergent regions are enriched in gyrase activity (rather than nonspecific binding) (Supplementary Figure S3A), but the distributions obtained after treatment with the transcription inhibitor rifampicin also directly demonstrate a contribution of adja-

cent transcription in this activity profile, in addition to possible sequence-dependent effects disregarded in our modeling (Supplementary Figure S3B).

In order to simulate and reproduce such *in vivo* genome-wide data (Figure 4B), we reasoned that simulating an entire chromosome would involve many arbitrary choices (definition of topological domains and transcription units, gene expression strengths...). Since most observed properties depend primarily on gene orientations, more than on precise gene positions and distances, we rather decided to simulate a 30-kb-long toy model circular genome involving three topological domains (Figure 4C), each of them carrying two identical 1-kb-long active genes (yellow) in different configurations, and a central inactive gene (gray) which does not experience any transcription elongation but is used as a regulatory probe. All promoters have the same basal initiation rate and SC-dependence: any observed difference in the initiation rate of inactive promoters thus directly reflects the effect of TSC induced by their neighbours, according to their orientation. The relatively small differences in topoisomerases binding observed between convergent and divergent regions (Figure 4A) are reproduced (Figure 4B) when the genes are moderately expressed in these simulations (one transcript every 2.5 min), as expected for most of them in the genome.

In bacteria, the regulatory effect of SC is generally analysed from transcriptomes obtained shortly after chromosome relaxation by antibiotics (novobiocin, norfloxacin) which inhibit the DNA gyrase (8–11). Based on the previous observations, the picture of a global relaxation might yet appear as slightly misleading, if the effect of gyrase inhibition is very different in convergent versus divergent regions. In this case, spatial topological heterogeneities resulting from TSC may play an important role in the differential response of genes to a global change in gyrase activity, which is itself a key actor in the cellular response to environmental variations (3,6). To test this hypothesis, we analysed
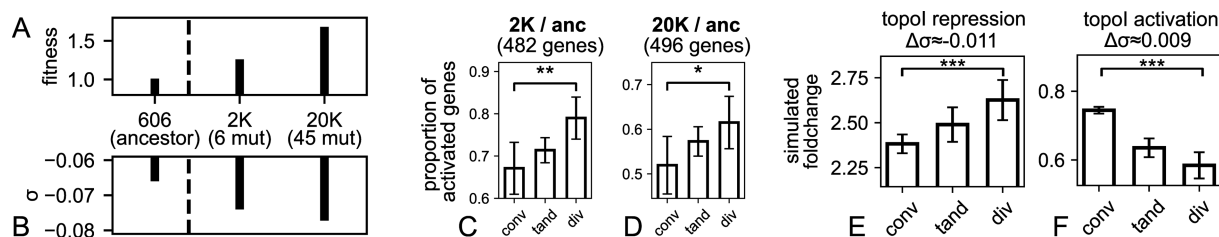
Figure 5. Gene expression profiles in clones from the long-term evolution experiment with *E. coli* (29,30) exhibits the signature of TSC-mediated regulation. (**A**) 2K and 20K clones from the Ara-1 population evolved increased fitness relative to their ancestor after 2000 and 20,000 generations, respectively (data from (7)). (**B**) The SC level sequentially increased, owing to the selection of two SC-modifying mutations (data from (7)), first in *topA* (among six mutations in 2K), then in *fis* (among 45 mutations in 20K, including the *topA* mutation) (32). (**C**) In 2K, divergent genes are more activated than convergent ones with respect to the ancestor ($P = 1.6 \times 10^{-3}$), as predicted by TSC. Differentially expressed genes were selected with a loose threshold (see text): we indicate the number of convergent+divergent responding genes. (**D**) Same for the 20K clone ($P = 0.016$). (**E**) We simulated the mutation in *topA* by reducing the activity of topoisomerase I by 2-fold: as expected and observed *in vivo*, divergent genes are favored by the resulting increase in negative SC. (**F**) Conversely, a 2-fold increase in topoisomerase I activity favours convergent genes, in the same manner as gyrase inhibition (Figure 4G).

