Adaptive Mutations in RNA Polymerase and the Transcriptional Terminator Rho Have Similar Effects on *Escherichia coli* Gene Expression

Andrea González-González,[1,2] Shaun M. Hug,[1] Alejandra Rodríguez-Verdugo,[3,4] Jagdish Suresh Patel,[5] and Brandon S. Gaut*,[1]

[1]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA
[2]Department of Biological Sciences, University of Idaho, Moscow, ID
[3]Department of Environmental Systems Sciences, ETH Zürich, Zürich, Switzerland
[4]Department of Environmental Microbiology, Eawag, Dübendorf, Switzerland
[5]Center for Modeling Complex Interactions, University of Idaho, Moscow, ID

*Corresponding author:** E-mail: bgaut@uci.edu.
Associate editor: Csaba Pal

Abstract

Modifications to transcriptional regulators play a major role in adaptation. Here, we compared the effects of multiple beneficial mutations within and between *Escherichia coli rpoB*, the gene encoding the RNA polymerase β subunit, and *rho*, which encodes a transcriptional terminator. These two genes have harbored adaptive mutations in numerous *E. coli* evolution experiments but particularly in our previous large-scale thermal stress experiment, where the two genes characterized alternative adaptive pathways. To compare the effects of beneficial mutations, we engineered four advantageous mutations into each of the two genes and measured their effects on fitness, growth, gene expression and transcriptional termination at 42.2 °C. Among the eight mutations, two *rho* mutations had no detectable effect on relative fitness, suggesting they were beneficial only in the context of epistatic interactions. The remaining six mutations had an average relative fitness benefit of ~20%. The *rpoB* mutations affected the expression of ~1,700 genes; *rho* mutations affected the expression of fewer genes but most (83%) were a subset of those altered by *rpoB* mutants. Across the eight mutants, relative fitness correlated with the degree to which a mutation restored gene expression back to the unstressed, 37.0 °C state. The beneficial mutations in the two genes did not have identical effects on fitness, growth or gene expression, but they caused parallel phenotypic effects on gene expression and genome-wide transcriptional termination.

Key words: transcriptional termination, protein structure, relative fitness, experimental evolution, adaptive pathways, phenotypic convergence.

Introduction

In 1975, King and Wilson famously speculated that evolutionary modifications in phenotype ". . . are more often based on changes in the mechanisms controlling the expression of genes than on sequence changes in proteins" (King and Wilson 1975). The relative frequency of changes in regulatory versus structural genes is still debated (Hoekstra and Coyne 2007; Carroll 2008; Wittkopp and Kalay 2011), but it has nonetheless become clear that regulatory genes do often play a crucial role in adaptation. For example, evolutionary modifications of transcriptional regulators have modified plant phenotypes during domestication (Doebley et al. 2006; Olsen and Wendel 2013), contributed to the shift from marine to freshwater environments for sticklebacks (Peichel and Marques 2017), and facilitated the adaptation of microbes during laboratory evolution experiments (Conrad et al. 2011; Long et al. 2015). Microbial evolution experiments have further shown that mutations in global transcriptional regulators are often highly beneficial, occur early in the course of adaptation, and appear under different selective regimes (Long et al. 2015). The last observation is especially true in *E. coli* evolution, because experiments over a wide range of strains and growth conditions have consistently identified adaptive genetic changes within the genes that encode the RNA polymerase (RNAP) complex (e.g., Herring et al. 2006; Conrad et al. 2009, 2010; Charusanti et al. 2010; Tenaillon et al. 2012; Sandberg et al. 2014; Deatherage et al. 2017) and the Rho transcriptional terminator (e.g., Freddolino et al. 2012; Tenaillon et al. 2012; Herron and Doebeli 2013; Le Gac et al. 2013; Haft et al. 2014; Lee and Helmann 2014; Dillon et al. 2016; Deatherage et al. 2017).

These observations raise at least three interesting topics about adaptive changes within *E. coli* transcriptional regulators. The first, which we do not address here, is why modifications of global regulators like RNAP and Rho are favored in

Open Access

Article

evolution experiments, despite the fact that RNAP function and sequence is often highly conserved in nature (Long et al. 2015). The second concerns potential pleiotropic effects, because some transcriptional regulators can, in theory, affect the expression of every gene in the genome. Indeed, single adaptive mutations in the *rpoB* gene, which encodes the RNA Polymerase β subunit, alter the expression of hundreds to thousands of downstream genes (Sandberg et al. 2014; Rodríguez-Verdugo et al. 2016; Utrilla et al. 2016). These findings raise the possibility that some of the shifts in gene expression are maladaptive. Consistent with this view, at least two studies have shown that fitness gains after the fixation of a beneficial *rpoB* mutation act to ameliorate negative pleiotropic effects (Sandberg et al. 2014; Rodríguez-Verdugo et al. 2016).

A third topic is about the relative effects of different adaptive mutations, both within the same transcriptional regulator (e.g., different mutations within *rpoB*) and between transcriptional regulators (e.g., *rpoB* vs. *rho*). These are potentially interesting contrasts because it has been shown that different mutations can have different effects on fitness, pleiotropy, and evolutionary contingency (Rosenblum et al. 2014), even for different mutations within the same gene (Linnen et al. 2009). It seems likely that there are large-scale differences in the effects of mutations to RNAP and Rho. A modification of RNAP has the potential to affect the expression of every gene, but Rho terminates transcription for a subset of ~20–50% of *E. coli* genes (Cardinale et al. 2008; Peters et al. 2009). This begs the questions: Do adaptive mutations in *rho* affect the expression of fewer genes? If so, are mutations in *rho* generally more beneficial, because they cause fewer negatively pleiotropic effects? And, to what extent are these effects consistent across different beneficial mutations within the same transcriptional regulator?

To our knowledge, the effects of adaptive mutations have not yet been compared explicitly between different transcriptional regulators. This may be due, in part, to the fact that a valid comparison requires that mutations in the two regulatory genes are adaptive in the same genetic background under identical culture conditions. Fortunately, our previous thermal adaptation experiment has yielded appropriate mutations. Our experiment evolved 115 populations of *E. coli* independently from the same ancestral clone (REL1206) for 2,000 generations at 42.2 °C (Tenaillon et al. 2012). At the end of the experiment, we isolated one clone from each population and found that fitness had improved ~40%, on average, relative to REL1206. We also sequenced the evolved clones and identified over 1,300 mutations. Using parallelism as a *prima facie* argument for adaptation (Wichman et al. 1999; Fong et al. 2005; Stern 2013), about half of the observed 1,300 mutations were adaptive, either because the same mutation occurred in more than one clone or because mutations occurred in the same gene in different clones.

Two of the most mutated genes in our thermal stress experiment were *rpoB*, which accumulated a total of 87 mutations in 76 of the 115 clones, and *rho*, which accumulated 24 different nonsynonymous mutations across 43 of 115 clones. Mutations in the two genes were found together less

often than expected at random, suggesting negative epistatic interactions (Tenaillon et al. 2012). More generally, the two genes define alternative pathways to thermal stress adaptation. The first includes mutations in *rpoB*, as well as mutations in the other RNA polymerase subunits (*rpoA*, *rpoC* and *rpoD*) and six *rod* genes that affect cell shape. In contrast, the *rho* pathway includes mutations of the cardiolipin synthase (*cls*) gene and the transcription factor gene *iclR*. We have shown that these two adaptive pathways differ in some aspects of their phenotypes, such as chemical sensitivities (Hug and Gaut 2015) and trade-offs for growth at low temperature (Rodríguez-Verdugo et al. 2014).

Before diving into contrasts between the effects of *rho* and *rpoB* mutations, it is useful to briefly consider the structure and function of RNAP and Rho. RNAP consists of a five-subunit core protein that has the catalytic activity for RNA synthesis (Ebright 2000; Ishihama 2000). The β subunit contributes to RNA synthesis activity and contains regions for nonsequence-specific interactions with DNA and nascent RNA. Some of the observed beneficial *rpoB* mutations have occurred in the contact region between the RNAP proximal active site and the downstream DNA duplex (Rodriguez-Verdugo et al. 2013). Mutations in the contact region are particularly likely to play key roles in modulating RNAP activity in the three stages of transcription: initiation, elongation, and termination (Ederth et al. 2006).

The Rho molecule acts to terminate RNAP transcription for a subset of *E. coli* genes. The protein is a hexameric ring composed of six monomers (Skordalakes and Berger 2003) (fig. 1). Given a mature hexamer, transcriptional termination proceeds, first, by the protein recognizing a *rho* utilization (*rut*) site on an elongating mRNA; second, by binding *rut* sites and translocating mRNA through the hexamer's central cavity (Skordalakes and Berger 2003); until, third, it contacts RNAP while it pauses, thereby disassociating the elongation complex (Skordalakes and Berger 2003; Peters et al. 2011). The *rut* sites remain difficult to predict bioinformatically, and so the genomic targets of Rho are characterized incompletely (Ciampi 2006; Hollands et al. 2014). However, a few hundred intergenic and intragenic Rho termination sites have been identified experimentally (Peters et al. 2009, 2012), and these have helped to evaluate whether *rho* mutants increase (Mori et al. 1989; Miwa et al. 1995) or decrease the efficiency of termination (Freddolino et al. 2012; Haft et al. 2014).

Here, we take advantage of a unique opportunity to compare and contrast the effect of adaptive mutations between the *rho* and *rpoB* genes that characterize alternative adaptive pathways. In the study, we engineer four beneficial *rho* mutations into the REL1206 ancestor and contrast effects on fitness, growth, gene expression and transcriptional termination to four previously studied *rpoB* mutants (Rodríguez-Verdugo et al. 2014, 2016). In doing so, we address three sets of questions. First, are fitness and growth characteristics similar between *rho* and *rpoB* mutations? Second, do the beneficial mutations influence the expression of similar numbers and sets of genes? Third, do *rho* mutations alter its termination efficiency and, if so, how do termination properties compare between *rho* and *rpoB* mutations? Finally, do the mutations
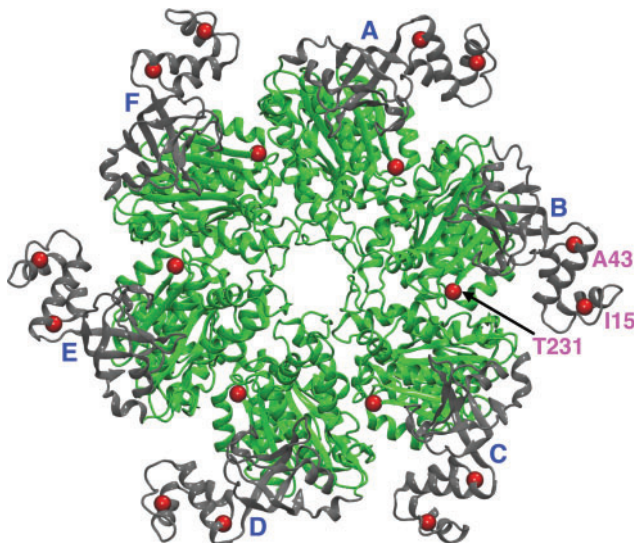
Fig. 1. A schematic of the Rho protein, showing each of its six hexamers (A–F) and the location of the single-mutations (red spheres) produced for this experiment. Mutations in *I15* and *A43* are located in the N-terminal domain (NTD) (gray); the mutation in *A231* in the C-terminal domain (CTD) (green).

have a consistent predicted effect on protein stability, which may be related to fitness (e.g., DePristo et al. 2005)?

Results

Rho Mutations and Relative Fitnesses

We engineered four *rho* mutations—*I15N, I15F, A43T,* and *T231A*—into the ancestral REL1206 background (see Materials and Methods). Each of these substitutions were adaptive by the criterion of parallelism, because they were observed in 15, 2, 3, and 2 independent populations, respectively, of our 42.2 °C thermal stress experiment (Tenaillon et al. 2012). *I15N, I15F,* and *A43T* point mutations are located in the NHB domain of the Rho monomer, whereas *T231A* is located in the central part of the C-terminal domain (CTD), specifically in the Loop-CTD (fig. 1).

