Laboratory evolution of anticipatory gene regulation in *Escherichia coli.*

Anjali Mahilkar, Akshat Mall, Supreet Saini*

Department of Chemical Engineering, Indian Institute of Technology Bombay, India 400076

* Corresponding author (saini@che.iitb.ac.in, 091 22 2576 7216)

Abstract.

Environmental cues in an ecological niche are often temporal in nature. For instance, in temperate climates, temperature is higher in daytime compared to during night. In response to these temporal cues, bacteria have been known to exhibit anticipatory regulation, whereby they trigger response to a yet to appear cue, anticipating its actual arrival in the near future. Such an anticipatory response in known to enhance Darwinian fitness, and hence, is likely an important feature of regulatory networks in microorganisms. However, the conditions under which an anticipatory response evolves as an adaptive response are not known. In this work, we develop a quantitative model to study response of a population to two temporal environmental cues, and predict variables which are likely important for evolution of anticipatory regulatory response. We follow this with experimental evolution of *E. coli* in alternating environments of a pentose sugar, rhamnose, and an oxidative stress molecule, paraquat for more than 800 generations. We demonstrate that growth in this cyclical environment leads to evolution of anticipatory regulation, whereby, exposure to rhamnose leads to partial induction of the oxidative stress response regulon. As a result, pre-exposure to rhamnose leads to a greater fitness in paraquat environment. Overall, we show that in niches where environmental stimuli have a cyclical nature, anticipatory regulation can evolve as an adaptive strategy, in a time course of a few hundred generations. This contributes to our understanding of how environment shapes the topology of regulatory networks in an organism.

Keywords: Adaptation, anticipatory regulation, *E. coli*, laboratory evolution, rhamnose, paraquat

Introduction.

Specificity is key in biological regulation. Biological systems expend considerable effort towards reducing non-specific substrate binding and reducing crosstalk [1-4]. However, despite this, regulatory crosstalk between cellular modules is ubiquitous in biology [5,6]. One of the forms of specificity and crosstalk in regulation is between transcription factors and their recognition site on the DNA. Although each transcription factor binds its cognate regulatory site on the DNA, significant non-specific binding, and hence regulatory crosstalk, exists [7,8].

A particular form of crosstalk between two functional modules is when gene expression changes upon anticipation of an upcoming environmental shift. For example, in *E. coli*, an increase in temperature elicits response to low oxygen conditions, mimicking how these two environments are sequentially encountered by the bacterium in the mammalian gastrointestinal tract [9]. Such anticipatory regulation was also demonstrated to occur between the sugars lactose and maltose. In this case, exposure to lactose leads to partial activation of maltose-utilization genes [10]. In the intestinal gut, exposure to lactose precedes that to maltose. As a result, this partial gene activation provides an adaptive benefit to the organism.

Anticipatory regulation is seen in yeast too. For example, different environmental isolates exhibit varying degrees of glucose-dependent catabolite repression on maltose-utilization genes [11]. When exposed to a mixture of glucose and maltose, while some strains exhibit strict hierarchical utilization of the two sugars, separated by a diauxic lag, others exhibit minimal lag, and co-utilization of the two sugars. Thus, the regulatory crosstalk between the glucose- and maltose-utilization modules varies among isolates. A similar variation, among natural isolates of yeast, exists in the degree of crosstalk between glucose- and galactose-utilization patterns [12]. Thus, fine-tuning of regulatory crosstalk between modules, leading to exhibition of anticipatory regulation, is done in accordance with the precise ecological niche of an organism. In a recent study, it was demonstrated that yeast anticipates exhaustion of a primary carbon source, and thus, switches to the secondary source in the environment, even before the primary source is exhausted [13].

Anticipatory gene regulation, as a strategy, is widespread in pathogens [14]. In *Mycobacterium*, the different two-component systems in the bacterium are wired so that activation of one leads to partial activation of the downstream TCSs [15]. This helps the bacterium respond to an upcoming environmental change faster. In *Salmonella*, flagella is assembled prior to assembly of a *Salmonella* Pathogenicity Island-1 (SPI1)-encoded Type 3

Secretion System (T3SS) in the course of infection [16]. Regulatory elements in the flagellar cascade are known to activate the SPI1-encoded T3SS genes [17]. *Burkholderia*, in response to high cell density, anticipates stationary phase stress and triggers anticipatory response [18]. Fungal pathogens have been shown to elicit anticipatory response for protection against attack from the immune system of the host [19].

Thus, repeated, temporal exposure of environmental cues may have led to evolution of regulatory crosstalk between cellular modules. Despite significant evidence that anticipatory gene expression provides an adaptive advantage to an organism, little is known about how anticipatory regulation can evolve in a population.

In this work, we ask the following question: if a population is exposed to two environmental signals, S1 and S2, sequentially and repeatedly, under what conditions can anticipatory regulation evolve as an adaptive strategy? To answer this question, we employ two approaches.

In the first, we use a simple mathematical model to represent exposure of a population to two temporal stimuli and identify the network and physiological parameters, which maximize fitness in the given conditions. Based on the inputs from our modelling results, in the second, we evolve *E. coli* under alternating exposure to a pentose sugar rhamnose (S1) and an oxidative stress molecule, paraquat (PQ) (S2).

Briefly, *E. coli* utilizes the transporter RhaT to internalize rhamnose [20,21]. Thereafter, the sugar is processed by metabolic enzymes RhaB, RhaA, and RhaD [22]. The regulon is controlled by two AraC-like transcription factors, RhaR and RhaS [23-25]. Both bind upstream of the coding sequence of the genes essential for rhamnose utilization. The fact that rhamnose utilization regulon has two regulators and that the coding regions of one of the transcription factors (RhaR) and the transporter (RhaT) overlap make the rhamnose utilization regulon unique in *E. coli*[23,26,27]. The rhamnose-utilization regulon is also under catabolite repression, when glucose is present in the environment [28].

Prokaryotes protect themselves against oxidative stress (H_2O_2, $^{\cdot}OH$ free radical, $^{\cdot}O^{2-}$) with the action of superoxide dismutases, catalases, and peroxiredoxin [29]. In the presence of oxidative stress, the expression of these enzymes is up-regulated by transcription factors, OxyR [30], SoxS, and SoxR [31]. In a recent study, more than 50 distinct binding sites were reported for OxyR, SoxS and SoxR, when *E. coli* was studied under paraquat stress [32]. Most of these binding sites were found to be RpoD (Sigma 70)-dependent, and combined, the

three transcription factors controlled more than 100 genes in response to oxidative stress. The OxyR, SoxRS regulon includes activation of *zwf* (which encodes glucose-6 phosphate-1-dehydrogenase) , increase in cellular NADPH pool [32,33], aromatic amino acid production, and upregulation in cell wall biosynthesis [34].

Repeated, alternating exposure to S1 and S2 for ~850 generations, leads to evolution of anticipatory regulation, where prior exposure to rhamnose provides an adaptive benefit when the population is shifted to paraquat. This benefit is observed only in lines which were exposed to alternating S1 and S2; and is only seen when the cells are pre-exposed to rhamnose. Gene expression studies demonstrate that one of the possible mechanisms of this adaptive benefit is partial activation of *soxS*, upon exposure to rhamnose. Our study thus demonstrates that, in controlled laboratory environment, anticipatory gene regulation can evolve in a short timeframe of a few hundred generations.

Methods.

Model system description.

Two regulatory modules and their possible interaction was studied as follows. A microorganism experiences an environmental signal S1. Upon sensing this signal, it modulates gene expression. As part of this response, a transcription factor R1 gets activated and triggers expression of a target protein T1. The protein T1 is assumed to serve a physiological function, which helps the bacterium survive in S1. The signal S1 is assumed to be present for time *t1*. After this time, S1 is replaced by another signal S2. In response to S2, the cell activates a transcription factor R2. The activated R2 then positively controls expression of the target protein T2, which helps the cell survive and replicate in S2. The signal S2 persists for time *t2* (**Figure 1**).

In the first regulatory design, it is assumed that there is no regulatory crosstalk between the two signal responses. That is, activated R1 is responsible for response to S1 (controlling expression of T1), and activated R2 is responsible for response to S2 (controlling expression of T2). We call this design a no-anticipatory regulation design (Figure 1A). In such a setting, when allowed to propagate in alternating S1 and S2 for a long time, the population will evolve and enhance fitness. In our mathematical framework, we solve for parameter values, which maximize fitness for the regulatory topology in Figure 1A.

On the other hand, in the design with anticipatory regulation, the promoter of T2 is controlled by active R1 as well as active R2 (Figure 1B). We solve for the values of these biochemical interaction that maximize fitness in this regulatory design.

In particular, we are interested in identification of parameter values such that the fitness conferred by design in Figure 1B (anticipatory regulation) exceeds that by design in Figure 1A (no anticipatory regulation).

The equations dictating the synthesis and degradation rate of proteins in the cell, upon sensing signal S1 or S2 can be written as below.

$$R1 + [s1] \leftrightarrow R1^*$$

and,

$$R2 + [s2] \leftrightarrow R2^*$$

is represented as:

$$k_{f1}[R1][s1] \leftrightarrow k_{r1}[R1^*]$$

and

$$k_{f2}[R2][s2] \leftrightarrow k_{r2}[R2^*]$$

(1)

The dynamics of target proteins T1 and T2 is represented as the following equations. *R1** is the transcription factor responsible for expression of T1, *Km* is the Michaelis-Menten kinetic constant for T1 production, b_o is the maximum production rate of T1, and k_d is the combined degradation and dilution rate of T1 in the cell [35]. In our analysis, we assume that expression of R1 is not contingent on the presence or absence of signal S1 (but its activation is).

$$\frac{dT1}{dt} = b_0 \frac{R1^*}{R1^* + K_{m0}} - k_{d1} T1$$

(2)

Similarly, for design in Figure 1A, expression of T2 is under the exclusive control of R2.

$$\frac{dT2}{dt} = b_2 \frac{R2^*}{R2^* + K_{m2}} - k_{d2} T2$$

(3)

However, in the case when anticipatory regulation controls gene expression (Figure 1B), the dynamics of T2 in the cell can be quantified as:

$$\frac{dT2}{dt} = b_1 \frac{R1^*}{R1^* + K_{m12}} + b_2 \frac{R2^*}{R2^* + K_{m2}} - k_{d2} T2$$

(4)

Where,

$$K_m = \frac{k_off}{k_on}$$

for each transcription factor-DNA interaction. In this manner, given the biochemical parameters of a network, we can quantify the dynamics of protein expression in a cell.

