

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Genetic bases and phenotypic consequences of high-temperature adaptation in Escherichia coli

Permalink

<https://escholarship.org/uc/item/92w0x92z>

Author

Rodriguez Verdugo, Alejandra

Publication Date

2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Genetic bases and phenotypic consequences of high-temperature adaptation in
Escherichia coli

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Alejandra Rodriguez Verdugo

Dissertation Committee:
Professor Brandon S. Gaut, Chair
Professor Anthony Long
Associate Professor Adam Martiny

2014

Chapter 1 © 2013 BioMed Central
Chapter 3 © 2014 National Academy of Sciences
All other materials © 2014 Alejandra Rodriguez Verdugo

DEDICATION

To

my parents, Leticia and Victor, and my brother Francisco;

three great researchers that nurtured my passion for science and knowledge.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	vi
ACKNOWLEDGMENTS	viii
CURRICULUM VITAE	x
ABSTRACT OF THE DISSERTATION	xiv
INTRODUCTION	1
CHAPTER 1: Evolution of <i>Escherichia coli</i> rifampicin resistance in an antibiotic-free environment during thermal stress	9
CHAPTER 2: First-step mutations restore the ancestral expression state during thermal stress adaptation	44
CHAPTER 3: Different trade-offs result from alternate genetic adaptations to a common environment	100
CONCLUSIONS	145

LIST OF FIGURES

	Page
Figure 1.1 MIC distribution for the high temperature adapted clones in rifampicin.	30
Figure 1.2 Temporal dynamics of the rifampicin-resistant individuals in 12 evolved populations during 2000 generations.	31
Figure 1.3 Relationship between the parameters of fixation.	32
Figure 1.4 Relative fitness of the <i>rpoB</i> mutants measured at 42.2°C in DM25 in different genetic backgrounds.	33
Figure S1.1 Parameters of fixation estimated from the trajectories of the rifampicin resistant individuals from the populations evolved at high temperature.	36
Figure S1.2 Three-dimensional structure of RNAP generated using Jmol from the Protein Data Bank	36
Figure 2.1 Phenotypic characterization of the ancestor at 37°C and 42°C.	75
Figure 2.2 Phenotypic characterization of the mutants compared to the ancestor at 42°C.	76
Figure 2.3 Global changes in GE during the acclimation and adaptive response.	77
Figure 2.4 Convergence of genes with restored expression in single mutants.	78
Figure S2.1 Transcription efficiency of the mutants and the ancestor at 42°C.	80
Figure S2.2 Volcano plots showing the differential expression of genes for the pairwise comparisons between mutants and high-temperature adapted clones.	81
Figure S2.3 Convergence of genes with novel expression in single mutants.	81
Figure S2.4 Phenotypic changes during thermal stress adaptation (clone 97).	82
Figure S2.5 Global changes in GE during the acclimation and adaptive response.	82
Figure S2.6 Similarities in global changes in GE between the mutation <i>I572L</i> and two high-temperature adapted clones.	83
Figure 3.1 Hypothetical evolutionary responses of adaptation to high temperature.	120

Figure 3.2	Mean relative fitnesses of high-temperature evolved clones at 20°C and at 37°C.	120
Figure 3.3	Mean absolute fitnesses \bar{w}_a of the ancestor and each of 114 high-temperature adapted clones at 17°C, 18°C and 19°C.	121
Figure S3.1	Population density trajectories of the 114 high-temperature adapted clones and the ancestor at 17°C, 18°C, 19°C and 45°C.	128
Figure S3.2.	“Lazarus effect” observed in the ancestor at 45°C.	129
Figure S3.3.	Distribution of the number of high-temperature adapted clones sharing the same identical mutation.	129
Figure S3.4	Kernel density plot from the absolute fitness data at 18°C.	130
Figure S3.5	Population density trajectories of the high-temperature evolved clones with the (A) <i>rho</i> I15N mutation, (B) <i>rpoB</i> I966S mutation, (C) 1-bp deletion, (D) 23164 bp large deletion and (E) 71416 bp large deletion.	131
Figure S3.6	Structural mapping of the codons analyzed.	132

LIST OF TABLES

	Page
Table 1.1 Non-synonymous mutations in the <i>rpoB</i> gene conferring rifampicin resistance.	35
Table 1.2 Two-way analysis of variance for relative fitness of mutants in three different genetic backgrounds.	35
Table 1.3 Two-way analysis of variance for relative fitness of mutants in five different environments with the environment and the genotype (mutants) treated as fixed effects, and a mixed effect model with the genotype treated as random and the temperature and glucose treated as fixed effects.	35
Table S1.1 Oligonucleotides and primers used in this study.	37
Table S1.2 Parameters of fixation estimated from the frequency trajectories.	37
Table S1.3 Relative fitness measured in different conditions and different genetic backgrounds.	38
Table 2.1 Growth parameters of the first-step mutants and the ancestor.	79
Table 2.2 Classification of the genes into four patterns of expression change.	79
Table 2.3 Growth parameters of the high-temperature adapted clones 27 and 97 at 42°C compare to the mutant <i>I572L</i> and the ancestor at 37°C and 42°C.	79
Table S2.1 Families of highly differentially expressed genes (and their GO significant enrichment categories) during the acclimation response.	85
Table S2.2 List of: <i>i</i>) previously reported up-regulated genes during the heat stress response, <i>ii</i>) previously reported up-regulated genes during the general stress response and, <i>iii</i>) genes encoding different subunits of RNAP.	86
Table S2.3 Criteria used to classify the genes differentially expressed during the acclimation and/or adaptive response.	89
Table S2.4 Mutations and GE changes of the high temperature adapted clone 27.	89
Table S2.5 Mutations and GE changes of the high temperature adapted clone 97.	93
Table S2.6 Primers used in this study.	93

Table 3.1	The five most common mutations and their associated effects on absolute fitness at 18.0°C.	122
Table 3.2	Fitness estimates for single mutations.	122
Table S3.1	Statistical analyses summary of the relative fitnesses of the 114 high-temperature adapted clones at 20°C, 37°C (this study) and 42.2°.	133
Table S3.2	Statistical analyses summary of the absolute fitnesses of the ancestor and the 114 high-temperature adapted clones at 17°C, 18°C, 19°C and 45°C.	136

ACKNOWLEDGMENTS

First of all, I would like to thank my fabulous advisor Dr. Brandon S. Gaut for his unconditional support throughout my PhD. Thank you for teaching me to think critically and to write properly. Thank you for contributing with brilliant ideas and invaluable insights. Thank you for trusting and believing in me. I learned from you values such as honesty, patience and hard work that not only made me a better scientist, but also made me a better person. It was an honor to have you as mentor, and it will be an honor to have you as a collaborator.

I would like to extend my sincere gratitude to Dr. Olivier Tenaillon, for all his help during my dissertation. Thank you for making me a better experimentalist and for teaching new techniques essential to my research. Thank you for sharing your passion for science and for inspiring me with fantastic ideas. Thanks to you my dissertation substantially improved.

I am really grateful to my committee members, Dr. Adam Martiny and Dr. Anthony Long for providing insightful feedback on my dissertations research. Thank you for volunteering some of your time attending committee meetings, sharing ideas and reading my dissertation. Adam, I would like to thank you one more time for hooding me at the graduation ceremony; it meant a lot to me.

This dissertation would not have been possible without the support of two wonderful technicians, Rebecca Gaut and Pamela McDonald, who help me with all the experiments. Thank you for your technical and moral support and thank you for spending an entire year (including week-ends) evolving 115 replicates lines of *Escherichia coli*.

I would like to give a special thank to David Carrillo for being so generous with his time and advising me on the statistical tests implemented in my research.

Also, I would like to thank all the members of the U722 research group in France. In particular, I would like to thank Dr. Erick Denamour and Dr. Mathilde Lescat for making my visit possible and for receiving me with welcoming arms. Finally, I would like to thank my “lab soulmates”, Florence and André, for making my nine months of research in Paris a one of the most fantastic experiences in my professional career.

This dissertation is also the result of a fruitfully interaction with former and current members of the Gaut lab. Thank you Andrea Gonzalez, Conception Muñoz, Kyria Roessler and Shaun Hug for providing me kind support and helpful feedback in the last couple of years.

I am extremely lucky to have had the opportunity to interact with talented graduate students from the Eco Evo department. I would like to thank Jenny Talbot, China Hanson, Cecillia Batmalle, Sandra Holden and Steve Hatosy from the Microbial group for providing me useful feedback. I am especially thankful to my friends, Adriana

Romero, Aide Macias, Jennifer Weber, Luis Abdala, Kyle McCulloch and Will Petry, for our weekly discussions of Science during Spanish Night.

In addition I would like to thank my dear UCI friends, Adrienne and Pawel for all their love and support. I would also like to thank Moni, Pale, Mari, Anke, Dani, Firo, Ado, Esteban, Eero, Vili, Ricardo and Rolf for being my second family and for making me feel at home. Lastly, I would like to thank my family and friend in Mexico for their long distance support.

Finally, I would like to acknowledge the funding sources that made my dissertation research possible. Fellowship support was provided by the University of California Institute for Mexico and the United States – Consejo Nacional de Ciencias y Tecnología (Mexico), UC MEXUS-CONACYT fellowship that covered tuitions, fees and stipend during five years of graduate studies. Additional fellowship support was provided by the UCI Graduate Division through the Miguel Velez Scholarship and the Dr. William F. Holcomb Scholarship. The nine months of research in Paris were possible thanks to the Chateaubriand fellowship offered by the Office for Science and Technology of the Embassy of France in the United States.

Chapter 1 was published with the permission of BioMed Central. The text of this chapter is a reprint of the material as it appears in BMC Evolutionary Biology. Chapter 3 was published with the permission of the National Academy of Sciences. The text of this chapter is a reprint of the material as it appears in Proceedings of the National Academy of Sciences.

CURRICULUM VITAE

Alejandra Rodriguez Verdugo

Department of Ecology and Evolutionary Biology
University of California, Irvine
321 Steinhaus Hall
Irvine, CA 92697-2525
Phone (949) 824-2963
e-mail alejanr1@uci.edu

EDUCATION

- 2009-2014 **Ph.D. in Biological Sciences**, University of California, Irvine
Dissertation topic: Genetic bases and phenotypic consequences of high-temperature adaptation in *Escherichia coli*
Advisor: Brandon S. Gaut
- 2003-2008 **Bachelor of Science in Biology**, National Autonomous University of Mexico (UNAM)
Thesis topic: Seasonal variation in *Pseudomonas* diversity within a fluctuating aquatic system
Advisor: Ana E. Escalante

RESEARCH EXPERIENCE

- 2012 **Visiting Researcher** with Dr. Olivier Tenaillon at the Institut National de la Santé et de la Recherche Médicale (National Institute of Health and Medical Research) INSERM - UMR-S U722 , Paris, France.

PUBLICATIONS

Peer-reviewed Articles and Chapters

- Rodríguez-Verdugo, A.**, D. Carrillo-Cisneros, A. Gonzalez-Gonzalez, B. S. Gaut and A. F. Bennett. 2014. Different trade-offs result from alternate genetic adaptations to a common environment. *Proceedings of the National Academy of Sciences of the United States of America*. 111: 12121-12126.
- Achaz G., **A. Rodríguez-Verdugo**, B.S. Gaut, and O. Tenaillon. 2014. The reproducibility of adaptation in the light of experimental evolution with whole genome sequencing. In: Landy C. R. and N. Aubin-Horth, editors. *Ecological Genomics: Ecology and the Evolution of Genes and Genomes*. Netherlands: Springer. p211-231. doi: 10.1007/978-94-007-7347-9_11.
- Rodríguez-Verdugo A.**, B.S. Gaut, and O. Tenaillon. 2013. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evolutionary Biology* 13:50. doi:10.1186/1471-2148-13-50

Rodríguez-Verdugo A., V. Souza, L.E. Eguiarte, and A.E. Escalante. 2012. Diversity across seasons of culturable *Pseudomonas* from a dessication lagoon in Cuatro Ciénegas, Mexico. *International Journal of Microbiology*. doi:10.1155/2012/201389

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

Cerritos R., L.E. Eguiarte, M. Avitia, J. Siefert, M. Travisano, **A. Rodríguez-Verdugo**, and V. Souza. 2010. Diversity of culturable thermo-resistant aquatic bacteria along an environmental gradient in Cuatro Ciénegas, Coahuila, México. *Antonie van Leeuwenhoek* 99: 303-318.

Escalante A.E., J. Caballero-Mellado, L. Martínez-Aguila, **A. Rodríguez-Verdugo**, A. González-González, J. Toribio-Jiménez, and V. Souza. 2009. *Pseudomonas cuatrocienegasensis* sp. nov., isolated from an evaporating lagoon in the Cuatro Ciénegas valley in Coahuila, México. *International Journal of Systematic and Evolutionary Microbiology (IJSEM)* 59: 1416-1420.

Publications in review or preparation

Rodríguez-Verdugo, A., O. Tenaillon, and B.S. Gaut. (*in preparation*) First-step mutations restore the ancestor's expression state during thermal stress adaptation.

FEATURED RESEARCH

"Predicting the evolution of antibiotic resistance" Schenk M.F. and A.GM. de Visser. *BMC Biology* 2013, 11:14

"Stressed bacteria become resistant to antibiotics" Feb 22, 2013
<http://www.sciencedaily.com/releases/2013/02/130221194045.htm>

Evaluation of "The molecular diversity of adaptive convergence" paper: Yuri Wolf and Natalya Yutin, Faculty of 1000 Biology, Feb 15, 2012, <http://f1000.com/13700958>

FELLOWSHIPS AND AWARDS

2014 Dr. William F. Holcomb Scholarship (\$2000)
AGS Travel Grant (\$400) and ASM Student Travel Grant (\$500)

2013 Miguel Velez Scholarship (\$10,000)

2012 Chateaubriand Fellowship (€12,600)

2009-2014 UC MEXUS-CONACYT Doctoral Fellowships for Mexican Students
(tuition, fees and \$65,400 total awarded over 5 years)

INVITED PRESENTATIONS

“Opportunities for American Students to come in France for a research experience: Chateaubriand fellowship.” French-American Workshop, MINATEC, Grenoble, France, June 2012.

ORAL PRESENTATIONS

“The genetic bases of fitness trade-offs in *Escherichia coli*.” 1st American Society of Microbiology (ASM) Conference on Experimental Microbial Evolution, Washington, DC, United States, June 2014.

“The genetic bases of fitness trade-offs in *Escherichia coli*.” Society of Molecular Biology and Evolution (SMBE) meeting, San Juan, Puerto Rico, June 2014.

“Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress adaptation.” Winter Ecology and Evolutionary Biology Graduate Student Symposium, UCI, Irvine CA, United States, February 2013.

“Emergence and evolution of antibiotic resistance as a collateral response to thermal stress.” Winter Ecology and Evolutionary Biology Graduate Student Symposium, UCI, Irvine CA, United States, January 2011.

POSTER PRESENTATIONS

“Fitness trade-offs and reduction of the lower thermal limit after thermal stress adaptation of *Escherichia coli*.” Society for Molecular Biology and Evolution (SMBE) meeting, Chicago IL, United States, July 2013.

“Parallel evolution of *rpoB* mutants of *Escherichia coli* during thermal stress adaptation.” Society for Molecular Biology and Evolution (SMBE) meeting, Dublin, Ireland, June 2012.

TEACHING EXPERIENCE

2014	Environmental Microbiology, University of California Irvine, <i>Teaching assistant</i>
2014	BioSci 94: From Organisms to Ecosystems, University of California Irvine, <i>discussion section leader (x3)</i>
2013	BioSci 94: From Organisms to Ecosystems, University of California Irvine, <i>discussion section leader (x3)</i>
2011	BioSci 94: From Organisms to Ecosystems, University of California Irvine, <i>discussion section leader (x3)</i>
2008	Biology of Prokaryotes, College of Science, National Autonomous University of Mexico (UNAM), Mexico <i>Laboratory teacher</i>
2007	Biology of Prokaryotes, College of Science, National Autonomous University of Mexico (UNAM), Mexico <i>Laboratory teacher</i>

ACADEMIC SERVICE

- 2014 Organizer of the 2014 Winter Ecology and Evolutionary Biology Graduate Student Symposium
- 2011 Graduate Student Representative, UCI EEB graduate admissions committee.

PROFESSIONLA SOCIETIES

- 2014 – present American Society for Microbiology
- 2012 – present Society for Molecular Biology and Evolution

ABSTRACT OF THE DISSERTATION

Genetic bases and phenotypic consequences of high-temperature adaptation in
Escherichia coli

By

Alejandra Rodriguez Verdugo

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2014

Professor Brandon S. Gaut, Chair

Connecting phenotype to genotype and fitness has been a major challenge when studying adaptation. Experimental evolution is a powerful method to facilitate the study of adaptation as an outcome and as a process. This dissertation used 114 clones of *Escherichia coli* that were evolved independently for 2000 generations under thermal stress (42.2°C). The goal of my dissertation was to link phenotypes and fitness with specific mutations and to identify the functional mechanisms leading to thermal adaptation.

Chapter 1 focused on a subset of populations that became resistant to an antibiotic, rifampicin, during thermal stress adaptation. Rifampicin resistance was caused by three mutations at codon position 572 of the *rpoB* gene, which encodes the beta subunit of RNA polymerase (RNAP). I used samples from 200 generation intervals to assess the frequency trajectory of rifampicin resistance. I found that resistant mutations typically appeared, and were fixed, early in the evolution experiment. Finally, I

confirmed that the three mutations conferred high advantage in glucose-limited medium at 42°C in the ancestral background.

Chapter 2 investigated the same *rpoB* mutations in codon 572 to explore the molecular mechanisms underlying their fitness improvement at high temperature. I measured their growth curves and gene expression (mRNAseq) at 42°C, and compared them to the growth and gene expression of the ancestor at 37°C and 42°C. Two *rpoB* mutations restored the gene expression back to an ancestral state, while one mutation, exaggerated the expression changes of the ancestor at 42°C. Lastly, I compared the phenotypic characteristics of one single mutant, *I572L*, to those of two high-temperature adapted clones with this mutation. I concluded that the *I572L* mutation contributed to most of the expression changes while later mutations did not substantially changed gene expression.

Chapter 3 explored the phenotypic consequences of high temperature adaptation in 114 clones. I measured the magnitude of fitness trade-offs across a thermal gradient. I identified two niche dynamics that I associated with two genetic pathways. Overall, my dissertation associates mutations to phenotypes in the context of thermal stress adaptation, and it highlights the difficulties of connecting genotype to phenotype, even if in a “simple” experimental system.

INTRODUCTION

Adaptation – the movement of a population towards a phenotype that best fits its present environment – is central to evolutionary change (Orr 2005). Yet, little is known about the diversity of mutations that contribute to adaptation and even less is known about their phenotypic and fitness effects. This inability to connect phenotype to genotype and fitness is a major failing of evolutionary biology (Barrett and Hoekstra 2011).

Historically, evolutionary biologists have compared living organisms to test hypotheses of adaptation. For example, if traits in different lineages correlate with a similar environment or selective pressure, evolutionary biologists often infer that there has been convergent evolution, which is then interpreted as evidence for adaptation. Occasionally, inferences based on comparative approaches are complemented by experimental efforts to test a hypothesized adaptive allele in the field. These two approaches – comparative and experimental – have been applied to studies on butterflies, deer mice and stickleback fish (Barrett and Hoekstra 2011). Unfortunately, these are rare examples linking genotype, phenotype, and adaptation in nature.

Another approach to studying adaptation is experimental evolution (Garland and Rose 2009), which has various advantages over comparative studies. In an experimental setting populations can be evolved for many generations with sufficient controls and replicates to make robust conclusions about the mechanisms and targets of selection. Moreover, evolution experiments allow us to observe intermediate steps of adaptation and to follow the frequency of new mutations through time (Lang et al. 2013).

Therefore experimental evolution facilitates the study of adaptation not only as an outcome but also as a process (Barrick and Lenski 2013).

In evolution experiments, bacteria are often the model of choice (Elena and Lenski 2003), which have at least three advantages. First, their large population sizes and short generations times allow the detection of evolutionary changes in a reasonable time scale. Second, bacteria can be frozen and later revived, permitting direct comparison between the ancestral and evolved types. Finally, as bacteria have relatively small genome sizes, it has become possible to sequence complete genomes with next-generation sequencing. By using bacterial model systems, such as *E. coli*, evolutionary experiments have the capability to directly link genetic change to its phenotypic and fitness consequences.

Experimental evolution in *E.coli*

E.coli is a model organism widely used in experimental evolution. One of the most famous evolution experiments using *E.coli* is The Long Term Experimental Evolution study (LTEE) started by Richard Lenski in 1988 at UC Irvine (Lenski et al. 1991). In this experiment 12 replicate populations, derived from a single ancestral genotype, were grown in a low glucose environment. The population was propagated through serial transfers, meaning that a proportion of the population was periodically transferred to fresh media and allowed to regrow until the glucose was exhausted from the media. This process was repeated daily. The 12 populations from the LTEE have been evolved for more than 60,000 generations and have provided crucial insights about adaptation.

Soon after the LTEE began, Bennett et al. used a similar experimental design and evolved *E. coli* at high temperature (~42°C; Bennett et al. 1990). Temperature is a crucial environmental variable because it governs the rates of biological reactions that underlie activities such as respiration, growth, and reproduction. Bennett et al. observed a rapid adaptive response of *E. coli* to high temperature (Bennett et al. 1992). After this finding, Riehle et al. (2001) characterized the genetic changes that mediate this adaptation (Riehle et al. 2001). They showed that gene duplication and deletions occurred, either as a correlate or a cause of adaptation. Unfortunately, however, they could not identify all of the mutations because complete genome sequencing was not feasible at the time.

Our large scale *E. coli* thermal stress experiment

In order to build on the work of Riehle et al. and to study the dynamic and genetic bases of adaptation to high temperature, we evolved 114 replicate populations of *E. coli* for 2000 generations at 42°C (Tenailon et al. 2012). The primary purpose of the project was to describe the genetic diversity of an adaptive response. The level of replication of this experiment is one order of magnitude higher than any other experimental study of bacterial adaptation, permitting inference about the reproducibility and novelty of adaptation to a complex environment.

Briefly, we began with a strain of Arabinose (Ara-) *E. coli* B (REL1206) and inoculated 114 independent cultures in 10 ml of Davis minimal medium supplemented with 25 mg/l glucose (DM25). The ancestor (REL1206) had evolved at 37°C for 2000 generations in DM25 prior to the initiation of the experiment and therefore was well

adapted to laboratory conditions (Lenski et al. 1991). Because the ancestor lacks any plasmids, all the genetic variation arose by *de novo* mutation.

The 114 populations were propagated daily at 42.2°C by transferring 0.1ml of each culture into 9.9 ml of DM25. The bacteria grew to quiescent stationary phase daily, representing $\log_2(100) = 6.64$ generation of binary fission. Each transfer included a minimum of ~1.2 million cells, and thus population were maintained at a large size. Population samples were saved at 100 generations and every 200 generations thereafter.

After 2000 generations, which corresponds roughly to a year in the lab, we measured the relative fitness of a single isolate clone from each population at 42.2°C. Fitness measures were based on competitions between each of the evolved lines against a newly-derived Ara⁺ mutant of REL1206. Relative fitness was estimated by comparing the ratio of growth rates between each line relative to the ancestor (Lenski et al. 1991). Fitness increased markedly in all experimental lines, to a mean of 1.42-fold (± 0.024 95% C.L.) relative the ancestral clone (Tenaillon et al. 2012).

We also sequenced the complete genome of one clone from each line, at an average coverage of 90x, ultimately identifying 1258 mutations relative to the ancestor (or ~11 mutations per line on average). With this unprecedented dataset of genetic changes – including single nucleotides mutations, insertions, deletions and duplication – we identified a list of candidate genes and biochemical pathways involved in the adaptation to high temperature (Tenaillon et al. 2012).

Briefly, we identified two “adaptive pathways” that were strongly negatively associated (Tenaillon et al. 2012). The first and most common pathway included

mutations in the RNA polymerase (RNAP) β subunit (*rpoB*) gene, along with associated changes in RNAP subunit genes (*rpoA*, *rpoC*, and *rpoD*) and the six *rod* genes that affect cell shape. The second adaptive pathway included mutations in the RNAP termination factor *rho*, which were positively associated with knockouts of the cardiolipin synthase (*c/s*) genes and the transcription factor gene *iclR* (Tenaillon et al. 2012). However, we still do not know how this genetic diversity translates to phenotypic diversity. What are the fitness effects of the numerous *rpoB* mutations? What are the molecular mechanisms underlying high-temperature adaptation? Do the two pathways defined by *rho* and *rpoB* lead to convergent or divergent phenotypes?

The main goal of my dissertation has been to link changes in phenotypes and fitness with specific mutations and ultimately to relate genotype to phenotype and fitness.

The first part of my dissertation focuses on the most common adaptive pathway, that is, the pathway enriched with *rpoB* mutations. I have focused on the *rpoB* gene because it exhibited interesting molecular patterns during our evolution experiment. First, it was the most mutated gene in our study, with 86 total mutations. Second, although there was little convergence at the level of single base mutations in the high-temperature adapted lines (Tenaillon et al. 2012), some *rpoB* mutations stood out. For example, one mutation – an isoleucine to leucine in codon 966 (*I966S*) – was shared among 15 different populations, suggesting that this mutation conferred a high fitness advantage in the conditions of our experiment. Third, *rpoB* was the only gene that accumulated more than one mutation within a single line (Tenaillon et al. 2012), suggesting that *rpoB* mutations might have pleiotropic effects. These observations,

coupled with the central importance of the *rpoB* gene in transcription and regulation, as part of the RNAP, has made it an object of interest for geneticists, biochemists, structural biologists and evolutionary biologists.

In Chapter 1, I characterize a subset of *rpoB* mutations located in the active site of RNAP (codon 572) and conferring rifampicin resistance. I used the rifampicin-resistance as a marker to follow their frequency trajectory within populations during our high-temperature adaptation experiment. I also reconstruct single mutants to evaluate their fitness effects in the ancestral background (REL1206). Finally, I test their fitness effects in different genetic backgrounds and different environmental conditions. In Chapter 2, I use the same subset of *rpoB* mutation in codon 572 to explore the molecular mechanisms of thermal adaptation. To characterize the phenotypic properties of the *rpoB* mutants, I measure phenotypes in two ways: growth curves and gene expression (mRNAseq). Finally, to assess the phenotypic changes through adaptation, I compare the phenotypic characteristics of one single mutant, *I572L*, to the phenotypic characteristics of two high-temperature adapted clones with this mutation.

The second part of my dissertation focuses on the phenotypic characteristics of the *rho* and the *rpoB* adaptive pathways, with the purpose of knowing if the two genetic pathways converge on the same phenotype. I address this issue by measuring a complex phenotype: the magnitude of fitness trade-offs across a thermal gradient. I measure both relative and absolute fitness over a range of temperatures. I identify two niche dynamics that I subsequently associate with different genotypes. Overall, my dissertation addresses questions that will have lasting impacts on evolutionary theory and our current views of the adaptive capabilities of organisms in altered climates.

Hopefully, it also brings us step closer to the ambitious goal of connecting genotype, phenotype and fitness.

REFERENCES

- Barrett, R.D.H., and H.E. Hoekstra. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet* 12:767-780.
- Barrick, J.E., and R.E. Lenski. 2013. Genome dynamics during experimental evolution. *Nat Rev Genet* 14:827-839.
- Bennett, A.F., K.M. Dao, and R.E. Lenski. 1990. Rapid evolution in response to high-temperature selection. *Nature* 346:79-81.
- Bennett, A.F., R.E. Lenski, and J.E. Mittler. 1992. Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution* 16-30.
- Elena, S.F., and R.E. Lenski. 2003. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nat Rev Genet* 4:457-469.
- Garland, T., and M.R. Rose. 2009. *Experimental evolution: concepts, methods, and applications of selection experiments*. University of California Press Berkeley, CA, USA.
- Lang, G.I., D.P. Rice, M.J. Hickman, E. Sodergren, G.M. Weinstock, D. Botstein, and M.M. Desai. 2013. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* 500:571-574.
- Lenski, R.E., M.R. Rose, S.C. Simpson, and S.C. Tadler. 1991. Long-term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2,000 generations. *Am Nat* 138:1315-1341.
- Orr, H.A. 2005. The genetic theory of adaptation: A brief history. *Nat Rev Genet* 6:119-127.
- Riehle, M.M., A.F. Bennett, and A.D. Long. 2001. Genetic architecture of thermal adaptation in *Escherichia coli*. *Proc Natl Acad Sci USA* 98:525-530.
- Tenaillon, O., A. Rodriguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennett, A.D. Long, and B.S. Gaut. 2012. The Molecular Diversity of Adaptive Convergence. *Science* 335:457-461.

CHAPTER 1

Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress

ABSTRACT

Beneficial mutations play an essential role in bacterial adaptation, yet little is known about their fitness effects across genetic backgrounds and environments. One prominent example of bacterial adaptation is antibiotic resistance. Until recently, the paradigm has been that antibiotic resistance is selected by the presence of antibiotics because resistant mutations are associated with fitness costs in antibiotic free environments. In this study we show that it is not always the case, documenting the selection and fixation of resistant mutations in populations of *Escherichia coli* B that had never been exposed to antibiotics but instead evolved for 2000 generations at high temperature (42.2°C). We found parallel mutations within the *rpoB* gene encoding the beta subunit of RNA polymerase. These amino acid substitutions conferred different levels of rifampicin resistance. The resistant mutations typically appeared, and were fixed, early in the evolution experiment. We confirmed the high advantage of these mutations at 42.2°C in glucose-limited medium. However, the *rpoB* mutations had different fitness effects across three genetic backgrounds and six environments. We describe resistance mutations that are not necessarily costly in the absence of antibiotics or compensatory mutations but are highly beneficial at high temperature and low glucose. Their fitness effects depend on the environment and the genetic background, providing glimpses into the prevalence of epistasis and pleiotropy.

INTRODUCTION

Mutations supply the genetic variation for adaptation, but their success depends on at least three different factors (Orr 2005). The first is the selective coefficient (s), which influences both the probability of fixation of a mutation and its frequency trajectory. Highly advantageous mutations have a higher probability of escaping loss by genetic drift and are also expected to reach high frequency more rapidly than mutations of smaller beneficial effect (Kimura 1983). The second factor is the genetic background; the effect of a mutation may change in amplitude or even in sign (shifting, for example, from beneficial to neutral or even deleterious) across genetic backgrounds due to epistatic effects (Schrag et al. 1997; Weinreich et al. 2005; Gros et al. 2009; Khan et al. 2011). Epistatic interactions may also limit the emergence and propagation of further beneficial mutations, thereby affecting long-term chances of survival (Woods et al. 2011). Finally, mutations may have differential fitness effects across environments (e.g. Remold and Lenski 2001; Ostrowski et al. 2005; Bataillon et al. 2011), including the possibility of genotype-by-environment (GXE) interactions. These differential effects may have a profound influence on the pattern of adaptation, because they may prevent a mutation from fixing across heterogeneous environments, which in turn leads to niche (or ecological) specialization (MacLean et al. 2004).

Antibiotic resistance is a particularly important class of beneficial mutation, both because of its potential implications for public health (Taubes 2008) and because resistance is easily studied in the laboratory, particularly in model systems like *Escherichia coli* (Andersson and Levin 1999; MacLean et al. 2010). Genetic resistance to antibiotics can result either from sequential accumulation of multiple beneficial

mutations – e.g. resistance to fluoroquinolones (Marcusson et al. 2009) – or from a single amino acid substitution – e.g. resistance to rifamycins (Tupin et al. 2010). This last mutational type is typically highly advantageous in the presence of antibiotics, leading to rapid fixation, often within hundreds of generations (Comas et al. 2012).

Despite their advantage in the presence of antibiotics, resistance mutations are usually deleterious in the absence of antibiotics, because they often modify vital cellular functions and are highly pleiotropic. For example, amino acid substitutions in the β subunit of the RNA polymerase (RNAP) that produce resistance to rifampicin (Campbell et al. 2001) diminish the transcription efficiency of RNAP and often entail a fitness cost in the absence of rifampicin (Reynolds 2000; Brandis et al. 2012). However, antibiotic resistance may not always be associated with fitness costs; resistance mutations sometimes appear to be neutral or even beneficial in the absence of antibiotics (Bataillon et al. 2011; Kassen and Bataillon 2006; Trindade et al. 2012), but for these cases the possibility of secondary advantageous mutations have not been precluded. Costly resistance mutations may also lead to the rapid selection of compensatory mutations that diminish or cancel the cost of resistance (Comas et al. 2012; Brandis et al. 2012; Levin et al. 2000; Hall and MacLean 2011). Nonetheless, the paradigm remains that antibiotic resistance is typically selected by the presence of the antibiotic and costly in its absence.

Here we characterize a series of mutations that confer antibiotic resistance but appeared as beneficial mutations in the absence of antibiotics. These mutations arose in the context of an experiment to adapt 114 lines of *E. coli* to thermal stress for 2000 generations (Tenaillon et al. 2012). At the end of the experiment, we identified rifampicin

resistant clones. Surprised to find their emergence and prevalence in the absence of antibiotics, we have explored the evolutionary context of their appearance, along with the fitness effect of single mutants under different environmental conditions and genetic backgrounds. To do so, we first monitor the trajectory of rifampicin-resistant clones within the evolution experiment, showing that the frequency trajectory varies with the time of first appearance and other factors. We then demonstrate that resistance is conferred by three previously characterized variants and confirm that these mutations confer an unprecedented level of fitness advantage under the conditions of the evolution experiment. Finally, we show that these same mutations can be highly deleterious with different genetic backgrounds and environmental condition, thereby providing glimpses into the prevalence of epistasis and pleiotropy for even well-characterized mutations.

MATERIALS AND METHODS

Bacterial strains and experimental design

We examined 114 replicated experimental lines from a previous experiment (Tenaillon et al. 2012). The high temperature adapted lines were founded from a common ancestral strain of *E. coli* B (genotype *REL1206*), which was descended from the strain *REL606* after 2000 generations at 37°C in Davis minimal medium supplemented with 25 µg/ml glucose (DM25). The ancestor was evolved in 114 replicate lines at 42.2°C for 2000 generations. Briefly, each population was founded from a single colony from an asexual clone (*REL1206*) stored in a glycerol-based suspension stored at -80°C. The lines were propagated by daily transfers of 0.1 ml of

each culture into 9.9 ml of DM25, allowing populations sizes to fluctuate daily between 5×10^6 cells at the transfer bottleneck to 5×10^8 cells, for a total of ~ 6.64 generations of binary fission per day (Lenski et al. 1991). Population samples of all 114 lines were taken at 100 generations, 200 generations and at 200-generation intervals thereafter (Tenailon et al. 2012).

Determining the level of rifampicin resistance

To assess the level of antibiotic resistance, we estimated the minimum inhibitory concentration (MICs) of rifampicin (Fisher Scientific, Fair Lawn, NJ) in the ancestral clone and the 114 evolved clones that were characterized genetically (Tenailon et al. 2012). Each isolate was grown in 5 ml of LB broth overnight at 37°C with constant shaking (120 rpm). After diluting the overnight cultures down 10^{-4} into MgSO_4 (10mM) in a 96 well microplate, we used a multichannel pipette to deliver 2 μL of culture on the surface of the LB agar with rifampicin. The dilution range for rifampicin was 0-800 $\mu\text{g/mL}$ (0, 0.25, 0.5, 1, 2.5, 5, 10, 25, 50, 100, 200, 400, 800). Plates were incubated overnight at 37°C. The MIC was defined as the lowest concentration of antibiotic inhibiting visible growth after overnight incubation. MIC values were confirmed in at least three separate experiments.

We constructed a histogram of the MIC distribution of the *REL1206* ancestor and the 114 evolved lines. A bimodal distribution is usually observed when strains have abnormally elevated MICs, with strains distributed above the upper end of the “susceptible” strains distribution defined as “resistant” strains (Laboratory Standards Institute, CLSI guidelines).

Fixation parameters and the time of appearance of resistance

We employed mixed population samples from 200-generation intervals to estimate the time of appearance of the rifampicin resistant phenotype. Briefly, cultures were inoculated from frozen stocks into 5 ml LB and incubated overnight at 37°C. 100 μ l of the culture was diluted 10⁶-fold and incubated on LB plates at 37°C for 12 hrs. From these plates, we randomly chose 100 colonies to streak on LB plates supplemented with rifampicin at one of three different rifampicin levels - low (12.5 μ g/mL), medium (50 μ g/mL) and high (100 μ g/mL) – corresponding to lower rifampicin concentration than the MIC of the characterized line. We incubated the rifampicin plates at 37°C for 12 hrs and estimated the frequency of resistant individuals in the population by dividing the number of colonies that grew in the LB + rifampicin plates by the number ($n=100$) of sampled colonies.

Based on the estimated frequency of rifampicin resistance over time, we estimated three parameters of fixation, as inspired by Lang et al. (Lang et al. 2011): τ_{up} , the time at which mutations reach 1% in the population, starting from the beginning of the experiment; τ_{fix} , the time from the beginning of the experiment to the time at which mutations reach 90% in the population; and s_{up} , the initial rate of increase of mutations (Figure S1.1). s_{up} is a proxy for the initial selection coefficient of a rifampicin resistant clone and was measured as the slope of the linear portion between the first two observations of rifampicin resistance in the populations. To estimate these parameters we fitted a linear regression to the natural logarithm of the ratio of resistant vs susceptible over time using the *lm* function in R (R Core Team 2013).

Strain construction and confirmation of recombinants

Single mutations that confer resistance to rifampicin were introduced into the ancestral strain *REL1206* using the pKD46 recombineering plasmid (Datsenko and Wanner 2000). The pKD46 plasmid supply the homologous recombination functions through the lambda Red genes. This plasmid has a temperature sensitive replication and carries an ampicillin resistant marker, so the plasmid can be cured from a strain when grown at 37°C without ampicillin. Briefly, we first introduced the pKD46 plasmid into the ancestral strain, electroporating 1 µl of plasmid (containing between 0.5 and 1 µg of plasmid) into 50 µl of competent cells using an Eppendorf Electroporator 2510 set at 1.8 kV. Following electroporation, we added 1 ml LB and incubated the cells at 30°C for 2 h with shaking. We then plated 100 µl of cells on LB agar plates containing 100 µg/ml ampicillin to select ampicillin-resistant (amp^{R}) transformants. The ancestral strain carrying the pKD46 plasmid was then grown overnight at 30°C in 5 ml of LB with 100 µg/ml of ampicillin. The overnight culture was 100 fold-diluted in 100 mL of LB with ampicillin and 1mM L-arabinose (Sigma) and grown at 30°C to an OD_{600} of 0.6. We made electrocompetent cells by washing the cultures 5 times with ice-cold water.

We designed three oligos of 70bp with the desired nucleotide change in the center of the oligo (Table S1.1) to introduce single point mutations that confer rifampicin resistance. 10 µM of each oligo was electroporated into 50 µl of cells. After electroporation we added 1 ml of LB and incubated cells at 30°C for 16 h with shaking and plated 100 µl in LB agar plates containing rifampicin. We selected single colonies and streaked them onto LB agar plates containing rifampicin. We then incubated the purified colonies on LB broth without antibiotic at 37°C and then tested for ampicillin

sensitivity to test for loss of the plasmid. Finally, the correct base replacement was confirmed by Sanger sequencing of ~420 bp of the *rpoB* gene, which was amplified by PCR (Table S1.1). The PCR thermal cycling conditions were 95°C for 4 min followed by 30 cycles of 95°C 30 sec and 51°C 30 sec; finally 72°C for 5 min.

Measurement of relative fitness effects of *rpoB* mutations

The fitness of the single mutant strains relative to the ancestral strain (*A*) was estimated from pairwise competition experiments following standard protocols (Lenski et al. 1991). Briefly, frozen samples of the mutated and ancestral strains were revived in LB broth and then grown separately for one day at 37°C and a second day at 42.2°C in DM25. The two competitors, a mutant line vs the *A* (*REL1206*) line, were mixed at a 1:1000 volumetric ratio and diluted 100-fold into 10 ml of DM25. We transferred 0.1 ml of each culture mixture daily into 9.9 ml of fresh DM and incubated at 42.2°C over a duration of two days. At the end of the daily growth cycle, we plated 100 µl of the culture on both LB agar plates and LB agar plates supplemented with rifampicin, in order to estimate the density of the total bacterial population ($A^{rif} + A$) and the resistance density strains (A^{rif}), respectively. The frequency of resistant strains ($freq A^{rif}$) was estimated as the density of A^{rif} divided by the density of total population. The relative fitness of the A^{rif} mutants, w_{mut} , was determined from the slope of the regression = $\ln [freq A^{rif} / (1 - freq A^{rif})]$ plotted against the time course in generations (Schrag et al. 1997; Lenski 1991).

Measurements of fitness effects across genotypes and environments

To measure the fitness effects of rifampicin resistance mutations in different genetic backgrounds, we introduced the *rpoB* mutations into two additional strains: *E. coli* B *REL606* and *E. coli* K12 *MG1655*. These strains are the most widely used laboratory strains and are genetically very similar with more than 99% sequence identity over approximately 92% of their genomes. The mutants derived from these strains were also competed against their original strains (*REL606* and *MG1655*), as described above. To test the differential effects of the mutations across genetic backgrounds, we performed a two-way analysis of variance (ANOVA) with genetic background (3 different *E. coli* strains) and genotype (mutations) as fixed effects.

To determine the differential effects of the mutations across environments, we competed mutants against their ancestors in four environments that differed in the temperature of incubation and/or the composition of the medium: 1) DM25 at 37°C, 2) DM1000 (Davis minimal medium supplemented with 1000µg/ml glucose) at 37°C, 3) DM1000 at 42.2°C and 4) LB at 42.2°C. To test for differential fitness effects across environments, we performed a two-way ANOVA with environment (5 environments) and genotype (3 mutations) as fixed effects. Finally, we used a mixed-effect model to assess the heterogeneity of fitness effects using genotype (mutations) as a random effect and temperature (37°C and 42.2°C) and glucose (25µg/ml and 1000µg/ml) as fixed effects. Statistical analyses were performed using the *lm* and *lmer* functions of R (R Core Team 2013) for the two-way ANOVAs and the mixed-effect model, respectively.

RESULTS

Parallel mutations in the *rpoB* gene conferred different levels of rifampicin resistance

We screened 114 evolved clones for the presence of rifampicin resistance. These clones represented single isolates from each of the replicate populations at the end of the 2000 generation experiment, and all 114 clones had been sequenced in their entirety (Tenaillon et al. 2012). Of these, 13 clones were resistant to rifampicin (Figure 1.1) at MIC concentrations corresponding to intermediate (25 to 50 $\mu\text{g/ml}$), high (100 $\mu\text{g/ml}$) and very high (more than 800 $\mu\text{g/ml}$) level of rifampicin resistance (Table 1.1).

Because resistance to rifampicin has been documented previously to be caused by single amino acid substitutions on *rpoB* (Campbell et al. 2001), we investigated the relationship between *rpoB* mutations and resistance. Overall, 46 non-synonymous *rpoB* mutations were observed in the original data set of 114 clones, but only 4 mutations were strong candidates for conferring rifampicin resistance in the 13 clones that exhibited resistance. The first three mutations were all found in codon 572 of *rpoB* (Table 1.1), which has been previously been shown to be both within the active site and the location of resistance mutations (Tupin et al. 2010). Twelve of the 13 lines had a non-synonymous mutation in codon 572, representing substitutions between Isoleucine and either Asparagine (*I572N*), Leucine (*I572L*) or Phenylalanine (*I572F*; Table 1.1). The level of resistance of these 12 clones was perfectly linked to their genotype. Mutations *I572N*, *I572L* and *I572F* corresponded, respectively, to an intermediate (25 to 50 $\mu\text{g/ml}$), high (100 $\mu\text{g/ml}$) and very high (more than 800 $\mu\text{g/ml}$) level of rifampicin resistance (Table 1.1). Assuming *I572F* is the sole cause of resistance (Campbell et al. 2001), this

single mutation in RNAP increased resistance by more than 320 fold relative to the average MIC of susceptible lines (2.5 $\mu\text{g/ml}$).

