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], Brandon S. Gaut[1,6]

[1] Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA.

[2] Department of Biological Sciences, University of Idaho, Moscow, ID, USA

[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, USA.

[6] Address for Correspondence: bgaut@uci.edu; 321 Steinhaus Hall, UC Irvine, 92617

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 two 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 altered the expression of ~1700 genes; *rho* mutations altered the expression of fewer genes, most of which were a subset of the genes altered by *rpoB*. Across our eight mutants, relative fitness correlated with the degree to which a mutation restored gene expression back to the unstressed, 37.0°C state. The *rho* mutations do not enhance transcriptional termination in known *rho*-terminated regions, but the genome-wide effects of mutations in both genes was to enhance termination. Although beneficial mutations in the two genes did not have identical effects on fitness, growth or gene expression, they acted predominantly through parallel phenotypic effects on gene expression and 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 vs. 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, Charusanti et al., 2010, Conrad 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 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

73 findings raise the possibility that some of the shifts in gene expression are maladaptive.
74 Consistent with this view, at least two studies have shown that fitness gains after the
75 fixation of a beneficial *rpoB* mutation act to ameliorate negative pleiotropic effects
76 (Sandberg et al., 2014, Rodríguez-Verdugo et al., 2016).
77 A third topic is about the relative effects of different adaptive mutations, both
78 within the same transcriptional regulator (e.g., different mutations within *rpoB*) and
79 between transcriptional regulators (e.g., *rpoB* vs. *rho*). These are potentially interesting
80 contrasts because it has been shown that different mutations can have different effects on
81 fitness, pleiotropy and evolutionary contingency (Rosenblum et al., 2014), even for
82 different mutations within the same gene (Linnen et al., 2009). It seems likely that there
83 are large-scale differences in the effects of mutations to RNAP and Rho. A modification
84 of RNAP has the potential to affect the expression of every gene, but Rho terminates
85 transcription for a subset of ~20% to 50% of *E. coli* genes (Cardinale et al., 2008, Peters
86 et al., 2009). This begs the questions: Do adaptive mutations in *rho* affect the expression
87 of fewer genes? If so, are mutations in *rho* generally more beneficial, because they cause
88 fewer negatively pleiotropic effects? And, to what extent are these effects consistent
89 across different beneficial mutations within the same transcriptional regulator?
90 To our knowledge, the effects of adaptive mutations have not yet been compared
91 explicitly between different transcriptional regulators. This may be due, in part, to the
92 fact that a valid comparison requires that mutations in the two regulatory genes are
93 adaptive in the same genetic background under identical culture conditions. Fortunately,
94 our previous thermal adaptation experiment has yielded appropriate mutations. Our
95 experiment evolved 115 populations of *E. coli* independently from the same ancestral
96 clone (REL1206) for 2000 generations at 42.2°C (Tenaillon et al., 2012). At the end of
97 the experiment, we isolated one clone from each population and found that fitness had
98 improved ~40%, on average, relative to REL1206. We also sequenced the evolved clones
99 and identified over 1300 mutations. Using parallelism as a *prima facie* argument for
100 adaptation (Wichman et al., 1999, Fong et al., 2005, Stern, 2013), about half of the
101 observed 1300 mutations were adaptive, either because the same mutation occurred in
102 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 (Rodriguez-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 non-sequence-specific interactions with DNA and nascent RNA. Some of the observed beneficial *rpoB* mutations have occurred in the contact region between the downstream DNA duplex and RNAP (proximal active site) (Rodríguez-Verdugo et al., 2014). 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 ut*ilization (*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).

134 However, a few hundred intergenic and intragenic Rho termination sites have been
135 identified experimentally (Peters et al., 2009, 2012), and these have helped to evaluate
136 whether *rho* mutants increase (Mori et al., 1989, Miwa et al., 1995) or decrease the
137 efficiency of termination (Freddolino et al., 2012, Haft et al., 2014).
138 Here we take advantage of a unique opportunity to compare and contrast the
139 effect of adaptive mutations between the *rho* and *rpoB* genes that characterize alternative
140 adaptive pathways. In the study, we engineer four beneficial *rho* mutations into the
141 REL1206 ancestor and contrast effects on fitness, growth, gene expression and
142 transcriptional termination to four previously-studied *rpoB* mutants (Rodriguez-Verdugo
143 et al., 2014, Rodríguez-Verdugo et al., 2016). In doing so, we address three sets of
144 questions. First, are the fitness and growth characteristics similar between *rho* and *rpoB*
145 mutations? Second, do the beneficial mutations influence the expression of similar
146 numbers and sets of genes? Third, do *rho* mutations alter its termination efficiency and, if
147 so, how do termination properties compare between *rho* and *rpoB* mutations? Finally, do
148 the mutations have a consistent predicted effect on protein stability, which is predicted to
149 be linked to fitness (e.g. DePristo et al., 2005)?
150
151 **RESULTS**
152 **Rho mutations and relative fitnesses**
153 We engineered four *rho* mutations - *I15N*, *I15F, A43T* and *T231A* – into the
154 ancestral REL1206 background (see Materials and Methods). Each of these substitutions
155 were adaptive by the criterion of parallelism, because they were observed in 15, 2, 3 and
156 2 independent populations, respectively, of our 42.2°C thermal stress experiment
157 (Tenaillon et al., 2012). *I15N*, *I15F* and *A43T* point mutations are located in the NHB
158 domain of the Rho monomer, while *T231A* is located in the central part of the C-terminal
159 domain (CTD), specifically in the Loop-CTD (fig. 1).
160 We assessed the relative fitness ($\overline{w}_r$) of the four *rho* mutants to REL1206 at
161 42.2˚C (supplementary table S1, Supplementary Material online). Relative fitness data
162 conveyed three observations (table 1). First, two of the four mutations (*A43T* and *T231A*)
163 yielded a $\overline{w}_r$ value significantly > 1.0, indicating an adaptive benefit under thermal stress.
164 In contrast, neither the *I15N* nor the *I15F* mutation produced a $\overline{w}_r$ value significantly

165 different from 1.0. *I15N* yielded an estimate < 1.0 (Rodriguez-Verdugo et al., 2014)
166 suggesting a slightly deleterious mutation (table 1). Second, the $\overline{w}_r$ estimates differed
167 significantly among the four mutations (ANOVA, $p = 2.4 \times 10^{-8}$), with *I15N* and *I15F*
168 forming one group and *A43T* and *T231A* forming a second, significantly different group
169 (Tukey-test; $p < 0.05$).
170 We contrasted $\overline{w}_r$ among four *rpoB* mutants (*I572L*, I*572F*, *I572N*, *I966S*)
171 (Rodriguez-Verdugo et al., 2014) and the four *rho* mutants. Each of the *rpoB* mutations
172 had $\overline{w}_r > 1.0$, with estimates ranging from $\overline{w}_r$=1.16 to 1.26 at 42.2°C (table 1). Across all
173 mutants, comparisons again indicated heterogeneity of $\overline{w}_r$ (ANOVA, $p = 1.91 \times 10^{-08}$;
174 table 1), and Tukey tests grouped *I15N* and *I15F* against the remaining six mutations ($p <$
175 0.05 for 10 of 12 comparisons between *I15* mutations and the remaining mutations).
176 Altogether, $\overline{w}_r$ values were >1.00 and indistinguishable among the four *rpoB* mutations
177 and two of the four *rho* mutations; the two remaining *rho* mutations had demonstrably
178 lower $\overline{w}_r$ values and little effect on fitness.
179
180 **Growth dynamics**
181 To gain additional insights about potential differences between *rho* and *rpoB*
182 mutants, we measured growth curves of mutant lines at 37.0˚C and 42.2˚C (fig. 2, table 2
183 and supplementary table S2, Supplementary Material online) and estimated maximal
184 growth rates and final yields at stationery phase. At 37.0˚C, all *rho* mutants had
185 significantly higher final yields than at 42.2˚C (supplementary table S2, Supplementary
186 Material online), indicating that high temperature remains stressful even after the
187 introduction of putatively beneficial mutations. At 42.2˚C, the *rho* mutants had a lower
188 final yield than the ancestor, and there were also differences in maximum growth rate
189 among mutants: *I15F* and *I15N* showed somewhat lower maximum growth rate than the
190 ancestor but *A43T* and *T231A* had higher growth rates (fig. 2 and supplementary table S2,
191 Supplementary Material online), consistent with their higher $\overline{w}_r$ values.
192 We then contrasted growth dynamics between *rho* and the three *I572 rpoB* mutants
193 (Rodríguez-Verdugo et al., 2016). Similar to results with *rho*, *rpoB* mutants grew more
194 slowly and had lower yields at 42.2˚C compared to 37.0˚C (fig. 2 and supplementary
195 table S2, Supplementary Material online). At 37.0˚C, both *rho* and *rpoB* mutants grew

196 with similar rates relative to the REL1206 ancestor (fig. 2 and supplementary table S2,
197 Supplementary Material online). However, a striking difference became apparent at
198 42.2˚C, because the *I572 rpoB* mutants reached significantly higher final yields than the
199 *rho* mutants (table 2 and supplementary table S2, Supplementary Material online). The
200 only significant difference in growth rates was the contrast between *rpoB I572F* vs. *rho*
201 *I15F* contrast (table 2). Overall, growth curves indicate that: *i*) *rho* and *rpoB* mutations
202 differ with respect to yield, and *ii*) growth rates tend to be lower for *rho I15F* compared
203 to the remaining clones.
204
205 **Gene expression phenotypes**
206 To gain further insight into the phenotypic effects of *rho* mutations, we measured
207 gene expression at 42.2˚C using replicated RNAseq data (see Materials and Methods).
208 We hypothesized that expression covaries with $\overline{w}_r$, so that *I15N* and *I15F* differ more
209 dramatically in expression from *A43T* and *T231A* than they do from each other. To test
210 this hypothesis, we measured the correlation in expression across all genes for each pair
211 of *rho* mutants (table 3). All pairwise correlations were high ($r^2 > 0.95$), but *A43T* and
212 *T231A* had the highest correlation ($r^2 = 0.992$), while contrasts between an *I15* mutant
213 with either *A43T* or *T231A* had lower correlations (table 3). We also identified
214 differentially expressed genes (DEGs) between *rho* mutants, based on a significance cut-
215 off of $q < 0.001$ (see Material and Methods). Only one gene was expressed differentially
216 between *I15F* and *I15N*, but both *I15* mutations had > 100 DEGs between *A43T* and
217 *T231A* (table 3). These data suggest that *A43T* and *T23A* mutations have quantitatively
218 different effects on GE relative to the two *I15* mutations, mimicking observed differences
219 in fitness among clones.
220 Given that single *rho* mutations can confer extensive GE shifts relative to an
221 ancestral background (Freddolino et al., 2012, Haft et al., 2014), we compared GE
222 between the REL1206 ancestor at 42.2˚C and each of the four *rho* mutants. Across the
223 four *rho* mutants, we detected a total of 1140 DEGs compared to the 42.2˚C ancestor, but
224 there were more DEGs for *T231A* (1028 genes) and *A43T* (656) than for *I15F* (372) and
225 *I15N* (195) (fig. 3*A*). Among the complete set of 1140 DEGs, 141 (or 12%) were shared
226 among all four *rho* mutants (fig. 3*A*). Gene Ontology (GO) analysis of these 141 genes

227 revealed significant enrichment for down-regulation of maltose transport genes (GO:
228 0015768, $p = 1.04 \times 10^{-3}$) and significant up regulation of genes involved in transcription
229 (GO:0010467, $p = 1.12 \times 10^{-12}$) and translation (GO: 0006412, $p = 9.17 \times 10^{-37}$)
230 (supplementary table S3, Supplementary Material online). Interestingly, the *A43T* and
231 *T231A* mutants shared 330 DEGs that were not identified in either *I15F* or *I15N*. These
232 330 genes are likely candidates to contribute to $\overline{w}_r$ differences among *rho* mutants, and
233 they were enriched for up-regulation of purine nucleotide biosynthetic processes (GO:
234 0006164, $p = 1.15 \times 10^{-12}$) and down-regulation for glycerol catabolic processes (GO:
235 0019563, $p = 1.15 \times 10^{-12}$) (supplementary table S4, Supplementary Material online).
236 Finally, to determine whether putatively beneficial mutations in *rpoB* or *rho* have
237 similar effects on gene expression, we identified DEGs in four *rpoB* mutants -- *I572N*,
238 *I572F*, *I572L* and *I966S* -- using previously published data (Rodríguez-Verdugo et al.,
239 2016). We took the union of the DEGs for the four *rpoB* mutations and found 1726 DEGs
240 relative to REL1206 at 42.2˚C. The *rpoB* and *rho* DEG sets were remarkably overlapping
241 (fig. 3*B*): of the 1726 DEGs identified with *rpoB* mutations, 957 genes (or 55.4%) were
242 also identified as DEGs in *rho* mutants. This similarity is more impressive when one
243 considers that *rho* mutations altered gene expression for significantly fewer genes than
244 *rpoB* mutations (1140 vs. 1726; binomial test: $p < 2.2 \times 10^{-16}$) so that 83% (=957/1140)
245 of DEGs caused by *rho* mutations were also identified as DEGs in *rpoB* mutants (fig.
246 3*B*). These results indicate that *rho* mutants generally affect the expression of a subset of
247 the genes affected by *rpoB* mutants.
248 The DEG overlap between *rho* and *rpoB* mutations left only 183 DEGs unique to
249 the set of *rho* mutations (fig. 3*B*). Among these 183 DEGs, GO analyses identified up-
250 regulated responses to temperature stimuli (e.g., *clpB*, *dnaK*, *dnaJ*, *degP, grpE*; GO:
251 0009408, $p = 3.35 \times 10^{-2}$) (supplementary data set S1, Supplementary Material online). In
252 contrast, the 769 genes uniquely modified by *rpoB* mutants were enriched for up-
253 regulation of iron ion transport (GO: 0006826, $p = 3.2 \times 10^{-05}$) and flagellum cell motility
254 (GO: 00048870, $p = 7.1 \times 10^{-06}$) genes. Monosaccharide catabolic processes (GO:
255 0005996, $p = 2.0 \times 10^{-04}$) and cell wall biogenesis (GO: 0042546, $p = 2.94 \times 10^{-4}$) were
256 also unique to *rpoB* mutants, but they were not consistently up- or down-regulated
257 (supplementary data set S1, Supplementary Material online).

