RESEARCH ARTICLE

Recombination and mutational robustness in neutral fitness landscapes

Alexander Klug [1], **Su-Chan Park** [2], **Joachim Krug** [1] *

1 Institute for Biological Physics, University of Cologne, Cologne, Germany, **2** Department of Physics, The Catholic University of Korea, Bucheon, Republic of Korea

* jkrug@uni-koeln.de

Abstract

Mutational robustness quantifies the effect of random mutations on fitness. When mutational robustness is high, most mutations do not change fitness or have only a minor effect on it. From the point of view of fitness landscapes, robust genotypes form neutral networks of almost equal fitness. Using deterministic population models it has been shown that selection favors genotypes inside such networks, which results in increased mutational robustness. Here we demonstrate that this effect is massively enhanced by recombination. Our results are based on a detailed analysis of mesa-shaped fitness landscapes, where we derive precise expressions for the dependence of the robustness on the landscape parameters for recombining and non-recombining populations. In addition, we carry out numerical simulations on different types of random holey landscapes as well as on an empirical fitness landscape. We show that the mutational robustness of a genotype generally correlates with its recombination weight, a new measure that quantifies the likelihood for the genotype to arise from recombination. We argue that the favorable effect of recombination on mutational robustness is a highly universal feature that may have played an important role in the emergence and maintenance of mechanisms of genetic exchange.

Author summary

Two long-standing and seemingly unrelated puzzles in evolutionary biology concern the ubiquity of sexual reproduction and the robustness of organisms against genetic perturbations. Using a theoretical approach based on the concept of a fitness landscape, in this article we argue that the two phenomena may in fact be closely related. In our setting the hereditary information of an organism is encoded in its genotype, which determines it to be either viable or non-viable, and robustness is defined as the fraction of mutations that maintain viability. Previous work has demonstrated that the purging of non-viable genotypes from the population by natural selection leads to a moderate increase in robustness. Here we show that genetic recombination acting in combination with selection massively enhances this effect, an observation that is largely independent of how genotypes are connected by mutations. This suggests that the increase of robustness may be a major driver underlying the evolution of sexual recombination and other forms of genetic exchange throughout the living world.

Introduction

The reshuffling of genetic material by recombination is a ubiquitous part of the evolutionary process across the entire range of organismal complexity. Starting with viruses as the simplest evolving entities, recombination occurs largely at random during the coinfection of a cell by more than one virus strain [1]. For bacteria the mechanisms involved in recombination are already more elaborate and present themselves in the form of transformation, transduction and conjugation [2, 3]. In eukaryotic organisms, sexual reproduction is a nearly universal feature, and recombination is often a necessary condition for the creation of offspring. Although its prevalence in nature is undeniable, the evolution and maintenance of sex is surprising since compared to an asexual population, only half of a sexual population is able to bear offspring and additionally a suitable partner needs to be found [4, 5]. Whereas the resulting two-fold cost of sex applies only to organisms with differentiated sexes [6], the fact that genetic reshuffling may break up favorable genetic combinations or introduce harmful variants into the genome poses a problem also to recombining microbes that reproduce asexually [7, 8]. Since this dilemma was noticed early on in the development of evolutionary theory, many attempts have been undertaken to identify evolutionary benefits of sex and recombination based on general population genetic principles [9–19].

In this article we approach the evolutionary role of recombination from the perspective of fitness landscapes. The fitness landscape is a mapping from genotype to fitness, which encodes the epistatic interactions between mutations and provides a succinct representation of the possible evolutionary trajectories [20]. Previous computational studies addressing the effect of recombination on populations evolving in epistatic fitness landscapes have revealed a rather complex picture, where evolutionary adaptation can be impeded or facilitated depending on, e.g., the structure of the landscape, the rate of recombination or the time frame of observation [21–26].

Here we focus specifically on the possible benefit of recombination that derives from its ability to enhance the mutational robustness of the population. A living system is said to be robust if it is able to maintain its function in the presence of perturbations [27–31]. In the case of mutational robustness these perturbations are genetic, and the robustness of a genotype is quantified by the number of mutations that it can tolerate without an appreciable change in fitness. Robust genotypes that are connected by mutations therefore form plateaux in the fitness landscape that are commonly referred to as neutral networks [32–35]. Mutational robustness is known to be abundant at various levels of biological organization, but its origins are not well understood. In particular, it is not clear if mutational robustness should be viewed as an evolutionary adaptation, or rather reflects the intrinsic structural constraints of living systems.

Arguments in favor of an adaptive origin of robustness were presented by van Nimwegen *et al.* [32] and by Bornberg-Bauer and Chan [33], who showed that selection tends to concentrate populations in regions of a neutral network where robustness is higher than average. Whereas this result is widely appreciated, the role of recombination for the evolution of robustness has received much less attention. An early contribution that can be mentioned in this context is due to Boerlijst *et al.* [36], who discuss the error threshold in a viral quasi-species model with recombination and point out in a side note that "*in sequence space recombination is always inwards pointing.*" This observation was picked up by Wilke and Adami [37] in a review on the evolution of mutational robustness, where they conjecture that the enhancement of robustness by selection should be further amplified by recombination, because "*recombination alone always creates sequences that are within the boundaries of the current mutant cloud.*" At about the same time, de Visser *et al.* discussed a mechanism based on the spreading of robustness modifier alleles in recombining populations [27] (see also [38]).

In fact indications of a positive effect of recombination on robustness had been reported earlier in computational studies of the evolution of RNA secondary structure [39] and 2D lattice proteins [40] in the presence and absence of recombination. In these systems the native folding structure of a given sequence is determined by its global free energy minimum. Due to the restricted number of attainable folds, most structures are degenerate in the sense that many sequences fold into the same structure. These sequences form neutral networks in sequence space. Xia and Levitt [40] consider two scenarios, in which the evolution of the lattice proteins is dominated by mutation and by recombination, respectively. The results show that in the latter case the concentration of thermodynamically stable protein sequences is enhanced, which is qualitatively explained by the fact that recombination tends to focus the sequences near the center of their respective neutral network. Therefore most often a single mutation does not change the folding structure.

More recently, Azevedo *et al.* [41] used a model of gene regulatory networks to investigate the origin of negative epistasis, which is a requirement for the advantage of recombination according to the mutational deterministic hypothesis [13]. In this study a gene network is encoded by a matrix of interaction coefficients. It is defined to be viable if its dynamics converges to a stable expression pattern and non-viable otherwise. Thus the underlying fitness landscape is again neutral. Based on their simulation results the authors argue that recombination of interaction matrices reduces the recombinational load, which in turn leads to an increase of mutational robustness and induces negative epistasis as a byproduct. In effect, then, recombination selects for conditions that favor its own maintenance. Other studies along similar lines have been reviewed in [42]. Taken together they suggest that the positive effect of recombination on robustness may be largely independent of the precise structure of the space of genotypes or the genotype-phenotype map. Indeed, a related scenario has also been described in the context of computational evolution of linear genetic programs [43].

Finally, in a numerical study that is similar to ours in spirit, Szöllősi and Derényi considered the effect of recombination on the mutational robustness of populations evolving on different types of neutral fitness landscapes [44]. Using neutral networks that were either generated at random or based on RNA secondary structure, they found that recombination generally enhances mutational robustness by a significant amount. Moreover, they showed that this observation holds not only for infinite populations but also for finite populations, as long as these are sufficiently polymorphic.

The goal of this article is to explain these scattered observations in a systematic and quantitative way. For this purpose we begin by a detailed examination of the simplest conceivable setting consisting of a haploid two-locus model with three viable and one lethal genotype [35]. We derive explicit expressions for the robustness as a function of the rates of mutation and recombination that demonstrate the basic phenomenon and guide the exploration of more complex situations. The two-locus results are then generalized to mesa landscapes with L diallelic loci, where genotypes carrying up to k mutations are viable and of equal fitness [45–48]. Using a communal recombination scheme and previous results for multilocus mutation-selection models, we arrive at precise asymptotic results for the mutational robustness for large L and small mutation rates. Subsequently two types of random holey landscape models are considered, including a novel class of sea-cliff landscapes in which the fraction of viable genotypes depends on the distance to a reference sequence. For the isotropic percolation landscape analytic upper and lower bounds on the robustness are derived.

As a first step towards a unified explanation for the effect of recombination on mutational robustness we introduce the concept of the recombination weight, which is a measure for the likelihood of a genotype to arise from a recombination event. In analogy to the classic fitness landscape concept in the context of selection [20], the recombination weight allows one to

identify genotypes that are favored by recombination without referring to any specific evolutionary dynamics. We show that recombination weight correlates with mutational robustness for the landscape structures used in this work, thus providing a mechanistic basis for the enhancement of robustness by recombination. Finally, using an empirical fitness landscape as an example, we quantify the competition between selection and recombination as a function of recombination rate. Throughout we describe the evolutionary dynamics by a deterministic, discrete time model that will be introduced in the next section.

Models and methods

Genotype space

We consider a haploid genome with L loci and the corresponding genotype is represented by a sequence $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_L)$ of length L. The index i labels genetic loci and each locus carries an allele specified by σ_i. Here we rely on binary sequences, which means that there are only two different alleles $\sigma_i \in \{0, 1\}$. This can be either seen as a simplification in the sense that only two alleles are assumed to exist, or in the sense that the genome consisting of all zeros describes the wild type, and the 1's in the sequence display mutations for which no further distinctions are made.

The resulting genotype space is a hypercube of dimension L, where the 2^L genotypes represent vertices, and two genotypes that differ at a single locus and are mutually reachable by a point mutation are connected by an edge. A metric is introduced by the Hamming distance

$$d(\sigma, \kappa) = \sum_i (1 - \delta_{\sigma_i \kappa_i}),$$

(1)

which measures the number of point mutations that separate two genotypes σ and κ. Here and in the following the Kronecker symbol is defined as $\delta_{xy} = 1$ if $x = y$ and $\delta_{xy} = 0$ otherwise. The genotype $\overline{\sigma}$ at maximal distance $d(\sigma, \overline{\sigma}) = L$ from a given genotype σ is called its antipodal, and can be defined by $\overline{\sigma}_i = 1 - \sigma_i$. Finally, in order to generate a fitness landscape, a (Wrightian) fitness value w_σ is assigned to each genotype.

Dynamics

The forces that drive evolution are selection, mutation and recombination. To model the dynamics we use a deterministic, discrete-time model with non-overlapping generations, which can be viewed as an infinite population limit of the Wright-Fisher model. Demographic stochasticity or genetic drift is thus neglected. Numerical simulations of evolution on neutral networks have shown that the infinite population dynamics is already observable for moderate population sizes, which justifies this approximation [32, 44]. We will return to this point in the Discussion.

Once the frequency $f_\sigma(t)$ of a genotype σ at generation t is given, the frequency at the next generation is determined in three steps representing selection, mutation, and recombination. After the selection step, the frequency $q_\sigma(t)$ is given as

$$q_\sigma(t) = \frac{w_\sigma}{\overline{w}(t)} f_\sigma(t),$$

(2)

where $\overline{w} \equiv \sum_\sigma w_\sigma f_\sigma(t)$ is the mean population fitness at generation t. After the mutation step, the frequency $p_\sigma(t)$ is given as

$$p_\sigma(t) = \sum_\kappa U_{\sigma\kappa} q_\kappa(t),$$

(3)

where $U_{\sigma\kappa}$ is the probability that an individual with genotype κ mutates to genotype σ in one generation. Here, we assume that alleles at each locus mutate independently, and the mutation probability μ is the same in both directions ($0 \rightarrow 1$ and $1 \rightarrow 0$) and across loci. This leads to the symmetric mutation matrix

$$U_{\sigma\kappa} = (1 - \mu)^{L-d(\sigma,\kappa)}\mu^{d(\sigma,\kappa)}. \tag{4}$$

In order to incorporate recombination we have to consider the probability that two parents with respective genotypes κ and τ beget a progeny with genotype σ by recombination. This is represented by the following equation:

$$f_\sigma(t + 1) = \sum_{\kappa\tau} R_{\sigma|\kappa\tau} p_\kappa(t) p_\tau(t). \tag{5}$$

Descriptively speaking, two genotypes (κ and τ) are taken to recombine with a probability that is equal to their frequency in the population (after selection and mutation). The probability for the offspring genotype σ is then given by $R_{\sigma|\kappa\tau}$. These probabilities depend of course on the parent genotypes κ and τ but also on the recombination scheme. Here we consider a uniform and a one-point crossover scheme; see Fig 1 for a graphical representation. These two represent extremes in a spectrum of possible recombination schemes. Nevertheless we will show that both lead to qualitatively similar results in the regimes of interest. In the case of uniform crossover the recombination probabilities are given by

$$R_{\sigma|\kappa\tau} = \frac{r}{2^L}\left(\prod_i^L (\delta_{\sigma_i\kappa_i} + \delta_{\sigma_i\tau_i})\right) + \frac{1-r}{2}(\delta_{\sigma\kappa} + \delta_{\sigma\tau}) \tag{6}$$

and in the case of one point crossover the probabilities can be written as

$$\begin{aligned}
R_{\sigma|\kappa\tau} = \quad & \frac{r}{2(L-1)}\sum_{n=1}^{L-1}\left[\left(\prod_{m=1}^{n}\delta_{\sigma_m\kappa_m}\right)\left(\prod_{m=n+1}^{L}\delta_{\sigma_m\tau_m}\right)\right. \\
& \left. + \left(\prod_{m=1}^{n}\delta_{\sigma_m\tau_m}\right)\left(\prod_{m=n+1}^{L}\delta_{\sigma_m\kappa_m}\right)\right] + \frac{1-r}{2}(\delta_{\sigma\kappa} + \delta_{\sigma\tau}).
\end{aligned} \tag{7}$$

In both equations a variable $r \in [0, 1]$ appears which describes the recombination rate. For $r = 0$ no recombination occurs and $f_\sigma(t + 1)$ is the same as $p_\sigma(t)$. For $r = 1$ recombination is a necessary condition for the creation of offspring (obligate recombination). But also intermediate values of r can be chosen as they occur in nature, e.g., for bacteria and viruses.

In the following we are mostly interested in the equilibrium frequency distribution f_σ^* of a population, which is determined by the stationarity condition

$$f_\sigma(t + 1) = f_\sigma(t) = f_\sigma^* \tag{8}$$

for all genotypes σ.

