Mapping the Peaks: Fitness Landscapes of the Fittest and the Flattest

Joshua Franklin
Michigan State University
  Department of Microbiology and
    Molecular Genetics

Thomas LaBar*
Harvard University
  Department of Molecular and
    Cellular Biology
  thomas_labar@fas.harvard.edu
Michigan State University
  BEACON Center for the Study of
    Evolution in Action

Christoph Adami
Michigan State University
  Department of Microbiology and
    Molecular Genetics
  Department of Ecology, Evolutionary
    Biology, and Behavior
  BEACON Center for the
    Study of Evolution in Action
Arizona State University
  Department of Physics and
    Astronomy

Abstract   Populations exposed to a high mutation rate harbor abundant deleterious genetic variation, leading to depressed mean fitness. This reduction in mean fitness presents an opportunity for selection to restore fitness through the evolution of mutational robustness. In extreme cases, selection for mutational robustness can lead to *flat* genotypes (with low fitness but high robustness) outcompeting *fit* genotypes (with high fitness but low robustness)—a phenomenon known as *survival of the flattest*. While this effect was previously explored using the digital evolution system Avida, a complete analysis of the local fitness landscapes of fit and flat genotypes has been lacking, leading to uncertainty about the genetic basis of the survival-of-the-flattest effect. Here, we repeated the survival-of-the-flattest study and analyzed the mutational neighborhoods of fit and flat genotypes. We found that the flat genotypes, compared to the fit genotypes, had a reduced likelihood of deleterious mutations as well as an increased likelihood of neutral and, surprisingly, of lethal mutations. This trend holds for mutants one to four substitutions away from the wild-type sequence. We also found that flat genotypes have, on average, no epistasis between mutations, while fit genotypes have, on average, positive epistasis. Our results demonstrate that the genetic causes of mutational robustness on complex fitness landscapes are multifaceted. While the traditional idea of the survival of the flattest emphasized the evolution of increased neutrality, others have argued for increased mutational sensitivity in response to strong mutational loads. Our results show that both increased neutrality and increased lethality can lead to the evolution of mutational robustness. Furthermore, strong negative epistasis is not required for mutational sensitivity to lead to mutational robustness. Overall, these results suggest that mutational robustness is achieved by minimizing *heritable* deleterious variation.

1   Background

All well-adapted populations receive an influx of deleterious variation every generation due to *de novo* mutations. This influx reduces the population's average fitness and causes a mutational load [2, 11]. For populations with high mutation rates, deleterious variation can alter the dynamics of natural

---

 * Corresponding author.

selection and lead to selection acting not just on individual genotypes, but on genotypes and their likely mutant offspring [5, 6, 31, 42]. This combination of a wild-type sequence and its mutants is known as a *quasispecies* and is thought to be a unit of selection in organisms such as the first RNA replicators [14, 15] and modern-day RNA viruses [27].

Since the quasispecies concept was first proposed [14, 15], many studies have analyzed the conditions and consequences of evolutionary dynamics in the strong mutational regime (e.g., [4, 23, 42]). One consequence of evolution under high mutation rates is that populations are expected to evolve mutational robustness by fixing mutations that decrease the likelihood (and/or decrease the average effect) of deleterious mutations [12, 44]. In the extreme case, this trend can lead to genotypes with lower fitness outcompeting genotypes with higher fitness if the lower-fitness genotypes are more mutationally robust [37]. This effect has been termed the *survival-of-the-flattest* effect, where *flat* genotypes, that is, those with lower fitness but greater mutational robustness, outcompete *fit* genotypes, meaning those with higher fitness but decreased mutational robustness [45]. This effect was first observed in the digital evolution system Avida; it was later demonstrated with experimental evolution of RNA viruses [8, 28, 36].

Even though the evolutionary dynamics of the survival-of-the-flattest effect are well understood, the exact genetic causes of robustness are unknown. Indeed, the differences in the distribution of fitness effects [17] (that is, the local fitness landscapes) between fit genotypes and flat genotypes have not been explored. While it is generally assumed that mutational robustness evolves due to the evolution of an increased number of neutral mutational neighbors [18, 39, 41], other authors have argued that increased mutational sensitivity may evolve [3, 23, 26, 32, 46]. Additionally, recent work has indicated that populations evolving under strong genetic drift are expected to evolve *drift-robust* genetic architectures, giving rise to genotypes with a decreased likelihood of small-effect deleterious mutations and an increased likelihood of neutral and/or strongly deleterious mutations [25]. A mapping of the local fitness landscapes of fit genotypes and flat genotypes is needed to compare the survival-of-the-flattest effect with these other proposed evolutionary mechanisms.

To explore the local fitness landscapes of fit genotypes and flat genotypes, we first repeated the original survival-of-the-flattest experiment. We then took genotypes adapted to either low or high mutation rates and obtained the distribution of fitness effects for these genotypes by examining mutants up to four substitutions away from the wild-type sequence. We found that flat genotypes had a greater likelihood of beneficial and neutral mutations (as expected for mutationally robust genotypes), while fit genotypes had a greater likelihood of deleterious mutations, as expected from mutationally sensitive genotypes. Contrary to expectations, we also found that flat genotypes had a greater likelihood of lethal mutations, suggesting that the survival-of-the-flattest effect relies on a reduction of heritable deleterious mutation, which can be achieved both by an increase in neutral variations and by an increase of lethal mutations. These results illustrate the complexity of mutationally robust genome architectures on nontrivial fitness landscapes.

