

# Localization of neutral evolution: selection for mutational robustness and the maximal entropy random walk

Matteo Smerlak<sup>1</sup>

<sup>1</sup>Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

January 28, 2020

## Abstract

If many mutations confer no immediate selective advantage, they can pave the way for the discovery of fitter phenotypes and their subsequent positive selection. Understanding the reach of neutral evolution is therefore a key problem linking diversity, robustness and evolvability at the molecular scale. While this process is usually described as a random walk in sequence space with clock-like regularity, new effects can arise in large microbial or viral populations where new mutants arise before old ones can fix. Here I show that the clonal interference of neutral variants shuts off the access to neutral ridges and thus induces localization within the robust cores of neutral networks. As a result, larger populations can be less effective at exploring sequence space than smaller ones—a counterintuitive limitation to evolvability which invalidates analogies between evolution and percolation. I illustrate these effects by revisiting Maynard Smith’s word-game model of protein evolution. Interestingly, the phenomenon of neutral interference connects evolutionary dynamics to a Markov process known in network science as the maximal-entropy random walk; its special properties imply that, when many neutral variants interfere in a population, evolution chooses mutational paths—not individual mutations—uniformly at random.

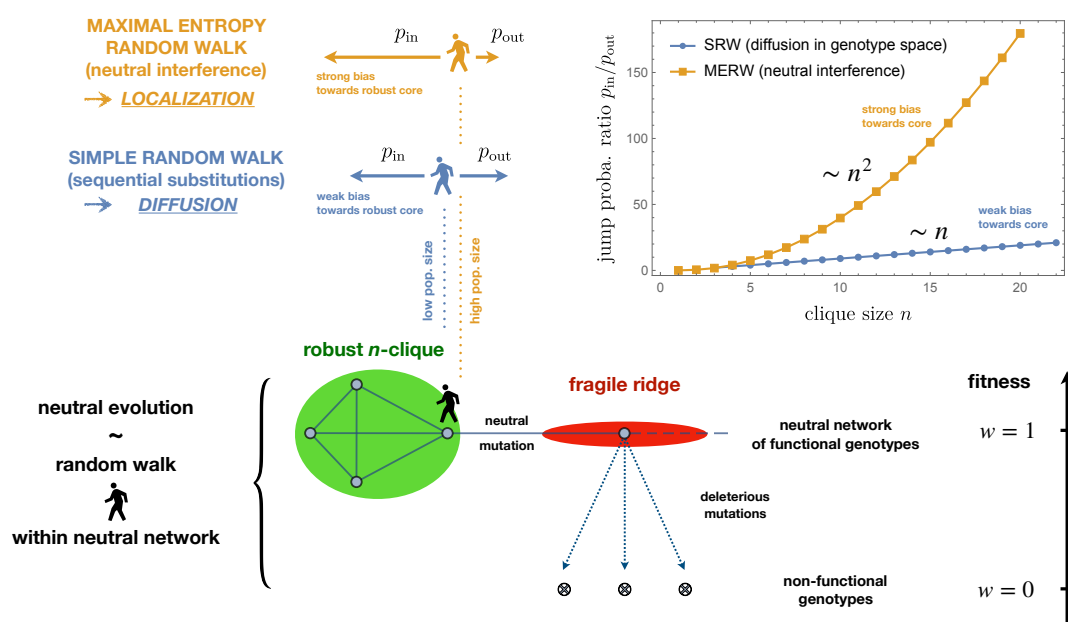


Figure 1: Two Markovian limits of neutral evolution with sharply different dynamics. When mutations are rare and populations mostly homogeneous, neutral evolution amounts to a simple random walk (SRW) along a neutral network  $G$ . In larger populations, the interference of multiple neutral mutants leads to the selection for mutational robustness, which can be described as a maximal entropy random walk (MERW) on  $G$ . The qualitative difference between the SRW and the MERW is illustrated here by a simple network configuration where an  $n$ -clique is connected to a one-dimensional ridge. In that case, the SRW on the edge of the clique with jump back into it with probability  $p_{in} \sim np_{out}$ , while the MERW will jump back with the much higher probability  $p_{in} \sim n^2 p_{out}$ . This stronger attraction towards robust cores can induce the localization of populations within  $G$ , invalidating commonly analogies between neutral evolution and diffusion (or percolation) in sequence space.

## Introduction

Kimura famously championed the view that a large part of all evolutionary change in genomes confer no selective advantage, *i.e.* molecular evolution is largely *neutral* (Kimura, 1983). Initially based on high observed substitution rates (Kimura, 1968; King and Jukes, 1969), this hypothesis was supported by the later discovery of extended neutral networks—sets of sequences connected by one-point mutations with equivalent phenotype or function (Smith, 1970)—in many molecular genotype-to-phenotype maps, e.g. in RNA secondary structure (Fontana et al., 1993), protein struc-

ture (Babajide et al., 1997) or transcriptional regulation networks (Ciliberti et al., 2007). While the exact rate of adaptive vs. neutral evolution remains under investigation (Eyre-Walker, 2006), neutralism is now widely understood as a central aspect of evolutionary dynamics (Nei et al., 2010). Besides ensuring a high level of mutational robustness (van Nimwegen et al., 1999), neutral evolution can enhance evolvability by providing access to novel—and possibly fitter—phenotypes (Wagner, 2008). In this way, the variation generated by neutral evolution enables the “arrival of the fittest” (Wagner, 2014) and can facilitate adaptation in new selective environments (Gibson and Dworkin, 2004).

Because mutations are random events, it is tempting to picture neutral evolution as a simple random walk (SRW) taking place within neutral networks, and indeed this is how it is usually described, both verbally (Smith, 1970) and in quantitative studies (Huynen et al., 1996). For example, Gavrillets and Gravner address the problem of speciation by considering the percolation of random subgraphs within the sequence hypercube (Gavrillets and Gravner, 1997), assuming that “after a sufficiently long time, the population is equally likely to be at any of the points of the [giant] component” (Gavrillets, 1997). Similarly, Crutchfield and van Nimwegen attempt to link evolutionary dynamics with statistical mechanics via the “maximum entropy” assumption that “[infinite] populations [have] equal probabilities to be in any of the microscopic states consistent with a given [neutral network]” (Crutchfield and van Nimwegen, 2002). These and many other works treat neutral evolution as though its effect were to wash out concentration gradients in genotype space the same way particle diffusion washes out concentration gradients in liquids or gases.

This picture breaks down when many neutral variants co-exist within the same population, *i.e.* when the number of new mutants per generation  $M$  is much larger than one. It is well known that clonal interference can lead to the loss of beneficial mutations and limits the speed of adaptation in asexual populations (Gerrish and Lenski, 1998; Park and Krug, 2007). Here I show that *neutral interference*—the competition of mutants with equal adaptive value but possibly different robustness through negative selection—has a similar, but perhaps more counter-intuitive, effect on neutral evolution. Instead of allowing a population to explore its neutral network at a faster rate, increasing  $M$  can confine it to a small, highly connected region within that network. This effect is conceptually and mathematically similar to the Anderson conductor-insulator localization transition in disordered metals (Anderson, 1958) and can be captured with a different kind of random walk model—a “maximal entropy random walk” (Burda et al., 2009).

