MICROBIAL EVOLUTION

Global epistasis makes adaptation predictable despite sequence-level stochasticity

Sergey Kryazhimskiy,[1,3]*† Daniel P. Rice,[1,3]* Elizabeth R. Jerison,[2,3] Michael M. Desai[1,2,3]†

Epistatic interactions between mutations can make evolutionary trajectories contingent on the chance occurrence of initial mutations. We used experimental evolution in *Saccharomyces cerevisiae* to quantify this contingency, finding differences in adaptability among 64 closely related genotypes. Despite these differences, sequencing of 104 evolved clones showed that initial genotype did not constrain future mutational trajectories. Instead, reconstructed combinations of mutations revealed a pattern of diminishing-returns epistasis: Beneficial mutations have consistently smaller effects in fitter backgrounds. Taken together, these results show that beneficial mutations affecting a variety of biological processes are globally coupled; they interact strongly, but only through their combined effect on fitness. As a consequence, fitness evolution follows a predictable trajectory even though sequence-level adaptation is stochastic.

Epistatic interactions between mutations are pervasive in microbial and viral systems (*1*–*6*). In some cases, a single mutation can open up previously unavailable opportunities for a population to colonize a new metabolic niche (*2*) or survive in a previously intolerable drug concentration (*3*). Such idiosyncratic epistasis makes evolutionary trajectories dependent on the chance occurrence of initial mutations that constrain or potentiate future adaptation. This historical contingency can render adaptation fundamentally unpredictable (*7*). However, recent work has also provided evidence for more systematic patterns of epistasis (*8*–*10*), which can drive convergent phenotypic evolution (*11*–*13*) or can lead to parallel adaptation at the sequence level (*14*). These observations suggest that evolutionary outcomes may be statistically predictable if mutations causing idiosyncratic changes in adaptability are rare, and if epistasis instead channels evolution into convergent phenotypic or genotypic pathways.

To test how epistasis and historical contingency affect the predictability of adaptation, we conducted a hierarchical laboratory evolution experiment in *S. cerevisiae* (fig. S1). In the first phase of the experiment ("diversification"), we created 432 independent lines from a single haploid clone (the diversification ancestor, DivAnc) isolated from an earlier long-term evolution experiment (*15*). We evolved each line independently, half at large and half at small population size, in rich media in 96-well microplates for 240 generations (*16*). We then selected 64 clones ("founders"), each from a different line, chosen to span a range of fitness relative to the DivAnc (*16*) (table S1). Founders differed from the DivAnc by 4.2 mutations on average (*16*). In the second phase of the experiment ("adaptation"), we founded 10 independent replicate populations with each founder and then allowed each of the resulting 640 lines to adapt at large population size for 500 generations. This enabled us to compare variation among lines descended from the same founder (which reflects the inherent stochasticity of evolution) to variation between lines descended from different founders, thereby providing an assessment of the extent to which genetic background influences evolution.

The competitive fitness of each population after 250 and 500 generations of the adaptation phase increased on average by 3.3% and 6.6%, respectively (Fig. 1A and table S2). However, not all populations adapted at the same rate. Instead, the initially large variation in fitness between lines declined with time (Fig. 1A). We carried out an analysis of variance (ANOVA) to partition observed variance in fitness increase during the adaptation phase into contributions from measurement noise, inherent stochasticity of the evolutionary process, and the identity of the founder (*16*, *17*). After 250 generations of adaptation, 49% of the variance in fitness increment was attributed to inherent evolutionary stochasticity, 17% to measurement error, and 34% to founder identity; after 500 generations, 29% of the variance in fitness increment was attributed to inherent stochasticity, 21% to measurement error, and 50% to founder identity (Fig. 1B and table S3). This demonstrates that genetic background is a key determinant of how rapidly a population will adapt.

These differences in adaptability are not random: Populations with lower initial fitness systematically adapt more rapidly than populations with higher initial fitness, driving the overall pattern of convergent evolution in fitness (Fig. 1C). We further partitioned the variation in fitness increment attributed to founder identity and found that after 250 generations of adaptation, 31% of this variation was explained by the fitness of the founder, whereas 3% was determined by its specific genotype; for 500 generations, these percentages rose to 46% and 4%, respectively (Fig. 1B and table S3). Thus, the differences in adaptability between founders are almost entirely predicted by their differences in fitness and are independent of the specific mutations underlying this fitness. The initial fitness of the founder therefore predicts the average rate of adaptation in its descendant lines (Fig. 1D). Although the effects of specific genotype on adaptability are rare or weak, they are significant (fig. S2 and tables S3 and S4).