Cost-benefit framework.

From the time-course data of protein expression, we use a cost-benefit model to understand the cellular fitness, as a result of the gene expression dynamics [36]. Such a framework has been used previously to understand cellular behavior and physiology [37-39]. In this formulation, the benefit conferred to the cell by the target protein T2 is given by the following expression:

$$Benefit, b = b_{max} \frac{[T2]}{[T2] + K_{mT2}} \tag{4}$$

In addition, protein production is known to have an energetic cost associated with it [40]. This cost is a linear function of the amount of protein being synthesized. Hence, the cost component associated with synthesis of protein T2 can be represented as:

$$Cost, c = c_o\,[T2] \tag{5}$$

The overall fitness of a cell is simply the difference of the benefit and cost associated with synthesis of T2. Hence, fitness is given by:

$$Fitness, f = b - c \tag{6}$$

Given a set of network topology and the biochemical parameters associated with it, this framework allows us to compute the benefit of production of T2 and track this quantity with time. In the context of this work, we compute the fitness conferred by T2 in the time window [*t1, t2*] and the total cost associated with synthesis of protein T2 in the window [*t0*, *t2*]. Note that synthesis of T2 in the regulatory design with anticipatory regulation starts before time *t1*.

Solving for optimal fitness.

Equations (1-6) have biochemical constants appear as equation parameters. Each parameter represents specific cellular processes. For instance, *Km* represents the interaction dynamics between the transcription factor and DNA. The parameter *b* represents the maximal rate of transcription from the promoter of gene *t2.*

In this work, we assume that the parameters associated with changing the regulatory patterns of gene expression (like, *Km* and *b*) are evolvable on a time scale much faster than the time needed to change parameters associated with protein property (like, enzyme activity,

degradation rate) [41]. This has been shown in several studies [42-44]. As a result, we let the parameters, which control regulation evolve, while keeping the others constant. Moreover, the parameters associated with these biochemical functions cannot take arbitrary values. Each of these parameters is constrained by the thermodynamics associated with biological processes. Hence, we let the parameters change within a defined range (given in **Table 1**). The range chosen for each parameter represents the biologically permissible window in which they can take a value [45-49].

Given the above, we introduce "mutations" (change parameters) in the network and for each of the two regulatory designs (with and without anticipatory regulation, as defined above), find the parameter set which optimizes fitness. We are particularly interested in finding whether fitness is optimized via topology in Figure 1A or 1B. Note that in this scheme, when anticipatory regulation is not the optimal response, b_{12} = 0 should be optimal design in the regulatory topology.

All simulations were performed in Matlab 7.0.

Strain used.

E. coli K12 MG1655 (ATCC 47076) (F- lambda-) was used in this study. The *soxS* mutant was generated using the primers 5'-CCCCAACAGATGAATTAACGAACTGAACACTGAAA AGAGG GTGTAGGCTGGAGCTGCTTC-3' and 5'- GAGCAATTACCCGCGCGGGAGTTAA CGCGCGGGCAATAAACATATGAATATCCTCCTTA-3' from the parent strain, as per the method described by Datsenko and Wanner [50].

Media and reagents.

M9 glycerol medium contained 0.2% glycerol unless otherwise stated. M9 salts, trace elements and casamino acids were prepared in concentrated stocks. Stock of a 5x M9 salts solution consisted of 64 g/L $Na_2HPO_4.7H_2O$, 15 g/L KH_2PO_4, 2.5 g/L NaCl, and 5 g/L NH_4Cl dissolved in Milli-Q filtered water. Casamino acids were prepared as 10x solution and were used at a final concentration of 0.05% (w/v) in the growth media. $MgSO_4$ and $CaCl_2$ were prepared at 1M stock solutions each. All stock solutions were sterilized by autoclaving. The stocks of rhamnose and paraquat was prepared at 20% and 5mM respectively and sterilized by filtration through 0.22 μm filters.

Laboratory evolution experiment.

Adaptive evolution experiments were performed by serial dilutions every 12 hours in M9 medium with appropriate stimulus. All cultures were grown at 37°C and 250rpm unless

otherwise stated. Every 12 hours, growing cultures of bacteria were diluted 1:100 in fresh M9 medium with appropriate stimulus yielding ~6-7 generations every 12 hours. The choice of sub-cultures every 12 hours ensured that cells were transferred to fresh media before they entered stationary phase. This was done to avoid the physiological changes that the bacteria undergo, upon entering the stationary phase [51]. Moreover, RpoS is also known to confer stress resistance [52]. The evolution experiment was carried out for a total of 850 generations. Freezer stocks of intermediate time points were made in the experiment. Analysis of the stocks at 300 and 600 generations are presented in this work. Starting from the ancestor, three independent lines were evolved for each environmental condition.

Conditions for evolution of anticipatory regulation.

We used rhamnose and paraquat at a concentration of 0.2% and 40 μM respectively, as the two stimuli. The evolution experiment was carried out in three replicate lines serially diluted 1:100 in alternating conditions of rhamnose and paraquat every 12 hours, as described above. As controls, we evolved three independent lines in either 0.2% rhamnose only or 40uM paraquat only. The control experiment were both the stimuli were added together would not grow beyond four dilutions, hence was dropped out of the experiment.

Analysis of evolved lines for anticipatory regulation.

All evolved lines and the ancestor were revived from freezer stocks into 2ml LB and incubated for 12 hours with shaking at 250 rpm at 37°C. The cultures were then sub-cultured 1:100 in 2ml M9 medium, containing 0.2% glycerol as the carbon source, for 12 hours. From each tube, cells were then sub-cultured 1:100 in 2ml M9 glycerol media (a) with and (b) without 0.2% rhamnose, and allowed to grow for 12 hours at 250 rpm at 37°C. After growth for 12 hours, all lines were sub-cultured 1:100 into 2ml M9 with PQ (40uM), containing glycerol as the carbon source. A volume of 150 μL of these cultures were transferred to a 96-well clear flat-bottom microplate (Costar) in triplicates. The cultures were grown at 37°C in an automatic microplate reader (Tecan Infinite M200 Pro), until they reached stationary phase. OD600 readings were taken every 30 minutes with 10 minutes of orbital shaking at 5mm amplitude before the readings. A gas permeable *Breathe-Easy* (Sigma-Aldrich) sealing membrane was used to seal the 96-well plates.

Gene expression measurements.

The 850 generation lines from each lineage and wild type were transformed with *soxS*-GFP, a plasmid-based promoter fusion from Thermo Scientific *E. coli* promoter collection (PEC3877) [53]. For estimating the promoter activity under different conditions all evolved lines with *soxS*-GFP plasmid and the wild-type ancestor were revived from freezer stocks from the

-80°C deep freezer into 2ml LB and incubated for 12 hours at 37ºC. All the media for these transformed lines contained 40µg/ml Kanamycin. The saturated cultures were then sub-cultured 1:100 in 2ml M9 glycerol medium for 12 hours. All lines were then sub-cultured 1:100 into fresh M9 glycerol medium without/with 0.2% rhamnose and allowed to grow for 12 hours at 37 deg C. After this period of growth, 150 µL of these cultures was transferred to a 96-well Black clear flat-bottom microplate (Costar) and Fluorescence (488/525 nm) and OD600 readings taken using a microplate reader (Tecan Infinite M200 PRO).

For *soxS*-GFP expression in paraquat, cells were transferred at a 1:100 dilution from M9 glycerol to M9 glycerol with /without 0.2% rhamnose. The culture were then allowed to grow for 12 hours at 37ºC. The resulting cultures were diluted to an initial OD600 of 0.1 in 2ml M9 glycero media with 40µM PQ. A volume of 150 µL of these cultures were transferred to a 96-well Black clear flat-bottom microplate (Costar) in triplicates. The cultures were grown at 37°C in an automatic microplate reader (Tecan Infinite M200 Pro), until they reached stationary phase. Fluorescence (488/525 nm) and OD600 readings were taken every 30 minutes with 10 minutes of orbital shaking at 5mm amplitude before the readings. A gas permeable *Breathe-Easy* (Sigma-Aldrich) sealing membrane was used to seal the 96-well plates.

Sequencing.

Promoters of four target genes of OxyR and SoxSR were sequenced. The primers used for sequencing were as follows: *zwf* (CCAGCATATTCATGATGTAA and GTCATTCTCCTTAAG TTAAC); *sodA* (TAATCGCGTTACTCATCTTC and ATTCATCTCCAGTATTGTCG); *fpr* (CTGATTGATTTGATCGATTG and GTTTTTCTCCTGTTTTGATT); and *fumC* (TGAAAT AAACAGAGCCGCCC and GACCTGCTCCTCACCTGATT). For sequencing, individual colonies of respective strains were collected from LB plates, and allowed to grow for 4-6 hours in liquid LB media at 37 deg C. The cells were harvested, and their genomic DNA isolated. The four promoter sequences were amplified by PCR. Sequencing was done by Eurofins Scientific.

Results

A framework to identify when anticipatory regulation can evolve in transcription networks.

We simulate a scenario where a bacterium is responding to two environmental cues encountered in temporal succession. The response to the first cue, S1, is expression of target protein T1 and the response to the second cue, S2, is expression of target protein T2 (Figure 1A). As a starting point, we assume that expression of T1 and T2 (in presence of the respective cues) is triggered by transcription factors R1 and R2, respectively. Under what conditions, however, is this regulatory design rewired such that transcription factor R1, in addition to activating expression of T1, also controls expression of T2 (Figure 1B)? In other words, under what conditions associated with the environments or properties of the proteins involved does the design represented by Figure 1B confer a greater fitness as compared to one in Figure 1A?

Anticipatory regulation is an adaptive response for intermediate values of degradation rates of the protein T2.

We simulate the networks in Figure 1A and 1B for a set of parameter values, chosen from a window typically representing each biochemical interaction in a cell [45-49,54,55]. We focus on control of transcription as the most flexible parameters of genetic rewiring of networks, as promoter mutations are thought to be accomplished much quicker than protein property changes (like, enzyme activity) [41]. For a given value of *kd* (degradation + dilution rate constant for protein T2), we find the regulatory design in Figure 1, which leads to a maximal fitness.

