



## Predictive Behavior Within Microbial Genetic Networks

Ilias Tagkopoulos *et al.*  
*Science* **320**, 1313 (2008);  
DOI: 10.1126/science.1154456

*This copy is for your personal, non-commercial use only.*

**If you wish to distribute this article to others**, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

**Permission to republish or repurpose articles or portions of articles** can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of October 14, 2013 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/320/5881/1313.full.html>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/content/suppl/2008/06/09/1154456.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/320/5881/1313.full.html#related>

This article **cites 23 articles**, 8 of which can be accessed free:

<http://www.sciencemag.org/content/320/5881/1313.full.html#ref-list-1>

This article has been **cited by** 27 article(s) on the ISI Web of Science

This article has been **cited by** 17 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/320/5881/1313.full.html#related-urls>

This article appears in the following **subject collections**:

Microbiology

<http://www.sciencemag.org/cgi/collection/microbio>

# Predictive Behavior Within Microbial Genetic Networks

Ilias Tagkopoulou,<sup>1,2\*</sup> Yir-Chung Liu,<sup>2,3\*</sup> Saeed Tavazoie<sup>2,3†</sup>

The homeostatic framework has dominated our understanding of cellular physiology. We question whether homeostasis alone adequately explains microbial responses to environmental stimuli, and explore the capacity of intracellular networks for predictive behavior in a fashion similar to metazoan nervous systems. We show that *in silico* biochemical networks, evolving randomly under precisely defined complex habitats, capture the dynamical, multidimensional structure of diverse environments by forming internal representations that allow prediction of environmental change. We provide evidence for such anticipatory behavior by revealing striking correlations of *Escherichia coli* transcriptional responses to temperature and oxygen perturbations—precisely mirroring the covariation of these parameters upon transitions between the outside world and the mammalian gastrointestinal tract. We further show that these internal correlations reflect a true associative learning paradigm, because they show rapid decoupling upon exposure to novel environments.

Although originally conceived in the context of human physiological adaptation (1), homeostasis has also become the de facto framework for understanding cellular behavior. In its most essential form, the homeostatic response is an attempt to maintain the “constancy of the internal state” in response to perturbations resulting from environmental fluctuations (e.g., expression of osmoprotectants in response to osmolarity stress). When such fluctuations are essentially random (unpredictable), the cell may directly or indirectly sense the perturbation and enact the appropriate response program. On the other hand, if such variations are perfectly predictable, such as periodic changes in photon flux due to Earth’s rotation, an internal model (circadian rhythm) can be used to anticipate relevant changes. The organism can then mount a “preemptive” response—for example, by gearing up the photosynthetic machinery before sunrise. The widespread use of internal circadian models, from unicellular cyanobacteria to humans (2), suggests that this predictive mode of behavior confers considerable fitness advantages to organisms that have evolved it.

Environmental variables that show purely random fluctuations or perfectly periodic rhythms define idealized extremes. In fact, some parameters whose fluctuations may seem random when viewed in isolation can nonetheless be highly “predictable” when considered in the temporal context of variation in other parameters (Fig. 1). As such, variations in one environmental variable can convey substantial information about variation in another. For example, a bacterium may experience strong covariation in temperature and photon flux as it traverses the upper layers of a

stratified marine ecosystem. Such temporally structured correlations can exist on multiple time scales, reflecting the highly structured (non-random) habitats of free-living organisms. Temporal delays are a typical feature of these correlations. For example, an increase in temperature may herald an impending decrease in O<sub>2</sub> levels some 20 min later. An organism that is capable of learning (internalizing) these correlations can then exploit them in order to anticipate vital changes in the environment—for example, preparing for resource fluctuations or mounting protective responses to extreme perturbations.

Within metazoans, the basic capacity for predictive behavior requires complex neural network architecture. Here we hypothesize that an analog of this capacity, implemented by networks of biochemical reactions, exists in unicellular microbes. To demonstrate this potential, we have developed a biochemically realistic computer simulation for evolving populations of organisms under precisely defined environments where multiple time-varying signals encode information about resource abundance. Randomly evolving biochemical networks of these organisms form internal representations of their dynamic environments that enable predictive behavior. We provide experimental evidence for this capacity by revealing strong correlations in genome-wide transcriptional responses of *E. coli* to transitions

in oxygen and temperature. These correlations do not reflect an essential biochemical coupling between oxygen and temperature because they are rapidly decoupled in the context of selection in a novel environment.