several transcriptomes obtained under different conditions of chromosomal relaxation shock. In *E. coli*, they were obtained by microarrays, either from mutant strains favouring a strong relaxation by norfloxacin (9) (Figure 4 D), or with novobiocin in the wild-type strain (8) (Supplementary Figure S4A in Supplementary Materials). In both cases, a significant difference is observed in the response of convergent vs divergent genes ($P = 2.3 \times 10^{-4}$ and $P = 0.02$ respectively, χ^2 test). Since we are investigating a mechanism relying on basal properties of transcription shared by many different bacterial species, we assessed the universality of this observation by conducting RNA-Seq experiments on *Dickeya dadantii*. Transcriptomes were collected 15 min after a novobiocin shock applied either in exponential phase (Figure 4E) or, for the first time in enterobacteria, at the transition to stationary phase (Supplementary Figure S4B); in both conditions, the same difference is observed as in *E. coli* ($P = 8 \times 10^{-3}$ and $P = 1.6 \times 10^{-3}$, respectively). Altogether, in spite of strong differences in organism, experimental conditions and methods, a systematic and coherent relation between gene orientation and transcriptional response is observed in all data. But interestingly, while we intuitively expected convergent genes to be more hampered by gyrase inhibition (since the latter is more active in these regions), the reverse effect is observed, as they systematically appear more activated by chromosome relaxation than divergent genes. This effect was already noted by Sobetzko (14), who invoked a putative difference in promoter sequences, suggesting that evolution has placed relaxation-activated promoters in convergent regions where DNA is more relaxed. However, in view of our previous observations, gyrase inhibition probably affects convergent vs divergent promoters in a different manner. If the *local* SC relaxation is different in these two regions, the latter effect might explain why RNAPs are redistributed in favor of the convergent promoters, even if these exhibit exactly the same 'SC-sensitivity' as divergent ones, i.e., without invoking any sequence effect. Because of the complexity of the process, this counter-intuitive scenario cannot be easily predicted, but it can be tested within our modeling where all promoters have the same response curve (Figure 1C) and gyrase inhibition can be simulated explicitly. Using the same parameters as in Figure 4B, we simulate one hour of stationary transcription, and then mimic the action of antibiotics by suddenly divid-

ing the gyrase concentration by 5; the initiation rate before and after the shock are then compared (Figure 4 F). As expected, all initiation rates are reduced by the global relaxation of the chromosome, but by different factors: more than 2.5 for divergent genes, against less than 2 for convergent genes (Figure 4G). Remarkably, the shock thus leads to a redistribution of RNAPs in (relative) favour of convergent genes: the model therefore predicts a stronger activation of the latter in transcriptomics experiments (where gene expression levels are normalized in each condition), precisely as systematically observed. Interestingly also, the simulations make it possible to analyse the mechanism responsible for this counter-intuitive behaviour, i.e., the different SC levels in the convergent vs divergent regions before and after the shock (Figure 4C). These should be compared with the promoter activation curve shown on Figure 1C, where the SC-sensitivity is strongest between –0.08 and –0.02, and much lower outside this range. Before the shock, the convergent gene was already located in a relaxed region because of TSC ($\sigma \simeq -0.03$), and was thus already partly repressed, whereas the divergent gene was in a strongly negative region and thus much more active (blue curve). In contrast, after the relaxation shock, both promoters have been shifted to a level around –0.02, thus to a similarly repressed state, explaining why the repressive effect is stronger for the divergent gene. Note that since gene expression is globally reduced after the shock, TSC-induced spatial superhelical variations are weaker than before, hence a slightly weaker amount of SC relaxation in the convergent configuration (vertical distance between the blue and orange curves).

Interplay between gene expression strength and topoisomerase activity in global regulation by TSC

Since generic rules of basal regulation by TSC seem well-reproduced by our model, we can now use it to explore how these rules and the whole system behaviour depend on the precise conditions of the simulation. We tested the influence of two biologically relevant parameters: (i) gene expression strength (of all active genes in our model genome) and (ii) DNA gyrase concentration. The latter parameter actually also mimics the controlled variations of gyrase activity in the cell in response to environmental constraints, in particular through variations of the ATP/ADP ratio

in exponentially growing vs stationary cells (44,45). Figure 4H shows the average superhelical density and the relative convergent/divergent activation factor due to gyrase inhibition, for different gyrase concentrations and basal expression rates, corresponding to effective waiting times of ∼1, 2.5 and 10 min between transcription events of each active gene respectively. Unsurprisingly, the negative SC level (upper panel) increases with gyrase concentration; but interestingly, this level also strongly depends on the gene expression strength, with differences as large as $\Delta\sigma \simeq 0.01$ for physiological levels of DNA gyrase. This difference is of the same order as the measured effects of novobiocin (4) or topoisomerase mutations (see below). Actually, such a behavior has already been observed experimentally on a plasmid, which was found more supercoiled in the cell when it carried a strong than a weak promoter (46). Our simulations show that it results from the nonlinear response curve of topoisomerases in presence of TSC (Figure 1D): since highly expressed domains experience strong transient SC inhomogeneities, topoisomerases sometimes bind very actively and non-uniformly in these domains, resulting in a stationary stronger SC level. Significant consequences result from this observation: (a) in a living cell, the local SC level is likely higher in strongly than weakly expressed topological domains (in addition to orientation-dependent effects already noted); (b) the SC level measured on a plasmid (e.g. in a chloroquine gel experiment) only semi-quantitatively reflects the chromosomal level, as both levels are affected by the presence/absence and activity of promoters.