## Results

### Neutral evolution in a holey landscape

The simplest setting to study neutral evolution is a *holey fitness landscape* (Gavrilets, 1997; van Nimwegen et al., 1999), understood as a set of genotypes with binary fitness  $w$ : a genotype is either fully functional ( $w = 1$ ), or it is unviable ( $w = 0$ ). Representing possible mutations as edges turns this landscape into a graph. Generically, the subset of functional genotypes in a holey landscape splits in multiple connected components, one of which (the “giant component”  $G$ ) contains the majority of all genotypes. We denote  $A$  the adjacency matrix of  $G$  and  $d(x) = \sum_{y \in G} A_{xy}$  the neutral degree (number of neutral mutants) of a genotype  $x$ .

In the classical description of neutral evolution (Kimura, 1983), a functional mutant has a probability  $1/N$  to fix and replace the wild type through genetic drift. Under this substitution process the entire population performs a SRW within  $G$  (with  $N$ -independent jump rate), whence the concept of populations “diffusing in a neutral network” evoked earlier. However, when the number of new mutants  $M = \mu N \gg 1$ , as in e.g. RNA viruses (Drake and Holland, 1999), the selection of mutational robustness becomes more important than genetic drift as a driver of evolution (Schuster and Swetina, 1988; van Nimwegen et al., 1999; Forster et al., 2006; Sanjuán et al., 2007). The dynamics of the distribution of viable genotypes  $p_t(x)$  is then better described by a replicator-mutator (or “quasi-species”) equation of the form (van Nimwegen et al., 1999)

$$p_{t+1}(x) = \frac{1}{\langle w \rangle_t} \left( \mu \frac{A}{L} + (1 - \mu) I \right) p_t(x). \quad (1)$$

where  $L$  is the total number of possible mutants for each sequence and the population mean fitness  $\langle w \rangle_t = \mu \langle d \rangle_t / L + (1 - \mu)$  ensures that  $p_t(x)$  is normalized in  $G$ . From 1 it is easy to see that the equilibrium distribution (or mutation-selection balance)  $Q(x)$  is an eigenvector of the adjacency matrix of the neutral network  $A$ ; by the Perron-Frobenius theorem  $Q(x)$  is in fact the dominant eigenvector of  $A$  and the equilibrium mean neutral degree  $\langle d \rangle_\infty$  is its spectral radius  $\rho$ , both of which only depend on the topology of  $G$ . Since the eigenvector  $Q(x)$  tends to concentrate in regions with high connectivity, these facts establish the “neutral evolution of mutational robustness” (van Nimwegen et al., 1999).

### Equivalence with the maximal entropy random walk

To understand the limits of neutral evolution in interfering populations we must study the full trajectory  $p_t(x)$  and not just its asymptotic equilibrium  $Q(x)$ . For this purpose we can use the



framework recently developed in (Smerlak, 2019), which maps the non-linear dynamics 1 onto a Markov chain on  $G$  via the change of variables  $q_t(x) \propto Q(x)p_t(x)$  (SI ). This new distribution satisfies the master equation

$$q_{t+1}(x) = Rq_t(x) \quad \text{with} \quad R_{xy} = \frac{Q(x)(\mu A_{xy}/L + (1 - \mu)\delta_{xy})Q(y)^{-1}}{\mu\rho/L + (1 - \mu)}. \quad (2)$$

This equation describes a stochastic walk on  $G$  such that, at each time step, the walker either stays in her current location  $x$  or jumps to a neighboring node  $y$  in  $G$  with probability proportional to  $Q(y)/Q(x)$ . The sequence of actual transitions is then a *maximal entropy random walk* (MERW) on  $G_0$  (SI ).

The MERW was introduced as a Markov process with optimal mixing properties by Burda *et al.* (Burda et al., 2009) and has found applications in complex networks theory, image analysis and other fields. In a nutshell, while a simple random walker is “blind” (or “drunk”) and therefore chooses the next node to visit uniformly at random among nearest neighbors, a maximal-entropy random walker is “curious”: her transition probabilities are such that the each new step is as surprising as possible, *i.e.* the MERW maximizes the entropy *rate* of the process. Somehow paradoxically, the blind walker is sure to eventually visit all nodes of a (finite, connected) graph with finite probability, but the curious walker may not: in irregular networks, the equilibrium probability of the maximum-entropy random walker is exponentially suppressed outside a small “localization island” (Burda et al., 2009), as in Fig. S1. In the context of neutral evolution, this implies that narrow neutral ridges are much more difficult to traverse than is usually appreciated—percolation is not enough.

## Revisiting Maynard Smith’s four-letter model

To illustrate this localization effect I reconsidered Maynard Smith’s famous toy model of protein evolution (Smith, 1970). In the set of all possible four-letter words, Maynard Smith used meaning as a binary measure of fitness: any meaningful word is considered functional. He gave the sequence of one-point mutations  $\sigma = (\text{word}, \text{wore}, \text{gore}, \text{gone}, \text{gene})$  as an example of a neutral path, arguing that, unless a large section of genotype space can be traversed through such neutral paths, molecular evolution is impossible. It is easy to come up with other examples of neutral paths; in the following we will focus on  $\sigma' = (\text{opus}, \text{onus}, \text{anus}, \text{ants}, \text{arts})$ .

Using the Wolfram dataset of English “KnownWords” we find that, out of 2405 meaningful four-letter words, 2268 belong to the giant component  $G$ , including both paths  $\sigma$  and  $\sigma'$ . However, due to the irregular structure of  $G_0$  with three communities separated by narrow ridges (Fig. 2A), the majority of these words have negligible equilibrium probability: a core of just 420 (resp. 1064) words concentrates 90% (resp. 99%) of the total probability (Fig. S2). In particular, if all the

words in Maynard Smith’s path  $\sigma$  (except **gene**) belong to the 99%-core, none of the words in  $\sigma'$  do. Note that the words with the largest equilibrium probability are poorly predicted from their neutral degree, as illustrated in Fig. 2B: while **says** and **seed** both have relatively high neutrality ( $d = 21$ ), the former is ten thousand times more frequent than the latter. This highlights that the “neutral evolution of mutational robustness” is not simply the evolutionary advantage of robust genotypes—it is a selection principle which singles out, on a logarithmic scale, a subset of robust (e.g. **bays**) and non-robust (e.g. **whys**) genotypes that are globally well-connected in  $G$ . These patterns generalize to words of different lengths (Fig. S4 and Table S1).