258

259 **Shifts in gene expression are partially restorative**

260 Taken in total, *rpoB* and *rho* mutants affect an overlapping gene set, but it is not a
261 foregone conclusion that they affect expression in the same direction. Here direction
262 refers to expression shifts in mutants relative to the stressed ancestral state (i.e., REL1206
263 at 42.2 ˚C) and the unstressed ancestral state (i.e., REL1206 at 37.0˚C) (Carroll and
264 Marx, 2013). We determined directionality in two ways. First, we applied PCA to all
265 expression data for *rho, rpoB* and ancestral replicates, and then plotted the first two
266 principal components, PC1 and PC2. Based on the PCA, the *rpoB* mutants were
267 intermediate on PC1 between the ancestor at 37.0˚C and 42.2 ˚C (fig. 3*C*). This position
268 is consistent with an overall tendency of *rpoB* mutants to restore expression patterns from
269 the stressed (42.2˚C) toward the unstressed (37.0˚C) ancestral state (Rodríguez-Verdugo
270 et al., 2016). Expression patterns from *rho* mutants were also intermediate between the
271 two ancestral states on PC1, but they were not restored as fully as the *rpoB* mutants (fig.
272 3*C*). Intriguingly, replicates from the two less fit *rho* mutants (*I15N* and *I15F*) clustered
273 more closely to REL1206 at 42.2˚C. These clustering patterns suggest that fitness is
274 related to restoration in PCA space; indeed, $\overline{w}_r$ and the average PC1 eigenvector were
275 correlated across the eight mutants ($r^2 = 0.4156$; $p = 0.042$).

276 Second, we categorized expression of single genes into categories. Following
277 previous studies (Carroll and Marx, 2013), genes in mutants were classified as: *i*)
278 "restored", if expression in the mutant shifts toward the unstressed ancestral state from
279 the stressed ancestral state; *ii*) "unrestored", if expression remains similar to the
280 REL1206 ancestor in its stressed state; *iii*) "reinforced" if expression in the mutant is
281 exaggerated in the mutant relative to REL1206 at both 37.0˚C and 42.2˚C and, finally, *iv*)
282 "novel" if the two ancestral treatments did not differ in GE but the mutant differed from
283 both. (The Materials and Methods section includes more precise quantitative definitions
284 of these four directional categories.) The majority of genes fell into the restored category
285 for all mutants except *rpoB I572L* (table 4). More importantly, the number of restored
286 genes was correlated with $\overline{w}_r$ across the eight mutants (*rho* and *rpoB* mutants) at a
287 borderline level of significance (Spearman's $r = 0.67$; $p = 0.055$). Altogether, directional

288 analyses show that $\overline{w}_r$ of a mutant is a function of the degree of restoration of gene
289 expression.
290
291 **Assessing potential effects on transcriptional termination**
292 We have shown that *rho* mutations differ in $\overline{w}_r$ and that $\overline{w}_r$ correlates with shifts
293 in the magnitude and direction of gene expression. However, the question of mechanism
294 remains – i.e. how do mutations in *rho* affect its function as a transcriptional terminator?
295 Previous studies have shown that *rho* mutations tend to decrease termination and therefor
296 increase read-through activity (Freddolino et al., 2012, Haft et al., 2014). Alternatively,
297 we have hypothesized that adaptive *rho* mutations increase transcriptional efficiency at
298 high temperature by enhancing termination (Rodriguez-Verdugo et al., 2014, Rodríguez-
299 Verdugo et al., 2016). We addressed this hypothesis both directly and indirectly.
300 *Direct effects on Rho-terminated regions*: The direct approach relied on a set of
301 178 Rho-terminated regions that were defined by Peters et al. (2009) and previously used
302 to analyze read-through activity of a *rho* mutation (Haft et al., 2014) (supplementary data
303 set S2, Supplementary Material online). The advantage of studying these regions is that
304 they can, in theory, examine direct (primary) effects of *rho* mutations, as opposed to
305 indirect (secondary) effects on downstream genes. The size of the *rho*-terminated regions
306 spanned from 100 to ≈ 4000 bp, with roughly half located at the 3' end of coding regions
307 (intergenic) and the other half within genes (intragenic). Following Haft et al. (2014), we
308 calculated γ, the $\log_2$ ratio of normalized RNAseq counts between each mutant at 42.2˚C
309 and REL1206 at 42.2˚C for each of the 178 regions (see Materials and Methods). Under
310 the null hypothesis that the mutant and ancestor have similar termination properties,
311 γ is distributed binomially. If a mutation enhances termination, we expected significantly
312 more than 50% of the regions to have $\gamma < 0.0$. In contrast, increased read-through results
313 in >50% of regions having $\gamma > 0.0$ (Haft et al., 2014). Contrary to our prediction, both
314 *I15N* and *I15F* have increased read-through activity, with 108 and 119 of 178 regions
315 having $\gamma > 0.0$ (binomial, $p < 0.01$). The two remaining *rho* mutations (*A43T, T231A*)
316 yielded no evidence of altered termination properties ($p = 1.00$; supplementary fig. S1,
317 Supplementary Material online).

318 We further examined expression dynamics of the Rho-terminated regions by
319 taking the ratio of expression counts for each *rho*-terminated region (RT) against the
320 counts from the immediate 5' upstream region (UP) of the same size. We then computed
321 a measure of termination efficiency for each mutant as the ratio of averages, r_{avg}, of read
322 counts in RT vs. UP across all 178 regions (see Materials and Methods). While not
323 without caveats (such as the fact that upstream and downstream regions are in different
324 operons in some cases), this approach allowed us to measure the strength of termination
325 for each mutant and for each ancestral treatment.

326 Bootstrapped estimates of r_{avg} are presented in fig. 4. The estimates suggest
327 weaker termination transcription (i.e., enhanced read-through) for REL1206 at 42.2˚C
328 compared to 37.0˚C; although the trend is notable, the two ancestral treatments do not
329 differ significantly (bootstrap; p =0.864). For the *rho* mutants, *I15N* was alone in having
330 a higher median estimate of r_{avg} than REL1206 at 42.2˚C (fig. 4), which verifies the
331 inference of increased read-through activity based on γ values. In contrast, the
332 r_{avg} estimates for *T231A* and *A43T* (and *I15F*, to a lesser extent) appeared to restore
333 termination to the level of the ancestor at 37.0˚C (fig. 4).

334 It has been shown previously that some rpoB mutants enhance transcriptional
335 termination (Jin et al. 1988). Hence for completeness we also applied analysis of Rho-
336 terminated regions to *rpoB* mutants. Based on γ analyses, none of the *rpoB* mutants
337 yielded evidence of altered termination in these regions ($p > 0.500$; supplementary fig.
338 S1, Supplementary Material online). However, r_{avg}estimates for *rpoB* were consistently
339 lower than the ancestral treatments and the *rho* mutants (fig. 4). Furthermore, the average
340 r_{avg} estimate between the set of *rho* and *rpoB* mutants differed significantly ($p < 0.001$),
341 indicating either that termination in these regions is more efficient in *rpoB* mutants or that
342 *rpoB* mutants had consistently higher expression in the upstream regions. The *rpoB*
343 mutations also differed significantly from the ancestors at 42.2˚C and 37.0˚C (p<0.003).
344 Altogether, $r_{avg}$and γ analyses indicate that the *rho I15N* and *I15F* mutations likely
345 decrease termination efficiency, but as a group the *rpoB* mutations have statistically
346 better termination for these same regions.

347 *Indirect effects:* In addition to assessing the properties of *rho* termination directly,
348 we assessed potential effects on termination indirectly by focusing on the expression of
349 all genes and intergenic regions (IRs). Focusing first on genes, we compared expression
350 data between REL1206 at 37.0˚C and 42.2˚C. Similar to Rodriguez-Verdugo et al.
351 (2016), we found that ~1700 of 4204 coding regions were differentially expressed ($q <$
352 0.001) during thermal stress, with a tendency toward lower expression at 42.2 ˚C (972
353 down-regulated genes vs. 695 up regulated genes; binomial $p = 1.243 \times 10^{-11}$) (fig. 5*A*).
354 Given this baseline, we then assessed the effect of *rho* mutants relative to REL1206 at
355 42.2˚C. GE was significantly biased toward higher expression in each of the mutants
356 relative to REL1206 at 42.2˚C at $q < 0.001$ (fig. 5*A*; binomial, $p = 1.41 \times 10^{-10}$ (*I15N*), $p =$
357 2.13×10^{-6} (*I15F*), $p = 8.6 \times 10^{-12}$ (*A43T*), $p = 2.2 \times 10^{-16}$ (*T231A*)). In effect, thermal
358 stress led to lower GE at 42.2˚C relative to 37.0˚C, and the four mutations reversed that
359 trend. These trends were quantitatively similar for *rpoB* mutants (fig. 5*C*).
360 The picture was, however, different for IRs. Among 2306 IRs (supplementary
361 data set S3, Supplementary Material online) we found more were up-regulated in
362 REL1206 at 42.2˚C relative to 37.0˚C (212 up regulated IR vs. 164 down; regulated;
363 binomial $p = 0.015$) (fig. 5*B*). This up-tick in the expression of IRs may indicate more
364 transcriptional read-through for REL1206 under thermal stress. This trend was reversed
365 by both *rho* and *rpoB* mutations, such that differentially expressed IRs tended to have
366 *lower* expression in the mutants relative to REL1206 at 42.2˚C [fig. 5*B*, binomial $p =$
367 0.0016 (*I15N*), $p = 0.0001$ (*I15F*), $p = 0.04164$ (*A43T*); fig 5*D*; $p = 0.0016$ (*I15N*), $p =$
368 0.0001 (*I15F*), $p = 0.04164$ (*A43T*)]. These results raise the possibility that all of the *rho*
369 and *rpoB* mutations enhance termination efficiency at 42.2˚C. For the *rho* mutations this
370 is not likely a direct effect (fig. 4) but an indirect, downstream effect.
371
372 **Mutations and protein stability**
373 Finally, we evaluated the predicted folding stability of the Rho and RpoB
374 proteins. Our principal goal was to assess whether changes in fitness correlated with the
375 direction and magnitude of shifts in the free energy of protein folding stability, as is
376 expected under some evolutionary models (DePristo et al., 2005, Wylie and
377 Shakhnovich, 2011). We hypothesized that $\overline{w}_r$ from specific mutations correlates with the

378 difference of the free energy of folding between ancestral and mutant proteins (ΔΔG) at
379 42.2^{o} C. To determine the free energy of folding (ΔG), we applied an approach that
380 combines classical Molecular Dynamics (MD) with the FoldX algorithm (see Materials
381 and Methods) and then measured the change in folding stability (ΔΔG) induced by each
382 of the four *rho* and four *rpoB* mutations. Positive values of ΔΔG suggest a decrease either
383 in the stability of the protein fold or in the ability of the protein to fold relative to the
384 ancestor; negative values indicate increased stability (Guerois et al., 2002, Miller et al.,
385 2016). Three of four *rho* mutations (*I15F*, *A43T* and *I15N*) were destabilizing at 42.2˚C,
386 with ΔΔG values ranging between 0.50 to 1.22 kcal/mol (table 1); in contrast, the *T231A*
387 substitution mildly increased folding stability at 42.2˚C (ΔΔG=-0.06). Two of the four
388 *rpoB* mutations were destabilizing at 42.2˚C (table 1). Although our sample size was
389 small (n = 4) for each gene, there was no clear trend between $\overline{w}_r$ and ΔΔG (table 1). For
390 completeness, we estimated ΔΔG for 16 additional *rho* mutations from our evolution
391 experiment, for a total of 20 ΔΔG estimates (supplementary table S5, Supplementary
392 Material online). Of these, 17 of 20 ΔΔG estimates were destabilizing, and there were no
393 differences in ΔΔG for mutations in different structural features (e.g., α-helices vs. β-
394 sheets, t-test, p = 0.45; NHB and CSD-like RBD, vs. non-defined regions, p = 0.36, see
395 supplementary table S5, Supplementary Material online). The only consistent effect of
396 mutations was toward destabilizing the wild type Rho protein at high temperature.
397
398 **DISCUSSION**
399 This study was designed to compare fitness and phenotype, as measured primarily
400 by gene expression, among and between adaptive mutations in *rho* and *rpoB*. We
401 undertook this study both because transcriptional regulators contribute substantially to
402 adaptive evolutionary change across many types of organisms and because mutations of
403 these two genes are particularly common in *E. coli* experimental evolution. However, to
404 our knowledge there have not yet been direct comparisons of the effect of beneficial
405 mutations within and between transcriptional regulators. Our comparisons reveal
406 substantial differences in relative fitness ($\overline{w}_r$) and growth characteristics among
407 mutations, but we also find broad similarities. Specifically, *rho* mutations primarily affect
408 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 extent of restoration correlates with $\overline{w}_r$.

Heterogeneity in fitness and growth characteristics

The first major difference among beneficial mutations is in $\overline{w}_r$. Of the eight mutations, six have $\overline{w}_r$ 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 (Rodriguez-Verdugo et al., 2013).