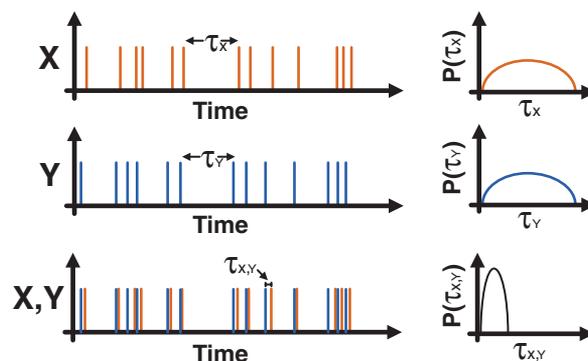
**Emergence of predictive behavior in simulated biochemical networks.** Computational simulations of biological systems are yielding unique insights into a variety of fundamental questions in biology (3–10). We developed a simulation framework, called Evolution in Variable Environment (EVE), aimed at exploring the capacity of biochemical networks to evolve predictive internal models of complex environments (11). Within EVE, biochemical networks are structured around the “central dogma” and they evolve in an asynchronous and stochastic manner, achieving the temporal dynamics of cascades of biochemical interactions/transformations (including transcription, translation, and protein modification) that are present in real cells (Fig. 2A). Each node in the biochemical network of an organism is parameterized by several continuous variables mapping to biological parameters such as basal expression, degradation, and regulatory strength. At any point during the simulation, random mutations (e.g., transcription rate change, node duplication, node deletion) may alter any of the organism’s parameters and consequently its phenotype (Fig. 2B). Environmental signals, conveying information about resources, may couple and stochastically regulate any action (e.g., transcription, translation, or transformation) in the organism’s internal network. A node can regulate any other node, although one may restrict unlikely interactions using a probabilistic model (for instance, we may choose to make RNA-mediated protein modification a rare event). Each organism possesses a generic and stable response pathway through which it can interact with the environment, whether for energy extraction or response to stress. Expression of the response pathway is modeled to have a high energy cost; thus, there is strong selection for organisms whose response correlates with the appropriate environmental event (e.g., presence of food). Organisms evolve in the context of a constant population size, where growth rate is directly proportional to the amount of energy per cell. The energy of each

<sup>1</sup>Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA. <sup>2</sup>Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA. <sup>3</sup>Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: tavazoie@genomics.princeton.edu

**Fig. 1.** Predictability emerges from temporal context of correlated random variables. When viewed in isolation, events X and Y have a random temporal structure (left) as manifested by the large uncertainty in the inter-event interval  $\tau_x$  and  $\tau_y$  (right). However, the occurrence of event X is highly predictable within the temporal context of event Y (left), with a relatively tight distribution of temporal delay between the two events  $\tau_{x,y}$  (right).



cell depends on the energy extraction rate from the environment as well as on the metabolic demands of expressing and maintaining network components (Fig. 2C).

Within this *in silico* ecology, evolving organisms with random initial networks compete against each other in structured environments where signals and resources fluctuate with a distinct temporal correlation structure. In a typical experiment, the combinatorial states of multiple signals convey information about the availability of extractable energy resources in the near future. Cells that can efficiently “learn” such correlations are able to express the energy-extracting metabolic pathway at the appropriate time, giving them sizable fitness advantage over their competitors.

We conducted large-scale simulations of nine temporally structured environments across mutation rates spanning four orders of magnitude. A variety of environments were imposed, including those that selected for one of five delayed logic gates, oscillators, bistable switches, and duration inference mechanisms. Despite the complex nature of many of these environments, high-fitness organisms emerged with variable success rate in every case (table S1 and figs. S4 and S5).

The full historical documentation of the evolutionary process in EVE allowed us to connect evolutionary dynamics, phenotypic behavior, and network structure in ways that are difficult to do in naturally evolved biological systems. In a particularly challenging environment, the abundance of resource (food) is related to the signals through a delayed dynamic exclusive OR (XOR) relationship, where resources become abundant shortly (within 1000 time units) after the presence of either signal S1 or signal S2 alone, but not when they co-occur. Fitness is defined as the Pearson correlation (PC) between the abundance of resource and response pathway expression. Interestingly, the fitness trajectory of the fittest organism displays nonmonotonic behavior (Fig. 3A). In the first  $5.4 \times 10^6$  time units (1100 epochs), there is no stable phenotype in the population whose response protein expression correlates well with the presence of resources in the environment (Fig. 3B, 1). However, several advantageous mutations in the next  $4.5 \times 10^5$  time units (100 epochs) lead to a phenotype where the expression of the response protein loosely couples to the presence of signal S1 and the absence of signal S2 (Fig. 3B, 2). Subsequent mutations give rise to a fit (PC > 0.8) but unstable phenotype (Fig. 3B, 3) where a noisy dynamic XOR is achieved. Instead of stabilization of the fittest phenotype at the time, we observe an abrupt decline of fitness and subsequent fixation to a suboptimal fitness peak (Fig. 3B, 4). The next fitness increase occurs after  $3.6 \times 10^5$  time units (800 epochs) when a gene duplication event (fig. S6) allows the newly formed gene and its protein and modified protein products to take on their own course in evolution, thus giving rise to an optimal final phenotype (Fig. 3B, 5). The