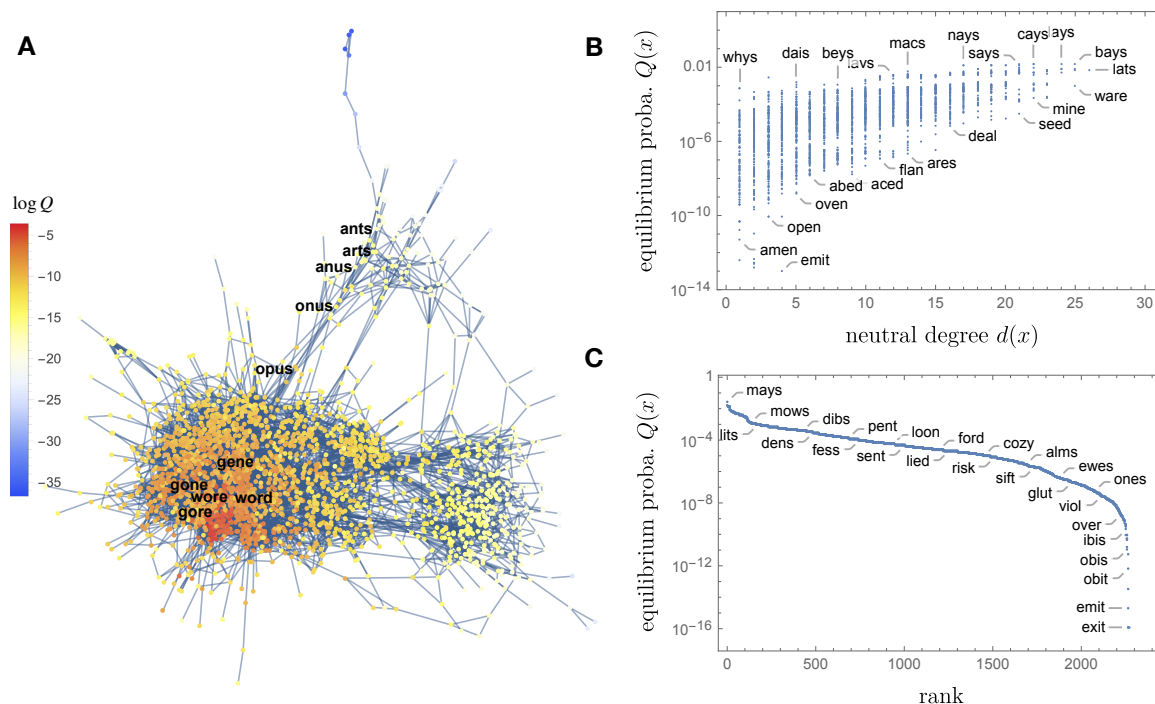


Figure 2: Localization in Maynard Smith’s holey landscape of meaningful four-letter words. A: The giant component with vertices colored by their logarithmic probability at mutation-selection balance, with the neutrals paths  $\sigma$  and  $\sigma'$  highlighted in boldface characters. While  $\sigma$  is deep in the core and easily evolvable,  $\sigma'$  belongs to a narrow ridge which can hardly be traversed. B: The evolutionary stability of four-letter words, measured by their probability at mutation-selection equilibrium, is correlated with their neutral degree (mutational robustness), but cannot be reliably inferred from it. C: The rank plot of  $Q(x)$  shows the approximately exponential decay of the equilibrium density away from the core typical of localization phenomena.

In multiple runs of a simple evolutionary simulation with a 10% mutation probability per genome per generation, population size  $N = 10^3$  and time horizon of  $10^4$  generations, **gene** almost always evolved from **word** (through Maynard Smith’s or, more often, some other path), but—because of the fragility of its intermediate forms—**arts** only rarely evolved from **opos** (Fig. 3A and S3).

When it did evolve, **arts** arose in the first few generations of the run, before the population got permanently trapped in the core. Strikingly, smaller (but not too small) populations had a *higher* probability to evolve **arts** at least once within the prescribed time horizon (Fig. 3B). This effect is due to the coupling between individuals through negative selective pressures and would be incomprehensible if we view neutral evolution as diffusion.

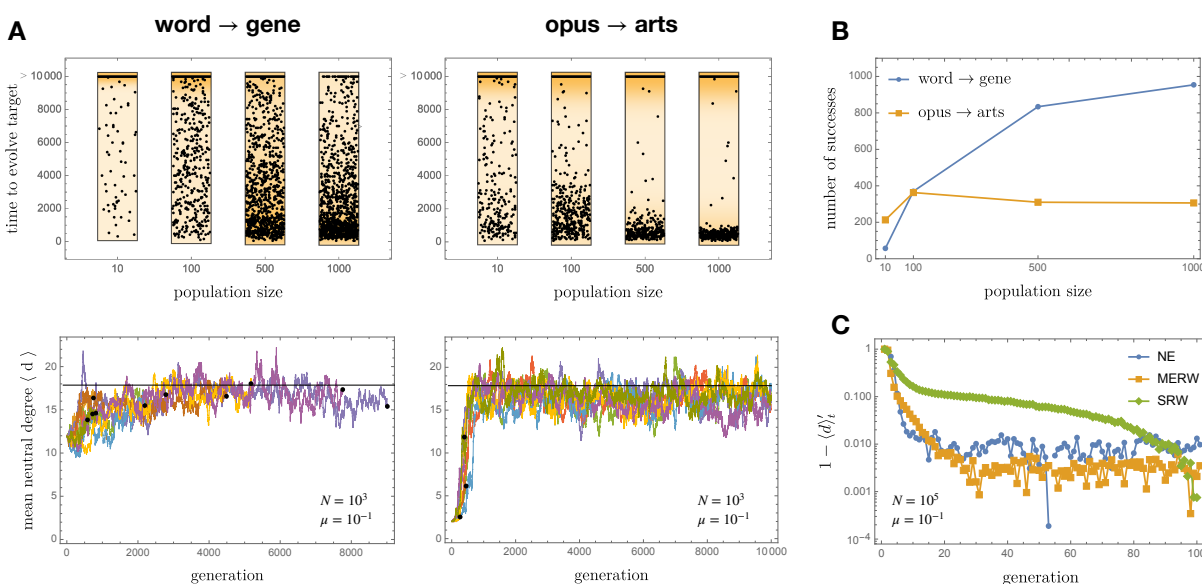


Figure 3: A: The evolution of arts from opus (left column) is a very different challenge than that of gene from word (right column), as is seen by comparing the time to evolve the target word at varying  $N$  (top row, with a cutoff after  $10^4$  generations) and the shape of  $\langle d \rangle_t$  trajectories at  $N = 10^3$  (bottom row, with the black horizontal line representing the equilibrium value  $\rho$  and black dots the endpoint of successful trajectories). While the likelihood of success of word → gene increases with the population size, it does *not* for opus → arts. If a population does not succeed to evolve arts in the first few generations of the run, before it can get trapped in the core, it never will. B: The likelihood to evolve a genotype from another through neutral evolution can depend non-monotonically on the population size. Here there were  $10^3$  attempts with  $\mu = 0.1$ , and a population with 100 individuals was more likely to evolve opus → arts at least once than one with 1000 individuals. C: The convergence to equilibrium, here measured by the rescaled mean degree  $\langle d \rangle'_t$ , is much faster in a neutrally evolving population (NE) than in diffusing one (SRW) of the same size; it is however consistent with the maximal mixing property of the MERW.

