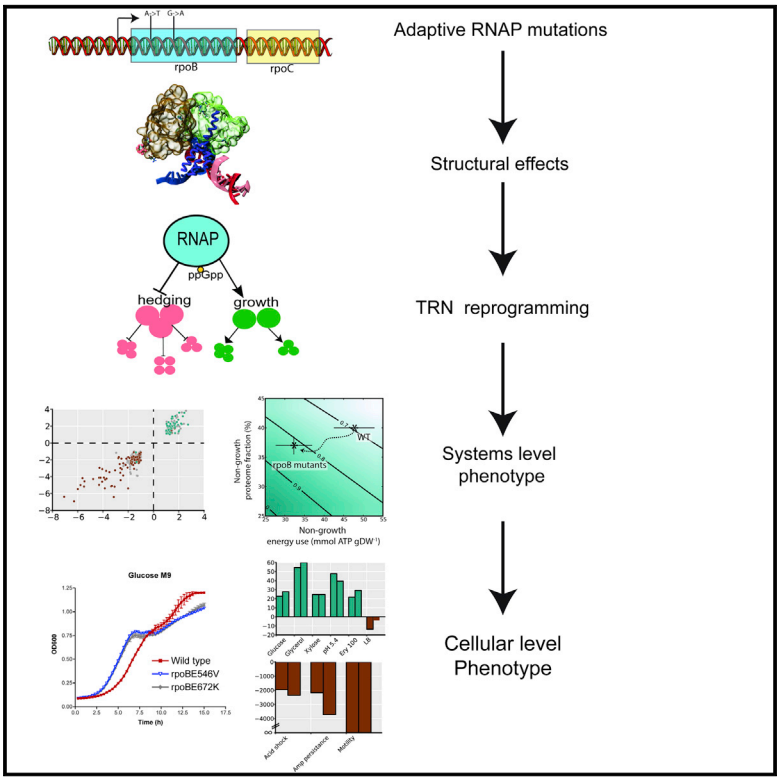
Article

Cell Systems

Global Rebalancing of Cellular Resources by Pleiotropic Point Mutations Illustrates a Multi-scale Mechanism of Adaptive Evolution

Graphical Abstract



Authors

Jose Utrilla, Edward J. O'Brien, Ke Chen, ..., Dagoberto Armenta-Medina, Adam M. Feist, Bernhard O. Palsson

Correspondence

palsson@eng.ucsd.edu

In Brief

Utrilla et al. present a multi-scale analysis of molecular, resource allocation, and physiological responses underlying the pleiotropic effects of adaptive regulatory mutations.

Highlights

- RNAP mutants show antagonistic pleiotropy for growth and hedging phenotypes
- Mutation effects on hedging function expression are independent of growth rate
- Structural analysis of the RNAP mutations suggests a common reprograming mechanism
- Model allows for quantification of relationships between fitness and resource allocation

Accession Numbers

GSE59377

Utrilla et al., 2016, Cell Systems 2, 260–271
April 27, 2016
http://dx.doi.org/10.1016/j.cels.2016.04.003

CellPress

Cell Systems
Article

Global Rebalancing of Cellular Resources by Pleiotropic Point Mutations Illustrates a Multi-scale Mechanism of Adaptive Evolution

Jose Utrilla,[1,2,8] Edward J. O'Brien,[1,3,8] Ke Chen,[1] Douglas McCloskey,[1] Jacky Cheung,[4] Harris Wang,[4] Dagoberto Armenta-Medina,[5] Adam M. Feist,[1,6] and Bernhard O. Palsson[1,6,7,*]

[1]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA
[2]Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, 62210 Cuernavaca, Morelos, México
[3]Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA 92093, USA
[4]Department of Systems Biology, Columbia University, New York, NY 10027, USA
[5]Departamento de Ingeniería Celular y Biocatálisis, Instituto de Biotecnología, UNAM, 62210 Cuernavaca, Morelos, Mexico
[6]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark
[7]Department of Pediatrics, University of California, San Diego, La Jolla, CA 92093, USA
[8]Co-first author
*Correspondence: palsson@eng.ucsd.edu
http://dx.doi.org/10.1016/j.cels.2016.04.003

SUMMARY

Pleiotropic regulatory mutations affect diverse cellular processes, posing a challenge to our understanding of genotype-phenotype relationships across multiple biological scales. Adaptive laboratory evolution (ALE) allows for such mutations to be found and characterized in the context of clear selection pressures. Here, several ALE-selected single-mutation variants in RNA polymerase (RNAP) of *Escherichia coli* are detailed using an integrated multi-scale experimental and computational approach. While these mutations increase cellular growth rates in steady environments, they reduce tolerance to stress and environmental fluctuations. We detail structural changes in the RNAP that rewire the transcriptional machinery to rebalance proteome and energy allocation toward growth and away from several hedging and stress functions. We find that while these mutations occur in diverse locations in the RNAP, they share a common adaptive mechanism. In turn, these findings highlight the resource allocation trade-offs organisms face and suggest how the structure of the regulatory network enhances evolvability.

INTRODUCTION

Many causal genetic variants across all forms of life are found in regulatory regions (Enard et al., 2014; Fraser, 2013; Jones et al., 2012; King and Wilson, 1975; Prud'homme et al., 2007; Wray, 2007). In addition to *cis*-regulatory variation, causal mutations are often found in *trans*-acting transcriptional regulators (Barrick et al., 2010; Ferenci, 2008; LaCroix et al., 2015; Sandberg et al., 2014; Saxer et al., 2014). Because of transcriptional regulators' involvement in multiple cellular processes, mutations in transcriptional regulators often affect multiple phenotypes (King et al., 2004; Solopova et al., 2014; Venturelli et al., 2015; Wang et al., 2015). Understanding the multi-scale and pleiotropic cascade of events resulting from regulatory mutations poses a challenge for systems biology.

Adaptive laboratory evolution (ALE) allows for adaptive mutations to be found and characterized in the context of clear selection pressures. ALE experiments repeatedly identify mutations in the RNA polymerase (RNAP) (Barrick et al., 2010; Dragosits et al., 2013; LaCroix et al., 2015; Sandberg et al., 2014; Tenaillon et al., 2012). Several of these mutations have been shown to result in fitness benefits in the selected environment and fitness deficits in several different environments (Cheng et al., 2014; Conrad et al., 2010; Dragosits et al., 2013). While the RNAP is not typically considered a transcription factor, it can be considered to lie at the top of the transcriptional regulatory network hierarchy. Presumably, these mutations alter the gene regulatory network to change the specific balance of cellular resources.

Here, we detail the multi-scale mechanism underlying several *trans*-acting adaptive regulatory mutations in RNAP of *Escherichia coli* (Cheng et al., 2014; Conrad et al., 2010; LaCroix et al., 2015). Using detailed phenotypic assays, we show consistent fitness effects in different environments. A multi-"omic" approach with key environmental controls reveals a systematic and consistent modulation of the transcriptional regulatory network (TRN) toward growth functions and away from functions that hedge against environmental change. "Econometric" analysis using a genome-scale model reveals that the resultant resource reallocation can quantitatively explain the fitness effects. Finally, structural dynamics of RNAP provide insight into how these mutations result in strikingly similar effects. This study moves the field forward by detailing the multi-scale molecular, phenotypic, and physiological responses underlying the pleiotropic effects of adaptive regulatory mutations. It provides insight into the evolutionary constraints and the mechanisms that govern resource allocation in simple organisms.

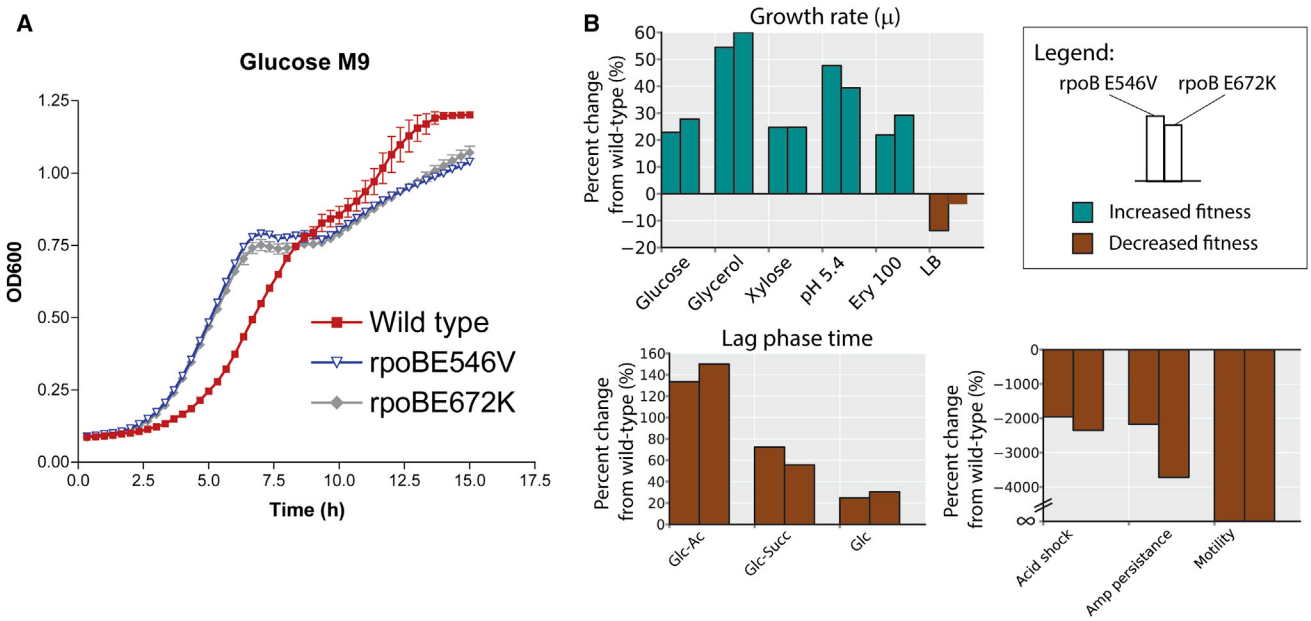


Figure 1. Growth versus Hedging Antagonistic Pleiotropy in Organismal Phenotypes

(A) Adaptive laboratory evolution (ALE)-selected *rpoB* mutations (E546V blue and E672K gray) grow faster in the glucose consumption phase but have a longer diauxic shift to grow on acetate than does wild-type (red) (Table S1).

