The Journal of Open Source Software

STEPS: Serially Transferred Evolving Population Simulator

Devin M. Lake [1,2*¶], **Zachary W. D. Matson** [3*], **Minako Izutsu** [2,4,5], **and Richard E. Lenski** [1,2,4]

1 Department of Integrative Biology, Michigan State University, USA **2** Ecology, Evolution, and Behavior Program, Michigan State University, USA **3** Department of Computer Science and Engineering, Michigan State University, USA **4** Department of Microbiology, Genetics, and Immunology, Michigan State University, USA **5** DeSC Healthcare, Shibuya, Tokyo, Japan ¶ Corresponding author * These authors contributed equally.

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor:

Submitted: 06 August 2025
Published: unpublished

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

DRAFT

Summary

Bacteria and other microbes are widely used in the field of experimental evolution (Ascensao & Desai, 2025; Lenski, 2017; McDonald, 2019). In many of these experiments, the populations are propagated by periodic serial dilutions into fresh media, with the microbes growing to a maximum population size between successive transfers. Here we present the Serially Transferred Evolving Population Simulator (STEPS) software, which simulates the dynamics of asexual populations as they grow and evolve in these experiments. The STEPS software allows new mutations to occur while a population grows, including beneficial, neutral, and deleterious mutations. The resulting lineages then grow at different rates according to the fitness effects of their accumulated mutations. After the total population reaches the maximum size, the dilution factor determines the proportion of individuals that are randomly chosen to start the next growth cycle. The differential growth rates of the lineages reflect the process of natural selection, while the periodic bottlenecks impose random genetic drift.

The underlying mechanics of the STEPS program are designed to simulate the biological processes that occur in actual microbial evolution experiments, while taking advantage of computational methods that allow the simulations to proceed quickly and efficiently. For example, the timesteps used in the simulations coincide roughly with the population's doubling time, providing meaningful timepoints for introducing mutations. The mutational categories, the effects of mutations on relative growth rates, and the serial transfer process have been implemented such that the dynamics of the simulated populations reflect the same interplay of mutation, selection, and drift that occurs in biological experiments.

Statement of need

The Long-Term Evolution Experiment (LTEE) with *E. coli* was started in 1988 (Lenski et al., 1991; Lenski, 2023) and it continues to this day (Barrick et al., 2023). The LTEE has been used to study many aspects of evolutionary dynamics (Blount et al., 2008; Couce et al., 2024; Tenaillon et al., 2016; Wiser et al., 2013), and it has inspired many other experiments that use similar approaches (Izutsu et al., 2024; Johnson et al., 2021; Stroud & Ratcliff, 2025). Various publications based on these projects include simulations to analyze and interpret the experimental results. The authors of these papers typically write new simulation software that is custom-built to a specific experimental system or result, and even publications based on the same system may use different simulation methods (e.g., Good et al., 2017; Wiser et al., 2013). We anticipate that future studies will be able to use STEPS to perform relevant quantitative analyses or interpret qualitative results. Researchers may also use STEPS to

design new experiments or as a framework for building customized simulations. We expect that employing the STEPS program in these various ways will improve consistency and efficiency across studies.

Key features

To enable STEPS to be accessible to as many researchers as possible, we have produced two different ways to run the simulations. The first approach requires using a command-line interface to run STEPS, and it is intended for advanced applications, such as generating data for large numbers of replicate populations for statistical analysis and comparison to experimental results. All features and options built into the STEPS software are available in this version. The second approach uses a web-based graphical user interface that provides immediate visualization of the results, helping researchers develop their intuition about the dynamics of evolving populations under various scenarios. This version of STEPS can also be used in educational settings, allowing students without computational backgrounds to obtain results quickly and easily. Some features of the command-line version are not available in the web-based version. However, both versions use the same underlying code while the simulations are running.

We show figures below to illustrate a few of the dynamical outputs that are available with the STEPS package. Figure 1 shows the trajectories of (A) average fitness and (B) average accumulated mutations for 12 simulated populations over the course of 2,000 generations. The essential biological parameter values are listed in the figure legend; they correspond to the approximate values for the LTEE. Both trajectories exhibit steplike dynamics that are typical of large microbial populations in which *de novo* mutations drive adaptation (Lenski & Travisano, 1994), and which inspired the name of the STEPS program.

Trajectories for (A) average fitness and (B) average accumulated mutations in 12 simulated populations. The data were produced using the command-line version of STEPS with the following key parameter values: number of transfers = 300, maximum population size = 5e8, dilution factor = 100, rate of beneficial mutations = 1.7e-6, average beneficial effect size = 0.012 (drawn from an exponential distribution), rate of neutral mutations = 0.001, rate of deleterious mutations = 0.001 (drawn from a uniform distribution), strength of epistasis = 6.0, and initial seed = 606. All other parameters and settings are default values, except that an optional metagenomic dataset was saved (mutation-summary-output) and used to produce the next figure.