2   Materials and Methods

2.1   Avida

Here we describe the relevant details of Avida (version 2.14; our version, with added code for data analysis, is available at https://github.com/joshf93/MappingThePeaks) for the current experiments (see [33] for a full overview of the Avida software). In Avida, a finite population of computer programs (*Avidians*) compete for the resources (memory, space, and processing time) required for reproduction. Each program consists of a circular genome of computer instructions that encode the ability to self-replicate. During this replication process, instructions may be copied inaccurately, resulting in mutations that are passed onto an Avidian's offspring. As Avidians with faster replication speeds enter the population, they reproduce faster than slower-replicating Avidians, leading to the spread of faster replicators and the extinction of slower replicators. In other words, there is selection for faster replicators. Therefore, because there is heritable variation and selection in Avida, populations of

Avidians undergo evolution by natural selection [34]. Avida has been used to test many evolutionary concepts difficult to test in biological systems [1, 10, 19, 20, 24, 29, 38].

The Avida world consists of a toroidal grid of N cells; each cell can be occupied by at most one Avidian, and thus N is the maximum population size. During reproduction, each offspring Avidian is placed in one of the nine adjacent cells of its parent, including the cell that contains the parent. If some of these cells are unoccupied, a random empty cell is chosen. If all neighboring cells are occupied, a random cell is chosen, its occupant is removed, and the new Avidian then occupies the cell. This random selection of replacement adds an element of genetic drift to Avida.

The limited number of cells in the Avida environment is one of the limiting resources over which Avidians compete. The other limiting resource is the opportunity to execute the instructions in an Avidian's genome. The resource required to execute one instruction is called a *single instruction processing* unit (SIP). During each unit of time in Avida (called an update), $30N$ SIPs are available to the population for execution. These SIPs are probabilistically distributed to Avidians according to a figure of merit that is related to their ability to perform certain computations. These computations (binary logic operation) are the *traits* that Avidians display. In a monoclonal population, all Avidians express the same traits and thus have the same merit and thus receive on average 30 SIPs (i.e., they execute 30 genome instructions during that update). However, Avidian lineages can evolve the ability to increase merit, and thus increase the number of instructions that can be executed per update.

Fitness in Avida is implicit and can only be estimated by executing the computer instructions within an Avidian's genome. In this regard, Avida differs from classical evolutionary simulations where genotypes are assigned a fitness and their frequency increases or decreases proportionally to this assigned value. Here, faster replicators outcompete slower replicators and fitness emerges from this dynamic. There are two means by which an Avidian lineage can evolve increased replication speed. First, genotypes that require fewer instruction executions to undergo reproduction will reproduce faster than Avidians that require a greater number of instruction executions. In environments where genome size can vary, this is often achieved by decreasing the genome size. In fixed-genome-size environments, this trajectory of fitness gain still occurs, often through the fixation of instructions that copy more genome information from parent to offspring within a single execution of the copy loop. The second method of fitness increase in Avida is through the evolution of complex adaptations: the ability to perform one- and two-input Boolean logic calculations (the traits discussed earlier). In the setup used here, nine of these functions are potentially rewarded with SIPs; the number of SIPs awarded is proportional to the complexity of the evolved trait [30].[1] In other words, the more calculations an Avidian can perform, the more SIPs it receives per update, and the larger the portion of its genome it can then execute per unit time. This increases the Avidian's replication speed and leads to selection for the ability to perform these calculations. When Avida estimates a genotype's fitness, it is estimated as the genotype's merit divided by the number of instruction executions required for replication (otherwise known as its gestation time).

Avidian populations evolve these complex adaptations by fixing a sequence of mutations. This genetic variation is introduced into a population when an Avidian copies one instruction inaccurately into its daughter's genome. The rate at which these mutations occur is set by the experimenter and is usually set as a rate per instruction copied (a copy-mutation rate). When mutations occur during this copying process, multiple mutations can occur per replication event. Additionally, in experiments where genome size can evolve, insertion and deletion mutations occur upon division (not during the replication process) at a prespecified rate. These *indel* mutations insert or delete one instruction. Finally, large-scale genome changes can occur, but these changes are caused by the specific replication algorithm encoded by an Avidian's genome; these mutations are inherent in the Avida genetic code (implicit mutations) and are not under the control of the experimenter.

---

1 The performance of a calculation increases an Avidian's merit as a power of the number of operations needed to perform the calculation, with increasing powers for calculations of increasing difficulty. These powers are the default Avida setting. The exact values appear in the config file that is available on the github repository.

2.2   Experimental Design

Here, we repeated the experimental protocol of the original survival-of-the-flattest study [45]. We differed from the original study in one aspect: the placement of offspring. Here, we used Avida's current default setting, where a new offspring randomly replaces one of the nine neighbors of its parent (including possibly the parent). The original article used mass-action reproduction, where new offspring could randomly replace any offspring in the population. We first generated a collection of genotypes during the *initial adaptation* step. We evolved 201 populations of 3,600 individuals for 5×10^4 updates.[2]

All populations initially started with the default Avida ancestor with a 100-instruction genome. Point mutations occurred at a rate of 0.0075 mutations per instruction copied, and insertion/deletion mutations occurred at a rate of 0.05 mutations per division each. At the end of these populations' evolution, we extracted the most abundant genotype.

For the *mutation-rate adaptation* phase of the experiment, we duplicated each of the 201 genomes and used one as progenitor of an experiment at a fixed *low* genomic mutation rate of 0.5 mutations per genome per generation, and one at a fixed *high* genomic rate of 2.0 mutations per genome per generation. To achieve this, we adjusted the per-site mutation rate of each genome so that the targeted genomic rate was achieved, and disallowed sequence length changes. In this manner, the initial adaptation created genetic variation used as input to the adaptation step at a fixed genomic rate. From this stage onward each genotype was limited to the tasks it could complete at the end of the initial adaptation phase—that is, novel traits were not rewarded (and as a consequence could not become fixed in populations of this size). However, they could lose traits they already possessed, and could re-evolve them subsequently. Each population adapting to a high or low mutation rate consisted of 3,600 individuals and evolved for 10^4 generations. After evolution we again selected the most abundant genotype from each population and classified genotypes from the low-mutation environment as fit genotypes, and genotypes from the high-mutation environment as flat genotypes.

