

SINGLE-CELL ANALYSIS

Mutation dynamics and fitness effects followed in single cells

Lydia Robert,^{1,2*} Jean Ollion,¹ Jerome Robert,¹ Xiaohu Song,³ Ivan Matic,³ Marina Elez^{1,4*}

Mutations have been investigated for more than a century but remain difficult to observe directly in single cells, which limits the characterization of their dynamics and fitness effects. By combining microfluidics, time-lapse imaging, and a fluorescent tag of the mismatch repair system in *Escherichia coli*, we visualized the emergence of mutations in single cells, revealing Poissonian dynamics. Concomitantly, we tracked the growth and life span of single cells, accumulating ~20,000 mutations genome-wide over hundreds of generations. This analysis revealed that 1% of mutations were lethal; nonlethal mutations displayed a heavy-tailed distribution of fitness effects and were dominated by quasi-neutral mutations with an average cost of 0.3%. Our approach has enabled the investigation of single-cell individuality in mutation rate, mutation fitness costs, and mutation interactions.

The pace of evolution and possible trajectories depend on the dynamics of mutation incidence and the effects of mutations on fitness. Mutation dynamics has never been analyzed directly. It is assumed to be Poissonian, but mutagenesis characterized by bursts, with transient mutation rate increases, has been demonstrated in several organisms (1–3). Mutation bursts may facilitate the acquisition of combinations of mutations, which is critical to the evolution of complex traits (4). Effects of mutations are known to range from beneficial to deleterious or lethal, but estimating the entire distribution of fitness effects (DFE), central to evolutionary modeling, has proven challenging (5). The DFE for spontaneous mutations has been

estimated with the mutation accumulation (MA) approach (6–8), whereby independent lines of individuals are propagated in conditions that minimize the effects of natural selection and the evolution of their fitness is tracked. The precision of such DFE estimations is limited because of the small number of lines that can be monitored, typically fewer than 100. In addition, only an upper boundary for the average mutational effect is obtained, or strong assumptions are made regarding the DFE shape (9, 10). Finally, in MA studies of microorganisms, the mutation sample is biased, as natural selection purges lethal and strongly deleterious mutations.

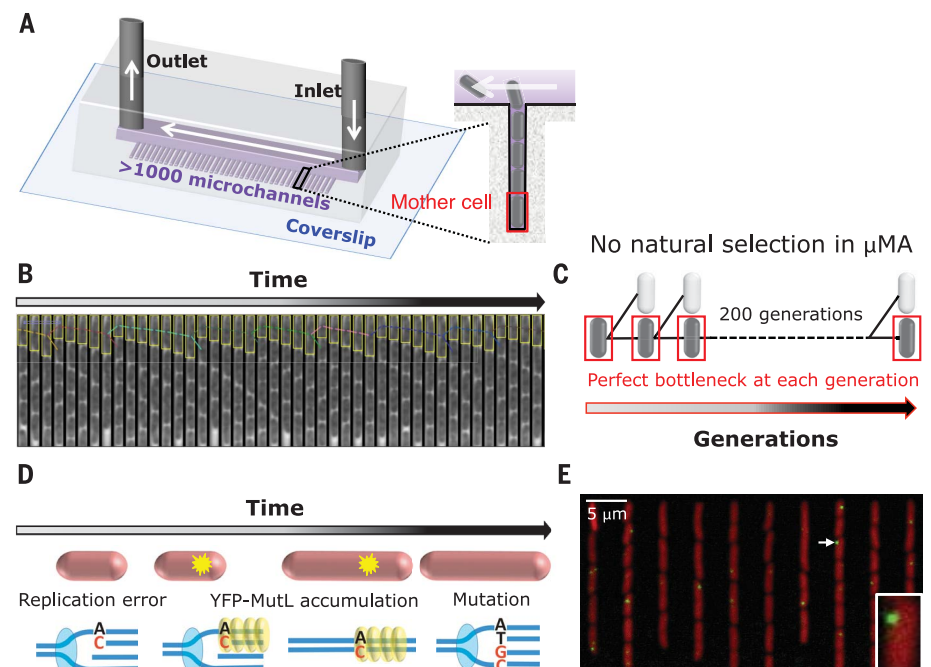
We quantitatively characterized the dynamics of spontaneous point mutations arising from

replication errors, a major source of mutations (11, 12), and their DFE in the bacterium *Escherichia coli*. For real-time detection of mutations in single living cells, we developed a mutation visualization (MV) experiment. We also designed a microfluidic MA (μ MA) experiment, which allowed us to track fitness in the complete absence of natural selection in a high-throughput fashion during the accumulation of mutations.

