

INCREASED GENE DOSAGE PLAYS A PREDOMINANT ROLE IN THE INITIAL STAGES OF EVOLUTION OF DUPLICATE TEM-1 BETA LACTAMASE GENES

Riddhiman Dhar,^{1,2,3} Tobias Bergmiller,^{4,5} and Andreas Wagner^{1,2,6,7}

¹*Institute of Evolutionary Biology and Environmental Studies, University of Zurich, CH-8057 Zurich, Switzerland*

²*The Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland*

³*Current address: Centre for Genomic Regulation (CRG), C/Dr. Aiguader 88, 08003 Barcelona, Spain*

⁴*ETH Zurich and Eawag, CH-8600 Dübendorf, Switzerland*

⁵*Current address: Institute of Science and Technology, Am Campus 1, 3400 Klosterneuburg, Austria*

⁶*The Santa Fe Institute, Santa Fe, New Mexico 87501*

⁷*E-mail: andreas.wagner@ieu.uzh.ch*

Received August 8, 2013

Accepted January 22, 2014

Gene duplication is important in evolution, because it provides new raw material for evolutionary adaptations. Several existing hypotheses about the causes of duplicate retention and diversification differ in their emphasis on gene dosage, subfunctionalization, and neofunctionalization. Little experimental data exist on the relative importance of gene expression changes and changes in coding regions for the evolution of duplicate genes. Furthermore, we do not know how strongly the environment could affect this importance. To address these questions, we performed evolution experiments with the TEM-1 beta lactamase gene in *Escherichia coli* to study the initial stages of duplicate gene evolution in the laboratory. We mimicked tandem duplication by inserting two copies of the TEM-1 gene on the same plasmid. We then subjected these copies to repeated cycles of mutagenesis and selection in various environments that contained antibiotics in different combinations and concentrations. Our experiments showed that gene dosage is the most important factor in the initial stages of duplicate gene evolution, and overshadows the importance of point mutations in the coding region.

KEY WORDS: Beta lactamase, duplicate genes, experimental evolution, gene dosage.

Gene duplication is one of the fundamental driving forces of molecular evolution. Although studied for a long time (Kuwada 1911; Bridges 1936; Serebrovsky 1938; Stephens 1951; reviewed in Taylor and Raes 2004), only the genomic era revealed its full extent, showing that about 20–60% of a genome's genes can have duplicates (Himmelreich et al. 1996; Klenk et al. 1997; Tomb et al. 1997; Arabidopsis Genome Initiative, 2000; Rubin et al. 2000; Li et al. 2001; Kellis et al. 2004). Gene duplication, which occurs approximately as frequently as mutation (Lynch and Conery 2000), is important for the evolution of new protein

functions (Ohno 1970; Hughes 1994; Walsh 1995), and may also facilitate speciation (Sidow 1996; Spring 1997; Lynch and Force 2000a; Bailey et al. 2002; Simillion et al. 2002; Bowers et al. 2003; Dehal and Boore 2005; Marques-Bonet et al. 2009). Most duplicate gene copies are lost by deletion or become pseudogenes (Rouquier et al. 1998, 2000; Balakirev and Ayala 2003), yet many duplicate genes have persisted in the genomes of organisms in all branches of life (Tekaia and Dujon 1999; Zhang 2003; Vavouri et al. 2008), further underscoring the importance of gene duplication.



Four main classes of hypotheses exist to explain the retention of duplicate genes (reviewed in Hahn 2009). The first hypothesis highlights the role of gene dosage itself. Increased expression level of a gene, due to duplication, can be beneficial and help retain both duplicates (Horiuchi et al. 1963; Kondrashov et al. 2002; Conant and Wolfe 2007). Examples include old duplicate genes in yeast metabolism that still show substantial functional overlap, suggesting that higher dosage is likely beneficial (Conant and Wolfe 2007; DeLuna et al. 2008). Gene duplication, a special case of gene amplification to high copy numbers, may also be beneficial in nutrient-limited environments, for adaptation to stressors, and for resistance to antibiotics, presumably due to increased gene dosage (Straus 1975; Otto et al. 1986; Matthews and Stewart 1988; Nichols and Guay 1989; Sonti and Roth 1989; Ives and Bott 1990; Kondratyeva et al. 1995; Brown et al. 1998; Riehle et al. 2001; Dunham et al. 2002; Reams and Neidle 2003; Toprak et al. 2011). Moreover, gene duplications may accelerate evolution in new or stressful environments by providing more genetic material for mutation, a hypothesis referred to as the amplification-mutagenesis hypothesis (Andersson et al. 1998; Hendrickson et al. 2002; Roth and Andersson 2004).

The second hypothesis for duplicate gene retention emphasizes the evolution of new functions—neofunctionalization (Ohno 1970; Hughes 1994; Walsh 1995; Zhang et al. 2002; Hooper and Berg 2003; Rodríguez-Trelles et al. 2003; Francino 2005; Hughes 2005; Hughes and Liberles 2007; Conant and Wolfe 2008). One copy of a duplicate gene pair experiences relaxed selection and can evolve a beneficial new function (Ohno 1970) that can help retain both gene copies. Several models for neofunctionalization such as the innovation–amplification–divergence (IAD; Bergthorsson et al. 2007; Soo et al. 2011; Näsvall et al. 2012), and the escape from adaptive conflict (EAC) model (Des Marais and Rausher 2008; Barkman and Zhang 2009; Deng et al. 2010) have been proposed.

The third hypothesis for duplicate retention suggests subfunctionalization as a possible mechanism. Force et al. (1999) proposed that two copies of a gene acquire mutations that change the genes' activity in such a way that the two copies become complementary to each other in function. In this duplication–degeneration–complementation (DDC) model, both copies are required to maintain the original function (Force et al. 1999; Lynch and Force 2000b; Lynch et al. 2001). Subfunctionalization may also allow retention of duplicate genes that could later evolve new functions (Rastogi and Liberles 2005).

A fourth candidate mechanism for duplicate gene retention involves novel sequences that are created at the junctions between new duplicates. Such novel sequences can themselves be beneficial for cells and help spread gene duplicates through a population (Glansdorff and Sand 1968; Jackson and Yanofsky 1973; Ahmed 1975; Anderson and Roth 1977, 1978a,b; Kugelberg et al. 2006,

2010). In addition to these main classes of hypotheses, theoretical studies have suggested various further scenarios for gene duplicate retention (Nowak et al. 1997; Wagner 2000a, 2002; Gu et al. 2003; Kafri et al. 2006, 2008).

Changes in gene dosage (gene expression) and in gene coding regions play a role in the first three hypotheses. Although comparative studies address the importance of gene dosage changes or coding sequence changes for the retention of duplicate genes (Kondrashov et al. 2002; Zhang et al. 2002; Tocchini-Valentini et al. 2005; Conant and Wolfe 2007; Lynch 2007; Tirosh and Barkai 2007; Woolfe and Elgar 2007; Des Marais and Rausher 2008; Kleinjan et al. 2008), very little pertinent experimental evidence exist (Holloway et al. 2007; Näsvall et al. 2012). Moreover, we do not know how environmental conditions can influence the retention of duplicate genes via one mechanism or the other. Although some experimental work studied the evolution of duplicate genes (Holloway et al. 2007; Näsvall et al. 2012), none of them examined the relative importance of these mechanisms in the retention process, nor did they study the role of the environment in the outcome.

To study the relative importance of changes in gene dosage or expression and in protein coding regions, we designed an experimental system in which a plasmid hosts a tandem duplication of the TEM-1 beta lactamase gene. In this system, TEM-1 gene dosage increase can occur through higher order amplification of the gene or through increased copy number of the plasmid. In addition, duplicate TEM-1 genes are free to accumulate coding mutations that change the activity of the protein and regulatory mutations that can tune the expression level of the TEM-1 gene. We subjected the two duplicates to repeated rounds of mutagenesis and selection on the native substrate ampicillin, on novel substrates, and on a combination of native and novel substrates, to ask what role the chemical environment plays in the retention of duplicate genes. We found that increased gene dosage plays the most important role in retaining duplicate genes, regardless of the antibiotic environment. Our results also suggest that elevated gene dosage is affected by four molecular processes that could alter TEM-1 gene expression in our evolved populations.

Materials and Methods

BACTERIAL STRAINS AND MEDIUM

We used the *E. coli* strain DH5 α for initial fitness measurements, for evolution experiment, and for the fitness measurements after the evolution experiment. Because the frequently used mutagenic polymerase chain reaction is highly recombinogenic, we used a mutator *E. coli* strain, TB90 for introducing genetic variation in TEM-1 genes (Methods S1). In TB90, mutagenesis can be induced with the addition of L-(+)-arabinose (Sigma; Methods S1; Table S6). The TB90 strain lacks the *recA* gene,

to reduce recombination and gene loss among duplicate TEM-1 copies as well as to minimize the incidence of higher order amplifications (Reams et al. 2010), which have already been studied (Wiebauer et al. 1981; Goldberg and Mekalanos 1986; Reams et al. 2012). For all *E. coli* cultures, we used LB medium (Becton-Dickinson). All *E. coli* strains were grown at 37°C except DY330 or strains harboring the plasmid pCP20, which were cultured at 30°C.

CONSTRUCTION OF DUPLICATE COPY

We cloned the gene TEM-1 beta lactamase along with its promoter from the pBR322 plasmid (Sutcliffe 1978) into the pUA66 plasmid (Zaslaver et al. 2006), using *Xho*I and *Bam*HI restriction sites. We transformed this plasmid into *E. coli* DH5 α cells, referring to the resulting strain as the ancestral S (AncS) strain, where S stands for “single copy” (Fig. 1A). We then inserted a second, identical copy of the TEM-1 gene into the same plasmid using the *Bam*HI restriction site, and transformed the resulting construct into *E. coli* DH5 α cells. We refer to this strain as the ancestral D (AncD) strain (Fig. 1A). We confirmed the sequences of the inserts and the orientation of the inserts using Sanger sequencing. Both genes had the same transcriptional orientation and independent transcription start sites. We performed initial fitness measurements on these strains. We then amplified the inserts of the AncS and AncD strains using PCR and recloned them into plasmid pUA67, a modified version of pUA66 (with more restriction enzyme sites for cloning, but otherwise identical to pUA66; Fig. S1) using *Eco*RI and *Hind*III (NEB) restriction sites. We chose 10 clones for each of the two strains as starting populations for our evolution experiment (Fig. S2). Because the pUA66 plasmid contains the kanamycin resistance gene, we used kanamycin in all growth media as a selective agent.

INITIAL FITNESS MEASUREMENTS

We measured fitness of the ancestral strains (AncS and AncD) at multiple concentrations of each of the following antibiotics: ampicillin (Sigma), the native substrate for the TEM-1 beta lactamase enzyme; the novel substrates cefotaxime (Sigma) and imipenem (Sigma); and combinations of native and novel substrates, that is, ampicillin with cefotaxime, and ampicillin with imipenem. We used 25 μ g/mL kanamycin in each growth condition for selection of the plasmid marker. For fitness measurements, we cultured the ancestral strains overnight (for 16 h) at 37°C in LB supplemented with 25 μ g/mL kanamycin. We measured the optical density of cells using a plate reader (Tecan infinite F200 pro) at 600 nm (OD(0) at time $t = 0$), diluted the cultures 100-fold into medium with appropriate antibiotics (total volume 2 mL) and let the cells grow for 7 further hours in a 96-well plate (Sarstedt, MegaBlock 96 well plate) in a plate shaker (VWR, Incubating Microplate Shaker) at 600 rpm. After 7 h, we again measured the optical

density of the cells (OD(7) at $t = 7$ h). We estimated fitness (f) by calculating the fold-change in cell density given by

$$f = \frac{\text{OD}(7)}{(\text{OD}(0)/100)}.$$

We used fold-change in cell density in 7 h of growth as a proxy for fitness for two reasons. First, because there are three main components of fitness in our experiment (Warringer et al. 2011), a fitness increase in our evolved lines could occur through changes in each (or any combination) of these three components: a reduction in lag phase duration, an increase in the maximal growth rate, or an increase in carrying capacity. Thus, it could be highly misleading to just consider one of these components for fitness analysis and hence, we decided on a measure that is affected by all the three components of fitness, namely cell density after 7 h of growth. Second, we could have measured fitness by performing direct competition experiments. However, results from such assays would be difficult to interpret, because the higher gene dosage in cells carrying duplicate genes can accelerate hydrolysis of antibiotics, which could benefit slower growing single-copy cells in a mixed culture.