For a given set of fixed parameters, maximal fitness is obtained via three distinct regulatory logics, depending on the degradation rate of the protein T2 (**Figure 2A**). For small values of the *kd* of T2, the maximal fitness is achieved when control of T2 is entirely under the regulator R1. That is, R2 does not play any role in controlling expression of T2. This can be understood as follows. Expression of T2 by R1 leads to T2 production even before signal S2 is present in the environment. As a result, the T2 so produced does not confer any advantage to the cell. Thus, this preemptive production of T2 does not give any benefit to the cell, but does come at an additional cost associated with protein production in the cell. This cost, however, is offset by the enhanced T2 levels inside the cell (and consequently the benefit conferred to the cell), from the time when signal B is introduced to the system. When T2 is stable, the steady state T2 in the cell is high, and as a result, the fitness of the cell at the moment signal B is introduced to the system is also high.

At the other extreme, when the *kd* of T2 is high, the regulatory design that optimizes cellular fitness is one where control of expression of T2 is not by R1, but is instead entirely by R2. In such a context, the energetic cost of T2 production in the “futile” time-period [0, *t1*] prior to introduction of signal B, because of the high protein turnover, is high. In such a scenario, the cellular adaptive strategy is control of each target protein by its cognate regulatory partner.

However, at intermediate values of the protein T2 degradation rate constant, for maximal cellular fitness, both b1 and b2 are non-zero. This window of the values of the degradation rate constant of T2 represents the scenario where conditioning is the adaptive response in cellular functioning. Such a set up identifies conditions for a distributed regulatory design for control of gene expression. Interestingly, distributed control strategies are known to be more robust in control systems, and have also been shown to be present in biological regulatory designs [56].

Among the three regimes, anticipatory regulation is only observed when b1 and b2 are both non-zero. The other two regimes are solutions to the optimization problem, which lead to different regulatory designs. Thus, one of the key variables which leads to anticipatory regulation as the optimal solution is the degradation rate of the target protein.

In the above result, we only compare the optimal solutions of the two regulatory designs. That is, we locate the parameter values of the two regulatory designs (with and without anticipatory regulation), which optimize fitness. In such a context, we note that anticipatory regulatory design can outperform the other under certain parameter regimes. Accessibility of these optimal solutions via an adaptive path on a landscape, however, is not known. Walks on adaptive landscapes is a multidimensional problem and little is known about structures of an actual fitness landscape [57]. In the context of the problem we discuss, how a mutation changes the parameter values is an open problem [58]. Therefore, from a starting parameter set, to what adaptive solution does a population evolve to is an open problem, and has been observed to depend on the details of the problem. Little is known about this facet of evolutionary dynamics [59].

Another scheme to incorporate changes in regulatory design would be to not change the maximum strength of the promoter, but change the binding affinity of the transcription factor with the promoter region of the DNA. In the context of the model, we change the parameter k_on, which represents the affinity of the DNA for the transcription factor.

Similar to the results obtained above, we note that anticipatory regulation could also outperform for intermediate values of the degradation rate of protein T2, when we permit k_on1 or k_on2 to vary (**Figure 2B**). Also similar with the previous results, while conditioning only outperforms the non-conditioning designs for intermediate values of *kd* of protein T2, positive fitness in regulatory designs with conditioning exist for all values of kd.

We demonstrate that anticipatory regulation offers the optimal fitness in certain scenarios, for the simplest gene regulatory circuit. It is likely that as we move towards regulatory topologies of greater complexities, the ability to incorporate newer features in the dynamical responses of networks also increases. These responses could include cell-cell heterogeneity, memory, speed of response among others. All these properties have shown to have fitness consequences for a cell, in different contexts [39,55,60,61]. In addition, the number of ways in which these responses could be incorporated into the cellular manifestations also increases as the size of the network increases. This facet of networks makes tracking evolutionary dynamics of a system quite daunting – not only in terms of <u>how</u> a design could incorporate a specific feature into cellular response, but also in terms of what additional features (and also future possibilities) are going to get incorporated in the cellular physiology as a fallout of the adaptive response.

Anticipatory regulation, though not optimal, confers a positive fitness to the cell.

Figure 2 shows that anticipatory regulation is not the optimal response for very low and very high values of the degradation and dilution rates for protein T2. This is further illustrated in **Figure 3**. Here, ΔF corresponds to the difference between the best fitness when anticipatory regulation is permitted (for a corresponding (b1, b2) or (kon1, kon2) pair) and the best value of fitness when it is not permitted. Thus, for a high degradation and dilution rate (which lies in the region where regulatory design in Figure 1A maximizes fitness), $\Delta F < 0$ for all combinations of (b12, b2) (**Figure 3A**) and (k_on12, k_on2) (**Figure 3B**).

However, for a large parts of the parameter scanned in these heat plots, the fitness when anticipatory regulation is permitted, is greater than zero (Figures 4C and 4D). Hence, even though, it is not the best possible response in these parameter regions, it may still be observed as the prevailing regulatory design. This is particularly so when strength of selection on the target protein is not very high, or in population regimes where drift has a significant role to play.

Similar results are also observed when parallel heat plots are analyzed for protein-DNA interaction of T2. We note that while anticipatory regulation might not be the best possible

response, it corresponds to greater than zero fitness values of the system, for a large part of the parameter regime (**Figure 3E-H**).

Can anticipatory regulation be evolved in a laboratory set-up?

Given the above, we speculate the requirements of an experimental design to demonstrate evolution of anticipatory regulation in laboratory conditions. Our model lays down conditions, which might facilitate design of experiments towards this goal.

We predict that anticipatory regulation is most likely to evolve if the target protein has an intermediate degradation and dilution rate. This is perhaps the single most important criteria for anticipatory regulation to emerge as an outcome of an evolutionary experiment. In addition to this, the target protein responding to condition/environment 2 should serve a physiological function, which enhances cellular fitness considerably. Alternatively, the absence of target protein T2 should cause a significant fitness cost to the cell (**Figure 4A**). In this context, for experimental evolution of anticipatory regulation, it makes sense if the design of experiment is such that the signal S2 is a stress, which, in the absence of the appropriate cellular response, has a significant fitness penalty to the cell.

The other variable which dictates the likelihood of evolution of anticipatory regulation is the relative timing of the two environmental signals. In this context, we note that the greater the relative time for which the signal S2 prevails in the environment, the greater the likelihood of conditioning evolving as the optimal fitness solution (**Figure 4B**). Thus, in the context of design of experimental evolution, the relative timing of the two signals is another key variable, which likely dictates the evolutionary fate of the population. From the context of bacteria in their ecological niches, cyclical nature of the environment is not too stringent a requirement to be fulfilled for anticipatory regulation to emerge. Organisms in ecological niches are exposed to cyclical daily environments, and hence, selection oscillates with time.

However, these considerations are unlikely to be sufficient for evolution of anticipatory regulation. Moreover, while we only analyze incorporation of anticipatory regulation via transcription, several other possibilities remain. Gene regulation can be incorporated at via post-transcription, translation, and/or post-translation mechanisms. These regulatory possibilities only extend the number of ways that anticipatory regulation can be built into a regulatory design of cellular networks, should it confer adaptive benefit to the cell.

Timescale of evolution of anticipatory regulation also remains unknown. It depends on the selection pressure we put, the choice of S1 and S2, and the background genotype on which the experiment is performed.

Evolution of anticipatory regulation in *E. coli*.

From our modeling exercise, we note that two key parameters for evolution of conditioning are (a) an intermediate degradation and dilution rate of the target T2, and (b) a large cost associated with not initiating a rapid response to signal S2. In view of these inputs, we chose a pair of environmental signals, which satisfy these conditions. We chose the pentose sugar rhamnose as signal S1 and an oxidative stress molecule, paraquat (PQ) as signal S2. Rhamnose utilization is controlled by two transcription factors, RhaS and RhaR[62,63]. In the presence of rhamnose, RhaR activates the expression from the *rhaSR* promoter[64]. RhaS, on the other hand, activates expression of the rhamnose transporter protein, RhaT, and the catabolic operon, *rhaBAD*[27]*.* Both, RhaR and RhaS, belong to the AraC/XylS family of transcriptional regulators, and share sequence similarity in the DNA binding regions of the proteins[65-67].

On the other hand, oxidative stress response is largely controlled by OxyR and SoxRS system in *E. coli*[68]*.* To the best of our knowledge, the two systems are not linked by any direct transcriptional regulation. A stress of 40 uM PQ induces a growth defect in *E. coli* (**Figure 5A**). Thus, any anticipated expression of the SoxRS regulon is likely to provide a large adaptive benefit to the individual. With this premise, we performed an evolutionary experiment where the cells were exposed to S1 and S2 alternatively for a total of about 850 generations (See experimental design in **Figure 5B**, and methods section). Alongside, we also performed experiments where populations were exposed to only S1 or only S2.

In such a context, conditioning can be said to have evolved when, after alternating exposure to rhamnose and PQ for ~850 generations, pre-exposure to rhamnose confers a growth advantage when the cells are exposed to PQ. Moreover, this advantage should only be present in the lines, which were evolved in an environment with alternating presence of rhamnose and PQ. This advantage should be absent in the lines which were exposed to only rhamnose, and the lines which were exposed to only PQ.

In **Figure 5C**, growth kinetics of one of the lines from the three experimental conditions (Lines 1 from alternating rhamnose-PQ; only rhamnose; and only PQ) are shown. When cells evolved in alternating rhamnose-PQ were pre-grown in rhamnose, they exhibit a growth advantage. This advantage is absent in the lines evolved in rhamnose-only, or PQ-only.

In **Figure 5D**, the growth rate for all three lines in each of the three environments is shown. Among all the lines, only rhamnose-PQ alternating evolved lines show a growth advantage, when pre-exposed to rhamnose. The relative amounts of the advantage vary between the three lines. As expected, the rhamnose-evolved lines show a considerably lower growth rate as compared to the PQ-evolved and the rhamnose-PQ alternating evolved lines.

At 850 generations, the rhamnose-PQ alternating evolved lines exhibit a growth advantage due to anticipatory regulation. To test the kinetics of the growth advantage phenotype in the course of our evolution experiment, two intermediate time-points (generation 300 and 600) in the evolutionary experiment were also checked for their growth rate. This was done as described in Figure 5C. As shown in **Figure 6**, the anticipatory regulation was not observed at the 300 generation time point. By 600 generations, anticipatory regulation was observed in two of the three rhamnose-PQ alternating lines. This trend was exaggerated by generations 850 (**Figure 6**). The growth advantage, associated with a prior exposure to rhamnose, is at both the intermediate points, is not observed in rhamnose-only and PQ-only evolved lines. Interestingly, in two of the three rhamnose-evolved lines, fitness decreases in the PQ environment. Trade-off between performance in two different environments has been studied extensively in recent years [69-71].

Molecular targets leading to evolution of anticipatory regulation.