In μ MA experiments, we grew cells in a “mother machine” microfluidic chip (13) (Fig. 1A) and imaged them by phase-contrast microscopy to monitor their growth and survival at the single-cell level (14) (Fig. 1B and movie S1). The mother machine contained a series of separate microchannels where cells grew in a single row (Fig. 1A). These microchannels were closed on one side and retained the “mother” cell abutting the dead end through consecutive divisions. Therefore, at each generation, one individual was kept independently of its fitness (Fig. 1C), thus eliminating natural selection. Malthusian fitness, the exponential growth rate at the population level, can also be defined at the single-cell level as the cell exponential elongation rate. We imaged more than 1000 microchannels in parallel, in controlled and constant conditions, every 4 min

¹Laboratoire Jean Perrin, UMR 8237 Sorbonne Universités, UPMC Université Paris 06, Paris, France. ²Micalis Institute, Institut National de la Recherche Agronomique, AgroParisTech, Université Paris Saclay, Jouy-en-Josas, France. ³INSERM U1001, Université Paris-Descartes, Sorbonne Paris Cité, Faculté de Médecine Paris Descartes, Paris, France. ⁴Institute of Systems and Synthetic Biology, UMR 8030, CNRS, Commissariat à l’Energie Atomique et aux Energies Alternatives, Genopole, Université d’Évry Val-d’Essonne, Université Paris Saclay, Evry, France. *Corresponding author. Email: lydia.robert@polytechnique.org (L.R.); marina.elez3@gmail.com (M.E.)

Fig. 1. Experimental setup. (A) The mother machine microfluidic chip, used for growing *E. coli* cells during μ MA and MV experiments. μ MA and MV experiments were performed independently, using phase-contrast microscopy and epifluorescence microscopy, respectively. (B) Phase-contrast images of one microchannel over time in a μ MA experiment. We developed software that measures the mother cell length (represented by yellow boxes) at each frame and computes its fitness at each generation (14). (C) The mother machine allowed us to keep the mother cell at each division, regardless of its fitness, thus blocking natural selection. (D) Tracking point mutation emergence in single cells with YFP-MutL in the *mutH* strain, where all replication errors are converted into mutations. (E) Overlay of red (from the constitutive expression of tdCherry for automated cell segmentation) and yellow (YFP-MutL) fluorescence images from an MV experiment. The inset shows a magnified image of a cell with a YFP-MutL focus (arrow).



over the course of 3 days, which corresponded to ~200 generations. We obtained ~10⁵ fitness measurements per experiment.

In MV experiments, cells grown in the mother machine were imaged by fluorescence microscopy (14). We visualized mutations using a fusion of yellow fluorescent protein (YFP) with MutL mismatch repair protein, which forms fluorescent foci at replication errors (15) (Fig. 1, D and E, and movie S2). To detect and track MutL foci, we developed an automated image analysis procedure (14). The dynamics of mutations originating from replication errors depends on both error production and repair. We examined the dynamics of error production by eliminating error repair through inactivation of the mismatch repair endonuclease *MutH* (14). Full inactivation was verified by control experiments (14). The YFP-MutL protein forms transient foci (15) (movie S2), and the average lifetime of the foci in the *mutH* strain corresponds to the time between the passage of two successive replication forks (14) (fig. S1). This pattern shows that in *mutH* cells, all MutL foci are converted into mutations by the next round of DNA replication.