(B) In addition to the increase of growth rate on glucose (the environment in which the mutants were selected), several additional organismal phenotypes are affected by the *rpoB* mutations. Bar charts show the percent change in measured phenotypes compared to wild-type. The steady-state growth rate increases (cyan), and the growth rate in LB medium, as well as fitness in environmental shifts and shocks, decreases (brown). LB, Luria broth; Glc, glucose, Succ, succinate; Ac, acetate; Ery 100, 100 μg/ml erythromycin; Amp, ampicillin.

See also Figures S1 and S2 and Tables S1 and S2.

RESULTS

Pleiotropy of Adaptive Mutations in RNA Polymerase

A recent adaptive laboratory evolution (ALE) experiment of *E. coli* in glucose minimal media (MM) identified recurring mutations in *rpoB* (the β subunit of RNAP), including *rpoB* E546V and *rpoB* E672K (LaCroix et al., 2015). We introduced these two ALE-selected mutations into the starting strain (i.e., the "wild-type" strain) and observed consistent physiological effects. Growth rate increased (by ~25%) as the result of increases in both biomass yield (by ~11%) and substrate uptake rate (by ~14%). The use of an automated plate reader to obtain frequent measurements revealed a diauxic shift of the mutant strains in glucose M9 mineral media (Figure 1A).

Since mutations often have positive and negative fitness effects across several environments (referred to as pleiotropy), we then assessed the growth rate of the *rpoB* E546V and *rpoB* E672K mutants under a variety of single carbon sources, mixtures of carbon sources, rich media, and stress conditions. Additionally, we performed motility, acid shock, and antibiotic persistence phenotypic tests (Figures 1B and S1; Data S1 and S2). These RNAP mutations show consistent fitness effects: they enable faster growth in several carbon sources, in low pH, and in the presence of erythromycin. However, they lead to lower motility, lower survival under acid shock, reduced antibiotic persistence, longer diauxic shifts, and lower growth rates in complex media.

Therefore, the mutants show increased fitness in conditions of steady-state growth, but a decreased fitness in changing environments. In other words, they show strong, consistent antagonistic pleiotropy for growth versus "hedging" functions.

Mutations in RNA Polymerase Are Highly Specific

To assess whether other amino-acid substitutions in the RNAP ALE-selected loci affect growth phenotypes, we used multiplex automated genome engineering (MAGE) (Wang et al., 2009) to generate variants in all possible 546 and 672 positions. After 8–12 rounds of MAGE, we isolated and verified six variants of E546X position and seven variants for the E672X position. Two amino-acid substitutions, E546K and E672R, which have similar chemical properties as those discovered by ALE, resulted in an increase in growth rate. These MAGE-selected mutants also exhibit longer diauxic shifts, showing similar pleiotropic effects as the ALE selected mutants (Figure S2). All other amino-acid substitutions generated by MAGE did not affect growth rate significantly.

These results suggest that the mutations in RNAP affecting fitness are specific. Namely, all faster growing RNAP mutants showed antagonistic pleiotropy for growth versus "hedging."

Genome-Scale Transcript Profiling Reveals Conserved Growth versus Hedging Response

To reveal the systems-level mechanism of the pleiotropic effects of the RNAP mutations, we obtained RNA sequencing (RNA-seq) and metabolomics data from mid-logarithmic growth phase in glucose minimal media for the wild-type, *rpoB* E546V, and *rpoB* E672K mutant strains (Figure S3). The concentrations

of ten metabolites, including pyrimidine, glycolytic, and TCA intermediates, changed significantly compared to wild-type (Student's t test, $p < 0.01$, Bonferroni corrected). But, overall, the metabolome remained fairly stable (Figure S3; Data S1). In contrast, the expression profiling data revealed 243 consistently differentially expressed genes. Like the pleiotropic fitness effects of the mutants, the differential gene expression is strikingly consistent (Figure 2A, left), suggesting a common underlying mechanism at the systems level.

Notably, we also found that the differential expression profiles of the two strains harboring *rpoB* mutants were similar to that of a previously profiled mutant strain that has a 27 amino-acid deletion in the β' subunit of RNAP (*rpoC*-del27, identified by ALE on glycerol) (Cheng et al., 2014; Conrad et al., 2010; Herring et al., 2006). The changes in expression of the *rpoC*-del27 mutant (Cheng et al., 2014) (compared to wild-type) grown in glycerol match those of the *rpoB* mutants grown in glucose (Figure 2A, right; χ^2 test for independence, $p < 10^{-8}$).

To obtain insight into the processes perturbed by the RNAP mutations, we classified the 243 consistently differentially expressed genes by function (Data S3). We found that the genes in the same functional category are often differentially expressed in a consistent direction. We used this observation to define upregulated and downregulated functions. The upregulated functions (defined as >80% of the genes being upregulated) were broadly related to cellular growth, including protein synthesis and folding, amino-acid biosynthesis and uptake, and carbohydrate transport and utilization. On the other hand, the downregulated functions (defined as >80% of the genes being downregulated) broadly hedge against environmental change and stress, including osmotic and oxidative stress, flagella, chemotaxis, acid resistance, and biofilm formation. Two categories of genes were not consistently up or downregulated; these are DNA repair and genes with unknown function. Some of the differentially expressed genes were transcription factors or small RNA regulators (Figure 4B). These regulators were differentially expressed in the direction expected based on the directionality of the regulator (i.e., activator or repressor) and the direction of the differential expression of the regulated functional category (Table S3; Data S3). Thus, at the molecular level, the differentially expressed genes reflect the growth versus hedging phenotypes observed at the organismal level.

Environmental Controls Disentangle Cause versus Effect of Mutations

As growth rate itself has a strong effect on gene expression (Klumpp and Hwa, 2014), we sought to identify the differential expression changes caused only by the mutation from those caused indirectly by increased growth. To disentangle these effects, we obtained RNA-seq data under conditions where the wild-type and mutant strains grow at the same rate (glucose limited chemostat culture) and under conditions where the mutants grow slower than wild-type (Luria broth [LB]-rich media). Regardless of the growth rate and environment, the hedging functions were downregulated in the mutant strain compared to wild-type (Figure 2C). Differential expression of the growth functions, however, was dependent on the growth rate: growth genes were not differentially expressed in chemostat and were downregulated in LB. Thus, these environmental controls disentangle the cause and effect of the mutations: the mutations directly result in the downregulation of hedging genes, whereas expression of the growth-related genes is coupled to the cell's growth rate.

Structural Dynamics of RNAP Suggests a Common Allosteric Mechanism

Both mutations, *rpoB* E546V and E672K, are located ~25 Å from the catalytic site of RNAP and ~25 Å from each other. How do they result in similar patterns in transcriptional reprogramming to downregulate hedging functions?

To investigate this question, we performed molecular dynamics simulations of the core RNAP open complex aiming to propose a common putative molecular mechanism for the pleiotropic fitness effects of the *rpoB* mutations. We found a strong correlation between the extent of increase in interaction energy between the β and β' subunits, and the increase in cell fitness for various E672 mutations generated by MAGE (both beneficial and neutral; Figure 3A). Such destabilization of subunit interaction is consistent with a previous study that showed a decrease in open complex half-life of the *rpoC*-del27 mutation, which has similar growth and transcriptional effects (Conrad et al., 2010).

To further explore the functional correlation among different mutations, we decomposed the RNAP complex into ~20 "structural communities," within which the molecular motions of residues are strongly correlated (Sethi et al., 2009). Despite the large spatial separation between E672 and E546, they belong to the same dynamical community (Figure 3B). Furthermore, many mutations detected in RNAP in other ALE experiments (Barrick et al., 2010; LaCroix et al., 2015; Tenaillon et al., 2012) can also be found in this and neighboring communities (Figure 3B; Table S4). This structural community consists of ~250 residues in *rpoB*, part of the bridge helix (BH) in *rpoC*, and nucleotides on the template DNA strand. As such, we may link the collective motions of these structural modules to the structural rearrangements of BH shown to coordinate catalysis and DNA translocation in the nucleotide addition cycle (NAC) (Bar-Nahum et al., 2005; Weinzierl, 2010).