EVOLUTION EXPERIMENT

Our evolution experiment had six steps (Fig. 2A; Methods S2). First, we transformed the plasmids containing single- or double-copy TEM-1 genes into mutator *E. coli* TB90 cells with a chemical transformation protocol (Hanahan 1983; Swords 2003), and grew cells for 16 h without inducing mutagenesis. Second, we induced mutagenesis with L-(+)-arabinose and allowed cells to grow for 20 more hours. In both these steps we used LB media (BD Biosciences) supplemented with 50 μ g/mL kanamycin to retain plasmids in the *E. coli* TB90 cells. Third, we isolated mutagenized plasmids from the TB90 cells. Fourth, we transformed these mutagenized plasmids into *E. coli* DH5 α strain using chemical transformation, and let the cells grow for 16 h in medium with 25 μ g/mL kanamycin. Fifth, we diluted cells 10-fold into appropriate selection medium (see next section), and allowed them to grow for an additional 7 h. Sixth, we isolated plasmids from *E. coli* cells after selection. We used a part of these isolated plasmids for transformation of the mutator TB90 cells for the next round of mutagenesis. We repeated these six steps 12 times. At each step of the experiment, we ensured that population sizes remained above 10^5 cells, in an effort to minimize the role of genetic drift. We transformed the plasmids isolated at the end of each round of selection experiment into *E. coli* DH5 α cells, let them grow for 16 h in medium with 25 μ g/mL kanamycin, and stored them at -80°C after the addition of 15% glycerol (Sigma; final concentration, w/v) for future use. We emphasize that kanamycin was present throughout our experiments, from the mutagenesis stage to the

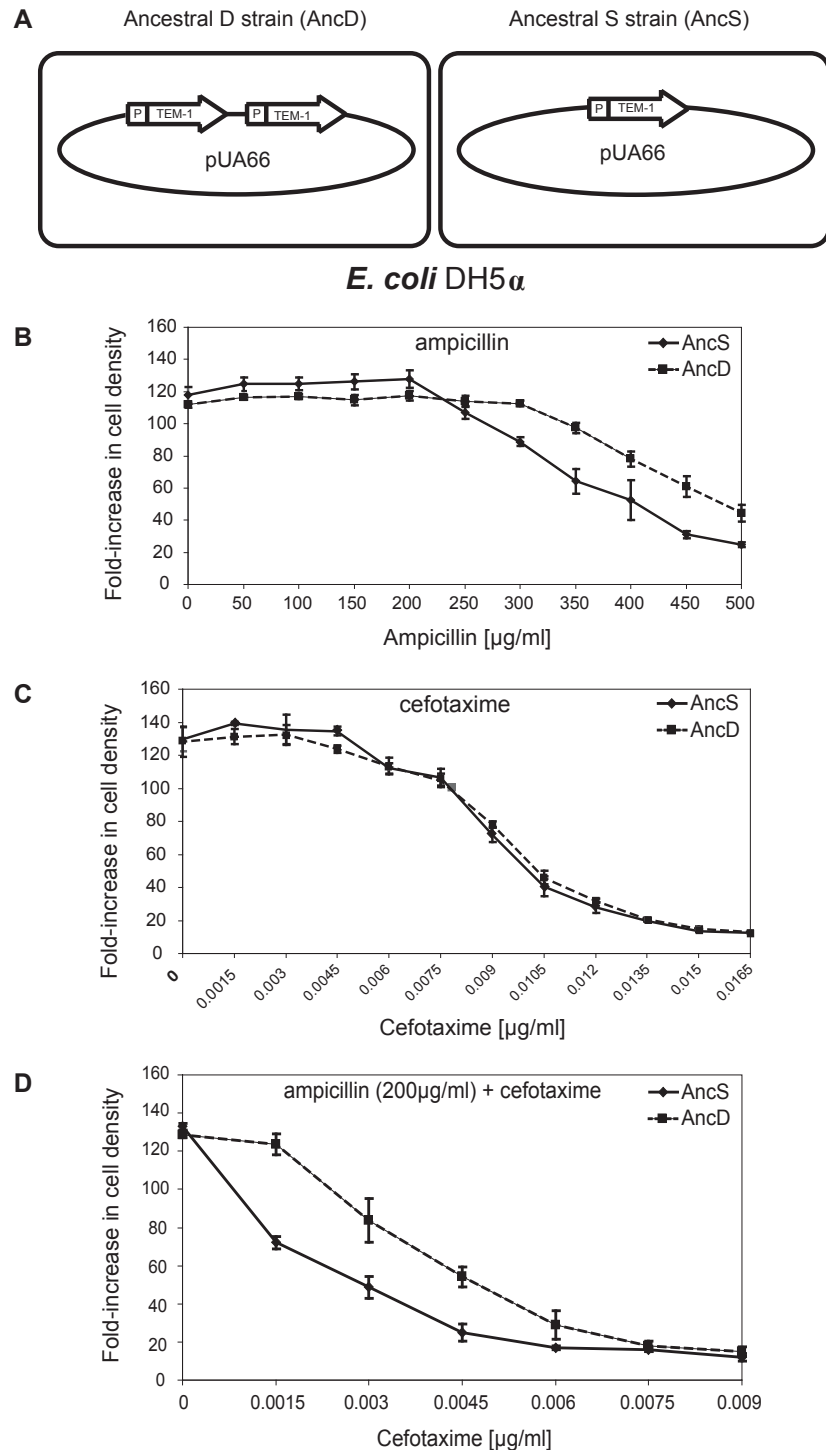


Figure 1. (A) Ancestral strains of our experiments. Our starting strains AncS and AncD are derivatives of *E. coli* DH5 α cells that contain one (AncS) or two (AncD) copies of the TEM-1 gene on plasmid pUA66 (Zaslaver et al. 2006). "P" refers to the promoter of the TEM-1 gene and "TEM-1" refers to the coding region. (B–D) Gene duplication confers a significant fitness advantage on some antibiotic concentrations (AncS and AncD). To measure fitness, we cultured cells from the ancestral strains for 7 h in medium supplemented with antibiotics, and calculated the fold-change in cell density (y-axis) on different concentrations (x-axis) of (B) ampicillin, (C) cefotaxime, and (D) cefotaxime with a constant 200 μ g/mL concentration of ampicillin. The dotted line shows the data for the AncD strain, and the solid line shows the data for the AncS strain. Note a general decrease in fitness as antibiotic concentrations increase. On ampicillin (panel a), the AncS strain declined significantly more in fitness than the AncD strain (ANOVA, Methods S9; $P = 0.03$). No such advantage of the AncD strain existed on cefotaxime (ANOVA, $P = 0.90$), but on cefotaxime supplemented with ampicillin (200 μ g/mL), this advantage became more pronounced than on ampicillin alone (ANOVA; $P = 0.0009$). Figure S4 contains fitness data on further antibiotic concentrations.

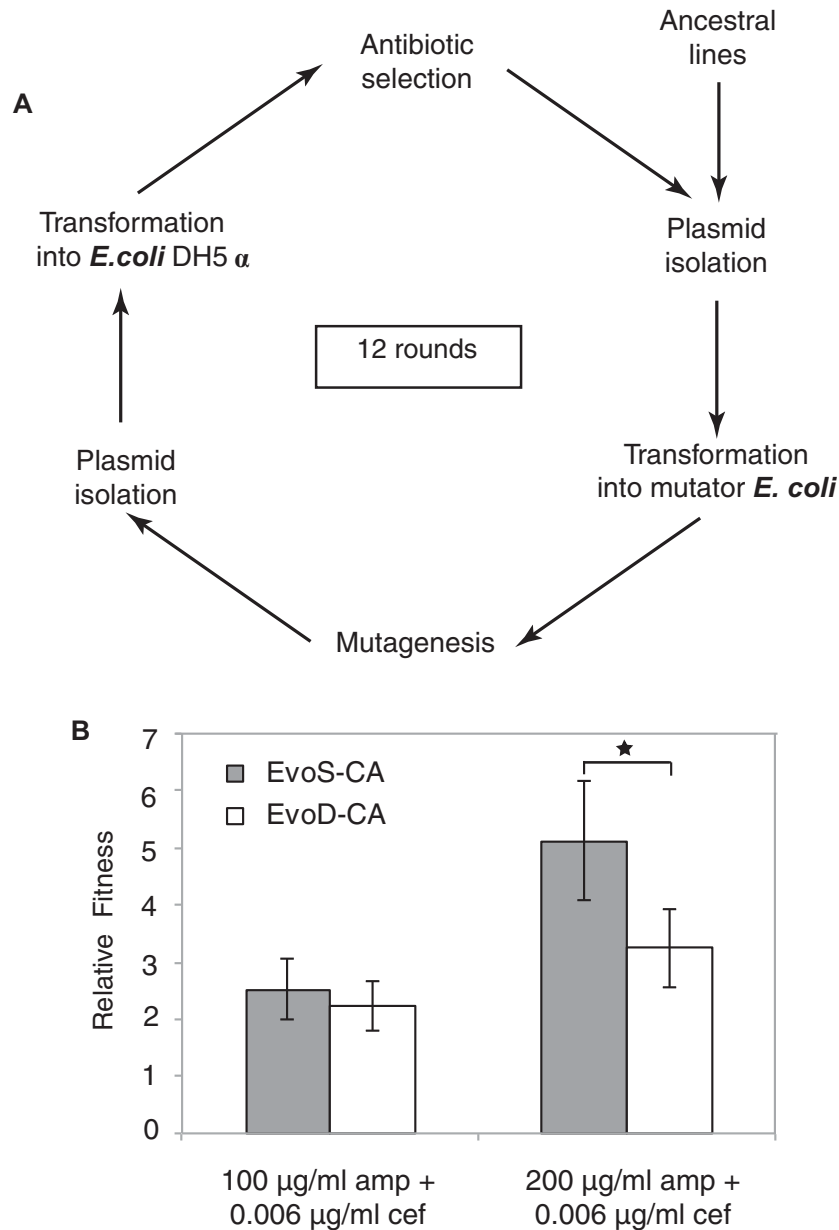


Figure 2. (A) Experimental evolution design. Our evolution experiment had six steps (see Methods). Briefly, we transformed the plasmids containing single- or double-copy TEM-1 inserts into the mutator *E. coli* TB90 cells. We then isolated the mutagenized plasmids and transformed them into the *E. coli* DH5 α strain. We then let these cells grow for 16 h in LB medium supplemented with 25 $\mu\text{g}/\text{mL}$ of kanamycin. In the final step, we diluted the cells 10-fold into appropriate antibiotic selection medium, let them grow in the selection medium for 7 h and subsequently isolated the plasmids. We repeated these steps 12 times, thus subjecting the ancestral strains to 12 rounds of mutagenesis and selection. We refer to the populations thus evolved as the EvoS and EvoD lines. We emphasize that only the plasmids evolved in our experiment and not the cells, because we changed the host cells at each step of the experiment. (B) Laboratory evolution significantly increases fitness. We measured the fitness of the CA lines (evolved on ampicillin and cefotaxime) relative to the AncD strain as a relative fold-change in cell density after 7 h of growth (see Methods). A relative fitness of 1 would mean that an evolved line and the ancestral strain have identical fitness. We measured fitness of the CA lines in 100 $\mu\text{g}/\text{mL}$ ampicillin + 0.006 $\mu\text{g}/\text{mL}$ cefotaxime and in 200 $\mu\text{g}/\text{mL}$ ampicillin + 0.006 $\mu\text{g}/\text{mL}$ cefotaxime. Both the EvoS-CA lines and EvoD-CA lines show a significant fitness increase relative to the AncD strain ($r_f = 2.52 \pm 0.53$ for EvoS-CA lines; $r_f = 2.22 \pm 0.44$ for EvoD-CA lines in 100 $\mu\text{g}/\text{mL}$ ampicillin + 0.006 $\mu\text{g}/\text{mL}$ cefotaxime; $r_f = 5.12 \pm 1.05$ for EvoS-CA lines; $r_f = 3.25 \pm 0.69$ for EvoD-CA lines in 200 $\mu\text{g}/\text{mL}$ ampicillin + 0.006 $\mu\text{g}/\text{mL}$ cefotaxime). Note that the fitness increase in the EvoS-CA lines was significantly higher than that in the EvoD-CA lines in 200 $\mu\text{g}/\text{mL}$ ampicillin + 0.006 $\mu\text{g}/\text{mL}$ cefotaxime ($P = 1.2 \times 10^{-5}$, Wilcoxon rank sum test). Bars show mean relative fitness values (averaged over measurements in triplicate for each of the replicate lines) whiskers show one standard deviation (SD), and asterisks indicate a significant fitness difference.

selection stage, to ensure continual maintenance of the plasmid in our populations.