Figure 1: Trajectories for (A) average fitness and (B) average accumulated mutations in 12 simulated populations. The data were produced using the command-line version of STEPS with the following key parameter values: number of transfers = 300, maximum population size = 5e8, dilution factor = 100, rate of beneficial mutations = 1.7e-6, average beneficial effect size = 0.012 (drawn from an exponential distribution), rate of neutral mutations = 0.001, rate of deleterious mutations = 0.001 (drawn from a uniform distribution), strength of epistasis = 6.0, and initial seed = 606. All other parameters and settings are default values, except that an optional metagenomic dataset was saved (mutation-summary-output) and used to produce the next figure.

Figure 2 plots the metagenomic data for one of the 12 populations from the optional file that recorded the trajectories of every mutation that reached a threshold frequency. One can see a clean selective sweep in the first few hundred generations, followed by more complex dynamics as beneficial mutations (alone and in combination) compete with one another. This competition leads to the phenomenon of clonal interference, whereby most beneficial mutations are outcompeted by lineages that have acquired mutations with even larger beneficial effects (Gerrish & Lenski, 1998; Levy et al., 2015).

DRAFT

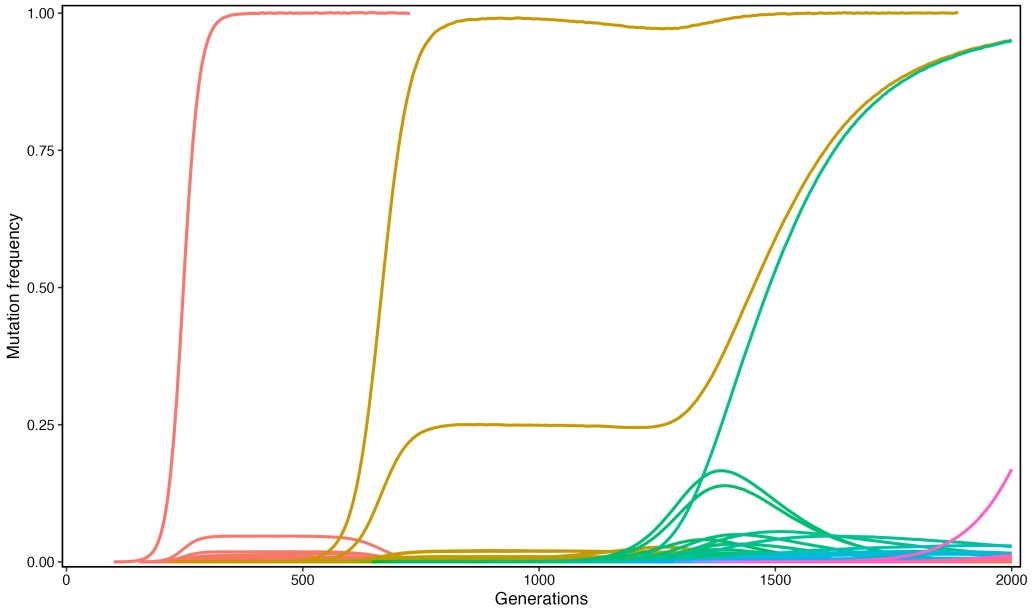

Figure 2: Trajectories for the frequencies of individual mutations in one of the populations in Figure 1 (brown trajectory, third population from left to show a conspicuous increase in fitness.) Only those mutations that reached a frequency of at least 0.01% were saved in the output file. See Figure 1 for parameter values and other details.

To facilitate the use of STEPS by users with diverse interests and skillsets in both research and educational settings, we have produced a User Manual (Lake et al., 2025) that explains: the scientific context and purpose of the STEPS software (Chapter 1); the use of the web-based version, including numerous exercises with figures (Chapter 2); the setup and full set of options available in the command-line version (Chapter 3); and the mechanics of the simulations that underpin both the command-line and web-based versions (Chapter 4). Figure 3 is a screenshot of the web interface along with results generated in a matter of seconds using the default settings.

DRAFT

Figure 3: Screenshot of the interface and results using the web-based version of STEPS. Default parameter values were used, except that the random seed was set to 606. These values are the same as used for Figure 1, except that the default values for the neutral and deleterious mutation rates are set to 0 to allow faster runtimes. Exercises presented in the User Manual (Lake et al., 2025) explore the effects of adding these types of mutations and varying population sizes, run durations, and various other parameters and options.