Three observations about the motions of structural modules in RNAP are notable. First, the bridge helix kinked at two hinge positions (BH-H_N and BH-H_C; Figure 3C) in discrete steps of ~10° to 20°, indicating the existence of transient metastable intermediates between the RNAP open and close conformations. Second, the boundary between the two neighboring structural communities switched from BH-H_N to BH-H_C when the BH went through a transition from its bent to relaxed state. Third, there was a strong correlation between the bending angle of the BH and the relative motion between neighboring communities along the direction of DNA translocation. Based on these observations, we hypothesize that the ALE-selected mutations may modulate the stability of the RNAP open complex by adjusting the coupling between the relative motions of neighboring communities and the bending-relaxing cycle of the BH (Figure 3C).

The stability of the RNAP open complex not only affects efficiency of the nucleotide elongation reactions but may also modulate the multi-step process of bacterial transcription initiation (Saecker et al., 2011). First, it has been shown that the competitive binding of different σ factors to the RNAP core

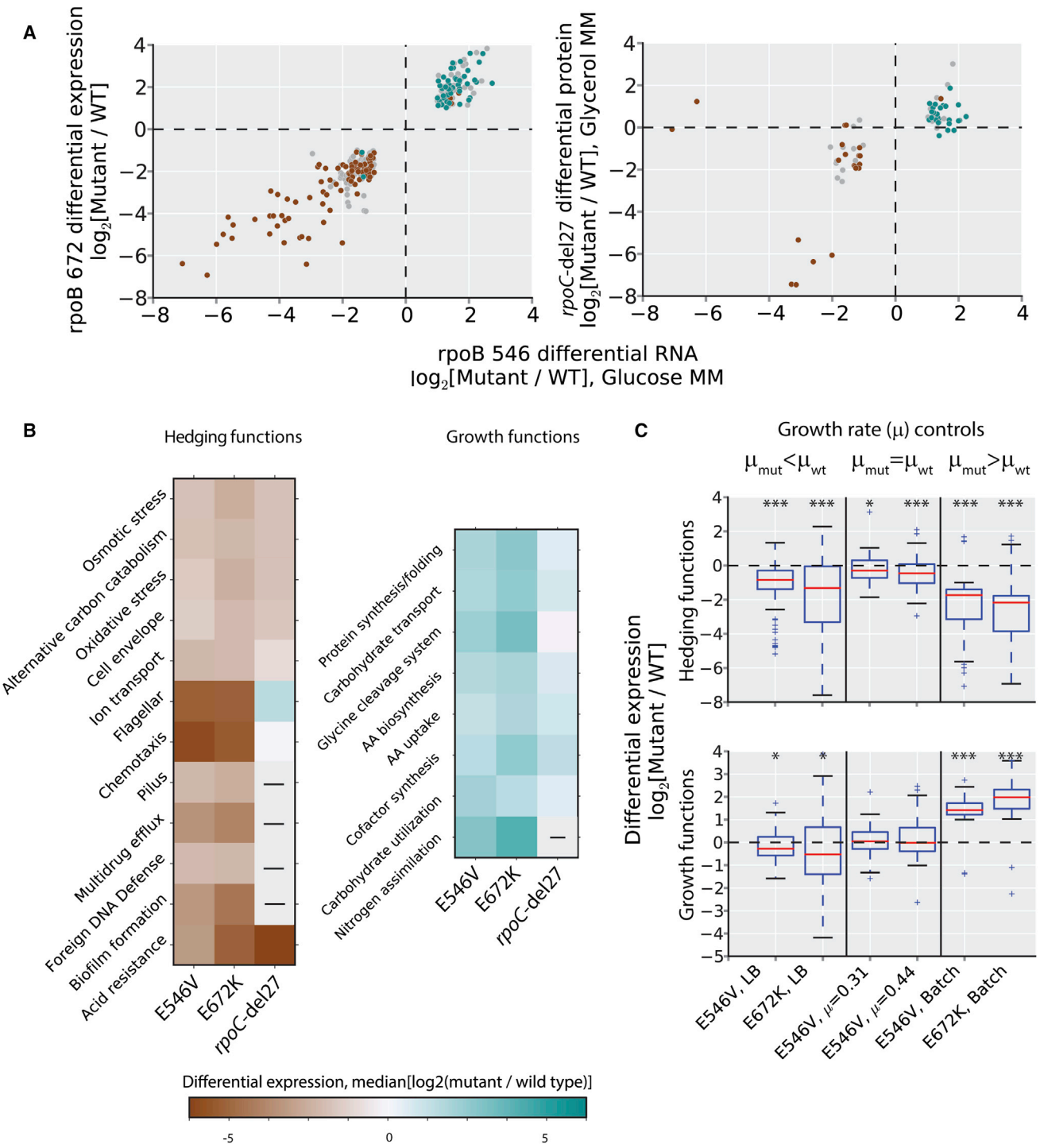


Figure 2. Consistent Molecular Growth versus Hedging Response

(A) The differential RNA expression in the ALE-selected *rpoB* mutants (E546V and E672K) is consistent (left). The differential RNA expression in glucose is also concordant with the differential protein expression measured in previous work on glycerol of an ALE-selected 27-amino acid deletion in β' compared to wild-type (*rpoC*-del27 [Cheng et al., 2014], right).

(B) Functional classification of differentially expressed genes reveals that genes with common functions are often differentially expressed in the same direction, segregating growth (upregulated, cyan) and hedging (downregulated, brown) functions. Gray dots are genes with functions not consistently differentially expressed. Median differential expression of genes in the functional categories is shown in the heat map; dashes indicate genes not detected in proteomics data (Cheng et al., 2014).

(legend continued on next page)

enzyme can be tuned by the stability of open complex in the presence of (p)ppGpp or σ^D mutations (Barker et al., 2001; Jishage et al., 2002; Österberg et al., 2011). Notably, many RNAP mutations with experimentally measured short-lived open complex are located in community 1, indicating that E672 and E546 (and other ALE-selected mutations in the same and neighboring community; Figure 3B) may achieve global transcriptional regulation in a similar way. Second, further along the transcription initiation process, lower stability of the open complex is shown to facilitate σ factor release and promoter escape, which, in turn, coordinate the probability of abortive and productive RNA synthesis (Cashel et al., 2003; Roberts and Roberts, 1996). It is worth noting that sequence and structural differences in the critical $\sigma_{3.2}$ region of σ^D and σ^S may be coupled to open complex stability in order to convey differential transcription regulation for growth-related versus hedging functions.

In summary, we observed a strong correlation between the open complex stability and the growth rates of the RNAP mutants. Detailed analysis of the structural communities suggested a common molecular mechanism, regarding how distantly located mutations may result in the same allosteric effect. Although previous findings indicated that open complex stability can modulate global transcription through alternative σ factors binding, coordination of abortive RNA synthesis, and speed of nucleotide addition, further experiments are required to draw conclusions about the exact mechanism how the ALE-selected mutations achieve these goals.

Transcriptional Regulatory Network Perturbation Explains Observed Molecular Response

Consistent with the perturbed structural properties of the mutated RNAP, the downregulated (hedging) genes tend to have promoters utilizing stress-related sigma factors (σ^S and σ^F), and the upregulated (growth) genes tend to have promoters utilizing growth-related sigma factors (σ^D, σ^N, and σ^H) (Figures 4A and S4). The sigma factors themselves are not detectably differentially expressed. However, the observed differential expression is more specific than that caused by sigma factors alone.

There are ten transcription factors (TFs) and regulatory small RNAs (sRNAs) differentially expressed in the mutant strains (Figure 4B). Each of these regulators can be associated with one or more of the differentially expressed functional categories identified (Table S3). Furthermore, across all of the strains (wild-type, *rpoB* E546V, and *rpoB* E672K) and environments (glucose excess, glucose limitation, and rich media) examined with RNA-seq, the differential expression of the identified growth and hedging functions was in a direction consistent with the differential expression of their regulators (based on known activation or repression relationships; Figure 4B). Taken together, these results suggest that the balance between growth and hedging functions is achieved through global modulation of the transcriptional regulatory network. The structure of the network may enable *E. coli* to rebalance its proteome in response to evolutionary pressures with single-point mutations in RNAP.

"Econometric" Analysis of Proteome and Energy Resource Allocation Explains the Fitness Trade-off

The molecular and regulatory effects of the *rpoB* mutations revealed that resource allocation underlies the observed growth versus hedging fitness effects. We used a recently developed genome-scale computer model of microbial growth (O'Brien et al., 2013), called an ME-model (Lerman et al., 2012; Liu et al., 2014; O'Brien et al., 2013; Thiele et al., 2009) (for metabolism and expression), to quantify the fitness effects associated with proteome and energy reallocation by these mutations (Figure 5A).