Altogether, these observations show that defining the 'supercoiling-sensitivity' of promoters purely from transcriptomics experiments is delicate, since the SC levels *locally* experienced by these promoters on the chromosome may differ quite strongly from the *global* levels measured from plasmids, and even more so when considering other factors disregarded here, such as nucleoid-associated proteins (3,6,25). Note that the SC levels observed in the simulations are somewhat lower than those usually measured *in vivo* on plasmids, which is consistent with our neglecting the writhe contribution (and keeping in mind the previous reservations). Simulations of gyrase inhibition (Figure 4H, lower panel) also show that the relative induction of convergent vs divergent genes (datapoints above 1) is observed at all gene expression levels (and more strongly for highly expressed genes), but not when the initial gyrase concentration is very high. In particular, strongly expressed convergent genes might even be repressed by gyrase inhibition in physiological conditions (right part of red curve), as one would expect if these genes specifically require DNA gyrase to stay active. This effect would probably be even stronger had we considered the stalling effect of positive supercoils on transcription elongation (19).

TSC-mediated regulation in experimental evolution

If TSC constitutes an ancestral mode of regulation, it might play an important role in genome evolution. An interesting example is that of synteny segments, i.e., genes remaining adjacent in evolution, which exhibit distinct orientation and co-expression features currently unexplained (47). Here, we focus on the longest-running evolution experiment, during which SC was indeed identified as a critical adaptive actor (7). In this experiment, 12 independent *E. coli* populations were grown by serial daily transfer in minimal medium for 50 000 generations (30). In one representative population, fitness already strongly increased after 2000 generations (Figure 5A) together with the SC level within the cells (by $\Delta\sigma \simeq -0.008$ as measured on a plasmid, Figure 5B). These changes result from only six mutations (32), including one beneficial SNP in the *topA* gene encoding topoisomerase I, which reduces its efficiency. A second increase in SC (of an additional $\Delta\sigma \simeq -0.003$) occurred before 20 000 generations, owing to the fixation of another beneficial mutation affecting *fis* among 45 mutations in total (7). It was proposed that SC-modifying mutations provided an efficient way to change the cell expression program globally (7). We then expect TSC to contribute in this transcriptional response, probably in an opposite direction to that observed in relaxation shocks (since the SC level is here increased).

To test this putative effect of TSC, we collected the global transcription profiles of the ancestral and evolved strains using microarrays (see Materials and Methods). Evolved strains exhibit differentially expressed genes with respect to the ancestor, yet their number is relatively low under standard statistical selection procedures compared to relaxation shock experiments (97 and 148 genes in the 2K and 20K strains, respectively; *q*-value <0.1). However, in contrast to the latter case, here the alterations in gene expression are not caused by a fast change in SC, but rather by an inheritable SC change, the consequences of which can be balanced by other actors of the regulatory network, including transcription factors. Since the few identified genes with altered expression are those where the activation/repression is strongest, most of these changes are probably due to the strong and localised action of 'digital' (on/off) transcription factors. Indeed, SC alone usually exhibits a more subtle and 'analogue' (more/less) regulatory effect on a larger number of genes (3,48). To analyse the latter effect, we therefore generated new datasets of differentially expressed genes, in numbers similar to those obtained in relaxation shock experiments, by reducing the threshold of statistical significance (Figure 4, *P*-value < 0.25). This operation implies that these larger datasets likely contain many false positives which might blur out the investigated effects of TSC, but the latter may still emerge as a statistical feature. We therefore looked for a statistical relationship between gene orientation and response (Figure 5C, D). Both evolved clones with increased SC level indeed exhibit a higher rate of activated divergent genes, as expected from our previous analysis. The confidence level is high for the 2K strain where there are few other mutations ($P = 1.6 \times 10^{-3}$), and also significant for the 20K strain ($P = 0.016$) which exhibits more mutations that could contribute to rewire the regulatory network independently of SC effects. Note that the observed behaviour does not depend on the precise threshold chosen above and is robust with more stringent selection of differentially expressed genes, although with weaker statistical significance owing to their lower number (data not shown). To crosscheck if it is indeed reproduced by the model, we ran simulations involving decreased (Figure 5E) or increased