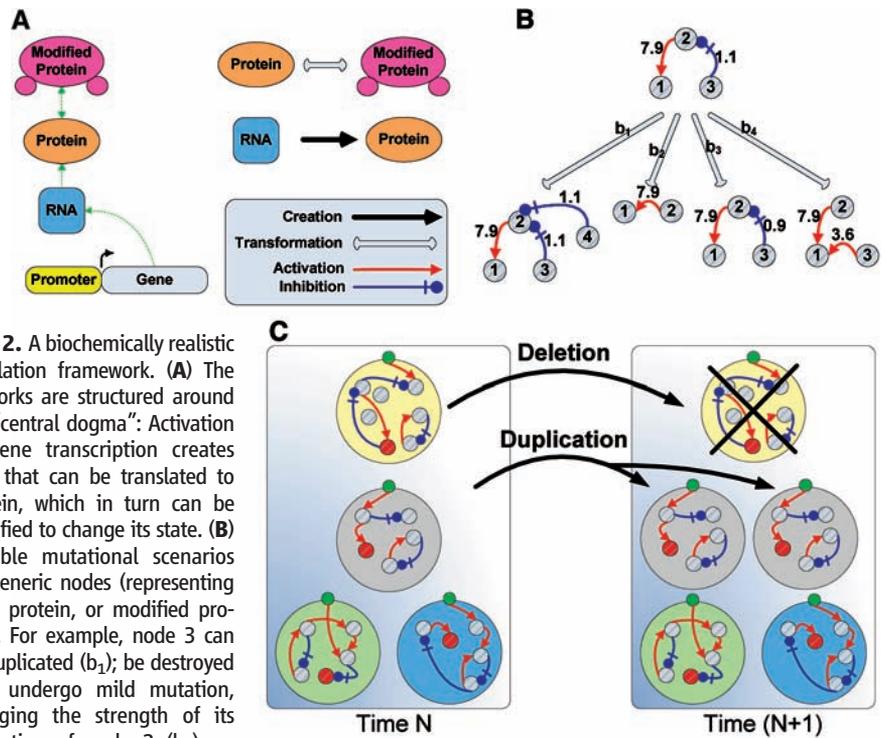
evolutionary impact of gene duplication was also observed in other experiments, most of which displayed monotonically increasing fitness trajectories (figs. S7 and S8).

The evolved biochemical networks displayed a high level of redundancy, as the majority of single node/link “knockouts” did not reveal a sizable fitness effect. In an attempt to identify a “minimal network” that was sufficient to express the evolved behavior, we devised a procedure (11) that sequentially knocks out links until the network hits a critical mass that cannot be reduced without a large (>5%) decrease in the organism’s fitness. We found that such minimal networks were considerably more informative about the relationship between network structure and function relative to the initial networks they were derived from.

The full and minimal networks belonging to the fittest organism evolved under the delayed XOR selection are displayed in Fig. 4A; the expression profile of all the nodes in the full network is shown in Fig. 4B. From a total of 31 links, only 11 were found to be essential as defined above. Signals S1 and S2 serve as direct observable environmental inputs, whereas resources can be harvested only when response protein (RP1) is expressed. The presence of S1 catalyzes the translation of RNA2 to P2, which in turn leads to the activation of G1 and subsequent translation of its RNA to P1. Once P1 is created, it can undergo modification and become response

protein RP1. However, this step is not completed immediately, as P2 also inhibits P1’s modification. Modification can occur after signal S1 goes down and the high degradation rate of P2 leads to its rapid drop. This in turn introduces a delay interval that is large enough to allow the expression of RP1 to coincide precisely with the presence of the resource. Once expressed, RP1, operating in a negative feedback loop, represses the expression of G1. The low degradation rate of RNA1 and self-activation of P1 contribute to the network’s ability to stay in standby mode where P1 molecules are constantly replenished but not modified to RP1. With this mechanism, the time delay is easily tuned through changes in the kinetic constants of one protein, namely P2. Similarly, when only S2 is present, the response pathway is activated through P0 in an analogous fashion. However, when both signals S1 and S2 are present, P0 and P2 are unable to activate G1 transcription because the two environmental signals cross-catalyze the modification of these proteins to their inactive states (MP0 and MP2). Additionally, the high degradation rate of MP0 and MP2 makes transitions from MP0 to P0 or from MP2 to P2 unlikely.

In general, careful inspection revealed how behavior emerges from the topology and kinetics of the evolved minimal networks (11). However, we found variable success in the ability of standard reverse-engineering methods (12, 13) to reveal network structure from behavior (figs. S14 to S16).



**Fig. 2.** A biochemically realistic simulation framework. (A) The networks are structured around the “central dogma”: Activation of gene transcription creates RNA that can be translated to protein, which in turn can be modified to change its state. (B) Possible mutational scenarios for generic nodes (representing RNA, protein, or modified protein). For example, node 3 can be duplicated ( $b_1$ ); be destroyed ( $b_2$ ); undergo mild mutation, changing the strength of its regulation of node 2 ( $b_3$ ); or mutate strongly, causing its decoupling from node 2 and coupling to node 1 ( $b_4$ ). In addition to network topology, mutations may change node parameters such as basal transcription rate, degradation rate, and interaction affinity. (C) Organisms that accumulate enough energy undergo duplication and increase their representation in the population. Duplication is accompanied by deletion of the organism with the least energy in the population.