We assessed the relative fitness ($\bar{w}_r$) of the four *rho* mutants to REL1206 at 42.2 °C (supplementary table S1, Supplementary Material online). Relative fitness data conveyed three observations (table 1). First, two of the four mutations (*A43T* and *T231A*) yielded a $\bar{w}_r$ value significantly > 1.0, indicating an adaptive benefit under thermal stress. In contrast, neither the *I15N* nor the *I15F* mutation produced a $\bar{w}_r$ value significantly different from 1.0, and *I15N* yielded an estimate < 1.0 (Rodríguez-Verdugo et al. 2014) suggesting a slightly deleterious mutation (table 1). Second, the $\bar{w}_r$ estimates differed significantly among the four mutations (ANOVA, $P = 2.4 \times 10^{-8}$), with *I15N* and *I15F* forming one group and *A43T* and *T231A* forming a second, significantly different group (Tukey-test; $P < 0.05$).

We contrasted $\bar{w}_r$ among four *rpoB* mutants (*I572L, I572F, I572N,* and *I966S*) (Rodríguez-Verdugo et al. 2014) and the four *rho* mutants. Each of the *rpoB* mutations had $\bar{w}_r > 1.0$, with estimates ranging from $\bar{w}_r = 1.16$ to 1.26 at 42.2 °C (table 1). Across all mutants, comparisons again indicated

heterogeneity of $\bar{w}_r$ (ANOVA, $P = 1.91 \times 10^{-08}$; table 1), and Tukey tests grouped *I15N* and *I15F* against the remaining six mutations ($P < 0.05$ for 10 of 12 comparisons between *I15* mutations and the remaining mutations). Altogether, $\bar{w}_r$ values were > 1.00 and indistinguishable among the four *rpoB* mutations and two of the four *rho* mutations by Tukey tests; the two remaining *rho* mutations had demonstrably lower $\bar{w}_r$ values and little effect on fitness.

Growth Dynamics

To gain additional insights about potential differences between *rho* and *rpoB* mutants, we measured growth curves of mutant lines at 37.0 and 42.2°C (fig. 2A and B, table 2, and supplementary table S2, Supplementary Material online) and estimated maximal growth rates and final yields at stationary phase. At 37.0 °C, all *rho* mutants had significantly higher final yields than at 42.2 °C (fig. 2A and B; supplementary table S2, Supplementary Material online), indicating that high temperature remains stressful even after the introduction of putatively beneficial mutations. At 42.2 °C, the *rho* mutants had a lower final yield than the ancestor, and there were also differences in maximum growth rate among mutants: *I15F* and *I15N* showed somewhat lower maximum growth rate than the ancestor but *A43T* and *T231A* had higher growth rates (fig. 2B and supplementary table S2, Supplementary Material online), consistent with their higher $\bar{w}_r$ values.

We then contrasted growth dynamics between *rho* and the three *I572 rpoB* mutants (Rodríguez-Verdugo et al. 2016). Similar to results with *rho*, *rpoB* mutants grew more slowly and had lower yields at 42.2 °C compared with 37.0 °C (fig. 2A and B and supplementary table S2, Supplementary Material online). At 42.2 °C, the *I572 rpoB* mutants reached significantly higher final yields than the *rho* mutants (table 2 and supplementary table S2, Supplementary Material online). The only significant difference in growth rates was the contrast between *rpoB I572F* versus *rho I15F* (table 2). Overall, growth curves indicate that: 1) *rho* and *rpoB* mutations differ with respect to yield, and 2) growth rates tend to be lower for *rho I15F* compared with the remaining clones.

Gene Expression Phenotypes

To gain further insight into the phenotypic effects of *rho* mutations, we measured gene expression at 42.2 °C using replicated RNAseq data (see Materials and Methods). We hypothesized that expression covaries with $\bar{w}_r$, so that *I15N* and *I15F* differ more dramatically in expression from *A43T* and *T231A* than they do from each other. To test this hypothesis, we measured the correlation in expression across all genes for each pair of *rho* mutants (table 3). All pairwise correlations were high ($r^2 > 0.95$), but *A43T* and *T231A* had the highest correlation ($r^2 = 0.992$), whereas contrasts between an *I15* mutant with either *A43T* or *T231A* had lower correlations (table 3). We also identified differentially expressed genes (DEGs) between *rho* mutants, based on a significance cut-off of $q < 0.001$ (see Materials and Methods). Only one gene was expressed differentially between *I15F* and *I15N*, but both *I15* mutations had > 100 DEGs between *A43T*

Table 1. Relative Fitness Measured at 42.2 °C for Four *rho* Mutants and Four *rpoB* Mutants, Along with Their Estimated ΔΔG Values, as Predicted by Molecular Modeling.

|	Relative Fitness			Estimated ΔΔG Values	
Mutant[a]	Mean	± 95% CI	P-Value[b]	Location[c]	ΔΔG at 42.2 °C(kcal/mol) ± SE[d]
rho I15F	1.008	0.117	0.837	α2 NHB*	1.22 ± 0.25
rho I15N	0.945	0.184	0.396	α2 NHB*	0.50 ± 0.35
rho A43T	1.237	0.093	<0.05	α3 NHB*	1.17 ± 0.39
rho T231A	1.266	0.104	<0.05	Loop - CTD**	−0.06 ± 0.45
rpoB I572F	1.155	0.047	<0.05	Rif[+]	−0.327 ± 0.59
rpoB I572L	1.229	0.099	<0.05	Rif[+]	−0.421 ± 0.26
rpoB I572N	1.184	0.067	<0.05	Rif[+]	0.803 ± 0.52
rpoB I966S	1.258	0.106	<0.05	Eco flap - βi9[++]	1.405 ± 0.34

[a]*rho* data from this study, based on nine replicated measures. *rpoB* data include that from (Rodríguez-Verdugo et al. 2014) plus three additional replicates, for a total of six replicates.
[b]$P < 0.05$ rejects the null hypothesis that $\bar{w}_r = 1.0$, indicating that the mutation is beneficial ($\bar{w}_r > 1.0$) relative to the REL1206 ancestor.
[c]Location of mutation in the Rho protein: NH B = N-terminal helix bundle *Mutations located in the NTD= N-terminal and **CTD= C-terminal domain. Location of mutations in the RNAP protein: [+]RNA polymerase, β subunit, Rif= Rif-binding regions in the fork domain, [++]RNA polymerase, β subunit, Eco flap domain, β subunit insert 9.
[d]Estimated changes in the free energy of folding are calculated with 95% confidence interval. ΔΔG value for *rho* and *rpoB* wild-type is zero.
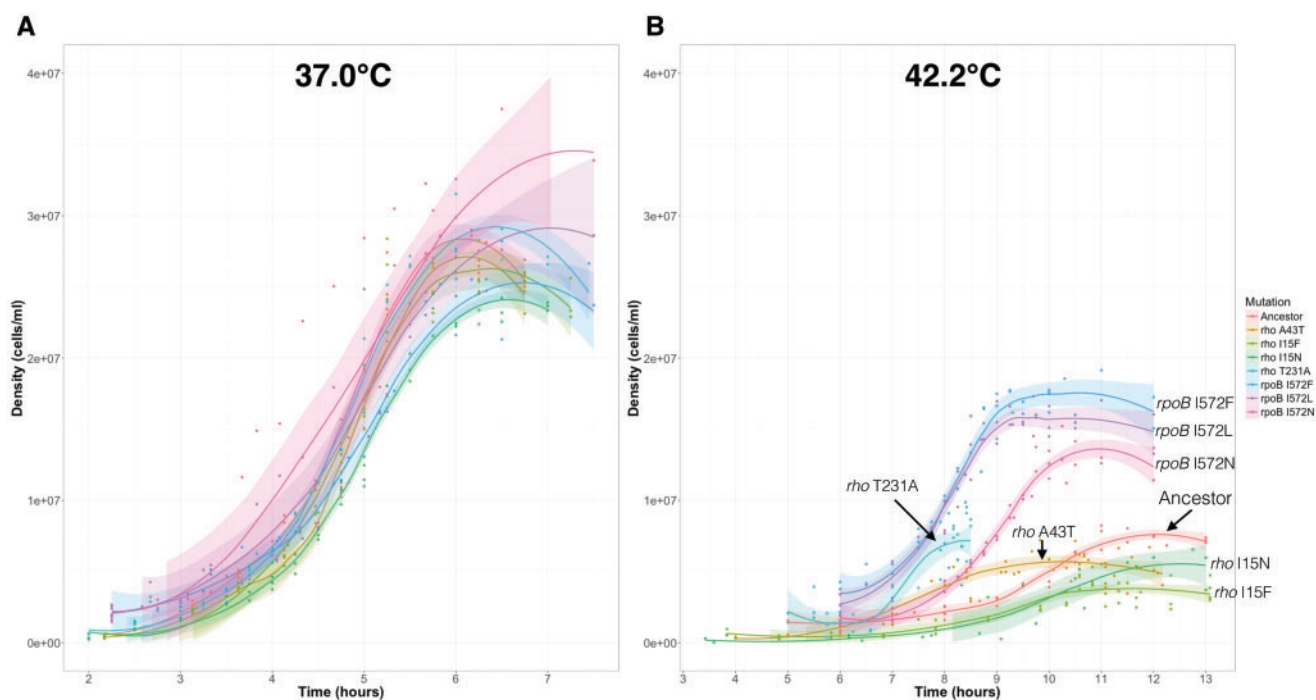
FIG. 2. Growth curves for the mutants and ancestor at 37.0 °C (*A*) and 42.2 °C (*B*). Each line corresponds to a local polynomial regression fitting of three replicates with its 95% confidence interval (see table 2 and supplementary table S2, Supplementary Material online).

and *T231A* (table 3). These data suggest that *A43T* and *T23A* mutations have quantitatively different effects on gene expression relative to the two *I15* mutations, mimicking observed differences in fitness among clones.

Given that single *rho* mutations can confer extensive shifts in expression relative to an ancestral background (Freddolino et al. 2012; Haft et al. 2014), we compared gene expression between the REL1206 ancestor at 42.2 °C and each of the four *rho* mutants. Across the four *rho* mutants, we detected a total of 1,140 DEGs compared with the 42.2 °C ancestor, but there were more DEGs for *T231A* (1,028 genes) and *A43T* (656) than for *I15F* (372) and *I15N* (195) (fig. 3A). Among the complete set of 1,140 DEGs, 141 (or 12%) were shared among all four *rho* mutants (fig. 3A). Gene Ontology (GO) analysis of

these 141 genes revealed significant enrichment for down-regulation of maltose transport genes (GO: 0015768, $P = 1.04 \times 10^{-3}$) and significant up regulation of genes involved in transcription (GO: 0010467, $P = 1.12 \times 10^{-12}$) and translation (GO: 0006412, $P = 9.17 \times 10^{-37}$) (supplementary table S3, Supplementary Material online). Interestingly, the *A43T* and *T231A* mutants shared 330 DEGs that were not identified in either *I15F* or *I15N*. These 330 genes are likely candidates to contribute to $\bar{w}_r$ differences among *rho* mutants, and they were enriched for up-regulation of purine nucleotide biosynthetic processes (GO: 0006164, $P = 1.15 \times 10^{-12}$) and down-regulation for glycerol catabolic processes (GO: 0019563, $P = 1.15 \times 10^{-12}$) (supplementary table S4, Supplementary Material online).

Table 2. P Values for Statistical Comparisons of Maximum Growth Rate and Yield of *rho* and *rpoB* Mutants at 42.2 °C.

rpoB	I15N[a]	I15F	A43T	T231A
Maximum growth rate				
I572N	0.138	0.471	0.440	0.127
I572F	0.177	*0.045*[b]	0.732	0.157
I572L	0.203	0.092	0.995	0.175
Final yield[c]				
I572N	*0.001*	*8×10^{-4}*	*0.002*	*0.010*
I572F	*3×10^{-4}*	*4×10^{-4}*	*0.001*	*0.004*
I572L	*7×10^{-5}*	*4×10^{-5}*	*1×10^{-4}*	*0.002*

The header row also shows *rho* label spanning the mutant columns.

[a]Entries correspond to significant values from two-sample *t*-tests. The null hypothesis is that the mean values of parameters are equal. Numbers in italics correspond to P values <0.05.