Next, we performed competition experiments between each pair of fit and flat genotypes (each pair has the same ancestor). We only performed these competitions with genotype pairs where the fit genotype was 1.5 times fitter than the flat genotype. For the competitions, we seeded populations of 3,600 individuals with 1,800 individuals of each genotype and ran the simulation for 200 generations. We repeated these competitions across a range of genomic mutation rates (0.5 to 3.0 mutations per genome per generation in increments of 0.5) and performed five replicates per mutation rate. We tracked the abundance of the fit genotypes over time to determine the outcome of the competition. We classified the flat genotype as the winner if it reached at least 95% of the population in three of five replicates at a given mutation rate.

2.3   Data Analysis

We performed mutational analyses to determine the fitness landscapes of the fit and flat genotypes used in the competitions, using Avida's *analyze* mode. In analyze mode, Avidians run through their life cycle in isolation (rather than in a population), and characteristics such as their fitness can be estimated. We first estimated the base fitness of all Avidian genotypes used in the experiment. Then, we took the genotypes used in the competition and generated all possible single mutants. For each genotype, there are $25L$ single mutants, where L is the genome size and 25 is the number of possible alternative alleles at each position in the genome, as there are 26 instructions in the Avida instruction set. We then repeated this procedure for all double mutants. For triple and quadruple mutants, we randomly sampled 10^8 mutants, as the number of all possible mutants made generating every mutant computationally prohibitive.[3]

---

2 While all experiments performed on the organisms obtained in this step are performed by measuring time in "generations," we chose 5×10^4 *updates* for the adaptation step, because that was the time for adaptation in [29], which gave rise to the organisms used in [45].

3 The number of possible *n*-mutants of a length L genome encoded in an alphabet of size D is $\binom{L}{n}(D-1)^n$, which translates to over 1.5×10^{12} possible mutants for $L = 100$, $n = 4$, and $D = 26$.

After generating these mutants, we calculated their fitness and compared them with their ancestor genotype. Mutants with greater fitness were designated as beneficial, and mutants with the same fitness were designated as neutral. Mutants that were viable but with decreased fitness were designated as deleterious. Mutations resulting in genomes that could not reproduce were classified as lethal. To estimate differences in epistasis between fit and flat genotypes, we fitted the mutational analysis data with the following equation:

$$\log\left(\frac{w(n)}{w_0}\right) = -\alpha n^{\beta} \tag{1}$$

Here, n is the genetic distance from the initial genotype, $w(n)$ is the mean fitness of all mutants at a given genetic distance n, and w_0 is the fitness of the initial genotype. This is a similar equation to that used in the original survival-of-the-flattest article [45], which used the mean fitness estimated at a given mutation rate instead of the mean fitness of all n-mutants (which is often used [29]). The quantities α and β can be thought of as parameters that measure a genotype's mutational robustness and epistasis, respectively. In the absence of epistasis ($\beta = 1$), mutational robustness is solely determined by α: The smaller α, the more robust the organism. When epistasis is present ($\beta \neq 1$), both α and β combine to establish robustness, which can be defined as the ability to maintain high fitness under severe mutational pressure, as compared to an organism that is not robust. A $\beta > 1$ indicates epistasis that is *negative* (i.e., the fitness effect of a double mutant decreases fitness more than expected from the fitness deficit of the two single mutants), while $\beta < 1$ represents *positive* epistasis due to the buffering of the effects of the individual mutations. To estimate α and β, we performed a nonlinear fit of $w(n)$ up to $n = 4$, which gives excellent results in that β never differs from the multiplicative value $\beta = 1$ by more than 5%, which implies that the fits are nearly linear. To assess goodness of fit, we provide histograms of the residual standard error of all fits (for both the fit and the flat cases; see Figure 6 in the Appendix), which shows that the residuals never exceed 0.03. Note, however, that the median of residuals was larger for the fit peaks than for the flat peaks (see Figure 7 in the Appendix). All statistical analyses were performed using R 3.4.3 [35], and figures were generated with either the ggplot2 R package [40] or the Matplotlib Python package [21].

3   Results

We first evolved 201 populations and isolated the most abundant genotype from each population. We then evolved these genotypes in both a high-mutation-rate environment and a low-mutation-rate environment and isolated the most abundant genotype from each environment. The genotypes from the low-mutation-rate environment did not significantly change in fitness compared to their ancestor (Figure 1(a); Wilcoxon rank sum test, $p = 0.411$). The genotypes from the high mutation rate environment significantly decreased in fitness compared to their ancestors (Figure 1(a); $p = 2.216 \times 10^{-7}$). By tracking the fitness of the high-mutation-rate populations through time, we observe that these genotypes first severely decrease in fitness upon transfer, but then recover fitness as they search out new, and presumably more mutationally robust, fitness peaks (Figure 1(b)).