Replication errors might occur at a constant rate with Poissonian dynamics, or alternatively, bursts of errors might be produced from fluctuations in intracellular composition, such as transient nucleotide pool imbalances (1, 3). To investigate these possibilities, we performed MV experiments using *mutH* cells, tracking more than 3000 MutL foci in mother cells and deriving their interarrival times (the times between two successive focus occurrences) (Fig. 2A). These interarrival times were exponentially distributed and uncorrelated, which is characteristic of a Poissonian process (14) (Fig. 2, B and C, and fig. S2). In agreement with this finding, the number of MutL foci appearing between two frames followed a Poissonian distribution (fig. S3). As our method is limited by optical resolution, we cannot detect bursts of replication errors if they occur at a single replication fork on a short genomic region (typically less than 80 kilobases). In contrast, we can detect transient hypermutator states if they affect several replication forks, as would be expected for metabolic fluctuations, because different forks occupy different positions in the cell.

Our data indicate Poissonian dynamics rather than dynamics characterized by bursts, but they do not exclude some fluctuations in the replication error rate. In particular, we expect the error rate to depend on the number of replication forks, which changes during the cell cycle (14). We found a global proportionality between cell size, a proxy for the stage in the cell cycle, and the replication error rate across a range of cell sizes (fig. S4). Further analysis of the relationship between cell size and error rate revealed steplike variations reminiscent of variations in the number of replication forks (Fig. 2, D and E), with error rates varying by a factor of ~3 during a single cell cycle.

When exposed to certain external stresses, bacteria trigger specific responses that transiently increase the mutation rate (16–18). In our

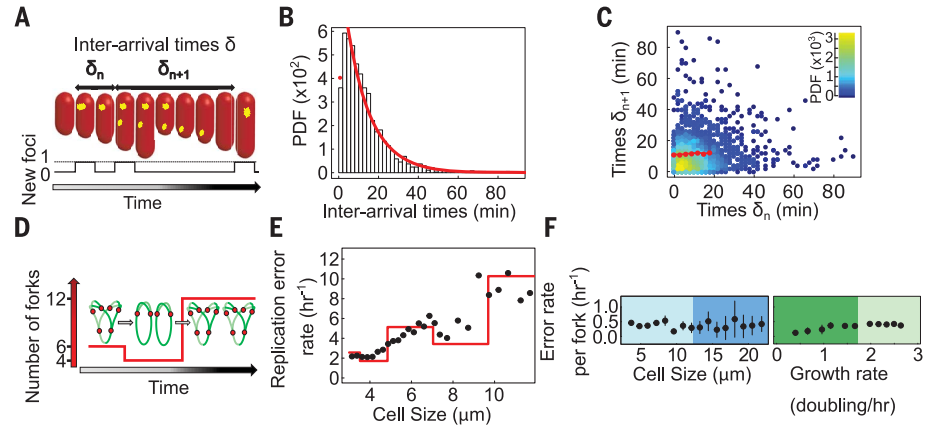


Fig. 2. Replication error dynamics. (A) The variables analyzed in MV experiments were the interarrival times for YFP-MutL foci (δ) and the number of new foci per mother cell. δ_n and δ_{n+1} indicate two successive interarrival times. (B) Histogram of focus interarrival times for a representative *mutH* experiment ($n = 1873$ foci). The red line and dot are predictions under the Poisson process hypothesis, which includes a correction accounting for discrete observation (14). PDF, probability density function. (C) Correlation between two successive focus interarrival times for the data in (B). Color indicates local density (inset). Red dots represent data binned according to δ_n . (D) Theoretical number of forks as a function of time for a doubling time of 26 min. Chromosomes are depicted in green, and the red dots denote replication forks. (E) Rate of replication error as a function of cell size in single cells. Black dots, experimental data; red line, prediction based on the number of forks (14). (F) Error rate per fork as a function of cell size and growth rate for single cells. Data were binned according to either cell size or growth rate. Results are means, and errors bars show ± 2 SEM.

MV experiments, a small fraction of cells experienced endogenous stress, as demonstrated by reduced growth rate and/or abnormally large cell size. We investigated whether such subpopulations exhibit higher rates of replication error. To take into account the different replication fork numbers for different subpopulations, we estimated the number of forks in single cells and the mean error rate per fork in each subpopulation (14). Abnormally large cells and slowly growing cells did not present elevated error rates per fork (Fig. 2F and fig. S5), suggesting that stress-induced mutagenesis (SIM) acts mainly at the level of error repair or that endogenous stress in a favorable environment causes limited SIM.