OxyR and SoxRS are the two major transcription regulatory systems, responsible for cellular response to oxidative stress. The combined regulon of these two systems is more than a 100 genes in *E. coli.* For identification of regions of DNA where a mutation could have facilitated evolution of anticipatory regulation, we sequenced the promoter regions of four target genes which respond to oxidative stress. These genes include *zwf* (enzyme in the pentose phosphate pathway and a response to oxidative stress) [72-75], *sodA* (one of three superoxide dismutases in *E. coli*) [76-78], *fumC* (one of the three fumarase isozymes participating in TCA and is expressed in response to oxidative stress) [79-81], and *fpr* (Flavodoxin/ferredoxin-NADP+ reductase) [82,83] promoters. The promoter regions of these four genes were sequenced in the three lines evolved in rhamnose only, PQ only, and rhamnose-PQ alternating lines.

In these sequenced parts of the genome, two mutations were identified when compared with the ancestor's genotype. Both these mutations were found in the lines evolving in rhamnose-PQ alternating conditions. In one of the lines (Alt1), in the region upstream of the coding sequence of *zwf* an adenine residue 57 bases downstream of the transcription start site of the gene was deleted. This nucleotide likely interferes with the mRNA structure and stability. The

mutation could also alter the ribosome binding site of the *zwf* mRNA, and hence change the translational efficiency of the gene. In another line evolved in the rhamnose-PQ alternating environment (Alt2), an insertion (adenine) was identified 46 bases downstream of the transcriptional start site of the *zwf* gene. As a result of this mutation, the inhibitory function of small RNAs, fnrS and rhyB, is likely compromised [84,85]. The rhamnose-only and PQ-only evolved lines did not have any mutations in the promoters sequenced in this work. Evolution of anticipatory regulation likely involves a large set of mutations. Compared to the ancestral growth rate in PQ (Figure 5D), the growth rate of the evolved lines by generation 600 is considerably higher. However, this increase in the growth rate is achieved without anticipatory regulation. Thus, our data suggests that a set of mutations are likely to have been acquired in the first phase of the evolution experiment. These mutations confer benefit by altering gene expression in a manner independent of anticipatory regulation. Only when these beneficial mutations are fixed in the population, does anticipatory regulation evolve in the culture.

Evolved lines lead to faster induction of SoxS.

SoxS is one of the primary proteins responsible for cellular response towards oxidative stress [86]. To test the possibility if induction of SoxS differs in the evolved and ancestral lines, when the cells are exposed to rhamnose, we performed the following experiment. *E. coli* is grown in M9 glycerol media in the absence and presence of 0.2% rhamnose. As shown in **Figure 7**, in all three lines evolved with alternating exposure to rhamnose and paraquat, *soxS* promoter is activated in presence of 0.2% rhamnose. This rhamnose-dependent activation of the *soxS* promoter is absent in the lines evolved in rhamnose only, or the lines evolved in paraquat only. In addition, rhamnose-dependent activation of the *soxS* promoter is absent in the ancestor strain.

Next, lines Alt1, Alt2, and Alt3 were grown in M9 glycerol media with and without rhamnose. Upon saturation, the cells were transferred to M9 glycerol media containing paraquat. As shown in **Figure 8**, the cells transferred from environments with rhamnose exhibited a faster induction of the *soxS* promoter. This phenomenon was absent in the lines evolved in cyanate, in lines evolved in rhamnose, and in the ancestor (Figure 8).

Put together, this data demonstrates that in the lines evolved in alternating rhamnose and paraquat, anticipatory regulation evolves. As we show in the manuscript, this anticipatory regulation leads to a growth phenotype. Our results here demonstrate that one of the manifestations of evolution of anticipatory regulation is partial induction of the *soxS* promoter,

when cells are grown in rhamnose. This partial induction leads to a faster transition to the ON state, when cells transition to an environment containing paraquat.

Discussion.

In this work, we present experimental evidence of evolution of anticipatory regulation in bacteria. We demonstrate that by evolving *E. coli* in alternating environments of rhamnose and PQ, anticipatory regulation evolves, where prior exposure to rhamnose enhances fitness of the population in PQ. In our experimental set-up this adaptation evolves in the time scale of a few hundred generations. Anticipatory regulation does not evolve in the first 600 generations, and only evolves thereafter. We show that the first mutations that fix in the evolved environment are the ones which directly enhance fitness. Only when these adaptive mutations are fixed in the population, do mutations that confer adaptive benefit because of anticipatory regulation fix in the population. While the molecular targets which confer advantage because of anticipatory regulation are unknown, the observed fitness advantage is likely because of reprogramming of transcription networks, which evolve faster than other biological networks[87]. This can happen via various events and processes such as gene duplication, horizontal gene transfer, and via point mutations in the transcription factor and/or target promoter[88-90].

The ecological forces leading to this anticipatory evolution of crosstalk are less clear. Molecular mechanisms for evolution of this phenomenon are not known at this point. How does this crosstalk evolve in a more complex, ecological niche? The relative benefits and costs of anticipatory regulation; and the regularity with which the second stimulus follows the first one are critical parameters, which likely dictate the evolution of anticipatory crosstalk. Our work provides a model to study these questions in a greater detail.

One of the cues in our study is oxidative stress. This was chosen since modelling suggested that anticipatory regulation could evolve if one of the cues caused a growth defect in the bacterium. But this also poses a challenge regarding the specificity of the response. Bacteria and yeast have previously been reported to elicit a general stress response, that is independent of the specific details of the stressor applied [91-94]. Thus, studying evolution of anticipatory regulation between two stresses poses a challenge. In such an experiment, where S1 and S2 are both stresses, disentangling the effects of cross-regulation and anticipatory behaviour becomes a challenge. When this was attempted in yeast, cross-regulation or anticipatory regulation was observed in a period of 300 generations. However, the authors in such a study were unable to clearly resolve these two effects [95]. Thus, we argue that one of the two signals should not be associated with a stress. While our modelling results indicate that one of the signals have a strong growth phenotype.

Sequencing parts of genome reveal little details of the molecular mechanisms of this adaptation. This is not surprising. The molecular details of the physiological changes that we characterize likely involve a large number of genes, and are complex in nature. In fact, previous work with yeast indicated has indicated that resistance to a particular stress could be acquired via a number of distinct molecular pathways [96]. This multiplicity of solutions to an environmental challenge has been predicted by theory too [97]. In the case of stress, resistance has also been shown to be dependent on the environment that cells were exposed to, prior to introduction of the stress [96]. Interestingly, stress response in yeast has also been shown to exhibit anticipatory regulation, consistent with how the organism encounters these stresses in its niche [98]. However, how gene regulation ensuring this evolves remains unknown. Our study is an attempt to answer these questions about evolution of anticipatory regulation.

Acknowledgements.

AM is supported by the Council of Scientific and Industrial Research (CSIR), Government of India, as a Senior Research Fellow (09/087(0873)/2017-EMR-I).

Conflict of interest.

The authors declare that they have no conflict of interest.

Figures

Figure 1

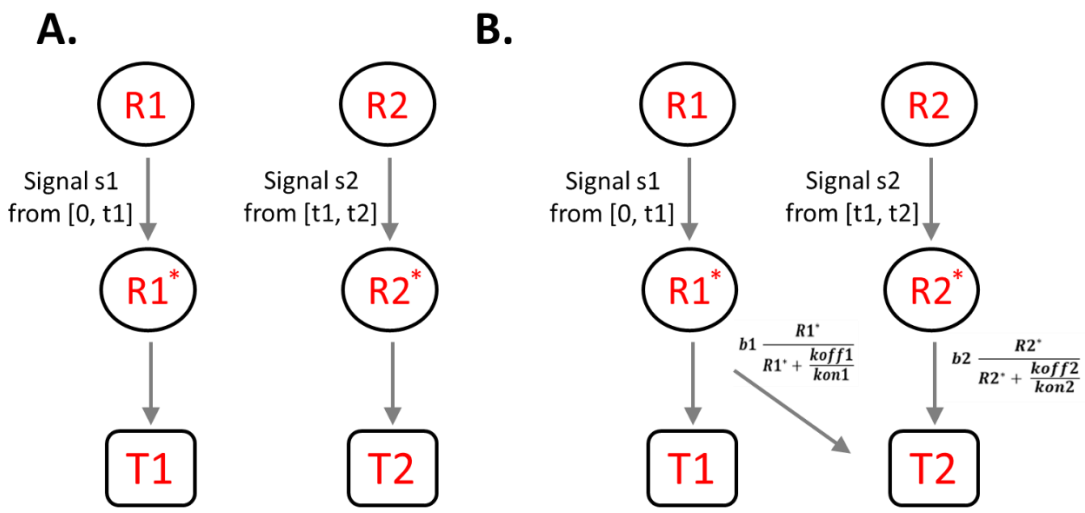


Figure 1. Regulatory designs without (A) and with (B) anticipatory regulation. In the prevailing environment, signal s1 is present from [0, t1] and signal s2 is present from [t1, t2]. In the first design, activated regulator R1* controls expression of target protein T1, and R2* controls expression of T2. In the second regulatory design, R1*, in addition to controlling expression of T1, also partially/fully controls expression of target protein T2.

Figure 2.

(A) **(B)**

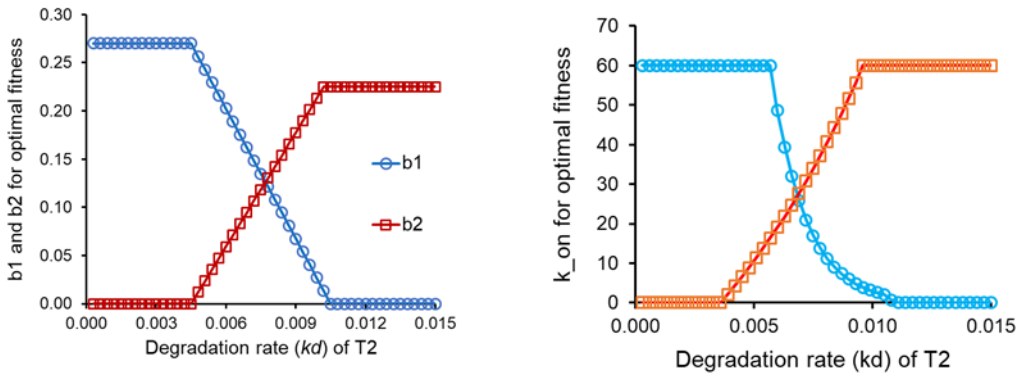


Figure 2. (A) For intermediate values of the degradation rate *kD* for target protein, T2, conditioning is the optimal response. For low values of the degradation constant, expression of T2 controlled entirely by T1 (i.e. b2 = 0) is the optimal response. For high degradation rates, optimal fitness is obtained when expression of T2 is controlled by R2 (i.e. b1 = 0). [b1 is the maximal promoter strength when R1* triggers expression from promoter of T2; b2 is the maximal promoter strength when R2* triggers expression from promoter of T2]. **(B) For intermediate values of the degradation rate *kD* for target protein, T2, conditioning is the optimal response.** For low values of the degradation constant, expression of T2 controlled entirely by T1 is the optimal response. For high degradation rates, optimal fitness is obtained when expression of T2 is controlled by R2 (i.e. k_on for b1 = 0). Blue circles represent the k_on for interaction between R1* and promoter of T2. Orange squares represent k_on for interaction between R2* and promoter of T2.