The two *rho I15* mutations have $\overline{w}_r \leq 1.0$ 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 30S 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 evolution experiments with *E. coli* that use glucose-limited medium (Herron and Doebeli, 2013; Deatherage et al., 2017), perhaps because they derepress the acetate operon improving the utilization of acetate after glucose is depleted (Spencer et

440 al., 2007). Interestingly, the *aceBAK* operon is up-regulated under thermal stress, but
441 down-regulated again in all of our *rho* and *rpoB* mutants (and significantly down-
442 regulated in *rho A43T*, *T231A* and *rpoB I152F*, *I572N* and *I966S*). Hence, mutations to
443 *iclR* may compensate for the effect of mutations on the *aceB*AK operon (Yamamoto and
444 Ishihama, 2003). It is worth noting that though *iclR* mutations are statistically more
445 likely to be found in the *rho* evolved lines from our thermal stress experiment, they are
446 also found in *rpoB* lines with the *I966*S and *I572N* mutations (Tenaillon et al., 2012). *iclR*
447 mutations may thus compensating for pleiotropic effects.
448 In contrast to the two *rho I15* mutations, the *A43T* and *T231A* mutations are
449 capable of driving adaptive change as solo mutations, but there is a twist: they too, are
450 often found in *cls* and *iclR* backgrounds. All three *A43T* clones and one of two *T231A*
451 clones from our evolution experiment also had *cls* and *iclR* mutations, again implying the
452 possibility of positively epistatic interactions. Potential functional interactions between
453 *cls* and *rho*, if any, are unknown. The *cls* gene has not been identified as *rho*-terminated
454 (Peters et al., 2009) and is not differentially expressed in any one of our four *rho* mutants.
455 One possible - but purely speculative - interaction between *cls* and *rho* is through
456 membrane signaling. Cardiolipin plays a key role in recruiting signaling molecules to the
457 membrane (Romantsov et al., 2007), and *E. coli* can perceive alterations in membrane
458 fluidity and transfer that information to signal-transduction pathways that alter gene
459 expression (Los and Murata, 2004, Los et al., 2013). The potential for, and causes of,
460 interactions between *rho* and *cls* require further study.
461 The effects of mutations on growth vary, too (fig. 2). Generally speaking, the
462 *rpoB* mutants grow better at 42.2˚C than the *rho* mutants, including higher maximum
463 growth rates and yields (table 2 and supplementary table S2, Supplementary Material
464 online). The effects on maximum growth rates are not statistically significant -- probably
465 due in part to high variance in growth characteristics for the *rho* mutants (fig. 2 and
466 supplementary fig. S2, Supplementary Material online) -- but the final yields for *rpoB*
467 mutations are statistically higher than that of *rho* mutants. Of these two parameters, yield
468 is not a component of fitness (Vasi et al., 1994) and maximum growth rate is at best a
469 proxy for fitness (Concepción-Acevedo et al., 2015, Durão et al., 2015, Wiser and
470 Lenski, 2015). Nonetheless, estimated growth rates corroborate $\overline{w}_r$ estimates by

471 indicating that the *I15 rho* mutations are not marked improvements (if at all) over the
472 ancestor. Another interesting facet of growth dynamics is that the *rpoB* mutants are
473 robust across 37.0˚C and 42.2˚C temperature regimes, but the same mutations are
474 antagonistically pleiotropic at lower (< 20.0˚C) temperatures (Rodriguez-Verdugo et al.,
475 2014).
476
477 **Mutations alter the expression of similar genes in similar directions**
478 Despite differences in relative fitness and growth dynamics between *rho* and *rpoB*
479 mutations, there are some clear similarities with respect to gene expression. One such
480 similarity is that mutations in both genes act in part by restoring expression from the
481 stressed condition (in this case, REL1206 under thermal stress at 42.2˚C) back toward the
482 wild type state (in this case, REL1206 at 37.0˚C). We have detected this restoration both
483 as a shift in PCA space (fig. 3*C*) and as the dominant directional shift among individual
484 genes (table 4). Several previous studies have also shown that adaptation proceeds
485 through phenotypic restoration (Fong et al., 2005, Carroll and Marx, 2013, Sandberg et
486 al., 2014, Carroll et al., 2015, Hug and Gaut, 2015, Rodríguez-Verdugo et al., 2016), but
487 our study is unique in establishing a quantitative correlation between restoration and $\overline{w}_r$.
488 There is a relevant counter-example. In their study of *Methylobacterium*
489 *extorquens*, Carroll and Marx (2013) show that adaptation was accompanied by the
490 restoration of ancestral gene expression levels, just as we find here. In their case, growth
491 rates are not correlated with restoration but with reinforcement (i.e., the number of genes
492 that exaggerate the acclimation response to physiological stress). We do not find a
493 correlation between the number of reinforced genes and $\overline{w}_r$ (r^2 = -0.48; p= 0.24). Taken
494 together these studies suggest that relationships among fitness and the direction of
495 evolutionary change may vary with experimental conditions. An important goal for future
496 research is to uncover generalities about relationships among fitness, restoration,
497 reinforcement and novelty.
498 A second similarity is that the set of genes affected by mutations in *rho* and *rpoB*
499 overlap to a remarkable extent, such that the *rho*-affected genes are a nearly perfect
500 subset of *rpoB*-affected genes (fig. 3*B*). There are, of course, caveats to this observation.
501 One is that to compare the *rho* and *rpoB* pathways in this way, we have taken the union

502 of DEGs among *rho* and *rpoB* mutants. In reality, different mutations are associated with
503 different DEG sets (fig. 3*A*); the pattern is recognizable primarily because we have
504 studied multiple mutants in each gene. A second caveat is to recognize that the number of
505 DEGs is an imperfect metric, because it varies with statistical power and can be non-
506 transitive across pairwise comparisons. Finally, RNAseq approaches are also not without
507 caveats. RNAseq measures relative RNA amounts of expression across genes within a
508 sample, making it difficult to contrast absolute levels of expression between samples.
509 Nonetheless, the overarching impression is that *rho* mutations mostly affect a smaller
510 subset of the genes altered by *rpoB* mutations.

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

522 Although there is a considerable overlap of genes affected by *rho* and *rpoB*
523 mutations, some functional categories were altered predominantly by one of the two
524 pathways. For example, *rho* mutants are enriched for DEGs that respond to temperature
525 stimuli, such as *clpB*, *dnaK*, *dnaJ*, *degP* and *grpE*. For reasons not fully understood, the
526 *rpoB* mutations do not up-regulate these temperature response genes (Gunasekera et al.,
527 2008, Rodríguez-Verdugo et al., 2016). In contrast, *rpoB* but not *rho* mutations enhance
528 the expression of iron transport and flagellum biogenesis. This last observation is
529 particularly interesting, because the up-regulation of flagellar genes is unlikely to have
530 been adaptive in our evolution experiment and because two studies have shown that these
531 are later down-regulated by subsequent adaptive mutations in populations (Gunasekera et
532 al., 2008). This example of flagellar genes raises an important point: for all of these

533 gross-scale effects on GE, it is difficult to distinguish between directly adaptive shifts vs.
534 negatively pleiotropic shifts.
535
536 **Potential impacts on transcription termination**
537 Genome sequencing uncovers the type and potential function of adaptive
538 mutations in evolution experiments (Long et al., 2015, Tenaillon et al., 2016), but the gap
539 from 'potential function' to an established molecular mechanism remains a chasm for all
540 but a few exceptions [e.g., (Blount et al., 2012, Quandt et al., 2015)]. With respect to
541 thermal stress it is known that the speed of RNAP increases with increasing temperature
542 (Ryals et al., 1982). We have therefore hypothesized that adaptive *rpoB* mutations slow
543 RNAP transcription under thermal stress, leading to increased termination efficiency.
544 *rpoB* mutations have been shown to increase termination previously (Jin et al., 1988).
545 The effects of our *rpoB* mutations on the expression of IRs (fig. 5c) and *rho*-terminated
546 regions (fig. 4) are consistent with our increased termination efficiency hypothesis.
547 Termination efficiency may require a balance (or 'kinetic coupling') between Rho
548 and RNAP activity (Jin et al., 1992). Under this model, when RNAP increases in speed,
549 Rho must match the increase to maintain efficient termination. We predicted, then, that
550 adaptive *rho* mutants translocate along RNA more quickly, leading to increased
551 termination efficiency relative to REL1206 at 42.2˚C. Improved termination efficiency is
552 possible, in theory, because at least one *rho* mutation (*L3F*) has been shown to increase
553 termination (Mori et al., 1989). Despite this prediction, transcriptomic analyses of *rho*
554 terminated regions indicate that two *rho* mutants (*A43T* and *T231A*) had little detectable
555 direct effect on termination, and the two without a detectable fitness benefit (*I15F* and
556 *I15N*) exhibited weaker termination (i.e., enhanced read-through) (fig. 4).
557 Our results have at least three implications. First, they imply that a direct shift in
558 Rho termination efficiency – at least as measured here – is not a requirement for
559 adaptation via the *rho* pathway. Second, they unfortunately provide few insights into
560 structure-function relationships, because the free energy of folding ($\Delta\Delta G$) bears no
561 discernible relationship with read-through properties or with fitness (table 1,
562 supplementary table S5), despite predictions regarding the latter (DePristo et al., 2005,
563 Tokuriki and Tawfik, 2009, Wylie and Shakhnovich, 2011, Bershtein et al., 2012,

564 Serohijos and Shakhnovich, 2014). The only discernible structural pattern was a tendency
565 toward destabilization of the Rho protein (supplementary table S5, Supplementary
566 Material online), which is not surprising given that ~70% of all possible mutations on
567 globular proteins yield $\Delta\Delta G > 1$ (Guerois et al., 2002, Schymkowitz et al., 2005) and that
568 beneficial *rpoB* mutations are known to destabilize the β subunit of RNAP (Utrilla et al.,
569 2016).

570 Third, there remains the remarkable result that all of our *rho* and *rpoB* mutations
571 tend torestore gene expression to that of the unstressed ancestor (fig. 2*C*), typically by
572 increasing gene expression (fig. 5*A* and 5*B*) in a similar set of genes (fig. 2*B*). At the
573 same time, all eight mutations tend to reduce expression of IR regions (fig. 5*B* and 5*D*).
574 The latter result is consistent with our hypothesis of enhanced termination efficiency in
575 *rho* mutants. Analyses of *rho*-terminated regions suggest that this is not a *direct* effect of
576 the *rho* mutations but rather an *indirect*, downstream effect. Given the substantial overlap
577 in differentially expressed genes (fig. 2*B*) and intergenic expression patterns (fig. 5*B* and
578 5*D*), *rpoB* mutations appear to cause a similar effect on transcription termination. For
579 *rpoB* we cannot easily determine whether the effect is direct – i.e., due to a modification
580 of RNAP - or downstream and indirect. Despite our observations, we still do not know
581 how, exactly, enhanced termination efficiency translates to increased fitness and the
582 restoration of gene expression. We can, however, draw several conclusions from our
583 study of *rho* and *rpoB* mutants: that these beneficial mutations differ in relative fitness,
584 that their fitness is correlated with their effects on gene expression, and that they all,
585 somehow, tend to effect the same sets of genes and simultaneously increase expression of
586 genes while decreasing expression of IR regions. Finally, and perhaps most importantly,
587 these observations show that alternative adaptive pathways encompass substantial
588 phenotypic convergence.
589

590 **MATERIALS and METHODS**

591 **The construction of *rho* mutants**

592 The *I15F, A43T and T231A rho* mutations were introduced into the REL1206
593 ancestral strain using the pJk611 recombineering plasmid, as previously described
594 (Rodriguez-Verdugo et al., 2014). Briefly, the pJk611 plasmid was introduced into

595 REL1206, electroporating 2 µl of plasmid (80 ng of plasmid in total) into 50 µl of
596 competent cells using an Eppendorf Electroporator 2510 set at 1.8 kV. After
597 electroporation, 1 ml LB was added to the electroporated cells, and they were incubated
598 at 30°C for 2 h under constant shaking (120 rpm). Thereafter 100 µl of cells were plated
599 on LB agar plates containing 100 µg/ml ampicillin to select ampicillin-resistant
600 transformants. The ancestral strain carrying the pJk611 plasmid was grown overnight at
601 room temperature (~20°C) in 25 ml LB with 100 µg/ml ampicillin and 1mM L-arabinose
602 (Sigma) until it reached and OD_{600} of 0.6. Electrocompetent cells were made by washing
603 the culture five times with ice-cold water.
604 To construct each mutant, two oligos of 70 bp were electroporated into cells
605 (supplementary table S6, Supplementary Material online); one oligo introduced the
606 desired single mutation in *rho* and the second oligo incorporated the change that produces
607 an Arabinose positive (Ara+) phenotype. Two µl of each oligo (10 µM) was
608 electroporated into 50 µl of cells. After electroporation 1 ml LB was added, and the cells
609 were incubated at 37.0˚C for 3 h with shaking, followed by 500 µl spread onto minimal
610 medium agar (MA) plates supplemented with L-arabinose. The plates were incubated 48 h
611 at 42.2°C, after which 188 single colonies were selected and grown on TA agar plates
612 overnight at 37.0˚C. We subsequently performed colony PCR, followed by Sanger
613 sequencing to amplify the *rho* gene fragment surrounding the putative mutation
614 (supplementary table S6, Supplementary Material online). PCR conditions were 10 min
615 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
616 with an extension step of 5 min at 68˚C. For each mutation, we isolated multiple lines
617 with the mutation of interest, such that biological replicates were independently derived.
618
619 **Relative fitness competitions**
620 We estimated $\overline{w}_r$ by competing mutants against the ancestral REL1206 strain.
621 Briefly, both competitors were revived in LB broth and grown separately overnight at
622 37.0ºC under constant shaking (120rpm). After cultures were revived, they were
623 subcultured by 1:10000 dilution in 10 ml DM25 media for 24 hours at 37.0°C and 120
624 rpm and inoculated 1:100 into 9.9 ml of fresh DM25 and incubated at 42.2°C for another
625 day in order to acclimate cells to high temperature. Although the clones had evolved at