The ME-model allows global energy accounting based on the physiological data from wild-type and RNAP mutant strains. This analysis showed that the RNAP mutations shift about one-third (28%–37%) of the unaccounted for energy (i.e., processes outside of metabolism and protein synthesis, often referred to as the "maintenance energy" [Pirt, 1982]) to growth-related processes (Figure 5B). Then, using the gene expression data, we estimated a 2%–5% reduction of the transcriptome allocated to non-ME genes (i.e., genes not included in the ME-model; non-growth functions) and a commensurate increase in allocation to ME genes (i.e., modeled, growth functions) in the RNAP mutants (Figure 5B). ME-model analysis thus shows a clear shift to a more growth-supporting energy and proteome allocation as a result of the observed RNAP mutations.

We used the ME-model to understand how these changes in resource allocation affect cellular physiology (i.e., growth rate, biomass yield, and uptake rate). The non-ME proteome and energy allocation are adjustable model variables; when varied in the model, the measured changes in non-ME energy and transcriptome allocation can quantitatively account for the measured physiological changes (biomass yield and uptake rate) in the mutant strains (Figures 5C and S5). Therefore, the growth increase can be accounted for by the measured change in resource allocation, suggesting that the expression of hedging functions restrains growth rate in the wild-type strain.

The ME-model allowed us to quantitatively elucidate the relationship between changes in overall physiological measures (i.e., growth rate, substrate uptake rate, and yield) and the changes in allocation of protein and energy (Figure 5). On the basis of this quantitative relationship, we suggest that the pleiotropic effects of the *rpoB* mutation are due to a fundamental constraint of limited proteome and energy resources, leading to an inherent trade-off in resource allocation.

DISCUSSION

Here, we elucidate the mechanistic multi-scale phenotypic effects of adaptive regulatory mutations. Single amino-acid changes in the RNAP reprogram the transcriptional regulatory network to reallocate resources toward growth and away from hedging

(C) Environmental controls disentangle the direct effects of the mutations and indirect effects of changes in growth. Box plots show differential expression of identified growth and hedging functions across environments, showing that hedging functions are consistently downregulated and the expression of growth functions depends on the growth rate. Stars indicate if the mean differential expression of the group of genes is significantly different than zero, based on a two-sided t test (* $p < 0.05$, *** $p < 0.0001$).
See also Figure S3, Table S3, and Data S1 and S2.

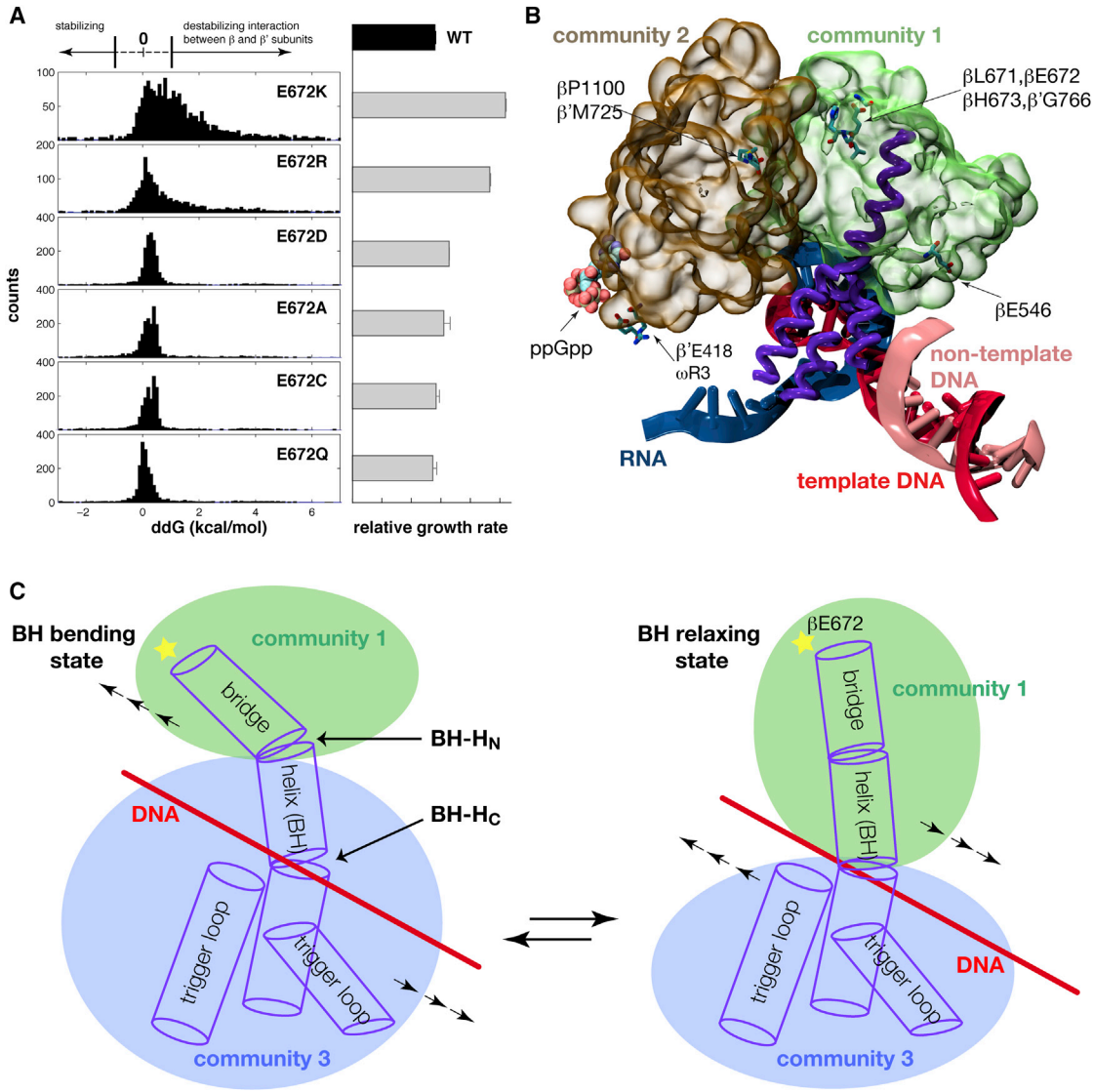


Figure 3. ALE-Selected *rpoB* Mutations Modulate Structural Dynamic of RNAP of *E. coli*

(A) Change in interaction energy between the β and β' subunits across six different E672 mutations, compared with their corresponding growth rates. To reduce bias from a single static crystal structure, interaction energy is calculated every 25 ps over a 60-ns molecular dynamic trajectory starting from the RNAP open complex.

(B) Dynamical community structures encompassing the ALE-selected mutations. As discussed in the text, Community 1 (green) includes the bridge helix in β' subunit (purple), βE672, and βE546, as well as a few other ALE-selected mutations in contact with βE672. Community 2 (brown) spans the interface between the β and β' subunits, interacting with community 1 on one side and the (p)ppGpp binding site on the other side.

(C) A schematic representation showing how relative movements between the dynamical communities modulate open complex stability. Components are color-coded as in (B). A third community (blue) is identified for calculating the correlated motions with respect to the mutation containing community 1. Continuous arrows show the direction of relative collective motions of the community structures. Effective allosteric communication between distantly located residues can be resolved from the optimal path calculated based on a dynamical correlation network. The result shows that βE672 and βE546 share the same optimal dynamical path (orange) toward the (p)ppGpp binding site in the ω subunit. Structural elements are shown from the same perspective and color-coded as in (B). See also Figure S2 and Table S4.

functions. The mutations result in antagonistic pleiotropy in which the organism is more fit in stable environments but less fit in environmental shifts and shocks (Futuyma and Moreno, 1988).

Antagonistic Pleiotropy due to a Fundamental Trade-off

Mutations that are beneficial or neutral in one environment, but often have negative fitness effects in other environments are referred to as pleiotropy. Pleiotropy shapes the evolution of organisms and is thought to underlie the evolution of specialist species. Several mechanisms can give rise to pleiotropy and some have been demonstrated (Cooper and Lenski, 2000; Leiby and Marx, 2014; Remold, 2012).

Fundamental biological constraints can result in antagonistic pleiotropy, though examples of these cases are lacking. Using