Figure 3.

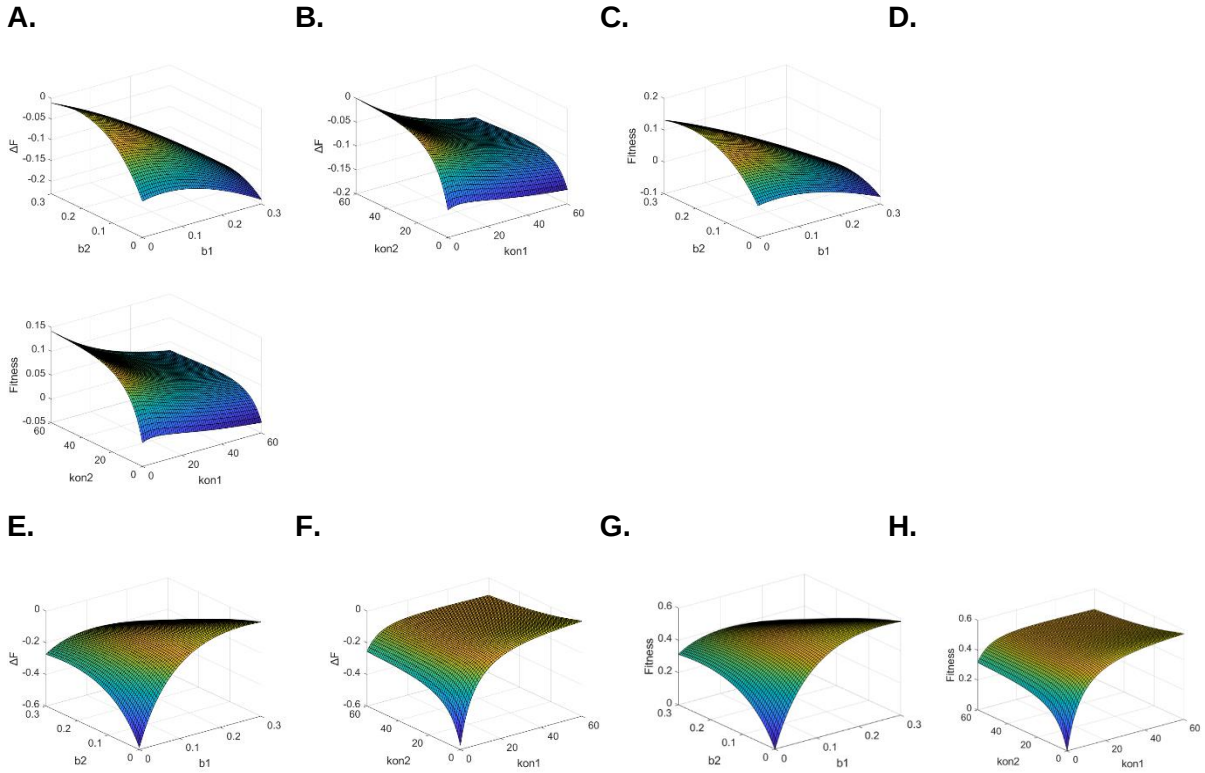
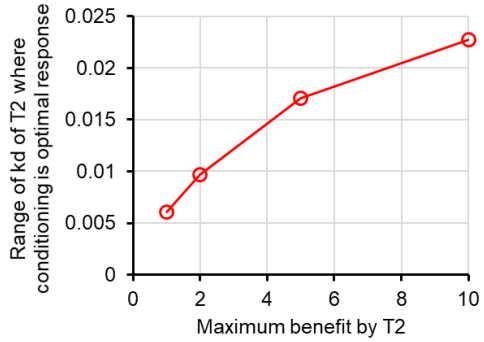


Figure 3. Anticipatory regulation yields positive fitness at high protein degradation and dilution rates (A-D). When the promoter strength **(A)** or the DNA-protein interaction **(B)** is allowed to change, the fitness of anticipatory regulatory logic is always less than of a non-anticipatory regulatory design in Figure 1A. However, for large parts of the range of parameter values, the absolute fitness conferred in the regulatory design corresponding to conditioning is greater than zero **(C-D)**. **Anticipatory regulation yields positive fitness at low protein degradation rates (E-H).** When the promoter strength **(E)** or the DNA-protein interaction **(F)** is allowed to change, the fitness of conditioning regulatory logic is always less than the design when R1* controls both T1 and T2 gene expression. However, for large parts of the range of parameter values, the absolute fitness conferred in the regulatory design corresponding to conditioning is greater than zero **(G-H)**. (A-D) are for *kd* (for T2) equal to 0.0144. (E-H) are for *kd* (for T2) equal to 0.0025.

Figure 4.

A.



B.

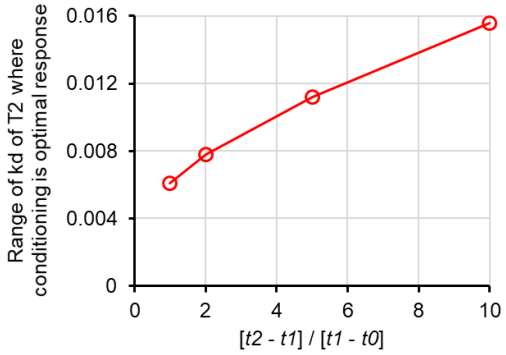


Figure 4. Anticipatory regulation emerges as optimal solution as (A) Maximum benefit conferred by T2 increases, and (B) Relative time of S2 with respect to S1 increases. The y-axis on the plots represents the range of the values of *kd* for protein T2, for which anticipatory regulation is the optimal regulatory design.

Figure 5.

A. **B.**

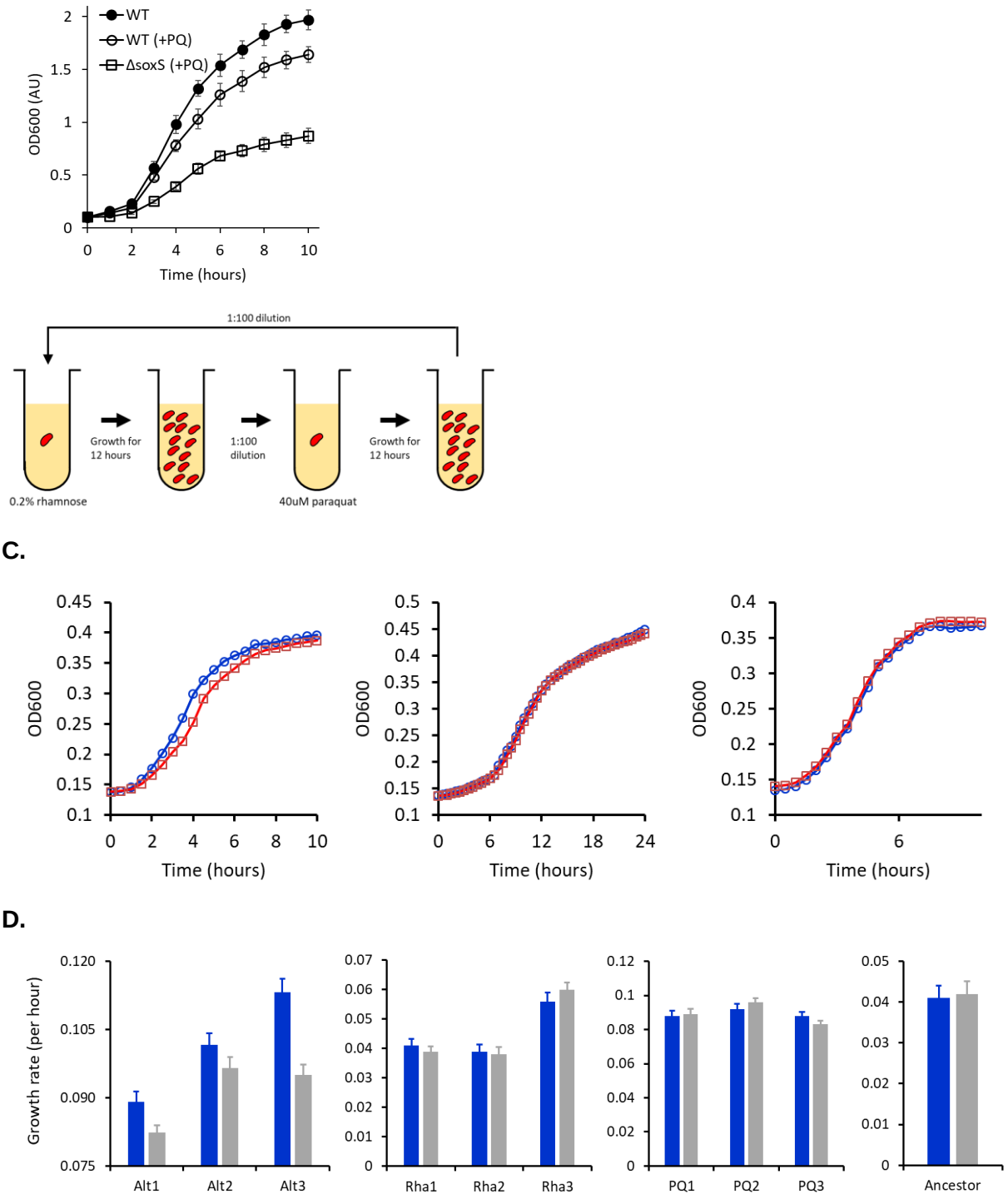


Figure 5. (A) *E. coli* exhibits a growth defect in M9 glucerol media, when grown in the presence of 40 μM PQ. **(B)** Experimental plan for the rhamnose-PQ alternating evolution. Cells were exposed to rhamnose and then PQ for 6-7 generations for a total of 850 generations. **(C)**

Anticipatory regulation evolves in rhamnose-PQ alternating lines. Cells from rhamnose-PQ evolved line (Line Alt1) (left), rhamnose-only evolved line (Line Rha1) (middle), and PQ-only evolved line (PQ1) (right) were grown in M9 minimal media with (blue) or without (red) rhamnose. The cells were transferred 1:100 to M9 glycerol media with 40 μM PQ, and their growth kinetics observed. **(D) Anticipatory regulation provides a growth advantage in all rhamnose-PQ alternating evolved lines.** Growth rates of the evolved lines in M9 glycerol media containing 40 μM PQ, when transferred from M9 glycerol media with (blue) or without (grey) 0.2% rhamnose.