SELECTION CONDITIONS

We used several antibiotic conditions for our selection experiment (Figs. S3, S24; Methods S3). For each condition, we evolved five replicate populations starting from the clones derived from the AncS strain, and five replicate populations starting from clones derived from the AncD strain. In general, we refer to these evolved lines as EvoS and EvoD lines, but we additionally labeled them with the selection conditions in which they evolved (Methods S3). For all selection conditions, we supplemented the medium with 25 $\mu\text{g}/\text{mL}$ kanamycin (for plasmid maintenance) in addition to any other antibiotics we used. We evolved “neutral” lines in medium supplemented only with 25 $\mu\text{g}/\text{mL}$ kanamycin.

FITNESS MEASUREMENT OF THE EVOLVED LINES

We calculated the fitness of the evolved lines relative to the fitness of the AncD strain for reasons explained in Methods S4. We grew cells of the AncD strain and the evolved lines (populations) from the glycerol stocks in LB medium with 25 $\mu\text{g}/\text{mL}$ kanamycin for 16 h and diluted cells 100 times into fresh LB medium with 25 $\mu\text{g}/\text{mL}$ kanamycin and appropriate antibiotics. We grew the cells for 7 h, measured optical density at 600 nm using a plate reader (TECAN), and calculated the relative fitness as follows. For an evolved line “Evo,”

$$\text{relative fitness (rf)} = \frac{\frac{\text{OD}_{\text{Evo}}(7)}{(\text{OD}_{\text{Evo}}(0)/100)}}{\frac{\text{OD}_{\text{AncD}}(7)}{(\text{OD}_{\text{AncD}}(0)/100)}} = \frac{\text{OD}_{\text{Evo}}(7)}{\text{OD}_{\text{AncD}}(7)},$$

where $\text{OD}_{\text{AncD}}(0)$ is the cell density of the AncD strain at $t = 0$; $\text{OD}_{\text{AncD}}(7)$ is the cell density of the AncD strain after 7 h; $\text{OD}_{\text{Evo}}(0)$ is the cell density of the evolved line at $t = 0$; and $\text{OD}_{\text{Evo}}(7)$ is the cell density of the evolved line after 7 h.

MEASUREMENT OF PLASMID COPY NUMBER CHANGE

To estimate the change in plasmid copy number during the evolution experiments, we isolated plasmids from population samples of the ancestral strain and of the evolved lines. We cultured cells from glycerol stocks in 5 mL LB medium with 25 $\mu\text{g}/\text{mL}$ of kanamycin for 16 h. Then, we estimated cell density by measuring optical density of the culture at 600 nm using a plate reader (TECAN) and performed plasmid isolation using the ChargeSwitch-Pro plasmid miniprep kit (Invitrogen). We separated plasmid DNA on a 0.8% agarose gel, and estimated its concentration with the Genetools software (Syngene; Methods S5).

QUANTIFICATION OF GENE LOSS

To estimate the extent of gene loss in the EvoD lines, we employed a simple PCR-based screen. Because of our plasmid design, individuals that contain duplicate copies of the TEM-1 gene have a unique sequence in the region where the two copies are joined. We designed a primer pair (Table S1) where one primer bound only to this unique region. Together, the two primers can amplify a 1 kbp long PCR product that occurs only in individuals with two TEM-1 copies. To perform this screen, we streaked out evolved lines onto LB-agar plates to isolate clones, performed colony PCR on these clones, and determined the number of clones with duplicate copies of the gene. We screened at least 20 clones in each EvoD line.

HIGH-THROUGHPUT SEQUENCING

To study mutations in evolved lines, we sequenced population samples of all evolved lines after 1, 4, 8, and 12 rounds of mutagenesis and selection, using high-throughput sequencing on the Roche 454 platform (Margulies et al. 2005). In addition, we sequenced samples from the neutral lines (EvoS-N and EvoD-N) after two and three rounds of mutagenesis and selection. We sequenced population samples instead of clones to obtain information about the spectrum and frequency of mutations in the population. For the EvoD lines, we only sequenced individuals that had retained two copies of the TEM-1 gene. The experimental details and the methods for the analysis of sequencing data are described in the Supporting Information (Methods S6, S7, S8, S14; Figs. S25, S29–S32; Table S7). We note that for several reasons ratios of nonsynonymous to synonymous substitutions are not suitable for inferring the strength of selection on the TEM-1 genes in our experiments (Methods S13; Figs. S27, S28; Results S8).

Results

DUPLICATION CONFERS A FITNESS ADVANTAGE

We first wanted to know if the two ancestral strains AncS and AncD, containing single and duplicate copies of the TEM-1 gene, differ in fitness. To this end, we first measured the fitness of these strains in medium supplemented with ampicillin. Fitness decreased in both ancestral strains as we increased the ampicillin concentration in the medium. However, this fitness reduction became smaller in the AncD strain than in the AncS strain at increasing ampicillin concentrations (Fig. 1B, $P = 0.03$, ANOVA—Methods S9). The advantage of the double-copy strain was especially large at high concentration of ampicillin (>250 $\mu\text{g}/\text{mL}$, Fig. 1B; Table S2a). This was not surprising, because the AncD strain has a higher dosage of the gene encoding TEM-1 beta lactamase, and we would thus expect it to hydrolyze its native substrate ampicillin more efficiently. This difference should

Table 1. Relative fitness values of the EvoS-CA, EvoD-CA, EvoD-CA-S, and EvoD-CA-D lines.

Line	Antibiotic concentrations ($\mu\text{g}/\text{mL}$)	Relative fitness (rf) mean \pm SD	<i>P</i> -value (Wilcoxon signed rank test)
EvoS-CA	Amp 100 + Cef 0.006	2.52 \pm 0.53	6.1×10^{-5}
EvoD-CA	Amp 100 + Cef 0.006	2.22 \pm 0.44	6.1×10^{-5}
EvoS-CA	Amp 200 + Cef 0.006	5.12 \pm 1.05	6.1×10^{-5}
EvoD-CA	Amp 200 + Cef 0.006	3.25 \pm 0.69	6.1×10^{-5}
EvoD-CA-S	Amp 100 + Cef 0.006	3.24 \pm 0.76	$<2.2 \times 10^{-16}$
EvoD-CA-D	Amp 100 + Cef 0.006	3.28 \pm 0.76	$<2.2 \times 10^{-16}$
EvoD-CA-S	Amp 100 + Cef 0.009	3.80 \pm 1.39	$<2.2 \times 10^{-16}$
EvoD-CA-D	Amp 100 + Cef 0.009	4.62 \pm 2.15	$<2.2 \times 10^{-16}$
EvoD-CA-S	Amp 200 + Cef 0.006	7.20 \pm 4.28	$<2.2 \times 10^{-16}$
EvoD-CA-D	Amp 200 + Cef 0.006	7.60 \pm 4.12	$<2.2 \times 10^{-16}$
EvoD-CA-S	Amp 200 + Cef 0.009	4.52 \pm 1.83	$<2.2 \times 10^{-16}$
EvoD-CA-D	Amp 200 + Cef 0.009	5.76 \pm 2.85	$<2.2 \times 10^{-16}$

The table shows the fitness of the evolved lines in medium supplemented with ampicillin and cefotaxime and relative to the fitness of the AncD strain (see Methods). A relative fitness significantly greater than 1 suggests a fitness increase in the evolved populations. We tested whether the relative fitness is significantly greater than 1 using a one-sample Wilcoxon signed rank test. The *P*-value resulting from this test is shown in the rightmost column.

become more important at higher, more detrimental concentrations of ampicillin, just as we observed.

In a next step, we asked whether this difference also exists in a novel substrate such as cefotaxime. We used cefotaxime because it has previously been shown that the TEM-1 gene can acquire activity toward cefotaxime by accumulating mutations (Stemmer 1994; Zacco and Gherardi 1999; Barlow and Hall 2002). As on ampicillin, increasing cefotaxime concentrations led to a fitness decrease in both ancestral strains (Fig. 1C; Table S2b). However, unlike in ampicillin, the AncD strain did not have a significant advantage over the AncS strain (ANOVA, $P = 0.90$).

We then asked whether the two antibiotics interact in reducing fitness, and whether this interaction favors the duplicate strain. We measured growth of the ancestral strains in medium supplemented with various concentrations of ampicillin and cefotaxime (Figs. 1D, S4). In low concentrations of ampicillin (50 and 100 $\mu\text{g}/\text{mL}$) supplemented with cefotaxime, the double-copy strain performed like the single-copy strain (ANOVA; $P = 0.37$ in 50 $\mu\text{g}/\text{mL}$ ampicillin supplement; $P = 0.08$ in 100 $\mu\text{g}/\text{mL}$ ampicillin supplement; Fig. S4; Table S2c). However, at higher ampicillin concentrations, the double-copy strain had higher fitness than the single-copy strain (ANOVA; $P = 0.0003$ in 150 $\mu\text{g}/\text{mL}$ ampicillin supplement; $P = 0.0009$ in 200 $\mu\text{g}/\text{mL}$ ampicillin supplement; Figs. 1D, S4; Table S2c).

Thus, we found that the AncD strain had a significant fitness advantage over the AncS strain, when faced with a combination of ampicillin and cefotaxime, even though the AncD strain did not show a fitness advantage only on cefotaxime. One candidate explanation is that the AncD strain can hydrolyze ampicillin more efficiently because it has a higher dosage of TEM-1. Thus, the AncD strain may mainly be subjected to the effect of cefotaxime,

in contrast to AncS, which is subjected to the combined effects of higher concentration of ampicillin and cefotaxime.

These observations also hold for another novel antibiotic, imipenem, and the combination of ampicillin and imipenem (Figs. S5, S6; Results S1). A different fitness measure, namely the rate at which cell density increases over time provides further evidence in support of our observations (Methods S11; Fig. S7; Results S2). Taken together, the results show that two copies of the TEM-1 gene can cause benefits due to the higher gene dosage they provide.

