On the significance of neutral spaces in adaptive evolution

Steffen Schaper
Exeter College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2012

To my parents

On the significance of neutral spaces in adaptive evolution
Submitted for the degree *Doctor of Philosophy*
Trinity 2012
Steffen Schaper
Exeter College

Abstract

Evolutionary dynamics arise from the interplay of mutation and selection. Fundamentally, these two processes operate at different levels: Mutations modify genetic information (the genotype), which is passed from parent to offspring. Selection is triggered by variation in reproductive success, which depends on the physical properties (the phenotype) of an organism and its environment. Thus the genotype-phenotype map determines if and how mutations can lead to selection. The aim of this dissertation is to incorporate this map explicitly into a theoretical description of evolutionary dynamics.

The first part of the analysis presented here is concerned with the static properties of simple models of these maps, which are studied using exhaustive enumeration. The two most important observations are phenotypic bias – some phenotypes are realized by many more genotypes than most other phenotypes – and the existence of neutral spaces – genotypes with the same phenotype can often be reached from each other by single mutational steps.

The remainder of the dissertation provides a theoretical description of evolutionary dynamics on and across neutral spaces. Two different mean-field approximations lead to simple analytic results for the first discovery of alternative phenotypes, highlighting the importance of phenotypic bias: Rare phenotypes are hard to find by evolutionary search. These results are used to discuss the relationship of robustness, the ability to withstand mutational change, and evolvability, the ability to create variation through mutation.

Several types of fluctuations beyond the mean-field limit are studied, both theoretically and in simulations. The discrete structure of genotype spaces can lead to strong correlations in the spectra of phenotypes produced, increasing the probability that a particular phenotype is fixed in the population quickly after its discovery. Structural correlations between genotypes can increase the effect of phenotypic bias, while the qualitative features of the mean-field description remain valid. All these results highlight that neutral spaces impact evolutionary dynamics in many non-trivial ways, in particular by favouring phenotypes of high accessibly, but comparably low fitness over those phenotypes that are highly fit, but very hard to discover.

Acknowledgements

This dissertation marks the result of three years of thought and work, and I am immensely grateful to all who made this ride a time I most thoroughly enjoyed.

I wish to express my sincere gratitude to Ard Louis, who has been a great supervisor, giving me freedom when I wanted, guidance when I needed, and insightful input throughout this time. Among my fellow DPhil students, I especially thank Kamal Dingle, Iain Johnston, Jake Scott, Christian Matek and Petr Šulc for all their feedback and encouragement and the many hours of discussions from which I have greatly benefited.
For financial support, I am grateful to EPSRC, Oxford Theoretical Physics and Exeter College.

This work presents my first larger foray beyond (a small stretch of) the edge of current knowledge. I am indebted to the many people which helped get close to this edge, especially my Physics tutors Andrew Steane and Frank Close, who first and foremost taught me to think before starting to calculate.

I am deeply grateful to my family and friends, without whom I would have been lost on this journey. This thesis is dedicated to my parents Ulli and Erika Schaper, who will always have my gratitude for their support in more ways than I could imagine. I thank my grandmother Else Kania and my late grandfather Rudolf Kania for all the encouragement and support they have constantly provided. My brothers Sebastian and Philip have been great sources of fun and motivation. Without Inge and Dieter Donner, I would not have set foot to the streets of Oxford and would have missed the great time I had here.

Last but certainly not least, I thank Laura Martin, who supported me in so many ways that I cannot possibly recount them all: From bearing with (and without) me through the ups and downs of this work, to being the shining example of self-discipline that kept me going, and providing me with the best laughs when they were least expected but most dearly needed – I dare not wonder how I would have arrived at the end of this journey without her.

Contents

Chapter 1

Introduction

> Natural selection may explain the survival of the fittest, but it cannot explain the arrival of the fittest.
> Hugo de Vries, *Species and Varieties – Their Origin by Mutation*, 1904 [29]

Just over 150 years ago, Charles Darwin put forth the theory of biological evolution [27]. His account explains how populations of living organisms change over generations: If the reproductive success, or *fitness*, is inherited from parent to offspring, and if there is variation in fitness within a population of individuals, then the lineages of those individuals which reproduce most will come to dominate the population. Competition between individuals in the population – for example due to limited resources – leads to *natural selection* and the survival of the fittest.

Selection alone cannot sustain evolution, and would indeed bring evolution to a standstill: When only the fittest individuals survive, variation in the population declines and finally disappears. The potential for further evolution is lost as natural selection is predicated on the existence of variation. In this thesis, we shall therefore largely be concerned with the following question: Where does variation, the fuel of natural selection, come from?

We will attack this question by means of simple toy models of evolutionary dynamics. In order to decide what these models should include, we cover background material in Chapter 2. After reviewing some of the classic models of evolutionary dynamics, we will meet the key object of our study, the *genotype-phenotype* (GP) map [4]. The GP map determines

how genetic information (the genotype) is expressed in physical properties (phenotype) of an organism. Therefore, the mapping provides the link between mutations, which introduce genotypic variation, and selection, which operates on differences in the fitness of phenotypes. The literature on theoretical descriptions of evolution, in particular through the framework of population genetics [41, 60, 62, 74, 143], is voluminous, but only a very small fraction takes into account the GP map explicitly.

For the purposes of this thesis, the most important property of GP maps is that they are many-to-one: Many different genotypes give rise to the same phenotype. What is more, these genotypes can often be inter-converted by single mutations. This proximity gives rise to the idea of *neutral spaces* or *neutral networks* [85, 115] which comprise all the genotypes of a particular phenotype. The existence of neutral networks, which has been demonstrated both computationally and experimentally in many different systems [125, 128], has profound consequences for the evolutionary process: In contrast to the classical picture of rugged *fitness landscapes* [145], in which selection often drives populations to local fitness optima from which further adaptation requires the crossing of fitness barriers [144], neutral networks provide flat ridges through the landscape along which populations can accumulate genotypic variation without any fitness penalties [51].

The landscape metaphor (in physics, we typically think of energy landscapes rather fitness landscapes) has for a long time been a powerful tool for the understanding and prediction of dynamical systems in many branches of science [46, 116, 117]. The existence of neutral networks leads us to study a special kind of dynamics that can be described in a two-step picture: First, classical adaptation corresponds to the transition from one network to another whose phenotype bestows greater fitness on the population. Second, there are phases of *neutral exploration*, during which the population moves within a neutral network in the search of a fitter phenotype, as illustrated in Fig. 1.1. Understanding evolutionary dynamics in terms of these processes is the main focus of this dissertation.

We conclude Chapter 2 by a review of the literature concerned with neutral spaces and neutral exploration. These descriptions are often cast in terms of *mutational robustness* and *evolvability*. Robustness refers to the ability of populations to withstand genetic change, and

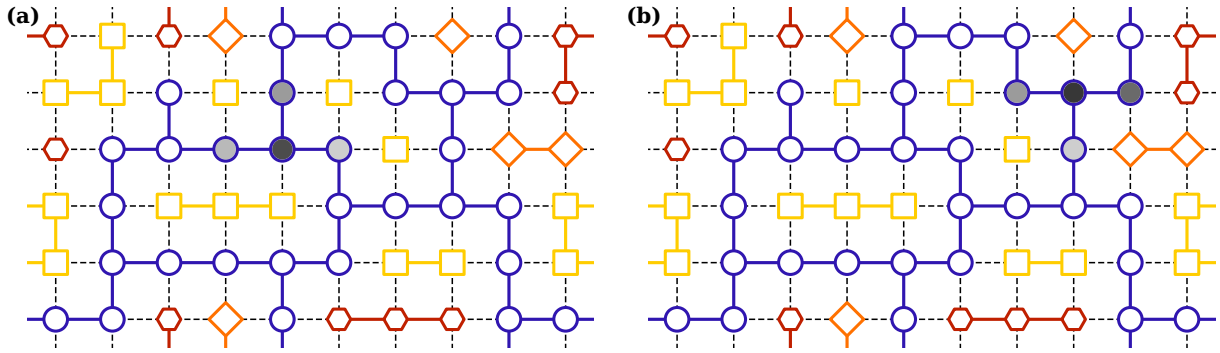

Figure 1.1: An illustration of neutral exploration. The pictures illustrate the exploration of a neutral space by an evolving population. Each node corresponds to a genotype, and the shapes and colours indicate the different phenotypes. Edges correspond to mutations that change one genotype into another: Neutral mutations preserve the phenotype and are drawn in the appropriate colour; non-neutral mutations are indicated by dashed black lines. The filling of each node shows how many individuals carry that particular genotype. Between **(a)** and **(b)**, the population has moved across the neutral space. The main theme of this thesis will be to understand evolutionary dynamics in terms of this simple picture.

is directly related to the availability of neutral mutations [125]. Evolvability refers to the ability to bring about phenotypic variation [4, 6, 28, 99, 127]. Intuitively, it appears that robustness should reduce evolvability: After all, a mutation can either be neutral (a sign of robustness) or not (contributing to evolvability) [127]. However, we will see throughout this thesis that the relation of robustness and evolvability is much more subtle and complex: Since robustness allows populations to explore a wide variety of genotypes, these populations may also be able to sample a greater variety of alternative phenotypes than is accessible from a single genotype [127]. Depending on different definitions of evolvability, its relation to robustness has been claimed to be positive [127], negative [24] and even non-monotonic [33], so that an intermediate degree of robustness would optimize evolvability.

In order to make sense of these results, we study the static properties of one of the most important model GP maps, namely the folding of RNA sequences [64], in Chapter 3. A striking findings in RNA (and also in other systems [32]) is a strong *phenotypic bias* [13]: The neutral networks of some phenotypes are larger by many orders of magnitude than those of most other phenotypes. So if we pick a genotype at random, the probability of obtaining one of the most frequent phenotypes is much greater than the probability to find a rare phenotype.

We then zoom in on the structure of neutral networks and show that the biophysics of RNA folding cause the fragmentation of neutral networks into many disjoint components. We explicitly confirm Wagner's result [127] that larger components endow populations with greater robustness and also have more alternative phenotypes in their mutational neighbourhood. As the potential for further evolution depends on the phenotypes that can be accessed from a population's neutral space, we compare the spectra of accessible phenotypes among these different components. We find a striking heterogeneity: The possible innovations are very sensitive to the current location of the population in genotype space, suggesting historical contingency in evolution [114].

In Chapter 4, we use these observations of the static properties of GP maps to define a microscopic toy-model of neutral exploration. The key ingredient into the model is phenotypic bias, and we show in a mean-field description how this bias influences the course of evolution: When a population explores a neutral network, frequent phenotypes (that is, phenotypes whose neutral networks are large) are produced earlier and more often than rare ones. The analytic expressions show a simple dependence of both the discovery time and the number of discoveries on phenotype frequency. We validate these results by simulations on toy landscapes without genotypic correlations; the qualitative agreement of simulations and theory is excellent, but for quantitative accuracy, we need to introduce a single fitting parameter to account for genetic drift. The limiting behaviour in the regimes of weak and strong drift of this factor can be predicted from our mean-field approach.

Chapter 5 proposes a dynamic measure of evolvability that counts the number of phenotypes actually produced by an evolving population. Using the model of Chapter 4, we obtain a slightly unwieldy expression for population evolvability in terms of population dynamic parameters and the details of the GP map. To obtain a more intuitive result, we approach population evolvability from the angle of phenotypic bias, building on the idea of diminishing returns. Thus we find that phenotypic bias can cause population evolvability to increase only logarithmically with search time, in contrast to a model without bias which predicts a linear increase.

When modelling neutral exploration, we assume that all alternative phenotypes are

strongly deleterious, so that adaptation can be ignored. We lift this assumption in Chapter 6, where we consider the fate of mutants with a beneficial phenotype. First, we show that the dynamics of neutral exploration at low mutation rates can lead to time-correlations in the phenotypes carried by non-neutral mutants. These correlations can be attributed purely to the discrete structure of neutral spaces, and we show that these correlations make it quite likely that a beneficial phenotype is fixed quickly after its first discovery. Importantly, we find that this probability of fast fixation is negatively related to the mutational robustness of the population. This result underlines the necessity to take neutral spaces into account when we think about evolutionary dynamics.

Finally, Chapter 7 connects our study of RNA neutral spaces and the analytic toy model of evolutionary dynamics. In contrast to the random GP map, neutral spaces in RNA show strong heterogeneity. Nonetheless, the toy model provides a good first-order explanation of the dynamics of neutral exploration in RNA: Rare phenotypes remain hard to find, and the majority of non-neutral mutants produces the small set of frequent phenotypes. We conclude our study of evolution in RNA with the question of adaptation in a fitness landscape with two peaks. One peak corresponds to a frequent phenotype of relatively low fitness, while the other peak has higher fitness, but is harder to discover due to its lower frequency. Comparing our simulations to a study that ignored neutral spaces [146], we find that the effect of phenotypic bias is increased by the need for neutral exploration to discover rare phenotypes.

Chapter 8 summarizes our results and puts them in a common perspective. We argue that the existence of neutral spaces may in principle simplify the introduction of novel phenotypes; in practice however, phenotypic bias leads to severe limitations to discover rare phenotypes quickly. Consequently, it may not be possible to resolve the apparent tension of robustness and evolvability in general. Neutral spaces and phenotypic bias are just two aspects that must be taken into account when reasoning about evolution, and this complexity alone provides a richness of different behaviours that cannot be captured by a simple 'yes' or 'no' answer to the question of robustness and evolvability. Similarly, the evident necessity to discover a fit phenotypes may place a limit on the degree of adaptation that evolving populations may achieve when the desired phenotypes are rare.

Chapter 2

Background

2.1 Heuristic evolutionary dynamics

Evolution is a dynamic process at the population level [92]: Individuals are born, mature, reproduce and eventually die. Rather than studying the life history of particular individuals, we are interested in the changes of the population. This change emerges from three principles: Inheritance, competition and variation. In this section, we develop simple theoretical models that highlight the basic ideas of evolutionary dynamics, loosely following Nowak's exposition [92].

2.1.1 Deterministic description

Inheritance ensures that genetic information is maintained even though the organisms carrying it are not. We account for inheritance in our model by categorizing individuals into *types*, where each type has a particular fitness. For the moment, we assume that the time to reproduction is distributed exponentially with average τ, and take fitness as $f = 1/\tau$: Fitter organisms reproduce faster. If we only have a single type and no competition, then the number of individuals in the population n is simply described by

$$\frac{dn}{dt} = fn \tag{2.1}$$

which yields an unbounded exponential population growth: $n(t) = n_0 \exp(ft)$ where $n_0 = n(t = 0)$ is the initial population size. The only steady state of this system ($dn/dt = 0$) is $n = 0$: Populations do not suddenly appear from nowhere. But this fixpoint is not stable ($d^2n/dt^2 = f > 0$): Even a single individual is sufficient to found an infinite population.

Of course, such an infinite growth is unrealistic. We can try to do better by imposing a death rate d (corresponding to the inverse average lifetime) such that $dn/dt = fn - dn$; in this model, we either get extinction ($d > f$) or again an ever-expanding population ($f > d$). When $d > f$, the fixpoint at $n = 0$ now becomes stable, so that the population will always go extinct.

If we want to limit the population to a finite size N, we can impose a logistic rule:

$$\frac{dn}{dt} = fn\left(1 - \frac{n}{N}\right) = fn - \frac{fn}{N}n \tag{2.2}$$

Effectively, we have introduced a size-dependent death rate $d(n) = fn/N$. In ecology, N is often called the carrying capacity and reflects the finite resources that are available to the population.

The dynamics described by Eq. (2.2) have a stable fixpoint at $n^* = N$ (we use the asterix throughout this section to denote fixpoints): Because of competition (!) for finite resources, the population size is limited. N is thus a natural scale for the population size and we define $x = n/N$ as the fraction of individuals in the population that the environment could support. We have

$$\frac{dx}{dt} = fx(1 - x) \tag{2.3}$$

with solution

$$x(t) = \frac{x_0 \exp(ft)}{1 + x_0(\exp(ft) - 1)} \tag{2.4}$$

where $x_0 = x(t = 0)$ is the initial population size, normalized by the carrying capacity.

Let us now add variation to the model. For simplicity, let us consider two alternative types a and A with fitness f_a, f_A respectively. We will usually think of a as the 'wild-type', which dominates the population at $t = 0$, and of A as some 'mutant' type in whose

fate we are interested. Each individual in the population belongs to one of these types, so $n(t) = n_a(t) + n_A(t)$. In the absence of competition, we have one equation like (2.1) for each type. These uncoupled equations indicate exponential growth for each type, which of course is still unrealistic.

Let us now impose a finite total population size, $n_a + n_A \leq N$, by imposing a death rate $\bar{f}$. We shall assume that this rate is the same for individuals both types – fitness affects only reproduction, but not individual survival. Thus we have

$$\frac{dn_i}{dt} = \Big(f_i - \bar{f}\Big)n_i \tag{2.5}$$

and we can enforce a fixpoint at $n_a + n_A = N$ by requiring

$$\sum_i \frac{dn_i}{dt} = \sum_i \Big(f_i - \bar{f}\Big)n_i = 0 \tag{2.6}$$

so that

$$\bar{f} = \frac{\sum_i f_i n_i}{N} \tag{2.7}$$

At the fixpoint, the average death rate $\bar{f}^*$ equals the average growth rate of the population.

Let us now study the composition of the population. For simplicity, we assume that the total population size is equal to the carrying capacity N and hence constant. We can thus write $x_i = n_i/N$ for the relative abundance of type i and we have $x_a + x_A = 1$ so that Eq. (2.5) can be written as

$$\frac{dx_A}{dt} = x_A(f_A - (f_A x_A + f_a(1 - x_A))) = (f_A - f_a)x_A(1 - x_A) \tag{2.8}$$

which is just Eq. (2.2) with the fitness replaced by $f_A - f_a$. For convenience, we shall write $f_A = f_a(1 + s)$ where s is the *selection coefficient* of A. If $s > 0$, the stable fix point of (2.8) is $x_A^* = 1$ and $x_a^* = 0$, while $s < 0$ yields $x_A^* = 0, x_a^* = 1$ as the stable steady state. This result is the *survival of the fittest*: The type of highest fitness is *fixed* in the population and the other type is driven to extinction. This result generalizes straightforwardly to a larger

number of types.

To summarize, we have just seen that inheritance, competition and variation lead to natural selection and the survival of the fittest. But we also observe that natural selection alone destroys variation: When only the fittest type survives, there is no variation left and evolutionary dynamics reach a steady state. In particular, our current model cannot explain how the individuals with the mutant type arise in the first place.

In order to solve this problem, we need to add another ingredient to the model: mutations. So far, we have assumed that inheritance has a perfect fidelity: The type of the parent completely determines the type of the offspring. Let us now allow for the possibility that parent and offspring differ. In our simple model with only two types, let us assume that mutations occur with probability m per reproduction and change one type into the other. For simplicity, we shall also assume that the population size is constant (that is, given by the carrying capacity), which allows us to write $x_a = 1 - x_A$. Then we have

$$\frac{dx_A}{dt} = (1-m) f_A x_A + m f_a (1 - x_A) - \bar{f} x_A \tag{2.9}$$

with $\bar{f} = x_a f_a + f_A x_A$ as before. The first term describes the increase of A-individuals through A-parents whose offspring does not mutate, the second term corresponds to A-mutant offspring from a-parents, and the third term ensures that the total population size remains constant.

For convenience, we again set $f_A = f_a(1+s)$ and renormalize time such that $f_a = 1$ to obtain

$$\frac{dx_A}{dt} = -s x_A^2 + (s(1-m) - 2m) x_A + m \tag{2.10}$$

By setting $dx_A/dt = 0$, we can solve for the fixpoints:

$$x_A^* = \frac{1-m}{2} - \frac{m}{s} \pm \sqrt{\left(\frac{1-m}{2} - \frac{m}{s}\right)^2 + \frac{m}{s}} \tag{2.11}$$

In the strong selection limit $m \ll s$, we expand to lowest order in m/s and obtain for a beneficial mutant type (that is, $s > 0$) the steady state frequency $x_A^* \approx (1-m)(1-m/s)$,

while if the mutant type A is deleterious ($s < 0$), we find $x_A^* \approx (1-m)m/|s|$.

At the other extreme, we have the limit of strong mutations $m \gg s$, when there is only one fixpoint with $x_A^* = 1/2 + s(1-m)/4m$ (which is stable): A strong mutational 'pressure' prevents a high degree of adaptation in the population.

More generally, we see that neither $x_A = 0$ nor $x_A = 1$ remain steady-states: Mutations lead to a co-existence between types of different fitness, and hence mutations preserve a certain amount of variation. This co-existence is also called mutation-selection balance [62, 92].

For later use, let us generalize these results for a bigger number of types and also allow for different probabilities of mutations between types. We write the state of the population as the vector $\mathbf{x}(t)$, whose components $x_i(t)$ correspond to the number of individuals of the different types i, normalized by the carrying capacity. We introduce the selection matrix F which has the fitness values of the types on the diagonal: $F_{ij} = f_i \delta_{ij}$ where δ_{ij} is the Kronecker delta. Mutations are encoded in the mutation matrix m whose element m_{ij} is the probability that a j-parent gives rise to an i-offspring; in particular, we have m_{ii} for the probability that no mutation occurs, and hence the normalization $\sum_j m_{ji} = 1$. The dynamics of the population are given by the *quasispecies equation* [35, 36, 37, 38]:

$$\frac{d\mathbf{x}}{dt} = Fm\mathbf{x} - \bar{f}\mathbf{x} \tag{2.12}$$

where $\bar{f} = \sum_i f_i x_i$ is the average fitness of the population, as before.

The steady state $\mathbf{x}^*$ of the population is determined by the eigenvalue equation[1]

$$Fm\mathbf{x}^* = \bar{f}\mathbf{x}^* \tag{2.13}$$

The leading eigenvector of the mutation-selection matrix (Fm) gives the stable distribution of the population; this distribution over the different types is also called the *quasispecies*

[1]Technically, $\bar{f}$ depends on $\mathbf{x}^*$, leading to a non-linear equation. However, we can always recover the desired $\bar{f}$ by normalizing $\mathbf{x}^*$ appropriately. Indeed, the proper normalization yields probabilities, as we need $\sum_i x_i^* = 1$ [92].

[92].

In summary, we have seen in a very simple model how evolutionary dynamics can arise from the interplay of mutations, which create variation, and natural selection acting on this variation. The quasispecies equation (Eq. (2.12), [35]) incorporates mutation and selection into a relatively simple deterministic framework, and is a very popular model in the literature [92, 123, 138, 139]. As we shall see in Section 2.3.1, quasispecies treatments are a valuable tool to generate insight into evolutionary dynamics. However, the deterministic description neglects fluctuations: In particular, treating the $x_i(t)$ as continuous variables (so that it makes sense to think of their derivatives) ignores the fact that individuals are in-dividual: In a finite population, the x_i can only take a discrete set of values. So the quasispecies model is applicable when populations are large and mutation rates are high [138]; in the next section, we shall turn to a classic model of population genetics which aims to incorporate the effects of fluctuations explicitly.

2.1.2 Stochasticity and genetic drift

So far, we have modeled evolutionary dynamics using deterministic differential equations. We have thus neglected the stochastic fluctuations of reproduction which arise from the fact that we can only have integer numbers of individuals (so that the x_i cannot vary continuously). We shall now introduce one of the classic models of population genetics that does account for these stochastic effects.

The Wright-Fisher model [41, 143] considers a population of constant, finite size N, evolving in discrete generations. At each time step, N parents are drawn from the population with replacement; whenever an individual is selected to be parent, it produces one offspring. After N offspring have been produced, the parents are removed and the offspring form the new population. The fitness of an individual determines its probability of reproduction: If individual α has fitness f_α, its selection probability p_α per reproduction is simply proportional to fitness, $p_\alpha = f_\alpha / \sum_\beta f_\beta$ where the sum includes all individuals in the population.

Several results for the Wright-Fisher model can be derived analytically using a diffusion

approximation [72, 74]. We shall not review the technicalities here, but instead focus on the results that are most important to our later developments. For simplicity, we will again consider a population of only two types a and A. Individuals of the wild-type a have fitness $f_a = 1$, while the mutant type A imparts fitness $f_A = f_a(1 + s) = 1 + s$. We do not allow for mutations between the types.

Most importantly, and in contrast to the the deterministic treatment, the survival of the fittest is no longer guaranteed even in the absence of mutations. If we have n individuals of type A in the population, the probability that eventually the entire population will be of type A is [71, 74]

$$P_n(s) = \frac{1 - \exp(-2ns)}{1 - \exp(-2Ns)} \tag{2.14}$$

The probability $P_1(s)$ that a single A-individual will take over the population is called the *fixation probability* of type A. We have

$$P_1(s) = \frac{1 - \exp(-2s)}{1 - \exp(-2Ns)} \tag{2.15}$$

Intuitively, $P_1(s)$ increases with s, but for any finite selective advantage, fixation is not guaranteed even in the limit $N \to \infty$, where $P_1(s) \to 1 - \exp(-2s) \approx 2s$. Due to the stochasticity of reproduction, beneficial mutants may be lost. This stochastic change in the composition of the population is called *genetic drift*.

The appearance of Ns in the fixation probability indicates a natural scale of measuring the selection coefficient s. If $Ns \ll 1$, we have

$$P_1(s) \approx \frac{1 - (1 - 2s + 2s^2)}{1 - (1 - 2Ns + 2N^2s^2)} = \frac{1 - s}{N(1 - Ns)} \tag{2.16}$$

In particular, we have the fixation probability of a neutral mutant, $P_1(s = 0) = 1/N$. This result is quite intuitive: When no individual has a selective advantage, each one has the same chance of becoming the common ancestor of the population. We also see that mutants with $Ns \ll 1$ are *effectively neutral* [93]: The (inverse) population size sets a natural scale against fitness benefits must be measured. In contrast to the deterministic approximation,

the fate of a mutant is one of two qualitatively different outcomes: It is 'destined' either to be fixed or to be lost.

In summary, the stochastic treatment of the Wright-Fisher model leads to qualitative differences from the deterministic models of large populations. While different stochastic models (such as the Moran model [87]) give different results for the fixation probability [96], the main result is consistent: The stochasticity of reproduction in finite population implies that beneficial mutants can be lost and deleterious mutants can be fixed. The survival of the fittest is no longer guaranteed; instead, genetic drift can cause the loss of beneficial mutants. Similarly, in finite populations the maintenance of standing neutral variation requires mutations.

Mutations are of paramount importance in producing the variation that fuels natural selection. In the next section, we consider the molecular mechanism of inheritance and mutations. We will use this background to later construct a more microscopic model of evolutionary dynamics.

2.2 Inheritance and the origin of variation

2.2.1 DNA - the molecular basis

Variation arises when offspring are different from their parents. The origin of variation is thus closely connected to the biophysical mechanism of inheritance. The experiments of Gregor Mendel indicated already in Darwin's time that inheritance was based on discrete entities (*genes*) [86], but these insights went largely unnoticed for some 40 years. In the 1920s and early 1930s, the work of Fisher [41], Haldane [60] and Wright [143, 145] founded the field of population genetics which brought together Mendelian genetics and Darwinian selection.

Population genetics laid the basis for a quantitative understanding of evolutionary dynamics. The genotype is described in terms of *genes*. Each gene sits at a particular *locus* in the genotype and can exist in different *alleles* – for example, a hypothetical gene deter-

mining eye-colour could exist in a 'blue' or 'brown' allele. The models we discussed above correspond to the simplest class which consider only two alleles a and A at a single locus.

During reproduction, one allele may be replaced by another through a *mutation*. In principle, mutations thus lead to variation and hence to natural selection. Maybe the most powerful conceptual tool for the study of evolution, developed in the context of population genetics, is provided by Wright's metaphor of *fitness landscapes* [145]. In this metaphor, fitness is pictured as the height of a mountain range over the space of all possible genotypes. Peaks of high fitness are separated by valleys of lower fitness. In Wright's shifting balance theory, populations evolve from one peak to another of even higher fitness by crossing such valleys [144]. It is important to note that Wright proposed the landscape metaphor well before the microscopic mechanism of inheritance were established.

In 1953, Watson and Crick solved the structure of DNA [134], ending their paper with the now famous remark 'It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material'. Their discovery finally allowed a microscopic understanding of the process of inheritance: Genetic information is stored in sequences of the four nucleotide bases adenine (A), cytosine (C), guanine (G) and thymine (T), which are arranged in double-helical DNA. When cells divide, the DNA strands are copied and distributed between mother and daughter cell.

Microscopically then, mutations are 'copying errors'[2] of various kinds: The simplest mutation is a *substitution*, whereby a single nucleotide is replaced by a different one. Nucleotides can also be lost (*deletion*) or added (*insertion*), but these changes are not limited to single bases. Indeed, entire genes and chromosomes can be lost, and the duplication of the entire genome may be one of the most important events in the evolution of higher organisms [70]. In general, mutations are relatively rare: In most species for which data is available, the rate of mutations is on the order of one per genome per generation [34]; only some viruses appear to have slightly elevated rates of a few mutations per genome per generation.

Another source of genetic variation occurs in sexually reproducing organisms, where

[2]The word error has a negative connotation which is not appropriate here. If it weren't for mutations, evolution could not proceed!

offspring receive genetic material from both parents. This recombination of genetic material can lead to new genotypes, provided the genotypes of the parents differ. While we shall focus on asexual organisms throughout this thesis, we will discuss recombination and cross-over mutations briefly in Section 3.4.

Mutations change genotypes and thus are the ultimate cause of selectable variation. But the extent to which mutations lead to selectable changes in phenotypes is a much more subtle question. In order to address this question, we need explicit models of mutational effects on genotypes and a way to translate genotypic change into phenotypic change.

2.2.2 Genotype space

The discovery of the structure of DNA and its role in the transmission of hereditary information called for an incorporation of these microscopic details into evolutionary theory, leading to the field of molecular evolution [54, 89]. For our purposes, the most important ingredient into our models is the idea of a genotype space in which all possible genotypes are arranged [85]. In this space, mutations induce a notion of distance: Two genotypes are close to each other and hence similar if only a few mutations are required to convert one genotype into the other. This leads to the idea of the mutational landscape [53].

To make these notions more precise, we will model genotypes as strings of length L where each letter is drawn from an alphabet of size K (for DNA, $K = 4$). Throughout most of this thesis, we will assume that mutations are rare and restrict our attention to only the simplest type of mutations, namely single base substitutions (also called point mutations). We can thus arrange all genotypes in an L-dimensional, generalized hypercube [108] in which two genotypes are nearest neighbours when they differ by a single mutation (see Fig. 2.1 for an illustration). In this space, the distance between two genotypes is given by their Hamming distance [61], which simply counts the number of bases at which the genotypes differ.

Due to the discrete nature of genotype space, it is useful to use a few concepts from network theory (Newman's book [90] gives a more complete introduction). Briefly, networks consists of *nodes*, or vertices, and *edges* that join these nodes. So for our purposes, nodes

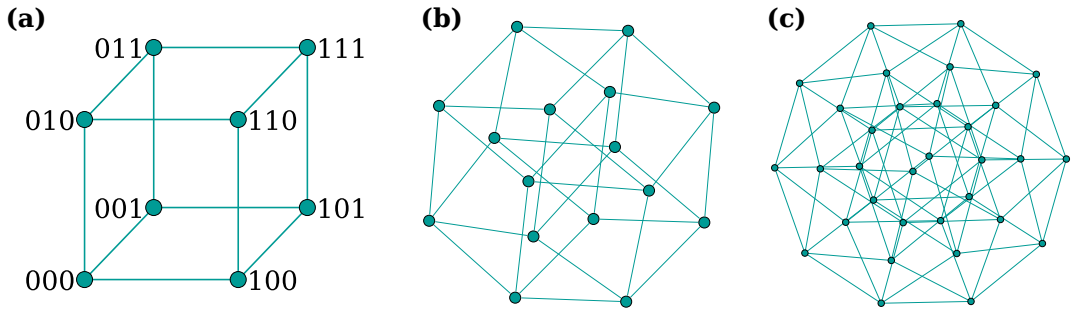

Figure 2.1: Genotypes are embedded on hypercubes. The diagram shows two-dimensional projections of genotype space for binary sequences (so the alphabet size is $K = 2$). Each vertex corresponds to a particular genotype, given as a binary sequence in **(a)**. Edges join genotypes that differ only by a single substution, a bit flip $0 \leftrightarrow 1$ in this case. The pictures correspond to **(a)** $L = 3$, **(b)** $L = 4$ and **(c)** $L = 5$.

correspond to genotypes, and edges connect genotypes that can be converted into each other by a single mutation. The structure of a network with n nodes (labelled by indices $i = 1, \ldots, n$) can be conveniently described in terms of the *adjacency matrix* $\mathcal{A}$, which is an $n \times n$ matrix with elements

$$\mathcal{A}_{ij} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

The *degree* k_i of node i gives the number of edges incident on it, or in our case the number of one-mutational neighbour genotypes:

$$k_i = \sum_j \mathcal{A}_{ji} \tag{2.18}$$

Since $\mathcal{A}$ is symmetric in the case of genotype space (because mutations are reversible), we need not distinguish between the sum over rows or columns.

Genotype spaces are vast networks: The number of genotypes is K^L. Under point mutations, each genotype has $(K-1)L$ *neighbours* which differ by a single point mutation, so $k_i = (K-1)L$ for all genotypes. As we shall see throughout this thesis, the high dimensionality of genotype space has profound consequences for evolution, since it determines the geometry underlying the fitness landscape [145]. One counter-intuitive feature is that the maximum distance between any two genotypes grows only logartihmically with the size of

the space, as it evidently never takes more than L mutational steps to convert any genotype into any other.

The most profound consequence of a high-dimensional genotype space, however, is that there need not necessarily be valleys separating genotypes of high fitness [21, 22, 51, 52, 108]. For example, if we assume that fitness is distributed randomly across genotypes, it is very unlikely that all $(K-1)L$ genotypes surrounding a potential peak do indeed have a lower fitness value [51]. Instead, we should expect 'extra-dimensional bypasses' [22] to provide ridge-like connections between high-fitness genotypes.

So because of the high-dimensionality of genotype space, the fitness landscape metaphor must be taken with care [100]. Another difficulty is that fitness derives from the phenotype, not the genotype. So the extent to which a mutation leads to a change in fitness is mediated by the effect that the mutation has on the phenotype. In order to understand the effect of mutations, we therefore need to understand the genotype-phenotype map.

2.2.3 The genotype-phenotype map and neutral spaces

In order to put to use the genetic information encoded in the genotype, this information must be expressed physically, giving rise to the phenotype. The relation of genotype and phenotype can be described in terms of the genotype-phenotype (GP) map [4]. In principle, the GP map takes as input a DNA sequence and environmental parameters (such as temperature or the availability of nutrients) and outputs the physical properties of the organism. In practice, predicting phenotypes from sequence is far beyond our current knowledge and computational capabilities; while we are getting better and better at reading genome sequences [11, 78, 137], we continue to have severe difficulty making sense of them [18, 55].

Consequently, we need to consider simplified models of GP maps to gain insight into their role in biology and evolution. One of the most fundamental GP map models describes how genetic information in DNA sequences is translated into protein molecules which then perform biological functions. Proteins are macromolecules consisting of polypeptide chains. Just as we can view DNA sequences just in terms of the 4 nucleotide letters, we can describe

protein sequences as a string of amino acids, of which there are 20 different types [98]. The segment of DNA that encodes the sequence of one particular protein is what we usually mean by a gene.

A protein molecule is created from its gene through the biochemical processes of transcription and translation. For our present purposes, we shall conceptually treat them as a black box and only consider their outcome. For a unique association of gene and protein, codons comprising 3 nucleotides are mapped into one amino acid. The genetic code, drawn in Fig. 2.2a, describes the assignment of codons to amino acids that is observed in almost all living organisms [91].

We can view the genetic code as a simple example of a GP map in which genotypes are codons and phenotypes are amino acids[3]. As there are $4^3 = 64$ codons, but only 20 amino acids (plus a stop signal that marks the end of the gene), the GP map is necessarily many-to-one (Fig. 2.2b). We note that the distribution of codons to amino acids is not quite uniform: Three amino acids – Arginine (Arg), Leucine (Leu), and Serine (Ser) – are associated with six codons each, while only a single codon maps into Methionine (Met).

What is more, the codons of the same amino acid are often connected by single mutations, such as GTA, GTC, GTG and GTT which all map into Valine (Val). Mutations between these codons are called synonymous in the case of the genetic code; more generally, we call a mutation *neutral* if it does not lead to a change in phenotype. At the molecular level, the existence of neutral mutations appears quite natural – it is not hard to imagine that a small change in the composition in a large molecule would go largely unnoticed.

The idea of neutral mutations at the molecular level is the corner-stone of Kimura's neutral theory of molecular evolution [74]. Kimura postulated that the vast majority of mutations leaves no selectable trace and that the course of evolution is largely determined by the random fixation of neutral mutations, that is genetic drift. One of the most important results of the neutral theory is the idea of a *molecular clock* [73]: Suppose neutral mutations arise at slow rate μ_n per generation, so that the total number of neutral mutants is small,

[3]In many studies of proteins, this mapping is completely ignored and instead protein sequences are viewed as genotypes that determing the shape of folded proteins and their function. Fundamentally however, information about proteins is transmitted genetically in DNA.

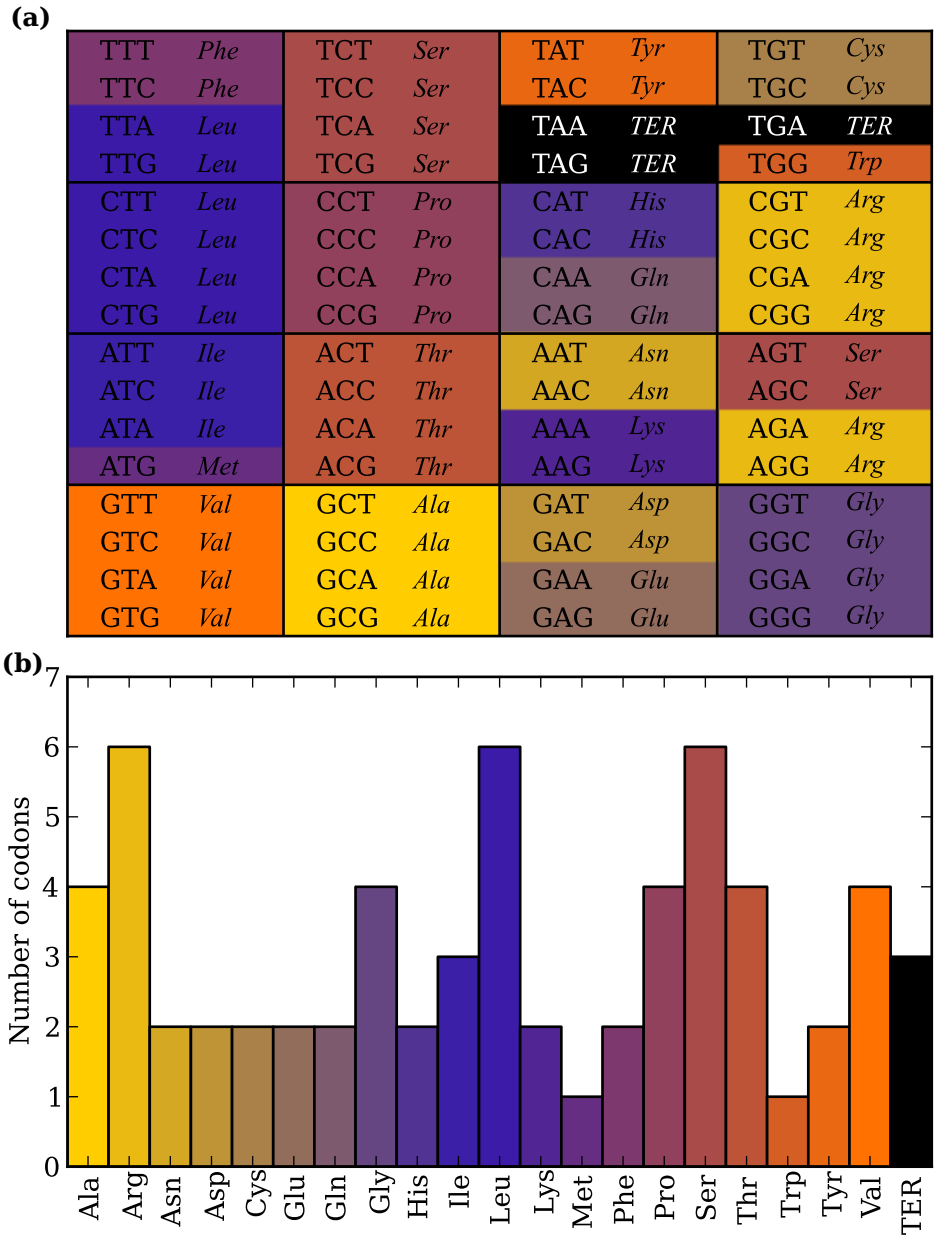

Figure 2.2: The Genetic Code is a many-to-one GP map. The genetic code describes how triplets of DNA nucleotides (codons) are translated into the 20 amino acids. Panel (a) shows the assignment of codons to amino acids. In (b), the number of codons mapping into each amino acid is shown. TER is the termination signal, indicating the end of translation.

so $N\mu_n \ll 1$ (as above, N is the population size). For neutral mutations ($s = 0$), we can see from Eq. (2.15) that their fixation probability is just $1/N$. Consequently, the rate at which neutral mutations go to fixation is simply $N\mu_n \times 1/N = \mu_n$, that is the rate of neutral mutations, independent of population size. As mutation rates are believed to be relatively constant in time (unlike population size, which can undergo strong variations) [74], the neutral theory suggests that we can measure time simply by counting mutations: If we find m differences between two species today, we can assume that they had a common ancestor $m/(2\mu_n)$ generatios ago (the factor of 2 accounts for the fact that mutations have accumulated in both species).

The neutral theory was extended by Ohta to *effectively neutral* mutations [93], which have fitness effects $s \ll 1/N$ (cf. Eq. (2.16)). Throughout this thesis, we appeal to this notion of effective neutrality when thinking about the relation of phenotypes and fitness: Ultimately, natural selection arises from differences in fitness, and it is not clear a priori which aspects of a phenotype are relevant to the fitness of an organism. We assume that genotypes mapping into the same phenotype are effectively neutral with respect to one another, while mutations that change phenotypes are never effectively neutral. In other words, we assume that the map from phenotype to fitness is one-to-effectively-one. We can turn this argument around and say that a particular system provides a useful level of abstraction if the variants of phenotypes in the system have appreciably different fitness.

This notion of phenotypic neutrality and the many-to-one nature of the GP map lead to the idea of *neutral spaces* or *neutral networks* [115]: When we view genotype space as a network, the GP map can be pictured as a colouring of the nodes, where colours correspond to phenotypes. All nodes (genotypes) of the same colour (phenotype) form a neutral space; edges between nodes of the same colour indicate neutral mutations (see Fig. 2.3 for an illustration). Technically, the neutral space of phenotype p is an induced subgraph on genotype space [108, 110]: All the genotypes mapping into p are the nodes of the neutral space, and its edges are all edges that connect genotypes with the same phenotype.

Apart from the genetic code, neutral spaces have been identified in many other GP maps, from similarly microscopic systems such as the folding of RNA molecules [56, 107, 115]

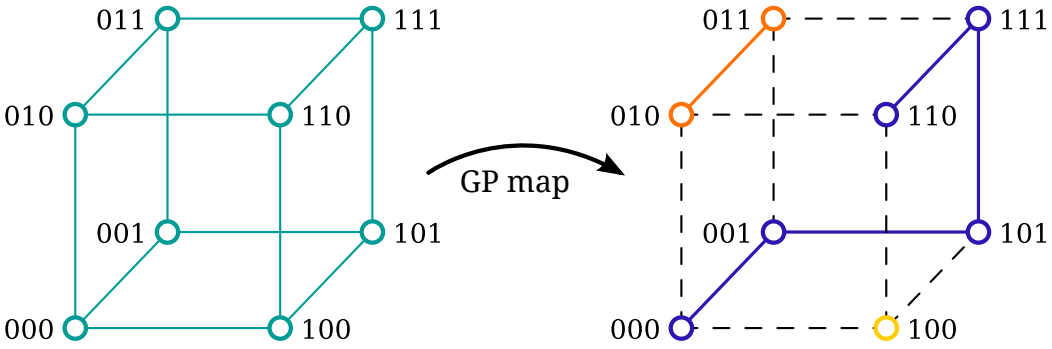

Figure 2.3: Many-to-one GP maps induce neutral spaces. The GP map can be pictured as a colouring of the genotype hypercube, so that colours represent phenotypes. Genotypes with the same phenotype form a neutral space. Note how the 'extra-dimensional bypass' along 001, 101 and 111 neutrally connects 000 and 110.

(we shall discuss RNA in detail in Chapter 3) and model proteins [14], up to cellular and organism-scale systems as gene regulatory networks [12, 19, 20] and metabolic networks [113]. Unfortunately, the nomenclature in the literature is not always consistent. Throughout the thesis, we shall borrow terminology from network theory. Therefore, each phenotype has one *neutral network* that comprises all genotypes mapping into that phenotype. A *neutral component* is a maximal subset of the neutral network in which all genotypes are connected by neutral mutations. When we do not make this technical distinction, we shall also use the more generic term *neutral space*.