A negative correlation between fitness and adaptability has also been observed in prokaryotes (*11*, *12*), and it is consistent with the common observation in evolution experiments that the rate of increase in fitness slows down over time (*13*, *18*). Combined with this earlier work, our results suggest a general "rule of declining adaptability" that holds for prokaryotes and eukaryotes adapting to rich and minimal media. Further, our observations support a stronger version of this rule: Genotypes with lower fitness are more adaptable than those with higher fitness, and distinct genotypes with identical fitness are equally adaptable (up to the rare or weak exceptions noted above). This is consistent with the argument recently presented in (*13*).

The rule of declining adaptability could arise for two non-exclusive reasons. First, there could be only a few ways to increase in fitness. In this model, high-fitness founders have lower adaptability because they have already acquired all or most of the possible strong-effect beneficial mutations: They are "running out" of beneficial mutations. In contrast, low-fitness founders adapt more quickly because they have not yet acquired these mutations. More generally, some groups of mutations may have redundant functional effects (e.g., those that knock out a given pathway). In this case, the number of nonredundant ways to increase fitness would be much smaller than the number of distinct beneficial mutations. We refer to this general form of the running-out-of-mutations hypothesis as the "modular epistasis" model [inspired by (*14*)]: Each beneficial mutation improves a single module, mutations within each module are redundant, and high-fitness founders adapt more slowly because they have fewer remaining modules to improve, especially among those modules that confer the largest fitness gains (*16*).

Alternatively, mutations arising in higher-fitness backgrounds may be less beneficial than those arising in lower-fitness backgrounds; that is, diminishing-returns epistasis may be pervasive among adaptive mutations, as suggested by (*8*–*10*). This epistasis could have two forms. If epistasis is idiosyncratic, mutations may often have widely different effects in different genetic backgrounds (possibly including sign epistasis, where the sign of the fitness effect depends on genetic background), but the average effect of a beneficial mutation is smaller in fitter backgrounds. On the

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [2]Department of Physics, Harvard University, Cambridge, MA 02138, USA. [3]FAS Center for Systems Biology, Harvard University, Cambridge, MA 02138, USA.
*These authors contributed equally to this work. †Corresponding author. E-mail: skryazhi@oeb.harvard.edu (S.K.); mdesai@oeb.harvard.edu (M.M.D.)

other hand, if epistasis is global, each individual beneficial mutation provides a smaller advantage in a fitter genetic background. This latter model implies that the effect of each mutation depends on all other mutations, but only through their combined effect on fitness.

In the modular and idiosyncratic epistasis models, different founders have different sets of beneficial mutations available to them. Hence, in both models we expect lines descended from the same founder to take more similar mutational trajectories than lines descended from different founders. In the modular model, we expect each mutation to either confer some fixed advantage (in genotypes lacking mutations in that module) or be neutral (in genotypes that already have a mutation in that module). In the idiosyncratic epistasis model, we expect individual mutations to have a variety of different fitness effects in different genetic backgrounds. In contrast to these two models, in the global epistasis model all genotypes acquire beneficial mutations from the same pool, but the advantage conferred by each mutation consistently declines with the fitness of the genetic background. In this model, founder identity should not affect subsequent mutational trajectories.

To assess the extent to which these three types of epistasis contribute to the rule of declining adaptability, we sampled one clone from each population descended from 15 founders at generation 500 of the adaptation phase and sequenced their complete genomes (*16*, *19*). We found that four sequenced clones acquired a mutator phenotype during the adaptation phase, and two founders and all their descendants became diploid (fig. S3). We excluded these from further analysis, leaving a total of 104 sequenced clones descended from 13 founders (*16*). We identified a total of 55 mutations that occurred in these founders during diversification and 1149 mutations that occurred in their descendants during adaptation. We annotated each mutation to a gene or intergenic region and classified coding mutations as synonymous or nonsynonymous (Fig. 2A and table S5). Because most synonymous and intergenic mutations are likely neutral hitchhikers, we restricted analysis to putatively functional nonsense, frameshift, nonsynonymous, and promoter mutations (818 total mutations).