Of our 201 pairs of genotypes adapted to low mutation rates and high mutation rates, 166 pairs evolved so that the low mutation rate genotype (the fit genotype) had a fitness greater than 1.5 times the high-mutation-rate genotype (the flat genotype), which is the threshold used in the original study [45]. We took these pairs and performed competition experiments between the two genotypes across a range of mutation rates. In 76 (45.8%) of these competition experiments we noted that the eventual winner of the competition depended on the mutation rate, and that a switch between winners occurred at a critical mutation rate, just as was observed in the initial study [45]. Fit genotypes win the competition at low mutation rates, but flat genotypes win the competition at high mutation rates even with the large fitness difference (Figure 2). It is worth noting that our definition of winning excluded many genotype pairs that showed the effect to a lesser degree.
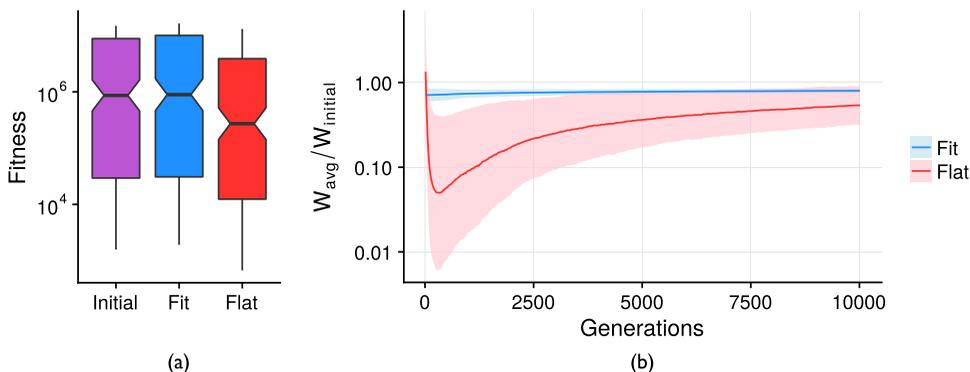
Figure 1. Populations adapt to different fitness peaks in high-mutation-rate environments. (a) Fitness values for the most abundant genotype from each population after the initial adaptation experiment (denoted by "Initial") and after adaptation to low ("Fit") and high ("Flat") mutation-rate environments. For boxplots, lines are the median values, boxes show the interquartile range, and whiskers are 1.5 times the interquartile range or the minimum or maximum value; points are outliers. Overlapping notches suggest that the difference in medians is not significant. (b) Relative average fitness values over the course of the mutation-rate adaptation experiment. Lines are the mean value across all populations, and the shaded regions are standard deviations. Blue represents fit genotypes and red represents flat genotypes.

Next, we analyzed the mutational neighborhood (the local fitness landscape) of all 166 pairs of fit and flat genotypes. We first generated all genotypes containing one point mutation and measured their fitness. There are significant differences in particular classes of mutations. Flat genotypes have a greater likelihood of beneficial (Figure 3(a); Wilcoxon rank sum test, $p = 1.35 \times 10^{-20}$), neutral
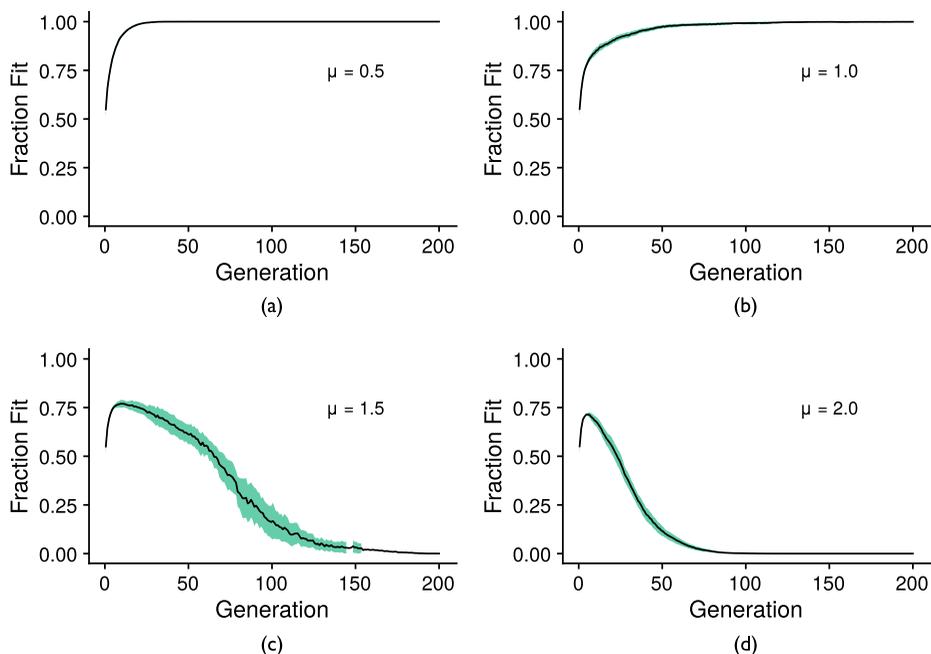

Figure 2. Competition outcome between fit and flat genotypes is determined by the mutation rate. Outcome from competition experiments between one fit-flat genotype pair. Black line is the mean fraction of the population consisting of the fit genotype as a function of time (measured in generations). Green shading represents one standard deviation. Each subplot shows the results for a given genomic mutation rate for five replicates. (a) $\mu = 0.5$, (b) $\mu = 1.0$, (c) $\mu = 1.5$, (d) $\mu = 2.0$.
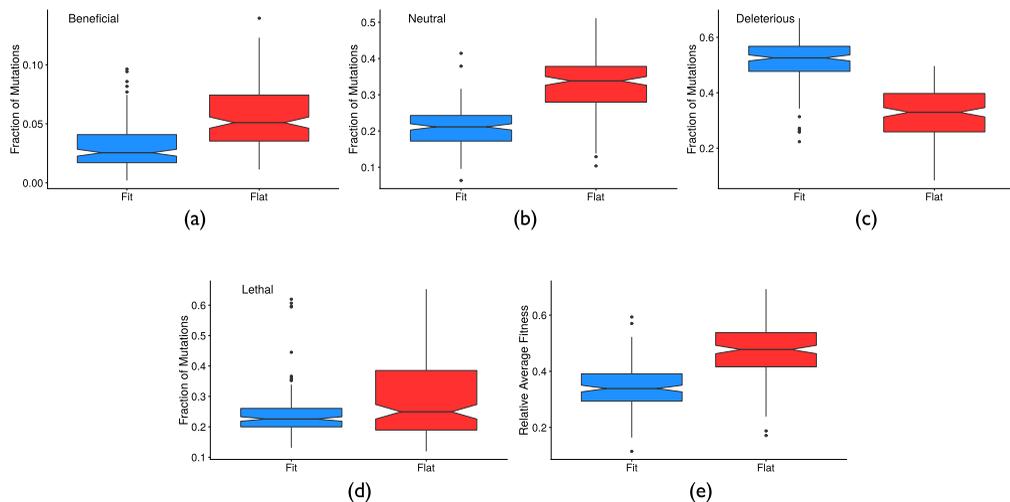
Figure 3. Description of local fitness landscapes for fit and flat genotypes. (a)–(d) Distribution of fitness effects with mutations grouped by their broad effect. Boxplots as previously described. Color refers to whether the data are for fit genotypes (blue) or flat genotypes (red). (e) Average relative fitness of all one-step mutants for each fit and flat genotype.