The existence of (at least partially connected) neutral spaces leads to the notion of *mutational robustness*: When a mutation from genotype g_1 to g_2 preserves the phenotype, we say that g_1 is robust with respect to that mutation. In our picture of the genotype hypercube, this mutation corresponds to an edge that is part of the neutral space (in the right panel of Fig. 2.3, neutral mutations share the colour of the phenotype they preserve). Mutational robustness ρ_g of a genotype g is commonly defined as the probability that a mutation applied to g is neutral [123, 125, 130]. Mathematically, this probability is given by g's degree in the neutral space $k_g^{(n)}$ (that is, the number of *neutral neighbours* of g) divided by its degree on the genotype hypercube:

$$\rho_g = \frac{k_g^{(n)}}{(K-1)L} \tag{2.19}$$

By averaging over all genotypes with a particular phenotype p, we thus obtain the (average) robustness ρ_p of the phenotype:

$$\rho_p = \frac{\sum_{g \in \mathcal{G}_p} \rho_g}{|\mathcal{G}_p|} \tag{2.20}$$

where $\mathcal{G}_p$ is the set of all genotypes that map into p, that is the set of nodes of p's neutral network.

Kimura's molecular clock relies on an absolute notion of neutrality, in the sense that a mutation is either always neutral or it always leaves a selectable effect. One problem with this idea is that whether a mutation triggers selection may depend on population size, and population size may vary over time. However, there is a possibly even more important objection to the notion of absolute mutational effects: The effect of a mutation could depend on the genetic background on which it arises. The interdependence of mutational effects is called *epistasis* [9, 97], and as we shall see, epistasis has important consequences for the structure of neutral spaces and the process of neutral evolution.

2.2.4 Epistasis

Mendel's interpretation of his experiments [86] lead him to hypothesize the existence of genes as units of hereditary information which determine aspects of the phenotype. In the theoretically most convenient model, each gene acts independently from the other genes, giving a one-to-one correspondence of genes and their effects and a somewhat trivial GP map. However, soon after the re-discovery of Mendel's work, Bateson demonstrated that genes do not act independently [9], but that the emergence of phenotypes is subject to epistasis: Through interactions between different genetic elements, genes influence each other's effects. The effect of a given mutation thus depends on the genetic background in which it occurs.

With modern molecular hindsight, it appears obvious that genotype-phenotype maps display epistasis. After all, these maps encompass complex biophysical processes across multiple scales, from the expression of DNA into proteins through protein interactions to the behaviour of cells and tissues [98]. Thus we should not think of 'one gene per phenotype' [8] or of 'genes as blueprints' [101].

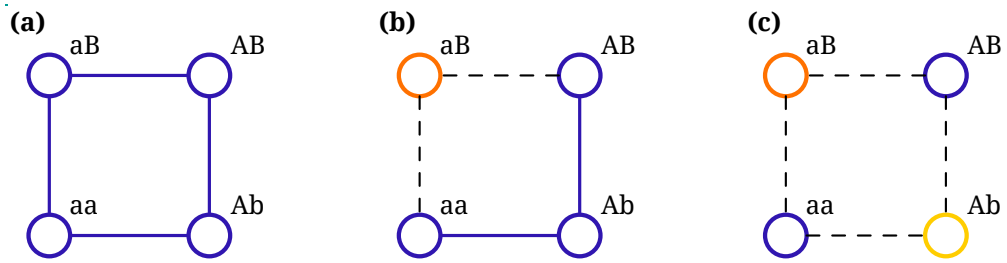

Figure 2.4: Epistasis shapes neutral spaces. Genotypes with the same phenotype (indicated by colour) belong to the same neutral space. Epistasis determines whether mutations are neutral (solid lines) or lead to a change in phenotype (dashed). **(a)** In the absence of epistasis, both $a \to A$ and $b \to B$ are neutral, and the neutral space is fully connected. **(b)** Under (phenotypic) sign epistasis, $b \to B$ is only conditionally neutral: on the a background, it leads to a change in phenotype, while on A it does not. The neutral space of the green phenotype is connected, but only the path $ab \to Ab \to AB$ is neutral. **(c)** Phenotypic reciprocal sign epistasis leads to a fragmented neutral space. Both $a \to A$ and $b \to B$ are individually deleterious, but compensate each other's effect.

In the literature, epistasis is often determined from the fitness effect of mutations [50, 97]. Since we are mainly interested in phenotypic variation and neutrality, we adapt the classification proposed by Weinreich et al. [135] to the description of phenotypes (Fig. 2.4). For simplicity, we consider a simple system of two loci with two alleles each, and we postulate that both the wild-type ab and the double-mutant AB share the same phenotype. So both phenotypes belong to the same neutral space. But will they be connected by a neutral path? This depends on the epistasis of mutations under the GP map.

In the absence of epistasis (Fig. 2.4a), both $a \to A$ and $b \to B$ are neutral so Ab and aB are in the same neutral space. Fig. 2.4b displays (phenotypic) *sign epistasis*, where $a \to A$ is always neutral but $b \to B$ is only neutral when it occurs with A. Thus aB has a different phenotype, and there is only one neutral path connecting ab and AB. The mutation $b \to B$ is thus *conditionally neutral*.

Finally, there is the possibility of (phenotypic) *reciprocal sign epistasis* (Fig. 2.4c): Individually both $a \to A$ and $b \to B$ change the phenotype, but together they are neutral. We can say that the two mutations are *compensatory*. In the two-allele, two-gene system, this type of epistasis leads to a fragmented neutral space; however, in a higher-dimensional system, there could be extra-dimensional bypasses connecting ab and AB by a neutral path.

On face value, (phenotypic) reciprocal sign epistasis may appear quite artificial. However,

in a classic paper on compensatory mutations, Kimura gave a very powerful analogy with a lock-and-key system [75]: A sufficiently sensitive lock can only be opened by a matching key. So if we have two configurations of a lock (say a and A) and two matching keys (b and B), then a change in the lock breaks the functionality of the system: Key b does not open lock A. So the system is only functional when a compensatory mutation changes the key (or the lock is mutated back).

Epistasis is manifest in the genotype-phenotype map, and it determines the shape and connectivity of neutral spaces. An important consequence of epistasis is the idea of conditional neutrality: A particular mutation may be neutral when applied to some genotypes, but change the phenotype when applied to other genotypes. Thus epistasis is of central importance for *evolvability*, which describes how phenotypic (and hence selectable) variation can arise. In the next section, we shall turn our attention to evolvability and thus complete our overview of the concepts we shall study in this dissertation.

2.2.5 Evolvability

Evolvability describes the propensity to introduce phenotypic variation and is the subject of much recent theoretical interest [28, 49, 76, 99, 127, 131]. Naively, we might expect that robustness and evolvability should be negatively correlated, because any particular mutation can either be neutral (increasing robustness) or change the phenotype (and hence contribute to evolvability). But as we shall see in this section, once we attempt to account for neutral spaces, evolvability becomes quite a subtle concept, and the relation of robustness and evolvability is far from clear.

Recent developments have tried to capture how the GP map affects evolvability. Several definitions of evolvability have been proposed in the literature, and we shall now discuss two approaches which both emphasize the importance of neutral space.

In an important paper, Wagner proposed to measure the evolvability of a genotype g by counting the number of distinct phenotypes in its one-mutational neighbourhood [127]. We shall call this the *potential evolvability* $\mathcal{E}_g$ of g, since this is the number of phenotypes

that can potentially be reached by following all possible single mutations away from g. Mathematically, if we write the set of these phenotypes as P_g, we have

$$\mathcal{E}_g = \left| P_g \right| \tag{2.21}$$

At the level of genotypes, we can make the antagonistic relation of robustness and evolvability quantitative through the inequality

$$\mathcal{E}_g \leq (1 - \rho_g)(K - 1)L \tag{2.22}$$

which states again that a mutation is either neutral or not, but not both at the same time.

In light of the existence of neutral spaces, however, we should take into account the possibility that populations either contain different genotypes simultaneously or explore the space over time [127]. As the population diffuses over the network by neutral mutations, non-neutral mutants 'fall off' and discover new phenotypes. Wagner proposed to account for these dynamics by measuring evolvability in terms of the number of distinct phenotypes that are accessible from an entire neutral space [127]. According to Wagner's definition, the potential evolvability $\mathcal{E}_p$ of phenotype p is given by the number of distinct phenotypes accessible from p's entire neutral space. Mathematically,

$$\mathcal{E}_p = \left| \bigcup_{g \in \mathcal{G}_p} P_g \right| \tag{2.23}$$

where $\mathcal{G}_p$ is the set of all genotypes mapping into p (that is, the genotypes of p's neutral network), just as in the definition of phenotype robustness (Eq. (2.20)).

Using RNA secondary structure folding as an example GP map and employing a sampling approach, Wagner demonstrated that at the level of phenotypes (or neutral networks) the correlation of robustness ρ_p and evolvability $\mathcal{E}_p$ is positive, while of course the correlation of ρ_g and $\mathcal{E}_g$ at the level of genotypes is negative (cf. Eq. (2.22)). With various collaborators, Wagner subsequently provided similar findings in a variety of different systems, from proteins

[40] to signalling circuits [106], metabolic networks [112, 113] and even evolvable computer hardware [105].

All these systems share the following characteristics [128, 130]: (1) The number of genotypes vastly exceeds the number of phenotypes. So there are multiple genotypes with the same phenotype. (2) Genotypes mapping into the same phenotype are often connected to each other by single neutral mutations. Thus, there are connected neutral spaces in these GP maps, and hence populations can be robust to mutations. (3) Neutral spaces typically extend far through genotype space: The Hamming distance of two genotypes in the same connected neutral space can be of the order of the genome length L. Such widely different genotypes typically have different sets of phenotypes in their direct mutational neighbourhood.

Together, these findings imply that larger neutral spaces will typically show a greater potential evolvability than smaller ones: As potential evolvability simply counts all accessible phenotypes, phenotypes with larger neutral spaces have larger $\mathcal{E}_p$, provided the sets of accessible phenotypes differ between genotypes. Intuitively, we would expect that larger spaces will also show a greater robustness (we will discuss this issue in more detail in Chapter 3). So larger neutral spaces typically show both greater robustness and potential evolvability [128].

This synergism of robustness and potential evolvability may arise simply because adding genotypes to a neutral space cannot decrease its potential evolvability – the union of two sets cannot be smaller than each set on its own. Cowperthwaite and co-workers have criticized this definition of evolvability because it does not account for the probability that the potentially accessible phenotypes will actually be produced by mutations [24]. They proposed a different measure of evolvability that aims to account for these probability. Their approach is based on the number of mutations n_{pq} that lead the neutral network of phenotype q to phenotype p. If we follow a random non-neutral mutation from q, we arrive at phenotype p with probability

$$c_{pq} = \frac{n_{pq}}{\sum_{a \neq q} n_{aq}} \tag{2.24}$$

We can then define *diversity evolvability* ϵ_q of phenotype q as

$$\epsilon_q = 1 - \sum_{p \neq q} c_{pq}^2 \tag{2.25}$$

which is the probability that two non-neutral mutations lead to two distinct phenotypes.

If all alternative phenotypes were equally well connected, we have $c_{pq} = 1/\mathcal{E}_q$ and hence $\epsilon_q = 1 - 1/\mathcal{E}_q$. Clearly, this provides an upper bound for ϵ_q: If any particular alternative phenotype a was favoured (in the sense that it had $c_{aq} > 1/\mathcal{E}_q$), this phenotype would be more likely to arise by two independent mutations, and hence would reduce the diversity evolvability[4]. So when all phenotypes are equally accessible, ϵ_q increases with $\mathcal{E}_q$. Again, it is clear that ϵ_q depends on the GP map, which determines the c_{pq}. When calculating diversity evolvability for short RNA secondary structures, Cowperthwaite et al. found that the diversity evolvability of a neutral space decreases with its size [24].

Our overview of static properties of GP map thus leaves us with an interesting conundrum: While potential evolvability $\mathcal{E}_q$ and robustness ρ_q are positively correlated, diversity evolvability ϵ_q appears to be negatively correlated with ρ_q. The studies of Wagner [127] and Cowperthwaite et al. [24] show that even the qualitative relation of robustness and evolvability is sensitive to the definition of evolvability. We shall revisit this problem in Chapter 5 where we introduce a population-dynamic measure of evolvability.

2.3 Evolutionary dynamics and the GP map

The concept of neutral spaces, arising from the GP map, leads to a relatively simple picture of the evolutionary process that we shall employ throughout this thesis: The population is viewed as a cloud in genotype space; under reproduction and mutation, the cloud moves (cf. Fig. 1.1). Some mutations are (effectively) neutral, and these neutral mutants may eventually become fixed and cause the centre of mass of the cloud to move; other mutations lead to a non-neutral change. If the resulting phenotype brings an increase in fitness, it

[4]It is easy to demonstrate this mathematically, by maximizing $1 - \sum_{i=1}^{N} c_i^2$ subject to the constraint $\sum_i c_i = 1$, for example using a Lagrange multiplier.

may go to fixation; deleterious mutants are quickly removed by selection. In short, adaptive evolution (transitions to fitter phenotypes) interrupts periods of neutral exploration.

As we shall see, the simplicity of this picture is deceptive, particularly when we deal with finite populations. For a more gentle start, we begin with the study of infinite populations.

2.3.1 Selection for robustness in quasispecies theory

As we have seen in Section 2.1.1 above, the evolution of infinite populations may be described in terms of deterministic differential equations in which fluctuations are ignored. The quasispecies equation (Eq. (2.12)) incorporates the effects of mutation and selection, and thus presents the ideal starting point for our present purposes. In order to understand the implications of mutational robustness on evolutionary dynamics, we will assume that only one phenotype p is viable while individuals with all other phenotypes cannot reproduce, so that the fitness of phenotype q is $f_q = \delta_{pq}$.

In this case, the dynamics of the population are completely determined by the structure of the neutral space of p [123]. We can describe this structure in terms of the adjacency matrix $\mathcal{A}$ (Eq. (2.17)). When μ_G is the probability of mutations per reproduction (the genomic mutation rate) and x_i is the fraction of individuals at genotype i, we have

$$\frac{dx_i}{dt} = \sum_j \frac{\mu_G}{(K-1)L}\mathcal{A}_{ij}x_j + (1-\mu)x_i - \bar{f}x_i \tag{2.26}$$

The first term accounts for mutant from adjacent genotypes, the second term describes the loss of i-individuals due to mutation, and $\bar{f}$ ensures that $\sum_i x_i = 1$. In matrix form, we have

$$\frac{d\mathbf{x}}{dt} = \left(\frac{\mu_G}{(K-1)L}\mathcal{A} - (\bar{f} + \mu - 1)\mathbb{1}\right)\mathbf{x} \tag{2.27}$$

where $\mathbb{1}$ is the identity matrix. The steady state distribution of the population is given by the leading eigenvector of $\mathcal{A}$, $\mathbf{x}^*$ with normalization $\sum_i x_i^* = 1$. From this distribution, we can calculate the mutational robustness of the population ρ^*, which is the occupation-weighted

average of the robustness of the genotypes in the population:

$$\rho^* = \sum_i x_i^* \rho_i = \frac{\sum_i x_i k_i^{(n)}}{(K-1)L}$$

where $k_i^{(n)}$ is the degree of genotype i in the neutral space (cf. Eq. (2.19)). By Eq. (2.18), we have $k_i^{(n)} = \sum_j \mathcal{A}_{ji}$ and therefore

$$\rho^* = \frac{\sum_{i,j} \mathcal{A}_{ji} x_i^*}{(K-1)L} = \frac{\mathbf{1}^T \mathcal{A} \mathbf{x}^*}{(K-1)L} = \frac{\lambda^*}{(K-1)L} \tag{2.28}$$

where $\mathbf{1}$ is a vector with each component being equal to 1, and λ^* is the leading eigenvalue of $\mathcal{A}$. This classic result – the population robustness is equal to the leading eigenvalue of the adjacency matrix of the neutral space, up to normalization $(K-1)L$ – was first derived by van Nimwegen et al. [123].

By the Frobenius-Perron theorem, λ^* is bounded from below by the average degree of nodes in the network $\langle k^{(n)} \rangle$ [90, 123]: $\lambda^* \geq \langle k^{(n)} \rangle$, where equality holds if and only if the number of neutral neighbours $k_g^{(n)}$ is the same for all genotypes g in the neutral space. Consequently, we have $\rho^* \geq \rho_p$, which we can view as an indication of selection for mutational robustness: Since individuals with more robust genotypes produce more viable offspring, the population will cluster in the regions of largest robustness in the neutral space. The variation necessary for this kind of selection stems from heterogeneity in robustness among the genotypes in the neutral space [123].

There is another scenario that highlights the importance of mutational robustness for evolving populations. Let us consider a coarse-grained picture of 3 phenotypes a, b, c of fitness $f_a = 1$, $f_b = 1 + s$ and $f_c = 0$: a is the wild-type, b is an adaptive type ($s > 0$), and c represents all unviable alternatives. Further, we denote by ρ_a and ρ_b the population robustness of a and b, and assume that mutations occur with probability μ_G per reproduction and only lead to c. Essentially, we have a fitness landscape with two plateaus, separated by

a valley that cannot be traversed. With x_i the frequency of type $i = a, b$, we have

$$\frac{dx_i}{dt} = f_i(1 - \mu_G(1 - \rho_i))x_i - \bar{f}x_i \tag{2.29}$$

where $\bar{f}$ as usual ensures that $x_a + x_b + x_c = 1$. Ignoring the small fraction of individuals of type c which are produced by mutations but cannot reproduce, we arrive at equations just like Eq. (2.8) but with fitness renormalized by $(1 - \mu_G(1 - \rho_i))$: Rather than counting the total expected number of offspring, the 'effective' fitness of individuals with type i measures the expected number of viable offspring, ie. those descendants that themselves will be able to reproduce. From our arguments in Section 2.1.1, we see that b goes to fixation when

$$\begin{aligned} (1 + s)(1 - \mu_G(1 - \rho_b)) &> 1 - \mu_G(1 - \rho_a) \\ s &> \frac{\mu_G(\rho_a - \rho_b)}{1 - \mu_G(1 - \rho_b)} \end{aligned} \tag{2.30}$$

Particularly when mutation rates are high, b can go extinct even if it is beneficial when it is not robust enough. Rather than the survival of the fittest, we then observe the 'survival of the flattest' [140]. When direct mutation from a to b (and vice versa) are also possible, there can be coexistence between the fittest and the flattest [10].

2.3.2 Finite populations

In practice, populations are finite. Just as for infinite populations, mutational robustness can allow finite populations to build up genotypic variation, sometimes also called 'cryptic' variation as it is not observable in the phenotype [63]. But in contrast to infinite populations, there is mechanism that can act to reduce this neutral diversity, namely genetic drift.

When populations are finite, they often cannot occupy the entire neutral space, simply because there might be more neutral genotypes than individuals. In addition, genetic drift leads to a localization in genotype space. In order to gain insight into evolutionary dynamics under drift, we ignore fitness effects and consider the evolution of a population in a completely flat fitness landscape, following Derrida and Peliti [31]. It can be shown (details are given

in Appendix A) that the average Hamming distance $\langle d \rangle$ between two individuals in the population is approximately given by

$$\langle d \rangle = d_\infty \left(1 + \frac{d_\infty}{2NL\mu}\right)^{-1} \tag{2.31}$$

where $d_\infty = L(K-1)/K$ is the average Hamming distance in an infinite population for genome length L and genome alphabet size K, population size N and mutation probability μ per base per reproduction. This result allows us to distinguish between *monomorphic* and *polymorphic* populations, based on the number of mutants per generation $NL\mu$. When mutants are rare $NL\mu \ll 1$, we have $\langle d \rangle = 0$: All individuals carry the same genotype. By contrast, when $NL\mu \gg 1$, the population contains many different genotypes at any one time. Sumedha et al. showed by numerical simulations that the scaling with $NL\mu$ is subject to corrections of order $1/N$ [118].

Selection for robustness arises from the variation in robustness. Of course, in monomorphic populations, there is no variation in robustness, and thus these populations do not evolve to greater robustness [123]. Instead, monomorphic populations perform a special kind of random walk on the neutral space, sometimes called 'blind ant' dynamics [123]. In a usual random walk on a network, where the walker takes one step per unit time, the time spent at each node is proportional to the node's degree [90]; intuitively, this is because the walker follows a random edge at each time step and is thus more likely to move to a node of high degree. However, this is not quite true for an evolving population, which spends (on average) equal amounts of time at each genotype [123]. We can understand this in terms of Kimura's molecular clock. The rate of neutral fixations is simply the rate of neutral mutations, and this is clearly proportional to mutational robustness. So the expected time spent at any one genotype is $\tau_g = 1/(L\mu\rho_g)$, provided $L\mu \ll 1$, so that we may ignore double mutations. So while it is still true that the population will often move to a robust genotype, these genotypes are also left faster than less robust ones. Averaged over long times, the robustness of the population will simply be the average robustness of the genotypes in the neutral space.

The dynamics of neutral exploration in finite populations has been the subject of several studies. van Nimwegen and Crutchfield compared the speed of neutral exploration to the rate at which fitness valleys are crossed [121]. Specifically, the authors compared the time taken for a population starting from a single genotype g_0 to a (single) target genotype at Hamming distance w in two cases: First, when all intermediate states had lower fitness than g_0, corresponding to the crossing of a fitness valley, and second, when all mutations are neutral. While the first case could be solved analytically, for the (to us more interesting) second case, only very loose, approximate bounds where given. Nonetheless, van Nimwegen and Crutchfield showed that for fixed w, exploration along a neutral ridge is generally faster than the traversal of a low-fitness valley.

In Section 2.2.5, we saw that two different static definitions of evolvability give qualitatively different results for the relation of robustness and evolvability. In both papers, the authors also carried out simulations under Wright-Fisher dynamics to support their claims. Wagner [127] showed that populations with more abundant phenotypes (corresponding to larger and more robust neutral spaces) had more different phenotypes in their one-mutational neighbourhood than populations with less abundant phenotypes. In contrast to our developments in Chapter 5, Wagner did not count the phenotypes actually produced by non-neutral mutations, but kept track of all phenotypes that potentially could have been produced by mutants stepping off the neutral network.

The study of Cowperthwaite et al. [24] followed a different approach. The authors studied the dynamics of adaptive evolution towards a predefined target phenotype over a million generations in a monomorphic populations. They noted that frequent phenotypes (realized by a very large number of genotypes) were often discovered within the simulation time, while the populations often did not reach the target if it was rare (that is, only a relatively small number of genotypes mapped into the target phenotype). At the end of such 'failed' simulations, the population was often found to occupy a large neutral space from which adaptation was possible in principle. These large (and thus more robust) neutral spaces thus acted as entropic barriers that could trap populations for very long times. Cowperthwaite et al. called this effect the 'ascent of the abundant': When observing a particular phenotype in

nature, this phenotype need not be a fitness peak, but may in fact more likely correspond to a large saddle in the landscape. Their work highlights an important misconception that is evident in the term 'evolutionary optimization': Competition only leads to selection of the fittest phenotype *present in the population*, but this does not guarantee global optimality.

Finally, Draghi et al. [33] were able to make analytic progress in the description of neutral exploration. Their calculations focused on the discovery time of a novel phenotype, when the population is otherwise restricted to a single neutral space. In their description of the GP map, mutations are neutral with a fixed probability (which describes robustness) and a single neutral mutation completely reshuffles the mutational neighbourhood. By very technical calculations, they showed a non-monotonic dependence of the first discovery time on robustness with an intermediate level that minimizes the discovery time. To put this result into perspective, it is important to note that under this model, the target phenotype is guaranteed to be eventually discovered by a single mutation. As we shall see in the next chapter, this need not be the case under more complex GP maps.

To summarize, the description of neutral exploration is complicated by several factors: Most importantly, the dynamics of evolution under genetic drift are highly intricate. In particular for this reason, purely static approaches can be misleading. On the other hand, the complex structure of the GP map cannot be ignored in a realistic treatment. To get an impression of the features that we may want to include in a dynamic model, we investigate the structural properties of an important model GP map in the next chapter.

Chapter 3

The static structure of neutral spaces

3.1 Introduction

Neutral spaces arise from GP maps under which different genotypes map into the same phenotype [115]. Both the internal structure of these spaces as well as the connections between them have profound consequences for the dynamics of evolving populations. The existence of mutational connections within neutral spaces is related to mutational robustness [125]. By exploring neutral spaces along these neutral mutations, populations have access to a much larger variety of phenotypes than what is found in the neighbourhood of a single genotype [65, 130].

The impact of neutral spaces rests crucially on their connectedness under *small* mutational steps. By contrast, a neutral 'space' consisting of many mutationally isolated genotypes would endow populations with hardly any robustness against mutations and would not allow for easy exploration (see Fig. 3.1). To distinguish between the connected and disconnected parts of a neutral space, we employ terminology from graph/network theory: A *neutral network* (NN) is formed by *all* genotypes with a given phenotype; a *neutral component* (NC) is a maximal set of genotypes that map into a fixed phenotype and are connected by single mutations.

Whether or nor not a given network contains several components depends on the details of the GP map. In this chapter, we will study two different model GP maps. First, we will

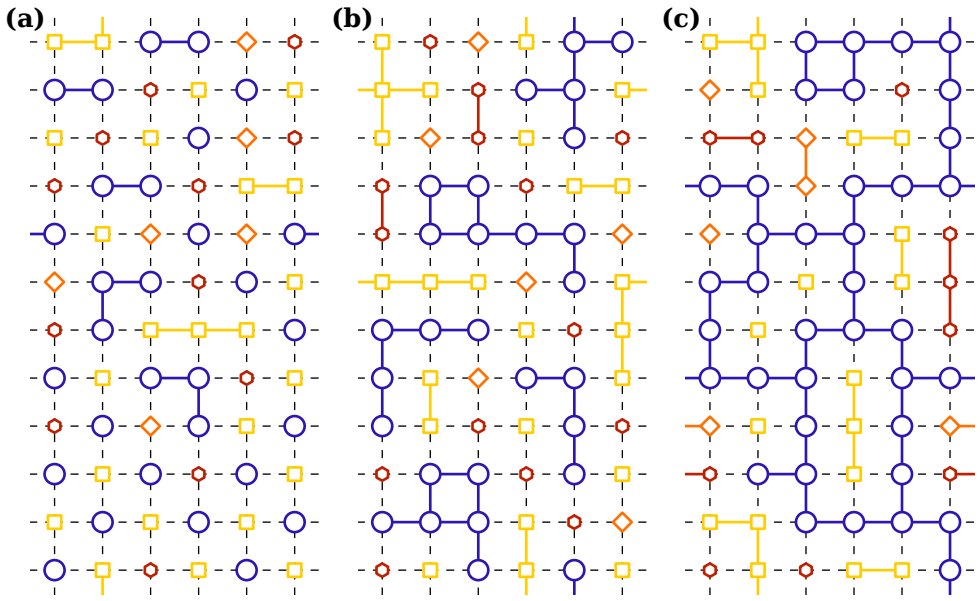

Figure 3.1: Neutral network structure constrains neutral exploration. The diagram contrasts three different types of structure for the neutral network of the blue phenotype. **(a)** The neutral network consists of mostly isolated genotypes; only a few neutral components comprise more than one genotype. Neutral exploration is hardly possible in this setting. **(b)** Here, the neutral network is fragmented into several neutral components, but all of these components contain several genotypes. The scope of neutral exploration is limited to a single component. **(c)** When the neutral network is fully connected, all genotypes with the blue phenotype can in principle be reached by neutral exploration.

briefly investigate a random assignment of genotypes to phenotypes, where the *frequency* F_p of a phenotype p (that is the fraction of all genotypes mapping into p) can be used an control parameter. This random model is conceptually straightforward and some analytic results have been derived, even though the high-dimensionality of genotype space leads to quite involved technicalities [108, 109]. As a function of F_p, we observe two transitions: First from a collection of isolated genotypes to a giant component surrounded by many small components [109], and then to a single connected network [108]. Both of these transitions occur at relatively high values of F_p: Under the random GP map, only a few phenotypes have NCs of considerable size.

We shall then turn our attention to a more biologically realistic GP map, namely RNA secondary structure folding. In this map, genotypes are RNA sequences, that is ordered sequences of the four nucleotides adenine (A), cytosine (C), guanine (G) and uracil (U), which replaces thymine (T) from the DNA alphabet. A and G belong to the chemical family of

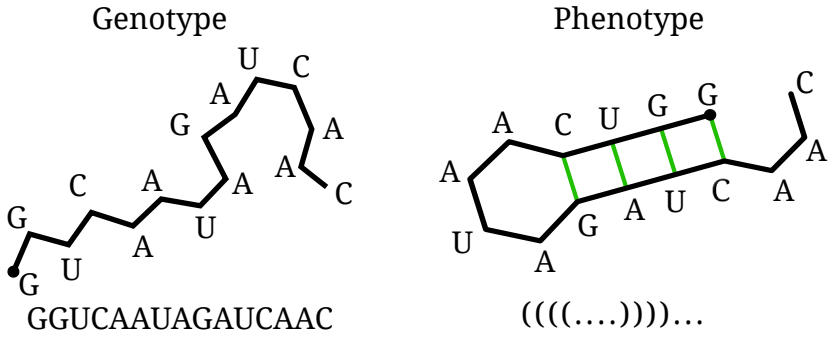

Figure 3.2: Illustration of RNA secondary structure folding. Genotypes in this map are RNA sequences, consisting of a ribose backbone (drawn in black) with nucleotide bases attached. Genotypes are uniquely represented by the sequence of nucleotides, which is read from the 5′ end (marked by the dot) to the 3′ end by convention. Phenotypes are the secondary structures of minimum free energy (throughout this thesis, we will consider only the case of physiological temperature 37°C). The secondary structure describes which bases form base-pairs (drawn in green). A unique representation of the secondary structure is provided by the 'dot-bracket' notation [64] (given below the phenotype), in which paired bases are written as matching parentheses and unpaired based are indicated by dots.

purines, while C and U are pyrimidines [98]. Just as in DNA, a purine and a complementary pyrimidine can form hydrogen bonds, leading to base-pairs. This base-pairing, together with stacking interactions of neighbouring nucleotides, leads to the folding of single RNA strands back onto themselves. In contrast to T in DNA, U not only forms the analogue of the Watson-Crick pairing with A, but also allows a 'wobble' pairing with G. In total, there are thus six different *compatible* nucleotide doublets, namely the purine-pyrimidine pairs GC,GU,AU and their pyrimidine-purine counterparts CG,UG,UA.

RNA strands fold into complex three-dimensional shapes. An important coarse-grained description of these shapes is the *secondary structure* which only describes which bases form pairs. Here, we will use the mapping from RNA sequence into the secondary structure of minimum free energy as a model GP map (see Fig. 3.2). This abstraction of RNA folding has been one of the most important workhorse model GP maps (e.g. the studies of evolvability by Wagner [127] and Cowperthwaite et al. [24] used simulations in RNA), as it combines biological relevance – the secondary structure of an RNA molecule is an important determinant of its biological function [39, 43], and is conserved over evolutionary timescales [125] – with reliable computational tractability [64].

The RNA GP map displays strong epistasis (cf. Section 2.2.4) which can be understood in terms of its base-pairing logic. The effect of a particular mutation is very sensitive to the other bases in the sequence. For example, a GU pair could be transformed into a GC pair, while the same mutation U→C could break an AU pair into a non-complementary AC configuration. It is particularly important that the effect of this mutation can be *compensated* by a mutation in the first nucleotide, in this case from A to G. As we shall see below, this manifestation of reciprocal sign epistasis causes the fragmentation of neutral spaces in RNA. This fragmentation has profound consequences: Different components of the same phenotype are very heterogeneous in their properties. In particular, the spectra of accessible alternative phenotypes can vary considerably between components. If populations are localized to just a single component, the heterogeneity of these components may cause contingency in evolution, giving a clear indication that the structure and interconnections of neutral spaces can have a strong effect on the course of evolution.

3.2 Random colouring of the genotype hypercube

Under point mutations, genotypes are organized into a generalized hypercube in L dimensions, where L is the length of the genotype. For the purpose of visualization, it is convenient to think of the GP map as a colouring of the hypercube: Genotypes with the same colour belong to the same neutral network. To build intuition about the nature of GP maps and neutral spaces, we shall begin by studying a random colouring in which each genotype is assigned one of N_P phenotypes.

In this random GP map, the internal structure of a neutral network can only depend on L, the alphabet size K ($K = 4$ for DNA and RNA) and the size of the network. We parametrize network size as phenotype frequencies F_p, which correspond to the number of genotypes in the network $\mathcal{S}_p$, normalized by the total number of genotypes K^L:

$$F_p = \frac{\mathcal{S}_p}{K^L} \tag{3.1}$$

Normalizing F_p by the size of genotype space implies that $\sum_{p=1}^{N_P} F_p = 1$.

The structure of neutral spaces in the random GP map can be studied using ideas from percolation theory. In a mean-field picture, it is clear that as a function of F_p we should expect a transition from largely isolated genotypes to a large connected component around a threshold

$$\delta = \frac{1}{(K-1)L} \tag{3.2}$$

when each genotype has on average one neutral neighbour, in analogy to the percolation transition in random Erdös-Rènyi networks [90]. We call δ the *giant component onset*. For $F_p < \delta$, we expect that the average size of connected components is small, while for $F_p > \delta$, the largest component size should scale with F_p. Reidys gives a technical proof for this value of δ [109].

In the limit $L \to \infty$, it can be shown analytically that there is another transition at

$$\lambda = 1 - \frac{1}{\sqrt[K-1]{K}} \tag{3.3}$$

such that if $F_p > \lambda$, the neutral network of p is in general completely connected [108].

We have subjected these predictions to numerical tests for $K = 4$ and $L = 12$, so that $\delta = 1/36 \approx 0.03$. In contrast to the 'canonical' setting (each genotype is part of the NN with a fixed probability) in which the analytic results have been derived, we constructed 100 realizations of the random GP map in a micro-canonical approach (the number of genotypes per phenotype is held fixed across the ensemble). The motivation for the micro-canonical setting is two-fold: First, our later models will feature a deterministic GP map in which the phenotype frequencies are fixed. Second, in the microcanonical setting we can evaluate the NN structure for all the phenotype frequencies in the distribution, which allows us to efficiently sample the neutral space structure over a wide range of phenotype frequencies.

For each realization, we evaluated the size of the largest NC and the total number of NCs per phenotype. The results show a good agreement with the analytic predictions: When the frequency F_p of phenotype p is below the giant component threshold δ (Eq. (3.2)), the largest NC is typically tiny compared to the entire NN (Fig. 3.3a). But when $F_p > \delta$, the size of the

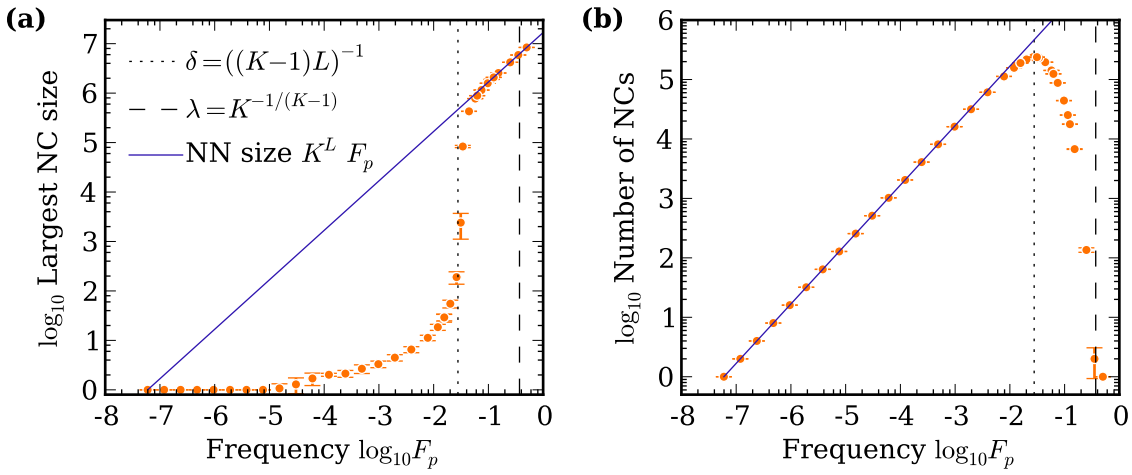

Figure 3.3: Percolation on genotype space. (a) shows the size of the largest NC in the NN of phenotype p against the frequency F_p of this phenotype. The blue line shows the total NN size, given by $K^L F_p$, and the vertical lines mark the gaint component onset δ (Eq. (3.2)) and the full connectivity threshold λ (Eq. (3.3)). **(b)** shows how the total number of NCs in an NN depends on its size. Each data point corresponds to the average of 100 realizations with $K = 4$ and $L = 12$; error bars indicate the sample standard deviation.

largest NC clearly scales with the NN size. It is worth noting that the mean-field argument δ = (degree of each genotype in genotype space)$^{-1}$ predicts to the actual transition point extremely well; this accuracy is a consequence of the high dimensionality of genotype space.

Fig. 3.3b underlines that below the giant component onset, most genotypes are completely isolated: The total number of NCs scales with the size of the NN, that is the total number of genotypes in the NN. Around the giant component threshold δ, the number of NCs drops because now a single component contains almost all genotypes. Finally, when F_p exceeds the connectivity threshold λ (Eq. (3.3)), the NN is completely connected.

It is important to realize that λ is independent of L. For $K = 4$, we obtain $\lambda \approx 0.37$. So in the random GP map, we should not expect more then 2-3 fully connected networks. Further, since $\delta \propto L^{-1}$, the upper bound on the total number of large NC (given by $1/\delta$) grows only linearly in L. While the total number of phenotypes will depend on the particular GP map, the exponential growth in the number of genotypes makes it plausible that the number of phenotypes could also be quite large. Our analysis shows that under the random GP map, most of these phenotypes could not have NCs of appreciable sizes, let alone fully connected NNs.

3.3 Neutral spaces of RNA secondary structures

Let us now turn our attention to RNA. As outlined above, RNA has been an important model GP map that has been studied widely in the past [23, 42]. The most detailed results can be obtained from exhaustive enumeration of sequences with fixed length L; due to the combinatorial explosion of the size of genotype space, these studies have been limited to relatively short sequences [2, 24, 114] or to reduced alphabets including only Guanine and Cytosine nucleotides [56].

Here, we will study both the internal and external connectivity of RNA neutral spaces, moving from questions about large-scale properties such as neutral space sizes to details of neutral space inter-connections. In particular, we shall demonstrate that neutral networks of RNA secondary structures are generally not fully connected; instead, due to the base-pairing logic, the network corresponding to a secondary structure with n base-pairs has typically at least 2^n disjoint components. We shall then compare the properties of the components with a fixed phenotype. Our most important result is that the spectrum of accessible, alternative phenotypes can vary significantly between different components. When evolving populations are well localized in genotype space, this fragmentation of neutral spaces into heterogeneous components suggests contingency in evolution.

3.3.1 Global properties of the RNA GP map

In contrast to the random GP map, the phenotype frequencies F_p in RNA are a fixed property of the GP map. In the random map, these frequencies are of fundamental importance. Thus we shall begin our exploration of RNA neutral spaces by determining the phenotype frequencies. As a by-product, we will learn another important large-scale property of the RNA map, namely the total number phenotypes N_P.

Secondary structures can be described in terms of the adjacency matrix of a network, in which nodes correspond to nucleotides and edges represent the ribose backbone and base-pairing. This construction allows for an enumeration of valid secondary structures for a

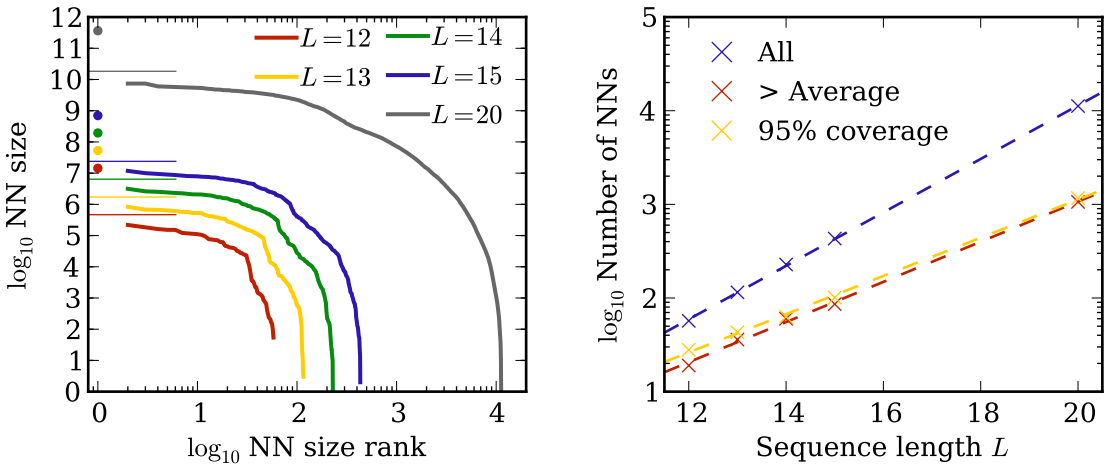

Figure 3.4: A minority of phenotypes covers the majority of genotype space. (a) For different sequence lengths L (indicated by colour), phenotypes are ranked by their size. The trivial structure without base-pairs has rank 1. The horizontal lines mark the giant component transition point, $4^L/3L$ (see Eq. (3.2)). **(b)** The number of phenotypes increases exponentially with L, but most of these phenotypes have tiny NNs. The number of frequent (greater than average) phenotypes and the number of phenotypes that contain 95% of all genotypes also increase exponentially, but at a slower rate. The trivial structure has been excluded in these calculations.

given length L, showing that the number of allowed structures scales approximately as [133]

$$S(L) \propto 2.21^L \tag{3.4}$$

An improved approximation (accounting for the fact that two the bases of a base-pair cannot be next to each other in the sequence) suggests a slightly different scaling of the number of structures: [115]:

$$S(L) \propto L^{-3/2} 1.85^L \tag{3.5}$$

Both of these scaling relations only determine the number of valid structures; but not all these structures are necessarily realized by at least one sequence. Nonetheless, we should expect to see a roughly exponential increase in the total number of actually observed phenotypes, $N_P(L)$. We also note that the rate of increase in $S(L)$ is smaller than the growth of the number of genotypes 4^L: Thus we expect that the average size of neutral spaces increases with L, while phenotype frequencies decrease.

Fig. 3.4a shows the distribution of phenotype frequencies in RNA, obtained through

exhaustive enumeration. The most interesting feature of this distribution, which has been observed in several earlier studies [24, 56, 107, 115], is the strong *phenotypic bias* [13]: Most phenotypes are rare, while the majority of all genotypes is covered by a small fraction of very frequent phenotypes. Fig. 3.5 shows the 30 most abundant (non-trivial) structures for $L = 15$, which contain 56% of the non-trivially folding genotypes, and the 10 least abundant structures.

For the relatively short sequences that can be studied using exhaustive enumeration (already for $L = 20$, there are more than 10^{12} sequences!), the trivial structure without any bonds is by far the most abundant. But if fitness is based on the secondary structure of an RNA molecule, an unfolded strand is unlikely to provide an adaptive advantage; therefore we ignore the trivial structure in the following discussion. Furthermore, we can see that the frequency F_0 of the trivial structure decreases with sequence length (see also Tab. 3.1 on page 44). To illustrate this, we randomly sampled 10^6 sequences of length $L = 50$; in the sample, the trivial structure had a frequency of 0.12%. Results for the trivial structure at small L are therefore unlikely to hold for longer (and thus biologically more relevant) sequences.

The heterogeneity of phenotype frequencies can be quantified in different ways. For example, we can calculate the number of largest phenotypes that together cover a certain fraction of genotype space. It is also convenient to be able to distinguish between rare and frequent phenotypes. To this end, we follow Grüner et al. [56] and use a cutoff given by the average frequency $\langle F\rangle = (1 - F_0)/N_P$ where N_P is the number of (non-trivial) phenotypes. If phenotype p has $F_p > \langle F\rangle$, we call p a 'frequent' phenotype; otherwise, we say p is 'rare'.