## RNA secondary structures

As a second illustrative example I reconsidered the neutral network of RNA secondary structures studied in (van Nimwegen et al., 1999). The secondary structure of an RNA molecule is the pattern of pairings between complementary bases along its sequence, and can be represented with brackets (paired bases) and dots (unpaired bases). Here we consider sequences of length  $\ell = 18$  with minimum-free-energy secondary structure “(((((((...))))))..))” and only purine-pyrimidine base pairs. With the RNA Vienna folding algorithm v2.4.8 (Hofacker, 2003), the giant component has size  $|G_0| = 17557$ , but its mutation-selection equilibrium distribution  $Q(x)$  varies over 7 orders of magnitude between core and periphery due to narrow ridges within  $G$  (Fig. 4); such ridges are also seen in the experimental assay of a small protein neutral network (Podgornaia and Laub, 2015), and may be a generic feature of biological genotype-to-phenotype mappings (Aguirre et al., 2011). Thus, in molecular evolution as well as in Maynard Smith’s toy model, neutral evolution may not efficiently sample neutral networks. Statements to the effect that “diffusion enables the search of vast areas in genotype space” (Huynen et al., 1996) must therefore be qualified accordingly.

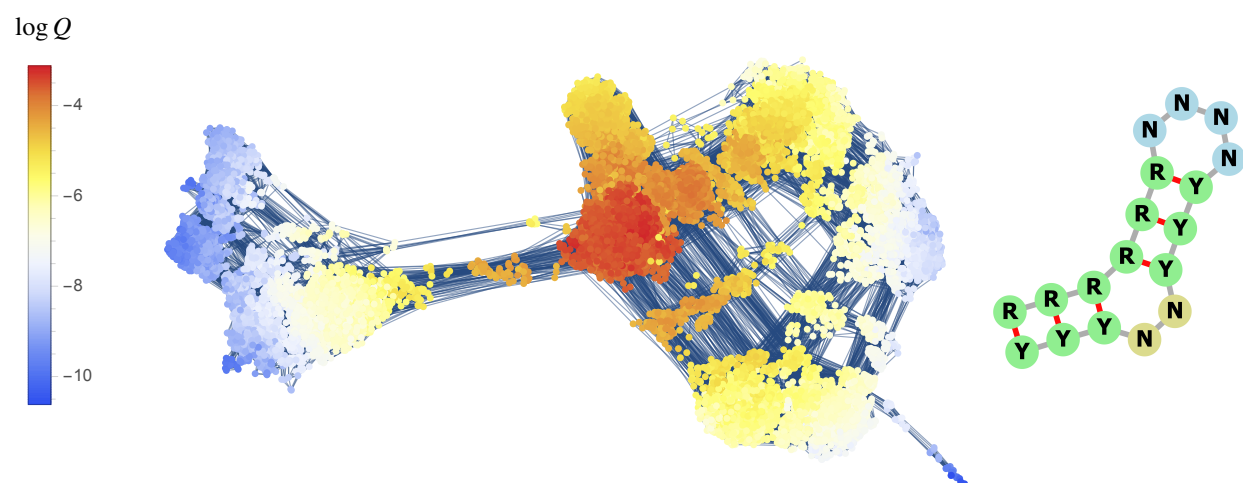


Figure 4: The neutral network of the RNA secondary structure “(((((((...))))))..))” with only purine (R) - pyrimidine (Y) base pairs (van Nimwegen et al., 1999), as predicted by the RNA Vienna algorithm v2.4.8. Its giant component consists of multiple communities separated by narrow ridges which induce the localization of the mutation-selection equilibrium  $Q$  and limits non-equilibrium dynamics accordingly. For this reason the notion that “diffusion enables the search of vast areas in genotype space” (Huynen et al., 1996) must be understood as holding within communities but not between them.

## Robustness and evolvability

The association of NE with the MERW rather than the SRW has notable implications for the classical issues of robustness and evolvability (Masel and Trotter, 2010). In the literature, the potential of (e.g. RNA) sequences to generate new phenotypes—their evolvability—has been related to the (linear or logarithmic) size  $|G_0|$  of their neutral network (Jörg et al., 2008), aka their versatility. The rationale for this hypothesis is that a large neutral network potentially communicates with many other neutral networks (other phenotypes) through “portal” sequences interfacing between them. But, just like the randomness of an information source should be quantified by its entropy and not merely its alphabet size (Shannon, 1948), the reach of neutral evolution should be quantified by its ability to generate new sequences, not by the number of all possible sequences. Information theory provides the correct language for this: in the  $M \gg 1$  regime, the versatility of a phenotype should be measured by the entropy  $H(Q) = -\sum_{x \in G} Q(x) \log Q(x)$  of its mutation-selection equilibrium, or equivalently by the effective size  $|G|_{\text{eff}} = 2^{H(Q)}$  of the giant component. In both cases considered above this quantity is much smaller than the naive value  $|G|$ : with four-letter words we have  $|G|_{\text{eff}} = 346.6 \ll 2268$ , and for the RNA secondary structure  $|G|_{\text{eff}} = 3971.8 \ll 17557$ .

## Conclusion

I have described the localization of populations within neutral networks induced by the competition for mutational robustness among neutral mutants. This “neutral interference” is interference in the two senses of the word: in the biological sense, because it involves the competition of clonal subpopulations, an effect usually referred to as clonal interference; and in the physical sense, because localization is an interference phenomenon normally encountered in (classical or quantum) wave mechanics. The link between these two seemingly different process—the evolution of molecular populations and the destructive interference of waves—is provided by the MERW, a Markov chain whose local transition probabilities depend on the global structure of the underlying graph.

These observations reveal sharper constraints on the navigability of neutral networks than previously appreciated: not only does evolution favor high mutational robustness and low genetic loads, it also positively refuses to engage in tightrope walking along narrow neutral ridges. This effect manifests itself in the surprising non-monotonic dependence of discovery rates of certain target genotypes on population size and implies that hopes to “infer the complete structure of the neutral network from accurate measurements of the transient population dynamics” (van Nimwegen et al., 1999) are unfounded. In this way we can only hope to infer the structure of the robust cores of neutral networks. In molecular fitness landscapes, walking a narrow ridge is not *a priori* easier than crossing a fitness valley.

## Acknowledgements

I thank the members of the Structure of Evolution group at MPI MiS, and especially A. Zadorin and I. H. Hatton, for useful critical comments. Funding for this work was provided by the Alexander von Humboldt Foundation in the framework of the Sofja Kovalevskaja Award endowed by the German Federal Ministry of Education and Research.