EXPERIMENTAL EVOLUTION

We next asked whether the inherent fitness difference we observed between the AncD and AncS strains is maintained in the long term. To do so, we subjected the ancestral strains to 12 cycles of mutagenesis and selection. We performed evolution experiments in seven selection conditions (Fig. S3) to ensure that our inferences are not specific to particular antibiotic condition, and to study the effect of different environmental selection conditions on the evolution of duplicate genes. However, we will mainly discuss results from one selection condition, that is, cefotaxime in combination with ampicillin, which gave rise to populations that we call CA lines. This selection condition is of major interest, because it has been suggested to accelerate divergence between duplicate copies, where one copy would maintain the original function while the second copy would evolve a new function (Ohno 1970; Bergthorsson et al. 2007).

EVOLVED LINES SHOW INCREASED FITNESS

At the end of the evolution experiment, we first measured the fitness of evolved populations to see whether these lines had

adapted to their respective selection conditions. For the CA lines, which had evolved in cefotaxime and ampicillin, we measured fitness on several ampicillin and cefotaxime concentrations. Both the EvoS-CA and EvoD-CA lines showed significant fitness increase in all combinations (Table 1; Figs. 2B, S8a), suggesting that these evolved lines had adapted to ampicillin and cefotaxime during the course of the experiment. However, in none of these conditions did the EvoD lines show a higher fitness increase than the EvoS lines. To the contrary, in two conditions, the EvoS lines showed a significantly higher fitness increase than the EvoD lines (Figs. 2B, S8a; Wilcoxon rank sum test; $P = 0.007$ in 200 $\mu\text{g}/\mu\text{L}$ ampicillin and 0.003 $\mu\text{g}/\mu\text{L}$ cefotaxime; $P = 1.2 \times 10^{-5}$ in 200 $\mu\text{g}/\mu\text{L}$ ampicillin and 0.006 $\mu\text{g}/\mu\text{L}$ cefotaxime). We observed similar phenomena in the other evolved lines (Table S3; Results S3; Figs. S9–S14). We note that even our “neutral lines”—they were not exposed to antibiotic except kanamycin required to maintain the plasmid—also showed an increase in fitness in ampicillin (see Results S4, S5, and Discussion).

PLASMID COPY NUMBER INCREASE IN THE EVOLVED LINES

We then turned our attention to investigate the changes at the molecular level that could lead to the increased fitness in these evolved lines, and that might explain the difference in fitness between the EvoS and the EvoD lines. Because we subjected the whole plasmid sequence to evolutionary change, the plasmid copy numbers could change in the evolved lines, and indeed they did. The ancestral plasmid contains the SC101 origin of replication and is present inside the bacterial cell in 5–10 copies (Hasunuma and Sekiguchi 1977; Armstrong et al. 1984). Plasmid copy numbers increased significantly in both the EvoS-CA and EvoD-CA lines (Fig. 3; 6.29 ± 2.08 (SD) fold in the EvoS-CA lines; 8.69 ± 1.31 (SD) fold in the EvoD-CA lines, based on measurements from five replicate lines), but did not differ significantly between these lines (Wilcoxon rank sum test, $P = 0.09$). We also observed plasmid copy number increases in all other evolved lines (Fig. S15; Results S5; Table S4). They suggest that the accompanying increase in gene dosage and the resulting increase in gene expression level (Methods S12; Fig. S16) are beneficial for the cells under all conditions employed in our experiment.

We note that the neutral lines also showed increases in plasmid copy number. There are two candidate explanations, both resulting from the possibility that our plasmids were subject to selection pressures in addition to that for an intact TEM-1 gene (see Discussion and Results S5).

We also observed high-frequency SNPs in the origin of replication of plasmids in all evolved lines where we sequenced this origin (Methods S15; Results S5). These SNPs are likely to be responsible for the increased plasmid copy number.

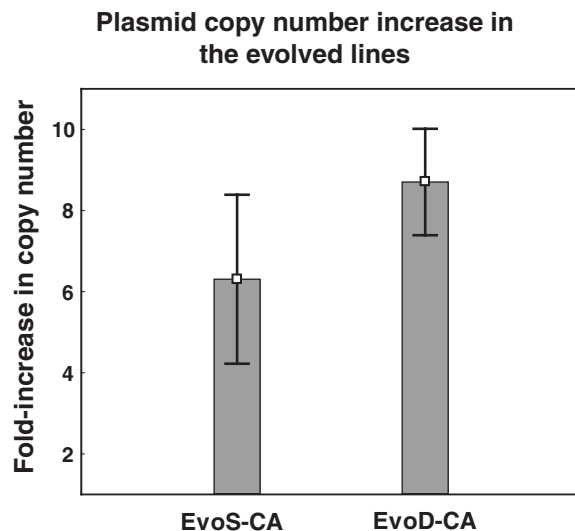


Figure 3. In CA lines plasmid copy number increased several fold. We estimated the change in plasmid copy number in the evolved lines compared to the ancestral strain (see Methods). The horizontal axis shows the names of the evolved lines in which the changes were measured, and the vertical axis shows the estimated fold-change in plasmid copy number. We observed a 6.29 ± 2.08 -fold increase in copy number in the EvoS-CA lines and a 8.69 ± 1.31 -fold increase in the EvoD-CA lines. The two lines showed no significant difference in copy number increase ($P = 0.09$). Columns represent means and whiskers indicate one standard deviation (SD).

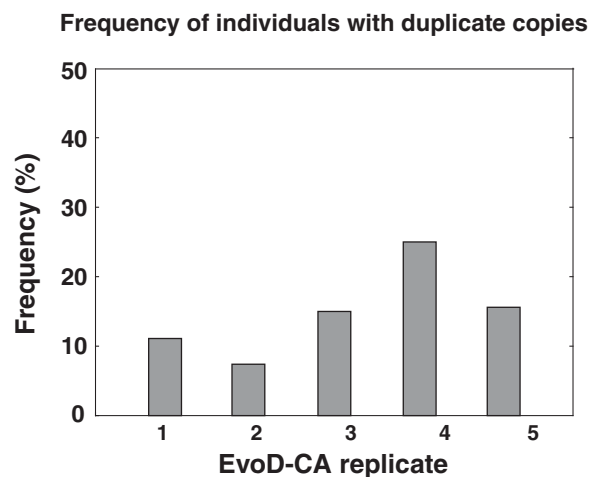


Figure 4. Some initially double-copy individuals lose one copy of the TEM1 gene in EvoD-CA lines. The vertical axis shows the percentage of individuals that still retained duplicate copies of the TEM-1 gene (see Methods), and the x-axis shows the biological replicate of the EvoD-CA lines for which we measured duplicate retention. The percentage of individuals with retained duplicates varied from 7% in the EvoD-CA2 replicate to 25% in EvoD-CA4 replicate.

Another possibility that we considered was that gene dosage could increase through higher order amplifications of the TEM-1 gene on the plasmid. However, agarose gel electrophoresis did not show any detectable higher order amplifications, neither in single-copy lines nor in lines with duplicate genes.

LOSS OF ONE COPY IN THE EvoD LINES COUNTERS INCREASE IN GENE DOSAGE

A second molecular change that we observed in our populations is a loss of one TEM-1 copy in many individuals of the double-copy (EvoD) lines. Even though we used recombination-deficient *E. coli* strains, avoided recombinogenic mutagenesis and cloning, gene loss had occurred in our EvoD lines. We estimated the extent of such loss using a PCR-based screen. It revealed that 7–25% of the individuals in the five replicate EvoD-CA lines retained duplicate copies of the gene (Fig. 4). Gene loss also occurred in all other EvoD lines (Fig. S17; Table S5). We note that such loss of tandem duplications has been observed before, even in recombinase-deficient strains (Matfield et al. 1985; Dianov et al. 1991; Lovett et al. 1993).

Gene loss would lead to reduced gene dosage in the EvoD lines, and its high incidence in the CA lines (>50% of individuals) could have two reasons. First, selection may favor individuals with a single TEM-1 copy. The relevant selection pressure may even be unrelated to selection for TEM-1 gene activity. For example, because of their smaller size, single-copy plasmids may replicate faster in host cells and may also transform more efficiently than larger plasmids (Hanahan 1983). A second candidate reason is that the loss of one gene copy from individuals with duplicate genes may simply be very frequent, such that more and more individuals lose gene copies in the course of the evolution experiment. In other words, gene copy loss could be driven by selection or by frequent recombination.

The following observations argue for the second scenario of frequent copy loss. First, a substantial fraction of clones that contained double-copy plasmids themselves harbored single-copy plasmids (Fig. S18), and such gene copy loss recurred in double-copy plasmids even after several rounds of separating clones with single-copy and double-copy plasmids from each other. Second, we measured the fitness of clones from EvoD-CA lines, and asked whether the clones with duplicate gene copies (referred to as EvoD-CA-D clones) had higher fitness than the clones that had lost a copy (EvoD-CA-S clones). Although the rapid plasmid loss we observed complicates this analysis, it can still give us some indication of the fitness difference between these two different types of clones. We measured the fitness of 33 clones each for the EvoD-CA-S and EvoD-CA-D lines in medium supplemented with ampicillin and cefotaxime (Fig. 5). The EvoD-CA-D lines showed significant higher fitness than the EvoD-CA-S lines in medium supplemented with 100 $\mu\text{g}/\text{mL}$ ampicillin and 0.009 $\mu\text{g}/\text{mL}$ ce-

fotaxime (Wilcoxon rank sum test, $P = 0.012$) and in medium with 200 $\mu\text{g}/\text{mL}$ ampicillin and 0.009 $\mu\text{g}/\text{mL}$ cefotaxime ($P = 0.003$; Fig. 5; Table 1). Taken together, all these observations suggest that the loss of one gene copy is caused by the high frequency of spontaneous loss in the EvoD lines, rather than by the selective advantage of a single TEM-1 gene copy in evolved lines.

A REGULATORY MUTATION IN THE EvoD LINES

Because we expected that point mutations in our TEM-1 genes might influence evolutionary adaptation, we sequenced the coding and near-upstream regions of TEM-1 genes in both single- and double-copy inserts using high-throughput sequencing. We were especially interested in mutations that rise to high frequency during the evolution experiment, because they indicate the action of positive selection. For the EvoD lines, we only sequenced the individuals that had retained duplicate genes. We observed several mutations at low population frequency (<5%) in all the evolved lines. However, we found only one high-frequency mutation, T-53C (T→C transition at position –53 relative to +1 at the transcription start site), which occurred only in the individuals with duplicate gene copies in the EvoD lines, and mapped to the upstream region of the gene copy on the left (Fig. 6A). This mutation reached a frequency of almost 80% in two of the five replicate EvoD-CA lines (EvoD-CA1 and EvoD-CA2; Fig. 6A) and in several other EvoD lines (Fig. S19). In addition, the mutation was present in one of the starting clones of the EvoD lines (AncD clone 9 that gave rise to EvoD-CA4; Fig. S2), but its frequency decreased in this population from 80% after round 1 to less than 1% at the end of round 12 (Fig. 6A). Similar trends existed in the other EvoD lines, except for the neutral lines, where this mutation persisted at high frequency at round 12 (Fig. S19). It is possible that this mutation was present in the starting populations of most of the EvoD lines. We cannot exclude this possibility, even though we sequenced plasmid populations at sufficiently high coverage to detect alleles with frequencies as low as 1%. In the EvoS-N3 line, the mutation appears at a low frequency (<1%) in round 4, suggesting that it might be a recurrent mutation, perhaps as a result of a mutational bias in our mutator strain.