(Figure 3(b); $p = 2.66 \times 10^{-35}$), and lethal (Figure 3(d); $p = 0.0168$) mutations. Fit organisms have a greater likelihood of deleterious mutations (Figure 3(c); $p = 1.31 \times 10^{-44}$). Additionally, fit genotypes have lower mean relative fitness than flat genotypes (Figure 3(e); $p = 2.24 \times 10^{-27}$).

Furthermore, we analyzed the mutational neighborhood of these genotypes for sequences that are two, three, and four mutations away from the wild-type sequence in order to estimate differences in epistasis between fit and flat genotypes. We found that the fraction of lethal mutants increased
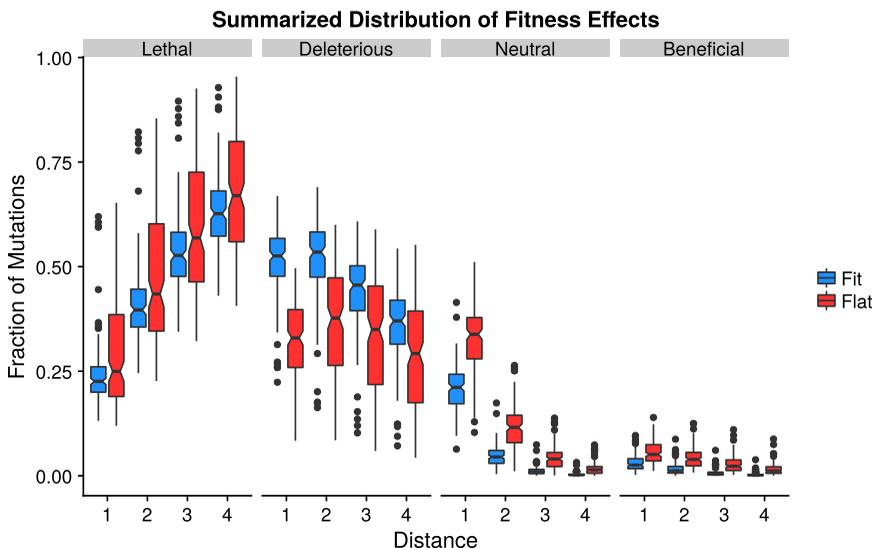

Figure 4. Summarized distribution of fitness effects for mutants with multiple mutations. Proportion of mutations with a mutational effect classification for mutants up to four mutations away from the wild-type sequence. See text for definitions of different classes of mutations. Distance is the number of mutations away from the wild-type sequence. Data for mutants with distance of one are the same as in Figure 3. Blue (red) boxplots are the proportions from the fit (flat) genotypes. Boxplots as previously described.
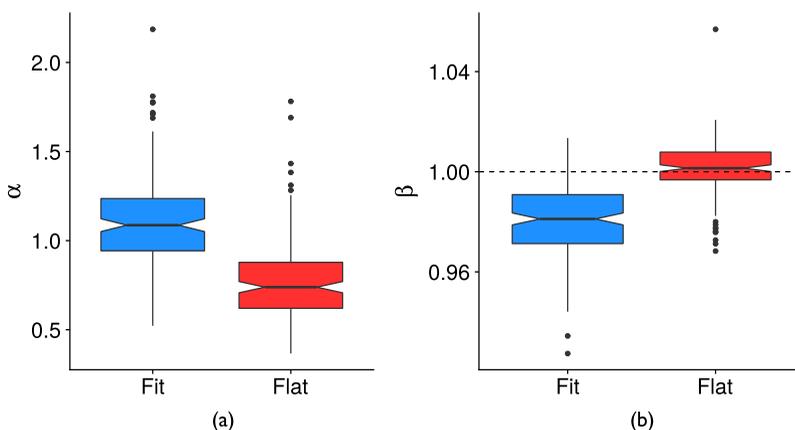
Figure 5. Distributions of α and β values for fit and flat genotypes for each treatment. α is a measure of single-mutant fitness, while β is a measure of epistasis (see Section 2 for further details). Boxplots as previously described.

with mutational distance for both the fit and the flat genotypes, with the flat genotypes always show-ing significantly more lethals than the fit types (see Figure 4). The fraction of deleterious mutations decreases with increasing mutational distance, with the fit types showing a higher fraction of dele-terious mutations than the flat types for all distances we tested.

As expected from the data on the relative fitness of one-step mutants (Figure 3(e)), fit genotypes have greater α values (Figure 5(a); $p = 1.06 \times 10^{-27}$), that is, they are less robust to mutations. Flat genotypes have greater β values (Figure 5(b); $p = 5.44 \times 10^{-34}$), indicating that flat genotypes have somewhat negative epistasis ($\beta > 1$) compared to the fit genotypes, whose average β is solidly in the positive epistasis territory ($\beta < 1$). However, it should be noted that flat genotypes have, on average, almost no epistasis (β very close to 1). These data demonstrate how the selection pressure for mutational robustness drives populations to alternative areas of the fitness landscape.