## Supplementary Information

### Selection-mutation dynamics as a Markov process

In (Smerlak, 2019) I showed that continuous-time replicator-mutator (or quasispecies) equations can be understood in terms of a derived Markov process, in which the logarithm of the selection-mutation equilibrium plays the role of an effective potential. With discrete generations, this scheme can be reformulated as follows. Given a discrete space  $X$ , consider a sequence of probability distributions  $p_t : X \rightarrow \mathbb{R}$  evolving under the dynamics

$$p_{t+1}(x) = \frac{Bp_t(x)}{\sum_y Bp_t(y)} \quad (3)$$

with  $B$  any irreducible non-negative matrix. (Here I identify a function on  $X$  with the vector of its values.) In replicator-mutator systems we have  $B = MW$  with  $W$  a diagonal matrix of Wrightian fitnesses and  $M$  a stochastic matrix of mutation probabilities. The process 3 is not Markovian due to the global interactions introduced by the normalization factor; for this reason it is difficult to interpret it in terms of “evolutionary trajectories”.

This can be remedied by means of the change of variable  $q_t(x) \propto S(x)p_t(x)$ , where  $S$  is the (left) eigenvector of  $B$  with largest eigenvalue; by the Perron-Frobenius theorem this vector is positive and its eigenvalue is the spectral radius  $\sigma$  of  $S$ . Via this transformation—which only depends on  $t$  through a global constant—we obtain a Markovian representation of the original dynamical problem with master equation

$$q_{t+1}(x) = \tilde{B}q_t(x) \quad \text{with} \quad \tilde{B}_{xy} = \frac{S(x)B_{xy}S(y)^{-1}}{\sigma}. \quad (4)$$

For this derived process the function  $U = -2 \log S$  plays the role of a potential; its analysis reveals the metastable states and preferred trajectories of the original (non-linear) process (Smerlak, 2019). The equilibrium distribution for 4 is  $q_\infty(x) \propto S^2(x)$ , and the original distribution can be reconstructed from  $q_t(x)$  with  $p_t(x) \propto S(x)^{-1}q_t(x)$ . When  $B$  happens to be the adjacency matrix  $A$  of a connected graph  $G$ , as in neutral evolution, the process 4 coincides with the MERW on  $G$ .



## The maximal entropy random walk

The SRW on a connected graph  $G$  with adjacency matrix  $A$  and degree matrix  $D = \text{diag}(d(x))_{x \in G}$  is the Markov chain with transition matrix

$$R_{xy} = \text{Prob}(x|y) = \frac{A_{xy}}{d(y)}. \quad (5)$$

By contrast, the MERW was defined in (Burda et al., 2009) as the Markov chain with transition matrix

$$\text{Prob}(x|y) = \frac{Q(x)A_{xy}Q(y)^{-1}}{\rho} \quad (6)$$

where  $Q$  is the dominant eigenvector of  $A$  with eigenvalue  $\rho$  (equal to the spectral radius of  $A$ ). This Markov chain maximizes the entropy rate among ergodic stationary processes on  $G$  by assigning equal probability to all paths connecting two given nodes. Rather paradoxically, this property of “maximal mixing” also leads to dynamical localization when the graph  $G$  is irregular, as in Fig. S1.

## Attraction to a clique in the SRW and MERW

The qualitative difference between the SRW and the MERW can be understood by considering the probabilities  $p_{\text{in,out}}$  for a walker to jump into (resp. out of) a fully connected subgraph with  $n$  nodes (an  $n$ -clique) along a ridge, as in Fig. 1. Since there are  $n-1$  edges going in and 1 edge going out, for the SRW these probabilities are simply  $p_{\text{in}}^{\text{SRW}} = (n-1)/n$  and  $p_{\text{out}}^{\text{SRW}} = 1/n$ , hence  $p_{\text{in}}^{\text{SRW}}/p_{\text{out}}^{\text{SRW}} = n$ .

To evaluate the same ratio for the MERW, we must compute the dominant eigenvector  $Q$  of the adjacency matrix  $A$  for the complete graph over  $n$  nodes  $\{1, \dots, n\}$  with one extra node, labelled 0, attached to vertex 1. By symmetry this eigenvector  $Q = (q_0, q_1, \dots, q_n)$  can be chosen such that  $q_2 = \dots = q_n = 1$ . Writing  $AQ = \rho Q$  then gives a system of two quadratic equations for  $(q_0, q_1)$ , from which we then obtain  $p_{\text{in}}^{\text{MERW}}/p_{\text{in}}^{\text{MERW}} = (n-1)/q_0$ . Because  $q_0 \sim 1/n$  when  $n \gg 1$ , this gives  $p_{\text{in}}^{\text{MERW}}/p_{\text{in}}^{\text{MERW}} \sim n^2$  as claimed in Fig. 1.

## Evolutionary simulations

The simulations displayed in Fig. 2 and S3 were performed with the following algorithm, which takes as input a connected graph  $G$ , an integer  $\delta \geq \max_{x \in G} d(x)$  representing the total number of possible (neutral or lethal) mutants a genotype can have, a population size  $N$ , a mutation probability  $\mu$ , a time horizon  $T$ , an initial genotype  $x_0 \in G$  and a target genotype  $x_1 \in G$ . Here we



denote for each  $x \in G$   $\nu(x)$  the set of its  $d(x)$  neighbors in  $G$ , and let  $\bar{\nu}(x)$  be the set obtained by adjoining  $\delta - d(x)$  copies of the symbol  $\dagger$  (meaning “dead”) to  $\nu(x)$ .

First, generate an initial population  $P_0$  consisting of  $N$  copies of  $x_0$ . Next, for each generation  $t$ , perform the two following steps:

- **mutation:** draw a number  $n \sim \text{Binom}(N, \mu)$  and replace  $n$  randomly chosen individuals  $x$  in  $P_t$  by random samples from the corresponding  $\bar{\nu}(x)$ ; these samples together with the elements of  $P_t$  not chosen for mutation form the mutated population  $P'_t$
- **selection:** sample with replacement  $N$  elements from  $P'_t \cap G$  and collect them in a new population  $P_{t+1}$

The algorithm terminates whenever  $x_1 \in P_t$  (the target genotype is found) or  $t = T$  (the time horizon is reached).