Mutational robustness

From the point of view of fitness landscapes the occurrence of mutational robustness implies that fitness values of neighboring genotypes are degenerate, giving rise to neutral networks in genotype space [29, 32–35]. In order to model this situation we use two-level landscapes that only differentiate between genotypes that are viable ($w_\sigma = 1$) or lethal ($w_\sigma = 0$). Any selective advantage between viable genotypes is assumed to be negligible. The mutational robustness of
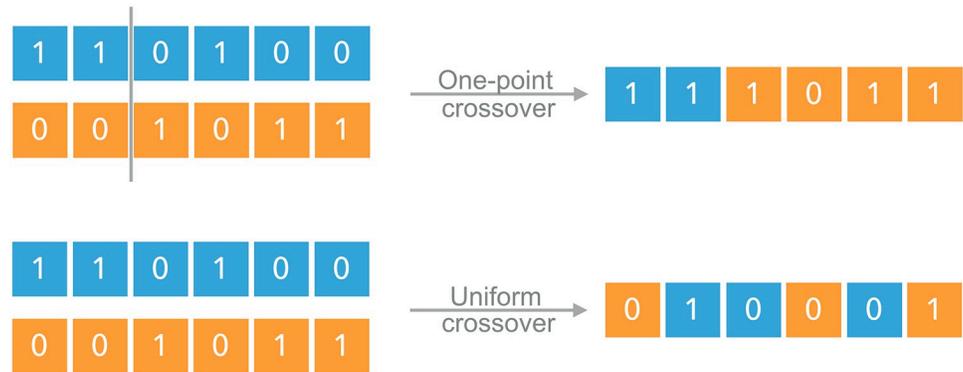
Fig 1. Recombination schemes. In the one-point crossover scheme, the parent genotypes are cut once between two randomly chosen loci and recombined to form the offspring. In the uniform crossover scheme, at each locus of the offspring, an allele present in one of the parents is chosen at random.

a population can then be measured by the average fraction of viable point mutations in an individual, which depends on the population distribution in genotype space [32–34]. It increases if the population mainly adapts to genotypes for which most point mutations are viable. Therefore we define mutational robustness m as the average fraction of viable point mutations of a population,

$$m \equiv \sum_{\sigma \in V} m_\sigma f_\sigma^* \quad \text{with} \quad m_\sigma \equiv \frac{n_\sigma}{L}. \tag{9}$$

Here the sum is over the set V of all viable genotypes and n_σ is the number of viable point mutations of genotype σ. We will refer to m_σ as the mutational robustness of the genotype. The expression is normalized by the total number of loci L, since in an optimal setting the entire population has L viable genotypic neighbors and $m_\sigma = 1$ for all $\sigma \in V$. The value of m is thus constrained to be in the range [0, 1]. We weight the genotypes by their stationary frequencies f_σ^*, since we want to determine the mutational robustness of populations that are in equilibrium with their environment.

Recombination weight

In order to elucidate the interplay of recombination and mutational robustness it will prove helpful to introduce a representation of how recombination can transfer genotypes into each other. The number of distinct genotypes that two recombining genotypes are able to create depends on their Hamming distance. In particular, the recombination of two identical genotypes does not create any novelty, whereas a genotype and its antipodal are able to generate all possible genotypes through uniform crossover.

Here we introduce a measure which expresses how many pairs of viable genotypes are able to recombine to a specific genotype. It is complementary to the mutational robustness, in the sense that instead of counting the viable mutation neighbors of a genotype, the size of its recombinational neighborhood of viable recombination pairs is determined. The recombinational neighborhood depends on the recombination scheme and the distribution of viable genotypes in the genotype space. For a given recombination scheme the probability for a genotype σ to be the outcome of recombination of two genotypes κ, τ is given by the recombination

tensor $R_{\sigma|\kappa\tau}$. The *recombination weight* λ_σ is therefore obtained by summing the recombination tensor over all ordered pairs of viable genotypes,

$$\lambda_\sigma = \frac{1}{2^L} \sum_{\kappa \in V, \tau \in V} R_{\sigma|\kappa\tau}. \tag{10}$$

It can be seen from (5) that $\lambda_\sigma = 1$ when all genotypes are viable, and hence the normalization by 2^L ensures that the recombination weight lies in the range [0, 1]. Under this normalization, the recombination weights sum to $\sum_\sigma \lambda_\sigma = |V|^2/2^L$, where $|V|$ stands for the number of viable genotypes. In the following the genotype maximizing λ_σ will be referred to as the *recombination center* of the landscape.

Since neutral landscapes only differentiate between viable (unit fitness) and lethal (zero fitness) genotypes, the recombination weight (10) can alternatively be written as a sum over all ordered pairs of genotypes whereby the recombination tensor is multiplied by the pair's respective fitness,

$$\lambda_\sigma = \frac{1}{2^L} \sum_{\kappa,\tau} R_{\sigma|\kappa\tau} w_\kappa w_\tau. \tag{11}$$

In this way the concept naturally generalizes to arbitrary fitness landscapes. In the absence of recombination ($r = 0$) the recombination weight (11) of a genotype is simply proportional to its fitness, $\lambda_\sigma = \tilde{w} w_\sigma$, where $\tilde{w} = 2^{-L} \sum_\sigma w_\sigma$ is the unweighted average fitness. Within our recombination schemes, the recombination tensor depends linearly on r and, by definition, so does the recombination weight. Accordingly, for general r the recombination weight interpolates linearly between the limiting values at $r = 0$ and $r = 1$. Since λ_σ for $r = 0$ is known, the remaining task will be to find λ_σ for $r = 1$.

Results

In the following sections we investigate how mutational robustness depends on the mutation and recombination rates. In order to test the generality of our results, we use, besides contrasting recombination schemes, also different neutral landscape models such as the mesa [45–48] and the percolation models [35, 49]. Additionally we introduce a more general landscape named sea-cliff model, which combines elements of both the landscape models and contains them as limiting cases. In the end, we discuss mutational robustness and its relation with recombination weight for an empirical landscape.

Two-locus models are commonly used in population genetics to gain a foothold in understanding evolutionary scenarios involving multiple recombining loci [35, 38, 50–57]. Following this tradition, we first discuss a two-locus model and then extend our results to multi-locus models.

Two-locus model

The simplest fitness landscape to study the mutational robustness of a population would be the haploid two-locus model in which all but one genotype are viable [35]; see Fig 2 for a graphical representation of the model. In this setting the population gains mutational robustness if the frequency of the genotype (0,0) for which both point mutations are viable increases relative to the genotypes (0,1) and (1,0). This model has been analyzed previously using a unidirectional mutation scheme where reversions $1 \rightarrow 0$ are suppressed [58, 59]. As a consequence, selection cannot contribute to mutational robustness because the genotype (0,0) goes extinct in the
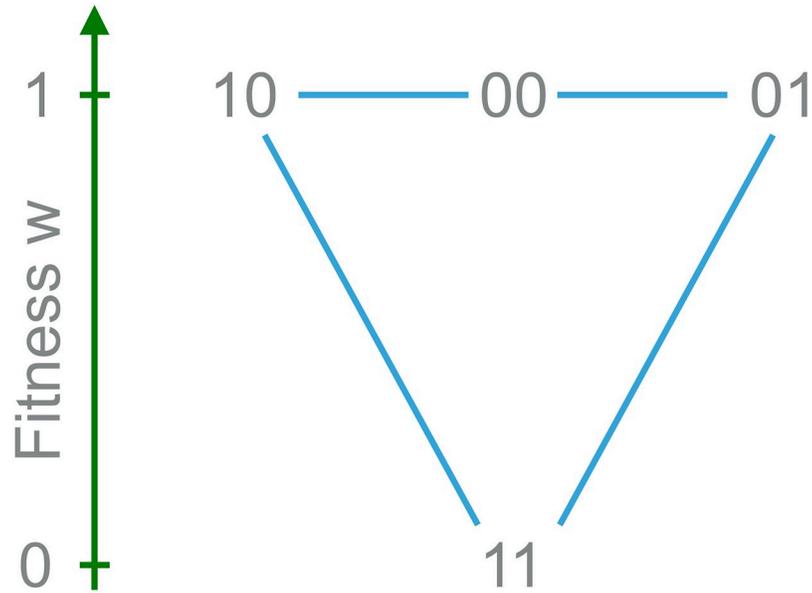
Fig 2. Two-locus model. Genotype (1,1) is lethal while the other three genotypes are viable with the same fitness. Here, genotype (0,0) is most robust since both its single mutants are viable.

absence of recombination. Here we consider the case of bidirectional, symmetric mutations in which both selection and recombination contribute to robustness. A comparison of the two mutation schemes is provided in S1 Appendix.

We proceed to solve the equilibrium condition Eq (8). Since the equilibrium genotype frequencies f_{01}^* and f_{10}^* are the same due to the symmetry of the landscape and the mutation scheme, the recombination step at stationarity reads

$$
\begin{aligned}
f_{00}^* &= p_{00} - \rho(p_{00}p_{11} - p_{10}p_{01}) &\Leftrightarrow\quad f_0 &= p_0 - \rho D, \\
f_{10/01}^* &= p_{10/01} + \rho(p_{00}p_{11} - p_{10}p_{01}) &\Leftrightarrow\quad f_1 &= p_1 + 2\rho D, \\
f_{11}^* &= p_{11} - \rho(p_{00}p_{11} - p_{10}p_{01}) &\Leftrightarrow\quad f_2 &= p_2 - \rho D,
\end{aligned}
\tag{12}
$$

where p_σ is the (equilibrium) frequency of genotype σ after the mutation step, f_i and p_i are the corresponding lumped frequencies [60] of all genotypes with i 1's, and $D \equiv p_{00}p_{11} - p_{10}p_{01} = p_0 p_2 - p_1^2/4$ is the linkage disequilibrium after the mutation step. Notice that the one-point and uniform crossover schemes give the same equation form except that the parameter ρ is given by $\rho = r$ in the case of one-point crossover and $\rho = r/2$ for uniform crossover. However, we would like to emphasize that this is a mere coincidence of the two-locus model which disappears as soon as L is larger than 2.

The lumped frequencies q_i of all genotypes with i 1's after the selection step are given by

$$
q_0 = \frac{f_0}{1 - f_2}, \quad q_1 = \frac{f_1}{1 - f_2}, \quad q_2 = 0.
\tag{13}
$$

Applying the mutation step we obtain

$$
\begin{aligned}
p_0 &= q_0(1-\mu)^2 + \mu(1-\mu)q_1 = \mu(1-\mu) + (1-\mu)(1-2\mu)q_0, \\
p_1 &= q_1[(1-\mu)^2 + \mu^2] + 2\mu(1-\mu)q_0 = 1 - 2\mu + 2\mu^2 - (1-2\mu)^2 q_0, \\
p_2 &= \mu(1-\mu)q_1 + \mu^2 q_0 = \mu(1-\mu) - \mu(1-2\mu)q_0, \\
D &= p_0 p_2 - p_1^2/4 = -\frac{1}{4}(1-2\mu)^2(1-q_0)^2,
\end{aligned}
\tag{14}
$$

where we have used the normalization $q_0 + q_1 = 1$ to express the right hand sides in terms of q_0. Putting everything together, the problem is reduced to solving the following third order polynomial equation for q_0,

$$
\begin{aligned}
0 &= q_0(1 - f_2) - f_0 = q_0(1 - p_2 + \rho D) - p_0 + \rho D \\
&= \frac{\rho}{4}(1-2\mu)^2(1-q_0)^2(1+q_0) + \mu\left[1 - 2q_0 - q_0^2 - \mu(1-2q_0)(1+q_0)\right],
\end{aligned}
\tag{15}
$$

from which we can in principle find exact analytic expressions for f_σ^*. However, it is difficult to extract useful information from the exact solution. In the following we will therefore provide approximate solutions.

If we neglect recombination ($\rho = 0$), we obtain the following equilibrium genotype frequency distribution:

$$
\begin{aligned}
f_{00}^*(\rho = 0) &= \frac{1-\mu}{2}\sqrt{8 - 16\mu + 9\mu^2} - \frac{1}{2}(2 - 5\mu + 3\mu^2) \\
&\approx (\sqrt{2} - 1) + \left(\frac{5}{2} - 2\sqrt{2}\right)\mu + O(\mu^2), \\
f_{01/10}^*(\rho = 0) &= \frac{1}{4}(4 - 9\mu + 6\mu^2) - \frac{1-2\mu}{4}\sqrt{8 - 16\mu + 9\mu^2} \\
&\approx \left(1 - \frac{1}{\sqrt{2}}\right) + \left(\frac{3}{\sqrt{2}} - \frac{9}{4}\right)\mu + O(\mu^2), \\
f_{11}^*(\rho = 0) &= \frac{\mu}{2}\left(4 - 3\mu - \sqrt{8 - 16\mu + 9\mu^2}\right) \approx \left(2 - \sqrt{2}\right)\mu + O(\mu^2).
\end{aligned}
\tag{16}
$$

When $\rho = 1$, which corresponds to the one-point crossover scheme with $r = 1$, linkage equilibrium ($f_{00} f_{11} = f_{10} f_{01}$) is restored after one generation [55]. Accordingly, we can treat each locus independently and get rather simple expressions for f_σ^* as

$$
\begin{aligned}
f_{00}^*(\rho = 1) &= \frac{1}{4}(2 + \mu - \sqrt{\mu^2 + 4\mu})^2 \approx 1 - 2\sqrt{\mu} + 2\mu + O(\mu^{3/2}), \\
f_{01/10}^*(\rho = 1) &= \frac{1}{4}\left(2 + \mu - \sqrt{\mu^2 + 4\mu}\right)\left(\sqrt{\mu^2 + 4\mu} - \mu\right) \approx \sqrt{\mu} - \frac{3\mu}{2} + O(\mu^{3/2}), \\
f_{11}^*(\rho = 1) &= \frac{1}{4}(\sqrt{\mu^2 + 4\mu} - \mu)^2 \approx \mu - O(\mu^{3/2}).
\end{aligned}
\tag{17}
$$

We depict the equilibrium solutions for the above two cases in Fig 3.