[b]*rpoB* I572F had a significantly higher maximum growth rate compared with the *rho* mutant I15F at 42.2 °C (see fig. 2B and supplementary table S2, Supplementary Material online).

[c]In every case, *rpoB* mutants had a significantly higher final yield compared with the *rho* mutants at 42.2 °C (fig. 2B and supplementary table S2, Supplementary Material online).

Table 3. Comparisons of Gene Expression between Single *rho* Mutants.

	I15N	I15F	A43T	T231A
I15N		0.991	0.962	0.955
I15F	1		0.969	0.964
A43T	149	190		0.992
T231A	419	552	16	

NOTE.—The cells above the diagonal report the Pearson correlation coefficient in gene expression across all genic regions ($n = 4,204$). The cells below the diagonal indicate the number of genes significantly different via DEseq ($q < 0.001$).

Finally, to determine whether putatively beneficial mutations in *rpoB* or *rho* have similar effects on gene expression, we identified DEGs in four *rpoB* mutants—I572N, I572F, I572L, and I966S—using previously published data (Rodríguez-Verdugo et al. 2016). We took the union of the DEGs for the four *rpoB* mutations and found 1,726 DEGs relative to REL1206 at 42.2 °C. The *rpoB* and *rho* DEG sets were remarkably overlapping (fig. 3B): of the 1,726 DEGs identified with *rpoB* mutations, 957 genes (or 55.4%) were also identified as DEGs in *rho* mutants. This similarity is more impressive when one considers that *rho* mutations altered gene expression for significantly fewer genes than *rpoB* mutations (1,140 vs. 1,726; binomial test: $P < 2.2 \times 10^{-16}$) so that 83% (=957/1,140) of DEGs caused by *rho* mutations were also identified as DEGs in *rpoB* mutants (fig. 3B). These results indicate that *rho* mutants generally affect the expression of a subset of the genes affected by *rpoB* mutants.

The DEG overlap between *rho* and *rpoB* mutations left only 183 DEGs unique to the set of *rho* mutations (fig. 3B). Among these 183 DEGs, GO analyses identified up-regulated responses to temperature stimuli (e.g., *clpB*, *dnaK*, *dnaJ*, *degP*, *grpE*; GO: 0009408, $P = 3.35 \times 10^{-2}$) (supplementary data set S1, Supplementary Material online). In contrast, the 769 genes uniquely modified by *rpoB* mutants were enriched for up-regulation of iron ion transport (GO: 0006826, $P = 3.2 \times 10^{-05}$) and flagellum cell motility (GO: 00048870, $P = 7.1 \times 10^{-06}$) genes. Monosaccharide catabolic processes

(GO: 0005996, $P = 2.0 \times 10^{-04}$) and cell wall biogenesis (GO: 0042546, $P = 2.94 \times 10^{-4}$) were also unique to *rpoB* mutants, but they were not consistently up- or down-regulated (supplementary data set S1, Supplementary Material online).

Shifts in Gene Expression Are Partially Restorative

Taken in total, *rpoB* and *rho* mutants affect an overlapping gene set, but it is not a foregone conclusion that they affect expression in the same direction. Here direction refers to expression shifts in mutants relative to the stressed ancestral state (i.e., REL1206 at 42.2 °C) and the unstressed ancestral state (i.e., REL1206 at 37.0 °C) (Carroll and Marx 2013). We determined directionality in two ways. First, we applied PCA to all expression data for *rho*, *rpoB*, and ancestral replicates, and then plotted the first two principal components, PC1 and PC2. Based on the PCA, the *rpoB* mutants were intermediate on PC1 between the ancestor at 37.0 and 42.2 °C (fig. 3C). This position is consistent with an overall tendency of *rpoB* mutants to restore expression patterns from the stressed (42.2 °C) toward the unstressed (37.0 °C) ancestral state (Rodríguez-Verdugo et al. 2016). Expression patterns from *rho* mutants were also intermediate between the two ancestral states on PC1, but they were not restored as fully as the *rpoB* mutants (fig. 3C). Intriguingly, replicates from the two less fit *rho* mutants (I15N and I15F) clustered more closely to REL1206 at 42.2 °C. These clustering patterns suggest that fitness is related to restoration in PCA space; indeed, $\bar{w}_r$ and the average PC1 eigenvector were correlated across the eight mutants ($r^2 = 0.4156$; $P = 0.042$).

Second, we categorized expression of single genes into categories. Following previous studies (Carroll and Marx 2013), genes in mutants were classified as: 1) "restored", if expression in the mutant shifts toward the unstressed ancestral state from the stressed ancestral state; 2) "unrestored", if expression remains similar to the REL1206 ancestor in its stressed state; 3) "reinforced" if expression in the mutant is exaggerated in the mutant relative to REL1206 at both 37.0 and 42.2 °C and, finally, 4) "novel" if the two ancestral treatments did not differ in GE but the mutant differed from both. (The Materials and Methods section includes more precise quantitative definitions of these four directional categories.) The majority of genes fell into the restored category for five of the eight mutants, and >45% of the genes were in the restored category for the *rho* I15N and I15F mutants (table 4). More importantly, the number of restored genes was correlated with $\bar{w}_r$ across the eight mutants (*rho* and *rpoB* mutants) at a borderline level of significance (Spearman's $r = 0.67$; $P = 0.055$). Altogether, directional analyses show that $\bar{w}_r$ of a mutant is a function of the degree of restoration of gene expression.

Assessing Potential Effects on Transcriptional Termination

We have shown that *rho* mutations differ in $\bar{w}_r$ and that $\bar{w}_r$ correlates with shifts in the magnitude and direction of gene expression. However, the question of mechanism remains—that is, how do mutations in *rho* affect its function as a
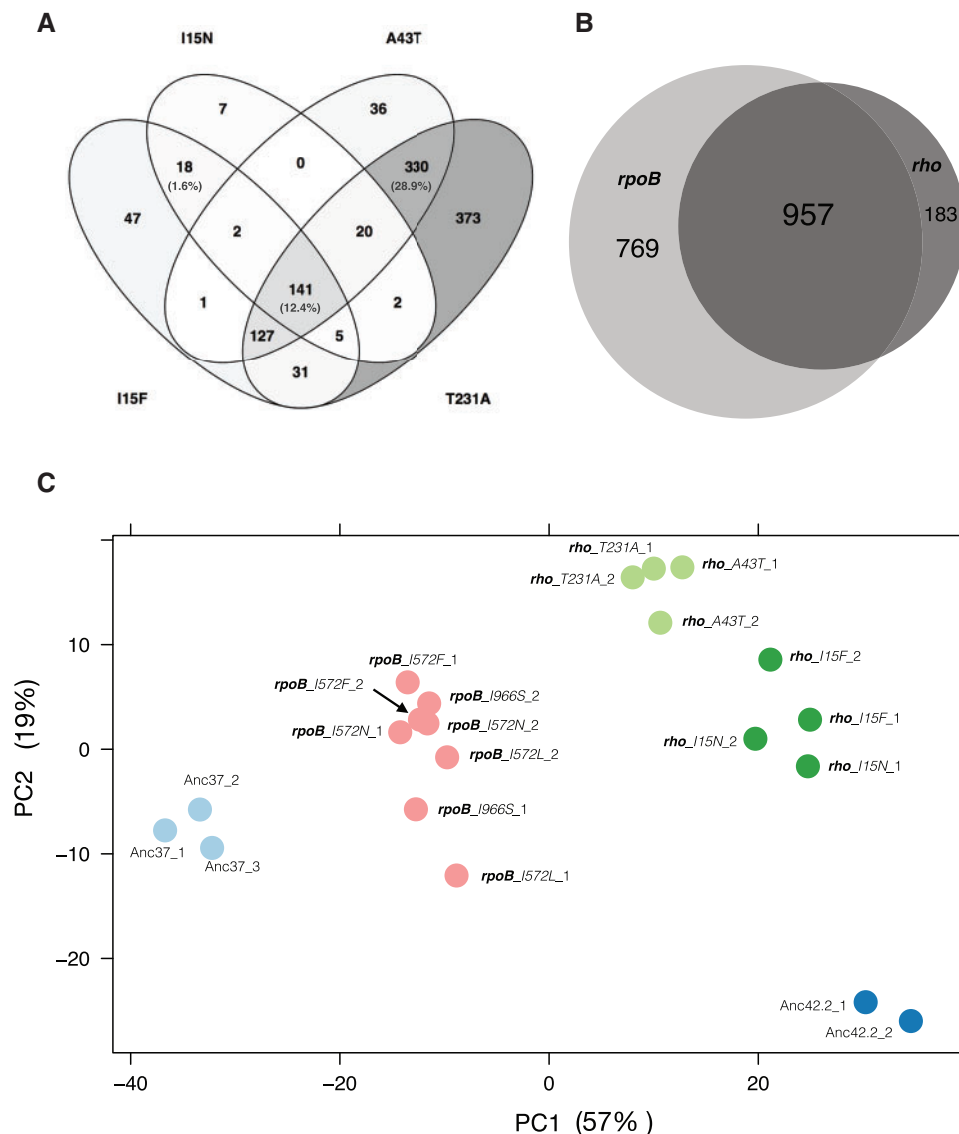
FIG. 3. (*A*) A Venn diagram of the number of genes that exhibit significant DEG between *rho* mutants. (*B*) A Venn diagram of the number of genes that differ in expression between the four *rho* mutants and REL1206 at 42.2 °C and all four *rpoB* mutants and REL1206 at 42.2 °C. (*C*) A PCA plot of the RNAseq data, including replicated data from REL1206 at 37.0 °C (light blue dots), REL1206 at 42.2 °C (dark blue dots); replicated data from each of the four *rho* mutations (light and dark green dots); and replicated data from each of the four *rpoB* mutations (pink dots). The *x*-axis represents restoration along PC1, where smaller values reflect greater restorative effects.

Table 4. The Direction of Gene Expression for Genes within *rho* and *rpoB* Mutants.

Mutant[a]	Restored[b]	Unrestored	Reinforced	Novel	Total[c]
rho I15F	762	886	19	4	1,671
rho I15N	802	857	8	16	1,683
rho A43T	1,115	544	7	32	1,698
rho T231A	1,250	405	12	46	1,713
rpoB I572F	1,092	644	1	59	1,796
rpoB I572L	563	1,174	0	6	1,743
rpoB I572N	1,165	569	3	53	1,790
rpoB I966S	1,308	428	1	39	1,776

[a]*rho* data from this study. *rpoB* data from (Rodríguez-Verdugo et al. 2014).
[b]The number of genes was counted as restored, unrestored, reinforced and novel using the rules elaborated in the Materials and Methods.
[c]The total tallies the number of genes that differ at *q* < 0.001 from either REL1206 at 37.0 °C or REL1206 at 42.2 °C or both. The total differs significantly between *rho* and *rpoB* mutations (*t*-test; *P* = 0.0016).

transcriptional terminator? Previous studies have shown that *rho* mutations tend to decrease termination efficiency, and therefore, increase read-through activity (Freddolino et al. 2012; Haft et al. 2014). Alternatively, we have hypothesized that adaptive *rho* mutations increase transcriptional efficiency at high temperature by enhancing termination (Rodríguez-Verdugo et al. 2014, 2016). We addressed this hypothesis both directly, by examining regions known to be terminated by Rho, and indirectly, by assessing global gene expression patterns.

Direct Effects on Rho-Terminated Regions
The direct approach relied on a set of 178 Rho-terminated regions that were defined by Peters et al. (2009) and previously used to analyze read-through activity of a *rho* mutation

(Haft et al. 2014) (supplementary data set S2, Supplementary Material online). The advantage of studying these regions is that they can, in theory, reflect direct (primary) effects of *rho* mutations, as opposed to indirect (secondary) effects on genes that then affect additional downstream genes. The size of the *rho*-terminated regions spanned from 100 to $\approx$ 4,000 bp, with roughly half located at the 3′ end of coding regions (intergenic) and the other half within genes (intragenic). Following Haft et al. (2014), we calculated γ, the $\log_2$ ratio of normalized RNAseq counts between each mutant at 42.2 °C and REL1206 at 42.2 °C for each of the 178 regions (see Materials and Methods). Under the null hypothesis that the mutant and ancestor have similar termination properties, γ is binomially distributed. If a mutation enhances termination, we expected significantly >50% of the regions to have $\gamma < 0.0$. In contrast, increased read-through results in >50% of regions having $\gamma > 0.0$ (Haft et al. 2014). Contrary to our prediction, both I15N and I15F have increased read-through activity according to this test, with 108 and 119 of 178 regions having $\gamma > 0.0$ (binomial, $P < 0.01$). The two remaining *rho* mutations (A43T, T231A) yielded no evidence of altered termination properties ($P = 1.00$; supplementary fig. S1, Supplementary Material online).