Fig. 3.4b shows how the number of phenotypes depends on L. From a simple fit to an exponential curve, we obtain for the number of non-trivial phenotypes N_P as a function of L

$$N_P(L) \propto 1.93^L \tag{3.6}$$

While the numerical results is somewhat in-between the scalings suggested by Eqs. (3.4) and (3.5), our findings underline that the number of phenotypes grows rapidly with L, while

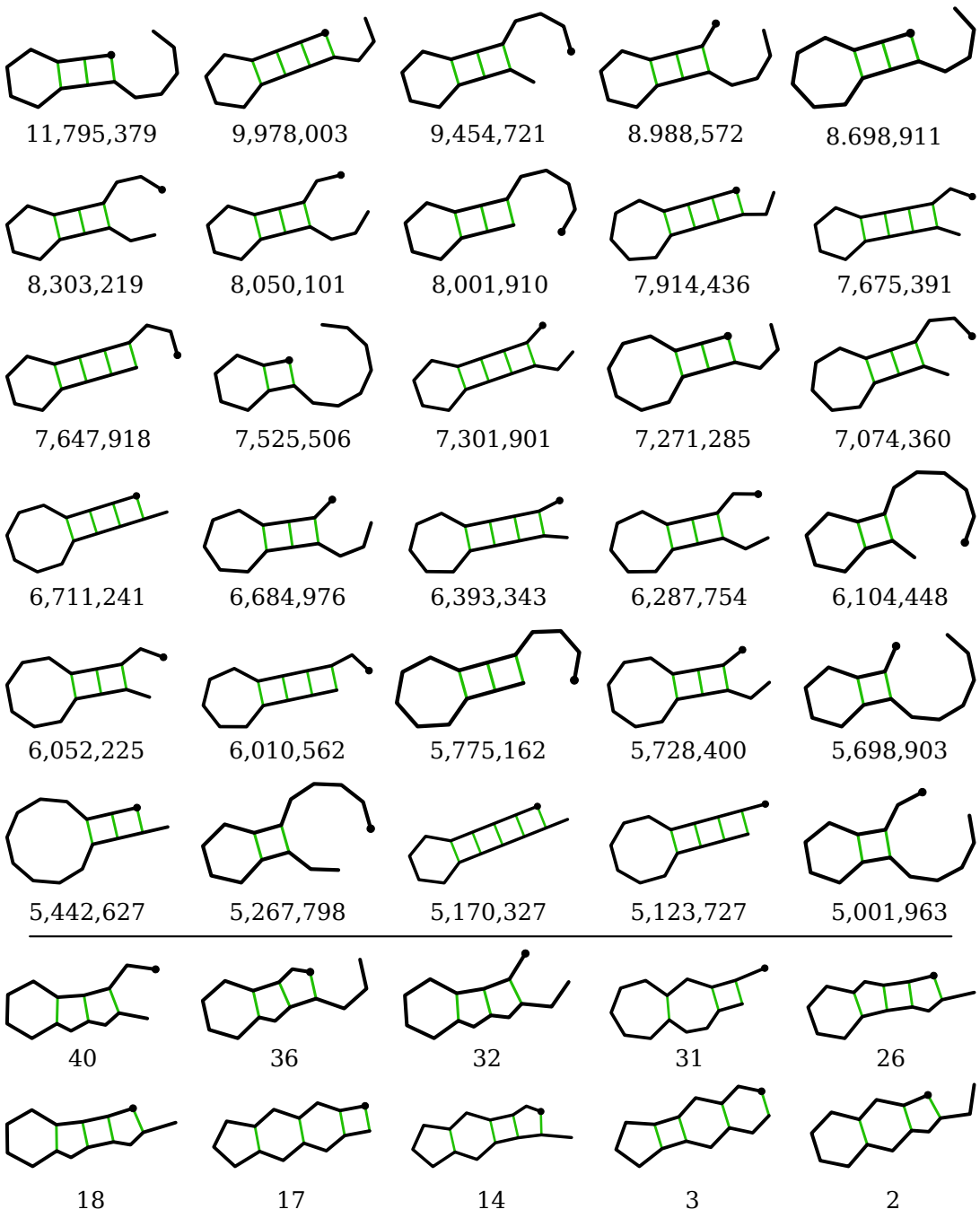

Figure 3.5: The most abundant RNA secondary structures for $L = 15$. The diagram shows 30 most abundant (top) structures at the top; these structures together contain 56% of all non-trivially folding genotypes; at the bottom, the 10 least abundant structures are shown. Below each structure, the exact number of sequences folding into the respective structure are shown. Just as in Fig. 3.2, the backbone is shown in black, base-pairs are drawn in green and the $5'$ end of the structure is marked by the dot.

of course we cannot confidently conclude that the scaling should be purely exponential. More interestingly, the number of frequent phenotypes (ie. the phenotypes of more than average

L	N_P	F_0	$N_P^{freq.}$	$N_P^{freq.}/N_P$	$N_P^{cov.}$	$N_P^{cov.}/N_P$
12	57	0.85	19	0.33	27	0.47
13	115	0.79	36	0.31	42	0.37
14	228	0.72	60	0.26	62	0.27
15	431	0.65	86	0.20	100	0.23
20	11218	0.33	1062	0.09	1168	0.10

Table 3.1: The RNA map is characterized by large-scale heterogeneity. For sequence length L, the table lists the number of non-trivial structures N_P, the frequency of the trivial structure F_0, the average (non-trivial) neutral network size $\langle V_S \rangle$, the number of frequent structures $N_P^{freq.}$, and the proportion r_{freq} of sequences in frequent structures to all sequences with a non-trivial structure.

frequency) also increases exponentially with L but at a smaller rate:

$$N_P^{freq.}(L) \propto 1.64^L \tag{3.7}$$

The growth of the number of phenotypes covering 95% of all non-trivial genotypes is slightly slower still:

$$N_P^{cov.}(L) \propto 1.60^L \tag{3.8}$$

So as sequence length L increases, genotype space becomes dominated by an exponentially decreasing fraction of phenotypes, although in absolute terms, the number of dominating phenotypes grows exponentially. Table 3.1 summarizes the large-scale properties of the RNA map for different values of L.

It is intuitively clear that this strong heterogeneity of phenotype frequencies will have important consequences for evolutionary dynamics; the implications of phenotypic bias will take centre stage in chapter 4. For now, we shall continue our investigation into the structure of neutral spaces.

3.3.2 The structure of neutral spaces

We can see in Fig. 3.4 that no non-trivial structure comes close in frequency to the giant component threshold δ (Eq. (3.2)), let alone the threshold λ (Eq. 3.3) for fully connected networks. So if the RNA sequences folding into a given structure were distributed randomly

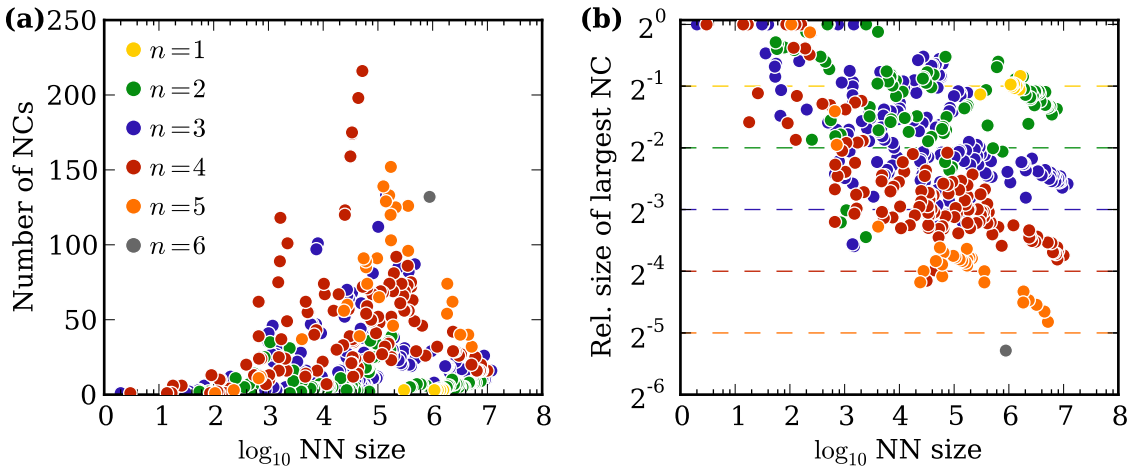

Figure 3.6: RNA neutral spaces are partially connected. **(a)** The number of NCs shows no simple relation to the size of an NN. The colour indicates the number of base-pairs n in the corresponding phenotype. **(b)** Here, the size of the largest NC is plotted against NN size, with the same colour coding as in (a). The data is for $L = 15$.

in genotype space, we would not expect to see any connected neutral spaces of appreciable sizes (apart from the NN of the trivial structure).

Fig. 3.6 shows the actual number of NCs in the RNA NNs for $L = 15$, along with the size of the largest component in each network. Neither of these two order parameters shows any reminiscence of the clear dependence on NN size (which is simply proportional to phenotype frequency) that we saw in the random map (Fig. 3.3). In addition to the trivial structure (whose NN is fully connected for $L = 12$ to $L = 15$), there are 19 phenotypes with fully connected NNs. All of these are small; the largest connected NN contains 2464 genotypes and is ranked 333^{rd} by size. There are 334 NCs of size 1, corresponding single genotypes completely isolated from the rest of their NN.

Fig. 3.7 displays the properties of the global NC size distribution for different sequence lengths. Since the NN size distribution shows a strong skew and the NN size bounds the size of the NCs, it is does not come as a surprise that NCs also have a skewed size distribution (see Fig. 3.7a). To quantify the heterogeneity of this distribution, we call an NC 'large' if its size is greater than the average NC size and count the number of large NCs $N_C^{lrg.}$. We also calculate the number of largest NCs $N_C^{cov.}$ that together cover 95% of all (non-trivially folding) genotypes. Fig. 3.7b shows that the total number of NCs $N_C(L)$ as well as $N_C^{lrg.}(L)$

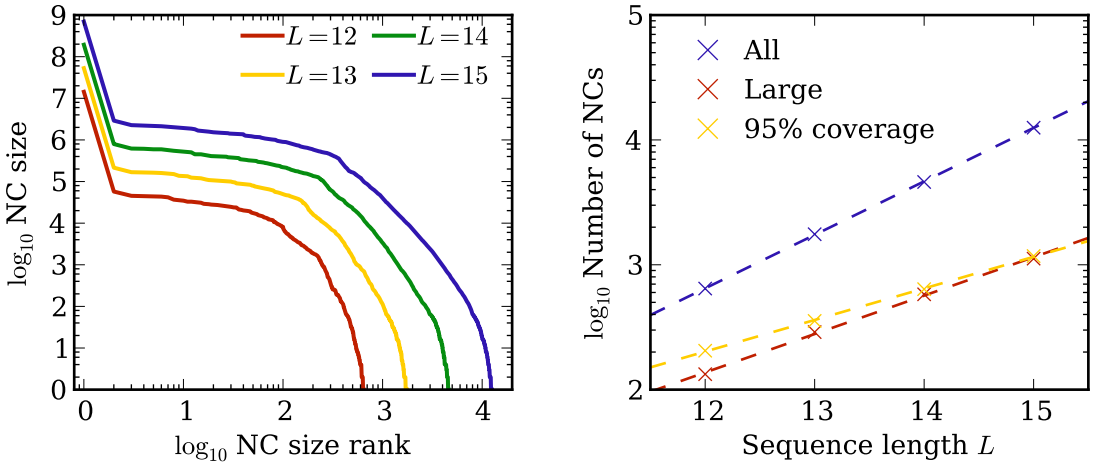

Figure 3.7: The distribution of RNA NC sizes is skewed. (a) The size of an NC is given by the number of genotypes in the NC. For each sequence length L, the NCs have been ranked by their size. At the short sequence lengths amenable to exhaustive enumeration, the NN of the trivial structure is fully connected and has been omitted from the diagram for clarity. **(b)** The total number of NCs increases exponentially with L. The number of large (greater than average) NCs and the number of NCs containing 95% of all (non-trivially folding) genotypes also increase exponentially, but at a slower rate.

and $N_C^{cov.}(L)$ increase roughly exponentially with sequence length L. From our data, we obtain the scaling behaviour

$$N_C(L) \propto 2.68^L \tag{3.9}$$

$$N_C^{lrg.}(L) \propto 2.03^L \tag{3.10}$$

$$N_C^{cov.}(L) \propto 1.79^L \tag{3.11}$$

These findings suggest that most NCs are tiny, and as sequence length increases, genotype space is covered by an ever-smaller fraction of the NCs. Just as for NNs, the absolute number of dominant NCs increases with L. And in terms of absolute size, NCs get larger as L increases. Even if we ignore the fact that more and more genotypes map into non-trivial phenotypes, the average NC size increases exponentially like $(4/2.68)^L$.

Comparing NNs and NCs, we find that the heterogeneity at the NC level is even more pronounced: The fraction of NCs covering 95% of genotype space decays as $(1.79/2.68)^L \approx (2/3)^L$ while the analogous fraction of NNs decays as $(1.60/1.93)^L \approx (5/6)^L$. Table 3.2 summarizes these results quantitatively.

L	N_C	$N_C^{freq.}$	$N_C^{freq.}/N_C$	$N_C^{cov.}$	$N_C^{cov.}/N_C$	N_1	N_1/N_C
12	645	133	0.21	204	0.32	14	0.02
13	1757	289	0.16	354	0.20	64	0.04
14	4603	580	0.13	637	0.14	84	0.02
15	12526	1120	0.09	1174	0.09	334	0.03

Table 3.2: RNA NCs have very different sizes. The table quantifies the total number of NCs N_C, the number of large (ie. greater than average) NCs $N_C^{lrg.}$, the number of NCs covering 95% of (non-trivial) genotype space $N_C^{cov.}$ and the number of isolated genotypes N_1 for different sequence length. While the absolute number of large and covering NCs increases with L, these NCs constitute a decreasing fraction of all NCs. Exhaustive determination of NCs at $L = 20$ is computationally unfeasible.

The evident discrepancy with the random map results leads us immediately to a profound insight: *In RNA, genotypes with the same phenotype are not distributed randomly in genotype space.* In the next section, we will use the biophysical details of the RNA folding process to understand these deviations from the random GP map.

3.3.3 The impact of epistasis on neutral space structure

The RNA GP map is simple enough that we can understand several features of the structure of its neutral spaces. For our purposes, the crucial biophysical feature of the folding process is the base-pairing logic. If a sequence folds into a particular structure, it necessarily has to be *compatible* with that structure: The bases that form pairs must be complementary – otherwise, the structure could not be formed. This simple biophysical necessity has fundamental consequences for the structure of neutral spaces.

A first biophysical argument suggests why neutral spaces in RNA might be more tightly connected than in the random GP map: Since only 6 out of the 16 nucleotide doublets allow the formation of a base-pair, there are $4^{L-2n}6^n$ compatible sequences of a structure comprising n base-pairs. While this is still an astronomically vast number, these sequences belong to a small fraction of sequence space: The relative size of the 'compatible set' is $(6/16)^n$.

For simplicity, let us consider an even more restricted set of compatible sequences where we completely fix the paired bases (say, to GC pairs throughout the sequence). This 'scaffold'

corresponds to an $(L-2n)$-dimensional sub-hypercube. If we assume that the genotypes realizing p are distributed equally among the 6^n scaffolds, the frequency of p *within each sub-hypercube* is given by $F_p \times (16/6)^n$. The exponential increase with n suggests it may be possible to reach a reduced threshold Δ_n for the formation of a giant component, given by

$$\Delta_n \equiv \left(\frac{16}{6}\right)^n \delta = \left(\frac{16}{6}\right)^n \frac{1}{3L} \tag{3.12}$$

in this sub-hypercube, even though the global frequency F_p is small. In addition to the trivial structure, for $L = 15$ there are 67 phenotypes with frequencies greater than Δ_n.

However, we can make a clear biophysical argument why RNA NNs in general cannot be fully connected; the argument is again based on the idea of compatible sets. It is easy to see that each such set falls into 2^n disjoint sub-spaces: For each base-pair, it is not possible to connect a purine-pyrimidine (e.g. GC) to a pyrimidine-purine (CG) by a mutational path that maintains compatibility. While may be possible to change GC into GU and then AU, a change in the order of a base-pair necessarily involves an intermediate state of non-complementary bases. So even if all compatible sequences did actually fold into a particular structure (so that $F_p = (6/16)^n$), the corresponding NN would still contain 2^n NCs. All of these NCs would have similar size, so that the relative size of the largest NC would only be 2^{-n}. In Fig. 3.8, we see that this biophysical argument which incorporates the epistatic interactions in RNA, does indeed explain the most salient features of the structure of the largest NNs: The NNs are covered either completely, or at least to a very large proportion, by 2^n NCs.

Another clear indication of the importance of epistasis on RNA neutral space structure is given in Fig. 3.9: When we allow for base-pair swaps (e.g. changing GC into CG) as single mutations, the number of NCs per NN drops drastically (compare the scale of the y-axis of Figs. 3.6a and 3.9a), and the most NNs are dominated by a single NC (Fig. 3.9b).

In practice, not all genotypes compatible with a phenotype do actually map into that phenotype. One way to see that this must be true in RNA is that compatible sets are nested: all sequences compatible with some phenotype p are also compatible with all phenotypes in

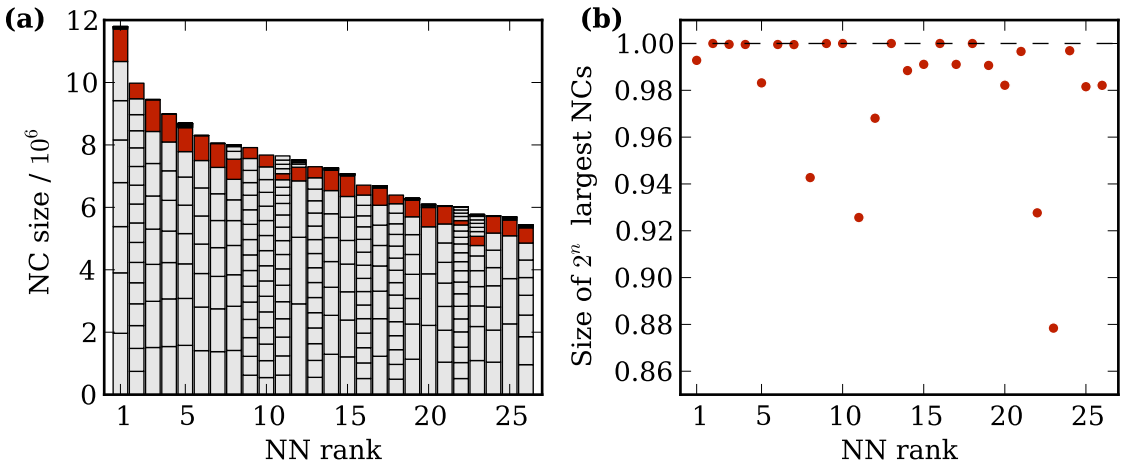

Figure 3.8: NNs are dominated by many large components. **(a)** shows the sizes of the NCs in the 26 largest NNs of $L = 15$ RNA; together, these NNs contain just over 50% of all non-trivially folding genotypes. Each column represents one NN, and the blocks indicate NC sizes, ranked from the largest NC (in the respective NN) at the bottom to the smallest NC at the top of the stack. The red block in each stack corresponds to the NC of rank 2^n where n is the number of bonds in the phenotype corresponding to the NN. **(b)** shows the fraction of genotypes contained in the 2^n largest NCs in each NN. Together, these components clearly dominate the NN.

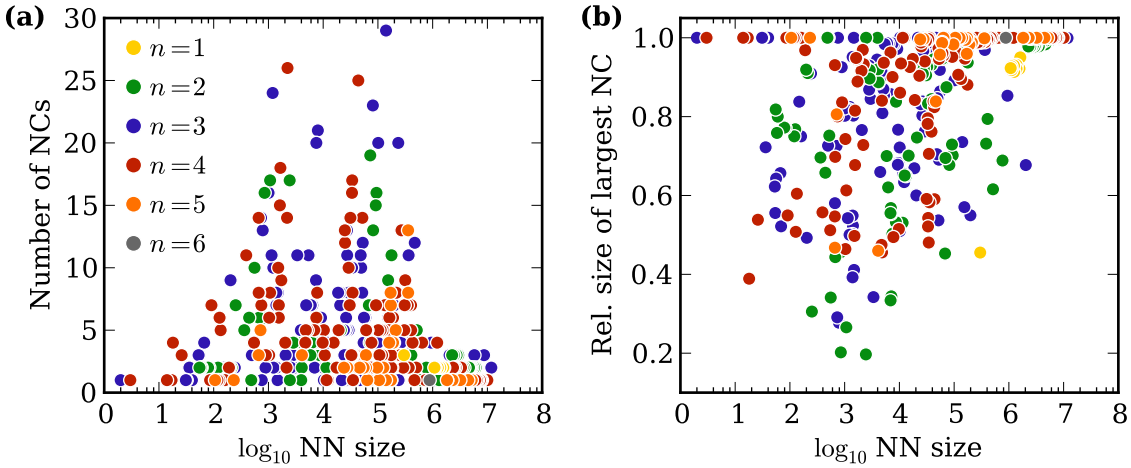

Figure 3.9: RNA neutral spaces are largely connected under base-pair swaps. The diagram shows analogous results to Fig. 3.6, except that here, we allow for base-pairs swaps (eg. GC $\leftrightarrow$ CG) as single mutational moves. **(a)** shows the number of NCs per NN, colours indicate the number n of base-pairs in the corresponding phenotype. Note that in comparison to Fig. 3.6a, the number of NCs is dramatically reduced. **(b)** shows the relative size of the largest NC for each NN. Again, in comparison to Fig. 3.6b, we note that the fraction of the NC covered by the largest component is typically much larger when we allow for base-pair swaps.

which some of p's base-pairs are replaced by unpaired bases (for example, all genotypes compatible with `(((....)))`..... are evidently compatible with `((......))`.....). Still,

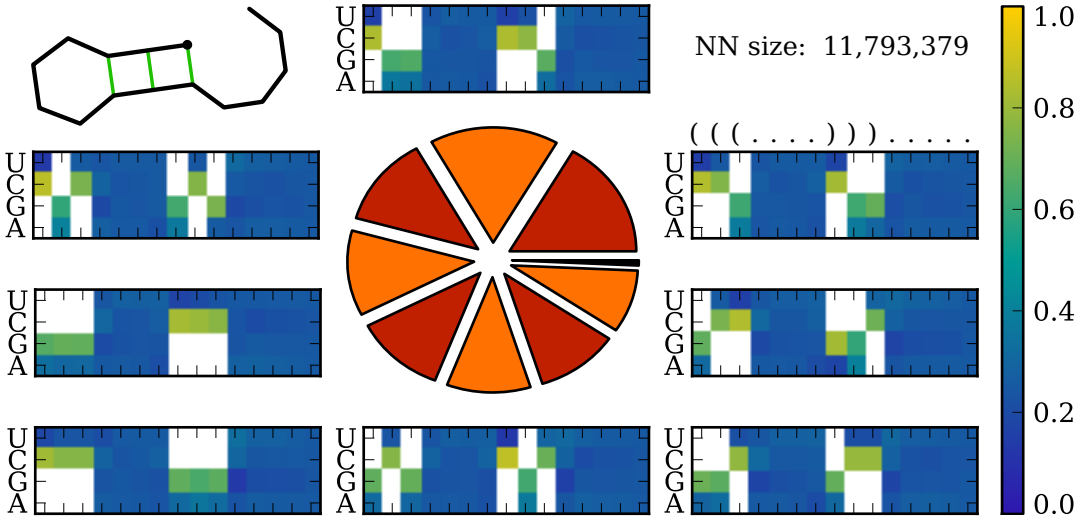

Figure 3.10: Different NCs correspond to different base-pair configurations. The most frequent structure for $L = 15$, shown at the top, has 3 base-pairs. The relative sizes of the $2^3 = 8$ largest NCs are indicated by the pie chart; the small wedge shows the size of the remaining 8 NCs combined. For each large NC, the relative proportion of each nucleotide is encoded by the colour of the respective square. A white square indicates that the nucleotide does not occur at all at this position.

we expect that sufficiently large NNs are dominated by 2^n NCs of roughly similar size. The size of the largest NC should thus be on the order of 2^{-n} – indeed, this is what we observe in Fig. 3.6b for the phenotypes of high frequency. Further, we expect that the 2^n largest NCs of these phenotypes correspond to different base-pair arrangements. In Fig. 3.10, we check these arguments for the most frequent (non-trivial) phenotype at $L = 15$, which has 3 base-pairs. In line with our expectations, the NN is dominated by $2^3 = 8$ NCs, each of which has a particular composition of the base-pairs.

Finally, we turn our attention to the rare phenotypes, which constitute the overwhelming majority of phenotypes. Because of their low frequency, they fall below even the reduced percolation threshold Δ_n. Nonetheless, their neutral spaces are far from completely fragmented. As Fig. 3.11 underlines, even the smallest NNs contain far fewer NCs than genotypes: Both rare and frequent phenotypes have NNs with a much higher degree of connectivity than expected from the random GP map.

In summary, we have seen that genotypes mapping into a particular phenotype are not randomly distributed in genotype space. We have exploited the simplicity of RNA folding to understand how epistasis arises from the biophysical processes underlying the GP map,

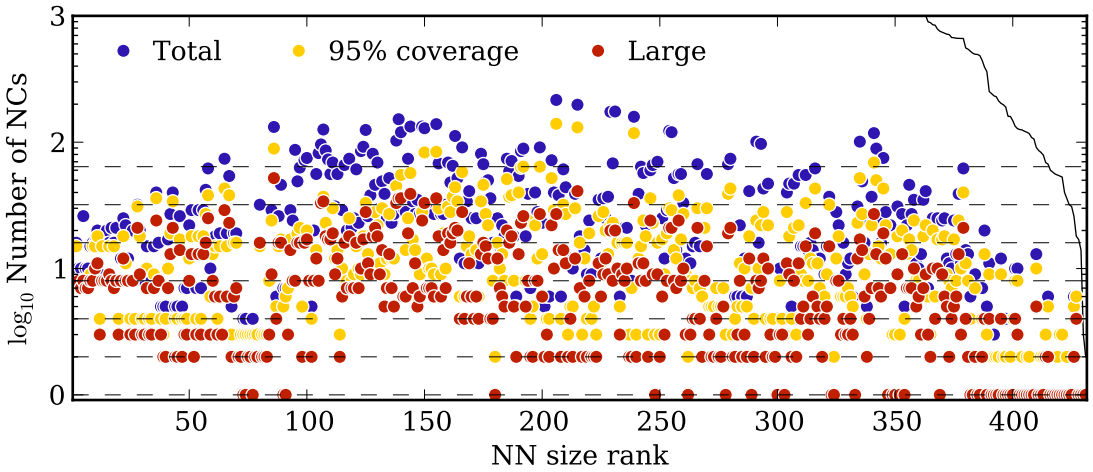

Figure 3.11: Even small neutral spaces are partially connected. The diagram shows the degree of fragmentation against NN rank (as before, rank 1 corresponds to the largest non-trivial NN). The total number of NCs is shown in blue. Yellow dots indicate the number of NCs that together cover 95% of the NN, while the number of large NCs (greater than average size *within their NN*) are drawn in red. The dashed horizontal lines are the integers powers of 2. The solid line gives the upper bound on the number of NCs, namely the NN size. The fact that this line is largely outside of the diagrams limits underlines that NNs show a relatively high degree of connectivity.

and how reciprocal sign epistasis can lead to the fragmentation of neutral spaces. We shall now turn our attention to the evolutionary implications of neutral space structure.

3.3.4 Properties of neutral components and evolutionary implications

Connected neutral spaces allow populations to explore large portions genotype space without fitness penalties [126]. But when neutral spaces are fragmented, not all genotypes with the same phenotype can be reached from each other by neutral mutations and the scope of neutral exploration may be limited (cf. Fig. 3.1). If phenotypic changes are strongly deleterious (ie. the current phenotype constitutes a (local) peak of the fitness landscape) and double mutations are very rare, genetic drift will cause an evolving population to be localized to a single NC.

Does this localization matter? In order to answer this question, we need to compare the properties of the NCs in the same neutral space. After all, we are assuming that only the phenotype (that is, the NN) determines fitness, not the NC. So the effect of natural selection

is to decide whether a population makes a transition to *any* part of an NN, but selection is indifferent to *which* part is chosen.

Here, we shall discuss the mutational robustness (cf. Eq. (2.20)), potential evolvability (cf. Eq. (2.23)) and diversity evolvability (cf. Eq. (2.25)) of individual NCs. To this end, we need to modify the definitions given above so that they consider only the genotypes in the NC of interest, not the entire NN. If we denote the set of genotypes in NC κ by $\mathcal{G}_\kappa$ and let q be the phenotype to whose NN κ belongs, we have

$$\rho_\kappa = \frac{1}{|\mathcal{G}_\kappa|} \sum_{g \in \mathcal{G}_\kappa} \rho_g \tag{3.13}$$

$$\mathcal{E}_\kappa = \left| \bigcup_{g \in \mathcal{G}_\kappa} P_g \right| \tag{3.14}$$

$$\epsilon_\kappa = 1 - \sum_{p \neq q} c_{p\kappa}^2 \tag{3.15}$$

where ρ_g is the robustness of genotype g (Eq. (2.19), P_g is the set of phenotypes accessible by single mutations from genotype g (as in Eq. (2.21)) and the $c_{p\kappa}$ are defined in analogy to the c_{pq} in Eq. (2.24):

$$c_{p\kappa} = \frac{n_{p\kappa}}{\sum_{a \neq q} n_{a\kappa}} \tag{3.16}$$

where $n_{p\kappa}$ is the total number of mutations from κ to phenotype p. The motivation for restricting our attention to a single NC derives from the idea that the scope of neutral exploration may typically be limited to a single NC, while the rest of the NN remains inaccessible to the population (at least for very long times, since crossing a fitness valley is typically slow compared to neutral exploration [121]).

From Fig. 3.12 we see that larger NCs generally show greater robustness. This trend holds on the global scale, where we ignore what phenotype an NC corresponds to (Fig. 3.12a), as well as within a particular NN (Fig. 3.12b). As the NC sizes in one NN can show large variations, we see that we should *not* think of robustness as a property of the NN – that is, the phenotype – but rather as a property of the particular NC.

Nonetheless, we can of course calculate the robustness ρ_p (Eq. (2.20)) of a particular

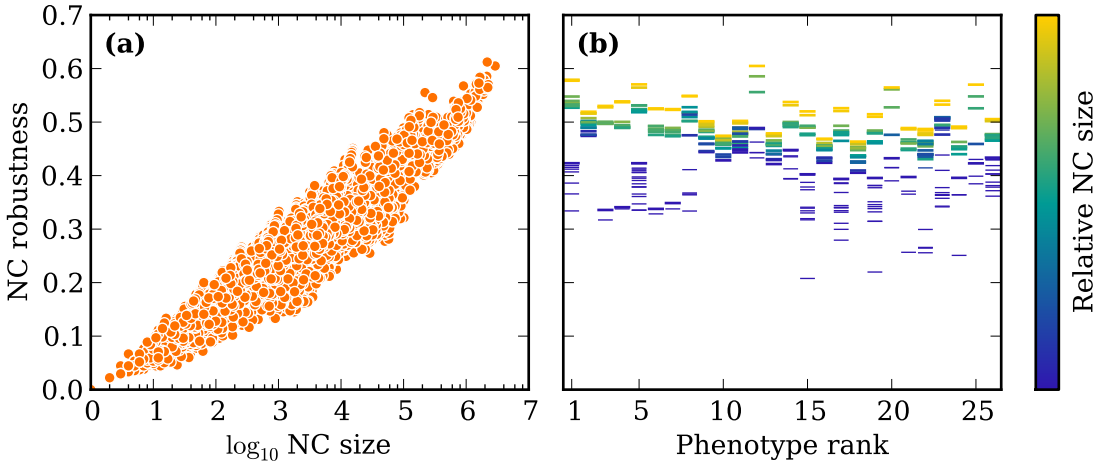

Figure 3.12: Robustness increases with NC size. **(a)** Larger NCs show a greater mean robustness (calculated using Eq. (3.13)). The Pearson correlation with log-size is $r = 0.964, p < 10^{-15}$. **(b)** Consequently, the larger NCs in each NN tend to have greater robustness. The data is for the 26 largest NCs only; each vertical stack of lines corresponds to one NN. Thick lines indicate the largest NCs that together cover 95% of the NN; thin lines correspond to the remaining NCs.

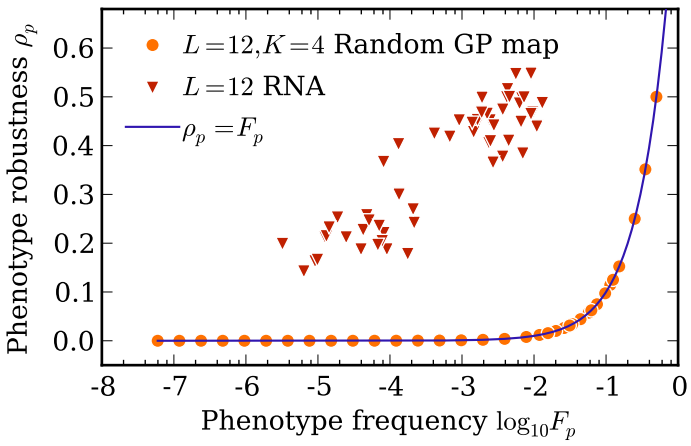

Figure 3.13: Robustness in RNA is much higher than in the random GP map. The diagram shows the dependence of the phenotype robustness ρ_p (Eq. (2.20)) on the phenotype frequency F_p (Eq. (3.1)) for the random GP map (orange circles) and RNA at $L = 12$ (red triangles; the trivial structure is not shown). The agreement of the random GP map simulations (errorbars are much smaller than the markers and omitted for clarity) and the prediction $\rho_p = F_p$ (blue line) is excellent.

phenotype p. The comparison ρ_p and the phenotype frequency F_p (Eq. (3.1)) underlines the non-randomness of the RNA GP map: If genotypes are assigned randomly to phenotypes, we expect (and observe to a high accuracy) $\rho_p = F_p$: Whether or not a given genotype maps into p does not depend on the phenotypes of g's neighbours. By contrast, ρ_p is much larger

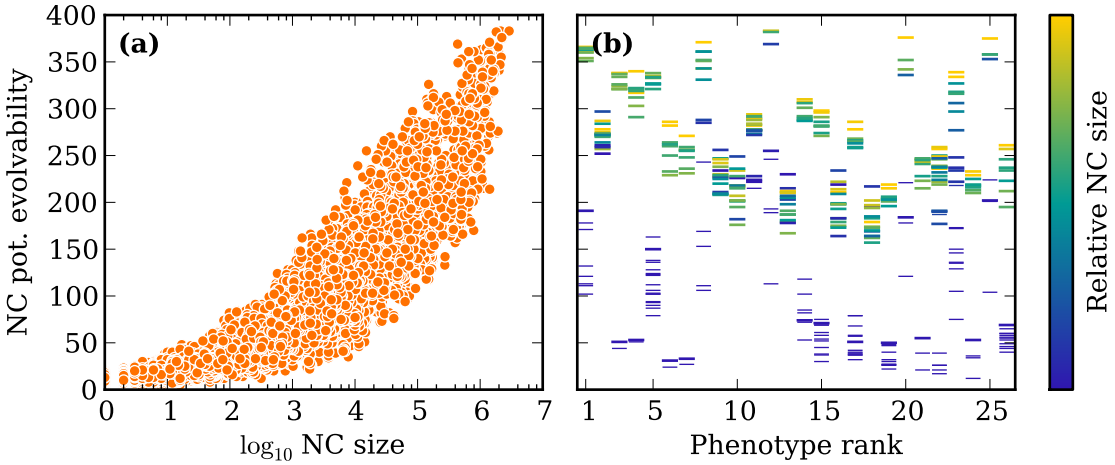

Figure 3.14: Potential evolvability increases with NC size. **(a)** Larger NCs typically allows access to a greater variety of alternative phenotypes. The correlation of $\mathcal{E}_\kappa$ (Eq. (3.14)) with log-size is $r = 0.877, p < 10^{-15}$. **(b)** Within one NN, the size-evolvability order is not always strict. The data is for the 26 largest NCs only; each vertical stack of lines corresponds to one NN. Thick lines indicate the largest NCs that together cover 95% of the NN; thin lines correspond to the remaining NCs. In both panels, the black dashed line marks the total number of phenotypes.

than F_p in RNA, and it appears that ρ_p increases only logarithmically with F_p (Fig. 3.13).

Just as robustness, the potential evolvability of an NC generally increases with its size (Fig. 3.14a). Compared to NC robustness (Fig. 3.12a), the increase is less stringent: NCs spanning 2-3 orders of magnitude in size can have very similar potential evolvability. Consequently, the largest NCs in one NN need not necessarily be the ones with largest evolvability. Note that no NC has $\mathcal{E}_\kappa = N_P - 1$: There are always some phenotypes which are inaccessible from a given NC.

Both robustness ρ_κ and potential evolvability $\mathcal{E}_\kappa$ thus increase (more or less strictly) with NC size, and hence ρ_κ and $\mathcal{E}_\kappa$ are positively correlated themselves: The Pearson correlation is $r = 0.81$ and the Spearman rank-correlation is $k = 0.86$, with p-values less than 10^{-15} in both cases. This result based on exhaustive enumeration confirms a similar result by Wagner, which was based on random samples of RNA sequences [127].

Potential evolvability measures how many phenotypes can be accessed, but not which ones. In light of the fragmentation of NNs, it is interesting to also study the identity of the accessible phenotypes. In particular, are the same phenotypes accessible from all NCs in the NN, or does each NC allow access to a different set of phenotypes? In order to address this

question, we define the joint and common (potential) evolvability for each NN as [114]:

$$\mathcal{E}_p^{(j)} = \left| \bigcup_{\kappa \in p} P_\kappa \right| \tag{3.17}$$

$$\mathcal{E}_p^{(c)} = \left| \bigcap_{\kappa \in p} P_\kappa \right| \tag{3.18}$$

where $\kappa \in p$ means all the NCs in the NN of p, and P_κ is the set of phenotypes accessible from NC κ. Joint evolvability thus counts the number of phenotypes accessible from *at least one* NC, while common evolvability is the number of phenotypes accessible from *all* NCs in a given NN.

To quantify the distinction of joint and common evolvability, we also calculate their ratio, averaged over the NNs. We contrast the unweighted average

$$R^{(u)} = \frac{1}{N_\mathcal{P}} \sum_p \frac{\mathcal{E}_p^{(c)}}{\mathcal{E}_p^{(j)}} \tag{3.19}$$

and the weighted average, in which the contribution of each NN is weighted by its size

$$R^{(w)} = \sum_p F_p \frac{\mathcal{E}_p^{(c)}}{\mathcal{E}_p^{(j)}} \tag{3.20}$$

From the definitions, it is clear that $\mathcal{E}_p^{(c)} \leq \mathcal{E}_p^{(j)}$. In particular, $\mathcal{E}_p^{(c)}$ is bounded from above by the least evolvable NC, while $\mathcal{E}_p^{(j)}$ is bounded from below by the most evolvable NC. Given the large variation of NC size (and hence evolvability) in many NNs, it is not surprising that we only see the equality of joint and common evolvability for the few NNs that are fully connected (Fig. 3.15a). However, the degree to which joint and common evolvability are different is quite striking: Only about 1 in 7 phenotypes in reach from at least one NC is also accessible from all NCs ($R^{(u)} = 0.14$). It is interesting to note that this result is largely independent of NN size ($R^{(w)} \approx R^{(u)}$).

This discrepancy between joint and common evolvability cannot be attributed entirely to the effect of the smallest NCs: For Fig. 3.15b, we calculate joint and common evolvability only from the large NCs, ie. those that are greater than average in their NN. This restriction

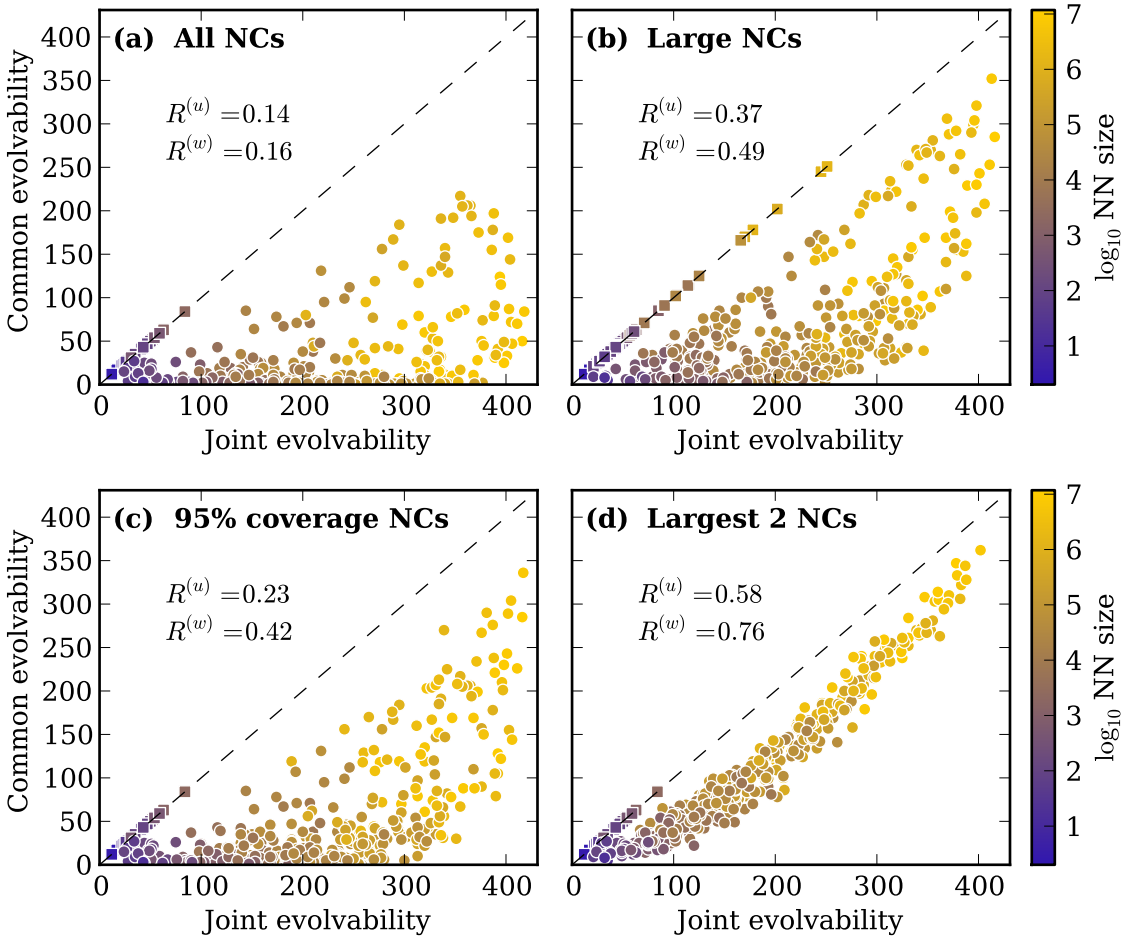

Figure 3.15: Different NCs in the same NN give access to different phenotypes. The diagrams show the joint and common evolvability for the NNs of $L = 15$ RNA secondary structures, calculated according to Eqs. (3.17) and (3.18) respectively. In addition, the mean and size-weighted averages (Eqs. (3.19) and (3.20)) are given. The different panels correspond to various sets of NCs that are used in the calculations: **(a)** All NCs in the NN are used. **(b)** Only the large NCs (greater than average size) are taken into account. **(c)** The largest NCs that together contain 95% of their NN are used. **(d)** Only the largest 2 NCs are compared. In all cases, square markers indicate the NNs with only a single (relevant) NC. These NCs trivially have $\mathcal{E}_p^{(j)} = \mathcal{E}_p^{(c)}$, as indicated by the black dashed lines. Colours indicate NN sizes.

does increase the common evolvability (on average by about 43 phenotypes). On the other hand, we also note a reduction in joint evolvability (on average by about 12 phenotypes): Some alternative phenotypes may only be accessible from the smaller NCs. Together, these effects lead to an increase in the average ratio of common to joint evolvability to $R^{(u)} = 0.37$. We also note that the spectra of accessible phenotypes vary less among the large NCs of frequent phenotypes ($R^{(w)} = 0.49$); yet still half of the accessible phenotypes can only be reached from some of the large NCs.

Other restrictions on the set of NCs that are used to calculate the joint and common evolvability have qualitatively similar effects: In Fig. 3.15c, we use the largest NCs that together cover 95% of their NN. Again, there is a clear difference between joint and common evolvability, which is less pronounced for the large NNs ($R^{(u)} = 0.23$, $R^{(w)} = 0.42$). Finally in Fig. 3.15d, we only use the 2 largest NCs of each NN (unless of course the NN is fully connected), but we still find that joint and common evolvability can be quite different: Overall, just over one half of the accessible phenotypes can be reached from both NCs ($R^{(u)} = 0.58$). Again, this value is higher for larger NNs, tending to about 3 in 4 phenotypes ($R^{(u)} = 0.76$).

Instead of considering only a subset of NCs, we can also narrow down the range of alternative phenotypes to include when calculating evolvability. A natural subset of phenotypes to consider are the frequent phenotypes, which have greater than average NNs. Fig. 3.16 compares joint and common evolvability with this restriction. As there are only 86 frequent phenotypes, joint and common evolvability are reduced for most phenotypes. Comparing the average ratios $R^{(u)}$ and $R^{(w)}$, we find clear evidence that the discrepancy between joint and common evolvability is smaller when we consider only frequent alternative phenotypes. On other hand, we note that it is not possible to access all frequent phenotypes from most phenotypes ($\mathcal{E}_p^{(j)} < 86$). And just as before, when we compare the sets of accessible phenotypes for only a subset of NCs in each NN, we find that the difference between joint and common evolvability decreases.

Counting the number of accessible phenotypes is not the only way to measure evolvability. We shall now turn to the diversity evolvability ϵ_κ (Eq. (3.15)) proposed by Cowperthwaite et al. [24], which gives the probability that two randomly chosen non-neutral mutations away from NC κ produce two distinct phenotypes (cf. Section 2.2.5).