## Supplementary Figures

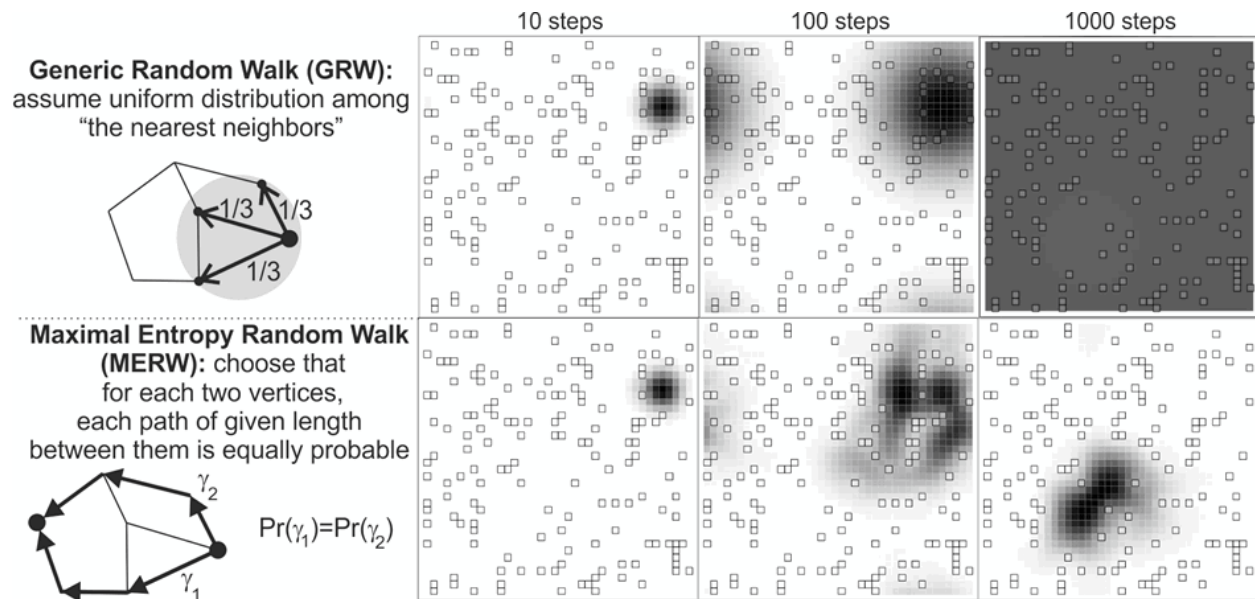


Figure S1: Illustration of the difference between the simple (or generic) random walk (top) and the maximal entropy random walk (bottom) in a two-dimensional lattice with defects. The SRW choose elementary steps uniformly at random among nearest neighbors and eventually fills the lattice uniformly (top); the MERW chooses paths between fixed nodes uniformly at random and ends up trapped in a defect-poor "localization island" (bottom). Illustration by Jarek Duda (CC BY-SA 4.0).

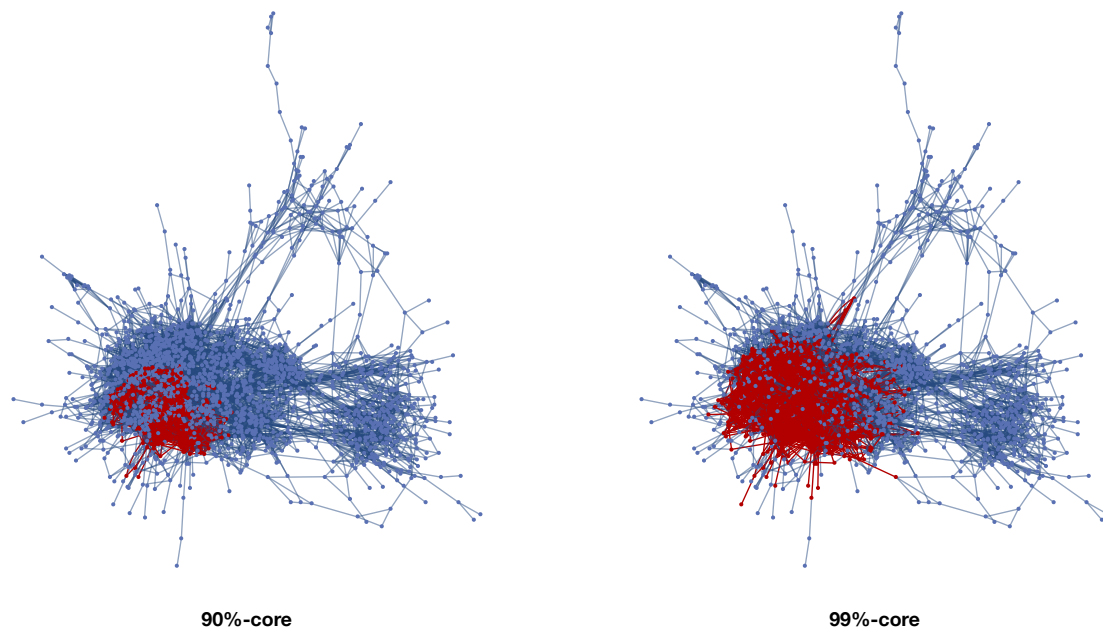


Figure S2: The 90%- and 99%-cores within the giant component of meaningful four-letter English words, with size 420 and 1024 respectively (out of 2268 words in the giant component and 2405 in total). These cores correspond to regions with higher mutational robustness ([van Nimwegen et al., 1999](#)).

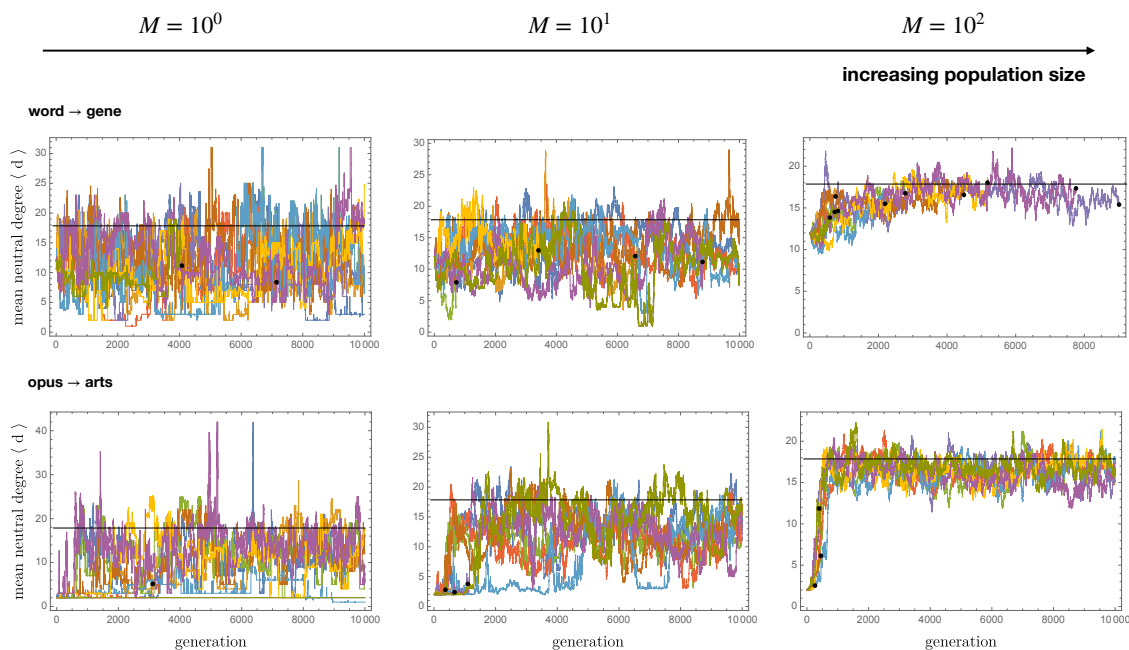


Figure S3: Sample evolutionary trajectories for the **word** → **gene** (top) and **opus** → **arts** (bottom) problems, with success shown as a black dot. When just a few mutants co-exist in the population ( $M = 1$ ), neutral evolution is insensitive to mutational robustness and behaves as a SRW, blindly exploring the whole neutral network. As  $M$  increases, the attraction towards the core of the giant component (where  $\langle d \rangle = \rho$ , indicated by the black horizontal line) becomes stronger and the likelihood to walk a tightrope such as  $\sigma'$  decreases accordingly.