In contrast to experiments in bacteria and viruses (*14*, *20*), all but four mutations are unique at the nucleotide level, consistent with earlier work in *S. cerevisiae* (*21*). However, we found significant gene-level convergent evolution. For example, 24 genes had mutations in at least three replicate lines [versus 2.7 genes expected by chance; multinomial test, $P < 0.01$ (*16*) (tables S6 and S7)], indicating that most mutations observed in these "multi-hit" genes are likely beneficial. Moreover, mutations in genes involved in negative regulation of Ras, cell cycle regulation, and filamentous growth were enriched (table S8), demonstrating convergence at higher levels of biological organization.

We next compared the total number of mutations observed in different evolved lines. Among lines descended from a given founder, the lines that increased most in fitness acquired more mutations on average in multi-hit genes, as we would expect if these mutations are beneficial (fig. S4). The modular epistasis model predicts that lines descended from high-fitness founders should acquire fewer beneficial mutations than those descended from low-fitness founders, because the former have fewer ways to improve. However, this is not the case: The numbers of putatively functional mutations in lines descended from different founders are not significantly different (Fig. 2B and table S9). This result is also surprising under the diminishing-returns epistasis models, although not strictly inconsistent with them (*22*). Because neutral hitchhiker mutations could mask differences in numbers of beneficial mutations between lines (*23*), we repeated this analysis on more restricted sets of "putatively beneficial mutations" [e.g., those in multi-hit genes (*24*)]. We found similar results in all cases (figs. S5 and S6).

In the modular and idiosyncratic epistasis models, many mutations are beneficial only in particular genetic backgrounds. Hence, these models predict that clones descended from the same founder should on average have more mutations in common (parallelism) than expected by chance, given the observed degree of overall convergence. However, this is not the case. Instead, clones descended from the same founder are not significantly more likely to share mutations than clones descended from different founders (Fig. 2, C and D, and fig. S7), as expected in the global diminishing-returns epistasis model. This pattern holds regardless of the level at which we define parallelism and convergence (genes or GO Slim categories).

We next selected three genes (*SFL1*, *WHI2*, and *GAT2*) in which we found putative loss-of-function (nonsense or frameshift) mutations in three or