Now, the mutational robustness

$$
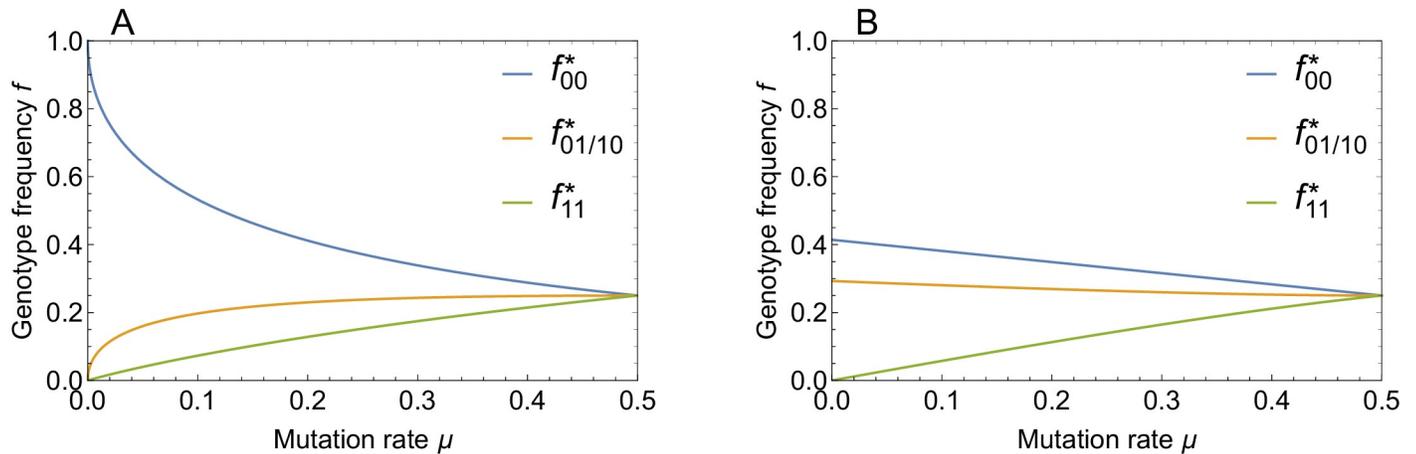m = \frac{1}{2}(2f_{00}^* + f_{10}^* + f_{01}^*) = f_{00}^* + f_{10}^* = f_0 + \frac{1}{2}f_1
\tag{18}
$$

Fig 3. Equilibrium genotype frequencies in the two locus model. Genotype frequencies in the stationary state are shown as a function of mutation rate for (A) strong recombination ($\rho = 1$) and (B) no recombination ($\rho = 0$).

for the above two cases is obtained as

$$m(\mu, \rho = 0) = \frac{1}{4}\left(\mu + \sqrt{8 - 16\mu + 9\mu^2}\right) \approx \frac{1}{\sqrt{2}} - \left(\frac{1}{\sqrt{2}} - \frac{1}{4}\right)\mu + O(\mu^2), \tag{19}$$

$$m(\mu, \rho = 1) = \frac{1}{2}\left(2 + \mu - \sqrt{\mu^2 + 4\mu}\right) \approx 1 - \sqrt{\mu} + \frac{\mu}{2} + O(\mu^{3/2}), \tag{20}$$

which is depicted in Fig 4. These results encapsulate in a simple form the main topic of this paper. Selection alone ($\rho = 0$) leads to a moderate increase of robustness from the baseline value $m = \frac{1}{2}$ corresponding to a random distribution over genotypes, which is attained at
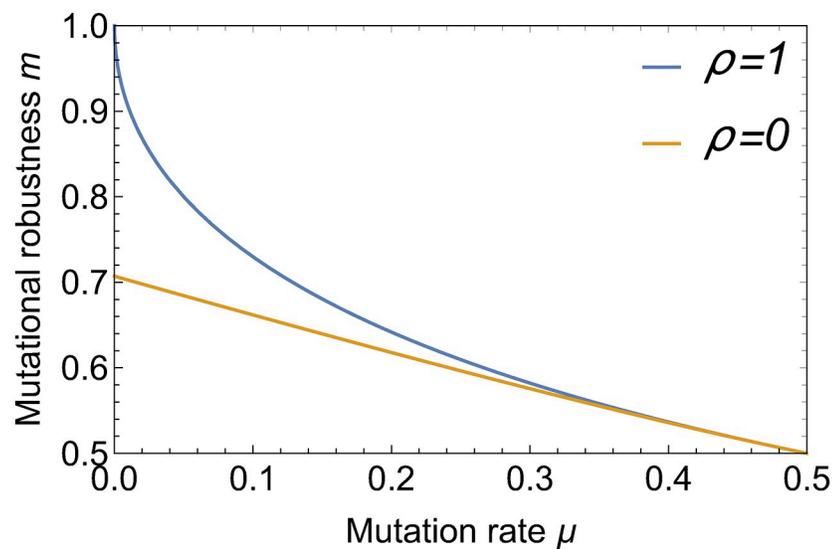

Fig 4. Mutational robustness as a function of mutation rate. The figure shows the robustness in the two-locus model at $\rho = 0$ and $\rho = 1$. Recombination leads to a massive enhancement of robustness for small mutation rates.

$\mu = \frac{1}{2}$, to $m = \frac{1}{\sqrt{2}}$ for $\mu \to 0$. In contrast, for recombining populations ($\rho = 1$) robustness is massively enhanced at small mutation rates due to the strong frequency increase of the most robust genotype (0,0) and reaches the maximal value $m = 1$ at $\mu = 0$. The underlying mechanism is analogous to Kondrashov's deterministic mutation hypothesis, which posits that recombination makes selection against deleterious mutations more effective when these interact synergistically [13]. In the present case recombination increases the frequency of the double mutant genotype (1, 1), which is subsequently purged by selection, and thereby effectively drives the frequency of the allele 1 at both loci to zero. The enhancement of the frequency of the genotype (0,0) by recombination is also reflected in the recombination weights, which take on the values

$$\lambda_{00} = \frac{3}{4} + \frac{\rho}{4}, \quad \lambda_{01} = \lambda_{10} = \frac{3}{4} - \frac{\rho}{4}, \quad \lambda_{11} = \frac{\rho}{4}. \tag{21}$$

Thus the genotype (0,0) is the recombination center of the two-locus landscape.

Next we investigate how mutational robustness varies with μ for intermediate recombination rates, assuming that μ is small. As can be seen from Eq (15), the asymptotic behavior of the solution for small ρ and μ depends on which of the two parameters is smaller. We first consider the case $\rho \ll \mu \ll 1$. Defining $l = \rho/(4\mu) \ll 1$, Eq (15) is approximated by

$$0 = l(1 - q_0)^2(1 + q_0) + 1 - 2q_0 - q_0^2 - \mu(1 - 2q_0)(1 + q_0), \tag{22}$$

where we kept terms up to $O(\mu)$, since we have not determined whether l is smaller than μ or not. Since $q_0 = \sqrt{2} - 1$ is the solution of Eq (22) for $l = \mu = 0$, we set $q_0 = \sqrt{2} - 1 + al + b\mu$ and solve the equation to leading order, which gives

$$q_0 \approx \sqrt{2} - 1 + \left(3 - 2\sqrt{2}\right)l - \left(\frac{3}{2} - \sqrt{2}\right)\mu. \tag{23}$$

The mutational robustness then follows as

$$m = f_0 + \frac{f_1}{2} = \frac{1}{2} + \frac{p_0 - p_2}{2} = \frac{1}{2} + (1 - 2\mu)\frac{q_0}{2} \approx \frac{1}{\sqrt{2}} + \frac{3 - 2\sqrt{2}}{2}l - \left(\frac{1}{\sqrt{2}} - \frac{1}{4}\right)\mu, \tag{24}$$

which is consistent with our previous result for $\rho = 0$; see Eq (19). We note that in this regime it is sufficient for the recombination rate to be of order $O(\mu^2)$ to compensate the negative effect of mutations on mutational robustness, as the two effects cancel when $\rho = \rho_c$ with

$$\rho_c = 2(5 + 4\sqrt{2})\mu^2 \approx 21.3 \times \mu^2. \tag{25}$$

In the regime $\rho \gg \mu$, Eq (15) is approximated as

$$(1 - 4\mu)(1 - q_0)^2(1 + q_0) + s(1 - 2q_0 - q_0^2) = 0, \tag{26}$$

with $s = 4\mu/\rho$. Again we have kept terms up to $O(\mu)$ because μ and s are of the same order if $\rho = O(1)$. Since the solution of Eq (26) for $\mu = s = 0$ is $q_0 = 1$, we set $q_0 = 1 - \alpha$ with $\alpha \ll 1$. Inserting this into Eq (26), we get $\alpha \approx \sqrt{s}$. Since $\alpha \gg \mu$, $q_0 = 1 - \sqrt{s}$ is the approximate solution to leading order. Hence

$$m = \frac{1}{2} + (1 - 2\mu)\frac{q_0}{2} \approx 1 - \sqrt{\frac{s}{4}} = 1 - \sqrt{\frac{\mu}{\rho}}, \tag{27}$$

which is again consistent with our previous result for $\rho = 1$ in Eq (20). The square root
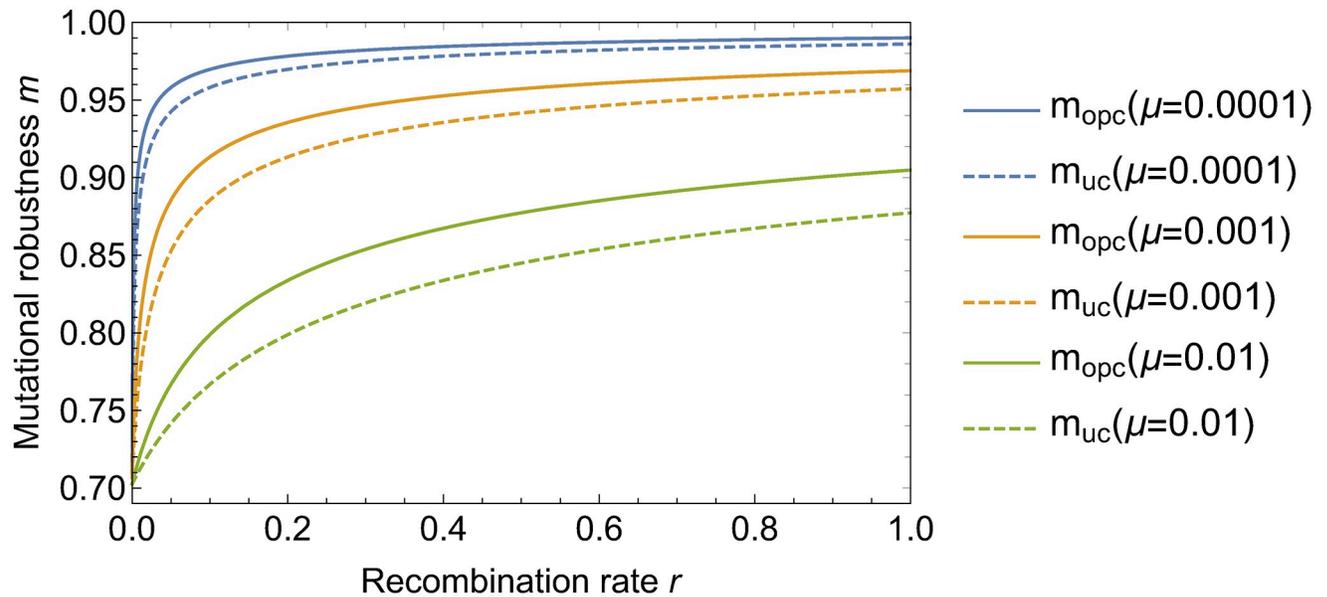
Fig 5. Mutational robustness as a function of recombination rate. The figure shows the mutational robustness for one-point crossover (m_{opc}) and uniform crossover (m_{uc}) and three different values of the mutation rate μ. When mutations are rare, a small amount of recombination is sufficient to significantly increase mutational robustness.

dependence on μ/ρ derives from the corresponding behavior of the genotype frequency f_{00}^* and has been noticed previously in the model with unidirectional mutations [58, 59].

For arbitrary ρ and μ, we have to use the full Eq (15). Fig 5 illustrates the behaviour of mutational robustness as a function of the recombination rate for different mutation rates and both recombination schemes. For small μ, a low rate of recombination suffices to bring the robustness close to its maximal value $m = 1$. More precisely, according to Eq (27), a robustness $m > 1 - \epsilon$ is reached for recombination rates $\rho > \mu/\epsilon^2$.

To summarize, we have seen that analytic results for the two-locus model are easily attainable. For multi-locus models it is much more challenging to derive analytical results, particularly in the presence of recombination. By way of contrast the dynamics induced only by mutation and selection are easier to understand: While mutations increase the genotype diversity in the population, fitter ones grow in frequency through selection, which reduces diversity. Although one might expect that recombination would increase diversity, a number of studies have shown that recombination is more likely to impede the divergence of populations. Recombining populations tend to cluster on single genotypes or in a limited region of a genotype space and furthermore the waiting times for peak shifts in multipeaked fitness landscapes diverge at a critical recombination rate [22, 26, 54–56, 61]. The results for the two-locus model presented above are consistent with this behaviour, as the genotype heterogeneity of the population decreases with increasing recombination rate (S1 Fig).

In the following we will investigate how the focusing effect of recombination enhances the mutational robustness of the population in three different multi-locus models.

Mesa landscape

In the mesa landscape it is assumed that up to a certain number k of mutations all genotypes are functional and have unit fitness, whereas genotypes with more than k mutations are lethal
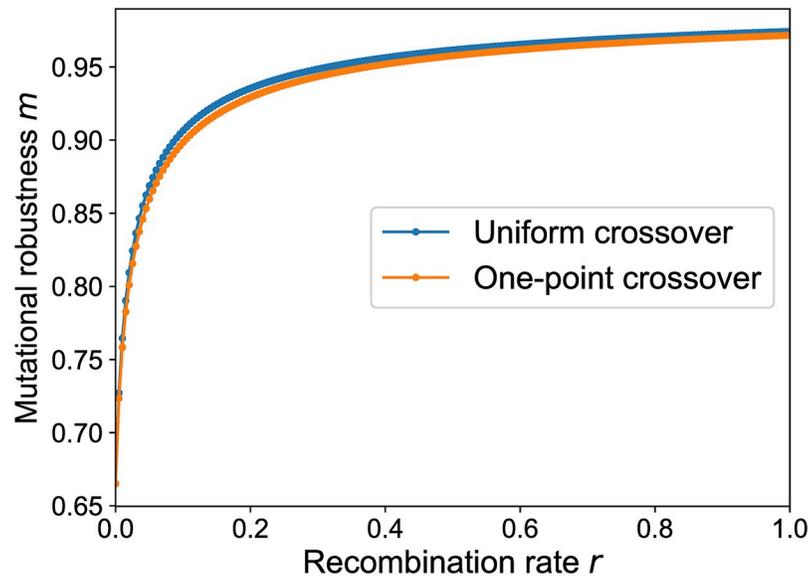
Fig 6. Mutational robustness in a mesa landscape as a function of recombination rate. Data points are obtained by numerically iterating the selection-mutation-recombination dynamics until the equilibrium state is reached. The parameters of the mesa landscape are $L = 6$, $k = 2$ and the mutation rate is $\mu = 0.001$.

and have fitness zero [48]. Hence the fitness landscape is defined as

$$
w_\sigma = \begin{cases} 1, & \text{if } d_\sigma \leq k, \\ 0, & \text{otherwise,} \end{cases}
\tag{28}
$$

where d_σ is the Hamming distance to the wild-type sequence $(0, 0, \ldots, 0)$ or, equivalently, the number of loci with allele 1. We will refer to k as the mesa width or as the critical Hamming distance.

