

# Tempo and mode of genome evolution in a 50,000-generation experiment

Olivier Tenaillon<sup>1\*</sup>, Jeffrey E. Barrick<sup>2,3\*</sup>, Noah Ribeck<sup>3,4</sup>, Daniel E. Deatherage<sup>2</sup>, Jeffrey L. Blanchard<sup>5</sup>, Aurko Dasgupta<sup>2†</sup>, Gabriel C. Wu<sup>2</sup>, Sébastien Wielgoss<sup>6,7</sup>, Stéphane Cruveiller<sup>8</sup>, Claudine Médigue<sup>8</sup>, Dominique Schneider<sup>7,9</sup> & Richard E. Lenski<sup>3,4\*</sup>

**Adaptation by natural selection depends on the rates, effects and interactions of many mutations, making it difficult to determine what proportion of mutations in an evolving lineage are beneficial. Here we analysed 264 complete genomes from 12 *Escherichia coli* populations to characterize their dynamics over 50,000 generations. The populations that retained the ancestral mutation rate support a model in which most fixed mutations are beneficial, the fraction of beneficial mutations declines as fitness rises, and neutral mutations accumulate at a constant rate. We also compared these populations to mutation-accumulation lines evolved under a bottlenecking regime that minimizes selection. Nonsynonymous mutations, intergenic mutations, insertions and deletions are overrepresented in the long-term populations, further supporting the inference that most mutations that reached high frequency were favoured by selection. These results illuminate the shifting balance of forces that govern genome evolution in populations adapting to a new environment.**

Comparative genomic studies have identified the molecular basis of adaptations including lactase permanence in humans<sup>1</sup>, domestication of plants<sup>2</sup> and animals<sup>3</sup>, and pathogenicity in bacteria<sup>4</sup>. Nevertheless, it is difficult to determine more generally what fraction of new mutations in an evolving lineage are beneficial. Answering this question is important for modelling sequence changes used in phylogenetic methods<sup>5</sup> and would inform debate about adaptive and non-adaptive modes of genome evolution<sup>6,7</sup>.

The combination of experimental evolution and genome sequencing provides a way forward that has been used with viruses, bacteria, yeast and flies<sup>8–13</sup>. In a study of bacteria, the diversity of mutations involved in adaptation to high-temperature stress was studied by sequencing >100 lineages after a 2,000-generation experiment<sup>10</sup>. In another study, sequencing a series of clones from one population over 40,000 generations showed the trajectory of genome evolution<sup>9</sup>. However, a short-term experiment reveals only the early steps of adaptation, and it is difficult to distinguish adaptive ‘driver’ and non-adaptive ‘passenger’ mutations when only one population is examined. Beneficial mutations can also be identified by lineage tracking<sup>14</sup> and genetic reconstruction<sup>15</sup> experiments, but these approaches become impractical after an initial selective sweep or when mutations become too numerous over time, respectively.

To overcome these limitations, we analysed complete genomes of 264 clones from 12 populations across 50,000 generations of the long-term evolution experiment (LTEE) with *E. coli*<sup>16,17</sup>. These populations have evolved in a defined medium with scarce resources since 1988. Mean fitness measured in competition with their ancestor increased by ~70% in that time<sup>17</sup>. The LTEE is a model system for studying many fundamental evolutionary questions<sup>9,15–23</sup>.

## Genome-wide mutations and hypermutability

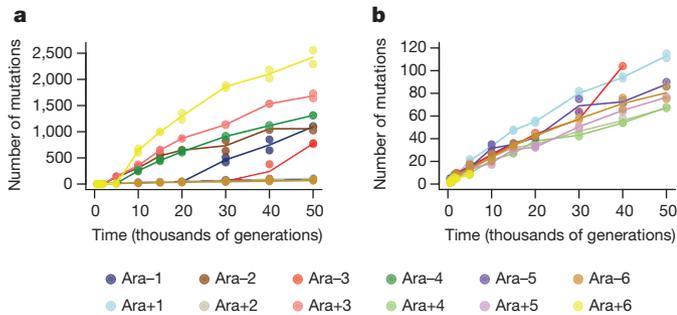
We sequenced the genomes of two clones from each population after 500, 1,000, 1,500, 2,000, 5,000, 10,000, 15,000, 20,000, 30,000, 40,000

and 50,000 generations using the Illumina platform (Supplementary Data 1). We called mutations, including structural variants, using the *breseq* pipeline<sup>24,25</sup>. In total, we found 14,572 point mutations; 500 insertions of insertion sequence (IS) elements; 726 deletions and 1,132 insertions each ≤ 50 base pairs (bp) (small indels); and 267 deletions and 45 duplications each > 50 bp (large indels). After 50,000 generations, average genome length declined by 63 kb (~1.4%) relative to the ancestor (Extended Data Fig. 1). Mutations were not distributed uniformly across the populations. Instead, six populations (Ara–1, Ara–2, Ara–3, Ara–4, Ara+3 and Ara+6) had 96.5% of the point mutations, having evolved hypermutable phenotypes caused by mutations that affect DNA repair or removal of oxidized nucleotides<sup>18,20</sup>. Figure 1a shows the trajectories for the total mutations in all 12 populations; Fig. 1b is rescaled for better resolution of those that did not become point-mutation mutators. Hypermutability tended to decline over time as the load of deleterious mutations favoured antimutator alleles<sup>20</sup>. All four populations that were hypermutable at 10,000 generations accumulated synonymous substitutions (a proxy for the underlying point-mutation rate) between generations 40,000 and 50,000 at much lower rates than from 10,000 to 20,000 generations (Extended Data Fig. 2).

Increased numbers of IS elements can also cause hypermutability<sup>26</sup>, with higher rates not only of transpositions but also deletions and duplications through homologous recombination. In population Ara+1, 31.8% of all mutations up to 50,000 generations were IS150 insertions, compared with 12.3% for the other populations that never evolved elevated point-mutation rates. This mode of hypermutability arose early in Ara+1; IS150 insertions are overrepresented in each Ara+1 clone from 5,000 generations onwards when compared individually to all other non-mutator clones from the same generation (Fisher’s exact test with Bonferroni correction,  $P < 0.05$ ). Two clones from other populations were also IS150 hypermutators by this test: 38.7% of the mutations in

<sup>1</sup>AME, UMR 1137, INSERM, Université Paris Diderot, Sorbonne Paris Cité, F-75018 Paris, France. <sup>2</sup>Department of Molecular Biosciences, Institute for Cellular and Molecular Biology, Center for Systems and Synthetic Biology, Center for Computational Biology and Bioinformatics, The University of Texas at Austin, Austin, Texas 78712, USA. <sup>3</sup>BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, Michigan 48824, USA. <sup>4</sup>Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824, USA. <sup>5</sup>Department of Biology, University of Massachusetts, Amherst, Massachusetts 01003, USA. <sup>6</sup>Institute of Integrative Biology, ETH Zürich, Universitätsstrasse 16, Zürich 8092, Switzerland. <sup>7</sup>Université Grenoble Alpes, Laboratoire Technologies de l’Ingénierie Médicale et de la Complexité — Informatique, Mathématiques et Applications (TIMC-IMAG), F-38000 Grenoble, France. <sup>8</sup>UMR 8030, CNRS, Université Évy-Val-d’Essonne, CEA, Institut de Génétique, Laboratoire d’Analyses Bioinformatiques pour la Génétique et le Métabolisme, F-91000 Évy, France. <sup>9</sup>Centre National de la Recherche Scientifique, TIMC-IMAG, F-38000 Grenoble, France. <sup>†</sup>Present address: Department of Internal Medicine, Washington University School of Medicine, St Louis, Missouri 63110, USA.

\*These authors contributed equally to this work.



**Figure 1 | Total number of mutations over time in the 12 LTEE populations. a**, Total mutations in each population. **b**, Total mutations rescaled to reveal the trajectories for the six populations that did not become hypermutable for point mutations, and for the other six before they evolved hypermutability. Each symbol shows a sequenced genome; some points are hidden behind others. Each line passes through the average of the genomes from the same population and generation.

a 30,000-generation clone from Ara-5 and 31.7% of the mutations in a 40,000-generation clone from Ara-3 were *IS150* insertions. The aberrant Ara-5 clone shares only one mutation with other sequenced Ara-5 clones, indicating early divergence; it does not share point mutations with any other population, excluding cross-contamination. The emergence of these various mutator types shows that evolution can alter the production of genetic diversity<sup>20,27</sup>, which in turn changes the tempo and mode of genome evolution.

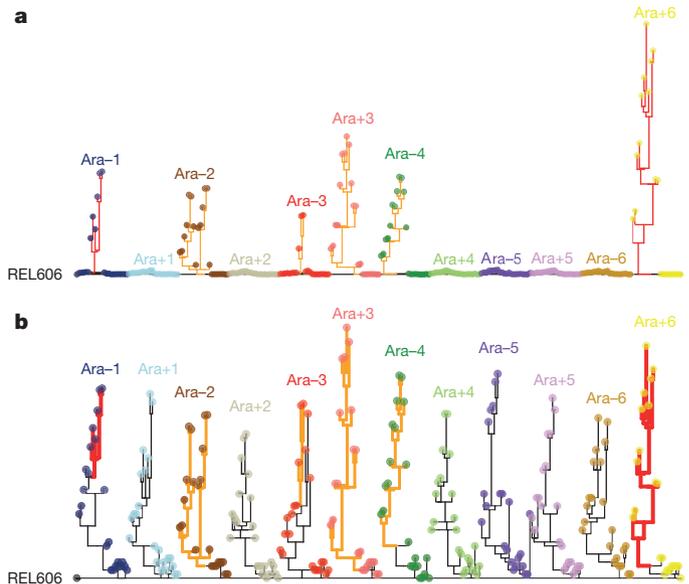
### Population phylogenies

Figure 2a shows phylogenetic trees constructed using point mutations for each population; Fig. 2b shows the trees with branches rescaled after mutators evolved. Some populations—including Ara-2, which became hypermutable early, and Ara-6, which never did—harbour lineages that coexisted for tens of thousands of generations. Some others—including Ara-4, which became hypermutable, and Ara+2, which did not—are more linear in structure, without deep branches among the sequenced clones. Deep branches were probably supported by the diversity-promoting effects of negative-frequency-dependent interactions, as shown in the Ara-2 population<sup>22,23</sup>. Sequencing whole-population samples would provide more detailed information on within-population diversity<sup>11,12</sup>.

### Dynamics of genome evolution

The accumulation of point mutations increased greatly in hypermutable populations<sup>9,19,20</sup>, potentially overwhelming the genomic signature of adaptation. Although mutator lineages may experience higher rates of fitness improvement<sup>17,27</sup>, the effect is usually small owing to clonal interference between competing beneficial mutations<sup>28,29</sup> and the increased load of deleterious mutations<sup>20,30</sup>. Therefore, beneficial mutations become harder to detect in a sea of unselected mutations in mutator lineages. To understand better the dynamic coupling between adaptation and genome evolution, we first analysed the populations that retained the ancestral mutation rate up to 50,000 generations and the others before they became point-mutation or *IS150* mutators.

It was previously found<sup>17</sup> that the mean-fitness trajectory of the LTEE is well described by a power-law relation, in which log fitness increases linearly with log time. Moreover, the power law accurately predicts fitness to 50,000 generations using data from only the first 5,000 generations. It was shown that a population-dynamical model that incorporates two phenomena known to be important in the LTEE—clonal interference<sup>29,31</sup> and diminishing-returns epistasis<sup>15,29</sup>—generates a power-law relation. This model in turn predicts that the number of beneficial mutations should increase with the square root of time<sup>17</sup>. However, not all mutations that accumulate are beneficial; neutral and nearly neutral mutations can spread by recurring mutation, random drift, and hitchhiking<sup>32–34</sup>. Selective sweeps will purge some neutral



**Figure 2 | Phylogenetic trees for LTEE populations. a**, Phylogenies for 22 genomes from each population, based on point mutations. **b**, The same trees, except branches are rescaled as follows: branches for lineages with mismatch-repair defects are orange and shortened by a factor of 25; branches for *mutT* mutators are red and shortened by a factor of 50. Strain REL606 (on the left) is the ancestor. No early mutations are shared between any populations, confirming their independent evolution. Most populations have multiple basal lineages that reflect early diversification and extinction; some have deeply divergent lineages with sustained persistence, most notably Ara-2.

mutations but cause others to increase; overall, the expected number of neutral mutations should increase linearly with time<sup>35</sup>.