(Figure 5F) topoisomerase I activity (by 2-fold, compatible with the measured SC variation $\Delta\sigma \simeq -0.011$). They confirmed a relationship between relative activities of topoisomerases and orientation-dependent response of genes. Altogether, these analyses show that (i) TSC defines robust rules that allow the cell to selectively activate convergent or divergent genes by tuning its global SC level, independently from more specific regulation of each promoter; (ii) experimentally evolved populations in which SC-modifying mutations were selected exhibit changes in global transcription profiles consistent with the TSC mode of regulation; (iii) at larger evolutionary timescales, TSC might also contribute to the selection of events such as insertions/deletions or inversions, which could modify the expression profile by changing the relative distances and orientations between adjacent genes, even without modifying the gene coding or promoter sequences.

CONCLUSION

Many bacterial genes exhibit 'supercoiling sensitivity' in transcriptomics experiments, the mechanisms of which remain unclear. Inspired by classical regulation models, most efforts have focused on promoter sequences to explain their response (3,8). Here, we propose alternatively that strong spatial variations of DNA supercoiling along the chromosome equally contribute to the complexity of this response, and allow the cell to activate its genes in a global but selective way depending on their local orientation. Such variations were measured at low resolution in the bacterial chromosome (42), and were already linked to gene orientations in eukaryotes (20), but these effects remained nonetheless as yet neglected in quantitative regulatory models. Those data obtained in very different organisms suggested that the fundamental mechanical properties of transcription give rise to an ancestral coupling between genome expression and torsion (here called TSC). This coupling underpins a basal form of regulation that might affect all organisms, albeit with different rules owing to differences in genome structures, topoisomerase enzymes, etc. In bacteria, we have analysed here some of these complex yet predictable rules by which global variations of DNA topology are distributed at gene promoters, and allow for a fine-tuned global regulation mode under the control of cell physiology. The model presented, which was kept voluntarily as simple as possible, is already remarkably predictive of how the genome architecture dictates this regulation by mechanically coupling adjacent genes as a result of the transcription process itself, and depending on their relative orientations. As a further step, the model may be used to directly simulate a larger chromosomal region *in vivo*, which will raise new substantial issues regarding the proper definition of topological barriers, transcription units, and other required components. An especially attractive application is the so-called 'pathogenicity islands' where virulence genes are co-localized and remain co-regulated when they are horizontally transferred from one pathogenic species to another (6). The model predictions should then be experimentally tested with more detail and in a wider range of bacterial organisms. Such analyses will likely exhibit more subtle regulatory effects involving nucleoid-associated proteins and other actors disregarded here, which contribute in the 3D organisation of the chromosome (25). We anticipate that this comparison will lead to an incremental refinement of the modeling (in particular by incorporating sequence-dependent properties of DNA and DNA-protein interactions), toward a more comprehensive description of the global transcriptional regulation embedded in the bacterial chromatin structure.

DATA AVAILABILITY

Transcriptomes have been deposited at ArrayExpress repository E-MTAB-7650 (*Dickeya dadantii* RNA sequencing dataset) and Gene Expression Omnibus repository GSE126600 (*Escherichia coli* LTEE microarray dataset).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Leblanc,B.P., Benham,C.J. and Clark,D.J. (2000) An initiation element in the yeast CUP1 promoter is recognized by RNA polymerase II in the absence of TATA box-binding protein if the DNA is negatively supercoiled. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 10745–10750.
2. Hatfield,G.W. and Benham,C.J. (2002) DNA topology-mediated control of global gene expression in Escherichia coli. *Annu. Rev. Gen.*, **36**, 175–203.
3. Travers,A. and Muskhelishvili,G. (2005) Bacterial chromatin. *Curr. Opin. Gen. Dev.*, **15**, 507–514.
4. Ouafa,Z.-A., Reverchon,S., Lautier,T., Muskhelishvili,G. and Nasser,W. (2012) The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium Dickeya dadantii. *Nucleic Acids Res.*, **40**, 4306–4319.
5. Gilbert,N. and Allan,J. (2014) Supercoiling in DNA and chromatin. *Curr. Opin. Gen. Dev.*, **25**, 15–21.
6. Dorman,C.J. and Dorman,M.J. (2016) DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys. Rev.*, **8**, 209–220.
7. Crozat,E., Philippe,N., Lenski,R.E., Geiselmann,J. and Schneider,D. (2005) Long-term experimental evolution in Escherichia coli. XII. DNA topology as a key target of selection. *Gen.*, **169**, 523–532.