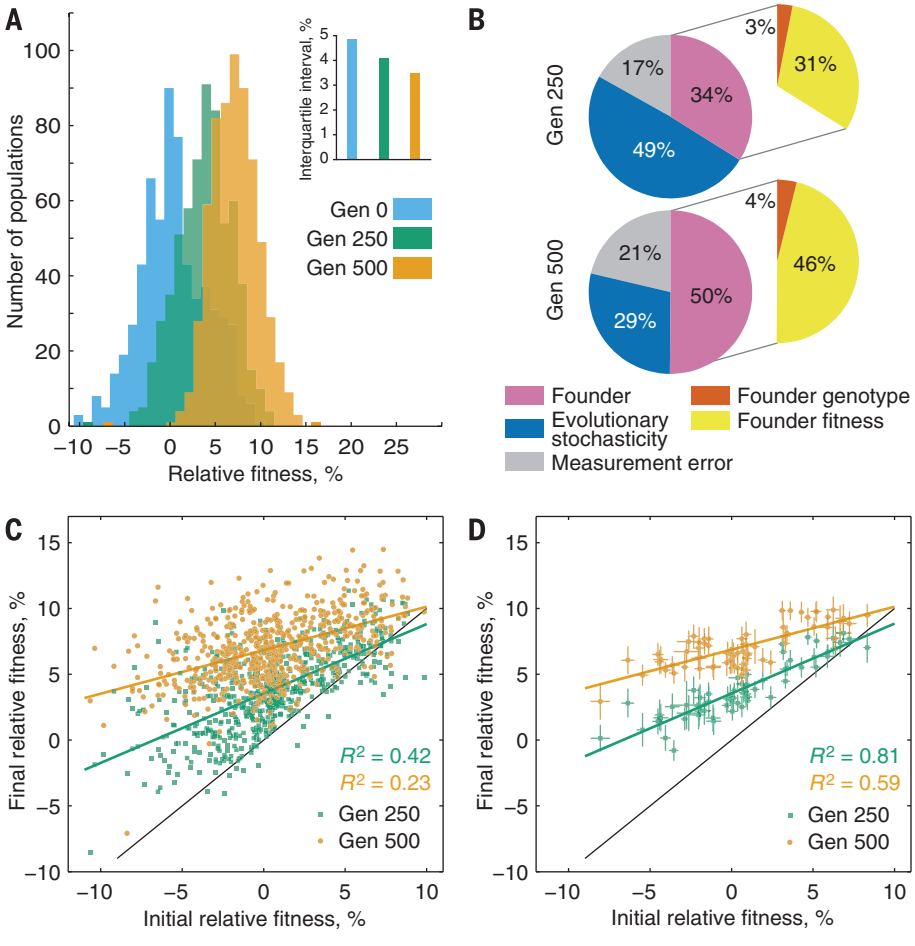

Fig. 1. Fitness evolution. (**A**) Distribution of mean population fitness over time, relative to the diversification ancestor (DivAnc). Inset shows interpopulation fitness variation over time. (**B**) Fraction of the variance between lines in fitness increment that is attributable to each indicated component after 250 and 500 generations of the adaptation phase. All variance components are significant (table S3). (**C**) Relationship between founder fitness and population fitness after 250 and 500 generations of adaptation. (**D**) Relationship between founder fitness and the mean fitness of the 10 independent lines descended from that founder, after 250 and 500 generations of adaptation. Error bars denote SEM.

more lines, suggesting that knockouts of these genes are beneficial in our system. *GAT2* displays the strongest signature of parallel evolution in our data (Fig. 2D) and therefore represents the strongest candidate for idiosyncratic epistasis. We constructed separate targeted knockouts of each of these genes, along with one control gene, *HO*, in several replicates into all 13 founders, DivAnc, and four additional clones (*16*) (table S1). We measured the fitness effects of each knockout in each background, and found a negative correlation between the fitness effect of the *gat2*Δ, *whi2*Δ, and *sfl1*Δ gene deletions and the fitness of the background strain (Fig. 3). Furthermore, there were no idiosyncratic epistatic interactions specific to particular genotypes: Up to small deviations, the fitness effect of each knockout depends only on the fitness of the genetic background and not on the specific mutations present in that background.

Taken together, these results support the global diminishing-returns epistasis model as the dominant explanation for declining adaptability with increasing fitness, and paint a surprisingly simple picture of adaptation in our system. Many mutations scattered across many biological processes appear to be beneficial. Yet despite their lack of apparent functional relationship, these mutations are globally coupled by diminishing-returns epistasis; their effects are strongly mediated by background fitness but are otherwise essentially independent of the specific identity of mutations present in the background. The biological basis of this global coupling remains unknown. Nonetheless, it

