EVOLUTION

Idiosyncratic epistasis leads to global fitness–correlated trends

Christopher W. Bakerlee[1,2,3]†, Alex N. Nguyen Ba[1,2,4,5]†, Yekaterina Shulgina[3], Jose I. Rojas Echenique[1,6], Michael M. Desai[1,2,7,8]*

Epistasis can markedly affect evolutionary trajectories. In recent decades, protein-level fitness landscapes have revealed extensive idiosyncratic epistasis among specific mutations. By contrast, other work has found ubiquitous and apparently nonspecific patterns of global diminishing-returns and increasing-costs epistasis among mutations across the genome. Here, we used a hierarchical CRISPR gene drive system to construct all combinations of 10 missense mutations from across the genome in budding yeast and measured their fitness in six environments. We show that the resulting fitness landscapes exhibit global fitness–correlated trends but that these trends emerge from specific idiosyncratic interactions. We thus provide experimental validation of recent theoretical work arguing that fitness-correlated trends can emerge as the generic consequence of idiosyncratic epistasis.

Epistatic interactions have important consequences for the design and evolution of genetic systems (*1*–*3*). Much work in recent decades has studied these interactions by measuring empirical fitness landscapes, most often at "shallow" depth for genome-scale studies (e.g., by quantifying pairwise but not higher-order epistasis between all gene deletions or mutations) or at "narrow" breadth (e.g., complete landscapes at the scale of small select regions in single genes such as by quantifying all orders of epistatic interactions among few amino acid residues) (*4*–*18*). These studies have often found many epistatic interactions among specific mutations at both lower orders (i.e., among few mutations) and higher orders (i.e., among many mutations). These reflect particular biological and physical interactions among the mutations involved; following recent work (*19*, *20*), we refer to them as "idiosyncratic" epistasis because they involve the specific details of these mutations. Overall, this body of work has highlighted the potential for epistasis to create historical contingency that tightly constrains the distribution of adaptive trajectories accessible to natural selection.

By contrast, other work examining adaptive trajectories that implicate loci across the genome has found patterns of apparently "global" epistasis, in which the fitness effect of a mutation varies systematically with the fitness of the genetic background on which it occurs (*21*–*28*). Typically, this manifests as either diminishing returns for beneficial mutations or increasing costs for deleterious mutations, with mutations having a less positive or more negative effect on fitter backgrounds. These consistent patterns of global epistasis may give rise to the dominant evolutionary trend of declining adaptability, and in contrast to the complexity of idiosyncratic interactions, they suggest that historical contingency could be less critical in constraining adaptive trajectories (*29*).

Despite their importance, these dual descriptions of epistasis have not been satisfactorily unified. In one view, global epistasis results from nonspecific fitness-mediated interactions among mutations (*24*). Such interactions may, for example, emerge from the topology of metabolic networks, which generates overall patterns of diminishing returns and increasing costs that eclipse the specific details of idiosyncratic interactions (*30*). By contrast, other recent theoretical work has proposed an alternative view, hypothesizing that apparent fitness-mediated epistasis can instead emerge as the generic consequence of idiosyncratic interactions if they are sufficiently numerous and widespread (*19*, *20*). These two models have substantially different implications for the structure of fitness landscapes, which in turn influence our expectations of the repeatability and predictability of evolution and of the effect of chance and contingency on adaptation at both the genotypic and phenotypic levels. Thus, this dichotomy plays a central role in our understanding of how epistasis affects evolutionary dynamics.

Thus far, however, empirical work has been unable to distinguish between these perspectives. The key difficulty is that testing these ideas requires both depth and breadth: We must analyze landscapes involving enough loci that we sample idiosyncratic interactions that can potentially give rise to overall fitness-mediated trends, and we must survey possible combinations of these mutations at sufficient depth to quantify the role of higher-order interactions (including potential "global" nonspecific fitness-mediated interactions). Larger landscapes are also necessary to reduce the influence of measurement error on the inference of epistasis and analysis of fitness-correlated trends (FCTs) (see the supplementary materials, section 6.3). Achieving this depth and breadth is technically challenging because it requires us to synchronize many mutations across the genome.

Here, we overcame this challenge by developing a method that exploits Cre-Lox recombination to create a combinatorially expanding CRISPR guide-RNA (gRNA) array in *Saccharomyces cerevisiae*, which allowed us to iteratively generate mutations at distant loci through a gene-drive mechanism (Fig. 1A). Briefly, strains of opposite mating type containing inducible Cre recombinase and *SpCas9* genes were mutated at one of two loci (*A* or *B*), and DNA encoding gRNAs specific to the wild-type (WT) alleles at these loci were integrated into their genomes (fig. S1). After mating to produce a diploid heterozygous at *A* and *B*, we induced a gene drive to make the loci homozygous. This began with expressing Cas9 and generating gRNA-directed double-strand breaks at the WT *A* and *B* alleles. These breaks were then repaired by the mutated regions of homologous chromosomes, making the diploid homozygous at these loci with at least 95% efficiency. Simultaneously, we expressed Cre to induce recombination that brought gRNAs into physical proximity on the same chromosome by way of flanking Lox sites using a strategy similar to that described previously (*31*) (Fig. 1B). We then sporulated diploids and selected haploids bearing the linked gRNAs from both parents. In parallel, we performed this process with "pseudo-WT" versions of these loci, which contain synonymous changes that abolish gRNA recognition but lack the nonsynonymous change of interest. This created a set of four strains, with all possible genotypes at loci *A* and *B*. Concurrently, we created separate sets of four strains with all possible genotypes at other pairs of loci (e.g., *C* and *D*).