To test these predictions, we fit three models to the trajectory for the total number of mutations in the non-mutator and premutator lineages:

$$m = at$$

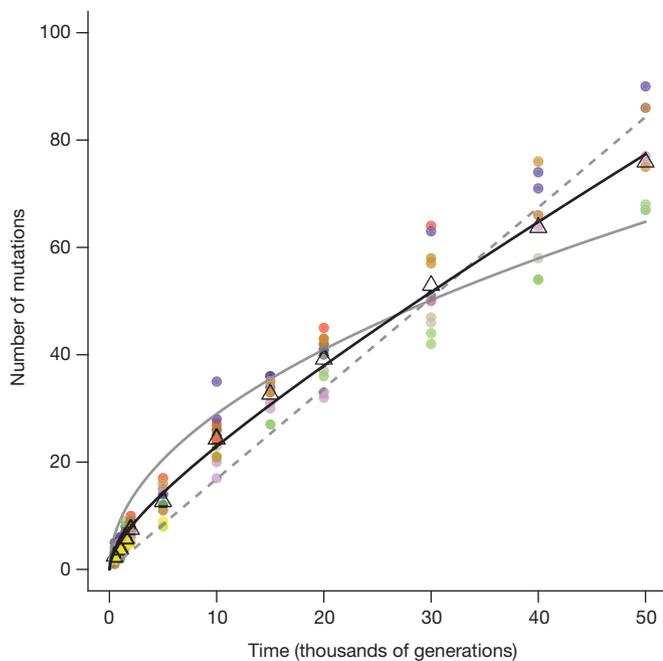
$$m = b\sqrt{t}$$

$$m = at + b\sqrt{t}$$

where  $m$  is the number of mutations,  $t$  is time (generations), and  $a$  and  $b$  govern the genome-wide rates of accumulation of neutral and beneficial mutations, respectively (Fig. 3). (Extended Data Fig. 3 shows the models fit to each population separately.) Using the Akaike information criterion (AIC), the two-parameter model fits the data much better than those with only the linear ( $\Delta\text{AIC} = -77.7$ ) or square-root ( $\Delta\text{AIC} = -99.7$ ) terms. Because the one-parameter models are nested within the two-parameter model, we can also assess the significance of adding the second parameter;  $P$  values are  $7.5 \times 10^{-5}$  and  $5.2 \times 10^{-7}$  relative to the linear and square-root models, respectively. The trajectory for genome evolution thus shows signatures of both adaptive and non-adaptive changes. However, the model that predicts the square-root trajectory of beneficial substitutions makes various assumptions (for example, about the form of epistasis), and both the predicted and observed trajectories have statistical uncertainties. (Extended Data Fig. 4 shows the uncertainty in estimating  $a$  and  $b$  from the observed trajectory.) Therefore, we examined additional evidence to shed light on the proportion and identity of beneficial mutations.

### Evidence for beneficial mutations

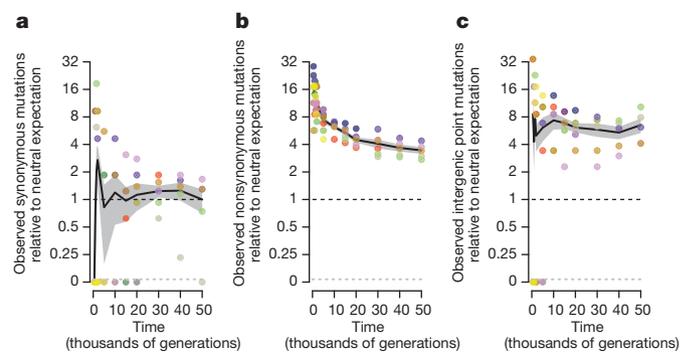
We sought to understand what proportion of the genomic changes in the non-mutator populations was adaptive, and how that proportion changed over time. One line of evidence derives from the expectation that synonymous substitutions—point mutations in protein-coding genes that do not affect the amino-acid sequence—are neutral and should therefore accumulate at a rate equal to the underlying



**Figure 3 | Alternative models fit to the trajectory of genome evolution.** Each symbol shows total mutations in a clone from five populations that never became mutators and seven before point mutation or IS150 hypermutability evolved. Colours are the same as in Fig. 1; open triangles indicate grand means. Dashed grey line shows the best fit to the linear model,  $m = at$ . Solid grey curve shows the fit to the square-root model,  $m = b\sqrt{t}$ . Black curve is fit to the composite model,  $m = at + b\sqrt{t}$ , where  $a = 0.000944$  and  $b = 0.134856$ . See text for statistical analysis.

mutation rate<sup>20,35</sup>. This expectation is not strictly true owing to selection on codon usage, RNA folding, and other effects, but it is generally thought that such selection is extremely weak, affects only a small fraction of sites at risk for synonymous mutations, or both<sup>36,37</sup>. We calculate whether nonsynonymous and intergenic point mutations are found in excess relative to synonymous mutations, given the number of sites at risk for each class. Figure 4a shows the number of synonymous mutations in non-mutator and pre-mutator populations, scaled so the mean at 50,000 generations is unity. As expected, synonymous mutations accumulated at an approximately constant rate (Extended Data Fig. 5). Figure 4b shows the number of nonsynonymous mutations relative to the neutral expectation based on synonymous mutations. Nonsynonymous mutations accumulated  $\sim 17.1$  times faster than synonymous ones during the first 500 generations and  $\sim 3.4$  times faster over 50,000 generations. Nonsynonymous mutations continued to accumulate at over twice the rate of synonymous mutations in the later generations (Extended Data Fig. 6), implying that most nonsynonymous mutations that reached high frequency were beneficial even after so long in a constant environment. The same approach applied to intergenic point mutations (Fig. 4c) also reveals a large excess relative to synonymous mutations, although the number of events is smaller and the uncertainty greater. This result implicates adaptive changes in noncoding regions that presumably affect the binding sites for regulatory proteins<sup>38–40</sup>.

Synonymous mutations provide an internal benchmark for nonsynonymous and intergenic point mutations. However, synonymous mutations are not directly informative for understanding how selection affects the accumulation of indels that comprise almost half the mutations in non-mutator clones at 50,000 generations (Extended Data Fig. 7). To estimate the proportion of beneficial changes for other types of mutation, we compare the LTEE and a mutation accumulation experiment (MAE) in which 15 lines were propagated via repeated single-cell bottlenecks<sup>41</sup>. Such bottlenecks eliminate the variation needed for natural selection, so that all types of mutations accumulate



**Figure 4 | Trajectories for synonymous, nonsynonymous and intergenic point mutations.** **a**, Synonymous mutations, scaled so that the mean of five non-mutator populations (excluding point mutation and IS150 hypermutators) is unity at 50,000 generations. **b**, Nonsynonymous mutations, scaled using the same rate as synonymous mutations after adjusting for sites at risk for both classes. **c**, Intergenic point mutations, scaled using the same rate as synonymous mutations after adjusting for sites at risk. Each symbol shows the mean for sequenced genomes from a non-mutator or pre-mutator lineage. Colours are as in Fig. 1. Note the discontinuous scale; populations with zero mutations are plotted below. Black lines connect grand means; shading shows standard errors calculated from replicate populations.

at the rates at which they happen, regardless of fitness effects, except for lethal or highly deleterious mutations that preclude cells from making colonies used to propagate lines<sup>29</sup>. MAE lines thus provide an external baseline for distinguishing beneficial and non-beneficial mutations. In fact, because more unselected mutations are deleterious than beneficial, MAE lines are expected to lose fitness over time, which they did (Extended Data Fig. 8).

To quantify the relative rates for all types of mutations in the absence of selection, we sequenced clones from the MAE lines after 550 daily bottlenecks (Supplementary Data 1). Consistent with the random accumulation of mutations, the number of nonsynonymous (including nonsense) mutations was similar to the expectation based on synonymous mutations (117 observed, 105.02 expected); the resulting ratio of 1.11 is well within the 95% confidence interval (0.70–1.50) obtained by a randomization test. Also, there was no among-line variation in total mutations ( $\chi^2 = 5.46$ , degrees of freedom (df) = 14,  $P = 0.978$ ). We can therefore reasonably use the MAE lines to estimate relative rates of different types of mutations, with synonymous ones providing a benchmark largely free of selection in both experiments. For example, LTEE population Ara–1 had 21 nonsynonymous mutations at 20,000 generations and the expected number of synonymous mutations based on the average non-mutator population was 1.08 (Extended Data Fig. 5); the 15 MAE lines in total had 117 nonsynonymous and 39 synonymous mutations; thus, the ratio of observed mutations to the neutral expectation is  $(21/1.08)/(117/39) = 6.5$ . These ratios show that all major classes of mutations—including various indels—are substantially overrepresented in the LTEE relative to the MAE (Extended Data Fig. 9), implying that many mutations in each class were adaptive during the LTEE.

### Parallel evolution at many gene loci

Parallel evolution occurs when similar changes arise independently in multiple lineages, and it is often used to discover putative targets of selection<sup>4,8,10–13,21</sup>. Genetic parallelism can be studied at the level of DNA sequence, affected genes, or integrated functions. Parallelism at the nucleotide level tends to be rare because different mutations in a gene often produce similar benefits<sup>4,10–12,21</sup>, although there are exceptions<sup>8</sup>. Parallelism at a functional level requires detailed understanding that may be unavailable, and it is difficult to interpret when there are many mutations. We therefore examined parallelism at the gene level.

**Table 1 | Protein-coding genes with the highest G scores**

Gene	Length	Observed	Expected	G	Annotation
<i>pykF</i>	1,413	19	0.16	181	Pyruvate kinase
<i>iclR</i>	825	13	0.10	128	Transcriptional repressor, glyoxylate bypass
<i>spoT</i>	2,109	14	0.25	113	Stringent response
<i>nadR</i>	1,233	12	0.14	106	Bifunctional transcriptional repressor and NMN adenylyltransferase
<i>hslU</i>	1,332	11	0.15	94	Molecular chaperone and ATPase component of protease
<i>yijC</i> (also known as <i>fabR</i> )	705	7	0.08	62	Transcriptional repressor, fatty acid and phosphatidic acid pathway
<i>topA</i>	2,598	8	0.30	52	DNA topoisomerase I subunit
<i>malT</i>	2,706	8	0.31	52	Transcriptional activator, maltotriose-ATP-binding
<i>mrda</i>	1,902	7	0.22	48	Transpeptidase in peptidoglycan synthesis
<i>mreB</i>	1,044	6	0.12	47	Longitudinal peptidoglycan synthesis
<i>infB</i>	2,673	7	0.31	44	Translation initiation factor IF-2
<i>arcA</i>	717	5	0.08	41	Response regulator in two-component system, anoxic redox control
<i>argR</i>	471	4	0.05	34	Repressor of arginine regulon
<i>rplF</i>	534	4	0.06	33	50S ribosomal subunit protein
<i>mreC</i>	1,104	4	0.13	28	Longitudinal peptidoglycan synthesis

Genes are ranked by G scores computed using observed independent nonsynonymous mutations relative to expected number given gene length (bp). Data are from populations with the ancestral point-mutation rate throughout and other populations before they evolved hypermutability.