We further examined direct effects on the 178 Rho-terminated regions by taking the ratio of expression counts for each *rho*-terminated region (RT) against the counts from the immediate 5′ upstream region (UP) of the same size. If *rho* mutations enhance termination at these sites, then RT/UP values across sites should be lower for *rho* mutants than for the Ancestor at 42.2 °C. To test this hypothesis, we computed a measure of termination efficiency for each mutant as the ratio of averages, r_{avg}, of read counts in RT versus UP across all 178 regions (see Materials and Methods). While not without caveats (such as the fact that upstream and downstream regions are in different operons in some cases), this approach allowed us to measure the strength of termination for each mutant and for each ancestral treatment.

Estimates of r_{avg}, based on bootstrapped samples across sites (see Materials and Methods), are presented in figure 4. The estimates suggest weaker termination transcription (i.e., enhanced read-through) for REL1206 at 42.2 °C compared with 37.0 °C; although the trend is notable, the two ancestral treatments do not differ significantly (bootstrapping test of homogeneity; $P = 0.864$). For the *rho* mutants, I15N had a notably higher mean estimate of r_{avg}(=1.29) than the other mutants, but not higher than REL1206 at 42.2 °C (r_{avg}=1.34; fig. 4). The r_{avg} estimates for T231A and A43T (and I15F, to a lesser extent) suggest that the mutations restored termination to the level of the ancestor at 37.0 °C (fig. 4). As a group the four *rho* mutants did not differ in r_{avg} relative to REL1206 at 37.0 °C (bootstrapping test of homogeneity; $P = 0.720$) or at 42.2 °C ($P = 0.077$). The latter was, however, borderline significant.

It has been shown previously that some *rpoB* mutants enhance transcriptional termination (Jin et al. 1988). Hence, for completeness we also applied analysis of Rho-terminated regions to *rpoB* mutants. Based on γ analyses, none of the *rpoB* mutants yielded evidence of altered termination in

these regions ($P > 0.500$; supplementary fig. S1, Supplementary Material online). However, r_{avg} estimates for *rpoB* were consistently lower than the ancestral treatments and the *rho* mutants (fig. 4). Furthermore, the average r_{avg} estimate between the set of *rho* and *rpoB* mutants differed significantly ($P < 0.001$), indicating either that termination in these regions is more efficient in *rpoB* mutants or that *rpoB* mutants had consistently higher expression in the upstream regions. The *rpoB* mutations also differed significantly from the ancestors at 42.2 °C and at 37.0 °C ($P < 0.003$; fig. 4).

To sum: γ analyses suggest that the *rho* I15N and I15F mutations decrease termination efficiency, with no evidence for improved termination efficiency. The second metric (r_{avg}) does suggest a direct effect on termination for three of the four mutants, but not *rho* I15N. The set of *rpoB* mutations, in contrast, have statistically better termination for these same regions based on r_{avg} analyses.

Indirect Effects

In addition to assessing the effect of mutations on regions known to be terminated by Rho, we also assessed potential effects on termination by focusing on the expression of all genes and intergenic regions (IRs). Focusing first on genes, we compared expression data between REL1206 at 37.0 and 42.2 °C. Similar to Rodríguez-Verdugo et al. (2016), we found that ~1,700 of 4,204 coding regions were differentially expressed ($q < 0.001$) during thermal stress, with a tendency toward lower expression at 42.2 °C (972 down-regulated genes vs. 695 up regulated genes; binomial $P = 1.243 \times 10^{-11}$) (fig. 5A). Given this baseline, we then assessed the effect of *rho* mutants relative to REL1206 at 42.2,°C. GE was significantly biased toward higher expression in each of the mutants relative to REL1206 at 42.2,°C at $q < 0.001$ (fig. 5A; binomial, $P = 1.41 \times 10^{-10}$ (I15N), $P = 2.13 \times 10^{-6}$ (I15F), $P = 8.6 \times 10^{-12}$ (A43T), $P = 2.2 \times 10^{-16}$ (T231A)). In effect, thermal stress led to lower GE at 42.2,°C relative to 37.0,°C, and the four mutations reversed that trend. These trends were quantitatively similar for *rpoB* mutants (fig. 5C).

The picture was, however, different for IRs. Among 2306 IRs (supplementary data set S3, Supplementary Material online) we found more were up-regulated in REL1206 at 42.2,°C relative to 37.0,°C (212 up regulated IR vs. 164 down; regulated; binomial $P = 0.015$) (fig. 5B). This up-tick in the expression of IRs may indicate more transcriptional read-through for REL1206 under thermal stress. This trend was reversed by both *rho* and *rpoB* mutations, such that differentially expressed IRs tended to have *lower* expression in the mutants relative to REL1206 at 42.2,°C [fig. 5B, binomial $P = 0.0016$ (I15N), $P = 0.0001$ (I15F), $P = 0.04164$ (A43T); fig 5D; $P = 0.0016$ (I15N), $P = 0.0001$ (I15F), $P = 0.04164$ (A43T)]. These results raise the possibility that all of the *rho* and *rpoB* mutations enhance termination efficiency at 42.2,°C either due to direct modifications on their protein products or due to secondary (or indirect and downstream) effect.
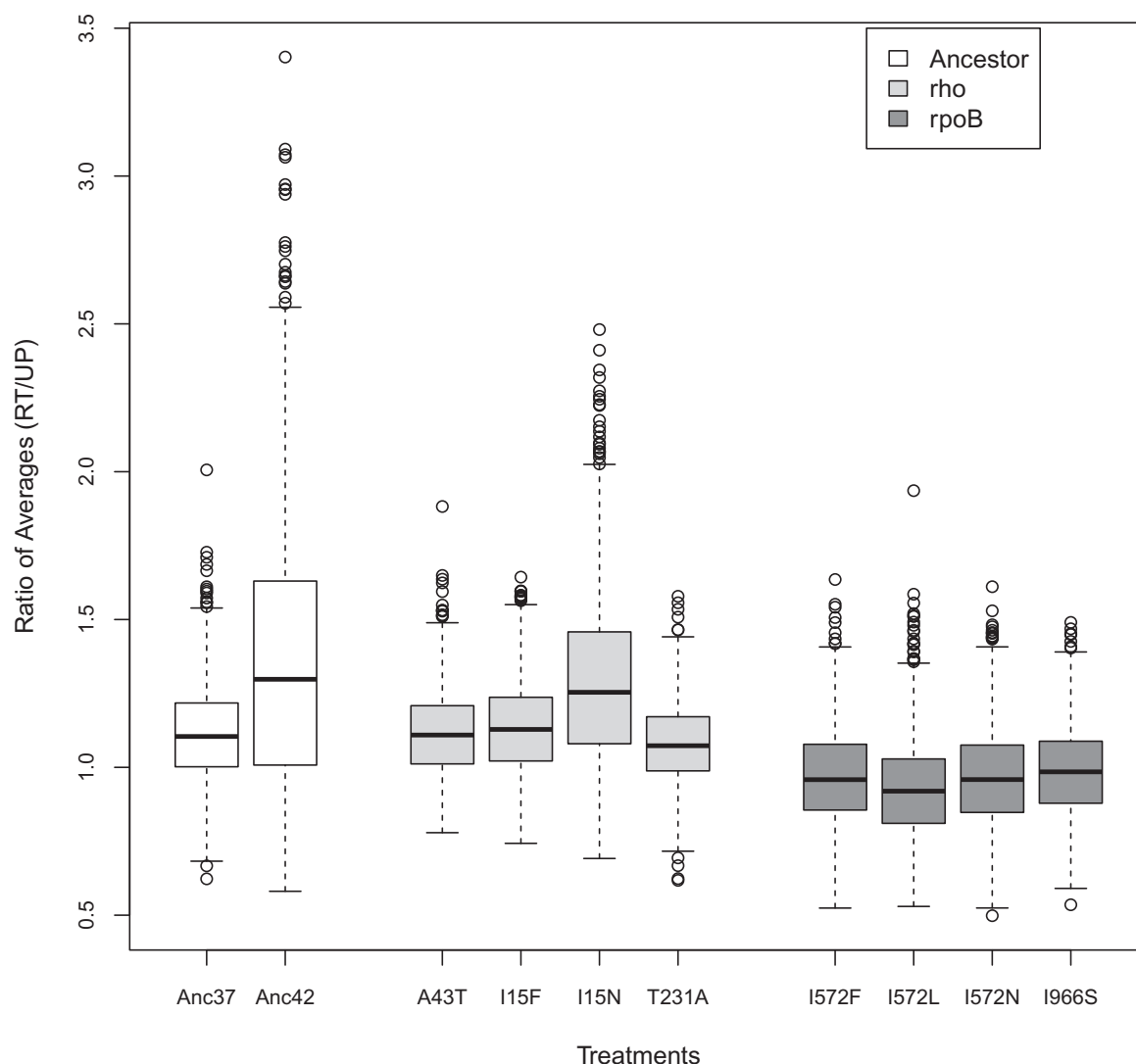
Fig. 4. Graph of bootstrap estimates of the ratio of averages ($\bar{r}_{avg}$) that measures the ratio of counts between *rho* terminated sites (RT) and their upstream 5′ regions (UP) across 178 loci. Lines in the boxplots represent the median bootstrap-based estimate of $\bar{r}_{avg}$ based on 1,000 resamplings (see Materials and Methods). The boxes represent the upper and lower quartiles of bootstrap replicates. Anc37 represents the REL1206 ancestor at 37.0 °C and Anc42 is the ancestor at 42.2 °C.

Mutations and Protein Stability

Finally, we evaluated the predicted folding stability of the Rho and RpoB proteins. Our principal goal was to assess whether changes in fitness correlated with the direction and magnitude of shifts in the free energy of protein folding stability, as is expected under some evolutionary models (DePristo et al. 2005; Wylie and Shakhnovich 2011). We hypothesized that $\bar{w}_r$ from specific mutations correlates with the difference of the free energy of folding between ancestral and mutant proteins ($\Delta\Delta G$) at 42.2 °C. To determine the free energy of folding (ΔG), we applied an approach that combines classical Molecular Dynamics (MD) with the FoldX algorithm (see Materials and Methods) and then measured the change in folding stability ($\Delta\Delta G$) induced by each of the four *rho* and four *rpoB* mutations. Positive values of $\Delta\Delta G$ suggest a decrease either in the stability of the protein fold or in the ability of the protein to fold relative to the ancestor; negative values indicate increased stability (Guerois et al. 2002;