The features of this mutation indicate how it affects fitness. First, its occurrence in the upstream region suggests a regulatory role. Second, it rose to high frequency only in the double-copy EvoD lines, which suggests that the single-copy EvoS lines do not tolerate it. Both features could be explained if the mutation caused a reduction in the expression of the TEM-1 copy next to it, a reduction that EvoD lines can tolerate, because they have a second, intact TEM-1 with normal expression. Third, the frequency of this mutation decreased with increasing antibiotic concentration beyond round 8 of the evolution

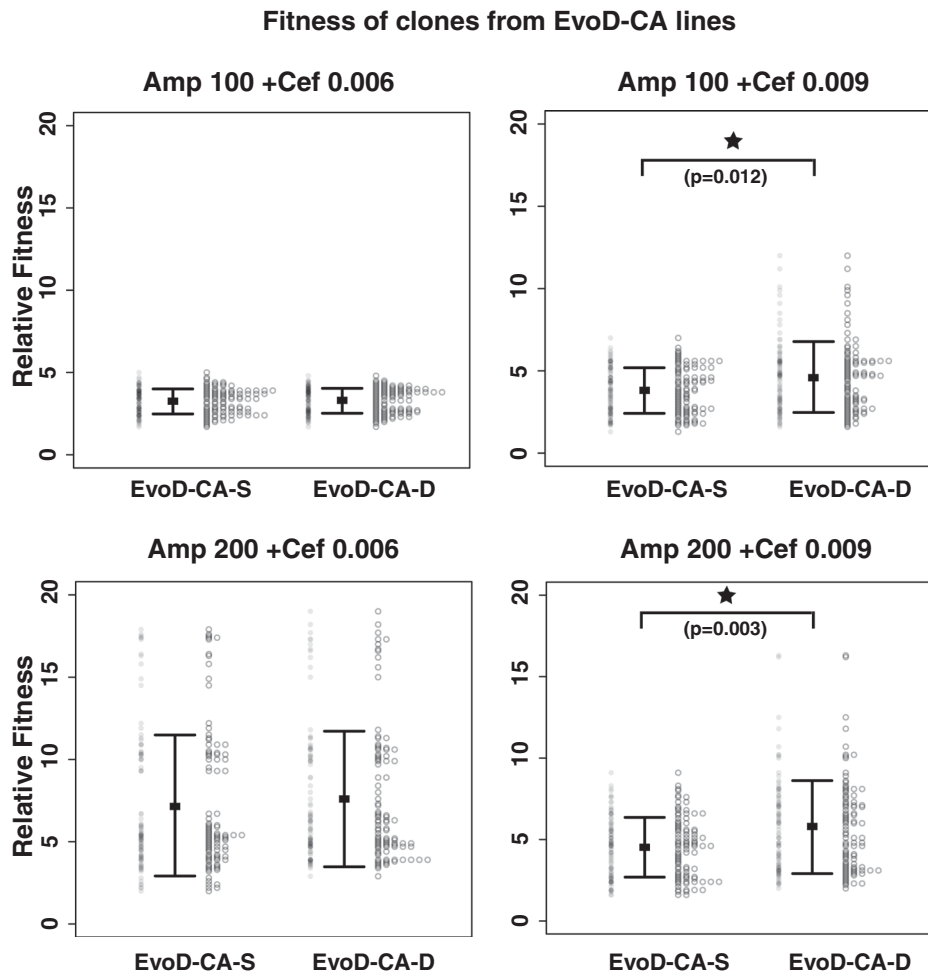


Figure 5. Fitness of the clones from the EvoD-CA lines. From the EvoD-CA lines we isolated clones that still contained duplicate genes (EvoD-CA-D) and clones that had lost one TEM-1 copy (EvoD-CA-S). We then measured fitness of these clones in various combinations of ampicillin and cefotaxime concentrations, as indicated on the horizontal axis. We performed this analysis for 33 clones containing duplicate genes and 33 clones containing a single gene copy and in triplicates. In medium containing 100 $\mu\text{g}/\text{mL}$ ampicillin and 0.009 $\mu\text{g}/\text{mL}$ cefotaxime, and in medium containing 200 $\mu\text{g}/\text{mL}$ ampicillin and 0.009 $\mu\text{g}/\text{mL}$ cefotaxime, the EvoD-CA-D and EvoD-CA-S clones showed a significant fitness increase relative to the ancestor, and this increase was significantly higher in the EvoD-CA-D clones compared to the EvoD-CA-S clones (Wilcoxon rank sum test, $P = 0.012$ and $P = 0.003$, respectively). The black rectangles represent means and whiskers indicate one standard deviation (SD). The asterisk indicates a significant difference in fitness. The open circles on the right of the mean \pm SD bars show individual relative fitness values for each of the three replicates for each of the 33 clones. The solid circles on the left of the mean \pm SD bars show total number of observations for a particular relative fitness value. The shades of gray indicate the total number of observations, with darkest shade being the highest number of observations. We used all the replicate values individually, instead of a mean value, because of the high heterogeneity in our EvoD clones such that multiple subpopulations from even a single EvoD clone are likely to be different from each other in terms of genetic makeup.

experiment (Fig. S19). Taken together, these observations suggest a role of this mutation in downregulating the expression of one of the copies.

To test the fitness effect of this mutation, we reconstructed it in the AncS strain using site-directed mutagenesis along with an appropriate control (Methods S10) and measured their fitness in medium supplemented with 0, 100, and 200 $\mu\text{g}/\text{mL}$ of ampicillin. As would be expected from a downregulating mutation, we saw a significant fitness decrease in the mutant compared to the an-

cestral control at 100 and 200 $\mu\text{g}/\text{mL}$ of ampicillin ($P = 0.0001$ in 100 $\mu\text{g}/\text{mL}$ ampicillin and $P = 1.6 \times 10^{-5}$ in 200 $\mu\text{g}/\text{mL}$ ampicillin; Fig. 6B), suggesting this mutation indeed lowers the expression of the TEM-1 gene. The observation that the mutation shows no fitness differences in the absence of ampicillin (Fig. 6B) further supports this assertion. We also analyzed the effect of the mutation on gene expression by quantitative PCR. The mean expression level in the mutant construct was lower (although not significantly so, Fig. S20) than the ancestral construct, suggesting

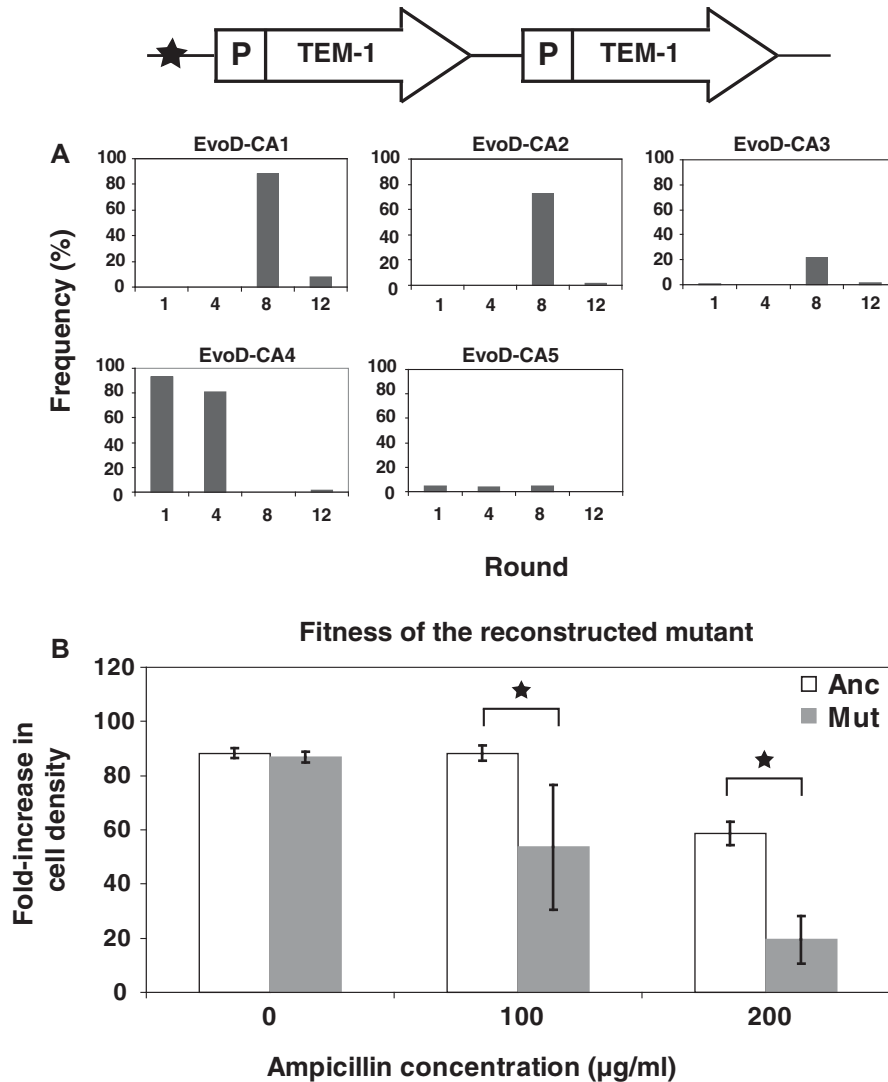


Figure 6. (A) Change in frequency of the regulatory mutation T-53C. We sequenced plasmid inserts in the evolved lines to identify mutations, and observed only one high-frequency mutation. It occurred only in the EvoD lines, and had high frequency only at some time points. It maps to the upstream region of the promoter of the left copy of the TEM-1 gene, as indicated by the asterisk to the left of the boxed region with the P (promoter) and TEM-1 (coding region) lettering. The graphs below show the mutation's frequency (vertical axes) in the EvoD-CA lines at different rounds of the experiment (horizontal axes). The frequency of this mutation was highest at round 8 in the EvoD-CA1, EvoD-CA2, and EvoD-CA3 lines (approximately 80, 60, and 20%, respectively) and had decreased dramatically by round 12. The mutation was present in the starting clone for the EvoD-CA4 line, persisted at high frequency after round 1, and decreased in frequency thereafter (from >80% in round 1 to <1% in round 12). The mutation occurred only in low frequency (<5%) in the EvoD-CA5 line at rounds 1, 4, and 8. (B) Fitness of the reconstructed mutant. To test the effect of the mutation T-53C on the fitness of the EvoD lines, we reconstructed this mutant ("Mut") in the ancestral strain (AncS) by site-directed mutagenesis (see Methods S10). We also reconstructed a control with the ancestral sequence, using the same method to ensure that the method of reconstruction did not bias our results ("Anc"). We then measured the fitness of the reconstructed mutant and the control on 0, 100, and 200 μg/mL of ampicillin. Fitness was significantly lower in the mutant than in the ancestral reconstruction on 100 and 200 μg/mL of ampicillin ($P = 0.0001$ and $P = 1.6 \times 10^{-5}$, respectively), but not in the absence of ampicillin. Columns represent means and whiskers indicate one standard deviation (SD). The asterisk indicates a significant fitness difference in fitness.

a modest reduction in expression level. We note that our results are consistent with an earlier observation that expression reduction may help retain some duplicate genes in yeast and mammals (Qian et al. 2010).

We also considered the possibility that this regulatory mutation reduces fitness by reducing plasmid stability or replicability, but this is not the case. In the EvoD-N3, EvoD-N4, and EvoD-N5 lines, this mutation was present at high frequency (>30%), and

the copy numbers of the plasmids were also very high compared to the ancestor (~14-fold increase; Figs. S15, S19; Table S4).