626 42.2°C in DM25, it is common practice in thermal stress studies to allow clones to
627 recover from freezing under less stressful conditions (Bennett and Lenski, 1993,
628 Rodriguez-Verdugo et al., 2014).
629 Once revived, the two competitors were mixed at a 1:1 volumetric ratio and
630 diluted 100-fold into 10 ml of fresh DM25 and incubated one day at the assay
631 temperature. The initial and final densities were estimated by plating the culture onto
632 tetrazolium-arabinose (TA) agar plates, counting the number of Arabinose+ colonies of
633 the REL1206 ancestor and the number of Arabinose- clones of the mutants and evolved
634 clones. Relative fitness assays were performed on three biological replicates for each
635 mutation and for three technical replicates for each biological replicate; thus each
636 mutation was tested against the ancestor a total of nine times.
637 ANOVA analyses were based on the *avo* package in R. With ANOVA, we first
638 tested for the effects specific to different biological replicates within a mutation using a
639 model of Fitness~Mutant*Replicate. Given that there were no differences among
640 biological replicates, the model was simplified to Fitness~Mutant.
641
642 **Growth curves**
643 To obtain the growth parameters of *rho* mutants, we first revived them in LB
644 broth and incubated them overnight at 37.0°C. The overnight cultures were then diluted
645 10,000 fold into 10 ml DM25 and incubated for 1 day at 37°C. The following day, the
646 acclimated cultures were dilute 100-fold into 10 ml DM25 and incubated for 1 day at the
647 assay temperature (either 37.0°C or 42.2°C). Cell densities were measured every hour
648 during the lag phase and every ~15 minutes during exponential phase. Cell densities were
649 quantified using an electronic particle counter (Coulter Counter model Multisizer 3
650 equipped with a 30 μm diameter aperture tube). For this, 50 μl of the growing culture was
651 diluted in 9.9 ml of Isoton II (Beckman Coulter) diluent and 50 μl of this dilution was
652 counted electronically. Three independent growth replicates were performed per mutant.
653 Estimations of the cell growth parameters were performed as per (Rodríguez-Verdugo et
654 al., 2016). Briefly, we estimated the maximum growth rate by fitting a linear regression
655 to the natural logarithm of the cell density over time during the exponential phase. We
656 estimated the final yield from the cell counts at the end of the exponential phase. Three

657 estimates of maximum growth rate and final yield were obtained for each mutant. The
658 exponential phase was confirmed by linear regression (Zwietering et al., 1990). Statistical
659 differences between *rho* and *rpoB* mutants were assessed by two sample *t*-tests.
660 Statistical analyses (fitting of the linear regression to the exponential phase using the *lm*
661 function and *t*-test analyses) were performed in R version 3.0.2 (Team, 2013). Graphs for
662 fig. 2 were constructed with the function stat-smooth in the ggplot2 R package version
663 3.2.3 (Team, 2013)
664
665 **RNA Harvesting and Sequencing**
666 To harvest cells for RNA sequencing (RNAseq), we revived cells as described
667 above and then cultivated populations at 42.2˚C until the mid-exponential phase of
668 population growth. The mid-exponential phase was determined by electronic counts
669 using an electronic particle counter (Coulter Counter model Multisizer 3 equipped with a
670 30 µm diameter aperture tube) and linear regression. For each of the single *rho*
671 mutations, at least two biological replicates were grown to mid-exponential for RNA
672 harvesting.
673 To harvest RNA, 80 ml of culture from the mid-exponential phase was filtered
674 through a 0.2 µm cellulose membrane (Life Science, Germany). Then cells were washed
675 with Qiagen RNA-protect Bacteria Reagent and pelleted for storage at -80˚C prior to
676 RNA extraction. Cell pellets were thawed and incubated with lysozyme (Sigma-Aldrich),
677 for 5 minutes at room temperature ~20˚C. Total RNA was isolated and purified using
678 RNeasy Mini Kit (Qiagen). An on-column DNase treatment was performed for 30 min at
679 room temperature. RNA quality was assessed by running an Agilent RNA-Nano chip on
680 a bioanalyzer. rRNA was depleted using the Ribo-Zero^{TM} rRNA Removal kit for Gram-
681 Negative Bacteria (Epicentre Biotechnologies, Medion, WI, USA). cDNA libraries were
682 constructed using TruSeq RNA v2 kit (Illumina, San Diego, CA, USA). After
683 amplification, the samples were quantified by qPCR using the KAPA Library
684 Quantification Kit code KK4822 for the Illumina Genome Analyzer platform (Kapa
685 Biosystems, Wilmington, MA, USA). Libraries were sequenced on an Illumina Hiseq
686 2000 platform to generate 100 bp single-end reads.

687 RNAseq data for *rpoB* mutations and the REL1206 ancestor were generated
688 previously, using identical protocols (Rodríguez-Verdugo et al., 2016).
689
690 **RNAseq data and DEG analysis**
691 We gathered three sets of RNAseq data. First, we generated data for four *rho*
692 mutants, with two replicates each. Second, we used the REL1206 data from Rodriguez-
693 Verdugo et al. (2016), representing states of both thermal stress (42.2˚C) and normal
694 laboratory conditions (37.0˚C). Finally, we also used data from *rpoB* single mutants from
695 (Rodríguez-Verdugo et al., 2016). The *rpoB* data can be accessed at the NCBI's
696 Sequence Read Archive (SRA) under the BioProject Number PRJNA291128. The *rho*
697 data generated for this study have BioProject Number PRJNA339971 and SRA accession
698 number SRP083065.
699 For all RNAseq data, sequence reads were filtered to a quality cut-off of 20.
700 Filtered reads were mapped to the *E. coli* B REL606 reference genome (NCBI:
701 NC_012967.1 GI:254160123) using BWA aligner (Li and Durbin, 2010) using default
702 parameters. Statistics about mapping are provided in supplementary table S7 of the
703 Supplementary Material online. Unique matching reads to the 4204 annotated coding
704 regions were retained for further analyses, as were reads that mapped to 2306 non-coding
705 intergenic regions (IRs) of the REL606 reference (supplementary data set S3,
706 Supplementary Material online). IRs were defined as the sequence that extends from the
707 3' end of a coding region to its neighboring coding region. IRs were analyzed if they
708 exceeded 1 bp in length and had a minimum coverage of at least 1 read in at least one of
709 the ancestors and *rho* mutants replicates. Note, however, that similar trends were
710 obtained when IRs were defined as ≥100 bp in length.
711 Coding regions and IRs counts were obtained from SAM files using HTSeq-count
712 tool (Anders et al., 2015) from HTSeq, at
713 http://www.huber.embl.de/users/anders/HTSeq/doc/count.html using intersection-
714 nonempty counting mode. Differential gene expression analysis was performed with
715 DESeq version 1.18.0 (Anders and Huber, 2010). The FDR was determined using
716 Benjamini-Hochberg adjusted *p*-values. Genes with *q* values <0.001 were considered as
717 differentially expressed.

718 Gene Ontology enrichment analysis was performed through the Gene Ontology
719 Consortium (http://geneontology.org/page/go-enrichment-analysis) with Bonferroni
720 correction for multiple testing. Principal component analysis (PCA) was applied to
721 RNAseq data using the *plotPCA* function in DESeq version 1.18.0 (Anders and Huber,
722 2010) with the default parameters (`ntop=500`).
723
724 **Identification and analysis of rho-terminated genome regions**
725 To evaluate the effects of each *rho* mutation on the transcription termination
726 activity, we evaluated the gene expression of a set of previously defined Rho dependent
727 terminated regions (Peters et al., 2009). Because these Rho-terminated regions were
728 originally defined in *E. coli* K-12, it was necessary to first determine their coordinates in
729 our REL606 genome. To do so, we utilized the start positions and lengths of the regions
730 as defined by Peters et al. (2009) located the closest gene to the start of the Rho-
731 terminated region in *E. coli* K-12, noting the position of the Rho-terminated region
732 relative to the closest gene. We then located this gene in the REL606 genome and used
733 the relative position information from K-12 to assign our Rho-terminated regions in
734 REL606. In the few cases where gene orientation was not preserved across strains,
735 coordinates were determined in an orientation-dependent manner (supplementary data set
736 S2, Supplementary Material online). The number of read counts of each region was
737 obtained from SAM files using the HTseq-count tool (Anders et al., 2015) and
738 normalization was performed with DEseq based on a complete gff file that included the
739 putatively Rho-terminated regions and their 5' upstream region of the same size (Anders
740 and Huber, 2010). Five of 183 regions had no reads and were discarded. Normalized gene
741 counts were averaged among replicates to calculate the ratio γ , the $\log_2$ ratio of mutant to
742 ancestral counts.
743 For each of the 178 *rho* termination regions, we also compared read counts
744 between the termination region and its immediate upstream 5' region of the same length.
745 Thus, for each of the i=10 treatments (i.e., 8 mutants and 2 control temperatures), the data
746 consisted of paired upstream and downstream counts from j=178 regions. For each clone
747 i, we estimated termination efficiency by bootstrap resampling across the j regions,
748 keeping the 5' and termination regions paired. For each of 1000 resamplings, we

749 calculated the ratio of averages, r_{avg}, across the *j* samples (fig. 4). To test for differences
750 among a set of *i* treatments, we bootstrapped under the null hypothesis of homogeneity
751 by resampling across the *i* treatments for each of the *j* regions. For each of 1000 datasets
752 resampled under the null hypothesis, we calculated r_{avg},for each of the *i* treatments.
753 Bootstrap resamplings were performed in R version 3.0.2 (Team, 2013).
754
755 **Directions of evolutionary change**
756 GE was classified into one of four directions of evolutionary change, as described
757 previously (Carroll and Marx, 2013, Rodríguez-Verdugo et al., 2016). The four directions
758 classify GE in a mutant relative to REL1206 at 42.2˚C and to REL1206 at 37.0˚C.
759 Briefly, a gene had *novel* expression if the mutant differed significantly ($q < 0.001$) from
760 both ancestral treatments (42.2˚C and 37.0˚C), but GE did not differ between the two
761 ancestral treatments (i.e. Anc42≈Anc37 & Anc37≠Mutant & Anc42≠Mutant).
762 *Restoration* occurred when the mutant GE levels fell between those of the two ancestral
763 treatments and when GE differed significantly between the ancestral treatments and also
764 between the mutant and REL1206 at 42.2˚C [i.e., (Anc42>Mut>Anc37 or
765 Anc42<Mut<Anc37) & Anc42 ≠Anc37 & Anc42 ≠Mutant]. *Reinforced* expression
766 occurred when GE differed significantly between the ancestral treatments and between
767 the mutant and REL1206 at 42.2˚C, but GE in the mutant did not fall between the
768 ancestral treatments [i.e., Anc42 ≠Anc 37 & Anc42 ≠Mutant & (Mut>Anc42>Anc37
769 or Mut<Anc42<Anc37)]. Finally, *unrestored* expression occurred when GE did not differ
770 significantly between the mutant and REL1206 at 42.2˚C, but the two ancestral
771 treatments differed significantly from each other (i.e., Anc42 ≠Anc 37 & Mut≈Anc42).
772
773 **Estimates of free energy of folding.**
774 We calculated two values: ΔG, the free energy of folding, and ΔΔG, the
775 difference of ΔG between the REL1206 protein and mutant proteins. We focused on
776 folding free energy differences of the Rho hexamer and *rpoB*. To calculate ΔΔG, we used
777 an approach that combines classical molecular dynamics with FoldX (Guerois et al.,
778 2002, Miller et al., 2016), which improves the calculation of ΔΔG compared to using

779 FoldX alone (Miller et al., 2016). To perform the calculations, we downloaded the X-ray
780 crystal structures from the Protein Data Brank (PDB): (1) Rho transcription protein
781 bound to RNA and ADP-BeF3 (PDB ID: 3ICE) (Thomsen and Berger, 2009) and (ii)
782 SigmaS-transcription initiation complex with 4-nt nascent RNA (PDB ID: 5IPL) (Liu et
783 al., 2016). The PDB file 3ICE was modified to remove all but the six chains of Rho
784 hexamer. Out of all the six chains in the X-ray crystal structure, chain C had least missing
785 residues. MODELLER v9.15 software (Sali and Blundell, 1993) was then used to rebuild
786 the coordinates of the missing residues for chain C and complete chain C was used as a
787 template to fill gaps in the other chains of the hexamer. The final hexamer structure had
788 2484 residues. The chain C of 5IPL PDB file was used to obtain the structure of *rpoB*
789 which had 1342 residues.

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

797 Fold X was used to analyze the starting structure and all 10 snapshots (total 11
798 snapshots) captured during molecular dynamics simulations (Guerois et al., 2002).
799 Before, analyzing all 11 snapshots using FoldX, we first extracted chain C from each of
800 the snapshots in case of Rho hexamer and *rpoB* and later only the monomer structures -
801 i.e. chain C - were used for folding free energy calculations. We started by minimizing
802 monomer structure using the *RepairPDB* command of FoldX to further minimize the
803 potential energy. All single amino acid mutations were then generated using *BuildModel*.
804 Finally, protein folding stabilities were estimated using *Stability* on the *rho* monomer and
805 *rpoB* structure. For each mutation we then estimated $\Delta\Delta G$ by averaging across all 11
806 individual snapshots estimates. The molecular graphics package VMD was used to
807 produce fig. 1(Humphrey et al., 1996).

808

809 **ACKNOWLEDGMENTS**

810 We thank R. Gaut and P. MacDonald for technical assistance. A.G.G. and A.R.V. were
811 supported by University of California Institute for Mexico and the United States-Consejo
812 Nacional de Ciencia y Tecnología (Mexico) Post-doctoral and Doctoral fellowships
813 respectively. S.H. was supported by the National Institute of Biomedical Imaging and
814 Bioengineering, National Research Service Award EB009418 from University of
815 California, Irvine, Center for Complex Biological Systems. J.S.P. was supported by the
816 National Institute Of General Medical Sciences of the National Institutes of Health under
817 Award Number P20GM104420 from Center for Modeling Complex Interactions,
818 University of Idaho, Moscow. BSG was supported by a fellowship from the Borchard
819 Foundation, and the work was also supported by the National Science Foundation grant
820 DEB-0748903. Computer resources to perform the Molecular dynamic simulations were
821 provided by the Institute for Bioinformatics and Evolutionary Studies Computational
822 Resources Core sponsored by the National Institute of Health (P30 GM103324) at
823 University of Idaho.