Figure 6.

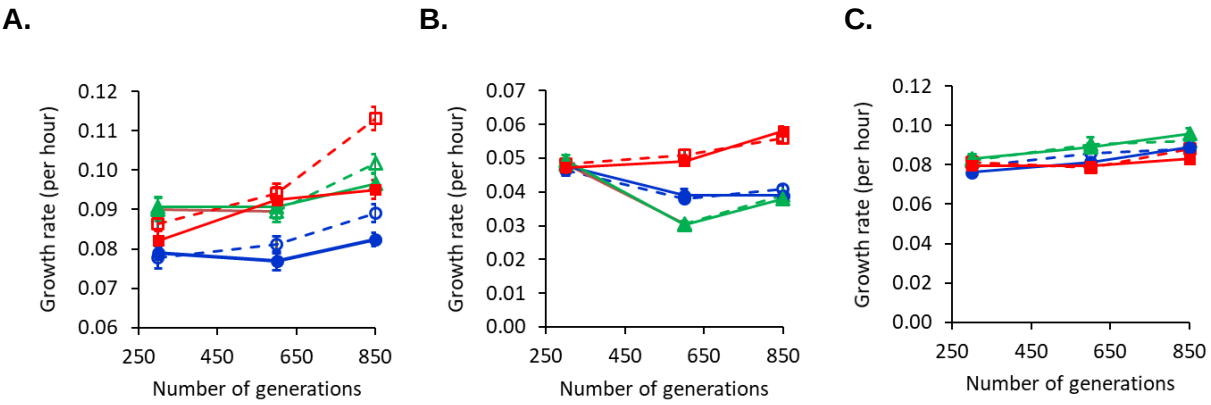


Figure 6. Growth advantage because of evolution of anticipatory regulation takes place around and after 600 generations into the experiment. Growth rate of Lines 1 (Blue), 2 (Green), and 3 (Red) in rhamnose-PQ alternating conditions (**A**), rhamnose-only (**B**), and PQ only (**C**). Solid and dashed lines represent growth curve of the lines, when transferred from M9 glycerol medium without or with rhamnose, respectively.

Figure 7.

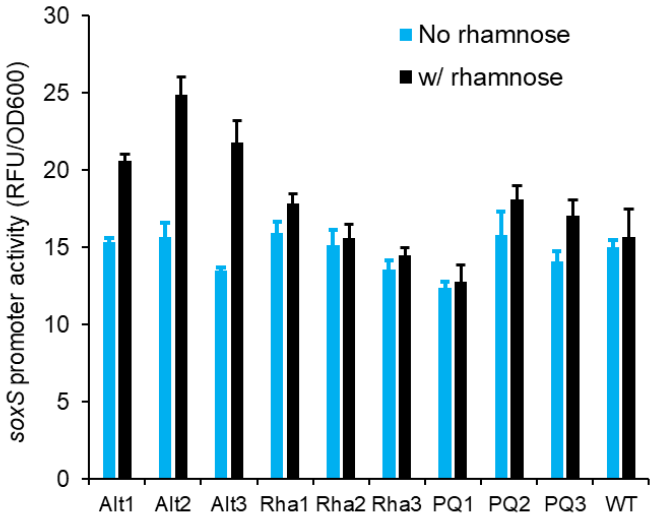


Figure 7. Lines evolved with alternating exposure to rhamnose and paraquat Alt1, Alt2, and Alt3) exhibit a statistically higher expression from the *soxS* promoter in the presence of rhamnose ($P<0.005$, t-test). This rhamnose-dependent activation of the *soxS* promoter is absent in the ancestor (WT), in rhamnose-evolved lines (Rha1, Rha2, and Rha3), and in paraquat-evolved lines (PQ1, PQ2, and PQ3).

Figure 8.

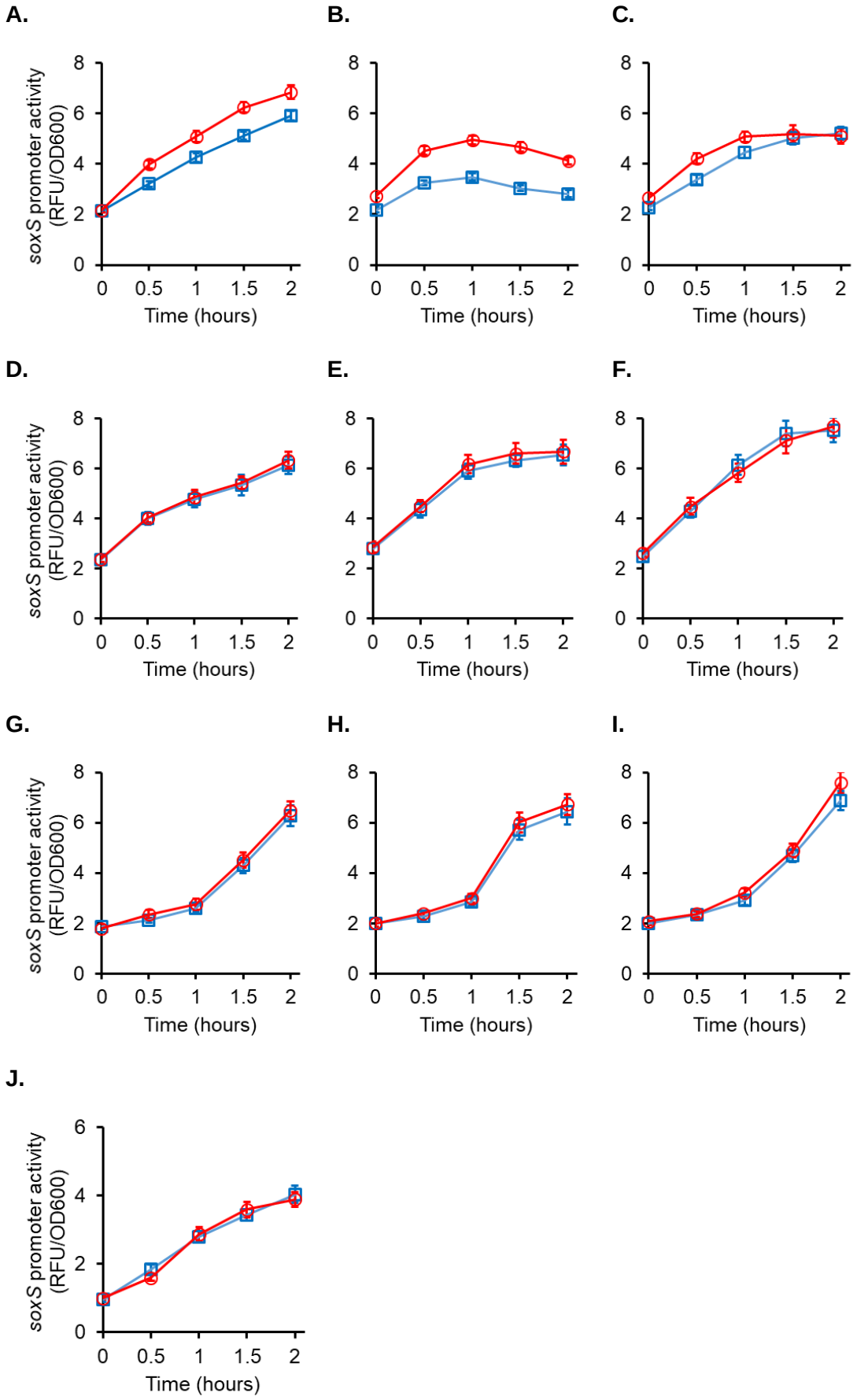

A.
B.
C.
D.
E.
F.
G.
H.
I.
J.
soxS promoter activity (RFU/OD600)
Time (hours)
0
0.5
1
1.5
2
2
4
6
8

Figure 8. In the presence of paraquat, lines Alt1 **(A)**, Alt2 **(B)**, and Alt3 **(C)** exhibited a faster induction of the *soxS* promoter when brought from M9 glycerol media containing rhamnose (red) compared to cells brought from M9 glycerol media (blue). The rhamnose-dependent faster induction of the *soxS* promoter is absent in the paraquat-evolved lines (PQ1 **(D)**, PQ2 **(E)**, and PQ3 **(F)**) and in rhamnose-evolved lines (Rha1 **(G)**, Rha2 **(H)**, and Rha3 **(I)**). All experiments were done in triplicate. The average of the three experiments and the standard deviation is reported. **(J)** Ancestor cells, when transitioned to M9 glycerol with paraquat from M9 glycerol media with (red) or without rhamnose (blue) do not exhibit any statistically significant difference in the induction kinetics from the *soxS* promoter.

Tables.

Table 1. Range of values of the parameters used in this study.

Parameter	Value/Range
b (for all promoters)	0 to 0.3
k_on (for all DNA-protein interactions)	0 to 60
k_off (for all DNA-protein interactions)	0.3
kd (degradation rate constant)	0 to 0.015
Signal (S1 or S2)	0 or 1
b_max (maximum benefit)	1.5
Km (half-maximum benefit)	15
co (cost of synthesis of one protein molecule)	0.01

References.

1 Johnson, M. E. & Hummer, G. Nonspecific binding limits the number of proteins in a cell and shapes their interaction networks. *P Natl Acad Sci USA* **108**, 603-608, doi:10.1073/pnas.1010954108 (2011).

2 Siryaporn, A. & Goulian, M. Cross-talk suppression between the CpxA-CpxR and EnvZ-OmpR two-component systems in E-coli. *Mol Microbiol* **70**, 494-506, doi:10.1111/j.1365-2958.2008.06426.x (2008).

3 Skerker, J. M. *et al.* Rewiring the specificity of two-component signal transduction systems. *Cell* **133**, 1043-1054, doi:10.1016/j.cell.2008.04.040 (2008).

4 Swain, P. S. & Siggia, E. D. The role of proofreading in signal transduction specificity. *Biophys J* **82**, 2928-2933, doi:Doi 10.1016/S0006-3495(02)75633-6 (2002).