Visualization of mutations and fitness measurements could in principle be combined into a single time-lapse microscopy experiment. However, simply linking one mutation to one fitness effect is impossible for two reasons: First, there can be a substantial lag between the occurrence of a mutation and its effect on fitness. Second, a MutL focus indicates the insertion of an erroneous base in one of two newly replicated DNA strands. Consequently, the progeny from a cell with a MutL focus will contain a mixture of mutant and nonmutant cells (14) (fig. S6). Therefore, we performed μ MA and MV experiments separately, which allowed us to increase the throughput of μ MA experiments.

Using MV, we estimated mutation rates for *E. coli* wild-type (WT) (~0.0022 mutation/hour) and *mutH* mutant (~0.32 mutation/hour) strains (14), which agree with estimations from whole-

genome sequencing of MA lines of WT and mismatch repair-deficient cells (11). We then performed a μ MA experiment with WT cells, which accumulated an average of ~0.1 mutation per mother cell line over ~60 hours. As expected, the growth rate distribution for WT cells was stable over time (Fig. 3A, left). In contrast, for *mutH* cells, which accumulated ~20 mutations per mother cell line over ~60 hours, the mean growth rate slightly decreased (Fig. 3A, middle), and some cells underwent a steplike decrease in growth rate (Fig. 3B, blue trajectory) or a total cessation of growth, indicating cell death (fig. S7). Surprisingly, the average decrease in the growth rate of *mutH* cells during μ MA (~0.1% per hour) was smaller than what would be expected from previous estimates of the average effect of mutations (7, 19, 20). We also performed μ MA experiments using the strain MF1 (14), which has a mutation rate ~20 times that of the *mutH* strain (14) (table S1) because of impaired proofreading (21). As expected, the decrease in fitness in MF1 cells was ~20 times that in *mutH* cells (Fig. 3A, right, and table S2).

We then characterized the DFE by developing a nonparametric method, with no assumptions about DFE shape (14). Using the probabilistic framework of MA studies, we derived all the moments of the DFE (14). For nonlethal mutations, our estimate of mean fitness cost (Fig. 3C) is $0.31\% \pm 0.04\%$, which is lower than previous estimates (7, 19, 20). We found that the DFE for nonlethal mutations has a coefficient of variation (CV) of 9.5 ± 1.2 , a skewness of 16.6 ± 0.7 , and a kurtosis of 360 ± 90 (Fig. 3C). We also

obtained moments of higher order for the DFE (table S3). These results indicate a distribution that is wide, skewed, and heavy tailed, dominated by quasi-neutral mutations relative to infrequent, strongly deleterious mutations, such as the beta distribution in Fig. 3D.

The estimated DFE for MF1 was similar to the DFE for the *mutH* strain (table S4), although these strains have different mutation spectra (11, 21). To further investigate the effect of the mutation spectrum on the DFE, we performed μ MA experiments using a *mutT* mutant (fig. S8), which has a mutation rate similar to that of the *mutH* strain but a specific mutation spectrum of >99% AT-to-CG transversions (11, 12, 14) (table S1). Despite the differences in the way mutagenesis occurs, we found that the DFE for the *mutT* strain did not differ substantially from the DFE for the *mutH* or the MF1 strain (table S4).

Steeplike growth rate decreases such as those shown in Fig. 3B (blue trajectory) occurred during μ MA in all mutator strains and not in the WT strain. They might be provoked by a single deleterious mutation, independently of previously acquired mutations, or by the accumulation of interacting mutations. To investigate this issue, we studied the rate of occurrence of growth rate decreases in *mutH* and *mutT* cells. From the analysis of all single-cell trajectories (14) (fig. S9), we defined two categories of cells: slow-growing cells that underwent a >30% decrease in growth rate (Fig. 4A, top) and fast-growing cells corresponding to the rest of the population. The percentage of fast-growing cells decreased exponentially over time (14) (Fig. 4A, bottom), demonstrating that >30% growth rate decreases are caused by single deleterious mutations with >30% fitness cost. These mutations occur at a constant rate, independently of the accumulation of previous mutations, and represent 0.3% of all mutations. As expected from the quasi-optimal growth rate of our *mutH* strain, our analysis of single-cell trajectories detected no strongly beneficial mutations (>20% effect) during μ MA (14).