We can measure diversity evolvability for individual NCs as well as for entire NNs. Fig. 3.17 shows the results of these different approaches. In agreement with the results of Cowperthwaite et al. (who studied $L = 12$), we find that the diversity evolvability ϵ_p (Eq. (2.25)) of the NN corresponding to phenotype p decreases with its size (and hence its robustness ρ_p). However, this decay is quite weak (Fig. 3.17a), and when considering indi-

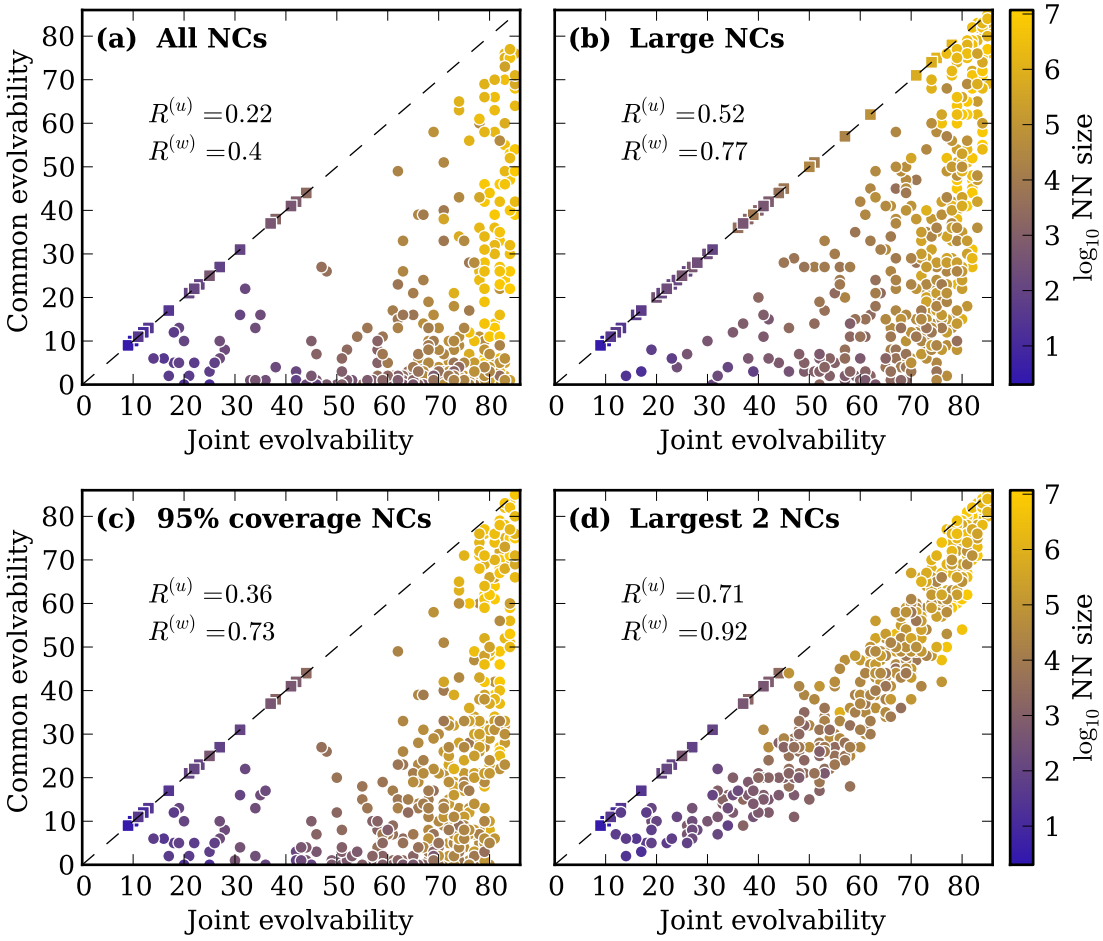

Figure 3.16: Joint and common evolvability towards frequent phenotypes. This diagram show the results analogous to Fig. 3.15; the difference is that here, joint and common evolvability are calculated only from those alternative phenotypes that are frequent (ie. their NN has greater than average size).

vidual NCs, we actually observe a small increase of diversity evolvability ϵ_κ (Eq. (3.15)) and NC size. Consequently, we cannot conclude with confidence at this stage whether or not robustness generally increases or decreases diversity evolvability.

In contrast to robustness, where the NN robustness is simply the (size-weighted) average of the NC robustness values, the non-linearity in the definition of ϵ_p (Eq. (2.25)) means that the relation of NC and NN diversity evolvability may be more complicated. When we average over the diversity evolvability of the NCs in an NN, we find that this average is usually smaller than the diversity evolvability of the NN; only for 28 out of the 431 NNs, the NC average is slightly greater (Fig. 3.18a). If we weight the contribution of each NC by its size, we find that the NN value is almost always greater than this weighted average (except

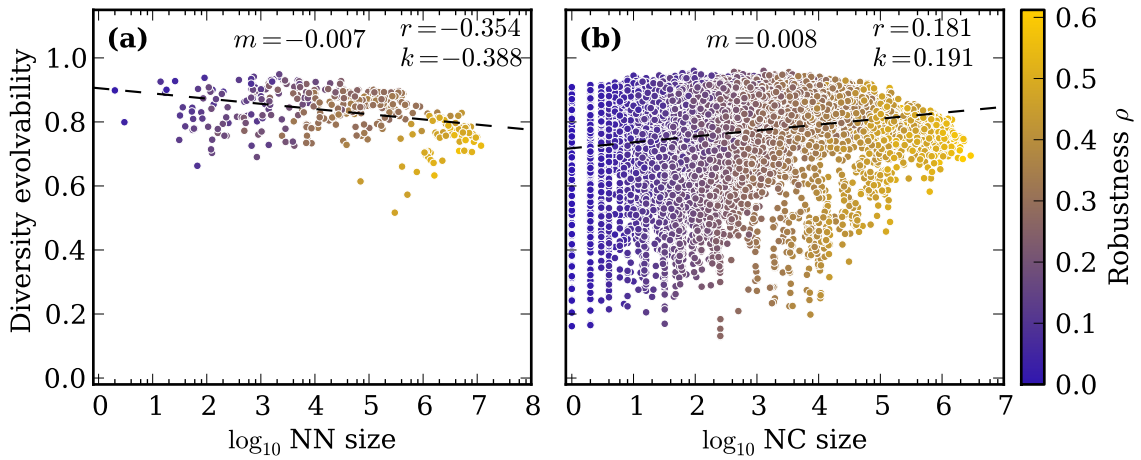

Figure 3.17: Diversity evolvability varies between neutral spaces, but shows no clear dependence on their size. **(a)** Diversity evolvability for entire NNs. **(b)** Diversity evolvability for individual NCs. The data is for $L = 15$ RNA. The black dashed lines are linear least-squares fits (with log-transformed size), the slope of the line is m. r is the correlation between log-size and diversity evolvability; k is the rank correlation coefficient of size and diversity evolvability.

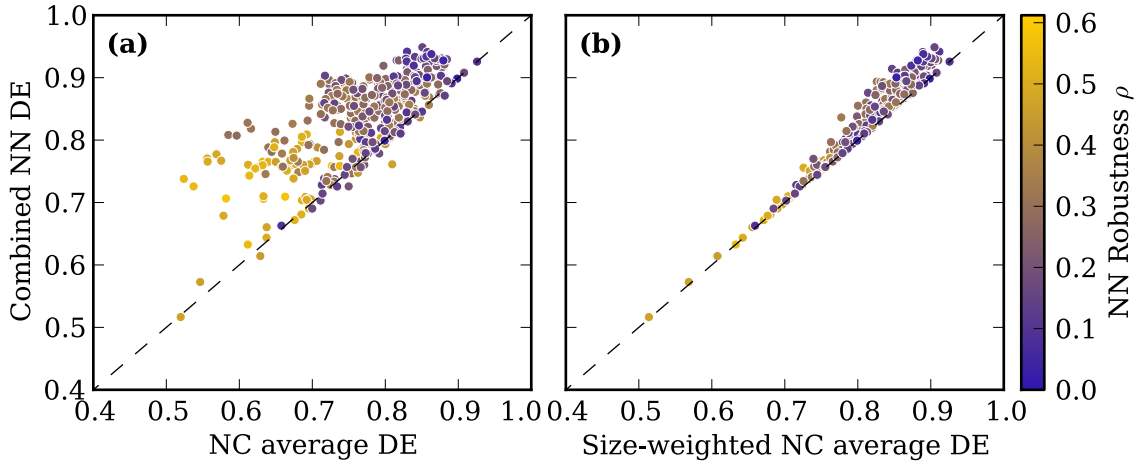

Figure 3.18: Diversity evolvability and neutral space fragmentation. The diagrams compare the diversity evolvability (DE) of NNs and their constituent NCs. The black dashed lines indicate equality of the NN value and the respective NC average. **(a)** The DE of the NN is plotted against the mean DE of its NCs. For 28 (out of 431) NNs, the NC average is greater than the NN value. **(b)** Here, the NN's DE is compared to the NC size-weighted average of the DEs. This average is greater than the NN value for 2 phenotypes.

for 2 NNs, see Fig. 3.18b).

This discrepancy again highlights the heterogeneity among the NCs – if all NCs were the same, there would of course be no distinction between NN values and NC averages (weighted or not). The fact that the diversity evolvability of the entire NN is usually greater than the NC average underlines our results above: Two mutations off different NCs tend to produce

more different phenotypes than two mutations off a single NC, even when we account for the fact that some phenotypes are easier to produce than others.

In summary, we find that the spectra of accessible alternative phenotypes can be very heterogeneous among the different NCs of the same NN. This result suggests a source of contingency in evolution: When selection can only steer population towards particular phenotypes (NNs), it may not be able to distinguish between the NCs. But the further potential for evolutionary innovations, which is constrained in principle by potential evolvability [129] and in practice also by diversity evolvability, is determined by the random fluctuations which cause the population to choose one particular NC over another.

3.3.5 Dependence on sequence length

So far, we have analysed the neutral spaces for RNA secondary structures of $L = 15$ nucleotides. Already for these small molecules, we find vast neutral spaces. A complete characterization of the GP map of longer sequences is therefore not practical. Instead, we have two options two investigate the L-dependence of our findings: Either we can exhaustively study shorter sequences, or we resort to a sampling approach. In this section, we follow both routes. While the exact enumeration for smaller L is conceptually straightforward, the sampling approach is slightly more involved.

By definition, a NC is a maximal set of neutrally connected genotypes. So given a single genotype of the NC, we can find all others by consecutively tracing out the available neutral mutations, until the NC is complete (cf. Appendix B). This approach guarantees that we obtain the entire NC. But how do we make sure to find all NCs? With limited computing resources, it is not feasible to achieve any such guarantee. And since genotype space grows exponentially fast, it is very easy to push any system to its limits. Yet, if we have an estimate of the total NN size, we can combine the sizes of the components we have found to see if any major NCs are likely to be missing.

For $L = 20$ results we discuss here, it is at least possible to work out phenotype frequencies by exact enumeration (the results are given in Fig. 3.4). For larger L, even this may no longer

be feasible – the $L = 20$ enumeration and folding required about 1 year of CPU time on present-day hardware. However, Jörg et al. have devised a sampling algorithm that can be used to determine NN sizes approximately [68]; the results of this procedure are quite reliable (see [68] and Table 3.3 below) and can be used as a guide when an exact enumeration approach fails.

So while we can check how good our coverage of the NN is, we still need a method to generate starting genotypes on different NCs. Here, the base-pairing logic of RNA plays out to our advantage: We have found consistently that the majority of a reasonably large NN is contained within 2^n NCs, where n as before is the number of base-pairs in the secondary structure to which the NN corresponds. So if we generate starting sequences for each of these 2^n NCs, we should be able to find the NCs of significant sizes. To create these sequences, we can employ an inverse folding routine of the Vienna package [64]: For this function, we can specify a starting sequence in which bases can be fixed, and the program then attempts to change the remaining nucleotides in such a way that the resulting sequence folds into the desired structure. If we feed this method with starting genotypes from the different compatible subsets of the sequence, we have a good change of obtaining starting sequences in the large NCs. As the compatible subsets are necessarily disjoint, the search for NCs can easily be run in parallel over the sets.

Table 3.3 shows the results of our sampling for 19 different structures at $L = 20$. Clearly, our sampling approach is quite successful: Exploiting the base-pairing logic, we can find genotypes on all major NCs. For sufficiently large neutral spaces, we can thereby recover the entire NN. For some of the less frequent phenotypes with more intricate structure, a small fraction of genotypes is not found. It appears that this issue may be resolved by creating a larger number of starting genotypes, as we can indeed find completely isolated genotypes (we used 1000 inversion attempts for each compatible subset).

For the large NNs, we find that the number of NCs is often exactly given by our expectation 2^n: In each subspace of compatible bases, the genotypes actually realizing the particular phenotype are connected. Only in some cases (rank 41 and 124), there are a few additional NCs, which are still within one order of magnitude of the largest NC. Nonetheless,

Structure	Rank	Exact size	Estimated size	Sampled size	Missing	NCs	C_n	Largest NC	Smallest NC
((((....))))....	3	7,401,668,230	$7.36 \times 10^9 \pm 7.35 \times 10^7$	7,401,668,230	0	16	1.0	559,226,190	364,799,345
.((((....))))....	9	5,459,848,003	$5.45 \times 10^9 \pm 6.31 \times 10^7$	5,459,848,003	0	16	1.0	425,643,201	264,979,132
..((((....))))....	15	4,770,625,062	$4.76 \times 10^9 \pm 6.84 \times 10^7$	4,770,625,062	0	16	1.0	374,094,656	217,865,887
(((((....)))))....	22	4,364,744,209	$4.42 \times 10^9 \pm 2.87 \times 10^7$	4,364,744,209	0	32	1.0	159,842,554	111,617,821
...((((....))))....	28	4,130,904,339	$4.18 \times 10^9 \pm 7.51 \times 10^7$	4,130,904,339	0	16	1.0	329,880,695	179,479,560
....((((....))))....	33	3,958,204,779	$3.92 \times 10^9 \pm 6.16 \times 10^7$	3,958,204,779	0	16	1.0	317,392,598	168,531,600
......(((((....)))))	41	3,638,723,916	$3.64 \times 10^9 \pm 2.49 \times 10^7$	3,638,723,916	0	40	0.92	146,033,386	34,806,893
...(((((....)))))....	86	2,465,108,974	$2.41 \times 10^9 \pm 2.10 \times 10^7$	2,465,108,974	0	32	1.0	91,074,611	51,390,638
((((((....))))))....	112	1,969,084,252	$1.97 \times 10^9 \pm 1.05 \times 10^7$	1,969,084,252	0	64	1.0	35,290,626	23,203,475
....((((((....))))))	124	1,715,071,950	$1.69 \times 10^9 \pm 8.01 \times 10^6$	1,715,071,950	0	80	0.91	34,044,285	7,989,916
..((((((....))))))..	165	1,274,645,721	$1.26 \times 10^9 \pm 9.55 \times 10^6$	1,274,645,721	0	64	1.0	22,422,835	14,591,824
((((..((....))..))))	403	272,339,190	$2.67 \times 10^8 \pm 3.20 \times 10^6$	272,338,729	461	115	0.93	6,469,383	6
..((((((....)).)))).	408	269,729,528	$2.81 \times 10^8 \pm 4.67 \times 10^6$	269,729,528	0	117	1.0	6,710,238	12
((....))..(((....)))	1278	44,469,174	$4.92 \times 10^7 \pm 2.33 \times 10^6$	44,468,912	262	138	0.92	3,150,436	1
((.((.((....)).)).))	1346	39,407,608	$3.39 \times 10^7 \pm 8.15 \times 10^5$	39,406,792	816	166	0.91	1,107,225	1
(((....)))..((....))	1696	22,488,827	$2.62 \times 10^7 \pm 2.12 \times 10^6$	22,488,072	755	150	0.87	1,260,330	10
((....))....((....))	2206	11,156,755	$1.08 \times 10^7 \pm 8.35 \times 10^5$	11,150,519	6,236	84	0.85	1,324,269	103
..((((((....)).)).))	2957	4,699,054	$4.96 \times 10^6 \pm 3.35 \times 10^5$	4,698,425	629	546	0.79	238,852	2
((.((.((....)))).))	4902	646,305	$7.28 \times 10^5 \pm 1.11 \times 10^5$	644,268	2,037	750	0.82	37,776	1

Table 3.3: Details of the sampled structures at $L = 20$. For each structure, the table shows its representation in the dot-bracket notation and its global rank (the trivial structure has rank 1, just as in Fig. 3.4), along with the exact, estimated (mean ± standard error on the mean for 10 independent sampling runs, using the algorithm in [68]) and sampled size; the latter is just the sum over the sizes of the discovered NCs. The column 'Missing' refers to the difference of the exact size and the sampled size. C_n is the fraction of the NN contained in the 2^n largest NCs. The size of the largest and smallest NC are also given.

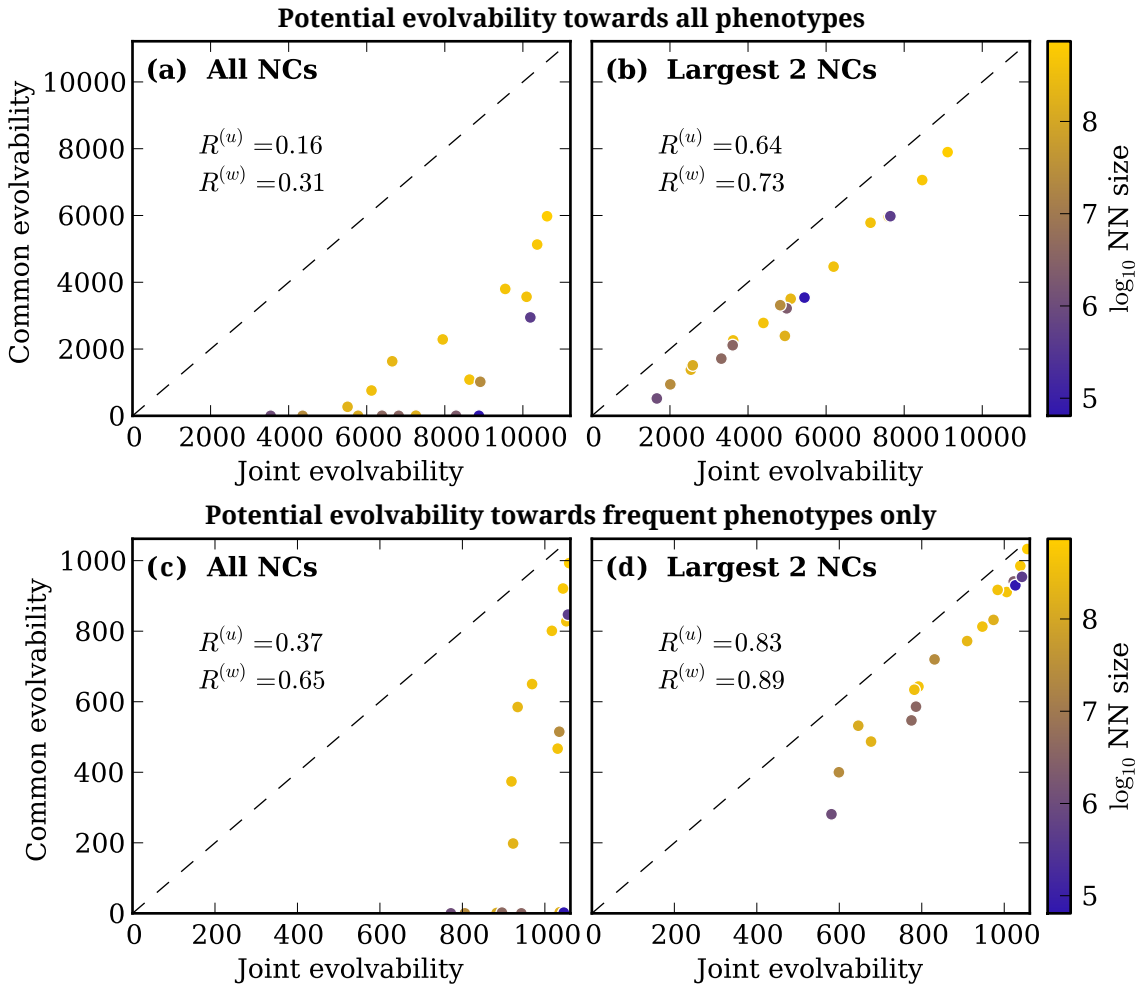

Figure 3.19: The spectra of mutants phenotypes differ strongly at $L = 20$. The diagram contrasts joint and common evolvability of the $L = 20$ NNs given in Table 3.3. In the top row, all alternative phenotypes are considered in the calculation, while the bottom row accounts only for the frequent phenotypes. **(a)** and **(c)** show results when all (discovered) NCs are taken into account, while in **(b)** and **(d)** only the largest 2 NC have been included. $R^{(u)}$ and $R^{(w)}$ are the unweighted and size-weighted averages of the ratio of common to joint evolvability, as defined in Eqs. (3.19) and (3.20).

the need for compensatory mutations clearly appears to be the main reason that drives the fragmentation of large NNs in RNA.

Let us now turn our attention to the properties of the neutral spaces at $L = 20$. Given the limited number of NNs for which we have data, it is clear that the following results should not be viewed as conclusive evidence, but rather to be indicative of trends.

Fig. 3.19 shows the joint and common evolvability calculated for the NNs we sampled. In interpreting the results for incompletely sampled NNs, we note that a subset of NCs bounds the actual $\mathcal{E}_p^{(j)}$ (Eq. (3.17)) from below, and the actual $\mathcal{E}_p^{(c)}$ (Eq. (3.18)) from above. A more

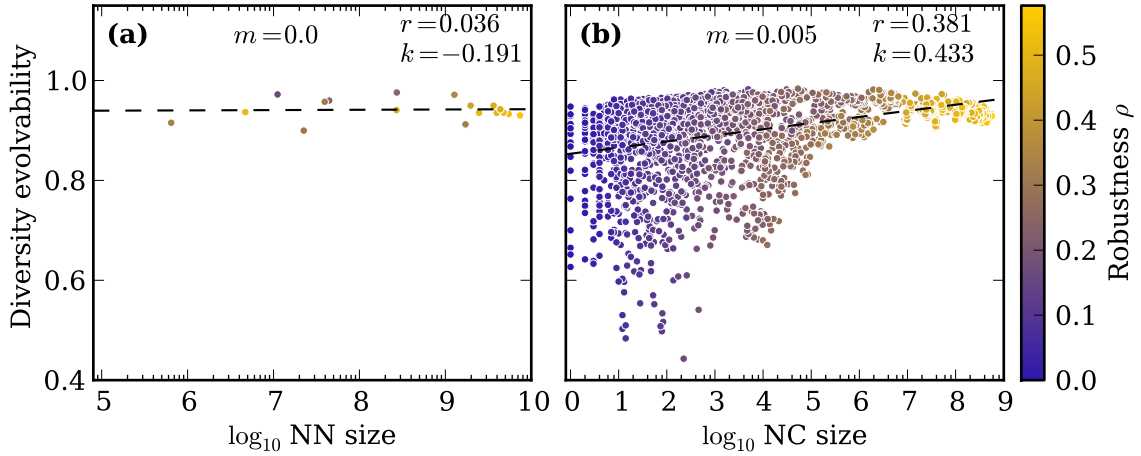

Figure 3.20: Diversity evolvability is large at $L = 20$. The diagram is analogous to Fig. 3.17: **(a)** shows the diversity evolvability (DE; Eq. (2.25)) of our $L = 20$ NNs against their size; r is the correlation coefficient of DE and log-size, k is the Spearman rank correlation of DE and size, and m is the slope of the linear fit DE against log-size, given to 3 digits of accuracy. **(b)** shows the analogous results for individual NCs, where DE is calculated according to Eq. (3.15).

complete picture of the NN can thus only increase the difference between joint and common evolvability. Even with incompletely discovered NNs, it is clear that the discrepancy between joint and common evolvability remains strong at $L = 20$. In fact, we even observe that for the phenotypes ranked 403rd, 1346th and 2206th, the common evolvability among all NCs vanishes: No non-trivial phenotype can be reached from all the NCs. As we have seen in our $L = 15$ data, restricting the source NCs and/or target phenotypes under consideration, the discrepancy between joint and common evolvability reduces, but clearly does not vanish completely.

Let us now turn to diversity evolvability ϵ. As we noted above, this measure of evolvability is dominated by the contribution of the most easily accessible phenotype. For the short sequences we can study, this phenotype is usually the trivial structure; but as its frequency decreases with L, we would expect that its impact on ϵ decreases, and hence that ϵ itself should grow. As Fig. 3.20 shows, our limited data at $L = 20$ confirms this intuition. The minimum diversity evolvability we find for NNs is 0.90, while at $L = 15$ this is only 0.52; the minima for NC diversity evolvability are 0.44 at $L = 20$ and 0.13 at $L = 15$. While larger NCs tend to have slightly greater diversity evolvability ϵ_κ (Eq. (3.15)), we do not find a clear dependence of ϵ_p (Eq. (2.25)) on NN size.

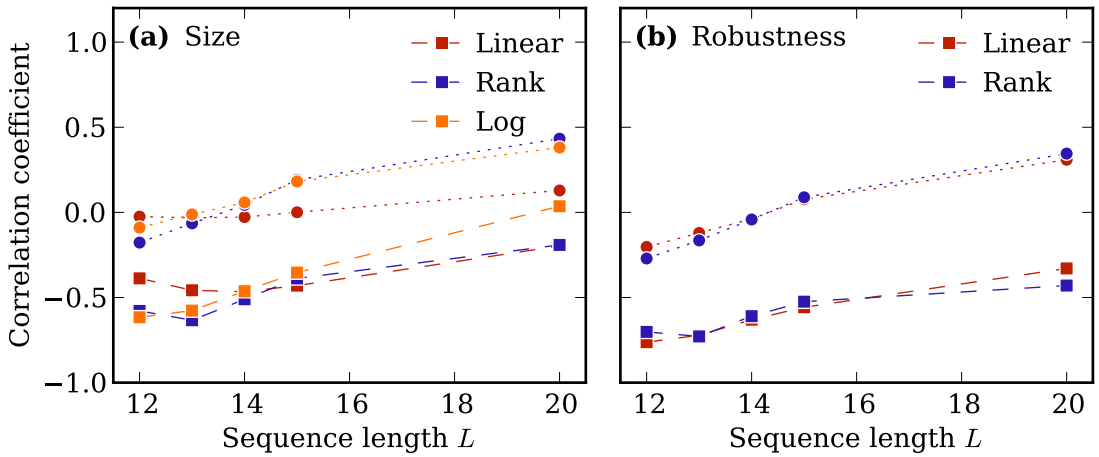

Figure 3.21: The dependence of diversity evolvability on sequence length. The diagram compares the correlation coefficients of diversity evolvability and **(a)** neutral space size, or **(b)** neutral space robustness. Square marker show the values calculated at the NN scale, circles correspond to the correlation at the level of individual NCs. Red markers show the Pearson correlation, blue marker indicate the Spearman rank correlation. In (a), the orange markers show the correlation of diversity evolvability and log-size.

When dealing with our $L = 15$ data, we could not make a clear case for a positive or negative correlation of neutral space size (or robustness) and diversity evolvability. In Fig. 3.21, we show how various correlation coefficients change with sequence length. Again, we stress that without a full characterization of the GP map, we should be cautious to draw strong conclusions. With the data at hand, it appears that for longer sequences, the correlation coefficients increase. Interestingly, while the correlation of diversity evolvability and NC size/robustness looks clearly positive, the case is not at all clear for the NNs. Here, the correlation coefficients get less negative as L grows. But whether this trend saturates at a positive or negative correlation is unclear.

In summary, our sampling results at $L = 20$ confirm our earlier findings: Reciprocal sign epistasis is the major source of neutral space fragmentation. The resulting NCs have very heterogeneous spectra of accessible alternative phenotypes. Finally, we note that diversity evolvability seems to level off close to unity as L increases. This suggests, at least in RNA, that smaller values diversity evolvability are largely an artifact of the extreme frequency of the trivial structure for short sequences. We conclude our discussion of the properties of RNA neutral spaces with the study of a different mutational move in genotype space:

Recombination or cross-over.

3.4 Cross-over mutations

So far, we have only considered point mutations to connect adjacent genotypes. In nature, other mutational moves are also observed (cf. Section 2.2): Deletions or insertions correspond to omitting or adding bases during reproduction of the genotype. These mutations change the sequence length which makes them hard to discuss in our framework.

There is one other kind of mutation that we can address, namely cross-over. This mutational move is particularly relevant to sexually reproducing organisms in which the offspring receives part of its genotype from each parent[1]. It is interesting to ask whether cross-over mutations can produce genotypes that belong to an NC that is different from the parental NCs, that is whether cross-over can connect different NCs.

The simplest case to study is when both parental genotypes belong to the same NC. Let us focus on a single base pair for simplicity – if any single pair is not maintained in the offspring phenotype, that phenotype is necessarily different from the parental one. For definiteness, we assume that the base pair of interest is made by a purine-pyrimidine pair in both parents, say GC in one and AU in the other[2]. Evidently, crossover between the parents will again result in a purine-pyrimidine pair. Depending on the parental base pairs, this could be GC, GU, AU or AC. The first three pairs are compatible and may thus lead to the parental phenotype. However, they will again be part of the same NC – a transition into a pyrimidine-purine pair (which would necessarily be part of a different NC) is not possible. Finally, an AC pair is incompatible and will lead to a different phenotype.

The case of parental genotypes from different NCs is slightly more involved. Let us again consider a purine-pyrimidine pair (say, GC) in the first parent, but now a pyrimidine-purine pair (e.g. CG) in the other parent. We now need to distinguish two cases depending on the

[1] Asexual organisms such as bacteria can also mutate under cross-over by horizontal gene transfer [77].

[2] Having a GC and a GU pair in the same NC is likely to occur only in frequent structures. For rare phenotypes, only GC may exist; this is not important to the argument.

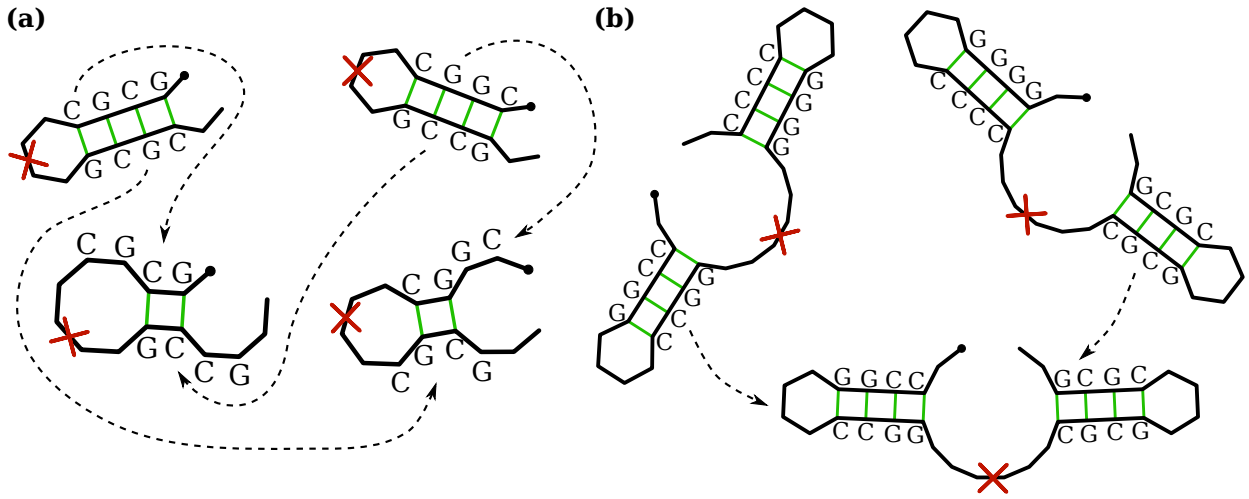

Figure 3.22: The impact of crossover mutations on neutral space connectivity. The diagram shows two scenarios of crossover where the two parental genotypes belong to different NCs. The point of crossover is marked by the red cross. **(a)** When crossover joins incompatible bases, the phenotype necessarily changes. **(b)** A crossover mutation can lead to an NC that is distinct from both parental NCs when mismatches are kept separate.

point of crossover. First, consider the case that the crossover point is within the stem-loop region enclosed by the base pair of interest (illustrated in Fig. 3.22a). This means that the resulting genotypes will have either a purine-purine (GG) or a pyrimidine-pyrimidine (CC) pair, neither of which can form a bond. Thus the offspring phenotype is necessarily different from the parent. Second, the point of crossover may be outside the stem-loop region of interest (Fig. 3.22b). In that case, the base pair is left intact. If the parental phenotype contains two separate stem-loop regions, this scenario of crossover may lead to a new NC: Individually incompatible stems are collectively 'shuffled'. Yet this way of exploring different NCs is limited to reusing the already existing stems; new variants of individual stems cannot be achieved in this manner.

In summary, crossover alone cannot lead a population to new NCs. In order to create genotypes on an NC that is different from the parental NCs, it is necessary to cross genotypes from different NCs in special positions. In other words, crossover mutations can lead to new areas of genotype space only if genetic variation is already present in the population. It is worth noting that under cross-over, a population takes longer to fix a base-pair change than under point mutations alone [25, 75].

3.5 Neutral spaces in other GP maps

In RNA, we can explicitly track the effects of epistatic interactions on the structure of neutral spaces. The base-pairing logic clearly motivates the necessity of compensatory mutations and thus of reciprocal sign epistasis, which causes the fragmentation of neutral spaces in RNA. Here, we study the NN structure of other GP map models.

3.5.1 The genetic code

We have already encountered the genetic code as an example GP map in Section 2.2.3. This is a particularly interesting system, because it is quite likely that the GP map itself may have been subject to evolution [47, 59, 124, 141, 142]. This conclusion has been reached by measuring several properties of the universal code and comparing them to ensembles of alternative codes. Examples include the change in chemical properties of the amino acid upon mutation [47, 59, 142] and the ability to include other information (e.g. the secondary structure of mRNA) in the DNA sequence encoding a particular protein [66].

It is interesting therefore to note that almost all amino acids have completely connected neutral spaces; only serine (Ser) has two NCs, one of size 4 and another of size 2 (see Fig. 2.2). Under the random GP map (where codons are assigned randomly to amino acids, maintaining the amino acid frequencies of the universal code), such well-connected NNs are extremely unlikely [114]. Whether the structure of NNs in the universal code is the product of direct selection for connected neutral spaces, comes as a by-product of selection for mutational robustness, or arises due to biochemical constraints of mRNA to protein translation is presently unclear. Attempting to decide between these (and other) options would give us a better understanding of the role of neutral spaces in evolution.

3.5.2 HP protein folding

Apart from RNA, one of the best studied model GP maps is HP protein folding [7, 14, 15, 82, 83]. Real proteins often fold into relatively compact 3-dimensional structures, with hydrophobic amino acids in the core and polar ones on the outside [98]. The HP model

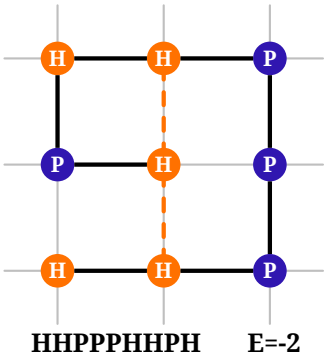

Figure 3.23: Illustration of the HP model. The HP model is a coarse-grained description of protein folding. Proteins are represented as sequences of hydrophobic (H, blue) and polar (P, red) residues. Folds are compact shapes on a regular lattice. The energy of a given sequence depends on the number of H-H neighbours in the fold. In this example, we have a 2-dimensional square lattice and a protein of length 9. The backbone of the protein is drawn in black, indicating the fold. There are two H-H neighbours in the fold which do not arise from adjacency in the sequence, indicated by the red dashed lines. In units of of the H-H bond energy, the given fold thus has an energy of -2.

provides a coarse-grained description of this system: Genotypes are sequences of amino acid residues which are usually described only in terms of the chemical polarity as hydrophobic (H) or polar (P). Phenotypes are compact structures on a regular lattice; sequences 'fold' into the structure that minimizes an energy function. In the simplest setting, the energy is given by the number of H-H neighbours in the structure which are not adjacent in the genotype; see Fig. 3.23 for an example.

While various modifications of the HP model have been studied, the general results are consistent with the simple model [14]: Neutral networks of HP proteins are often completely connected or dominated by a giant component. Reciprocal sign epistasis does not play a role in the HP model: As only H-H interactions lower the energy, swapping the order of a pair is an ill-defined concept in this system. And if we want to go from a non-bonding H-P to a P-H, this is always possible via P-P.

These results in the HP model are consistent with the general trend we have seen in RNA: Unless there is reciprocal sign epistasis in the GP map, neutral spaces tend to be largely connected.

3.5.3 Transcription factor binding sites

Another instructive system to study is the interaction of proteins and DNA. These interactions are important in the regulation of gene expression. Some proteins predominantly act as *transcription factors* (TFs): They bind to specific regulatory regions of DNA and thus either enhance or promote the expression of genes that are physically close to the binding region [79].

TFs and their target sites in the genome co-evolve. Usually, the rate of evolution (as measured by mutations being fixed) is much smaller in transcription factors than in their binding sites [119]. Over short time-scales, we can thus assume that only the binding sites evolve.

For a fixed TF, the interaction with the binding site can be described quite reliably in terms of a simple energy model, where the nucleotides in the binding site contribute additively to the overall binding energy [79]. The binding energy $E(g)$ of genotype $g = (b_1, \ldots, b_L)$, where b_i is the nucleotide base at the ith position along the sequence and L is the length of the binding site, is thus given by

$$E(g) = \sum_{i=1}^{L} E_i(b_i) \tag{3.21}$$

where E_i determines the contribution of the ith site. This energy gives an indication of how tightly a given TF binds to a particular site. Following the convention Mustonen et al.[88], the E_i are normalized such that each has a minimum of 0, and E is measured in units of the chemical potential of the bound state. A simple coarse-graining can be achieved by introduced a cutoff E' such that a factors bind only when $E \leq E'$ [88]. Thus we have two phenotypes, binding and non-binding.

We can see straightaway that the neutral space of the genotypes that allow binding of a particular TF is completely connected. The key idea is that of a 'master' genotype g_0, which is the one of the lowest binding energy which is $E(g_0) = 0$ by convention. Now let us consider another genotype g_1 that also allows binding of the TF, ie. $E(g_1) \leq E'$. Evidently, all the mutations that reduce the Hamming distance from g_1 to g_0 do not increase the binding

energy, so that $E(g) \leq E(g_1) \leq E'$ where g is a genotype closer to g_0 in genotype space than g_1. Any path from g_1 that reduces the Hamming distance to g_0 at each step is thus neutral. So all binding genotypes are neutrally connected to g_0, and thus the neutral space is fully connected.

The mutational paths towards the master genotype indicate the absence of reciprocal sign epistasis: A mutation that reduces the distance to the master genotype cannot increase the binding energy. On the other hand, sign epistasis is quite likely to exist in this system. To see this, we consider two distinct binding genotypes g, g' at Hamming distance d from g_0. As the genotypes are distinct, there must be at least one position where g and g_0 are identical but g' is not, and vice versa. So g and g' share at least two neighbours g_+, g_- at Hamming distance $d \pm 1$ from g_0. Clearly, g_- is also binding, but g_+ need not be, depending on the energies involved. If only $g \leftrightarrow g_- \leftrightarrow g'$ is a neutral path but $g \leftrightarrow g_+ \leftrightarrow g'$ is not, we have an instance of sign epistasis (cf. Fig. 2.4).

The structure of neutral spaces becomes more interesting if we impose another condition on the binding energy, namely that it should not be too large (in absolute value). The motivation of this requirement is that TFs should not only be able to bind, but also to unbind (this is similar to the notion of marginal stability in proteins [30]). Otherwise a single molecule of the TF may control the expression of its target gene indefinitely long, and the ability to actually manipulate gene expression is severely hampered. Thus we require for a *functional* binding site g that $E^* \leq E(g) \leq E'$, where E^* causes sites with too strong binding to be discarded from the set of functional sites. When $E^* > 0$, the master genotype is no longer functional, and our argument above no longer applies. Depending on the band width $E' - E^*$, the neutral space of functional binding sites may thus no longer be connected: The enforcement of intermediate binding energies can lead to reciprocal sign epistasis.

Mustonen and co-workers analysed high-throughput data for two TFs in yeast [88]. The authors suggest $E' = 0.9$ to ensure that a random genotype is typically not functional [88]. Using their data for one of the TFs called *abf1*, we investigate the structure of the NN of functional sites as a function of the energy cutoffs E^* and E'. As Fig. 3.24a shows, E' predominantly determines the number of functional genotypes; provided the band of

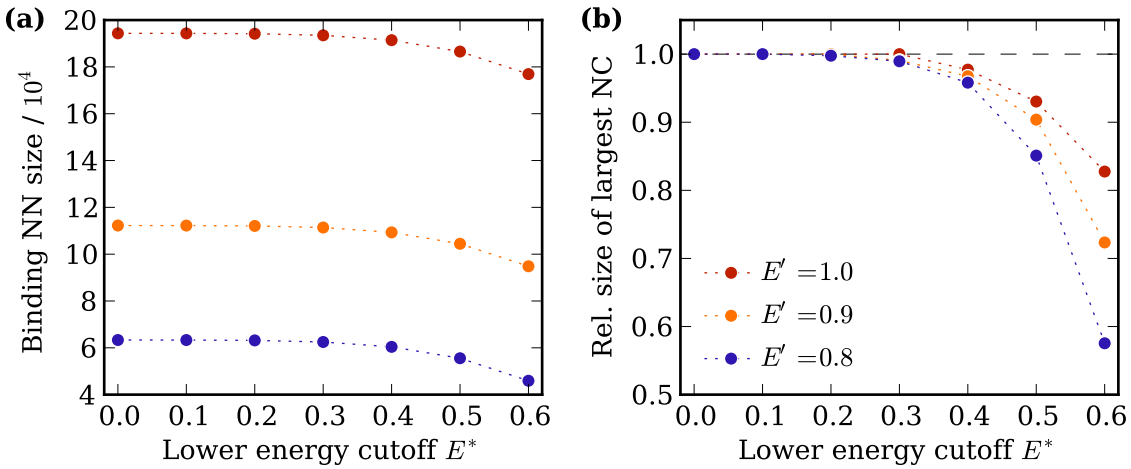

Figure 3.24: Neutral networks of functional binding sites are largely connected. **(a)** The size of the NN of functional genotypes depends largely on the upper energy bound E', but decreases only slowly with the energy band width $E' - E^*$. **(b)** As the energy band width decreases, the NN of functional genotypes starts to fragment. However, even for the narrow band $E' - E^* = 0.2$, the majority of the NN is still contained in a single NC. The results have been obtained using the energy data for the TF *abf1* in yeast [88].

functional binding energies is not too narrow, E^* has little effect on the size of the functional NN. For small E^*, we find that the NN is completely connected (Fig. 3.24b), while at some point (depending on E'), a fragmentation of the NN sets in. The effect of this fragmentation is relatively weak: the majority of functional genotypes can still be accessed from each other by neutral mutations. We note that these results are for binding sites of length $L = 14$, giving a total genotype space size of $4^{14} \approx 2.7 \times 10^8$; under the random GP map, we would therefore expect that the NN of the functional genotypes was completely fragmented.

Whether or not this results pertain for other TFs with different binding energies, is an open question. More generally, it is important to ask whether the simple binding energy model of Eq. (3.21) is genuinely valid. It is clearly the additive nature of the model that allowed our argument about neutral paths to the master genotype. This question can only be answered with more in-depth studies of the interactions of TFs and their binding sites.

3.5.4 Gene regulatory networks

The setting of one TF and a single binding site forms the basis of gene regulatory networks (GRNs). In these networks, each node corresponds to a gene, and a (usually directed) edge

from gene A to gene B indicates that the protein produced by expressing A regulates the expression of B. The study of GRNs is one of the central topics of systems biology [5, 8].

Models of GRNs can be defined at various levels of resolution. When the aim is to make quantitative predictions for a particular experiment, a common approach is to employ systems of coupled differential equations. These models contain large numbers of parameters, for example the binding energies we discussed above, but also rates of transcription, translation and degradation of proteins. The entire set of parameters is usually viewed as the genotype, and the behaviour of the GRN over time defines a phenotype.

Due to vast (and usually continuous) parameter space, an exhaustive characterization of the systems in not feasible (or even impossible in principle). Yet, it appears that an analogue of neutral spaces exists in these highly parametrized models. Work by Sethna and various collaborators lead to the identification of *sloppiness* in models with many parameters [57, 58, 120, 132]. Broadly speaking, a model is called sloppy if changes in some parameter combinations (the so-called stiff modes) have much greater effects on the behaviour of the model than in most other parameter combinations (the sloppy modes). Changes along sloppy modes are thus likely to be effectively neutral, leading to a notion of neutral spaces [26]. However, little is known about the structure of these spaces.

To facilitate the study of neutral spaces of GRNs, several approximations can be made. A common model is boolean threshold dynamics [19, 20], where the state of each gene is either 1 (the gene is being expressed) or 0 (not expressed). The states of all genes are updated in synchrony in discrete time steps by multiplying the current state vector into the adjacency matrix describing the GRN, and thresholding the resulting vector. One approach to view this system as a GP map considers the adjacency matrix as the genotype and the time series of the activity of some (or all) genes as the phenotype. Ciliberti and co-workers demonstrate that under this definition, neutral networks tend to be fully connected or dominated by a giant NC [19].