5 Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabasi, A. L. Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551-1555, doi:DOI 10.1126/science.1073374 (2002).

6 Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **31**, 64-68, doi:10.1038/ng881 (2002).

7 Wunderlich, Z. & Mirny, L. A. Different gene regulation strategies revealed by analysis of binding motifs. *Trends Genet* **25**, 434-440, doi:10.1016/j.tig.2009.08.003 (2009).

8 Burger, A., Walczak, A. M. & Wolynes, P. G. Abduction and asylum in the lives of transcription factors. *P Natl Acad Sci USA* **107**, 4016-4021, doi:10.1073/pnas.0915138107 (2010).

9 Tagkopoulos, I., Liu, Y. C. & Tavazoie, S. Predictive behavior within microbial genetic networks. *Science* **320**, 1313-1317, doi:10.1126/science.1154456 (2008).

10 Mitchell, A. *et al.* Adaptive prediction of environmental changes by microorganisms. *Nature* **460**, 220-U280, doi:10.1038/nature08112 (2009).

11 New, A. M. *et al.* Different Levels of Catabolite Repression Optimize Growth in Stable and Variable Environments. *Plos Biol* **12**, doi:ARTN e100176410.1371/journal.pbio.1001764 (2014).

12 Wang, J. *et al.* Natural Variation in Preparation for Nutrient Depletion Reveals a Cost-Benefit Tradeoff. *Plos Biol* **13**, doi:ARTN e100204110.1371/journal.pbio.1002041 (2015).

13 Venturelli, O. S., Zuleta, I., Murray, R. M. & El-Samad, H. Population Diversification in a Yeast Metabolic Program Promotes Anticipation of Environmental Shifts. *Plos Biology* **13**, doi:ARTN e100204210.1371/journal.pbio.1002042 (2015).

14 Brunke, S. & Hube, B. Adaptive Prediction As a Strategy in Microbial Infections. *Plos Pathog* **10**, doi:ARTN e100435610.1371/journal.ppat.1004356 (2014).

15 Agrawal, R., Pandey, A., Rajankar, M. P., Dixit, N. M. & Saini, D. K. The two-component signalling networks of Mycobacterium tuberculosis display extensive cross-talk in vitro. *Biochem J* **469**, 121-134, doi:10.1042/Bj20150268 (2015).

16 Saini, S., Slauch, J. M., Aldridge, P. D. & Rao, C. V. Role of Cross Talk in Regulating the Dynamic Expression of the Flagellar Salmonella Pathogenicity Island 1 and Type 1 Fimbrial Genes. *J Bacteriol* **192**, 5767-5777, doi:10.1128/Jb.00624-10 (2010).

17 Chubiz, J. E. C., Golubeva, Y. A., Lin, D. X., Miller, L. D. & Slauch, J. M. FliZ Regulates Expression of the Salmonella Pathogenicity Island 1 Invasion Locus by Controlling HilD Protein Activity in Salmonella enterica Serovar Typhimurium. *J Bacteriol* **192**, 6261-6270, doi:10.1128/Jb.00635-10 (2010).

18 Goo, E. *et al.* Bacterial quorum sensing, cooperativity, and anticipation of stationary-phase stress. *P Natl Acad Sci USA* **109**, 19775-19780, doi:10.1073/pnas.1218092109 (2012).

19 Pradhan, A. *et al.* Anticipatory Stress Responses and Immune Evasion in Fungal Pathogens. *Trends Microbiol*, doi:10.1016/j.tim.2020.09.010 (2020).

20 Garciamartin, C., Baldoma, L., Badia, J. & Aguilar, J. Nucleotide-Sequence of the Rhar-Soda Interval Specifying Rhat in Escherichia-Coli. *J Gen Microbiol* **138**, 1109-1116, doi:Doi 10.1099/00221287-138-6-1109 (1992).

21 Baldoma, L., Badia, J., Sweet, G. & Aguilar, J. Cloning, Mapping and Gene-Product Identification of Rhat from Escherichia-Coli K12. *Fems Microbiol Lett* **72**, 103-108, doi:DOI 10.1111/j.1574-6968.1990.tb03870.x (1990).

22 Moralejo, P., Egan, S. M., Hidalgo, E. & Aguilar, J. Sequencing and Characterization of a Gene-Cluster Encoding the Enzymes for L-Rhamnose Metabolism in Escherichia-Coli. *J Bacteriol* **175**, 5585-5594, doi:Doi 10.1128/Jb.175.17.5585-5594.1993 (1993).

23 Wickstrum, J. R. & Egan, S. M. Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. *J Bacteriol* **186**, 6277-6285, doi:10.1128/JB.186.18.6277-6285.2004 (2004).

24 Holcroft, C. C. & Egan, S. M. Interdependence of activation at rhaSR by cyclic AMP receptor protein, the RNA polymerase alpha subunit C-terminal domain, and rhaR. *J Bacteriol* **182**, 6774-6782, doi:10.1128/jb.182.23.6774-6782.2000 (2000).

25 Wickstrum, J. R. *et al.* Transcription activation by the DNA-binding domain of the AraC family protein RhaS in the absence of its effector-binding domain. *J Bacteriol* **189**, 4984-4993, doi:10.1128/JB.00530-07 (2007).

26 Baldoma, L., Badia, J., Sweet, G. & Aguilar, J. Cloning, mapping and gene product identification of rhaT from Escherichia coli K12. *FEMS Microbiol Lett* **60**, 103-107, doi:10.1016/0378-1097(90)90353-r (1990).

27 Via, P., Badia, J., Baldoma, L., Obradors, N. & Aguilar, J. Transcriptional regulation of the Escherichia coli rhaT gene. *Microbiology* **142 (Pt 7)**, 1833-1840, doi:10.1099/13500872-142-7-1833 (1996).

28 Egan, S. M. & Schleif, R. F. A Regulatory Cascade in the Induction of Rhabad. *J Mol Biol* **234**, 87-98, doi:DOI 10.1006/jmbi.1993.1565 (1993).

29 Lushchak, V. I. Oxidative stress and mechanisms of protection against it in bacteria. *Biochemistry-Moscow*+ **66**, 476-489, doi:Doi 10.1023/A:1010294415625 (2001).

30 Zheng, M., Aslund, F. & Storz, G. Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**, 1718-1721, doi:DOI 10.1126/science.279.5357.1718 (1998).

31 Nunoshiba, T., Hidalgo, E., Amabile Cuevas, C. F. & Demple, B. Two-stage control of an oxidative stress regulon: the Escherichia coli SoxR protein triggers redox-inducible expression of the soxS regulatory gene. *J Bacteriol* **174**, 6054-6060, doi:10.1128/jb.174.19.6054-6060.1992 (1992).

32 Seo, S. W., Kim, D., Szubin, R. & Palsson, B. O. Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in Escherichia coli K-12 MG1655. *Cell Rep* **12**, 1289-1299, doi:10.1016/j.celrep.2015.07.043 (2015).

33 Henard, C. A., Bourret, T. J., Song, M. & Vazquez-Torres, A. Control of Redox Balance by the Stringent Response Regulatory Protein Promotes Antioxidant Defenses of Salmonella. *J Biol Chem* **285**, 36785-36793, doi:10.1074/jbc.M110.160960 (2010).

34 Pomposiello, P. J. & Demple, B. Global adjustment of microbial physiology during free radical stress. *Adv Microb Physiol* **46**, 319-341, doi:Doi 10.1016/S0065-2911(02)46007-9 (2002).

35 Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. 1st edition edn, (Chapman and Hall/CRC, 2006).

36 Dekel, E. & Alon, U. Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588-592, doi:10.1038/nature03842 (2005).

37 Fritz, G., Walker, N. & Gerland, U. Heterogeneous Timing of Gene Induction as a Regulation Strategy. *J Mol Biol* **431**, 4760-4774, doi:10.1016/j.jmb.2019.05.020 (2019).

38 Roselius, L. *et al.* Modelling and analysis of a gene-regulatory feed-forward loop with basal expression of the second regulator. *J Theor Biol* **363**, 290-299, doi:10.1016/j.jtbi.2014.08.043 (2014).

39 Zaslaver, A., Mayo, A., Ronen, M. & Alon, U. Optimal gene partition into operons correlates with gene functional order. *Phys Biol* **3**, 183-189, doi:10.1088/1478-3975/3/3/003 (2006).

40 Kafri, M., Metzl-Raz, E., Jona, G. & Barkai, N. The Cost of Protein Production. *Cell Rep* **14**, 22-31, doi:10.1016/j.celrep.2015.12.015 (2016).

41 Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M. & Teichmann, S. A. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14**, 283-291, doi:10.1016/j.sbi.2004.05.004 (2004).

42 Rubinstein, M. & de Souza, F. S. Evolution of transcriptional enhancers and animal diversity. *Philos Trans R Soc Lond B Biol Sci* **368**, 20130017, doi:10.1098/rstb.2013.0017 (2013).

43 Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet* **13**, 59-69, doi:10.1038/nrg3095 (2011).

44 Wray, G. A. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* **8**, 206-216, doi:10.1038/nrg2063 (2007).

45 Mitrophanov, A. Y. & Groisman, E. A. Positive feedback in cellular control systems. *Bioessays* **30**, 542-555, doi:10.1002/bies.20769 (2008).

46 Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nat Genet* **31**, 69-73, doi:10.1038/ng869 (2002).

47 Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. Gene regulation at the single-cell level. *Science* **307**, 1962-1965, doi:10.1126/science.1106914 (2005).

48 Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. Accurate prediction of gene feedback circuit behavior from component properties. *Mol Syst Biol* **3**, 143, doi:10.1038/msb4100185 (2007).

49 Sneppen, K., Krishna, S. & Semsey, S. Simplified models of biological networks. *Annu Rev Biophys* **39**, 43-59, doi:10.1146/annurev.biophys.093008.131241 (2010).

50 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).

51 Navarro Llorens, J. M., Tormo, A. & Martinez-Garcia, E. Stationary phase in gram-negative bacteria. *FEMS Microbiol Rev* **34**, 476-495, doi:10.1111/j.1574-6976.2010.00213.x (2010).

52 Chiang, S. M. & Schellhorn, H. E. Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Arch Biochem Biophys* **525**, 161-169, doi:10.1016/j.abb.2012.02.007 (2012).

53 Zaslaver, A. *et al.* A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nat Methods* **3**, 623-628, doi:10.1038/nmeth895 (2006).

54 Suel, G. M., Garcia-Ojalvo, J., Liberman, L. M. & Elowitz, M. B. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**, 545-550, doi:10.1038/nature04588 (2006).