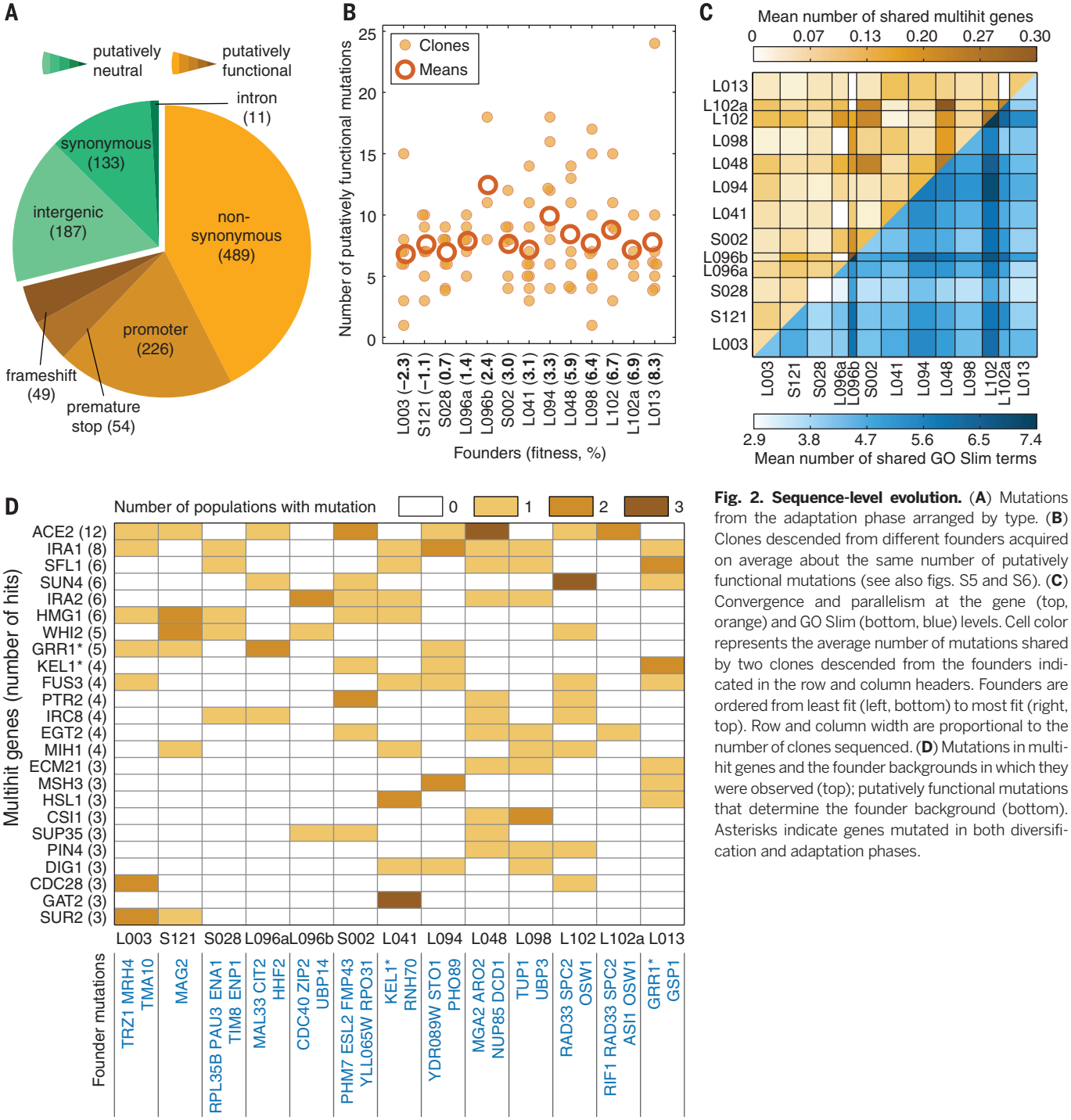

Fig. 2. Sequence-level evolution. (**A**) Mutations from the adaptation phase arranged by type. (**B**) Clones descended from different founders acquired on average about the same number of putatively functional mutations (see also figs. S5 and S6). (**C**) Convergence and parallelism at the gene (top, orange) and GO Slim (bottom, blue) levels. Cell color represents the average number of mutations shared by two clones descended from the founders indicated in the row and column headers. Founders are ordered from least fit (left, bottom) to most fit (right, top). Row and column width are proportional to the number of clones sequenced. (**D**) Mutations in multihit genes and the founder backgrounds in which they were observed (top); putatively functional mutations that determine the founder background (bottom). Asterisks indicate genes mutated in both diversification and adaptation phases.

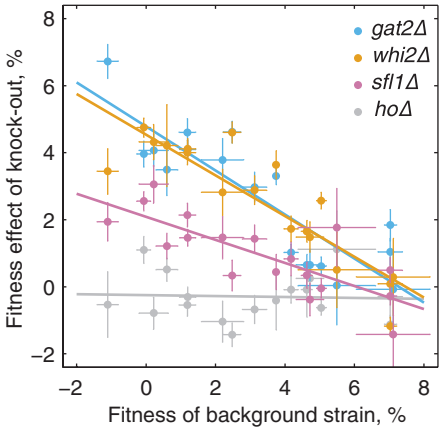

Fig. 3. Diminishing-returns epistasis among specific mutations. The fitness effect of knocking out genes *gat2*, *whi2*, and *sfl1* declines with the fitness of the background strain. The *ho* knockout is a negative control. Error bars are SEM over biological replicates.

leads to a striking pattern of convergent evolution, making fitness evolution relatively predictable. Despite this fitness-level convergence, evolution remains highly stochastic at the genotype level, likely because many distinct mutational paths can lead a population to any given fitness.