Miller et al. 2016). Three of four *rho* mutations (I15F, A43T and I15N) were destabilizing at 42.2,°C, with $\Delta\Delta G$ values ranging between 0.50 and 1.22 kcal/mol (table 1); in contrast, the *T231A* substitution mildly increased folding stability at 42.2 °C ($\Delta\Delta G = -0.06$). Two of the four *rpoB* mutations were destabilizing at 42.2 °C (table 1). Although our sample size was small ($n = 4$) for each gene, there was no clear trend between $\bar{w}_r$ and $\Delta\Delta G$ (table 1). For completeness, we estimated $\Delta\Delta G$ for 16 additional *rho* mutations from our evolution experiment, for a total of 20 $\Delta\Delta G$ estimates (supplementary table S5, Supplementary Material online). Of these, 17 of 20 $\Delta\Delta G$ estimates were destabilizing, and there were no differences in $\Delta\Delta G$ for mutations in different structural features (e.g., α-helices vs. β-sheets, *t*-test, $P = 0.45$; NHB and CSD-like RBD, vs. nondefined regions, $P = 0.36$, see supplementary table S5, Supplementary Material online). The only consistent effect of mutations was toward destabilizing the wild type Rho protein at high temperature.
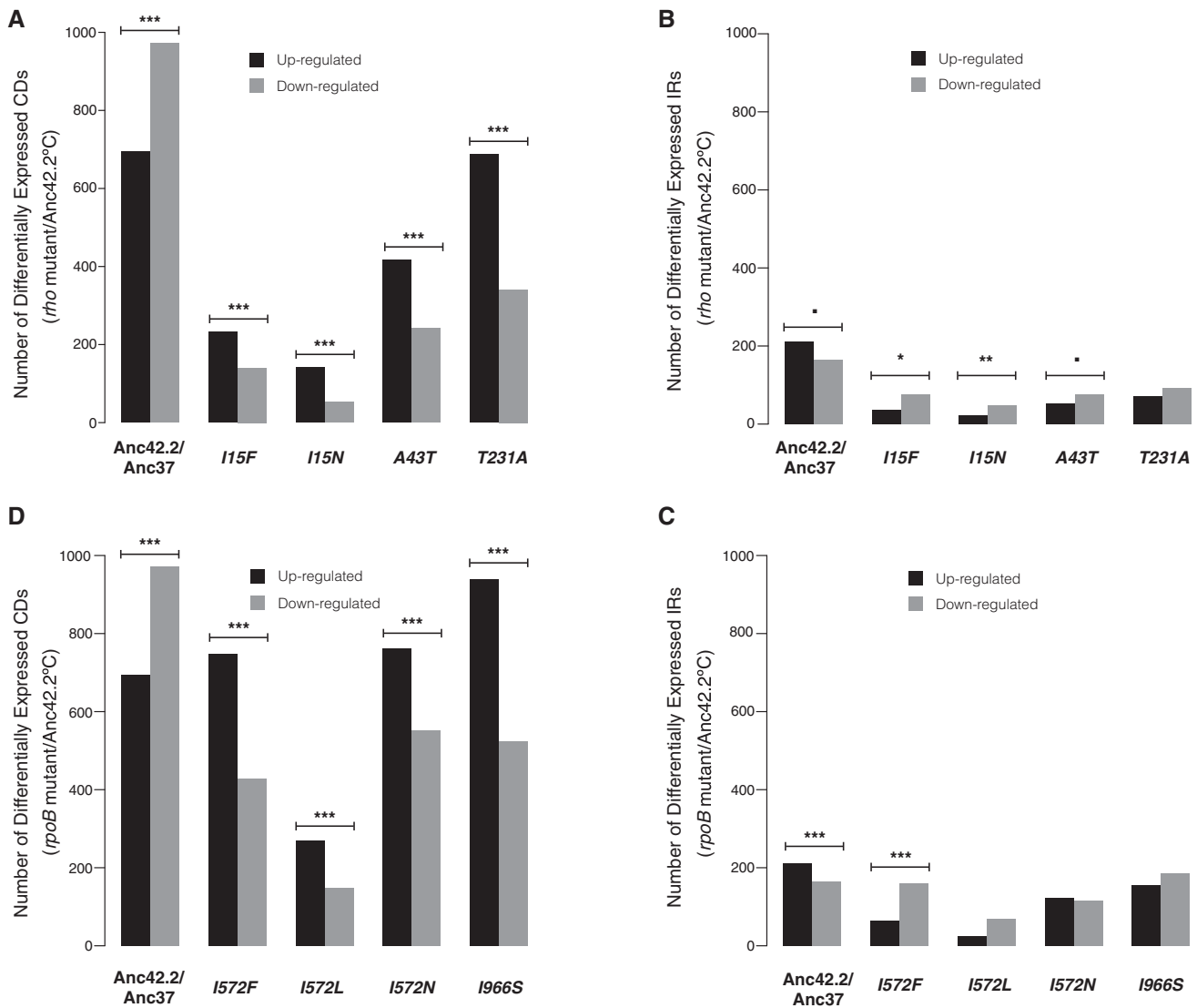
Fig. 5. Genome-wide expression patterns of genic and intergenic regions at 42.2 °C. Histograms showing the number of coding regions that are up or down regulated compared with REL1206. (A) The first pair of histograms shows REL1206 at 37.0 °C versus 42.2 °C. Up- or down-regulation is relative to the first of the listed pair. For example, the first pair of histograms show that REL1206 at 42.2 °C has ~700 and ~1,000 genes that are expressed at higher and lower levels, respectively, than REL1206 at 37.0 °C. The remaining histograms compare *rho* mutants at 42.2 °C to REL1206 at 42.2 °C. The bars with three, two and one asterisks indicate $P < 0.0001$, $P < 0.001$ and $P < 0.01$, respectively, a dot indicates $P < 0.05$. (B) Similar histograms as in A, but based on intergenic regions (IRs). (C) and (D) are identical to (A) and (B) but based on the *rpoB* mutants.

Discussion

This study was designed to compare fitness and phenotype, as measured primarily by gene expression, among and between adaptive mutations in *rho* and *rpoB*. We undertook this study both because transcriptional regulators contribute substantially to adaptive evolutionary change across many types of organisms and because mutations of these two genes are particularly common in *E. coli* experimental evolution. However, to our knowledge there have not yet been direct comparisons of the effect of beneficial mutations within and between transcriptional regulators. Our comparisons reveal substantial differences in relative fitness ($\bar{w}_r$) and growth characteristics among mutations, but we also find broad similarities. Specifically, *rho* mutations primarily affect the expression of a subset of the genes affected by *rpoB* mutations,

mutations in both genes tend to restore gene and IR expression to the unstressed state of the REL1206 ancestor, and the magnitude of this restoration correlates with $\bar{w}_r$.

Heterogeneity in Fitness and Growth Characteristics

The first major difference among beneficial mutations is in $\bar{w}_r$. Of the eight mutations, six have $\bar{w}_r$ significantly > 1.00 and are statistically indistinguishable (table 1). Of the six advantageous mutations, the three *rpoB* codon 572 mutations are unambiguous drivers of adaptive change, both because they have high fitness values and because they often fix rapidly in experimental populations (Rodríguez-Verdugo et al. 2013).

The two *rho I15* mutations have $\bar{w}_r \leq 1.0$ and thus confer no apparent advantage under thermal stress (table 1). These two mutations are unlikely to drive adaptation as solo

mutations. They thus illustrate the potential importance of linkage and epistasis, because I15N was one of the two most common nucleotide substitutions in the entire thermal stress experiment and therefore must confer a substantial benefit in the proper genetic background. But what is the proper background? Two possibilities stand out: mutations in the cls gene or in the iclR gene. Knock-downs or knock-outs of the cls gene, which encodes a cardiolipin synthase, are present in 14 of 17 of the clones with rho I15F and I15N mutations; mutations within iclR, a transcriptional repressor, are in 10 of 17 of clones; and either cls or iclR was mutated in all 17 clones (Tenaillon et al. 2012). rho mutations have been shown to interact epistatically with mutations in other genes—such as rpsL, metJ, and rpsQ—that have a role in transcription or translation (Freddolino et al. 2012; Haft et al. 2014). Both rpsL and rpsQ are critical to the 30 S ribosomal unit (Cabezon et al. 1975), and metJ represses the expression of genes involved in methionine biosynthesis (Marincs et al. 2006).

The association between rho and iclR mutations are particularly interesting, because iclR encodes a transcriptional repressor of the aceBAK operon, which encodes the enzymes involved in the glyoxylate bypass (Nègre et al. 1992). Mutations in iclR commonly arise in E. coli evolution experiments that use glucose-limited medium (Herron and Doebeli 2013; Deatherage et al. 2017), perhaps because the mutations derepress the acetate operon, thereby improving the utilization of acetate after glucose is depleted (Spencer et al. 2007). Interestingly, the aceBAK operon is up-regulated under thermal stress, but down-regulated again in all of our rho and rpoB mutants (and significantly down-regulated in rho A43T, T231A and rpoB I152F, I572N and I966S). Hence, mutations to iclR may compensate for the effect of mutations on the aceBAK operon (Yamamoto and Ishihama 2003). Note that this scenario assumes that rho mutations occurred prior to iclR mutations, but further studies are needed to assess this assumption. It is also worth noting that though iclR mutations are statistically more likely to be found in the rho evolved lines from our thermal stress experiment, they are also found in rpoB lines with the I966S and I572N mutations (Tenaillon et al. 2012). iclR mutations may thus compensating for pleiotropic effects of both rho and rpoB mutations.

In contrast to the two rho I15 mutations, the A43T and T231A mutations are capable of driving adaptive change as solo mutations, but there is a twist: they too, are often found in cls and iclR backgrounds. All three A43T clones and one of two T231A clones from our evolution experiment also had cls and iclR mutations, again implying the possibility of positively epistatic interactions. Potential functional interactions between cls and rho, if any, are unknown. The cls gene has not been identified as rho-terminated (Peters et al. 2009) and is not differentially expressed in any one of our four rho mutants. One possible—but purely speculative—interaction between cls and rho is through membrane signaling. Cardiolipin plays a key role in recruiting signaling molecules to the membrane (Romantsov et al. 2007), and E. coli can perceive alterations in membrane fluidity and transfer that information to signal-transduction pathways that alter

gene expression (Los and Murata 2004; Los et al. 2013). The potential for, and causes of, interactions between rho and cls require further study.

The effects of mutations on growth vary, too (fig. 2A and B). Generally speaking, the rpoB mutants grow better at 42.2 °C than the rho mutants, including higher maximum growth rates and yields (table 2 and supplementary table S2, Supplementary Material online). The effects on maximum growth rates are not statistically significant—probably due in part to high variance in growth characteristics for the rho mutants (fig. 2B and supplementary fig. S2, Supplementary Material online) —but the final yields for rpoB mutations are statistically higher than that of rho mutants. Of these two parameters, yield is not a component of fitness (Vasi et al. 1994) and maximum growth rate is at best a proxy for fitness (Concepción-Acevedo et al. 2015; Durão et al. 2015; Wiser and Lenski 2015). One interesting facet of growth dynamics is that the rpoB mutants are robust across 37.0 and 42.2 °C temperature regimes, but the same mutations are antagonistically pleiotropic at lower (< 20.0 °C) temperatures (Rodríguez-Verdugo et al. 2014). Moreover, estimated growth rates corroborate $\bar{w}_r$ estimates by indicating that the I15 rho mutations are not marked improvements (if at all) over the ancestor.

Mutations Alter the Expression of Similar Genes in Similar Directions

Despite differences in relative fitness and growth dynamics between rho and rpoB mutations, there are some clear similarities with respect to gene expression. One such similarity is that mutations in both genes act in part by restoring expression from the stressed condition (in this case, REL1206 under thermal stress at 42.2 °C) back toward the wild type state (in this case, REL1206 at 37.0 °C). We have detected this restoration both as a shift in PCA space (fig. 3C) and as the dominant directional shift among individual genes (table 4), although we do note that >50% of genes in rho I15N and I15F are unrestored and therefore exhibit no directional shift (table 4). Several previous studies have also shown that adaptation proceeds through phenotypic restoration (Fong et al. 2005; Carroll and Marx 2013; Sandberg et al. 2014; Carroll et al. 2015; Hug and Gaut 2015; Rodríguez-Verdugo et al. 2016), but our study is unique in establishing a quantitative correlation between restoration and $\bar{w}_r$.

There is a relevant counter-example. In their study of Methylobacterium extorquens, Carroll and Marx (2013) show that adaptation was accompanied by the restoration of ancestral gene expression levels, just as we find here. In their case, growth rates are not correlated with restoration but with reinforcement (i.e., the number of genes that exaggerate the acclimation response to physiological stress). We do not find a correlation between the number of reinforced genes and $\bar{w}_r$ ($r^2 = -0.48$; $P = 0.24$). Taken together these studies suggest that relationship among fitness and the direction of evolutionary change may vary with experimental conditions. An important goal for future research is to uncover generalities about relationships among fitness, restoration, reinforcement and novelty.