55 Prajapat, M. K., Jain, K., Choudhury, D., Raj, N. & Saini, S. Revisiting demand rules for gene regulation. *Mol Biosyst* **12**, 421-430, doi:10.1039/c5mb00693g (2016).

56 Saini, S., Slauch, J. M., Aldridge, P. D. & Rao, C. V. Role of cross talk in regulating the dynamic expression of the flagellar Salmonella pathogenicity island 1 and type 1 fimbrial genes. *J Bacteriol* **192**, 5767-5777, doi:10.1128/JB.00624-10 (2010).

57 Fragata, I., Blanckaert, A., Dias Louro, M. A., Liberles, D. A. & Bank, C. Evolution in the light of fitness landscape theory. *Trends Ecol Evol* **34**, 69-82, doi:10.1016/j.tree.2018.10.009 (2019).

58 Brajesh, R. G., Dutta, D. & Saini, S. Distribution of fitness effects of mutations obtained from a simple genetic regulatory network model. *Sci Rep* **9**, 9842, doi:10.1038/s41598-019-46401-7 (2019).

59 Blount, Z. D., Lenski, R. E. & Losos, J. B. Contingency and determinism in evolution: Replaying life's tape. *Science* **362**, doi:10.1126/science.aam5979 (2018).

60 Lan, G. & Tu, Y. Information processing in bacteria: memory, computation, and statistical physics: a key issues review. *Rep Prog Phys* **79**, 052601, doi:10.1088/0034-4885/79/5/052601 (2016).

61 van Hoek, M. & Hogeweg, P. The effect of stochasticity on the lac operon: an evolutionary perspective. *PLoS Comput Biol* **3**, e111, doi:10.1371/journal.pcbi.0030111 (2007).

62 Egan, S. M. & Schleif, R. F. DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at rhaBAD. *J Mol Biol* **243**, 821-829, doi:10.1006/jmbi.1994.1684 (1994).

63 Egan, S. M. & Schleif, R. F. A regulatory cascade in the induction of rhaBAD. *J Mol Biol* **234**, 87-98, doi:10.1006/jmbi.1993.1565 (1993).

64 Tobin, J. F. & Schleif, R. F. Transcription from the rha operon psr promoter. *J Mol Biol* **211**, 1-4, doi:10.1016/0022-2836(90)90003-5 (1990).

65 Egan, S. M. Growing repertoire of AraC/XylS activators. *J Bacteriol* **184**, 5529-5532, doi:10.1128/jb.184.20.5529-5532.2002 (2002).

66 Gallegos, M. T., Michan, C. & Ramos, J. L. The XylS/AraC family of regulators. *Nucleic Acids Res* **21**, 807-810, doi:10.1093/nar/21.4.807 (1993).

67 Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K. & Ramos, J. L. Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* **61**, 393-410 (1997).

68 Jain, K. & Saini, S. MarRA, SoxSR, and Rob encode a signal dependent regulatory network in Escherichia coli. *Mol Biosyst* **12**, 1901-1912, doi:10.1039/c6mb00263c (2016).

69 Ferenci, T. Trade-off Mechanisms Shaping the Diversity of Bacteria. *Trends Microbiol* **24**, 209-223, doi:10.1016/j.tim.2015.11.009 (2016).

70 Litchman, E., Edwards, K. F. & Klausmeier, C. A. Microbial resource utilization traits and trade-offs: implications for community structure, functioning, and biogeochemical impacts at present and in the future. *Front Microbiol* **6**, 254, doi:10.3389/fmicb.2015.00254 (2015).

71 Cooper, V. S. & Lenski, R. E. The population genetics of ecological specialization in evolving Escherichia coli populations. *Nature* **407**, 736-739, doi:10.1038/35037572 (2000).

72 Greenberg, J. T., Monach, P., Chou, J. H., Josephy, P. D. & Demple, B. Positive control of a global antioxidant defense regulon activated by superoxide-generating agents in Escherichia coli. *Proc Natl Acad Sci U S A* **87**, 6181-6185, doi:10.1073/pnas.87.16.6181 (1990).

73 Griffith, K. L. & Wolf, R. E., Jr. Systematic mutagenesis of the DNA binding sites for SoxS in the Escherichia coli zwf and fpr promoters: identifying nucleotides required for DNA binding and transcription activation. *Mol Microbiol* **40**, 1141-1154, doi:10.1046/j.1365-2958.2001.02456.x (2001).

74 Jair, K. W., Fawcett, W. P., Fujita, N., Ishihama, A. & Wolf, R. E., Jr. Ambidextrous transcriptional activation by SoxS: requirement for the C-terminal domain of the RNA polymerase alpha subunit in a subset of Escherichia coli superoxide-inducible genes. *Mol Microbiol* **19**, 307-317, doi:10.1046/j.1365-2958.1996.368893.x (1996).

75 Fawcett, W. P. & Wolf, R. E., Jr. Genetic definition of the Escherichia coli zwf "soxbox," the DNA binding site for SoxS-mediated induction of glucose 6-phosphate dehydrogenase in response to superoxide. *J Bacteriol* **177**, 1742-1750, doi:10.1128/jb.177.7.1742-1750.1995 (1995).

76 Compan, I. & Touati, D. Interaction of six global transcription regulators in expression of manganese superoxide dismutase in Escherichia coli K-12. *J Bacteriol* **175**, 1687-1696, doi:10.1128/jb.175.6.1687-1696.1993 (1993).

77 Carlioz, A. & Touati, D. Isolation of superoxide dismutase mutants in Escherichia coli: is superoxide dismutase necessary for aerobic life? *EMBO J* **5**, 623-630 (1986).

78 Touati, D. Transcriptional and posttranscriptional regulation of manganese superoxide dismutase biosynthesis in Escherichia coli, studied with operon and protein fusions. *J Bacteriol* **170**, 2511-2520, doi:10.1128/jb.170.6.2511-2520.1988 (1988).

79 Woods, S. A., Schwartzbach, S. D. & Guest, J. R. Two biochemically distinct classes of fumarase in Escherichia coli. *Biochim Biophys Acta* **954**, 14-26, doi:10.1016/0167-4838(88)90050-7 (1988).

80 Liochev, S. I. & Fridovich, I. Fumarase C, the stable fumarase of Escherichia coli, is controlled by the soxRS regulon. *Proc Natl Acad Sci U S A* **89**, 5892-5896, doi:10.1073/pnas.89.13.5892 (1992).

81 Gruer, M. J. & Guest, J. R. Two genetically-distinct and differentially-regulated aconitases (AcnA and AcnB) in Escherichia coli. *Microbiology (Reading)* **140 (Pt 10)**, 2531-2541, doi:10.1099/00221287-140-10-2531 (1994).

82 Liochev, S. I., Hausladen, A., Beyer, W. F., Jr. & Fridovich, I. NADPH: ferredoxin oxidoreductase acts as a paraquat diaphorase and is a member of the soxRS regulon. *Proc Natl Acad Sci U S A* **91**, 1328-1331, doi:10.1073/pnas.91.4.1328 (1994).

83 Manchado, M., Michan, C. & Pueyo, C. Hydrogen peroxide activates the SoxRS regulon in vivo. *J Bacteriol* **182**, 6842-6844, doi:10.1128/jb.182.23.6842-6844.2000 (2000).

84 Argaman, L., Elgrably-Weiss, M., Hershko, T., Vogel, J. & Altuvia, S. RelA protein stimulates the activity of RyhB small RNA by acting on RNA-binding protein Hfq. *Proc Natl Acad Sci U S A* **109**, 4621-4626, doi:10.1073/pnas.1113113109 (2012).

85 Boysen, A., Moller-Jensen, J., Kallipolitis, B., Valentin-Hansen, P. & Overgaard, M. Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. *J Biol Chem* **285**, 10690-10702, doi:10.1074/jbc.M109.089755 (2010).

86 Seo, S. W., Kim, D., Szubin, R. & Palsson, B. O. Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in Escherichia coli K-12 MG1655. *Cell Rep* **12**, 1289-1299, doi:10.1016/j.celrep.2015.07.043 (2015).

87 Shou, C. *et al.* Measuring the evolutionary rewiring of biological networks. *PLoS Comput Biol* **7**, e1001050, doi:10.1371/journal.pcbi.1001050 (2011).

88 Teichmann, S. A. & Babu, M. M. Gene regulatory network growth by duplication. *Nat Genet* **36**, 492-496, doi:10.1038/ng1340 (2004).

89 Price, M. N., Dehal, P. S. & Arkin, A. P. Horizontal gene transfer and the evolution of transcriptional regulation in Escherichia coli. *Genome Biol* **9**, R4, doi:10.1186/gb-2008-9-1-r4 (2008).

90 Price, M. N., Dehal, P. S. & Arkin, A. P. Orthologous transcription factors in bacteria have different functions and regulate different genes. *PLoS Comput Biol* **3**, 1739-1750, doi:10.1371/journal.pcbi.0030175 (2007).

91 Gasch, A. P. *et al.* Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11**, 4241-4257, doi:10.1091/mbc.11.12.4241 (2000).

92 Causton, H. C. *et al.* Remodeling of yeast genome expression in response to environmental changes. *Mol Biol Cell* **12**, 323-337, doi:10.1091/mbc.12.2.323 (2001).

93 Carrasco, B., Fernandez, S., Petit, M. A. & Alonso, J. C. Genetic recombination in Bacillus subtilis 168: effect of DeltahelD on DNA repair and homologous recombination. *J Bacteriol* **183**, 5772-5777, doi:10.1128/JB.183.19.5772-5777.2001 (2001).

94 Weber, H., Polen, T., Heuveling, J., Wendisch, V. F. & Hengge, R. Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* **187**, 1591-1603, doi:10.1128/JB.187.5.1591-1603.2005 (2005).

95 Dhar, R., Sagesser, R., Weikert, C. & Wagner, A. Yeast adapts to a changing stressful environment by evolving cross-protection and anticipatory gene regulation. *Mol Biol Evol* **30**, 573-588, doi:10.1093/molbev/mss253 (2013).

96 Berry, D. B. *et al.* Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. *PLoS Genet* **7**, e1002353, doi:10.1371/journal.pgen.1002353 (2011).

97 Brajesh, R. G., Raj, N. & Saini, S. Optimal parameter values for the control of gene regulation. *Mol Biosyst* **13**, 796-803, doi:10.1039/c6mb00765a (2017).

98 Mitchell, A. *et al.* Adaptive prediction of environmental changes by microorganisms. *Nature* **460**, 220-224, doi:10.1038/nature08112 (2009).