We focused on lineages that retained the ancestral point-mutation rate (including clones from populations that later became hypermutable) because, as shown earlier, most mutations are drivers in those cases; we expect hypermutability to make the analysis less informative because many more mutations are passengers. We first calculated the expected number of nonsynonymous mutations for each single-copy protein-coding gene based on its length as a fraction of all such genes and the total number of nonsynonymous mutations in the relevant lineages (Supplementary Data 2). We computed G scores for goodness of fit between observed and expected values; the total score is 2,593.7. We compared that total with simulated data sets in which positions of mutations in the coding genome were randomized, and the observed total significantly exceeded the simulations (mean simulated  $G = 1,933.7$ ,  $Z = 25.5$ ,  $P < 10^{-143}$ ). Fifty-seven genes had two or more mutations; these genes had 50.1% of the nonsynonymous mutations

but constituted only 2.1% of the coding genome. (Only one gene had multiple synonymous changes.) Table 1 shows the 15 genes that contribute the most to the total G score. Several encode proteins with core metabolic or regulatory functions, including three involved in peptidoglycan synthesis.

We ran the same analysis for lineages that evolved hypermutability (Supplementary Data 3), and the randomization test indicates significant parallelism ( $G$  statistic = 5,098.4, mean simulated  $G = 4,581.1$ ,  $Z = 5.745$ ,  $P < 10^{-8}$ ). As expected, however, the signal-to-noise ratio reflected in the significance level is much weaker than for the non-mutator lineages. Most genes with the highest scores in mutator lineages differ from those in non-mutators, in part because those genes often had beneficial mutations before hypermutability evolved.

Table 2 lists the 16 genes with the most deletions, duplications, insertions and intergenic point mutations in non-mutator lineages

**Table 2 | Genes with the most mutations of other types**

Genes	Mutations	Number	IS	MAE	Annotation
<i>rhsD</i>	Mostly large deletions	41	Yes	No	D-Ribose utilization; most deletions affect entire <i>rhs</i> operon
<i>nupC</i>	Various intergenic	19	Yes	Yes	Nucleoside transporter
<i>iap</i>	Mostly large indels	19	Yes	No	Alkaline-phosphatase isozyme conversion; most indels affect tens of adjacent genes including <i>rpoS</i> , which encodes stationary-phase $\sigma$ factor
<i>mokB</i>	Various indels	17	Yes	Yes	Enables <i>hokB</i> toxin expression
<i>yhgI/gntT</i>	Intergenic point mutations	16	No	No	Gluconate transport
<i>mokC</i>	Various indels	15	Yes	Yes	Enables <i>hokC</i> toxin expression
<i>ybcU</i> (also known as <i>borD</i> )	Large indels	14	Yes	No	Indels affect this and adjacent remnants of DLP12 prophage
ECB_02013	Various indels	14	No	Yes	Indels affect this and adjacent remnants of P2-like prophage
ECB_02816 (also known as <i>kpsD</i> )	Various indels	14	Yes	No	Polysialic-acid transport protein precursor
<i>acs/nrfA</i>	Various intergenic	14	No	No	Acetyl-CoA synthase; nitrite reductase
<i>hokE</i>	Large indels	12	Yes	No	Toxin in plasmid-derived toxin-antitoxin system; most indels affect several adjacent genes involved in iron acquisition
<i>ybeB/phpB</i>	Various intergenic	11	Yes	No	Unknown functions, but adjacent to genes involved in cell-wall synthesis
<i>ydiJ/ydiK</i>	Various intergenic	11	No	No	Predicted FAD-linked oxidoreductase; putative inner membrane protein
<i>ldrC</i>	Various indels	10	Yes	Yes	Small toxic polypeptide
<i>menC</i>	IS insertions	10	Yes	Yes	Menaquinone biosynthesis
<i>fimA</i>	Mostly IS insertions	10	Yes	No	Component of fimbrial complex

Genes are ranked by total mutations excluding nonsynonymous and synonymous point mutations. When two genes are separated by a solidus, the affected sequence includes the intergenic region between them. IS column indicates whether the majority of mutations involve IS elements. MAE column indicates whether the same or nearly identical mutations occurred in one or more MAE lines. Data are from populations with the ancestral point-mutation rate throughout and others before they evolved hypermutability.

(Supplementary Data 2). For mutations that impact multiple genes, we show the most frequently affected gene (or adjacent pair when most events are intergenic). In 12 cases, the majority of the mutations were mediated by IS elements; these include insertions as well as deletions and duplications that appear to involve homologous recombination. In six cases (five with IS insertions), the same or nearly identical mutations occurred in one or more MAE lines, suggesting mutational hotspots. These changes may indicate high-frequency events, but recall that IS insertions and large indels are enriched in the LTEE relative to the MAE (Extended Data Fig. 9), implying that many are also beneficial. Indeed, the IS-mediated *rhsD* deletions occur at a high rate and are beneficial in the LTEE environment<sup>42</sup>, and some IS-mediated mutations appear to be beneficial in other studies as well<sup>43,44</sup>.

The parallelisms involving nonsynonymous substitutions and other mutations in the LTEE, coupled with their high rates of accumulation relative to the MAE, indicate that many observed mutations were drivers of adaptation. For indels, however, the specific target genes are difficult to identify owing to the multiplicity of genes affected and the potentially confounding effect of mutational hotspots.

## Discussion

Adaptation by natural selection sits at the heart of phenotypic evolution. However, the random processes of spontaneous mutation and genetic drift often overwhelm and obscure genomic signatures of adaptation. We overcame this difficulty by analysing genomes from 12 bacterial populations that evolved for 50,000 generations under identical culture conditions. Even so, six populations evolved hypermutable phenotypes that increased point-mutation rates ~100-fold, and another evolved hypermutability caused by a transposable element. By focusing on populations that retained the ancestral mutation rate, we identified several key features of the tempo and mode of their genome evolution. First, a population-genetic model with two terms—one for beneficial drivers, the other for neutral hitchhikers—fits the dynamics much better than models without both terms. Second, the great majority of mutations observed during the early generations were beneficial drivers. Third, the proportion of observed mutations that were beneficial declined over time but remained substantial even after 50,000 generations. The second and third findings follow from the population-genetic model. Both are also strongly supported by the excess of nonsynonymous to synonymous substitutions in the LTEE and by the excess of several classes of mutations, including indels, in comparison to mutation-accumulation lines. Fourth, there was strong gene-level parallel evolution across the replicate LTEE populations.

Our analyses also show a contrast between the contributions of beneficial mutations to molecular evolution and to the fitness trajectory in a stable environment. In particular, beneficial mutations continued to constitute a large fraction of genetic changes throughout the 50,000 generations of the LTEE, whereas the resulting fitness gains were only a few per cent in the last 10,000 generations<sup>17</sup>. Beneficial mutations with very small selection coefficients are nonetheless visible to natural selection<sup>17</sup>. Hence, adaptation can remain a major driver of molecular evolution long after an environmental shift. Our experimental results thus support a selectionist view of molecular evolution, complementing indirect evidence based on comparative genomics in bacteria, *Drosophila* and humans<sup>45–47</sup>. Of course, the LTEE may differ from many natural populations in important respects including its low mutation rate, the absence of sex or horizontal gene transfer, and a stable environment. As we showed, high mutation rates tend to obscure the role of selection in molecular evolution. The effects of horizontal gene transfer<sup>48</sup> and variable environments<sup>49,50</sup> on the dynamic coupling of genomic and adaptive evolution should also be examined further. Long-term experiments with microorganisms provide opportunities for rigorous analyses of these issues.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 January; accepted 23 June 2016.

Published online 1 August 2016.

- Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
- Hufford, M. B. *et al.* Comparative population genomics of maize domestication and improvement. *Nature Genet.* **44**, 808–811 (2012).
- Vonholdt, B. M. *et al.* Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**, 898–902 (2010).
- Lieberman, T. D. *et al.* Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genet.* **43**, 1275–1280 (2011).
- Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
- Lynch, M. & Conery, J. S. The origins of genome complexity. *Science* **302**, 1401–1404 (2003).
- Whitney, K. D. & Garland, T. Jr Did genetic drift drive increases in genome complexity? *PLoS Genet.* **6**, e1001080 (2010).
- Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. Different trajectories of parallel evolution during viral adaptation. *Science* **285**, 422–424 (1999).
- Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
- Tenaillon, O. *et al.* The molecular diversity of adaptive convergence. *Science* **335**, 457–461 (2012).
- Lang, G. I. *et al.* Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**, 571–574 (2013).
- Kvitek, D. J. & Sherlock, G. Whole genome, whole population sequencing reveals that loss of signaling networks is the major adaptive strategy in a constant environment. *PLoS Genet.* **9**, e1003972 (2013).
- Burke, M. K. *et al.* Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* **467**, 587–590 (2010).
- Levy, S. F. *et al.* Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* **519**, 181–186 (2015).
- Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E. & Cooper, T. F. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* **332**, 1193–1196 (2011).
- Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations. *Am. Nat.* **138**, 1315–1341 (1991).
- Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-term dynamics of adaptation in asexual populations. *Science* **342**, 1364–1367 (2013).
- Snigowski, P. D., Gerrish, P. J. & Lenski, R. E. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**, 703–705 (1997).
- Blount, Z. D., Barrick, J. E., Davidson, C. J. & Lenski, R. E. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* **489**, 513–518 (2012).
- Wielgoss, S. *et al.* Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc. Natl Acad. Sci. USA* **110**, 222–227 (2013).
- Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. & Lenski, R. E. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **103**, 9107–9112 (2006).
- Rozen, D. E. & Lenski, R. E. Long-term experimental evolution in *Escherichia coli*. VIII. Dynamics of a balanced polymorphism. *Am. Nat.* **155**, 24–35 (2000).
- Plucain, J. *et al.* Epistasis and allele specificity in the emergence of a stable polymorphism in *Escherichia coli*. *Science* **343**, 1366–1369 (2014).
- Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using *breseq*. *Methods Mol. Biol.* **1151**, 165–188 (2014).
- Barrick, J. E. *et al.* Identifying structural variation in haploid microbial genomes from short-read resequencing data using *breseq*. *BMC Genomics* **15**, 1039 (2014).
- Chao, L., Vargas, C., Spear, B. B. & Cox, E. C. Transposable elements as mutator genes in evolution. *Nature* **303**, 633–635 (1983).
- Tenaillon, O., Taddei, F., Radman, M. & Matic, I. Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Res. Microbiol.* **152**, 11–16 (2001).
- Gerrish, P. J. & Lenski, R. E. The fate of competing beneficial mutations in an asexual population. *Genetica* **102–103**, 127–144 (1998).
- Barrick, J. E. & Lenski, R. E. Genome dynamics during experimental evolution. *Nature Rev. Genet.* **14**, 827–839 (2013).
- Good, B. H. & Desai, M. M. Deleterious passengers in adapting populations. *Genetics* **198**, 1183–1208 (2014).
- Maddamsetti, R., Lenski, R. E. & Barrick, J. E. Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with *Escherichia coli*. *Genetics* **200**, 619–631 (2015).
- Gillespie, J. H. Genetic drift in an infinite population. The pseudohitchhiking model. *Genetics* **155**, 909–919 (2000).