A rather different result was reported by Boldhaus and Klemm [12] who studied a simplified description of the yeast cell-cycle [81]. In their study, the phenotype is the pattern of activity of 11 genes over 13 time steps. By exploiting the structure of the model [80],

they could demonstrate the existence of about 4.7×10^8 NCs of sizes between 6.1×10^{24} and 4.4×10^{26}.

All these results for neutral space structure in higher-level GP maps come with an important caveat (which is often not even mentioned): As we have stressed throughout, neutral spaces are induced by the GP map on the mutational network between genotypes. Fundamentally, all genotypes are transmitted as DNA sequences, and hence mutations affect these sequences. In GRN models, mutations are usually assumed to occur at a higher level, changing parameters (in the case of ODE models) or elements of the GRN adjacency matrix; whether single mutations can have just these effects is presently not clear, and hence the results concerning the structure of neutral spaces in these systems must be viewed with care.

3.5.5 Antibiotic resistance

So far, we have only discussed computational studies of neutral spaces in various GP maps. The main experimental problem of similar studies is simply the problem of scale: If we cannot exploit simple biophysical reasons to suppose that a particular neutral space must be fully connected or fragmented, an exhaustive characterization of the entire GP map is necessary. Performing such exhaustive approaches experimentally is generally prohibitive.

However, progress is under way. In a important study, Weinreich and co-workers characterized a small part of the GP map associated with resistance against a particular antibiotic in *E. coli* [136]. By five mutations in a single protein, the bacteria can increase their resistance by a factor of about 10^5; resistance is measured in terms of the minimum inhibitory concentration (MIC) of the antibiotic that is required to prevent the bacteria from growing. Weinreich et al. measured the MIC for all $2^5 = 32$ combinations of genotypes with any of the five mutations present or absent.

We analysed the landscape and found that several neutral spaces are not fully connected but consist of several NCs (Fig. 3.25). This result comes with two qualifications. First, resistance was measured on a rather coarse scale: starting from a base concentration of antibiotic, the concentration was increased by a factor of about 1.4 between subsequent

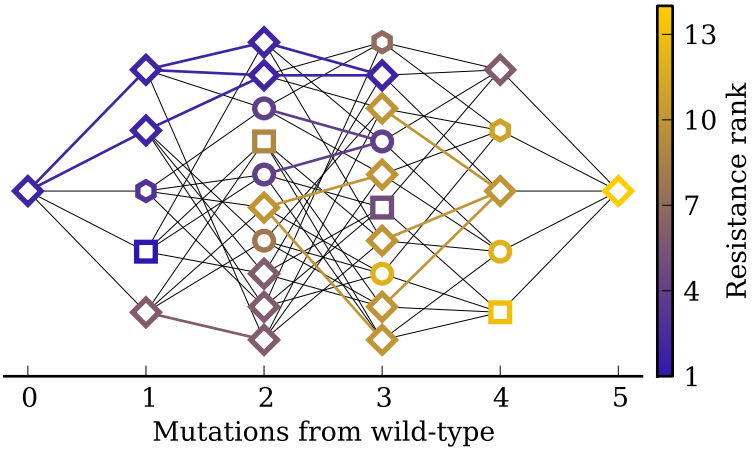

Figure 3.25: Fragmented neutral spaces in an experimental system. The diagram shows the mutational landscapes constructed by Weinreich et al. [136]: Each vertex corresponds to a genotype, and markers and colours indicate phenotypes. In this case, the phenotype is the minimum inhibitory concentration (see Text for details), ranked 1 for the wild-type and 14 for the maximally resistant strain. Solid black lines show non-neutral mutations; coloured lines mark neutral mutations. Several neutral spaces are not fully connected. Genotypes are aranged by their Hamming distance from the wild-type.

experiments; genotypes that appear neutral on this scale may be non-neutral on a finer resolution. Second, and more importantly, it may be that neutral mutations outside the scope of the experiment connect different NCs. Nonetheless, it is interesting to find this fragmentation of neutral spaces in experimental data.

3.6 Summary and discussion

In this chapter, we have investigated the structure of neutral spaces that arise under different GP maps.

We first studied the random GP map, where the number of genotypes for each phenotype is fixed and the genotypes are distributed randomly in genotype space. In agreement with analytic results [108, 109], we found evidence of two percolation transitions as functions of phenotype frequency (that is, the fraction of genotypes mapping into that phenotype). The main result is that only excessively frequent phenotypes have largely connected neutral spaces. By inverting the percolation thresholds given in Eqs. (3.3) and (3.2), we see that the total number of neutral spaces comprising more than a handful of connected genotypes

is severely limited.

The majority of this chapter was dedicated to the study of neutral spaces of the RNA secondary structure GP map [64], which is one of the central workhorse models in the field [23, 42]. We saw an explicit example of strong phenotypic bias: Most phenotypes have very small frequencies, while a minority of frequent phenotypes covers almost all of genotype space (Fig. 3.4). It is interesting to note that the most frequent phenotypes appear quite regular or 'simple' (Fig. 3.5). By comparing results for different sequence lengths L, we provided clear evidence for the exponential increase in the absolute number of frequent phenotypes with L; at the same time, the proportion of frequent phenotypes becomes exponentially small.

We identified the base-pairing logic of RNA folding as a source of (phenotypic) reciprocal sign epistasis (cf. Fig. 2.4). Consequently, we cannot expect fully connected neutral networks in RNA. However, we have clearly seen that RNA neutral spaces are much more connected than the random GP map would suggest. The number of neutral spaces of appreciable size is large and appears to grow exponentially with sequence length, while the random GP map would only allow for a linear growth.

Due to the biophysics of the RNA GP map, there are many disjoint neutral components (NCs) in each neutral network. By investigating their properties, we demonstrated explicitly the positive correlation of neutral space size and robustness. We also showed that potential evolvability increases with neutral space size, thereby confirming Wagner's sampling results [127]. In light of the fragmentation of neutral networks, we compared the spectra of accessible phenotypes among the NCs with the same phenotype. Here we found a striking heterogeneity that may cause contingency in evolution: When selection can steer populations only towards a particular phenotype, but is blind to the choice of NC, then the potential for future innovations is determined by the fluctuations of mutation and possibly also by memory effects [114].

In addition to Wagner's potential evolvability, we also studied the diversity evolvability proposed by Cowperthwaite et al. [24]. For individual NCs, we could not confirm their results of a decrease in diversity evolvability with neutral space size; even for entire neutral networks, there does not appear a strong, or even particularly clear, trend. Instead, our

results suggest that diversity evolvability in RNA is very sensitive to finite size effects. Nonetheless, we could use the discrepancy between NC-averaged and NN-based diversity evolvability to further underline the heterogeneity of NCs.

Finally we went beyond RNA and reviewed other GP maps studied in the literature. The HP model of protein folding shows no reciprocal sign epistasis and typically has (almost) completely connected neutral spaces [14]. For a simple model describing the binding of transcription factors to regulatory sequences [79, 88], we found that reciprocal sign epistasis can be enforced by requiring an intermediate binding strength, leading to a fragmentation of the neutral space of functional sites. For the transcription factor we studied, the fragmentation typically produces a giant NC and a few very small NCs.

For higher level systems such as gene regulatory networks, the picture is more complex. Owing to different choices in the definition of neutrality, largely connected [19] as well as strongly fragmented [12] neutral networks have been reported. We have argued that such results should be taken with caution: In these models, genotypes are typically defined at a more abstract level than DNA or RNA sequences, and hence the effect biologically reasonable mutations is far from clear. So it is presently unclear whether or not evolving populations can traverse the neutral 'spaces' defined in these high-level sytems.

One general point emerges from all the models we have discussed in this chapter: *Genotypes mapping into a particular phenotype are typically not distributed randomly in genotype space.* Instead, these genotypes tend to be close to each other. Consequently, neutral spaces are much more connected than the random GP map suggests.

In the next chapter, we will move beyond the static description of neutral spaces. After all, evolution is a dynamic process, and we would like to study the interplay of neutral space structure and evolutionary dynamics. We will take the two key ingredients of our model from the observations in this chapter: First, neutral spaces tend to be connected and large. Second, some phenotypes occur much more frequently than others. As we will see in the next two chapters, in particular this last point has profound consequences for the introduction of variation into evolving populations.

Chapter 4

A mean-field model of evolutionary dynamics

4.1 Introduction

The previous chapter gave a glimpse of the complexity that characterizes even simple GP maps like RNA. In this chapter, we will study how the static structure of neutral spaces impacts the dynamics of evolving populations.

In RNA, we have seen strong phenotypic bias (Fig. 3.4): A few phenotypes are realized by vast proportions of genotype space, while most other phenotypes are rare. Similar results have been found in many other model GP maps: In the genetic code, where genotypes are nucleotide triplets and phenotypes are amino acids (cf. Fig. 2.2); in the HP-model of protein folding (cf. Section 3.5.2), where genotypes are sequences of hydrophobic and polar amino acids, and phenotypes are compact arrangements of these sequences on a lattice (see Fig. 3.23) [14, 82, 83]; in the mapping of protein sequences to structures and functions [40]; in the polyomino model of self-assembly, which has been applied to study aspects of the formation of protein quaternary structure [3], where genotypes specify different building blocks and phenotypes are the shapes into which these blocks assemble [67]; in a model of gene regulatory networks [13]; in a signalling circuit, where genotypes are circuit topologies and phenotypes are activity profiles [106]; and in a model of neural development, where

genotypes are DNA sequences and phenotypes are neural connection patterns [104]. A thorough review of phenotypic bias in GP maps is in preparation [32].

Natural selection can only act on the variation that is present in the population. As the introduction of variation is mediated by the GP map, phenotypic bias is quite likely to have an impact on the variation that evolving populations produce, and hence on the course of evolution. The model presented in this chapter attempts to describe systematically how evolutionary dynamics are affected by phenotypic bias, or the GP map more generally. To this end, we explicitly include the GP map into the Wright-Fisher model.

Our central conceptual advance is to distinguish between mutations of genotypes and the outcome of mutations at the level of phenotypes. The phenotypic effects of mutations are quantified in terms of *phenotype production rates* (PPRs) ν_p which give for each phenotype p the probability that it arises in a particular reproduction event. In general, PPRs depend both on the state of the population and on the GP map.

For analytical insight, we perform several approximations that allow us to neglect some of the complexity of evolutionary dynamics under genetic drift that we encountered in Section 2.3.2. The motivation for these approximations is our focus on the introduction of phenotypic variation: We are not interested primarily in the dynamics of neutral exploration, but rather how this exploration impacts the discovery of novel phenotypes.

Therefore, we consider only two limiting cases of exploration dynamics: First, we study extremely polymorphic populations which are spread out widely over their neutral space. Such populations can access a large variety of alternative phenotypes without the necessity of neutral exploration. For reasons that will become clear in Section 4.2.2, we call this limit the continuity approximation. Second, we investigate the opposite limit of monomorphic populations which are localized to a single genotype at any one time. As we have seen in Section 2.3.2, the dynamics of neutral exploration in monomorphic populations can be described in terms of a random walk on the neutral space. In both limits, we treat the GP map in a mean-field approximation and find that the first discovery time T_p for a phenotype p is inversely proportional its connectivity c_{pq} (Eq. (2.24)) where q is the current phenotype of the population. Thus, (locally) rare phenotypes are hard to discover.

The inverse proportionality of T_p with c_{pq} emerges in both dynamic limits. This result suggests that genetic drift, whose strength determines whether populations are monomorphic or polymorphic regimes, primarily affects the constant of proportionality. So when phenotypic bias is strong (such that the c_{pq} vary over orders of magnitude), connectivity appears to be the main determinant of discovery times. Through extensive numerical simulations under the random GP map with phenotypic bias, we find a good agreement of this prediction with the simulation results; in a population between the monomorphic and polymorphic extremes, a single fitting factor appears sufficient to capture the effect of genetic drift. Phenotypic bias can thus provide an a posteriori justification for our simplistic treatment of neutral exploration dynamics.

Our microscopic approach integrates the population dynamic parameters N (population size) and μ (mutation rate per base per reproduction) with the GP map details represented by L (genome length), ρ_q (mutational robustness of the population's phenotype, Eq. (2.20)) and the set of c_{pq} (phenotype connectivities, Eq. (2.24)). The value of our approximate treatment stems from the simplicity of its results: For example, we show that the ratio N/L of population size and genome length is an important measure of the role of fluctuations. Furthermore, we can re-derive the non-monotonic scaling of the discovery time T_p of an alternative phenotype p on the robustness ρ_q of the current phenotype; in contrast to the rather involved calculations by Draghi et al. [33], our model gives an intuitive argument: More robust neutral spaces allow for faster exploration, but higher robustness reduces the probability that mutations produce novel phenotypes.

4.2 Theoretical description

In this section, we study the most basic question regarding the introduction of phenotypic variation: How long does it take for an evolving population to discover a particular phenotype? Given the large skew in phenotype frequencies, we anticipate that the discovery time can be very different for phenotypes that differ widely in frequency.

4.2.1 General exposition

We study the evolution of a population of N individuals under Wright-Fisher dynamics (cf. Section 2.1.2). Each individual i carries a string of length L over an alphabet of K letters as its genotype. The phenotype p_i of the individual is determined via the GP map, and fitness f_i is determined from the phenotype. In each discrete generation, we draw N parents from the population with replacement for reproduction; the probability to select individual i is $f_i / \sum_j f_j$ per reproduction event. The offspring receives a copy of the parental genotype; each letter in this copy is mutated independently with probability μ to a different letter. Thus the number of mutations d (equal to the Hamming distance between parent and offspring) is distributed binomially.

Since we are interested in neutral exploration, we assign fitness to phenotypes as $f(p) = \delta_{pq}$ where q is the 'source' phenotype. Pictorially, we place the population on the peak of the fitness landscape. At $t = 0$, we assume that the environment changes and another phenotypes becomes more fit, and our first aim is to calculate the time when this alternative phenotype arises for the first time in the population.

We can write down the probability $\nu_p(t)$ that a reproduction event at time t leads to an individual with phenotype p as

$$\nu_p(t) = \sum_{d=1}^{L} \binom{L}{d} \mu^d (1-\mu)^{L-d} \phi_p(d,t) \tag{4.1}$$

where the *local frequency* $\phi_p(d,t)$ is the probability that a d-mutation leads to p at time t. In general, $\phi_p(d,t)$ depends on the GP map and the current state of the population. We call ν_p the *phenotype production rate* (PPR) of p.

If we know $\phi_p(d,t)$, we can easily calculate several quantities of interest. The number of p-mutants $m_p(t)$ produced at time t follows a binomial distribution:

$$\Pr(m_p(t) = m) = \binom{N}{m} \nu_p(t)^m (1 - \nu_p(t))^{N-m} \tag{4.2}$$

Approximating the binomial by a Poisson distribution, the probability that p is *not* produced

in the current generation is simply $\exp(-N\nu_p(t))$. Over T generations, the expected total number M_p of p-mutants is

$$M_p(T) = N \sum_{t=1}^{T} \nu_p(t) \tag{4.3}$$

Thus the probability $\alpha_p(T)$ of discovering p within T generations is

$$\alpha_p(T) = 1 - \exp\Big(- N \sum_{t=1}^{T} \nu_p(t) \Big) \tag{4.4}$$

This equation can be inverted to obtain an implicit equation for the time $T_p(\alpha)$ at which p has been discovered with probability α (so $\alpha = 1/2$ corresponds to the median discovery time):

$$\sum_{t=1}^{T_p(\alpha)} \nu_p(t) = \frac{-\log(1-\alpha)}{N} \tag{4.5}$$

But how do we obtain the local frequencies $\phi_p(d,t)$? Ideally, we would like to write down some sort of differential equation that incorporates details of the GP map and governs the time-dependent dynamics of the population. Unfortunately, such results are hard to come by (cf. Section 2.3.2). In principle, we could enumerate all states of the population, where each state corresponds to a particular distribution of individuals over the neutral space, and use the Wright-Fisher model to write down the transition probabilities among the states, which simply follow from a multinomial sampling rule [122]. These probabilities are sufficient to describe the dynamics of the population. In practice, such an approach will often be futile, simply because the size of the state space is vast beyond imagination. In the limit where the population is much smaller than the neutral space of phenotype q, which the population currently occupies (whose number of genotypes we denote by $\mathcal{S}_q$, so that we are in the limit $N \ll \mathcal{S}_q$), a very crude lower bound for number of distinct states is given by the number of states S_u in which each individual has a unique genotype. Clearly

$$S_u = \binom{\mathcal{S}_q}{N}$$

and we can use Stirling's approximation $\log n! \approx n \log n - n$ to obtain

$$S_u \approx \left(\frac{\mathcal{S}_q}{N}\right)^N$$

For example, with $\mathcal{S}_q = 1000$ genotypes and $N = 10$ individuals we have much more than 10^{100} states.

To make progress, it is necessary to approach the problem from a different angle. We make several approximations that lead to substantial simplifications. Throughout, we shall ignore the possibility of higher-order mutations and restrict our attention to $d = 1$ (and we shall drop d in our notation); clearly, this is valid when $L\mu \ll 1$.

4.2.2 Continuity approximation

The complexity of neutral exploration is rooted both in the structure of neutral spaces and in the stochastic dynamics of genetic drift in finite populations. A much simpler picture emerges when we ignore the discrete network structure of the neutral space and the population. This idea leads us to the *continuity approximation*, which is illustrated in Fig. 4.1: Rather than dealing with intricate discrete systems, we treat a much simpler continuous description.

Technically, the continuity approximation can be viewed as a mean-field description of the exploration dynamics: We replace the state of the population by its long-time average. Thus the population constantly explores its average mutational neighbourhood. So we can replace $\phi_p(t)$ by its time average $\langle\phi_p\rangle$. Here, we shall assume that the time average $\langle\phi_p\rangle$ is equal to the average over the neutral space ϕ_{pq}, where the second index q indicates that the continuity approximation reduces the explicit time-dependence to a dependence on the population's neutral space[1]. Consequently, we have the time-independent PPR ν_{pq} given by

$$\nu_{pq} = L\mu\phi_{pq} \tag{4.6}$$

If we write $n_{p,g}$ for the number of mutations from genotype g that lead to phenotype p and

[1]For now, we equate neutral space and phenotype, ignoring the possibility of fragmented neutral spaces. We shall return to this issue in Chapter 7.

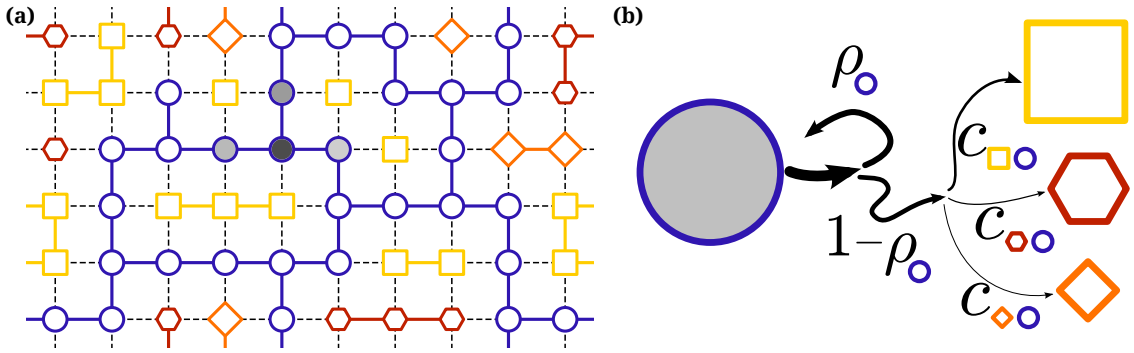

Figure 4.1: Illustration of the continuity approximation (a) An example genotype space, drawn in analogy to Fig. 1.1. Each point corresponds to a unique genotype; shapes indicate the phenotype. Lines connect genotypes that differ by a single mutation. The filling of each node shows what fraction of the population carries the corresponding genotype. **(b)** Under our approximation, the landscape is described by weighted connectivities between phenotypes, neglecting the local structure of the GP map. The arrows illustrate the outcome of mutations: A single mutation can either be neutral (with probability ρ_q, where q is the source phenotype and ρ_q is the robustness defined in Eq. (2.20)), or non-neutral with probability $1-\rho_q$. The probability that a non-neutral mutation leads to phenotype p is given by the connectivity c_{pq} (Eq. (2.24)).

denote by $\mathcal{G}_q$ the set of genotypes of the current neutral space, we have

$$\phi_{pq} = \frac{1}{\mathcal{M}} \sum_{g \in \mathcal{G}_q} n_{p,g} \tag{4.7}$$

where the normalization constant $\mathcal{M}$ is given by the total number of mutations that can be applied to all genotypes in the neutral space: $\mathcal{M} = (K-1)L\,|\mathcal{G}_q|$.

In what follows, we shall often use a slightly different notation that makes it easier to see the significance of neutral spaces, whose existence is intimately linked to mutational robustness. Thus, we split the probability ϕ_{pq} that a mutation leads from phenotype q to phenotype $p \neq q$ into the probability that a mutation is not neutral $(1-\rho_q)$ (where $\rho_q \equiv \phi_{qq}$ is just the mutational robustness of phenotype q, defined in Eq. (2.20)) and the probability c_{pq} that a non-neutral mutation away from q leads to p, which is simply the phenotype connectivity of Eq. (2.24). Mathematically, we have

$$\phi_{pq} = (1-\rho_q)c_{pq} \tag{4.8}$$

In the continuity approximation, we thus arrive at

$$\nu_{pq} = L\mu(1-\rho_q)c_{pq} \tag{4.9}$$

$$M_p(t) = N\nu_{pq}t \tag{4.10}$$

$$T_p(\alpha) = \frac{-\log(1-\alpha)}{N\nu_{pq}} \tag{4.11}$$

So in this description, robustness monotonically slows down the discovery of new phenotypes. The ease with which a particular phenotype p is discovered depends in a simple way on p's connectivity c_{pq} from the population's neutral space q. Locally frequent phenotypes, for which c_{pq} is large, are found much faster and are produced much more often than locally rare phenotypes with small values of c_{pq}.

We can view the continuity approximation as a mean-field description of the state of the population. From a technical perspective, this approach deserves criticism: By averaging, we ignore the fluctuations in the state of the population. But it is exactly these fluctuations (due to mutations) that we are chiefly interested in. The magnitude of the error made in the continuity approximation depends on the time-scale over which we need to average $\phi_p(t)$ to obtain ϕ_{pq}: The faster the population explores the neutral space, the better is our mean-field treatment of the exploration dynamics. Consequently, we expect that the results of the continuity approximation to be most applicable to the polymorphic limit (cf. Section (2.3.2)) with $NL\mu \gg 1$, that is when many mutants are produced in every single generation.

Finally, we remark that equating the time-average $\langle\phi_p\rangle$ of the local frequency $\phi_p(t)$ with the space-average ϕ_{pq} is only valid if the population explores the neutral uniformly over sufficiently long times. But as we have seen in the quasispecies treatment of neutral evolution (Section 2.3.1), large polymorphic populations tend to evolve greater robustness and cluster in the more robust regions of the neutral space. For a better prediction in the polymorphic regime, we should thus use the population robustness, and reweight the $n_{p,g}$ in Eq. (4.7) by the stationary distribution of the population.

To summarize, this section introduced the continuity approximation which neglects the internal structure of the neutral space as well as the internal dynamics of the population

associated with genetic drift. In the next section, we will consider the opposite extreme of strong genetic drift, when the population is well localized in a small part of the neutral space.

4.2.3 Strong genetic drift

In finite populations, genetic drift can lead to a loss of genetic diversity. We can picture the effect of drift as a localization in genotype space. This localization is most significant in the monomorphic limit $NL\mu \ll 1$, when the population is typically localized to a single genotype g at any one time and performs a random walk on the neutral space by successive fixation of neutral mutants (cf. Section 2.3.2). Between two fixations, the population explores only the immediate mutational neighbourhood of the current genotype g.

The monomorphic regime is thus characterized by two distinct types of dynamics, namely local exploration of the current neighbourhood and transitions to new neighbourhoods. Since each neutral mutant has the same chance of going to fixation (cf. Section 2.1.2), the time τ_f between two fixations is distributed geometrically with mean [74]

$$\tau_f = \frac{1}{L\mu\rho_q} \tag{4.12}$$

which is the just the inverse of the rate of neutral mutations. To be precise, we should calculate τ_f based on the robustness of the current genotype ρ_g instead of the average ρ_q. Since our first-order treatment generally assumes homogeneous distributions, we ignore this distinction here. We note that neutral spaces of more robust phenotypes are explored faster ($\tau_f \propto 1/\rho_q$).

Between fixations, mutations can only produce one of the $(K-1)L$ neighbours of the current genotype g. Thus the PPRs are determined by the neighborhood of the current genotype. If this neighborhood contains $n_{p,g}$ genotypes mapping into phenotype p, we have a local PPR $\nu_{p,g}$ given by

$$\nu_{p,g} = L\mu n_{p,g}/((K-1)L) = \mu n_{p,g}/(K-1) \tag{4.13}$$

Under uniformly random mutations of the genotype, each neighbour of g has the same chance to be discovered. The time to discovery is thus also distributed geometrically; the mean τ_e is given by

$$\tau_e = \frac{(K-1)L}{NL\mu} = \frac{K-1}{N\mu} \tag{4.14}$$

τ_e sets the time-scale of local exploration.

It is instructive to compare the ratio ξ of the two time-scales τ_f and τ_e, defined via

$$\xi = \frac{\tau_f}{\tau_e} = \frac{N}{(K-1)L\rho_q} \approx \frac{N}{L} \tag{4.15}$$

We can use ξ to distinguish between different dynamic regimes. If $\xi \gg 1$, fixation takes much longer than exploration. In this limit, which we call the large population limit, we typically expect that all phenotypes accessible from the current genotype are discovered before the population moves to a different genotype. In the opposite extreme $\xi \ll 1$, which we call the large genome limit, exploration is slow and the current genotype is left before all accessible mutants have been explored.

To make these ideas precise, we need to keep in mind that actual discovery and neutral fixation times can show strong fluctuations. In a continuous time approximation, it is straightforward to calculate the probability that a particular mutant is not produced before the population moves to a different genotype: Let τ be the time the population spends at the current genotype, which follows an exponential distribution with mean τ_f. The probability that a particular phenotype p is found during this time is $1 - \exp(-n_{p,g}\tau/\tau_e)$. Integrating over the distribution of τ, we have the probability $P^{(d)}$ that phenotype p is discovered from g, given by

$$P^{(d)}(n_{p,g}) = \int_0^\infty \frac{d\tau}{\tau_f}(1 - e^{-n_{p,g}\tau/\tau_e})e^{-\tau/\tau_f} = 1 - \frac{1}{1 + n_{p,g}\xi} \tag{4.16}$$

If fixations are the rate-limiting step (ie. $\xi \gg 1$), $P^{(d)} \to 1$ whenever $n_{p,g} > 0$, as each neighborhood is searched exhaustively before the population moves on. On the other hand, if fixation is fast ($\xi \ll 1$), the introduction of alternative phenotypes is completely determined by fluctuations; many accessible mutants may not be produced. Then, the probability that

phenotype p is found from g is to leading order simply given by

$$P^{(d)}(n_{p,g}) \approx n_p \xi = \frac{N n_{p,g}}{(K-1)L\rho_q} = N \frac{n_{p,g}}{n_{q,g}} \tag{4.17}$$

The inverse on dependence on ρ_q arises from τ_f: More robust neutral spaces are explored faster, but therefore less thoroughly.

We have now dealt with the period of exploration between two neutral fixations. Next, we need to account for the structure of the neutral space. This structure determines the change in the available phenotypes upon a neutral fixation. To lowest order, we treat the neutral space structure in a mean-field approximation, assuming that phenotypes are homogeneously distributed around the neutral space. So the expected number n_{pq} of neighbor genotypes mapping into phenotype p is simply

$$n_{pq} = (K-1)L(1-\rho_q)c_{pq} \tag{4.18}$$

where c_{pq} is the fraction of mutations that lead from q to p as defined in Eq. (2.24), and the index q again indicates that n_{pq} is an average over the genotypes in q's neutral space. Thus the probability that p is accessible after a neutral fixation is $1-\exp(-n_{pq}) \approx n_{pq}$, where the approximation is valid provided $n_{pq} \ll 1$, that is p is not accessible from every genotype in the neutral space of q.

The probability that p is discovered from the next genotype is thus[2] $P^{(d)}(1)n_{pq}$. So the first discovery time is

$$T_p(\alpha) = \frac{-\tau_f \log(1-\alpha)}{P^{(d)}(1)n_{pq}} \tag{4.19}$$

Let us first consider the small population limit $\xi \ll 1$. Using $P(1) \approx \xi = \tau_f/\tau_e$, we find

$$T_p(\alpha) = \frac{-\tau_e \log(1-\alpha)}{(K-1)L(1-\rho_q)c_{pq}} = \frac{-\log(1-\alpha)}{NL\mu(1-\rho_q)c_{pq}} \tag{4.20}$$

which is identical to our result in the continuity approximation, Eq. (4.11). At first sight,

[2]If n_{pq} is small, actual values of $n_{p,g}$ will practically only be 0 or 1, and hence we can use $P^{(d)}(1)$.

this is quite a surprise: After all, we have argued that the continuity approximation should fail when populations are well localized in genotype space, and now we recover its results in the limit of the smallest populations. But we can make sense of this finding. Just as in the continuity limit, in small populations the production of a particular *genotype* is completely governed by the fluctuations of random mutations, and in the mean-field model, the probability that a non-neutral mutation leads to phenotype p is simply c_{pq}.

In the opposite limit $\xi \gg 1$, however, the role of fluctuations is different. As we have seen, the population will almost certainly produce all neighbours of the currently fixed genotype. In this regime then, fluctuations only determine which neutral mutations are fixed. We obtain for the first discovery time

$$T_p(\alpha) = \frac{-\log(1-\alpha)}{L^2(K-1)\mu\rho_q(1-\rho_q)c_{pq}} \tag{4.21}$$

which shows a non-monotonic scaling with the source phenotype's robustness. This non-monotonicity has also been obtained by Draghi et al. [33]. Our model provides a very simple explanation for this scaling: The rate at which the neutral space q is traversed by a population is proportional to ρ_q, but the probability of producing non-neutral mutants decreases as $(1-\rho_q)$.

In summary, a well-localized population under strong genetic drift is sensitive to the local structure of the neutral space. In the next section, we compare these results to the continuous approximation and try to extrapolate our findings to the middle ground where population may occupy more than a single genotype, but less than the entire available neutral space.

4.2.4 Between the extremes

In the preceding two sections, we have studied the two simplest limits of completely polymorphic and completely monomorphic populations. Let us now compare the results. Most importantly, we always find $T_p \propto 1/c_{pq}$: Rare phenotype are hard to find.

In the monomorphic limit, our microscopic approach has revealed a novel distinction of two regimes depending on the ratio $\xi \approx N/L$. When populations are small ($N \ll L$), we

recover the results of the continuity approximation – Eqs. (4.11) and (4.20) are identical. In both cases, each non-neutral mutation has the same chance of producing a particular phenotype. However, this is not true when a monomorphic population is very large ($N \gg L$ while $NL\mu \ll 1$): Over a time-scale of τ_f, the outcomes of non-neutral mutations are correlated because they arise on the same genetic background. The localization in genotype space induces short-term memory. We follow up on this result in Chapter 6.

Let us now turn to our results for large populations in the monomorphic limit and in the continuous approximation. The only difference between Eqs. (4.11) and (4.21) is that N in the polymorphic regime is replaced by $L(K-1)\rho_q$ in the monomorphic limit. This difference is intuitive: When the population is diverse, every new individual helps exploration and reduces discovery times. But if all individuals have the same genotype, simply having 'more of the same' does not make neutral exploration faster.

These results suggest that for intermediate $NL\mu$ and $N \ll L$, there should be a smooth transition between the extremes. We can quantify the crossover by introducing a factor γ:

$$T_p(\alpha) = \frac{-\log(1-\alpha)}{N\gamma\nu_p} \tag{4.22}$$

Our theory then predicts that $\gamma \to 1$ as either $NL\mu$ becomes very large or $N \ll L$, and that $\gamma \to (K-1)L\rho_q/N$ as $NL\mu \ll 1$ and $N \gg L$.

γ accounts for the reduction of genetic diversity in the population under drift. Consequently, calculating γ from first principles brings us back to the problem of explicitly treating genetic drift, and we shall continue to side-step this problem. Instead, we use γ as a fitting parameter and obtain its numerical value by simple linear regression:

$$\gamma = \frac{-\log(1-\alpha)\ \sum_p N\nu_p T_p(\alpha)}{\sum_p \left(N\nu_p T_p(\alpha)\right)^2} \tag{4.23}$$

where the T_p are taken from simulation data.

To complete our theoretical discussion of first discovery times, we need to consider the state of the population at the time of environmental change ($t = 0$), which we turn to in the

next section.

4.2.5 Initial conditions

When the population does not simultaneously cover the entire neutral space, neutral exploration will often be necessary to discover a particular phenotype. However, some phenotypes will be accessible immediately at $t = 0$, and provided $N \gg L$, they will be discovered before the fixation of a neutral mutant. So for these phenotypes, the discovery time is still given by Eq. (4.20). For a random initial condition and a well-localized population, the probability that p is accessible at $t = 0$ is simply $1 - \exp(-(K-1)L(1-\rho_q)c_{pq}) \approx (K-1)L(1-\rho_q)c_{pq}$; in terms of the PPRs, we thus have a threshold

$$\nu^*(\alpha) = -\frac{\mu \log(1-\alpha)}{(K-1)} \tag{4.24}$$

such that we should use Eq. (4.20) when $\nu_{pq} \gg \nu^*(\alpha)$, and Eq. (4.21) if $\nu_{pq} \ll \nu^*(\alpha)$; since the time of fixation of the first neutral mutant has a wide distribution, we should expect a smooth cross-over between the extremes. We also note that ν^* will be smaller when the population is genetically diverse, simply because multiple genotypes are occupied at $t = 0$ so that the initial neighbourhood of the population is bigger than $(K-1)L$ genotypes.

Finally, we can also consider fixed initial conditions. In this case, the initial neighborhood of the population can be determined from the GP map, and for the first discovery times, we use an adapted PPR

$$\hat{\nu}_{pq} = L\mu \left(\hat{\phi}_p^{(0)} + \gamma \phi_{pq} \right) \tag{4.25}$$

where $\hat{\phi}_p^{(0)}$ is the fraction of p-genotypes in the initial neighborhood.

4.3 Simulations

4.3.1 Homogeneous neutral spaces

Our analytic description is based on a mean-field like approximation that ignores local correlations between genotypes. As a first test, we will perform simulations on neutral spaces which come as close as possible to meeting this assumption. The absence of correlations is easily engineered in the random GP map; furthermore, this system allows us to choose the phenotype frequencies $\{F_p\}$. We fix $K = 4$ and, for computational reasons, $L = 12$.

We introduce varying strengths of phenotypic bias by considering 3 sets of phenotype frequencies. As a reference, we use the frequencies of $L = 12$ RNA secondary structures (landscape A), which has $N_{\mathcal{P}} = 58$ distinct phenotypes. For comparison, we choose two different distributions in which the source phenotype q has frequency $F_q = 0.5$. Keeping $N_{\mathcal{P}} = 58$ fixed, we reduce the frequency from each phenotype to the next by a constant factor e^k. Thus if we label phenotypes by their frequency rank, F_p decays exponentially with p for $p \geq 2$: $F_p \propto e^{-kp}$ (rank 1 corresponds to the source phenotype). We use $e^k = 1.1$ (landscape B) and $e^k = 1.25$ (landscape C). The 3 distributions are shown in Fig. 4.2a.

By choosing F_q well above the threshold for a completely connected neutral network $\lambda \approx 0.37$ (see Section 3.2, in particular Eq. (3.3)), we ensure that the source neutral space is large (several million genotypes). And since we have $N_{\mathcal{P}} - 1 = 57 > (K-1)L = 36$, we are in the regime where a single genotype cannot allow access to all alternative phenotypes. So except for correlations between genotypes, the random GP map produces landscapes that feature the main ingredients of our investigation: Large, robust neutral spaces, strong phenotypic bias, and the necessity of neutral exploration to reach all alternative phenotypes.

For each of these three distributions, we created a single realization of the random GP map. This realization serves as the frozen landscape on which we set the population to evolve. As the source neutral space contains a vast part of genotype space, and since the assignment of genotypes to phenotypes (according to the phenotype frequencies) is random, we have $\phi_{pq} \approx F_p$ to a very high accuracy (Fig. 4.2b); however, due to frozen fluctuations in the realization of the random GP map, in landscape C two phenotypes of very low frequency

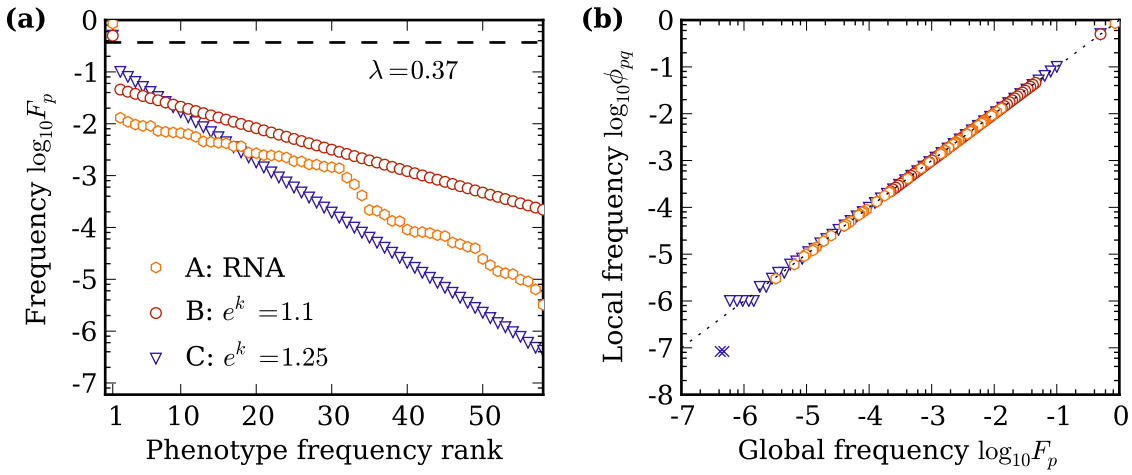

Figure 4.2: Structure of the random GP map landscapes. (a) Phenotype frequencies used in the simulations. A are the phenotype frequencies of $L = 12$ RNA secondary structures; B and C have the same total number of phenotypes, but frequencies decrease exponentially with rank p as $\exp^{-kp}$. **(b)** Comparison of global frequencies F_p and the local frequencies ϕ_{pq} around the source neutral space q with phenotype rank 1, defined in Eq. (4.7). Crosses mark the two phenotypes (in landscape C) for which $\phi_{pq} = 0$.

are inaccessible by single mutations from the source neutral space.

On each landscape, we simulate the evolutionary process defined above for different population dynamic parameter sets (N, μ). The initial state at $t = 0$ is reached by setting the entire population onto a single, randomly selected genotype in the source neutral space and then evolving for $10N$ generations to reach a mutational equilibrium. Discovery times and the number of mutants are then measured over $T = 10000/(N\mu)$ generations; the scaling $T \propto (N\mu)^{-1}$ ensures that the total number of mutants is roughly similar across the different parameter sets. 100 runs are performed with a fixed initial state, and another 100 simulations each start from different initial states.

4.3.2 Total number of mutants

Our theory predicts that over sufficiently long times T, the total number of mutants $M_p(T)$ with phenotype p is simply proportional to the mean-field PPR ν_{pq} (Eq. (4.10)). The left-hand panels of Fig. 4.3 illustrate that this is indeed the case to a very high accuracy. We note that in Fig. 4.3e, we have $N = 10^4$ and $\mu = 10^{-2}$ which yields $NL\mu = 1200$. In this highly polymorphic regime, we should expect the population to cluster in regions of higher

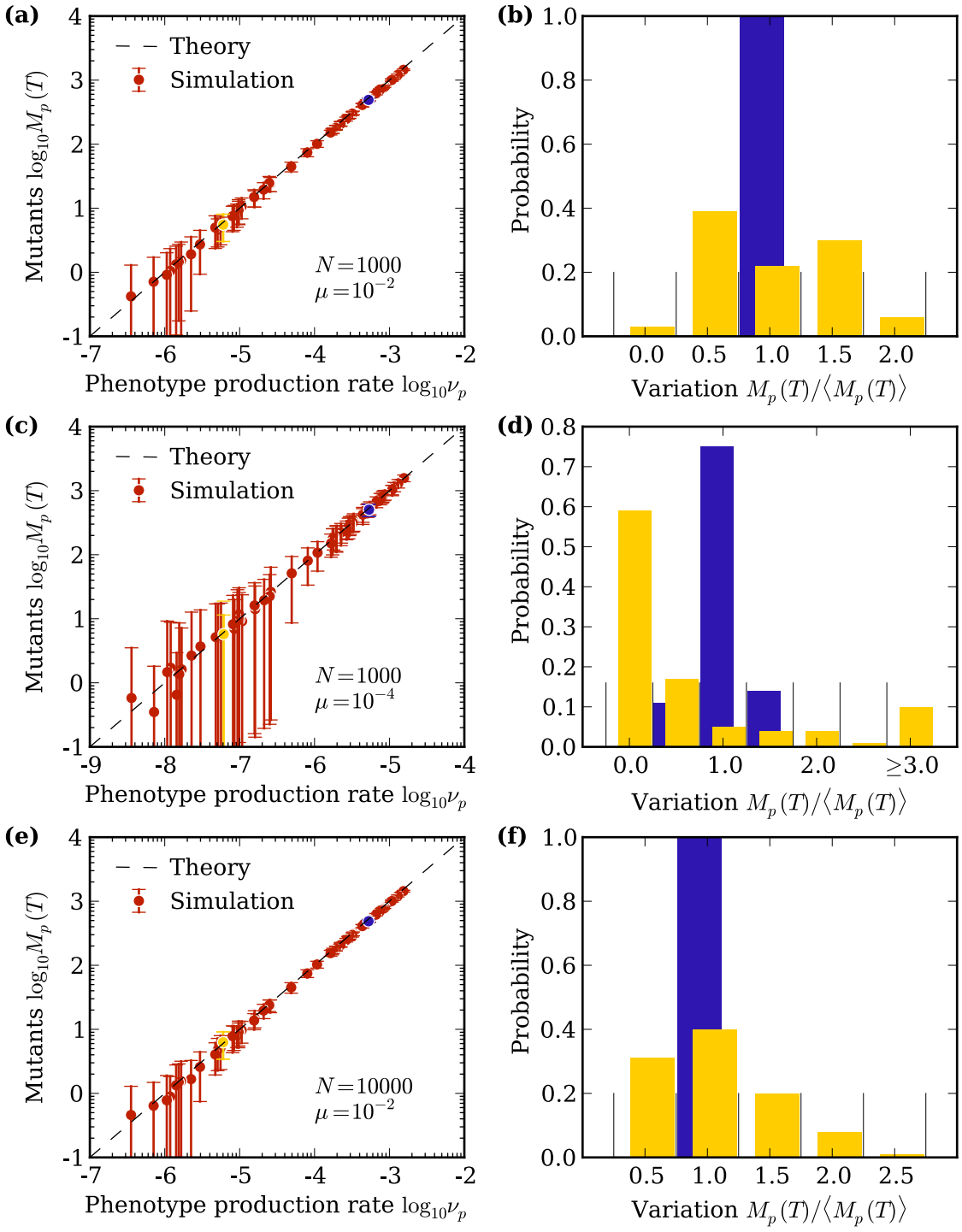

Figure 4.3: Frequent phenotypes are produced more often than rare ones. **(a)** Total number of mutants per phenotype, averaged over 100 simulations with $N = 1000, \mu = 10^{-4}$ over $T = 10^5$ generations on landscape A. Error bars are 1 standard error on the mean. The black dashed line corresponds to Eq. (4.10). Panel **(b)** shows the distribution of $M_p(T)$ in the individual simulations for the two phenotypes marked in blue and yellow. The values of $M_p(T)$ are normalized by their average, and then binned with a bin-width of 0.5. The black vertical lines indicate the bin edges. **(c)** and **(d)** are analogous, but for $N = 1000, \mu = 10^{-4}$ and hence $T = 1000$; **(e)** and **(f)** correspond to $N = 10000, \mu = 10^{-2}$ and $T = 100$.

robustness in the neutral space [44, 123], leading to a population robustness greater than the average robustness of the neutral space. The fact that this effect is hardly noticeable underlines the homogeneity of the source neutral space.

Thus over long times, homogeneous neutral spaces are explored uniformly. However, we can see from the panels on the right of Fig. 4.3 that it can take very long indeed to achieve this, particularly for rare phenotypes (small ν_p) and when $NL\mu$ is small. For example, with $N = 1000, \mu = 10^{-4}$ (Fig. 4.3d), 41 out of 100 simulations do not produce a single mutant of the phenotype marked in yellow; on the other hand, one simulations overshoots the average by a factor of almost 20. We will return to the observation in Chapter 6; for the time being, we note that Eq. (4.10) gives a very good agreement with simulations provided the number of generations T, or more accurately the expected number of mutants $M_p(T) = NL\mu(1 - \rho_q)c_{pq}T$ (Eq. (4.10)), is large.

4.3.3 First discovery times

This success of the mean-field model is not too surprising, given that the total number of mutants is the sum over a large number of generations, so that the effect of fluctuations is reduced. This contrasts with the first discovery time T_p, which is very sensitive to fluctuations – indeed, T_p directly measures the fluctuations due to mutations. The simulation results for first discovery times thus provides a more challenging test for our calculations.