8. Peter,B.J., Arsuaga,J., Breier,A.M., Khodursky,A.B., Brown,P.O. and Cozzarelli,N.R. (2004) Genomic transcriptional response to loss of chromosomal supercoiling in Escherichia coli. *Genome Biol.*, **5**, R87.
9. Blot,N., Mavathur,R., Geertz,M., Travers,A. and Muskhelishvili,G. (2006) Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep.*, **7**, 710–715.
10. Jiang,X., Sobetzko,P., Nasser,W., Reverchon,S. and Muskhelishvili,G. (2015) Chromosomal "stress-response" domains govern the spatiotemporal expression of the bacterial virulence program. *mBio*, **6**, e00353-15.
11. Junier,I., Unal,E.B., Yus,E., Lloréns-Rico,V. and Serrano,L. (2016) Insights into the mechanisms of basal coordination of transcription using a genome-reduced bacterium. *Cell Syst.*, **2**, 391–401.
12. Brinza,L., Calevro,F. and Charles,H. (2013) Genomic analysis of the regulatory elements and links with intrinsic DNA structural properties in the shrunken genome of Buchnera. *BMC Gen.*, **14**, 73.
13. Liu,L.F. and Wang,J.C. (1987) Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 7024–7027.
14. Sobetzko,P. (2016) Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. *Nucleic Acids Res.*, **44**, 1514–1524.
15. Opel,M.L. and Hatfield,G. (2001) DNA supercoiling-dependent transcriptional coupling between the divergently transcribed promoters of the ilvYC operon of Escherichia coli is proportional to promoter strengths and transcript lengths. *Mol. Microbiol.*, **39**, 191–198.
16. Moulin,L., Rahmouni,A.R. and Boccard,F. (2005) Topological insulators inhibit diffusion of transcription-induced positive supercoils in the chromosome of Escherichia coli. *Mol. Microbiol.*, **55**, 601–610.
17. Meyer,S. and Beslon,G. (2014) Torsion-Mediated Interaction between Adjacent Genes. *PLoS Comput. Biol.*, **10**, e1003785.
18. Brackley,C., Johnson,J., Bentivoglio,A., Corless,S., Gilbert,N., Gonnella,G. and Marenduzzo,D. (2016) Stochastic model of supercoiling-dependent transcription. *Phys. Rev. Lett.*, **117**, 018101.
19. Sevier,S.A. and Levine,H. (2018) Properties of gene expression and chromatin structure with mechanically regulated elongation. *Nucleic Acids Res.*, **46**, 5924–5934.
20. Naughton,C., Corless,S. and Gilbert,N. (2013) Divergent RNA transcription. *Transcr.*, **4**, 2154.
21. Bohrer,C.H. and Roberts,E. (2016) A biophysical model of supercoiling dependent transcription predicts a structural aspect to gene regulation. *BMC Biophys.*, **9**, 2.
22. Chong,S., Chen,C., Ge,H. and Xie,X.S. (2014) Mechanism of Transcriptional Bursting in Bacteria. *Cell*, **158**, 314–326.
23. Nollmann,M., Crisona,N.J. and Arimondo,P.B. (2007) Thirty years of Escherichia coli DNA gyrase: from in vivo function to single-molecule mechanism. *Biochimie.*, **89**, 490–499.
24. Schoeffler,A.J. and Berger,J.M. (2008) DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Q. Rev. Biophys.*, **41**, 41–101.
25. Lioy,V.S., Cournac,A., Marbouty,M., Duigou,S., Mozziconacci,J., Espéli,O., Boccard,F. and Koszul,R. (2018) Multiscale structuring of the E. coli chromosome by nucleoid-associated and condensin proteins. *Cell*, **172**, 771–783.
26. Travers,A. and Muskhelishvili,G. (2007) A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep.*, **8**, 147–151.
27. Scott,S., Xu,Z.M., Kouzine,F., Berard,D.J., Shaheen,C., Gravel,B., Saunders,L., Hofkirchner,A., Leroux,C., Laurin,J. et al. (2018) Visualizing structure-mediated interactions in supercoiled DNA molecules. *Nucleic Acids Res.*, **46**, 4622–4631.
28. Maes,M. and Messens,E. (1992) Phenol as grinding material in RNA preparations. *Nucleic Acids Res.*, **20**, 4374.