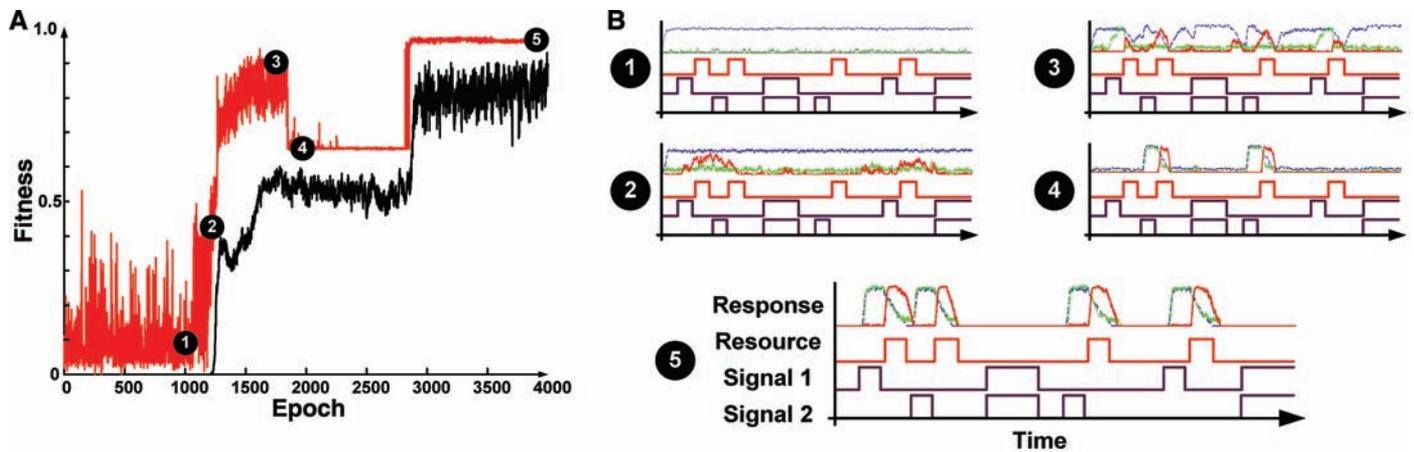
### Correlated transcriptional responses of *E. coli* to oxygen and temperature perturbations.

The evolution of predictive behavior in simulated networks suggested that this capacity might exist in naturally evolved microbial networks. Looking for evidence of such predictive modeling requires knowledge of microbial habitat structures well beyond what currently exists for most organisms. However, for some microbes, recurrent transitions in and out of a dominant niche are accompanied by deterministic correlations in important environmental variables. For example, transition from the outside environment into the oral cavity exposes the bacterium *E. coli* to an immediate increase in temperature from ambient ( $<30^{\circ}\text{C}$ ) to  $37^{\circ}\text{C}$ . This

is followed by an impending drop in oxygen levels as the bacterium transitions into the gastrointestinal tract (Fig. 5A). The complex ecology of the mammalian GI tract imposes strong competitive selection for colonization (14). In this setting, appropriate expression of adaptive functions confers strong fitness advantages. From a purely homeostatic perspective, physiological transition from aerobic to anaerobic respiration should only take place immediately after a drop in oxygen levels. However, the fitness benefits of such a reflexive response will not be optimal, in terms of shutting down superfluous capacities (such as aerobic respiration) as well as the time delay required to fully express beneficial functions (such as anaerobic respiration). On the other

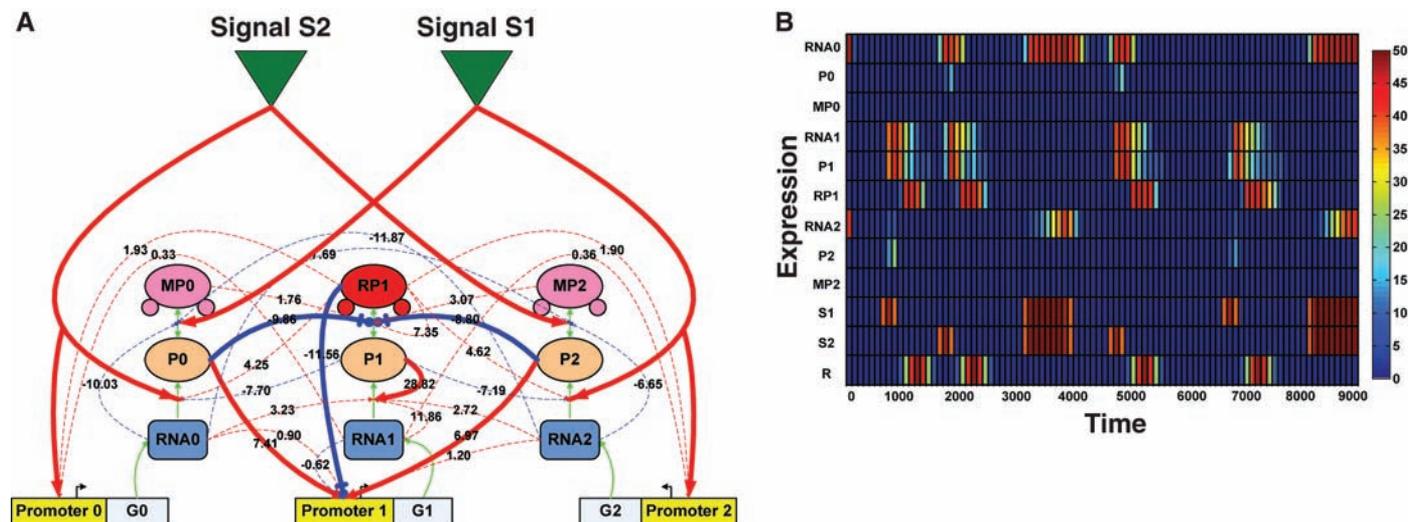
hand, if bacteria use the immediate increase in temperature as a predictive signal of impending oxygen drop, they could respond in an anticipatory fashion and be in the optimal physiological state at the time oxygen levels drop. If this were true, we would expect that an increase in temperature may lead to a similar physiological response to a decrease in oxygen—even in the presence of maximal oxygen levels.