33. Neher, R. A. & Shraiman, B. I. Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics* **188**, 975–996 (2011).
34. Kosheleva, K. & Desai, M. M. The dynamics of genetic draft in rapidly adapting populations. *Genetics* **195**, 1007–1025 (2013).
35. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
36. Sharp, P. M., Emery, L. R. & Zeng, K. Forces that influence the evolution of codon bias. *Phil. Trans. R. Soc. Lond. B* **365**, 1203–1212 (2010).
37. Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Rev. Genet.* **12**, 32–42 (2011).
38. Stern, D. L. Evolutionary developmental biology and the problem of variation. *Evolution* **54**, 1079–1091 (2000).
39. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
40. Oren, Y. *et al.* Transfer of noncoding DNA drives regulatory rewiring in bacteria. *Proc. Natl Acad. Sci. USA* **111**, 16112–16117 (2014).
41. Kibota, T. T. & Lynch, M. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature* **381**, 694–696 (1996).
42. Cooper, V. S., Schneider, D., Blot, M. & Lenski, R. E. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J. Bacteriol.* **183**, 2834–2841 (2001).
43. Miskinyte, M. *et al.* The genetic basis of *Escherichia coli* pathoadaptation to macrophages. *PLOS Pathog.* **9**, e1003802 (2013).
44. Wielgoss, S., Bergmiller, T., Bischofberger, A. M. & Hall, A. R. Adaptation to parasites and costs of parasite resistance in mutator and nonmutator bacteria. *Mol. Biol. Evol.* **33**, 770–782 (2016).
45. Charlesworth, J. & Eyre-Walker, A. The rate of adaptive evolution in enteric bacteria. *Mol. Biol. Evol.* **23**, 1348–1356 (2006).
46. Sawyer, S. A., Kulathinal, R. J., Bustamante, C. D. & Hartl, D. L. Bayesian analysis suggests that most amino acid replacements in *Drosophila* are driven by positive selection. *J. Mol. Evol.* **57** (suppl. 1), S154–S164 (2003).
47. Bustamante, C. D. *et al.* Natural selection on protein-coding genes in the human genome. *Nature* **437**, 1153–1157 (2005).
48. Cooper, T. F. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS Biol.* **5**, e225 (2007).
49. Satterwhite, R. S. & Cooper, T. F. Constraints on adaptation of *Escherichia coli* to mixed-resource environments increase over time. *Evolution* **69**, 2067–2078 (2015).
50. Paterson, S. *et al.* Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank N. Hajela for assistance, R. Maddamsetti and Z. Blount for discussions, and M. Lynch for starting the MAE lines. This research was supported by the US National Science Foundation (DEB-1451740 to R.E.L.), BEACON Center for the Study of Evolution in Action (DBI-0939454), European Research Council (FP7 grant 310944 to O.T.), European Union (FP7 grant 610427 to D.S.), French National Funding Agency (ANR-08-GENM-023-001 to D.S., O.T. and C.M.), French CNRS International Associated Laboratory (to D.S. and R.E.L.), and US National Institutes of Health (R00-GM087550 to J.E.B.). D.E.D. was supported by a traineeship from the Cancer Prevention and Research Institute of Texas. We acknowledge the use of high-performance computing resources at the Texas Advanced Computing Center.

**Author Contributions** O.T., J.E.B., D.S. and R.E.L. conceived the project; R.E.L. and J.L.B. provided strains; O.T., J.E.B., D.E.D., A.D., G.C.W., S.W., S.C. and C.M. analysed genomes and generated other data; N.R. developed theory; R.E.L., O.T. and J.E.B. wrote the paper. All authors approved the submitted version.

**Author Information** All sequencing data sets are available in the NCBI BioProject database under accession number PRJNA294072. The *breseq* analysis pipeline is available at GitHub (<http://github.com/barricklab/breseq>). Other analysis scripts are available at the Dryad Digital Repository (<http://dx.doi.org/10.5061/dryad.6226d>). R.E.L. will make strains available to qualified recipients, subject to a material transfer agreement. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.E.L. ([lenski@msu.edu](mailto:lenski@msu.edu)).

**Reviewer Information** *Nature* thanks M. Desai, G. Sherlock and C. Zeyl for their contribution to the peer review of this work.

## METHODS

**Long-term evolution experiment.** The LTEE has 12 populations founded from two almost identical strains of *Escherichia coli*. Six populations, designated Ara–1 to Ara–6, started from REL606, a descendant of the B strain of Luria and Delbrück<sup>51–53</sup>. The other six, Ara+1 to Ara+6, derive from REL607, which differs from REL606 by point mutations in *araA* and *recD*. The mutation in *araA* was selected before starting the LTEE; it confers the ability to grow on L-arabinose, which provides a marker in competition assays used to measure fitness<sup>16,17</sup>. The *recD* mutation arose inadvertently before starting the LTEE. The LTEE began in 1988, and the populations have been propagated (with occasional interruptions) at 37°C by daily 100-fold dilutions in 10 ml Davis minimal medium with 25 µg/ml glucose (<http://lenski.mmg.msu.edu/ecoli/dm25liquid.html>). The regrowth allows ~6.67 generations per day; the population size fluctuates between  $\sim 3 \times 10^6$  and  $\sim 3 \times 10^8$  cells except in population Ara–3, which has had a population size several times larger since ~33,000 generations, when cells gained the ability to consume the citrate that is also present in the medium<sup>19,54</sup>. Whole-population samples are taken every 75th transfer (500 generations) and stored with glycerol as a cryoprotectant at –80°C, where they are available for later analysis. Here we analysed the genomes of two clones sampled from each population at 500, 1,000, 1,500, 2,000, 5,000, 10,000, 15,000, 20,000, 30,000, 40,000 and 50,000 generations (Supplementary Data 1). We deliberately included clones from the deeply diverged lineages in population Ara–2 from 20,000 generations onwards and both the majority Cit<sup>+</sup> lineage and the minority Cit<sup>–</sup> lineage in population Ara–3 at generation 40,000. This sampling scheme does not affect inferences about the rates and patterns of genome evolution because both populations were hypermutable at these time points and thus excluded from the main analyses. These clones were included to illustrate diversity within populations, although we also found previously unknown cases of divergent lineages. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded during experiments and outcome assessment.

**Mutation-accumulation experiment.** The 15 MAE lines analysed here started from strain REL1207, which is an Ara<sup>+</sup> mutant of a clone sampled from LTEE population Ara–1 at 2,000 generations. REL1207 differs from REL606 by a total of eight mutations, including one in *araA* that confers the Ara<sup>+</sup> marker phenotype. Each line was propagated through 550 single-cell bottlenecks by picking a colony at random from a Davis minimal agar plate with glucose at 200 µg/ml and streaking the cells onto a fresh plate. Given ~25 cell doublings to produce a typical colony<sup>41</sup>, the 550 cycles represent ~13,750 generations. The bottlenecks imposed by this procedure eliminate the genetic variation that fuels adaptation by natural selection; as a consequence, mutations accumulate at rates that depend on their underlying mutation rate but not their fitness effects, except for highly deleterious mutations that preclude sufficient growth to form a colony<sup>29</sup>. Because more mutations are deleterious than are beneficial, fitness declined under this regime (Extended Data Fig. 8). The 15 sequenced clonal isolates, each from a different MAE line, are JEB807–JEB821 (Supplementary Data 1). None of the lineages became hypermutable based on their mutational signatures and the absence of significant heterogeneity in the total mutations accumulated (see main text). However, the mean per-generation rate at which synonymous mutations arose was ~3.5-fold higher in the MAE lines than in the five LTEE populations that remained non-mutators for all 50,000 generations (Supplementary Data 4;  $t_s = 3.0755$ ,  $P = 0.0065$ ). This difference may reflect the different conditions in liquid and agar media, including the glucose concentration and local cell density, which might affect the reactive oxygen species that cells experience. The comparisons between the LTEE and MAE (Extended Data Fig. 9) would change if the underlying rates of the various types of mutation responded disproportionately to the different conditions in the MAE. That possibility seems implausible for the different classes of point mutation (Extended Data Fig. 9a, b), and the differences would have to be substantially larger than the different rates of synonymous mutations to produce the excess IS150 insertions (Extended Data Fig. 9c) and large indels (Extended Data Fig. 9f) observed in the LTEE relative to the MAE.

**Genome sequencing.** Frozen samples from the LTEE and MAE were revived via overnight growth at 37°C in either LB or Davis minimal medium supplemented with 1,000 µg/ml glucose. Genomic DNA was isolated from each culture using the Qiagen Genomic-tip 100/G kit or equivalent. The DNA samples were sequenced at Genoscope or Integragen SA (Évry, France), the Michigan State University Research Technology Support Facility (East Lansing, USA), or the University of Texas at Austin Genome Sequencing and Analysis Facility (Austin, USA). Illumina Genome Analyzer and HiSeq instruments were used to generate single-end or paired-end reads ranging in length from 35 to 150 bases according to standard procedures, with median coverage of 80-fold and 95-fold for the 264 LTEE and 15 MAE clones, respectively (Supplementary Data 1). Of the 264 LTEE genomes in this study, 40 were previously analysed in other studies<sup>9,19,20,55–57</sup>. Supplementary Data 4 shows the number of every type of mutation inferred after performing

the analyses described below on each of the LTEE and MAE genomes used in this study.

**Mutation calling.** We used *breseq* (versions 0.26.0 to 0.27.0) to predict both single-nucleotide and structural differences<sup>24,25</sup> based on how the Illumina reads for each sample mapped to the genome sequence of *E. coli* B REL606 (GenBank accession NC\_012967.1)<sup>52</sup>. We counted and classified mutations using an updated version of the REL606 reference genome with improved feature annotations. The updated genome file (in both GenBank and GFF3 formats) and lists of predicted mutations in each evolved genome (in the Genome Diff format described in an appendix to the *breseq* manual) are freely available online (<http://github.com/barricklab/LTEE-Ecoli>).

Most types of single-step mutations, including large deletions and transposition events leading to copies of IS elements at new positions in the genome, are directly predicted by *breseq* when they occur in non-repetitive genomic regions. The initial lists of predicted mutations were curated and refined as previously described<sup>24</sup>. Briefly, complex mutations involving multiple steps (such as a new IS insertion followed by a flanking deletion) and structural mutations that overlap repetitive regions of the genome were manually resolved from unassigned new junction and missing coverage evidence in the *breseq* output. Large duplications and amplifications were detected by examining the coverage depth of mapped reads across the reference genome and comparing this information with the positions of repeat sequences and unassigned junctions. Owing to limitations of short-read DNA sequencing data, we could not fully predict point mutations and indels of one to a few base pairs within repeat regions (for example, IS elements) or gene conversions, in which intragenomic recombination between nearly identical copies of a large repeat region (for example, the seven copies of the rRNA operon) converts a minor variation in one copy to match exactly the sequence of another copy. Instead, all such genetic changes in repetitive regions of the genome were uniformly ignored in downstream analyses, as described later.

To validate the final lists of mutations predicted in each clone, we applied these changes to the ancestral REL606 sequence and used *breseq* to compare the Illumina reads against this simulated evolved genome to verify there were no further, unexplained discrepancies. This step of applying mutations to the reference genome was also used to estimate the final genome size of each evolved clone, with the assumption that new IS insertions were of the most common size for that IS element in the reference genome.

For 6 of the 264 LTEE samples, there was evidence of non-clonality in the sequence data. Some samples appeared to be mixtures of two very closely related clones that shared nearly all mutations but had one to several mutations specific to each type, together adding to a frequency of 100% (for example, sets of mutations at frequencies of 35% and 65%). This situation might result from inadvertently sampling two adjacent colonies on an agar plate when picking clones from an LTEE population. In other cases, only one or two mutations were found at an intermediate frequency. This type of heterogeneity might arise from strong selection favouring new mutations during colony outgrowth, subculturing and revival of samples before DNA extraction, as these conditions differ from the LTEE. In each case, we reconstructed the major genotype in the sample, as noted in Supplementary Data 1.

We also ignored putative genome variation associated with a cryptic 186-like prophage element (REL606 genome coordinates 880528–904682). In ten of the LTEE populations, we observed clones with increased read-coverage depth of this region and reads spanning a new sequence junction consistent with either tandem head-to-tail amplifications of this region or the production of circular DNA molecules joined at these exact nucleotides. The changes in the apparent copy number of this region often deviated from the integer values expected for a stable duplication or amplification. The prophage-related changes in coverage appeared most often in genomes isolated from 2,000 generations or earlier in the LTEE. There is no evidence of infective phage production in the LTEE, but it is possible that replication of DNA encoding a defective phage occurs stochastically at some low level in the ancestral strain REL606 or that production of this DNA is induced by stress when culturing samples for DNA isolation.