Fig. 4.4 shows the results of the simulations with $N = 100, \mu = 10^{-5}$ on landscape B. With these parameters, we have $NL\mu = 0.012$ and our predictions for the monomorphic limit describe the data reasonably well.

Fig. 4.5 displays discovery times over a wide range of mutation rates, crossing from the polymorphic regime ($NL\mu \gg 1$) into the monomorphic limit ($NL\mu \ll 1$). In the large genome limit ($L \gg N$, Fig. 4.5a), rescaling the discovery times by the mutation rate (which of course is a natural timescale in the system) shows a very good collapse of the data. We note that using factor of $\gamma \approx 0.5$ in Eq. (4.22) is necessary for good quantitative agreement with the simulations. In large populations ($N \gg L$, Fig. 4.5b), we do observe this collapse only for

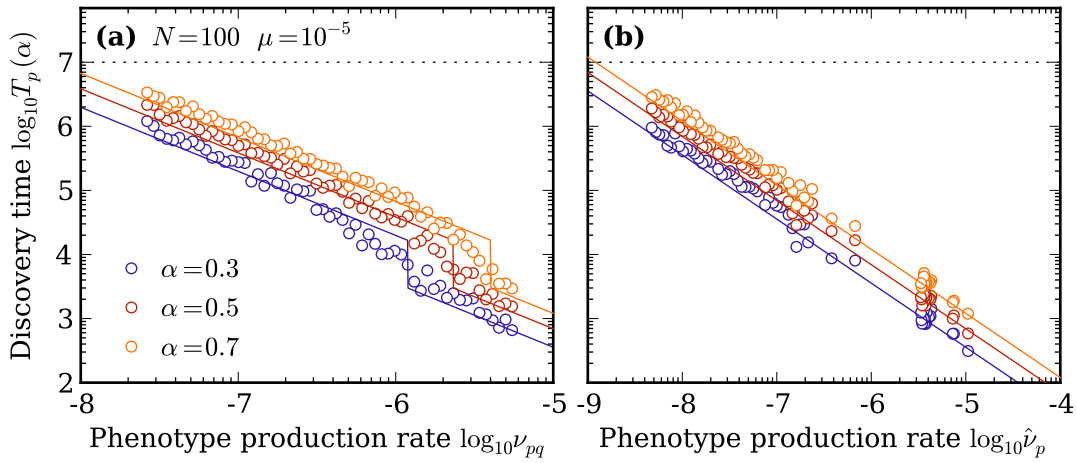

Figure 4.4: Frequent phenotypes are discovered earlier than rare ones. Simulations with $N = 100, \mu = 10^{-5}$ on landscape B. **(a)** Random initial conditions. Dots correspond to the simulation results for different discovery fractions α. The lines are Eq. (4.11) for phenotypes that are immediately accessible ($\nu_p > \nu_p^*(\alpha)$, Eq. (4.24)), and Eq. (4.21) for the phenotypes whose discovery requires neutral exploration. **(b)** With fixed initial conditions, phenotype production rates are calculated using Eq. (4.25).

the phenotypes of high connectivity c_{pq}, which can typically be found in the instantaneous neighbourhood of the population at $t = 0$ (cf. Section 4.2.5). By contrast, rare phenotypes p with low c_{pq}, which are discovered only after neutral exploration, receive an additional 'penalty': As the population moves from the polymorphic regime into the monomorphic regime as μ decreases, the discovery time T_p decreases faster than $1/\mu$. This effect is caused by the loss of diversity under strong genetic drift: Most individuals do not contribute to the discovery of novel phenotypes in monomorphic populations. In agreement with our discussion in Section 4.2.5, the necessity of neutral exploration applies to more and more frequent phenotypes (higher c_{pq}) as μ is reduced: The point of departure from the continuity approximation moves further to the right as μ decreases in Fig. 4.5b.

Fig. 4.6 complements our discussion the μ-dependence of the first discovery time T_p. In the large genome limit (Fig. 4.6a), we see that T_p scales as $1/\mu$ independent of phenotype connectivity c_{pq}. By contrast, when the population is large (as in Fig. 4.6b) we observe a departure from this $1/\mu$-scaling around the transition from the polymorphic to the monomorphic regime: As the population becomes monomorphic, we need to replace the population size N in Eq. (4.11) by the smaller value $(K-1)L\rho_q$ to obtain Eq. (4.21). When we fol-

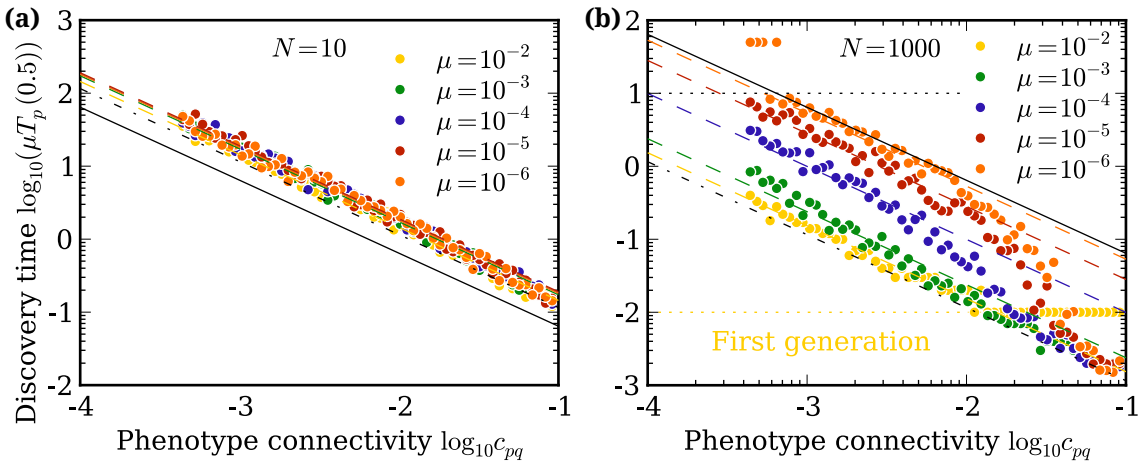

Figure 4.5: Genetic drift slows down neutral exploration. **(a)** shows median discovery times of the alternative phenotypes in landscape B, rescaled by the respective mutation rates. To facilitate the comparison, the x-axis gives phenotype connectivities c_{pq} rather than the μ-dependent mean-field PPRs $\nu_{pq} = L\mu(1-\rho_q)c_{pq}$. Dashed lines show Eq. (4.22), including the fitting factor γ (calculated according to Eq. (4.23)). The dash-dotted black line shows the continuity approximation (Eq. (4.20)), and solid black line marks the monomorphic limit of large populations ($NL\mu \ll 1$, $N \gg L$; Eq. (4.21)). Note that the diagram shows results in the large genome limit $N \ll L$ (since $N = 10, L = 12$). **(b)** shows analogous result in the large population regime $N \gg L$. Here, the horizontal black dotted line indicates the total simulation time, and points above the line have been discovered in less than 50% of the simulations. The horizontal yellow dotted line marks the generation one, rescaled by $\mu = 10^{-2}$.

low this transition as a function of μ, we thus obtain a transient regime where T_p increases faster than $1/\mu$ as the mutation rate drops. Again, the exact point of the transition depends on the connectivity c_{pq} of the phenotype p under consideration since phenotypes of higher connectivity are more likely to be immediately accessible to the population, so that the slow speed of neutral exploration affects these phenotypes only at smaller values of μ.

Let us now turn to the scaling of discovery times with population size. Our theory predicts that in the monomorphic limit, an increase of N beyond $(K-1)L\rho_q$ will have little effect, as the discovery time becomes independent of N in this limit (see Eq. (4.21)). On the other hand, when there is standing neutral diversity in the population, we should expect that adding more individuals will reduce discovery times.

Fig. 4.7 compares these expectations to the outcomes of our simulations. Just as in the dependence on mutation rate, we observe a dichotomy between the frequent phenotypes which are usually found in the population's immediate neighbourhood at $t = 0$, and the rare

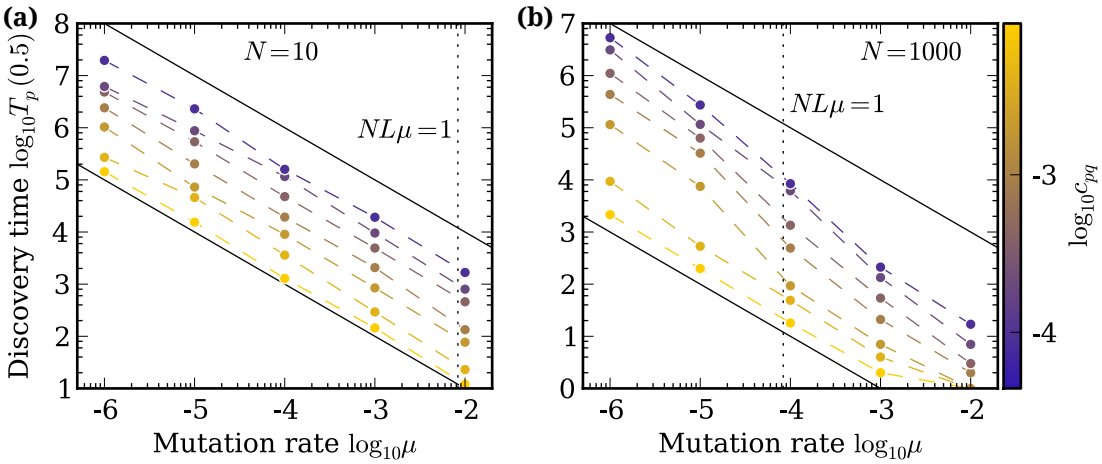

Figure 4.6: The effect of mutation rates on discovery times. (a) shows the median discovery times, plotted against mutation rate for different values of phenotype connectivity c_{pq}, in the large genome limit ($N = 10, L = 12$) on landscape B. The vertical dotted line corresponds $NL\mu = 1$, marking roughly the transition between the monomorphic ($NL\mu \ll 1$) and polymorphic ($NL\mu \gg 1$) regimes. The solid black lines are drawn for illustration only, and show the scaling $T_p \propto \mu^{-1}$ that both Eqs. (4.20) and (4.21) suggest. **(b)** shows analogous results for $N = 1000$.

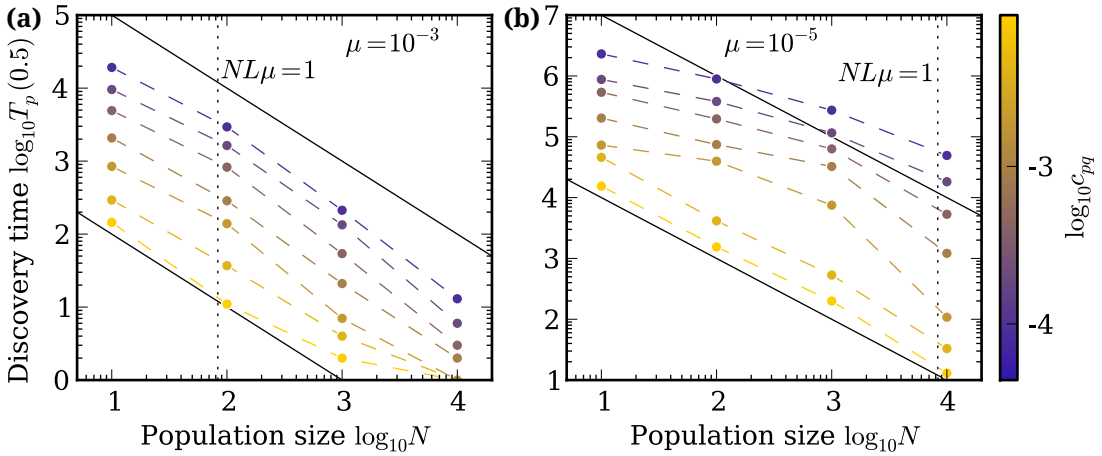

Figure 4.7: Dependence of discovery time on population size. The diagrams show how the median discovery time of phenotypes scales with population size. Colours indicate the connectivity of the target phenotypes to the source neutral space. The simulations were performed on landscape B. Panel (a) shows results for $\mu = 10^{-3}$ and panel (b) corresponds to $\mu = 10{-}5$.

phenotypes whose discovery is predicated on exploration of the neutral space. For frequent phenotypes, the scaling $T_p \propto N^{-1}$ can be seen clearly in the data.

The discovery times of rare phenotypes show a more complex behaviour: As we have argued in Section 4.2.5, it depends on N and μ whether or not a phenotype is immediately accessible. Consequently, the point of departure from the $T_p \propto N^{-1}$-scaling (towards a slower reduction of discovery times with N) depends on the connectivity c_{pq} of the target phenotype. The lower c_{pq}, the more diverse the population needs in order for p to be in the initial neighbourhood of the population, and hence to reduce the discovery time of p. It is interesting to note that for rare phenotypes, the slow decrease of T_p with N for small N is accelerated when $NL\mu$ becomes of order unity: As the population gains standing genotypic diversity, relatively rare phenotypes are more likely to occur in the initial neighbourhood of the population, so that their discovery is no longer predicated on neutral exploration.

We conclude our study of the uncorrelated landscapes by considering the behaviour of the fit parameter γ which accounts for genetic drift. Fig. 4.8 shows the values of γ calculated from our simulation according to Eq. (4.23). In agreement with our arguments in Section 4.2.4, we find that $\gamma \to 1$ in the polymorphic limit $NL\mu \gg 1$, and also in the large genome limit $N \ll L$, where γ is practically independent of μ (Fig. 4.8a,b). In Fig. 4.8c, we also see that $N\gamma \to (K-1)L\rho_q$ in the monomorphic limit $NL\mu \ll 1$, as the comparison of Eqs. (4.20) and (4.22) suggests.

In large populations ($N \gg L$) the impact of genetic drift becomes more severe when μ is reduced at fixed N ((ie. γ reduces, Fig. 4.8a). In the monomorphic limit $NL\mu \ll 1$ of large populations, γ decreases with N at fixed μ (Fig. 4.8b): The 'penalty' of localization becomes more significant when the population is large. As N is increased at fixed μ, γ decreases at the transition from small to large population, but eventually seems to increase as the population becomes more and more polymorphic (Fig. 4.8b). For the range of $NL\mu$ shown in Fig. 4.8c, γ shows a relatively clear scaling with ρ_q, indicating that genetic drift contributes to the question to what extent robustness aids in the discovery of new phenotypes.

In summary, we find that our analytic treatment agrees well with the simulations on homogeneous neutral spaces. Most importantly, phenotypic bias has a clear, intuitive effect:

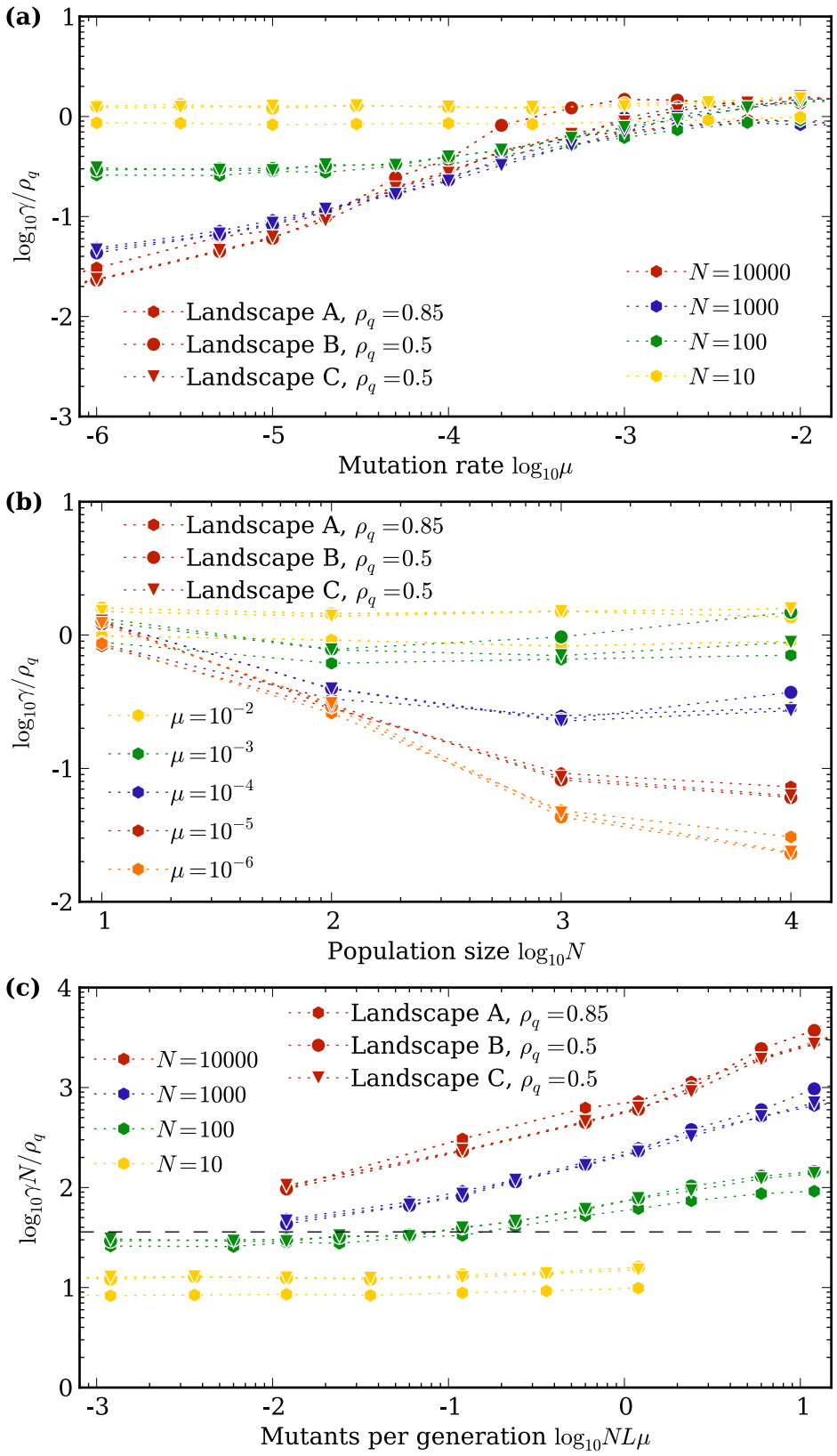

Figure 4.8: The impact of genetic drift. The diagrams show the dependence the drift correction factor γ (calculated according to Eq. (4.23)) on dynamic parameters. For comparison between the different landscape, γ is rescaled by $1/\rho$. **(a)** shows γ as a function of mutation rate, **(b)** displays the dependence of γ on population size. In **(c)**, γ is also rescaled by population size and plotted versus the number of mutants per generation. The black dashed line shows $N\gamma/\rho_q = (K-1)L$, our prediction for the monomorphic limit of large populations (cf. Section 4.2.4.

Frequent phenotypes are discovered earlier and more often than rare ones.

4.4 Summary and discussion

In this chapter, we have introduced and studied a simple model of neutral exploration. To account for the the GP map, our model distinguishes between the genotypic mutation rate μ and phenotypic production rates (PPRs) ν_p, which correspond to the probability that phenotype p is produced in a given reproduction event. This distinction allows us to coherently include phenotypic bias into our description of evolutionary dynamics. The main aim of our calculations in this chapter has been to predict the first discovery time of a particular alternative phenotype. As an advantageous phenotype can only be fixed after it has been discovered, the discovery time places a fundamental limit on the speed of evolution.

In general, PPRs depend on the GP map and the time-dependent state of the population. If the PPRs are given, it is straightforward to calculate the first discovery time. Yet to describe the time-dependence of the PPRs requires a full treatment of genetic drift, which – as we have seen in Section 2.3 – is a highly complex problem. Instead of attacking this problem head-on, we have attempted to gain analytic insights through simplifying approximations.

Our lowest-order treatment, the continuity approximation of Section 4.2.2, gave a mean-field description of neutral exploration. Replacing the time-dependent state of population by its long-time average, we ignored generation-to-generation fluctuations. We thus arrived at time-independent PPRs, corresponding to a population that is spread out widely over the neutral space. Consequently, the internal structure of the neutral space could be ignored and we saw that (local) phenotypic bias translates directly into first discovery times: Frequent phenotypes are quickly discovered, but rare phenotypes are very hard to find. As the continuity approximation assumes a population of large genotypic diversity, we argued that the predictions of continuity approximation would be most valid when populations are large and highly polymorphic (that is, $NL\mu \gg 1$).

In Section 4.2.3, we presented a calculation for the dynamics of neutral exploration in the opposite limit of monomorphic populations with $NL\mu \ll 1$. In this regime, genetic drift

causes a localization of the entire population onto a single genotype, and the fixation of neutral mutants allows the population to explore the neutral space. Between two fixations, PPRs are thus completely determined by the mutational neighbourhood of the genotype currently fixed in the population. Rather than explicitly accounting for this local structure, which is induced by the GP map, we treated the influence of the GP map in another mean-field approximation, whereby the probability of a phenotype becoming accessible after a neutral fixation is simply proportional to its connectivity. This approach is similar to the model of Draghi et al. [33], except that we explicitly allow for phenotypic bias.

In the monomorphic regime, we observe different dynamics depending on the size of the population N compared to the genotype length L. When $N \ll L$, we recovered the results of the continuity approximation. We have argued that this surprising finding is due to the fact that in both of these two limiting cases, the production of a particular genotype is completely dependent on the random fluctuations of mutations. By contrast, if $N \gg L$, we found that genetic drift severely delays the discovery of rare phenotypes; in particular, as long as $NL\mu \ll 1$, the discovery time becomes independent of N. Nonetheless, we again observed that the discovery time T_p of phenotype p is simply inverse proportional its connectivity c_{pq} from the neutral space corresponding to the population's phenotype q. Independent of the dynamic regime, rare phenotypes are hard to discover.

An important question concerns the biological interpretation of these regimes, in particular: What is L? In principle, of course, L can be viewed as the entire genome length, which is of the order of millions to billions of base-pairs in most organisms, except viruses [92]. However, neutral exploration is concerned with conditionally neutral mutations which can allow for later non-neutral changes; so instead of the whole genome, where most mutations may not have an observable effect [74], we should consider only those sites which may contribute to the discovery of novel phenotypes. Such an 'effective genome length' L_{eff} may typically comprise only a small number of genes which could respond to a particular environmental change. So L_{eff} may be of the order of thousands of bases. Equally, N should be replaced by an effective population size N_{eff}, which is a well-understood concept in population genetics which accounts for deviation from the idealized Wright-Fisher reproduction in actual

populations, for example due to age-structure [17]. In principle, both regimes $L_{eff} \gg N_{eff}$ and $L_{eff} \ll N_{eff}$ are conceivable. It is worth noting that even though L_{eff} may be relatively small, neutral spaces will typically still be vast, since their size depends exponentially on L_{eff}.

We have tested our predictions against simulations under the random GP map. As we have seen in Section 3.2, we need to choose a very large frequency for the population's current phenotype in order to ensure a large neutral network. For our purposes, this restriction is not a problem, since we are interested in the limit of large neutral spaces anyway. Furthermore, the random GP map allows us to choose the frequencies of the alternative phenotypes directly, so that we can explore a wide range of frequencies in a single landscape. The qualitative agreement with our prediction is satisfactory, and a single fitting parameter is sufficient for a good quantitative agreement under general drift scenarios.

Up to frozen fluctuations in the realization of the landscape, the random GP map conforms to our mean-field approximation. Yet we have seen in Chapter 3 that more biologically motivated GP maps lead to clustering of genotypes with the same phenotypes. It is quite conceivable that this leads to correlations between genotypes which the mean-field model cannot capture. We will address this issue with simulations on RNA neutral spaces in Chapter 7. But before that, we will apply of our mean-field calculations to the question of evolvablity and its relation to robustness and neutral spaces.

Chapter 5

A dynamic approach to evolvability

5.1 Introduction

In the previous chapter, we studied the discovery time of particular phenotypes. Here, we focus our attention on a related, but subtly different question: How many distinct phenotypes does a population encounter over time? We call this number of discovered phenotypes the *population evolvability* E.

In Section 5.2.1, we use our microscopic framework from the previous chapter to calculate the average population evolvability for a general GP map. We see how population evolvability relates to other evolvability measures (see Section 2.2.5): Wagner's potential evolvability [127] emerges as the infinite time limit, while the diversity-measure studied by Cowperthwaite et al. [24] appears in the first correction to a linear increase of population evolvability with time.

To understand the implications of phenotypic bias, we calculate the population evolvability using a slightly different approach in Section 5.2.2. The key idea is that of *diminishing returns*: Mutations typically lead to already discovered, frequent phenotypes, while rare phenotypes remain undiscovered for very long times. This effect has important implications for our understanding of evolvability. In Section 5.3, we check our analytic predictions against simulations in the random GP map (cf. Sections 3.2 and 4.3.1).

In general, we show that predictions about evolvability require detailed knowledge of the

GP map. Thus the relationship between robustness and evolvability can be quite complex. It may not be possible to resolve their apparent tension in a simple way.

5.2 Theory

5.2.1 Evolvability and the GP map

Using our mean-field description of the GP map introduced in Chapter 4, it is straightforward to calculate the average population evolvability $E(t)$, which is simply the expected number of phenotypes discovered by the population within t generations. We invert the expression for the first discovery time $T_p(\alpha)$ (Eq. (4.22)) to obtain the probability $\alpha_p(t)$ that phenotype p has been discovered by a population of size N with t generations:

$$\alpha_p(t) = 1 - \exp(-N\gamma\nu_{pq}t) \tag{5.1}$$

Again, we need to be careful about the application of γ: For short times, we should set γ to unity as neutral exploration is not important (cf. Section 4.2.5). Ignoring these complications for the moment, we can simply sum over alternative phenotypes at fixed t and obtain the expected number of phenotypes discovered:

$$E(t) = \sum_{p\neq q} \Big(1 - \exp(-N\gamma\nu_{pq}t)\Big) = \mathcal{E}_q - \sum_{p\neq q} \exp(-NL\mu\gamma(1-\rho_q)c_{pq}t) \tag{5.2}$$

where $\mathcal{E}_q$ is the potential evolvability of the current neutral space.

We can make several observations concerning the population evolvability. First, $E(t)$ converges to $\mathcal{E}_q$ in the limits of very large populations ($N \to \infty$) and at very long times ($t \to \infty$). Since c_{pq} can easily be tiny for some phenotypes, both of these limits should viewed with care: The infinite population limit requires the *entire* neutral space to be simultaneously occupied, while the infinite time limit requires that the population has explored the entire neutral space over time.

Second, we can expand Eq. (5.2) for short times to obtain

$$E(t) \approx NL\mu\gamma(1-\rho_q)t - \Big(NL\mu\gamma(1-\rho_q)t\Big)^2 \sum_{p\neq q} c_{pq}^2 + O(t^3) \tag{5.3}$$

where the linear term is independent of the distribution of c_{pq} due to the normalization $\sum_{p\neq q} c_{pq} = 1$, which follows directly from the definition of the c_{pq} in Eq. (2.24). For very short times, we see that $E(t)$ grows linearly with time, but the growth is reduced by $(1-\rho_q)$: Robustness does appear to impede evolvability. On the other hand, we should keep in mind that γ may increase with ρ_q (cf. Fig. 4.8), suggesting again an optimal value of ρ_q to boost evolvability. The second order contribution is proportional to $\sum_{p\neq q} c_{pq}^2$, which is reminiscent of the diversity measure proposed by Cowperthwaite et al. [24] (cf. Chapter 2.2.5).

Eq. (5.2) allows us to calculate the population evolvability from the GP map. In light of our findings about γ, it is not obvious whether robustness aids or impedes evolvability. By contrast, our calculations clearly highlight that population evolvability is sensitive to the details of the GP map. In the next section, we will follow a different approach to calculating $E(t)$ that further highlights the influence of the GP map.

5.2.2 The role of phenotypic bias

The developments above are general in the sense that they can be applied to any GP map, and (provided we know γ) are expected to hold for various population dynamic regimes. However, the description in terms of a sum of exponentials in Eq. (5.2) provides limited analytic insight, particularly concerning the role of phenotypic bias. For a deeper insight, we shall now return to the simpler limit of large, monomorphic populations ($NL\mu \ll 1, N \gg L$) and calculate population evolvability following a different route.

As we have discussed above, the regime of large monomorphic populations ($NL\mu \ll 1$ and $N \gg L$) can be pictured as a random walk over the neutral space. The immediate neighbourhood of the current genotype is explored quickly, while transitions from one genotype to another are slow by comparison: The exploration timescale Δt is given by the average

time τ_f (Eq. (4.12)) between the fixations of two neutral mutants:

$$\Delta t = \tau_f = \frac{1}{L\mu(1-\rho_q)} \tag{5.4}$$

Over Δt generations, we denote by the evolvability gain ΔE the increase in $E(t)$. In large monomorphic populations, naively this gain is simply given by the number of alternative phenotypes around the genotype currently occupied by the population. Averaging over the variations in robustness among genotypes, we have

$$\Delta E = (K-1)L(1-\rho_q) \equiv \langle \bar{n}_q \rangle \tag{5.5}$$

where $\langle \bar{n}_q \rangle$ is simply the number of non-neutral neighbour genotypes around a single genotype in the neutral space of phenotype q. In these simple terms, we obtain a naive differential equation for $E(t)$:

$$\frac{dE}{dt} = \frac{\Delta E}{\Delta t} = (K-1)L^2\rho_q(1-\rho_q)\mu \tag{5.6}$$

Together with the initial condition $E(t=0) = 0$, we obtain the straightforward solution

$$E(t) = \frac{\Delta E}{\Delta t} t \tag{5.7}$$

This naive picture reproduces the non-monotonic dependence on ρ_q that we have already encountered in Section 4.2.3. Again, this scaling stems from the trade-off between exploration speed ($1/\Delta t \propto \rho_q$) and ease of non-neutral mutations ($\Delta E \propto (1-\rho_q)$).

However, Eq. (5.7) shows no dependence on the details of the GP map, in contrast to our results in the previous section. The description ignores that the re-discovery of phenotypes already encountered does not increase the population evolvability. Consequently, the increase in $E(t)$ upon a neutral fixation is reduced:

$$\frac{dE}{dt} = \frac{\Delta E}{\Delta t} f(t) \tag{5.8}$$

where $f(t)$ is the probability that a non-neutral mutation leads to a phenotype not previously

discovered. Clearly, as time progresses, $f(t)$ decreases until it reaches 0 when $E(t) = \mathcal{E}_q$. So $f(t)$ measures the diminishing returns of long exploration when most non-neutral mutations lead to phenotypes that have already been discovered.

In the absence of phenotypic bias, we have $c_{pq} = 1/\mathcal{E}_q$ (cf. Section 2.2.5): All alternative phenotypes are equally accessible on average. Consequently, $f(t) = 1 - E(t)/\mathcal{E}_q$ and we have

$$\frac{dE}{dt} = \frac{\Delta E}{\Delta t}\left(1 - \frac{E(t)}{\mathcal{E}_q}\right) \tag{5.9}$$

which can solved straightforwardly to give

$$E(t) = \mathcal{E}_q\left(1 - \exp\left(-\frac{\Delta E}{\mathcal{E}_q}\frac{t}{\Delta t}\right)\right) \tag{5.10}$$

Thus in the absence of phenotypic bias, $E(t)$ increases linearly with time until an exponential cut-off ensures that $E(t)$ does not increase beyond $\mathcal{E}_q$. Substituting our expressions for Δt (Eq. (5.4)) and ΔE (Eq. (5.5)), we obtain for large monomorphic populations

$$E(t) = \mathcal{E}_q\left(1 - \exp\left(-(K-1)L^2\rho_q(1-\rho_q)\mu t/\mathcal{E}_q\right)\right) \tag{5.11}$$

We will now add phenotypic bias to our description. For convenience, we rank alternative phenotypes according to the connectivity coefficient (just as we rank phenotypes globally by their frequency); we can thus label phenotype p by its (local) rank, going from $p = 1$ for the highest connectivity to $p = \mathcal{E}_q$ for the lowest connectivity.

From our results for first discovery times (Eq. (4.21)), it is plausible make the 'sequential approximation': We assume that phenotypes are discovered in the order given by their ranks in local frequency, which should be valid if the differences in local phenotype frequencies are large. In the sequential approximation, the chance of discovering a new phenotype is reduced by the connectivity coefficients of the most easily accessible phenotypes:

$$f(t) = 1 - \sum_{p=1}^{E(t)} c_{pq} \tag{5.12}$$

To illustrate the importance of phenotypic bias, let us use a particularly simple form for the connectivity coefficients c_{pq}, assuming that connectivity decreases by a constant factor for each rank:

$$c_{pq} = C \exp(-k_q p) \tag{5.13}$$

Here C is a normalization constant and the constant $k_q \geq 0$ measures the strength of phenotypic bias around the source neutral space q ($k_q = 0$ corresponds to the absence of bias). We enforce the normalization $\sum_{p=1}^{\mathcal{E}_q} = 1$ to obtain C:

$$\begin{aligned} C \sum_{p=1}^{\mathcal{E}_q} e^{-k_q p} &= C e^{-k_q} \frac{1 - \exp(-k_q \mathcal{E}_q)}{1 - \exp(-k_q)} = 1 \\ \Leftrightarrow \quad C &= e^{k_q} \frac{1 - \exp(-k_q)}{1 - \exp(-k_q \mathcal{E}_q)} \end{aligned} \tag{5.14}$$

Using this form in Eq. (5.12) yields

$$\begin{aligned} f(t) &= 1 - C \sum_{p=1}^{E(t)} \exp(-k_q p) \\ &= 1 - C e^{-k_q} \frac{1 - \exp(-k_q E(t))}{1 - \exp(-k_q)} \\ &= 1 - \frac{1 - \exp(-k_q E(t))}{1 - \exp(-k_q \mathcal{E}_q)} \\ &\approx \exp(-k_q E(t)) \end{aligned} \tag{5.15}$$

where the last approximation is valid if $E(t) \ll \mathcal{E}_q$ and $\exp(-k_q \mathcal{E}_q) \ll 1$. The latter condition is the requirement of strong phenotypic bias, since $\exp(-k_q \mathcal{E}_q)$ is approximately the ratio of the smallest to the largest connectivity coefficient.

When we apply this approximation in Eq. (5.8), we obtain

$$\frac{dE}{dt} = \frac{\Delta E}{\Delta t} \exp(-k_q E(t)) \tag{5.16}$$

With exponential (local) phenotypic bias, we observe exponentially diminishing returns.

With the initial condition $E(t=0)=0$, Eq. (5.16) has the solution

$$E(t) = \frac{1}{k_q} \log\left(1 + \frac{\Delta E}{\Delta t} k_q t\right) \tag{5.17}$$

The complete solution of Eq. (5.8), using the exact form of Eq. (5.15), is

$$E(t) = \frac{1}{k_q} \log\left(1 + \mathcal{N}\left(1 - \exp\left(-\frac{k_q\, \Delta E}{\mathcal{N}} \frac{t}{\Delta t}\right)\right)\right) \tag{5.18}$$

where $\mathcal{N} = \exp(k_q \mathcal{E}_q) - 1$. At short times (that is, when $k_q \Delta E t/\Delta t \ll \exp(k_q \mathcal{E}_q)$), Eq. (5.18) converges to Eq. (5.17), provided $\exp(k\mathcal{E}_q) \gg 1$. So population evolvability increases only logarithmically with search time, even when $E(t) \ll \mathcal{E}_q$.

Under exponential phenotypic bias, we thus find a notably different scaling of evolvability with time than in the absence of bias, when only the finite potential evolvability $\mathcal{E}_q$ places a limit the linear growth of $E(t)$. We can compare our expressions with and without bias to understand when phenotypic bias will significant for population evolvability. For short times, we have in the absence of bias

$$E(t) \approx \frac{\Delta E}{\Delta t} t \left(1 - \frac{\Delta E\, t}{2\, \mathcal{E}_q\, \Delta t}\right) + O(t^3) \tag{5.19}$$

while under exponential bias

$$E(t) \approx \frac{\Delta E}{\Delta t} t \left(1 - \frac{k_q\, \Delta E\, t}{2\, \Delta t}\right) + O(t^3) \tag{5.20}$$

The first-order terms are clearly identical: For very short times, all non-neutral mutations lead to different phenotypes. But what do we mean by short times? The answer depends on the GP map: Without phenotypic bias, corrections are reduced by $1/\mathcal{E}_q$, which can well be of the order of the total number of phenotypes in the system, and hence the regime of short times can be rather long. By contrast, under exponential bias, the first correction to the linear increase with time scales with the bias parameter k_q, and short times may only mean the time taken for a single neutral fixation. Comparing Eqs. (5.19) and (5.20) shows

that the effect of phenotypic bias is important if $k_q \mathcal{E}_q \gg 1$; and again, since $\mathcal{E}_q$ is often large, even relatively small values of the bias parameter k_q may be significant.

Of course, the assumption of an exponential decay of phenotype connectivity with rank is quite idealized and will not be followed exactly in practice. On the other hand, it may often be possible to approximate real data by linear segments with different values of k_q. And since the sequential approximation only requires knowledge of the phenotypes discovered so far to predict the further increase of E, a piece-wise linear distribution of connectivity against rank will translate into segments of $E(t)$ according to Eq. (5.18) with different values of k_q for each segment. We will see an example of this result below.

We motivated our calculations in the limit of large monomorphic populations ($NL\mu \ll 1$, $N \gg L$). The results translate to other regimes if we use appropriate expressions for the exploration timescale Δt and the expected number of non-neutral mutants ΔE produced within Δt generations. The explicit expression for $E(t)$ in the regime of large, monomorphic populations follows from substituting Eqs. (5.4) and (5.5) into Eq. (5.18):

$$E(t) = \frac{1}{k_q} \log \left(1 + \mathcal{N} \left(1 - \exp \left(-k_q \frac{L^2 (K-1) \rho_q (1-\rho_q)}{\mathcal{N}} \mu t \right) \right) \right) \tag{5.21}$$

where $\mathcal{N} = \exp(k_q \mathcal{E}_q) - 1$ as before.

In the large genome limit $N \ll L$, the population will typically not sample all accessible mutants in a given mutational neighbourhood (cf. Section 4.2.3). While the exploration timescale is still given by the neutral fixation time, $\Delta t = \tau_f$, the evolvability gain is reduced by a factor of ξ (Eq. (4.15)) which accounts for the incomplete exploration of the current neighbourhood of the population. So we have $\Delta E = \langle \bar{n}_q \rangle \xi = N(1-\rho_q)/\rho_q$ and hence

$$\frac{\Delta E}{\Delta t} = NL\mu(1-\rho_q) \tag{5.22}$$

which is simply the rate of production of non-neutral mutants. Thus we obtain

$$E(t) = \frac{1}{k_q} \log \left(1 + \mathcal{N} \left(1 - \exp \left(-k_q \frac{NL\mu(1-\rho_q)t}{\mathcal{N}} \right) \right) \right) \tag{5.23}$$

Of course, this is also what we would expect in the polymorphic limit $NL\mu \gg 1$, where $\Delta t = 1$ (neutral exploration is fast) and $\Delta E = NL\mu(1-\rho_q)$ (all non-neutral mutants can in principle produce novel phenotypes).

Between these limiting regimes, we appeal to using γ (cf. Eq. (4.23)) to write the general expression

$$E(t) = \frac{1}{k_q} \log \left(1 + \mathcal{N} \left(1 - \exp \left(-k_q \frac{N\gamma L\mu(1-\rho_q)t}{\mathcal{N}} \right) \right) \right) \tag{5.24}$$

To summarize, we have calcuated the population evolvability based on the assumption that the more frequent phenotypes are discovered earlier than the rare phenotypes. This sequential approximation highlights the importance of phenotypic bias, which leads to quickly diminishing returns of neutral exploration: After the relatively fast discovery of the easily accessible phenotypes, it can take very long to find the rarer phenotypes.

5.3 Simulations

We test our predictions for the population evolvability using the simulation results of Section 4.3.1. Fig. 5.1 shows a good qualitative agreement with the simulations. In particular, we find unequivocal evidence for the reduction of population evolvability due to phenotypic bias. Fig. 5.1b shows results for landscape C where phenotypic bias is indeed exponential with rank ($k_q = -\log 1.25 \approx 0.22$), and the slow logarithmic increase of population evolvability with time is clearly visible. Fig. 5.1d displays data from landscape A with RNA phenotype frequencies. While phenotypic bias is not log-linear over the entire distribution (cf. Fig. 4.2), we still find that evolvability increases only logarithmically with time, but that the rate of increase varies over time.

Fig. 5.1b shows that for short times, $E(t)$ follows the prediction of the continuity equation, even when μ is very small and the population is monomorphic. This effect is again due to the exploration of the initial mutational neighbourhood of the population, as discussed in Section 4.2.5: This neighbourhood contains on the order of $(K-1)L(1-\rho_q)$ genotypes with alternative phenotypes. Once all phenotypes in the immediate neighbourhood have been discovered, no novel phenotypes are found before a neutral mutant is fixed. The sequential

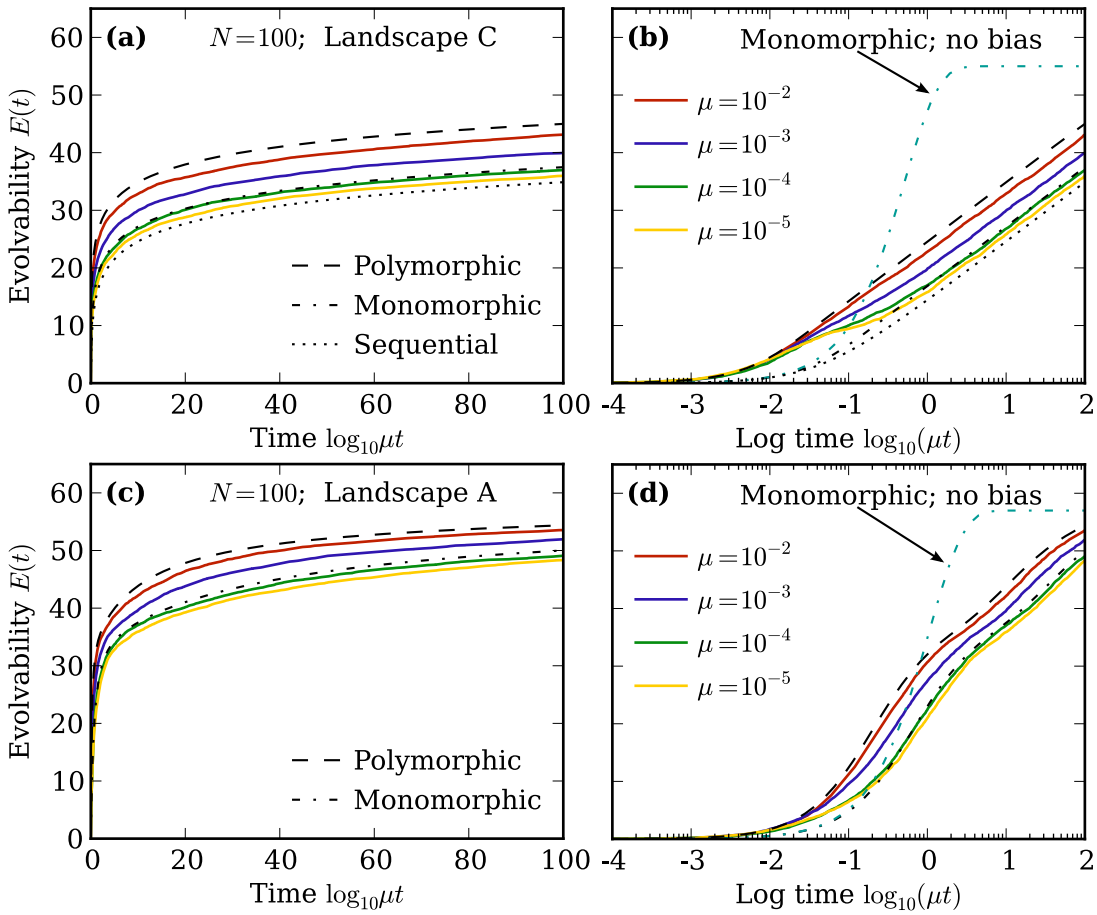

Figure 5.1: Phenotypic bias reduces evolvability. The diagrams show the population evolvability for population size $N = 100$ for landscape C (panels **(a)** and **(b)**) and A (panels **(c)** and **(d)**). Solid lines correspond to simulation data, with colours indicating mutation rates. Black lines show our predictions: The polymorphic limit (Eq. (5.2) with $\gamma = 1$) is dashed; the monomorphic limit (Eq. (5.2) with $\gamma = (K-1)L\rho/N$) is dash-dotted; in **(a)** and **(b)**, the dotted lines correspond to the sequential approximation in the monomorphic limit (Eq. (5.21)) with $k_q = -\log 1.25 \approx 0.22$. The grey dash-dotted line in **(b)** and **(d)** corresponds to the monomorphic limit in the absence of phenotypic bias (Eq. (5.11)). All results are averages over varying initial conditions.

approximation uses the neutral fixation time as the smallest dynamic time-scale, and hence its predictions are not accurate at very short times.