Such a scenario can for example be observed in the evolution of regulatory motifs, where the fitness depends on the binding affinity of the regulatory proteins and d_σ corresponds to the number of mismatches compared to the original binding motif [45, 47]. The two-locus model discussed in the preceding section corresponds to the mesa landscape with critical Hamming distance $k = 1$ and sequence length $L = 2$. Here we ask to what extent the behavior observed for the two-locus model generalizes to longer sequences and variable k. Numerical simulations suggest that the strong increase of mutational robustness with recombination rate indeed persists in the general setting, and the particular recombination scheme seems to have only a minor influence; see Fig 6.

Whereas an analytical treatment for general L, k and intermediate recombination rates appears to be out of reach, accurate approximations are available in the limiting case of strong recombination or of no recombination, assuming mutation rate is small. The full derivations for both cases can be found in S1 Appendix. In the following we summarize the main results.

Strong recombination. In the limit of strong recombination we demand linkage equilibrium after each recombination step. This is satisfied if we use the so-called communal recombination scheme [62]. In this scheme an individual is not the offspring of a pair of parents. Rather, its genotype is aggregated by choosing the allele at each locus from a randomly selected parent. Hence the probability of occurrence of an allele at each locus in the offspring genotype after recombination is given by the corresponding allele frequency of the whole population,

which is precisely the definition of linkage equilibrium. In order to obtain an approximation for the mutational robustness we further assume that the mutation rate μ is small, which in turn implies a low frequency of mutant alleles. Following the derivation in S1 Appendix this leads us to the expression

$$m_{cr} \approx 1 - \binom{L-1}{k}^{1/(k+1)} \mu^{k/(k+1)} + \frac{k}{k+1}\mu, \tag{29}$$

which can be approximated as

$$m_{cr} \approx 1 - U^{k/(k+1)}(k!)^{-1/(k+1)} + \frac{k}{k+1}\mu \tag{30}$$

for $L \gg k$, where $U = L\mu$ is the genome-wide mutation rate and the subscript signifies the communal recombination scheme. Using Eq (29) and setting $L = 2$ and $k = 1$ we retrieve the result (20) for the two-locus model. Furthermore comparing Eqs (29) and (30) to numerical simulations of communal recombination illustrates their validity for large L (S2 Fig). If we use uniform crossover and one-point crossover instead of communal recombination, the numerical simulations suggest that the leading behaviour of $1 - m$ is still a function of $U = L\mu$ with the same exponent $k/(k + 1)$, which supports the universality of our findings with respect to the recombination scheme; see S3 Fig.

No recombination. In order to obtain analytical results in the absence of recombination we assume that the mutation rate is small enough that only a single point mutation occurs in one generation. This condition is fulfilled if $U = L\mu \ll 1$. Interestingly, we observe that in this regime the equilibrium frequencies after selection are independent of U. Therefore also the mutational robustness after selection, denoted by M_{nr}, is independent of U. The relation between mutational robustness after selection (M_{nr}) and after mutation (m_{nr}) is given by

$$m_{nr} = M_{nr}(1 - U) + M_{nr}^2 U, \tag{31}$$

which makes it suffice to find M_{nr}.

Assuming $k/L \ll 1$ it is possible to link the set of stationarity conditions to the Hermite polynomials $H_n(x)$. This yields an approximation for the mutational robustness after selection as

$$M_{nr} = \sqrt{\frac{y_k}{L}} + o(L^{-1/2}), \tag{32}$$

where $\sqrt{y_k/2}$ is the largest zero of $H_{k+1}(x)$. Correspondingly, the mutational robustness after mutation is

$$m_{nr} = \sqrt{\frac{y_k}{L}}(1 - U) + \frac{y_k}{L}U. \tag{33}$$

A comparison to the exact solutions for M_{nr}, which have been obtained up to $k = 4$, confirms this approximation. If we further assume that $1 \ll k \ll L$, we find $y_k \sim 4k$, which leads to

$$m_{nr} = 2\sqrt{\frac{k}{L}}(1 - U) + 4\frac{k}{L}U. \tag{34}$$

Results for the joint limit $k, L \to \infty$ at fixed ratio $x = k/L$ can be obtained from the analysis of Ref. [48], which yields

$$M_{\mathrm{nr}} = \begin{cases} 2\sqrt{x(1-x)}, & \text{if } x < 1/2, \\ 1, & \text{if } x \geq 1/2 \end{cases} \tag{35}$$

and therefore

$$m_{\mathrm{nr}} = \begin{cases} 2\sqrt{x(1-x)}(1-U) + 4x(1-x)U, & \text{if } x < 1/2, \\ 1, & \text{if } x \geq 1/2. \end{cases} \tag{36}$$

The leading behaviour for small x coincides with Eq (34). A comparison of the approximations to numerical solutions is given in S4 Fig.

Comparison of the two cases. It is instructive to compare the results obtained above to the mutational robustness m_0 of a uniform population distribution. For the latter we assume that all viable genotypes have the same frequency and all lethal genotypes have frequency zero. For the mesa model this yields

$$m_0(L,k) = \frac{1}{\sum_{i=0}^{k}\binom{L}{i}} \left[\binom{L}{k}\frac{k}{L} + \sum_{i=0}^{k-1}\binom{L}{i} \right] \approx \min\left[2k/L, 1\right], \tag{37}$$

where the last approximation is valid for $L \to \infty$. In S5 Fig the behavior of m_0, m_{nr} and m_{cr} is depicted as a function of various model parameters. Similar to the results obtained for the two-locus model, we see that selection gives rise to a moderate increase of robustness (from $2k/L$ to $2\sqrt{k/L}$ for $1 \ll k \ll L$), but recombination has a much stronger effect and leads to values close to the maximal robustness $m = 1$ for a broad range of conditions.

To elucidate the underlying mechanism, it is helpful to consider the shape of the equilibrium frequency distributions in genotype space (Fig 7). The combinatorial increase of the number of genotypes with increasing d_σ generates a strong entropic force that selection alone cannot efficiently counteract. As a consequence, the non-recombining population distribution is localized near the brink of the mesa at $d_\sigma = k$ [48]. In contrast, the contracting property of recombination [44] allows it to localize the population in the interior of the fitness plateau where most genotypes are surrounded by viable mutants.

S6 Fig shows the corresponding recombination weight profile. Similar to the genotype frequencies in Fig 7(B) the recombination weight decays rapidly with increasing Hamming distance for $r > 0$, but the decay appears to be faster than exponential. Interestingly, at $d = k$ the recombination weight decreases with increasing r [see also Eq (21)]. The method used to compute λ_σ for large mesa landscapes is explained in S1 Appendix.

Percolation landscapes

In the percolation landscape genotypes are randomly chosen to be viable ($w_\sigma = 1$) with probability p and lethal ($w_\sigma = 0$) with probability $1 - p$. An interesting property of the percolation model is the emergence of two different landscape regimes [49, 63–65]. Above the percolation threshold p_c, viable genotypes connected by single mutational steps form a cluster that extends over the whole landscape, whereas below p_c only isolated small clusters appear. Since the percolation threshold depends inversely on the sequence length, $p_c \approx \frac{1}{L}$, for large L a small fraction of viable genotypes suffices to create large neutral networks. This allows a population to evolve to distant genotypes without going through lethal regions, and correspondingly the percolation model is often used to study speciation [35, 49]. A network representation of the
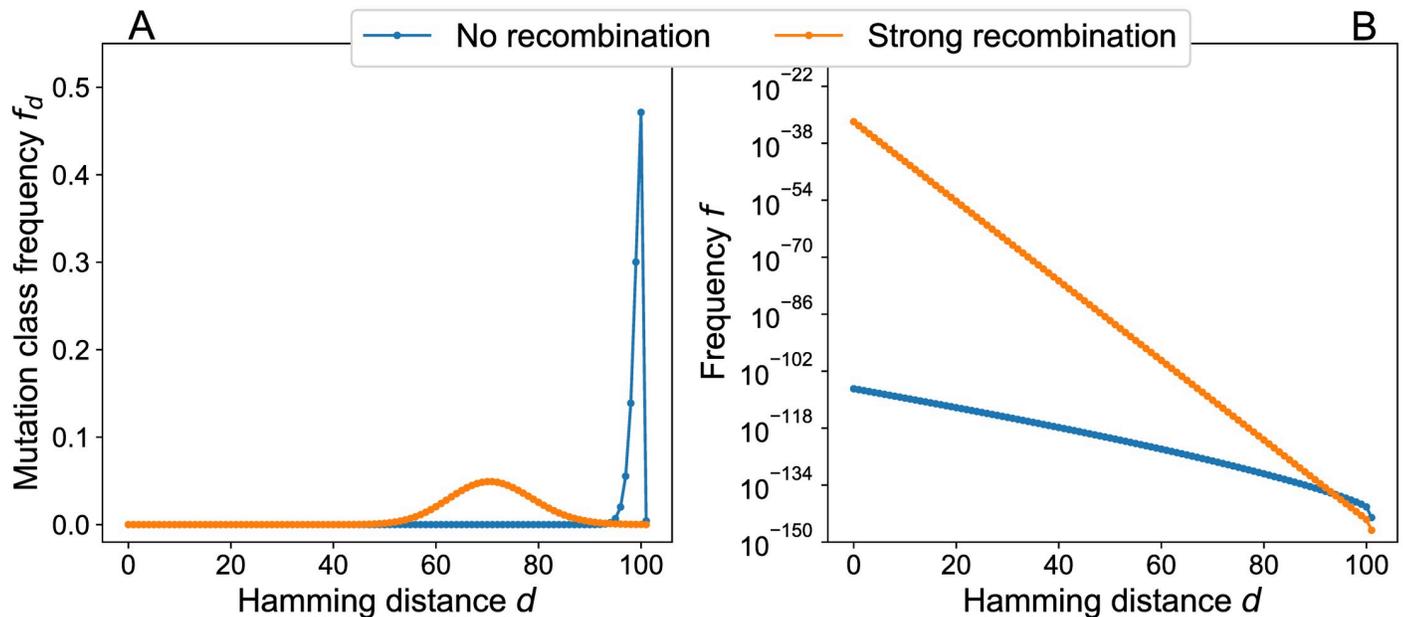
Fig 7. Equilibrium genotype distributions in a mesa landscape for strongly and non-recombining populations. Stationary states for populations with communal recombination and no recombination have been computed by assuming that only single point mutations occur with $U = 0.01$. Landscape parameters are $L = 1000$ and $k = 100$. The resulting mutational robustness is $m_{\mathrm{nr}} \approx 0.572$ for the non-recombining population and $m_{\mathrm{cr}} \approx 1.000$ for communal recombination. (A) Lumped mutation class frequencies on linear scales. In the absence of recombination the majority of the population is located at the critical Hamming distance $d = k$, whereas in the case of strong recombination the distribution is broader and shifted away from the brink of the mesa. (B) Genotype frequencies on semi-logarithmic scales. In both cases the genotype frequencies decrease exponentially with the Hamming distance to the wild type, but the distribution has much more weight at small distances in the case of recombination.

percolation model is shown in Fig 8. The algorithm used to generate this visual representation is explained in S1 Appendix.

Fig 9 shows three exemplary stationary genotype frequency distributions on the landscape depicted in Fig 8. In the absence of recombination the equilibrium frequency distribution is unique, but in the presence of recombination the non-linearity of the dynamics implies that multiple stationary states may exist [54, 55, 61]. Fig 9 displays two stationary distributions for $r = 1$ which are accessed from different initial conditions. It is visually apparent that the recombining populations are concentrated on a small number of highly connected genotypes, leading to a significant increase of mutational robustness.

To quantify this effect, the average mutational robustness $\overline{m}$ is calculated as a function of the recombination rate according to the following numerical protocol:

- A percolation landscape for given L and p is generated and the initial population is distributed uniformly among all genotypes.

- The population is evolved in the absence of recombination ($r = 0$) until the unique equilibrium frequency distribution is reached, for which the mutational robustness m is calculated.

- Next the recombination rate is increased by predefined increments. After increasing r, the population is again evolved using the stationary state obtained before the increment of r as the initial condition, until it reaches a new stationary state for which the mutational robustness is measured.

- When the recombination rate has reached $r = 1$, a new percolation landscape is generated and the process starts all over again. This is done for an adjustable number of runs over which the average is taken.
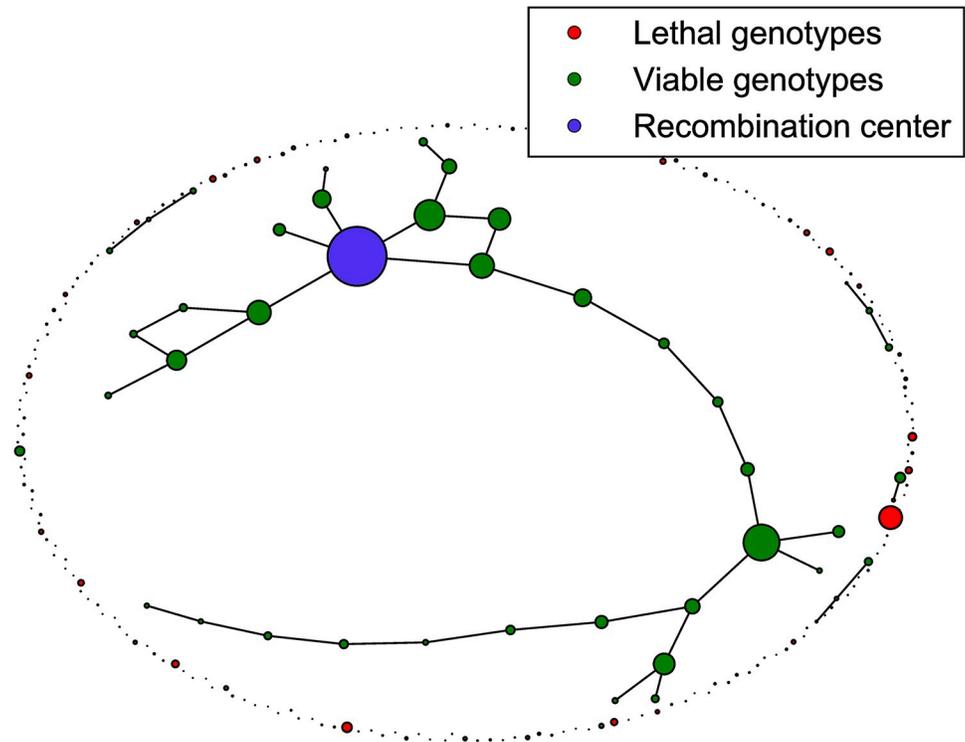
Fig 8. Network representation of a percolation landscape. The figure shows a percolation landscape with $L = 8$ loci and a fraction $p = 0.2$ of viable genotypes. Viable genotypes at Hamming distance $d = 1$ are connected by edges, and the node area of a genotype σ is proportional to λ_σ^6, where the recombination weight λ_σ is defined in Eq (10). The recombination center is the genotype with the largest recombination weight.