**Phylogenetic consistency.** Owing to the long duration of the LTEE and the evolution of mutators in several lineages, some mutations may be hidden or initially grouped with other mutations into a single change when comparing a late-generation evolved genome with the ancestral genome. For example, a point mutation might occur early in the experiment and then the region containing that mutation is later deleted. Similarly, the deletion of one base early and the subsequent deletion of an adjacent base would be called as a single two-base deletion in later samples. To obtain more accurate counts in light of these issues, we used each population's inferred phylogeny to split or add mutations, as appropriate, so that the mutation list for each clone reflects the most parsimonious set of mutational steps between that clone and its ancestor. Specifically, we chose histories with the fewest total mutations, the fewest mutations on early branches (in case of ties), and the fewest

total nucleotide changes summed over all mutations. Because this procedure is conservative in adding mutations to achieve phylogenetic consistency, it might underestimate the number of mutations on branches leading to an evolved genome when intermediate states are not resolved by the relationships of the sequenced clones.

**Final mutation lists.** We performed two final filtering steps to enable the sets of mutations to be uniformly compared across all genomes. In doing so, we classified as 'small mutations' all single-nucleotide substitutions, insertions and deletions of 20 or fewer bp, substitutions replacing 20 or fewer bp in the reference genome with 20 or fewer other bp, and all simple sequence repeat (SSR) mutations regardless of their size. SSR mutations add or remove one or more copies of a tandem-repeat unit consisting of one or a few bp. We defined SSR mutations as containing at least two copies of the repeat unit and having a total length of at least five bp when including all copies of the tandem repeat in the reference genome. For example, the genetic changes GGGGG→GGGG, TATATA→TATATATA and TACGTTACGT→TACGT would all be classified as SSR mutations, but GGGG→GGGGG, TATA→TATATA and TACGT→TACGTTACGT would not. All other genomic changes were considered 'large mutations' for purposes of filtering.

The ability to call small mutations located in repetitive regions of the genome is dependent on read length, so we removed all such mutations in regions where it would be a problem to uniformly detect them from the mutation lists before further analyses. To do this, we enumerated all regions of  $\geq 20$  bp that had an exact match elsewhere in the genome of the ancestral strain REL606 using MUMmer v.3.23 (ref. 58). We then merged regions from this list that were separated by five or fewer bp. All resulting regions that were now  $\geq 35$  bp were included in a list of masked genomic intervals. We also added to this list a hypervariable SSR consisting of seven copies of a tetranucleotide sequence that could not be reliably called in data sets with short reads (coordinates 2103889–2103919). Any small mutations contained in these masked regions were excluded from all downstream analyses.

Finally, we flagged all nucleotide substitutions or small indels occurring within 20 bp of the end of an IS element. The sequences directly adjacent to IS elements appear to experience an unusually high mutation rate, possibly due to frequent transposase cleavage and DNA repair. Mutations at these IS-adjacent sites probably have no effect on cellular phenotypes and fitness. We excluded them from the final lists of mutations used in all further analyses because they could bias the inferred mutational spectra and rates.

**Phylogenetic analyses.** To produce the phylogenetic trees shown in Fig. 2, we used the point mutations associated with each clone. A minimum-evolution tree was built using the Jukes–Cantor one-parameter model<sup>59</sup>. We used this model for two reasons. First, the mutator lineages had very different mutational spectra from the non-mutators<sup>9,20,55,57</sup>. Second, many mutations seen in non-mutator lineages were under positive selection, and so it is appropriate to give the mutations equal weight and not, for instance, reduce the importance of transitions relative to transversions. The trees were plotted with the R package APE<sup>60</sup>. The composite tree has the star-like structure expected for independent evolution of the populations. Therefore, trees were made separately for each population and then combined in Fig. 2, which allowed multiple basal branches to be placed with the appropriate populations.

**Parallel evolution in non-mutator lineages.** For genomes that did not come from point-mutation hypermutator lineages (Supplementary Data 1), we examined the extent of parallelism at the gene level in two ways. The first approach was based only on nonsynonymous mutations, because it is straightforward to quantify the overall extent of parallelism, determine the statistical significance of the parallelism, and rank genes based on their contributions to the significance. For each protein-coding gene  $i$ , we know its length,  $L_i$ , and the number of independent nonsynonymous mutations observed in that gene across all clones from non-mutator and premutator lineages,  $N_i$ . We summed the lengths and relevant mutations over all single-copy protein-coding genes in the ancestral genome to obtain  $L_{\text{tot}}$  (3,920,306) and  $N_{\text{tot}}$  (457, including two mutations that each affected overlapping

reading frames), respectively. We computed the expected number of mutations in each gene,  $E_i$ , as follows:

$$E_i = N_{\text{tot}} (L_i/L_{\text{tot}})$$

We then computed a  $G_i$  score for each gene for which  $N_i > 0$  as follows:

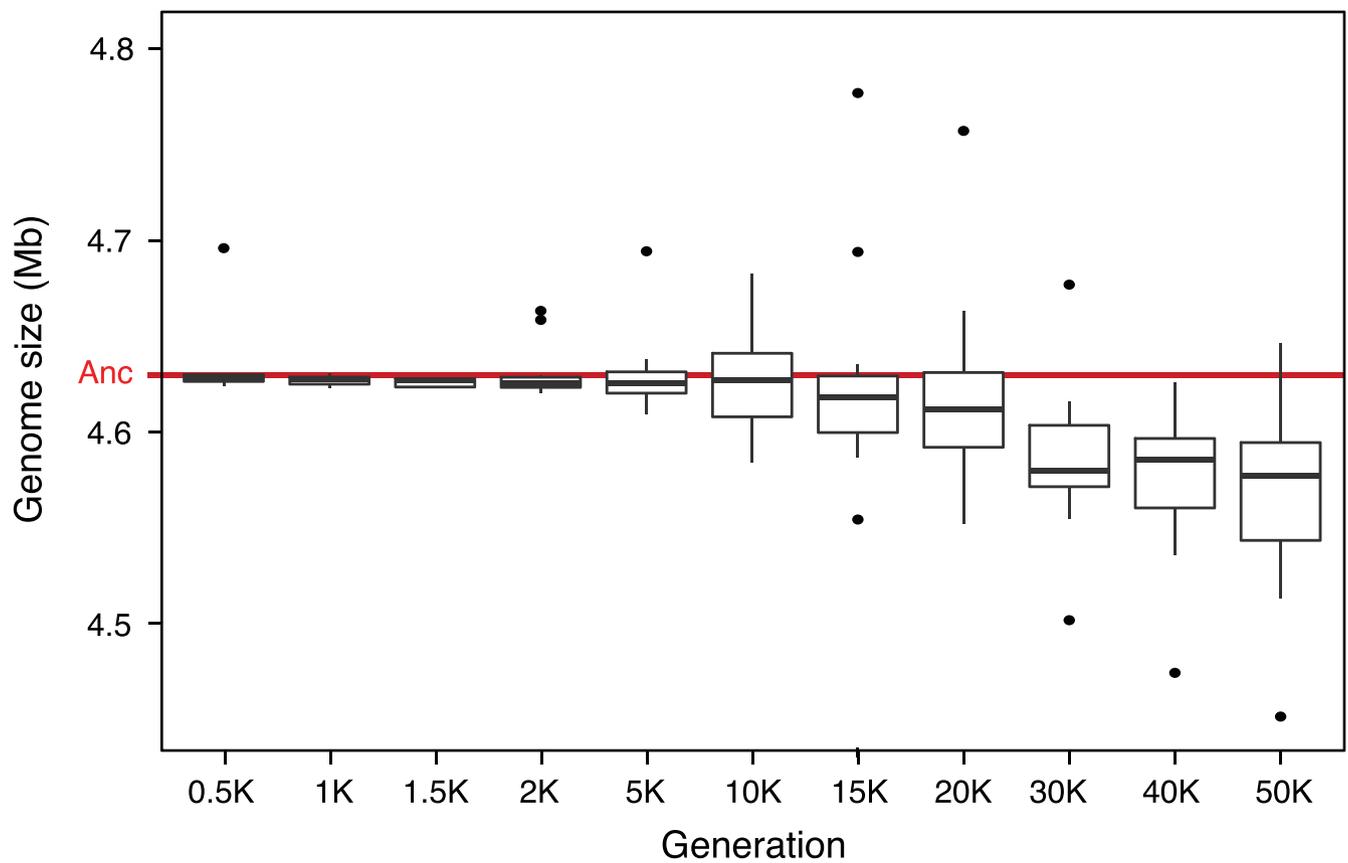
$$G_i = 2N_i \log_e(N_i/E_i)$$

We set  $G_i = 0$  for those genes for which  $N_i = 0$ . This analysis ignores variability among genes in the proportion of sites at risk for nonsynonymous mutations. However, such differences are small and should hardly affect the analysis. The total  $G$  statistic equals the sum of the scores over all genes. To compute the expected  $G$  statistic under the null hypothesis of a random distribution of mutations, we generated 1,000 simulated data sets in which  $N_{\text{tot}}$  mutations were randomly placed throughout the coding genome. We computed the total  $G$  statistic for each simulated data set, and we calculated its mean and standard deviation across the 1,000 simulations. To assess the significance of the observed  $G$  statistic, we computed the  $Z$  score as the difference between the observed and mean simulated values, divided by the standard deviation of the simulated values. Supplementary Data 2 lists each gene and the information used to calculate its  $G$  score. Table 1 shows the 15 genes with the highest  $G$  scores.

Supplementary Data 2 also shows other categories of mutation in or near each protein-coding gene including synonymous mutations, intergenic point mutations (between any particular gene and one of its immediately adjacent genes), IS insertions, small indels ( $\leq 50$  bp), large deletions ( $> 50$  bp) and long duplications ( $> 50$  bp). Table 2 shows the 16 genes that had the most total deletions, duplications, insertions and intergenic point mutations (that is, all mutations except synonymous and nonsynonymous mutations in the coding gene itself).