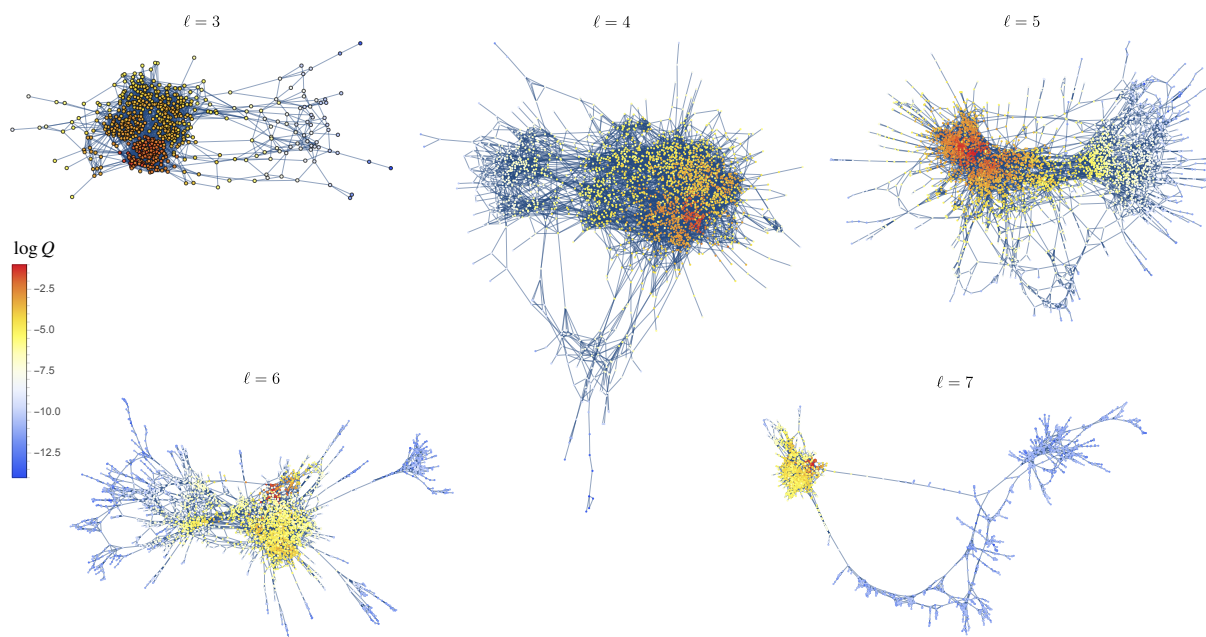


Figure S4: Giant components in the holey landscapes of meaningful English words of different lengths  $\ell$ , with words  $x$  colored by their logarithmic mutation-selection equilibrium probability  $Q(x)$ . In all cases the latter displays exponential localization in a robust core.

# Supplementary Table

| Word length $l$ | Number of meaningful words | Giant component size —G— | Giant component effective size $ G _{\text{eff}}$ | Sizes ratio $ G _{\text{eff}}/ G $ |
|-----------------|----------------------------|--------------------------|---------------------------------------------------|------------------------------------|
| 3               | 621                        | 603                      | 265.4                                             | 0.97                               |
| 4               | 2403                       | 2268                     | 346.6                                             | 0.94                               |
| 5               | 4753                       | 3598                     | 588.9                                             | 0.76                               |
| 6               | 7763                       | 3353                     | 64.93                                             | 0.43                               |
| 7               | 10926                      | 2081                     | 42.67                                             | 0.19                               |

Table S1: Holey landscapes of meaningful English words of different lengths, see Fig. S4.

# References

- Jacobo Aguirre, Javier M. Buldú, Michael Stich, and Susanna C. Manrubia. Topological Structure of the Space of Phenotypes: The Case of RNA Neutral Networks. *PLoS ONE*, 6(10):e26324, oct 2011. doi: 10.1371/journal.pone.0026324. URL <https://doi.org/10.1371%2Fjournal.pone.0026324>.
- P. W. Anderson. Absence of Diffusion in Certain Random Lattices. *Physical Review*, 109(5):1492–1505, mar 1958. doi: 10.1103/physrev.109.1492. URL <https://doi.org/10.1103%2Fphysrev.109.1492>.
- Aderonke Babajide, Ivo L Hofacker, Manfred J Sippl, and Peter F Stadler. Neutral networks in protein space: a computational study based on knowledge-based potentials of mean force. *Folding and Design*, 2(5):261–269, oct 1997. doi: 10.1016/s1359-0278(97)00037-0. URL <https://doi.org/10.1016%2Fs1359-0278%2897%2900037-0>.
- Z. Burda, J. Duda, J. M. Luck, and B. Waclaw. Localization of the Maximal Entropy Random Walk. *Physical Review Letters*, 102(16), apr 2009. doi: 10.1103/physrevlett.102.160602. URL <https://doi.org/10.1103%2Fphysrevlett.102.160602>.
- Stefano Ciliberti, Olivier C. Martin, and Andreas Wagner. Robustness Can Evolve Gradually in Complex Regulatory Gene Networks with Varying Topology. *PLoS Computational Biology*, 3(2):e15, 2007. doi: 10.1371/journal.pcbi.0030015. URL <https://doi.org/10.1371%2Fjournal.pcbi.0030015>.
- James P. Crutchfield and Erik van Nimwegen. The Evolutionary Unfolding of Complexity. In *Natural Computing Series*, pages 67–94. Springer Berlin Heidelberg, 2002. doi: 10.1007/978-3-642-55606-7\_4. URL [https://doi.org/10.1007%2F978-3-642-55606-7\\_4](https://doi.org/10.1007%2F978-3-642-55606-7_4).
- J. W. Drake and J. J. Holland. Mutation rates among RNA viruses. *Proceedings of the Na-*