The last resistant clone had an intermediate level of resistance and a mutation in codon 143, which is part of the N-terminus of the β -subunit (Table 1.1). Mutation *R143L* was found in only one clone, while each of the three separate mutations in codon 572 were all found in at least two clones. Since we were interested in evolutionary aspects of resistance, we focused our attention on the three mutations with a clear signal of selection – i.e., those found to have occurred independently in more than one line – namely mutations *I572N*, *I572L* and *I572F*.

Rifampicin-resistant clones appeared early

The 12 rifampicin-resistant clones were chosen randomly for sequencing from their population at the end of the 2000-generation experiment. It was thus unclear if rifampicin resistance was fixed in each of the 12 populations, and it was also unknown when rifampicin resistance appeared during the experiment. To characterize the frequency trajectory of rifampicin resistance, we screened the 12 populations throughout 200 generation intervals (see Materials and Methods). As detailed in Figure 1.2, rifampicin resistance appeared before 500 generations for all 12 lines except line 77, which acquired the resistant phenotype after 800 generations. The resistance phenotype was eventually fixed ($f > 0.90$) in 10 of the 12 populations; in contrast, resistance appeared early in lines 56 and 131 but did not fix by generation 2000 (1.2).

We estimated parameters of the fixation process from the frequency trajectory of rifampicin resistance assuming that the resistance is caused by the same single

mutation observed at the end of the experiment (Figure S1.1; Table S1.2). We observed that the time of appearance of resistant mutations (τ_{up}) was correlated with the speed and dynamics of fixation (Figure 1.3). Later-occurring mutations (larger τ_{up}) had a smaller initial rate of increase (s_{up} , Figure 1.3 A), ultimately taking longer to fix (τ_{fix}) than early-occurring mutations (Figure 1.3 C). For example, the resistance phenotype in lines 35, 43, 97 and 112 - in which the rifampicin mutants reached a frequency higher than 0.2 at generation 200 (Figure 1.2) - fixed more rapidly than lines 4, 92 and 142 (Figure 1.2). Not surprisingly, s_{up} was negatively correlated with τ_{fix} , so that lines with slower initial rate of increase of the resistance phenotype took longer to fix the phenotype (Figure 1.3 B).

Mutations in the *rpoB* gene confer both resistance and a selective advantage

To measure the phenotype and selective advantage of *rpoB* non-synonymous mutations in codon 572, we introduced single nucleotide substitutions into the ancestral background (*REL1206*). With these genetic constructions we confirmed that the single amino acid substitutions in codon 572 fully explained the level of rifampicin resistance; that is the *I572N* mutation resulted in intermediate resistance (25 to 50 $\mu\text{g/ml}$), with high resistance (100 $\mu\text{g/ml}$) for *I572L* and the highest resistance (800 $\mu\text{g/ml}$) for *I572F*. Thus, as expected (Campbell et al. 2001), single base mutations in codon 572 of *rpoB* are sufficient to explain the rifampicin resistance phenotype.

We measured the fitness effect of each of the three mutations in competition experiments at 42.2°C in DM25. The three amino acid substitutions conferred (individually) a fitness advantage ranging from 0.182 to 0.246 (Figure 1.4) relative to the

REL1206 ancestral line. The selective advantage was significantly different between mutations ($P = 0.0493$), with mutation *I572F* being the most advantageous. Thus, the mutations confer resistance to rifampicin as well as a fitness advantage in the absence of rifampicin at 42.2°C in DM25.

The *rpoB* mutations have differential effects across genotypes and environments

Resistance mutations are usually thought to incur a cost in the absence of antibiotic, but that is not the case for our mutations under the conditions of the original thermal stress experiment. Given the large ~20% selective advantage of these mutations, why is rifampicin resistance not fixed throughout *E. coli sensu lato*? Because previous work has demonstrated a cost to rifampicin resistance in the absence of antibiotic (Reynolds 2000; Barrick et al. 2010; Brandis et al. 2012), we suspected differential effects of the codon 572 mutations with respect to genetic background and environmental conditions. We thus assessed the fitness of the mutations in different genetic backgrounds and environments.

To test the effect of genetic background, we inserted the three codon 572 mutations in two additional strains: *E. coli* B *REL606* and *E. coli* K12 *MG1655*. Both are commensal *E. coli* laboratory strains from phylogenetic subgroup A that were isolated a century ago. All three mutations (*I572N*, *I572L* and *I572F*) conferred resistance to rifampicin in the two new backgrounds (*REL606* and *MG1655*; data not shown). However, under thermal stress (42.2°C) and low glucose (DM25) conditions in the absence of rifampicin, the mutations had differential fitness effects depending on genetic background. The mutations were strongly beneficial in *REL606*, with a net

fitness benefit similar to that of the *REL1206* background (Figure 1.4, Table S1.3) but deleterious in the K12 background, with a ~2.5% to 10% fitness cost in the single mutants *I572L* and *1572F* relative to the non-mutated K12 *MG1655* competitor. In fact, despite several attempts, we were unable to introduce the *I572N* mutation into the *MG1655* background, suggesting that the *I572N* mutation may have a fitness of zero (lethality) in the *MG1655* background. Consistent with these fitness observations, an analysis of variance detected a significant effect of genetic background (Table 1.2). Although we did not detect a significant background-by-mutation interaction, the low *P*-value (0.053) suggests that the three mutations have differential effects across genetic backgrounds.

We also assessed relative fitness for mutations in *REL1206* background for five different environments, including two glucose treatments (DM25 and DM1000), two temperatures (37.0°C and 42.2° C), and a different medium (LB). These experiments revealed that fitness effects varied across environments (Figure 1.5, Table 1.3). The three *rpoB* mutations were costly at 37°C in DM25, at 37°C in DM1000 and at 42.2°C in LB medium (Table S1.3), but the relative fitness did not differ from neutrality (i.e., a relative fitness that differs from 1.0) at 42.2°C in DM1000. There was also a significant non-additive interaction between temperature and the concentration of glucose in DM media (Table 1.3); that is, the deleterious effect on fitness of both conditions (DM1000, 37°C) was not the sum of the deleterious effect on fitness of each condition separately.

DISCUSSION

The basis for our study is the observation that rifampicin resistance arose in the absence of an antibiotic during an evolutionary experiment. After 2000 generations of thermal stress, 13 of 114 *E. coli* clones exhibited resistance to rifampicin. Twelve of these 13 clones included a mutation in codon 572 of the *rpoB* gene, with three different mutations observed in that codon (Table 1.1). These three mutations have been noted previously to confer rifampicin resistance (Garibyan et al. 2003), a finding we have reconfirmed. Moreover, each of these three mutations occurred independently in more than one population, providing strong evidence by the criterion of evolutionary convergence (Christin et al. 2010) that the mutations are beneficial under the experimental conditions. Concerning the 13th and final clone, a mutation in codon 143 has been previously described to confer low resistance to rifampicin (*R143W*; Severinov et al. 1994), but the mechanistic causes of resistance for this clone remain unclear. However, our analysis of the RNAP 3-D structure suggests that codon 143 folds into the vicinity of the active site of the RNAP (Figure S1.2). It is possible, then, that mutations in this codon alter rifampicin binding, thus leading to resistance.

We used both direct and indirect evidence to confirm that all three mutations in codon 572 result in a fitness advantage within a thermal stress / low glucose environment. For direct evidence, we introduced single mutations into the ancestral *REL1206* background and assessed the relative fitness of mutants to unmutated *REL1206*. The measured fitness effect varied statistically among the three mutations, with relative fitnesses ranging from 1.18 to 1.25. Perhaps the most notable feature of these measurements is the magnitude of the effect. In the experimental evolution literature, it is rare to find single mutations with relative fitness benefits above ~15%

(Maisnier-Patin et al. 2002; Rozen et al. 2002). Thus, with the exception of mutations that compensate the cost of antibiotic resistance (Maisnier-Patin et al. 2002), the measured fitness benefit of the single *I572N*, *I572L* and *I572F* mutations are unprecedented (Khan et al. 2011; Chou et al. 2011). We note, however, that these high fitness values still explain only a fraction of the total realized relative fitness benefit of the twelve evolved clones, which have accumulated an average of 8 mutations compared to *REL1206* and an average relative fitness increase of ~40% (mean relative fitness 1.396; stdev 0.122; Tenaillon et al. 2012).

Indirect evidence for the benefit of these mutations comes from the assessment of the frequency trajectory of rifampicin resistance over the course of the full 2000-generation experiment. Generally, rifampicin resistance evolved early – within 500 generations - and swept to fixation within a few hundred generations (Figure 1.2). This steep increase in frequency is consistent with a high selection coefficient for the haplotypes that carry the resistance marker. We have measured the selection coefficient for these haplotypes by estimating s_{up} , which ranges between 0.015 and 0.077 (Figure 1.3A, Table S1.2). While these are high selection coefficients, they are not directly comparable to our relative fitness estimates, for several reasons (see below). What s_{up} does, however, is confirm that the capability for antibiotic resistance may be highly beneficial even in the absence of antibiotic.

Fixation dynamics of the resistance mutations

The frequency trajectories also provide crucial insights into the fixation dynamics of beneficial mutations. One interesting observation is that the relative fitnesses of

single rif^R mutations do not correlate with the estimated selective coefficient (s_{up}) of the populations that harbor these mutations (Figure 1.3D). This may reflect a lack statistical power to detect a correlation - since there are only three relative fitness measures - or may reflect the possibility that the resistant individuals observed at intermediate time points harbor different resistant mutations than the one observed at the end of the experiment. Nonetheless, we believe the lack of a relationship is meaningful. For example, the *I572F* rifampicin-resistant mutation found in lines 56 and 61 has the highest relative fitness as a single mutation (Figure 1.5), but rifampicin resistance was not fixed rapidly in these two lines. Instead, we find that early-occurring rif^R mutations take less time to reach fixation than late-occurring mutations (Figure 1.3C); this pattern suggests either that epistasis, clonal interference or frequency dependent fitness interactions influences s_{up} (Lang et al. 2011).

For the former (epistatic interactions), diminishing-returns epistasis is expected theoretically (Weinreich et al. 2005) and has been observed empirically as more and more mutations accumulate over the time-course of an experiment (Khan et al. 2011) (Chou et al. 2011). Under diminishing-returns, a relatively late occurring *rpoB* mutation may have a smaller fitness effect, conditional on the occurrence of previous beneficial mutations. For clonal interference, competition between beneficial haplotypes will slow the process of fixation (Gerrish and Lenski 1998; de Visser and Rozen 2006). Finally, complex dynamics such those observe in lines 56 and 131, might be due to frequency-dependent selection. In any case, such competition may be more common in the later stages of an experiment when multiple mutations have accrued (Gerrish and Lenski 1998). In contrast, early rif^R mutations likely occurred in a *REL1206* background that

was fairly devoid of other new mutations, thus minimizing possibilities for either clonal interference or epistatic interactions with other new mutations.

The possibility that frequency trajectories have been shaped in part by epistasis (whether as diminishing-returns or one of several other possible forms; Phillips 2008) is not surprising given the study of Tenaillon et al. (Tenaillon et al. 2012). This study detected statistical associations among mutations that were consistent with extensive and varied epistatic effects. These associations shaped the adaptive response to thermal stress into one of two distinct genetic solutions typified by mutations either in *rpoB* or in the termination factor *rho*, but rarely in both genes. To investigate the potential relationship of s_{up} to these statistical associations, we examined genetic data from Tenaillon et al. (Tenaillon et al. 2012). Clones from lines 43, 61 and 131, all of which had high τ_{fix} values (> 400 generations; Figure 1.2) carried mutations in both *rho* and *rpoB*, a combination statistically highly disfavored among the full dataset of 114 clones. This observations suggests that the long fixation time in these lines could be due in part to negative epistatic interactions between *rho* and *rpoB* mutations that reduces beneficial effects of both mutations. The strength and mechanism of these interactions need to be characterized more fully, however.

Previous studies have identified potential epistatic interactions with mutations in codon 572 (Trindade et al. 2009), and it is likely that epistasis also contributes to varying fitness effects among genetic backgrounds (Figure 1.4). In the high temperature and low glucose condition, our three codon 572 mutations conferred a slightly (but not significantly) higher relative fitnesses in the *REL606* background than in the ancestral *REL1206* background. The similar effects in these two backgrounds may not be

surprising, however, given that *REL1206* and *REL606* differ by only a handful of mutations: *REL1206* differs from *REL606* in 3 SNPs, an IS element and a large deletion (Tenailon et al. 2012; Barrick et al. 2009). In contrast, the *rif^R* mutations are detrimental in the K12 *MG1655* background (Table 1.2), even though K12 and B are genetically similar (> 99% sequence identity over ~92% of their genomes (Studier et al. 2009)).

The specificity of adaptation

The effects of the rifampicin resistance mutations also vary as a function of environment. In our study, the only environment in which the mutations are demonstrably beneficial is that of the original evolution experiment (high temperature and low glucose). In contrast, the effects of *rif^R* mutations are indistinguishable from neutrality in a high temperature and rich glucose environment and demonstrably detrimental at 37°C in poor and rich glucose environment (Reynolds 2000; Brandis et al. 2012; Barrick et al. 2010; Figure 1.5; Table S1.3).

Other studies have demonstrated that the fitness advantage conferred by a mutation is maintained across environments and conditions (Ostrowski et al. 2005) (Bataillon et al. 2011). In other words, they have found that beneficial mutations are generally not severely compromised in other environments (Bataillon et al. 2011). The logical extension of these observations is that a single beneficial mutation is unlikely to result in niche specialization, because it will not lead to drastic fitness differences across environments.

In stark contrast to these studies, we do observe the potential for the evolution of ecological specialization in a single mutational step, because all three mutations in

codon 572 of *rpoB* confer a selective advantage in the conditions of the original evolution experiment but significant disadvantages in other environments (Table 1.3) and genetic backgrounds (Table 1.2). In this context, it is important to repeat that this potential for niche specialization is not a function of antibiotic resistant, for which niche specialization is well known, but rather due to fitness effects across antibiotic-free environments.

The question remains as to whether our single *rpoB* mutations are rare or instead cast doubt on previous conclusions that niche specialization is "... unlikely to occur through the substitution of a single mutation" (Bataillon et al. 2011). The degree of ecological specialization for our single mutations could be due in part to the drastic selection pressure (high temperature) in the original experiment or to *rpoB* itself. Because mutations within *rpoB* can be highly pleiotropic, they can affect a series of downstream traits like gene expression (Conrad et al. 2010; Freddolino et al. 2012) that may be fine-tuned for specific selective regimes. We note that pleiotropic mutations have been observed in early stages of adaptation to ethanol stress (Goodarzi et al. 2009) and glycerol minimal media (Applebee et al. 2008), suggesting that early mutations in adaptation are commonly involved in transcriptional regulation with large fitness and pleiotropic effects (Levert et al. 2010; Hindre et al. 2012). As such, our *rpoB* mutations may not be uncommon, either in their effects or in their potential for ecological specialization. Thus, in our opinion, the frequency and occurrence of niche-specialization by single beneficial mutations is still an open question worthy of further study.

Mechanism

The mechanistic basis for the beneficial effect of *rpoB* mutations remains unclear. Since temperature affects the stability and activity of proteins (Singleton.R and Amelunxen.Re 1973; Ryals et al. 1982; Mejia et al. 2008), *rpoB* mutations may modify the stability and/or activity of RNAP at high temperatures. For example, previous studies have shown that mutation *I572F* increase transcription termination (Jin et al. 1988), and mutation *I572L* reduces transcription efficiency at 37°C (Reynolds 2000). Another (but not mutually-exclusive) hypothesis is that *rpoB* mutations cause changes in gene expression through the redistribution of RNAP in manner that favors adaptation to new environments (Conrad et al. 2010). The unique challenge here is explaining how these mechanistic effects can be advantageous in *REL1206* but (for example) disadvantageous in K12 (Figure 1.4). Fortunately, questions of mechanism are amenable to future experimental investigation.

CONCLUSIONS

Numerous studies have investigated rifampicin resistance in bacterial populations. The general tenor of these studies is that rifampicin resistance is deleterious in the absence of an antibiotic, and thus compensatory mutations are required for resistance to persist (e.g. Brandis et al. 2012). Our study differs from most previous in demonstrating the origin of resistance in the absence of antibiotics and also in demonstrating that the resistance mutations can be highly beneficial in the absence of antibiotic, depending on both the background of the mutation and the environment.

FIGURES

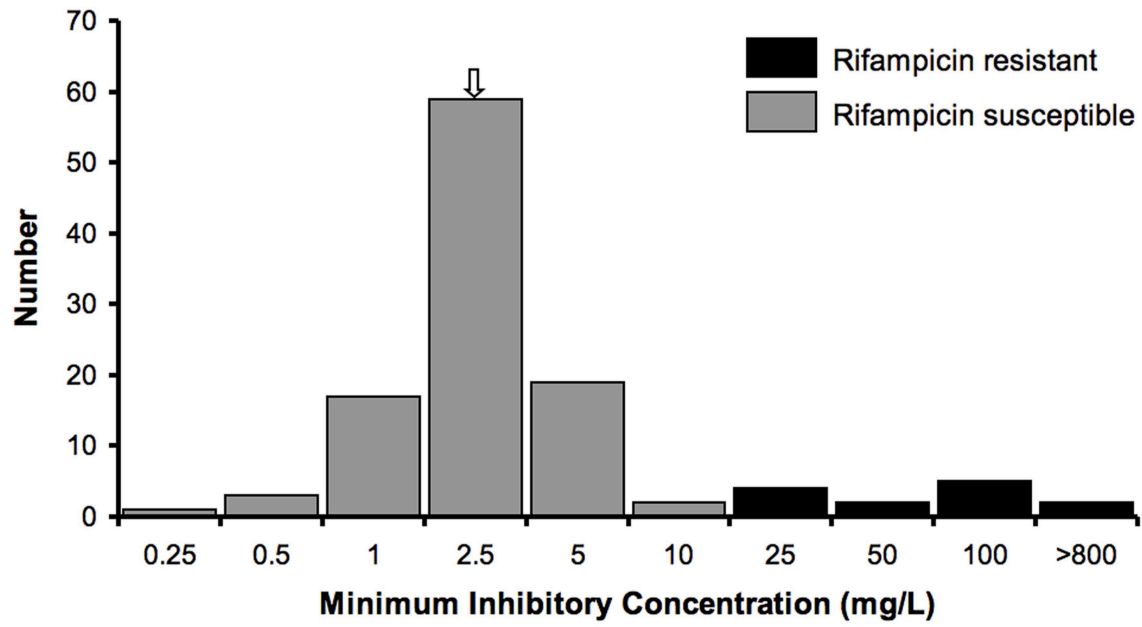


Figure 1.1. MIC distribution for the high temperature adapted clones in rifampicin. The MIC of the ancestral strain is indicated with an arrow.

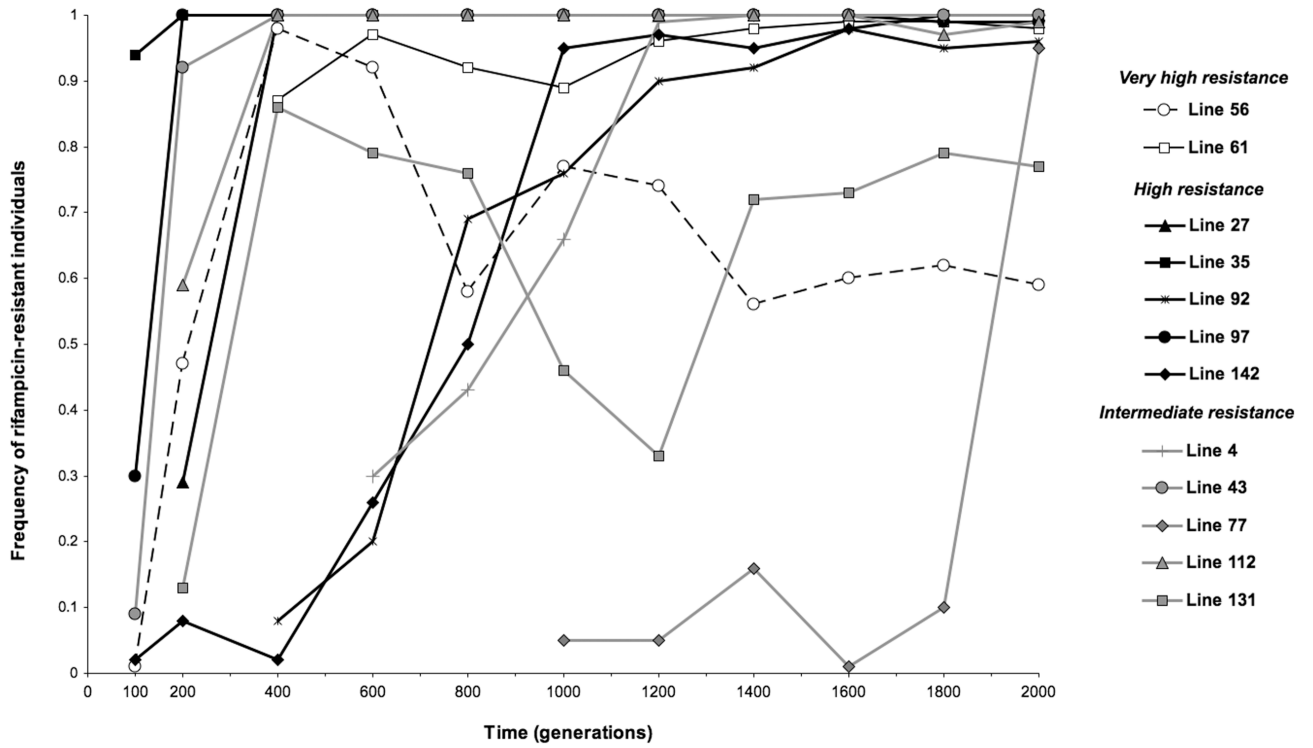


Figure 1.2. Temporal dynamics of the rifampicin-resistant individuals in 12 evolved populations during 2000 generations. The line numbers refer to the high temperature adapted clones in (Tenailon et al. 2012).

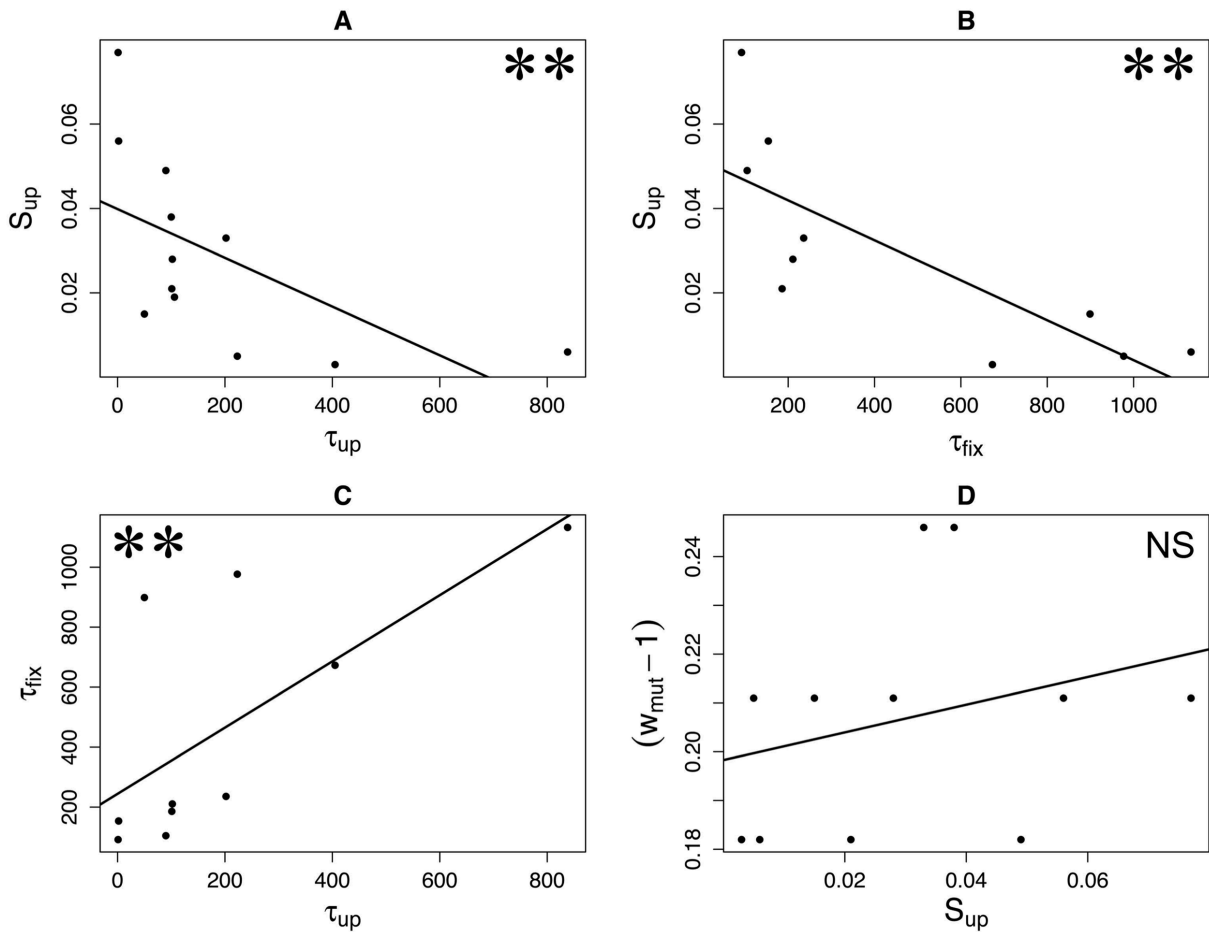


Figure 1.3. Relationship between the parameters of fixation. A) The time at which mutations initially occur (τ_{up}) negatively correlates with the initial rate of increase, S_{up} ($P = 0.0026$, Spearman's rank correlation). B) Mutations with lower initial rate of increase (S_{up}) take more time to fix ($P = 0.0016$, Spearman's rank correlation). C) Late-occurring mutations (larger τ_{up}) take more time to fix ($P = 0.0092$, Spearman's rank correlation). D) No correlation found between the selective advantage measured by direct competition experiments ($w_{mut} - 1$) and the initial rate of increase ($P = 0.3678$ Spearman's rank correlation). For all panels, double asterisks denote significance at $P < 0.01$ and 'NS' conveys non-significance ($P > 0.05$).

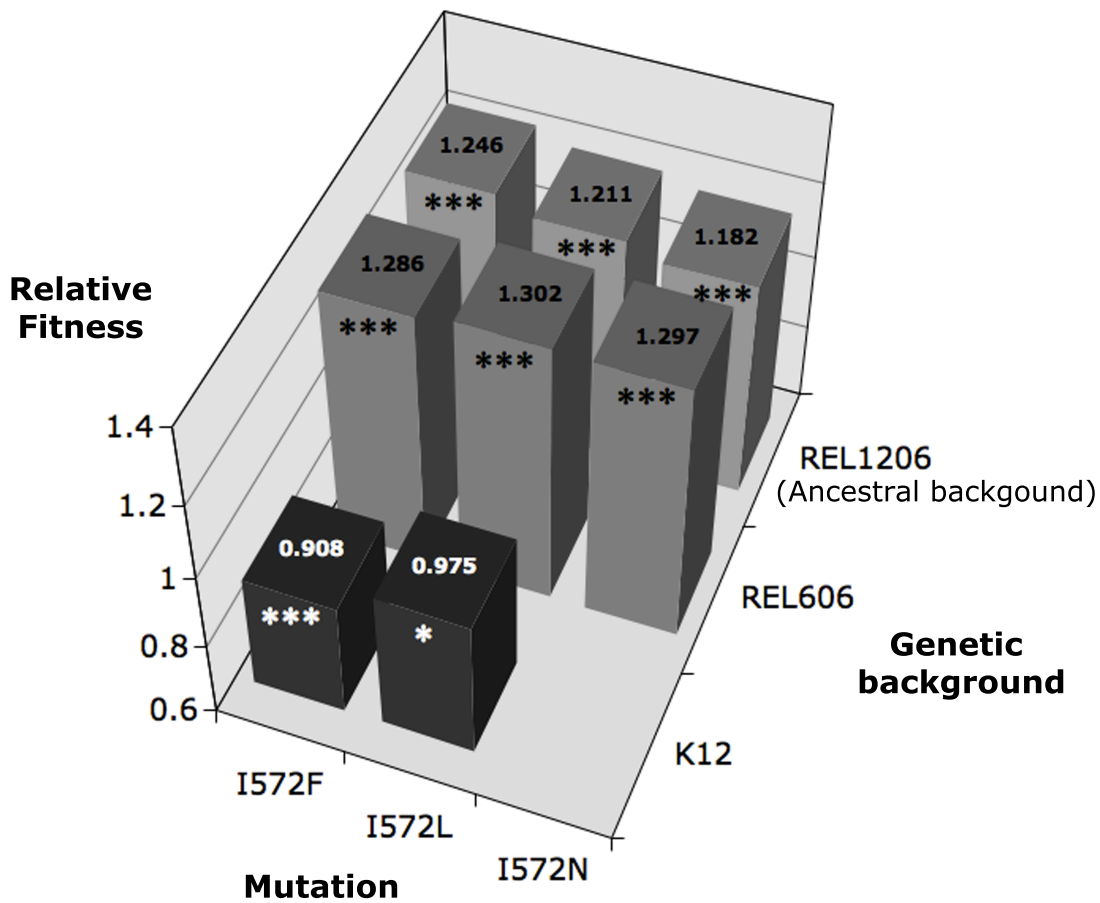


Figure 1.4. Relative fitness of the *rpoB* mutants measured at 42.2°C in DM25 in different genetic backgrounds. The mean relative fitness values, calculated from 6 replicates, are indicated on top of the bars. The gray color corresponds to an advantageous fitness effect of the mutations in relation to the ancestor and the dark color corresponds to a deleterious fitness effect of the mutations in relation to the ancestor. The asterisks represent significant deviation from the null hypothesis that mean fitness equals 1.0, with one and three asterisks denoting significance at $P < 0.05$ and $P < 0.001$, respectively.

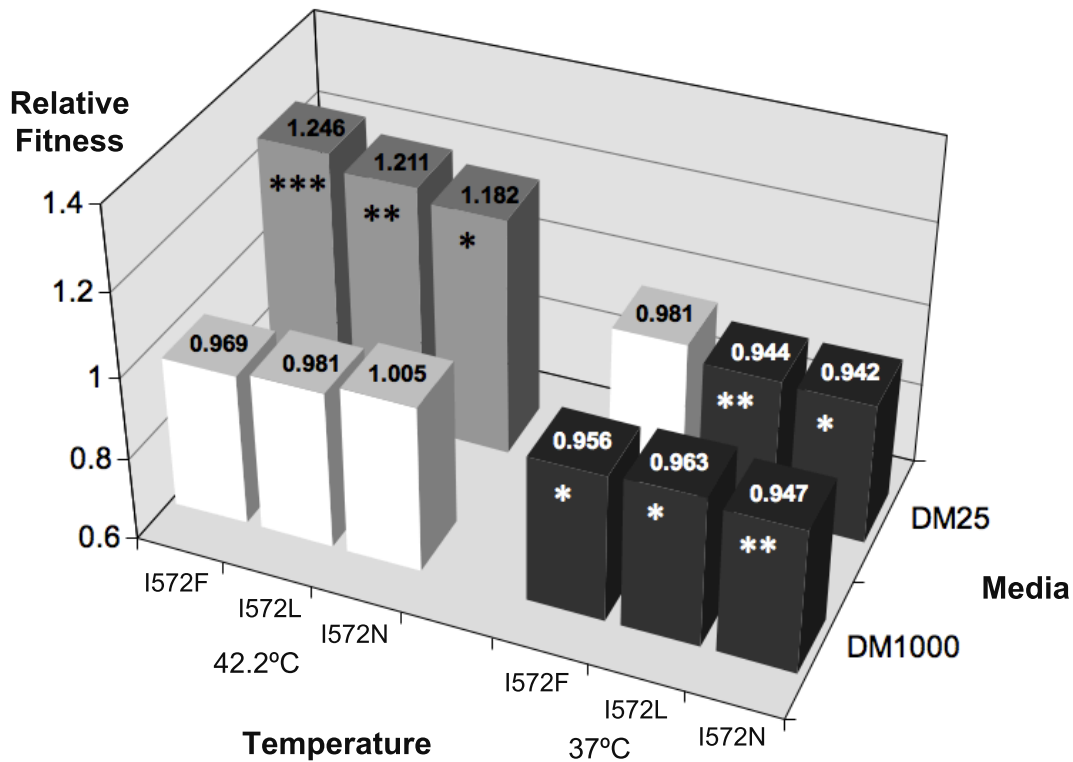


Figure 1.5. Relative fitness of the *rpoB* mutants measured in different conditions. The mean relative fitness values are indicated in the top of the bars. The gray color corresponds to an advantageous fitness effect of the mutations in relation to the ancestor, the dark color corresponds to a deleterious fitness effect of the mutations in relation to the ancestor, and the white color corresponds to a neutral fitness effect of the mutations in relation to the ancestor. The asterisks represent a significant difference from a mean fitness of 1.0, based on one-tailed t-distribution with n-1 degrees of freedom. One asterisk represents significance at $P < 0.05$; two denote significance at $P < 0.01$; three convey significance at $P < 0.001$.

TABLES

Table 1.1. Non-synonymous mutations in the *rpoB* gene conferring rifampicin resistance.

High temperature adapted clones	Nucleotide change	Amino acid change	Mutation change	Codon change	Phenotype (MIC in Rifampicin)	Level of Rifampicin resistance
56	ATC→ITC	Ile (I)→Phe (F)	A1714T	I572F	800 µg/mL	VERY HIGH
61	ATC→ITC	Ile (I)→Phe (F)	A1714T	I572F	800 µg/mL	
27	ATC→CTC	Ile (I)→Leu (L)	A1714C	I572L	100 µg/mL	HIGH
35	ATC→CTC	Ile (I)→Leu (L)	A1714C	I572L	100 µg/mL	
92	ATC→CTC	Ile (I)→Leu (L)	A1714C	I572L	100 µg/mL	
97	ATC→CTC	Ile (I)→Leu (L)	A1714C	I572L	100 µg/mL	
142	ATC→CTC	Ile (I)→Leu (L)	A1714C	I572L	100 µg/mL	
4	ATC→AAC	Ile (I)→Asn (N)	T1715A	I572N	50 µg/mL	INTERMEDIATE
43	ATC→AAC	Ile (I)→Asn (N)	T1715A	I572N	25 µg/mL	
77	ATC→AAC	Ile (I)→Asn (N)	T1715A	I572N	25 µg/mL	
112	ATC→AAC	Ile (I)→Asn (N)	T1715A	I572N	50 µg/mL	
131	ATC→AAC	Ile (I)→Asn (N)	T1715A	I572N	25 µg/mL	
59	CGT→CTT	Arg (R)→Leu (L)	G428T	R143L	25 µg/mL	

Table 1.2. Two-way analysis of variance for relative fitness of mutants in three different genetic backgrounds.

Analysis of Variance						
Source	df	SS	MS	F values	P	
Background	2	1.23909	0.61955	207.9335	<0.0001***	
Genotype (Mutation)	1	0.00737	0.00737	2.4734	0.12367	
Background x Genotype	2	0.01881	0.00941	3.1568	0.05334	
Residuals	40	0.11918	0.00298			

We treated the genetic background and the genotype (mutation) as fixed effects.

Table 1.3. Two-way analysis of variance for relative fitness of mutants in five different environments with the environment and the genotype (mutants) treated as fixed effects, and a mixed effect model with the genotype treated as random and the temperature and glucose treated as fixed effects.

Analysis of Variance					
Source	df	SS	MS	F values	P
Environment	4	1.18349	0.295871	49.0517	<0.0001***
Genotype	2	0.00386	0.001930	0.3200	0.7268
Environment x Genotype	8	0.05311	0.006639	1.1006	0.3679
Residuals	117	0.70572	0.006032		
Mixed effects model					
Source	df	F values	P		
Temperature	1	49.171	<0.0001***		
Glucose	1	49.333	<0.0001***		
Temperature x Glucose	1	52.560	<0.0001***		

SUPPORTING INFORMATION

Supporting figures

Parameters of fixation

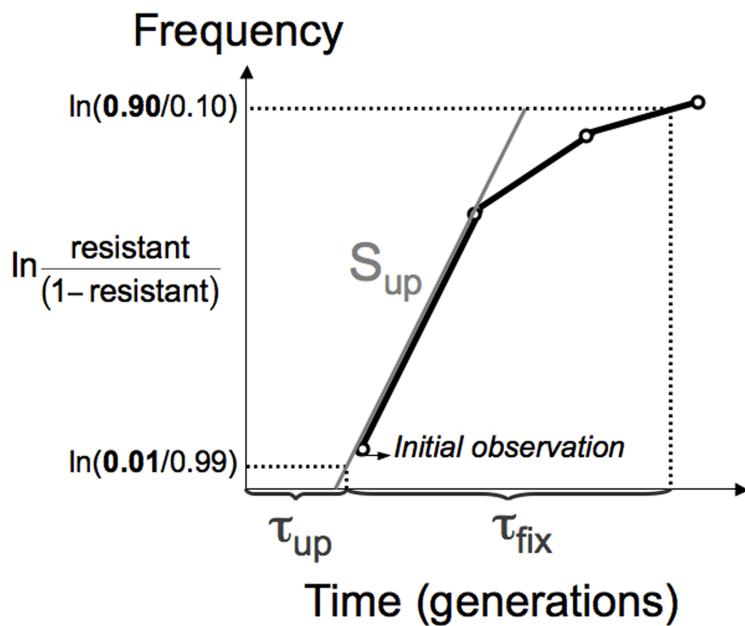


Figure S1.1. Parameters of fixation estimated from the trajectories of the rifampicin resistant individuals from the populations evolved at high temperature.

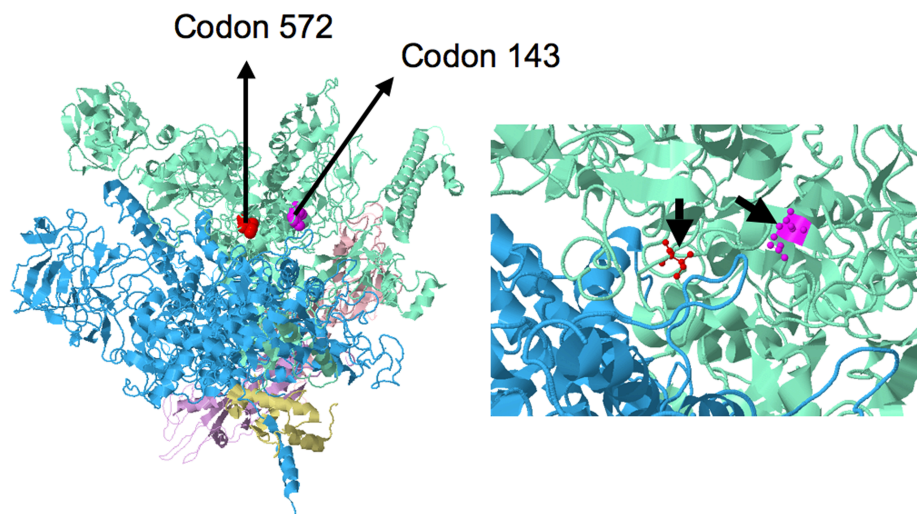


Figure S1.2. Three-dimensional structure of RNAP generated using Jmol from the Protein Data Bank (www.rcsb.org). The residues that are mutated (I572 and R143) are indicated with black arrows. These residues are predicted to prevent the binding of the rifampicin to the active site of the RNAP.

Supporting Tables

Table S1.1. Oligonucleotides and primers used in this study.

Oligos used for recombineering to create <i>rpoB</i> mutants	
I572F	5'- ATC GAA ACC CCT GAA GGT CCG AAC ATC GGT CTG <u>I</u> TC AAC TCT CTG TCC GTG TAC GCA CAG ACT AAC GAA TA -3'
I572L	5'- ATC GAA ACC CCT GAA GGT CCG AAC ATC GGT CTG <u>C</u> TC AAC TCT CTG TCC GTG TAC GCA CAG ACT AAC GAA TA -3'
I572N	5'- ATC GAA ACC CCT GAA GGT CCG AAC ATC GGT CTG <u>A</u> AC AAC TCT CTG TCC GTG TAC GCA CAG ACT AAC GAA TA -3'
Primer pair used for PCR amplification of <i>rpoB</i> region	
I572 Forward	5'- ACA ACC CGC TGT CTG AGA TT -3'
I572 Reverse	5'- TGG GTG GAT ACG TCC ATG TAG -3'

Table S1.2. Parameters of fixation estimated from the frequency trajectories.

High temperature adapted lines	Codon change	Phenotype (MIC in Rifampicin)	S_{up}	τ_{up} (generations)	τ_{fix} (generations)
35	I572L	100 μ g/mL	0.077	1	92
97	I572L	100 μ g/mL	0.056	2	154
43	I572N	25 μ g/mL	0.049	90	105
56	I572F	800 μ g/mL	0.038	100	Not fixed
61	I572F	800 μ g/mL	0.033	202	236
27	I572L	100 μ g/mL	0.028	102	211
112	I572N	50 μ g/mL	0.021	101	186
131	I572N	25 μ g/mL	0.019	106	Not fixed
142	I572L	100 μ g/mL	0.015	50	899
77	I572N	25 μ g/mL	0.006	838	1133
92	I572L	100 μ g/mL	0.005	223	977
4	I572N	50 μ g/mL	0.003	405	673

Table S1.3. Relative fitness measured in different conditions and different genetic backgrounds.

Genetic background	Fitness assay at	Mutant	Codon Change	Relative fitness ($\pm 95\%CI$) ¹	p-value ²	n	Phenotypic effect
REL 1206 (Ancestral background)	42.2°C in DM25	A ^{Rif} 1	I572F	1.246 \pm 0.012	<0.0001***	6	24.6% Advantage
		A ^{Rif} 2	I572L	1.211 \pm 0.018	<0.0001***	6	21.1% Advantage
		A ^{Rif} 3	I572N	1.182 \pm 0.020	0.0001***	6	18.2% Advantage
	37°C in DM25	A ^{Rif} 1	I572F	0.981 \pm 0.013	0.0893	12	Neutral
		A ^{Rif} 2	I572L	0.944 \pm 0.017	0.011*	6	5.6% Cost
		A ^{Rif} 3	I572N	0.942 \pm 0.016	0.0064**	6	5.8% Cost
	42.2°C in DM1000	A ^{Rif} 1	I572F	0.969 \pm 0.055	0.2944	12	Neutral
		A ^{Rif} 2	I572L	0.981 \pm 0.015	0.1155	12	Neutral
		A ^{Rif} 3	I572N	1.005 \pm 0.017	0.6065	12	Neutral
	37°C in DM1000	A ^{Rif} 1	I572F	0.956 \pm 0.017	0.0135*	12	4.4% Cost
		A ^{Rif} 2	I572L	0.963 \pm 0.015	0.0147*	12	3.7% Cost
		A ^{Rif} 3	I572N	0.947 \pm 0.017	0.0051**	12	5.3% Cost
	42.2°C in LB	A ^{Rif} 1	I572F	0.868 \pm 0.014	0.0001***	6	13% Cost
		A ^{Rif} 2	I572L	0.836 \pm 0.033	0.002**	6	13% Cost
		A ^{Rif} 3	I572N	0.928 \pm 0.030	0.03*	6	9.8% Cost
REL 606	42.2°C in DM25	R ^{Rif} 1	I572F	1.286 \pm 0.030	<0.0001***	6	28.6% Advantage
		R ^{Rif} 2	I572L	1.302 \pm 0.036	0.0002***	6	30.2% Advantage
		R ^{Rif} 3	I572N	1.297 \pm 0.038	0.0003***	6	29.7% Advantage
K12 (MG1655)	42.2°C in DM25	K ^{Rif} 1	I572F	0.908 \pm 0.018	0.0004***	6	9.2% Cost
		K ^{Rif} 2	I572L	0.975 \pm 0.010	0.0218*	6	2.5% Cost

¹Mean relative fitness and 95% confidence interval based on n replicates.