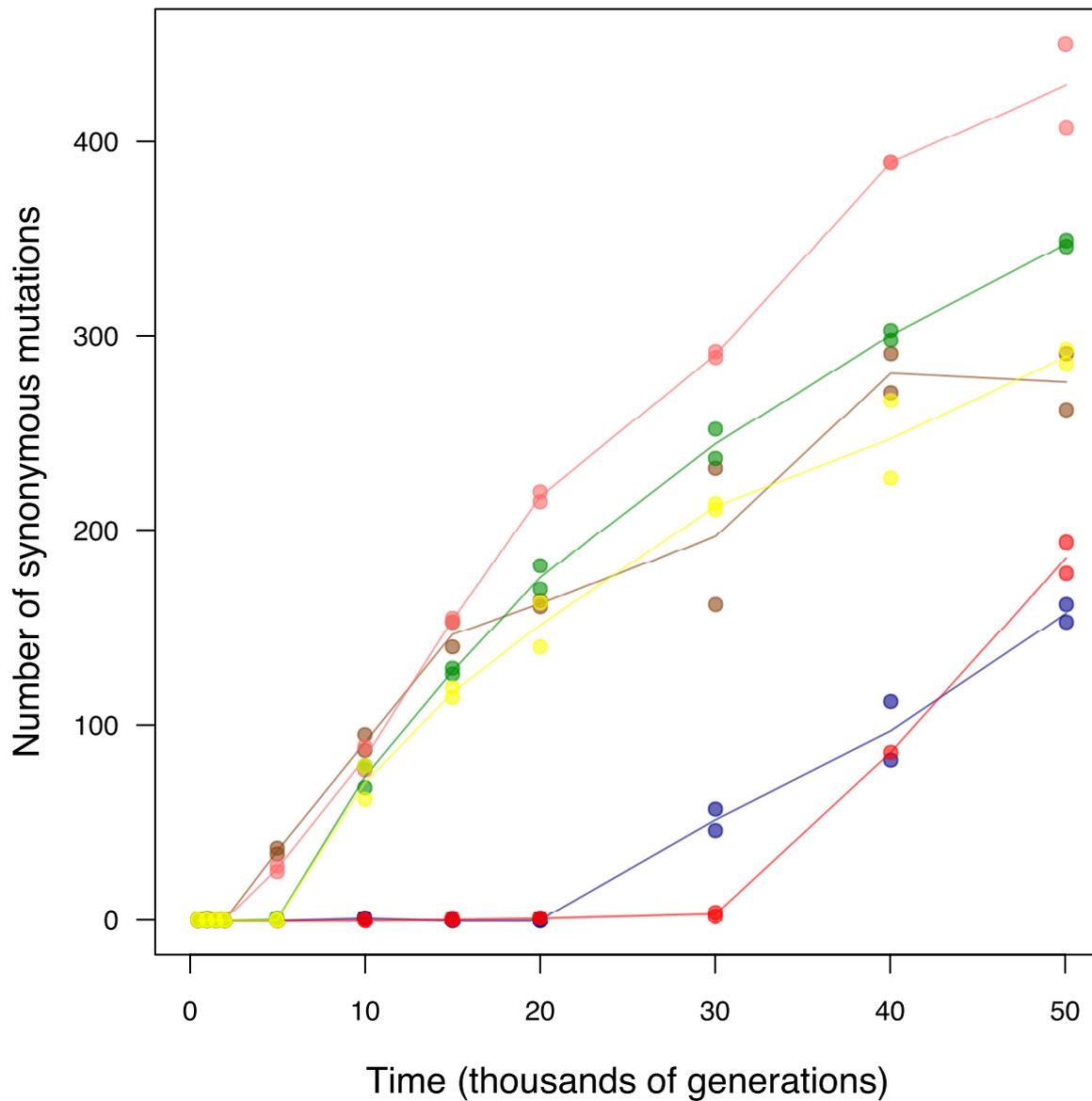
**Parallel evolution in mutator lineages.** We examined parallel changes in lineages that evolved point-mutation hypermutability by analysing nonsynonymous substitutions as above. To identify mutations that occurred after a lineage became hypermutable (Supplementary Data 3), we subtracted the mutations that occurred on non-mutator branches from the total mutations. This approach may result in a few mutations that arose before hypermutability being included in the counts for mutator lineages, but given the large increases in the point-mutation rate in the mutators (Fig. 1) it provides a reasonable approximation.

51. Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S. & Kim, J. F. Tracing ancestors and relatives of *Escherichia coli* B, and the derivation of B strains REL606 and BL21(DE3). *J. Mol. Biol.* **394**, 634–643 (2009).
52. Jeong, H. et al. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J. Mol. Biol.* **394**, 644–652 (2009).
53. Luria, S. E. & Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511 (1943).
54. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **105**, 7899–7906 (2008).
55. Wielgoss, S. et al. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. G3 (Bethesda) **1**, 183–186 (2011).
56. Raeside, C. et al. Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. *mBio* **5**, e01377–14 (2014).
57. Maddamsetti, R. et al. Synonymous genetic variation in natural isolates of *Escherichia coli* does not predict where synonymous substitutions occur in a long-term experiment. *Mol. Biol. Evol.* **32**, 2897–2904 (2015).
58. Kurtz, S. et al. Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12 (2004).
59. Desper, R. & Gascuel, O. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *J. Comput. Biol.* **9**, 687–705 (2002).
60. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).



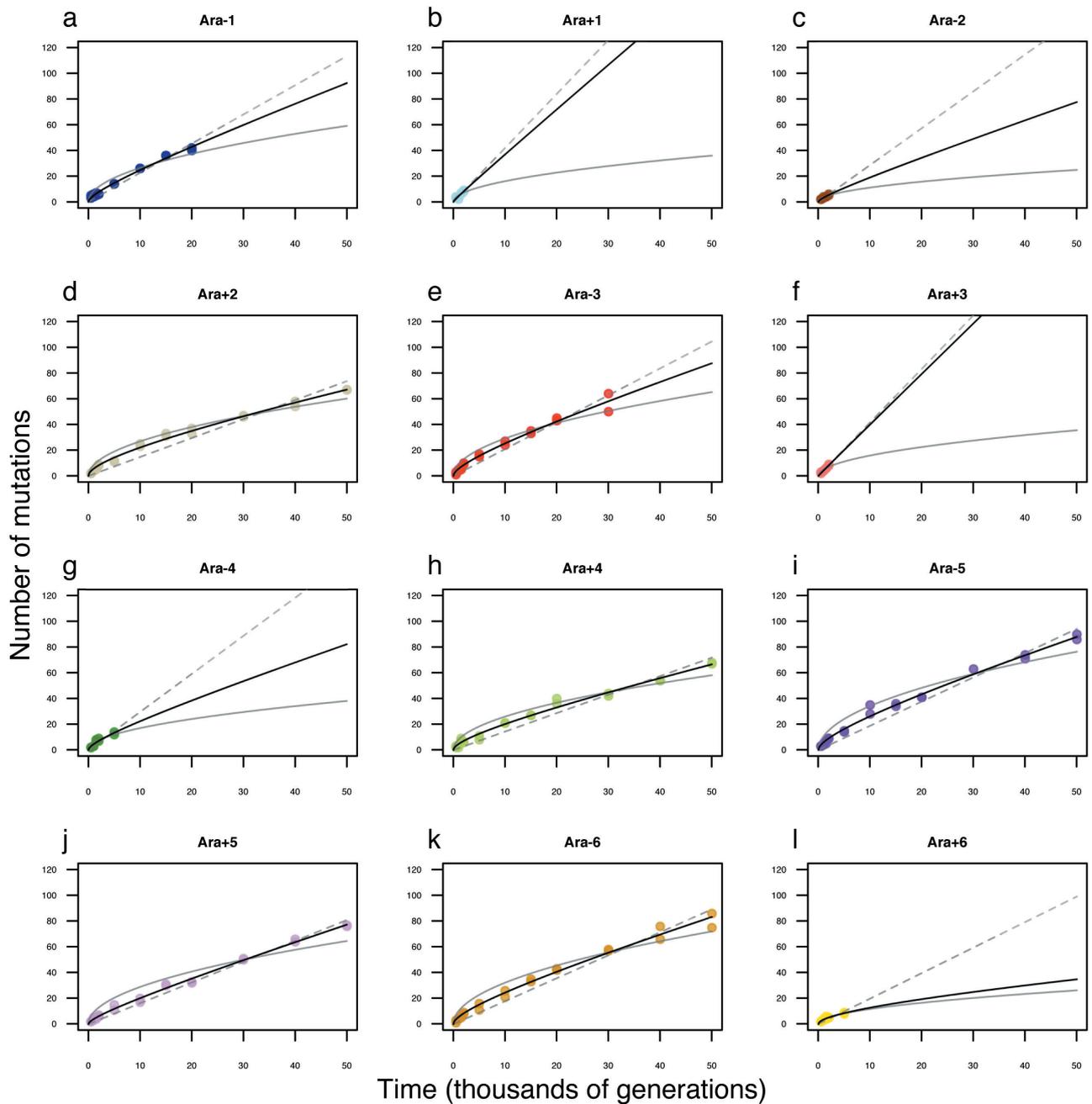
**Extended Data Figure 1 | Changes in genome size during the LTEE.** Box-and-whiskers plot showing the distribution of average genome length (Mb) for each of the 12 LTEE populations based on the two clones sequenced at each time point shown from 500 to 50,000 generations. The red line shows the length of the ancestral genome. The boxes are the

interquartile range (IQR), which spans the second and third quartiles of the data (25th to 75th percentiles); the thick black lines are medians; the whiskers extend to the outermost values that are within 1.5 times the IQR; and the points show all outlier values beyond the whiskers.



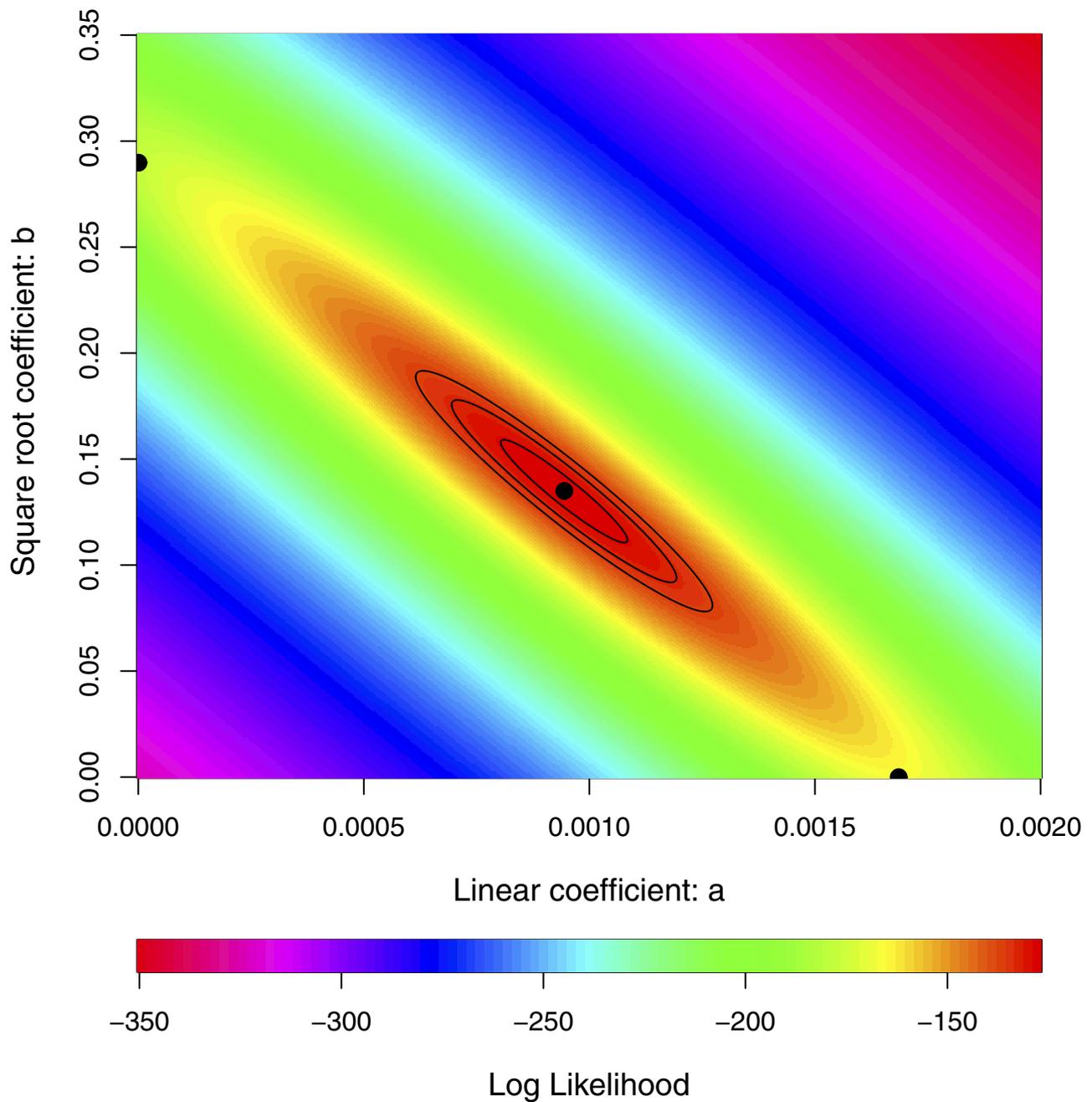
**Extended Data Figure 2 | Accumulation of synonymous mutations in populations that evolved point-mutation hypermutability.** Each symbol shows a sequenced genome from a hypermutable lineage. Colours are the same as those in Fig. 1. The accumulation of synonymous substitutions serves as a proxy for the underlying point-mutation rate. All four of

the populations that became hypermutable before 10,000 generations accumulated synonymous mutations at higher rates between 10,000 and 20,000 generations than between 40,000 and 50,000 generations, indicating the evolution of reduced mutability.