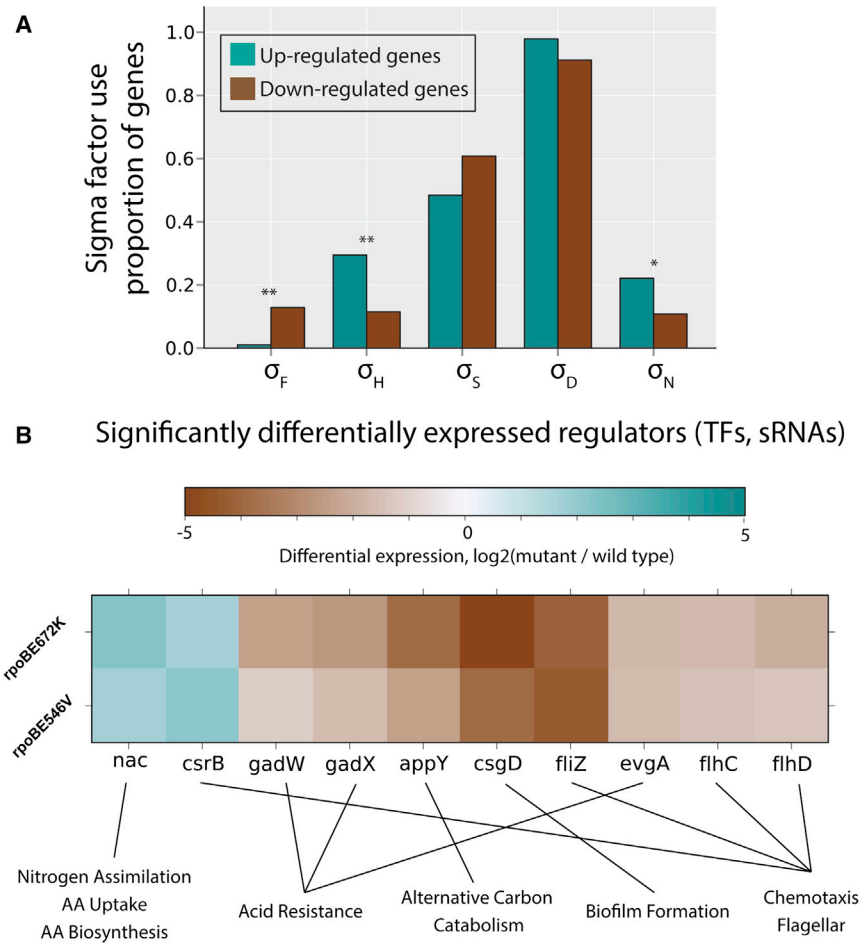


Figure 4. Reprograming of the Regulatory Network

(A) The σ factor usage of differentially expressed genes in mutant strains is shown. Bars indicate the fraction of upregulated (cyan) and downregulated (brown) genes that have a promoter regulated by a given σ factor. Only σ factors with greater than 10% of promoters regulated among either upregulated or downregulated genes are shown. Significant differences in the proportion between σ factor use in upregulated and downregulated genes are indicated with asterisks; one asterisk indicates $p < 0.05$, and two asterisks indicate $p < 0.005$.

(B) The fold change for transcription factors and sRNA that are significantly differentially expressed in both mutant strains compared to wild-type are shown.

See also Figure S4, Table S3, and Data S3.

a systems biology approach, we show that the growth rate difference in wild-type and mutant strains can be quantitatively explained by changes in proteome and energy allocation. These resources are limited, resulting in an inherent trade-off between growth and hedging functions. Such proteome and energy allocation constraints likely result in pervasive evolutionary trade-offs and likely underlie several recent examples of antagonistic pleiotropy (Solopova et al., 2014; Venturelli et al., 2015; Wang et al., 2015).

Bacterial Evolvability through Regulatory Network Structure

Mounting evidence supports that much of the functional divergence between organisms occurs in regulatory regions (Enard et al., 2014; Fraser, 2013; Jones et al., 2012; King and Wilson, 1975; Prud'homme et al., 2007; Wray, 2007). The detailed example of the RNAP mutations here suggests why (in part) this may be the case.

As regulatory networks are "aligned" with particular functional subsystems, mutations that perturb them change phenotypes in a functionally coherent manner (Innocenti and Chenoweth, 2013; Saxer et al., 2014; Wagner et al., 2007). The regulatory rebalancing detailed here occurs along a coherent growth versus hedging trajectory. On the other hand, mutations that are inconsistent or imbalanced in the molecular changes they cause would likely not be selected. Therefore, in addition to enabling short-term responses to environmental change, the structure of the regulatory network also enables longer-term productive evolutionary change. Remarkably, single, but non-unique, point mutations allow such adaptation.

Multi-scale Characterization of Genotype to Phenotype

Sequencing of many individual genomes has led to the identification of the genomic regions under selection (Grossman et al., 2013) and enabled the association of variants with organismal (McCarthy et al., 2008) and molecular (Cookson et al., 2009) phenotypes. However, there is a large gap between identifying causal variants and mechanistically understanding their phenotypic consequences. The mutations studied here are some of the most comprehensively phenotyped to date, with environmental controls to separate cause and effect. We employ state-of-the-art structural and systems biology modeling approaches to help bridge the gap between genotype and phenotype. Together, these analysis approaches enabled us to infer links from mutation to biophysical effects on protein function to systems-level molecular and regulatory response and, finally, to organismal phenotype (Figure 6). Therefore, this study outlines how we might begin to understand the multi-scale genotype-phenotype relationship at a true systems level.

EXPERIMENTAL PROCEDURES

Strains and Cultivations

E. coli MG1655 was used as wild-type. The ALE-selected rpoBE564V and rpoBE672K knock-in strains were previously constructed by allelic replacement (LaCroix et al., 2015). To generate additional variants of *rpoB* 546 and 672 positions, MAGE was performed on the wild-type strain first by transformation of recombineering plasmid pKD46 (Datsenko and Wanner, 2000) and then by inactivation of *mutS* with two nonsense mutations at residues 189 and 191 using an oligo (mutS_MUT). Two oligos (rpoB_E546X and rpoB_E672X) that resulted in NNS codon mutations at *rpoB* residues 546 and 672

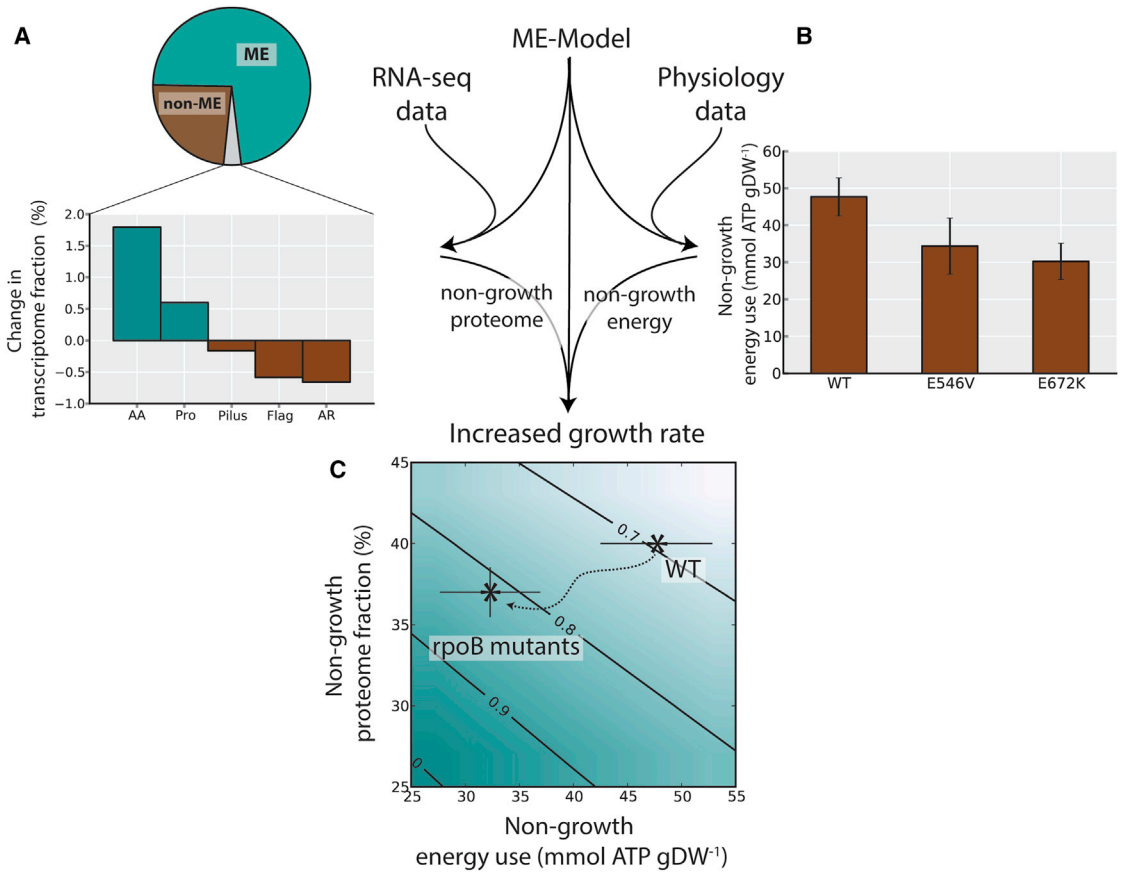


Figure 5. The Changes and Effects of Proteomic and Energetic Resource Allocation

(A) A genome-scale model of metabolism and gene expression (ME-model) is used to integrate the RNA-sequencing and physiological data. The transcriptome fraction devoted to ME and non-ME (i.e., not included in the ME-model) genes is calculated for the wild-type and mutant strains. The gray area of the pie chart indicates the fraction of the transcriptome reallocated from non-ME to ME genes. The bar chart shows the functional categories that reduced or increased in expression by more than 0.1% of the total transcriptome. All percentages are shown as the average for E546V and E672K. AA, amino acid biosynthesis; Pro, protein synthesis/folding; AR, acid resistance; Fla, flagellar.

(B) The physiological data were used to calculate the energy use not accounted for by the ME-model (see the Experimental Procedures, Computation of maximum unaccounted for energy), showing a reduction in unaccounted for energy use in *rpoB* mutants compared to wild-type. Error bars indicate SE across biological replicates.