https://doi.org/10.1371/journal.pcbi.1006884.g008

The results of such a computation are shown in Fig 10. Similar to the mesa landscapes, a strong increase of mutational robustness is observed already for small rates of recombination, and the effect is largely independent of the recombination scheme. However, in contrast to the mesa landscape the robustness does not reach its maximal value $m = 1$ for $r = 1$ and small μ. This reflects the fact that maximally connected genotypes with $m_\sigma = 1$ are very rare at this particular value of p.
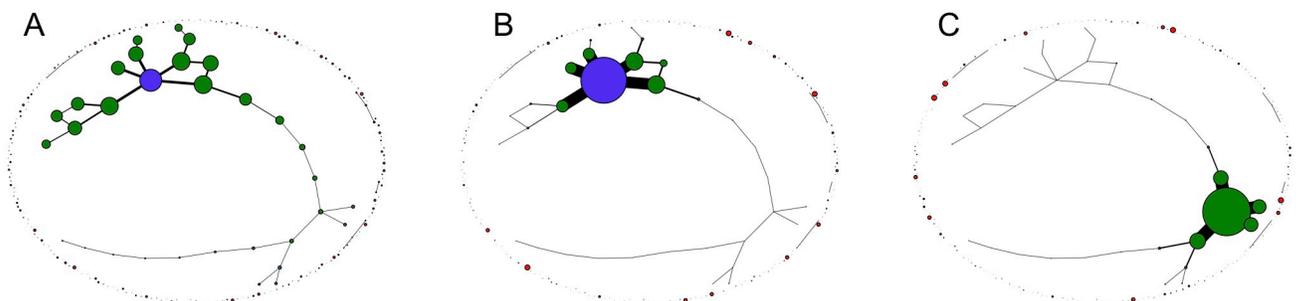

Fig 9. Stationary states in a percolation landscape. The figure shows three different stationary population distributions in the percolation landscape depicted in Fig 8. Node areas are proportional to the stationary frequency of the respective genotype in the population, and the edge width $e_{\sigma,\tau}$ between neighboring genotypes is proportional to the frequency of the more populated one, $e_{\sigma,\tau} \propto \max[f_\sigma^*, f_\tau^*]$. (A) Unique stationary state of a non-recombining population. (B,C) Stationary states for recombining populations undergoing uniform crossover with $r = 1$. The recombination center (purple) is the most populated genotype in (A,B), but not in (C). In all cases the mutation rate is $\mu = 0.01$.
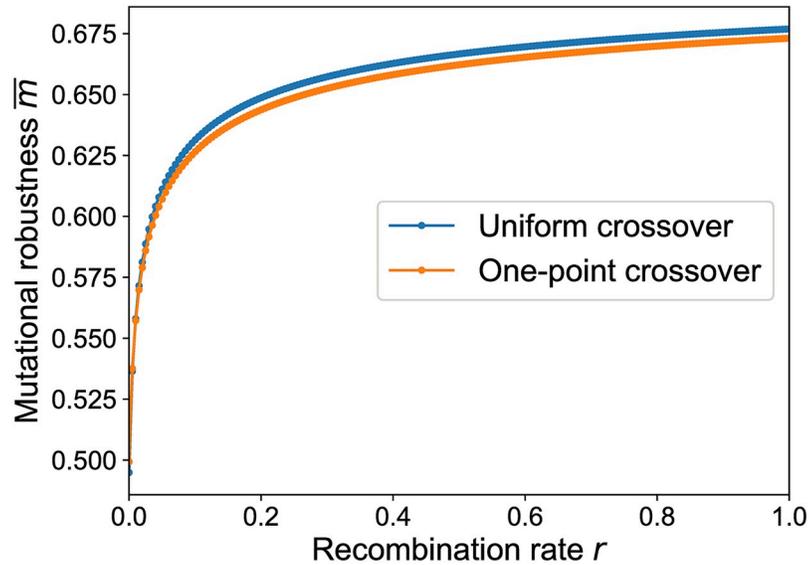
https://doi.org/10.1371/journal.pcbi.1006884.g009

Fig 10. Average mutational robustness in the percolation landscape as a function of recombination rate.
Mutational robustness is computed for 250 randomly generated percolation landscapes with $L = 6$ and $p = 0.4$, and the
results are averaged to obtain $\overline{m}(r)$. The mutation rate is $\mu = 0.001$.

For the purpose of comparison we also determined the average mutational robustness $\overline{m}_0$
of a uniform population distribution for the percolation model. Conditioned on the number
$v = |V|$ of viable genotypes and assuming that $v \geq 1$, we have $m_0(v, L) = n(v, L)/L$, where $n(v, L)$
is the average number of viable neighbors of a viable genotype. The latter is given by the
expression

$$n(L, v) = \frac{(v - 1)L}{2^L - 1}, \tag{38}$$

since for a given viable genotype there are $v - 1$ remaining genotypes, each of which has the
probability $L/(2^L - 1)$ to be a neighboring one. Taking into account that the number of viable
genotypes is binomially distributed with parameter p and that the empty hypercube ($v = 0$)
should yield $m_0 = 0$ we obtain

$$\overline{m}_0 = \sum_{v=1}^{2^L} \frac{(v - 1)}{2^L - 1} \binom{2^L}{v} p^v (1 - p)^{2^L - v} = \frac{2^L p - 1 + (1 - p)^{2^L}}{2^L - 1}, \tag{39}$$

which simplifies to $\overline{m}_0 = p$ when $2^L p \gg 1$. Note that the condition $2^L p \gg 1$ is naturally satis-
fied beyond the percolation threshold.

Fig 11 illustrates that the dynamics induced by mutation and selection already increase
mutational robustness compared to $\overline{m}_0$ and that the addition of recombination even further
increases mutational robustness for all values of p. The figure also displays the expected maxi-
mum number of viable neighbors of any genotype in the landscape, $\overline{m}_{\max}$, which provides an
upper bound on the robustness. The fact that the numerically determined robustness remains
below this bound for all p shows that the ability of recombination to locate the most connected
genotype is limited. In S1 Appendix it is shown that $\lim_{L \to \infty} \overline{m}_{\max} = 1$ for $p > \frac{1}{2}$.
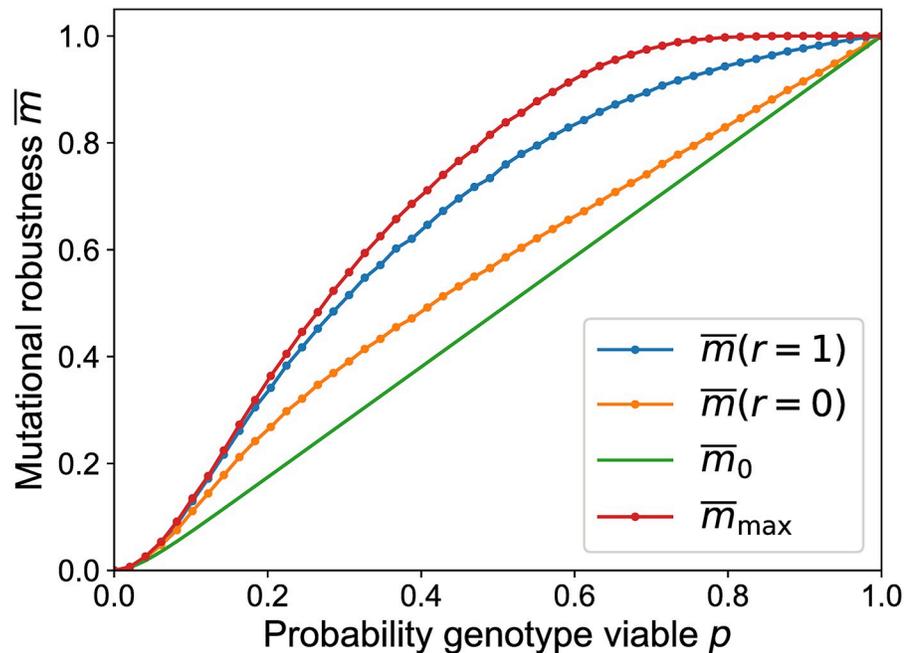
Fig 11. Mutational robustness in the percolation landscape as a function of the fraction of viable genotypes. The robustness for recombining ($\overline{m}(r = 1)$) and non-recombining ($\overline{m}(r = 0)$) populations is obtained by averaging over 6800 randomly generated landscapes with $L = 6$ and $\mu = 0.001$. In the same way the average maximal robustness $\overline{m}_{max}$ is estimated. The full line shows the analytic expression (39) for the robustness of a uniformly distributed population.

https://doi.org/10.1371/journal.pcbi.1006884.g011

As outlined above, the algorithm used to generate Figs 10 and 11 computes the mutational robustness of a particular stationary frequency distribution of the recombining population which is smoothly connected to the unique non-recombining stationary state. Although one expects this state to be representative in the sense of being reachable from many initial conditions, for large enough r there can be multiple stationary states that will generally display different robustness (see Fig 9). To illustrate this point, S7 Fig shows the results of a simulation of the percolation model where all stationary states were identified using localized initial conditions, and the mutational robustness was computed separately for each state. Whereas on average the mutational robustness is always enhanced by recombination, there are rare instances when recombination reduces the robustness compared to the non-recombining case. This may happen, for example, if recombination traps the population on a small island of viable genotypes [22, 26, 55, 56].

Sea-cliff landscapes

In this section we introduce a novel class of fitness-landscape models (to be called sea-cliff landscapes) that interpolates between the mesa and percolation landscapes. Similar to the mesa landscape, the fitness values of the sea-cliff model are determined by the distance to a reference genotype κ^*. The model differs from the mesa landscape in that it is not assumed that all genotypes have zero fitness beyond a certain number of mutations. Instead, the likelihood for a mutation to be lethal (to "fall off the cliff") is taken to increase with the Hamming distance from the reference genotype. This is mathematically realized by a Heaviside step function $\theta(x)$ that contains an uncorrelated random contribution η_σ and the distance measure

$d(\sigma, \kappa^*)$,

$$w_\sigma = \theta[\eta_\sigma - d(\sigma, \kappa^*)] = \begin{cases} 1, & \text{if } \eta_\sigma > d(\sigma, \kappa^*), \\ 0, & \text{if } \eta_\sigma < d(\sigma, \kappa^*). \end{cases} \tag{40}$$

This construction is similar in spirit to the definition of the Rough-Mount-Fuji model [66, 67].

The average shape of the landscape can be tuned by the mean c and the standard deviation s of the distribution of the random variables η_σ, which we assume to be Gaussian in the following. The average fitness at distance d from the reference sequence is then given by

$$\overline{w}(d) = \text{Prob}(w_\sigma = 1) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{d-c}{s\sqrt{2}}\right)\right], \tag{41}$$

where $\text{erf}(x)$ is the error function. Note that the mesa landscape is reproduced if we take the limit $s \to 0$ for fixed c in the range $k < c < k + 1$ and the percolation landscape is reproduced if we take a joint limit $s, |c| \to \infty$ with c/s fixed.

To fix c and s we introduce two distances $d_<$ and and $d_>$ such that $\overline{w}(d_<) = 0.99$ and $\overline{w}(d_>) = 0.01$, which leads to the relations

$$c = \frac{1}{2}(d_< + d_>) \quad \text{and} \quad s \approx 0.215(d_> - d_<). \tag{42}$$

The model can be generalized to include several predefined reference sequences,

$$w(\sigma) = \theta\left\{\sum_{\kappa^*} \theta[\eta_{\sigma,\kappa^*} - d(\sigma, \kappa^*)]\right\}, \tag{43}$$

which allows to create a genotype space with several highly connected clusters. Depending on the Hamming distance between the reference sequences and the variables c and s, clusters can be isolated or connected by viable mutations.

Fig 12 shows stationary states in the absence and presence of recombination for two different sea-cliff landscapes with one and two reference genotypes, respectively. Similar to the other landscape models, mutational robustness increases strongly with recombination, due to a population concentration within a neutral cluster. In the presence of two reference genotypes the recombining population should be concentrated within a single cluster. Otherwise lethal genotypes would be predominantly created through recombination of genotypes on different clusters. This observation can also be interpreted in the context of speciation due to genetic incompatibilities [49, 61]. Without recombination genotypes on both clusters have a nonvanishing frequency, but still the larger cluster is more populated. In contrast to the percolation landscape, robustness reaches a value close to unity for large r, because highly connected genotypes are abundant close to the reference sequence (S8 Fig).

Mutational robustness and recombination weight

Comparing Figs 6 and 10 and S8 Fig, the dependence of mutational robustness on the recombination rate is seen to be strikingly similar. Despite the very different landscape topographies, in all cases a small amount of recombination gives rise to a massive increase in robustness compared to the non-recombining baseline. For the mesa landscape this effect can be plausibly attributed to the focusing property of recombination, which counteracts the entropic spreading towards the fitness brink and localizes the population inside the plateau of viable genotypes. In the case of the holey landscapes, however, it is not evident that focusing the
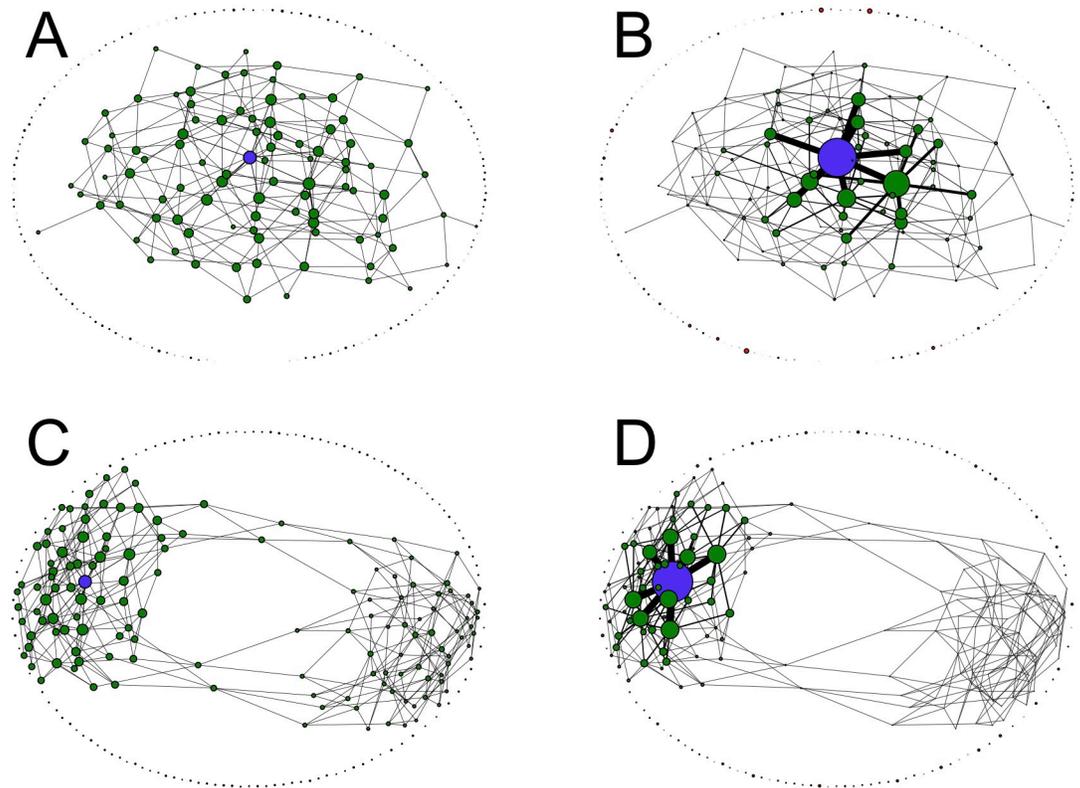
Fig 12. Stationary states in two different sea-cliff landscapes with and without recombination. (A,B) A single reference genotype with landscape parameters $L = 8$, $d_< = 1$ and $d_> = 6$. (C,D) Two reference genotypes which are antipodal to each other with landscape parameters $L = 8$, $d_< = 2$ and $d_> = 4.2$. (A,C) Stationary frequency distribution in the absence of recombination. (B,D) Stationary frequency distribution with uniform crossover and $r = 1$. In all cases node areas are proportional to genotype frequencies, and the recombination center is marked in blue. The edge width between neighboring genotypes is proportional to the frequency of the more populated one. The mutation rate is $\mu = 0.01$.

population towards the center of its genotypic range will on average increase robustness, since viable and lethal genotypes are randomly interspersed.