To test this hypothesis, we used microarray transcriptional profiling (15) to observe the global cellular state correlates of such physiological responses. Oxygen and temperature were precisely controlled in the context of growth within bioreactors (11). These experiments included temperature transitions between



**Fig. 3.** Emergence of a delayed XOR phenotype. **(A)** Fitness trajectory of an experiment where the presence of one (and only one) signal indicates the future availability of resources. Red and black lines correspond to the highest and mean fitness in the population at each epoch (4500 time units), respectively. **(B)** The phenotypic behavior of the fittest organism at

different points along the evolutionary trajectory. Each subplot consists of four rows. The first row depicts the abundance profiles of the RNA (blue), protein (green), and modified protein (red) of the response pathway. The second, third, and fourth rows correspond to resource abundance and environmental signals S1 and S2, respectively.



**Fig. 4.** Network topology and expression profiles of molecular components. **(A)** The regulatory network of an organism evolved under low mutation rate within a delayed dynamic XOR environment. Each node represents mRNA/siRNA (RNA), protein (P), or modified protein (MP). The environmental resource harvested at each time point is proportional to the number of response protein molecules (red RP1 node) and resource

abundance at the time. Regulatory interactions can be positive (activating; red arrows) or negative (repressing; blue arrows). Solid lines represent essential links. **(B)** Expression profile of all nine nodes, environmental signals (S1 and S2), and resource (R) during two epochs (9000 time units). The color scale refers to the number of molecules present.

25°C and 37°C, and shifts between anaerobic (0% dissolved O<sub>2</sub>) and aerobic (16 to 21% dissolved O<sub>2</sub>) growth. Duplicate experiments revealed highly reproducible temperature and oxygen profiles (fig. S17) and gene expression patterns (fig. S18).

Consistent with the predictive behavioral framework above, we found a striking correlation between global transcriptional responses to temperature upshift and oxygen downshift, corresponding to the transition of *E. coli* into the GI tract (Fig. 5A). For instance, we found a highly significant overlap between the set of genes down-regulated by temperature upshift and those down-regulated by oxygen downshift (hypergeometric  $P < 10^{-58}$ ). As predicted, these genes are highly enriched for aerobic respiration functions such as the tricarboxylic acid cycle (TCA) and glyoxylate cycle ( $P < 10^{-8}$ ) and cytochrome *bo*<sub>3</sub> oxidase complex ( $P < 10^{-4}$ ), the dominant electron donor under aerobic conditions (16).

Likewise, temperature downshift and oxygen upshift—corresponding to the transition out of the GI tract—induced strikingly similar gene expression responses ( $P < 10^{-28}$ ) (fig. S19). We also compared the global similarity in transcriptional responses through Pearson correlation (Fig. 5B). As expected, oxygen downshift and oxygen upshift are highly anticorrelated (Pearson  $r = -0.50$ ;  $P < 10^{-148}$ ), as are reverse perturbations in temperature ( $r = -0.56$ ;  $P < 10^{-171}$ ). The strong correlation between temperature upshift and oxygen downshift was most striking ( $r = +0.39$ ;  $P < 10^{-79}$ ), consistent with the similarity in differentially regulated genes. The relatively small correlation between oxygen downshift and unrelated perturbations such as ultraviolet exposure (17) ( $r = -0.06$ ;  $P < 0.02$ ) and osmolarity stress (18) ( $r = 0.11$ ;  $P < 10^{-7}$ ) makes it unlikely that these correlated behaviors may be due to a generic response to external perturbations (Fig. 5B), as seen, for example, in the common “stress” response in yeast (19). What is remarkable is the rapid transcriptional reprogramming from aerobic to anaerobic, during temperature increase, as reflected by the strong repression of genes encoding components of the TCA cycle and cytochrome *bo*<sub>3</sub> oxidase complex (Fig. 5C). These changes are accompanied by concomitant induction of genes encoding components of cytochrome *bd* oxidase complex—the preferred electron donor in low-oxygen environments (16) (Fig. 5C). Strikingly, this anaerobic transcriptional reprogramming occurs under a highly aerobic environment (18% dissolved O<sub>2</sub>), representing a seemingly maladaptive response—at least from the perspective of a homeostatic behavioral framework. Alternatively, the anticipatory repression of oxidative respiration seems adaptive when viewed in the context of the ecologically crucial transition of *E. coli* from the external world into the mammalian GI tract.