A second similarity is that the set of genes affected by mutations in *rho* and *rpoB* overlap to a remarkable extent, such that the *rho*-affected genes are a nearly perfect subset of *rpoB*-affected genes (fig. 3B). There are, of course, caveats to this observation. One is that to compare the *rho* and *rpoB* pathways in this way, we have taken the union of DEGs among *rho* and *rpoB* mutants. In reality, different mutations are associated with different DEG sets (fig. 3A); the pattern is recognizable primarily because we have studied multiple mutants in each gene. A second caveat is to recognize that the number of DEGs is an imperfect metric, because it varies with statistical power and can be nontransitive across pairwise comparisons. Finally, RNAseq approaches are also not without caveats. RNAseq measures relative RNA amounts of expression across genes within a sample, making it difficult to contrast absolute levels of expression between samples. Nonetheless, the overarching impression is that *rho* mutations mostly affect a smaller subset of the genes altered by *rpoB* mutations.

The common set of genes affected by the *rho* and the *rpoB* mutations provides insights into the physiological response to thermal stress. Mutations in *rho* and *rpoB* up-regulate functions related to cellular growth, including purine and pyridine biosynthesis, amino acid biosynthesis, ribosome biogenesis, nitrogen assimilation, translation and the regulation of translation. Mutations in *rho* and *rpoB* also down-regulate a common set of genes, particularly those associated with carbohydrate transport and amino acid degradation. The shared patterns between *rho* and *rpoB* mutations are consistent with the view that long-term acclimation to thermal stress leads to the down-regulation of growth-related pathways, potentially resulting in energy conservation and enhanced survival (Rodríguez-Verdugo et al. 2016) and that adaptation adjusts to support growth in the stressful environment (Gunasekera et al. 2008).

Although there is a considerable overlap of genes affected by *rho* and *rpoB* mutations, some functional categories were altered predominantly by one of the two pathways. For example, *rho* mutants are enriched for DEGs that respond to temperature stimuli, such as *clpB*, *dnaK*, *dnaJ*, *degP* and *grpE*. For reasons not fully understood, the *rpoB* mutations do not up-regulate these temperature response genes (Gunasekera et al. 2008; Rodríguez-Verdugo et al. 2016). In contrast, *rpoB* but not *rho* mutations enhance the expression of iron transport and flagellum biogenesis. This last observation is particularly interesting, because the up-regulation of flagellar genes is unlikely to have been adaptive in our evolution experiment and because two studies have shown that these are later down-regulated by subsequent adaptive mutations in populations (Gunasekera et al. 2008; Rodríguez-Verdugo et al. 2016). This example of flagellar genes raises the point that for all of these gross-scale effects on GE, it is difficult to distinguish between directly adaptive shifts versus negatively pleiotropic shifts.

Finally, we caution that it is difficult to separate cause from effect with respect to the overlapping gene set (fig. 2B). For example, some of this overlapping gene set likely drives enhanced growth rates in the mutants. All of our mutants have higher point estimates of maximal growth rates than REL1206 at 42.2 °C, although the difference is significant for only *rpoB* I572F and I572L (supplementary table S2, Supplementary Material online). In turn, enhanced growth probably shifts expression for a common suite of genes (Matsumoto et al. 2013). Hence, a subset of the overlapping gene set likely causes enhanced growth, whereas a second subset is an effect of enhanced growth. Unfortunately, we see no effective way to differentiate the two subsets.

Potential Impacts on Transcription Termination

Genome sequencing uncovers the type and potential function of adaptive mutations in evolution experiments (Long et al. 2015; Tenaillon et al. 2016), but the gap from "potential function" to an established molecular mechanism remains a chasm for all but a few exceptions (e.g., Blount et al. 2012; Quandt et al. 2015). With respect to thermal stress it is known that the speed of RNAP increases with increasing temperature (Ryals et al. 1982). We have therefore hypothesized that adaptive *rpoB* mutations slow RNAP transcription under thermal stress, leading to increased termination efficiency. Termination efficiency is crucial because it prevents inappropriate expression of downstream genes (Platt and Bear 1983; Platt 1986), limits potential interference from antisense transcripts (Peters et al. 2012), helps ensure that RNAP is utilized efficiently (von Hippel 1998; Ray-Soni et al. 2016), potentially minimizes collisions with replication complexes that can cause damaging double-strand breaks, and may be important for proper RNA structure (Leela et al. 2013).

Termination efficiency may require a balance (or "kinetic coupling") between Rho and RNAP activity (Jin et al. 1992). Under this model, when RNAP increases in speed, Rho must match the increase to maintain efficient termination. We predicted, then, that adaptive *rho* mutants translocate along RNA more quickly, leading to increased termination efficiency relative to REL1206 at 42.2 °C. Improved termination efficiency due to *rho* mutations is possible, in theory, because at least one *rho* mutation (L3F) has been shown to increase termination (Mori et al. 1989). To assess direct evidence for enhanced termination efficiency, we focused on known regions of *rho* termination (Peters et al. 2009) and assessed their expression dynamics. The evidence for improved termination was mixed. Using one widely utilized test (Haft et al. 2014), none of the rho mutants yielded evidence of increased termination efficiency, but two mutants (I15F and I15N) did seem to have reduced termination efficiency. Based on a new approach (fig. 4), we find that three of the four *rho* mutations have improved termination efficiency relative to REL1206 at 42.2 °C, but at a borderline level of significance. In contrast, I15N had no detectable effect.

Our results have at least three implications. First, they imply that improvement in Rho termination efficiency—at least as measured here—is not a requirement for adaptation via the *rho* pathway, because I15N provides no evidence for such a shift (fig. 4). Second, they unfortunately provide few insights into structure–function relationships, because the free energy of folding ($\Delta\Delta G$) bears no discernible relationship with read-through properties or with fitness (table 1 and supplementary

table S5, Supplementary Material online), despite predictions regarding the latter (DePristo et al. 2005; Tokuriki and Tawfik 2009; Wylie and Shakhnovich 2011; Bershtein et al. 2012; Serohijos and Shakhnovich 2014). The only discernible structural pattern was a tendency toward destabilization of the Rho protein (supplementary table S5, Supplementary Material online), which is not surprising given that ∼70% of all possible mutations on globular proteins yield $\Delta\Delta G > 1$ (Guerois et al. 2002; Schymkowitz et al. 2005) and that beneficial rpoB mutations are known to destabilize the β subunit of RNAP (Utrilla et al. 2016).

Finally, even though we could not detect direct effects of rho mutations on termination for I15N, we do note that all eight mutations tend to reduce expression of IR regions relative to REL1206 at 42.2 °C (fig. 5B and D). For the rpoB mutations, we cannot easily determine whether the effect is direct—that is, due to a modification of RNAP or downstream and indirect (Rodríguez-Verdugo et al., 2016) —but the rpoB results were consistent across analyses (figs. 4 and 5C). Our results are consistent with previous studies demonstrating that rpoB mutations can increase termination (Jin et al. 1988). Taken together, all of this information suggests that termination efficiency is an important factor for temperature adaptation.

Unfortunately, we still do not know how, exactly, enhanced termination efficiency translates to increased fitness and contributes to the restoration of gene expression. We can, however, draw several conclusions from our study of rho and rpoB mutants: that these beneficial mutations differ in their relative fitness effects, that their fitness is correlated with their restoration of gene expression, and that they tend to effect the same sets of genes while simultaneously increasing gene expression but decreasing expression of IR regions. Finally, these observations suggest that the alternative adaptive pathways typified by rho and rpoB mutations encompass substantial phenotypic convergence.

Materials and Methods

The Construction of rho Mutants

The I15F, A43T, and T231A rho mutations were introduced into the REL1206 ancestral strain using the pJk611 recombineering plasmid, as previously described (Rodríguez-Verdugo et al. 2014). Briefly, the pJk611 plasmid was introduced into REL1206, electroporating 2 µl of plasmid (80 ng of plasmid in total) into 50 µl of competent cells using an Eppendorf Electroporator 2510 set at 1.8 kV. After electroporation, 1 ml LB was added to the electroporated cells, and they were incubated at 30 °C for 2 h under constant shaking (120 rpm). Thereafter 100 µl of cells were plated on LB agar plates containing 100 µg/ml ampicillin to select ampicillin-resistant transformants. The ancestral strain carrying the pJk611 plasmid was grown overnight at room temperature (∼20 °C) in 25 ml LB with 100 µg/ml ampicillin and 1 mM L-arabinose (Sigma) until it reached and OD$_{600}$ of 0.6. Electrocompetent cells were made by washing the culture five times with ice-cold water.

To construct each mutant, two oligos of 70 bp were electroporated into cells (supplementary table S6, Supplementary Material online); one oligo introduced the desired single mutation in rho and the second oligo incorporated the change that produces an Arabinose positive (Ara+) phenotype. Two µl of each oligo (10 µM) was electroporated into 50 µl of cells. After electroporation 1 ml LB was added, and the cells were incubated at 37.0 °C for 3 h with shaking, followed by 500 µl spread onto minimal medium agar (MA) plates supplemented with L-arabinose. The plates were incubated 48 h at 42.2 °C, after which 188 single colonies were selected and grown on TA agar plates overnight at 37.0 °C. We subsequently performed colony PCR, followed by Sanger sequencing to amplify the rho gene fragment surrounding the putative mutation (supplementary table S7, Supplementary Material online). PCR conditions were 10 min at 94 °C followed by 30 cycles of 20 s at 94 °C, 30 s at 60 °C, and 1 min at 68 °C ending with an extension step of 5 min at 68 °C. For each mutation, we isolated multiple lines with the mutation of interest, such that biological replicates were independently derived.

Relative Fitness Competitions

We estimated $\bar{w}_r$ by competing mutants against the ancestral REL1206 strain. Briefly, both competitors were revived in LB broth and grown separately overnight at 37.0 °C under constant shaking (120 rpm). After cultures were revived, they were subcultured by 1:10,000 dilution in 10 ml DM25 media for 24 h at 37.0 °C and 120 rpm and inoculated 1:100 into 9.9 ml of fresh DM25 and incubated at 42.2 °C for another day in order to acclimate cells to high temperature. Although the clones had evolved at 42.2 °C in DM25, it is common practice in thermal stress studies to allow clones to recover from freezing under less stressful conditions (Bennett and Lenski 1993, Rodríguez-Verdugo et al. 2014).

Once revived, the two competitors were mixed at a 1:1 volumetric ratio and diluted 100-fold into 10 ml of fresh DM25 and incubated 1 day at the assay temperature. The initial and final densities were estimated by plating the culture onto tetrazolium-arabinose (TA) agar plates, counting the number of Arabinose+ colonies of the REL1206 ancestor and the number of Arabinose—clones of the mutants and evolved clones. Relative fitness assays were performed on three biological replicates for each mutation and for three technical replicates for each biological replicate; thus each mutation was tested against the ancestor a total of nine times.

ANOVA analyses were based on the avo package in R. With ANOVA, we first tested for the effects specific to different biological replicates within a mutation using a model of Fitness∼Mutant*Replicate. Given that there were no differences among biological replicates, the model was simplified to Fitness∼Mutant.

Growth Curves

To obtain the growth parameters of rho mutants, we first revived them in LB broth and incubated them overnight at 37.0 °C. The overnight cultures were then diluted 10,000 fold

into 10 ml DM25 and incubated for 1 day at 37 °C. The following day, the acclimated cultures were dilute 100-fold into 10 ml DM25 and incubated for 1 day at the assay temperature (either 37.0 or 42.2 °C). Cell densities were measured every hour during the lag phase and every ∼15 min during exponential phase. Cell densities were quantified using an electronic particle counter (Coulter Counter model Multisizer 3 equipped with a 30 µm diameter aperture tube), as this method has been used to measure the number of cells in low-density bacterial cultures (e.g., Julou et al. 2012). For this, 50 µl of the growing culture was diluted in 9.9 ml of Isoton II (Beckman Coulter) diluent and 50 µl of this dilution was counted electronically. Three independent growth replicates were performed per mutant. Estimations of the cell growth parameters were performed as per (Rodríguez-Verdugo et al. 2016). Briefly, we estimated the maximum growth rate by fitting a linear regression to the natural logarithm of the cell density over time during the exponential phase. We estimated the final yield from the cell counts at the end of the exponential phase. Three estimates of maximum growth rate and final yield were obtained for each mutant. The exponential phase was confirmed by linear regression (Zwietering et al. 1990). Statistical differences between *rho* and *rpoB* mutants were assessed by two sample *t*-tests. Statistical analyses (fitting of the linear regression to the exponential phase using the *lm* function and *t*-test analyses) were performed in R version 3.0.2 (Team 2013). Graphs for figure 2A and B were constructed with the function stat-smooth in the ggplot2 R package version 3.2.3 (Team 2013).