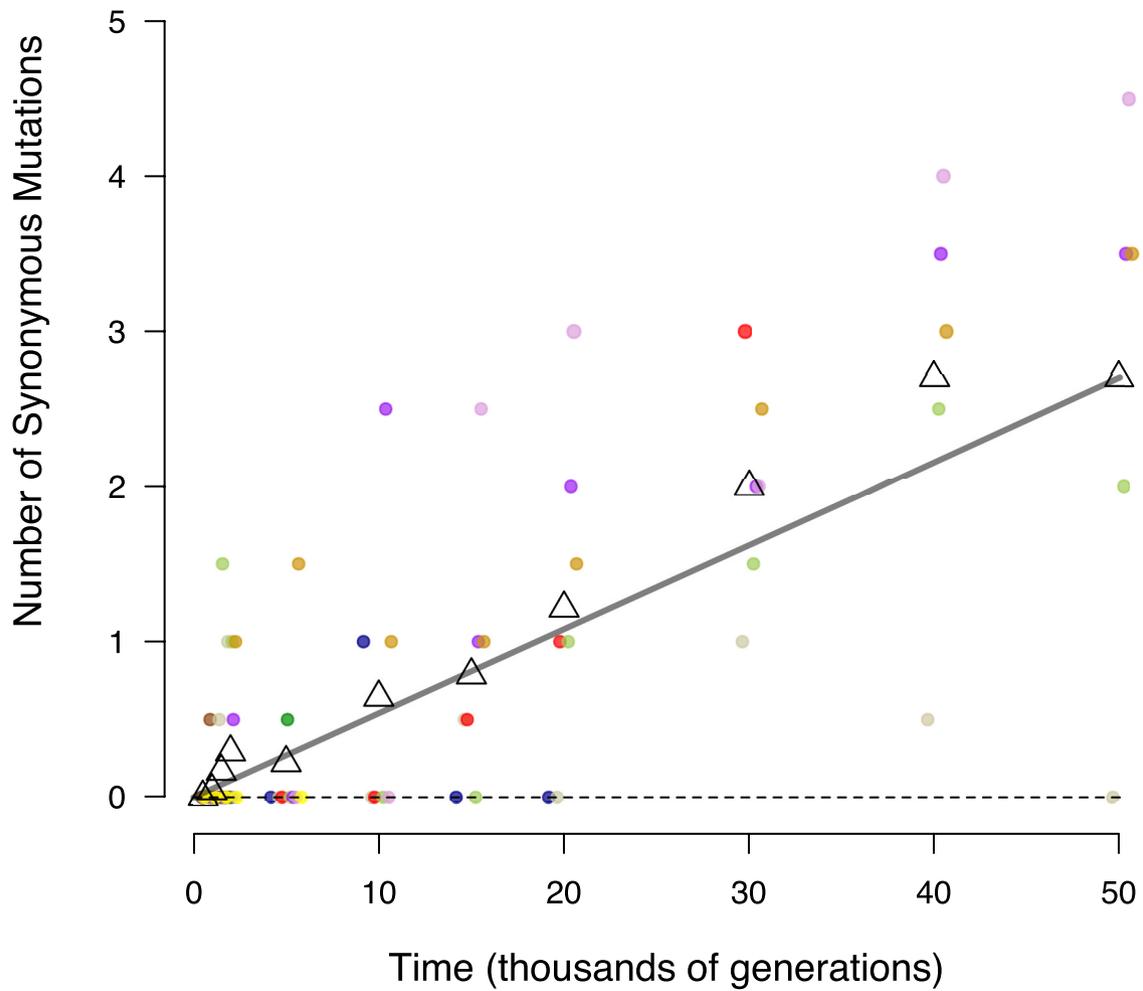
**Extended Data Figure 3 | Alternative models fit to trajectory of genome evolution for each LTEE population. a, Ara-1. b, Ara+1. c, Ara-2. d, Ara+2. e, Ara-3. f, Ara+3. g, Ara-4. h, Ara+4. i, Ara-5. j, Ara+5. k, Ara-6. l, Ara+6.** Each symbol shows the total mutations in a sequenced genome; in many cases, the symbols for the two genomes from the same population and generation are not distinguishable because they have

the same, or almost the same, number of mutations. For the populations that evolved hypermutability, data are shown only for time points before mutators arose. In each panel, the dashed grey line shows the best fit to the linear model; the solid grey curve shows the best fit to the square-root model; and the solid black curve shows the best fit to the composite model with both linear and square-root terms.



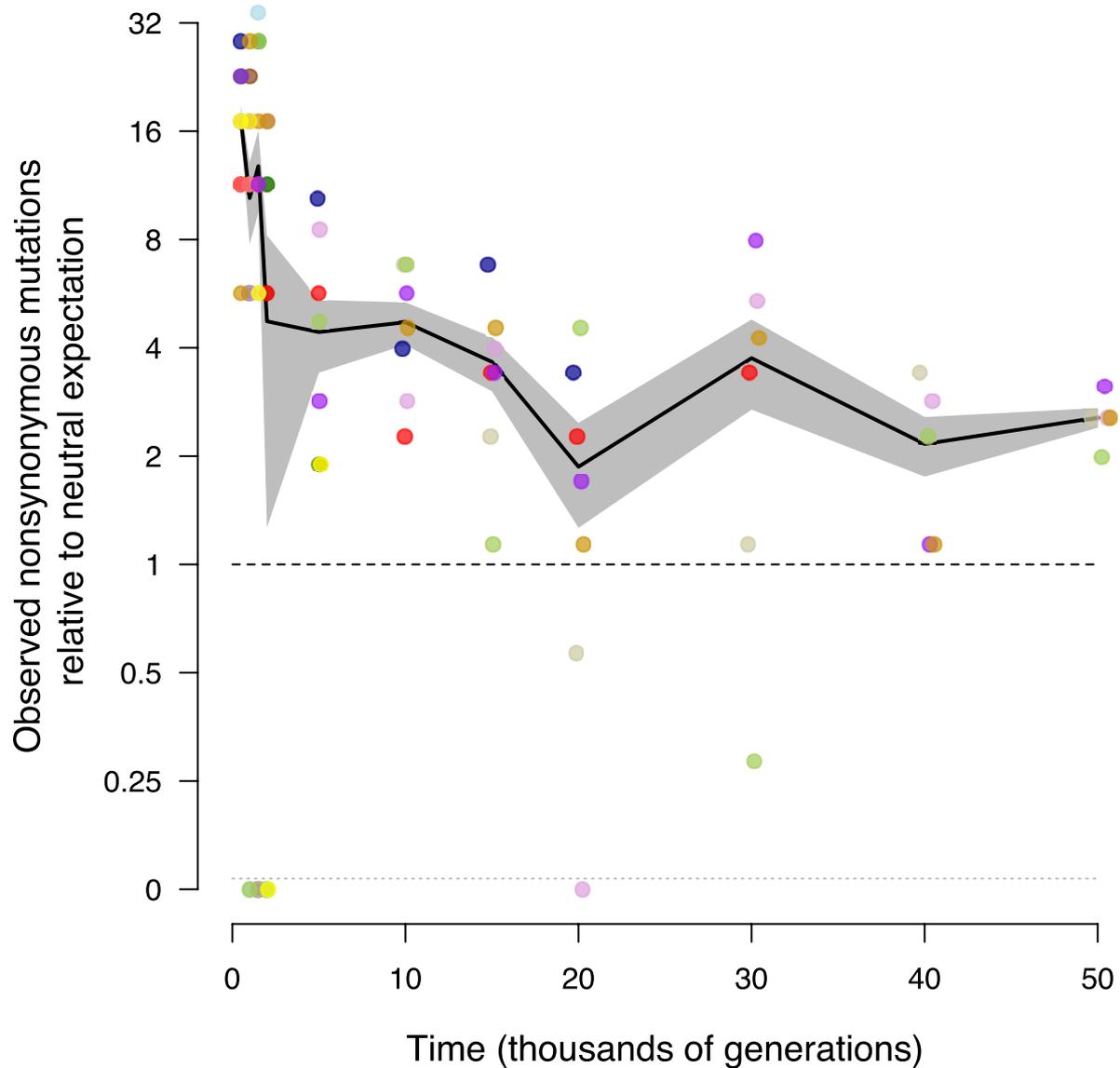
**Extended Data Figure 4 | Uncertainty in parameter estimation for the model describing the rates of accumulation for neutral and beneficial mutations.** Contours show relative likelihoods for simultaneously estimating the linear and square-root coefficients from the observed numbers of mutations that accumulated over time in non-mutator and

premutator lineages (Fig. 3). The black central point shows the maximum likelihood estimates, and the three black contours show solutions 2, 6 and 10 log units away. The points on the horizontal and vertical axes show values for the best one-parameter models.



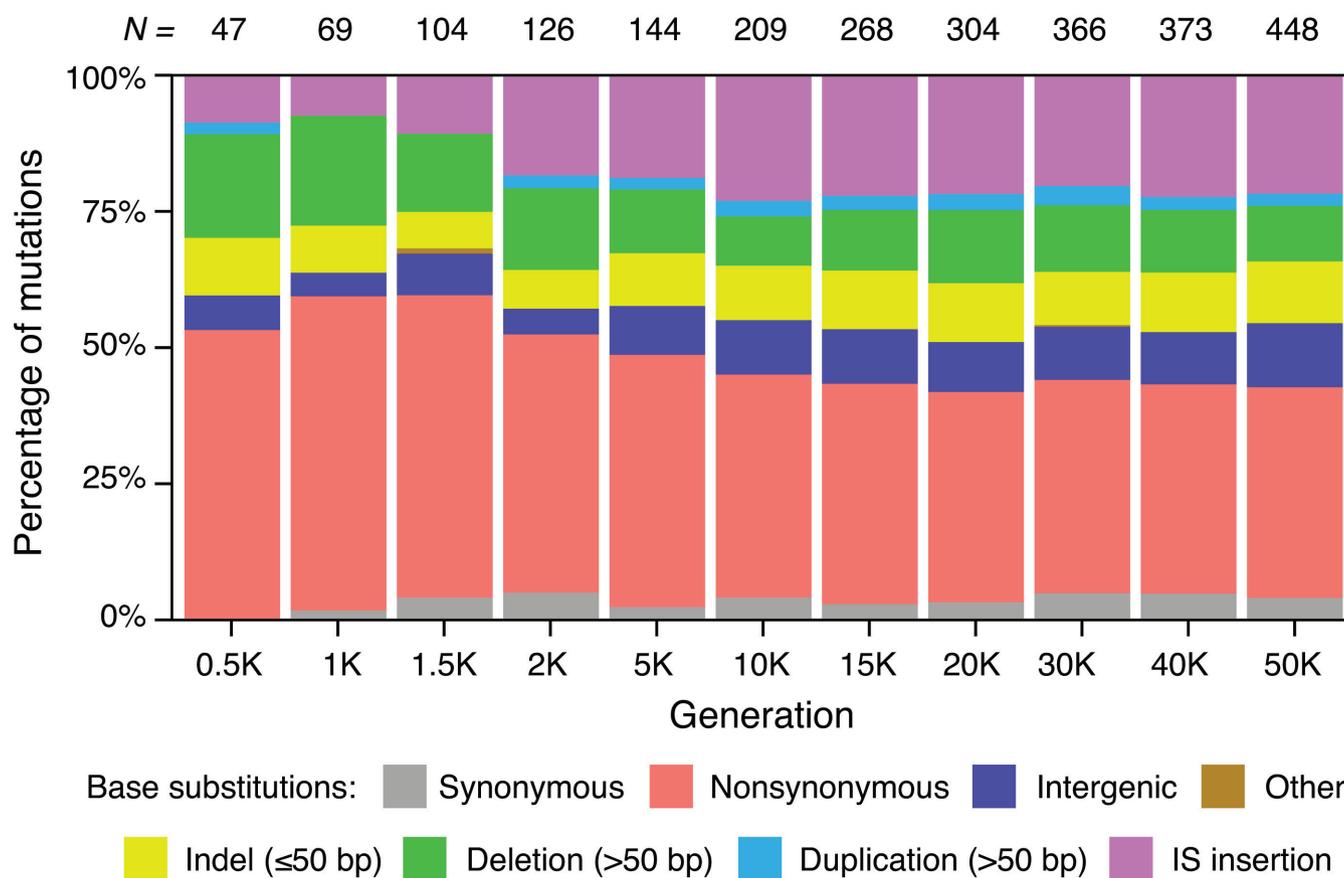
**Extended Data Figure 5 | Accumulation of synonymous substitutions in non-mutator lineages.** Each filled symbol shows the mean number of synonymous mutations in the (usually two) non-mutator genomes from an LTEE population that were sequenced at that time point; non-integer values can occur if the two genomes have different numbers.