²Significance values based on one-tailed t-distribution with n-1 degrees of freedom; the null hypothesis is that the mean fitness equals 1.

REFERENCES

- Andersson, D.I., and B.R. Levin. 1999. The biological cost of antibiotic resistance. *Curr Opin Microbiol* 2:489-493.
- Applebee, M.K., M.J. Herrgard, and B.O. Palsson. 2008. Impact of individual mutations on increased fitness in adaptively evolved strains of *Escherichia coli*. *J Bacteriol* 190:5087-5094.
- Barrick, J.E., M.R. Kauth, C.C. Streliaoff, and R.E. Lenski. 2010. *Escherichia coli* rpoB mutants have increased evolvability in proportion to their fitness defects. *Mol Biol Evol* 27:1338-1347.
- Barrick, J.E., D.S. Yu, S.H. Yoon, H. Jeong, T.K. Oh, D. Schneider, R.E. Lenski, and J.F. Kim. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243-U74.
- Bataillon, T., T.Y. Zhang, and R. Kassen. 2011. Cost of Adaptation and Fitness Effects of Beneficial Mutations in *Pseudomonas fluorescens*. *Genetics* 189:939-U317.
- Brandis, G., M. Wrande, L. Liljas, and D. Hughes. 2012. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. *Mol Microbiol* 85:142-151.
- Campbell, E.A., N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, and S.A. Darst. 2001. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* 104:901-912.
- Chou, H.H., H.C. Chiu, N.F. Delaney, D. Segre, and C.J. Marx. 2011. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. *Science* 332:1190-1192.
- Christin, P.A., D.M. Weinreich, and G. Besnard. 2010. Causes and evolutionary significance of genetic convergence. *Trends Genet* 26:400-405.
- Comas, I., S. Borrell, A. Roetzer, G. Rose, B. Malla, M. Kato-Maeda, J. Galagan, S. Niemann, and S. Gagneux. 2012. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* 44:106-U147.
- Conrad, T.M., M. Frazier, A.R. Joyce, B.K. Cho, E.M. Knight, N.E. Lewis, R. Landick, and B.O. Palsson. 2010. RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc Natl Acad Sci USA* 107:20500-20505.
- Datsenko, K.A., and B.L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640-6645.

- de Visser, J.A., and D.E. Rozen. 2006. Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. *Genetics* 172:2093-2100.
- Freddolino, P.L., H. Goodarzi, and S. Tavazoie. 2012. Fitness Landscape Transformation through a Single Amino Acid Change in the Rho Terminator. *Plos Genetics* 8: e1002744
- Gariyban, L., T. Huang, M. Kim, E. Wolff, A. Nguyen, T. Nguyen, A. Diep, K.B. Hu, A. Iverson, H.J. Yang, and J.H. Miller. 2003. Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair* 2:593-608.
- Gerrish, P.J., and R.E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* 102-3:127-144.
- Goodarzi, H., A.K. Hottes, and S. Tavazoie. 2009. Global discovery of adaptive mutations. *Nature Methods* 6:581-U44.
- Gros, P.A., H. Le Nagard, and O. Tenaillon. 2009. The Evolution of Epistasis and Its Links With Genetic Robustness, Complexity and Drift in a Phenotypic Model of Adaptation. *Genetics* 182:277-293.
- Hall, A.R., and R.C. MacLean. 2011. Epistasis buffers the fitness effects of rifampicin-resistance mutations in *Pseudomonas aeruginosa*. *Evolution* 65:2370-2379.
- Hindre, T., C. Knibbe, G. Beslon, and D. Schneider. 2012. New insights into bacterial adaptation through in vivo and in silico experimental evolution. *Nature Reviews Microbiology* 10:352-365.
- Jin, D.J., W.A. Walter, and C.A. Gross. 1988. Characterization of the termination phenotypes of rifampicin-resistant mutants. *J Mol Biol* 202:245-253.
- Kassen, R., and T. Bataillon. 2006. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nat Genet* 38:484-488.
- Khan, A.I., D.M. Dinh, D. Schneider, R.E. Lenski, and T.F. Cooper. 2011. Negative Epistasis Between Beneficial Mutations in an Evolving Bacterial Population. *Science* 332:1193-1196.
- Kimura, M. 1983. *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge
- Lang, G.I., D. Botstein, and M.M. Desai. 2011. Genetic Variation and the Fate of Beneficial Mutations in Asexual Populations. *Genetics* 188:647-U236.

- Lenski, R.E. 1991. Quantifying fitness and gene stability in microorganisms. In *Assessing ecological risks of biotechnology*. Edited by Ginzburg LR Bosto: Butterworth-Heinemann, Boston, pp 173-192
- Lenski, R.E., M.R. Rose, S.C. Simpson, and S.C. Tadler. 1991. Long-term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2,000 generations. *Am Nat* 138:1315-1341.
- Leveret, M., O. Zamfir, O. Clermont, O. Bouvet, S. Lespinats, M.C. Hipeaux, C. Branger, B. Picard, C. Saint-Ruf, F. Norel, T. Balliau, M. Zivy, H. Le Nagard, S. Cruvellier, B. Chane-Woon-Ming, S. Nilsson, I. Gudelj, K. Phan, T. Ferenci, O. Tenaillon, and E. Denamur. 2010. Molecular and Evolutionary Bases of Within-Patient Genotypic and Phenotypic Diversity in *Escherichia coli* Extraintestinal Infections. *Plos Pathogens* 6
- Levin, B.R., V. Perrot, and N. Walker. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* 154:985-997.
- MacLean, R.C., A.R. Hall, G.G. Perron, and A. Buckling. 2010. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nat Rev Genet* 11:405-414.
- MacLean, R.C., G. Bell, and P.B. Rainey. 2004. The evolution of a pleiotropic fitness trade-off in *Pseudomonas fluorescens*. *Proc Natl Acad Sci U S A* 101:8072-8077.
- Maisnier-Patin, S., O.G. Berg, L. Liljas, and D.I. Andersson. 2002. Compensatory adaptation to the deleterious effect of antibiotic resistance in *Salmonella typhimurium*. *Mol Microbiol* 46:355-366.
- Marcusson, L.L., N. Fridmodt-Moller, and D. Hughes. 2009. Interplay in the Selection of Fluoroquinolone Resistance and Bacterial Fitness. *Plos Pathogens* 5
- Mejia, Y.X., H.B. Mao, N.R. Forde, and C. Bustamante. 2008. Thermal probing of E-coli RNA polymerase off-pathway mechanisms. *J Mol Biol* 382:628-637.
- Orr, H.A. 2005. The genetic theory of adaptation: A brief history. *Nat Rev Genet* 6:119-127.
- Ostrowski, E.A., D.E. Rozen, and R.E. Lenski. 2005. Pleiotropic effects of beneficial mutations in *Escherichia coli*. *Evolution* 59:2343-2352.
- Phillips, P.C. 2008. Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* 9:855-867.

- R Core Team 2013. *R: A language and environment for statistical computing*. from <http://www.R-project.org/>
- Remold, S.K., and R.E. Lenski. 2001. Contribution of individual random mutations to genotype-by-environment interactions in *Escherichia coli*. *Proc Natl Acad Sci USA* 98:11388-11393.
- Reynolds, M.G. 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics* 156:1471-1481.
- Rozen, D.E., J.A. de Visser, and P.J. Gerrish. 2002. Fitness effects of fixed beneficial mutations in microbial populations. *Curr Biol* 12:1040-1045.
- Ryals, J., R. Little, and H. Bremer. 1982. Temperature-dependence of RNA-synthesis parameters in *Escherichia coli*. *J Bacteriol* 151:879-887.
- Schrag, S.J., V. Perrot, and B.R. Levin. 1997. Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proceedings of the Royal Society of London Series B-Biological Sciences* 264:1287-1291
- Severinov, K., M. Soushko, A. Goldfarb, and V. Nikiforov. 1994. Rif^R mutations in the beginning of the *Escherichia coli* rpoB gene. *Molecular & General Genetics* 244:120-126.
- Singleton, R., Amelunxen, R. 1973. Proteins from thermophilic microorganisms. *Bacteriol Rev* 37:320-342.
- Studier, F.W., P. Daegelen, R.E. Lenski, S. Maslov, and J.F. Kim. 2009. Understanding the Differences between Genome Sequences of *Escherichia coli* B Strains REL606 and BL21(DE3) and Comparison of the E-coli B and K-12 Genomes. *J Mol Biol* 394:653-680.
- Taubes, G. 2008. The bacteria fight back. *Science* 321:356-361.
- Tenaillon, O., A. Rodriguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennett, A.D. Long, and B.S. Gaut. 2012. The Molecular Diversity of Adaptive Convergence. *Science* 335:457-461.
- Trindade, S., A. Sousa, and I. Gordo. 2012. Antibiotic resistance and stress in the light of fisher's model. *Evolution* 66:3815-3824.
- Trindade, S., A. Sousa, K.B. Xavier, F. Dionisio, M.G. Ferreira, and I. Gordo. 2009. Positive Epistasis Drives the Acquisition of Multidrug Resistance. *Plos Genetics* 5:9.

- Tupin, A., M. Gualtieri, F. Roquet-Baneres, Z. Morichaud, K. Brodolin, and J.P. Leonetti. 2010. Resistance to rifampicin: at the crossroads between ecological, genomic and medical concerns. *Int J Antimicrob Agents* 35:519-523.
- Weinreich, D.M., R.A. Watson, and L. Chao. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59:1165-1174.
- Woods, R.J., J.E. Barrick, T.F. Cooper, U. Shrestha, M.R. Kauth, and R.E. Lenski. 2011. Second-Order Selection for Evolvability in a Large *Escherichia coli* Population. *Science* 331:1433-1436.

CHAPTER 2

First-step mutations restore the ancestral expression state during thermal stress adaptation

INTRODUCTION

Organisms are often exposed to stressful environments. To cope, they have evolved responses based on the duration of the stress. For example, bacteria that are exposed to increased temperatures display a transient heat-shock response, which involves up-regulation of genes encoding heat stress-proteins (Richter et al. 2010), followed by a period of phenotypic acclimation (Gunasekera et al. 2008). If the environmental stress is sustained over a long period of time, individuals may eventually accumulate mutations that can result in long-term adaptation of the population to the stressful environment. Although acute responses to stress have been well characterized (Richter et al. 2010), the mechanisms of stress acclimation and adaptation are less understood (Riehle et al. 2001; Gunasekera et al. 2008).

One of the possible mechanisms underlying adaptation to stressful conditions involves genetic changes that produce novel traits or new physiological functions, such as antibiotic resistance or the ability to use new metabolic pathways (Blount et al. 2012; Quandt et al. 2014). Another mechanism leading to stress adaptation may be the restoration of cellular functions to a pre-stressed state. Rather than creating new functions, restorative mutations revert some aspect of the individuals' altered physiological state back to an unstressed or "normal" state (Carroll and Marx 2013). This pattern of restoration has been recently observed during metabolic perturbation

and high-temperature adaptation (Fong et al. 2005; Carroll and Marx 2013; Sandberg et al. 2014). Although both novelty and restoration likely drive long-term adaptation, it is not clear which one prevails.

Another aspect that remains largely unexplored is the temporal process of phenotypic change during adaptation to stress. After an environment becomes stressful, the acclimated state becomes the initial phenotypic state upon which natural selection acts. Thereafter, each adaptive mutation moves the population towards a phenotypic optimum (i.e. to a phenotype that best fits the present environment (Fisher 1930; Orr 2005)). For historical and methodological reasons (Orr 2005), most evolutionary studies have focused on the end product of adaptation, leaving intermediate steps of adaptation largely unexplored.

Fortunately, studies in experimental evolution may provide insight into the sequential magnitude of fitness changes during adaptation. To date, these studies have shown that the first beneficial mutation, or the “first-step mutation”, generally confers the largest fitness advantage, perhaps because it avoids clonal interference. In contrast, subsequent mutations confer more moderate fitness gains (Chou et al. 2011; Khan et al. 2011), in part due to diminishing returns epistasis.

While our knowledge about the trajectory of fitness continues to grow, few studies have examined shifts in phenotypes during the adaptive process. One exception is the study of Fong et al. (2005) in which they used microarrays to follow the phenotype of gene expression (GE) after a shift in growth from glucose to lactate and from glucose to glycerol (Fong et al. 2005). Fong et al. observed that 39% of the total genes were differentially expressed during the process of acclimation to glycerol medium; these

changes due to acclimation were followed by decreases in the number of differentially expressed genes at an intermediate point of adaptation and at the endpoint of the experiment. Most of the GE changes during adaptation were restorative (Fong et al. 2005) –i.e., they restored GE to normal, pre-stressed levels. Unfortunately, however, the genetic bases of these phenotypic changes were not resolved; it was unclear if the changes in GE during the intermediate point of adaptation were caused by an early adaptive mutation or were caused by combinations of mutations that accumulated during the experiment. Thus, many questions regarding phenotypic adaptation remain unresolved. What are the molecular mechanisms underlying the large fitness advantage conferred by first-step mutations? What is the phenotypic contribution of the first-step mutation compared to phenotypic variation accumulated during an adaptive walk?

We have decided to explore these questions based on our recent, large-scale evolution experiment (Tenaillon et al. 2012). In this experiment a strain of *E. coli* B was evolved at 42°C in 114 replicate populations. After 2000 generations, we isolated single clones from each evolved population and identified the genetic changes that had accumulated during the experiment. Overall, the *rpoB* gene, which encodes the β subunit of RNA polymerase (RNAP), was the most mutated gene (with a total of 87 mutations in the 114 genomes). Three *rpoB* mutations were especially interesting because they resulted in amino acid substitutions in the active site of RNAP and conferred rifampicin resistance (Rodríguez-Verdugo et al. 2013). These mutations, which are located in the same codon, codon 572, were driven to high frequency in 12 populations and were also beneficial at high temperature in the low glucose medium (Rodríguez-Verdugo et al. 2013). Furthermore, these mutations typically appeared early

in the evolution experiment. In one population, for example, the *rpoB* I572L mutation swept to fixation before 100 generations, strongly suggesting that these are early, large-effect mutations in the adaptive process (Rodríguez-Verdugo et al. 2013). Since these *rpoB* mutations are located in the contact region between the downstream DNA duplex and RNAP (proximal active-site), they could play key roles in modulating the enzyme's activity in all three stages of transcription: initiation, elongation and termination (Ederth et al. 2006).

Yet, some questions remain. First, we have not identified the mechanistic bases of their fitness advantage. Second, although we know that these parallel mutations converge on having a fitness advantage at 42°C (Rodríguez-Verdugo et al. 2013), with a disadvantages at low temperatures (Rodríguez-Verdugo et al. 2014), we have yet to explore if they converged on other phenotypes. That is, it is still unclear if these mutations converged phenotypically with other *rpoB* mutations found in other regions of RNAP. In the present study, we address all these questions. Solving these questions is relevant for understanding the adaptive mechanisms prevalent in our experimental conditions and also in the other evolutionary experiments, because mutations that affect transcriptional regulators (such as RNAP and the Rho termination factor) typically appear in the early stages of stress adaptation (Applebee et al. 2008; Goodarzi et al. 2009; Kishimoto et al. 2010). Numerous observations point to the possibility that highly pleiotropic mutations in transcriptional regulators could be the first step of a general mechanism of adaptation (Fong et al. 2005; Hindre et al. 2012).

With the previous considerations in mind, we first describe the growth characteristics and GE during acclimation stress response to 42°C. We then explore

first-step mutations in terms of their effect on growth and GE. In this respect, we explore whether three amino acid substitutions in the RNAP active site converge on the same expression phenotype and also whether their effects are similar to another *rpoB* mutation that affects an amino acid substitution far from the RNAP active site. Finally, we evaluated the phenotypic contribution of one of these first-step mutations relative to the end product of adaptation, as represented by two of our high-temperature adapted clones.

RESULTS

Acclimation to thermal stress involved many changes in GE.

To explore the phenotypic effect of high temperature on the ancestor, we first characterized the ancestor's growth at 37°C, which is the ancestral optimal temperature, and at 42°C, the stress condition. The ancestor had a significantly longer lag phase and a significantly lower final yield when grown at 42°C than when grown at 37°C (Figure 2.1A, Table 2.1), reflecting the fact that 42°C is a stressful temperature.

Next, we explored the global GE profile after acclimation to 42°C. We obtained RNA-seq data from three replicates of the ancestor at the mid-exponential growth phase at both 37°C and 42°C. Even under a stringent criterion of significance ($q < 0.001$), we identified 1984 genes that were differentially expressed at 42°C relative to 37°C (Figure 2.1B). Of these differentially expressed genes, we identified 279 genes that were highly down-regulated (\log_2 fold change < -2 ; blue dots in Figure 2.1B) and 289 genes that were highly up-regulated (\log_2 fold change > 2 ; red dots in Figure 2.1B) at 42°C relative to 37°C. Based on an enrichment analysis of GO functional categories, we identified

significant down-regulation of genes involved in translation during heat acclimation (GO: 0006412; $P = 6.32 \times 10^{-47}$). Of the genes involved in translation, 54 genes transcribed products that were structural constituents of ribosomes (GO:0003735), including *rpl*, *rpm* and *rps* genes (Table S2.1). Other significantly down-regulated biological processes involved: *i*) amino acid biosynthesis (GO: 0008652; $P = 1.98 \times 10^{-7}$), particularly genes involved in the biosynthesis of methionine (*met* genes); *ii*) biosynthesis of ribonucleosides (GO: 0042455; $P = 3.15 \times 10^{-8}$), including purines and pyrimidine (*pur* and *pyr* genes); and *iii*) flagellum-dependent cell motility (GO: 0001539; $P = 6.25 \times 10^{-3}$), including *flg* and *flh* genes (Table S2.1). Surprisingly, the heat-shock inducible genes (Nonaka et al. 2006), were mostly down regulated during heat acclimation (Table S2.2). Also, surprisingly, the gene synthesizing subunits of the core RNAP (*rpoA*, *rpoB*, *rpoC* and *rpoZ*) were significantly down regulated (Table S2.2).

Among the up-regulated genes, we identified a significant enrichment of genes involved in: *i*) amino acid degradation (GO: 0009063; $P = 6.16 \times 10^{-4}$), in particular degradation of arginine to succinate and glutamate (*ast* genes; GO:0006527; $P = 8.76 \times 10^{-3}$); *ii*) fatty acid beta-oxidation (*fad* genes; GO:0006635; $P = 4.71 \times 10^{-3}$); *iii*) transport of glycerol-3-phosphate (*ugp* genes, GO:0015794; $P = 1.73 \times 10^{-2}$); and *iv*) dipeptide transport (*dpp* genes; GO: 0008643; $P = 1.07 \times 10^{-2}$; Table S2.1). Interestingly, most of the genes previously identified during general stress response were significantly up-regulated in our study (Table S2.2).

These changes in GE during thermal acclimation response served as a baseline to compare to the changes in GE that occurred during thermal adaptation.

Fist-step adaptive mutations drastically changed GE at 42°C.

We investigated the effect of three potential first-step mutations – *rpoB I572F*, *I572L* and *I572N* – on cell growth and GE. Previous estimates of relative fitness indicated that these mutations are advantageous at 42°C (Rodríguez-Verdugo et al. 2013). We complemented this observation by characterizing their growth curves, both at 42°C and at 37°C. The mutants had a significant longer lag phase and lower final yield at 42°C compared to their growth at 37°C (Table 2.1). Nevertheless, the mutants had a significant shorter lag phase and a significant higher final yield compared to the ancestor grown at 42°C, indicating that they performed better under thermal stress (Table 2.1, Figure 2.2).

We explored two hypotheses about the molecular mechanisms that may underlie the growth improvement of the mutants at 42°C. Our first hypothesis was the mutants grew better at 42°C because the mutated RNAP was more efficient than the ancestral RNAP at transcribing DNA to RNA (Jin et al. 1992; Reynolds 2000). Our second, non-exclusive hypothesis was both that the mutated RNAP transcribed a different set of genes than the ancestral RNAP at 42°C and that this differential expression underlies the growth improvement (Conrad et al. 2010).

Regarding the first hypothesis, we specifically predicted that the *rpoB* mutations slowed the RNAP complex (Rodríguez-Verdugo et al. 2014), which is otherwise accelerated at high temperatures (Ryals et al. 1982; Mejia et al. 2008). In turn, we reasoned that the reduced speed of RNAP increases both transcription fidelity and termination efficiency (Jin et al. 1988; Zhou et al. 2013), resulting in an overall higher transcription efficiency (Jin et al. 1992). To measure the transcription efficiency of the

mutants and the ancestor at 42°C, we measured the relative abundance of an inducible fluorescent gene (*YFP*) inserted in our strains, at different points post induction (see *Material and Methods*). Using this method (Reynolds 2000; Brandis et al. 2012), we could not detect statistical differences in the transcription efficiency of the mutants relative to the ancestor at 42°C (Figure S2.1). Within the limits of our experiment, these data suggest that the growth improvement of the mutants at 42°C is not caused by higher transcriptional efficiency of RNAP.

Therefore, we addressed our second hypothesis that the *rpoB* mutations caused changes in GE that presumably result in growth improvements at 42°C. To test this hypothesis, we obtained RNAseq data from each mutant, with two replicates per mutant at mid-exponential phase after growth at 42°C and then contrasted the GE profile from each mutant against that of the ancestor grown at 42°C. All three mutations displayed hundreds to thousands of differentially expressed genes ($q < 0.001$; Figure 2.2). The mutant *I572F* had 1674 differentially expressed genes at 42°C with 161 highly down-regulated (\log_2 fold change < -2) and 195 highly up-regulated (\log_2 fold change > 2) relative to the 42°C ancestor. The mutant *I572L* had 987 differentially expressed genes with 148 highly down-regulated and 156 highly up-regulated. Finally, the mutant *I572N* had 1567 differentially expressed genes with 142 highly down-regulated genes and 101 highly up-regulated genes.

Two mutations restored GE, while the third reinforced the acclimation response.

To investigate the general trend in GE changes from the ancestor to acclimation and then to first-step mutations, we plotted the \log_2 fold expression change during the

acclimation response (ancestor grown at 42°C vs ancestor grown at 37°C, x-axis) against the log₂fold GE change during the adaptive response (mutant grown at 42°C vs ancestor grown at 42°C, y-axis; Figure 2.3). In a hypothetical situation in which the expression for all the genes of a mutant were restored by an adaptive mutation – such that the expression of each genes changes from a stressed state back to a pre-stressed state – the slope of the fitted regression between the acclimation response and the adaptive response would be close to -1.0. For two mutants we observed a highly significant negative correlation, with a slope of -0.746 for *I572F* and -0.672 for *I572L* (Figure 2.3A and B), indicating that the main phenotypic effect of these mutations was to restore global GE back toward the pre-stressed state. The third mutant (*I572N*) showed a drastically different pattern of expression change (Figure 2.3C). Instead of restoring GE back to an ancestral state, it reinforced (or exaggerated) the direction of GE changes during acclimation, so that the correlation had a significantly positive correlation (slope of 0.439; Figure 2.3C).

We next examined the GE changes in more detail. We characterized the expression of approximately half of the total number of genes in *E.coli* (4202 genes in total) into one of four patterns of change that denote the direction of the effect on the mutated RNAP (see *Materials and Methods*, Table S2.3). First, the expression of an individual gene could be *restored* back toward the pre-stressed state. Second, the expression of a gene could be *reinforced* into an exaggeration of the acclimated state, such that the mutated RNAP produced more transcripts (in the case of up-regulated genes) or fewer transcripts (in the case of down-regulated genes) than the acclimated state. Third, a gene was *unrestored* if the mutated RNAP did not change significantly

GE relative to the ancestral RNAP at 42°C. Finally, GE was *novel* if it was not differentially expressed during the acclimation response, but was expressed significantly differently during the adaptive response.

Following this categorization, we observed that most (72%) of the genes that were differentially expressed during acclimation, were restored by mutant *I572F* (blue dots in Figure 2.3A; Table 2.2), as expected from the strong overall negative correlation in Figure 2.3A. For mutant *I572L*, most genes were unrestored (53%), followed by an important proportions of genes (46%) with a restored expression (Table 2.2). For the *I572N* mutant, most (51%) GE was reinforced (red dots in Figure 2.3C), followed by substantial proportion (46%) of unrestored genes (Table 2.2). Interestingly, this mutant also had the highest number of genes with novel expression (yellow dots in Figure 2.3C; Table 2.2), further suggesting that *I572N* affects a different mechanism of adaptation compared to the *I572F* and *I572L*.

An *rpoB* mutation away from the active site of the RNAP also restores GE.

Given that two of our first-step mutations (*I572F* and *I572L*) converged toward restorative GE, we sought to know if additional mutations in *rpoB* had similar effects. To address this question, we constructed a single mutant, *rpoB I966S*, which alters one of the two parallel α -helices of the Eco flap domain of RNAP (Opalka et al. 2010). The *I966S* mutation was found in 15 of our high-temperature adapted clones (Tenailon et al. 2012), and its fitness advantage at 42°C was confirmed by competition experiments (Rodríguez-Verdugo et al. 2014).

We obtained RNAseq data from two replicates of this mutant at 42°C and performed the same GE analyses. We observed a global pattern of GE change similar to that for the *I572F* and *I572L* mutants. That is, the *I966S* mutation tended to restore GE from the acclimated state toward the pre-stressed state (Figure 2.3D).

To explore the convergent effects of the *I966S*, *I572F* and *I572L* mutations further, we determined the number of genes with restored GE shared among the three mutants. Half of the restored genes were shared between the three mutants, and 84% of the genes were shared by at least two mutants (Figure 2.4). The mutants *I572F* and *I966S* shared the highest proportion of genes (34%, or 478 shared restored genes), indicating a high level of expression convergence between them. This high level of phenotypic convergence was also highlighted by the pairwise comparisons of differential expression between the mutants (Figure S2.2).

Finally, we performed an enrichment analysis of GO assignments for the 860 restored genes that were shared among the three mutants. Not surprisingly, given the overall pattern of restoration (Figure 2.3), the restored genes represented the same sets of genes that were enriched for the acclimation response. For example, the significantly down regulated genes during the acclimation response, such as the genes involved in translation (*rpl*, *rpm* and *rps* genes), were significantly up regulated in the three mutants. In conclusion, the phenotypic convergence observed between the mutants *I572F*, *I572L* and *I966S* suggests that: *i*) restoring the altered physiological state back to a pre-stressed state was advantageous in the conditions of our experiment, *ii*) much of that restoration can be achieved by single, highly pleiotropic mutations and *iii*) similar effects

can result from the different amino acid mutations at the same site (codon 572 for mutants *I572F* and *I572L*) or at a different site (codon 966 for mutant *I966S*).

Reinforcement and novel GE

Puzzled by the drastically different pattern of GE observed for the mutant *I572N*, we also performed a GO enrichment analysis for the 1005 genes with reinforced expression. As expected, the significantly enriched GO categories matched the significant categories observed during the acclimation response (Table S2.1). That is, the *I572N* mutation tended to further increase expression of genes up-regulated during acclimation (e.g. fatty acid beta-oxidation; GO:0006635; $P = 1.12 \times 10^{-2}$) and further decreases expression of genes that were down-regulated during acclimation (e.g. biosynthesis of ribonucleosides; GO: 0042455; $P = 3.62 \times 10^{-4}$).

The *I572N* has the most genes (405) with novel expression phenotypes. Among these, we identified: i) a significant up-regulation of genes involved in iron-sulfur cluster assembly (GO: 0016226; $P = 1.12 \times 10^{-2}$), including the *suf* genes; ii) a significant up-regulation of genes involved in glycogen metabolic process (GO: 0005977; $P = 3.59 \times 10^{-2}$), including the *glg* genes; and iii) a significant down-regulation of genes involved in proteolysis (GO: 0006508; $P = 1.34 \times 10^{-2}$).

Finally, we hypothesized that if the genes with novel expression occurred in parallel between mutants and had the same direction of change (up-regulated or down-regulated), they may be functionally adaptive. We thus determined how many genes with novel GE were shared between the four single mutants: *I572F*, *I572L*, *I572N* and *I966S*. Although the mutant *I572N* had 405 genes with novel expression, there was

almost no overlap between the mutation *I572N* and the other mutations (Figure S2.3A), suggesting that the *I572N* mutation initiated a different mechanism of adaptation compared to the other three mutants. In contrast, the three restorative mutations (*I572F*, *I572L* and *I966S*) exhibited a high level convergence in GE, with 46% of novel genes shared by at least two mutants (Figure S2.3B).

Mutations fixed during adaptation contributed few changes in GE.

Finally, we compared the GE phenotype of one first-step mutation (*I572L*) to an end-product of our adaptation experiment (i.e. clones evolved 2000 generations at 42°C; Figure 2.5A). We chose two high-temperature adapted clones, clone 27 and clone 97 (clone numbers correspond to reference; Tenaillon et al. 2012), isolated from two populations in which the mutation *I572L* swept to fixation before 400 generations (Rodríguez-Verdugo et al. 2013). Although sharing the same *rpoB* mutation, these two high-temperature adapted clones differed in their genetic background (Table S2.4 and Table S2.5). Clone 27 had two large deletions of 2,896 and 71,416 bp affecting > 65 genes, a 138 bp deletion disrupting the tRNA-Met gene and seven point mutations in seven genes. In contrast, clone 97 had only one 4 bp small deletion, an IS insertion and eight point mutations in eight different genes (Tenaillon et al. 2012).

We again characterized the clones' growth curves at 42°C and compared them to both the ancestor grown at 37°C, the ancestor grown at 42°C and the mutant *I572L* grown at 42°C (Table 2.3, Figure 2.5 and Figure S2.4). The high-temperature adapted clones 27 and 97 both had significantly shorter lag phases compared to *I572L* mutant. In addition, clone 27 had a significant higher final yield than *I572L* mutant (Table 2.3). It

thus appears that the mutations accumulated during the adaptation experiment contribute to better growth at 42°C. We note, however, that none of the adapted clones grow as well as the ancestor at 37°C, which had a significantly shorter lag phase and a higher final yield (Table 2.3 and Figure 2.5B).

Knowing that the evolved clones grew better than the single mutant *I572L* at 42°C, we explored whether the mutations accumulated during later steps of adaptation caused changes in GE that could explain their growth improvement relative to the mutant *I572L*. We obtained RNA-seq data from each high-temperature adapted clone (two replicates per clone) grown at 42°C. We then contrasted the GE from each clone against the GE of the first-step mutation *I572L* at 42°C. We observed very few genes that were differentially expressed ($q < 0.001$): 63 for clone 27 (Figure 2.5C) and 16 for clone 97 (Figure S2.4). Furthermore, the high-temperature adapted clones maintained the general pattern of restoration back to the ancestral physiological state previously observed for the mutant *I572L* (slope of -0.705 for clone 27 and slope of -0.692 for clone 97; Figure S2.5). These observations suggested that the mutations accumulated in later steps of thermal stress adaptation did not substantially change the expression profile caused by the first-step mutation; that is, most shifts of GE were caused by the *I572L* mutation.

To emphasize this point, we plotted the \log_2 fold expression change during the first-step adaptive response (mutant *I572L* grown at 42°C vs ancestor grown at 42°C, x-axis; Figure S2.6) against the \log_2 fold expression change during the complete adaptive response (high-temperature adapted clone grown at 42°C vs ancestor grown at 42°C, y-axis; Figure S2.6). We observed a highly significant positive correlation (slope of 0.985

for clone 27 and slope of 1.025 for clone 97; Figure S2.6A and B), confirming high similarity in the overall GE pattern between the mutant *I572L* and the high-temperature adapted clones. Finally, when we contrasted the GE from the two high-temperature adapted clones at 42°C (clone 97 vs clone 27), there were only 70 differentially expressed genes ($q < 0.001$), of which 52 genes were part of the two large deletions in clone 27 (Figure S2.6C). Therefore the GE profile for the two high-temperature adapted clones was almost identical within the limitations of our experiment, despite their differences in genetic background. This observation confirms that the first step mutation *I572L* contributed to most of the changes in GE.

DISCUSSION

Two aspects of adaptation that have been largely unexplored are the temporal change of phenotypes during the adaptive process and the molecular mechanisms underlying these changes. Here we have focused on the phenotypic effects of first-step mutations during adaptation of *E. coli* to high temperature (42°C). One of the most significant findings of our study is that mutations in RNAP changed the expression of thousands of genes – most of which were differentially expressed during acclimation to 42°C – and conferred large fitness advantages. Subsequent mutations also increased fitness but did not substantially change GE. We have also discovered that the main phenotypic effect of three first-step mutations – *I572F*, *I572L* and *I966S* – was to restore global gene expression back towards the pre-stressed state. Remarkably, mutation *I572N* had a different phenotypic effect: it reinforced the GE changes during acclimation.

Heat-stress acclimation involves a balance between energy conservation and stress resistance.

The acute response to thermal stress, known as the heat-shock response, is universal and occurs in diverse organisms (Richter et al. 2010). In *E.coli*, the heat-shock response is transient (on the order of magnitude of minutes) and is characterized by the induction of stress related proteins mediated by the alternative σ^{32} factor (Nonaka et al. 2006). The σ^{32} regulon encodes: *i*) molecular chaperons (i.e. ClpB, DnaK, DnaJ, IbpB, GrpE, GroEL, GroES) that promote protein folding; *ii*) cytosolic proteases (i.e. ClpP, ClpX) that clear misfolded and aggregated proteins; *iii*) metabolic enzymes; *iv*) DNA/RNA repair enzymes; *v*) regulatory proteins; *vi*) proteins involved in maintaining cellular integrity; and *vii*) proteins involved in transport and detoxification (Nonaka et al. 2006; Richter et al. 2010). Although the heat-shock response has been widely studied, a question remains as to what happens to the expression of the heat-shock genes after hours or days of thermal stress (i.e. thermal acclimation).

Here we have characterized GE changes associated with phenotypic acclimation to mid-exponential phase at 42°C. We have found that most heat-shock induced genes (Nonaka et al. 2006; Richter et al. 2010) were not differentially expressed during acclimation and were, in fact, down regulated. For example, most of the heat-shock genes encoding chaperones, such as *clpB*, *dnaJ*, *dnaK*, *groEL* and *groES*, were down regulated during the acclimation response (Table S2.2). However, one exception is *spy*, which encodes a periplasmic chaperone and was up regulated during acclimation (\log_2 fold change = 5.0; $q < 0.001$; DatasetS1). Previous studies in *Saccharomyces*

cerevisiae and *E.coli* have reported a decrease in the expression of molecular chaperones after ~15 minutes of heat-stress induction (Eisen et al. 1998; Zhao et al. 2005; Jozefczuk et al. 2010; Richter et al. 2010). Therefore, it is possible that the heat-shock genes were initially expressed during our experiment, immediately after the transfer at 42°C, but were subsequently down-regulated.

Nevertheless, two studies exploring the physiological acclimation of *E.coli* to high temperature reported an up-regulation of heat-shock genes (Gunasekera et al. 2008; Sandberg et al. 2014). This discrepancy might be explained by differences in genetic backgrounds (*E.coli* K-12 strain vs *E.coli* B strain), by differences in temperature (42°C vs 43°C in Gunasekera et al. 2008) or by differences in culture conditions. For example, in our study bacterial clones were allowed to acclimate at 42°C for one day before being sampled the next day during mid-exponential phase at 42°C. Thus, the bacteria spent one complete cycle of growth (lag, exponential and stationary phase) at 42°C, in addition to the ~ half cycle of growth (lag and half exponential phase) at 42°C (in total ~ 1.5 days). For other studies, the time that the bacteria spent at high temperature before the RNA extraction is unclear (Sandberg et al. 2014).

Microorganisms often resist stressful conditions by modulating GE to limit growth (López-Maury et al. 2008). As a consequence, genes with growth-related functions, which are energy demanding, are down-regulated, allowing a redistribution of resources and energy to the expression of genes related to stress resistance (López-Maury et al. 2008; Jozefczuk et al. 2010; Jin et al. 2012). Surprisingly, we have observed a down-regulation of genes encoding different subunits of RNAP during the acclimation response. Assuming that lower expression also reflects protein abundance, this

observation implies that the ancestor contains fewer RNAP molecules when grown at 42°C than when grown at 37°C. The reduction of RNAP molecules can have important physiological consequences, given that, RNAP is limiting for genome-wide transcription (Ishihama 2000). Therefore, a reduction in the total number of RNAP would limit the transcription rate and favors our hypothesis that mutated RNAP have higher transcriptional efficiency at 42°C. That said, we did not detect any differences in transcription efficiency among wild type and mutated RNAPs (Figure S2.1). A limitation in RNAP molecules could also indirectly affect bacterial growth (Jin et al. 2012). For example, when growth conditions are unfavorable, RNAP molecules are released from the rRNA operons, reducing rRNA synthesis (i.e. reducing growth), so that more RNAP molecules become free and available for genome-wide transcription (Jin et al. 2012). Therefore, the reduction in growth at 42°C may be explained in part by the down-regulation of genes involved in translation and ribosome biogenesis, but also by a potential limitation of RNAP.

Other down-regulated genes involved in energy demanding processes were genes associated to biosynthesis of amino acids, nucleotides, ribonucleotides and constituents of the flagella (Table S2.1). A similar pattern of expression has been observed during acute exposure to thermal stress (Jozefczuk et al. 2010). In addition, metabolomic studies have reported a sharp decline in the levels of nucleotides in *E.coli* cultures exposed to heat stress (Jozefczuk et al. 2010; Ye et al. 2012). Therefore, our study suggests that the pattern of energy conservation mediated through down-regulation of energy demanding processes not only occurs during the acute response to stress but also occurs during the acclimation response.

We hypothesize that some of the resources and energy are redistributed to express genes controlled by the sigma S factor or σ^S , the master regulator of the stress response (Battesti et al. 2011). Using previous observations (Weber et al. 2005; Keseler et al. 2013), we generated a list of 66 genes expressed under several stressful conditions (including high temperature) and under regulation of σ^S . Of these, 62 genes (94%) were significantly up-regulated during acclimation to 42°C (Table S2.2). σ^S induced genes were related to: *i*) the synthesis molecules responsible to deal with the detrimental effects of stress, *ii*) transport systems, and *iii*) the production of metabolic enzymes, mostly involved in the central energy metabolism (Weber et al. 2005).

A large proportion of up-regulated genes (including the σ^S -regulated genes) were related to central energy metabolism (Hasan and Shimizu 2008; Jozefczuk et al. 2010; Ye et al. 2012). At 42°C the oxygen solubility is lower than at 37°C, which may result in less aerobic conditions and maybe even anoxic conditions (Hasan and Shimizu 2008; Jozefczuk et al. 2010). Given these conditions, we hypothesize that oxygen reduction contributed to the up-regulation of genes involved in energy metabolism (Ye et al. 2012). For example, we observed an up-regulation of genes encoding the fumarate reductase (*frdABCD*), which allows fumarate to serve as a terminal electron acceptor (Gunsalus and Park 1994). Conversely, we observed the down-regulation of genes involved in aerobic growth, such as the genes synthesizing respiratory enzymes (e.g. *cyoABCDE* encoding the cytochrome oxidase; Dataset 1). Therefore it is possible that the ancestor adjusts its physiology and metabolism in response to high-temperature and low oxygen by using alternative electron acceptors and by engaging in alternative modes of carbon catabolism, such as mixed-acid fermentation (Ye et al. 2012).

In conclusion, the ancestor acclimates to thermal stress by launching alternative GE programs (López-Maury et al. 2008). One involves the down-regulation of growth-related pathways, resulting in energy conservation. The other involves the up-regulation of stress genes involved in repair and metabolic adjustments to high temperature.

Restoration of the ancestral physiological state is advantageous.

The acclimated state was the starting point for adaptive evolution and the initial phenotypic state upon which natural selection could act. Our study shows that the first adaptive change during thermal stress adaptation moved to restore the GE to an ancestral state.

A single amino acid substitution in RNAP mediated global restoration of GE (Figure 2.3). Previous studies report that small deletions in RNAP change GE of hundreds of genes (Conrad et al. 2010). We have shown that a single amino acid substitution impacts GE for ~1500 genes. This observation suggests that *rpoB* mutations have important downstream effects that affect complex networks of interacting genes and their products (i.e. global regulators of GE or “hubs”; (Barabasi and Oltvai 2004)). Therefore, our study adds to previous observations indicating that genes encoding global regulators are a main target of natural selection in experiments on microbial laboratory evolution (Philippe et al. 2007; Hindre et al. 2012).

There are at least two pieces of evidence that suggest that the restoration of GE was advantageous in our experiment, leading to the rapid fixation of *rpoB* mutations in evolved populations (Rodríguez-Verdugo et al. 2013). First, the general trend of restoration occurred in parallel in three mutants (*I572F*, *I572L* and *I966S*). Phenotypic

convergence is commonly interpreted as sign of adaptive evolution (Christin et al. 2010). Not only did we observe convergence in global GE (Figure 2.3), but we also observe an important overlap of genes with restored expression (Figure 2.4). Interestingly, the three mutations show an asymmetric restoration in GE. That is, these mutations more often restored the expression of genes that were down-regulated during acclimation than genes that were up-regulated. For example, if we take all the genes with restored expression in *I572F* (1429 genes), 60% of these genes were down-regulated during the acclimation response and 40% were up-regulated during the acclimation response. These proportions are the same for the other two mutants (*I572L and I966S*). Convergence also suggests similar molecular mechanisms of action; for example, they might have similar binding affinities to similar regions of the genome (Haugen et al. 2008).

Second, restoration in GE is most likely advantageous because it “reactivates” the growth-related functions that were down-regulated during the acclimation response. The GO categories related to growth, such as translation and ribosomal biogenesis, were significantly enriched in the three mutants. The up-regulation of growth-related genes might explain why the mutants have shorter lag phase and the longer exponential phase than the ancestor (Figure 2.2).

Shortening the duration of the lag phase prior to growth is one the life history-traits that increases fitness in a selective environment (Vasi et al. 1994; Lenski et al. 1998).

Therefore, the differences in the mutants (Figure 2.2) probably explain their higher fitness relative to the ancestor when competed in the same environment (i.e. relative fitness; Rodríguez-Verdugo et al. 2013; Rodríguez-Verdugo et al. 2014).

Reinforcement and novelty: an alternative evolutionary strategy.

Restorative changes in GE seem to be a general evolutionary trend in laboratory adaptation experiments, while novel and reinforced changes are less common (Fong et al. 2005; Carroll and Marx 2013; Sandberg et al. 2014). Although we observed a similar trend for three of our mutants (*I572F*, *I572L* and *I966S*), the mutant *I572N* had a different pattern of GE. Most of its expression changes were reinforced (red dots in Figure 2.3C), meaning that the shift in GE from the ancestor to the acclimated state was further exaggerated by the *I572N* mutant. In addition, this mutant caused the highest number of novel GE changes (yellow dots in Figure 2.3C). This was a surprising result because the mutation affects the same codon as the mutations *I572F* and *I572L*. These differences in GE might be caused by differences in the chemical characters of the amino acid side chains. Phenylalanine (*F*) and Leucine (*L*) have hydrophobic side chains, as does Isoleucine (*I*), while Asparagine (*N*) has hydrophilic side chain (Petsko and Ringe 2004).