4   Discussion

We used digital experimental evolution to map the local fitness landscape of high-fitness, but mutationally fragile, fit genotypes adapted under low mutation rates, and low-fitness, but mutation-ally robust, flat genotypes adapted under high mutation rates. After repeating the original survival-of-the-flattest effect [45], we calculated the distribution of fitness effects for *de novo* point mutations up to four substitutions from the wild-type genotypes. We found that, as expected, flat genotypes were more mutationally robust than fit genotypes. Flat genotypes had a lower likelihood of deleterious mutations and a greater likelihood of neutral and beneficial mutations than fit genotypes. Surpris-ingly, flat genotypes also had a greater likelihood of lethal mutations and more negative epistasis than fit genotypes, suggesting that the factors responsible for mutational robustness involve only limiting *heritable* deleterious variation.

Around the time when the original survival-of-the-flattest article first appeared, mutational robustness was largely conceptualized as a decrease in the likelihood of a deleterious mutation [39, 45]. And while these changes in genetic architecture can lead to mutational robustness, work since that original publication has suggested that it is *increases* in the effect of deleterious mutations, not decreases, that lead to robustness [3, 32]. Those authors argue that increased severity of dele-terious mutations can lead to population-level robustness because purifying selection is more effec-tive against strongly deleterious mutations. It has also been proposed that increased negative (or synergistic) epistasis can evolve as a mechanism to increase the strength of purifying selection and hence robustness [26, 46].

Our results suggest that *all* of these genetic mechanisms (increased neutrality, increased severity of deleterious mutations, and increased negative epistasis) can contribute to mutational robustness (although in this study the "increased negative epistasis" really is more of an absence of positive epistasis). These data support a previously proposed relationship between increased neutrality (low α) and negative epistasis (β > 1), based on data from Avida, RNA folding, and neutral-network-fitness landscapes [43]. In other words, a population can only maximize immediate mutational robustness at the cost of increased deleterious severity of mutants farther away in the fitness landscape. We should also note that mutations on a flat genetic background are not, on average, negatively epistatic. Flat genotypes evolve to have, on average, no epistasis between mutations, as previously shown to occur in high-mutation-rate Avida populations [13]. However, they are more negatively epistatic than fit genotypes, which have positive epistasis between mutations. The lack of negative epistatic interactions in flat genotypes is likely related to the increased likelihood of lethal mutations. Additional mutations on a genetic background with a lethal mutation cannot, by definition, lead to a greater decline in fitness (negative epistasis).

Theoretical arguments suggest that increasing the severity of mutations not only protects against high mutation rates but also protects against fitness loss due to genetic drift [22], although this depends on the specific distribution of fitness effects [7]. Recently, LaBar and Adami proposed the concept of *drift robustness*: the idea that small populations tend to populate drift-robust fitness peaks with a low likelihood of slightly deleterious mutations (and a high likelihood of both neutral and strongly deleterious mutations), while large populations would adapt to drift-fragile fitness peaks with high likelihood of slightly deleterious mutations [25]. Both flat genotypes and drift-robust genotypes have some noticeable similarities in their distribution of fitness effects (increased neutrality, in particular), so it is worth asking whether drift robustness and survival of the flattest are really the same phenomenon.

The main difference between drift robustness and survival of the flattest is the role of natural selection in the two phenomena. The evolution of drift robustness does not require competition between drift-robust and drift-fragile genotypes [25]. Instead, small populations tend to evolve towards drift robustness because they can only maintain fitness on drift-robust genetic backgrounds. Survival of the flattest, however, occurs when a flat genotype outcompetes a fit genotype, and is driven by selection for mutational robustness. It has also been established in survival of the flattest that the critical high mutation rate above which this selective advantage occurs is independent of population size [9], while drift robustness (almost by definition) is strongly dependent on population size. Furthermore, our previous work demonstrated that drift-robust genotypes are not more mutationally robust than drift-fragile genotypes, if one measures mutational robustness as the mean relative fitness of all single mutants [25]. Single mutants of flat genotypes are more fit, on average, than those of fit genotypes, again illustrating a difference between drift robustness and survival of the flattest. Thus, while similarities exist between drift-robust genetic architecture and flat genetic architecture, it appears that they represent distinct evolutionary responses to different population-genetic environments.

It would be worthwhile to determine if the results observed to hold here can be replicated with RNA viruses, the biological model for the survival-of-the-flattest effect [8, 28, 36]. It would be illuminating to test if the likelihood of lethal mutations does increase in these evolved viruses, especially in light of studies arguing that RNA viruses have high mutational *fragility*, not high mutational robustness [16]. Such a study would lead to a better understanding of the mutational mechanisms that lead to robust organisms in nature.

5 Conclusions

In high-mutation-rate environments, populations will evolve alternative genetic architectures that increase their robustness to deleterious mutations. Classically, this is viewed as a survival-of-the-flattest effect [45], where populations adapt to fitness peaks with increased neutrality [39]. Here, we repeated

the original experiments to finely map the fitness landscapes of fit and flat genotypes. Flat, mutationally robust genotypes adapt to fitness peaks with an increased likelihood of both neutral genotypes and lethal genotypes, a decreased likelihood of deleterious mutations, and greater negative epistasis compared to fit genotypes. These results demonstrate the genetic complexity of mutationally robust genomes that minimize the impact of heritable deleterious genetic variation.

Acknowledgments

References

1. Adami, C., Ofria, C., & Collier, T. C. (2000). Evolution of biological complexity. *Proceedings of the National Academy of Sciences of the U.S.A.*, *97*, 4463–4468.