824

825

826 **FIGURE LEGENDS**

827

828 **FIG. 1.**

829 A schematic of the Rho protein, showing each of its six hexamers (A-F) and the location
830 of the single-mutations (red spheres) produced for this experiment. Mutations in *I15* and
831 *A43* are located in the N-terminal domain (NTD) (gray); the mutation in *A231* in the C-
832 terminal domain (CTD) (green).

833

834 **FIG. 2.**

835 Growth curves for the mutants and ancestor at 37.0˚C and 42.2˚C. Each line corresponds
836 to a local polynomial regression fitting of three replicates with its 95% confidence
837 interval (see table 2 and supplementary table S2).

838

839 **FIG. 3.**

840 *(A)* A Venn diagram of the number of genes that exhibit significant DEG between *rho*
841 mutants. *(B)* A Venn diagram of the number of genes that differ in expression between
842 the four *rho* mutants and REL1206 at 42.2˚C and all four *rpoB* mutants and REL1206 at
843 42.2˚C. *(C)* A PCA plot of the RNAseq data, including replicated data from REL1206 at
844 37.0˚C (light blue dots), REL1206 at 42.2˚C (dark blue dots); replicated data from each
845 of the four *rho* mutations (light and dark green dots); and replicated data from each of the
846 four *rpoB* mutations (pink dots). The *x*-axis represents restoration along PC1, where
847 smaller values reflect greater restorative effects.
848
849 **FIG. 4.**
850 A graph of bootstrap estimates of the ratio of averages ($\bar{r}_{avg}$) that measures the ratio of
851 counts between *rho* terminated sites (RT) and their upstream 5' regions (UP) across 178
852 loci. Lines in the boxplots represent the median bootstrap-based estimate of $\bar{r}_{avg}$ based on
853 1,000 resamplings (see Material and Methods). The boxes represent the upper and lower
854 quartiles of bootstrap replicates. Anc37 represents the REL1206 ancestor at 37.0°C and
855 Anc42 is the ancestor at 42.2°C.
856
857 **FIG. 5.**
858 Genome-wide expression patterns of genic and intergenic regions at 42.2°C. Histograms
859 showing the number of coding regions that are up or down regulated compared to
860 REL1206. *(A)* The first pair of histograms shows REL1206 at 37.0˚C vs. 42.2˚C. Up- or
861 down-regulation is relative to the first of the listed pair. For example, the first pair of
862 histograms show that REL1206 at 42.2˚C has ~700 and ~1000 genes that are expressed at
863 higher and lower levels, respectively, than REL1206 at 37.0˚C. The remaining
864 histograms compare *rho* mutants at 42.2˚C to REL1206 at 42.2˚C. The bars with three,
865 two and one asterisks indicate $p < 0.0001$, $p < 0.001$ and $p < 0.01$ respectively, a dot
866 indicates $p < 0.05$. *(B)* Similar histograms as in *A*, but based on intergenic regions (IRs).
867 *(C)* and *(D)* are identical to *(A)* and *(B)* but based on the *rpoB* mutants.
868
869

870

871

872

873 **Table 1.** Relative fitness measured at 42.2°C for four *rho* mutants and four *rpoB* mutants,
874 along with their estimated ΔΔG values, as predicted by molecular modeling.

875

	Relative Fitness			Estimated ΔΔG values	
Mutant[1]	**Mean**	**± 95% CI**	***p*-value[2]**	**Location[3]**	**ΔΔG at 42.2°C (kcal/mol) ± S.E.[4]**
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

876

877 [1] *rho* data from this study, based on nine replicated measures. *rpoB* data include that from
878 (Rodriguez-Verdugo et al., 2014) plus three additional replicates, for a total of six
879 replicates

880 [2] $p < 0.05$ rejects the null hypothesis that $\overline{w}_r = 1.0$, indicating that the mutation is
881 beneficial ($\overline{w}_r > 1.0$) relative to the REL1206 ancestor.

882 [3] Location of mutation in the Rho protein: NH B=N-terminal helix bundle * Mutations
883 located in the NTD= N-terminal and **CTD= C-terminal domain. Location of mutations
884 in the RNAP protein: + RNA polymerase, β subunit, Rif= Rif-binding regions in the fork
885 domain, ++ RNA polymerase, β subunit, Eco flap domain, β subunit insert 9.

886 [4] Errors are calculated with 95% confidence interval. ΔΔG value for *rho* and *rpoB* W.T.
887 is zero.
888
889
890

891 **Table 2.** *P*-values for statistical comparisons of maximum growth rate and yield of *rho*
892 and *rpoB* mutants at 42.2°C
893
894

rho / ***rpoB***	*I15N* [1]	*I15F*	*A43T*	*T231A*
		Maximum growth rate		
I572N	0.138	0.471	0.440	0.127
I572F	0.177	**0.045** [2]	0.732	0.157
I572L	0.203	0.092	0.995	0.175
		Final Yield [3]		
I572N	**0.001**	**8 x 10^{-4}**	**0.002**	**0.010**
I572F	**3 x 10^{-4}**	**4 x 10^{-4}**	**0.001**	**0.004**
I572L	**7 x 10^{-5}**	**4 x 10^{-5}**	**1 x 10^{-4}**	**0.002**

895
896 [1] Entries correspond to significant values from two-sample *t*-tests. The null hypothesis is
897 that the mean values of parameters are equal. Numbers in bold correspond to *P*-values
898 <0.05.
899 [2] *rpoB I572F* had a significantly higher maximum growth rate compared with the *rho*
900 mutant *I15F* at 42.2°C (see fig. 2 and supplementary table S2, Supplementary Material
901 online).
902 [3] In every case, r*poB* mutants had a significantly higher final yield compared with the *rho*
903 mutants at 42.2°C (fig. 2 and supplementary table S2, Supplementary Material online).
904
905

906 **Table 3.** Comparisons of gene expression between single *rho* mutants. The cells above
907 the diagonal report the Pearson correlation coefficient in gene expression across all genic
908 regions (n=4204). The cells below the diagonal indicate the number of genes significantly
909 different via DEseq (q <0.001).
910

	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	

911
912
913

914
915
916
917 **Table 4.** The direction of gene expression for genes within *rho* and *rpoB* mutants.
918

Mutant [1]	Restored [2]	Unrestored	Reinforced	Novel	Total [3]
rho I15F	762	886	19	4	1671
rho I15N	802	857	8	16	1683
rho A43T	1115	544	7	32	1698
rho T231A	1250	405	12	46	1713
rpoB I572F	1092	644	1	59	1796
rpoB I572L	563	1174	0	6	1743
rpoB I572N	1165	569	3	53	1790
rpoB I966S	1308	428	1	39	1776

919
920 [1] *rho* data from this study. *rpoB* data from (Rodriguez-Verdugo et al., 2014).
921 [2] The number of genes was counted as restored, unrestored, reinforced and novel using
922 the rules elaborated in the Material and Methods.
923 [3] The total tallies the number of genes that differ at $q < 0.001$ from either REL1206 at
924 37.0˚C or REL1206 at 42.2˚C or both. The total differs significantly between *rho* and
925 *rpoB* mutations (t-test; $p = 0.0016$).

926 **CITATIONS**

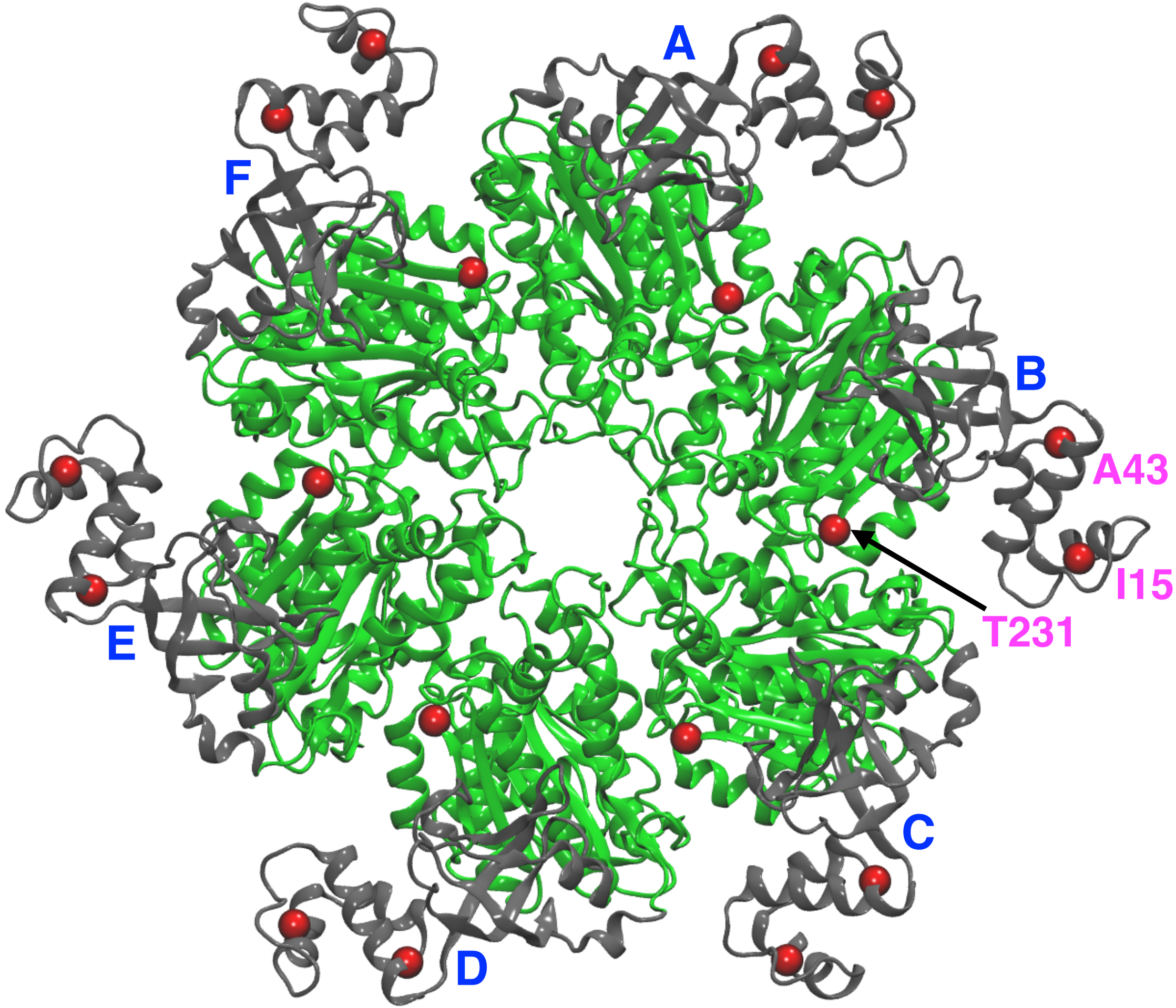
927 Anders S, Huber W. 2010. Differential expression analysis for sequence count data.
928 Genome Biol. 11:R106.
929 Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-
930 throughput sequencing data. Bioinformatics. 31:166-169.
931 Bennett AF, Lenski RE. 1993. Evolutionary adaptation to temperature II. Thermal niches
932 of experimental lines of Escherichia coli. Evolution. 47:1-12.
933 Bershtein S, Mu W, Shakhnovich EI. 2012. Soluble oligomerization provides a beneficial
934 fitness effect on destabilizing mutations. Proc Natl Acad Sci U S A. 109:4857-4862.
935 Blount ZD, Barrick JE, Davidson CJ, Lenski RE. 2012. Genomic analysis of a key
936 innovation in an experimental Escherichia coli population. Nature. 489:513-518.
937 Cabezon T, Faelen M, De Wilde M, Bollen A, Thomas R. 1975. Expression of ribosomal
938 protein genes in Escherichia coli. Mol Gen Genet. 137:125-129.
939 Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, Nudler E. 2008.
940 Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E.
941 coli. Science. 320:935-938.
942 Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of
943 morphological evolution. Cell. 134:25-36.
944 Carroll SM, Chubiz LM, Agashe D, Marx CJ. 2015. Parallel and divergent evolutionary
945 solutions for the optimization of an engineered central metabolism in
946 Methylobacterium extorquens AM1. Microorganisms.
947 Carroll SM, Marx CJ. 2013. Evolution after introduction of a novel metabolic pathway
948 consistently leads to restoration of wild-type physiology. PLoS Genet. 9:e1003427.
949 Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y,
950 Palsson BO. 2010. Genetic basis of growth adaptation of Escherichia coli after
951 deletion of pgi, a major metabolic gene. PLoS Genet. 6:e1001186.
952 Ciampi MS. 2006. Rho-dependent terminators and transcription termination.
953 Microbiology. 152:2515-2528.
954 Concepción-Acevedo J, Weiss HN, Chaudhry WN, Levin BR. 2015. Malthusian
955 Parameters as Estimators of the Fitness of Microbes: A Cautionary Tale about the
956 Low Side of High Throughput. PLoS One. 10:e0126915.
957 Conrad TM, Frazier M, Joyce AR, Cho BK, Knight EM, Lewis NE, Landick R, Palsson
958 BO. 2010. RNA polymerase mutants found through adaptive evolution reprogram
959 Escherichia coli for optimal growth in minimal media. P Natl Acad Sci Usa.
960 107:20500-20505.
961 Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y, Palsson BO. 2009.
962 Whole-genome resequencing of Escherichia coli K-12 MG1655 undergoing short-
963 term laboratory evolution in lactate minimal media reveals flexible selection of
964 adaptive mutations. Genome Biol. 10:R118.
965 Conrad TM, Lewis NE, Palsson BO. 2011. Microbial laboratory evolution in the era of
966 genome-scale science. Mol Syst Biol. 7:509.
967 Deatherage DE, Kepner JL, Bennett AF, Lenski RE, Barrick JE. 2017. Specificity of
968 genome evolution in experimental populations of Escherichia coli evolved at
969 different temperatures. Proc Natl Acad Sci U S A. 114:E1904-E1912.
970 DePristo MA, Weinreich DM, Hartl DL. 2005. Missense meanderings in sequence space:
971 a biophysical view of protein evolution. Nat Rev Genet. 6:678-687.