To establish the relation between recombination and mutational robustness on the level of individual genotypes, in Fig 13 we plot the recombination weight of each genotype against its robustness m_σ. A clear positive correlation between the two quantities is observed both for percolation and sea-cliff landscapes. Additionally we differentiate between viable and lethal genotypes. In the percolation landscape viable genotypes are uniformly distributed in the genotype space, which implies that lethal and viable genotypes have on average the same number of viable point-mutations. Nevertheless the recombination weight of viable genotypes is larger. The fitness of a genotype influences its own recombination weight, because the genotype itself is a possible parental genotype in the recombination event.

In non-neutral fitness landscapes the redistribution of the population through recombination competes with selection responding to fitness differences, and the generalized definition (11) of the recombination weight captures this interplay. To exemplify the relation between recombination weight and mutational robustness in this broader context, we use an empirical fitness landscape for the filamentary fungus *Aspergillus niger* originally obtained in [68]. In a nutshell, two strains of *A. niger* (N411 and N890) were fused to a diploid which is unstable and creates two haploids by random chromosome arrangement. Both strains are isogenic to each
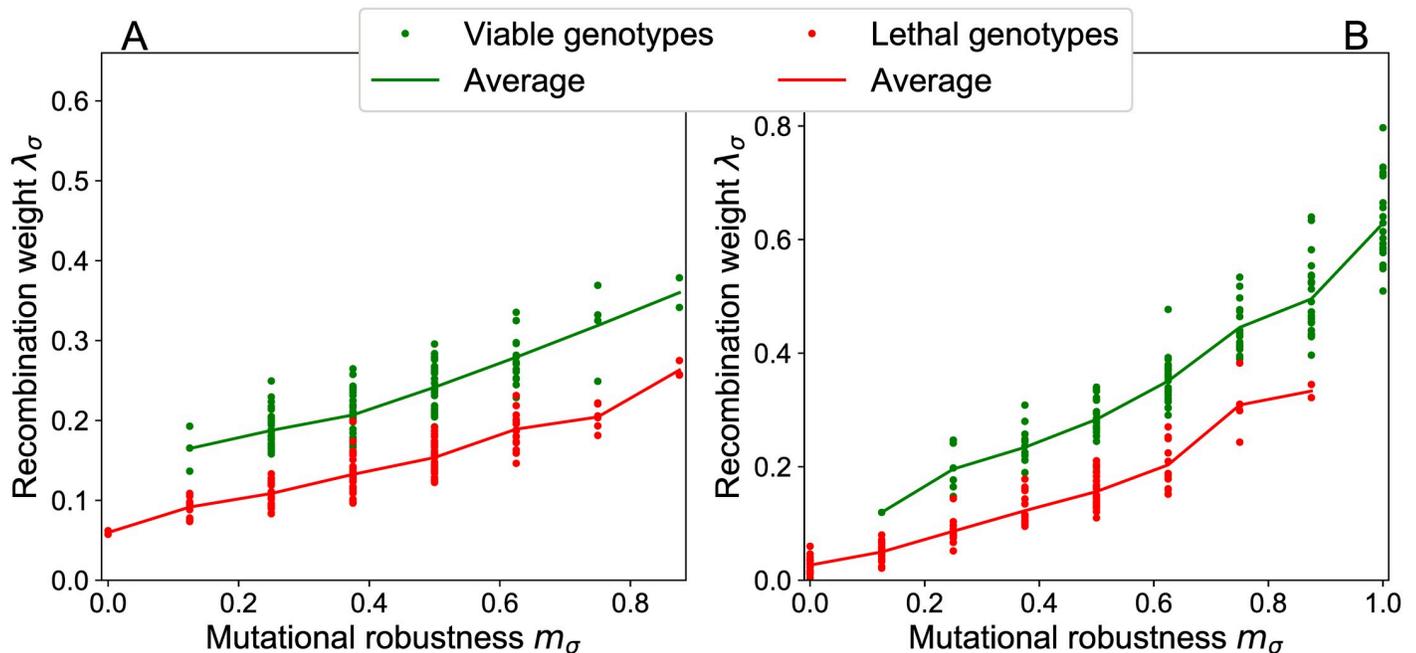
Fig 13. Mutational robustness correlates with recombination weight. The recombination weight of genotypes is plotted against their mutational robustness for (A) a percolation landscape with parameters $L = 8$, $p = 0.4$ and (B) a sea-cliff landscape with parameters $L = 8$, $d_< = 2$, $d_> = 6$. For the evaluation of the recombination weight (10), uniform crossover at rate $r = 1$ is assumed.

other, except that N890 has 8 marker mutations on different chromosomes, which were induced by low UV-radiation. Through this process $2^8 = 256$ haploid segregants can theoretically be created of which 186 were isolated in the experiment. As a result of a statistical analysis it was concluded that the missing 70 haploids have zero fitness [69].

In order to illustrate the fitness landscape, a network representation is employed where genotypes are arranged in a plane according to their fitness and their Hamming distance to the wild type, which in this case is the genotype of maximal fitness. In Fig 14A and 14B node sizes are adjusted to the recombination weights and mutational robustness of genotypes, respectively, in order to display the distribution of these quantities. In accordance with the analyses for neutral fitness landscapes, a clear correlation between the recombination weights and mutational robustness is shown in Fig 14C. Since fitness values are not binary we further consider the correlation between the recombination weights and fitness values (Fig 14D). The recombination center is one of the maximally robust genotypes with $m_\sigma = 1$, but it is not the fittest within this group. The wild type has maximal fitness but, by comparison, lower robustness ($m_\sigma = 7/8$).

Fig 15 highlights how the recombination weights change as a function of the recombination rate and how this affects the stationary state of a population. For small recombination rates the recombination weight of each genotype mainly depends on its own fitness, and therefore the wild type coincides with the recombination center. With increasing recombination rate the connectivity of the surrounding genotype network becomes more important and the recombination center switches to a genotype at Hamming distance $d = 2$. In contrast to the numerical protocol described previously, in the simulations used to generate Fig 15D–15F the population is reset to a uniform distribution before the recombination rate is increased. Otherwise the population would continue to adapt to the wild type, which has the highest fitness and from
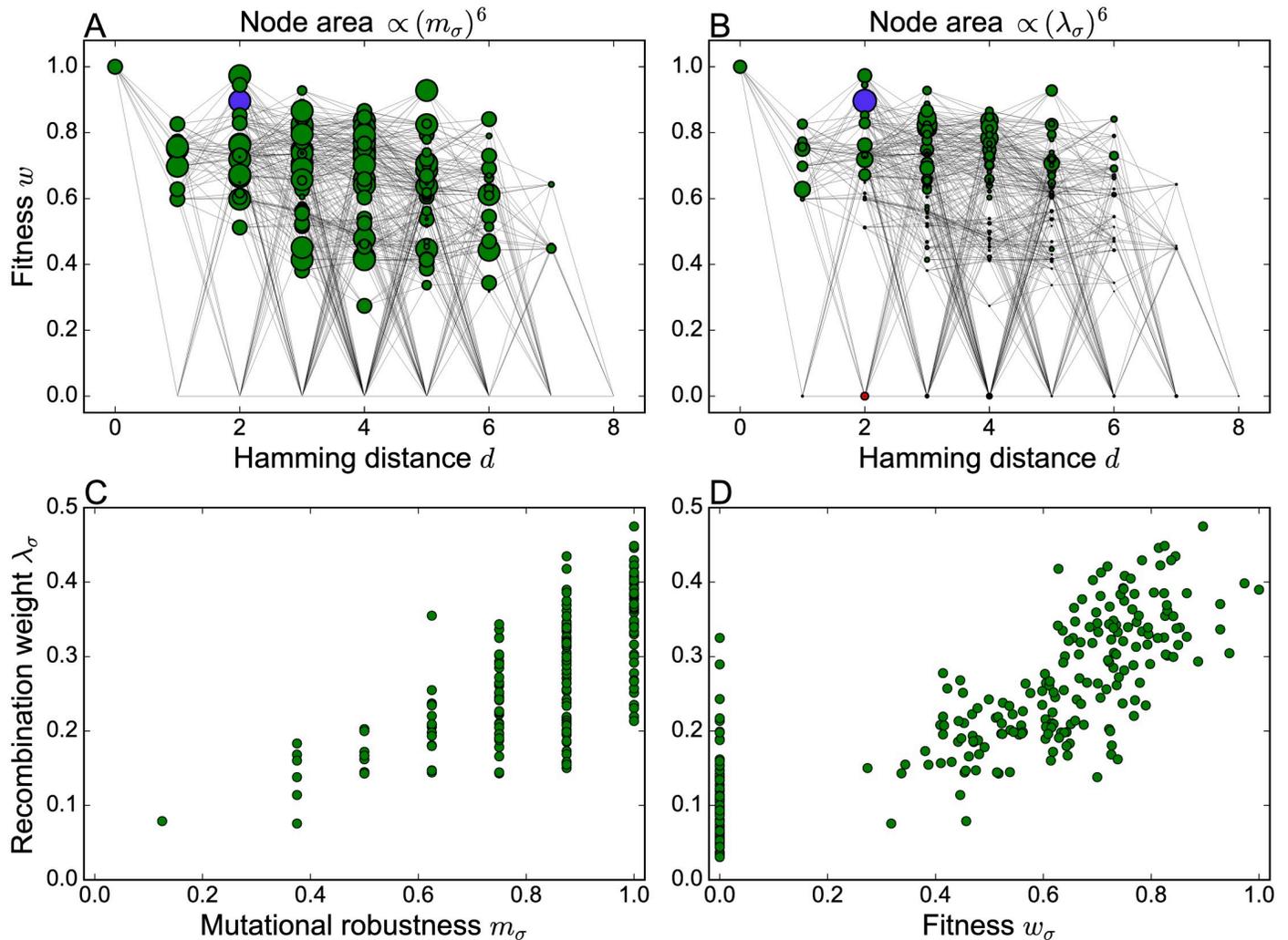
Fig 14. The empirical *A. niger* fitness landscape. (A,B) Two-dimensional network representation of the fitness landscape with node sizes determined by the mutational robustness m_σ and the recombination weight λ_σ, respectively. In order to make the differences between genotypes more conspicuous, the node area is chosen proportional to the sixth power of these quantities. The recombination weight is evaluated for uniform crossover with $r = 1$, and the recombination center is highlighted in purple. (C,D) Recombination weight plotted against mutational robustness and genotype fitness, respectively. Lethal genotypes with $w_\sigma = 0$ appear only in panel D.