One question that remains is how this mutation, even though it is deleterious in the single-copy TEM-1 gene, can rise to high frequency in multiple EvoD lines. Because this mutation accumulates specifically in cells with double-copy TEM-1 genes, it provides some benefit for such cells. The experimental observation that an optimal level of gene expression or gene dosage may exist (Dekel and Alon 2005) provides a candidate explanation. If increased dosage or expression of TEM-1 is beneficial only up to a certain level, then a regulatory mutation like T-53C can help achieve this level by reducing expression of one copy of the TEM-1 gene. One of our observations supports this hypothesis. In the neutral EvoD lines that show very high plasmid copy numbers (EvoD-N3, EvoD-N4, and EvoD-N5) the regulatory mutation occurs at high frequency (>30%), whereas this is not the case for the neutral EvoD line (EvoD-N1) without high plasmid copy number. This difference suggests that this mutation helps reduce the cost of expression associated with duplicate genes present on high copy number plasmids.

LOSS OF ONE COPY ALONG WITH LOW FREQUENCY MUTATIONS CAN EXPLAIN THE LOWER FITNESS OF THE EvoD LINES

We had expected the EvoD lines to evolve higher fitness than the EvoS lines, in accordance with existing hypotheses about duplicate gene evolution (Ohno 1970; Bergthorsson et al. 2007). Surprisingly, the opposite had occurred in several lines (Figs. 2B, S8). One candidate explanation relates to the ability of duplicate genes to buffer the effects of otherwise deleterious mutations: Loss-of-function mutations in one gene copy should be neutral or tolerable as long as the other copy provides the needed function (Ohno 1970; Wagner 2000b; Gu et al. 2003; Kellis et al. 2004; Hughes 2005). However, if one of the copies is lost after such mutations have occurred, the deleterious effects of these mutations would be exposed and would reduce fitness (Fig. 7). Over time, such gene deletions would disappear from the population except in either (or both) of the following two circumstances. First, if deletions occur at a very high rate, which is the case in our experiment (Figs. 4, S18). Second, in a heterogeneous population that contains fitter individuals, which hydrolyze antibiotic at a higher rate than less fit individuals, selection can be weakened in the later stages of each selection cycle. Such weaker selection may lead to a slower elimination of single-copy individuals with deleterious mutations.

With these observations in mind, we suggest that the double-copy lines have lower fitness, because their frequent gene copy loss exposes deleterious mutations that have occurred since the experiment's beginning. One prediction of this hypothesis is that deleterious mutations can be tolerated better in the EvoD lines. If

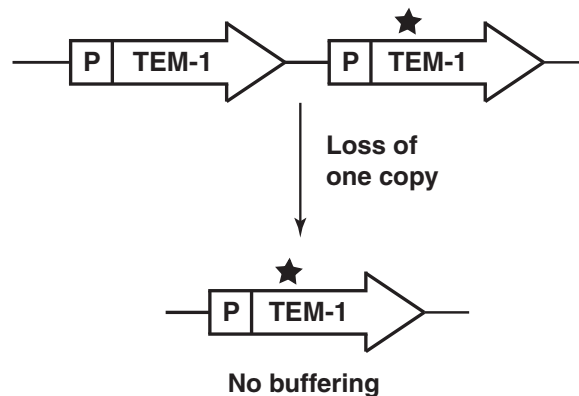


Figure 7. Loss of buffering in individuals with only one copy of the TEM-1 gene in the EvoD lines. Individuals with duplicate copies of the TEM-1 gene may be able to withstand some mutations that would otherwise be deleterious in a single copy of the gene. If one of the copies accumulates a deleterious mutation (indicated by the asterisk) the second copy would be able to buffer its deleterious effects. However, this deleterious effect would be exposed when the second copy is lost.

double-copy lines can indeed buffer deleterious mutations better, then we would expect a significantly higher sequence diversity in EvoD than in EvoS lines, at least at some time points during the evolution experiment. This is indeed the case at round 8 in EvoD-CA and EvoS-CA, even after normalizing for sequence length (Wilcoxon rank sum test, $P = 0.03$; Fig. S21f), as well as in several other EvoD lines (Fig. S21).

A second relevant line of evidence, which we discussed earlier (Fig. 5), is that the EvoD-CA-D clones, which have retained two TEM-1 copies have higher fitness than the single-copy EvoD-CA-S clones, just as the ancestral single-copy strains had at the beginning of the experiment (Figs. 1B–D, 5, S4). Taken together, these observations suggest that the lowered fitness of EvoD lines is caused by a combination of deleterious mutations and frequent gene copy loss in these lines.

Discussion

To determine the relative importance of genetic changes in gene expression and in protein coding regions for the evolution of duplicate genes, we constructed a tandem duplication of the TEM-1 beta lactamase gene on a plasmid in *E. coli*. Initial fitness measurements suggested that cells with two copies of the gene (AncD strain) have greater fitness than cells with only one copy (AncS strain) in some of the antibiotic concentrations used in our experiment. This indicates an advantage of increased TEM-1 gene dosage in the AncD strain. In these initial measurements, we also observed that the concentrations of ampicillin and cefotaxime required to reduce fitness to a given level was considerably lower when ampicillin and cefotaxime were both present in the medium, suggesting a synergistic effect of these two antibiotics (Gunnison

et al. 1955; Moellering et al. 1971; Farrar and Newsome 1973). The ancestral strain with double-copy TEM-1 genes (AncD) had a significant fitness advantage over the ancestral strain with single-copy TEM-1 genes (AncS) in these conditions, likely due to higher dosage.

We then subjected our plasmids to 12 rounds of experimental evolution in the form of mutagenesis and selection. We used several antibiotic conditions to investigate the effect of the selection environment on the outcome of the evolution experiment, and observed a fitness increase in the evolved lines. In some conditions, the EvoS lines carrying only a single copy of the TEM-1 gene had higher fitness than the EvoD lines with two TEM-1 gene copies.

While investigating the molecular mechanisms behind this fitness increase, we encountered four factors that could affect fitness by changing TEM-1 gene dosage and, as a result, gene expression. First, we observed that gene duplication can increase fitness because it increases gene dosage (Fig. 1). Second, during the evolution experiment we saw an average six- to eight-fold beneficial increase in plasmid copy number (Fig. 3), which also increases gene dosage. We observed high-frequency SNPs in the origin of replication of plasmids in the evolved lines. These SNPs are likely to explain the increased plasmid copy number in the evolved lines, although none of them are identical to previously known copy number mutants for this origin of replication (Armstrong et al. 1984). Third, we saw a frequent loss of one gene duplicate in the EvoD lines (Fig. 4), which reduced not only gene dosage but also fitness (Fig. 5). Finally, we found a high-frequency, putative regulatory mutation in the EvoD lines that may have downregulated expression of one TEM-1 gene copy (Fig. 6). It is also relevant here that no point mutations in coding regions (Figs. S21, S22, S26; Results S6) rose to high frequency during the experiment, suggesting that their contributions to improving fitness were small. Taken together, these observations suggest that an increase in gene dosage (resulting in increased gene expression level) through increased plasmid copy number was the most important factor behind the fitness increase in our evolved lines.

We observed that the “neutral” lines, which experienced no antibiotic selection to maintain TEM-1, and only selection on kanamycin to maintain their plasmids, also showed increased fitness in ampicillin. This increase is probably due to increased plasmid copy numbers in these lines, and can be caused by selection at the plasmid level that is unrelated to selection on the TEM-1 gene. For example, increased plasmid copy numbers and thus an increase in dosage of the kanamycin resistance gene could be beneficial under selection on kanamycin. Additionally, plasmids with higher copy number can be preferentially transformed and propagated (Results S5). However, we did not see any difference in growth between the ancestral strains and some of the evolved lines (EvoS-CA and EvoD-CA) on kanamycin (25 μ g/mL;

Fig. S23a), suggesting that the plasmid copy number increase in these lines is unlikely to be only due to selection in kanamycin. In addition, such selection at the plasmid level is not likely to be responsible for all the adaptations that we saw in our evolved lines. The reason is that several distinct changes occurred in the lines selected in antibiotics compared to the neutral lines (Results S7). First, plasmid copy number increases in the EvoD lines selected in novel antibiotics, and in combinations of native and novel antibiotics were different in extent to copy number increases in the neutral EvoD lines (Fig. S15). This could be due to detrimental effects of either high cost of TEM-1 gene expression, or high metabolic cost associated with high plasmid copy numbers in the selection conditions. Second, apart from the EvoD-A \uparrow lines, in all other EvoD lines selected in antibiotics, one gene copy was lost at a higher rate than in the EvoD-N lines (Fig. S17), which may be caused by higher instability of double-copy inserts in antibiotic medium. In this regard, it is worthwhile to note that in an earlier study, an increased antibiotic concentration led to an increased maintenance cost of a duplicate gene on a second plasmid and to the loss of this plasmid from a bacterial population (Holloway et al. 2007). Finally, in almost all the EvoD lines, the regulatory mutation T-53C rose to high frequency at round 8, and subsequently dropped to low frequency at round 12, except in the neutral lines (Fig. S19). We speculate that this regulatory mutation might downregulate expression of one TEM-1 gene copy, and if so, it was beneficial until round 8, as it helped reach an optimal expression level in increased plasmid copy number but became detrimental with increasing antibiotic concentration beyond round 12. In some of the neutral (EvoD-N) lines, this mutation was retained at high frequency. It may have benefitted these lines by reducing the high gene expression costs that the high increase in plasmid copy number had caused in these lines.

We next discuss how changes that increase gene dosage, such as an increasing plasmid copy number, and changes that reduce gene dosage, such as gene loss and our putative regulatory mutation, can occur simultaneously in our experiment. One possible reason is that an increase in gene dosage is beneficial up to a certain extent but can be detrimental beyond that, such that dosage level reduction can become beneficial (Dekel and Alon 2005; Poelwijk et al. 2011). Two observations are consistent with this possibility. First, the putative regulatory mutation T-53C decreased in population frequency, as we increased the antibiotic concentration after round 8. The mutation may still have been beneficial for the cells at round 8, but became detrimental as we increased antibiotic concentrations after round 8. Second, we saw a very high copy number increase only in the EvoD lines without selection, and in the lines selected in ampicillin (Fig. S15), but not in EvoD lines selected in novel substrates or in combinations of native and novel substrates. Taken together, these observations suggest that multiple changes influencing gene dosage occur in

our evolved lines, because they contribute to adjusting dosage levels to some optimum in the EvoD lines.

Among the limitations of our study is that it does not address how gene duplicates go to fixation in a population. The duplicate genes we studied were already “fixed” in our populations, because we were interested in their subsequent evolution. However, our experiments show that duplicate genes can confer a significant fitness advantage in some environments over single-copy gene. This advantage can help in fixation of a duplicate, because such fixation may require positive selection (Clarke 1994; Moore and Purugganan 2003).

Another limitation is that we did not study the extent of clonal interference, and how it contributes to the fate of duplicate genes. Clonal interference could be a reason behind the decrease in frequency of the T-53C mutation in several EvoD lines beyond round 8 (Miralles et al. 1999; Kao and Sherlock 2008). We did not observe any high-frequency mutation in TEM-1 genes, which would be required for such clonal interference, but note that such mutations could have occurred in the origin of replication or elsewhere in the plasmid.

A third limitation is that we used a plasmid to study the divergence of gene duplicates. This plasmid occurred at 5–10 copies per cell, implying that we compared the fitness and the evolution of 5–10 gene copies (in our single-copy TEM-1 lines) with that of 10–20 gene copies (for our double-copy TEM-1 lines). In contrast, gene duplicates in nature often occur on a single copy in the bacterial chromosome. Our experimental design may have given gene dosage and expression changes an intrinsic advantage over coding region changes, because the former may be easier to achieve for plasmid-borne genes. In contrast, gene copy number increases of chromosomal genes may have deleterious pleiotropic effects, because they often occur as part of segmental duplications of a chromosome. As a result, changes in the protein coding region may contribute more readily to evolutionary adaptation for chromosomally duplicated genes.