By iterating this process, we could rapidly assemble an exponentially expanding, combinatorially complete genotype library. We mated separate sets of four genotypes bearing all combinations of mutations at two loci each in an all-against-all cross, drove their mutations, recombined their gRNAs, and sporulated to produce a 16-strain library bearing all four-locus mutation combinations. Repeating these steps in a third cycle with two four-locus libraries of opposite mating type yielded a 256-strain, eight-locus library, and a complete landscape of up to 16 mutations (2^{16} strains) could be constructed in just four cycles.

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. [2]Quantitative Biology Initiative, Harvard University, Cambridge, MA, USA. [3]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA. [4]Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada. [5]Department of Biology, University of Toronto Mississauga, Mississauga, Ontario, Canada. [6]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [7]NSF-Simons Center for Mathematical and Statistical Analysis of Biology, Harvard University, Cambridge, MA, USA. [8]Department of Physics, Harvard University, Cambridge, MA, USA.
*Corresponding author. Email: mdesai@oeb.harvard.edu
†These authors contributed equally to this work

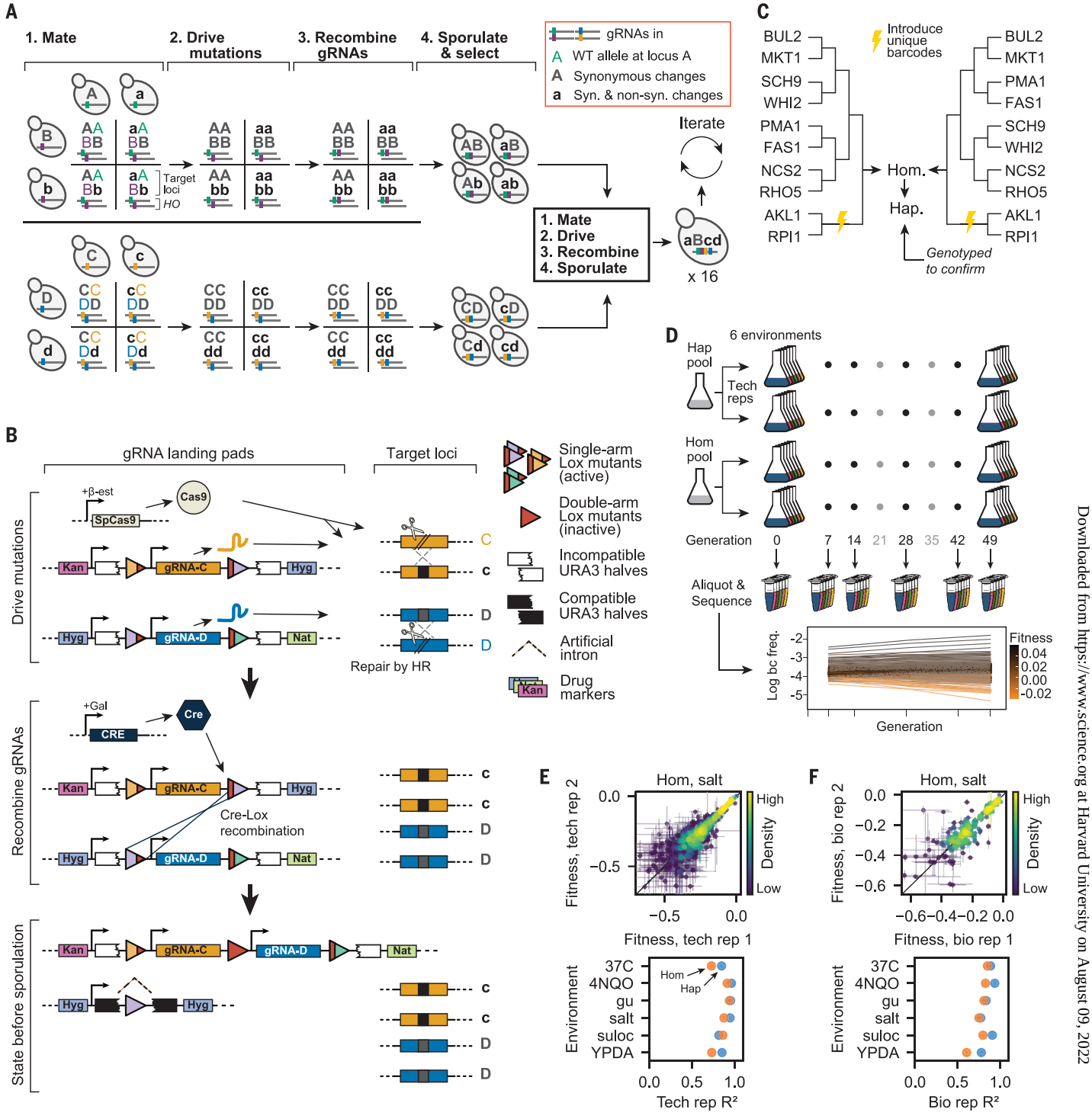

Fig. 1. Recombining CRISPR-gene drive system. (**A**) Experimental design. Strains of opposite mating type carrying known mutations and corresponding gRNAs mate to form heterozygous diploids. Cas9 expression "drives" these mutations, and site-specific recombination links gRNAs. Homozygous diploids are sporulated, haploids with linked gRNAs are selected, and the process repeats, incorporating exponentially increasing numbers of mutations. (**B**) Recombining gene drive system. gRNAs targeting heterozygotic loci are flanked by selection markers and two of three orthogonal Lox sites (colored triangles), which are inactivated through recombination (red triangles). Cas9 "drives" targeted mutations, whereas Cre-Lox recombination brings like markers to the same chromosome