Likewise, to assess whether lethal mutations occur at a constant rate, we studied survival during μ MA (14). The WT strain, accumulating ~0.1 mutation over the course of μ MA, showed age-related mortality (Fig. 4B, left), in agreement with previous observations (13). The high-mutation-rate *mutH*, *mutT*, and MF1 strains exhibited additional mutation-related mortality (Fig. 4B, left). During the first ~35 hours of μ MA, when the mortality rate for the WT was low and approximately constant, the survival of the *mutH*, *mutT*, and MF1 mutator strains exponentially decreased (Fig. 4B, right, and fig. S10). To further disentangle age- and mutation-related mortality, we performed a μ MA experiment using the strain pMQ, which has increased proofreading deficiency compared with MF1 (21). pMQ cells died rapidly, and their survival decreased exponentially (Fig. 4B, right, and fig. S11). These analyses demonstrate that lethal mutations occur at a constant rate in μ MA, independently of previously acquired mutations, and that ~1% of mutations are lethal.

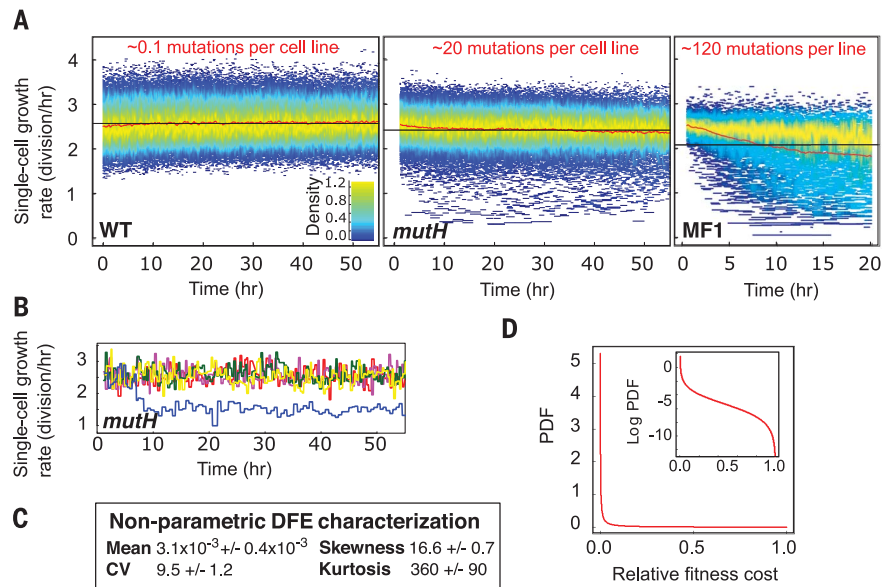


Fig. 3. μ MA analysis allows nonparametric DFE characterization. (A) Representative μ MA results for WT, *mutH*, and MF1 (*dnaQ926*) strains. For each time point, all single-cell growth rates for all mother cells are plotted (WT strain, $n = 1.8 \times 10^5$ cells; *mutH* strain, $n = 1.1 \times 10^5$ cells; MF1 strain, $n = 2.6 \times 10^4$ cells). Color depicts growth rate density. The red line represents the mean growth rate evolution. The black line indicates the mean of all growth rates. (B) Examples of single-cell growth rate trajectories for the *mutH* strain (one color per cell line). (C) Nonparametric estimation for the first four moments of the DFE allows calculation of the mean, the coefficient of variation (CV), skewness, and kurtosis (for each calculation, the mean ± 2 SEM for three independent *mutH* experiments is given). (D) Example of distribution with similar moments [beta distribution; parameters, $\alpha = 0.0074$ and $\beta = 2.4$, calculated to give the mean and CV as for the DFE in (C), giving skewness = 17 and kurtosis = 370]. The inset shows the same distribution on a log scale.