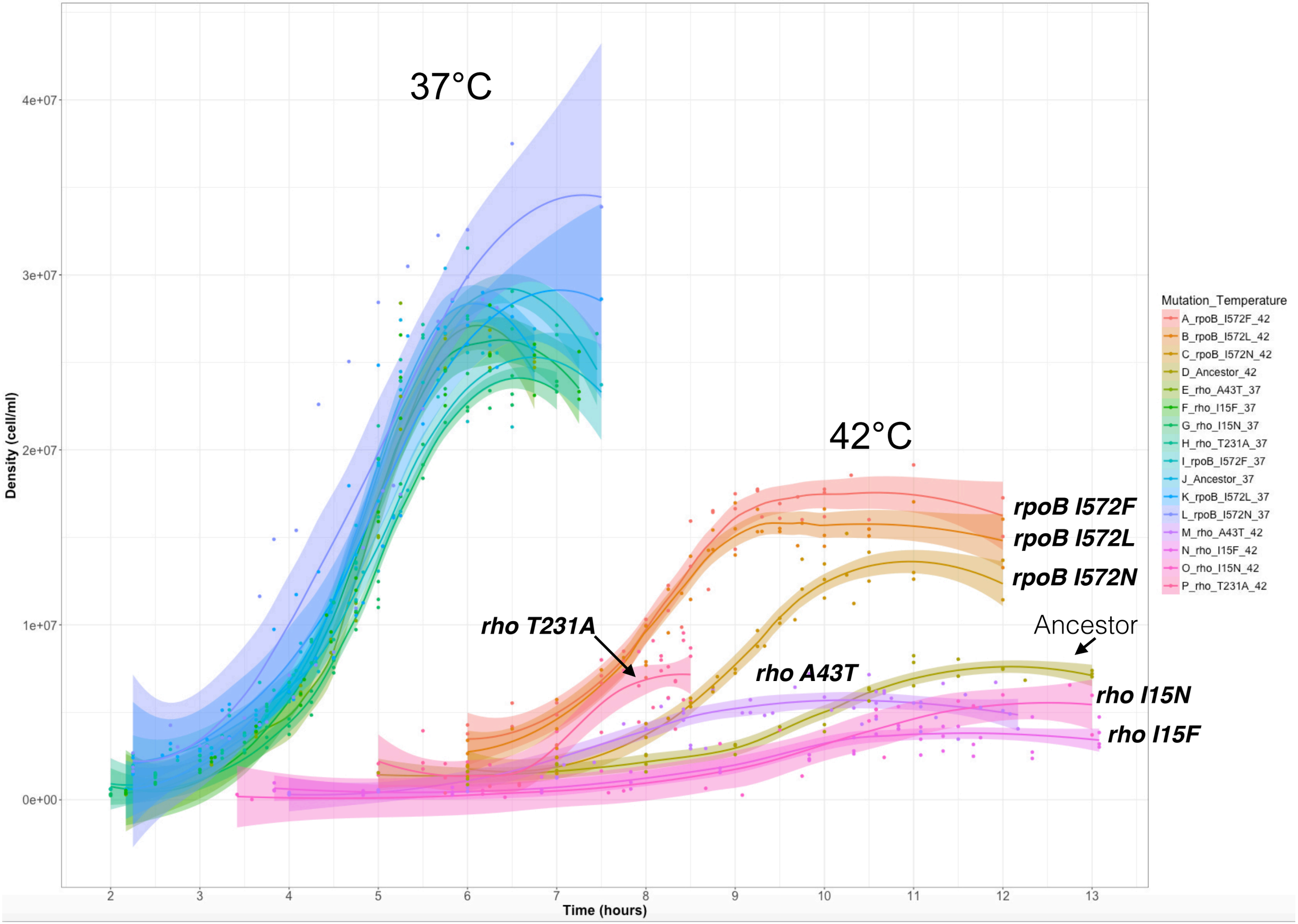
972 Dillon MM, Rouillard NP, Van Dam B, Gallet R, Cooper VS. 2016. Diverse phenotypic
973 and genetic responses to short-term selection in evolving Escherichia coli
974 populations. Evolution. 70:586-599.
975 Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication.
976 Cell. 127:1309-1321.
977 Durão P, Trindade S, Sousa A, Gordo I. 2015. Multiple Resistance at No Cost:
978 Rifampicin and Streptomycin a Dangerous Liaison in the Spread of Antibiotic
979 Resistance. Mol Biol Evol. 32:2675-2680.
980 Ebright RH. 2000. RNA polymerase: structural similarities between bacterial RNA
981 polymerase and eukaryotic RNA polymerase II. J Mol Biol. 304:687-698.
982 Ederth J, Mooney RA, Isaksson LA, Landick R. 2006. Functional interplay between the
983 jaw domain of bacterial RNA polymerase and allele-specific residues in the product
984 RNA-binding pocket. J Mol Biol. 356:1163-1179.
985 Fong SS, Joyce AR, Palsson BO. 2005. Parallel adaptive evolution cultures of
986 Escherichia coli lead to convergent growth phenotypes with different gene
987 expression states. Genome Res. 15:1365-1372.
988 Freddolino PL, Goodarzi H, Tavazoie S. 2012. Fitness landscape transformation through
989 a single amino acid change in the rho terminator. PLoS Genet. 8:e1002744.
990 Guerois R, Nielsen JE, Serrano L. 2002. Predicting changes in the stability of proteins
991 and protein complexes: a study of more than 1000 mutations. J Mol Biol. 320:369-
992 387.
993 Gunasekera TS, Csonka LN, Paliy O. 2008. Genome-wide transcriptional responses of
994 Escherichia coli K-12 to continuous osmotic and heat stresses. J Bacteriol. 190:3712-
995 3720.
996 Haft RJ, Keating DH, Schwaegler T et al. 2014. Correcting direct effects of ethanol on
997 translation and transcription machinery confers ethanol tolerance in bacteria. Proc
998 Natl Acad Sci U S A. 111:E2576-85.
999 Herring CD, Raghunathan A, Honisch C et al. 2006. Comparative genome sequencing of
1000 Escherichia coli allows observation of bacterial evolution on a laboratory timescale.
1001 Nat Genet. 38:1406-1412.
1002 Herron MD, Doebeli M. 2013. Parallel evolutionary dynamics of adaptive diversification
1003 in Escherichia coli. PLoS Biol. 11:e1001490.
1004 Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of
1005 adaptation. Evolution. 61:995-1016.
1006 Hollands K, Sevostiyanova A, Groisman EA. 2014. Unusually long-lived pause required
1007 for regulation of a Rho-dependent transcription terminator. Proc Natl Acad Sci U S
1008 A. 111:E1999-2007.
1009 Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. 2006. Comparison
1010 of multiple Amber force fields and development of improved protein backbone
1011 parameters. Proteins. 65:712-725.
1012 Hug SM, Gaut BS. 2015. The phenotypic signature of adaptation to thermal stress in
1013 Escherichia coli. BMC Evol Biol. 15:177.
1014 Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. J Mol
1015 Graph. 14:33-8, 27.
1016 Ishihama A. 2000. Functional modulation of Escherichia coli RNA polymerase. Annu
1017 Rev Microbiol. 54:499-518.

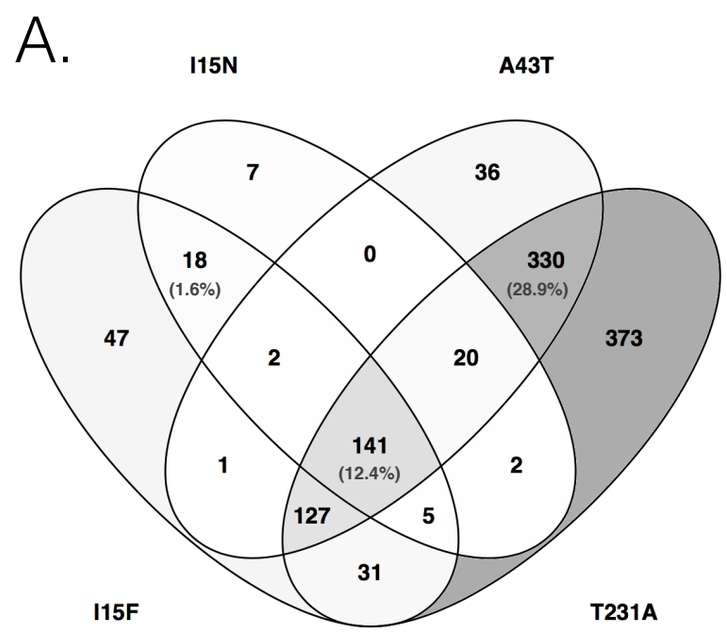
1018 Jin DJ, Burgess RR, Richardson JP, Gross CA. 1992. Termination efficiency at rho-
1019 dependent terminators depends on kinetic coupling between RNA polymerase and
1020 rho. Proc Natl Acad Sci U S A. 89:1453-1457.
1021 Jin DJ, Walter WA, Gross CA. 1988. Characterization of the termination phenotypes of
1022 rifampicin-resistant mutants. J Mol Biol. 202:245-253.
1023 King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees.
1024 Science. 188:107-116.
1025 Le Gac M, Cooper TF, Cruveiller S, Medigue C, Schneider D. 2013. Evolutionary history
1026 and genetic parallelism affect correlated responses to evolution. Mol Ecol. 22:3292-
1027 3303.
1028 Lee YH, Helmann JD. 2014. Mutations in the primary sigma factor sigmaA and
1029 termination factor rho that reduce susceptibility to cell wall antibiotics. J Bacteriol.
1030 196:3700-3711.
1031 Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler
1032 transform. Bioinformatics. 26:589-595.
1033 Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE. 2009. On the origin and spread of an
1034 adaptive allele in deer mice. Science. 325:1095-1098.
1035 Liu B, Zuo Y, Steitz TA. 2016. Structures of E. coli σS-transcription initiation complexes
1036 provide new insights into polymerase mechanism. Proc Natl Acad Sci U S A.
1037 113:4051-4056.
1038 Long A, Liti G, Luptak A, Tenaillon O. 2015. Elucidating the molecular architecture of
1039 adaptation via evolve and resequence experiments. Nat Rev Genet. 16:567-582.
1040 Los DA, Mironov KS, Allakhverdiev SI. 2013. Regulatory role of membrane fluidity in
1041 gene expression and physiological functions. Photosynth Res. 116:489-509.
1042 Los DA, Murata N. 2004. Membrane fluidity and its roles in the perception of
1043 environmental signals. Biochim Biophys Acta. 1666:142-157.
1044 Marincs F, Manfield IW, Stead JA, McDowall KJ, Stockley PG. 2006. Transcript
1045 analysis reveals an extended regulon and the importance of protein-protein co-
1046 operativity for the Escherichia coli methionine repressor. Biochem J. 396:227-234.
1047 Miller CR, Johnson EL, Burke AZ, Martin KP, Miura TA, Wichman HA, Brown CJ,
1048 Ytreberg FM. 2016. Initiating a watch list for Ebola virus antibody escape mutations.
1049 PeerJ. 4:e1674.
1050 Miwa Y, Horiguchi T, Shigesada K. 1995. Structural and functional dissections of
1051 transcription termination factor rho by random mutagenesis. J Mol Biol. 254:815-
1052 837.
1053 Mori H, Imai M, Shigesada K. 1989. Mutant rho factors with increased transcription
1054 termination activities. II. Identification and functional dissection of amino acid
1055 changes. J Mol Biol. 210:39-49.
1056 Nègre D, Cortay JC, Galinier A, Sauve P, Cozzone AJ. 1992. Specific interactions
1057 between the IclR repressor of the acetate operon of Escherichia coli and its operator.
1058 J Mol Biol. 228:23-29.
1059 Olsen KM, Wendel JF. 2013. A bountiful harvest: genomic insights into crop
1060 domestication phenotypes. Annu Rev Plant Biol. 64:47-70.
1061 Peichel CL, Marques DA. 2017. The genetic and molecular architecture of phenotypic
1062 diversity in sticklebacks. Philos Trans R Soc Lond B Biol Sci. 372