and activates a *URA3* gene interrupted by an artificial intron. After sporulation, the chromosome with gRNAs is selected using the markers of interest and the other is counterselected using 5-fluoroorotic acid (5-FOA). (**C**) Cross-design. A complete fitness landscape is produced in parallel by distinct cross designs that yield final homozygous diploids and haploids in biological replicates with unique DNA barcodes. (**D**) Bulk-fitness assays. Pooled strains are assayed in replicate for competitive fitness in several environments by sequencing barcodes to obtain strain frequencies over time. (**E**) Repeatability of technical replicate competitive fitness measurements. (**F**) Repeatability of biological replicate competitive fitness measurements.

We sought to use this method to construct a complete fitness landscape that would shed light on the structure of epistasis: Are FCTs primarily the product of a global coupling of mutations through fitness or do they emerge as the consequence of idiosyncratic epistasis? We surveyed studies of natural variation [e.g., (*32*–*36*)] and experimental evolution [e.g., (*37*–*39*)] to identify mutations potentially relevant to adaptation in the laboratory strain. We selected a set of mutations that sample a wide range of cellular functions, such as membrane stress response, mitochondrial stability, and nutrient sensing. Our goal in making this choice was to maximize fitness variance while minimizing pathway-specific idiosyncratic interactions. Alternative choices of mutations, particularly if they were focused on a specific protein or pathway (or limited to those that accumulated along the line of descent in a single lineage), might exhibit very different patterns of epistasis, which would be characteristic of the particular details of that specific protein or pathway (or that specific adaptive trajectory). However, our goal here was to analyze potentially global patterns of epistasis among mutations across the genome that are relevant to fitness in a variety of conditions and thus represent an overall fitness landscape for the laboratory strain.