https://doi.org/10.1371/journal.pcbi.1006884.g014

which it cannot escape because of peak trapping [22, 26]. Starting from an initially uniform distribution the population will adapt to one of three possible final genotypes which depend on the recombination rate. For small and large recombination rates the most abundant genotype coincides with the recombination center (Fig 15D and 15F), whereas for intermediate recombination rates the population chooses another genotype that is also located at Hamming distance $d = 2$ but has higher fitness (Fig 15E). The recombination center ultimately dominates the population, not only because it is maximally connected ($m_\sigma = 1$), but also because the genotypes that it is connected to have high fitness. In this sense the sequence of transitions in the most abundant genotype that occur with increasing recombination rate is akin to the scenario described previously in non-recombining populations as the "survival of the flattest" [48, 70]. Along this sequence mutational robustness increases monotonically whereas the average fitness of the population actually declines (S9 Fig).
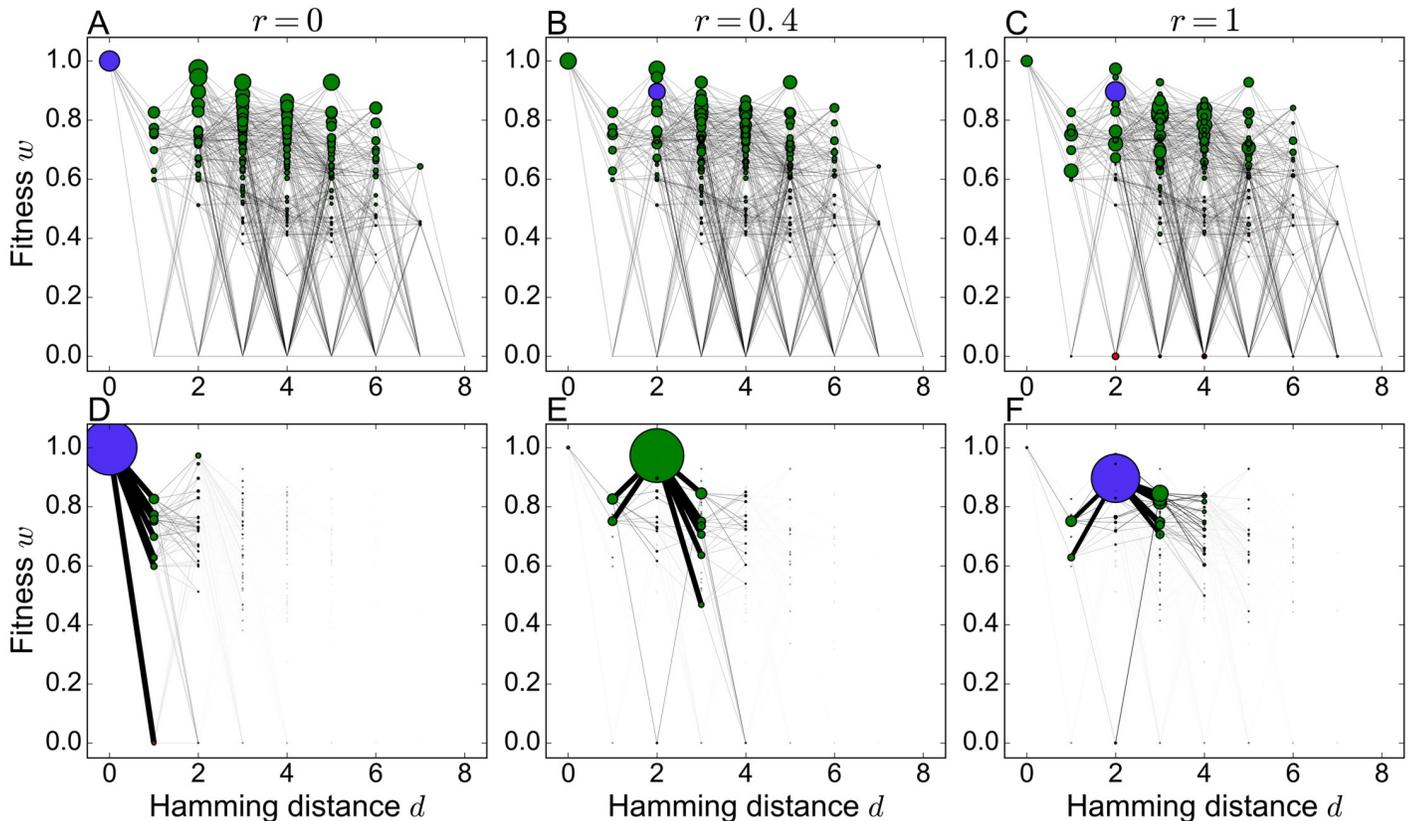
Fig 15. Recombination weights and stationary states at different recombination rates. (A-C) Two-dimensional network representation of the *A. niger* fitness landscape with node areas proportional to the sixth power of the recombination weight for recombination rates $r = 0$, $r = 0.4$ and $r = 1$, respectively. (D-F) Two-dimensional network representation of the *A. niger* fitness landscape with node areas proportional to the stationary genotype frequency at the same recombination rates and mutation rate $\mu = 0.005$. The edge width between neighboring genotypes is proportional to the frequency of the more populated one.

Discussion

Despite a century of research into the evolutionary bases of recombination, a general mechanism explaining the ubiquity of genetic exchange throughout the domains of life has not been found [17, 18]. Even within the idealized scenario of a population evolving in a fixed environment, whether or not recombination speeds up adaptation and leads to higher fitness levels depends in a complicated way on the structure of the fitness landscape and the parameters of the evolutionary dynamics [21–26].

The most important finding of the present work is that, by comparison, the effect of recombination on mutational robustness is much simpler and highly universal. Irrespective of the number of loci, the structure of the fitness landscape or the recombination scheme, recombination leads to a significant increase of robustness that is usually much stronger than the previously identified effect of selection [32–34]. This suggests that the evolution of recombination may be closely linked to the evolution of robustness, and that similar selective benefits are involved in the two cases. Although the relation of robustness to evolutionary fitness is subtle and not fully understood [27], it has been convincingly argued that robustness enhances evolvability and hence becomes adaptive in changing environments [29, 31, 71, 72]. A common perspective on recombination, robustness and evolvability can help to develop novel hypotheses about the evolutionary origins of these phenomena that can be tested in future computational or empirical studies.

On a quantitative level, we have shown that robustness generally depends on the ratio of recombination to mutation rates, and that the robustness-enhancing effect saturates when $r \gg \mu$. This observation highlights the importance of r/μ as an evolutionary parameter. Interestingly, even in bacteria and archaea, which have traditionally been regarded as essentially non-recombining, the majority of species displays values of r/μ that are significantly larger than one [73–75]. Similarly, a recent study of the evolution of *Siphoviridae* phages revealed a ratio of recombination events to mutational substitutions of about 24 [76]. In eukaryotes this ratio is expected to be considerably higher [40]. This indicates that most organisms maintain a rate of recombination that is sufficient to reap its evolutionary benefits in terms of increased robustness.

In order to clarify the mechanism through which recombination enhances robustness, we have introduced the concept of the recombination weight, which is a measure for the likelihood of a genotype to arise from the recombination of two viable parental genotypes. The recombination weight defines a "recombination landscape" over the space of genotypes which is similar in spirit to, but distinct from, previous mathematical approaches to conceptualizing the way in which recombining populations navigate a fitness landscape [77]. It is complementary to the more commonly used notion of a recombination load, which refers to the likelihood for a viable genotype to recombine to a lethal one [41, 42]. In many cases the maximum of the recombination weight correctly predicts the most populated genotype in a recombining population at low mutation rate. Moreover, the concept generalizes to non-neutral landscapes and thus permits to address situations where selection and recombination compete.

Provided recombination weight is correlated with mutational robustness for the individual genotypes, this explains the positive effect of recombination on the population-level robustness. Whether or not such a correlation exists will generally depend on the structure of the fitness landscapes. For simple neutral landscapes such as the mesa landscape it is an immediate consequence of the focusing property of recombination, but for more complex neutral networks the relationship between the two quantities is nontrivial and needs to be studied on a case-by-case basis. Although a positive correlation was observed numerically both for the holey landscapes and the empirical landscape considered in this work, it is not difficult to construct landscapes where the genotypes with high recombination weight are not highly robust. As a simple but instructive example, in S10 Fig we show results for an 'atoll' landscape where a ring of viable genotypes surrounds a central hole of lethals.

Throughout this work the effects of genetic drift have been neglected. We expect that our results will be applicable to finite populations as long as the population is sufficiently diverse rather than being monomorphic. This requires the population-wide mutation rate $N\mu L$ to be much larger than unity [32, 44]. If $N\mu L \ll 1$ the population is almost always monomorphic and recombination has no effect. In this regime the population explores the fitness landscape as a random walker and the observed mutational robustness is the uniform robustness m_0. In S11 Fig we present the results of finite population simulations on a mesa landscape, which show a sharp transition from the random walk regime to the behavior predicted by the deterministic theory when $N\mu L \sim 1$.

Future work should be directed towards extending the present investigation to more realistic genotype-phenotype maps arising, for example, from the secondary structures of biopolymers such as RNA or proteins [39, 40, 44], or from simple genetic, metabolic or logical networks [29, 41, 43, 78]. There is ample evidence from numerical studies that a favorable effect of recombination on mutational robustness is present also in these more complex systems, but a detailed analysis of the underlying mechanism has not been carried out. This would entail, in particular, the generalization to genotype spaces composed of sequences

carrying more than two alleles per site. We expect that at least part of the analysis for the mesa landscapes carries over to this setting, and in fact some results for the non-recombining case have already been obtained [48]. More importantly, the role of the topology of the corresponding neutral networks in shaping the correlation between recombination weight and robustness needs to be explored systematically. Research along these lines will help to corroborate the relationship between recombination and robustness that we have sketched, and to further elucidate the origins of these two pervasive features of biological evolution.

Supporting information

S1 Appendix. This appendix contains detailed derivations of analytic results presented in the main text.
(PDF)

S1 Fig. Population heterogeneity decreases with increasing recombination rate. The figure shows the entropy of the genotype frequency distribution in the two-locus model defined as $S = -\sum_\sigma f_\sigma^* \ln \left(f_\sigma^* \right)$. For small mutation rates the strongly recombining population primarily consists of a single genotype, which implies that $S \to 0$.
(PDF)

S2 Fig. Mutational robustness for the mesa landscape with communal recombination. The figure compares the analytic approximations in Eqs (29) and (30) to the numerical solution of the stationary genotype frequency distribution for the communal recombination scheme. The two panels show the mutational robustness as a function of the genome-wide mutation rate in linear (A) and double-logarithmic (B) scales, respectively. The parameters of the mesa landscape are $L = 30$ and $k = 3$.
(PDF)

S3 Fig. Mutational robustness in a mesa landscape with different recombination schemes. The figure compares the analytic results for communal recombination (m_{cr}) with numerical data obtained using uniform crossover (m_{uc}) and one-point crossover (m_{opc}) at $r = 1$. The landscape parameters are $L = 5$, $k = 2$ and robustness is plotted as a function of the genome-wide mutation rate $L\mu$. (A) Mutational robustness on linear scales. (B) Double-logarithmic plot of $1 - m$ vs. $L\mu$, illustrating the power-law behavior $1 - m \sim (L\mu)^b$ with the exponent $b = k/(k + 1) = 2/3$ predicted by the analysis of the communal recombination model.
(PDF)

S4 Fig. Mutational robustness for the mesa landscape in the absence of recombination. The figure compares the analytic predictions in Eqs (35) and (36) to the numerical solution for the genotype frequency distribution in the absence of recombination. The two panels show the mutational robustness (A) after selection and (B) after mutation as a function of the scaled mesa width $x_0 = k/L$ for $L = 1000$ and $U = 0.01$.
(PDF)

S5 Fig. Mutational robustness in mesa landscapes with and without recombination. Numerical results for communal recombination (m_{cr}) and no recombination (m_{nr}) are shown as dots. The mutational robustness m_0 of a uniformly distributed population, given by Eq (37), as well as the analytic expressions Eqs (30) and (36) are depicted as lines. (A) Robustness as a function of mutation rate $U = L\mu$ for a landscape with $L = 1000$ and $k = 10$. (B) Robustness as a function of mesa width k at fixed $L = 1000$ and $U = L\mu = 0.01$. (C) Robustness as a function of

genome length L at fixed $k = 10$ and $U = 0.01$. (D) Robustness as a function of genome length L at fixed $k = 10$ and $\mu = 0.001$.
(PDF)

S6 Fig. Recombination weight in a mesa landscape. The parameters of the mesa landscape are $L = 100$ and $k = 10$. For $r = 0$ the recombination weight is directly proportional to the fitness and hence equal for all viable genotypes. Already small rates of recombination are sufficient to redistribute the recombination weight such that the weight of genotypes with small Hamming distance is strongly enhanced. Beyond $d = 20$ the recombination weight is identically zero, since the recombinant of two viable genotypes cannot carry more than $2k$ mutations.
(PDF)

S7 Fig. Mutational robustness for different stationary states within a percolation landscape. The figure compares the mutational robustness of non-recombining ($r = 0$) and recombining ($r = 1$) populations on individual realizations of the percolation model with $L = 6$ and three values of p. In order to obtain different stationary states we used localized initial population distributions of the form $f_\tau(0) = \delta_{\tau\sigma}$ for all genotypes with mutational robustness $m_\sigma \neq 0$ and propagated them until stationarity. Since the stationary populations are usually highly concentrated for large r and small μ, this is a natural choice in order to access all stationary states. Each data point represents the robustness of the recombining population $m(r = 1)$ for a particular stationary state. Data points within the same landscape are plotted above the corresponding unique robustness of the non-recombining population $m(r = 0)$ and connected by a vertical line. The orange crosses show the average over all initial conditions.
(PDF)

S8 Fig. Average mutational robustness in the sea-cliff landscape as a function of recombination rate. Mutational robustness is computed for 200 randomly generated sea-cliff landscapes with parameters $L = 6$, $d_< = 1$ and $d_> = 5$, and the results are averaged to obtain $\overline{m}(r)$. The mutation rate is $\mu = 0.001$.
(PDF)

S9 Fig. Mutational robustness and average fitness in the empirical *A. niger* fitness landscape. The mutational robustness and the population-averaged fitness in the stationary state are computed as a function of recombination rate by evolving the population from a uniform initial genotype distribution at mutation rate $\mu = 0.005$. Jumps mark changes in the most populated genotype.
(PDF)

S10 Fig. Recombination on an atoll landscape. This landscape is similar to the mesa landscape but includes an inner critical radius within which genotypes are lethal. In this example the inner radius is chosen to be 1 such that only the wild type is lethal. The outer radius is 2 and the sequence length is $L = 7$. The recombination rate is $r = 1$ and the mutation rate is $\mu = 0.001$. The frequencies f_n of the stationary state at the same Hamming distance n are lumped together. The population is concentrated at distance 1 which is most robust since only one point mutation is lethal, but the recombination center coincides with the lethal wild type. This example shows that the correlation between recombination weight and mutational robustness depends on the topology of the neutral network.
(PDF)

S11 Fig. Finite population size effects. The figure shows the mutational robustness in a mesa landscape with parameter $L = 6$, $k = 2$ as a function of mutation rate. The finite population

results were obtained using Wright-Fisher dynamics for $N = 1000$ individuals. For small mutation rates such that $N\mu L \ll 1$ the monomorphic population performs a random walk among viable genotypes, which leads to the uniform mutational robustness m_0 given by Eq (37) (green dashed line). In this regime recombination cannot have any effect. For $N\mu L > 1$ the robustness rises sharply to the value predicted by the infinite population approach. At the maximal mutation rate $\mu = 0.5$ the population is uniformly distributed among all (lethal or viable) genotypes after the mutation step and recombination has again no effect.
(PDF)

Acknowledgments

Author Contributions

References

1. Simon-Loriere E, Holmes EC. Why do RNA viruses recombine? Nat Rev Microbiol. 2011; 9:617–626. https://doi.org/10.1038/nrmicro2614 PMID: 21725337

2. Redfield RJ. Do bacteria have sex? Nat Rev Genet. 2001; 2:634–639. https://doi.org/10.1038/35084593 PMID: 11483988

3. Feil EJ, Spratt BG. Recombination and the Population Structures of Bacterial Pathogens. Annu Rev Microbiol. 2001; 55:561–590. https://doi.org/10.1146/annurev.micro.55.1.561 PMID: 11544367

4. Maynard Smith J. The Evolution of Sex. Cambridge, UK: Cambridge University Press; 1978.

5. Michod RE, Levin BR, editors. The Evolution of sex: an examination of current ideas. Sunderland, Massachusetts: Sinauer Associates Inc.; 1988.