(C) The effects of non-ME protein and energy use on maximal growth rates in the ME-model are computed and shown in the contour plot (see the Experimental Procedures). The wild-type and mutant strains are indicated on the plot, showing how lower non-ME protein and energy use can cause increased growth.

See also Figure S5.

were introduced into the strain through 8–12 rounds of MAGE, followed by colony isolation of mutants, PCR verification, and Sanger sequencing. To perform each cycle of MAGE, the λ-Red system was induced with 0.5% arabinose 45 min prior to generation of electrocompetent cells and oligo. Batch cultures were done in a flask with M9 minimal media and 4 g/l of glucose at 37°C or LB-rich media. Glucose-limited chemostats were carried out in a Bioflo 110 fermentor (New Brunswick Scientific). Glucose-supplemented M9 was added to the reactor at 0.31 and 0.44 h^{-1} dilution rates controlled by a peristaltic pump. Steady state was achieved after 3–5 residence times and was verified by biomass measurements. For all cases, phenotypic tests were performed by inoculation of the media with an overnight pre-culture of glucose M9 media. Erythromycin was added to the media to the indicated concentration. The pH of M9 was adjusted to the indicated value with 6 M HCl. Different substrates and mixtures were added to M9 to test growth in the indicated conditions. All growth curves were inoculated to a 0.02 OD, and 200 μl were cultured by triplicate in a Bioscreen C device at 37°C for 15–24 hr. Growth rates were calculated by determining the slope of the log-linear region using a linear regression; the first growth phase was used to do calculations when diauxic growth was observed.

Motility Test

Cells were grown to mid-log phase, and 10 μl of cells in suspension were spotted onto 0.3% agar plate with glucose M9 media. Plates were photographed, and motility was determined by halo expansion between 24 and 48 hr.

Acid Shock

Cells were harvested in mid-log phase and normalized to 1 × 10^8 cells/ml. 50 μl of cells in suspension were diluted in 950 μl of pH 2.6 glucose M9 media. After 3 hr of incubation, cells were diluted and plated on LB agar plates for cell counts (Tucker et al., 2003).

Antibiotic Persistence

Cells were harvested in mid-log phase and normalized to 1 × 10^8 cells/ml. Different dilutions were plated on LB-ampicillin plates. After 24 hr, a sterile solution of 25 U of penicinillase was plated, and plates were re-incubated for 24 hr. The appearance of colonies was determined, and persistence frequency was determined using the base of the initial cell counts (Korch et al., 2003).

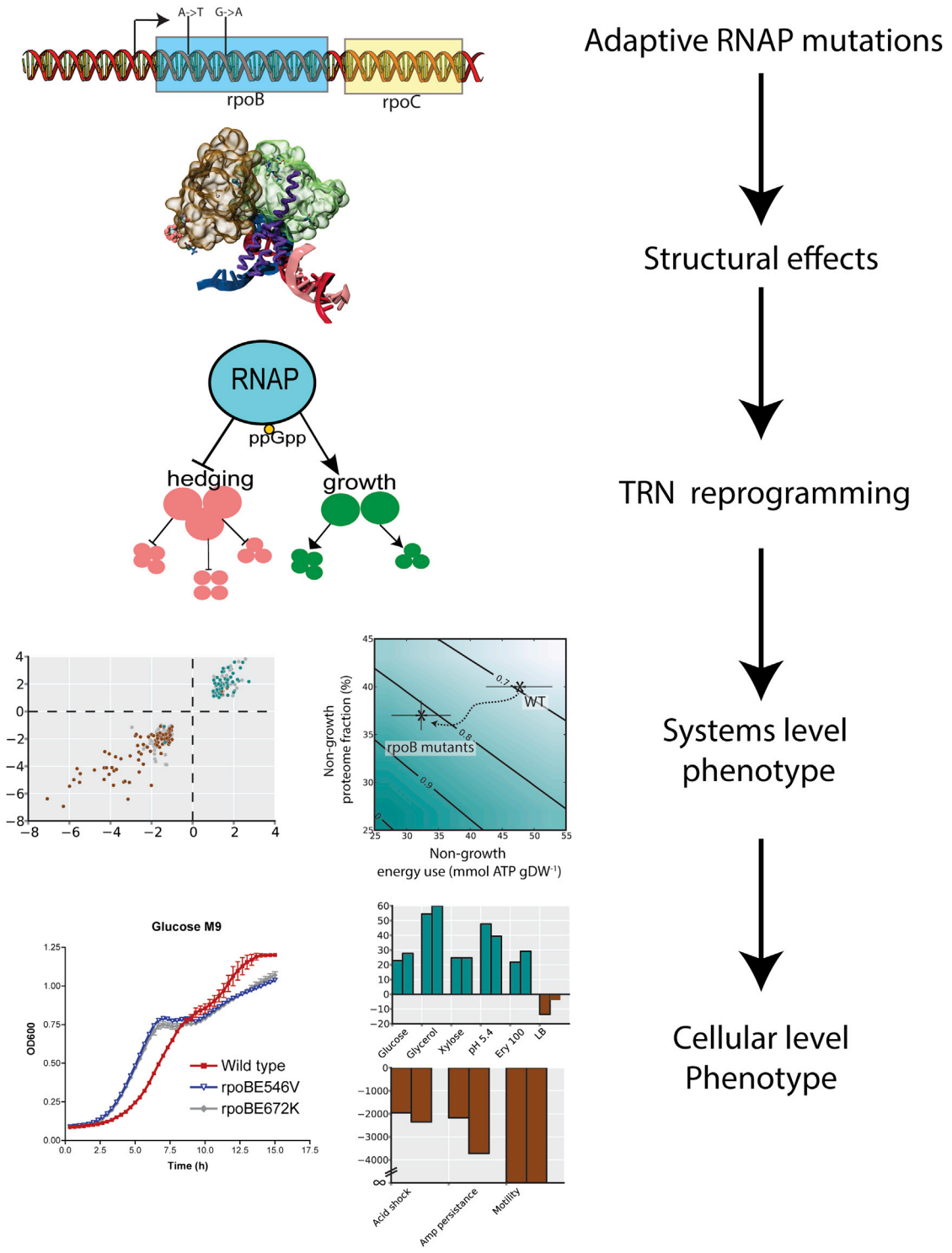


Figure 6. Multi-scale Characterization from Genotype to Phenotype

The multi-scale effects of the studied adaptive regulatory mutations in RNAP are summarized. The mutations alter the structural dynamics of RNAP, perturbing the transcriptional regulatory network through the action of key transcription factors. The decrease in expression of hedging functions lowers the proteome and energy allocation toward hedging functions and increases cellular growth. In turn, the cell can grow faster in conditions of steady-state growth, but is less fit under environmental shifts and shocks.

Analytics

Biomass was determined by measuring the absorbance of the culture at 600 nm using an equivalence of 0.429 gDW/l per optical density 600 unit. Glucose and acetate were measured by high-pressure liquid chromatography (HPLC) using refractive index (RI) detection by high-performance liquid chromatography (HPLC) (Waters) with a Bio-Rad Aminex HPX87-H ion-exclusion column (injection volume, 10 μl) and 5 mM H_2SO_4 as the mobile phase (0.5 ml/min, 45°C). Metabolomic sampling, extraction, and analysis were carried out as described earlier by our group (McCloskey et al., 2014, 2015). t tests were performed on log-normalized data with MetaboAnalyst (v.3.0) (Xia et al., 2015), comparing wild-type to each of the mutants; to determine significance between metabolite concentration levels, a p value cutoff of 0.01 after Bonferroni correction was used.

RNA-Seq Libraries

Samples for RNA-sequencing were taken in mid-log phase of batch cultures or during the steady state in chemostats. Cells were collected with QIAGEN RNA-protect Bacteria Reagent and pelleted for storage at −80°C prior to RNA extraction. Cell pellets were thawed and incubated with Readylyse Lysozyme, SuperaseIn, Protease K, and 20% SDS for 20 min at 37°C. Total RNA was isolated and purified using the QIAGEN RNeasy Mini Kit columns and following vendor procedures. An on-column DNase treatment was performed for 30 min at room temperature. RNA was quantified using a NanoDrop and quality assessed by running an RNA-nano chip on a bioanalyzer. The rRNA was removed using Epicenter's Ribo-Zero rRNA Removal Kit for Gram-negative bacteria. Paired-end, strand-specific RNA-seq was performed following a modified deoxy uridine triphosphate method (Latif et al., 2013).

Transcriptome Analyses

The obtained reads were mapped to the *E. coli* MG1655 genome (NC_000913.2) using the short-read aligner Bowtie (http://bowtie-bio.sourceforge.net) (Langmead, 2010) with two mismatches allowed per read alignment. To estimate gene expression, fragments per kilobase of transcript per million of mapped reads (FPKM) values were calculated using cufflinks tool and differential expression analysis was carried out using cuffdiff feature of the same package using the upper quartile normalization (http://cufflinks.cbcb.umd.edu/) (Trapnell et al., 2010).

Regulatory Network

Sigma factor use at promoters was obtained by combining annotations in Cho et al. (2014) and EcoCyc (Keseler et al., 2013). The list of all transcription factors and sRNAs was obtained from RegulonDB (Salgado et al., 2013). A two-proportion z test with two-tailed comparisons was used to determine significant differences in sigma factor usage among upregulated and downregulated genes.