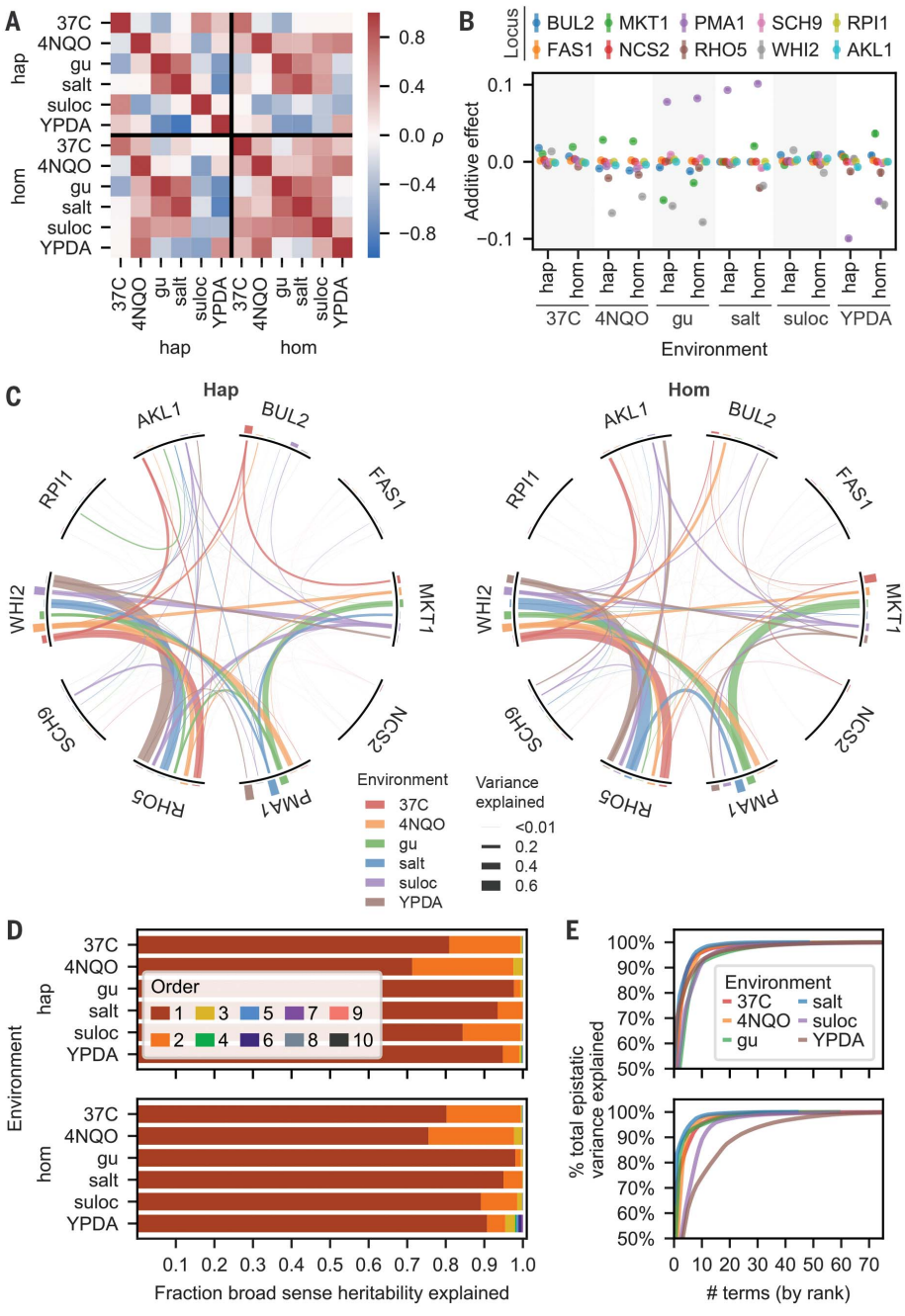

Fig. 2. Fitness landscapes. (**A**) Correlation in observed fitness (upper right) and predicted fitness (from inferred model, lower left; see the supplementary materials, section 5.1) across ploidies and environments. (**B**) Background-averaged additive effect of each locus across ploidies and environments. Error bars represent 95% confidence intervals. (**C**) Background-averaged pairwise epistatic effects between loci across ploidies and environments. Weights of edges connecting loci represent the proportion of pairwise variance explained by each interaction. Heights of bars on the perimeter correspond to the proportion of additive variance explained by each locus in each environment. (**D**) Variance partitioning of broad-sense heritability from additive and epistatic orders across ploidies and environments. (**E**) Cumulative distribution of the epistatic variance explained by rank-ordered epistatic terms of all orders.

We thus implemented our gene-drive system to construct a near-complete landscape spanning 10 missense mutations in 10 genes (including essential genes) on eight chromosomes: *AKL1* (S176P), *BUL2* (L883F), *FAS1* (G588A), *MKT1* (D30G), *NCS2* (H71L), *PMA1* (S234C), *RHO5* (G10S), *RPI1* (E102D), *SCH9* (P220S), and *WHI2* (L262S) (Fig. 1C and table S1). We found that a landscape of about this size is required to distinguish the two models (see the supplementary materials, section 6.3). Immediately before the final mating cycle, all strains were transformed with a unique DNA barcode next to the *LYS2* locus to enable high-throughput, sequencing-based competitive fitness assays (figs. S2 and S3). All strains in each replicate haploid library were genotyped at all 10 loci to confirm the presence of the desired alleles (this step also ensures presence in the diploid libraries). We excluded strains in which gene drive failure led to improper genotypes, such that ultimately 875 out of 1024 (85.4%) genotypes remained (468 were present in just one library, and 407 were present in both libraries). We also performed whole-genome sequencing of 96 randomly selected strains to rule out pervasive aneuploidies or influential but spurious background mutations. One aneuploidy was identified, and three spurious background mutations were observed at >5% frequency. Subsequent analysis showed that these were unlikely to systematically influence our findings (table S2 and supplementary materials, section 5.1).

To obtain fitness landscapes, we conducted replicate bulk barcode–based fitness assays on both pooled haploid and homozygous diploid versions of the genotype library in six distinctly stressful media environments: yeast extract peptone dextrose (YPD) + 0.4% acetic acid, YPD + 6 mM guanidium chloride, YPD + 35 μM suloctidil, YPD at 37°C, YPD + 0.8 M NaCl, and synthetic dextrose (SD) + 10 ng/ml 4-nitroquinoline 1-oxide (4-NQO) (Fig. 1D). For each of 7 days, pools were allowed seven