2. Agrawal, A. F., & Whitlock, M. C. (2012). Mutation load: The fitness of individuals in populations where deleterious alleles are abundant. *Annual Review of Ecology, Evolution, and Systematics*, *43*, 115–135.

3. Archetti, M. (2009). Survival of the steepest: Hypersensitivity to mutations as an adaptation to soft selection. *Journal of Evolutionary Biology*, *22*, 740–750.

4. Beardmore, R. E., Gudelj, I., Lipson, D. A., & Hurst, L. D. (2011). Metabolic trade-offs and the maintenance of the fittest and the flattest. *Nature*, *472*, 342–346.

5. Biebricher, C., & Eigen, M. (2006). What is a quasispecies? *Quasispecies: Concept and Implications for Virology*, 1–31.

6. Bull, J. J., Meyers, L. A., & Lachmann, M. (2005). Quasispecies made simple. *PLoS Computational Biology*, *1*, e61.

7. Butcher, D. (1995). Muller's ratchet, epistasis and mutation effects. *Genetics*, *141*, 431–437.

8. Codoñer, F. M., Darós, J.-A., Solé, R. V., & Elena, S. F. (2006). The fittest versus the flattest: Experimental confirmation of the quasispecies effect with subviral pathogens. *PLoS Pathogens*, *2*, e136.

9. Comas, I., Moya, A., & González-Candelas, F. (2005). Validating viral quasispecies with digital organisms: A re-examination of the critical mutation rate. *BMC Evolutionary Biology*, *5*, 5.

10. Covert, A. W., Lenski, R. E., Wilke, C. O., & Ofria, C. (2013). Experiments on the role of deleterious mutations as stepping stones in adaptive evolution. *Proceedings of the National Academy of Sciences of the U.S.A.*, *110*, E3171–E3178.

11. Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology*, *30*, 1–13.

12. de Visser, J. A. G. M., Hermisson, J., Wagner, G. P., Meyers, L. A., Bagheri-Chaichian, H., et al. (2003). Perspective: Evolution and detection of genetic robustness. *Evolution*, *57*, 1959–1972.

13. Edlund, J. A., & Adami, C. (2004). Evolution of robustness in digital organisms. *Artificial Life*, *10*, 167–179.

14. Eigen, M. (1971). Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, *58*, 465–523.

15. Eigen, M., & Schuster, P. (1977). A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften*, *64*, 541–565.

16. Elena, S. F., Carrasco, P., Daròs, J.-A., & Sanjuán, R. (2006). Mechanisms of genetic robustness in RNA viruses. *EMBO Reports*, *7*, 168–173.

17. Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, *8*, 610–618.

18. Forster, R., Adami, C., & Wilke, C. O. (2006). Selection for mutational robustness in finite populations. *Journal of Theoretical Biology*, *243*, 181–190.

19. Goldsby, H. J., Dornhaus, A., Kerr, B., & Ofria, C. (2012). Task-switching costs promote the evolution of division of labor and shifts in individuality. *Proceedings of the National Academy of Sciences of the U.S.A.*, *109*, 13686–13691.

20. Goldsby, H. J., Knoester, D. B., Ofria, C., & Kerr, B. (2014). The evolutionary origin of somatic cells under the dirty work hypothesis. *PLoS Biology*, *12*, e1001858.

21. Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*, 90–95.

22. Kondrashov, A. S. (1994). Muller's ratchet under epistatic selection. *Genetics*, *136*, 1469–1473.

23. Krakauer, D. C., & Plotkin, J. B. (2002). Redundancy, antiredundancy, and the robustness of genomes. *Proceedings of the National Academy of Sciences of the U.S.A.*, *99*, 1405–1409.

24. LaBar, T., & Adami, C. (2016). Different evolutionary paths to complexity for small and large populations of digital organisms. *PLoS Computational Biology*, *12*, e1005066.

25. LaBar, T., & Adami, C. (2017). Evolution of drift robustness in small populations. *Nature Communications*, *8*, 1012.

26. Lan, Y., Trout, A., Weinreich, D. M., & Wylie, C. S. (2017). Natural selection can favor the evolution of ratchet robustness over evolution of mutational robustness. bioRxiv, 121087.

27. Lauring, A. S., & Andino, R. (2010). Quasispecies theory and the behavior of RNA viruses. *PLoS Pathogens*, *6*, e1001005.

28. Lauring, A. S., Frydman, J., & Andino, R. (2013). The role of mutational robustness in RNA virus evolution. *Nature Reviews Microbiology*, *11*, 327–336.

29. Lenski, R. E., Ofria, C., Collier, T. C., & Adami, C. (1999). Genome complexity, robustness and genetic interactions in digital organisms. *Nature*, *400*, 661–664.

30. Lenski, R. E., Ofria, C., Pennock, R. T., & Adami, C. (2003). The evolutionary origin of complex features. *Nature*, *423*, 139–144.

31. Nowak, M. A. (1992). What is a quasispecies? *Trends in Ecology & Evolution*, *7*, 118–121.

32. O'Fallon, B. D., Adler, F. R., & Proulx, S. R. (2007). Quasi-species evolution in subdivided populations favours maximally deleterious mutations. *Proceedings of the Royal Society of London B: Biological Sciences*, *274*, 3159–3164.

33. Ofria, C., Bryson, D. M., & Wilke, C. O. (2009). Avida: A software platform for research in computational evolutionary biology. In A. A. Maciej Komosinski (Ed.), *Artificial life models in software* (pp. 3–35). London: Springer.

34. Pennock, R. T. (2007). Models, simulations, instantiations, and evidence: The case of digital evolution. *Journal of Experimental & Theoretical Artificial Intelligence*, *19*, 29–42.