We observed that a temperature upshift (25°C to 37°C) led to the induction of the heat shock response regulon, a set of operons activated

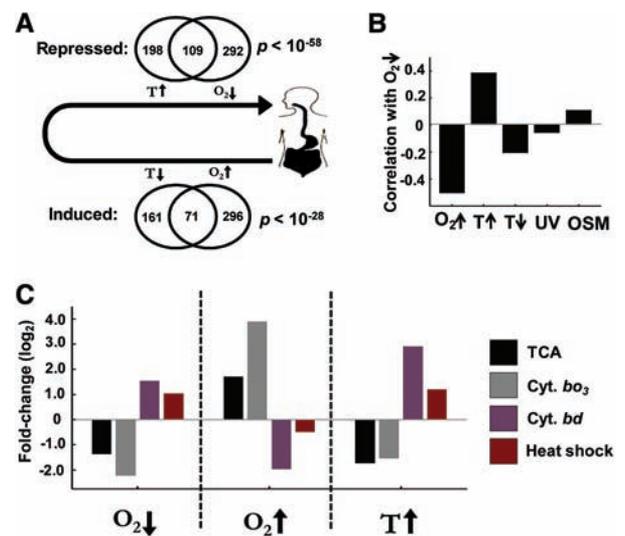
through the expression and activity of the  $\sigma^{32}$  alternative sigma factor (20). This robust response allowed us to explore the possibility of a reciprocal associative coupling, in which oxygen downshift leads to the induction of heat shock response in much the same way as temperature upshift causes respiratory repression. Remarkably, this is exactly what was observed (Fig. 5C).

**Novel environment decouples correlated transcriptional responses.** Strong correlations in the expression of distinct biochemical pathways may reflect compatibility or mutual dependence between pathway operations—as seen in the spatiotemporal separation of photosynthesis and nitrogen fixation in cyanobacteria (21)—rather than reflecting ecological structure per se. To show that the observed correlations are due to an “associative learning” paradigm, we evolved a population of *E. coli* under a dynamic environment where temperature and oxygen fluctuations had a temporal relationship counter to that expected in nature (Fig. 6A). Wild-type *E. coli* should perform poorly in this environment, because a temperature upshift will cause repression of aerobic respiratory pathways, just when oxygen saturation has achieved maximum levels. This inverted environment imposes strong selection for bacteria that fully or partially decouple the native behavior because it is highly maladaptive. If such a reprogramming can occur, then the originally observed correlated responses to temperature upshift and oxygen downshift cannot be due to hard biochemical constraints, but rather is a reflection of a common response to correlated changes in temperature and oxygen that has evolved over geological time scales.

**Fig. 5.** Global responses to oxygen and temperature perturbations reflect ecological correlation structure. (A) Transition of *E. coli* between the outside environment and the mammalian GI tract is accompanied by anticorrelated changes in temperature and oxygen. Correspondingly, highly significant overlap is seen in sets of genes down-regulated by temperature upshift and oxygen downshift (transition into the GI tract;  $P < 10^{-58}$ ). Similarly, a highly significant overlap is seen in sets of genes induced by temperature downshift and oxygen upshift (transition out of the GI tract;  $P < 10^{-28}$ ). (B) Pearson correlation of global changes in gene expression between oxygen downshift and other perturbations: oxygen upshift ( $r = -0.50$ ;  $P < 10^{-148}$ ), temperature upshift ( $r = +0.39$ ;  $P < 10^{-79}$ ), ultraviolet exposure ( $r = -0.06$ ;  $P < 0.02$ ), and osmolarity stress ( $r = 0.11$ ;  $P < 10^{-7}$ ). (C) The average relative change (log<sub>2</sub>) of TCA (*acnA*, *acnB*, *fumB*, *gltA*, *mdhA*, *sdhA*, *sdhB*, *sucB*, *sucC*, and *sucD*), cytochrome *bo*<sub>3</sub> (*cyoA*, *cyoB*, *cyoC*, *cyoE*), cytochrome *bd* (*cydA* and *cydB*), and heat shock-regulated (*clpB*, *dnaJ*, *grpE*, *hslU*, *hslV*, *htpX*, and *lon*) genes displaying reciprocal cross-regulation of aerobic respiration and heat shock in response to temperature and oxygen perturbations.

Previous evolution experiments in bacteria have focused mostly on adaptation to “steady-state” conditions, with sizable fitness increases (~30%) occurring only after thousands of generations (22). Remarkably, we witnessed large increases in reproductive fitness occurring in fewer than a hundred generations (Fig. 6B). Population growth rate increased only marginally in the 25°C and 0% O<sub>2</sub> regime, but showed more than a 50% increase within the 37°C and 21% O<sub>2</sub> regime. We attribute this rapid increase in fitness to the strength of recurring selection occurring over many cycles. Isolation of individual bacteria and competition with the parental strain (23) confirmed the fitness advantage of the evolved strain (fig. S21).