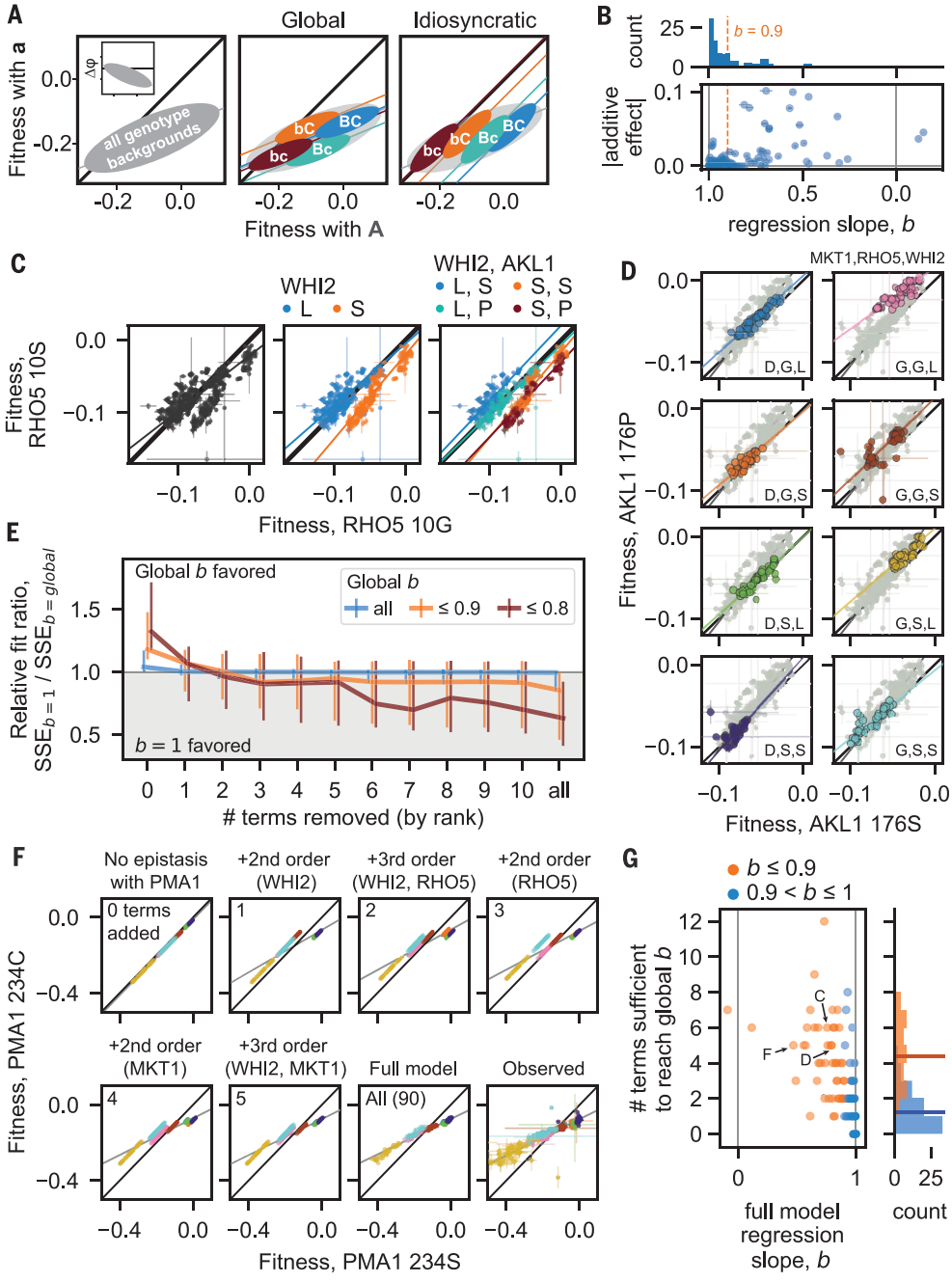

Fig. 3. FCTs. (**A**) Schematic contrasting how global or idiosyncratic epistasis could produce FCTs. Inset shows FCT analyzed as the effect of a mutation ($\Delta\varphi$) on backgrounds of different fitnesses. (**B**) Histogram and scatterplot of regression slopes (b) between φ_{Mut} and φ_{WT}, and corresponding absolute additive effects of mutations. Polarity was adopted such that $b \leq 1$. Total error bar length is twice the standard error of the slope. (**C**) Fitness effect of *RHO5* mutation (G10S) (φ_{Mut} versus φ_{WT}) in all haploid backgrounds at 37°C (left) and partitioned by genotypes at *WHI2* (L262S) (middle) and *WHI2* and *AKL1* (S176P) (right). Initial $SSE_{b=1}/SSE_{b=global}$ is 1.21. (**D**) Fitness effect of *AKL1* mutation in all homozygote backgrounds in the suloctidil environment, partitioned by genotypes at *MKT1* (D30G), *RHO5*, and *WHI2*. Initial $SSE_{b=1}/SSE_{b=global}$ = 1.31. (**E**) Median relative fit ratio between regressions with fixed slope of b = 1 and b = global as function of number of epistatic terms removed from observed phenotypes. Vertical lines represent interquartile range. Polarity was adopted such that $b \leq 1$. (**F**) Inferred fitness effect of *PMA1* S234C mutation in 4-NQO environment across all haploid backgrounds. Epistatic terms interacting with *PMA1* are completely removed from genotype fitnesses, then added back sequentially (from largest to smallest). Bottom-right: full-model (inferred) and observed genotype fitnesses, respectively. Gray line is regression slope. (**G**) Scatterplot and histograms of FCT regression slopes for all data and the number of epistatic terms sufficient to recapitulate them. Horizontal lines in the histogram indicate means. Arrows and letters indicate populations presented in previous panels. Polarity was adopted such that $b \leq 1$.

generations of growth, and aliquots were sampled and sequenced at the barcode locus at generations 7, 14, 28, 42, and 49. We estimated the relative fitness of each genotype from changes in barcode frequencies through time, achieving consistent measurements across technical and biological replicates (Fig. 1, E and F, and fig. S4). From these data, we inferred the background-averaged additive and epistatic effects of each mutation and combination of mutations, respectively (using LASSO regularization; see the supplementary materials).