REFERENCES AND NOTES

1. R. J. Woods *et al.*, *Science* **331**, 1433–1436 (2011).
2. Z. D. Blount, C. Z. Borland, R. E. Lenski, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7899–7906 (2008).
3. J. D. Bloom, L. I. Gong, D. Baltimore, *Science* **328**, 1272–1275 (2010).
4. D. R. Rokyta *et al.*, *PLOS Genet.* **7**, e1002075 (2011).
5. C. L. Burch, L. Chao, *Nature* **406**, 625–628 (2000).
6. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, *Science* **312**, 111–114 (2006).
7. S. J. Gould, *Wonderful Life* (Norton, New York, 1989).
8. A. I. Khan, D. M. Dinh, D. Schneider, R. E. Lenski, T. F. Cooper, *Science* **332**, 1193–1196 (2011).
9. H.-H. Chou, H.-C. Chiu, N. F. Delaney, D. Segrè, C. J. Marx, *Science* **332**, 1190–1192 (2011).
10. H.-H. Chou, J. Berthet, C. J. Marx, *PLOS Genet.* **5**, e1000652 (2009).
11. J. E. Barrick, M. R. Kauth, C. C. Strelioff, R. E. Lenski, *Mol. Biol. Evol.* **27**, 1338–1347 (2010).
12. L. Perfeito, A. Sousa, T. Bataillon, I. Gordo, *Evolution* **68**, 150–162 (2014).
13. M. J. Wiser, N. Ribeck, R. E. Lenski, *Science* **342**, 1364–1367 (2013).
14. O. Tenaillon *et al.*, *Science* **335**, 457–461 (2012).
15. G. I. Lang, D. Botstein, M. M. Desai, *Genetics* **188**, 647–661 (2011).
16. See supplementary materials on *Science* Online.
17. M. Travisano, J. A. Mongold, A. F. Bennett, R. E. Lenski, *Science* **267**, 87–90 (1995).
18. S. Kryazhimskiy, G. Tkačik, J. B. Plotkin, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18638–18643 (2009).
19. Note that this includes two founders inadvertently picked from the same diversified population (*16*).
20. J. P. Bollback, J. P. Huelsenbeck, *Mol. Biol. Evol.* **24**, 1397–1406 (2007).
21. G. I. Lang *et al.*, *Nature* **500**, 571–574 (2013).
22. This result is also surprising in the diminishing-returns models because we expect fewer beneficial mutations to fix in high-fitness backgrounds where they provide a smaller selective advantage. This puzzle is related to the observation (*25*) that fixation rates in long-term evolution of *E. coli* are constant through time despite a declining rate of fitness increase. However, our result would be consistent with the diminishing-returns models if the beneficial mutation rate is also higher in high-fitness backgrounds.
23. See (*16*) for a power analysis.
24. These are likely enriched for the most strongly beneficial mutations. Hence, if modular epistasis is prevalent, it is among these mutations that we expect the strongest trend.
25. J. E. Barrick *et al.*, *Nature* **461**, 1243–1247 (2009).

ACKNOWLEDGMENTS

We thank A. Murray, Q. Justman, B. Good, D. van Dyken, M. McDonald for useful discussions; A. Subramaniam, G. Lang, M. Müller, and J. Koschwanez for experimental advice and strains; and P. Rogers and C. Daly for technical support. Supported by the Burroughs Wellcome Foundation (S.K.), NSF graduate research fellowships (D.P.R., E.R.J.), and the James S. McDonnell Foundation, the Alfred P. Sloan Foundation, the Harvard Milton Fund, NSF grant PHY 1313638, and NIH grant GM104239 (M.M.D.). Sequence data have been deposited to GenBank under BioProject identifier PRJNA242140.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/344/6191/1519/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S12
References (*26*–*34*)

16 January 2014; accepted 28 May 2014
10.1126/science.1250939

NONHUMAN GENETICS

Genomic basis for the convergent evolution of electric organs

Jason R. Gallant,[1,2]* Lindsay L. Traeger,[3,4]* Jeremy D. Volkening,[4,5] Howell Moffett,[6,7] Po-Hao Chen,[6,7,8] Carl D. Novina,[6,7,8] George N. Phillips Jr.,[9] Rene Anand,[10] Gregg B. Wells,[11] Matthew Pinch,[12] Robert Güth,[12] Graciela A. Unguez,[12] James S. Albert,[13] Harold H. Zakon,[2,14,15]† Manoj P. Samanta,[16]† Michael R. Sussman[4,5]†