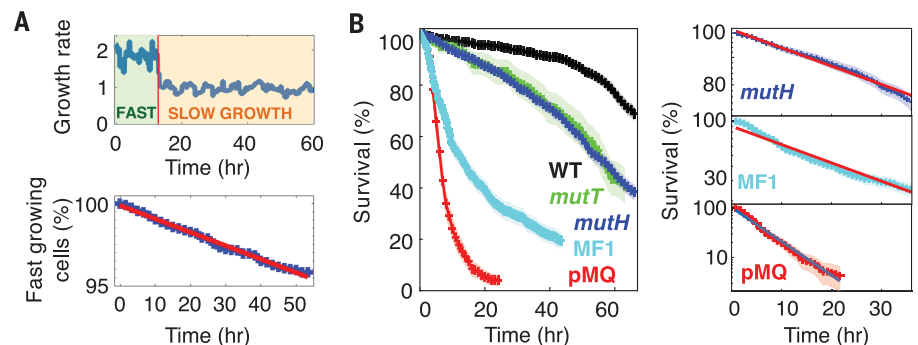


Fig. 4. Strongly deleterious (>30% cost) and lethal mutations occur at a constant rate during μ MA. (A) The top panel shows an example of a *mutH* growth rate trajectory (blue line) corresponding to the occurrence of a strongly deleterious mutation in a mother cell. To analyze such events, we defined two categories of mother cells, slow-growing cells that underwent an abrupt, >30% growth rate decrease and fast-growing cells corresponding to the rest of the population. The bottom panel shows the temporal evolution of the percentage of fast-growing cells (represented by blue dots, on a log scale; we analyzed 5272 single-cell trajectories and found 156 strongly deleterious mutations). The 95% confidence interval is narrower than the line thickness. The red curve is an exponential fit (coefficient of determination $R^2 = 0.99$). (B) The left panel shows survival curves of the population for the WT ($n = 1283$ single-cell trajectories), *mutH* ($n = 1480$), *mutT* ($n = 216$), MF1 ($n = 838$), and pMQ ($n = 472$) strains. The 95% confidence intervals are represented by the shaded areas (they are narrower than the line thickness for $n > 400$). The right panel shows the same survival curves on a log scale for *mutH* (data in blue, exponential fit in red; $R^2 = 0.99$), MF1 (data in cyan, exponential fit in red; $R^2 = 0.99$), and pMQ (data in red, exponential fit in blue; $R^2 = 0.99$) strains. The plot is limited to ~35 hours, during which age-related mortality was low and approximately constant. The 95% confidence intervals are represented by the shaded areas.

Our results show that the accumulation of mutations with <30% fitness cost does not change the rate of lethal or strongly deleterious mutations (>30% cost), as would be expected from the saturation or induction of a buffering mechanism (22, 23). We anticipate that examination of various suboptimal genotypes in μ MA experiments would allow further investigation of mutation interactions and provide a complete picture of epistasis, central to the evolution of sexual reproduction and recombination (24). Our approach may also allow for investigation of how the DFE and mutation dynamics change in response to the environment, shedding light on the adaptation potential of populations.