Computation of Maximum Non-growth Energy Use

The *E. coli* ME-model with all parameters as published in O'Brien et al. (2013) was used. For all replicate cultivations, the measured growth rate, glucose uptake rate, and acetate secretion rate were fixed in the model. The maximum unaccounted for energy use was then computed by maximizing the flux through ATP maintenance reaction, which hydrolyzes ATP. For a given strain, the unaccounted for energy use is reported as the average across biological replicates.

Computation of Non-ME Transcriptome

The (protein coding) ME and non-ME transcriptome fractions were estimated using FPKM and gene length. A gene's transcriptome fraction was taken to be the product of FPKM and the gene length, divided by the sum of this product over all genes. The ME and non-ME transcriptome fractions were then calculated by summing the transcriptome fractions of all ME and non-ME genes, respectively. Ranges are determined from the estimated lower and upper FPKM values across different samples.

Computation of the Effects of Changes in Resource Allocation

Protein and energy not used toward cell growth are changeable variables in the ME-model. These are varied to determine the growth rate, biomass yield, and substrate uptake rate contours (Figures 5C and S5). The points and error bars for wild-type and *rpoB* mutants are placed according to the unaccounted for energy (Figure 5C) and change in non-ME transcriptome (Figure 5B). As we do not explicitly know the proteome fraction devoted to growth in each strain, we determine these values with two assumptions. First, we assume the change in non-growth proteome is equal to the change in the non-ME transcriptome. Second, we infer the non-growth proteome in the wild-type strain based on its measured growth (which is why there is no y axis error bar for wild-type), resulting in a value consistent with previous estimates (Scott et al., 2010).

Molecular Dynamics Simulations

The molecular model of the *E. coli* RNAP elongation complex (EC) was created using the crystal structure of the *E. coli* RNAP core enzymes (PDB: 3LU0 [Opalka et al., 2010]), the template and non-template DNA strands, and the DNA:RNA hybrid helix (PDB: 2O5J [Vassylyev et al., 2007]). The model represented the open conformation of the RNAP EC, and we model selected substrates into the active site to mimic the pre-translocated state (with ADE), the post-translocated state (no additional substrate), and the post-translocated pre-insertion state (with ATP) during the nucleotide-elongation cycle.

All three systems were neutralized with Mg^{2+} and K^+ ions, initially placed in positions occupied by metal ions in the crystal structure or according to the electrostatic potential. The complexes were then solvated by well-equilibrated water molecules with periodic boundary conditions. 200 mM KCl was added to the final solution. Minimization and equilibration were conducted by releasing constraints step-by-step according to the protocol described in Eargle and Luthey-Shulten (2012) and Lai et al. (2013). This procedure was shown to be critical in maintaining stability of a complex molecular system with interactions between protein, RNA, and DNA. Production runs were performed with 1-fs time step under constant pressure (1 atm) and constant temperature (25°C) using NAMD (v.2.9) (Phillips et al., 2005) and the CHARMM36 force field (Best et al., 2012) A total of 140-ns molecular dynamics (MD) simulations of the RNAP open complex were obtained for analysis, including 50 ns from the pre-translocated state, 60 ns from the pre-translocated state, and 30 ns from the pre-translocated, pre-inserted state.

Analysis of the Molecular Dynamics Trajectory

Analysis of the three MD trajectories showed consistent results. Therefore, we present as the representative the results from the longest simulation of the post-translocated state.

Interaction energy change between the β and β' subunits after mutation was calculated with the alanine scan script using PyRosetta (Chaudhury et al., 2010; originally distributed by the Gray lab) (http://graylab.jhu.edu/pyrosetta/downloads/pyrosetta_scripts/ala_scan.zip). We applied modifications of the score function parameterized according to recently reported protocols (Gavenonis et al., 2014; Kortemme and Baker, 2002). To reduce the bias introduced by a single static crystal structure, we performed the computational alanine scan every 25 ps through the entire trajectories, resulting in a broad distribution of the ddG values. Although this ddG value was taken to be qualitative conventionally (with ddG > 1 kcal/mol to be destabilizing), we emphasized that it was the observed trend over the dynamical trajectory that correlated with phenotypic fitness of the MAGE mutants.

Dynamic community analysis was done using algorithms described in Sethi et al. (2009), with the NetworkView plugin (Eargle and Luthey-Shulten, 2012) in VMD (Humphrey et al., 1996). The analysis was done for two chosen windows (3–23 and 35–55 ns) in the 60-ns trajectory of the post-translocated state, because the overall root-mean-square deviation of the RNAP complex suggested that they might represent two different conformations during the RNAP EC open-close conformational transition.

ACCESSION NUMBERS

The accession number for the RNAseq data reported in this paper is GEO: GSE59377.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, four tables, and three data files and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2016.04.003.

AUTHOR CONTRIBUTIONS

J.U. and B.O.P. conceived of the study. J.U. performed the physiological and RNA-seq experiments. D.M. performed metabolomics experiments. J.C. and H.W. generated mutation variants by MAGE. J.U. and E.J.O. performed data processing and computational analysis. K.C. and D.A.-M. performed structural analysis. B.O.P. supervised the study. A.M.F. provided critical feedback. J.U., E.J.O., K.C., and B.O.P. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Richard Szubin for technical assistance and Markus Herrgard and Elizabeth Brunk for valuable discussions. The Novo Nordisk Foundation supported this work. E.J.O. was supported by the NIH (GM057089). This research used the resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy (DE-AC02-05CH11231).

Received: August 28, 2015
Revised: December 10, 2015
Accepted: April 6, 2016
Published: April 27, 2016

REFERENCES

Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A., and Nudler, E. (2005). A ratchet mechanism of transcription elongation and its control. Cell *120*, 183–193.

Barker, M.M., Gaal, T., and Gourse, R.L. (2001). Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. J. Mol. Biol. *305*, 689–702.

Barrick, J.E., Kauth, M.R., Strelioff, C.C., and Lenski, R.E. (2010). Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects. Mol. Biol. Evol. *27*, 1338–1347.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E., Mittal, J., Feig, M., and Mackerell, A.D., Jr. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. J. Chem. Theory Comput. *8*, 3257–3273.

Cashel, M., Hsu, L.M., and Hernandez, V.J. (2003). Changes in conserved region 3 of Escherichia coli sigma 70 reduce abortive transcription and enhance promoter escape. J. Biol. Chem. *278*, 5539–5547.

Chaudhury, S., Lyskov, S., and Gray, J.J. (2010). PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. Bioinformatics *26*, 689–691.

Cheng, K.K., Lee, B.S., Masuda, T., Ito, T., Ikeda, K., Hirayama, A., Deng, L., Dong, J., Shimizu, K., Soga, T., et al. (2014). Global metabolic network reorganization by adaptive mutations allows fast growth of Escherichia coli on glycerol. Nat. Commun. *5*, 3233.

Cho, B.K., Kim, D., Knight, E.M., Zengler, K., and Palsson, B.O. (2014). Genome-scale reconstruction of the sigma factor network in Escherichia coli: topology and functional states. BMC Biol. *12*, 4.

Conrad, T.M., Frazier, M., Joyce, A.R., Cho, B.K., Knight, E.M., Lewis, N.E., Landick, R., and Palsson, B.O. (2010). RNA polymerase mutants found through adaptive evolution reprogram Escherichia coli for optimal growth in minimal media. Proc. Natl. Acad. Sci. USA *107*, 20500–20505.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex disease traits with global gene expression. Nat. Rev. Genet. *10*, 184–194.

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

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. USA *97*, 6640–6645.

Dragosits, M., Mozhayskiy, V., Quinones-Soto, S., Park, J., and Tagkopoulos, I. (2013). Evolutionary potential, cross-stress behavior and the genetic basis of acquired stress resistance in Escherichia coli. Mol. Syst. Biol. *9*, 643.

Eargle, J., and Luthey-Shulten, Z. (2012). Simulating dynamics in RNA-protein complexes. In RNA 3D Structure Analysis and Prediction, N.L.E. Westhof, ed. (Springer-Verlag), pp. 213–238.

Enard, D., Messer, P.W., and Petrov, D.A. (2014). Genome-wide signals of positive selection in human evolution. Genome Res. *24*, 885–895.

Ferenci, T. (2008). The spread of a beneficial mutation in experimental bacterial populations: the influence of the environment and genotype on the fixation of rpoS mutations. Heredity (Edinb) *100*, 446–452.

Fraser, H.B. (2013). Gene expression drives local adaptation in humans. Genome Res. *23*, 1089–1096.

Futuyma, D.J., and Moreno, G. (1988). The evolution of ecological specialization. Annu. Rev. Ecol. Syst. *19*, 207–233.

Gavenonis, J., Sheneman, B.A., Siegert, T.R., Eshelman, M.R., and Kritzer, J.A. (2014). Comprehensive analysis of loops at protein-protein interfaces for macrocycle design. Nat. Chem. Biol. *10*, 716–722.

Grossman, S.R., Andersen, K.G., Shlyakhter, I., Tabrizi, S., Winnicki, S., Yen, A., Park, D.J., Griesemer, D., Karlsson, E.K., Wong, S.H., et al.; 1000 Genomes Project (2013). Identifying recent adaptations in large-scale genomic data. Cell *152*, 703–713.