It is not clear why the *I572N* mutation confers a fitness advantages in our experiment given that most of its genes with growth related functions are down-regulated compared to the ancestor grown at 42°C (i.e. reinforced). One hypothesis is that the mutation reinforced the expression of genes related to hypoxic metabolism conferring an advantage in our experimental conditions.

Pleiotropy and side effects

Even if four *rpoB* mutations (*I572F*, *I572L*, *I572N* and *I966S*) were advantageous in the conditions of our experiment, they had pleiotropic effects that lead to changes in other phenotypes, such as fitness trade-offs at low temperatures (Rodríguez-Verdugo et al. 2014). Pleiotropy can also affect GE; for example, mutations in global regulator often have maladaptive side effects (Hindre et al. 2012). Therefore, later adaptive mutations might compensate for these maladaptive side effects (Hindre et al. 2012). In our study we observed two examples of possible compensatory changes.

First, we observed that the genes involved in flagellum-dependent cell motility (e.g. *flg* genes), which were highly down-regulated during the acclimation response ($\log_2\text{fold} < -3$ and $q < 0.001$), were up-regulated in the single mutants *I572F* and *I572L* ($q < 0.001$; Database 1). In the conditions of our evolution experiment — a well-mixed environment lacking physical structure — motility seems unnecessary. Given that the biosynthesis of flagella is costly (Soutourina and Bertin 2003), reducing the expression *flg* genes might be beneficial (Cooper et al. 2003; Fong et al. 2005). We therefore posit that the restoration in GE of *flg* genes might be costly and have deleterious effects. Interestingly, some of the *flg* genes were down-regulated ($q < 0.05$) in clone 97 when compared to the first-step mutant *I572L*. Therefore, later adaptive mutations might contribute to the fine-tuning of GE by compensating the side effects of restoration. We note, however, that we have not yet identified the mutation that causes the down-regulation of *flg* genes in clone 97. That being said, the up-regulation of flagellar genes after restorative shifts in GE and the fixation of subsequent mutations that down-regulate them has been observed previously (Sandberg et al. 2014).

Restorative changes in GE are not the only changes with maladaptive side effects; novel changes in GE might also have detrimental effects. One example is the case of the novel expression of genes involved in enterobactin biosynthesis (*ent* genes), observed in the mutants *I572F*, *I572L* and *I966S* (Figure S2.3B). Enterobactin is a type of siderophore used to scavenge iron from the environment (Crosa and Walsh 2002). The functional benefit of producing siderophores in our culture conditions is not obvious, and it seems that this may be energetically costly to the cells. Interestingly, the *ent* operon forms part of the large deletion of clone 27 (Table S2.4), which was observed in 35 of the 114 high-temperature adapted clones (Tenailon et al. 2012). It seems possible that the deletion of iron acquisition genes could compensate the cost of producing siderophores, but a robust test of this hypothesis requires further experimentation.

We observed that the first-step mutation *I572L* contributed to most of the GE variation during thermal stress adaptation (987 differentially expressed genes), while later mutations contributed to fewer changes in GE (63 and 16 differentially expressed genes in clones 27 and 97, respectively; Figure 2.5 and Figure S2.4). Interestingly, these few changes in GE contributed to significant changes in growth parameters (Table 2.3). Therefore, the number of genes differentially expressed was not proportional to the fitness gains.

This “disconnect” between the number of genes differentially expressed and the magnitude of the fitness advantage might be caused by the pleiotropic side effects of *I572L*. Under this model, we presume that some of the hundred of differentially expressed genes in *I572L* have beneficial effects, while others have deleterious effects,

netting an overall beneficial change in GE. If true, it is likely that subsequent mutations change the expression of only few genes, but most of these changes are beneficial. For example, the large deletion in clone 27 contained genes involved in iron acquisition (*fep* and *ent* operons; Crosa and Walsh 2002), and copper and silver efflux system (*cus* operons; Long et al. 2010). Costly, non-functional pathways are often “shut-down” in order to save energy that would be otherwise used to produce unnecessary proteins and metabolisms (Cooper et al. 2001; Lewis et al. 2010). Therefore the large deletion of 71 kb in length may be an energetic benefit, explaining its occurrence in 35 high-temperature adapted clones (Tenaillon et al. 2012).

Clone 97, which lacked the large deletions of clone 27, displayed fewer changes in gene expression than clone 27. In this clone, one of the significantly down-regulated genes ($q < 0.001$) was the *rmf* gene, which encodes the ribosome modulation factor (RMF). RMF has been associated with decreased translation activity and is expressed during slow growth conditions, such as stationary phase (Polikanov et al. 2012). Therefore, the down-regulation of *rmf*, occurring in parallel in clone 27 and 97, might increase protein synthesis and thus growth.

Concluding remarks

Mutations in global regulators of GE are observed recurrently in laboratory evolution experiments (Applebee et al. 2008; Goodarzi et al. 2009; Kishimoto et al. 2010). It is not always clear if these mutations represent the first step of an adaptive walk or later step, but at least in the case of *rpoB* and *rpoC* mutations, it seems they are often first-step mutations (Herring et al. 2006). Therefore, the pattern that we have observed in our

study may not be specific to our system but instead a more general phenomenon. Based on our results, we propose a general, two-step adaptive process. First, a mutation affecting global transcriptional regulator appears in the population and changes global GE. This expression change is mostly restorative, so that the stressed state returns to a pre-stressed (ancestral) state. These changes in GE confer a high advantage promoting the rapid fixation of the mutation in the population. Once the cell recovers its “normal” homeostatic state, other mutations accumulate and contribute to novel functions (fine-tuning of adaptive traits) or/and compensate for the side effects of the first-step pleiotropic mutation. Future directions would be to confirm this pattern by performing time course studies of GE (including heat shock and acclimation response, as well as all the intermediate steps of adaptation) coupled with genomic data.

MATERIALS AND METHODS

Growth conditions

Unless otherwise noted, the culture conditions used for the physiological assays (growth curves, transcription efficiency and RNA-seq assays), were the same used during the high-temperature evolution experiment (Tenaillon et al. 2012). Briefly, strains were revived in LB and incubated at 37°C with constant shaking (120 rpm). Overnight cultures were diluted 10⁴-fold into 10 ml Davis minimal medium supplemented with 25 µg/ml glucose (DM25) and incubated 1 d at 37°C to allow the strains to acclimate to the culture conditions. The following day, we transferred 100 µl of the overnight culture in 9.9 ml of fresh DM25 (100-fold dilution) and we incubated them at 42°C for one day to allow the strains to acclimate to high temperature.

Growth curves

Acclimated strains were grown at the assay temperature (either 37°C or 42°C) and the densities were measured approximately every hour during the lag phase and every ~20 minutes during the exponential phase. Population densities were measured using an electronic particle counter (Coulter Counter model Multisizer 3 equipped with a 30- μ m-diameter aperture tube). To measure density, 50 μ l of culture was diluted in 9.9 mL IsotonII diluent (Beckman Coulter), and 50 μ l of the resulting dilution was counted electronically. The maximum growth rate was estimated from the log-linear phase of growth for three replicate cultures of each strain using R version 3.0.2 (Team 2013).

RNA extraction and preparation for sequencing

To investigate the genes expression profile prior to thermal stress, we grew the ancestor (REL1206), previously acclimated to the growth conditions, at 37°C until it reached mid-exponential phase (three replicates). To investigate the GE profile during acclimation, we grew the ancestor at 42°, previously acclimated to 42°C (three replicates), until mid-exponential phase. The remaining strains (single first-step mutants and high-temperature adapted clones) were grown at 42°C until mid-exponential phase, with two replicates for each.

Briefly, bacterial cultures were grown in DM25 medium until they reached mid-exponential phase, which we determined by electronic counts. 80 ml of culture was filtered through a 0.2 μ m cellulose membrane (Life Science, Germany). Cells, concentrated in the filter, were stabilized with Qiagen RNA-protect Bacteria Reagent

and pellet for storage at -80°C prior to RNA extraction. Total RNA was extracted using RNeasy Mini Kit (Qiagen). Total RNA was DNase treated using Turbo DNA-free kit (Ambion) and rRNA was depleted using the Ribo-Zero rRNA Removal kit for Gram-Negative Bacteria (Epicentre Biotechnologies, Medion, WI, USA). cDNA library was constructed using TruSeq RNA v2 kit (Illumina, San Diego, CA, USA). Libraries were multiplexed 8-fold and sequenced on an Illumina HiSeq 2000 platform. 100-bp single-end reads were generated.

mRNAseq data analyses

Reads were mapped to the *E.coli* B REL606 genome reference (CP000819.1) using *bwa* 0.6.2 (Li and Durbin 2009), using default parameters (<http://bio-bwa.sourceforge.net/bwa.shtml>). Only unique, perfectly matching reads to the 4204 annotated coding regions were retained for further analyses. Differential expression analysis was performed using the DESeq R package (Anders and Huber 2010). We used the *P*-values adjusted by the Benjamini and Hochberg approach (*q* values), which controls for false discovery rate. Genes with *q* less than 0.001 were considered significantly differentially expressed. Differentially expressed genes (DEG) were classified in one of the four categories (restored, reinforced, unrestored, and novel) based on the contrast for which they were significant and the direction and value of their fold change (Table S2.3).

For the GO term enrichment analyses, we used the Enrichment analysis tool from <http://geneontology.org/page/go-enrichment-analysis>.

Construction of the fluorescently labeled strains for the transcription efficiency assay

We generated a ~4 kb-long linear DNA fragments carrying the *CAT*, *tetR* and *YFP* genes from an *E.coli* strain carrying the *CAT:tetR:YFP* genomic cassette (Fehér et al. 2012) kindly provided by Csaba Pál (Biological Research Center, Szeged, Hungary). We amplified the *CAT;tetR;YFP* cassette by PCR using the primers ARV19 and ARV20 (Table S2.6) and *Pfu* DNA polymerase (Promega). The purified PCR product was integrated into the ancestral strain REL1206 carrying the pKD46 recombineering plasmid as previously described by Datsenko and Wanner (Datsenko and Wanner 2000). Briefly, the ancestral strain carrying the pKD46 plasmid was grown overnight at 30°C in 5 ml of LB with 100 µg/ml of ampicillin. The overnight culture was 100 fold-diluted in 100 ml of LB with ampicillin and 1 mM L-arabinose (Sigma) and grown at 30°C to an OD₆₀₀ of 0.6. Electrocompetent cells were made by washing the culture 5 times with ice-cold water. ~200 ng of linear DNA was electroporated into 25 µl of cells. After electroporation, 1 ml of LB was added, and the cells were incubated at 30°C for 2 h with shaking, then plated 100 µl on LB agar plates with chloramphenicol (20 µg/ml). We selected a single colony and purified it in LB agar plate containing chloramphenicol. Correct integration was verified by PCR using primers ARV34 and ARV35.

Transcription efficiency assay

To measure the transcription efficiency of the ancestor and the mutants we performed a quantitative reverse transcriptase RT-PCR assay (Reynolds 2000; Brandis et al. 2012). In brief, we grew the fluorescently labeled strains in the same culture

conditions previously described except that we supplemented the DM medium with 100 µg/ml glucose (DM100). We confirmed the advantage of these mutations to high temperature despite the higher amount of glucose based on growth curves of the ancestor and the mutants at 42°C in DM100. Acclimated cultures were grown at 42.2°C until they reached mid-exponential phase. We took 1 ml of uninduced cells (sample T₀) and stabilized them in RNAprotect Bacteria Reagent (Qiagen). Immediately after, we induced cells by adding 10 µl of anhydrotetracycline (66 µg/ml) to the medium. Samples of 1 ml of culture were taken at 1, 2, 3 and 4 min after induction (samples T₁, T₂, T₃ and T₄) and stabilized with RNA protect Bacteria Reagent and pelleted for storage at -80°C prior to RNA extraction. Total RNA was prepared using RNeasy Mini Kit (Qiagen). RNA was DNase treated using Turbo DNA-free kit (Ambion). We used 300 ng of DNA-free RNA to produce cDNA with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). From each reverse transcribed product we quantified the abundance of cDNA of the gene reference *gst* (Gst, glutathione transferase; (Pfaffl 2001)) and the target gene YFP (Yellow fluorescent protein) using Fast SYBR Green Master Mix (Applied Biosystems) quantified on a Stratagene MX3005P QPCR System (Agilent Technologies). For each PCR reaction we used 0.625 µM forward primer (ARV48 for *gst* or ARV50 for YFP, Table S2.6), 0.625 µM reverse primer (ARV49 for *gst* or ARV51 for YFP, Table S2.6), 3 µl CDNA template, 4.5 µl RNase-free water and 10 µl of Fast SYBR Green Master Mix, to have a final reaction volume of 20 µl. To control for the intra-assay variation (repeatability), we prepared three replicates of each reaction. The PCR thermal cycling conditions were 95°C for 20 sec followed by 40 cycles of 95°C 3 sec and 60°C 30 sec.

The efficiency of the amplifications for each pair of primers was determined from a standard curve using the formula $E = 10^{[-1/s]}$, where s is the slope of the standard curve (Pfaffl 2001). To calculate the *Relative Expression Ratio* (i.e. the relative change in GE of the target gene YFP normalized to the reference gene *gst* and relative to the uninduced control sample T_0), we used the mathematical model for relative quantification in real-time RT-PCR developed by Pfaffl, 2001 (Pfaffl 2001; Brandis et al. 2012). Transcription efficiency was calculated as the slope of the fitted linear regression between the *Relative Expression Ratio* against time, based on three replicates.

FIGURES

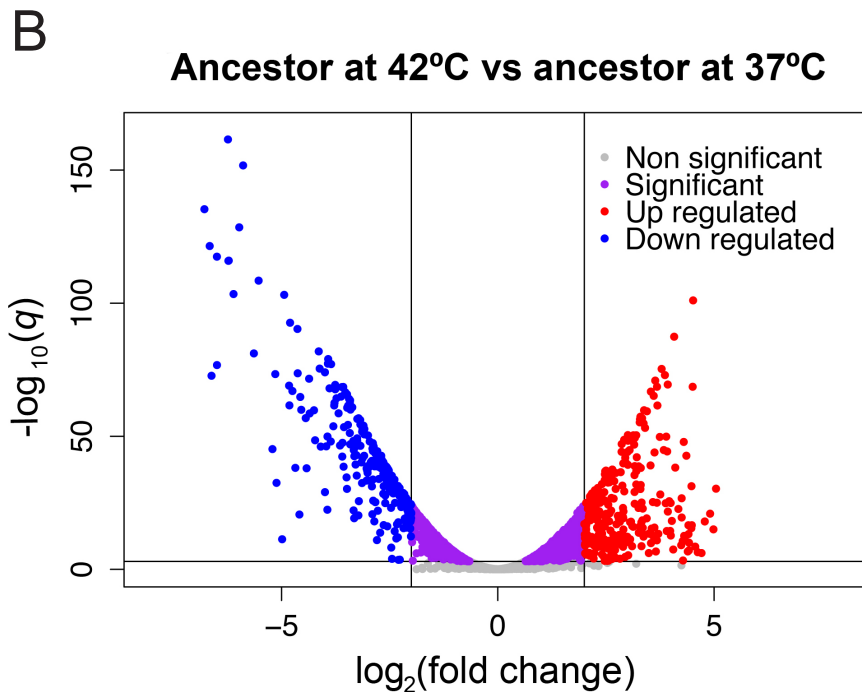
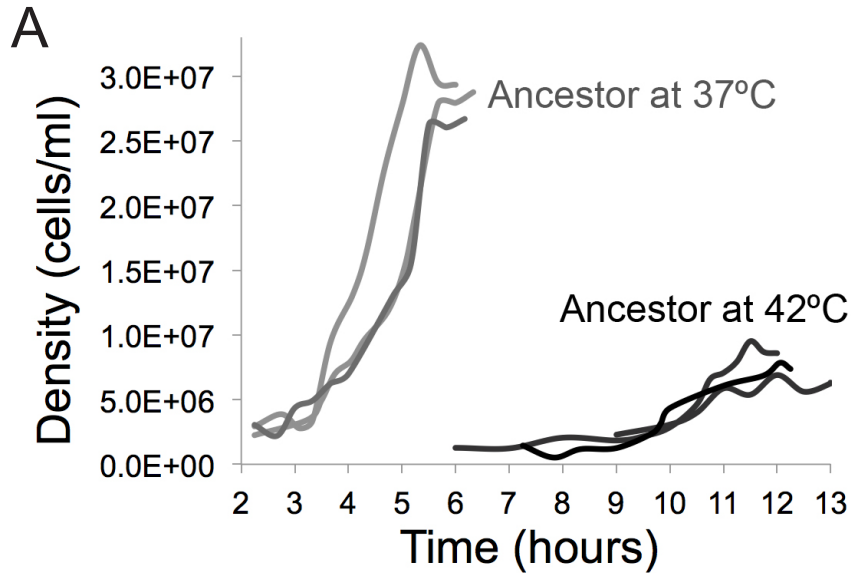


Figure 2.1. Phenotypic characterization of the ancestor at 37°C and 42°C. A) Growth curves of the ancestor grown at 37°C and 42°C (three replicates at each temperature). B) Volcano plot showing the global differential expression of genes (represented as dots) between the ancestor grown at 42°C compare to the ancestor grown at 37°C. Colors represent status with respect to 2-fold expression difference, represented by two vertical lines, and a significance at $q = 0.001$, represented by an horizontal line.

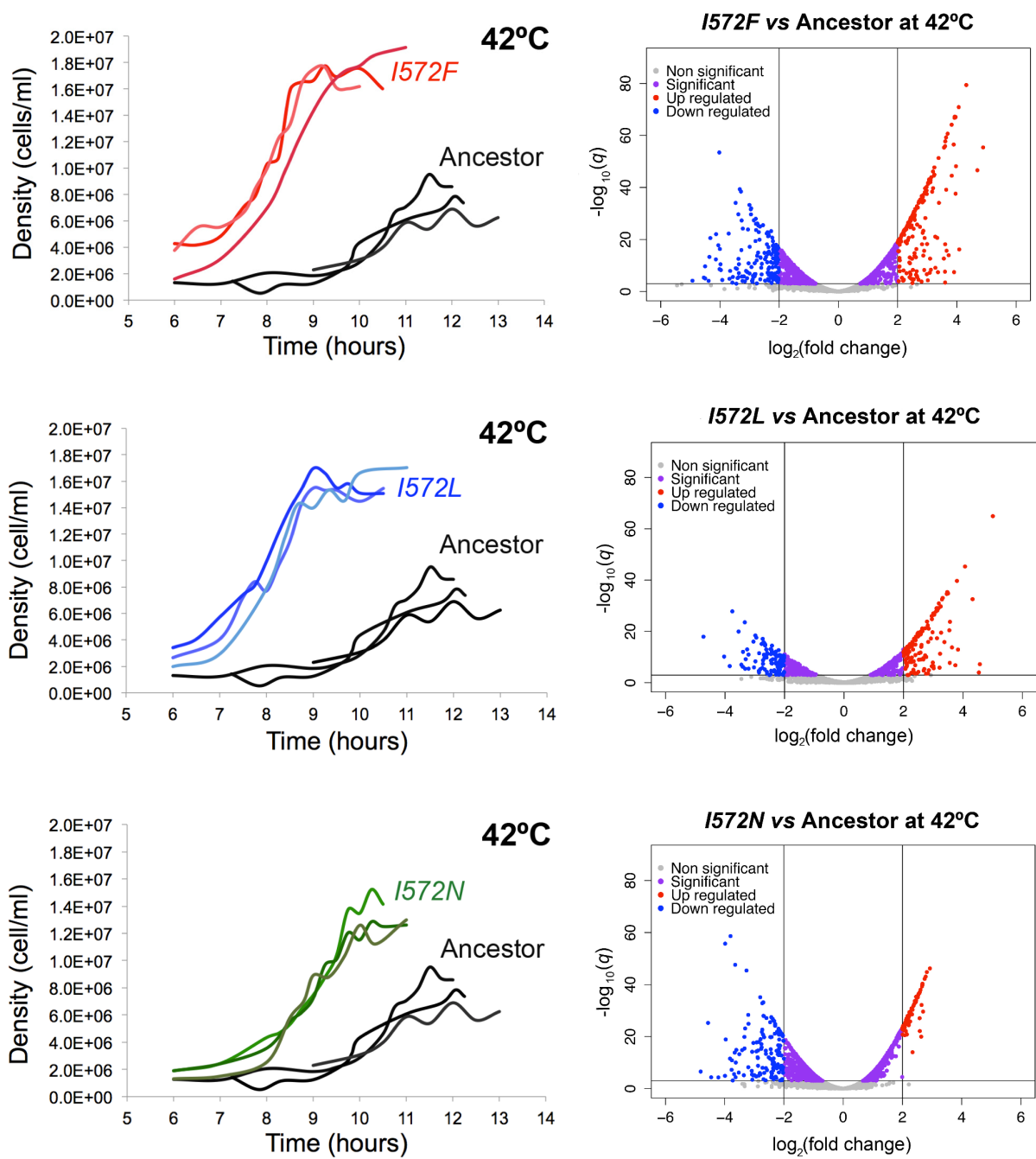


Figure 2.2. Phenotypic characterization of the mutants compared to the ancestor at 42°C.

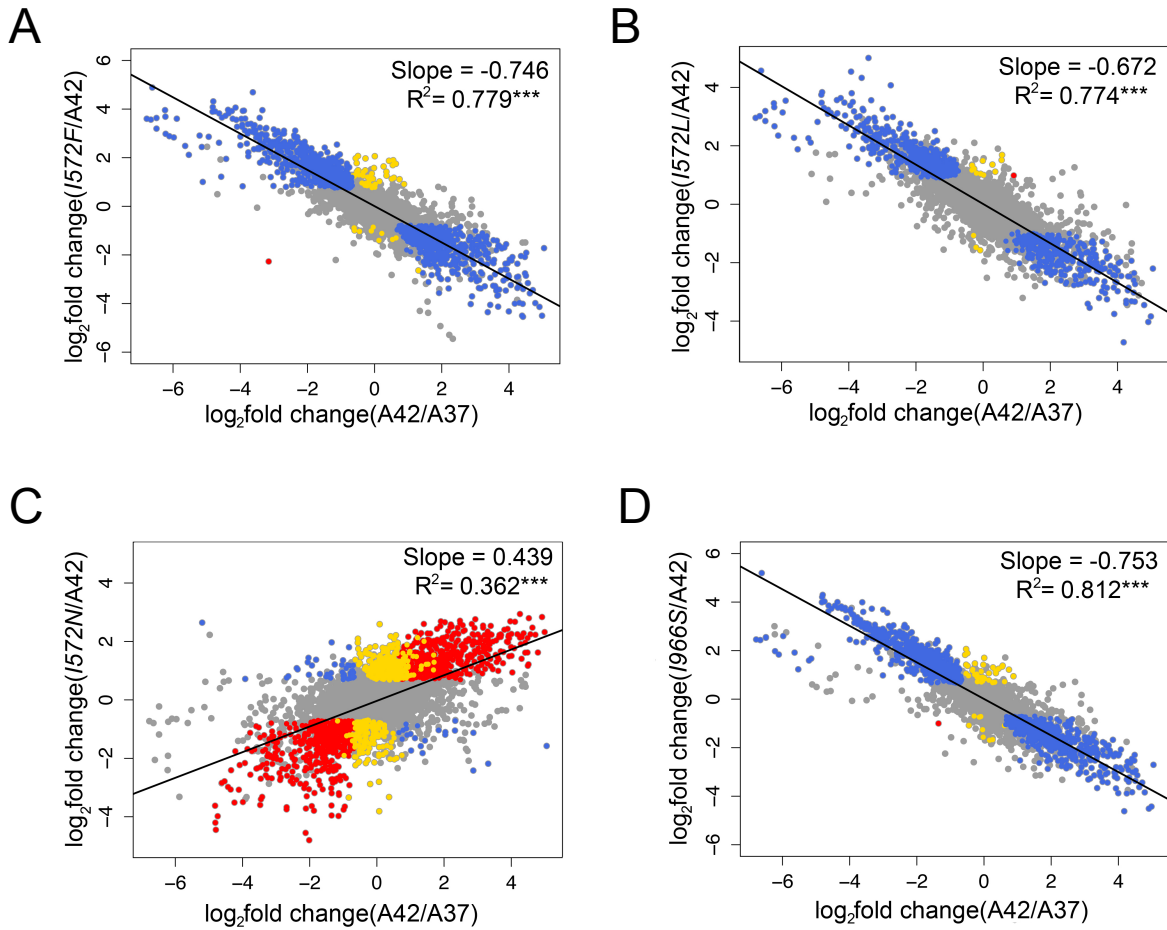


Figure 2.3. Global changes in GE during the acclimation and adaptive response. In all the graphs the x-axis represents the acclimation changes (ancestor grown at 42°C vs ancestor grown at 37°C). The y-axis represents the changes at different steps of the adaptive walk: first-step adaptive mutations (*I572F*, *I572L*, *I572N*) and a single mutant (*I966S*) vs ancestor grown at 42°C. Changes in expression were categorized and colored as follows: restored (blue), reinforced (red) and novel (yellow). Both unrestored and uninformative genes are colored in grey. The black line represents the linear regression fitted to ~ 4163 dots in each graph.

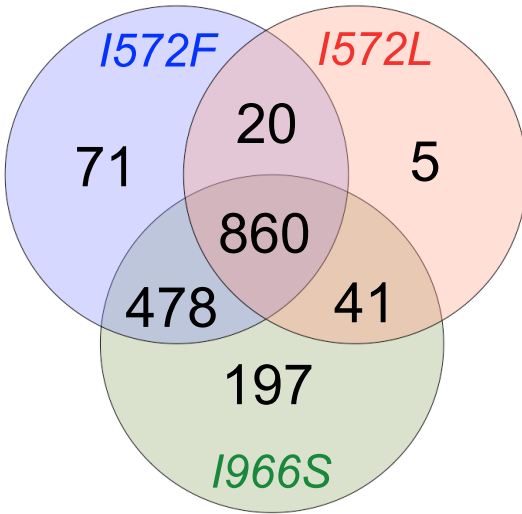


Figure 2.4. Convergence of genes with restored expression in single mutants. Number of shared genes with restored expression.

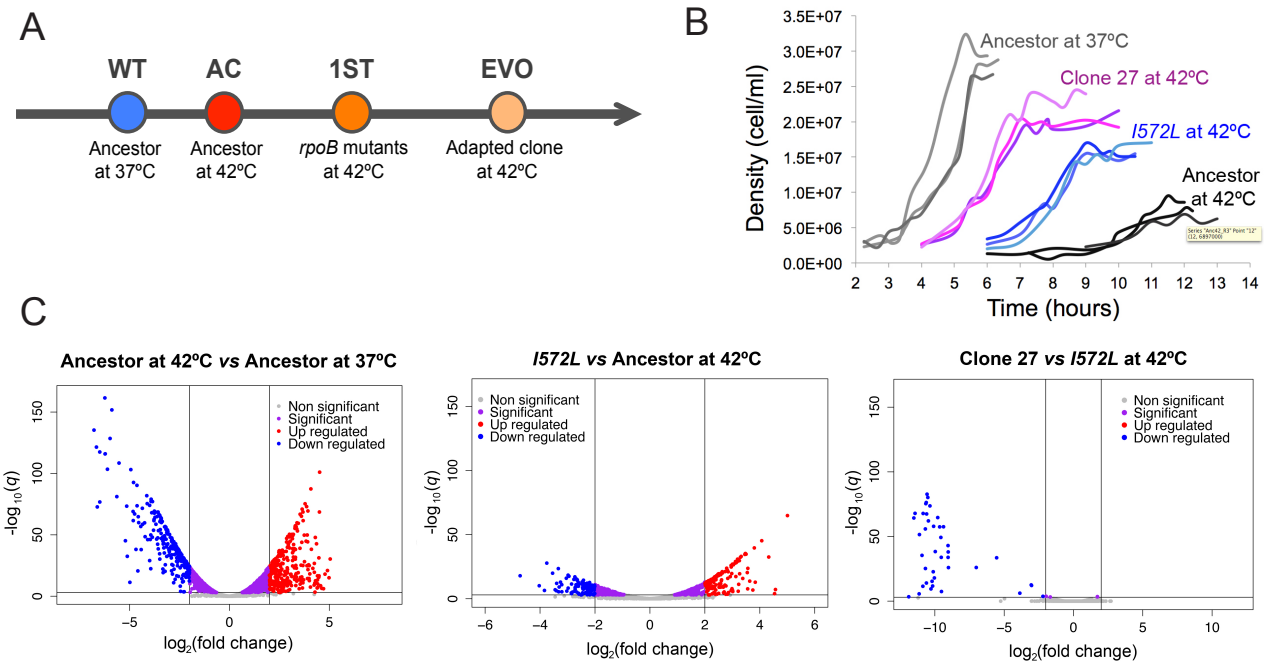


Figure 2.5. Phenotypic changes during thermal stress adaptation of clone 27. A) Thermal stress adaptation. B) Growth improvements during the heat stress adaptive walk. C) Changes in GE during the acclimation and adaptive response.

TABLES

Table 2.1. Growth parameters of the first-step mutants and the ancestor.

Strain	μ_{\max}^1			τ^3			Yield ⁴		P^2
	37°C	42°C	P^2	37°C	42°C	37°C	42°C		
Ancestor	1.160	0.923	0.391	3.2	9.9	5.6e10⁻⁷	2.81*10 ⁷	8.42*10 ⁶	0.002
SE	(0.213)	(0.100)		(0.1)	(0.1)		(9.89*10 ⁵)	(1.56*10 ⁵)	
<i>I572F</i>	0.898	0.734	0.136	4.1	6.7	0.011	2.44*10 ⁷	1.73*10 ⁷	0.009
SE	(0.071)	(0.021)		(0.3)	(0.4)		(1.08*10 ⁶)	(6.45*10 ⁵)	
P^5	0.344	0.196		0.083	0.015		0.065	0.003	
<i>I572L</i>	0.923	0.844	0.428	3.9	6.8	0.001	2.71*10 ⁷	1.57*10 ⁷	0.004
SE	(0.032)	(0.079)		(0.1)	(0.2)		(1.05*10 ⁶)	(4.24*10 ⁵)	
P^5	0.382	0.571		0.002	5.4e10⁻⁴		0.526	0.001	
<i>I572N</i>	1.399	0.616	0.042	3.5	8.2	2.2e10⁻⁵	3.09*10 ⁷	1.29*10 ⁷	0.013
SE	(0.189)	(0.066)		(0.1)	(0.1)		(2.39*10 ⁶)	(5.49*10 ⁵)	
P^5	0.451	0.073		0.116	0.002		0.363	0.010	

¹ Maximum growth rate with standard error in parenthesis.

² Significance value from a two-sample t-test. The null hypothesis is that the mean at 37°C and the mean at 42°C are equal.

³ Duration of the lag phase with standard error in parenthesis.

⁴ Final yield (cells/ml) with standard error in parenthesis.

⁵ Significance value from a two-sample t-test. The null hypothesis is that the mean of the ancestor and the mean of the mutant are equal. Numbers in bold correspond to P -values <0.05.

Table 2.2. Classification of the genes into four patterns of expression change.

Category	<i>I572F</i>	<i>I572L</i>	<i>I572N</i>	<i>I966S</i>
1) Restored	1429	926	70	1576
2) Reinforced	1	1	1005	1
3) Unrestored	554	1057	909	407
4) Novel	54	14	405	49

Table 2.3. Growth parameters of the high-temperature adapted clones 27 and 97 at 42°C compare to the mutant *I572L* and the ancestor at 37°C and 42°C (values in Table 2.1).

Strain	μ_{\max}^1	τ^2	Yield ³
27	0.938 (0.222)	5.0 (0.03)	2.09*10 ⁷ (1.22*10 ⁶)
97	0.935 (0.062)	5.2 (0.12)	1.87*10 ⁷ (2.20*10 ⁶)
Comparison	Significance ⁴		
27 vs ancestor at 42°C	0.694	4.4e10 ^{-5***}	0.001**
97 vs ancestor at 42°C	0.521	1.5e10 ^{-5***}	0.024*
27 vs <i>I572L</i> at 42°C	0.720	0.007**	0.039*
97 vs <i>I572L</i> at 42°C	0.418	0.002**	0.302
27 vs ancestor at 37°C	0.510	5.5e10 ^{-4***}	0.011*
97 vs ancestor at 37°C	0.403	3.3e10 ^{-4***}	0.035

¹ Maximum growth rate with standard error in parenthesis.

² Duration of the lag phase with standard error in parenthesis.

³ Final yield (cells/ml) with standard error in parenthesis.

⁴ Significance value from a two-sample t-test. The null hypothesis is that the means are equal.

SUPPORTING INFORMATION

Supporting Figures

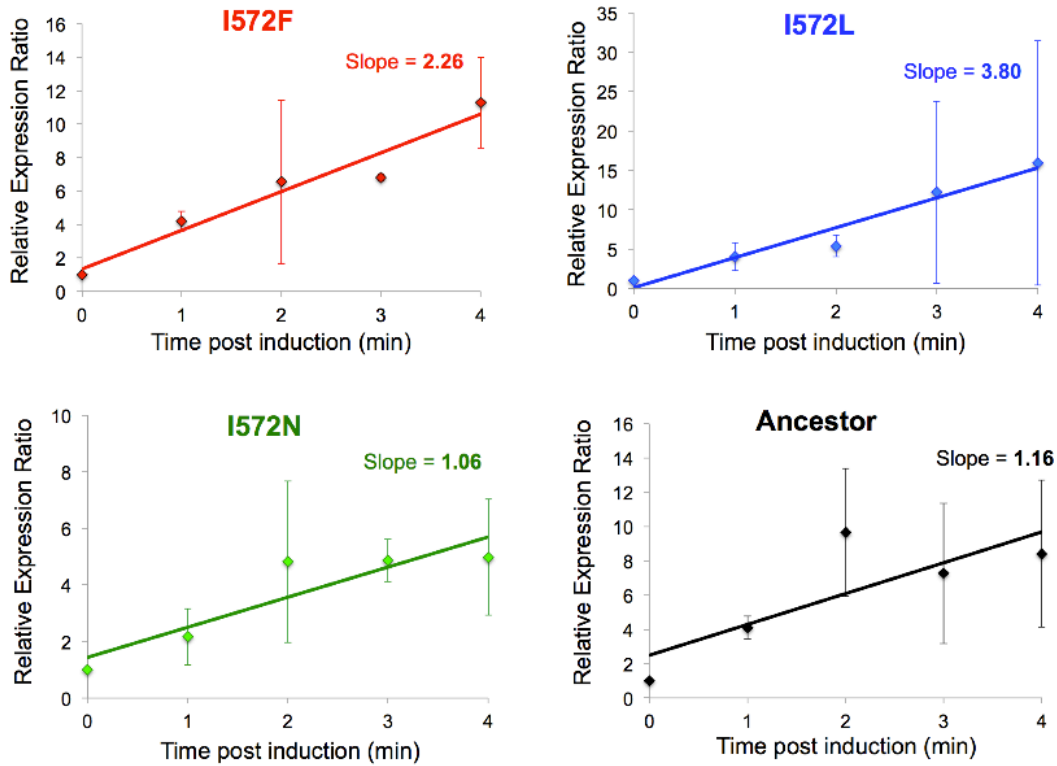


Figure S2.1. Transcription efficiency of the mutants and the ancestor at 42°C. Each diamond represents the mean of three independent measurements and the bar indicates the associated standard deviation. The slope of the fitted regression represents the overall transcription efficiency of RNAP. We didn't observed statistical differences between the transcription efficiency of each mutants and the transcription efficiency of the ancestor based on a test of equality of regression parameters.

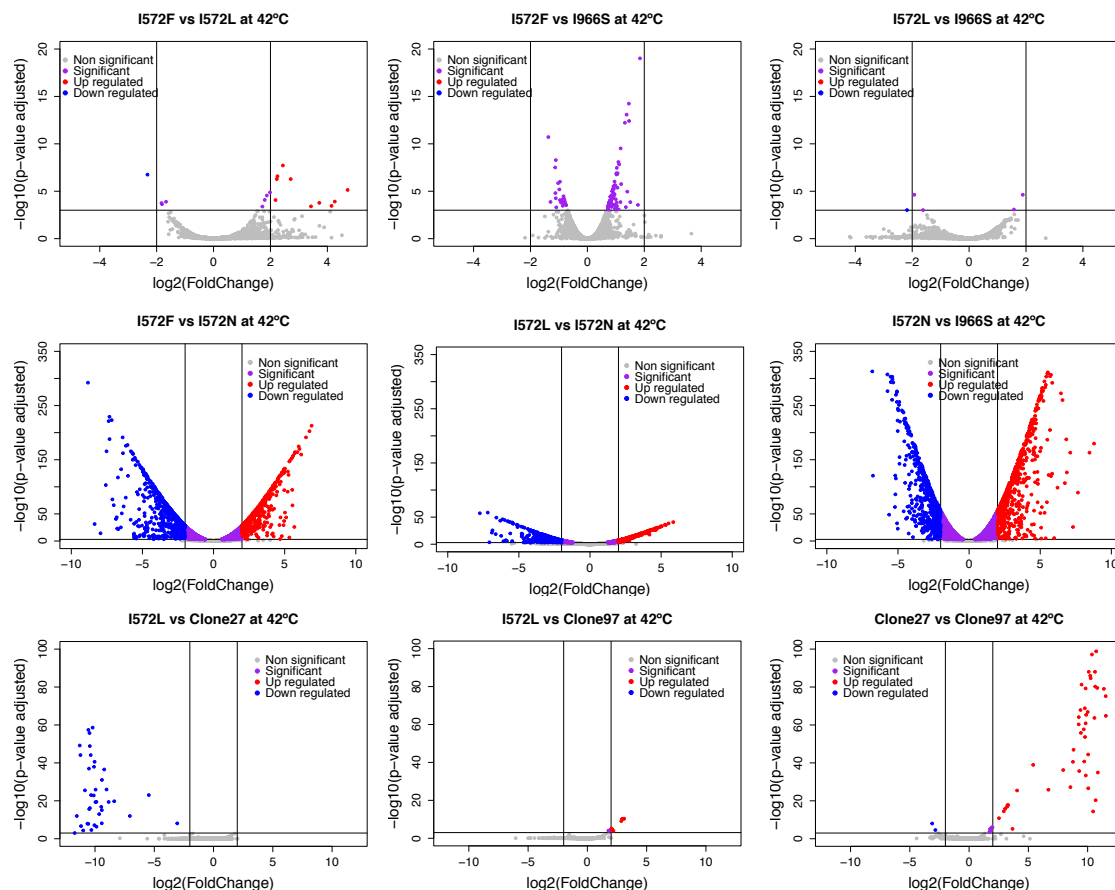


Figure S2.2. Volcano plots showing the differential expression of genes for the pairwise comparisons between mutants and high-temperature adapted clones.

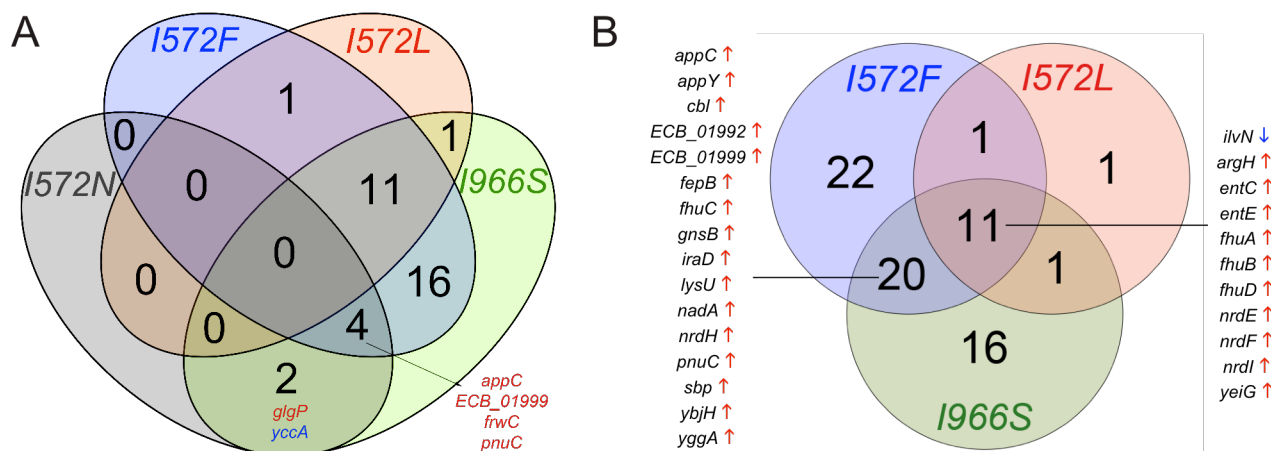


Figure S2.3. Convergence of genes with novel expression in single mutants. Number of shared genes with novel expression, in the four mutants (A), or in only three mutants (B).

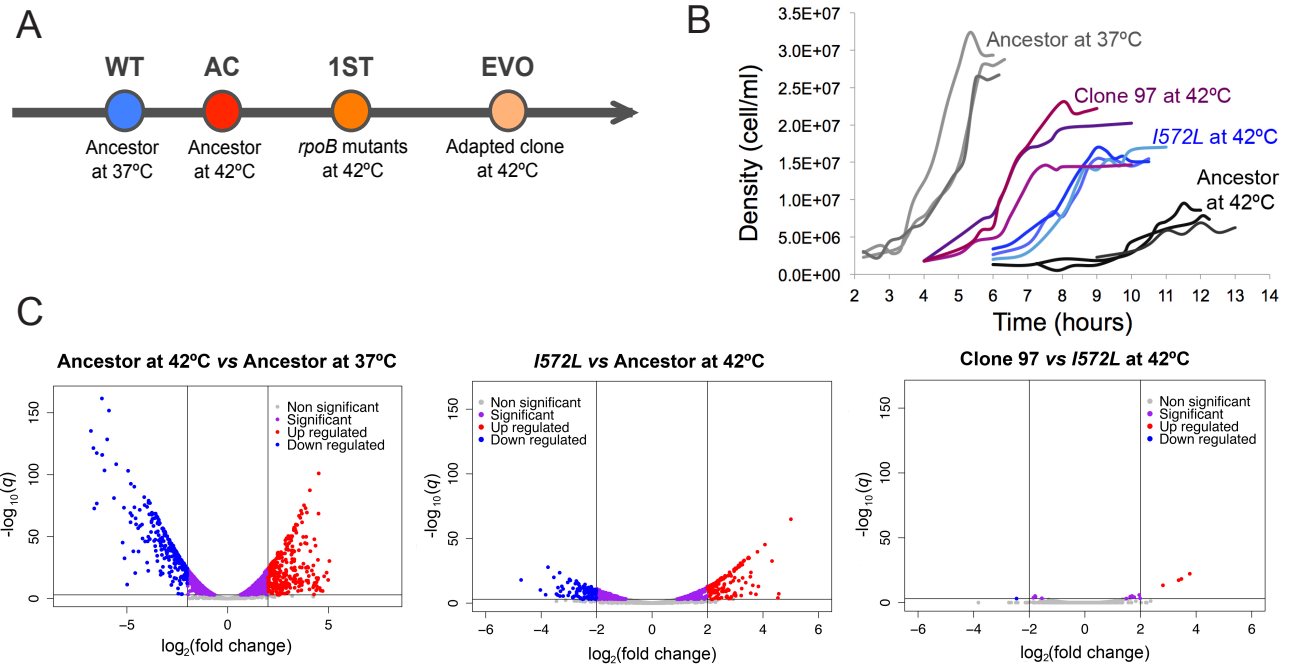


Figure S2.4. Phenotypic changes during thermal stress adaptation (clone 97). A) Thermal stress adaptation. B) Growth improvements during the heat stress adaptive walk. C) Changes in GE during the acclimation and adaptive response.