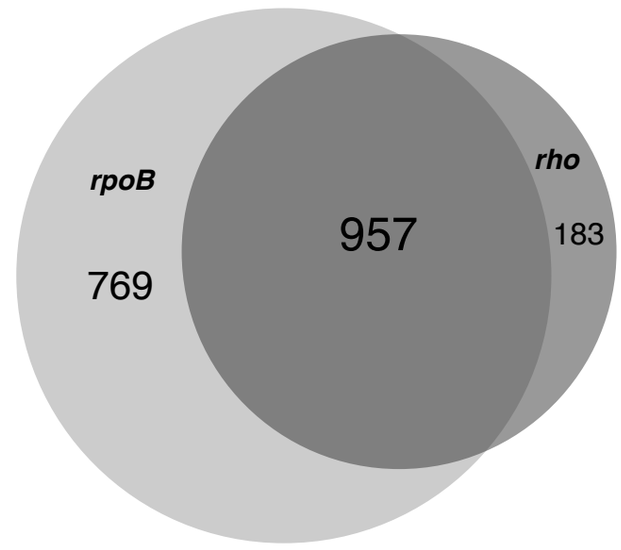
1063 Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, Landick R. 2012. Rho and NusG
1064 suppress pervasive antisense transcription in Escherichia coli. Genes Dev. 26:2621-
1065 2633.
1066 Peters JM, Mooney RA, Kuan PF, Rowland JL, Keles S, Landick R. 2009. Rho directs
1067 widespread termination of intragenic and stable RNA transcription. Proc Natl Acad
1068 Sci U S A. 106:15406-15411.
1069 Peters JM, Vangeloff AD, Landick R. 2011. Bacterial transcription terminators: the RNA
1070 3'-end chronicles. J Mol Biol. 412:793-813.
1071 Quandt EM, Gollihar J, Blount ZD, Ellington AD, Georgiou G, Barrick JE. 2015. Fine-
1072 tuning citrate synthase flux potentiates and refines metabolic innovation in the
1073 Lenski evolution experiment. Elife. 4
1074 R Core Team. 2013. R: A language and environment for statistical computing. R
1075 Foundation for Statistical Computing, Vienna, Austria. http://wwwR-projectorg/.
1076 Rodriguez-Verdugo A, Carrillo-Cisneros D, Gonzalez-Gonzalez A, Gaut BS, Bennett
1077 AF. 2014. Different tradeoffs result from alternate genetic adaptations to a common
1078 environment. Proc Natl Acad Sci U S A. 111:12121-12126.
1079 Rodriguez-Verdugo A, Gaut BS, Tenaillon O. 2013. Evolution of Escherichia coli
1080 rifampicin resistance in an antibiotic-free environment during thermal stress. BMC
1081 Evol Biol. 13:50.
1082 Rodríguez-Verdugo A, Tenaillon O, Gaut BS. 2016. First-Step Mutations during
1083 Adaptation Restore the Expression of Hundreds of Genes. Mol Biol Evol. 33:25-39.
1084 Romantsov T, Helbig S, Culham DE, Gill C, Stalker L, Wood JM. 2007. Cardiolipin
1085 promotes polar localization of osmosensory transporter ProP in Escherichia coli. Mol
1086 Microbiol. 64:1455-1465.
1087 Rosenblum EB, Parent CE, Brandt EE. 2014. The molecular basis of phenotypic
1088 convergence. Ann Review of Ecology, Evolution and Systematics. 45:203-226.
1089 Ryals J, Little R, Bremer H. 1982. Temperature dependence of RNA synthesis
1090 parameters in Escherichia coli. J Bacteriol. 151:879-887.
1091 Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial
1092 restraints. J Mol Biol. 234:779-815.
1093 Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, Palsson
1094 BO, Sommer M, Feist AM. 2014. Evolution of Escherichia coli to 42 degrees C and
1095 subsequent genetic engineering reveals adaptive mechanisms and novel mutations.
1096 Mol Biol Evol. 31:2647-2662.
1097 Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. 2005. The FoldX web
1098 server: an online force field. Nucleic Acids Res. 33:W382-8.
1099 Serohijos AW, Shakhnovich EI. 2014. Contribution of selection for protein folding
1100 stability in shaping the patterns of polymorphisms in coding regions. Mol Biol Evol.
1101 31:165-176.
1102 Skordalakes E, Berger JM. 2003. Structure of the Rho transcription terminator:
1103 mechanism of mRNA recognition and helicase loading. Cell. 114:135-146.
1104 Spencer CC, Bertrand M, Travisano M, Doebeli M. 2007. Adaptive diversification in
1105 genes that regulate resource use in Escherichia coli. PLoS Genet. 3:e15.
1106 Stern DL. 2013. The genetic causes of convergent evolution. Nat Rev Genet. 14:751-764.
1107 Tenaillon O, Barrick JE, Ribeck N et al. 2016. Tempo and mode of genome evolution in
1108 a 50,000-generation experiment. Nature. 536:165-170.

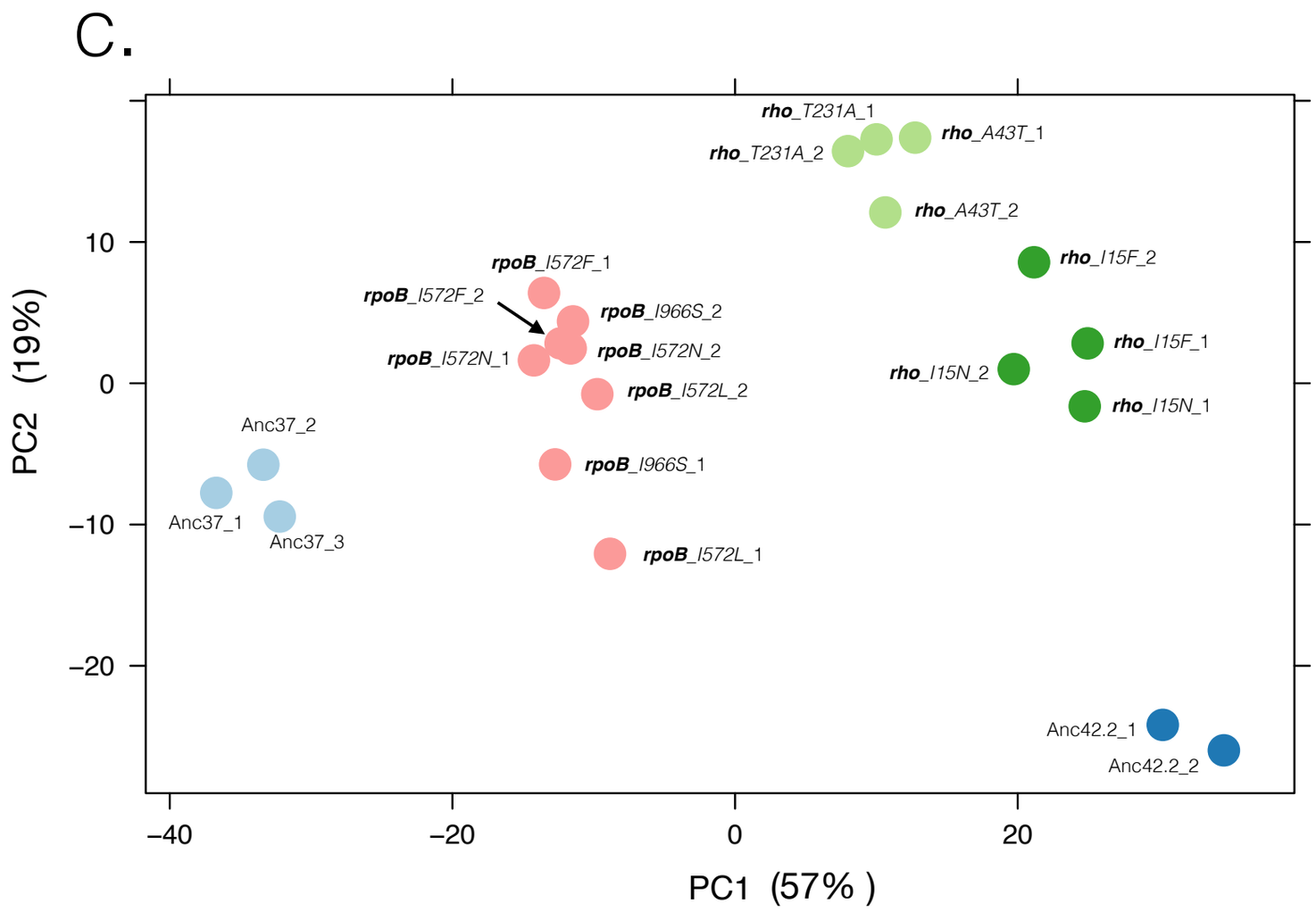
1109 Tenaillon O, Rodríguez-Verdugo A, Gaut… RL. 2012. The molecular diversity of
1110 adaptive convergence. Science.
1111 Thomsen ND, Berger JM. 2009. Running in reverse: the structural basis for translocation
1112 polarity in hexameric helicases. Cell. 139:523-534.
1113 Tokuriki N, Tawfik DS. 2009. Stability effects of mutations and protein evolvability.
1114 Curr Opin Struct Biol. 19:596-604.
1115 Utrilla J, O'Brien EJ, Chen K, McCloskey D, Cheung J, Wang H, Armenta-Medina D,
1116 Feist AM, Palsson BO. 2016. Global Rebalancing of Cellular Resources by
1117 Pleiotropic Point Mutations Illustrates a Multi-scale Mechanism of Adaptive
1118 Evolution. Cell Syst. 2:260-271.
1119 Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. 2005.
1120 GROMACS: fast, flexible, and free. J Comput Chem. 26:1701-1718.
1121 Vasi F, Travisano M, Lenski RE. 1994. Long-term experimental evolution in Escherichia
1122 coli. II. Changes in life history traits during adaptation to a seasonal environments.
1123 American Naturalist. 144:432-456.
1124 Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ. 1999. Different
1125 trajectories of parallel evolution during viral adaptation. Science. 285:422-424.
1126 Wiser MJ, Lenski RE. 2015. A Comparison of Methods to Measure Fitness in
1127 Escherichia coli. PLoS One. 10:e0126210.
1128 Wittkopp PJ, Kalay G. 2011. Cis-regulatory elements: molecular mechanisms and
1129 evolutionary processes underlying divergence. Nat Rev Genet. 13:59-69.
1130 Wylie CS, Shakhnovich EI. 2011. A biophysical protein folding model accounts for most
1131 mutational fitness effects in viruses. Proc Natl Acad Sci U S A. 108:9916-9921.
1132 Yamamoto K, Ishihama A. 2003. Two different modes of transcription repression of the
1133 Escherichia coli acetate operon by IclR. Mol Microbiol. 47:183-194.
1134 Zwietering MH, Jongenburger I, Rombouts FM, van 't Riet K. 1990. Modeling of the
1135 bacterial growth curve. Appl Environ Microbiol. 56:1875-1881.
1136