RNA Harvesting and Sequencing

To harvest cells for RNA sequencing (RNAseq), we revived cells as described above and then cultivated populations at 42.2 °C until the mid-exponential phase of population growth. The midexponential phase was determined by electronic counts using an electronic particle counter (Coulter Counter model Multisizer 3 equipped with a 30 µm diameter aperture tube) and linear regression. For each of the single *rho* mutations, at least two biological replicates were grown to midexponential for RNA harvesting.

To harvest RNA, 80 ml of culture from the midexponential phase was filtered through a 0.2 µm cellulose membrane (Life Science, Germany). Then cells were washed with Qiagen RNA-protect Bacteria Reagent and pelleted for storage at −80 °C prior to RNA extraction. Cell pellets were thawed and incubated with lysozyme (Sigma-Aldrich), for 5 min at room temperature ∼20 °C. Total RNA was isolated and purified using RNeasy Mini Kit (Qiagen). An on-column DNase treatment was performed for 30 min at room temperature. RNA quality was assessed by running an Agilent RNA-Nano chip on a bioanalyzer. rRNA was depleted using the Ribo-Zero rRNA Removal kit for Gram-negative bacteria (Epicentre Biotechnologies, Medion, WI, USA). cDNA libraries were constructed using TruSeq RNA v2 kit (Illumina, San Diego, CA, USA). After amplification, the samples were quantified by qPCR using the KAPA Library Quantification Kit code KK4822 for the Illumina Genome Analyzer platform (Kapa

Biosystems, Wilmington, MA, USA). Libraries were sequenced on an Illumina Hiseq 2000 platform to generate 100 bp single-end reads.

RNAseq data for *rpoB* mutations and the REL1206 ancestor were generated previously, using identical protocols (Rodríguez-Verdugo et al. 2016).

RNAseq Data and DEG Analysis

We gathered three sets of RNAseq data. First, we generated data for four *rho* mutants, with two replicates each. Second, we used the REL1206 data from Rodríguez-Verdugo et al. (2016), representing states of both thermal stress (42.2 °C) and normal laboratory conditions (37.0 °C). Finally, we also used data from *rpoB* single mutants from (Rodríguez-Verdugo et al. 2016). The *rpoB* data can be accessed at the NCBI's Sequence Read Archive (SRA) under the BioProject Number PRJNA291128. The *rho* data generated for this study have BioProject Number PRJNA339971 and SRA accession number SRP083065.

For all RNAseq data, sequence reads were filtered to a quality cut-off of 20. Filtered reads were mapped to the *E. coli* B REL606 reference genome (NCBI: NC_012967.1 GI: 254160123) using BWA aligner (Li and Durbin 2010) using default parameters. Statistics about mapping are provided in supplementary table S8 of the Supplementary Material online. Unique matching reads to the 4204 annotated coding regions were retained for further analyses, as were reads that mapped to 2306 noncoding intergenic regions (IRs) of the REL606 reference (supplementary data set S3, Supplementary Material online). IRs were defined as the sequence that extends from the 3′ end of a coding region to its neighboring coding region. IRs were analyzed if they exceeded 1 bp in length and had a minimum coverage of at least 1 read in at least one of the ancestors and *rho* mutants replicates. Note, however, that similar trends were obtained when IRs were defined as ≥100 bp in length.

Coding regions and IRs counts were obtained from SAM files using HTSeq-count tool (Anders et al. 2015) from HTSeq, at https://bioconductor.riken.jp/packages/3.0/bioc/html/DESeq.html (last accessed August 12, 2017) using intersection-nonempty counting mode. Differential gene expression analysis was performed with DESeq version 1.18.0 (Anders and Huber 2010). The FDR was determined using Benjamini-Hochberg adjusted *P* values. Genes with *q* values < 0.001 were considered as differentially expressed.

Gene Ontology enrichment analysis was performed through the Gene Ontology Consortium (http://geneontology.org/page/go-enrichment-analysis; last accessed August 12, 2017) with Bonferroni correction for multiple testing. Principal component analysis (PCA) was applied to RNAseq data using the *plotPCA* function in DESeq version 1.18.0 (Anders and Huber 2010) with the default parameters (ntop = 500).

Identification and Analysis of Rho-Terminated Genome Regions

To evaluate the effects of each *rho* mutation on the transcription termination activity, we evaluated the gene expression of

a set of previously defined Rho dependent terminated regions (Peters et al. 2009). Because these Rho-terminated regions were originally defined in *E. coli* K-12, it was necessary to first determine their coordinates in our REL606 genome. To do so, we utilized the start positions and lengths of the regions as defined by Peters et al. (2009) located the closest gene to the start of the Rho-terminated region in *E. coli* K-12, noting the position of the Rho-terminated region relative to the closest gene. We then located this gene in the REL606 genome and used the relative position information from K-12 to assign our Rho-terminated regions in REL606. In the few cases where gene orientation was not preserved across strains, coordinates were determined in an orientation-dependent manner (supplementary data set S2, Supplementary Material online). The number of read counts of each region was obtained from SAM files using the HTseq-count tool (Anders et al. 2015) and normalization was performed with DEseq based on a complete gff file that included the putatively Rho-terminated regions and their 5′ upstream region of the same size (Anders and Huber 2010). Five of 183 regions had no reads and were discarded. Normalized gene counts were averaged among replicates to calculate the ratio γ, the $\log_2$ ratio of mutant to ancestral counts.

For each of the 178 *rho* termination regions, we also compared read counts between the termination region and its immediate upstream 5′ region of the same length. Thus, for each of the $i = 10$ treatments (i.e., eight mutants and two control temperatures), the data consisted of paired upstream and downstream counts from $j = 178$ regions. For each clone i, we estimated termination efficiency by bootstrap resampling across the j regions, keeping the 5′ and termination regions paired. For each of 1,000 resamplings, we calculated the ratio of averages, r_{avg}, across the j samples (fig. 4). To test for differences among a set of i treatments, we bootstrapped under the null hypothesis of homogeneity by resampling across the i treatments for each of the j regions. For each of 1,000 data sets resampled under the null hypothesis, we calculated r_{avg}, for each of the i treatments. Bootstrap resamplings were performed in R version 3.0.2 (Team 2013).

Directions of Evolutionary Change

GE was classified into one of four directions of evolutionary change, as described previously (Carroll and Marx 2013, Rodríguez-Verdugo et al. 2016). The four directions classify GE in a mutant relative to REL1206 at 42.2 °C and to REL1206 at 37.0 °C. Briefly, a gene had *novel* expression if the mutant differed significantly ($q < 0.001$) from both ancestral treatments (42.2 and 37.0 °C), but GE did not differ between the two ancestral treatments (i.e., Anc42 ≈ Anc37 & Anc37 ≠ Mutant & Anc42 ≠ Mutant). *Restoration* occurred when the mutant GE levels fell between those of the two ancestral treatments and when GE differed significantly between the ancestral treatments and also between the mutant and REL1206 at 42.2 °C [i.e., (Anc42 > Mut > Anc37 or Anc42 < Mut < Anc37) & Anc42 ≠ Anc37 & Anc42 ≠ Mutant]. *Reinforced* expression occurred when GE differed significantly between the ancestral treatments and between the mutant and REL1206 at 42.2 °C, but GE in the mutant did

not fall between the ancestral treatments [i.e., Anc42 ≠ Anc 37 & Anc42 ≠ Mutant & (Mut > Anc42 > Anc37 or Mut < Anc42 < Anc37)]. Finally, *unrestored* expression occurred when GE did not differ significantly between the mutant and REL1206 at 42.2 °C, but the two ancestral treatments differed significantly from each other (i.e., Anc42 ≠ Anc 37 & Mut ≈ Anc42).

Estimates of Free Energy of Folding

We calculated two values: ΔG, the free energy of folding, and $\Delta\Delta G$, the difference of ΔG between the REL1206 protein and mutant proteins. We focused on folding free energy differences of the Rho hexamer and *rpoB*. To calculate $\Delta\Delta G$, we used an approach that combines classical molecular dynamics with FoldX (Guerois et al. 2002; Miller et al. 2016), which improves the calculation of $\Delta\Delta G$ compared with using FoldX alone (Miller et al. 2016). To perform the calculations, we downloaded the X-ray crystal structures from the Protein Data Brank (PDB): 1) Rho transcription protein bound to RNA and ADP-BeF3 (PDB ID: 3ICE) (Thomsen and Berger 2009) and 2) SigmaS-transcription initiation complex with 4-nt nascent RNA (PDB ID: 5IPL) (Liu et al. 2016). The PDB file 3ICE was modified to remove all but the six chains of Rho hexamer. Out of all the six chains in the X-ray crystal structure, chain C had least missing residues. MODELLER v9.15 software (Sali and Blundell 1993) was then used to rebuild the coordinates of the missing residues for chain C and complete chain C was used as a template to fill gaps in the other chains of the hexamer. The final hexamer structure had 2,484 residues. The chain C of 5IPL PDB file was used to obtain the structure of *rpoB* which had 1,342 residues.

The software package GROMACS 5.0.7 (Van Der Spoel et al. 2005) was used for 10 ns long Rho hexamer and *rpoB* Molecular Dynamics simulations with the AMBER99SB forcefield (Hornak et al. 2006). We followed the standard energy minimization, thermalization, and equilibration protocol (Miller et al. 2016) before carrying out the final production run. During the 10 ns production simulation snapshots were saved every 1 ns giving ten snapshots of Rho hexamer protein and *rpoB* for the further analysis.

FoldX was used to analyze the starting structure and all ten snapshots (total 11 snapshots) captured during molecular dynamics simulations (Guerois et al. 2002). Before, analyzing all 11 snapshots using FoldX, we first extracted chain C from each of the snapshots in case of Rho hexamer and *rpoB* and later only the monomer structures—that is, chain C—were used for folding free energy calculations. We started by minimizing monomer structure using the *RepairPDB* command of FoldX to further minimize the potential energy. All single amino acid mutations were then generated using *BuildModel*. Finally, protein folding stabilities were estimated using *Stability* on the *rho* monomer and *rpoB* structure. For each mutation, we then estimated $\Delta\Delta G$ by averaging across all 11 individual snapshots estimates. The molecular graphics package VMD was used to produce figure 1 (Humphrey et al. 1996).

References

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11(10):R106.

Anders S, Pyl PT, Huber W. 2015. HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2):166–169.

Bennett AF, Lenski RE. 1993. Evolutionary adaptation to temperature II. Thermal niches of experimental lines of *Escherichia coli*. *Evolution* 47(1):1–12.

Bershtein S, Mu W, Wu W, Shakhnovich EI. 2012. Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations. *Proc Natl Acad Sci U S A.* 109(13):4857–4862.

Blount ZD, Barrick JE, Davidson CJ, Lenski RE. 2012. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 489(7417):513–518.

Cabezon T, Faelen M, De Wilde M, Bollen A, Thomas R. 1975. Expression of ribosomal protein genes in *Escherichia coli*. *Mol Gen Genet.* 137(2):125–129.

Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, Nudler E. 2008. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* 320(5878):935–938.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis:a genetic theory of morphological evolution. *Cell* 134(1):25–36.

Carroll SM, Marx CJ. 2013. Evolution after introduction of a novel metabolic pathway consistently leads to restoration of wild-type physiology. *PLoS Genet* 9(4):e1003427.

Carroll SM, Chubiz LM, Agashe D, Marx CJ. 2015. Parallel and divergent evolutionary solutions for the optimization of an engineered central metabolism in Methylobacterium extorquens AM1. *Microorganisms* 3:152–174.

Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson BØ, Casadesús J. 2010. Genetic basis of growth adaptation of *Escherichia coli* after deletion of pgi, a major metabolic gene. *PLoS Genet.* 6(11):e1001186.