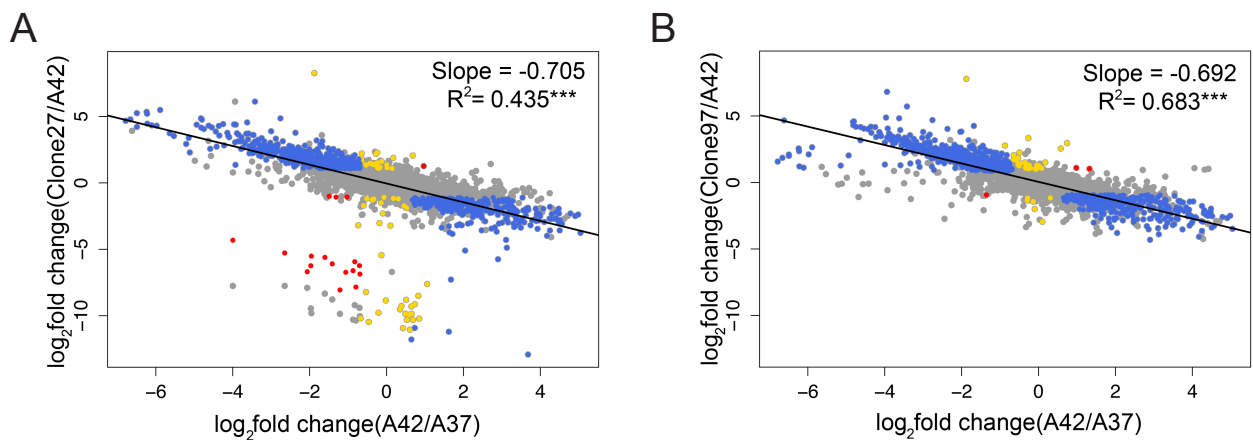


Figure S2.5. Global changes in GE during the acclimation and adaptive response. In all the graphs the x-axis represents the acclimation changes (ancestor grown at 42°C vs ancestor grown at 37°C). The y-axis represents the changes during thermal adaptation: clone 27 vs ancestor grown at 42°C (A) and clone 97 vs ancestor grown at 42°C (B). We used the same color than in Figure 2.3.

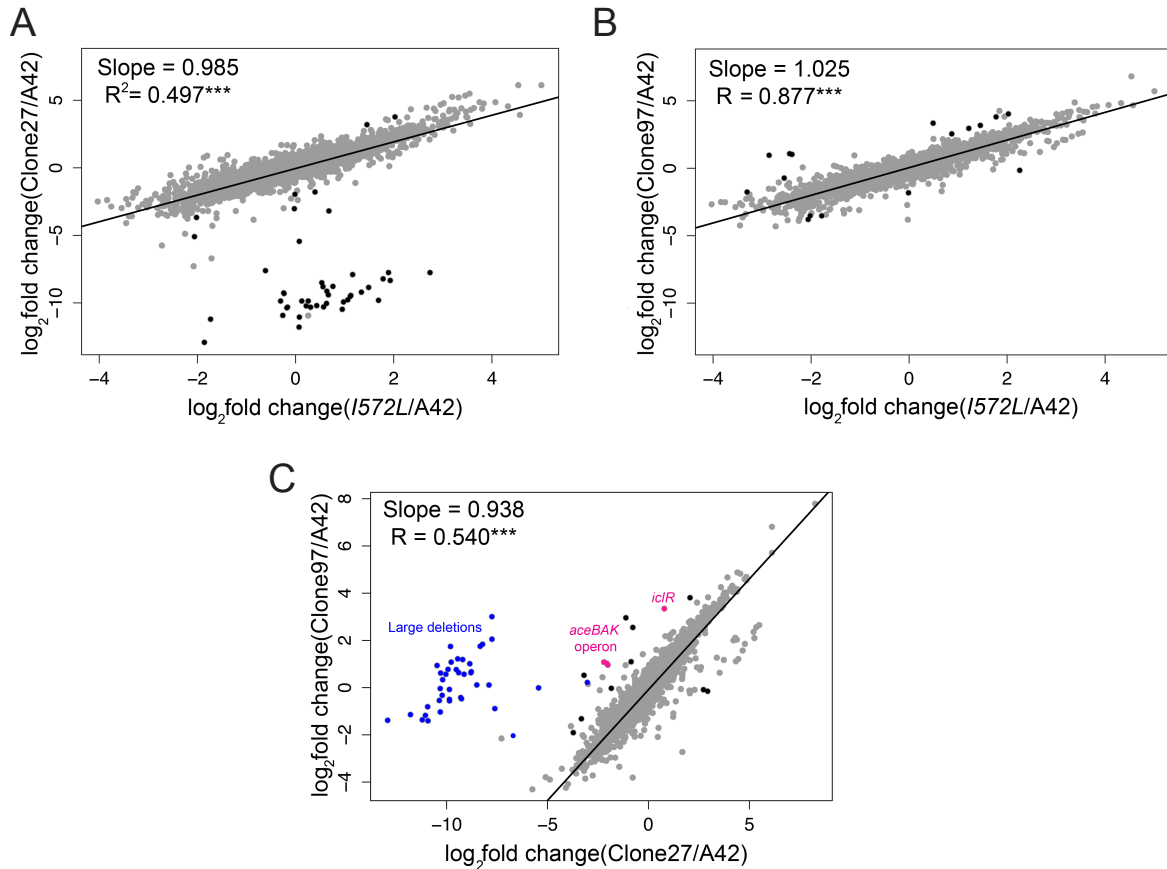


Figure S2.6. Similarities in global changes in GE between the mutation *I572L* and two high-temperature adapted clones. The black dots in A) and B) correspond to the differentially expressed genes between the clone 27 or 97 vs the mutation *I572L*. C) Similarities in global changes in GE between the two high-temperature adapted clones. The blue dots correspond to the differentially expressed genes (clone 97 vs clone 27) caused by the two large mutations in clone 27. The pink dots correspond to the differentially expressed genes (clone 97 vs clone 27) possibly explained by the *iclR* mutation in clone 97 (Table S2.5). Finally, the black dots correspond to the rest of the genes differentially expressed between clone 97 and clone 27.

Supporting Tables

Table S2.1. Families of highly differentially expressed genes (and their GO significant enrichment categories) during the acclimation response.

Target function/Pathway (GO enrichment category)	Gene	Products/putative cellular functions	log ₂ fold change ¹
DOWN-REGULATED GENES			
Flagellum (Bacterial-type flagellum-dependent cell motility, GO:0071973)	<i>flgA</i>	Flagellar biosynthesis; hook-filament junction protein	-4.9
	<i>flgB</i>	Flagellar basal-body rod protein FlgB	-5.9
	<i>flgC</i>	Flagellar basal-body rod protein FlgC	-6.2
	<i>flgD</i>	Flagellar biosynthesis	-6.2
	<i>flgE</i>	Flagellar hook protein FlgE	-6.5
	<i>flgF</i>	Flagellar basal-body rod protein FlgF	-6.7
	<i>flgG</i>	Flagellar basal-body rod protein FlgG	-6.8
	<i>flgH</i>	Flagellar L-ring protein FlgH	-6.2
	<i>flgI</i>	Flagellar P-ring protein FlgI	-6.1
	<i>flgJ</i>	Flagellar protein FlgJ	-5.5
	<i>flgK</i>	Flagellar biosynthesis; hook-filament junction protein	-4.6
	<i>flgL</i>	Flagellar biosynthesis; hook-filament junction protein	-3.9
	<i>flgN</i>	Flagellar biosynthesis protein FlgN	-3.4
	<i>flhA</i>	Flagellar biosynthesis protein FlhA	-5.1
	<i>flhB</i>	Flagellar biosynthesis protein FlhB	-6.0
<i>flhE</i>	Flagellar protein	-5.6	
Ribosomal proteins (Structural constituent of ribosomes; GO:0003735)	<i>rplA</i>	50S ribosomal subunit protein L1	-3.9
	<i>rplB</i>	50S ribosomal subunit protein L2	-4.4
	<i>rplC</i>	50S ribosomal subunit protein L3	-4.8
	<i>rplD</i>	50S ribosomal subunit protein L4	-4.7
	<i>rplE</i>	50S ribosomal subunit protein L5	-3.3
	<i>rplF</i>	50S ribosomal subunit protein L6	-3.3
	<i>rplI</i>	50S ribosomal subunit protein L9	-4.2
	<i>rplJ</i>	50S ribosomal subunit protein L10	-3.7
	<i>rplK</i>	50S ribosomal subunit protein L11	-4.1
	<i>rplL</i>	50S ribosomal subunit protein L12	-3.8
	<i>rplM</i>	50S ribosomal subunit protein L13	-3.8
	<i>rplN</i>	50S ribosomal subunit protein L14	-3.1
	<i>rplO</i>	50S ribosomal subunit protein L15	-3.0
	<i>rplP</i>	50S ribosomal subunit protein L16	-3.9
	<i>rplQ</i>	50S ribosomal subunit protein L17	-3.3
	<i>rplR</i>	50S ribosomal subunit protein L18	-3.4
	<i>rplS</i>	50S ribosomal subunit protein L19	-3.5
	<i>rplT</i>	50S ribosomal subunit protein L20	-2.4
	<i>rplU</i>	50S ribosomal subunit protein L21	-3.2
	<i>rplV</i>	50S ribosomal subunit protein L22	-4.5
	<i>rplW</i>	50S ribosomal subunit protein L23	-4.8
	<i>rplX</i>	50S ribosomal subunit protein L24	-3.3
	<i>rplY</i>	50S ribosomal subunit protein L25	-3.2
	<i>rpmA</i>	50S ribosomal subunit protein L27	-3.1
	<i>rpmB</i>	50S ribosomal subunit protein L28	-3.0
	<i>rpmC</i>	50S ribosomal subunit protein L29	-3.6
	<i>rpmD</i>	50S ribosomal subunit protein L30	-3.1

	<i>rpmE</i>	50S ribosomal subunit protein L31	-3.3
	<i>rpmF</i>	50S ribosomal subunit protein L32	-2.0
	<i>rpmG</i>	50S ribosomal subunit protein L33	-3.0
	<i>rpmH</i>	50S ribosomal subunit protein L34	-3.0
	<i>rpmI</i>	50S ribosomal subunit protein L35	-2.5
	<i>rpmJ</i>	50S ribosomal subunit protein L36	-2.6
	<i>rpsA</i>	30S ribosomal subunit protein S1	-3.1
	<i>rpsB</i>	30S ribosomal subunit protein S2	-2.9
	<i>rpsC</i>	30S ribosomal subunit protein S3	-4.1
	<i>rpsD</i>	30S ribosomal subunit protein S4	-3.3
	<i>rpsE</i>	30S ribosomal subunit protein S5	-3.3
	<i>rpsF</i>	30S ribosomal subunit protein S6	-4.2
	<i>rpsG</i>	30S ribosomal subunit protein S7	-3.2
	<i>rpsH</i>	30S ribosomal subunit protein S8	-3.5
	<i>rpsI</i>	30S ribosomal subunit protein S9	-3.6
	<i>rpsJ</i>	30S ribosomal subunit protein S10	-4.8
	<i>rpsK</i>	30S ribosomal subunit protein S11	-3.3
	<i>rpsL</i>	30S ribosomal subunit protein S12	-3.0
	<i>rpsM</i>	30S ribosomal subunit protein S13	-3.1
	<i>rpsN</i>	30S ribosomal subunit protein S14	-3.2
	<i>rpsO</i>	30S ribosomal subunit protein S15	-2.9
	<i>rpsP</i>	30S ribosomal subunit protein S16	-3.4
	<i>rpsQ</i>	30S ribosomal subunit protein S17	-3.5
	<i>rpsR</i>	30S ribosomal subunit protein S18	-4.3
	<i>rpsS</i>	30S ribosomal subunit protein S19	-4.6
	<i>rpsT</i>	30S ribosomal subunit protein S20	-3.1
	<i>rpsU</i>	30S ribosomal subunit protein S21	-2.9
Methionine biosynthesis (Sulfur amino acid biosynthetic process GO:0000097)	<i>metA</i>	homoserine O-succinyltransferase	-3.2
	<i>metB</i>	O-succinylhomoserine lyase	-1.6
	<i>metC</i>	Cystathionine β -lyase	-2.2
	<i>metE</i>	Cobalamin-independent homocysteine transnethylase	-6.6
	<i>metF</i>	5,10-methylenetetrahydrofolate reductase	-3.0
	<i>metG</i>	methionyl-tRNA synthetase	-2.1
	<i>methH</i>	cobalamin-dependent methionine synthetase	-1.6
	<i>metI</i>	L/D-methionine ABC transporter membrane subunit	-1.9
	<i>metJ</i>	MetJ transcriptional repressor	-1.4
	<i>metK</i>	S-adenosylmethionine synthetase	-3.6
	<i>metL</i>	Aspartate kinase	-1.8
	<i>metN</i>	L/D-methionine ABC transporter ATP binding subunit	-2.5
	<i>metQ</i>	L/D-methionine ABC transporter periplasmic binding subunit	-2.6
<i>metR</i>	MetR DNA-binding transcriptional dual regulator	-1.6	
Pathway for de novo biosynthesis of purine nucleosides ('de novo' IMP biosynthetic process; GO:0006189; $P=1.63 \times 10^{-8}$)	<i>purB</i>	Adenylosuccinate lyase	-2.9
	<i>purC</i>	PurC	-3.6
	<i>purD</i>	Phosphoribosylamine-glycine ligase	-3.9
	<i>purE</i>	PurE	-3.9
	<i>purF</i>	PurF	-3.2
	<i>purH</i>	AICAR transformylase/ IMP cyclohydrolase	-4.1
	<i>purK</i>	N5-carboxyaminoimidazole ribonucleotide synthetase monomer	-3.9
	<i>purL</i>	Phosphoribosylformylglycinamide synthetase	-3.4
	<i>purM</i>	PurM	-3.7
	<i>purN</i>	Phosphoribosylglycinamide formyltransferase 1	-3.2
	<i>purR</i>	PurR	-2.9
	<i>purT</i>	Phosphoribosylglycinamide formyltransferase 2	-3.5

Pathway for de novo biosynthesis of pyrimidine nucleosides ('de novo' UMP biosynthetic process; GO:0044205; $P=2.05 \times 10^{-5}$)	<i>pyrB</i>	Aspartate carbamoyltransferase, catalytic subunit	-3.4
	<i>pyrC</i>	PyrC	-2.1
	<i>pyrD</i>	Dihydroorotate dehydrogenase, type 2	-3.1
	<i>pyrE</i>	Orotate phosphoribosyltransferase (PyrE)	-3.6
	<i>pyrF</i>	Orotidine-5'-phosphate-decarboxylase (PyrF)	-2.9
	<i>pyrI</i>	Aspartate carbamoyltransferase, PyrI subunit	-2.8
Up-regulated genes			
Arginine degradation pathway (Arginine catabolic process; GO:0006527)	<i>astA</i>	Arginine succinyltransferase (AstA)	4.7
	<i>astB</i>	Succinylarginine dihydrolase (AstB)	4.2
	<i>astC</i>	Succinylornithine transaminase (AstC)	4.6
	<i>astD</i>	Succinylglutamate semialdehyde dehydrogenase	4.6
	<i>astE</i>	Succinylglutamate desuccinylase	4.3
Dipeptide transport (GO:0042938)	<i>dppA</i>	dipeptide ABC transporter – periplasmic binding protein	2.6
	<i>dppB</i>	dipeptide ABC transporter – putative membrane subunit	2.7
	<i>dppC</i>	dipeptide ABC transporter – putative membrane subunit	2.7
	<i>dppD</i>	dipeptide ABC transporter – putative ATP binding subunit	2.6
	<i>dppF</i>	dipeptide ABC transporter – putative ABC binding subunit	2.6
Fatty acid beta-oxidation (GO:0019395)	<i>fadA</i>	Fatty acid oxidation complex	4.0
	<i>fadB</i>	Fatty acid oxidation complex	4.0
	<i>fadE</i>	FadE acyl-CoA dehydrogenase enzyme	4.1
	<i>fadH</i>	2,4-dienoyl-CoA reductase	3.8
	<i>fadI</i>	Fad I component of anaerobic fatty acid oxidation complex	2.7
	<i>fadJ</i>	FadJ component of anaerobic fatty acid oxidation complex	2.7
glycerol-3-phosphate transport (GO:0015794)	<i>ugpA</i>	glycerol-3-phosphate ABC transporter - membrane subunit	3.6
	<i>ugpB</i>	glycerol-3-phosphate ABC transporter - periplasmic binding protein	3.9
	<i>ugpC</i>	glycerol-3-phosphate ABC transporter - ATP binding subunit	3.1
	<i>ugpE</i>	glycerol-3-phosphate ABC transporter - membrane subunit	3.6
	<i>glpT</i>	glycerol-3-phosphate: phosphate antiporter	1.4

[†] The logarithm (to basis 2) of the fold change

Table S2.2. List of: *i*) previously reported up-regulated genes during the heat stress response, *ii*) previously reported up-regulated genes during the general stress response and, *iii*) genes encoding different subunits of RNAP.

Gene	Protein/Product (Function)	Counts 37°C ¹	Counts 42°C ¹	log ₂ fold change ²	<i>q</i> ³
Heat shock genes⁴					
<i>rpoH</i>	RpoH σ 32 Sigma factor controlling the heat shock response	21281	26251	0.3	0.118
<i>clpA</i>	ClpA	12811	45837	1.8	5.92E-22
<i>clpB</i> (<i>htpM</i>)	ClpB (Hsp100)	37544	18417	-1.0	3.55E-06
<i>clpP</i>	ClpP (protease)	19570	18357	-0.1	0.726
<i>clpX</i>	ClpX (protease)	54230	56497	0.1	0.749
<i>creB</i>	CreB (catabolic response regulator)	3549	2527	-0.5	0.022
<i>cspD</i>	CspD (cold shock protein)	12812	72493	2.5	8.03E-38
<i>dnaJ</i>	DnaJ (Hsp40)	33998	5459	-2.6	2.36E-17
<i>dnaK</i>	DnaK (Hsp70)	176165	26756	-2.7	1.72E-14
<i>fkpA</i>	FkpA (heat shock peptidyl-prolyl isomerase)	76660	28916	-1.4	2.45E-13
<i>fxsA</i> (<i>yjeG</i>)	FxsA overproduction inhibits F exclusion of bacteriophage T7	6479	4082	-0.7	0.006
<i>gapA</i>	GapA, glyceraldehyde-3-phosphate	614632	89123	-2.8	4.22E-45

	dehydrogenase				
<i>groEL</i>	GroEL (Hsp60)	365814	73776	-2.3	9.58E-13
<i>groES</i>	GroES (Hsp10)	74591	10315	-2.9	9.52E-19
<i>grpE</i>	GrpE	33596	9971	-1.8	2.30E-11
<i>hflB (ftsh)</i>	HflB (ATP-dependent protease)	159483	81076	-1.0	6.16E-07
<i>hflX</i>	HflX (protease)	30953	36905	0.3	0.203
<i>hslO (yrfI)</i>	Hsp33	13534	6747	-1.0	1.00E-06
<i>hslR (yrfH)</i>	YrfH (hsp15)	4275	2140	-1.0	2.99E-05
<i>hslU</i>	HslU (protease)	42831	12805	-1.7	4.68E-09
<i>hslV</i>	HslV (protease)	9544	3334	-1.5	1.26E-07
<i>htpG</i>	HtpG (Hsp90)	71618	7101	-3.3	7.22E-23
<i>htpX</i>	HtpX (membrane protein)	40332	31977	-0.3	0.120
<i>ibpA (hslT)</i>	IbpA (sHsp)	2831	3039	0.1	0.838
<i>ibpB (hslS, htpE)</i>	IbpB (sHsp)	449	2685	2.6	0.001
<i>Int (cutE)</i>	CutE (catalyses the last step in lipoprotein maturation)	7767	5132	-0.6	0.007
<i>lon</i>	Lon (protease)	2857	2157	-0.4	0.061
<i>macB (ybjZ)</i>	YbjZ (putative component transport system)	6229	5001	-0.3	0.144
<i>miaA (trpX)</i>	MiaA (tRNA-transferase)	37385	44031	0.2	0.237
<i>mutM</i>	MutM (DNA glycosylase)	2569	548	-2.2	7.60E-26
<i>narP</i>	NarP (nitrate response regulator)	1684	1602	-0.1	0.820
<i>phoB</i>	PhoB (DNA-binding response regulator)	1623	2647	0.7	0.001
<i>phoR</i>	PhoR (histidine kinase)	1323	2590	1.0	2.061E-06
<i>prlC (opdA)</i>	PrlC (peptidase)	24403	19689	-0.3	0.148
<i>rpoD</i>	RpoD σ 70	54718	56050	0.0	0.841
<i>rrmJ (ftsJ)</i>	FtsJ (cell division)	22243	20277	-0.1	0.574
<i>sdaC</i>	SdaC (serine transporter)	10232	2392	-2.1	5.25E-17
<i>topA</i>	TopA (topoisomerase)	82407	48834	-0.8	1.416E-04
<i>yafD</i>	YafD	6715	9133	0.4	0.0235
<i>yafE</i>	YafE (putative biotin synthesis)	888	1633	0.9	2.53E-05
<i>ybbN</i>	YbbN (putative thioredoxin like)	16757	5350	-1.6	4.93E-17
<i>ybeY</i>	YbeY (metal binding, required for translation at 42°C)	6870	5997	-0.2	0.407
<i>ybeZ</i>	YbeZ (pho regulon protein)	16480	16968	0.0	0.821
<i>yccV (hspQ)</i>	HspQ (Hsp, hemimethylated DNA-binding protein)	6815	8422	0.3	0.117
<i>ycjF</i>	YcjF (putative membrane protein)	3627	1794	-1.0	9.46E-06
<i>yhdN</i>	YhdN (DUF1992 family protein)	1257	1119	-0.2	0.497
<i>yheL</i>	YheL (required for tRNA synthesis)	2425	1682	-0.5	0.016
<i>yhgH</i>	YhgH (protein required for utilization of DNA as carbon source)	3391	1490	-1.2	6.62E-09
<i>yrfG</i>	YrfG (phosphatase)	4809	3948	-0.3	0.209
General stress genes ⁵					
<i>rpoS</i>	RpoS σ S Master regulator of the general stress response	21887	36567	0.7	0.001
<i>aidB</i>	Putative acyl-CoA dehydrogenase (flavoprotein)	889	5574	2.6	3.18E-38
<i>artM</i>	Arginine transport protein (ABC superfamily, membrane)	6598	6500	0.0	0.971
<i>artP</i>	Arginine transport protein (ABC superfamily, ATP-binding subunit)	6194	9075	0.6	0.005
<i>bic</i>	Outer membrane lipoprotein (lipocalin)	687	7271	3.4	7.13E-54
<i>bolA</i>	Transcriptional activator of morphogenic pathway (BolA family)	5985	22283	1.9	4.97E-23
<i>cbpA</i>	Curved DNA-binding protein, cochaperone of DnaK (Hsp40 family)	2388	7906	1.7	9.54E-19
<i>csiD</i>	Conserved protein with clavaminic synthase-like domain	1776	17757	3.3	2.00E-09
<i>dps</i>	Stress response DNA-binding protein with ferritin-like domain	4557	34511	2.9	2.13E-49
<i>fbaB</i>	Fructose-bisphosphate aldolase class I	1520	9057	2.6	8.87E-19

<i>fic</i>	Stationary-phase protein	817	7975	3.3	4.18E-57
<i>gabD</i>	Succinate-semialdehyde dehydrogenase NADP dependent	2396	16754	2.8	1.19E-13
<i>gabP</i>	Gamma-aminobutyrate transport protein	695	5788	3.1	1.83E-18
<i>gadA</i>	Glutamate decarboxylase, isozyme A	244	555	1.2	3.26E-06
<i>gadB</i>	Glutamate decarboxylase, isozyme B	288	1479	2.4	9.97E-26
<i>gadX</i>	GadX DNA-binding transcriptional dual regulator	1598	9330	2.5	3.42E-37
<i>hdeA</i>	HdeA dimer, inactive form of acid-resistance protein	289	341	0.2	0.389
<i>hdeB</i>	Acid stress chaperon	65	131	1.0	0.012
<i>hnr</i>					
<i>(rssB)</i>	Response regulator involved in protein turnover	2770	11403	2.0	1.08E-25
	Catalase; hydroperoxidase HP II (III), RpoS dependent	2158	15234	2.8	2.91E-10
<i>katE</i>					
<i>ldcC</i>	Lysine decarboxylase 2, constitutive	2202	7648	1.8	8.23E-20
<i>mscL</i>	Mechanosensitive channel	4031	5327	0.4	0.053
	Acidic protein suppresses mutants lacking function of protein export				
<i>msyB</i>		1629	18919	3.5	1.66E-67
<i>narU</i>	Nitrate extrusion protein	272	1068	2.0	4.22E-07
<i>osmB</i>	Lipoprotein osmotically inducible	486	7065	3.9	1.05E-73
<i>osmC</i>	Resistance protein, osmotically inducible	1241	12585	3.3	3.80E-19
<i>osmY</i>	Hyperosmotically inducible periplasmic protein	1623	6038	1.9	0.000
<i>otsA</i>	Trehalose-6-phosphate synthase	2200	10866	2.3	1.98E-12
<i>otsB</i>	Trehalose-6-phosphate phosphatase, biosynthetic	777	6190	3.0	5.67E-13
	Transcriptional repressor for pyruvate dehydrogenase complex				
<i>pdhR</i>		7837	3989	-1.0	1.85E-06
	Pyruvate dehydrogenase/oxidase: FAD and thiamine PPi -binding				
<i>poxB</i>		2108	20176	3.3	1.80E-14
<i>rpsV</i>					
<i>(sra)</i>	30S ribosomal subunit protein S22	811	17193	4.4	2.70E-17
<i>rssA</i>	Putative transmembrane protein	3834	9298	1.3	1.69E-10
<i>talA</i>	Transaldolase A	2182	10790	2.3	4.31E-16
<i>tam</i>	Trans-aconitate methyltransferase	505	4351	3.1	3.47E-49
<i>tktB</i>	Transketolase 2, thiamon binding, isozyme	3659	20731	2.5	1.05E-19
<i>treA</i>	Trehalase, periplasmic	976	14891	3.9	2.00E-22
<i>treF</i>	Trehalase, cytoplasmic	1122	3862	1.8	1.26E-18
<i>wrbA</i>	Flavodoxin-like protein, trp repressor-binding protein	2281	28426	3.6	1.22E-71
<i>xasA</i>					
<i>(gadC)</i>	glutamic acid:4-aminobutyrate antiporter	676	2271	1.7	6.74E-17
	Putative transport protein, integral membrane location				
<i>ybiO</i>		651	2359	1.9	3.79E-14
<i>ycgB</i>	Conserved protein	1490	34076	4.5	9.15E-102
<i>ycgZ</i>	Predicted protein	219	913	2.1	5.76E-13
<i>yciF</i>	Putative structural protein	94	700	2.9	3.43E-09
<i>yciG</i>	Predicted protein	69	1525	4.5	3.51E-16
<i>ydaM</i>	diguanylate cyclase	1500	8179	2.4	3.16E-10
	Putative transport protein (ABC superfamily, periplasmic)				
<i>ydcS</i>		1318	39568	4.9	1.21E-21
<i>ydgA</i>	Conserved protein	461	3204	2.8	7.13E-20
<i>yedU</i>					
<i>(hchA)</i>	glyoxalase III, Hsp31 molecular chaperone	214	425	1.0	0.000
<i>ygaF</i>	Enzyme: L-2-hydroxylutarate oxidase	1387	10736	3.0	4.48E-11
<i>ygaM</i>	Predicted protein	1286	16593	3.7	2.78E-62
<i>ygaU</i>	Predicted protein	1617	16188	3.3	7.93E-26
<i>yggE</i>	Conserved protein	4095	16085	2.0	2.07E-24
<i>ygjG</i>	Putrescine aminotransferase	2098	32721	4.0	2.51E-26
<i>ygjG</i>	Putrescine aminotransferase	2098	32721	4.0	2.51E-26
<i>yhbo</i>	Protein involved in stress response	249	475	0.9	0.000
<i>yhiD</i>	Predicted Mg(2+) transport ATPase	59	135	1.2	0.011
<i>yhiE</i>					
<i>(gadE)</i>	Transcriptional regulator for gasABC operon	101	996	3.3	6.74E-31
<i>yhiO</i>					
<i>(uspB)</i>	Universal stress protein B	794	8268	3.4	1.69E-60

<i>yhiU</i>	MdtEF-TolC multidrug efflux transport system	254	719	1.5	8.32E-10
<i>yhiW</i>					
(<i>gadW</i>)	GadW DNA-binding transcriptional dual regulator	346	1667	2.3	9.71E-25
<i>yhjG</i>	Predicted outer membrane biogenesis protein	1341	8284	2.6	1.47E-27
	Putative transcriptional regulator with DNA-binding domain				
<i>yiaG</i>		654	3939	2.6	5.93E-36
<i>yjbJ</i>	Predicted stress response protein	707	5870	3.1	2.65E-49
<i>yjgB</i>	Ahr aldehyde reductase, NADPH-dependent	556	3439	2.6	6.04E-36
<i>ymgA</i>	Protein involved in biofilm formation	74	375	2.3	5.42E-11
<i>ynhG</i>	L,D-transpeptidase YnhG	1834	8559	2.2	2.24E-29
RNA polymerase genes and sigma factors					
<i>rpoA</i>	α subunit of the RNA polymerase	707881	66764	-3.4	9.10E-48
<i>rpoB</i>	β subunit of the RNA polymerase	306203	105179	-1.5	1.16E-15
<i>rpoC</i>	β' subunit of the RNA polymerase	323516	108182	-1.6	2.66E-14
<i>rpoZ</i>	ω subunit of the RNA polymerase	14040	7378	-0.9	2.38E-06
<i>rpoD</i>	RNA polymerase σ 70	54718	56050	0.0	0.841
<i>rpoE</i>	RNA polymerase σ E	39126	41095	0.1	0.786
<i>rpoN</i>	RNA polymerase σ 54	27627	24116	-0.2	0.385

¹ Mean normalized counts (sequence reads) from the indicated temperature

² The logarithm (to basis 2) of the fold change

³ Significance (adjusted by the Benjamini-Hochberg procedure).

⁴ Heat shock genes (and synonyms) expressed in *E.coli* during transient up-shift in temperature (Riehle et al. 2003; Nonaka et al. 2006; Gunasekera et al. 2008).

⁵ General stress response genes induced by σ^S (Weber et al. 2005).

Table S2.3. Criteria used to classify the genes differentially expressed during the acclimation and/or adaptive response.

Category	Anc42 vs Anc37	Mut42 vs Anc42	Mut42 vs Anc37
1) Restored	Significant \uparrow	Significant \downarrow	-
	Significant \downarrow	Significant \uparrow	-
2) Reinforced	Significant \uparrow	Significant \uparrow	Significant \uparrow
	Significant \downarrow	Significant \downarrow	Significant \downarrow
3) Unrestored	Significant \uparrow	Non-Significant	-
	Significant \downarrow	Non-Significant	-
4) Novel	Non-Significant	Significant \uparrow	Significant \uparrow
	Non-Significant	Significant \downarrow	Significant \downarrow

\uparrow Up-regulated genes

\downarrow Down-regulated genes

Significant: significantly differentially expressed gene ($q < 0.001$)

Non-Significant: not significantly differentially expressed gene ($q > 0.001$)

Table S2.4. Mutations and GE changes of the high temperature adapted clone 27.

Mutation I event	Gene	Product	Position ¹	Genetic effect	\log_2 fold change ²		
					A42 vs A37	I572L vs A42	c27 vs A42
Point mutation I250N	<i>secF</i>	Polypeptide SecF involved in translocation (inner membrane)	398676	Amino acid substitution	-1.5 ***	1.2 ***	1.4 ***
2,896 bp Large deletion From	<i>ybaL</i>	Polypeptide YbaL CPA2 transporter	473,629 - 475,305	Inactivation	-0.2	-0.0	-3.0 ***
	<i>fsr</i>	Transporter: fosmidomycin efflux transporter	475,543 - 476,763	Deletion	-1.2 ***	0.1	-9.9 ***

position 475,290 to position 478,186.	<i>ushA</i>	Enzyme: 5' - deoxyribonucleotidase	476,981 - 478,633	Inactivation	-0.1 0.0	0.0	-5.4 ***
71,416 bp Large deletion From position 547,700 to position 619,116.	<i>ybcR</i>	Polypeptide: DLP12 prophage	548,867 - 549,082	Deletion	1.9 ***	-1.8 ***	NA ***
	<i>ybcS</i>	Enzyme: DLP12 prophage; lysozyme	549,082 - 549,579	Deletion	1.2 ***	-1.0 **	NA ***
	<i>ybcT</i>	Polypeptide: DLP12 prophage; predicted murein endopeptidase	549,576 - 550,037	Deletion	-2.6 ***	1.9 ***	-7.8 ***
	<i>ybcU</i>	Polypeptide: putative lipoprotein	550,069 - 550,362	Deletion	-4.0 ***	2.3 ***	-7.8 ***
	<i>ECB_0</i> <i>0510</i>	Polypeptide: hypothetical protein	550,723 - 550,917	Deletion	0.3	-0.3	NA ***
	<i>nohB</i>	Polypeptide: DLP12 prophage; DNA packaging protein	551,306 - 551,881	Deletion	0.4	-0.3	NA ***
	<i>ECB_0</i> <i>0512</i>	Polypeptide: putative tail component of prophage	551,841 - 554,657	Deletion	0.6 **	-0.5	-9.9 ***
	<i>ECB_0</i> <i>0513</i>	Polypeptide: orf	554,716 - 557,061	Deletion	0.4	0.0	-9.3 ***
	<i>ECB_0</i> <i>0514</i>	Polypeptide: orf	557,058 - 557,339	Deletion	0.8	0.0	NA ***
	<i>ECB_0</i> <i>0515</i>	Polypeptide: orf	557,349 - 558,053	Deletion	1.1 **	-0.7	-7.6 ***
	<i>ECB_0</i> <i>0516</i>	Polypeptide: orf	558,064 - 558,357	Deletion	1.5 **	-0.5	NA ***
	<i>ECB_0</i> <i>0517</i>	Polypeptide: orf	558,112 - 558,447	Deletion	1.3 *	-1.0	NA ***
	<i>appY</i>	Polypeptide: APPY DNA- binding transcriptional activator	559,033 - 559,782	Deletion	-0.5	1.2	-8.2 ***
	<i>ompT</i>	Enzyme: outer membrane protease VII	560,031 - 560,984	Deletion	-1.6 ***	1.4 ***	-8.3 ***
	<i>envY</i>	Polypeptide: EnvY DNA- binding transcriptional activator	561,498 - 562,259	Deletion	0.5 *	-0.2	NA ***
	<i>ybcH</i>	Polypeptide: predicted protein	562,442 - 563,332	Deletion	0.7 **	-0.4	NA ***
	<i>nfrA</i>	Polypeptide: bacteriophage N4 receptor, outer membrane protein	563,333 - 566,305	Deletion	-0.6 **	-0.4	-9.3 ***
	<i>yhhl-2</i>	Enzyme: putative transposase	568,800 - 569,936	Deletion	0.1	1.0	-6.7 ***
	<i>ECB_0</i> <i>0526</i>	Polypeptide: orf	570,615 - 570,776	Deletion	-0.0	1.1	NA ***
	<i>ECB_0</i> <i>0527</i>	Polypeptide: orf	570,901 - 571,191	Deletion	-0.3	1.1	NA ***
	<i>ECB_0</i> <i>0528</i>	Polypeptide: orf	571,188 - 571,448	Deletion	1.0	-0.1	NA ***
	<i>ECB_0</i> <i>0529</i>	Polypeptide: orf	571,481 - 572,038	Deletion	0.5	0.2	NA ***
	<i>ECB_0</i> <i>0530</i>	Polypeptide: orf	571,890 - 573,488	Deletion	-0.0	0.9	NA ***
	<i>cusS</i>	Polypeptide: CusS sensory histidine kinase	574,225 - 575,673	Deletion	0.5 *	0.0	-9.9 ***
	<i>cusR</i>	Polypeptide: CusR	575,663 - 576,346	Deletion	0.5 **	-0.4	-10.3 ***
	<i>cusC</i>	Polypeptide: copper / silver efflux transport system - outer membrane porin	576,503 - 577,876	Deletion	0.7 ***	-0.4	-10.9 ***

<i>ylcC</i>	Polypeptide: copper / silver efflux transport system – periplasmic binding protein and metallochaperone	578,034 – 578,366	Deletion	0.6 **	0.0	-11.0 ***
<i>cusB</i>	Polypeptide: copper / silver efflux system – membrane fusion protein	578,382 – 579,605	Deletion	0.6 ***	-0.2	-11.8 ***
<i>cusA</i>	Polypeptide: copper / silver efflux transport system – membrane subunit	579,617 – 582,760	Deletion	0.4 *	-0.1	-10.9 ***
<i>pheP</i>	ransporter: phenylalanine: H+ symporter PheP	582,862 – 584,238	Deletion	-0.9 ***	0.0	-10.3 ***
<i>ybdG</i>	Polypeptide: mechanosensitive channel of miniconductance YbdG monomer	584,306 – 585,553	Deletion	-1.4 ***	0.5	-8.8 ***
<i>nfnB</i>	Enzyme: NfsB monomer	585,661 – 586,314	Deletion	-2.0 ***	0.8 *	-9.4 ***
<i>ybdF</i>	Polypeptide: conserved protein	586,408 – 586,776	Deletion	0.2	0.0	NA ***
<i>ybdJ</i>	Polypeptide: predicted inner membrane protein	586,841 – 587,089	Deletion	0.2	-0.3	NA ***
<i>ybdK</i>	Polypeptide: carboxylate-amine ligase	587,155 – 588,273	Deletion	3.7 ***	-1.5 ***	-12.9 ***
<i>insB-7</i>	Polypeptide: IS1 protein InsB	589,570 – 590,073	Deletion	0.4	0.2	NA ***
<i>insA-7</i>	Polypeptide: IS1 protein InsA	589,992 – 590,267	Deletion	0.1	0.0	NA ***
<i>hokE</i>	Polypeptide: HokE	590,543 – 590,695	Deletion	1.1 *	-0.5	NA ***
<i>insL-2</i>	Polypeptide: IS186/IS421 transposase	590,772 – 591,884	Deletion	-0.1	0.8	0.2
<i>entD</i>	Enzyme: phosphopantetheinyl transferase	592,166 – 592,795	Deletion	-0.2	0.6	-9.8 ***
<i>fepA</i>	Polypeptide: ferric enterobactin / colicin B / colicin D outer membrane porin FepA	592,961 – 595,201	Deletion	-0.8 ***	1.0 *	-9.2 ***
<i>fes</i>	Enzyme: enterochelin esterase	595,523 – 596,647	Deletion	0.9	0.0	-10.2 ***
<i>ybdZ</i>	Polypeptide: MbtH-like protein that enhances the catalytic function of EntF	596,650 – 596,868	Deletion	0.8	0.2	-8.5 ***
<i>entF</i>	Polypeptide: apo-serine activating enzyme	596,865 – 600,746	Deletion	0.7 *	0.1	-10.3 ***
<i>fepE</i>	Polypeptide: ferric enterobactin (enterochelin) transport	600,962 – 602,095	Deletion	0.3	-0.2	NA ***
<i>fepC</i>	Polypeptide: ferric enterobactin ABC transporter – ATP binding subunit	602,092 – 602,907	Deletion	-0.7 ***	0.5	-9.4 ***
<i>fepG</i>	Polypeptide: ferric enterobactin ABC transporter – membrane subunit	602,904 – 603,896	Deletion	-0.7 ***	0.3	-10.0 ***
<i>fepD</i>	Polypeptide: ferric enterobactin ABC transporter – membrane subunit	603,893 – 604,897	Deletion	-0.7 **	0.2	-10.2 ***
<i>ybdA</i>	Transporter: enterobactin efflux transporter EntS	605,008 – 606,258	Deletion	-2.0 ***	1.3 ***	-9.8 ***
<i>fepB</i>	Polypeptide: ferric enterobactin ABC transporter	606,262 – 607,218	Deletion	-0.5	0.7	-10.5 ***

	<i>entC</i>	– periplasmic binding protein Enzyme: isochorismate synthase 1	607,593 – 608,768	Deletion	0.0	1.1	-8.9 ***
	<i>entE</i>	Enzyme: EntE	608,778 – 610,388	Deletion	0.4	0.8	-9.5 ***
	<i>entB</i>	Polypeptide: EntB monomer	610,402 – 611,259	Deletion	0.6 *	0.7	-9.9 ***
	<i>entA</i>	Enzyme: EntA	611,259 – 612,005	Deletion	0.7 **	0.6	-9.1 ***
	<i>ybdB</i>	Enzyme: proofreading thioesterase in enterobactin biosynthesis	612,008 – 612,421	Deletion	0.5	0.5	-8.8 ***
	<i>ybdD</i>	Polypeptide: conserved protein	614,879 – 615,076	Deletion	1.6 ***	-1.0	-11.2 ***
	<i>ybdH</i>	Polypeptide: predicted oxidoreductase	615,086 – 616,174	Deletion	-0.8 ***	-0.6	-10.4 ***
	<i>ybdL</i>	Enzyme: methionine-oxo-acid transaminase, PLP-dependent	616,283 – 617,443	Deletion	-2.1 ***	0.3	-7.9 ***
	<i>ybdM</i>	Polypeptide: conserved protein	617,444 – 618,073	Deletion	-0.4	-0.1	NA ***
	<i>ybdN</i>	Polypeptide: orf	618,046 – 619,107	Deletion	-0.1	0.1	NA ***
Point mutation <i>E271G</i>	<i>mrdB</i>	Polypeptide: rod shape-determining membrane protein	648086	Amino acid substitution	-0.7 **	0.3	0.9 **
Point mutation <i>H30R</i>	<i>ydaE</i>	Polypeptide: Rac prophage; zinc-binding protein	1,415,090	Amino acid substitution	-0.0	1.3	NA ***
Point mutation <i>A35E</i>	<i>ECB_02828</i>	Transporter: Polysialic acid transporter protein kpsM	3,028,366	Amino acid substitution	-1.4 ***	0.6	0.9 **
Point mutation <i>D2E</i>	<i>glpE</i>	Enzyme: thiosulfate sulfurtransferase	3,490,427	Amino acid substitution	-0.7 ***	0.7	1.8 ***
Point mutation <i>I572L</i>	<i>rpoB</i>	Enzyme: DNA-directed RNA polymerase subunit beta	4,162,571	Amino acid substitution	-1.5 ***	0.4	1.4 ***

¹ Position relative to the genome of *E.coli* B REL606

² The logarithm (to basis 2) of the fold change. The asterisks represent the significance (q), with one asterisk denoting significance at $q < 0.05$, two asterisks denoting significance at $0.001 < q < 0.01$, and three asterisks denoting significance at $q < 0.001$. Bold values represent a significant differentiation of genes with red values representing up-regulation of genes and blue values representing down-regulation of genes.

Table S2.5. Mutations and GE changes of the high temperature adapted clone 97.