It is also worth noting that although the source phenotype q has higher robustness in landscape A ($\rho_q = 0.85$) than in landscape C ($\rho_q = 0.5$), even in the monomorphic limit (in which $\Delta E/\Delta t \propto \rho_q(1-\rho_q)$) the impact of robustness is clearly secondary to the influence of the GP map through the connectivity distribution c_{pq}: On landscape C, evolvability is consistently higher than on landscape A (cf. Figs. 5.1a and 5.1c) when $E(t)$ is greater than about 20, since the c_{pq} are larger in landscape A than in landscape C beyond rank 20.

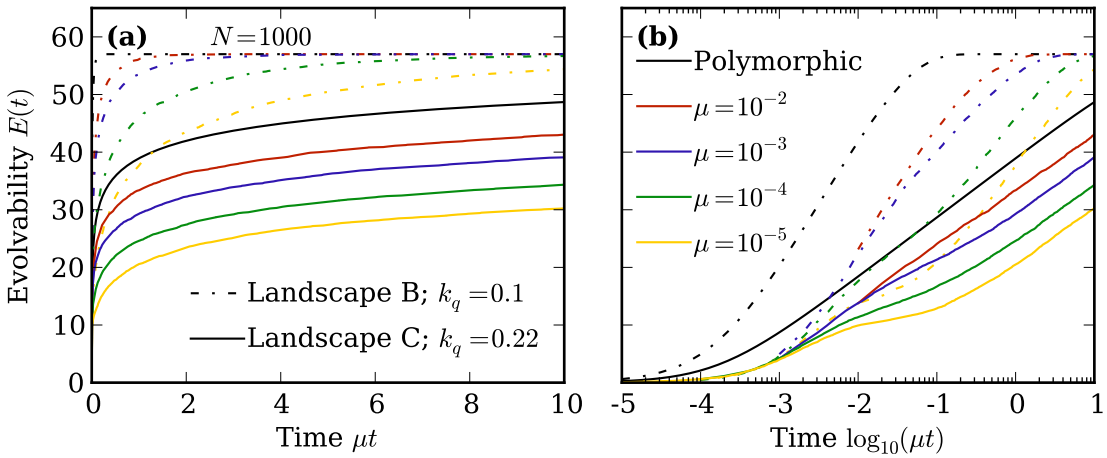

Figure 5.2: Evolvability depends on the GP map. The diagrams compare evolvability on landscapes B (dashed lines, $k_q = 0.10$) and C (solid lines, $k_q = 0.22$), see Fig. 4.2. Both landscapes have $\rho = 0.5$, and the population sizes was held fixed at $N = 1000$. Colours indicate mutation rates; the black lines correspond to the continuity approximation, Eq. (5.23). **(a)** shows time rescaled by mutation rate μ, and in **(b)** μt is shown on a log-scale.

All these findings underline that evolvability depends on the entire GP map. A case in point is provided by Fig. 5.2 which compares the results on landscapes B and C, which have practically identical values of ρ_q but show different strengths of phenotypic bias. Clearly, a population evolving on landscape B ($k_q = 0.1$) can produce a large variety of alternative phenotypes more quickly than if it were to evolve on landscape C (where $k_q = 0.22$).

Finally, we can compare the evolvability of populations of different sizes in Fig. 5.3. Measured in absolute time, larger populations are of course more evolvable (see Fig. 5.3a). But if we measure time in reproduction events Nt, Fig. 5.3b shows that smaller populations tend to be more evolvable in the long run. The disadvantage of large population stems from the fact that strong genetic drift can cause a greater loss of genotypic diversity: The maximum diversity of the population is bounded by the population size, and hence drift causes a greater reduction when the population is large. Also, as we shall discuss more explicitly in the next chapter, large monomorphic populations tend to produce the same mutants repeatedly. But these secondary mutants do not contribute to the population evolvability, thus reproduction events producing such mutants are 'wasted effort' from the viewpoint of evolvability.

Fig. 5.3b also allows us to study the effect of standing genotypic variation for the discovery

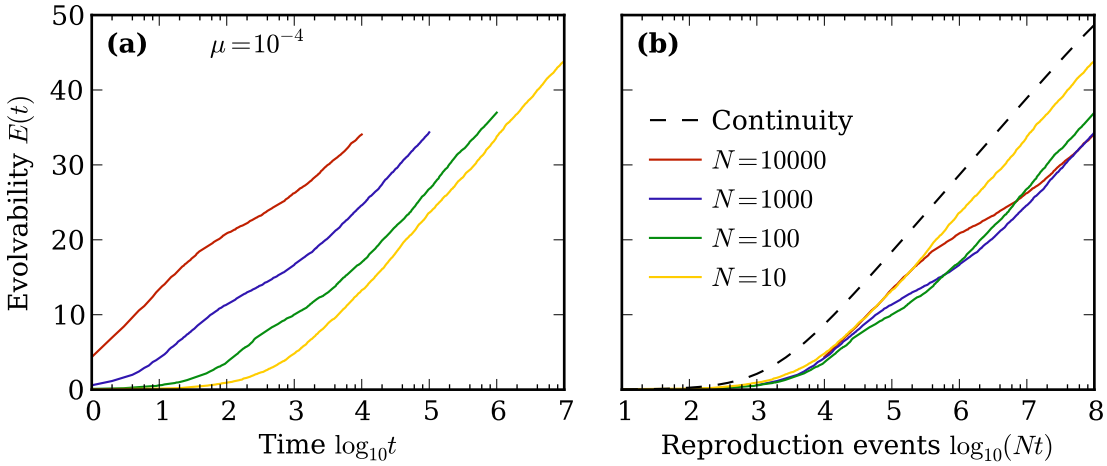

Figure 5.3: Effect of population size on evolvability. The diagrams compare evolvability for populations of different sizes. Simulations were performed on landscape C with mutation rate $\mu = 10^{-4}$ and varying initial conditions. Solid lines show population evolvability for different population sizes, ranging from polymorphic populations (red line, $N = 10000$, so $NL\mu = 12$) to large monomorphic populations (green, $N = 100$, so $NL\mu = 0.12$) and the large genome limit (yellow, $N = 10$). Panel **(a)** shows evolvability against bare time, while in **(b)** the x-axis shows the number of reproduction events Nt; here, the black dashed line corresponds to the sequential solution in the continuity approximation, Eq. (5.23) with $k_q = 0.22$.

of novel phenotypes. We can view the large genome limit $N \ll L$ (illustrated by the yellow curve with $N = 10$) as a benchmark, since in this regime neutral exploration is so fast that the localization under genetic drift does not delay the discovery of novel phenotypes (cf. Section 4.2.3). Shortly after the environmental shift at $t = 0$, we see that the probability of a given reproduction event to discover a new phenotype is independent of N: The curves E against Nt overlap closely. But as time progresses, the evolvability of large populations deviates from the large genome limit. The point of departure marks the time when the initial neighbourhood of the population has been completely explored and the discovery of novel phenotypes requires exploration. As we can see from Fig. 5.3b, the phase of initial exploration is longer (in terms of reproduction events) for larger populations, since at fixed mutation rate, they show a larger standing genotypic variation (cf. Eq. (2.31) and Section 4.2.5). In the long run, however, we find that smaller monomorphic populations are more efficient (per reproduction) at discovering novel phenotypes, simply because they produce less (wasted) mutants between neutral fixations.

In summary, our theoretical arguments concerning population evolvability do capture the most interesting features of actual simulation results. In particular, we can demonstrate that (local) phenotypic bias reduces evolvability: Most non-neutral mutations lead to phenotypes that have already been discovered. As our theory does not describe the general implications of genetic drift, the quantitative accuracy of our predictions is not perfect, and could be improved by a more explicit treatment of the genotypic dynamics of the population.

5.4 Summary and discussion

In this chapter, we proposed a measure of evolvability that accounts of the dynamics neutral exploration: Instead of counting the potentially accessible phenotypes, the population evolvability counts the number of phenotypes actually discovered.

We employed the mean-field description of Chapter 4 to derive an analytic prediction for the population evolvability. The resulting expression incorporates the details of the GP map and relates to Wagner's potential evolvability [127] as well as the diversity evolvability defined by Cowperthwaite et al. [24]. Intuitively, potential evolvability emerges as the long-time limit of the population evolvability: When the population explores the entire neutral space, all accessible phenotypes are discovered eventually. By contrast, diversity evolvability relates to the short-time behaviour of population evolvability, describing the deviation from an initial linear growth when all non-neutral mutants produce novel phenotypes, to a phase of slower growth when most non-neutral mutations lead to phenotypes already discovered.

Our prediction for the population evolvability under a general GP map, Eq. (5.2), contains a sum over exponentials decaying with different scales, which are determined by the GP map. While accurate, this is provides limited analytic insight, particularly into the role of phenotypic bias. For an explicit account of the impact of this bias, we calculated the population evolvability using the sequential approximation, where we assumed that phenotypes are discovered in the order of their connectivity. In this scheme, the discovery of novel phenotypes suffers from quickly diminishing returns. In the special case of exponential phenotypic bias (connectivity decays by a constant factor from one phenotype to the next), we

showed that population evolvability increases only logarithmically in time (Eq. (5.18)). This contrasts with GP maps which do not show phenotypic bias, where population evolvability increases linearly with time up to an exponential saturation at very long times (Eq. (5.10)).

The main message to take away from these results is that population evolvability is a property of the entire GP map. It therefore appears that we cannot simply predict whether robustness aids or hinders evolvability in general. It is always true that a single mutation is less likely to produce a novel phenotype when the parental genotype is more robust to mutations. But this argument fails to capture that neutral exploration is generally faster at finding new phenotypes than the crossing of a fitness valley [121, 126]. However, phenotypic bias adds another perspective to these arguments: Rare phenotypes are hard to find for neutrally searching populations; only the frequent phenotypes are likely to be discovered. So the potential benefit of robustness may not necessarily play out in practice.

Of course, it is important to ask how these results change with increasing L to more biologically realistic values. As we saw in Chapter 3, the total number of phenotypes can grow exponentially in L; on the other hand, the size of the mutational neighbourhood of each genotype only increases linearly in L. Together, these scaling relations imply that the majority of phenotypes may typically *not* be immediately accessible to the population. Therefore, it seems quite likely in systems with larger genome length L (or effective genome length L_{eff}, as discussed in Section 4.4), the slow speed of neutral exploration discussed in this chapter will pose an even more important obstacle for the introduction of phenotypic variation.

So far, we have only been concerned with the first discovery of alternative phenotypes. Non-neutral mutants with an already discovered phenotype therefore do not contribute to population evolvability; this is the origin of the diminishing returns on neutral exploration. This neglect of secondary mutants is valid when their phenotype is strongly deleterious, as we have assumed so far. In the next chapter, we will alleviate this assumption and instead consider how neutral spaces affect the fixation of beneficial mutants.

Chapter 6

Neutral spaces and adaptation

6.1 Introduction

The previous two chapters revolved around the question how the GP map determines the introduction of phenotypic variation. This variation provides the fuel for natural selection. So far, we have assumed that selection is only stabilizing – alternative phenotypes are unviable and hence are quickly removed from the population. But if selection was only ever stabilizing, there would be no phenotypic change. In this chapter, we relax the assumption of stabilizing selection and study how the existence of neutral spaces affects adaptation.

In agreement with many earlier studies, we have seen that a population can sample a greater variety of alternative phenotypes by exploring a neutral space than it could reach from a single genotype [65, 127, 128]. But how does the availability of neutral mutations (that is, mutational robustness) affect the chances of a beneficial phenotype actually being fixed? This is the question we will address here.

As we have seen in Section 2.1, the fixation of a single adaptive mutant in a finite population is not guaranteed. In particular, when the selective advantage s is rather small, it is much more likely that a single mutant is lost than that it goes to fixation. On the other hand, we found in Chapter 4 that mutants with particular phenotypes are produced repeatedly. Clearly, with repeated attempts the chance of fixation increases.

We begin our investigation by studying the time-scales over which mutants with a par-

ticular phenotype are produced, in the limit of monomorphic populations. This regime is characterized by long periods of genotypic stasis: All individuals carry the same genotype, and mutants explore the immediate genetic neighbourhood of that genotype. A neutral mutant only rarely goes to fixation, thereby taking the population from one genotype to another.

In Section 6.2, we study the spectra of mutants produced during a period of stasis. Since only phenotypes in the immediate neighbourhood of the population can be reached, phenotypic variation is limited to this neighbourhood. In large populations, this implies that there are typically many more mutants than there are alternative phenotypes, and the same phenotypes are thus produced in relatively quick succession. The magnitude of these 'bursts' of mutants with the same phenotype depends on the population size: In small populations, which are dominated by fluctuations (cf. Section 4.2.3), bursts occur only rarely. But in large populations, the signature of bursts can be very strong.

We then investigate implications of bursts for the fixation of a beneficial phenotype in Section 6.3. We quantify the intuitive notion that if there are more attempts at fixing the phenotype due to a burst of mutants, the overall chance of fixation can increase significantly. By now, we are familiar with the fact that stasis times are shorter at genotypes of higher robustness (cf. Eq. (4.12)). Consequently, we find that the probability of fixing an accessible phenotype decreases with the robustness of the 'portal' genotype. Especially in small populations, it is possible that the population leaves the portal genotype too quickly for a successful adaptation. On the other hand, in large populations the probability of fixing an accessible adaptive phenotype is quite large, even if the selective benefit of the phenotype is relatively small.

These results show that the existence of a neutral space (which enters the calculations via its robustness) can act to prevent, or at least significantly slow down, adaptation. Overall, the existence of bursts and their implications for fixation highlight the impact of neutral spaces not only for neutral exploration, but also for adaptive evolution.

6.2 Bursts of similar mutants

6.2.1 Theory

We have seen in Chapter 4 that the total number of mutants with some fixed phenotype p produced over sufficiently long times is well-described by our continuity approximation. In particular, we do not need to include the drift-factor γ to obtain the cumulative number of mutants. On the other hand, the first discovery of all but the most frequent phenotypes is delayed by the slow exploration under genetic drift, and introducing γ is necessary.

This distinction in the necessity to include γ (or not) is intriguing. As the first discovery is delayed by drift, the time to produce the remaining $M_p(T)-1$ secondary mutants is shorter than the continuity approximation suggests. As a consequence, the rate of production of p-mutants after the first discovery must be larger than the continuity approximation's value $NL\mu(1-\rho_q)c_{pq}$: Genetic drift delays the first discovery of rare phenotypes, but it speeds up the production of secondary mutants.

Let us formalize this argument. As usual, the analytic treatment is simplest in the monomorphic limit $NL\mu \ll 1$, when the population is localized on a single genotype g. The time spent at g – the *stasis* time – is distributed exponentially[1] with mean $\tau_f = (L\mu\rho_g)^{-1}$ (cf. Eq. (4.12)) where ρ_g is simply the robustness of g.

Between two neutral fixations, we evidently have constant PPRs, determined from the mutational neighbourhood of g. Thus the time taken to produce a p-mutant from g is simply given by the local analogue of the continuity approximation (Eq. (4.11))

$$\tau_{p,g} = \frac{1}{NL\mu\frac{n_{p,g}}{(K-1)L}} = \frac{K-1}{N\mu n_{p,g}} \tag{6.1}$$

where $n_{p,g}$ is the number of p-genotypes accessible from g. So when $\tau_{p,g} \ll \tau_f$, we expect that many mutants with phenotype p are produced in relatively quick succession. The expected

[1] As usual, this corresponds to a continuous time approximation. In discrete time, the distributed is geometrical with the same mean as in the continuous case.

total number of p-mutants produced while the population is at g is simply

$$M_{p,g}(T) = \frac{T}{\tau_{p,g}} = N\mu\frac{n_{p,g}}{K-1}T \tag{6.2}$$

where T is the time the population spends at g. The average $\langle M_{p,g}\rangle$ of $M_{p,g}$ over the distribution of stasis times T (whose average is the time between neutral fixation τ_f, Eq. (4.12)) is simply given by

$$\langle M_{p,g}\rangle = M_{p,g}(\tau_f) = \frac{N\mu n_{p,g}}{L\mu\rho_g(K-1)} = N\frac{n_{p,g}}{n_{q,g}} \tag{6.3}$$

where $n_{q,g}$ is the number of neutral neighbours of the current genotype, analogous to $n_{p,g}$. The ratio $n_{p,g}/n_{q,g}$ is simply the mean number of p-mutants that arise while a single q-mutant is produced, and a single neutral fixation requires on average N mutants with the source phenotype q (cf. Section 2.1.2). We note that $\langle M_{p,g}\rangle$ decreases with ρ_g (or $n_{q,g}$): The more neutral mutations are available, the faster a neutral mutant is fixed, and when the population spends less time at a fixed genotype g, it produces less of g's neighbour genotypes by mutation.

Finally, we note two useful identities involving $\langle M_{p,g}\rangle$. First, since $M_{p,g}(T)$ scales linearly with T, we have

$$M_{p,g}(T) = \langle M_{p,g}\rangle\frac{T}{\tau_f} \tag{6.4}$$

Second, we have the relation

$$\langle M_{p,g}\rangle = \xi n_{p,g} \tag{6.5}$$

where $\xi = N/n_{p,g} = N/((K-1)L\rho_g)$ is the ratio of the fixation time to the local exploration time, in analogy to Eq. (4.15) (except that we are dealing here with the current genotype's robustness ρ_g instead of the average ρ_q). Just as in Section 4.2.3, we see that in the large genome limit $\xi \ll 1$, the production of any particular mutant is highly unlikely, while in the large population limit $\xi \gg 1$ all accessible mutants are copiously produced.

As genotype space is discrete, $n_{p,g}$ can only take integer values. This contrasts with the mean-field connectivity c_{pq}, which is practically continuous: Since c_{pq} is an average over the entire neutral space, the discretisation scale of c_{pq} is on the order of $1/|\mathcal{G}_q|$, where $\mathcal{G}_q$ is the

(typically vast) set of genotypes in the neutral space of the source phenotype q. So provided p is accessible from the current genotype g (that is $n_{p,g} > 0$) and p is not accessible from every genotype in the neutral space (that is $\langle n_{pq} \rangle = (K-1)L(1-\rho_q)c_{pq} \ll 1$), we have $n_{p,g} \gg \langle n_{pq} \rangle$. Thus the instantaneous production rate of p is much greater than its average: Drift leads to strong correlations ('bursts') in the spectrum of phenotypes produced over relatively short times. In other words, the time taken to produce the second mutant $\tau_{p,g}$ *given that a first mutant has been produced* is usually much shorter than the discovery time of the first mutant T_p.

Bursts are fundamentally a consequence of the discreteness of genotype space. Thus the prediction of the continuity approximation, in which this discrete structure is ignored, provides a useful null model against which we can compare our simulations. Since in the continuity approximation, the expected number $M_p(T)$ of p-mutants over T generations follows a binomial (or approximately a Poisson) distribution, we expect fluctuations of order $\sigma_p = \sqrt{M_p(T)}$ as the standard deviation of a Poisson distributed random variable is just the square-root of its expectation value. Since we consider large values of T (on the order of millions of generations), we further approximate the Poisson distribution by a Gaussian; thus we expect that with probability of 95%, the observed number of p-mutants over time T should fall into the range $M_p(T) \pm 2\sqrt{M_p(T)}$. The number $P_{2\sigma}$ of phenotypes for which the observed number of mutants is outside this 2σ-interval provides a simple measure of the severity of bursts. Note that we should not expect $P_{2\sigma} < \mathcal{E}_q/20$: The potential evolvability $\mathcal{E}_q$ (Eq. (2.23)) counts the number of accessible phenotypes, and for each phenotype we allow for a 5% chance of falling outside the 2σ range.

6.2.2 Simulations

As before, we simulate evolutionary dynamics on the uncorrelated landscapes of Section 4.3.1. Throughout this section, we shall study only landscape B where – albeit clear phenotypic bias – even the least frequent phenotypes have a reasonable chance of being discovered.

A difficulty in identifying bursts has to do with timescales. Bursts are a time correlation

in the number of mutants produced with a particular phenotype, and our theory suggests that bursts are most pronounced in the monomorphic limit $NL\mu \ll 1$. But in this regime, there are no mutation events in most generations. So when a p-mutant has been produced in generation t, we would not expect another p-mutant in generation $t+1$. Rather, we would expect to find the next mutant with phenotype p around $\tau_{p,g}$ generations later. Thus in order to identify bursts, we need to analyse the number of p-mutants produced over a coarse-grained time-scale τ with $\tau_{p,g} \ll \tau \ll \tau_f$. Clearly, this is only sensible when $\tau_{p,g} \ll \tau_f$, that is when genetic drift is strong and populations are large.

Fig. 6.1 shows the number of mutants of each phenotype produced, taken from a single simulation. In both panels, $\mu = 10^{-6}$ so that $\tau_f \approx (L\mu\rho_q)^{-1} \approx 1.6 \times 10^5$. Fig. 6.1a has $N = 1000$ and hence $\tau_{p,g} = (K-1)/(N\mu) = 3000$ (when $n_{p,g} = 1$; greater $n_{p,g}$ only decreases this). So when coarse-graining the time series data into bins of width $\tau = 10000$ generations, we observe clear signatures of bursts: For $P_{2\sigma} = 48$ out of 57 accessible alternative phenotypes, the observed number of mutants over 10^6 generations differs by more than two standard deviations from the predictions of the continuity approximation: The discreteness of genotype space leads to strong correlations between mutants produced over relatively long times in large monomorphic populations.

By contrast, we have $N = 100$ in Fig. 6.1b, which increases $\tau_{p,g}$ to 30000. On a coarse-graining scale of $\tau = 10^5$, we still observe bursts. But both stasis times and local exploration times are distributed exponentially, and hence fluctuations are large (the standard deviation is equal to the mean). So the visual signature of bursts is less clear. Nonetheless, we find $P_{2\sigma} = 33$: For more than half of the accessible 57 phenotypes, the continuity approximation does not give a very reliable prediction for the number of mutants produced in 10^7 generations.

Fig. 6.1 compares populations of different sizes, but with the same mutation rate. Thus τ_f is the same for both populations while $\tau_{p,g}$ varies. Instead in Fig. 6.2, we keep $N\mu$ fixed (at 10^{-3}, the value used in Fig. 6.1a) and thus vary N and μ simultaneously. So $\tau_{p,g}$ remains constant and we use a coarse-graining timescale $\tau = 10000$. In Fig. 6.2a, we have $N = 100$ and $\mu = 10^{-5}$. This implies $\tau_f \approx 16000$ which is still greater than τ, albeit

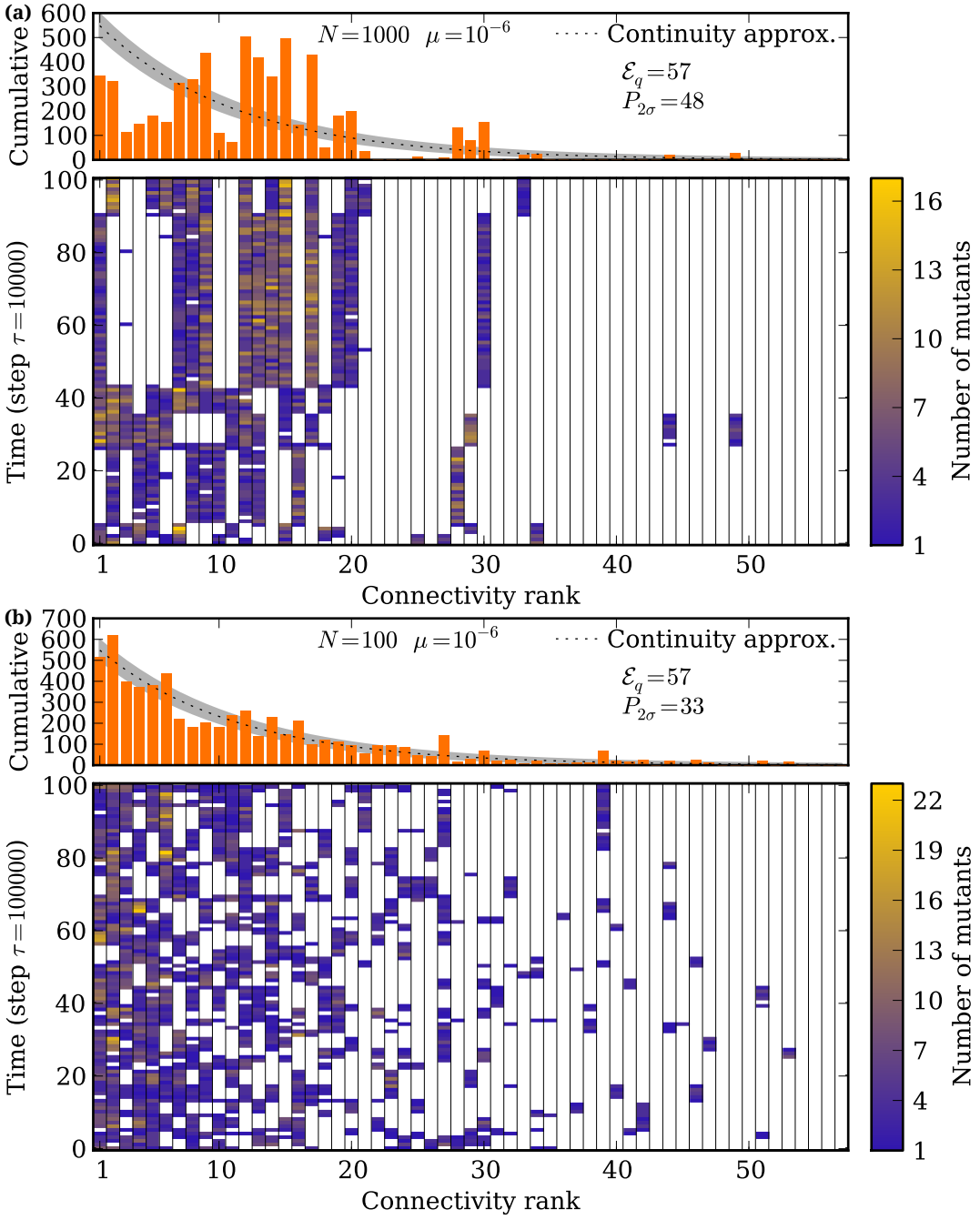

Figure 6.1: Genetic drift induces correlation in the spectrum of mutant phenotypes. (a) The upper panel shows the total number $M_p(T)$ of mutants for each phenotype p, produced over $T = 10^6$ generations. The black dotted line shows the prediction of the continuity approximation (Eq. (4.10)), and the grey shaded region shows the 95% confidence interval $M_p(T) \pm 2\sqrt{M_p(T)}$. $P_{2\sigma}$ is the number of phenotypes for which the total number of mutants is not within this interval, and $\mathcal{E}_q$ is the total number of accessible phenotypes. The lower panel shows the time-series of mutant production during the simulation. The colour of each block indicates the number of mutants with a particular phenotype, produced over $\tau = 10^4$ generations; white blocks indicate that no mutant with the respective phenotype has been produced. The data corresponds to a single simulation with $N = 1000$, $\mu = 10^{-6}$ on landscape B. **(b)** Analogous results for $N = 100$, $\mu = 10^{-6}$. Due to the smaller N, $\tau = 10^5$ is used for the length of the time interval. Mutations to the source phenotype q are not shown.

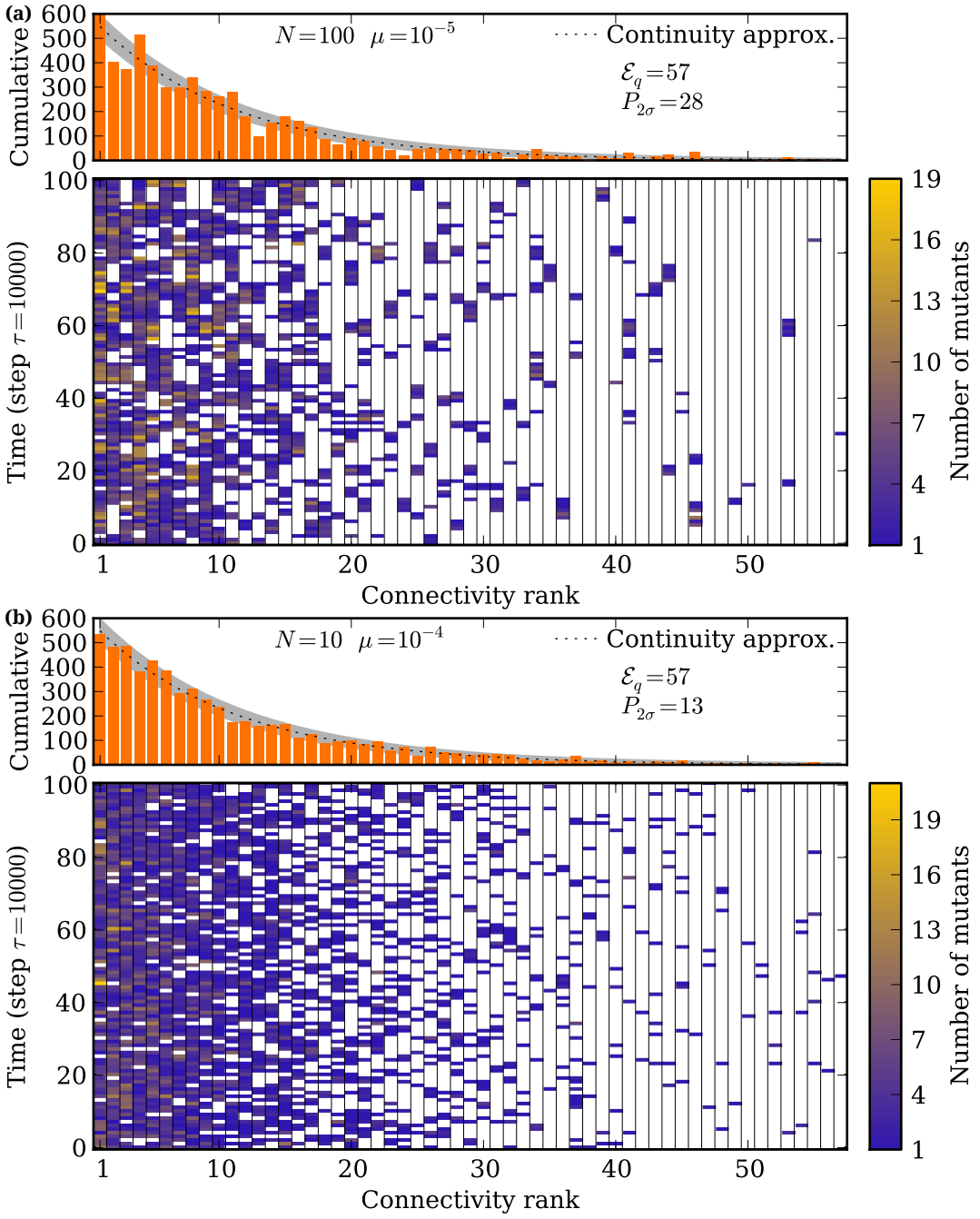

Figure 6.2: Bursts occur more frequently in large populations. The diagrams are analogous to Fig. 6.1, except that the population dynamic parameters are different, but $N\mu = 10^{-3}$ is fixed. **(a)** shows data for $N = 100$, $\mu = 10^{-5}$. **(b)** shows results for $N = 10$, $\mu = 10^{-4}$. The time interval length is $\tau = 10^4$ throughout.

not very much. Since the distribution of stasis times is exponential, $1 - e^{-\tau/\tau_f} \approx 45\%$ of fixations occur within a single time-step on our coarse-grained scale. As a consequence, the visual signature of bursts is less pronounced. Still, over the 10^6 generations displayed in the diagram, the agreement with the continuity approximation is good only for half of the 57 accessible phenotypes ($P_{2\sigma} = 28$).

Finally, in Fig. 6.2b, we have $N = 10$ and $\mu = 10^{-4}$ so that $\tau_f \approx 1600 < \tau_{p,g}$. This is the large-genome limit, where fluctuations dominate and the exploration of the current mutational neighbourhood is typically incomplete. Here, bursts of rare phenotypes are almost completely absent in a visual observation. Instead, the time between individual mutants with the same phenotype is usually quite large, on the order of the first discovery time. Again, in the fluctuation-dominated regime, discovery times are distributed exponentially/geometrically and hence fluctuate strongly. Over the complete 10^6 generations shown, the agreement with the continuity approximation is good for 44 phenotypes ($P_{2\sigma} = 13$). Nonetheless, $P_{2\sigma}$ is still relatively large: Observations should fall into a 2σ with 95% probability, and hence we should expect $P_{2\sigma} = 57/20 \approx 3$. So even in the large genome limit, the discrete nature of genotype space can lead to non-trivial dynamic correlations.

The better agreement with the continuity approximation with reduced N is not due to an increased number of mutants over the simulation (which should always bring us closer to the expectation): We have kept $NL\mu T$ fixed through Figs. 6.1a, 6.2a and 6.2b. Instead, the closer match of theory and simulation is caused by the faster time-scale of neutral exploration as we increase μ (to keep $N\mu$ fixed): The smaller, faster mutating populations explore a larger part of the neutral space. This is also reflected by the evolvability of the populations: While the large population ($N = 1000$) only discovered 30 alternative phenotypes, the intermediate population ($N = 100$) found 50 of them and only the small population produced all 57.

So as long as we are only interested in the discovery of alternative phenotypes, correlations between mutants are clearly a disadvantage: During the bursts, the same phenotype is produced repeatedly in quick succession and most non-neutral mutants do not discover novel phenotypes. So far, we treated all alternative phenotypes as strongly deleterious and were interested in the diversity of phenotypes produced. Per reproduction event then, small and

quickly mutating populations have an advantage. But if some alternative phenotype were to impart a selective benefit, a large flux of mutants of that phenotype should aid its fixation. We turn to the implications of bursts for fixation in the next section.

6.3 Fixation

6.3.1 Theory

Due to the stochasticity of reproduction in finite populations, the fixation of a single adaptive mutant is not guaranteed (cf. Section 2.1.2). But as we have just seen, genetic drift leads to correlations in the production of mutants over relatively short time-scales: There may be multiple attempts at fixing p in relatively quick succession.

To understand the impact of bursts, let us calculate the probability that a mutant 'destined to be fixed' is created during a single burst. For simplicity, we shall again assume $NL\mu \ll 1$: This means that the fate of each mutant is decided independently, and we may ignore the possible complications of clonal interference. Thus the probability that a particular mutant will go to fixation is $P_1(s) = (1 - \exp(-2s)/(1 - \exp(2Ns)$ (Eq. (2.15)) where the fitness of the advantageous phenotype p is $f_p = 1 + s$. With a total of m mutants with phenotype p during a burst, the probability that this burst will *not* lead to the fixation of p is

$$\Pr(\text{None of } m \text{ mutants fixed}) = \Big(1 - P_1(s)\Big)^m$$

For a given stasis time T_s, the probability that exactly m mutants are produced follows a Poisson distribution with mean $M_{p,g}(T_s) = NL\mu(1 - \rho_g)c_{p,g}T_s$ (Eq. (6.2)):

$$\Pr(m|T_s) = \frac{M_{p,g}(T_s)^m}{m!} \exp(-M_{p,g}(T_s))$$

Finally, we can use the exponential stasis time distribution $\Pr(T_s = T)dT = \exp(-T/\tau_f)dT/\tau_f$

to obtain the fixation probability

$$\begin{aligned}\Phi(s) &= 1 - \int_0^\infty dT \Pr(T_s = T) \Pr(m|T) \Pr(\text{None of } m \text{ mutants fixed}) \\ &= 1 - \int_0^\infty \frac{dT}{\tau_f} \exp\left(-\frac{T}{\tau_f}\right) \sum_{m=0}^{\infty} \frac{M_{p,g}(T)^m}{m!} \exp(-M_{p,g}(T)) \Big(1 - P_1(s)\Big)^m\end{aligned}$$

which gives the probability that p is fixed, *provided the phenotype is accessible* from the current genotype. Note that this is quite different in spirit from $P_1(s)$, which gives the probability that a single mutant of phenotype p is fixed, *provided the mutant exists* in the population.

The summation over m simply yields $\exp\Big((1-P_1(s))M_{p,g}(T)\Big)$ which cancels the $\exp(-M_{p,g}(T))$ factor. Thus we obtain

$$\begin{aligned}\Phi(s) &= 1 - \int_0^\infty \frac{dT}{\tau_f} \exp\left(-\frac{T}{\tau_f}\right) \exp\Big(-P_1(s)M_{p,g}(T)\Big) \\ &= 1 - \int_0^\infty \frac{dT}{\tau_f} \exp\left(-\frac{T}{\tau_f}\right) \exp\Big(-P_1(s)\langle M_{p,g}\rangle \frac{T}{\tau_f}\Big) \\ &= 1 - \frac{1}{1 + P_1(s)\langle M_{p,g}\rangle} \\ &= \left(1 + \frac{1}{P_1(s)\langle M_{p,g}\rangle}\right)^{-1} \\ &= \left(1 + \frac{1}{P_1(s)\xi n_{p,g}}\right)^{-1} \\ &= \left(1 + \frac{n_{q,g}}{P_1(s)N n_{p,g}}\right)^{-1} \qquad (6.6)\end{aligned}$$

This result shows several interesting features: First, the fixation probability increases with population size, because in a larger population, more beneficial mutants will be produced before a neutral mutant is fixed. For adaptive phenotypes that have a selectable effect ($Ns > 1$), this increase means that the actual value selection coefficient s may be of relatively little consequence for the fixation of the phenotype.

Second, robustness $\rho_g = n_{q,g}/((K-1)L)$ of the source phenotype reduces the probability of adaptation, as the population may move away from the access genotype g. It is important to realize that this penalty due to robustness arises in addition to the fact that a high

robustness necessarily lowers the chance of non-neutral mutations, which we can express here in terms of the inequality

$$n_{p,g} \leq (K-1)L - n_{q,g}$$

Third, we see that the number of available beneficial mutations $n_{p,g}$ is just as important as the probability of fixing a single mutant. This implies that more accessible phenotypes may be easier to fix. Finally, we observe that for neutral mutants ($s = 0$ so $P_1(s) = 1/N$), we recover $\Phi(s) = n_{p,g}/(n_{p,g} + n_{q,g})$ which is just the fraction of neutral neighbours of g which map into p (since $s = 0$, the genotypes mapping into p are also neutral neighbours of g).

It is worth noting that $\Phi(s)$ need not necessarily be greater than $P_1(s)$. It is easy to see that $\Phi(s) > P_1(s)$ only if

$$P_1(s) < 1 - \langle M_{p,g} \rangle^{-1} = 1 - \frac{1}{\xi n_{p,g}} \tag{6.7}$$

So when $\xi \gg 1$ (that is, in the regime of large monomorphic populations), bursts typically will increase the probability of fixation, particularly when the adaptive advantage of the alternative phenotype p is relatively small (but still greater than $1/N$). But if $\xi \ll 1$ (that is in the fluctuation-dominated regime of large genomes), the conditional fixation probability $\Phi(s)$ is typically smaller than the probability $P_1(s)$ of fixing an single existing mutant, simply because it is quite likely that no advantageous mutant is actually produced.

6.3.2 Simulations

In order to check our prediction for the conditional fixation probability, we need to seed the population on a portal genotype. As we have two control static parameters – the number of accessible target genotypes $n_{p,g}$ and the number of accessible source genotypes $n_{q,g}$ – we construct our landscapes in the following way: First, we calculate the desired robustness $\rho = n_{p,g}/((K-1)L)$ and distribute $K^L\rho$ source genotypes randomly in genotype space.

Next we pick one of these source genotypes as the portal genotype g. Among its non-neutral neighbours, we randomly place the $n_{p,g}$ target genotypes, which all carry fitness $1+s$. Finally, if g does not have the desired $n_{q,g}$ neutral neighbours exactly, we add the missing neutral neighbours randomly from the set of non-neutral, non-target neighbours (or remove the surplus genotypes randomly from the set neutral neighbours).

Once the landscape is fixed, we seed N individuals with genotype g – the advantageous phenotype is not initially present in the population but must be discovered before it may be fixed. To facilitate comparison for populations of different sizes, we scale mutation rates and fitness advantages with $1/N$ by setting $N\mu = 5 \times 10^{-3}$ and considering $Ns = 2, 5, 10$. As before, we use $L = 12$ and $K = 4$.

The population evolves either until fixation (all individuals express the adaptive phenotype), or until the adaptive phenotype is no longer in the 1-mutational neighbourhood of the population. This second condition is slightly different to our theoretical calculations, where we assumed that the advantageous phenotype would no longer be accessible after a neutral fixation. For practical reasons, we stop simulations after $5\tau_f$ generations; if p has not been fixed within this time, we count the simulation as one where p did not fix. The probability of a stasis time greater than $5\tau_f$ is $\exp(-5) \approx 0.7\%$, and thus the error associated with the cutoff is negligible. In practice, less than 0.1% of our simulations stopped because they ran out of time.

Fig. 6.3 shows the outcomes of 1000 simulations per parameter set for populations of size 10 (large genome limit, $\xi \ll 1$) and 100 (large population limit, $\xi \gg 1$). Generally, the agreement of our theory with the simulations is quite satisfactory. In particular, the negative impact of the robustness of the portal genotype is clear: For all landscape parameter combinations, the correlation of $\Phi(s)$ with ρ is $r < -0.8$, with p-values consistently below 3%.

In the large population, the positive effect of bursts on the fixation probability is clearly visible (Fig. 6.3a): We always find $\Phi(s) > P_1(s)$. This is because for $N = 100$, it is very likely that several mutants with the adaptive phenotype arise while the phenotype is accessible. By contrast, in the smaller population ($N = 10 < L = 12$), it can be quite likely that no

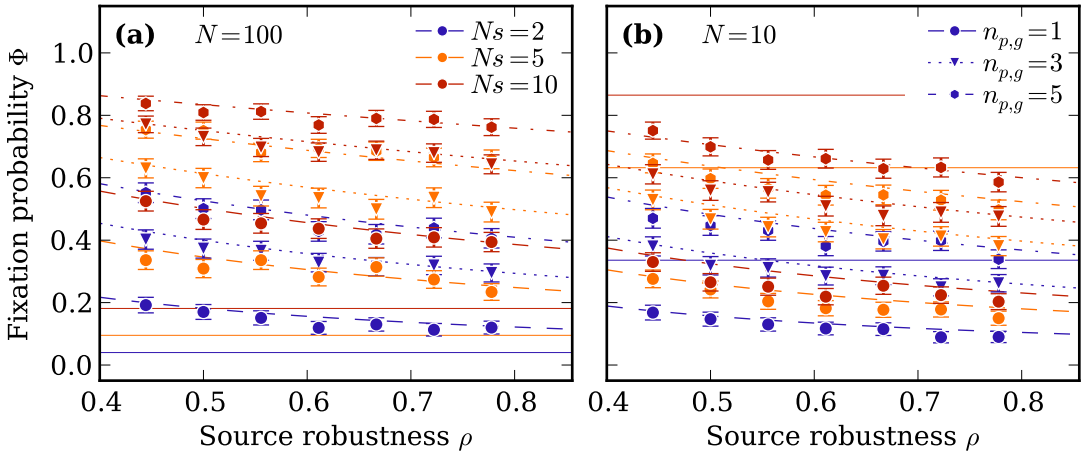

Figure 6.3: Robustness impedes adaptation. The diagram compares our predictions for the fixation probability from a portal genotype, Eq. (6.6), to simulations (described in the text). The panels correspond to population sizes **(a)** $N = 100$ (a large monomorphic population) and **(b)** $N = 10$ (a population in the large genome regime). Colour encodes the selective advantage Ns of the adaptive mutants. The solid horizontal lines mark the fixation probability $P_1(s)$ of a single adaptive mutant, Eq. (2.15). Broken lines correspond to Eq. (6.6) and markers show the outcome of 1000 simulations. Errorbars indicate 2 standard errors on the mean.

adaptive mutant is produced. Therefore we find $\Phi(s) < P_1(s)$, in particular when $n_{p,g}$ is small.

6.4 Summary and discussion

In this chapter, we studied the mutants produced by monomorphic populations ($NL\mu \ll 1$) over the duration of several neutral fixations (whose time-scale is given by $\tau_f = 1/(L\mu\rho_q)$, Eq. (4.12)). As the populations undergo long periods of genotypic stasis, the spectrum of accessible phenotypes is limited; provided the populations are large enough, we predicted correlations among the phenotypes of non-neutral mutants over a time-scale given by the stasis time, which is just the inverse of the rate of neutral fixations (Eq. (4.12)). The spectra of mutants in small populations, by contrast, is much more dominated by the fluctuations of random mutations, in agreement with our earlier results (cf. Section 4.2.3). Simulations under the random GP map confirmed these predictions: Large, well-localized populations produce mutants with a particular phenotype in bursts, rather than continuously.

The existence of bursts follows from the discreteness of genotype space. As long as the number of accessible phenotypes $\mathcal{E}_q$ is greater than $(K-1)L$, the number of single-mutation neighbours of one genotype, a localization in genotype space necessarily implies that only a subset of all accessible phenotypes can be produced. And when the genotypic stasis time τ_f is large enough that the population produces the accessible genotypic mutants repeatedly, the phenotypes carried by the mutants are correlated. We stress that bursts do not necessarily require phenotypic bias, but only that the different phenotypes are accessible from different genotypes in the same neutral space.

These dynamic correlations delay the discovery of alternative phenotypes under genetic drift, as we have seen in Chapters 4 and 5. The secondary mutants produce phenotypes that the population has already discovered and were removed by selection. The potential of these mutants to generate 'useful' variation has gone to waste.