6. Holmes CM, Nemenman I, Weissman DB. Increased adaptability to sudden environmental change can more than make up for the two-fold cost of males. Europhys Lett. 2018; 123:58001. https://doi.org/10.1209/0295-5075/123/58001

7. D J P Engelmoer ID, Rozen DE. Conservative Sex and the Benefits of Transformation in Streptococcus pneumoniae. PLOS Pathogens. 2013; 9:e1003758. https://doi.org/10.1371/journal.ppat.1003758

8. Moradigaravand D, Engelstädter J. The Evolution of Natural Competence: Disentangling Costs and Benefits of Sex in Bacteria. Amer Nat. 2013; 182:E112–E126. https://doi.org/10.1086/671909

9. Weismann A. Essays Upon Heredity and Kindred Biological Problems. Oxford University Press; 1889.

10. Muller HJ. Some genetic aspects of sex. Amer Nat. 1932; 66:118–138. https://doi.org/10.1086/280418

11. Muller HJ. The relation of recombination to mutational advance. Mutat Res. 1964; 1:2–9. https://doi.org/10.1016/0027-5107(64)90047-8

12. Felsenstein J. The evolutionary advantage of recombination. Genetics. 1974; 78:737–756. PMID: 4448362

13. Kondrashov AS. Deleterious mutations and the evolution of sexual reproduction. Nature. 1988; 336:435–440. https://doi.org/10.1038/336435a0 PMID: 3057385

14. Kondrashov AS. Classification of hypotheses on the advantage of amphimixis. J Hered. 1993; 84:372–387. https://doi.org/10.1093/oxfordjournals.jhered.a111358 PMID: 8409359

15. Feldman MW, Otto SP, Christiansen FB. Population genetic perspectives on the evolution of recombination. Annu Rev Genet. 1997; 30:261–295. https://doi.org/10.1146/annurev.genet.30.1.261

16. Burt A. Perspective: Sex, recombination, and the efficacy of selection—was Weismann right? Evolution. 2000; 54:337–351. https://doi.org/10.1111/j.0014-3820.2000.tb00038.x PMID: 10937212

17. Otto SP, Lenormand T. Resolving the paradox of sex and recombination. Nat Rev Genet. 2002; 3:252–261. https://doi.org/10.1038/nrg761 PMID: 11967550

18. de Visser JAGM, Elena SF. The evolution of sex: empirical insights into the roles of epistasis and drift. Nat Rev Genet. 2007; 8:139–149. https://doi.org/10.1038/nrg1985 PMID: 17230200

19. Otto SP. The evolutionary enigma of sex. Amer Nat. 2009; 174:S1–S14. https://doi.org/10.1086/599084

20. de Visser JAGM, Krug J. Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet. 2014; 15:480–490. https://doi.org/10.1038/nrg3744 PMID: 24913663

21. Kondrashov FA, Kondrashov AS. Multidimensional epistasis and the disadvantage of sex. Proc Nat Acad Sci USA. 2001; 98:12089–12092. https://doi.org/10.1073/pnas.211214298 PMID: 11593020

22. de Visser JAGM, Park SC, Krug J. Exploring the Effect of Sex on Empirical Fitness Landscapes. Amer Nat. 2009; 174:S15–S30. https://doi.org/10.1086/599081

23. Misevic D, Kouyos RD, Bonhoeffer S. Predicting the Evolution of Sex on Complex Fitness Landscapes. PLOS Comp Biol. 2009; 5:e1000510. https://doi.org/10.1371/journal.pcbi.1000510

24. Moradigaravand D, Engelstädter J. The Effect of Bacterial Recombination on Adaptation on Fitness Landscapes with Limited Peak Accessibility. PLOS Comp Biol. 2012; 8:e1002735. https://doi.org/10.1371/journal.pcbi.1002735

25. Moradigaravand D, Kouyos R, Hinkley T, Haddad M, Petropoulos CJ, Engelstädter J, et al. Recombination Accelerates Adaptation on a Large-Scale Empirical Fitness Landscape in HIV-1. PLOS Genetics. 2014; 10:e1004439. https://doi.org/10.1371/journal.pgen.1004439 PMID: 24967626

26. Nowak S, Neidhart J, Szendro IG, Krug J. Multidimensional Epistasis and the Transitory Advantage of Sex. PLOS Comp Biol. 2014; 10:e1003836. https://doi.org/10.1371/journal.pcbi.1003836

27. de Visser JAGM, Hermisson J, Wagner GP, Meyers LA, Bagheri-Chaichian H, Blanchard JL, et al. Perspective: Evolution and detection of genetic robustness. Evolution. 2003; 57:1959–1972. https://doi.org/10.1554/02-750R PMID: 14575319

28. Kitano H. Biological robustness. Nat Rev Genet. 2004; 5:826–837. https://doi.org/10.1038/nrg1471 PMID: 15520792

29. Wagner A. Robustness and Evolvability in Living Systems. Princeton: Princeton University Press; 2005.

30. Lenski RE, Barrick JE, Ofria C. Balancing robustness and evolvability. PLOS Biol. 2006; 4:e428. https://doi.org/10.1371/journal.pbio.0040428 PMID: 17238277

31. Masel J, Trotter MV. Robustness and evolvability. Trends Genet. 2010; 26:406–414. https://doi.org/10.1016/j.tig.2010.06.002 PMID: 20598394

32. van Nimwegen E, Crutchfield JP, Huynen M. Neutral evolution of mutational robustness. Proc Nat Acad Sci USA. 1999; 96:9716–9720. https://doi.org/10.1073/pnas.96.17.9716 PMID: 10449760

33. Bornberg-Bauer E, Chan HS. Modeling evolutionary landscapes: Mutational stability, topology, and superfunnels in sequence space. Proc Nat Acad Sci USA. 1999; 96:10689–10694. https://doi.org/10.1073/pnas.96.19.10689 PMID: 10485887

34. Wilke CO. Adaptive evolution on neutral networks. Bull Math Biol. 2001; 63:715–730. https://doi.org/10.1006/bulm.2001.0244 PMID: 11497165

35. Gavrilets S. Fitness Landscapes and the Origin of Species. Princeton: Princeton University Press; 2004.

36. Boerlijst MC, Bonhoeffer S, Nowak MA. Viral Quasi-Species and Recombination. Proc R Soc Lond B. 1996; 263:1577–1584. https://doi.org/10.1098/rspb.1996.0231

37.   Wilke CO, Adami C. Evolution of mutational robustness. Mutat Res. 2003; 522:3–11. https://doi.org/10.1016/s0027-5107(02)00307-x PMID: 12517406

38.   Gardner A, Kalinka AT. Recombination and the evolution of mutational robustness. J Theor Biol. 2006; 241:707–715. https://doi.org/10.1016/j.jtbi.2006.01.011 PMID: 16487979

39.   Huynen MA, Hogeweg P. Pattern Generation in Molecular Evolution: Exploitation of the Variation in RNA Landscapes. J Mol Evol. 1994; 39:71–79. https://doi.org/10.1007/BF00178251 PMID: 7520506

40.   Xia Y, Levitt M. Roles of mutation and recombination in the evolution of protein thermodynamics. Proc Nat Acad Sci USA. 2002; 99:10382–10387. https://doi.org/10.1073/pnas.162097799 PMID: 12149452

41.   Azevedo RBR, Lohaus R, Srinivasan S, Dang KK, Burch CL. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. Nature. 2006; 440:87–90. https://doi.org/10.1038/nature04488 PMID: 16511495

42.   Singhal S, Gomez SM, Burch CL. Recombination drives the evolution of mutational robustness. Curr Opin Syst Biol. 2019; 13:142–149. https://doi.org/10.1016/j.coisb.2018.12.003

43.   Hu T, Banzhaf W, Moore JH. The effects of recombination on phenotypic exploration and robustness in evolution. Artificial Life. 2014; 20:457–470. https://doi.org/10.1162/ARTL_a_00145 PMID: 25148550

44.   Szöllősi GJ, Derényi I. The effect of recombination on the neutral evolution of genetic robustness. Math Biosci. 2008; 214:58–62. https://doi.org/10.1016/j.mbs.2008.03.010 PMID: 18490032

45.   Gerland U, Hwa T. On the Selection and Evolution of Regulatory DNA Motifs. J Mol Evol. 2002; 55:386–400. https://doi.org/10.1007/s00239-002-2335-z PMID: 12355260

46.   Peliti L. Quasispecies evolution in general mean-field landscapes. Europhys Lett. 2002; 57:745–751. https://doi.org/10.1209/epl/i2002-00526-5

47.   Berg J, Willmann S, Lässig M. Adaptive evolution of transcription factor binding sites. BMC Evol Biol. 2004; 4:42. https://doi.org/10.1186/1471-2148-4-42 PMID: 15511291

48.   Wolff A, Krug J. Robustness and epistasis in mutation-selection models. Phys Biol. 2009; 6:036007. https://doi.org/10.1088/1478-3975/6/3/036007 PMID: 19411737

49.   Gavrilets S. Evolution and speciation on holey adaptive landscapes. Trends Ecol Evol. 1997; 12:307–312. https://doi.org/10.1016/S0169-5347(97)01098-7 PMID: 21238086

50.   Bürger R. The Mathematical Theory of Selection, Recombination and Mutation. Chichester, UK: John Wiley & Sons; 2000.

51.   Crow JF, Kimura M. Evolution in sexual and asexual populations. Amer Nat. 1965; 99:439–450. https://doi.org/10.1086/282389

52.   Eshel I, Feldman MW. On the evolutionary effect of recombination. Theor Pop Biol. 1970; 1:88–100. https://doi.org/10.1016/0040-5809(70)90043-2

53.   Kimura M. The role of compensatory neutral mutations in molecular evolution. J Genet. 1985; 64:7–19. https://doi.org/10.1007/BF02923549

54.   Higgs PG. Compensatory neutral mutations and the evolution of RNA. Genetica. 1998; 102/103:91–101. https://doi.org/10.1023/A:1017059530664

55.   Park SC, Krug J. Bistability in two-locus models with selection, mutation, and recombination. J Math Biol. 2011; 62:763–788. https://doi.org/10.1007/s00285-010-0352-x PMID: 20617437

56.   Altland A, Fischer A, Krug J, Szendro IG. Rare Events in Population Genetics: Stochastic Tunneling in a Two-Locus Model with Recombination. Phys Rev Lett. 2011; 106:088101. https://doi.org/10.1103/PhysRevLett.106.088101 PMID: 21405603

57.   Weinreich DM, Sindi S, Watson RA. Finding the boundary between evolutionary basins of attraction, and implications for Wright's fitness landscape analogy. J Stat Mech. 2013; 2013:P01001. https://doi.org/10.1088/1742-5468/2013/01/P01001

58.   Nowak MA, Boerlijst MC, Cooke J, Maynard Smith J. Evolution of genetic redundancy. Nature. 1997; 388:167–171. https://doi.org/10.1038/40618 PMID: 9217155

59.   Phillips PC, Johnson NA. The Population Genetics of Synthetic Lethals. Genetics. 1998; 150:449–458. PMID: 9725860

60.   Baake E, Georgii HO. Mutation, selection, and ancestry in branching models: a variational approach. J Math Biol. 2007; 54:257–303. https://doi.org/10.1007/s00285-006-0039-5 PMID: 17075709

61.   Paixão T, Bassler KE, Azevedo RBR. Emergent speciation by multiple Dobzhansky-Muller incompatibilities. bioRxiv. 2015;Available from: https://www.biorxiv.org/content/early/2015/07/07/008268.

62.   Neher RA, Shraiman BI, Fisher DS. Rate of Adaptation in Large Sexual Populations. Genetics. 2010; 184:467–481. https://doi.org/10.1534/genetics.109.109009 PMID: 19948891

63. Gavrilets S, Gravner J. Percolation on the Fitness Hypercube and the Evolution of Reproductive Isolation. J Theor Biol. 1997; 184:51–64. https://doi.org/10.1006/jtbi.1996.0242 PMID: 9039400

64. Reidys CM. Random Induced Subgraphs of Generalized *n*-Cubes. Adv Appl Math. 1997; 19:360–377. https://doi.org/10.1006/aama.1997.0553

65. Reidys CM. Large components in random induced subgraphs of *n*-Cubes. Discrete Mathematics. 2009; 309:3113–3124.

66. Aita T, Uchiyama H, Inaoka T, Nakajima M, Kokubo T, Husimi Y. Analysis of a Local Fitness Landscape with a Model of the Rough Mt. Fuji-Type Landscape: Application to Prolyl Endopeptidase and Thermolysin. Biopoly. 2000; 54:64–79. https://doi.org/10.1002/(SICI)1097-0282(200007)54:1%3C64::AID-BIP70%3E3.0.CO;2-R

67. Neidhart J, Szendro IG, Krug J. Adaptation in Tunably Rugged Fitness Landscapes: The Rough Mount Fuji Model. Genetics. 2014; 198:699–721. https://doi.org/10.1534/genetics.114.167668 PMID: 25123507

68. de Visser JAGM, Hoekstra RF, van den Ende H. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. Evolution. 1997; 51:1499–1505. https://doi.org/10.1111/j.1558-5646.1997.tb01473.x PMID: 28568630

69. Franke J, Klözer A, de Visser JAGM, Krug J. Evolutionary accessibility of mutational pathways. PLOS Comp Biol. 2011; 7:e1002134. https://doi.org/10.1371/journal.pcbi.1002134

70. Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. Nature. 2001; 412:331–333. https://doi.org/10.1038/35085569 PMID: 11460163

71. Draghi JA, Parsons TL, Wagner GP, Plotkin JB. Mutational robustness can facilitate adaptation. Nature. 2010; 463:353–355. https://doi.org/10.1038/nature08694 PMID: 20090752

72. Payne JL, Wagner A. The causes of evolvability and their evolution. Nat Rev Genet. 2019; 20:24–38. https://doi.org/10.1038/s41576-018-0069-z PMID: 30385867

73. Guttman DS, Dykhuizen DE. Clonal divergence in Escherichia-coli as a result of recombination, not mutation. Science. 1994; 266:1380–1383.

74. Vos M, Didelot X. A comparison of homologous recombination rates in bacteria and archaea. ISME J. 2009; 3:199–208. https://doi.org/10.1038/ismej.2008.93 PMID: 18830278

75. Didelot X, Maiden MCJ. Impact of recombination on bacterial evolution. Trends Microbiol. 2010; 18:315–322. https://doi.org/10.1016/j.tim.2010.04.002 PMID: 20452218

76. Kupczok A, Neve H, Huang KD, Hoeppner MP, Heller KJ, Franz CMAP, et al. Rates of Mutation and Recombination in *Siphoviridae* Phage Genome Evolution over Three Decades. Mol Biol Evol. 2018; 35:1147–1159. https://doi.org/10.1093/molbev/msy027 PMID: 29688542

77. Stadler PF, Wagner G. Algebraic Theory of Recombination Spaces. Evol Comp. 1997; 5:241–275. https://doi.org/10.1162/evco.1997.5.3.241

78. García-Martín JA, Catalán P, Manrubia S, Cuesta JA. Statistical theory of phenotype abundance distributions: A test through exact enumeration of genotype spaces. Europhys Lett. 2018; 123:28001. https://doi.org/10.1209/0295-5075/123/28001