35. R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

36. Sanjuán, R., Cuevas, J. M., Furió, V., Holmes, E. C., & Moya, A. (2007). Selection for robustness in mutagenized RNA viruses. *PLoS Genetics*, *3*, e93.

37. Schuster, P., & Swetina, J. (1988). Stationary mutant distributions and evolutionary optimization. *Bulletin of Mathematical Biology*, *50*, 635–660.

38. Strona, G., & Lafferty, K. D. (2016). Environmental change makes robust ecological networks fragile. *Nature Communications*, *7*, 12462.

39. Van Nimwegen, E., Crutchfield, J. P., & Huynen, M. (1999). Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences of the U.S.A.*, *96*, 9716–9720.

40. Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

41. Wilke, C. O. (2001). Adaptive evolution on neutral networks. *Bulletin of Mathematical Biology*, *63*, 715–730.

42. Wilke, C. O. (2005). Quasispecies theory in the context of population genetics. *BMC Evolutionary Biology*, *5*, 44.

43. Wilke, C. O., & Adami, C. (2001). Interaction between directional epistasis and average mutational effects. *Proceedings of the Royal Society of London B: Biological Sciences*, *268*, 1469–1474.

44. Wilke, C. O., & Adami, C. (2003). Evolution of mutational robustness. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *522*, 3–11.

45. Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., & Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, *412*, 331–333.

46. Wilkins, J. F., McHale, P. T., Gervin, J., & Lander, A. D. (2016). Survival of the curviest: Noise-driven selection for synergistic epistasis. *PLoS Genetics*, *12*, e1006003.

Appendix

List of Abbreviations
RNA: Ribonucleic acids
SIP: Single Instruction Processing

Availability of Data and Material
All Avida configuration files and data analysis scripts to recreate these experiments, analyze the data, and recreate the figures are available at https://github.com/joshf93/MappingThePeaks.

Competing Interests
The authors declare that they have no competing interests.

Authors' Contributions
J.F. performed the experiments and analyzed the data. J.F., T.L., and C.A. designed the study and wrote the manuscript.
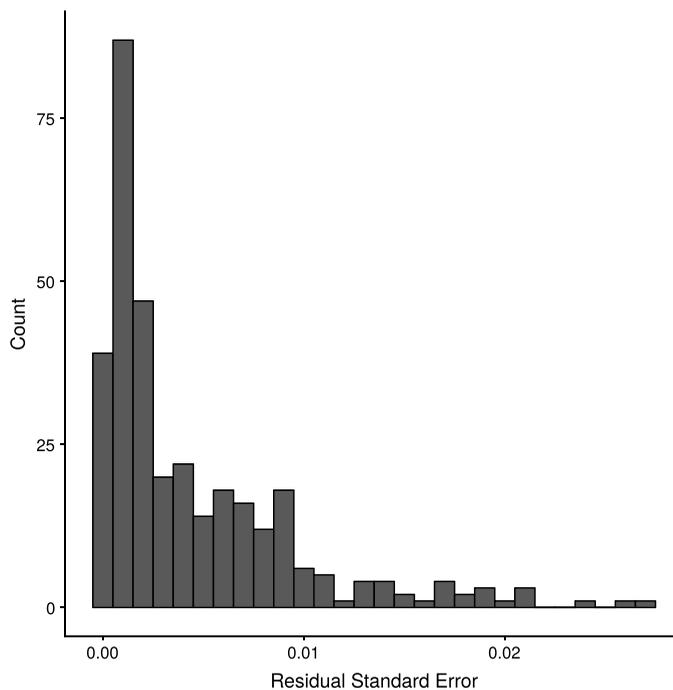
Supplementary Figures

Figure 6. Histogram of the residual standard error (RSE) of the fits of Equation 1 to all fit and all flat peaks.
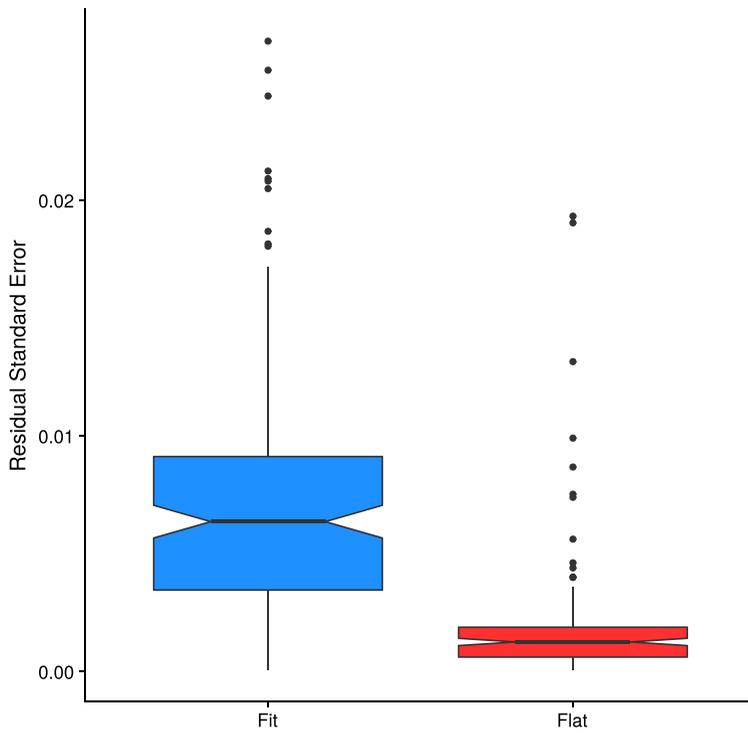
Figure 7. Boxplot of the residuals of the fit to Equation 1 for the fit peaks (blue) and flat peaks (red) separately. As in the boxplots of the main text, lines are the median values, boxes show the interquartile range, and whiskers are 1.5 times the interquartile range or the minimum or maximum value; points are outliers.