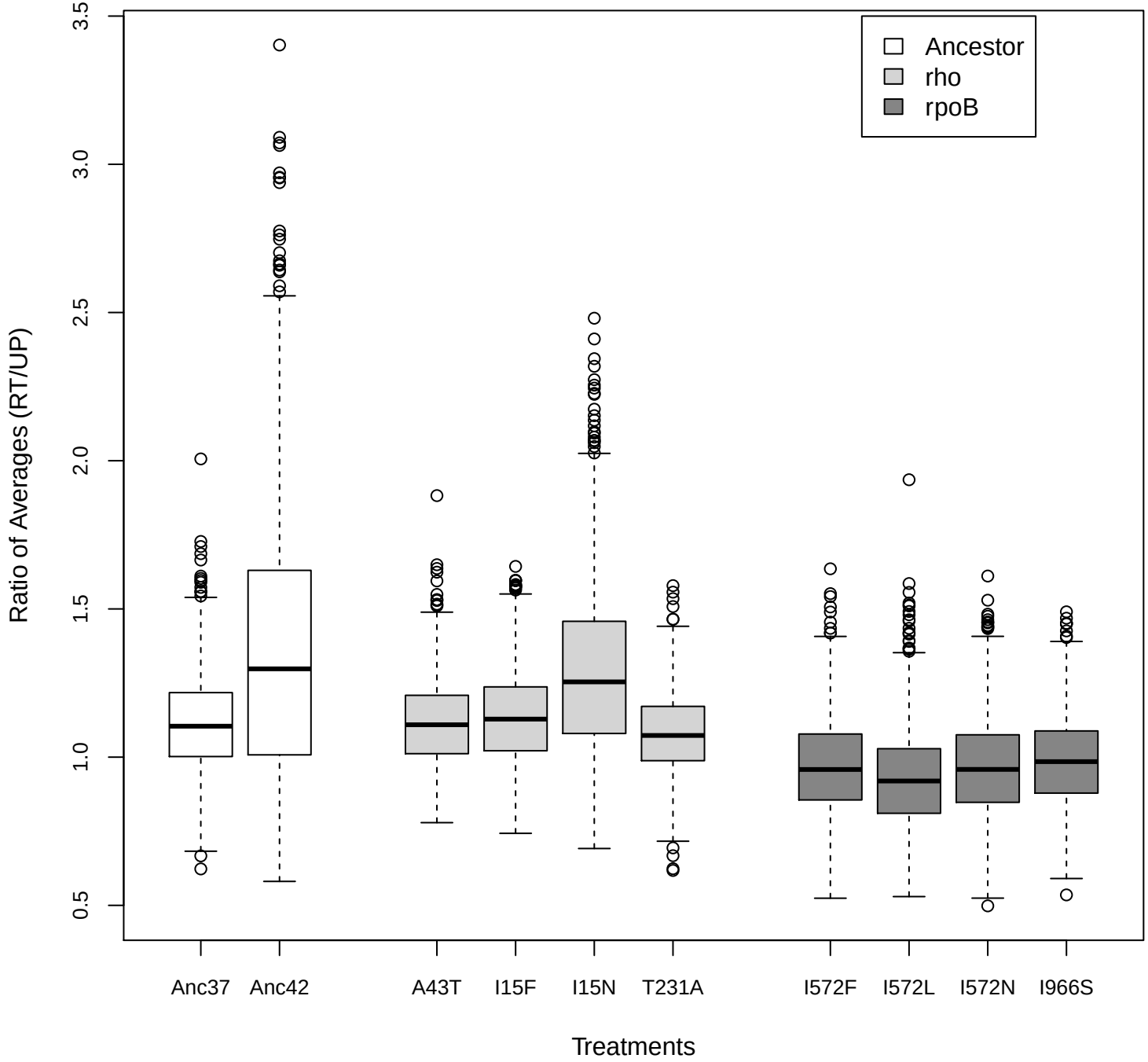

A
F
B
A43
I15
T231
E
C
D

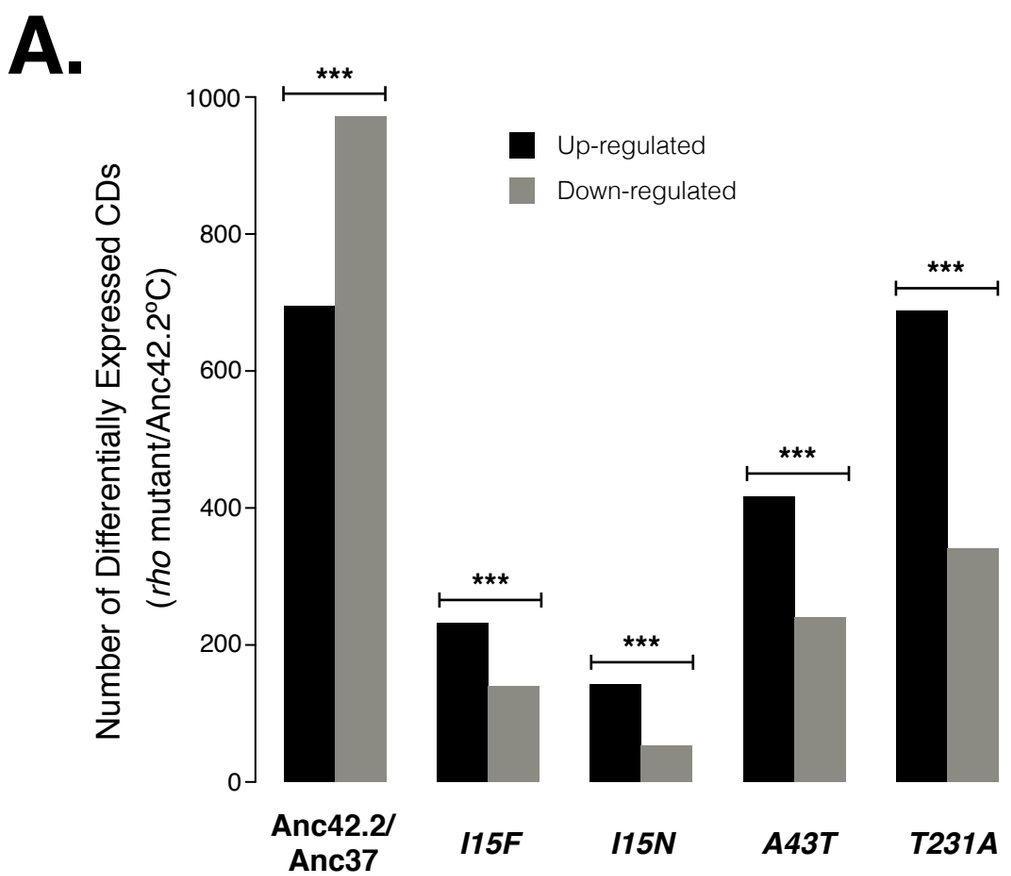

37°C
42°C
4e+07
3e+07
2e+07
1e+07
0e+00
Density (cell/ml)
2
3
4
5
6
7
8
9
10
11
12
13
Time (hours)
rpoB I572F
rpoB I572L
rpoB I572N
rho T231A
rho A43T
Ancestor
rho I15N
rho I15F
Mutation_Temperature
A_rpoB_I572F_42
B_rpoB_I572L_42
C_rpoB_I572N_42
D_Ancestor_42
E_rho_A43T_37
F_rho_I15F_37
G_rho_I15N_37
H_rho_T231A_37
I_rpoB_I572F_37
J_Ancestor_37
K_rpoB_I572L_37
L_rpoB_I572N_37
M_rho_A43T_42
N_rho_I15F_42
O_rho_I15N_42
P_rho_T231A_42

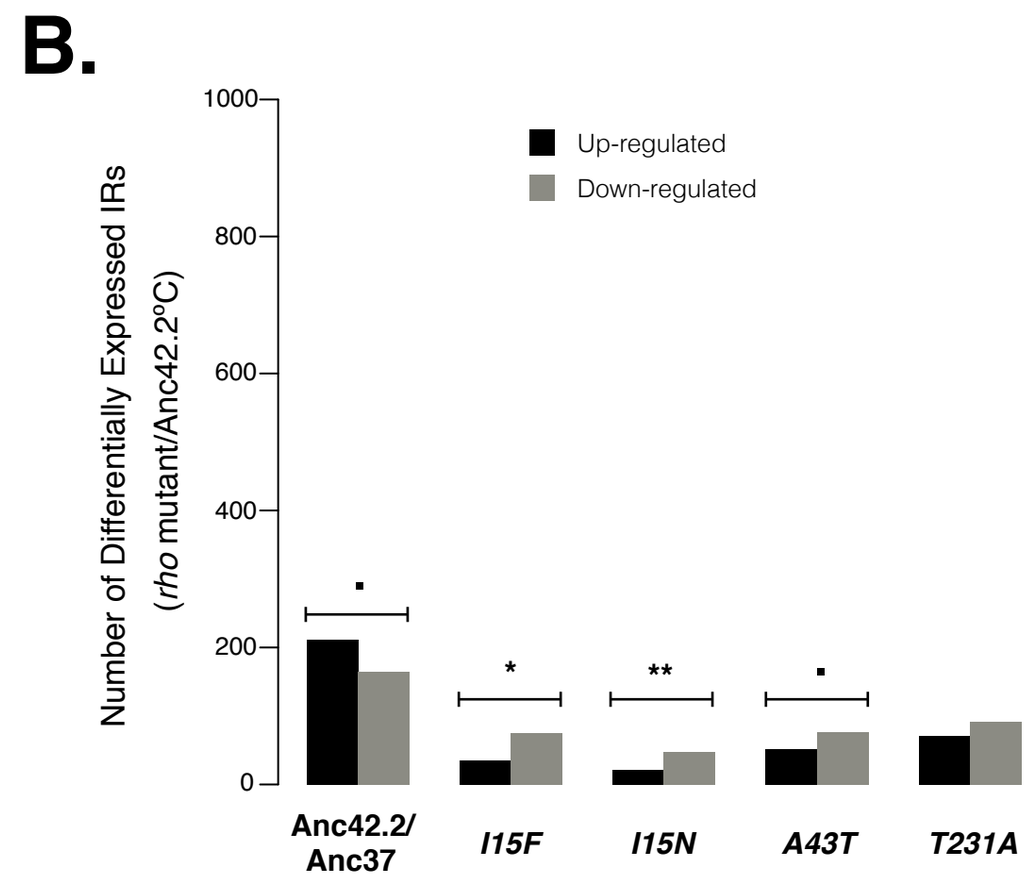

A.
I15N
A43T
I15F
T231A
7
36
18
(1.6%)
0
330
(28.9%)
47
2
20
373
1
141
(12.4%)
2
127
5
31

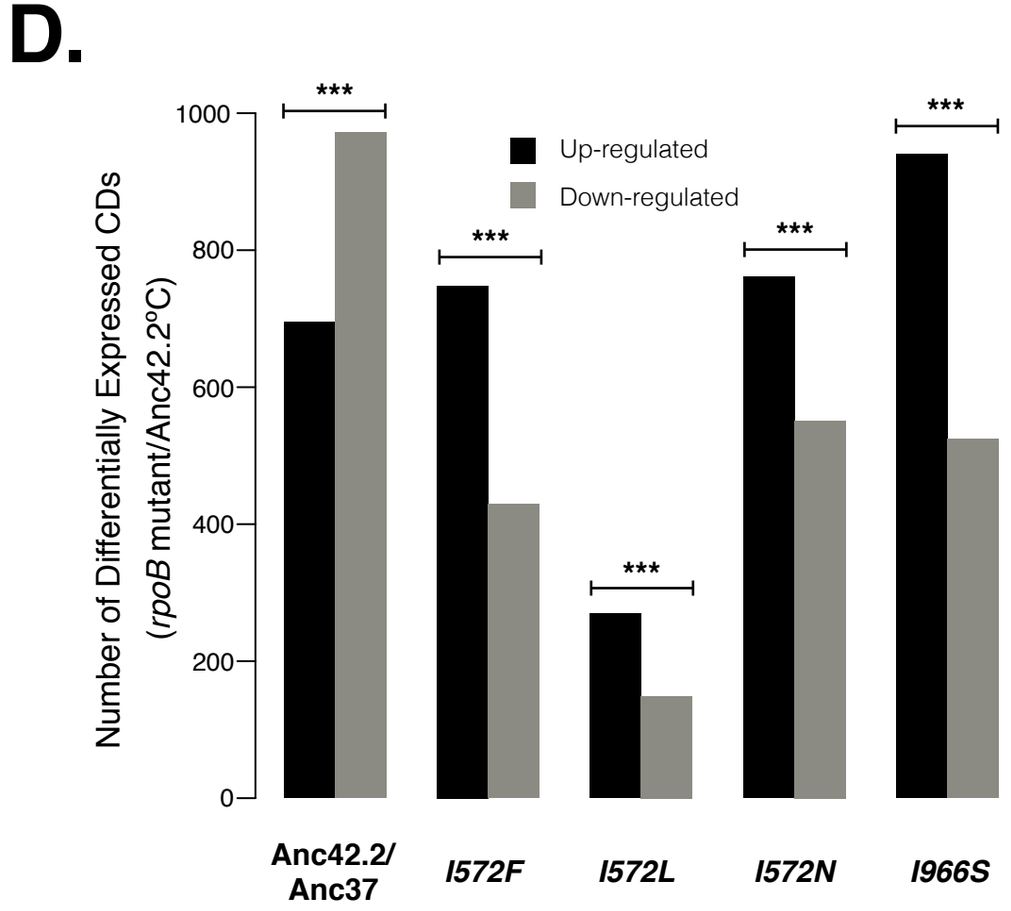

B.
rpoB
rho
957
769
183

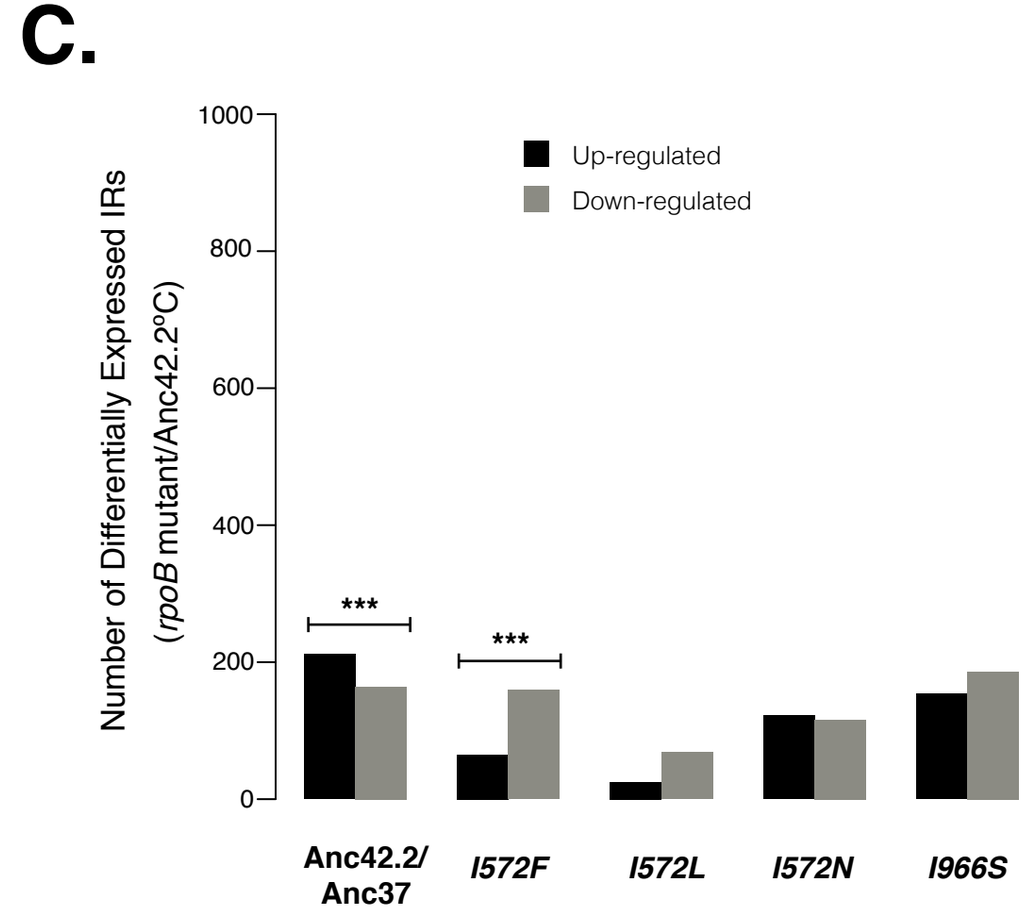

C.
rho_T231A_1
rho_A43T_1
rho_T231A_2
rho_A43T_2
rho_I15F_2
rho_I15F_1
rho_I15N_2
rho_I15N_1
rpoB_I572F_1
rpoB_I572F_2
rpoB_I966S_2
rpoB_I572N_1
rpoB_I572N_2
rpoB_I572L_2
rpoB_I966S_1
rpoB_I572L_1
Anc37_2
Anc37_1
Anc37_3
Anc42.2_1
Anc42.2_2
10
0
−10
−20
PC2 (19%)
−40
−20
0
20
PC1 (57%)


Ancestor
rho
rpoB
Ratio of Averages (RT/UP)
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Anc37
Anc42
A43T
I15F
I15N
T231A
I572F
I572L
I572N
I966S
Treatments


A.
Number of Differentially Expressed CDs
(rho mutant/Anc42.2°C)
1000
800
600
400
200
0
Up-regulated
Down-regulated
***
Anc42.2/
Anc37
I15F
I15N
A43T
T231A


B.
Number of Differentially Expressed IRs
(rho mutant/Anc42.2°C)
1000
800
600
400
200
0
Up-regulated
Down-regulated
*
**
Anc42.2/
Anc37
I15F
I15N
A43T
T231A


D.
Number of Differentially Expressed CDs
(rpoB mutant/Anc42.2°C)
1000
800
600
400
200
0
Up-regulated
Down-regulated
***
Anc42.2/
Anc37
I572F
I572L
I572N
I966S


C.
Number of Differentially Expressed IRs
(rpoB mutant/Anc42.2°C)
1000
800
600
400
200
0
Up-regulated
Down-regulated
***
Anc42.2/
Anc37
I572F
I572L
I572N
I966S