We found that our six environments yielded substantially different landscapes, as demonstrated by the relatively low between-environment correlations of genotype fitnesses (Fig. 2A), the additive effects of each mutation (Fig. 2B), and the pairwise interactions between them (Fig. 2C). Haploid and homozygous diploid landscapes were largely correlated, but there were several notable exceptions, particularly in the suloctidil environment (Fig. 2, A and B). Although some pairwise interactions remain roughly constant in strength even as the corresponding additive effects varied considerably (e.g., *RHO5* and *WHI2*), most waxed and waned across environments (Fig. 2C). Nevertheless, the overall contribution from different epistatic orders showed some similarities across ploidies and environments (the magnitudes did differ; Fig. 2D), with additive and pairwise terms explaining most of the variance in the data, third-order terms contributing minorly, and the remaining orders making little difference, consistent with earlier studies (*40*). Across all epistatic orders, inferred effects were highly skewed, with a small number of terms explaining disproportionate variance (Fig. 2E).

We next sought to investigate potential patterns of global fitness–mediated epistasis. To do so, for each locus in each ploidy and environment, we plotted the fitness of a genotype

with the mutated allele (φ_{Mut}) against the fitness of the same genotype with the WT allele (φ_{WT}). A regression slope (b) different from 1 in these plots signifies an FCT (Fig. 3A, left; see the supplementary materials). Some previous work instead plotted the fitness effect of a mutation ($\Delta\varphi$) as a function of background fitness (φ_{WT}). The advantage of our formulation here is that it does not privilege a specific allele as the "WT." Instead, regression in our plots translates intuitively when reversing direction to treat the reversion as the mutation: $b_{rev} = 1/b_{orig}$ by weighted-total least squares (see the expanded discussion in the supplementary materials; figs. S5 to S8).

We found that FCTs were common in our landscapes: Across all ploidies, environments, and loci, ~44% of regression slopes deviated substantially from 1 (i.e., $b \leq 0.9$ or $b \geq 0.9^{-1}$; these deviations were all significant; Fig. 3B, histogram, and figs. S13 and S14). However, FCTs were not universal for fitness-affecting mutations: Of the 49 examples across ploidies and environments of mutations with additive effects of magnitude ≥ 0.5%, 18 were associated with $0.9 < b < 0.9^{-1}$ (Fig. 3B).

By partitioning background genotypes by the presence or absence of specific mutations, we could determine whether FCTs were truly "global" (i.e., whether they transcended these partitions and any corresponding idiosyncratic interactions; Fig. 3A, middle) or were instead fundamentally idiosyncratic (i.e., they emerged from regression across partitions shifted in φ_{Mut} versus φ_{WT} space by sparse interactions with specific background loci; Fig. 3A, right). When we partitioned FCTs by the presence or absence of interacting mutations in the background, we found several instances in which the idiosyncratic model clearly explained the FCT. For example, the effect of the G10S mutation in *RHO5* at 37°C exhibited a clear FCT (b = 0.76) (Fig. 3C). However, we could partition points by the presence of interacting *WHI2* and *AKL1* alleles in the background. Doing so showed that pairwise interactions with these alleles caused systematic shifts in φ_{10S} versus φ_{10G} space, with each partition assuming a slope near 1. Thus, over a range of background fitnesses, an FCT in the effect of the G10S emerged from these specific idiosyncratic interactions (Fig. 3C and fig. S11). In the case of the homozygous *AKL1* S176P mutation in suloctidil, we observed a similar decomposition of an FCT (b = 1.29) when partitioning genotypes according to the presence of three interacting loci in the background (*MKT1*, *RHO5*, and *WHI2*) (Fig. 3D and fig. S11). However, in other cases, it was less clear whether the FCT could be partitioned in this way, and because deeper partitions tend to reduce background fitness variance and limit our confidence in regression slopes, a different approach was required to characterize the extent to which idiosyncratic terms caused FCTs across our data.

To investigate this, we analyzed the effect of removing specific idiosyncratic epistatic terms on the overall FCTs. To do so, for each focal locus (in each ploidy and environment), we first calculated the weighted sum of squared errors (SSE) (*41*) of observed fitnesses from the global regression line ($SSE_{b=global}$) and from a fitted line of slope 1 ($SSE_{b=1}$, which corresponds to no FCT). We then set the largest epistatic term to zero and recalculated the expected fitness of each resulting genotype (assuming that all other terms and residuals were nonzero), again obtaining both $SSE_{b=global}$ and $SSE_{b=1}$. If the FCT arose from a global effect, then we would expect that $SSE_{b=global}$ would be less than $SSE_{b=1}$ even as terms were removed. Instead, we found that after removing the effect of just a few terms, a regression with a fixed slope of b = 1 typically fit the data better than the b = global FCT slope (with the FCT threshold set to $b \leq 0.9$ or 0.8; Fig. 3E and fig. S11), approaching the fit of an unconstrained regression that minimizes SSE (i.e., the final slope approaches 1; fig. S10). This indicates that the apparent FCT arises from these few idiosyncratic interactions, even for global slopes very different from 1. Although we also documented cases in which b = global fit the data better than b = 1 even after removing many terms, we expect that most, if not all, of these instances may have been due to measurement error because they tended to arise in ploidies and environments in which the data were noisier (fig. S17).