29. Lenski,R.E., Rose,M.R., Simpson,S.C. and Tadler,S.C. (1991) Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. *Am. Nat.*, **138**, 1315–1341.
30. Tenaillon,O., Barrick,J.E., Ribeck,N., Deatherage,D.E., Blanchard,J.L., Dasgupta,A., Wu,G.C., Wielgoss,S., Cruveiller,S., Médigue,C. et al. (2016) Tempo and mode of genome evolution in a 50,000-generation experiment. *Nat.*, **536**, 165–170.
31. Le Gac,M., Plucain,J., Hindré,T., Lenski,R.E. and Schneider,D. (2012) Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 9487–9492.
32. Barrick,J.E., Yu,D.S., Yoon,S.H., Jeong,H., Oh,T.K., Schneider,D., Lenski,R.E. and Kim,J.F. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli. Nat*, **461**, 1243–1247.
33. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinform*, **31**, 166–169.
34. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Gen. Biol.*, **15**, 550.
35. Storey,J.D., Dai,J.Y. and Leek,J.T. (2007) The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. *Biostat.*, **8**, 414–432.
36. Leek,J.T., Monsen,E., Dabney,A.R. and Storey,J.D. (2006) EDGE: extraction and analysis of differential gene expression. *Bioinform*, **22**, 507–508.
37. Chen,H., Shiroguchi,K., Ge,H. and Xie,X.S. (2015) Genome-wide study of mRNA degradation and transcript elongation in Escherichia coli. *Mol. Syst. Biol.*, **11**, 808.
38. Sutormin,D., Rubanova,N., Logacheva,M., Ghilarov,D. and Severinov,K. (2018) Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the Escherichia coli genome. *Nucleic Acids Res.*, **47**, 1373–1388.
39. Lesne,A., Victor,J.-M., Bertrand,E., Basyuk,E. and Barbi,M. (2018) The Role of Supercoiling in the Motor Activity of RNA Polymerases. In: Lavelle,C (ed). *Molecular Motors: Methods and Protocols.* Springer, NY, pp. 215–232.
40. Postow,L., Hardy,C.D., Arsuaga,J. and Cozzarelli,N.R. (2004) Topological domain structure of the Escherichia coli chromosome. *Gen. Dev.*, **18**, 1766–1779.
41. Gerganova,V., Berger,M., Zaldastanishvili,E., Sobetzko,P., Lafon,C., Mourez,M., Travers,A. and Muskhelishvili,G. (2015) Chromosomal position shift of a regulatory gene alters the bacterial phenotype. *Nucleic Acids Res.*, **43**, 8215–9226.
42. Lal,A., Dhar,A., Trostel,A., Kouzine,F., Seshasayee,A.S.N. and Adhya,S. (2016) Genome scale patterns of supercoiling in a bacterial chromosome. *Nat. Commun.*, **7**, 11055.
43. Ahmed,W., Sala,C., Hegde,S.R., Jha,R.K., Cole,S.T. and Nagaraja,V. (2017) Transcription facilitated genome-wide recruitment of topoisomerase I and DNA gyrase. *PLoS Gen.*, **13**, e1006754.
44. Hsieh,L.S., Rouviere-Yaniv,J. and Drlica,K. (1991) Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. *J. Bact.*, **173**, 3914–3917.
45. van Workum,M., van Dooren,S.J., Oldenburg,N., Molenaar,D., Jensen,P.R., Snoep,J.L. and Westerhoff,H.V. (1996) DNA supercoiling depends on the phosphorylation potential in Escherichia coli. *Mol. Microbiol.*, **20**, 351–360.
46. Geertz,M., Travers,A., Mehandziska,S., Sobetzko,P., Chandra-Janga,S., Shimamoto,N. and Muskhelishvili,G. (2011) Structural coupling between RNA polymerase composition and DNA supercoiling in coordinating transcription: a global role for the omega subunit? *mBio*, **2**, e00034-11.
47. Junier,I. and Rivoire,O. (2016) Conserved units of co-expression in bacterial genomes: an evolutionary insight into transcriptional regulation. *PLoS One*, **11**, e0155740.
48. Marr,C., Geertz,M., Hütt,M.-T. and Muskhelishvili,G. (2008) Dissecting the logical types of network control in gene expression profiles. *BMC Syst. Biol.*, **2**, 18.