DRAFT

Acknowledgments

Development of the STEPS program has been supported by the U.S. National Science Foundation (DEB-1951307), a USDA Hatch Grant (MICL12143), and the John A. Hannah professorial endowment at Michigan State University. We thank Ben Good for valuable discussions during the design of STEPS, and Kyle Card and Nkrumah Grant for testing earlier versions of STEPS. The LTEE inspired us to create STEPS, and we thank everyone who has worked on that experiment over the several decades that it has been running.

References

Ascensao, J. A., & Desai, M. M. (2025). Experimental evolution in an era of molecular manipulation. *Nature Reviews Genetics*. https://doi.org/10.1038/s41576-025-00867-6

Barrick, J. E., Blount, Z. D., Lake, D. M., Dwenger, J. H., Chavarria-Palma, J. E., Izutsu, M., & Wiser, M. J. (2023). Daily transfers, archiving populations, and measuring fitness in the long-term evolution experiment with Escherichia coli. *JoVE*, *198*, e65342. https://doi.org/10.3791/65342

Blount, Z. D., Borland, C. Z., & Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences*, *105*(23), 7899–7906. https://doi.org/10.1073/pnas.0803151105

Couce, A., Limdi, A., Magnan, M., Owen, S. V., Herren, C. M., Lenski, R. E., Tenaillon, O.,

& Baym, M. (2024). Changing fitness effects of mutations through long-term bacterial evolution. *Science*, *383*(6681), eadd1417. https://doi.org/10.1126/science.add1417

Gerrish, P. J., & Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. *Genetica*, *102*, 127–144. https://doi.org/10.1023/A:1017067816551

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., & Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature*, *551*(7678), 45–50. https://doi.org/10.1038/nature24287

Izutsu, M., Lake, D. M., Matson, Z. W. D., Dodson, J. P., & Lenski, R. E. (2024). Effects of periodic bottlenecks on the dynamics of adaptive evolution in microbial populations. *Microbiology*, *170*(9). https://doi.org/10.1099/mic.0.001494

Johnson, M. S., Gopalakrishnan, S., Goyal, J., Dillingham, M. E., Bakerlee, C. W., Humphrey, P. T., Jagdish, T., Jerison, E. R., Kosheleva, K., Lawrence, K. R., Min, J., Moulana, A., Phillips, A. M., Piper, J. C., Purkanti, R., Rego-Costa, A., McDonald, M. J., Nguyen Ba, A. N., & Desai, M. M. (2021). Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. *eLife*, *10*, e63910. https://doi.org/10.7554/eLife.63910

Lake, D. M., Matson, Z. W. D., & Lenski, R. E. (2025). STEPS user manual (v2.0.0). *Zenodo*. https://doi.org/10.5281/zenodo.16690374

Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME Journal*, *11*(10), 2181–2194. https://doi.org/10.1038/ismej.2017.69

Lenski, R. E. (2023). Revisiting the design of the long-term evolution experiment with Escherichia coli. *Journal of Molecular Evolution*, *91*(3), 241–253. https://doi.org/10.1007/s00239-023-10095-3

Lenski, R. E., Rose, M. R., Simpson, S. C., & Tadler, S. C. (1991). Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. *The American Naturalist*, *138*(6), 1315–1341. https://doi.org/10.1086/285289

Lenski, R. E., & Travisano, M. (1994). Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences*, *91*(15), 6808–6814. https://doi.org/10.1073/pnas.91.15.6808

Levy, S. F., Blundell, J. R., Venkataram, S., Petrov, D. A., Fisher, D. S., & Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, *519*(7542), 181–186. https://doi.org/10.1038/nature14279

McDonald, M. J. (2019). Microbial experimental evolution – a proving ground for evolutionary theory and a tool for discovery. *EMBO Reports*, *20*(8), e46992. https://doi.org/10.15252/embr.201846992

Stroud, J. T., & Ratcliff, W. C. (2025). Long-term studies provide unique insights into evolution. *Nature*, *639*(8055), 589–601. https://doi.org/10.1038/s41586-025-08597-9

Tenaillon, O., Barrick, J. E., Ribeck, N., Deatherage, D. E., Blanchard, J. L., Dasgupta, A., Wu, G. C., Wielgoss, S., Cruveiller, S., Médigue, C., Schneider, D., & Lenski, R. E. (2016). Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature*, *536*(7615), 165–170. https://doi.org/10.1038/nature18959

Wiser, M. J., Ribeck, N., & Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science*, *342*(6164), 1364–1367. https://doi.org/10.1126/science.1243357

DRAFT