To more precisely characterize the nature of the fitness gains, we monitored growth rates of the parental and evolved strains at a higher temporal resolution across a single cycle of selection (Fig. 6C). As expected, the increase in growth rate of the evolved strain was most pronounced immediately after the temperature upshift, with maximal difference occurring immediately after the upshift in O<sub>2</sub>. This suggests that at least part of the adaptation may have been due to the expected decoupling between temperature upshift and repression of aerobic respiration. To test this possibility, we performed transcriptional profiling of parental and evolved strains in the context of a temperature upshift from 25°C to 37°C. We used Pearson correlation to measure the global similarity in transcriptional responses between the original oxygen downshift perturbation and temperature upshift. As can be seen, the strong correlation between oxygen downshift and temperature upshift is considerably

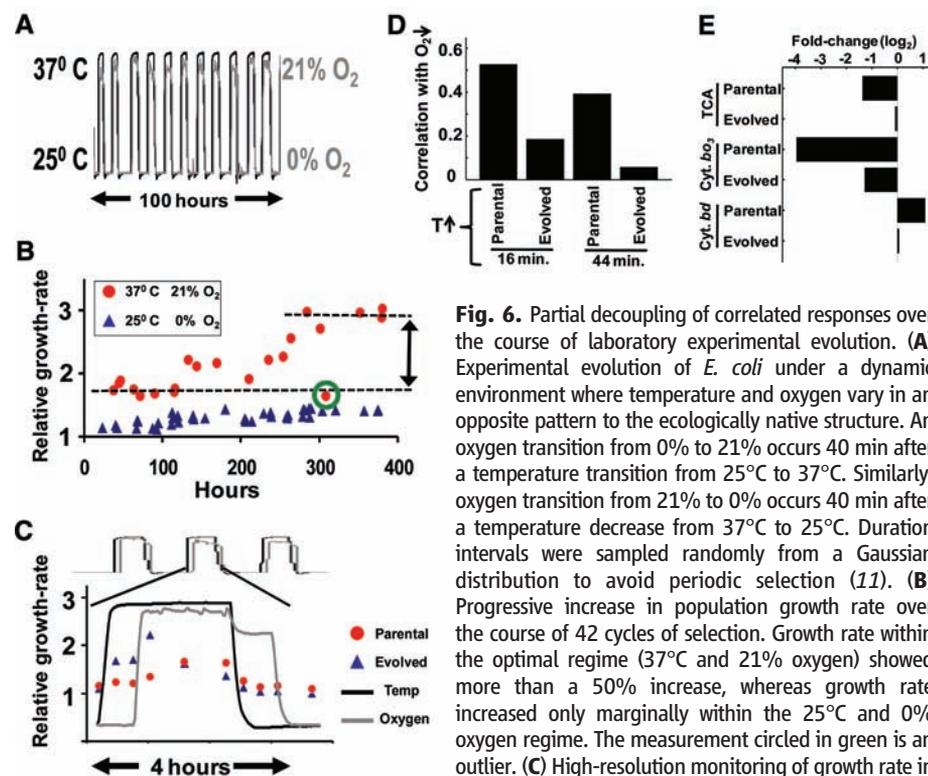


reduced in the evolved strain relative to the parental strain, both at early (16 min; parental,  $r = 0.53$ ; evolved,  $r = 0.19$ ) and late (44 min; parental,  $r = 0.39$ ; evolved,  $r = 0.06$ ) time points (Fig. 6D). This global decrease in correlation in the evolved strain was accompanied by a marked reduction in repression of TCA and cytochrome  $bo_3$  oxidase genes, as well as reduction in the activation of genes encoding cytochrome  $bd$  oxidase (Fig. 6E).

**Dynamic representations and predictive behavior.** Molecular interactions and catalytic transformations are the fundamental building blocks in all cellular processes. We have shown that randomly evolving networks composed of these basic elements are able to internalize dynamic representations of their complex environments, enabling predictive behavior, and that this ability explains the seemingly maladaptive responses of *E. coli* to transitions in oxygen and temperature. Although the correlated responses we observe precisely correspond to transitions of *E. coli* between the outside world and the mammalian GI tract, it is formally possible that they may be due to some other unknown ecological structure in the wild. What is critical, however, is that these correlations show “plasticity” over the course of laboratory experimental evolution. Our findings motivate an alternative interpretation of cellular responses to nominally stressful stimuli

(24); such “stresses” may be important to the organism not because of their immediate and direct fitness consequences, but in the information that they convey about the overall state of the environment and its likely trajectory.

Here, we have focused on the utility of learning temporally phased correlations for predicting sequential events in the environment (e.g., temperature upshift followed by oxygen downshift). However, the reciprocal cross-regulation of heat shock response and respiratory repression (Fig. 5C) suggests that *E. coli* also uses the simultaneous co-occurrence of events to reinforce perception of its immediate environment. More generally, the correlation structure of the environment can be internalized as a probabilistic model in the high-dimensional space of an organism’s complete sensory perception. As such, the very organization of microbial regulatory networks may, in large part, represent the physical instantiation of this probabilistic model. From this perspective, inferences regarding the functional utility of biological networks, including notions of modularity and optimality, may be incomplete, or even inaccurate, without considering habitat structure. Experiments of the flavor we have presented here may allow biologists to essentially go in the opposite direction and “reverse-engineer” ecological structure from physiological observations in the laboratory.