Ciampi MS. 2006. Rho-dependent terminators and transcription termination. *Microbiology* 152:2515–2528.

Concepción-Acevedo J, Weiss HN, Chaudhry WN, Levin BR. 2015. Malthusian parameters as estimators of the fitness of microbes:a cautionary tale about the low side of high throughput. *PLoS One* 10(6):e0126915.

Conrad TM, Frazier M, Joyce AR, Cho BK, Knight EM, Lewis NE, Landick R, Palsson BO. 2010. RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc Natl Acad Sci U S A.* 107(47): 20500–20505.

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

Conrad TM, Lewis NE, Palsson BO. 2011. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol.* 7:509.

Deatherage DE, Kepner JL, Bennett AF, Lenski RE, Barrick JE. 2017. Specificity of genome evolution in experimental populations of *Escherichia coli* evolved at different temperatures. *Proc Natl Acad Sci U S A.* 114(10):E1904–E1912.

DePristo MA, Weinreich DM, Hartl DL. 2005. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet.* 6(9):678–687.

Dillon MM, Rouillard NP, Van Dam B, Gallet R, Cooper VS. 2016. Diverse phenotypic and genetic responses to short-term selection in evolving *Escherichia coli* populations. *Evolution* 70(3):586–599.

Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* 127(7):1309–1321.

Durão P, Trindade S, Sousa A, Gordo I. 2015. Multiple resistance at no cost: rifampicin and streptomycin a dangerous liaison in the spread of antibiotic resistance. *Mol Biol Evol.* 32(10):2675–2680.

Ebright RH. 2000. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J Mol Biol.* 304(5):687–698.

Ederth J, Mooney RA, Isaksson LA, Landick R. 2006. Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. *J Mol Biol.* 356(5):1163–1179.

Fong SS, Joyce AR, Palsson BO. 2005. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res.* 15(10): 1365–1372.

Freddolino PL, Goodarzi H, Tavazoie S. 2012. Fitness landscape transformation through a single amino acid change in the rho terminator. *PLoS Genet.* 8(5):e1002744.

Guerois R, Nielsen JE, Serrano L. 2002. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol.* 320(2):369–387.

Gunasekera TS, Csonka LN, Paliy O. 2008. Genome-wide transcriptional responses of *Escherichia coli* K-12 to continuous osmotic and heat stresses. *J Bacteriol.* 190(10):3712–3720.

Haft RJF, Keating DH, Schwaegler T, Schwalbach MS, Vinokur J, Tremaine M, Peters JM, Kotlajich MV, Pohlmann EL, Ong IM, et al. 2014. Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. *Proc Natl Acad Sci U S A.* 111(25):E2576–E2585.

Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR, Albert TJ, Blattner FR, van den Boom D, Cantor CR, Palsson BØ. 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet.* 38(12):1406–1412.

Herron MD, Doebeli M. 2013. Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli. PLoS Biol.* 11(2):e1001490.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61(5):995–1016.

Hollands K, Sevostiyanova A, Groisman EA. 2014. Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator. *Proc Natl Acad Sci U S A.* 111(19):E1999–E2007.

Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65(3):712–725.

Hug SM, Gaut BS. 2015. The phenotypic signature of adaptation to thermal stress in *Escherichia coli. BMC Evol Biol.* 15:177.

Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J Mol Graph.* 14(1):33–38.

Ishihama A. 2000. Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol.* 54:499–518.

Jin DJ, Burgess RR, Richardson JP, Gross CA. 1992. Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proc Natl Acad Sci U S A.* 89(4):1453–1457.

Jin DJ, Walter WA, Gross CA. 1988. Characterization of the termination phenotypes of rifampicin-resistant mutants. *J Mol Biol.* 202(2):245–253.

Julou T, Desprat N, Bensimon D, Croquette V. 2012. Monitoring microbial population dynamics at low densities. *Rev Sci Instrum.* 83(7):074301.

King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188(4184):107–116.

Le Gac M, Cooper TF, Cruveiller S, Medigue C, Schneider D. 2013. Evolutionary history and genetic parallelism affect correlated responses to evolution. *Mol Ecol* 22(12): 3292–3303.

Lee YH, Helmann JD. 2014. Mutations in the primary sigma factor sigmaA and termination factor rho that reduce susceptibility to cell wall antibiotics. *J Bacteriol.* 196(21):3700–3711.

Leela KJ, Syeda AH, Anupama K, Gowrishankar J. 2013. Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli. Proc Natl Acad Sci U S A.* 110(1):258–273.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5):589–595.

Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE. 2009. On the origin and spread of an adaptive allele in deer mice. *Science* 325(5944):1095–1098.

Liu B, Zuo Y, Steitz TA. 2016. Structures of *E. coli* σS-transcription initiation complexes provide new insights into polymerase mechanism. *Proc Natl Acad Sci U S A.* 113(15):4051–4056.

Long A, Liti G, Luptak A, Tenaillon O. 2015. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat Rev Genet.* 16(10):567–582.

Los DA, Mironov KS, Allakhverdiev SI. 2013. Regulatory role of membrane fluidity in gene expression and physiological functions. *Photosynth Res.* 116(2–3):489–509.

Los DA, Murata N. 2004. Membrane fluidity and its roles in the perception of environmental signals. *Biochim Biophys Acta* 1666(1–2):142–157.

Matsumoto Y, Murakami Y, Tsuru S, Ying B-W, Yomo R. 2013. Growth rate-coordinated transcriptome reorganization in bacteria. *BMC Genomics* 14:808.

Marincs F, Manfield IW, Stead JA, McDowall KJ, Stockley PG. 2006. Transcript analysis reveals an extended regulon and the importance of protein-protein co-operativity for the *Escherichia coli* methionine repressor. *Biochem J.* 396(2):227–234.

Miller CR, Johnson EL, Burke AZ, Martin KP, Miura TA, Wichman HA, Brown CJ, Ytreberg FM. 2016. Initiating a watch list for Ebola virus antibody escape mutations. *PeerJ* 4:e1674.

Miwa Y, Horiguchi T, Shigesada K. 1995. Structural and functional dissections of transcriptio termination factor rho by random mutagenesis. *J Mol Biol.* 254(5):815–837.

Mori H, Imai M, Shigesada K. 1989. Mutant rho factors with increased transcription termination activities. II. Identification and functional dissection of amino acid changes. *J Mol Biol.* 210(1):39–49.

Nègre D, Cortay JC, Galinier A, Sauve P, Cozzone AJ. 1992. Specific interactions between the IclR repressor of the acetate operon of *Escherichia coli* and its operator. *J Mol Biol.* 228(1):23–29.

Olsen KM, Wendel JF. 2013. A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol.* 64:47–70.

Peichel CL, Marques DA. 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philos Trans R Soc Lond B Biol Sci.* 372:

Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, Landick R. 2012. Rho and NusG suppress pervasive antisense transcription in *Escherichia coli. Genes Dev.* 26(23):2621–2633.

Peters JM, Mooney RA, Kuan PF, Rowland JL, Keles S, Landick R. 2009. Rho directs widespread termination of intragenic and stable RNA transcription. *Proc Natl Acad Sci U S A.* 106(36):15406–15411.

Peters JM, Vangeloff AD, Landick R. 2011. Bacterial transcription terminators: the RNA 3'-end chronicles. *J Mol Biol.* 412(5):793–813.

Platt T. 1986. Transcription termination and the regulation of gene expression. *Annu Rev Biochem.* 55:339–372.

Platt T, Bear DG. 1983. Role of RNA polymerase, rho factor, and ribosomes in transcription termination. p. 123–161. *In* Beckwith J, Davies J, Gallant A (ed.). Gene function in prokaryotes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Quandt EM, Gollihar J, Blount ZD, Ellington AD, Georgiou G, Barrick JE. 2015. Fine-tuning citrate synthase flux potentiates and refines metabolic innovation in the Lenski evolution experiment. *Elife* 4:

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org.

Ray-Soni A, Bellecourt MJ, Landick R. 2016. Mechanisms of bacterial transcription termination: all good things must end. *Annu Rev Biochem.* 85:319–347.

Rodríguez-Verdugo A, Carrillo-Cisneros D, Gonzalez-Gonzalez A, Gaut BS, Bennett AF. 2014. Different tradeoffs result from alternate genetic adaptations to a common environment. *Proc Natl Acad Sci U S A.* 111(33):12121–12126.

Rodríguez-Verdugo A, Gaut BS, Tenaillon O. 2013. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol Biol.* 13:50.

Rodríguez-Verdugo A, Tenaillon O, Gaut BS. 2016. First-Step Mutations during Adaptation Restore the Expression of Hundreds of Genes. *Mol Biol Evol.* 33(1):25–39.

Romantsov T, Helbig S, Culham DE, Gill C, Stalker L, Wood JM. 2007. Cardiolipin promotes polar localization of osmosensory transporter ProP in *Escherichia coli. Mol Microbiol.* 64(6):1455–1465.

Rosenblum EB, Parent CE, Brandt EE. 2014. The molecular basis of phenotypic convergence. *Ann Rev Ecol Evol Syst.* 45(1):203–226.

Ryals J, Little R, Bremer H. 1982. Temperature dependence of RNA synthesis parameters in *Escherichia coli. J Bacteriol.* 151(2):879–887.

Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 234(3):779–815.

Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, Palsson BO, Sommer M, Feist AM. 2014. Evolution of *Escherichia coli* to 42 degrees C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. *Mol Biol Evol.* 31(10):2647–2662.

Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. 2005. The FoldX web server: an online force field. *Nucleic Acids Res.* 33(Web Server issue):W382–W388.

Serohijos AW, Shakhnovich EI. 2014. Contribution of selection for protein folding stability in shaping the patterns of polymorphisms in coding regions. *Mol Biol Evol.* 31(1):165–176.

Skordalakes E, Berger JM. 2003. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* 114(1):135–146.

Spencer CC, Bertrand M, Travisano M, Doebeli M. 2007. Adaptive diversification in genes that regulate resource use in *Escherichia coli*. *PLoS Genet.* 3(1):e15.

Stern DL. 2013. The genetic causes of convergent evolution. *Nat Rev Genet.* 14(11):751–764.

Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, Wu GC, Wielgoss S, Cruveiller S, Médigue C, et al. 2016. Tempo and mode of genome evolution in a 50, 000-generation experiment. *Nature* 536(7615):165–170.

Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS. 2012. The molecular diversity of adaptive convergence. *Science* 335(6067):457–461.

Thomsen ND, Berger JM. 2009. Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* 139(3):523–534.

Tokuriki N, Tawfik DS. 2009. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol.* 19(5):596–604.

Utrilla J, O'Brien EJ, Chen K, McCloskey D, Cheung J, Wang H, Armenta-Medina D, Feist AM, Palsson BO. 2016. Global rebalancing of cellular resources by pleiotropic point mutations illustrates a multi-scale mechanism of adaptive evolution. *Cell Syst.* 2(4):260–271.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. 2005. GROMACS: fast, flexible, and free. *J Comput Chem.* 26(16):1701–1718.

Vasi F, Travisano M, Lenski RE. 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life history traits during adaptation to a seasonal environments. *Am Nat* 144(3):432–456.

von Hippel PH. 1998. An integrated model of the transcription complex in elongation, termination, and editing. *Science* 281(5377):660–665.

Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ. 1999. Different trajectories of parallel evolution during viral adaptation. *Science* 285(5426):422–424.

Wiser MJ, Lenski RE. 2015. A comparison of methods to measure fitness in *Escherichia coli*. *PLoS One* 10(5):e0126210.

Wittkopp PJ, Kalay G. 2011. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet.* 13(1):59–69.

Wylie CS, Shakhnovich EI. 2011. A biophysical protein folding model accounts for most mutational fitness effects in viruses. *Proc Natl Acad Sci U S A.* 108(24):9916–9921.

Yamamoto K, Ishihama A. 2003. Two different modes of transcription repression of the *Escherichia coli* acetate operon by IclR. *Mol Microbiol.* 47(1):183–194.

Zwietering MH, Jongenburger I, Rombouts FM, van't Riet K. 1990. Modeling of the bacterial growth curve. *Appl Environ Microbiol.* 56(6):1875–1881.