To further evaluate whether idiosyncratic interactions between these mutations were sufficient to generate FCTs, we performed the converse analysis, this time with genotype fitnesses as predicted by our model of additive and idiosyncratic epistatic terms. Instead of removing the effects of epistatic terms one at a time, we first stripped from the model all interactions involving the focal locus, yielding perfectly linear points of slope 1 when plotting φ_{Mut} versus φ_{WT}. We then added interactions one-by-one to our fitness prediction, from largest to smallest, and examined the resulting slopes. As shown in Fig. 3F for the haploid *PMA1* S234C mutation in 4-NQO, adding just a few terms associated with three background loci recapitulated a strong FCT. Repeating this analysis with all our mutations showed that, on average, just four idiosyncratic interactions (primarily pairwise) were sufficient to recapitulate the full-model FCTs (a slope within 0.01 of the global slope; Fig. 3G, orange; see the supplementary materials), which is far lower than the total number of inferred terms (median of 53) but represents on average 89% of the potential variance explained that could have been added (fig. S12). Thus, although FCTs are real and likely have important biological consequences, our data demonstrate that apparent fitness-mediated epistasis can readily emerge from very few low-order idiosyncratic interactions.

Because the landscapes that we studied here have no natural polarization (i.e., neither allele is the assumed WT), we cannot comment directly on why earlier studies of global epistasis have more commonly found negative than positive FCTs (when plotting $\Delta\varphi$ versus φ_{WT}). However, this distribution of FCT directions is important because it may underly the ubiquitous trend of declining adaptability observed across laboratory evolution experiments (*29*). The observed bias toward negative trends may arise from asymmetries in the average sign of epistatic interactions between mutations away from extant high-fitness genotypes relative to their reversions, which theory has predicted should arise from idiosyncratic interactions (*19*, *20*). In addition, choosing polarizations at random will lead to more negative than positive FCTs across the full parameter space (see the extended discussion in the supplementary materials).

Regardless of the cause of any asymmetry in the direction of FCTs, our results support recent theoretical arguments that fitness-mediated epistasis can emerge as the generic consequence of widespread idiosyncratic interactions, rather than reflecting a global fitness–mediated coupling of mutations. Indeed, at least in our system, FCTs could arise even from a relatively small number of low-order interactions. We note that landscapes involving other types of variation [e.g., within a single protein or pathway or along the line of descent in a single lineage (*21*)] may exhibit different patterns, although we may expect these scenarios to involve an even stronger role for idiosyncratic interactions. More generally, we emphasize that idiosyncratic epistasis and global fitness–mediated effects are not mutually exclusive, and although FCTs can be explained by the former in our system, in other cases, both effects may contribute. However, our results suggest that nonspecific global epistasis may not be the primary driver of patterns of declining adaptability in laboratory evolution experiments, and this has general implications for the ways in which epistasis constrains evolutionary trajectories.

REFERENCES AND NOTES

1. D. J. Kvitek, G. Sherlock, *PLOS Genet.* **7**, e1002056 (2011).
2. P. A. Romero, F. H. Arnold, *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).
3. K. M. Flynn, T. F. Cooper, F. B.-G. Moore, V. S. Cooper, *PLOS Genet.* **9**, e1003426 (2013).
4. M. Costanzo *et al.*, *Science* **353**, aaf1420 (2016).
5. A. D. Norris, X. Gracida, J. A. Calarco, *eLife* **6**, e28129 (2017).
6. M. A. Horlbeck *et al.*, *Cell* **174**, 953–967.e22 (2018).
7. K. S. Sarkisyan *et al.*, *Nature* **533**, 397–401 (2016).
8. C. A. Olson, N. C. Wu, R. Sun, *Curr. Biol.* **24**, 2643–2651 (2014).