In sum, our experiments show that an increase in gene dosage is initially the most important factor in the evolution of duplicate TEM-1 beta lactamase genes under the influence of antibiotics. In these experiments, the dosage increase occurs through changed plasmid copy numbers. Both selection on plasmid copy number itself and on the ability of TEM-1 to neutralize antibiotics play a role in these experiments. An increase in gene dosage following gene duplication can provide an initial benefit to cells, and could pave the way for neofunctionalization under appropriate selection conditions through an accumulation of mutations (Andersson and Hughes 2009; Näsvall et al. 2012).

ACKNOWLEDGMENTS

We thank the Functional Genomics Center Zurich for its service in generating sequencing data, M. Ackermann and E. Hayden for helpful discus-

sions, A. de Visser for comments on earlier versions of this manuscript, and M. Moser for help with quantitative PCR. This work was supported by Swiss National Science Foundation (grant 315230–129708), as well as through the YeastX project of SystemsX.ch, and the University Priority Research Program in Systems Biology at the University of Zurich. RD acknowledges support from the Forschungskredit program of the University of Zurich. The authors declare no conflict of interest.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.jc402.

LITERATURE CITED

- Ahmed, A. 1975. Mechanism of reversion of the gal3 mutation of *Escherichia coli*. *Mol. Gen. Genet.* 136:243–253.
- Andersson, D. I., and D. Hughes. 2009. Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.* 43:167–195.
- Andersson, D. I., E. S. Slechta, and J. R. Roth. 1998. Evidence that gene amplification underlies adaptive mutability of the bacterial lac operon. *Science* 282:1133–1135.
- Anderson, R. P., and J. R. Roth. 1977. Tandem genetic duplications in phage and bacteria. *Annu. Rev. Microbiol.* 31:473–505.
- . 1978a. Tandem chromosomal duplications in *Salmonella typhimurium*: fusion of histidine genes to novel promoters. *J. Mol. Biol.* 119:147–166.
- . 1978b. Tandem genetic duplications in *Salmonella typhimurium*: amplification of the histidine operon. *J. Mol. Biol.* 126:53–71.
- Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815.
- Armstrong, K. A., R. Acosta, E. Ledner, Y. Machida, M. Pancotto, M. McCormick, H. Ohtsubo, and E. Ohtsubo. 1984. A 37×10(3) molecular weight plasmid-encoded protein is required for replication and copy number control in the plasmid pSC101 and its temperature-sensitive derivative pHS1. *J. Mol. Biol.* 175:331–348.
- Bailey, J. A., Z. Gu, R. A. Clark, K. Reinert, R. V. Samonte, S. Schwartz, M. D. Adams, E. W. Myers, P. W. Li, and E. E. Eichler. 2002. Recent segmental duplications in the human genome. *Science* 297:1003–1007.
- Balakirev, E. S., and F. J. Ayala. 2003. Pseudogenes: are they “Junk” or functional DNA? *Annu. Rev. Genet.* 37:123–151.
- Barkman, T., and J. Zhang. 2009. Evidence for escape from adaptive conflict? *Nature* 462:E1; discussion E2–3.
- Barlow, M., and B. G. Hall. 2002. Predicting evolutionary potential: in vitro evolution accurately reproduces natural evolution of the tem beta-lactamase. *Genetics* 160:823–832.
- Bergthorsson, U., D. I. Andersson, and J. R. Roth. 2007. Ohno’s dilemma: evolution of new genes under continuous selection. *Proc. Natl. Acad. Sci. USA* 104:17004–17009.
- Bowers, J. E., B. A. Chapman, J. Rong, and A. H. Paterson. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422:433–438.
- Bridges, C. B., 1936. The bar “gene” a duplication. *Science* 83:210–211.
- Brown, C. J., K. M. Todd, and R. F. Rosenzweig. 1998. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol. Biol. Evol.* 15:931–942.
- Clark, A. G. 1994. Invasion and maintenance of a gene duplication. *Proc. Natl. Acad. Sci. USA* 91:2950–2954.
- Conant, G. C., and K. H. Wolfe. 2007. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol. Syst. Biol.* 3:1–12.
- . 2008. Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* 9:938–950.

- Dehal, P., and Boore, J. L. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3:1700–1708.
- Dekel, E., and U. Alon. 2005. Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436:588–592.
- DeLuna, A., K. Vetsigian, N. Shores, M. Hegreness, M. Colon-Gonzalez, S. Chao, and R. Kishony. 2008. Exposing the fitness contribution of duplicated genes. *Nat. Genet.* 40:676–681.
- Deng, C., C. H. Cheng, H. Ye, X. He, and L. Chen. 2010. Evolution of an antifreeze protein by neofunctionalization under escape from adaptive conflict. *Proc. Natl. Acad. Sci. USA* 107:21593–21598.
- Des Marais, D. L., and M. D. Rausher. 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* 454:762–765.
- Dianov, G. L., A. V. Kuzminov, A. V. Mazin, and R. I. Salganik. 1991. Molecular mechanisms of deletion formation in *Escherichia coli* plasmids. I. Deletion formation mediated by long direct repeats. *Mol. Gen. Genet.* 228:153–159.
- Dunham, M. J., H. Badrane, T. Ferea, J. Adams, P. O. Brown, F. Rosenzweig, and D. Botstein. 2002. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 99:16144–16149.
- Farrar, W. E. Jr., and J. K. Newsome. 1973. Mechanism of synergistic effects of beta-lactam antibiotic combinations on gram-negative bacilli. *Antimicrob. Agents Chemother.* 4:109–114.
- Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan, and J. Postlethwait. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–1545.
- Francino, M. P. 2005. An adaptive radiation model for the origin of new gene functions. *Nat. Genet.* 37:573–577.
- Glansdorff, N., and G. Sand. 1968. Duplication of a gene belonging to an arginine operon of *Escherichia coli* K-12. *Genetics* 60:257–268.
- Goldberg, I., and J. J. Mekalanos. 1986. Effect of a recA mutation on cholera toxin gene amplification and deletion events. *J. Bacteriol.* 165:723–731.
- Gu, Z., L. M. Steinmetz, X. Gu, C. Scharfe, R. W. Davis, and W. H. Li. 2003. Role of duplicate genes in genetic robustness against null mutations. *Nature* 421:63–66.
- Gunnison, J. B., E. Kunishige, V. R. Coleman, and E. Jawetz. 1955. The mode of action of antibiotic synergism and antagonism: the effect in vitro on bacteria not actively multiplying. *J. Gen. Microbiol.* 13:509–518.
- Hahn, M. W. 2009. Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.* 100:605–617.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557–580.
- Hasunuma, K., and M. Sekiguchi. 1977. Replication of plasmid pSC101 in *Escherichia coli* K12: requirement for dnaA function. *Mol. Gen. Genet.* 154:225–230.
- Hendrickson, H., E. S. Slechts, U. Bergthorsson, D. I. Andersson, and J. R. Roth. 2002. Amplification-mutagenesis: evidence that “directed” adaptive mutation and general hypermutability result from growth with a selected gene amplification. *Proc. Natl. Acad. Sci. USA* 99:2164–2169.
- Himmelreich, R., H. Hilbert, H. Plagens, E. Pirkel, B. C. Li, and R. Herrmann. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.* 24:4420–4449.
- Holloway, A. K., T. Palzkill, and J. J. Bull. 2007. Experimental evolution of gene duplicates in a bacterial plasmid model. *J. Mol. Evol.* 64:215–222.
- Hooper, S. D., and O. G. Berg. 2003. On the nature of gene innovation: duplication patterns in microbial genomes. *Mol. Biol. Evol.* 20:945–954.
- Horiuchi, T., S. Horiuchi, and A. Novick. 1963. The genetic basis of hyper-synthesis of β -galactosidase. *Genetics* 48:157–169.
- Hughes, A. L., 1994. The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond. B* 256:119–124.
- . 2005. Gene duplication and the origin of novel proteins. *Proc. Natl. Acad. Sci. USA* 102:8791–8792.
- Hughes T., and D. A. Liberles. 2007. The pattern of evolution of smaller-scale gene duplicates in mammalian genomes is more consistent with neo- than subfunctionalisation. *J. Mol. Evol.* 65:574–588.
- Ives, C. L., and K. F. Bott. 1990. Characterization of chromosomal DNA amplifications with associated tetracycline resistance in *Bacillus subtilis*. *J. Bacteriol.* 172:4936–4944.
- Jackson, E. N., and C. Yanofsky. 1973. Duplication-translocations of tryptophan operon genes in *Escherichia coli*. *J. Bacteriol.* 116:33–40.
- Kafri, R., M. Levy, and Y. Pilpel. 2006. The regulatory utilization of genetic redundancy through responsive backup circuits. *Proc. Natl. Acad. Sci. USA* 103:11653–11658.
- Kafri, R., O. Dahan, J. Levy, and Y. Pilpel. 2008. Preferential protection of protein interaction network hubs in yeast: evolved functionality of genetic redundancy. *Proc. Natl. Acad. Sci. USA* 105:1243–1248.
- Kao, K. C., and G. Sherlock. 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat. Genet.* 40:1499–1504.
- Kellis, M., B. E. Birren, and E. S. Lander. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428:617–624.
- Kleinjan, D. A., R. M. Bancewicz, P. Gautier, R. Dahm, H. B. Schonhaler, G. Damante, A. Seawright, A. M. Hever, P. L. Yeyati, V. van Heyningen, et al. 2008. Subfunctionalization of duplicated zebrafish pax6 genes by cis-regulatory divergence. *PLoS Genet.* 4:e29.
- Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. E. Ketchum, R. J. Dodson, M. Gwinn, E. K. Hickey, J. D. Peterson, et al. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* 390:364–370.
- Kondrashov, F. A., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin. 2002. Selection in the evolution of gene duplications. *Genome Biol.* 3:1–8.
- Kondratyeva, T. F., L. N. Muntyan, and G. I. Karavaiko. 1995. Zinc- and arsenic-resistant strains of *Thiobacillus ferrooxidans* have increased copy numbers of chromosomal resistance genes. *Microbiology* 141:1157–1162.
- Kugelberg, E., E. Kofoid, A. B. Reams, D. I. Andersson, and J. R. Roth. 2006. Multiple pathways of selected gene amplification during adaptive mutation. *Proc. Natl. Acad. Sci. USA* 103:17319–17324.
- Kugelberg, E., E. Kofoid, D. I. Andersson, Y. Lu, J. Mellor, F. P. Roth, and J. R. Roth. 2010. The tandem inversion duplication in *Salmonella enterica*: selection drives unstable precursors to final mutation types. *Genetics* 185:65–80.
- Kuwada, Y. 1911. Meiosis in the pollen mother cells of *Zea Mays* L. *Bot. Mag.* 25:1633.
- Li, W. H., Z. Gu, H. Wang, and A. Nekrutenko. 2001. Evolutionary analyses of the human genome. *Nature* 409:847–849.
- Lovett, S. T., P. T. Drapkin, V. A. Suter Jr., and T. J. Gluckman-Peskind. 1993. A sister-strand exchange mechanism for recA-independent deletion of repeated DNA sequences in *Escherichia coli*. *Genetics* 135:631–642.
- Lynch, M., and J. S. Conery. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–1155.
- Lynch, M., and A. Force. 2000a. The origin of interspecies genomic incompatibility via gene duplication. *Am. Nat.* 156:590–605.
- . 2000b. The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154:459–473.
- Lynch, M., M. O’Hely, B. Walsh, and A. Force. 2001. The probability of preservation of a newly arisen gene duplicate. *Genetics* 159:1789–1804.
- Lynch, V. J. 2007. Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes. *BMC Evol. Biol.* 7:2.

- Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380.
- Marques-Bonet, T., J. M. Kidd, M. Ventura, T. A. Graves, Z. Cheng, L. W. Hillier, Z. Jiang, C. Baker, R. Malfavon-Borja, L. A. Fulton, et al. 2009. A burst of segmental duplications in the genome of the African great ape ancestor. *Nature* 457:877–881.
- Matfield, M., R. Badawi, and W. J. Brammar. 1985. Rec-dependent and Rec-independent recombination of plasmid-borne duplications in *Escherichia coli* K12. *Mol. Gen. Genet.* 199:518–523.
- Matthews, P. R., and P. R. Stewart. 1988. Amplification of a section of chromosomal DNA in methicillin-resistant *Staphylococcus aureus* following growth in high concentrations of methicillin. *J. Gen. Microbiol.* 134:1455–1464.
- Miralles, R., P. J. Gerrish, A. Moya, and S. F. Elena. 1999. Clonal interference and the evolution of RNA viruses. *Science* 285:1745–1747.
- Moellering, R. C. Jr., C. Wennersten, and A. N. Weinberg. 1971. Studies on antibiotic synergism against enterococci. I. Bacteriologic studies. *J. Lab. Clin. Med.* 77:821–828.
- Moore, R. C., and M. D. Purugganan. 2003. The early stages of duplicate gene evolution. *Proc. Natl. Acad. Sci. USA* 100:15682–15687.
- Näsvall, J., L. Sun, J. R. Roth, and D. I. Andersson. 2012. Real-time evolution of new genes by innovation, amplification, and divergence. *Science* 338:384–387.
- Nichols, B. P., and G. G. Guay. 1989. Gene amplification contributes to sulfonamide resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* 12:2042–2048.
- Nowak, M. A., M. C. Boerlijst, J. Cooke, and J. M. Smith. 1997. Evolution of genetic redundancy. *Nature* 388:167–171.
- Ohno, S. 1970. *Evolution by gene duplication*. Springer-Verlag, Heidelberg, Germany.
- Otto, E., J. E. Young, and G. Maroni. 1986. Structure and expression of a tandem duplication of the *Drosophila* metallothionein gene. *Proc. Natl. Acad. Sci. USA* 83:6025–6029.
- Poelwijk, F. J., P. D. Heyning, M. G. de Vos, D. J. Kiviet, and S. J. Tans. 2011. Optimality and evolution of transcriptionally regulated gene expression. *BMC Syst. Biol.* 5:128.
- Qian, W., B. Y. Liao, A. Y. Chang, and J. Zhang. 2010. Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet.* 26:425–430.
- Rastogi, S., and D. A. Liberles. 2005. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol. Biol.* 5:28.
- Reams, A. B., and E. L. Neidle. 2003. Genome plasticity in *Acinetobacter*: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. *Mol. Microbiol.* 47:1291–1304.
- Reams, A. B., E. Kofoid, M. Savageau, and J. R. Roth. 2010. Duplication frequency in a population of *Salmonella enterica* rapidly approaches steady state with or without recombination. *Genetics* 184:1077–1094.
- Reams, A. B., E. Kofoid, E. Kugelberg, and J. R. Roth. 2012. Multiple pathways of duplication formation with and without recombination (RecA) in *Salmonella enterica*. *Genetics* 192:397–415.
- 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-Trelles, F., R. Tarrío, and F. J. Ayala. 2003. Convergent neofunctionalization by positive Darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proc. Natl. Acad. Sci. USA* 100:13413–13417.
- Roth, J. R., and Andersson, D. I. 2004. Amplification-mutagenesis—how growth under selection contributes to the origin of genetic diversity and explains the phenomenon of adaptive mutation. *Res. Microbiol.* 155:342–351.
- Rouquier, S., S. Taviaux, B. Trask, V. Brand-Arpon, G. van den Engh, J. Demaille, and D. Giorgi. 1998. Distribution of olfactory receptor genes in the human genome. *Nat. Genet.* 18:243–250.
- Rouquier, S., A. Blancher, and D. Giorgi. 2000. The olfactory receptor gene repertoire in primates and mouse: evidence for reduction of the functional fraction in primates. *Proc. Natl. Acad. Sci. USA* 97:2870–2874.
- Rubin, G. M., M. D. Yandell, J. R. Wortman, G. L. Gabor Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, P. W. Li, R. Apweiler, W. Fleischmann, et al. 2000. Comparative genomics of the eukaryotes. *Science* 287:2204–2215.
- Serebrovsky, A. S. 1938. Genes scute and achaete in *Drosophila melanogaster* and a hypothesis of gene divergency. *C. R. Acad. Sci. URSS* 19:77–81.
- Sidow, A. 1996. Gen(om)e duplications in the evolution of early vertebrates. *Curr. Opin. Genet. Dev.* 6:715–722.
- Simillion, C., K. Vandepoele, M. C. Van Montagu, M. Zabeau, and Y. Van de Peer. 2002. The hidden duplication past of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 99:13627–13632.
- Sonti, R. V., and J. R. Roth. 1989. Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* 123:19–28.
- Soo, V. W., P. Hanson-Manful, and W. M. Patrick. 2011. Artificial gene amplification reveals an abundance of promiscuous resistance determinants in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 108:1484–1489.
- Spring, J. 1997. Vertebrate evolution by interspecific hybridisation—are we polyploids? *FEBS Lett.* 400:2–8.
- Stemmer, W. P. 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370:389–391.
- Stephens, S. G. 1951. Possible significance of duplications in evolution. *Adv. Genet.* 4:247–265.
- Straus, D. S. 1975. Selection for a large genetic duplication in *Salmonella typhimurium*. *Genetics* 80:227–237.
- Sutcliffe, J. G. 1978. Nucleotide sequence of the ampicillin resistance gene of *Escherichia coli* plasmid pBR322. *Proc. Natl. Acad. Sci. USA* 75:3737–3741.
- Swords, W. E. 2003. Chemical transformation of *E. coli*. Pp 49–53 in N. Casali and A. Preston, eds. *E. coli* plasmid vectors methods and applications, methods in molecular biology, Vol. 235. Humana Press, Totowa, NJ.
- Taylor, J. S., and J. Raes. 2004. Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.* 38:615–643.
- Tekaia, F., and B. Dujon. 1999. Pervasiveness of gene conservation and persistence of duplicates in cellular genomes. *J. Mol. Evol.* 49:591–600.
- Tirosh, I., and N. Barkai. 2007. Comparative analysis indicates regulatory neofunctionalization of yeast duplicates. *Genome Biol.* 8:R50.
- Tocchini-Valentini, G. D., P. Fruscoloni, and G. P. Tocchini-Valentini. 2005. Structure, function, and evolution of the tRNA endonucleases of *Archaea*: an example of subfunctionalization. *Proc. Natl. Acad. Sci. USA* 102:8933–8938.
- Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, et al. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388:539–547.
- Toprak, E., A. Veres, J. B. Michel, R. Chait, D. L. Hartl, and R. Kishony. 2011. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genet.* 44:101–105.
- Vavouri, T., J. L. Semple, and B. Lehner. 2008. Widespread conservation of genetic redundancy during a billion years of eukaryotic evolution. *Trends Genet.* 24:485–488.

- Wagner, A. 2000a. The role of population size, pleiotropy and fitness effects of mutations in the evolution of overlapping gene functions. *Genetics* 154:1389–1401.
- . 2000b. Robustness against mutations in genetic networks of yeast. *Nat. Genet.* 24:355–361.
- . 2002. Selection and gene duplication: a view from the genome. *Genome Biol.* 3:1–3.
- Walsh, J. B. 1995. How often do duplicated genes evolve new functions? *Genetics* 139:421–428.
- Warringer, J., E. Zörgö, F. A. Cubillos, A. Zia, A. Gjuvsland, J. T. Simpson, A. Forsmark, R. Durbin, S. W. Omholt, E. J. Louis, et al. 2011. Trait variation in yeast is defined by population history. *PLoS Genet.* 7:e1002111.
- Wiebauer, K., S. Schraml, S.W. Shales, and R. Schmitt. 1981. Tetracycline resistance transposon Tn1721: recA-dependent gene amplification and expression of tetracycline resistance. *J. Bacteriol.* 147:851–859.
- Woolfe, A., and G. Elgar. 2007. Comparative genomics using Fugu reveals insights into regulatory subfunctionalization. *Genome Biol.* 8:R53.
- Zaccolo, M., and E. Gherardi. 1999. The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.* 285:775–783.
- Zaslaver, A., A. Bren, M. Ronen, S. Itzkovitz, I. Kikoin, S. Shavit, W. Liebermeister, M. G. Surette, and U. Alon. 2006. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* 3:623–628.
- Zhang, J. 2003. Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18:292–298.
- Zhang, J., Y. Zhang, and H. F. Rosenberg. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nat. Genet.* 30:411–415.

Associate Editor: A. De Visser

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

- Figure S1.** Plasmid pUA67.
- Figure S2.** Clones derived from AncS and AncD strains that were used in our experiment.
- Figure S3.** Nomenclature of the evolved lines based on the selection condition.
- Figure S4.** Fitness of the ancestral strains (AncS and AncD) in ampicillin + cefotaxime.
- Figure S5.** Fitness of the ancestral strains in imipenem.
- Figure S6.** Initial fitness measurement of the ancestral strains (AncS and AncD) in ampicillin + imipenem.
- Figure S7.** Growth curves of the ancestral strains (AncS and AncD).
- Figure S8.** Fitness of the evolved CA lines.
- Figure S9.** Fitness of the evolved N lines.
- Figure S10.** Fitness of the evolved A lines.
- Figure S11.** Fitness of the evolved A↑ lines.
- Figure S12.** Fitness of the evolved C lines.
- Figure S13.** Fitness of the evolved I lines.
- Figure S14.** Fitness of the evolved IA lines.
- Figure S15.** Increase of plasmid copy number in the evolved lines.
- Figure S16.** Analysis of TEM-1 gene expression level in AncS, AncD, EvoS-CA3, and EvoD-CA3 lines using quantitative PCR.
- Figure S17.** Frequency of individuals retaining duplicate copies of the TEM-1 gene in the EvoD lines.
- Figure S18.** Analysis of the inserts in the plasmids isolated from the clones of the EvoD-CA1 line.
- Figure S19.** Change in population frequency of the T-53C mutation in the evolved lines over the rounds of the experiment.
- Figure S20.** Analysis of TEM-1 mRNA expression level of the reconstructed T-53C mutant and the wild-type control using quantitative PCR.
- Figure S21.** Change in the sequence diversities of the evolved lines over the course of the experiment.
- Figure S22.** Change in the sequence diversities in replicate lines.
- Figure S23.** Growth curves of the evolved lines EvoS-CA3 and EvoD-CA3.
- Figure S24.** Antibiotics concentrations used for selection experiments.
- Figure S25.** Correlation between the read coverage and the number of polymorphisms detected.
- Figure S26.** Change in frequency of the mutations that were present in the starting lines over the rounds of the experiment.
- Figure S27.** Average number of nucleotide differences between sequence reads in pairwise comparisons.
- Figure S28.** K_a/K_s ratio in the evolved lines.
- Figure S29.** Barcodes in the amplicons for 454 sequencing (not to scale).
- Figure S30.** Amplicons used for 454 sequencing and how they map to TEM-1 regions (not to scale).
- Figure S31.** Error rates in the amplicons from the next generation sequencing data.
- Figure S32.** Error rates in two sets of reads in the 454 sequencing data.
- Table S1.** Primer pairs for screening of individuals containing duplicate genes in the EvoD lines.
- Table S2.** Fitness measurements of the ancestral strains.
- Table S3.** Relative fitness values of the evolved lines.
- Table S4.** Fold-increase in plasmid copy number in the evolved lines.
- Table S5.** Percentage of individuals that retained duplicate genes in the EvoD lines.
- Table S6.** Average mutation rate in the mutator *E. coli* strain (TB90).
- Table S7.** Primers with barcodes for sample multiplexing in 454 sequencing.