Mutational event	Gene	Product	Position ¹	Genetic effect	log ₂ fold change ²		
					A42 vs A37	I572L vs A42	c27 vs A42
IS Insertion IS1	<i>ykgE</i>	Polypeptide: predicted oxidoreductase	294,445	Inactivation ?	0.7 *	-1.4 ***	-1.1 ***
Point mutation <i>Q526P</i>	<i>dnaX</i>	Enzyme: DNA polymerase III subunits gamma and tau	465,735	Amino acid substitution	-1.1 ***	1.1 ***	1.3 ***
Point	<i>ybaL</i>	Polypeptide: YbaL CPA2	473,924	Amino acid	-0.2	-0.1	0.2

mutation		transporter	substitution				
V461G							
Point mutation	<i>mrdB</i>	Polypeptide: rod shape-determining membrane protein	648,692	Amino acid substitution	-0.7 **	0.6	0.8 **
Point mutation	<i>dinG</i>	Polypeptide: ATP-dependent helicase	825,980	Amino acid substitution	0.2	-0.2	-0.1
Deletion of 4bp	<i>glpT</i>	Transporter: glycerol-3-phosphate antiporter	2,298,766	Inactivation ?	1.4 ***	-1.6 **	-3.0 ***
Point mutation	<i>ECB_02828</i>	Transporter: Polysialic acid transport protein kpsM	3,028,381	Amino acid substitution	-1.4 ***	2.2 ***	1.0 ***
Point mutation	<i>rpoB</i>	Enzyme: DNA-directed RNA polymerase subunit beta	4,162,571	Amino acid substitution	-1.5 ***	0.9	1.5 ***
Point mutation	<i>iclR</i>	Polypeptide: IclR transcriptional repressor	4,202,435	Amino acid substitution	-0.3	0.1	3.3 ***
A45V							

Table S2.6. Primers used in this study

Name	Function	Sequence
ARV19	Amplify the CAT:tetR:YFP genomic cassette (forward primer)	5'- CGAAACATCCGGCAATAGAT -3'
ARV20	Amplify the CAT:tetR:YFP genomic cassette (reverse primer)	5'- CCAAAGCGACTTTTTTCAGC -3'
ARV34	External verification primer (forward primer)	5'- AACGCAACTGGAAACAGAGG -3'
ARV35	External verification primer (reverse primer)	5'- TGCCGGTAATACCCTGAAAC -3'
ARV48	Amplify the 117 bp fragment of the <i>gst</i> reference gene (forward primer)	5'- CTGAAGGATGAGCACTGGATC -3'
ARV49	Amplify the 117 bp fragment of the <i>gst</i> reference gene (reverse primer)	5'- AATGTGCTCTAACCCTTCCAG -3'
ARV50	Amplify the 120 bp fragment of the YFP target gene (forward primer)	5'- TGTGCTTTGCTAGATACCCAG -3'
ARV51	Amplify the 120 bp fragment of the YFP target gene (reverse primer)	5'- GTGTCTTGTAGTTCCCGTCATC -3'

REFERENCES

- Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biol* 11:R106.
- Applebee, M.K., M.J. Herrgard, and B.O. Palsson. 2008. Impact of individual mutations on increased fitness in adaptively evolved strains of *Escherichia coli*. *J Bacteriol* 190:5087-5094.
- Barabasi, A.-L., and Z.N. Oltvai. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5:101-113.
- Battesti, A., N. Majdalani, and S. Gottesman. 2011. The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* 65:189-213.
- Blount, Z.D., J.E. Barrick, C.J. Davidson, and R.E. Lenski. 2012. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 489:513-518.
- Brandis, G., M. Wrande, L. Liljas, and D. Hughes. 2012. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. *Mol Microbiol* 85:142-151.
- Carroll, S.M., and C.J. Marx. 2013. Evolution after introduction of a novel metabolic pathway consistently leads to restoration of wild-type physiology. *PLoS genetics* 9:e1003427.
- Chou, H.H., H.C. Chiu, N.F. Delaney, D. Segre, and C.J. Marx. 2011. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. *Science* 332:1190-1192.
- Christin, P.-A., D.M. Weinreich, and G. Besnard. 2010. Causes and evolutionary significance of genetic convergence. *Trends Genet* 26:400-405.
- Conrad, T.M., M. Frazier, A.R. Joyce, B.K. Cho, E.M. Knight, N.E. Lewis, R. Landick, and B.O. Palsson. 2010. RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc Natl Acad Sci USA* 107:20500-20505.
- Cooper, T.F., D.E. Rozen, and R.E. Lenski. 2003. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc Natl Acad Sci USA* 100:1072-1077.
- Cooper, V.S., D. Schneider, M. Blot, and R.E. Lenski. 2001. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J Bacteriol* 183:2834-2841.

- Crosa, J.H., and C.T. Walsh. 2002. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev* 66:223-249.
- Datsenko, K.A., and B.L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640-6645.
- Ederth, J., R.A. Mooney, L.A. Isaksson, and R. Landick. 2006. Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. *J Mol Biol* 356:1163-1179.
- Eisen, M.B., P.T. Spellman, P.O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863-14868.
- Fehér, T., B. Bogos, O. Méhi, G. Fekete, B. Csörgő, K. Kovács, G. Pósfai, B. Papp, L.D. Hurst, and C. Pál. 2012. Competition between transposable elements and mutator genes in bacteria. *Mol Biol Evol* 29:3153-3159.
- Fisher, R.A. 1930. The genetical theory of natural selection. Univ. Press.
- Fong, S.S., A.R. Joyce, and B.Ø. Palsson. 2005. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res* 15:1365-1372.
- Goodarzi, H., A.K. Hottes, and S. Tavazoie. 2009. Global discovery of adaptive mutations. *Nature Methods* 6:581-U44.
- Gunasekera, T.S., L.N. Csonka, and O. Paliy. 2008. Genome-wide transcriptional responses of *Escherichia coli* K-12 to continuous osmotic and heat stresses. *J Bacteriol* 190:3712-3720.
- Gunsalus, R.P., and S.-J. Park. 1994. Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons. *Res Microbiol* 145:437-450.
- Hasan, C.M.M., and K. Shimizu. 2008. Effect of temperature up-shift on fermentation and metabolic characteristics in view of gene expressions in *Escherichia coli*. *Microb Cell Fact* 7
- Haugen, S.P., W. Ross, and R.L. Gourse. 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature Reviews Microbiology* 6:507-519.
- Herring, C.D., A. Raghunathan, C. Honisch, T. Patel, M.K. Applebee, A.R. Joyce, T.J. Albert, F.R. Blattner, D. Van den Boom, and C.R. Cantor. 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 38:1406-1412.

- Hindre, T., C. Knibbe, G. Beslon, and D. Schneider. 2012. New insights into bacterial adaptation through in vivo and in silico experimental evolution. *Nature Reviews Microbiology* 10:352-365.
- Ishihama, A. 2000. Functional modulation of Escherichia coli RNA polymerase. *Annual Reviews in Microbiology* 54:499-518.
- Jin, D.J., W.A. Walter, and C.A. Gross. 1988. Characterization of the termination phenotypes of Rifampicin-resistant mutants. *J Mol Biol* 202:245-253.
- Jin, D.J., R.R. Burgess, J.P. Richardson, and C.A. Gross. 1992. Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proceedings of the National Academy of Sciences* 89:1453-1457.
- Jin, D.J., C. Cagliero, and Y.N. Zhou. 2012. Growth rate regulation in Escherichia coli. *FEMS Microbiol Rev* 36:269-287.
- Jozefczuk, S., S. Klie, G. Catchpole, J. Szymanski, A. Cuadros-Inostroza, D. Steinhauser, J. Selbig, and L. Willmitzer. 2010. Metabolomic and transcriptomic stress response of Escherichia coli. *Molecular systems biology* 6
- Keseler, I.M., A. Mackie, M. Peralta-Gil, A. Santos-Zavaleta, S. Gama-Castro, C. Bonavides-Martínez, C. Fulcher, A.M. Huerta, A. Kothari, and M. Krummenacker. 2013. EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res* 41:D605-D612.
- Khan, A.I., D.M. Dinh, D. Schneider, R.E. Lenski, and T.F. Cooper. 2011. Negative Epistasis Between Beneficial Mutations in an Evolving Bacterial Population. *Science* 332:1193-1196.
- Kishimoto, T., L. Iijima, M. Tatsumi, N. Ono, A. Oyake, T. Hashimoto, M. Matsuo, M. Okubo, S. Suzuki, and K. Mori. 2010. Transition from positive to neutral in mutation fixation along with continuing rising fitness in thermal adaptive evolution. *PLoS genetics* 6:e1001164.
- Lenski, R.E., J.A. Mongold, P.D. Sniegowski, M. Travisano, F. Vasi, P.J. Gerrish, and T.M. Schmidt. 1998. Evolution of competitive fitness in experimental populations of E. coli: what makes one genotype a better competitor than another? *Antonie van Leeuwenhoek* 73:35-47.
- Lewis, N.E., K.K. Hixson, T.M. Conrad, J.A. Lerman, P. Charusanti, A.D. Polpitiya, J.N. Adkins, G. Schramm, S.O. Purvine, and D. Lopez-Ferrer. 2010. Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. *Molecular systems biology* 6

- Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754-1760.
- Long, F., C.-C. Su, M.T. Zimmermann, S.E. Boyken, K.R. Rajashankar, R.L. Jernigan, and W.Y. Edward. 2010. Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature* 467:484-488.
- López-Maury, L., S. Marguerat, and J. Bähler. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet* 9:583-593.
- Mejia, Y.X., H.B. Mao, N.R. Forde, and C. Bustamante. 2008. Thermal probing of E-coli RNA polymerase off-pathway mechanisms. *J Mol Biol* 382:628-637.
- Nonaka, G., M. Blankschien, C. Herman, C.A. Gross, and V.A. Rhodius. 2006. Regulon and promoter analysis of the E. coli heat-shock factor, σ_{32} , reveals a multifaceted cellular response to heat stress. *Genes Dev* 20:1776-1789.
- Opalka, N., J. Brown, W.J. Lane, K.-A.F. Twist, R. Landick, F.J. Asturias, and S.A. Darst. 2010. Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. *PLoS biology* 8:e1000483.
- Orr, H.A. 2005. The genetic theory of adaptation: A brief history. *Nat Rev Genet* 6:119-127.
- Petsko, G.A., and D. Ringe. 2004. From Sequence to Structure Protein structure and function. New Science Press, 2-46
- Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res* 29:e45-e45.
- Philippe, N., E. Crozat, R.E. Lenski, and D. Schneider. 2007. Evolution of global regulatory networks during a long-term experiment with Escherichia coli. *BioEssays* 29:846-860.
- Polikanov, Y.S., G.M. Blaha, and T.A. Steitz. 2012. How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science* 336:915-918.
- Quandt, E.M., D.E. Deatherage, A.D. Ellington, G. Georgiou, and J.E. Barrick. 2014. Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in Escherichia coli. *Proc Natl Acad Sci USA* 111:2217-2222.
- R Core Team 2013. *R: A language and environment for statistical computing*. from <http://www.R-project.org/>

- Reynolds, M.G. 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics* 156:1471-1481.
- Richter, K., M. Haslbeck, and J. Buchner. 2010. The heat shock response: life on the verge of death. *Mol Cell* 40:253-266.
- Riehle, M.M., A.F. Bennett, and A.D. Long. 2001. Genetic architecture of thermal adaptation in *Escherichia coli*. *Proc Natl Acad Sci USA* 98:525-530.
- Rodríguez-Verdugo, A., D. Carrillo-Cisneros, A. Gonzalez-Gonzalez, B.S. Gaut, and A.F. Bennett. 2014. Different trade-offs result from alternate genetic adaptations to a common environment. *Proc Natl Acad Sci USA* 111:12121-6.
- Rodríguez-Verdugo, A., B.S. Gaut, and O. Tenailon. 2013. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol Biol* 13:50.
- Ryals, J., R. Little, and H. Bremer. 1982. Temperature-dependence of RNA-synthesis parameters in *Escherichia coli*. *J Bacteriol* 151:879-887.
- Sandberg, T.E., M. Pedersen, and R.A. LaCroix, et al. 2014. Evolution of *Escherichia coli* to 42° C and Subsequent Genetic Engineering Reveals Adaptive Mechanisms and Novel Mutations. *Mol Biol Evol* 31:2647-2662
- Soutourina, O.A., and P.N. Bertin. 2003. Regulation cascade of flagellar expression in Gram-negative bacteria. *FEMS Microbiol Rev* 27:505-523.
- Tenailon, O., A. Rodriguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennett, A.D. Long, and B.S. Gaut. 2012. The Molecular Diversity of Adaptive Convergence. *Science* 335:457-461.
- Vasi, F., M. Travisano, and R.E. Lenski. 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. *Am Nat* 432-456.
- Weber, H., T. Polen, J. Heuveling, V.F. Wendisch, and R. Hengge. 2005. Genome-wide analysis of the general stress response network in *Escherichia coli*: σ S-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* 187:1591-1603.
- Ye, Y., L. Zhang, F. Hao, J. Zhang, Y. Wang, and H. Tang. 2012. Global metabolomic responses of *Escherichia coli* to heat stress. *J Proteome Res* 11:2559-2566.
- Zhao, K., M. Liu, and R.R. Burgess. 2005. The global transcriptional response of *Escherichia coli* to induced σ 32 protein involves σ 32 regulon activation followed by inactivation and degradation of σ 32 in vivo. *J Biol Chem* 280:17758-17768.

Zhou, Y.N., L. Lubkowska, M. Hui, S. Chen, J. Strathern, D.J. Jin, and M. Kashlev.
2013. Isolation and characterization of RNA Polymerase rpoB mutations that alter transcription slippage during elongation in Escherichia coli. *J Biol Chem* 288:2700-2710.

CHAPTER 3

Different trade-offs result from alternate genetic adaptations to a common environment

ABSTRACT

Fitness trade-offs are often assumed by evolutionary theory, yet little is known about the frequency of fitness trade-offs during stress adaptation. Even less is known about the genetic factors that confer these trade-offs and whether alternative adaptive mutations yield contrasting trade-off dynamics. We addressed these issues using 114 clones of *Escherichia coli* that were evolved independently for 2000 generations under thermal stress (42.2°C). For each clone, we measured their fitness relative to the ancestral clone at 37°C and 20°C. Trade-offs were common at 37°C but more prevalent at 20°C, where 56% of clones were outperformed by the ancestor. We also characterized the upper and lower thermal boundaries of each clone. All clones shifted their upper boundary to at least 45°C; roughly half increased their lower niche boundary concomitantly, representing a shift of thermal niche. The remaining clones expanded their thermal niche by increasing their upper limit without a commensurate increase of lower limit. We associated these niche dynamics with genotypes and confirmed associations by engineering single mutations in the *rpoB* and *rho* genes. Single mutations in the *rpoB* gene exhibit antagonistic pleiotropy, with fitness trade-offs at 18°C and fitness benefits at 42.2°C. In contrast, a mutation within the *rho* transcriptional terminator, which defines an alternative adaptive pathway from that of *rpoB*, had no demonstrable effect on fitness at 18°C. This study suggests that two different genetic

pathways toward high temperature adaptation have contrasting effects with respect to thermal trade-offs.

INTRODUCTION

Despite the centrality of adaptation to evolution, surprisingly little is known about the diversity of mutations that contribute to adaptation or about their phenotypic and fitness effects (Orr 2005). There are, in fact, only a few well-known examples linking genotype, phenotype and adaptation in nature (Peichel et al. 2001; Hoekstra et al. 2006; Reed et al. 2011). In nature, this connection is often complicated by factors such as varying selection pressures or underlying genetic heterogeneities. Although the task is difficult, the general inability to connect phenotype to genotype in the context of environmental adaptation has been a major failing in the field of evolution (Barrett and Hoekstra 2011).

Experimental evolution provides a more tractable approach to study relationships among fitness, genotype and phenotype (Rose et al. 1996; Barrett and Hoekstra 2011). Here we explore these relationships based on our recent, large-scale evolutionary experiment (Tenaillon et al. 2012). The experiment began with an ancestral strain of *E. coli* B that was inoculated into ~115 independent replicates. Each replicate was grown at high temperature (42.2°C) for 2000 generations. At the end of the experiment, fitness was measured at 42.2°C for a single clone from each of 114 lineages; on average, fitness increased ~40% during the yearlong experiment.

We sequenced the genome of these 114 clones, identifying 1258 new mutations relative to the ancestral genome (Tenaillon et al. 2012). Broadly speaking, the mutations

fell into one of two 'adaptive pathways'. The first and most common pathway included mutations in the RNA polymerase (RNAP) β subunit (*rpoB*) gene, along with associated changes in RNAP subunit genes (*rpoA*, *rpoC* and *rpoD*) and the six *rod* genes that affect cell shape. The second adaptive pathway included mutations in the RNAP termination factor *rho*, which were positively associated with knockouts of the cardiolipin synthase (*cls*) gene and the transcription factor gene *icIR*. Mutations in the *rpoB* and *rho* adaptive pathways were not mutually exclusive, but mutations in the two pathways were strongly negatively associated (Tenailon et al. 2012).

Our thermal stress experiment has identified many putatively beneficial mutations that lead to higher fitness under thermal stress. But we still do not know the phenotypic consequences of these mutations or their relationship with fitness. Do the apparently distinct adaptive pathways converge on similar phenotypes? Or might the two pathways defined by *rho* and *rpoB* lead to alternative phenotypic solutions to a common selective pressure?

Here we begin to address these questions by measuring a complex phenotype: the magnitude of fitness trade-offs across a thermal gradient. Evolutionary trade-offs, which are defined as reduced fitness in a non-selected environment, are of great interest in their own right; they are widely observed and frequently assumed to govern and constrain trait evolution (Roff and Fairbairn 2007; Shoal et al. 2012). For example, trade-offs are commonly assumed in models of reaction norms niche specialization (Levins 1968; Lynch and Gabriel 1987; Futuyma and Moreno 1988; Huey and Kingsolver 1989; Angilletta et al. 2003).

Trade-offs have been examined previously in the context of experimental evolution, particularly trade-offs with respect to thermal niche (Bennett and Lenski 1993; Holder and Bull 2001; Bennett and Lenski 2007; Alto et al. 2013). Thermal niche has been a focus because temperature is a fundamental environmental property that affects physiological traits and often defines species' distributions (Somero 1978; Cooper et al. 2001). Most of the experimental studies of thermal niche have revealed, somewhat surprisingly, that thermal trade-offs are general but not universal. For example, of 24 *E. coli* lineages adapted to low temperature (20°C), 15 (62%) exhibit reduced fitness at high temperature (40°C) relative to their ancestor (Bennett and Lenski 2007). Such observations are not limited to *E. coli*, because studies of the vesicular stomatitis virus also suggest that fitness trade-offs are not universal across thermal gradients (Alto et al. 2013).

At least two questions remain about thermal trade-offs. The first is whether previous results – i.e. that fitness trade-offs are common but not universal – are accurate. The results may be inaccurate when there has been incomplete characterization of a thermal niche, which is defined as the range of temperatures over which an organism or genotype can maintain a stable population. To see this crucial point, it is helpful to visualize a thermal performance curve and some of its potential shifts during evolution to higher temperatures (Figure 3.1; Huey and Kingsolver 1989). In the niche-shift model, the organism adapts to high temperature by a horizontal shift of its niche (Figure 3.1B). In a specialist-generalist model, the organism adapts to high temperature by reducing niche breadth and increasing maximal performance (Figure 3.1C). Both of these models entail thermal trade-offs, but in the latter model the trade-

off is visible only near the lower niche limit of the ancestral strain (Huey and Kingsolver 1989). Thus careful characterization of niche limits is a necessary precursor to studying thermal trade-offs.

The second question concerns the underlying genetic causes of trade-offs. In theory, trade-offs may be caused either by antagonistic pleiotropy, in which a beneficial mutation in the selected environment has deleterious effects in non-selected environments (Levins 1968; Lynch and Gabriel 1987; Elena and Lenski 2003; MacLean et al. 2004), or by the accumulation of mutations that are neutral in a selected environment but deleterious in other environments. Whatever the cause, trade-offs have rarely been linked to specific genetic variants. One exception is a study of bacteriophage, in which a single adaptive mutation caused an increase in the breadth and height of the thermal reaction norm (Knies et al. 2006).

Here we characterize the thermal niche of 114 *E. coli* high-temperature adapted clones. To characterize thermal niche, we have measured both relative and absolute fitnesses over a range of temperatures. With these fitness data, we address the following sets of questions: First, are fitness trade-offs common and, if so, are they universal? That is, do thermal niches shift during adaptation to thermal stress, or do they follow alternative dynamics? Second, are any genetic variants associated with particular thermal growth dynamics? If so, do the two alternative 'adaptive pathways' have distinct phenotypic properties? Finally, can associations be confirmed with single, engineered mutations? If not, what might this imply about the underlying genetic complexities that contribute both to the evolution of thermal niche and to links between phenotype to genotype?

RESULTS

Relative fitnesses: To characterize thermal niches and fitness trade-offs, we measured both absolute and relative fitnesses. Relative fitnesses (w_r) have been estimated previously in the context of thermal trade-offs; we use them here to facilitate comparisons to previous work. In contrast, absolute fitnesses (w_a) can be assessed in a high-throughput manner, thus providing a tool to carefully measure the thermal boundaries of growth.

Relative fitnesses were measured against the ancestral clone REL1206 (Bennett et al. 1992) using standard competition assays (Lenski et al. 1991). The assays were performed at both 20°C and 37°C, two temperatures that have been assessed in previous studies of *E. coli* thermal trade-offs. Each of the 114 clones was tested in triplicate, with six-fold replication of a random subset of ~30 clones. In total, we performed > 800 fitness competitions, making this one of the largest studies of its kind.

At 37°C the mean of w_r estimates across all 114 high-temperature adapted clones was 0.973 (± 0.008 95% CI), representing a significant and general 2.7% decline in fitness across the entire experiment ($P=4.0 \times 10^{-9}$). For each clone, we also calculated the average of w_r estimates across replicates (\bar{w}_r) and tested the null hypothesis of $w_r = 1.0$ (Figure 3.2). At 37°C, 31% of clones had significant fitness deficits ($\bar{w}_r < 1.0$) relative to the ancestor (two-tailed t, $df=2$ and $P<0.05$). In contrast, one clone [clone #75; clone numbers correspond to reference (Tenailon et al. 2012)] had a fitness improvement of $\bar{w}_r = 1.085$ ($P<0.05$). The remaining clones (68%)

exhibited no significant difference in fitness compared to the ancestor at 37°C (Table S3.1).

Fitness trade-offs became more evident at 20°C (Figure 3.2). Across all 114 clones, the mean of w_r estimates was 0.910 (± 0.015 95% CI), representing a 9.0% decline in relative fitness across the entire experiment ($P < 10^{-15}$). For individual clones, 56% of clones had $\bar{w}_r < 1.0$ at 20°C. For one evolved clone (#107), the fitness impairment at 20°C was so severe that the bacteria did not grow, yielding a fitness estimate of 0.0; another clone (#66) had a significantly higher fitness than the ancestor, with $\bar{w}_r = 1.033$. The remaining 42% of clones had \bar{w}_r values that were not detectably different from 1.0.

The availability of \bar{w}_r values for each clone at 42.2°C (Tenailon et al. 2012), 37°C and 20°C permitted quantitative analyses of fitness across temperatures. The \bar{w}_r values were not strongly correlated between 42.2°C and either 20°C ($R^2 = -0.009$, $P=0.820$) or 37°C ($R^2 = 0.015$, $P=0.101$), suggesting that increases in w_r at 42.2°C are not necessarily associated with commensurate decreases in fitness at lower temperatures. In contrast, \bar{w}_r values were positively and significantly correlated between 20°C and 37°C ($R = 0.208$, $P=0.027$; Figure 3.2).

Characterization of niche boundaries: The \bar{w}_r results suggest that trade-offs are common but not universal, but \bar{w}_r does not provide information about the boundaries of the thermal niche. We therefore measured growth of the 114 clones at temperatures characteristic of the upper and lower niche boundaries. At each temperature, we measured bacterial density at the end of a daily growth cycle for four consecutive days and replicated the experiment three times. From these data, we

estimated the absolute fitness (\bar{w}_a) of each clone as the slope of fitted linear regression between time (day) and the natural logarithm of the density. We concluded that the bacterial populations ‘persisted’ when the slope was not significantly < 0.00 .

Under our assay conditions, the ancestor persisted at 18°C and 19°C but not 17°C (Table S3.2, Figure S3.1). Thus, the lower boundary of the ancestor’s thermal niche was between 17°C and 18°C. This lower boundary varied among the 114 high-temperature adapted clones (Figure 3.3, Table S3.2): 52% of clones were like the ancestor in their ability to persist at 18°C; 35% had a lower growth limit at 19°C; and 13% had a lower limit of 20°C. As mentioned, clone #107 was unable to grow at 20°C, representing a shift of $> 2^\circ\text{C}$ in its lower thermal niche. Note that \bar{w}_r at 20°C was correlated with \bar{w}_a at 18°C ($R^2=0.44$, $P=4 \times 10^{-16}$) and 19°C ($R^2=0.40$; $P=2 \times 10^{-14}$), indicating that \bar{w}_r at 20°C provides indirect information about lower niche boundaries.

For completeness, we also explored the upper niche boundary. All of the high-temperature adapted clones persisted at 43°C, and 96% (109 of 114) persisted at 45°C (Table S3.2, Figure S3.1). There was, however, no significant correlation between the relative fitness at 42.2°C and the absolute fitness at 45°C ($R^2=-0.008$, $P=0.730$). This lack of correlation may partially reflect difficulties in measuring the upper niche limit. These difficulties were especially prevalent for the ancestral clone, which persisted at 42°C but not at 43°C ($\bar{w}_a = -0.133$; 95% upper CI = -0.088) or 44°C ($\bar{w}_a = -0.297$; 95% upper CI = -0.116). Based on this information, we concluded that the upper thermal boundary of the ancestral REL1206 clone is $\sim 42^\circ\text{C}$. Surprisingly, however, \bar{w}_a of the ancestor was negative ($\bar{w}_a = -0.290$) but not significantly < 0.00 at 45.0°C (Table S3.2, Figure S3.1). This unexpected behavior was caused by the occasional sudden recovery

of populations whose densities had initially declined markedly, a phenomenon known as the “Lazarus effect” (Figure S3.2; *Discussion*; Bennett and Lenski 1993; Mongold et al. 1999).

Genotype and Thermal Niche: Our fitness assays characterize a complex phenotypic outcome to thermal stress. All of the evolved clones have expanded their upper niche boundary relative to the ancestor, and 49% exhibit a commensurate upward shift in their lower niche boundary. A remaining question is whether phenotypic variation in growth and fitness at low temperatures is associated with specific mutations.

To begin to address this question, we conducted an observational analysis to associate \bar{w}_a with genotypic variation. Recall that the genotypic data included full genomes with a total of >1000 mutations, most of which were identified in only a single clone. Low frequency mutations contain little information for associations, so we focused our analyses on the five most common mutations, which were found ≥ 14 clones (Table 3.1; Figure S3.3). We associated the genotype for each mutation with \bar{w}_a at 18°C. We chose 18°C both because this temperature had the most variability in \bar{w}_a among our assay temperatures (Figure S3.1, Table S3.2) and because it clearly delineated two groups of clones: those that survive and those that decline to extinction (Figure S3.4).

For all five mutations, we fitted a linear regression model that controls for the presence of co-occurring mutations (i.e. background effects; *SI Material and Methods*). Applying this approach, we found that the *rpoB* I966S mutation contributed to the model; its presence was associated with a significant decrease in \bar{w}_a relative to lines that do not have this mutation (Table 3.1). Indeed, 10 of 15 clones that harbored this

mutation had \bar{w}_a significantly < 0.0 at 18°C (Figure S3.5). Interestingly, three *rpoB* I966S clones contained a second mutation in *rpoB* (T539P). Two of these lines were persistent at 18°C, and as a group the three *rpoB* I966S/T539P clones had a higher average \bar{w}_a , at -0.233, than the remaining 11 *rpoB* I966S lines, at -0.538. These observations suggest that the *rpoB* T539P mutation has a compensatory effect on *rpoB* I966S.

Of the remaining four mutations, none yielded a significant association with decreased \bar{w}_a at 18°C (Table 3.1, Figure S3.5). However, we believe the lack of association to be meaningful for *rho* I15N, because 13 of 14 clones that harbored this mutation persisted at 18°C (Figure S3.5). The probability of randomly choosing a set of 14 clones – without respect to genotype and for which 13 or more do not exhibit a 18°C trade-off – is small ($P \sim 0.001$).

Direct tests of genotype – phenotype associations: Association analyses suggest that fitness trade-offs at 18°C are associated with the *rpoB* I966S mutation but not with the *rho* I15N mutation. Can a single mutation lead to a thermal shift? And do mutations that define alternative adaptive pathways lead to different thermal trade-off dynamics? To address these questions, we engineered *rpoB* I966S and *rho* I15N mutations into the REL1206 background. For each of the two mutants, we assessed \bar{w}_r at 20°C, 37°C and 42.2°C, and \bar{w}_a at 18°C (Table 3.2). The results indicate that the *rpoB* I966S mutant has: *i*) a lowered and borderline significant inability to persist at 18°C ($\bar{w}_a = -0.035$; upper CI = 0.043), as suggested by our association analyses, *ii*) decreased \bar{w}_r at 20°C ($\bar{w}_r = 0.929$; $P = 0.046$), *iii*) a strong benefit ($\bar{w}_r = 1.373$) at 42.2°C, and *iv*) no detectable effect on relative fitness at 37°C ($\bar{w}_r = 0.990$; $P = 0.793$). These observations

are consistent with this single mutation conferring a shift in thermal niche. In contrast, we could not detect a thermal trade-off for the *rho* I15N mutation at any temperature nor, in fact, could we detect a fitness benefit at 42.2°C (Table 3.2).

Because the *rpoB* I966S mutation yielded trade-off dynamics, we questioned whether the effect was specific to the I966S mutation or perhaps a general property of mutations within the *rpoB* gene. To address this issue, we measured fitnesses for three additional *rpoB* single mutants: *rpoB* I527F, *rpoB* I527N and *rpoB* I572L (Rodríguez-Verdugo et al. 2013). As a group, these three mutations were found in 12 of the 114 clones, with the two most common found in 5 clones (Tenailon et al. 2012). Each of the three clones with a mutation in codon 572 exhibited decreased persistence at 18°C as well as $\bar{w}_r < 1.0$ at 20°C (Table 3.2).

DISCUSSION

Trade-offs are often assumed to be a ubiquitous feature of adaptation (Levins 1968; Lynch and Gabriel 1987; Futuyma and Moreno 1988). Of course ubiquity is difficult to test precisely, because in theory trade-offs may affect a wide range of unknown or unsuspected phenotypes. The characterization of trade-offs is nonetheless important, both because they may constrain evolutionary trajectories and because they also potentially affect the ‘evolvability’ of a system (Pigliucci 2008; Barrick et al. 2010; Woods et al. 2011). Here we have examined fitness trade-offs across a thermal gradient, based on 114 *E. coli* clones that are adapted to high temperature (42.2°C). These clones have been shown to adapt through mutations in two adaptive pathways, one defined by mutations in RNAP subunits and another typified by mutations in the

Rho termination factor (Tenailon et al. 2012). An open question is whether these alternative adaptive pathways converge on the same phenotypes, including the same types and magnitudes of trade-offs.

Shifts and expansions of thermal niche: To characterize trade-offs and thermal niches, we have examined both relative and absolute fitnesses across a range of temperatures. We detect 2.7% and 9.0% decreases in relative fitness at 37°C and 20°C, respectively, across the combined sample of 114 clones. While these values signal a general trade-off effect, our results also suggest variance among clones. At 37°C, for example, trade-offs in relative fitness are common but not universal; 31% of clones exhibit \bar{w}_r values significantly < 1.0. The results at 20°C are similar to those at 37°C but exaggerated, in that a higher proportion of clones (56%) exhibit statistically significant reductions in relative fitness. Both of these proportions could be underestimates because our assays were based on a number of replicates ($n=3$) that may limit statistical power. When we assess this proportion with more replication ($n=6$) for a subset of 30 lines, we detect slightly higher proportions, at 32% and 73% for 37°C and 20°C, respectively. Overall, however, our results support previous conclusions that thermal trade-offs are general but not universal (Bennett and Lenski 2007).

Full categorization of trade-offs requires characterization of niche boundaries, because models of niche evolution predict trade-offs close to these boundaries (Levins 1968; Lynch and Gabriel 1987; Futuyma and Moreno 1988; Huey and Kingsolver 1989; Kingsolver 2009) (Figure 3.1). Empirical data have demonstrated this as well. For example, when the performance of bacteriophages ϕ X174 and G4 were measured over a wider temperature range than initial work, additional trade-offs were discovered

(Holder and Bull 2001). We have therefore assessed the thermal boundaries of the 114 clones.

At the upper end of the thermal niche, most (> 95%) of the clones persist at 45°C, signaling an expansion of their niche at least 2°C beyond that of the ancestor (Figure 3.1B). This observation contrasts with a previous study in which only one of six 42°C-adapted lines expanded their upper thermal limit (Bennett and Lenski 1993) but suggests a degree of 'pre-adaptation' to temperatures beyond the clones' immediate experience. Above 45°C the analyses become complicated by the Lazarus effect, in which declining populations suddenly recover, presumably due to major effect mutations. Indeed, the ancestral clone, which is habituated to lab conditions of 37°C, does not persist at 43°C but often recovers at 45°C (Figure S3.2). We do not yet know the molecular processes underlying the Lazarus effect, but two seem possible: either the fitness effects of mutations change as a function of the intensity of stress (Trindade et al. 2012a; Hietpas et al. 2013) or the mutation rate increases under high stress (Al Mamun et al. 2012; Hietpas et al. 2013). We do not yet know which of these two mechanisms predominates.

We also assessed persistence at lower temperatures. We have found, similar to a previous study (Bennett and Lenski 1993), that our REL1206 ancestor persists at 18°C but not 17°C, and thus its lower niche boundary is ~18°C in our culture conditions. Like the ancestor, none of the 114 clones persist at 17°C, but 52% persist at 18°C. This group of evolved clones exhibits an expansion of their thermal range relative to the ancestor, because their lower limit is unchanged but their upper limit has shifted by at least two degrees.

Niche expansions in constant environments are generally not predicted by evolutionary theory (Huey and Kingsolver 1989). Models such as the niche-shift model and the specialist-generalist model (Figure 3.1) assume that the total area under the fitness function is constant (Levins 1968; Lynch and Gabriel 1987). Therefore, if there is an expansion, it should be associated with decreased performance throughout much of the thermal gradient (Lynch and Gabriel 1987; Huey and Kingsolver 1989; Angilletta Jr et al. 2002). We do observe some such decreases in \bar{w}_r at 20°C, which is consistent with these predictions, but we do not detect fitness decreases at 37°C for the majority of clones. Moreover, the subset of clones with an expanded thermal range did not consistently have the lowest relative fitness values at 20°, 37° or 42.2°C, as is expected under these models (Lynch and Gabriel 1987). Finally, niche models predict a quantitative relationship between the magnitude of improvement at high temperature (42.2°C) and the magnitude of trade-offs at lower temperatures. We tested for correlations in fitness between high and lower temperatures but did not detect the expected correlations.

The remaining subset of 48% clones has shifted, rather than expanded, their thermal niche by increasing both their upper and lower thermal limits (Figure 3.1B). Such niche shifts have been predicted by evolutionary theory (Huey and Kingsolver 1989) and imply the existence of trade-offs, which we demonstrated directly by showing that a shift in lower thermal limit is correlated with relative fitness trade-offs at 20°C. Surprisingly, although most clones have either shifted or expanded their thermal niche, only one clone (#107) has reduced thermal range (Figure 3.1C).

Genetics, adaptive pathways and thermal trade-offs: Broadly speaking, we have found two phenotypes among our 114 clones: niche-shift vs. niche-expansion. For a subset of clones the difference between these two phenotypes is explained by single mutations. Niche-shift is associated with mutations in *rpoB*, and niche-expansion may be associated with at least one mutation in *rho*. As mentioned, these two genes typify alternative adaptive pathways in our thermal stress experiment (Tenailon et al. 2012). Taken together, these results imply that the two adaptive pathways define (and constrain) alternative fitness trajectories. Consistent with this conjecture, fitnesses at low temperatures vary significantly between the set of 60 and 26 clones containing *rpoB* and *rho* mutations, respectively. For example, w_a at 18°C has an average value of -0.370 across *rpoB* clones and -0.138 for *rho* clones ($P = 0.014$); similarly, average w_r at 20°C is 0.892 for the set of *rpoB* mutants and 0.946 for the *rho* mutants ($P = 0.015$).

Thus, the two alternative adaptive pathways confer different trade-off effects, but questions about mechanisms remain. Since the elongation speed of RNAP increases with increasing temperature (Ryals et al., 1982), we hypothesize that adaptive mutations in *rpoB* slow RNAP transcription under thermal stress, leading to increased termination efficiency (Jin et al., 1992). We have no direct evidence of this effect for the *rpoB* I966S mutation, which is located in one of the two parallel α -helices of the *Eco* flap domain (Figure S3.6) (Opalka et al., 2010). However, mutations in other locations of RNAP, including residues in *rpoB* codon 572, have been linked to changes in termination efficiency (Jin et al., 1988; Zhou et al., 2013).

The *rho* I15N mutation is located in the second α -helix from the N-terminus (Figure S3.6). No mutations have been observed in this residue, but a mutation in a

nearby residue (L3F) has been linked to an increase of termination efficiency of the Rho protein (Mori et al., 1989). We therefore hypothesize that I15N also increases termination efficiency. If true, this implies that increased termination efficiency can be achieved either by slowing RNAP (via mutations in *rpoB*) or by enhancing Rho termination. Under this scenario, negative epistasis between *rho* and *rpoB* mutations (Tenaillon et al. 2012) may be explained by the sets of mutations ‘over-tuning’ termination (Jin et al., 1988).

This conjecture does not readily explain the difference in trade-off dynamics between *rpoB* and *rho* mutants (Table 3.2). The explanation may lie with pleiotropic effects. RNAP mutants have the capacity to affect the expression of every gene, but *rho* influences termination in a subset of 20% to 50% of *E. coli* genes (Peters et al., 2009; Hollands et al., 2014). Since *rpoB* has the potential for more pleiotropic interactions, it likely also has a higher probability to generate fitness trade-offs via antagonistic pleiotropy.

Background Effects and Epistasis: While the precise mechanism of adaptation remains uncertain, our studies show that adaptive trajectories are profoundly affected by the identity of individual mutations and by interactions among these mutations. For example, the three alternative amino acid replacements within *rpoB* codon 572 vary in fitness effects among one-another, across nutrient conditions and against genetic backgrounds (Reynolds 2000; Barrick et al. 2010; Trindade et al. 2012b; Rodríguez-Verdugo et al. 2013). The background effect is particularly dramatic: when these mutations are introduced into *E. coli* K12, they are strongly disadvantageous at 42.2°C

(Rodríguez-Verdugo et al. 2013) but confer ~20% fitness benefits in REL1206 (Table 3.2). Background effects need not be driven by differences as pronounced as those between *E. coli* B and K12. For example, our association analyses suggest that the trade-off bestowed by *rpoB* I966S may be compensated by a single additional *rpoB* mutation (T539P). While the potential compensatory effects of the T539P mutation have not yet been confirmed by functional analyses, our results suggest that the presence, absence and potential compensation of thermal trade-offs may be encoded by single nucleotide mutations.

It is somewhat puzzling that we do not detect a beneficial effect of the *rho* I15N mutation (Table 3.2), because the presence of this mutation across 14 of 114 lines argues strongly that it is advantageous under the conditions of the thermal stress experiment. One explanation may be statistical power— i.e, the mutation may be beneficial but at a level too low to be detected in a single competition assay. However, given that we detect \bar{w}_r differences of < 3% in this study and also that the estimate of \bar{w}_r at 42.2°C is < 1.00 (Table 3.2), we do not believe that statistical power is the sole explanation. Instead, we hypothesize that *rho* I15N may not be beneficial by itself, as studied here, but rather beneficial only in the presence of other interacting mutations. In short, we predict sign epistasis between *rho* I15N and associated mutations. A likely candidate for positive epistatic interactions is the *cls* gene, which was mutated in 11 of the 14 *rho* I15N containing lines. Mutations within *cls* significantly increase membrane fluidity (Sleight et al., 2008, *Genetics*, 180, 431-43), but it is difficult to hypothesize how this phenotype might interact with *rho* mutations. Although the presence of sign epistasis between mutations is common (Silva et al. 2011; Chou et al. 2014), our

predictions about epistatic interactions between *rho* and *cls* need to be followed by functional studies.

Overall, our experiments demonstrate dramatic phenotypic variation with respect to thermal niche and fitness trade-offs in 114 clones that have been subjected to thermal stress for 2000 generations. We have also shown that some of these phenotypes can be recapitulated with single base substitutions and that therefore at least some trade-offs are generated by antagonistic pleiotropy. While it thus seems that the genetics underlying fitness trade-offs are straightforward in this system, our data hint at much additional genetic complexity that includes: *i*) similar but slightly different trade-offs encoded by different mutations in the same gene (e.g., *rpoB* I572N vs. I966S; Table 3.2), *ii*) mutations that may compensate for pleiotropic trade-offs, *iii*) widespread epistatic interactions and *iv*) alternative fitness trajectories defined by different adaptive pathways. Given the breadth of these complexities in a well-controlled experimental system, it is no wonder that the mapping of genotype, phenotype and fitness continues to be a daunting task in natural populations.

MATERIALS AND METHODS

Fitness estimates: We examined the 114 adapted clones from (Tenaillon et al. 2012) and estimated relative fitness for each clone at 20°C and 37°C, following reference (Lenski et al. 1991). We performed two-tailed *t* tests to test for a fitness difference relative to the ancestor. Absolute fitness (w_a) was estimated at 17°C, 18°C, 19°C, 43°C and 45°C using a protocol similar to reference (Bennett and Lenski 1993) (*SI Materials and Methods*). Clones were propagated daily by serial transfer (100-fold

dilution) for four days at the assay temperature. Bacterial density was measured each day by Coulter count. Absolute fitness was calculated as the slope of the fitted linear regression between the natural logarithm of the density against time, based on three replicates. If the w_a was significantly <0.0 , based on a one-sided 95% CI, we concluded that the clone was declining towards extinction. In contrast, if w_a was not significantly <0.0 , we concluded that the clone could persist at that temperature.

Strain construction and confirmation of recombinants: Single mutations were introduced into the *rpoB* and *rho* genes of the ancestral strain REL1206 using a recombineering plasmid similar to pKD46 (Datsenko and Wanner 2000). Because the *rpoB* (I966S) and the *rho* (I15N) mutations lack a selectable phenotype, we co-introduced one selectable marker (Ara+) with the non-selectable mutation and used the Ara+ marker as a first-pass selective screen on minimal medium supplemented with arabinose. The presence of the second, non-selectable mutation in the *rpoB* or *rho* gene was monitored with Sanger sequencing. Additional details are provided in *SI Materials and Methods*.

Genetic associations: Associations were based on the dataset of mutations from reference (Tenailon et al. 2012). Our approach was first to identify the clones that harbor our focal set of the most common mutations. We then defined clusters of mutations that co-occur with the focal set. Given these clusters, we fit a fixed effects linear model by combining a cluster incidence matrix with absolute fitness data from 18°C. The model fitted a regression line for each clone, with the slope of the regression given by the sum of the individual slope coefficients of the co-occurring mutations in that clone. We selected a final model based the lowest Akaike Information Criterion score.

The model yielded both β_{1,m_i} , a measure of the effect of the presence of mutation m_i on absolute fitness, and a 95% CI of the estimate. All numerical and statistical analyses were performed using R version 3.0.2. Additional details are provided in the *SI Materials and Methods*.