9. S. Bershtein, M. Segal, R. Bekerman, N. Tokuriki, D. S. Tawfik, . *Nature* **444**, 929–932 (2006).
10. N. C. Wu *et al*., *Nat. Commun.* **11**, 1233 (2020).
11. V. O. Pokusaeva *et al*., *PLOS Genet.* **15**, e1008079 (2019).
12. J. I. Jiménez, R. Xulvi-Brunet, G. W. Campbell, R. Turk-MacLeod, I. A. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14984–14989 (2013).
13. B. Madan *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011653118 (2021).
14. N. C. Wu, L. Dai, C. A. Olson, J. O. Lloyd-Smith, R. Sun, *eLife* **5**, e16965 (2016).
15. C. Bank, S. Matuszewski, R. T. Hietpas, J. D. Jensen, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14085–14090 (2016).
16. J. Domingo, G. Diss, B. Lehner, *Nature* **558**, 117–121 (2018).
17. J. Teyra, A. Ernst, A. Singer, F. Sicheri, S. S. Sidhu, *Protein Sci.* **29**, 433–442 (2020).
18. T. V. Lite *et al*., *eLife* **9**, e60924 (2020).
19. D. M. Lyons, Z. Zou, H. Xu, J. Zhang, *Nat. Ecol. Evol.* **4**, 1685–1693 (2020).
20. G. Reddy, M. M. Desai, *eLife* **10**, e64740 (2021).
21. A. I. Khan, D. M. Dinh, D. Schneider, R. E. Lenski, T. F. Cooper, *Science* **332**, 1193–1196 (2011).
22. H.-H. Chou, H.-C. Chiu, N. F. Delaney, D. Segrè, C. J. Marx, *Science* **332**, 1190–1192 (2011).
23. L. Perfeito, A. Sousa, T. Bataillon, I. Gordo, *Evolution* **68**, 150–162 (2014).
24. S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, *Science* **344**, 1519–1522 (2014).
25. M. S. Johnson, A. Martsul, S. Kryazhimskiy, M. M. Desai, *Science* **366**, 490–493 (2019).
26. X. Wei, J. Zhang, *Mol. Biol. Evol.* **36**, 1008–1021 (2019).
27. S. Schoustra, S. Hwang, J. Krug, J. A. G. M. de Visser, *Proc. Biol. Sci.* **283**, 20161376 (2016).
28. R. C. MacLean, G. G. Perron, A. Gardner, *Genetics* **186**, 1345–1354 (2010).
29. A. Couce, O. A. Tenaillon, *Front. Genet.* **6**, 99 (2015).
30. S. Kryazhimskiy, *eLife* **10**, e60200 (2021).
31. A. N. Nguyen Ba *et al*., *Nature* **575**, 494–499 (2019).
32. H. Sinha *et al*., *Genetics* **180**, 1661–1670 (2008).
33. S. W. Doniger *et al*., *PLOS Genet.* **4**, e1000183 (2008).
34. E. X. Kwan, E. Foss, L. Kruglyak, A. Bedalov, *PLOS Genet.* **7**, e1002250 (2011).
35. J. S. Bloom, I. M. Ehrenreich, W. T. Loo, T.-L. V. Lite, L. Kruglyak, *Nature* **494**, 234–237 (2013).
36. D. M. Wloch-Salamon, K. Tomala, D. Aggeli, B. Dunn, *G3* **7**, 1899–1911 (2017).
37. B. Szamecz *et al*., *PLOS Biol.* **12**, e1001935 (2014).
38. L. M. Kohn, J. B. Anderson, *Eukaryot. Cell* **13**, 1200–1206 (2014).
39. S. Venkataram *et al*., *Cell* **166**, 1585–1596.e22 (2016).
40. D. M. Weinreich, Y. Lan, J. Jaffe, R. B. Heckendorn, *J. Stat. Phys.* **172**, 208–225 (2018).
41. D. York, N. M. Evensen, M. L. Martínez, J. De Basabe Delgado, *Am. J. Phys.* **72**, 367–375 (2004).
42. C. W. Bakerlee, A. N. Nguyen Ba, Fitness-correlated trends analysis pipeline publication archive (Zenodo, 2022; https://zenodo.org/record/6352707).

ACKNOWLEDGMENTS

We thank the Bauer Core facility at Harvard; G. Reddy, B. Shraiman, and members of the Desai laboratory for experimental assistance and comments on the manuscript; and A. Rego-Costa for help in creating Fig. 2C. Computational work was performed on the Odyssey cluster supported by the Research Computing Group at Harvard University. **Funding:** This work was supported by the Department of Defense (DoD), National Defense Science & Engineering Graduate (NDSEG) Fellowship Program (C.W.B.); the Natural Sciences and Engineering Research Council of Canada (postdoctoral fellowship to A.N.N.B.); Discovery Grant RGPIN-2021-02716 and Discovery Launch supplement DGECR-2021-00117 (A.N.N.B.); the National Human Genome Research Institute of the National Institutes of Health (award F31-HG010984 to Y.S.); the National Science Foundation (grant PHY-1914916 to M.M.D.); and the National Institutes of Health (grant GM104239 to M.M.D.). **Author contributions:** Conceptualization: A.N.N.B., C.W.B., Y.S., J.I.R., M.M.D.; Formal analysis: C.W.B., A.N.N.B.; Funding acquisition: M.M.D.; Investigation: C.W.B., A.N.N.B.; Methodology: A.N.N.B., C.W.B., Y.S., J.I.R.; Resources: M.M.D.; Supervision: M.M.D.; Visualization: C.W.B., A.N.N.B.; Writing – original draft: C.W.B., A.N.N.B., M.M.D.; Writing – review and editing: C.W.B., A.N.N.B., M.M.D. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw sequencing data are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (accession no. PRJNA815849), and analysis code is available from Github (*42*). All other data are presented in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm4774
Methods and Supplementary Analysis
Figs. S1 to S17
Table S1
References (*43*–*59*)
Data S1 to S3
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

Submitted 22 September 2021; resubmitted 9 January 2022
Accepted 1 April 2022
10.1126/science.abm4774

Idiosyncratic epistasis leads to global fitness–correlated trends

Christopher W. BakerleeAlex N. Nguyen BaYekaterina ShulginaJose I. Rojas EcheniqueMichael M. Desai

Science, 376 (6593), • DOI: 10.1126/science.abm4774

Gene interactions and global fitness

The effect of a particular mutation in a given environment often varies systematically with the fitness of the genetic background in which it arises, but there are different theories as to how such systematic trends manifest. Using a CRISPR gene drive system, Bakerlee *et al*. generated a combinatorial yeast library that allowed them to test the effects of different combinations of missense mutations in 10 genes, including essential ones. Using this system, the authors found that fitness-correlated trends were better explained by an idiosyncratic model, in which they emerged from just a few interactions between specific loci, than they were by a global epistasis model, in which the effects of mutations were generally mediated by background fitness. —MAF