Little is known about the genetic basis of convergent traits that originate repeatedly over broad taxonomic scales. The myogenic electric organ has evolved six times in fishes to produce electric fields used in communication, navigation, predation, or defense. We have examined the genomic basis of the convergent anatomical and physiological origins of these organs by assembling the genome of the electric eel (*Electrophorus electricus*) and sequencing electric organ and skeletal muscle transcriptomes from three lineages that have independently evolved electric organs. Our results indicate that, despite millions of years of evolution and large differences in the morphology of electric organ cells, independent lineages have leveraged similar transcription factors and developmental and cellular pathways in the evolution of electric organs.

Electric fishes use electric organs (EOs) to produce electricity for the purposes of communication; navigation; and, in extreme cases, predation and defense (*1*). EOs are a distinct vertebrate trait that has evolved at least six times independently (Fig. 1A). The taxonomic diversity of fishes that generate electricity is so profound that Darwin specifically cited them as an important example of convergent evolution (*2*). EOs benefit as a model for understanding general principles of the evolution of complex traits, as fish have evolved other specialized noncontractile muscle-derived organs (*3*). Furthermore, EOs provide a basis to assess whether similar mechanisms underlie the evolution of other specialized noncontractile muscle derivatives, such as the cardiac conduction system (*4*).

Electric organs are composed of cells called electrocytes (Fig. 1B). All electrocytes have an innervated surface enriched in cation-specific ion channels and, on the opposite surface, an invaginated plasma membrane enriched in sodium pumps, and, in some species, ion channels as well. The functional asymmetry of these cells, and their "in-series" arrangement within each organ, allows for the summation of voltages, much like batteries stacked in series in a flashlight. Although EOs originate developmentally from myogenic precursors, they are notably larger than muscle fibers (*5*). Further, they either lack the contractile machinery clearly evident in electron

[1]Department of Zoology, Michigan State University, East Lansing, MI 48824, USA. [2]BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA. [3]Department of Genetics, University of Wisconsin, Madison, WI 53706, USA. [4]Biotechnology Center, University of Wisconsin, Madison, WI 53706, USA. [5]Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA. [6]Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, MA 02115, USA. [7]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA. [8]Broad Institute of Harvard and MIT, Cambridge, MA 02141, USA. [9]Department of Biochemistry and Cell Biology and Department of Chemistry, Rice University, Houston, TX 77005, USA. [10]Department of Pharmacology and Department of Neuroscience, College of Medicine, The Ohio State University Wexner Medical Center, Columbus, OH 43210, USA. [11]Department of Molecular and Cellular Medicine, Texas A&M University, College Station, TX 77843, USA. [12]Department of Biology, New Mexico State University, Las Cruces, NM 88003, USA. [13]Department of Biology, University of Louisiana, Lafayette, LA 70503, USA. [14]University of Texas, Austin, TX 78712, USA. [15]The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, The Marine Biological Laboratory, Woods Hole, MA 02543, USA. [16]Systemix Institute, Redmond, WA 98053, USA.
*These authors contributed equally to this work. **†Corresponding author. E-mail: msussman@wisc.edu (M.R.S.); manoj.samanta@systemix.org (M.P.S.); h.zakon@austin.utexas.edu (H.H.Z.)**

Global epistasis makes adaptation predictable despite sequence-level stochasticity

Sergey Kryazhimskiy, Daniel P. Rice, Elizabeth R. Jerison and Michael M. Desai

Science **344** (6191), 1519-1522.
DOI: 10.1126/science.1250939

Clouding evolution's crystal ball

Because of a sort of mutation buffering process, different starting mutations can tend to end up with similar overall affects on an organism's fitness. Kryazhimskiy *et al.* evolved lines of yeast, each originating from distinct single genotypes, under the same selective regimen. A subset of clones from these adapted populations was subjected to fitness assays and sequenced. Populations with lower initial fitness, adapted more rapidly than populations with higher initial fitness, so that in the end the fitness levels were similar.

Science, this issue p. 1519