- tional Academy of Sciences*, 96(24):13910–13913, nov 1999. doi: 10.1073/pnas.96.24.13910. URL <https://doi.org/10.1073%2Fpnas.96.24.13910>.
- Adam Eyre-Walker. The genomic rate of adaptive evolution. *Trends in Ecology & Evolution*, 21(10):569–575, oct 2006. doi: 10.1016/j.tree.2006.06.015. URL <https://doi.org/10.1016%2Fj.tree.2006.06.015>.
- Walter Fontana, Peter F. Stadler, Erich G. Bornberg-Bauer, Thomas Griesmacher, Ivo L. Hofacker, Manfred Tacker, Pedro Tarazona, Edward D. Weinberger, and Peter Schuster. RNA folding and combinatorial landscapes. *Physical Review E*, 47(3):2083–2099, mar 1993. doi: 10.1103/physreve.47.2083. URL <https://doi.org/10.1103%2Fphysreve.47.2083>.
- Robert Forster, Christoph Adami, and Claus O. Wilke. Selection for mutational robustness in finite populations. *Journal of Theoretical Biology*, 243(2):181–190, nov 2006. doi: 10.1016/j.jtbi.2006.06.020. URL <https://doi.org/10.1016%2Fj.jtbi.2006.06.020>.
- Sergey Gavrilets. Evolution and speciation on holey adaptive landscapes. *Trends in Ecology & Evolution*, 12(8):307–312, aug 1997. doi: 10.1016/s0169-5347(97)01098-7. URL <https://doi.org/10.1016%2Fs0169-5347%2897%2901098-7>.
- Sergey Gavrilets and Janko Gravner. Percolation on the Fitness Hypercube and the Evolution of Reproductive Isolation. *Journal of Theoretical Biology*, 184(1):51–64, jan 1997. doi: 10.1006/jtbi.1996.0242. URL <https://doi.org/10.1006%2Fjtbi.1996.0242>.
- Philip J. Gerrish and Richard E. Lenski. The fate of competing beneficial mutations in an asexual population. In *Mutation and Evolution*, pages 127–144. Springer Netherlands, 1998. doi: 10.1007/978-94-011-5210-5\_12. URL [https://doi.org/10.1007%2F978-94-011-5210-5\\_12](https://doi.org/10.1007%2F978-94-011-5210-5_12).
- Greg Gibson and Ian Dworkin. Uncovering cryptic genetic variation. *Nature Reviews Genetics*, 5(9):681–690, sep 2004. doi: 10.1038/nrg1426. URL <https://doi.org/10.1038%2Fnrg1426>.
- I. L. Hofacker. Vienna RNA secondary structure server. *Nucleic Acids Research*, 31(13):3429–3431, jul 2003. doi: 10.1093/nar/gkg599. URL <https://doi.org/10.1093%2Fnar%2Fgkg599>.
- M. A. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: the role of neutrality in adaptation. *Proceedings of the National Academy of Sciences*, 93(1):397–401, jan 1996. doi: 10.1073/pnas.93.1.397. URL <https://doi.org/10.1073%2Fpnas.93.1.397>.
- Thomas Jörg, Olivier C Martin, and Andreas Wagner. Neutral network sizes of biological RNA molecules can be computed and are not atypically small. *BMC Bioinformatics*, 9(1):464, 2008. doi: 10.1186/1471-2105-9-464. URL <https://doi.org/10.1186%2F1471-2105-9-464>.
- Motoo Kimura. Evolutionary Rate at the Molecular Level. *Nature*, 217(5129):624–626, feb 1968. doi: 10.1038/217624a0. URL <https://doi.org/10.1038%2F217624a0>.

- Motoo Kimura. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, 1983. doi: 10.1017/cbo9780511623486. URL <https://doi.org/10.1017%2Fcbo9780511623486>.
- J. L. King and T. H. Jukes. Non-Darwinian Evolution. *Science*, 164(3881):788–798, may 1969. doi: 10.1126/science.164.3881.788. URL <https://doi.org/10.1126%2Fscience.164.3881.788>.
- Joanna Masel and Meredith V. Trotter. Robustness and Evolvability. *Trends in Genetics*, 26(9): 406–414, sep 2010. doi: 10.1016/j.tig.2010.06.002. URL <https://doi.org/10.1016%2Fj.tig.2010.06.002>.
- Masatoshi Nei, Yoshiyuki Suzuki, and Masafumi Nozawa. The Neutral Theory of Molecular Evolution in the Genomic Era. *Annual Review of Genomics and Human Genetics*, 11(1):265–289, sep 2010. doi: 10.1146/annurev-genom-082908-150129. URL <https://doi.org/10.1146%2Fannurev-genom-082908-150129>.
- S.-C. Park and J. Krug. Clonal interference in large populations. *Proceedings of the National Academy of Sciences*, 104(46):18135–18140, nov 2007. doi: 10.1073/pnas.0705778104. URL <https://doi.org/10.1073%2Fpnas.0705778104>.
- Anna I. Podgornaia and Michael T. Laub. Pervasive degeneracy and epistasis in a protein-protein interface. *Science*, 347(6222):673–677, feb 2015. doi: 10.1126/science.1257360. URL <https://doi.org/10.1126%2Fscience.1257360>.
- Rafael Sanjuán, José M. Cuevas, Victoria Furió, Edward C. Holmes, and Andrés Moya. Selection for Robustness in Mutagenized RNA Viruses. *PLoS Genetics*, 3(6):e93, 2007. doi: 10.1371/journal.pgen.0030093. URL <https://doi.org/10.1371%2Fjournal.pgen.0030093>.
- Peter Schuster and Jörg Swetina. Stationary mutant distributions and evolutionary optimization. *Bulletin of Mathematical Biology*, 50(6):635–660, nov 1988. doi: 10.1007/bf02460094. URL <https://doi.org/10.1007%2Fbf02460094>.
- C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3): 379–423, jul 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x. URL <https://doi.org/10.1002%2Fj.1538-7305.1948.tb01338.x>.
- Matteo Smerlak. Effective potential reveals evolutionary trajectories in complex fitness landscapes. *bioRxiv:869883*, 2019. doi: <https://doi.org/10.1101/869883>.
- John Maynard Smith. Natural Selection and the Concept of a Protein Space. *Nature*, 225(5232): 563–564, feb 1970. doi: 10.1038/225563a0. URL <https://doi.org/10.1038%2F225563a0>.
- E. van Nimwegen, J. P. Crutchfield, and M. Huynen. Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences*, 96(17):9716–9720, aug 1999. doi: 10.1073/pnas.96.17.9716. URL <https://doi.org/10.1073%2Fpnas.96.17.9716>.



Andreas Wagner. Neutralism and selectionism: a network-based reconciliation. *Nature Reviews Genetics*, 9(12):965–974, dec 2008. doi: 10.1038/nrg2473. URL <https://doi.org/10.1038%2Fnrg2473>.

Andreas Wagner. *Arrival of the Fittest: Solving Evolution's Greatest Puzzle*. Current, 2014.