**Fig. 6.** Partial decoupling of correlated responses over the course of laboratory experimental evolution. (A) Experimental evolution of *E. coli* under a dynamic environment where temperature and oxygen vary in an opposite pattern to the ecologically native structure. An oxygen transition from 0% to 21% occurs 40 min after a temperature transition from 25°C to 37°C. Similarly, oxygen transition from 21% to 0% occurs 40 min after a temperature decrease from 37°C to 25°C. Duration intervals were sampled randomly from a Gaussian distribution to avoid periodic selection (11). (B) Progressive increase in population growth rate over the course of 42 cycles of selection. Growth rate within the optimal regime (37°C and 21% oxygen) showed more than a 50% increase, whereas growth rate increased only marginally within the 25°C and 0% oxygen regime. The measurement circled in green is an outlier. (C) High-resolution monitoring of growth rate in the parental and evolved strains shows that the fitness

differential is maximal immediately after the oxygen upshift. (D) Pearson correlation of global changes in gene expression between the oxygen downshift and temperature upshift perturbations in the parental and evolved strains at 16 and 44 min after the temperature upshift. (E) Average relative change of gene expression for the TCA, cytochrome  $bo_3$ , and cytochrome  $bd$  gene sets at 16 min after a temperature increase in both the parental and evolved strains.

Although our work on *E. coli* focused primarily on parameter changes that reflect macroscopic transitions in the environment, the dominant forces shaping microbial regulatory networks are likely to arise from microscopic correlation structure present in the complex chemistry of intra- and interspecies interactions (25, 26).

## References and Notes

- W. B. Cannon, *The Wisdom of the Body* (Norton, New York, 1932).
- D. Bell-Pedersen *et al.*, *Nat. Rev. Genet.* **6**, 544 (2005).
- A. Bergman, M. L. Siegal, *Nature* **424**, 549 (2003).
- P. Francois, V. Hakim, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 580 (2004).
- P. Francois, V. Hakim, E. D. Siggia, *Mol. Syst. Biol.* **3**, 154 (2007).
- N. Kashtan, U. Alon, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13773 (2005).
- R. E. Lenski, C. Ofria, T. C. Collier, C. Adami, *Nature* **400**, 661 (1999).
- G. Rodrigo, J. Carrera, A. Jaramillo, *Cent. Eur. J. Biol.* **2**, 233 (2007).
- O. S. Soyer, T. Pfeiffer, S. Bonhoeffer, *J. Theor. Biol.* **241**, 223 (2006).
- A. Wagner, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11775 (2005).
- See supporting material on Science Online.
- K. Basso *et al.*, *Nat. Genet.* **37**, 382 (2005).
- D. Segre, A. Deluna, G. M. Church, R. Kishony, *Nat. Genet.* **37**, 77 (2005).
- C. Dunne, *Inflamm. Bowel Dis.* **7**, 136 (2001).
- M. Schena, D. Shalon, R. W. Davis, P. O. Brown, *Science* **270**, 467 (1995).
- F. C. Neidhardt *et al.*, *Escherichia coli and Salmonella: Cellular and Molecular Biology* (American Society for Microbiology, Washington, DC, 1996).
- J. Courcelle, A. Khodursky, B. Peter, P. O. Brown, P. C. Hanawalt, *Genetics* **158**, 41 (2001).
- K. J. Cheung, V. Badarinarayana, D. W. Selinger, D. Janse, G. M. Church, *Genome Res.* **13**, 206 (2003).
- A. P. Gasch *et al.*, *Mol. Biol. Cell* **12**, 2987 (2001).
- T. Yura, K. Nakahigashi, *Curr. Opin. Microbiol.* **2**, 153 (1999).
- I. Berman-Frank *et al.*, *Science* **294**, 1534 (2001).
- R. E. Lenski, M. Travisano, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6808 (1994).
- S. F. Elena, R. E. Lenski, *Nat. Rev. Genet.* **4**, 457 (2003).
- G. Giaever *et al.*, *Nature* **418**, 387 (2002).
- B. L. Bassler, R. Losick, *Cell* **125**, 237 (2006).
- S. S. Branda, S. Vik, L. Friedman, R. Kolter, *Trends Microbiol.* **13**, 20 (2005).
- We thank the members of the Tavazoie laboratory for helpful comments on the manuscript, Princeton’s CS&ES team for supercomputer support, and S. Tavazoie, M. Tavazoie, and G. Vassilakis for crucial feedback throughout the preparation of this manuscript. Supported by a Burroughs Wellcome fellowship (I.T.) and by grants from the NSF Faculty Early Career Development (CAREER) Program, Defense Advanced Research Projects Agency, National Human Genome Research Institute, and National Institute of General Medical Sciences (P50 GM071508) (S.T.). The microarray data are deposited at NCBI Gene Expression Omnibus with accession number GSE10855.

## Supporting Online Material

www.sciencemag.org/cgi/content/full/1154456/DC1  
SOM Text  
Figs. S1 to S21  
Table S1  
References

20 December 2007; accepted 25 March 2008  
Published online 8 May 2008;  
10.1126/science.1154456  
Include this information when citing this paper.