FIGURES

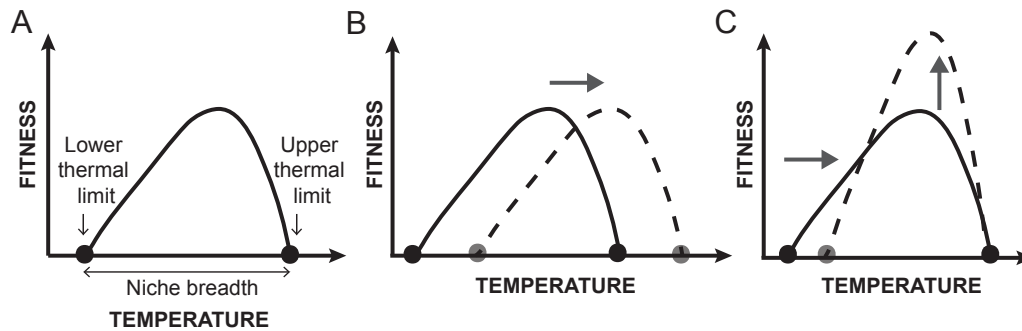


Figure 3.1. Hypothetical evolutionary responses of adaptation to high temperature. (A) The thermal dynamics of an ancestral genotype (solid line). (B) The thermal dynamics of a clone (broken line) for which adaptation to high temperature includes a shift in thermal niche (niche shift model). (C) Adaptation to high temperature (broken line) for which adaptation to high temperature includes a reduction in thermal niche (specialist-generalist model).

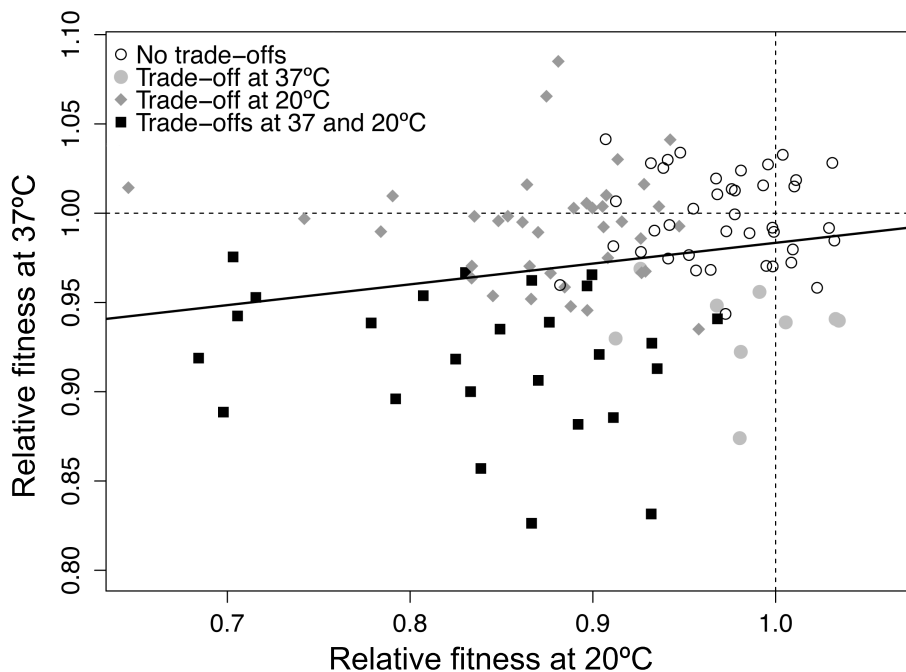


Figure 3.2. Mean relative fitnesses of high-temperature evolved clones at 20°C and at 37°C. Each point represents the average of three replicate relative fitness estimates for each clone. The dotted lines in each axis represent a relative fitness equal to 1.0 (i.e., no difference between the evolved clone and ancestor). Empty circles represent clones with w_r not significantly different than 1.0 at 20°C and 37°C. Filled symbols indicate w_r significantly < 1.0 at the indicated temperature(s), based on a two-tailed t tests ($df=2$, $P<0.05$). The black line represents the linear regression fitted to the data.

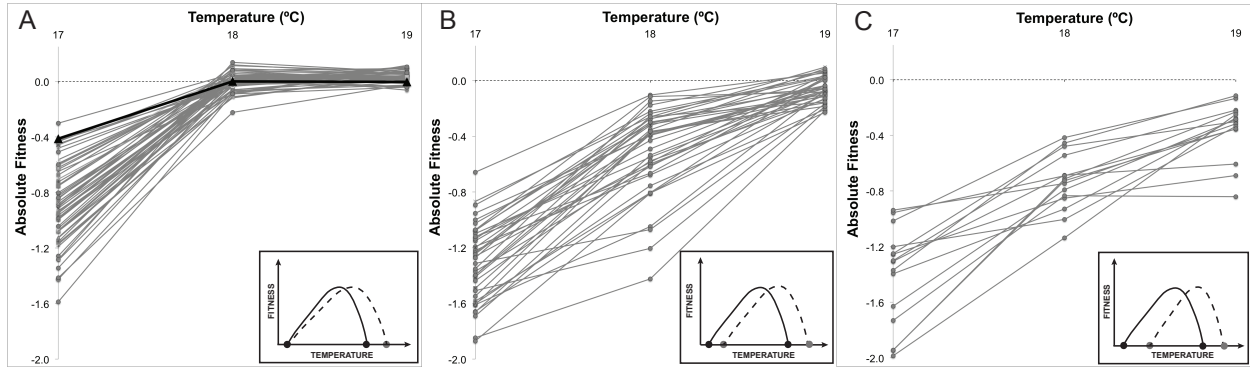


Figure 3.3. Mean absolute fitnesses \bar{w}_a of the ancestor and each of 114 high-temperature adapted clones at 17°C, 18°C and 19°C. (A) The ancestor (represented as a solid black line) and 59 evolved clones persist at 18°C and 19°C but not at 17°C (absolute fitnesses <0.00). (B) The subset of clones that do not persist at 18°C but maintain a stable population 19°C. (C) The 15 clones with absolute fitnesses < 0.00 at all three temperatures. The insets represent the inferred thermal niche for each set of clones, based on the schema defined in Figure 3.1.

TABLES

Table 3.1. The five most common mutations and their associated effects on absolute fitness at 18.0°C.

Mutation	Genomic Position ^a	Gene(s) affected	No ^b	Coefficient β_{1,m_i} (95% CI) ^c
<i>rho</i> I15N ^a	3921335	<i>rho</i>	14	0.011 (-0.156, 0.177)
<i>rpoB</i> I966S ^a	4156827	<i>rpoB</i>	15	-0.845 (-1.191, -0.500)
23164 bp deletion	2032562	Multigenic	21	0.016 (-0.203, 0.235)
1 bp deletion	2131465	Intergenic	25	-0.004 (-0.240, 0.241)
71416 bp deletion	547700	Multigenic	35	-0.011 (-0.351, 0.328)

^a Nonsynonymous point mutation

^b The number of times the mutation was found in 114 independently evolved clones.

^c The coefficient β_{1,m_i} represents the contribution of the mutation m_i to the absolute fitness; it is provided along with an associated 95% confidence interval.

Table 3.2. Fitness estimates for single mutations.

Mutant	Fitness estimate			
	18°C ^a	20°C ^b	37°C ^b	42°C ^b
<i>rpoB</i> I966S	-0.035 (0.043)	0.929 (± 0.068)*	0.990 (± 0.155)	1.373* (± 0.188)
<i>rpoB</i> I572F	-0.206 (-0.089)*	0.910 (± 0.060)*	0.956 (± 0.117)	1.170 (± 0.135)*
<i>rpoB</i> I572L	-0.308 (-0.132)*	0.928 (± 0.047)*	0.954 (± 0.067)*	1.176 (± 0.138)*
<i>rpoB</i> I572N	-0.229 (-0.073)*	0.946 (± 0.074)*	1.003 (± 0.104)	1.172 (± 0.154)*
<i>rho</i> I15N	0.022 (0.091)	1.018 (± 0.079)	1.000 (± 0.010)	0.949 (± 0.272)

^a Mean absolute fitness with 95% upper limit confidence interval. An asterisk corresponds to an absolute fitness \bar{w}_a significantly < 0.00.

^b Mean relative fitness \pm 95% confidence interval. The asterisks represent significant deviation from the null hypothesis that mean fitness equals 1.0, with one asterisk denoting significance at $P < 0.05$ and a dot denoting $0.05 < P < 0.1$.

SUPPORTING INFORMATION

SI Materials and Methods

Estimation of relative fitness

Briefly, both competitors (the Ara- evolved line and the ancestral Ara+ REL1207 strain) were revived in LB broth and grown separately for one day at 37°C and a second day at the temperature of interest (20°C or 37°C) in 10 ml of DM25. Then the two competitors were mixed at a 1:1 volumetric ratio and diluted 100-fold into 10 ml of fresh DM25 and incubated one day at the assay temperature. The initial and final densities were estimated by plating the culture onto tetrazolium-arabinose (TA) agar plates.

Estimation of absolute fitness

Briefly, the assays began by reviving the clones in 1 ml of LB and incubating overnight at 37°C with constant shaking (100 rpm). The following day, cultures were diluted 100-fold into MgSO₄ (10mM solution) and transferred 10 µl of this dilution into 990 µl of DM25. The 96-well plates (Megatiter™ plates, Neptune) were covered with gas permeable seals (Thermo Scientific) and incubated one day at 37°C (day 0) to allow the clones to acclimate to culture conditions. The following days (days 1 to 3), we transferred 10 µl of the overnight culture in 990 µl of fresh DM25 (100-fold dilution) and we incubated the plates at the assay temperature. The plates were incubated in a shaking incubator (Innova 4300) with an accuracy of ±0.25°C. At the end of each day we measured population densities using an electronic particle counter (Coulter Counter model Multisizer™ 3 equipped with a 30 micron diameter aperture tube). To measure density, 50 µl of culture was diluted in 9.9 ml of Isoton®II diluent (Beckman Coulter),

and 50 μ l of the resulting dilution was counted electronically. We subtracted the background noise to the counts by measuring the number of particle in 50 μ l of sterile DM25. Our pilot study showed a high correlation between the densities estimated with viable cell counts (colony-forming units) and estimated with electronic counts ($R^2=0.867$, $P=7.55^{-15***}$).

To estimate the absolute fitness we used a fixed effects linear model, in which we defined the natural logarithm of the cell density as the dependent variable, and the time (daily transfers) as the independent variable. The cell density from the first day of acclimation (day 0) was excluded from the regression.

Let $y_{r,d}^l$ be the natural logarithm of the bacterial density, in which d is the day of transfer, l is the line analyzed (high-temperature adapted clones or the ancestor) and r is the replicate. We used the following model:

$$y_{r,d}^l = (\beta_{o,1} * 1_{r=1} + \beta_{o,2} * 1_{r=2} + \beta_{o,3} * 1_{r=3}) + \beta_1 * d + \varepsilon_{r,d}^l$$

where β_0 is the intercept for each replicate, β_1 is the regression coefficient and $\varepsilon_{r,d}^l$ is the error term. The slope of the fitted linear regression β_1 is equivalent to the absolute fitness of a given genotype at a given temperature.

Strain construction and confirmation of recombinants

We used the pJk611 recombineering plasmid, kindly provided by M. Raffatellu. The pJk611 plasmid is identical to the pkD46 plasmid, with the addition of a *sacB* gene used to eliminate the plasmid when counter-selected on LB with sucrose.

Briefly, we first introduced the pJk611 plasmid into the ancestral strain, electroporating 2 μ l of plasmid (containing between 0.5 and 1 μ g of plasmid) into 50 μ l

of competent cells using an Eppendorf Electroporator 2510 set at 1.8 kV. Following electroporation, we added 1 ml LB and incubated the cells at 30°C for 2 h with shaking. We then plated 100 µl of cells on LB agar plates containing 100 µg/ml ampicillin to select ampicillin-resistant (amp^{R}) transformants. The ancestral strain carrying the pJk611 plasmid was then grown overnight at room temperature (~20°C) in 25 ml of LB with 100 µg/ml of ampicillin and 1 mM L-arabinose (Sigma) until it reached an OD_{600} of 0.6. We made electrocompetent cells by washing the cultures 5 times with ice-cold water. We simultaneously introduce two oligos of 70 bp with the desired nucleotide change in the center of the oligo (Table S3.3). The first oligo was used to introduce single mutation in *rpoB* or *rho*; the second oligo was to introduce the mutation that produces the Ara⁺ phenotype. 2 µl of each oligo (10 µM) was electroporated into 50 µl of cells. After electroporation we added 1 ml of LB and incubated cells at 37°C for 3 h with shaking and spread 500 µl in minimal medium agar (MA) plates supplemented with L-arabinose. The remainder was grown overnight at room temperature and 100 µl was spread in MA plates. The plates were incubated 48 h at 42°C. We selected 94 single colonies and streaked them onto tetrazolium-arabinose (TA) agar plates, incubated overnight at 37°C. We screened for the mutations by doing PCR on single colonies to amplify the region of the *rpoB* or *rho* gene with the mutation and Sanger sequencing the fragments of ~370 bp (Table S3.3). The PCR thermal cycling conditions were 94°C for 10 min followed by 30 cycles of 94°C 20 sec, 60°C 30 sec and 68°C 1 min; finally 68°C for 5 min.

Genetic associations

We first created a list of unique mutations (these mutations can be any kind of molecular change, from point mutations to large deletions) occurring in 114 high-temperature adapted clones. We then selected five of the most commonly shared mutations (mutations shared by 14 or more clones), which represent a disjoint set at the right tail of the frequency distribution (Figure S3.3), to explore the effect of each mutation on the absolute fitness (Table 3.1).

We selected L_m , which is the list of the high-temperature adapted clones in which the mutation m occurred (see Figure S3.7 for a simplified example). We then found all the mutations occurring in L_m , denoted by M_m (including m), and included them in the analysis.

Except for the mutations that are shared in more than one clone, most of the mutations are co-occurring in the same clone, which prevents us from drawing conclusions regarding the effect of single mutations on the phenotype. To overcome this problem, we partitioned the mutations into clusters C . A cluster is a subset of mutations that always occur together, and it contains a mutation m if, and only if $m \in C$ for all lines in L_m . If a mutation has no co-occurring mutation, then it forms its own cluster. We created a cluster incidence matrix showing the presence or absence of a cluster in a clone (Figure S3.7). We then combined the cluster incidence matrix with the data of absolute fitness at 18°C to fit the following fixed effects linear model:

$$y_{r,d}^l = (\beta_0^{l=1} * 1_{l=1} + \dots + \beta_0^{l=n} * 1_{l=n}) + \beta_1 * d + (\beta_{1,m_1} * d * 1_{m_1} + \dots + \beta_{1,m_M} * d * 1_{m_M}) + \varepsilon_{r,d}^l$$

where m_1, \dots, m_M are the groups of mutations that are being analyzed. In our model, $\beta_0^{l=1} * 1_{l=1} + \dots + \beta_0^{l=n} * 1_{l=n}$ are the intercepts coefficients associate to each clone,

β_1 is the absolute fitness of the overall population (All clones in L_m) without the effect of the mutations m_1 to m_M . In our model, the effect of the presence of the mutation m_i on the absolute fitness is given by the coefficient β_{1,m_i} .

In short, the model fitted a regression line for each one of the clones, where the slope of the regression for a given clone is given by the sum of the individual slope coefficient of the occurring mutations in that clone.

Finally, for each one of the most commonly shared mutations we selected a final model based on the lowest AIC (Akaike Information Criterion) and goodness of fit, to obtain a parsimonious model that quantify the contribution of each mutation group to the overall population fitness.

Supporting Figures

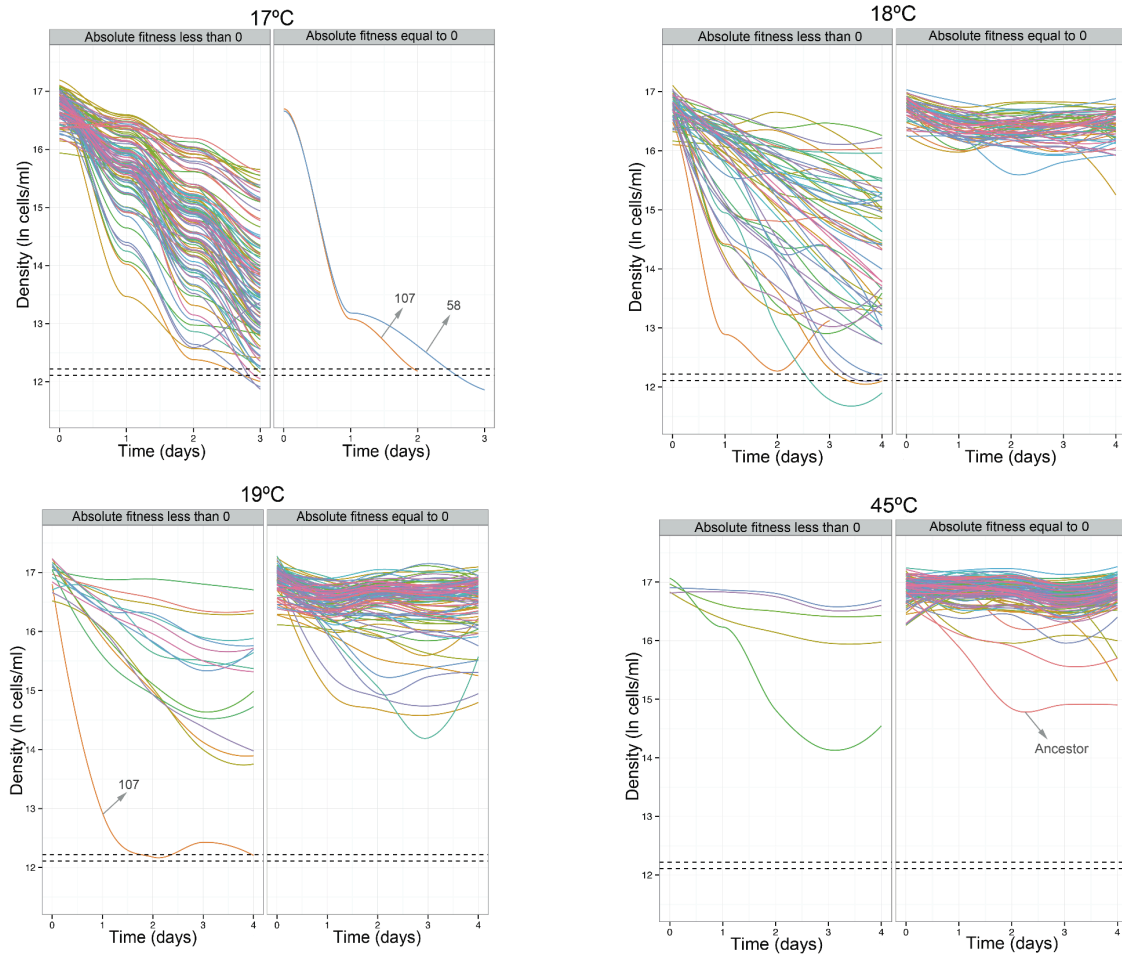


Figure S3.1. Population density trajectories of the 114 high-temperature adapted clones and the ancestor at 17°C, 18°C, 19°C and 45°C. For each temperature, the clones are divided in two groups: the panel on the left contains the clones in which the slope of the regression is significantly less than zero and the panel on the right contains the clones in which the slope of the regression is not significantly different from zero (Table S3.2). Each line corresponds to a local polynomial regression fitting of 3 replicates. The two horizontal dotted lines correspond to background noise (standard error limits estimated from the mean number of particles present in 71 samples of sterile DM25). Populations crossing the dotted lines are considered extinct. When the background noise was higher of the population density (negative value), the point was excluded from the analysis.

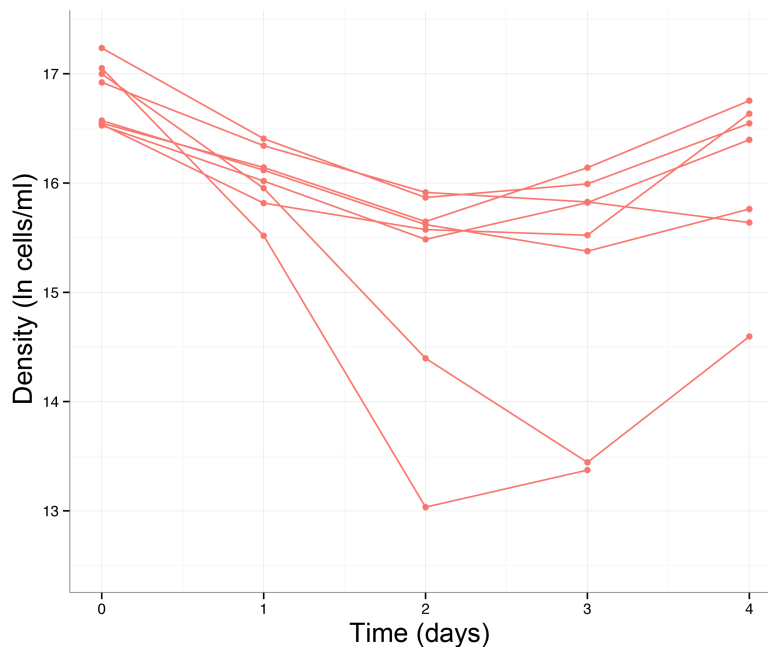


Figure S3.2. “Lazarus effect” observed in the ancestor at 45°C. Each point corresponds to the bacterial density measured at the end of a growth cycle (24 hours). The 8 solid lines connecting the circles correspond to 8 replicate measurements of the ancestor at 45°C.

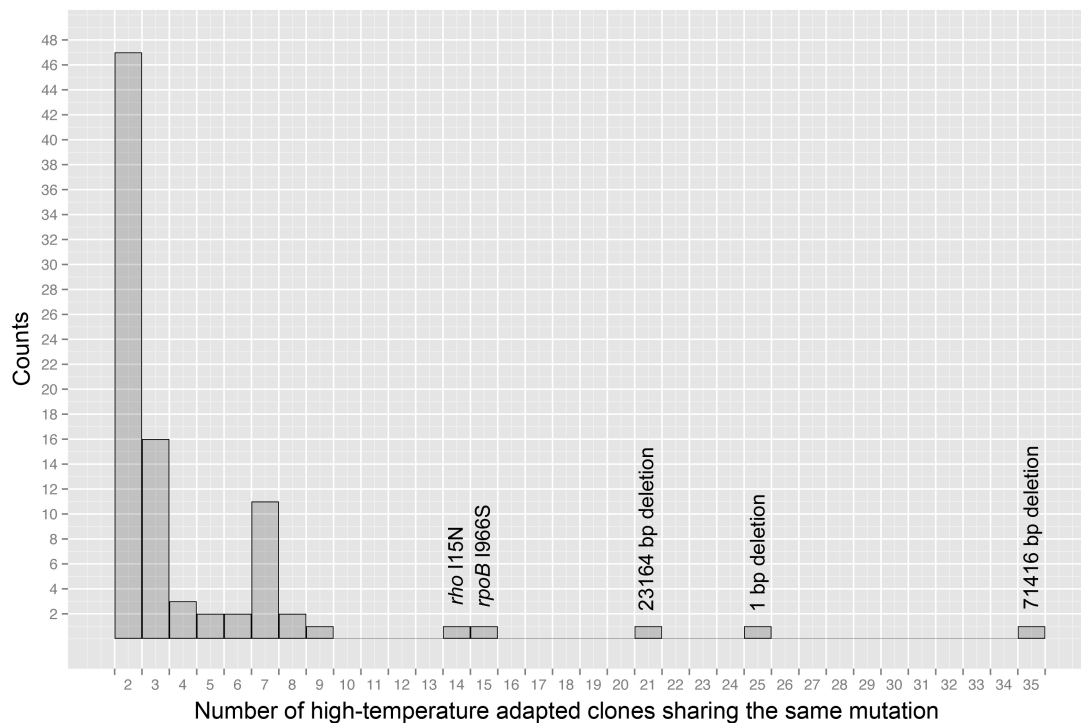


Figure S3.3. Distribution of the number of high-temperature adapted clones sharing the same identical mutation.

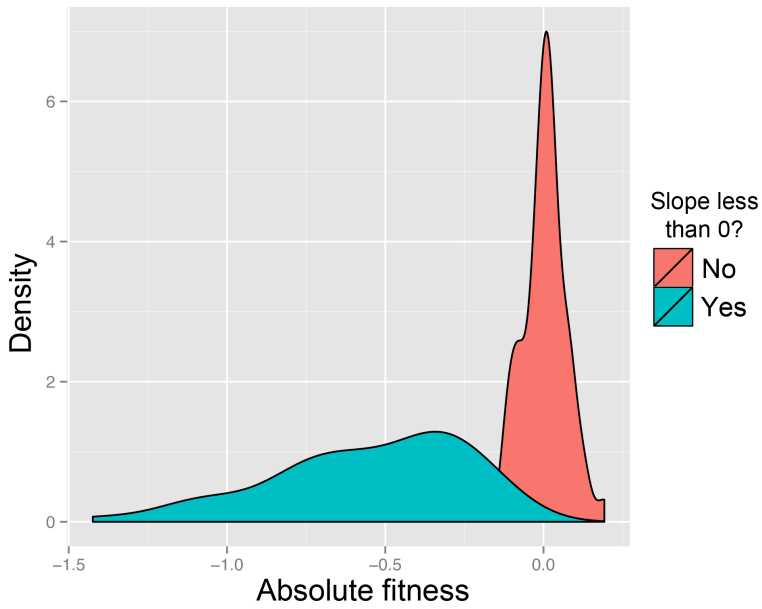


Figure S3.4. Kernel density plot from the absolute fitness data at 18°C. The area under the curve represents the empirical probability of occurrence of an absolute fitness value at 18°C, for the population of 114 high-temperature adapted clones.

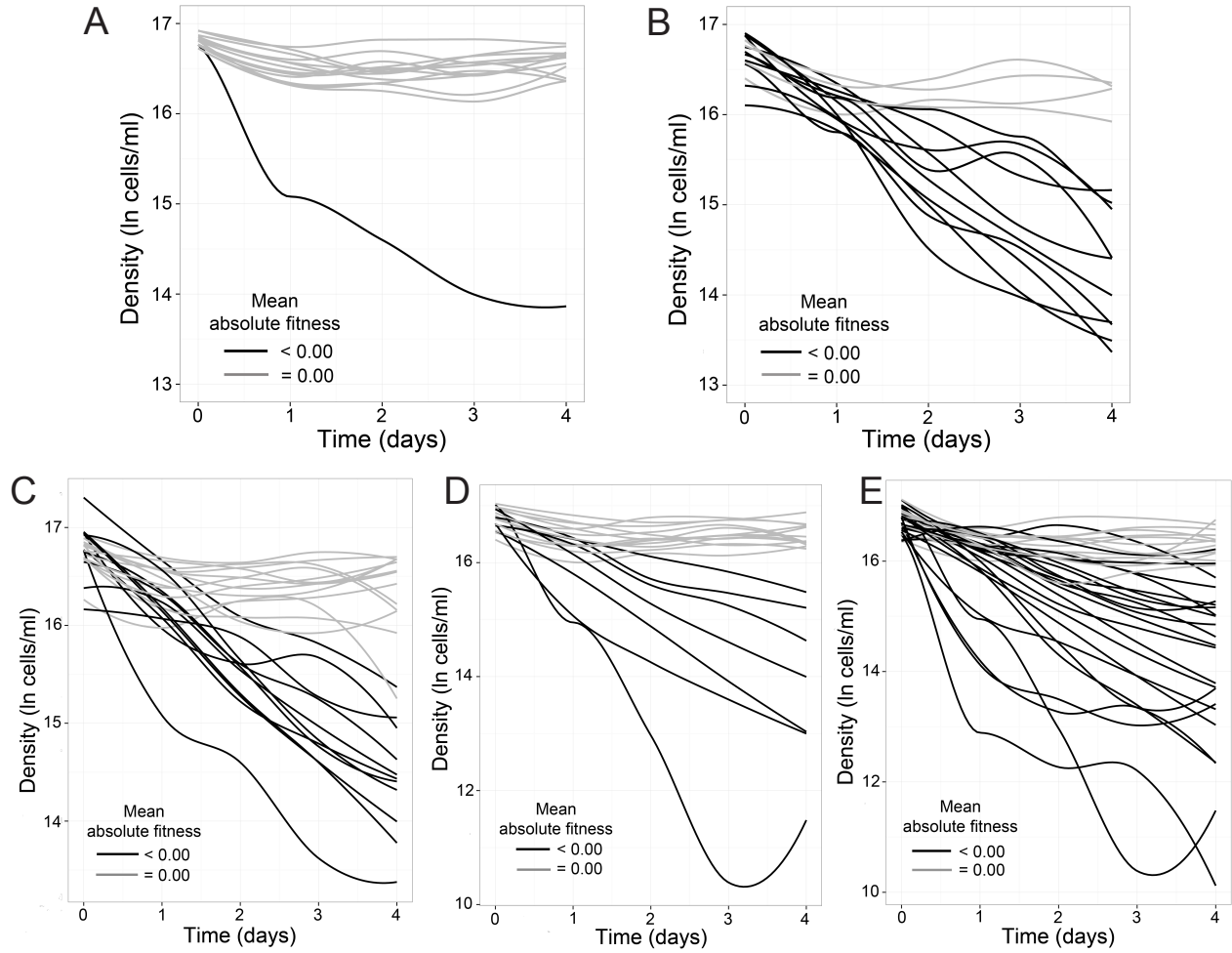


Figure S3.5. Population density trajectories of the high-temperature evolved clones with the (A) *rho* I15N mutation, (B) *rpoB* I966S mutation, (C) 1-bp deletion, (D) 23164 bp large deletion and (E) 71416 bp large deletion.

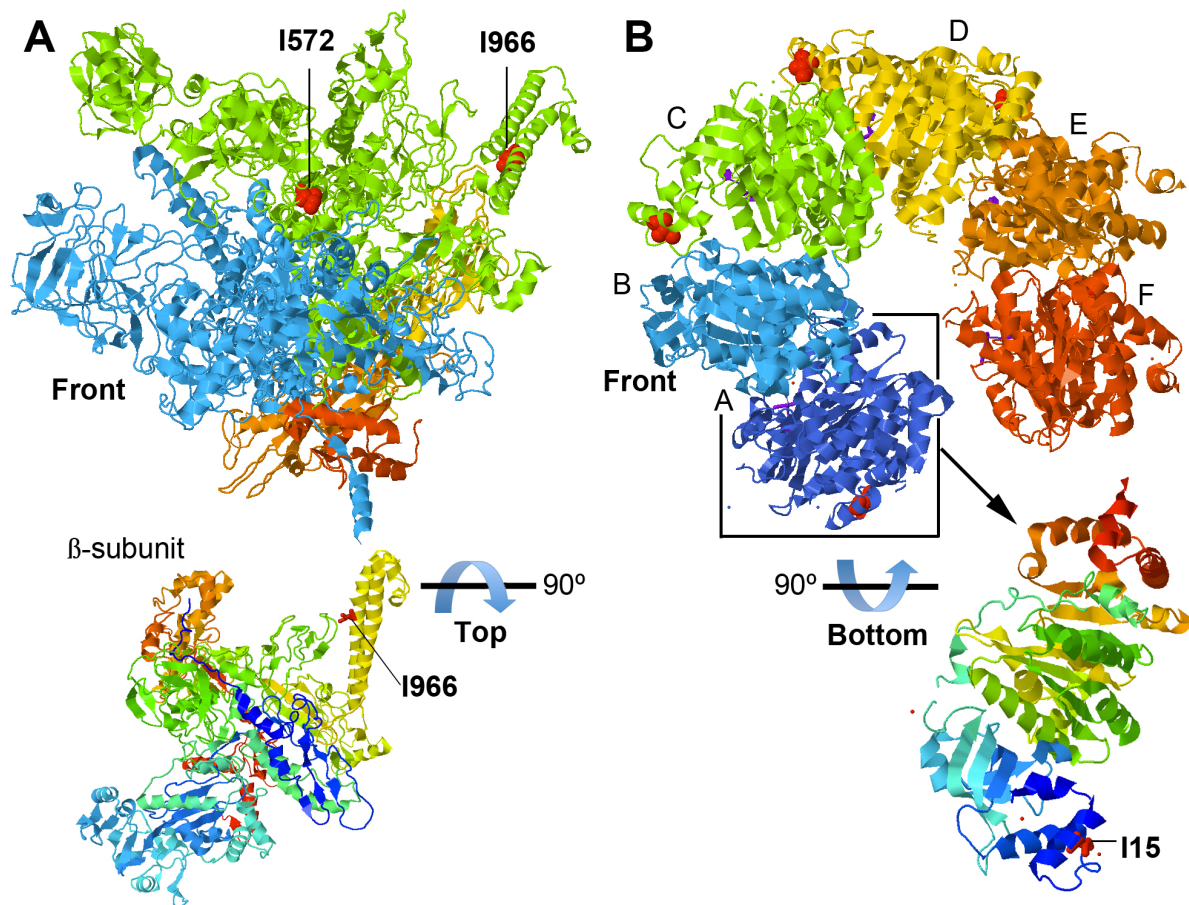


Figure S3.6. Structural mapping of the codons analyzed. (A) Front view of the RNAP from *E. coli* (PDB# 3LUO), color-coded as follows: α I and α II, yellow; β , green; β' , blue; ω , orange. Residues I572 (in the active site of the RNAP) and I966 (at the base of the flap domain) are shown as red spheres. Bellow is the top view of a ribbon diagram of the β -subunit colored from the N-terminus (blue) to the C-terminus (red) using a rainbow color gradient. Solvent-exposed residues in the back face of the ladder, including residue I966, are believed to interact with an unidentified regulatory factor (Opalka et al. 2010). (B) Front view of the Rho hexamer from *E. coli* (PDB# 1PV4). Residue I15 is shown as red sphere. Bellow is the bottom view of a Rho protomer colored from the N-terminus (blue) to the C-terminus (red) using a rainbow color gradient.

Supporting Tables

Table S3.1. Statistical analyses summary of the relative fitnesses of the 114 high-temperature adapted clones at 20°C, 37°C (this study) and 42.2°.

Clones	Mean Fitness at 20°C			Significance (P value) ³	Mean Fitness at 37°C			Significance (P value) ³	Mean Fitness at 42.2°C (Tenailon et al. 2012)			Significance (P value) ³
	Mean Fitness ±SE ¹	±95% CI ²			Mean Fitness ±SE ¹	±95% CI ²			Mean Fitness ±SE ¹	±95% CI ²		
1	0.853	0.013	0.058	0.008**	0.998	0.016	0.067	0.924	1.409	0.024	0.063	<0.001***
2	0.986	0.035	0.151	0.723	0.989	0.031	0.134	0.753	1.484	0.062	0.160	0.001**
3	0.981	0.019	0.050	0.370	0.922	0.013	0.057	0.028*	1.257	0.065	0.167	0.011*
4	0.936	0.023	0.058	0.038*	1.004	0.014	0.062	0.820	1.327	0.079	0.202	0.009**
5	1.023	0.024	0.062	0.386	0.958	0.032	0.139	0.326	1.559	0.109	0.279	0.004**
7	0.980	0.028	0.120	0.553	0.874	0.029	0.123	0.048*	1.529	0.054	0.139	<0.001***
8	0.958	0.013	0.033	0.023*	0.935	0.030	0.130	0.164	1.525	0.032	0.082	<0.001***
9	0.834	0.013	0.032	0.000***	0.970	0.036	0.153	0.493	1.472	0.051	0.130	<0.001***
10	1.011	0.009	0.040	0.355	1.019	0.033	0.085	0.598	1.667	0.072	0.186	<0.001***
11	0.866	0.029	0.123	0.043*	0.826	0.025	0.108	0.020*	1.244	0.032	0.083	0.001**
12	0.882	0.028	0.121	0.053	0.960	0.019	0.048	0.084	1.323	0.048	0.124	0.001**
13	0.926	0.011	0.046	0.020*	0.986	0.018	0.076	0.504	1.566	0.078	0.200	0.001**
14	0.897	0.003	0.014	0.001**	1.006	0.023	0.060	0.819	1.603	0.090	0.231	0.001**
15	0.973	0.021	0.089	0.321	0.990	0.045	0.192	0.841	1.499	0.070	0.180	0.001**
16	0.888	0.016	0.068	0.019*	0.948	0.028	0.120	0.201	1.272	0.070	0.180	0.012*
17	0.973	0.025	0.106	0.381	0.944	0.031	0.132	0.207	1.460	0.044	0.113	<0.001***
18	0.698	0.065	0.167	0.043*	0.889	0.026	0.110	0.049*	1.561	0.104	0.267	0.003**
20	0.861	0.028	0.120	0.038*	0.995	0.023	0.097	0.845	1.278	0.053	0.137	0.003**
21	0.967	0.023	0.097	0.286	1.019	0.016	0.040	0.272	1.478	0.118	0.303	0.010*
22	0.867	0.021	0.053	0.001**	0.962	0.003	0.014	0.008**	1.325	0.042	0.107	0.001**
23	0.996	0.018	0.076	0.837	1.027	0.030	0.131	0.464	1.129	0.035	0.089	0.014*
24	0.900	0.025	0.065	0.011*	1.003	0.024	0.103	0.913	1.570	0.067	0.172	<0.001***
25	0.897	0.011	0.027	0.000***	0.946	0.035	0.151	0.261	1.184	0.038	0.098	0.005**
26	1.006	0.025	0.106	0.842	0.939	0.011	0.048	0.031*	1.503	0.037	0.102	<0.001***
27	0.900	0.014	0.036	0.001**	0.966	0.007	0.030	0.039*	1.243	0.058	0.150	0.009**
28	0.790	0.009	0.022	0.000***	1.010	0.044	0.188	0.846	1.503	0.083	0.214	0.002**
31	0.742	0.037	0.161	0.020*	0.997	0.041	0.175	0.947	1.330	0.049	0.126	0.001**
32	0.948	0.016	0.068	0.080	1.034	0.021	0.054	0.169	1.290	0.042	0.109	0.001**
33	0.646	0.020	0.084	0.003**	1.014	0.009	0.038	0.247	1.402	0.036	0.092	<0.001***
34	0.978	0.029	0.124	0.521	1.013	0.047	0.202	0.812	1.510	0.059	0.152	<0.001***
35	0.942	0.010	0.044	0.030*	1.041	0.019	0.049	0.084	1.473	0.062	0.159	0.001**
38	0.964	0.041	0.176	0.476	0.968	0.022	0.096	0.292	1.483	0.054	0.140	<0.001***
39	0.956	0.015	0.066	0.104	0.968	0.018	0.076	0.210	1.437	0.068	0.174	0.001**
40	0.706	0.050	0.216	0.028*	0.942	0.005	0.020	0.006**	1.368	0.037	0.095	<0.001***
41	0.897	0.012	0.051	0.013*	0.959	0.008	0.032	0.032*	1.306	0.061	0.157	0.004**

42	0.834	0.038	0.164	0.049*	0.964	0.036	0.156	0.422	1.561	0.053	0.137	<0.001***
43	0.865	0.018	0.047	0.001**	0.970	0.041	0.176	0.544	1.496	0.048	0.124	<0.001***
44	0.932	0.052	0.224	0.320	1.028	0.019	0.048	0.192	1.516	0.078	0.201	0.001**
45	0.942	0.025	0.109	0.149	0.993	0.022	0.095	0.793	1.293	0.061	0.156	0.005**
46	0.941	0.034	0.145	0.222	1.030	0.013	0.035	0.077.	1.473	0.063	0.161	0.001**
47	0.875	0.025	0.106	0.036*	1.065	0.080	0.346	0.501	1.546	0.077	0.197	0.001**
48	0.929	0.014	0.061	0.038*	0.967	0.020	0.088	0.253	1.298	0.065	0.180	0.010*
51	0.905	0.020	0.086	0.042*	1.004	0.013	0.056	0.798	1.349	0.078	0.200	0.006**
52	1.009	0.061	0.265	0.901	0.972	0.027	0.115	0.411	1.415	0.053	0.136	0.001**
53	0.911	0.029	0.125	0.093.	0.982	0.028	0.122	0.581	1.096	0.043	0.110	0.074.
54	0.916	0.012	0.050	0.018*	0.995	0.035	0.149	0.904	1.384	0.079	0.204	0.005**
55	0.792	0.034	0.146	0.026*	0.896	0.023	0.100	0.046*	1.390	0.091	0.235	0.008**
56	1.035	0.018	0.047	0.117	0.940	0.015	0.039	0.011*	1.319	0.070	0.179	0.006**
57	0.779	0.050	0.129	0.007**	0.939	0.010	0.026	0.002**	1.279	0.043	0.110	0.001**
58	0.932	0.008	0.019	0.012*	0.832	0.020	0.087	0.014*	1.419	0.092	0.235	0.006**
59	0.845	0.024	0.102	0.023*	0.954	0.024	0.102	0.189	1.383	0.066	0.170	0.002**
60	0.993	0.017	0.072	0.723	1.016	0.021	0.053	0.487	1.277	0.043	0.110	0.001**
61	0.835	0.030	0.130	0.032*	0.998	0.026	0.112	0.952	1.536	0.107	0.276	0.004**
64	0.870	0.004	0.018	0.001**	0.906	0.008	0.022	0.000***	1.495	0.073	0.187	0.001**
65	0.914	0.004	0.018	0.002**	1.030	0.049	0.213	0.604	1.405	0.052	0.134	0.001**
66	1.033	0.011	0.028	0.029*	0.941	0.009	0.040	0.024*	1.430	0.059	0.151	0.001**
67	0.926	0.025	0.106	0.095.	0.969	0.006	0.027	0.037*	1.378	0.032	0.082	<0.001***
68	0.892	0.013	0.055	0.014*	0.882	0.022	0.097	0.034*	1.283	0.060	0.155	0.005**
69	0.907	0.032	0.138	0.102	1.041	0.032	0.137	0.322	1.541	0.047	0.121	<0.001***
70	1.029	0.015	0.063	0.183	0.992	0.025	0.064	0.752	1.308	0.035	0.091	<0.001***
71	0.833	0.023	0.099	0.018*	0.900	0.029	0.073	0.017*	1.270	0.060	0.154	0.006**
72	0.998	0.046	0.196	0.973	0.992	0.021	0.090	0.736	1.463	0.095	0.244	0.005**
73	1.031	0.016	0.069	0.192	1.028	0.019	0.084	0.285	1.419	0.062	0.160	0.001**
74	0.839	0.009	0.037	0.003**	0.857	0.026	0.110	0.030*	1.314	0.041	0.114	0.002**
75	0.881	0.012	0.053	0.011*	1.085	0.025	0.064	0.019*	1.357	0.062	0.158	0.002**
76	0.703	0.022	0.095	0.005**	0.976	0.005	0.020	0.034*	1.377	0.035	0.089	<0.001***
77	0.912	0.022	0.094	0.057.	0.930	0.004	0.016	0.003**	1.430	0.096	0.248	0.007**
78	1.004	0.013	0.056	0.790	1.033	0.062	0.269	0.653	1.570	0.096	0.247	0.002**
79	0.953	0.014	0.062	0.081.	0.977	0.018	0.078	0.326	1.238	0.048	0.132	0.008**
80	0.955	0.030	0.078	0.196	1.003	0.038	0.165	0.953	1.270	0.079	0.203	0.019*
81	0.864	0.025	0.064	0.003**	1.016	0.015	0.037	0.321	1.480	0.078	0.201	0.002**
82	0.876	0.019	0.050	0.001**	0.939	0.017	0.043	0.015*	1.498	0.053	0.137	<0.001***
83	0.995	0.030	0.128	0.881	0.970	0.026	0.112	0.374	1.178	0.036	0.092	0.004**
84	0.978	0.019	0.080	0.353	0.999	0.017	0.044	0.971	1.180	0.070	0.181	0.051.
85	0.848	0.020	0.085	0.017*	0.996	0.013	0.034	0.758	1.662	0.087	0.222	0.001**
86	0.866	0.014	0.061	0.011*	0.952	0.019	0.050	0.055.	1.310	0.077	0.198	0.010*
87	1.010	0.032	0.139	0.781	1.015	0.032	0.139	0.689	1.703	0.137	0.351	0.004**
89	0.941	0.044	0.189	0.313	0.975	0.041	0.177	0.600	1.387	0.046	0.117	<0.001***

91	0.784	0.022	0.095	0.010*	0.990	0.023	0.098	0.695	1.472	0.057	0.146	<0.001***	
92	0.825	0.009	0.040	0.003**	0.918	0.016	0.068	0.035*	1.446	0.093	0.239	0.005**	
93	0.907	0.011	0.049	0.015*	1.010	0.019	0.081	0.650	1.396	0.040	0.103	<0.001***	
94	0.968	0.001	0.005	0.001**	0.941	0.007	0.028	0.012*	1.767	0.134	0.344	0.002**	
95	0.903	0.001	0.003	0.000***	0.921	0.013	0.034	0.002**	1.488	0.109	0.280	0.007**	
96	0.934	0.048	0.124	0.302	0.990	0.018	0.077	0.641	1.439	0.105	0.270	0.009**	
97	0.926	0.105	0.269	0.512	0.978	0.019	0.049	0.310	1.277	0.096	0.247	0.034*	
101	0.976	0.018	0.047	0.248	1.014	0.016	0.041	0.438	1.404	0.031	0.079	<0.001***	
105	0.939	0.020	0.088	0.095.	1.025	0.020	0.087	0.338	1.397	0.038	0.097	<0.001***	
106	0.928	0.008	0.035	0.012*	1.016	0.031	0.079	0.620	1.383	0.100	0.258	0.012*	
107	Could not grow at 20°C					0.982	0.011	8	0.168	1.635	0.061	0.157	<0.001***
108	0.870	0.003	0.014	0.001**	0.989	0.020	0.085	0.642	1.671	0.112	0.289	0.002**	
110	0.913	0.028	0.072	0.089.	1.007	0.006	0.024	0.361	1.469	0.073	0.187	0.001**	
112	0.906	0.004	0.018	0.002**	0.992	0.019	0.048	0.693	1.230	0.085	0.218	0.042*	
114	0.830	0.012	0.052	0.005**	0.967	0.005	0.022	0.023*	1.217	0.063	0.162	0.018*	
118	0.849	0.022	0.056	0.001**	0.935	0.004	0.017	0.004**	1.612	0.039	0.101	<0.001***	
119	0.716	0.015	0.038	0.003**	0.953	0.009	0.038	0.033*	1.626	0.098	0.253	0.001**	
120	0.807	0.020	0.086	0.011*	0.954	0.015	0.039	0.029*	1.300	0.055	0.153	0.005**	
122	0.908	0.015	0.066	0.027*	0.975	0.044	0.189	0.626	1.381	0.033	0.084	<0.001***	
124	0.935	0.017	0.044	0.012*	0.913	0.015	0.066	0.030*	1.427	0.048	0.122	<0.001***	
126	0.998	0.017	0.043	0.929	0.970	0.025	0.108	0.358	1.544	0.061	0.157	<0.001***	
127	0.890	0.025	0.063	0.006**	1.003	0.012	0.031	0.822	1.387	0.047	0.120	<0.001***	
130	0.999	0.011	0.028	0.935	0.990	0.010	0.041	0.392	1.284	0.068	0.188	0.014*	
131	0.968	0.026	0.114	0.348	0.948	0.012	0.050	0.047*	1.345	0.036	0.093	<0.001***	
132	1.009	0.015	0.066	0.601	0.980	0.053	0.227	0.737	1.432	0.082	0.212	0.003**	
133	0.684	0.012	0.053	0.002**	0.919	0.004	0.015	0.002**	1.379	0.070	0.180	0.003**	
134	1.032	0.017	0.044	0.118	0.985	0.017	0.044	0.408	1.380	0.053	0.136	0.001**	
135	0.947	0.018	0.046	0.031*	0.993	0.017	0.074	0.715	1.473	0.056	0.143	<0.001***	
136	0.911	0.007	0.018	0.000***	0.886	0.000	0.002	0.000***	1.295	0.028	0.072	<0.001***	
137	0.932	0.006	0.025	0.008**	0.927	0.008	0.036	0.013*	1.349	0.041	0.104	<0.001***	
138	0.968	0.021	0.088	0.261	1.011	0.017	0.043	0.558	1.433	0.081	0.209	0.003**	
139	0.877	0.015	0.039	0.000***	0.966	0.024	0.104	0.299	1.375	0.053	0.135	0.001**	
140	0.981	0.024	0.104	0.516	1.024	0.031	0.134	0.522	1.648	0.039	0.101	<0.001***	
141	0.991	0.025	0.064	0.736	0.956	0.009	0.038	0.038*	1.460	0.055	0.141	<0.001***	
142	0.885	0.015	0.038	0.001**	0.959	0.033	0.142	0.336	1.609	0.067	0.172	<0.001***	
143	0.927	0.017	0.043	0.007**	0.967	0.029	0.127	0.375	1.540	0.081	0.208	0.001**	

Table S3.2. Statistical analyses summary of the absolute fitnesses of the ancestor and the 114 high-temperature adapted clones at 17°C, 18°C, 19°C and 45°C.