REFERENCES AND NOTES

1. J. W. Drake, *Crit. Rev. Biochem. Mol. Biol.* **42**, 247–258 (2007).
2. S. Uphoff *et al.*, *Science* **351**, 1094–1097 (2016).
3. J. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 8403–8408 (2007).
4. M. R. Stratton, P. J. Campbell, P. A. Futreal, *Nature* **458**, 719–724 (2009).
5. T. Bataillon, S. F. Bailey, *Ann. N.Y. Acad. Sci.* **1320**, 76–92 (2014).
6. T. Mukai, *Genetics* **50**, 1–19 (1964).
7. T. T. Kibota, M. Lynch, *Nature* **381**, 694–696 (1996).
8. P. D. Keightley, A. Caballero, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3823–3827 (1997).
9. A. Eyre-Walker, P. D. Keightley, *Nat. Rev. Genet.* **8**, 610–618 (2007).
10. D. L. Halligan, P. D. Keightley, *Annu. Rev. Ecol. Evol. Syst.* **40**, 151–172 (2009).
11. H. Lee, E. Popodi, H. Tang, P. L. Foster, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2774–E2783 (2012).
12. P. L. Foster, H. Lee, E. Popodi, J. P. Townes, H. Tang, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5990–E5999 (2015).
13. P. Wang *et al.*, *Curr. Biol.* **20**, 1099–1103 (2010).
14. Materials and methods and supplementary text are available as supplementary materials.
15. M. Elez *et al.*, *Curr. Biol.* **20**, 1432–1437 (2010).
16. R. S. Galhardo, P. J. Hastings, S. M. Rosenberg, *Crit. Rev. Biochem. Mol. Biol.* **42**, 399–435 (2007).
17. P. L. Foster, *Crit. Rev. Biochem. Mol. Biol.* **42**, 373–397 (2007).
18. R. Maharjan, T. Ferenci, *Mol. Biol. Evol.* **32**, 380–391 (2015).
19. S. Trindade, L. Perfeito, I. Gordo, *Philos. Trans. R. Soc. London Ser. B* **365**, 1177–1186 (2010).
20. L. Perfeito, A. Sousa, T. Bataillon, I. Gordo, *Evolution* **68**, 150–162 (2014).
21. K. M. Esvelt, J. C. Carlson, D. R. Liu, *Nature* **472**, 499–503 (2011).
22. S. Maisnier-Patin *et al.*, *Nat. Genet.* **37**, 1376–1379 (2005).
23. T. A. Sangster, S. Lindquist, C. Queitsch, *Bioessays* **26**, 348–362 (2004).
24. R. D. Kouyos, O. K. Silander, S. Bonhoeffer, *Trends Ecol. Evol.* **22**, 308–315 (2007).

ACKNOWLEDGMENTS

We thank M. F. Bredeche, S. Fleurier, L. Becerra, and M. Rosticher for technical assistance; O. Tenaillon, A. Amir, N. Henry, P. Thomen, G. Debregeas, A. F. Bitbol, V. Vargas, R. Robert, D. Chatenay, and M. De Paepe for discussions and comments on the manuscript; and N. Kuperwasser for manuscript editing. **Funding:** This study was funded by the Agence Nationale de Recherche (grant ANR-14-CE09-0015-01). **Author contributions:** M.E. and L.R. designed research and performed experiments. L.R., J.O., and M.E. analyzed the data. M.E. and L.R. wrote the manuscript with input from all authors. L.R. performed modeling. J.O. developed image analysis software. J.R., I.M., and X.S. contributed tools. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data described in the paper have been archived at Dryad (doi:10.5061/dryad.75625), and software codes and programs are available at Github (<https://github.com/LabJeanPerrin/BOA>).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6381/1283/suppl/DC1
 Materials and Methods
 Supplementary Text
 Figs. S1 to S18
 Tables S1 to S12
 References (25–46)
 Movies S1 and S2

1 March 2017; resubmitted 19 October 2017
 Accepted 30 January 2018
 10.1126/science.aan0797

Mutation dynamics and fitness effects followed in single cells

Lydia Robert, Jean Ollion, Jerome Robert, Xiaohu Song, Ivan Matic and Marina Elez

Science **359** (6381), 1283-1286.

DOI: 10.1126/science.aan0797

Mutation rates and effects in single cells

Understanding the dynamics of mutations and the distribution of fitness effects is critical for most evolutionary models. Robert *et al.* used a single-cell technology to visualize the accumulation of new mutations. The method identifies DNA sequences with mispaired bases and small insertions or deletions caused by DNA replication errors in living *Escherichia coli* cells. Following the fates of cells after mutation allowed for a precise quantification of the effects of new mutations. A smaller fraction of mutations were found to be deleterious than predicted previously from indirect observations.

Science, this issue p. 1283

ARTICLE TOOLS

<http://science.sciencemag.org/content/359/6381/1283>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2018/03/14/359.6381.1283.DC1>

REFERENCES

This article cites 43 articles, 11 of which you can access for free
<http://science.sciencemag.org/content/359/6381/1283#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)