Small horizontal offsets were added so that overlapping points are visible. Colours are the same as in Fig. 1. Open triangles show the grand means of the replicate populations. The grey line extends from the intercept to the final grand mean. The slope of that line was used to scale the relative rates of synonymous, nonsynonymous and intergenic point mutations in Fig. 4.



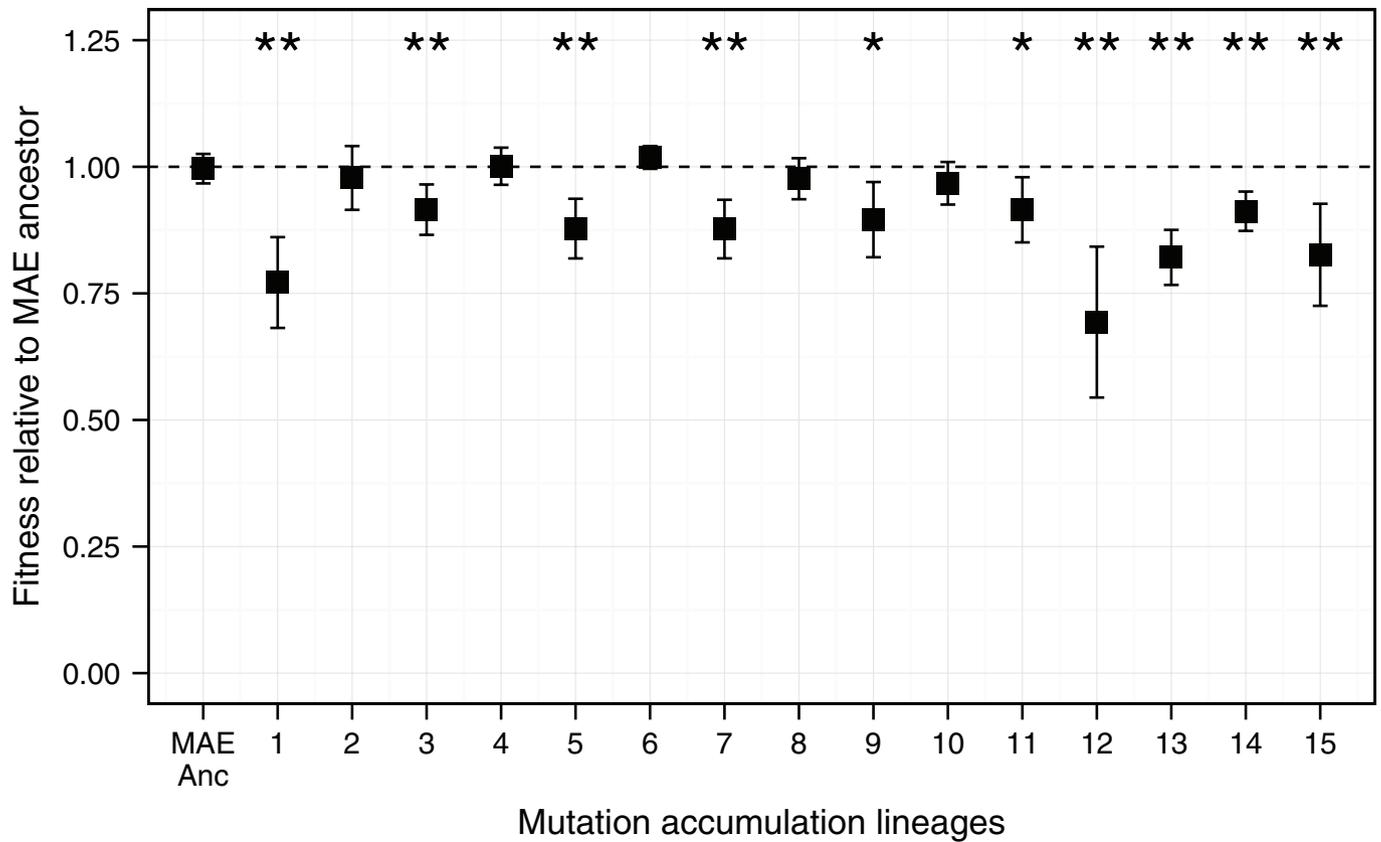
**Extended Data Figure 6 | Temporal trend in accumulation of nonsynonymous mutations relative to the neutral expectation in non-mutator lineages.** Interval-specific accumulation of nonsynonymous mutations calculated from changes in the total number of nonsynonymous mutations between successive samples. As with the cumulative data in Fig. 4b, values are scaled by the average rate of accumulation for synonymous mutations over 50,000 generations, after adjusting for the

numbers of genomic sites at risk for nonsynonymous and synonymous mutations. Each point shows the average rate calculated for a non-mutator or premutator population; small horizontal offsets were added so that overlapping points are visible. Note the discontinuous scale; populations with no additional mutations over an interval are plotted below. Colours are the same as in Fig. 1. Black lines connect grand means; the grey shading shows standard errors calculated from the replicate populations.



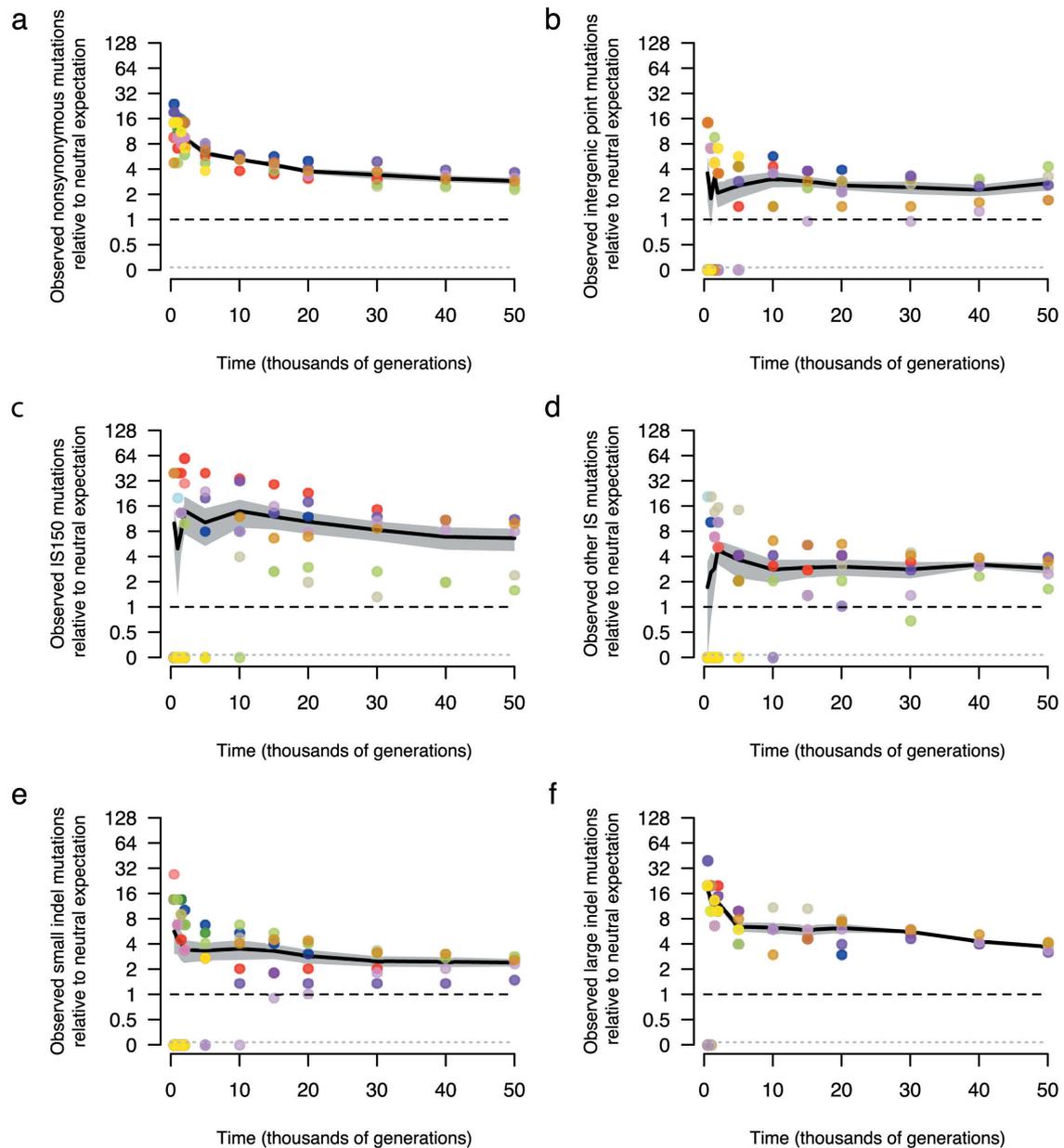
**Extended Data Figure 7 | Mutational spectrum for non-mutator lineages in the LTEE.** Shaded bars show the distribution of different types of genetic change for all independent mutations found in the set of non-mutator clones that were sequenced at each generation. The total number of mutations in this set at each time point (*N*) is shown above each

column. Base substitutions are divided into synonymous, nonsynonymous, intergenic, and other categories; the nonsynonymous category includes nonsense mutations, and the 'other' category includes rare point mutations in noncoding RNA genes and pseudogenes.



**Extended Data Figure 8 | Changes in fitness of MAE lines after 550 single-cell bottlenecks and ~13,750 generations.** Each point shows the mean fitness based on nine competition assays between the MAE ancestor (REL1207) or one of the 15 MAE lineages (JEB807–JEB821) and the Ara<sup>-</sup> variant of the MAE ancestor (REL1206). One-day competition

assays were performed using the standard procedures and same conditions as for the LTEE<sup>16,17</sup>. Error bars show 95% confidence intervals. \* $P < 0.05$ , \*\* $P < 0.01$ , based on two-tailed  $t$ -tests of the null hypothesis that relative fitness equals 1. Ten of the fifteen MAE lines experienced significant fitness declines, while none had significant gains.



**Extended Data Figure 9 | Trajectories for mutations by class in the LTEE in comparison with neutral expectations based on the MAE. a–f**, Accumulation of nonsynonymous mutations (a), intergenic point mutations (b), *IS150* insertions (c), all other IS-element insertions (d), small indels (e) and large indels (f). Colours are the same as in Fig. 1. All values are expressed relative to the rate at which synonymous mutations accumulated in non-mutator LTEE lineages over 50,000

generations (Fig. 4a), and then scaled by the ratio of the number of the indicated class of mutation relative to the number of synonymous mutations in the MAE lines. In all panels, each symbol shows a non-mutator or pre-mutator population. Note the discontinuous scale, in which populations with no mutations of the indicated type are plotted below. Black lines connect grand means over the replicate LTEE populations; the grey shading shows the corresponding standard errors.