Herring, C.D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M.K., Joyce, A.R., Albert, T.J., Blattner, F.R., van den Boom, D., Cantor, C.R., and Palsson, B.Ø. (2006). Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale. Nat. Genet. *38*, 1406–1412.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graphics *14*, 33–38.

Innocenti, P., and Chenoweth, S.F. (2013). Interspecific divergence of transcription networks along lines of genetic variance in Drosophila: dimensionality, evolvability, and constraint. Mol. Biol. Evol. *30*, 1358–1367.

Jishage, M., Kvint, K., Shingler, V., and Nyström, T. (2002). Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev. *16*, 1260–1270.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., et al.; Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team (2012). The genomic basis of adaptive evolution in threespine sticklebacks. Nature *484*, 55–61.

Keseler, I.M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., Fulcher, C., Huerta, A.M., Kothari, A., Krummenacker, M., et al. (2013). EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res. *41*, D605–D612.

King, M.C., and Wilson, A.C. (1975). Evolution at two levels in humans and chimpanzees. Science *188*, 107–116.

King, T., Ishihama, A., Kori, A., and Ferenci, T. (2004). A regulatory trade-off as a source of strain variation in the species Escherichia coli. J. Bacteriol. *186*, 5614–5620.

Klumpp, S., and Hwa, T. (2014). Bacterial growth: global effects on gene expression, growth feedback and proteome partition. Curr. Opin. Biotechnol. *28*, 96–102.

Korch, S.B., Henderson, T.A., and Hill, T.M. (2003). Characterization of the hipA7 allele of Escherichia coli and evidence that high persistence is governed by (p)ppGpp synthesis. Mol. Microbiol. *50*, 1199–1213.

Kortemme, T., and Baker, D. (2002). A simple physical model for binding energy hot spots in protein-protein complexes. Proc. Natl. Acad. Sci. USA *99*, 14116–14121.

LaCroix, R.A., Sandberg, T.E., O'Brien, E.J., Utrilla, J., Ebrahim, A., Guzman, G.I., Szubin, R., Palsson, B.O., and Feist, A.M. (2015). Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of Escherichia coli K-12 MG1655 on glucose minimal medium. Appl. Environ. Microbiol. *81*, 17–30.

Lai, J., Chen, K., and Luthey-Schulten, Z. (2013). Structural intermediates and folding events in the early assembly of the ribosomal small subunit. J. Phys. Chem. B *117*, 13335–13345.

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. Curr. Protoc. Bioinformatics *Chapter 11*. Unit 11.17.

Latif, H., Lerman, J.A., Portnoy, V.A., Tarasova, Y., Nagarajan, H., Schrimpe-Rutledge, A.C., Smith, R.D., Adkins, J.N., Lee, D.H., Qiu, Y., and Zengler, K. (2013). The genome organization of Thermotoga maritima reflects its lifestyle. PLoS Genet. *9*, e1003485.

Leiby, N., and Marx, C.J. (2014). Metabolic erosion primarily through mutation accumulation, and not tradeoffs, drives limited evolution of substrate specificity in Escherichia coli. PLoS Biol. *12*, e1001789.

Lerman, J.A., Hyduke, D.R., Latif, H., Portnoy, V.A., Lewis, N.E., Orth, J.D., Schrimpe-Rutledge, A.C., Smith, R.D., Adkins, J.N., Zengler, K., and Palsson, B.O. (2012). In silico method for modelling metabolism and gene product expression at genome scale. Nat. Commun. *3*, 929.

Liu, J.K., O'Brien, E.J., Lerman, J.A., Zengler, K., Palsson, B.O., and Feist, A.M. (2014). Reconstruction and modeling protein translocation and compartmentalization in Escherichia coli at the genome-scale. BMC Syst. Biol. *8*, 110.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat. Rev. Genet. *9*, 356–369.

McCloskey, D., Utrilla, J., Naviaux, R.K., Palsson, B.O., and Feist, A.M. (2014). Fast Swinnex filtration (FSF): a fast and robust sampling and extraction method suitable for metabolomics analysis of cultures grown in complex media. Metabolomics *11*, 198–209.

McCloskey, D., Gangoiti, J.A., Palsson, B.O., and Feist, A.M. (2015). A pH and solvent optimized reverse-phase ion-paring-LC-MS/MS method that leverages multiple scan-types for targeted absolute quantification of intracellular metabolites. Metabolomics *11*, 1338–1350.

O'Brien, E.J., Lerman, J.A., Chang, R.L., Hyduke, D.R., and Palsson, B.O. (2013). Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. Mol. Syst. Biol. *9*, 693.

Opalka, N., Brown, J., Lane, W.J., Twist, K.A., Landick, R., Asturias, F.J., and Darst, S.A. (2010). Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. PLoS Biol. *8*, 8.

Österberg, S., del Peso-Santos, T., and Shingler, V. (2011). Regulation of alternative sigma factor use. Annu. Rev. Microbiol. *65*, 37–55.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Pirt, S.J. (1982). Maintenance energy: a general model for energy-limited and energy-sufficient growth. Arch. Microbiol. *133*, 300–302.

Prud'homme, B., Gompel, N., and Carroll, S.B. (2007). Emerging principles of regulatory evolution. Proc. Natl. Acad. Sci. USA *104* (*Suppl 1*), 8605–8612.

Remold, S. (2012). Understanding specialism when the jack of all trades can be the master of all. Proc. Biol. Sci. *279*, 4861–4869.

Roberts, C.W., and Roberts, J.W. (1996). Base-specific recognition of the nontemplate strand of promoter DNA by E. coli RNA polymerase. Cell *86*, 495–501.

Saecker, R.M., Record, M.T., Jr., and Dehaseth, P.L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. J. Mol. Biol. *412*, 754–771.

Salgado, H., Peralta-Gil, M., Gama-Castro, S., Santos-Zavaleta, A., Muñiz-Rascado, L., García-Sotelo, J.S., Weiss, V., Solano-Lira, H., Martínez-Flores, I., Medina-Rivera, A., et al. (2013). RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. Nucleic Acids Res. *41*, D203–D213.

Sandberg, T.E., Pedersen, M., LaCroix, R.A., Ebrahim, A., Bonde, M., Herrgard, M.J., Palsson, B.O., Sommer, M., and Feist, A.M. (2014). Evolution of Escherichia coli to 42°C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. Mol. Biol. Evol. *31*, 2647–2662.

Saxer, G., Krepps, M.D., Merkley, E.D., Ansong, C., Deatherage Kaiser, B.L., Valovska, M.T., Ristic, N., Yeh, P.T., Prakash, V.P., Leiser, O.P., et al. (2014). Mutations in global regulators lead to metabolic selection during adaptation to complex environments. PLoS Genet. *10*, e1004872.

Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. Science *330*, 1099–1102.

Sethi, A., Eargle, J., Black, A.A., and Luthey-Schulten, Z. (2009). Dynamical networks in tRNA:protein complexes. Proc. Natl. Acad. Sci. USA *106*, 6620–6625.

Solopova, A., van Gestel, J., Weissing, F.J., Bachmann, H., Teusink, B., Kok, J., and Kuipers, O.P. (2014). Bet-hedging during bacterial diauxic shift. Proc. Natl. Acad. Sci. USA *111*, 7427–7432.

Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R.L., McDonald, P., Bennett, A.F., Long, A.D., and Gaut, B.S. (2012). The molecular diversity of adaptive convergence. Science *335*, 457–461.

Thiele, I., Jamshidi, N., Fleming, R.M., and Palsson, B.O. (2009). Genome-scale reconstruction of Escherichia coli's transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. PLoS Comput. Biol. *5*, e1000312.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Tucker, D.L., Tucker, N., Ma, Z., Foster, J.W., Miranda, R.L., Cohen, P.S., and Conway, T. (2003). Genes of the GadX-GadW regulon in Escherichia coli. J. Bacteriol. *185*, 3190–3201.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007). Structural basis for substrate loading in bacterial RNA polymerase. Nature *448*, 163–168.

Venturelli, O.S., Zuleta, I., Murray, R.M., and El-Samad, H. (2015). Population diversification in a yeast metabolic program promotes anticipation of environmental shifts. PLoS Biol. *13*, e1002042–e1002042.

Wagner, G.P., Pavlicev, M., and Cheverud, J.M. (2007). The road to modularity. Nat. Rev. Genet. *8*, 921–931.

Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., and Church, G.M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. Nature *460*, 894–898.

Wang, J., Atolia, E., Hua, B., Savir, Y., Escalante-Chong, R., and Springer, M. (2015). Natural variation in preparation for nutrient depletion reveals a cost-benefit tradeoff. PLoS Biol. *13*, e1002042.

Weinzierl, R.O. (2010). The nucleotide addition cycle of RNA polymerase is controlled by two molecular hinges in the Bridge Helix domain. BMC Biol. *8*, 134.

Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. *8*, 206–216.

Xia, J., Sinelnikov, I.V., Han, B., and Wishart, D.S. (2015). MetaboAnalyst 3.0–making metabolomics more meaningful. Nucleic Acids Res. *43*, W251–W257.