Clones	Mean Absolute fitness at 17°C ¹	95% CI (upper limit) ²	Less than 0? ³	Mean Absolute fitness at 18°C ¹	95% CI (upper limit) ²	Less than 0? ³	Mean Absolute fitness at 19°C ¹	95% CI (upper limit) ²	Less than 0? ³	Mean Absolute fitness at 45°C ¹	95% CI (upper limit) ²	Less than 0? ³
1	-1.015	-0.751	YES	-0.415	-0.352	YES	-0.134	-0.006	YES	-0.092	0.095	NO
2	-0.301	-0.059	YES	0.057	0.134	NO	0.018	0.091	NO	-0.916	-0.312	YES
3	-0.896	-0.684	YES	-0.024	0.048	NO	0.022	0.121	NO	0.015	0.062	NO
4	-1.300	-1.084	YES	-0.543	-0.405	YES	-0.219	-0.067	YES	-0.038	0.017	NO
5	-0.931	-0.818	YES	-0.089	0.095	NO	0.115	0.216	NO	-0.036	0.041	NO
7	-0.480	-0.303	YES	-0.034	0.082	NO	0.073	0.126	NO	0.016	0.064	NO
8	-0.867	-0.551	YES	-0.260	-0.017	YES	0.096	0.213	NO	-0.062	0.021	NO
9	-1.627	-1.180	YES	-0.791	-0.696	YES	-0.285	-0.049	YES	-0.032	0.035	NO
10	-0.893	-0.572	YES	-0.221	-0.054	YES	0.034	0.100	NO	-0.109	0.024	NO
11	-1.254	-0.530	YES	-0.852	-0.325	YES	-0.688	-0.549	YES	-0.003	0.055	NO
12	-1.099	-0.456	YES	-0.550	-0.314	YES	-0.076	0.130	NO	-0.061	0.033	NO
13	-0.752	-0.507	YES	0.056	0.164	NO	0.033	0.089	NO	-0.032	0.042	NO
14	-1.025	-0.774	YES	-0.314	-0.177	YES	-0.114	0.061	NO	0.022	0.080	NO
15	-0.458	-0.171	YES	0.033	0.088	NO	0.039	0.131	NO	0.052	0.095	NO
16	-0.952	-0.770	YES	-0.393	-0.294	YES	-0.072	0.014	NO	-0.029	0.037	NO
17	-0.658	-0.485	YES	-0.104	-0.020	YES	0.055	0.130	NO	0.004	0.064	NO
18	-0.937	-0.492	YES	-0.746	-0.174	YES	-0.316	-0.068	YES	0.035	0.076	NO
20	-1.221	-0.775	YES	-0.609	-0.447	YES	-0.072	0.105	NO	-0.004	0.065	NO
21	-0.964	-0.705	YES	-0.059	0.027	NO	0.008	0.110	NO	0.018	0.080	NO
22	-1.262	-1.024	YES	-0.586	-0.485	YES	0.043	0.132	NO	-0.025	0.037	NO
23	-0.434	-0.190	YES	0.048	0.105	NO	-0.063	-0.002	YES [§]	-0.011	0.052	NO
24	-1.250	-0.863	YES	-0.478	-0.287	YES	-0.300	-0.120	YES	-0.069	0.006	NO
25	-1.000	-0.790	YES	-0.388	-0.258	YES	-0.057	0.008	NO	-0.081	0.016	NO
26	-1.155	-1.033	YES	0.079	0.188	NO	0.006	0.074	NO	0.029	0.115	NO
27	-1.236	-0.981	YES	-0.272	-0.188	YES	-0.052	0.051	NO	0.028	0.057	NO
28	-1.245	-0.666	YES	-0.306	-0.231	YES	-0.036	0.024	NO	-0.012	0.052	NO
31	-1.200	-0.893	YES	-1.003	-0.725	YES	-0.334	-0.127	YES	-0.016	0.037	NO
32	-1.343	-0.795	YES	0.013	0.094	NO	-0.037	0.021	NO	-0.038	0.034	NO
33	-1.847	-1.259	YES	-1.424	-0.674	YES	-0.216	0.220	NO	-0.013	0.044	NO
34	-1.071	-0.907	YES	-0.113	-0.047	YES	-0.082	0.020	NO	0.000	0.076	NO
35	-1.077	-0.741	YES	-0.236	-0.125	YES	0.023	0.173	NO	-0.024	0.035	NO
38	-0.991	-0.827	YES	0.029	0.131	NO	-0.016	0.062	NO	-0.061	0.005	NO
39	-0.825	-0.686	YES	-0.005	0.078	NO	-0.042	0.024	NO	-0.056	0.014	NO
40	-1.395	-1.192	YES	-0.928	-0.823	YES	-0.276	-0.060	YES	-0.038	0.012	NO
41	-1.078	-0.746	YES	-0.067	0.029	NO	-0.026	0.057	NO	-0.015	0.058	NO
42	-0.948	-0.680	YES	0.020	0.083	NO	0.043	0.115	NO	-0.016	0.052	NO

43	-1.071	-0.753	YES	-0.276	-0.165	YES	0.066	0.146	NO	0.012	0.084	NO
44	-0.724	-0.522	YES	0.004	0.088	NO	0.027	0.131	NO	0.014	0.075	NO
45	-1.151	-0.913	YES	-0.428	-0.232	YES	-0.034	0.118	NO	-0.002	0.072	NO
46	-1.165	-0.925	YES	-0.112	0.015	NO	0.037	0.103	NO	0.028	0.078	NO
47	-1.691	-1.274	YES	-0.598	-0.449	YES	-0.105	0.077	NO	0.010	0.050	NO
48	-0.704	-0.456	YES	-0.022	0.100	NO	0.050	0.159	NO	-0.001	0.083	NO
51	-0.815	-0.583	YES	0.011	0.090	NO	0.049	0.146	NO	0.015	0.071	NO
52	-0.867	-0.613	YES	0.020	0.089	NO	0.027	0.110	NO	0.000	0.045	NO
53	-1.124	-0.966	YES	-0.018	0.070	NO	0.069	0.176	NO	-0.007	0.053	NO
54	-0.621	-0.392	YES	0.024	0.122	NO	0.085	0.129	NO	-0.016	0.039	NO
55	-1.944	-1.444	YES	-0.713	-0.637	YES	-0.231	-0.070	YES	-0.018	0.082	NO
56	-1.274	-0.956	YES	-0.176	-0.065	YES	0.070	0.125	NO	0.003	0.108	NO
57	-1.984	-1.181	YES	-1.137	-0.805	YES	-0.248	-0.011	YES	-0.017	0.041	NO
58	-1.509	0.613	NO*	-1.205	-0.582	YES	-0.185	0.143	NO	0.021	0.088	NO
59	-1.610	-1.273	YES	-1.048	-0.924	YES	-0.070	0.043	NO	-0.027	0.006	NO
60	-0.990	-0.676	YES	0.043	0.134	NO	0.047	0.149	NO	-0.074	-0.017	YES
61	-1.409	-1.165	YES	-0.378	-0.211	YES	-0.084	0.102	NO	-0.026	0.030	NO
64	-1.396	-1.151	YES	-0.353	-0.193	YES	0.085	0.179	NO	-0.042	0.006	NO
65	-1.425	-1.057	YES	-0.079	0.030	NO	-0.005	0.068	NO	-0.046	0.020	NO
66	-0.904	-0.545	YES	0.120	0.223	NO	0.053	0.111	NO	-0.017	0.030	NO
67	-1.369	-0.894	YES	-0.617	-0.453	YES	0.005	0.109	NO	-0.054	0.017	NO
68	-1.437	-0.679	YES	-0.576	-0.427	YES	-0.135	0.023	NO	-0.042	0.078	NO
69	-1.123	-0.894	YES	-0.490	-0.348	YES	0.003	0.079	NO	0.012	0.078	NO
70	-0.885	-0.670	YES	0.093	0.211	NO	0.054	0.169	NO	-0.001	0.031	NO
71	-1.313	-0.828	YES	-1.070	-0.845	YES	-0.166	0.059	NO	-0.016	0.038	NO
72	-1.352	-1.068	YES	-0.148	0.000	YES	-0.060	0.098	NO	-0.010	0.038	NO
73	-0.648	-0.367	YES	-0.091	0.034	NO	0.092	0.166	NO	-0.048	0.019	NO
74	-1.193	-0.245	YES	-0.677	-0.419	YES	-0.129	0.063	NO	0.041	0.093	NO
75	-1.871	-1.511	YES	-0.808	-0.426	YES	-0.093	0.097	NO	-0.006	0.075	NO
76	-1.305	-1.026	YES	-0.686	-0.478	YES	-0.606	-0.447	YES	-0.012	0.079	NO
77	-1.199	-0.947	YES	-0.321	-0.228	YES	0.021	0.114	NO	-0.092	-0.008	YES
78	-0.634	-0.546	YES	0.040	0.144	NO	0.039	0.116	NO	-0.027	0.021	NO
79	-1.189	-1.035	YES	-0.077	0.020	NO	0.077	0.156	NO	-0.068	0.018	NO
80	-1.058	-0.818	YES	0.033	0.128	NO	0.012	0.081	NO	-0.007	0.057	NO
81	-0.885	-0.714	YES	0.043	0.136	NO	0.049	0.104	NO	-0.028	0.047	NO
82	-0.674	-0.416	YES	0.071	0.132	NO	-0.003	0.079	NO	-0.044	0.006	NO
83	-1.128	-0.716	YES	-0.034	0.042	NO	0.043	0.108	NO	-0.034	0.006	NO
84	-0.913	-0.540	YES	-0.009	0.109	NO	0.062	0.139	NO	-0.013	0.046	NO
85	-1.602	-0.979	YES	-0.756	-0.605	YES	-0.206	0.039	NO	-0.019	0.017	NO
86	-1.587	-1.280	YES	-0.089	0.268	NO	-0.041	0.061	NO	-0.005	0.032	NO
87	-0.607	-0.510	YES	0.074	0.162	NO	0.044	0.111	NO	-0.009	0.062	NO
89	-0.995	-0.802	YES	-0.003	0.037	NO	0.040	0.122	NO	-0.001	0.057	NO
91	-1.584	-1.189	YES	-0.806	-0.491	YES	0.020	0.094	NO	-0.026	0.023	NO

92	-1.943	-1.370	YES	-0.725	-0.440	YES	-0.353	-0.101	YES	-0.074	0.023	NO
93	-1.079	-0.735	YES	-0.118	0.019	NO	0.048	0.094	NO	-0.037	0.057	NO
94	-0.844	-0.570	YES	0.113	0.215	NO	0.070	0.141	NO	0.006	0.054	NO
95	-0.981	-0.822	YES	0.015	0.126	NO	0.022	0.081	NO	-0.009	0.026	NO
96	-0.747	-0.573	YES	0.086	0.175	NO	-0.015	0.084	NO	-0.004	0.043	NO
97	-1.270	-0.675	YES	-0.119	0.011	NO	0.032	0.075	NO	-0.157	0.038	NO
101	-0.838	-0.539	YES	0.009	0.189	NO	0.097	0.182	NO	0.043	0.089	NO
105	-0.591	-0.343	YES	0.023	0.119	NO	0.036	0.094	NO	-0.008	0.051	NO
106	-1.258	-0.974	YES	0.081	0.151	NO	0.092	0.159	NO	-0.016	0.031	NO
107	-0.956	1.094	NO*	-0.690	-0.334	YES	-0.354	-0.072	YES	-0.032	0.023	NO
108	-1.043	-0.782	YES	0.080	0.162	NO	0.050	0.107	NO	0.026	0.068	NO
110	-1.283	-1.033	YES	0.136	0.380	NO	0.069	0.141	NO	0.028	0.066	NO
112	-1.289	-0.858	YES	-0.024	0.045	NO	-0.002	0.120	NO	0.039	0.074	NO
114	-1.658	-0.794	YES	-0.808	-0.680	YES	-0.228	0.002	NO	-0.380	0.058	NO
118	-1.609	-1.256	YES	-0.372	-0.249	YES	-0.159	0.040	NO	0.044	0.088	NO
119	-1.137	-0.759	YES	-0.298	-0.124	YES	-0.100	0.085	NO	0.005	0.076	NO
120	-1.414	-1.239	YES	-0.223	0.005	NO	-0.021	0.277	NO	0.041	0.083	NO
122	-1.120	-0.975	YES	-0.664	-0.430	YES	0.023	0.100	NO	0.023	0.099	NO
124	-1.662	-1.415	YES	-0.536	-0.360	YES	-0.036	0.021	NO	-0.008	0.084	NO
126	-0.599	-0.371	YES	0.012	0.037	NO	0.065	0.132	NO	-0.052	0.042	NO
127	-1.398	-1.121	YES	-0.257	-0.083	YES	-0.153	0.018	NO	0.014	0.097	NO
130	-1.369	-1.058	YES	-0.454	-0.303	YES	-0.114	-0.004	YES	0.060	0.157	NO
131	-0.426	-0.274	YES	-0.035	0.031	NO	0.065	0.111	NO	-0.137	-0.066	YES
132	-0.449	-0.292	YES	-0.010	0.099	NO	0.039	0.124	NO	0.200	0.500	NO
133	-1.731	-1.523	YES	-0.834	-0.669	YES	-0.840	-0.595	YES	0.049	0.157	NO
134	-0.563	-0.263	YES	0.004	0.089	NO	0.070	0.146	NO	-0.319	0.172	NO
135	-1.491	-1.251	YES	-0.387	-0.234	YES	-0.107	0.098	NO	0.038	0.095	NO
136	-1.548	-1.160	YES	-0.364	-0.286	YES	-0.185	0.062	NO	0.024	0.072	NO
137	-1.172	-0.933	YES	-0.111	0.028	NO	0.046	0.156	NO	-0.015	0.024	NO
138	-0.986	-0.700	YES	-0.061	0.047	NO	0.043	0.179	NO	-0.033	0.033	NO
139	-1.074	-0.803	YES	0.018	0.118	NO	0.032	0.101	NO	0.016	0.054	NO
140	-0.939	-0.657	YES	0.020	0.088	NO	0.070	0.148	NO	0.014	0.063	NO
141	-0.801	-0.463	YES	-0.005	0.068	NO	0.105	0.197	NO	0.000	0.072	NO
142	-1.078	-0.784	YES	0.001	0.129	NO	0.106	0.181	NO	-0.067	0.000	NO
143	-0.508	-0.174	YES	-0.023	0.041	NO	0.062	0.122	NO	-0.062	-0.012	YES
REL												
1206	-0.413	-0.264	YES	0.000	0.075	NO	-0.004	0.070	NO	-0.290	0.097	NO

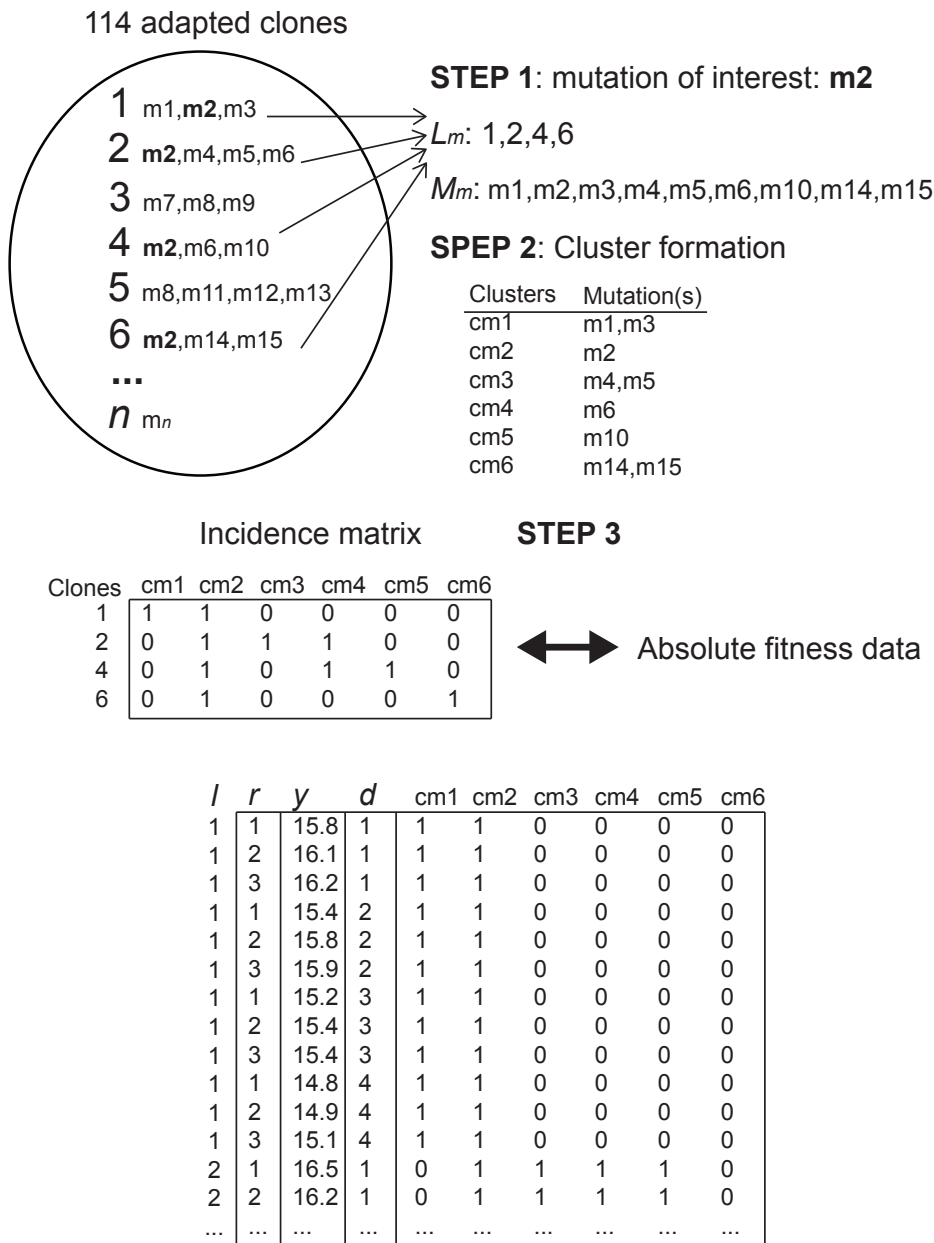


Figure S3.7. Schematic representation of the procedure followed to prepare the data that were analyzed in the genetic association study. In this simplified example, only 6 high temperature adapted clones with few mutations are considered. The first step consists in the selection of the mutations and the high-temperature clones to be analyzed. The second step consists in forming clusters of mutations co-occurring in the same clone. In the last step, the mutation clusters incidence matrix is combined with the absolute fitness data, containing the values of bacterial density (y) of the high-temperature adapted clones (l), measured in different days (d) and with several replicates (r).

REFERENCES CITED

- Al Mamun, A.A.M., M.-J. Lombardo, C. Shee, A.M. Lisewski, C. Gonzalez, D. Lin, R.B. Nehring, C. Saint-Ruf, J.L. Gibson, and R.L. Frisch. 2012. Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science* 338:1344-1348.
- Alto, B.W., B.R. Wasik, N.M. Morales, and P.E. Turner. 2013. Stochastic temperatures impede RNA virus adaptation. *Evolution* 67:969-979
- Angilletta Jr, M.J., P.H. Niewiarowski, and C.A. Navas. 2002. The evolution of thermal physiology in ectotherms. *J Therm Biol* 27:249-268.
- Angilletta, M.J., R.S. Wilson, C.A. Navas, and R.S. James. 2003. Trade-offs and the evolution of thermal reaction norms. *Trends Ecol Evol* 18:234-240.
- Barrett, R.D.H., and H.E. Hoekstra. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet* 12:767-780.
- Barrick, J.E., M.R. Kauth, C.C. Streliaoff, and R.E. Lenski. 2010. Escherichia coli rpoB Mutants Have Increased Evolvability in Proportion to Their Fitness Defects. *Mol Biol Evol* 27:1338-1347.
- Bennett, A.F., and R.E. Lenski. 1993. Evolutionary adaptation to temperature II. Thermal niches of experimental lines of Escherichia coli. *Evolution* 1-12.
- Bennett, A.F., R.E. Lenski, and J.E. Mittler. 1992. Evolutionary adaptation to temperature. I. Fitness responses of Escherichia coli to changes in its thermal environment. *Evolution* 16-30.
- Bennett, A.F., and R.E. Lenski. 2007. An experimental test of evolutionary trade-offs during temperature adaptation. *Proc Natl Acad Sci USA* 104:8649-8654.
- Chou, H.H., N.F. Delaney, J.A. Draghi, and C.J. Marx. 2014. Mapping the Fitness Landscape of Gene Expression Uncovers the Cause of Antagonism and Sign Epistasis between Adaptive Mutations. *PLoS genetics* 10:e1004149
- Cooper, V.S., A.F. Bennett, and R.E. Lenski. 2001. Evolution of thermal dependence of growth rate of Escherichia coli populations during 20,000 generations in a constant environment. *Evolution* 55:889-896.
- Datsenko, K.A., and B.L. Wanner. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640-6645.
- Elena, S.F., and R.E. Lenski. 2003. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nat Rev Genet* 4:457-469.

- Futuyma, D.J., and G. Moreno. 1988. The evolution of ecological specialization. *Annu Rev Ecol Syst* 19:207-233.
- Hietpas, R.T., C. Bank, J.D. Jensen, and D.N.A. Bolon. 2013. Shifting fitness landscapes in response to altered environments. *Evolution* 67:3512-3522.
- Hoekstra, H.E., R.J. Hirschmann, R.A. Bunday, P.A. Insel, and J.P. Crossland. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313:101-104.
- Holder, K.K., and J.J. Bull. 2001. Profiles of adaptation in two similar viruses. *Genetics* 159:1393-1404.
- Hollands, K., A. Sevostiyanova, E.A. Groisman. 2014. Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator. *Proc Natl Acad Sci USA*. 111:E1999-2007
- Huey, R.B., and J.G. Kingsolver. 1989. Evolution of thermal sensitivity of ectotherm performance. *Trends Ecol Evol* 4:131-135.
- Jin D.J., W.A. Walter, C.A. Gross. 1988. Characterization of the termination phenotypes of rifampicin-resistant mutants. *J Mol Biol* 202:245-253.
- Jin D.J., R.R. Burgess, J.P. Richardson, C.A. Gross. 1992. Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proc Natl Acad Sci USA*, 89, 1453-1457.
- Kingsolver, J.G. 2009. The Well-Tempered Biologist. *Am Nat* 174:755-768.
- Knies, J.L., R. Izem, K.L. Supler, J.G. Kingsolver, and C.L. Burch. 2006. The genetic basis of thermal reaction norm evolution in lab and natural phage populations. *Plos Biology* 4:1257-1264.
- Lenski, R.E., M.R. Rose, S.C. Simpson, and S.C. Tadler. 1991. Long-term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2,000 generations. *Am Nat* 138:1315-1341.
- Levins, R. 1968. *Evolution in Changing Environments: Some Theoretical Explorations*. 2. (Princeton Univ Press, Princeton).
- Lynch, M., and W. Gabriel. 1987. Environmental tolerance. *Am Nat* 283-303.
- MacLean, R.C., G. Bell, and P.B. Rainey. 2004. The evolution of a pleiotropic fitness trade-off in *Pseudomonas fluorescens*. *Proc Natl Acad Sci USA* 101:8072-8077.

- Mongold, J.A., A.F. Bennett, and R.E. Lenski. 1999. Evolutionary adaptation to temperature. VII. Extension of the upper thermal limit of *Escherichia coli*. *Evolution* 53:386-394.
- Mori H., M. Imai, and K. Shigesada. 1989. Mutant rho factors with increased transcription termination activities. II. Identification and functional dissection of amino acid changes. *J Mol Biol* 210:39-49.
- Opalka N. et al. 2010. Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol* 8:e1000483.
- Orr, H.A. 2005. The genetic theory of adaptation: A brief history. *Nat Rev Genet* 6:119-127.
- Peichel, C.L., K.S. Nereng, K.A. Ohgi, B.L.E. Cole, P.F. Colosimo, C.A. Buerkle, D. Schluter, and D.M. Kingsley. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* 414:901-905.
- Peters J.M. et al. 2009. Rho directs widespread termination of intragenic and stable RNA transcription. *Proc Natl Acad Sci USA*. 106:15406-15411.
- Pigliucci, M. 2008. Is evolvability evolvable? *Nat Rev Genet* 9:75-82.
- Reed, R.D., R. Papa, A. Martin, H.M. Hines, B.A. Counterman, C. Pardo-Diaz, C.D. Jiggins, N.L. Chamberlain, M.R. Kronforst, and R. Chen. 2011. Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333:1137-1141.
- Reynolds, M.G. 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics* 156:1471-1481.
- Rodríguez-Verdugo, A., B.S. Gaut, and O. Tenailon. 2013. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol Biol* 13:50.
- Roff, D.A., and D.J. Fairbairn. 2007. The evolution of trade-offs: where are we? *J Evol Biol* 20:433-447.
- Rose, M.R., T.J. Nusbaum, and A.K. Chippendale. 1996. Laboratory evolution: the experimental wonderland and the Cheshire cat. Adaptation. Academic Press, San Diego, Calif 221-241.
- Ryals, J., R. Little, H. Bremer. 1982. Temperature dependence of RNA synthesis parameters in *Escherichia coli*. *J Bacteriol* 151:879-887.

- Shoval, O., H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and U. Alon. 2012. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* 336:1157-1160.
- Silva, R.F., S.C.M. Mendonça, L.M. Carvalho, A.M. Reis, I. Gordo, S. Trindade, and F. Dionisio. 2011. Pervasive sign epistasis between conjugative plasmids and drug-resistance chromosomal mutations. *PLoS genetics* 7:e1004149
- Sleight, S.C., C. Orlic, D. Schneider, R.E. Lenski. 2008. Genetic basis of evolutionary adaptation by *Escherichia coli* to stressful cycles of freezing, thawing and growth. *Genetics* 180:431-443.
- Somero, G.N. 1978. Temperature adaptation of enzymes: biological optimization through structure-function compromises. *Annu Rev Ecol Syst* 9:1-29.
- Tenaillon, O., A. Rodriguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennett, A.D. Long, and B.S. Gaut. 2012. The Molecular Diversity of Adaptive Convergence. *Science* 335:457-461.
- Trindade, S., A. Sousa, and I. Gordo. 2012a. Antibiotic resistance and stress in the light of fisher's model. *Evolution* 66:3815-3824.
- Woods, R.J., J.E. Barrick, T.F. Cooper, U. Shrestha, M.R. Kauth, and R.E. Lenski. 2011. Second-Order Selection for Evolvability in a Large *Escherichia coli* Population. *Science* 331:1433-1436.
- Zhou Y.N. *et al.* 2013. Isolation and characterization of RNA polymerase rpoB mutations that alter transcription slippage during elongation in *Escherichia coli*. *J Biol Chem* 288:2700-2710.

CONCLUSIONS

Mutations can create or modify a trait, potentially leading to increased fitness of an organism and ultimately adaptation of a population. This simplified explanation of adaptation implies that the connection between genotype, phenotype and fitness is straightforward. Unfortunately, we know that this is not the case. Adaptation is exceptionally difficult to observe and study, particularly in nature. In natural environments the selecting agents that act on a potentially adaptive trait are often unclear or complex due to interactions of biotic and abiotic factors. Moreover, even if an adaptive trait is successfully identified and associated with a selecting agent, the challenge becomes to identify the genetic bases underlying such trait (Barrett and Hoekstra 2011).

These problems can be partially overcome by studying adaptation in the laboratory. In an experimental setting the agent of selection can be chosen and controlled. In addition, the task of identifying the genetic bases of adaptation can be simplified by choosing asexual organisms with small genomes. For example, working with asexual bacteria lacking any plasmid vectors to mediate genetic exchange assures that all genetic variation arises by *de novo* mutations and not by standing variation or lateral gene transfer. In addition whole genomes sequencing in small genomes is cheap and feasible.

With this consideration in mind, we adapted 114 independent replicates of *E.coli* for 2000 generations in a glucose-limited medium at 42°C. The selective agent was high temperature, which was relatively easy to manipulate and control. By evolving the bacteria in a low resource condition, we limited the potential emergence of ecological

interactions within each population. Finally, we sequenced the complete genomes of one clone from each population to identify all the genetic mutations associated with adaptation to thermal stress for these isolates.

The first major goal of the large-scale evolution experiment was to describe the genetic diversity of mutations underlying thermal stress adaptation. In short, we uncovered more beneficial mutations than any previous *E.coli* evolution experiment. We also identified two “adaptive pathways”, which are defined by mutations in the RNAP and by mutations in the Rho termination factor (Tenailon et al. 2012). The second major goal of the large-scale evolution experiment was to associate these mutations with phenotypes. My dissertation contributed to this second major goal by linking changes in phenotypes and fitness with specific mutations.

In the first part of my dissertation, I focused on the adaptive pathway enriched with *rpoB* mutations. My first chapter studied a subset of populations that became resistant to rifampicin. In summary, I found that rifampicin resistance was caused by three mutations at codon position 572 of the *rpoB* gene, which modified the active site of RNAP. I assessed the frequency trajectory of rifampicin resistance using samples from 200 generation intervals and found that resistant mutations typically appeared and were fixed early during the evolution experiment. Furthermore, when I tested their individual effects on the ancestral background, I confirmed that the three mutations conferred high advantage in glucose-limited medium at 42°C. Finally I observed that the *rpoB* mutations had different fitness effects across three genetic backgrounds and six environments (Rodríguez-Verdugo et al. 2013).

Overall, my first chapter focused on two phenotypes associated to *rpoB* mutations: rifampicin resistance and relative fitness. Rifampicin resistance was a collateral trait of thermal adaptation, meaning that the selective pressure was not the presence of rifampicin but rather high temperature stress. Even if not selected for, rifampicin resistance is a relevant trait and has implications for the evolution of antibiotic resistance (Schenk and de Visser 2013). Another important result of this chapter was that *rpoB* mutations were, in most of the populations, the first mutation fixed during thermal stress adaptation. This conclusion set up the “conceptual framework” for my second chapter.

In my second chapter I explored the molecular mechanisms underlying the large fitness advantage conferred by first-step mutations (i.e. *rpoB* mutations in codon 572). In this chapter I focused on two phenotypes associated with *rpoB* mutations: growth and gene expression (mRNAseq). First, I observed that the acclimation to high temperature was characterized by thousands of differentially expressed genes. Most of the expression changes were down-regulation of growth related processes and up-regulation of stress genes involved in repair and metabolic adjustments to high temperature. Second, I observed that the three *rpoB* mutations changed the expression of hundred to thousands of genes when compare to the gene expression of the ancestor at 42°C. By including measurements of the ancestor at 37°C – state prior to the thermal stress – I determined that most of the expression changes caused by the mutations *I572F* and *I572L* restored the stressed expression state (acclimation state) back to an ancestral state. The exception was the mutation *I572N*, which reinforced (or exaggerated) the expression changes of the acclimation response. Lastly, I determined

the phenotypic contribution of a first-step mutation, *I572L*, compared to the phenotypic variation accumulated during an adaptive walk (i.e. high-temperature adapted clones with the *I572L* mutation). I concluded that the *I572L* mutation contributed to most of the expression changes while later mutations did not substantially changed gene expression. This chapter has two potential implications for evolution. First, it suggests that many of the gene expression changes fixed during thermal stress adaptation were adaptive and therefore were selected by Darwinian evolution. This is an interesting result given that there is an intense debate of whether the majority of changes in gene expression are fixed by natural selection or by stochastic processes (i.e. selectively neutral; (Khaitovich et al. 2004; López-Maury et al. 2008; Wang and Zhang 2011). A second implication of this chapter relates to the possibility that the mechanisms of thermal adaptation observed in our study might follow a general mechanism of adaptation and might not be specific to high-temperature adaptation (Hindre et al. 2012).

In contrast to the first part of my dissertation, which started by investigating a specific subset of mutations that were later associated with phenotypes equivalent to a bottom-up approach; (Barrett and Hoekstra 2011), the second part of my dissertation started by screening a phenotype – the magnitude of fitness trade-offs across a thermal gradient – in 114 high-temperature adapted clones. The presence or absence of trade-offs at low temperature was later associated with specific mutations. In summary, in my third chapter I measured the relative fitness of 114 high-temperature adapted clones at 37°C and at 20°C, and characterized the thermal niche of each clone. Although trade-offs were uncommon at 37°C, I observed that more than a half of the clones performed

worse than the ancestor at 20°C. I also observed that half of the clones shifted their thermal niche while the remaining expanded their thermal niche. When I investigated the genetic bases of these two niche dynamics, I detected that mutations in the *rpoB* gene were associated with fitness trade-offs at 18°C. In contrast, mutations in the *rho* gene were associated with survival at 18°C. These results are significant because trade-offs are a basic assumption in evolutionary models, yet, demonstrating their prevalence has been difficult. In this chapter I showed that the presence or absence of trade-offs was associated in part with two different adaptive genetic pathways to thermal stress adaptation. Therefore I provided a possible explanation of why trade-offs are not universal.

Overall, my dissertation revealed that even in a “simple” setting, connecting genotype to phenotype is not trivial. First, I observed that the diversity of phenotypes resulting from apparently similar mutations is higher than expected. For example, I observed that a point mutation, *I572N*, had different phenotypic effects than other two mutations, *I572F* and *I572L*, located at the same codon position (Chapters 1 and 2). This observation suggests that the diversity of phenotypes resulting from *rpoB* mutations can be higher than expected. Second, I observed that background effects complicate the task of connecting genotypes to phenotypes. This was the case of the *rpoB I966S* mutation, which was associated with fitness trade-offs at low temperatures. Nevertheless, our association analyses suggested that in certain cases the trade-off might have been compensated by a second single mutation, *rpoB T539P* (Chapter 3).

Furthermore, the task of connecting genotypes to phenotypes was complicated by pleiotropy and epistasis. Pleiotropy occurs when a single mutation affects more than one phenotypic trait, and epistasis when two or more mutations at different loci interact in a non-additive manner such that their combined effect on a phenotype deviates from the sum of their individual effects (Orr 2005). Several observations lead me to conclude that both pleiotropy and epistasis were prevalent in our system.

Pleiotropy was evident at different phenotypic “levels”. First, pleiotropy was evident at the level of gene expression given that a single *rpoB* mutation changed the gene expression of hundreds to thousands of genes (Chapter 2). Pleiotropy was also evident at the level of growth in a non-selective environment. We observed that *rpoB* mutations were only beneficial in the conditions of our experiment (high temperature and low glucose). In all other environments, the mutations were neutral or deleterious (Chapter 1). This observation highlights the importance of being cautious when attempting to make general conclusions about the fitness consequences of particular mutations.

Concerning epistatic interactions, my results suggested at least two cases of sign epistasis. First, we observed that the *rpoB* mutations in codon 572 were advantageous in the REL1206 and REL606 background but deleterious in the K12 background (Chapter 1). Second, we observed that the *rho 115N* mutation was neutral in the condition of our experiment. This was surprising because it was observed in 14 high-temperature adapted clones. This high level of parallelism would have suggested that it conferred high advantage to the conditions of our experiment (such as the *rpoB* mutations in codon 572). We hypothesized that the *rho 115N* mutation might not be

beneficial by itself, but rather beneficial only in the presence of other interacting mutations. In conclusion, both pleiotropy and epistasis were prevalent features of our system and complicated the task of connecting genotype to phenotype.

Despite these complications, the findings from my dissertation research have opened a number of new research questions amenable to future experimental investigation. For example, given that we observed signals of pervasive epistasis and background effects, the next step would be to introduce single and double mutations into genes that are potentially in epistasis. There are at least three combinations of mutations that might be relevant: *i) rpoB I966S* and *rho I15N*, that are in complete repulsion (Tenaillon et al. 2012), might have negative epistatic interactions; *ii) rho I15N* and *cls* might have sign epistatic interactions (Rodríguez-Verdugo et al. 2014); and *iii) rpoB T539P* might compensate the deleterious side effects of *rpoB I966S* at low temperatures. Constructing these single and double mutants is a reasonable task, given that we implemented a method to construct double mutants using simultaneously two separate oligos (see Material and Methods in Chapter 3).

An interesting observation that requires further investigation was the occasional sudden recovery of populations whose densities had initially declined markedly at high temperature (i.e. “Lazarus effect”, Chapter 3). The frequency of these phenomenon and the molecular causes of the Lazarus effect are not fully understood. Does the frequency and speed of recovery increase with increasing temperature? Is the Lazarus effect caused by changes in fitness effects of mutations with increasing selective pressure or by an increase in mutation rate at high temperatures? Is the Lazarus effect only observed at high temperatures or it also occurs in other stressful conditions, such as

low temperatures, high salinity, low or high pH? More broadly, these questions relate to the study of evolvability and the limits of adaptation (Pigliucci 2008). How “far” can we adapt *E.coli*? For example, if we set up an experiment in which we gradually raise the temperature, instead of keeping a constant temperature, how far can the upper thermal limit be expanded? Does the magnitude of the upper thermal niche expansions vary between species, such that some species are more evolvable than other? Answering these questions will have lasting impacts on evolutionary theory and could help to predict the faith of organisms under global climate change.

REFERENCES

- Barrett, R.D.H., and H.E. Hoekstra. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet* 12:767-780.
- Hindre, T., C. Knibbe, G. Beslon, and D. Schneider. 2012. New insights into bacterial adaptation through in vivo and in silico experimental evolution. *Nat Rev Microbiol* 10:352-365.
- Khaitovich, P., G. Weiss, M. Lachmann, I. Hellmann, W. Enard, B. Muetzel, U. Wirkner, W. Ansorge, and S. Pääbo. 2004. A neutral model of transcriptome evolution. *PLoS biology* 2:e132.
- López-Maury, L., S. Marguerat, and J. Bähler. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet* 9:583-593.
- Orr, H.A. 2005. The genetic theory of adaptation: A brief history. *Nat Rev Genet* 6:119-127.
- Pigliucci, M. 2008. Is evolvability evolvable? *Nat Rev Genet* 9:75-82.
- Rodríguez-Verdugo, A., D. Carrillo-Cisneros, A. Gonzalez-Gonzalez, B.S. Gaut, and A.F. Bennett. 2014. Different trade-offs result from alternate genetic adaptations to a common environment. *Proc Natl Acad Sci USA* 111:12121-6.
- Rodríguez-Verdugo, A., B.S. Gaut, and O. Tenaillon. 2013. Evolution of Escherichia coli rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol Biol* 13:50.
- Schenk, M.F., and J.A.G.M. de Visser. 2013. Predicting the evolution of antibiotic resistance. *BMC biology* 11:14.
- Tenaillon, O., A. Rodriguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennett, A.D. Long, and B.S. Gaut. 2012. The Molecular Diversity of Adaptive Convergence. *Science* 335:457-461.
- Wang, Z., and J. Zhang. 2011. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc Natl Acad Sci USA* 108:E67-E76.