This negative view of secondary mutants relies on the detrimental effect of their phenotypes. In the second part of this chapter, we have endowed a target phenotype with an adaptive advantage, and studied the probability of fixing that phenotype once it becomes accessible. By incorporating the production of mutants into our analysis, we predicted that the robustness of the portal genotype (from which the adaptive mutants can be produced) reduces the probability of adaptation (Eq. (6.6)). Our theory highlights that the ease of producing adaptive mutants is just as important as their individual fixation probability for the chance of taking over the population.

Finally, our results show that in a large monomorphic population, the probability of fixing an accessible, adaptive phenotype, comes close to one even if the probability of fixing *one* mutant is relatively small. A single burst of mutants is often large enough to produce one mutant which is destined to be fixed. The existence of bursts is a simple consequence of the discrete nature of neutral spaces. Thus, the existence and consequences of bursts underline the importance of neutral spaces when we reason about adaptive evolution.

Our result that accessible adaptive mutants are quite likely to be fixed, serves as a theoretical foundation for the model of adaptive walks [53, 69, 94, 95], in which the evolution of a monomorphic population is modeled as an fitness-increasing walk over genotype space.

The traditional adaptive walk model assumes that neutral mutations cannot be fixed; our calculations show that this assumption is not valid in the limit of small populations (again, measured against the genotype length). Here, the fixation of a random mutant can be faster than the discovery of a directly accessible, beneficial mutant.

These results complement recent findings by Rendel [111], who demonstrated (in RNA) that allowing for neutral steps can allow adaptive walks to reach higher fitness peaks. Clearly, we need to add a caveat: Unless populations are very large, the fixation of neutral mutants may also prevent adaptive transitions, or at least slow them down significantly.

We have calculated the fixation probability of a potentially accessible phenotype, rather than for a single mutant (which, as we have seen in Section 2.1.2, is a well-studied problem). While our result is just as sensitive to the dynamic model as the single-mutant fixation probability [96], our approach arguably gives a more complete picture: We do not rely on pre-existing variation, but incorporate the introduction of variation into our description.

At first glance, the negative impact of robustness on adaptation is reminiscent of the survival of the flattest [140] (cf. Section 2.3.1). However, our result emerges in the opposite dynamic limit of monomorphic populations. Since mutations are rare, there is no adaptive advantage of robustness: There is no dependence on the robustness of the target phenotype in Eq. (6.6). Instead, our findings can be traced back simply to the speed of neutral exploration: More robust neutral spaces are explored more quickly, but their neighbourhood is searched less thoroughly.

The calculations of this chapter focused on the mutational neighbourhood of an individual genotype, rather than an entire neutral space. This zoom-in showed how the local structure of the GP map can induce correlations into the dynamics of evolving populations. In the random GP map, this local structure is determined by frozen fluctuations, but there are no other correlations between adjacent genotypes. But as we have seen in explicitly in Chapter 3, more biologically motivated GP maps do show correlations between genotypes: At least, genotypes with the *same* phenotype are more likely to be neighbours than the global frequency of their phenotype would suggest. In the next chapter, we study how the correlations of more complicated GP maps affect the applicability of our mean-field results.

Chapter 7

The dynamic consequences of genotypic correlations

7.1 Introduction

In the last three chapters, we have developed a simple mean-field model for the dynamics of neutral exploration. We have repeatedly tested our theory on the random GP map, which – up to frozen fluctuations in the realization of the mapping – fulfills the mean-field requirement: The phenotypes of adjacent genotypes are not correlated with each other (except of course for the null-expectation based on the frequencies of these phenotypes). On the other hand, we have seen in Chapter 3 that biologically motivated models of GP maps do lead to correlations between genotypes. Specifically, genotypes mapping into the same phenotype tend to be closer to each other than a random distribution over all of genotype space would achieve, and some pairs of phenotypes are not accessible from each other at all by single mutations.

The random GP map provided a useful toy model, but as with every toy model, we need to check how well the insights we gained apply to more complex situations. To this end, we will study evolutionary dynamics under the RNA GP map in this chapter. As we have seen throughout the earlier chapters, RNA has been an important model in which many insights about GP maps have been developed [23, 42].

We begin by a largely hand-waving theoretical discussion of the impact of correlations. By reviewing where we used the mean-field assumption in our theory, we argue that departures from homogeneity will on average lead to a delay in the discovery of novel phenotypes, and to stronger bursts of mutants with the same phenotype. We will then compare the outcome of simulations on RNA neutral spaces to the prediction of our theory. The effect of genotypic correlations is as we expect: The discovery of rare phenotypes is delayed beyond the expectation of the mean-field model, and bursts are even more pronounced than in the random GP map. Nonetheless, the qualitative features of the mean-field theory are consistent with simulations in RNA. In particular, rare phenotypes remain much harder to discover than frequent ones, and hence population evolvability increases only slowly with time.

These results underline that some phenotypes – the locally frequent ones – are much more readily discovered by the population than most others. This immediately raises the question if and to what extent phenotypic bias affects which phenotypes go to fixation in an evolving population. For example, we have seen in Section 2.3.1 how at high mutation, the importance of mutational robustness can lead to the 'survival of the flattest' [10, 140], and as we saw in Chapter 3, more frequent phenotypes generally tend to have higher robustness[1] (cf. Fig. 3.12).

In Section 7.4, we use our microscopic approach to study the impact of neutral spaces and phenotypic bias on adaptation over a wide range of mutation rates. To highlight the importance of the phenotypic bias, we consider a fitness landscape with two peaks: One adaptive phenotype is more easily accessible by mutations and has higher robustness, while the other phenotype has a higher fitness. At high mutation rates, we find the expected survival of the flattest (cf. Section 2.3.1): While both phenotypes are discovered by the population, the more robust phenotype goes to fixation unless the fitness advantage of the less robust phenotype is large. When mutation rates are small, slow neutral exploration is often

[1]Note that in the random GP map, the validity of this statement is limited: As long as a phenotype is too rare to have a large neutral component, the robustness of a typical genotype with this phenotype will increase only very weakly with the phenotype's frequency. It is largely for this reason that we do not study adaptation at high mutation rates in the random GP map.

necessary to discover rare phenotypes; instead the more frequent phenotype is discovered quickly and hence often goes to fixation before the rare phenotype has any chance of playing out its fitness advantage. Thus our simulations highlight that the existence of neutral spaces puts more emphasis on accessibility than on fitness. These finding underline the importance of mutations to generate selectable variation and add support to the 'ascent of the abundant' hypothesis [24].

7.2 Theoretical remarks

Going from a mean-field approximation to a more faithful account of correlations is hard, and a general treatment (as far as this is possible) is beyond the scope of this thesis. Instead, we discuss here the impact of one particular class of correlations on evolutionary dynamics, namely local correlations between adjacent genotypes.

In general, the scale at which correlations matter is set by the dispersion of the population in genotype space. For example, in the continuity approximation, we need not worry about local correlations, simply because the population is widely spread out over the neutral space and cannot 'resolve' local structure. Consequently, it is clear that monomorphic populations will be most sensitive to local structure, simply because these populations are well localized in genotype space. So to get a rough understanding of the influence of local correlations, we can revisit our treatment of monomorphic populations in Section 4.2.3.

When the population is localized to a single genotype g, the spectrum of accessible phenotypes is determined by the mutational neighbourhood of g. The neighbourhood can be described in terms of the number $n_{p,g}$ of genotypes mapping into phenotype p. The mean-field approximation implied that we could make two simplifications: First, we argued that $n_{p,g}$ would almost always be only 0 or 1, provided the mean $n_{pq} = (K-1)L(1-\rho_q)c_{pq}$ is small[2]. Second, we assumed that the probability that p would become accessible upon a neutral fixation is constant in time, and simply given in terms of the Poisson distribution as $1 - \exp(-n_{pq}) \approx n_{pq}$.

[2]The case $n_{pq} \geq 1$ implies that p is virtually always accessible, and hence neutral exploration typically does not play a role in the discovery of p

When there are strong local correlations, neither of these assumptions needs to hold. It may well be that $n_{p,g} \geq 2$ even for a rare phenotype p; this means that p-genotypes are clustered around a particular g. If this is the case, it may also be quite likely that p remains accessible after a mutation from g to ones of its neutral neighbours: p might be accessible from a small, connected cluster of genotypes in a neutral space, rather than from roughly equally spaced portal genotypes.

We can immediately see that the consequence of such clustering is that *on average*, the discovery of p is delayed: When the populations is far from the cluster, a single neutral fixation is unlikely to make p accessible. On the other hand, once the population moves into the cluster, it is very likely that p will be discovered multiple times; the burst length should increase, and population evolvability will grow only very slowly. It is important to realize that the strength of such correlations can differ among the alternative phenotypes. So in addition to the c_{pq}, there may be further, phenotype-specific parameters that affect the discovery time and burst duration.

Alternatively, we can say that correlations between adjacent genotypes increase the timescale over which the population obtains a representative sample of the phenotypic neighbourhood of its current neutral space. A case in point is the fragmentation of neutral spaces in RNA, which we have studied in detail in Chapter 3: As we have seen there, some alternative phenotypes are not accessible from all components of the neutral space. If the population is to start out on an NC from which some particular phenotype p is not accessible, it may take the population very long indeed to discover p.

More generally, if neutral spaces are heterogeneous across multiple scales, it may be the appropriate mean-field connectivities c_{pq} change over (very long) times. We have seen an example of this effect already in the random GP map (cf. Section 4.2.5): In the monomorphic regime, the initial neighbourhood of the population is explored over a much shorter timescale than the rest of the neutral space. In RNA, we certainly expect that if the population is initially localized to a single NC κ in the NN of phenotype q, the appropriate mean-field connectivity over the time-scales we consider here (which may still be of the order of millions of generations) will be $c_{p\kappa}$ (Eq. (3.16)) rather than c_{pq} (Eq. (2.24)). Similarly, we need to

consider local frequencies $\phi_{p\kappa} = c_{p\kappa}/(1-\rho_\kappa)$ for individual NCs instead of the ϕ_{pq} defined for entire NNs (cf. Eq. (4.7)).

7.3 Results

7.3.1 Correlations in RNA neutral spaces

Let us begin by a study of the static properties of neutral spaces that were relevant in the mean-field theory. There, we saw that the GP map can be encoded in the local phenotype frequencies ϕ_{pq}, defined in Eq. (4.7). In the random GP map, we found that (for a sufficiently large neutral space), local connectivities are equal to the global frequencies F_p to a high degree of accuracy (cf. Fig. 4.2): We would expect to find $\phi_{pq} = F_p \pm \Delta_p$, where $\Delta_p \ll F_p$ measures the magnitude of fluctuations; in the random GP, the relative size of these fluctuations Δ_p/F_p should decrease with the square-root of the neutral space size.

In RNA, it is clear that the situation will be more involved: As we have seen in Chapter 3, even frequent phenotypes need not necessarily be accessible at all from a given neutral space. Thus the 'noise term' Δ_p can be as large as the 'signal' F_p. However, it is important to note that the magnitude of fluctuations has asymmetric bounds: While it is perfectly possible that the local frequency $\phi_{p\kappa}$[3] around the NC κ vanishes, there is an upper bound on $\phi_{p\kappa}$ depending on F_p: There cannot be more mutations leading towards p than leading away from p, and the latter number is bounded by F_p. To be precise, let $\mathcal{S}_\kappa$ be the size of the current neutral space κ. Then we have

$$\phi_{p\kappa} \leq \frac{\text{mutations away from } p}{\text{mutations away from neutral space}} = \frac{(K-1)LK^L F_p}{(K-1)L\mathcal{S}_\kappa} = \frac{K^L F_p}{\mathcal{S}_\kappa} \tag{7.1}$$

Of course, this bound is only meaningful (that is, less than 1) for phenotypes whose number of genotypes is less than the size of the current neutral space (note that $\mathcal{S}_\kappa$ is not necessarily equal to the total number of genotypes $K^L F_q$ mapping into q, since (particularly in RNA)

[3]From now on, we shall use the local frequencies $\phi_{p\kappa}$ around NC κ instead of the ϕ_{pq} around the NN of phenotype q

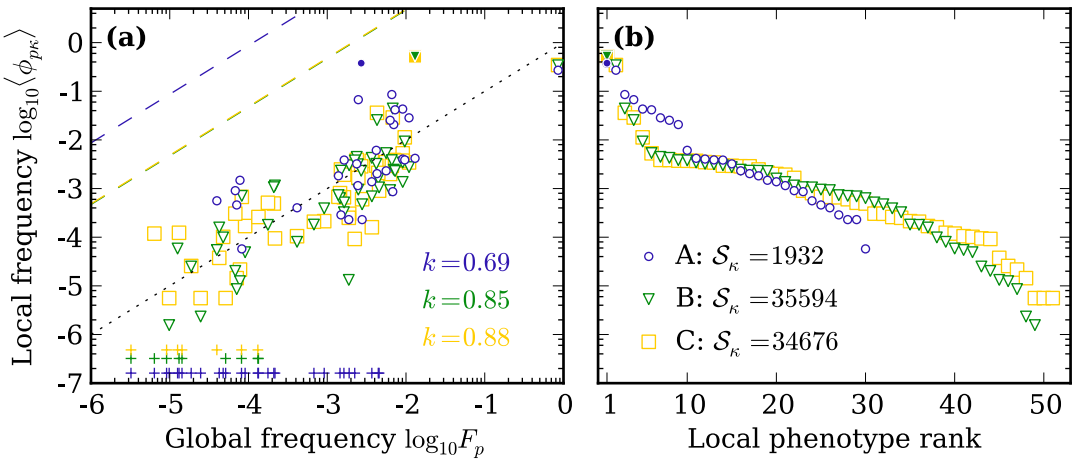

Figure 7.1: Complexity of RNA neutral space connections. (a) The diagram shows the local frequencies $\phi_{p\kappa}$ of phenotypes around three NCs (distinguished by colour) in RNA at $L = 12$. The filled markers indicate the phenotype of the respective NC, thus showing their mutational robustness. The black dotted line indicates the equality of local and global connectivity; the dashed lines show the upper bound of the local frequency, given by Eq. (7.1). For each space (colour), k is the Spearman rank correlation coefficient of global and local frequencies; the correlations are highly significant, with $p < 10^{-15}$ in all cases. **(b)** Here, the non-zero local frequencies are shown by rank. For all 3 spaces, the locally most frequent phenotype is that of the NC. $\mathcal{S}_\kappa$ is the size of each space. An example genotype and the phenotype of each space are shown in Fig. 7.2.

neutral spaces can be fragmented). Nonetheless, this result suggests that for large neutral spaces, the majority of non-neutral mutations will lead towards frequent phenotypes.

Fig. 7.1a shows $\phi_{p\kappa}$ against F_p for three RNA NCs that we shall study below. Clearly, the fluctuations in local frequencies are large, in particular for the relative small space A with only 1932 genotypes. Nonetheless, globally frequent phenotypes are also more abundant locally. The most important exception is the phenotype of the respective neutral space (indicated by the filled markers in Fig. 7.1), which has the highest local frequency. In other words, the most likely outcome of a mutation is to be neutral. Fig. 7.1b shows the (non-zero) local frequencies by rank. Note that although the global frequency of the source phenotype q is relatively small, its local frequency (that is the robustness ρ_κ) is greater than even the local frequency of the trivial structure whose global frequency is greater than 80%.

These results speak to the difference between the global structure of the genotype space and the neighbourhood of individual neutral spaces. Since the mean-field description is already cast in terms of the local frequencies ϕ_{pq} (or the connectivities $c_{pq} = \phi_{pq}/(1 - \phi_{qq})$

respectively), this kind of heterogeneity can easily be accounted for. As discussed in Section 7.2, we use $c_{p\kappa}$ for the NC κ on which the population is located initially, rather than c_{pq} for the source phenotype q.

A more interesting question is whether the genotypes mapping into a particular alternative phenotype p are distributed homogeneously around the source neutral space, like the mean-field description assumes. Fig. 7.2 suggests that this may not be the case: First, the source neutral space appears to have non-trivial internal structure; and second, the genotypes connecting to a particular phenotype are clustered. Clearly, such heterogeneity of the neutral space will increase the sensitivity to initial conditions: If the population was localized in the encircled cluster at $t = 0$, the discovery of the relatively rare phenotype p_2 would typically require several neutral fixations, even along the shortest possible neutral path. But of course, the population could spend a very long exploring the 'wrong' part (the right hand side of the diagram) of the neutral space. Again, this argument underlines that structural correlations will further add to the sensitivity of particular evolutionary trajectories on the random fixation of neutral mutants, and hence dynamic correlations may be visible over much longer timescales than in uncorrelated neutral spaces.

7.3.2 Number of mutants and bursts

We begin by studying of the dynamic implications of such correlations with the number of mutants $M_p(T)$ of phenotype p produced over long times T. As Fig. 7.3 shows, there is a relatively good agreement of the mean-field predictions and the simulations: Over long times, the neutral space is explored uniformly. Yet, we note that relative error defined by $1 - M_p(T)/(NT\nu_p)$ increases as μ decreases, even though the total number of mutants is the approximately the same in both panels of Fig. 7.3. This increase shows that the time-scale over which even a small neutral space is explored uniformly can be very long.

When we consider a single simulation over a shorter time-scale, we find that the static correlations lead to an increase in the correlations among mutants. Even over 10^6 generations, the population explores only a small part of the neutral space's mutational neighbourhood:

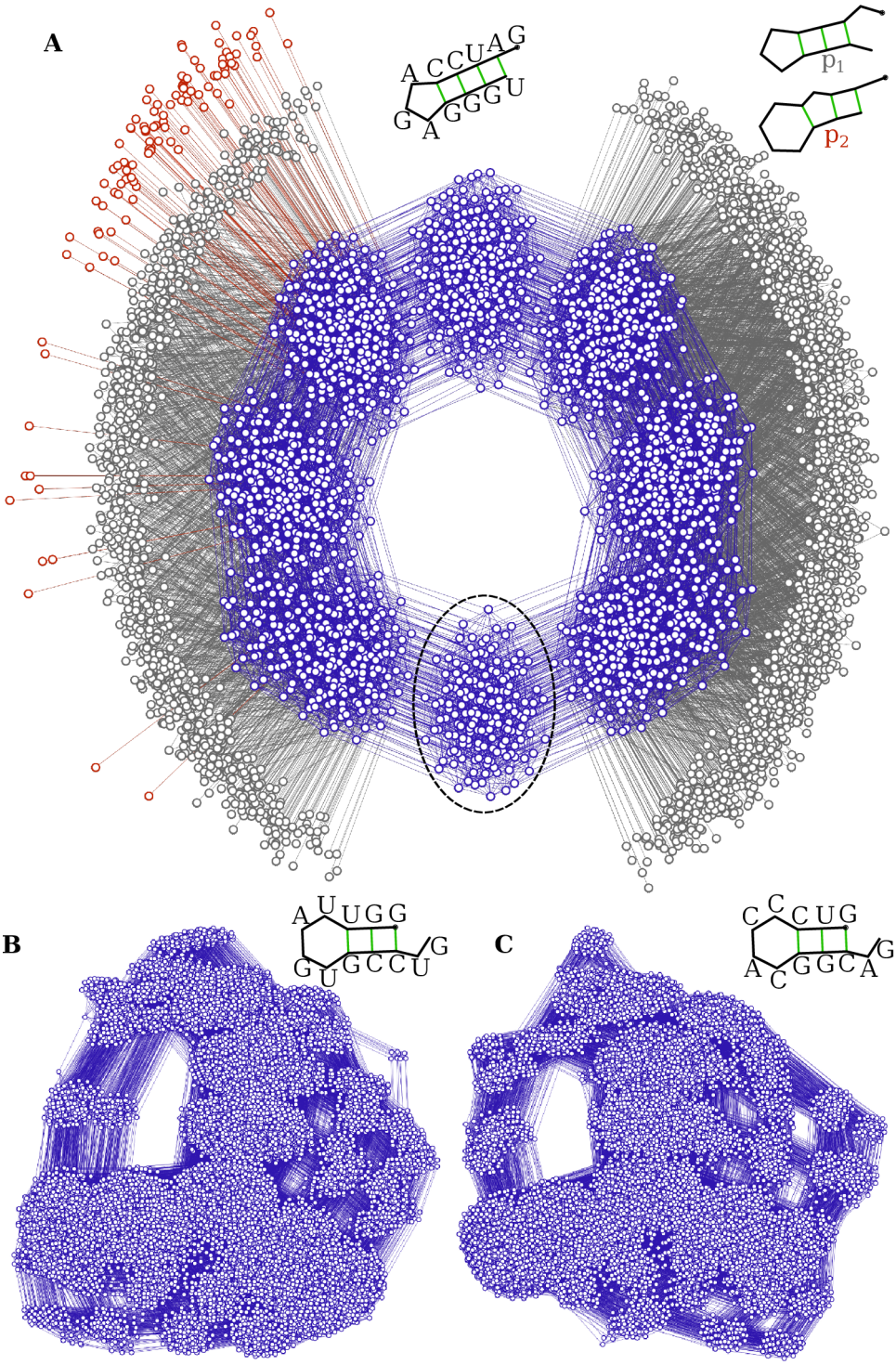

Figure 7.2: Heterogeneity of neutral spaces in RNA. The diagram shows the structure of neutral spaces of Fig. 7.1. Each circle marks a unique genotype (RNA sequence); edges connect one-mutant neighbour genotypes. Blue nodes correspond to genotypes in the neutral space; the respective phenotypes are shown by an example sequence. The diagram for space space A also shows the non-neutral neighbour genotypes mapping into phenotypes p_1 (grey) and p_2 (red), which will be discussed in Section 7.4. The layout of the neutral space was obtained by the Fruchterman-Reingold algorithm [48]: Each node is represented by an electric charge and each edge is replaced by an elastic spring, and the layout minimizes the energy of the system.

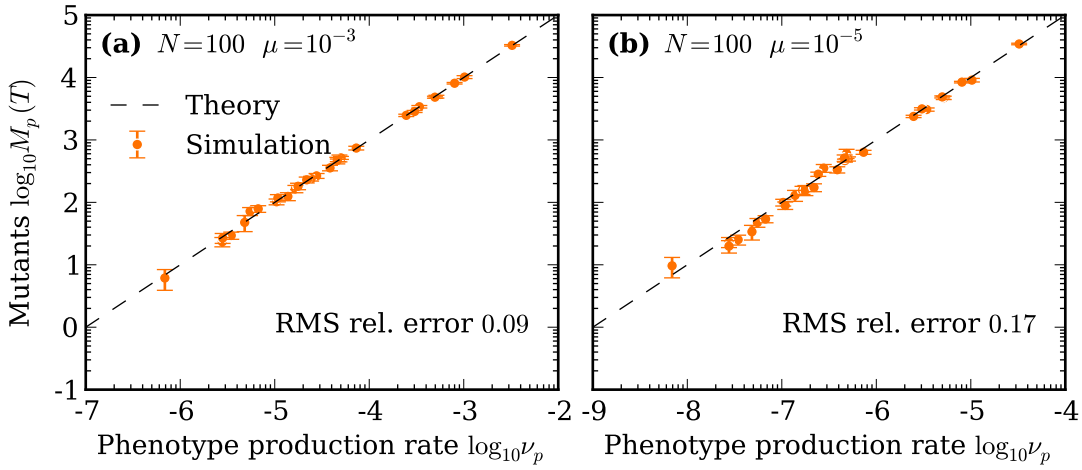

Figure 7.3: Over long times, static correlations are averaged out. The diagrams show the number of mutants produced for phenotype over T generations, averaged over 100 simulations on landscape A. Error bars indicate 2 standard errors on the mean, and the black dashed lines correspond to Eq. (4.10). In both panels, the population size is $N = 100$. In **(a)**, $\mu = 10^{-3}$ and $T = 10^5$, while in **(b)** $\mu = 10^{-5}$ and $T = 10^7$, giving the approximately the same number of mutation events in both cases.

11 of the 29 phenotypes in the mutational neighbourhood are not produced at all, while for some other phenotypes, the number of mutants overshoots the prediction by a large margin. We also note that if a phenotype was discovered by the population, it was subsequently produced at least 10 times.

The impact of genotypic correlation becomes clear if we compare these findings to the analogous results under the random GP map, which are shown in Fig. 6.2a. In comparing the two landscapes, we need to bear in mind that in the results for the random GP map, robustness $\rho_q = 0.5$ is slightly higher than in RNA NC A ($\rho_\kappa \approx 0.38$), which means that in RNA more non-neutral mutants will be produced, and hence the agreement of predictions and simulations ought to be (slightly) better in RNA. We also note that the smallest connectivity values are roughly similar in the random GP map results ($\min_p c_{pq} \approx 4 \times 10^{-4}$ for landscape B of the random GP map), except that the RNA NC A has one neighbour with smaller connectivity $c_{p\kappa} = 9 \times 10^{-5}$. Thus the mean-field structural properties of the two landscapes are relatively similar, and we may attribute differences in simulations results to the effects of genotypic correlations in RNA.

The comparison of Figs. 7.3 and 6.2a shows that structural correlations can have a strong

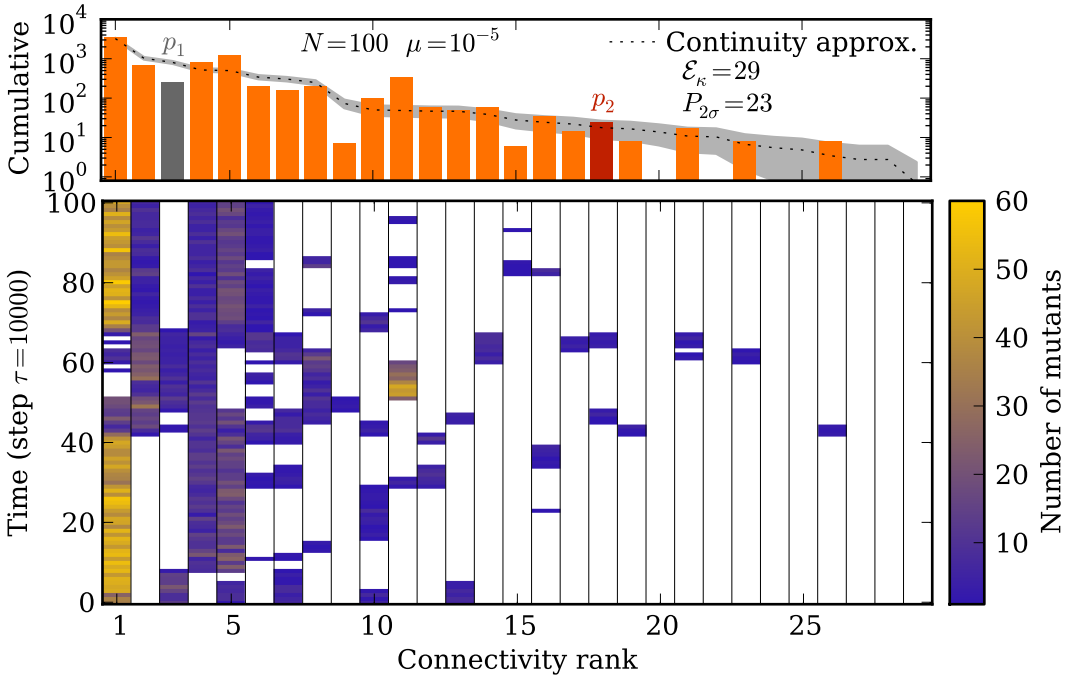

Figure 7.4: Structural correlations aggravate dynamic correlations. In analogy to Figs. 6.1 and 6.2, this diagram shows the number of mutants for each phenotype, produced over 10^7 generations in a single simulation with $N = 100$ and $\mu = 10^{-5}$. Due to the large variation in connectivity of the alternative phenotypes, the cumulative number of mutants per phenotype in the top-panel is shown on a log scale.

dynamic effect. In the random GP map, the number of mutants for 29 out of 57 accessible phenotypes were within the 2σ-region of the continuity approximation. By contrast, in RNA we find this agreement only for 6 of the 29 accessible phenotypes. It is also worth noting that the number of mutants produced with a particular phenotype over a short amount of time ($\tau = 10^4$ generations) is often much larger in RNA than in the random GP map (compare the scales of the colourbars in Figs. 6.2a and 7.3): Since the number of mutants produced from a particular genotype g over τ generations is simply $M_{p,g}(\tau) = NL\mu(1-\rho_g)n_{p,g}$, we can take the larger scale in RNA as evidence of the fact that the number of distinct p-mutants $n_{p,g}$ around g often exceeds than 1 in RNA, while $n_{p,g} > 1$ does not typically occur in the random GP map.

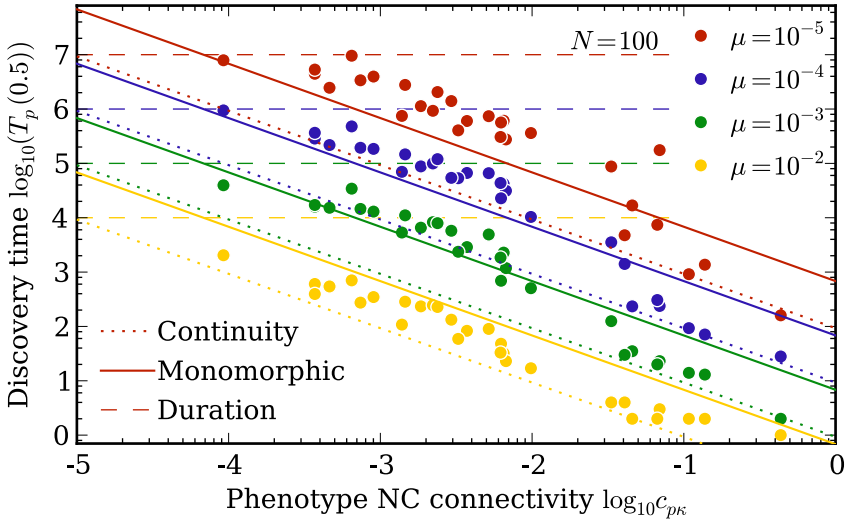

Figure 7.5: Correlations further delay the discovery of novel phenotypes. The diagram shows the median discovery time of alternative phenotypes p as a function of the connectivity $c_{p\kappa}$ around the NC κ (landscape A in Fig. 7.2). Mutation rates are encoded by colour. Dotted lines show the continuity approximation (Eq. (4.11)) and solid lines indicate the mean-field monomorphic limit (Eq. (4.21)). The dashed horizontal lines mark the simulation time $T = 100/\mu$. The data is for population size $N = 100$ on landscape A.

7.3.3 First discovery times

As we have argued in Section 7.2, heterogeneity in the distribution of portal genotypes will on average delay the discovery of alternative phenotypes. Indeed, this is what we find in Fig. 7.5: As the mutation rates becomes smaller, the median discovery times increase even beyond the predictions of the mean-field approximation in the monomorphic limit (solid lines). So correlations do indeed delay the discovery time. It is important to note that the magnitude of the delay, which can serve as a measure of heterogeneity, varies between phenotypes: Rather than following a simple inverse relation with the mean-field phenotype production rate, there is some scatter in the discovery times. Nonetheless, rare phenotypes remain harder to discover than frequent ones: The rank correlations are strong ($k < -0.95$ throughout) and highly significant ($p < 10^{-15}$).

As we have seen in Chapter 3, the structure of neutral spaces in RNA shows heterogeneity across different scales. In particular, not all phenotypes are accessible from a given NC, and the sets of accessible phenotypes vary among the NCs in a given NN. We study the dynamic importance of this multi-scale heterogeneity in Fig. 7.6, which shows the median

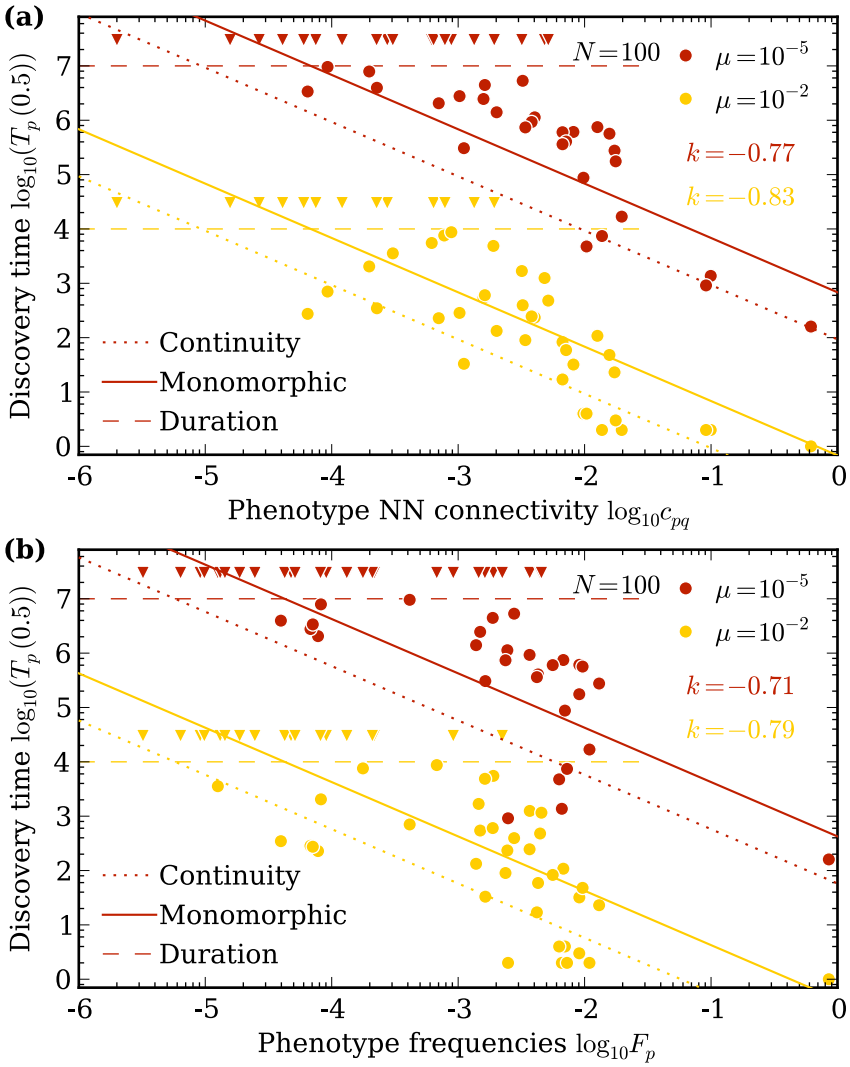

Figure 7.6: Comparison of predictors for discovery times. The diagram shows median discovery times T_p of alternative phenotypes against measures of phenotypic bias at different scales. **(a)** uses connectivities c_{pq} (Eq. (2.24)) which include all NCs of the source phenotype q. **(b)** shows T_p against the global phenotype frequencies F_p. In both panels, triangle markers indicate phenotypes whose median discovery time is beyond the simulation time which is marked by the dashed horizontal line. The respective connectivities can be used to predict discovery times according to the continuity approximation (dotted lines, Eq. (4.11)) and the mean-field theory for large monomorphic populations (solid lines, Eq. (4.21)). k gives the Spearman rank-correlation of discovery time and the respective connectivity parameters.

discovery times T_p against the NN connectivities c_{pq} (Eq. (2.24)) for the source phenotype q (Fig. 7.6a) and against the global phenotype frequencies F_p (Fig. 7.6b). In both cases, the most important weakness of these more large-scale connectivities is that they fail to predict which phenotypes are *not* discovered: There are quite a few phenotypes which are not discovered during the simulations (marked by the triangles), even though their connectivity c_{pq} or F_p is orders of magnitude larger than for some other phenotypes which are discovered by the population.

We can also consider the Spearman rank-correlation k of the respective connectivity predictors and the discovery times, where we can observe two effects: First, 'zooming out' (going from $c_{p\kappa}$ to c_{pq} and then F_p) decreases the absolute correlation strength $|k|$, even though the correlations remain relatively good ($|k| > 0.7$ throughout). Second, reducing mutation rate leads to a decrease in $|k|$. Both of these scalings suggests that neutral exploration is typically constrained to a single NC. Consequently, for quantitatively accurate predictions, it is clearly necessary to have a good understanding of the local structure of neutral spaces. We note that at low mutation rates, the correlation of T_p and F_p ($k = -0.71$, Fig. 7.6b) is not much greater in magnitude than the correlation of F_p and $\phi_{p\kappa}$ ($k = 0.69$, Fig. 7.1a), which is the predictor of T_p in the mean-field calculations.

7.3.4 Evolvability

Finally, we turn to the population evolvability. In addition to the question regarding the applicability of the mean-field results, we can use simulations also to study how evolutionary dynamics differ between different NCs of the same phenotype. To this end, we study in this section the larger NCs B and C of Fig. 7.1 which both belong to the same NN.

Fig. 7.7a compares the population evolvability measured from simulations to our predictions from Section 5.2.1. As the diagram shows results for $N = 1000, \mu = 10^{-3}$, it is not very surprising to find that the continuity approximation (Eq. (5.2) with $\gamma = 1$) is only valid for relatively short times. At long times, our mean-field model agrees quite well with the simulations, provided we use an appropriate value for γ, which we obtain from a fit according

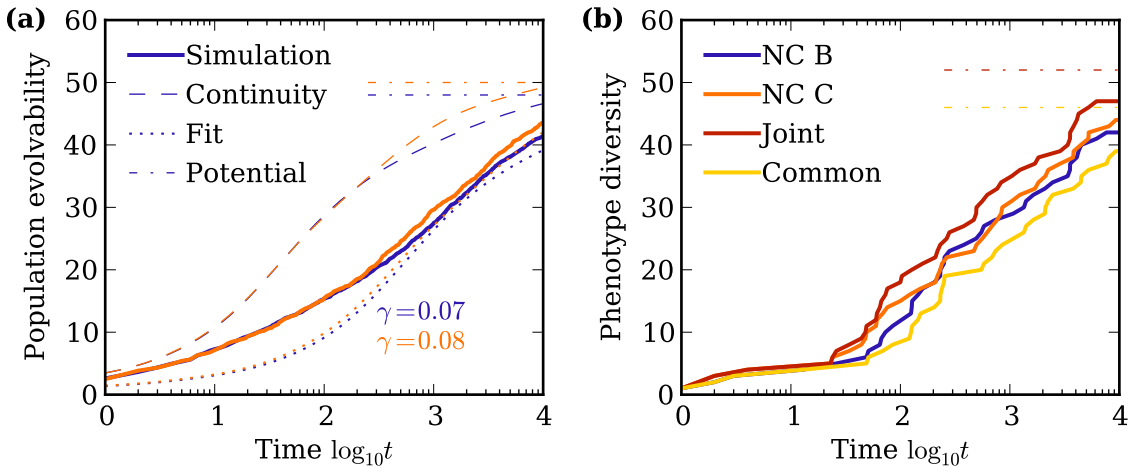

Figure 7.7: Evolvability on RNA neutral spaces. (a) Population evolvability increases only slowly with time. The solid lines show simulation results for $N = 1000, \mu = 10^{-3}$ for NCs B (blue) and C (orange). The broken lines indicate the mean-field approximation (Eq. (5.2)) for a continuity description ($\gamma = 1$, dashed) and using a fitting factor (γ calculated from Eq. (4.23) with $\alpha = 1 - e^{-1}$, dotted) to account for drift. **(b)** Populations on different NCs discover different phenotypes. The diagram shows the phenotypic diversity (Eq. (7.2)) on NCs B (blue) and C (orange), as well as the joint diversity (Eq. (7.3), red) and common diversity (Eq. (7.4), yellow).

to Eq. (4.23). This agreement suggests that heterogeneity in the connections to alternative phenotypes balance each other. It is also interesting to note that $E(t)$ increases roughly logarithmically in time, just as the sequential approximation (Eq. (5.18)) suggests. This finding underlines the importance of phenotypic bias even when the mean-field assumption is violated.

To compare *which* phenotypes have been discovered by populations evolving on different NCs, we count the number of phenotypes discovered by at least 50% of the simulations. We call this the number of phenotype diversity $D(t)$. Using the Heaviside function

$$\Theta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

we can define $D(t)$ mathematically as

$$D(t) = \sum_{p \neq q} \Theta\Big(t - T_p\Big) \tag{7.2}$$

where T_p is the median discovery time of phenotype p (the argument $\alpha = 0.5$ is omitted for clarity), and we do not include the source phenotype q in the calculation.

We can now compare the phenotypic diversity of population evolving on different NCs i in the same NN in terms of the median discovery times T_p^i. Let us define the joint and common diversity as

$$D_j(t) = \sum_{p \neq q} \Theta\Big(t - \min_i(T_p^i)\Big) \tag{7.3}$$

$$D_c(t) = \sum_{p \neq q} \Theta\Big(t - \max_i(T_p^i)\Big) \tag{7.4}$$

So $D_j(t)$ counts the number of phenotypes discovered from at least one NC, while $D_c(t)$ counts only the phenotypes that were found from all NCs.

As Fig. 7.7b shows, populations on two different NCs in the same NN (that is, the populations have the same phenotype) discover different sets of phenotypes. At short times, when the most frequent phenotypes are found, D_j and D_c are equal, but around generation $t = 20$, the curves begin to diverge: The phenotypes discovered become sensitive to the NC on which the population evolves. Even though the number of phenotypes discovered (measured by both $E(t)$ and $D(t)$) is similar across the NCs, there is a difference of about $D_j(t) - D_c(t) \approx 10$ that persists from $t = 100$ to $t = 10,000$. This finding supports our claim made in Chapter 3: The spectrum of phenotypes discovered by evolving populations is sensitive not only to the phenotype of the population, but also the location of the population in genotype space.

7.4 Fixation in a two-peaked fitness landscape

With the exception of Section 6.3, we have so far only studied how neutral spaces impact the introduction of variation. And in our study of fixation, we restricted ourselves to the case of a single advantageous phenotype that was within reach of the population. In this section, we will consider a more complex scenario in which two alternative phenotypes are

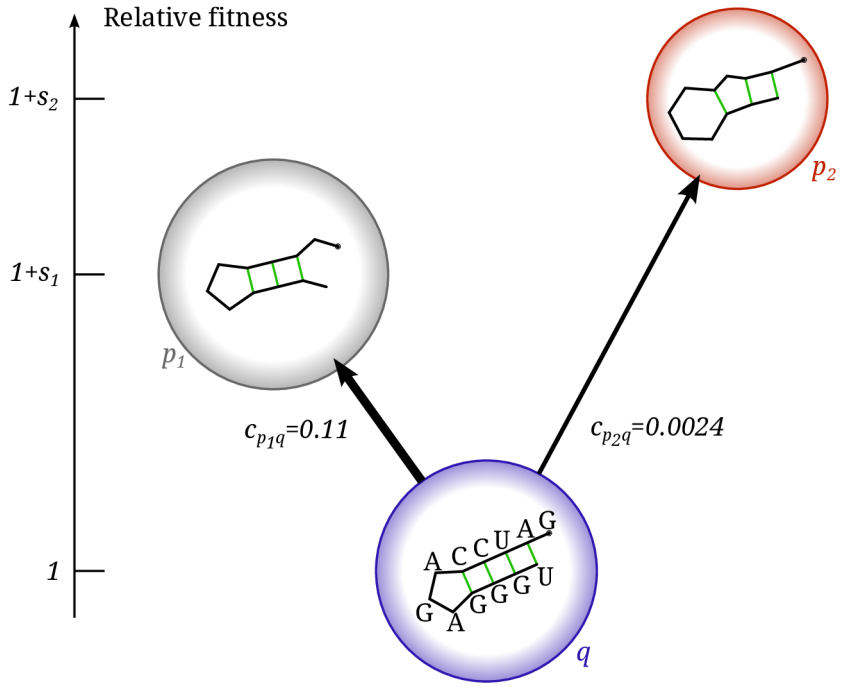

Figure 7.8: Illustration of the double-peaked fitness landscape. The diagram illustrates the doubled-peaked fitness landscape in the spirit of the continuity approximation, where we ignore the internal structure of connected neutral spaces (cf. Fig. 4.1). The source phenotype q has fitness 1 and the population starts out on NC κ (space A of Fig. 7.2). The two alternative phenotypes p_1 and p_2 differ in their connectivities $c_{p_i\kappa}$ by a factor of nearly 50. The fitness values s_1 and s_2 can be chosen independently of the connectivities. Typically, we consider the case where the less accessible phenotype p_2 has higher fitness.

advantageous; by choosing phenotypes that are mutually inaccessible, we obtain a fitness landscape with two distinct peaks, and we study to which of these peaks the population evolves.

We have stressed throughout the thesis that a phenotype p cannot be fixed before it is discovered, independent of the fitness advantage s it confers. Instead, the discovery of p depends on its connectivity $c_{p\kappa}$, as we have seen in Chapter 4, and possibly on correlations (cf. Section 7.3.3). In principle, $c_{p\kappa}$ and s are independent parameters: $c_{p\kappa}$ is determined by the GP map, while s depends on the phenotype-fitness relation. Here, we study a double-peaked landscape (see Fig. 7.8), in which the source neutral space is NC A and there are two alternative phenotypes p_1 and p_2 which differ both in their $c_{p_i\kappa}$ (cf. Fig. 7.2 and Table 7.1) and their selection coefficients s_i. Using different selection coefficients, we study the relative importance of accessibility and fitness.

Yampolski and Stoltzfus [146] studied a scenario that superficially looks very similar to ours: They considered a genetic system with $L = 2$ loci and $K = 2$ alleles (which we write as 0 and 1, where 1 indicates a mutation away from the wild-type allele 0) at each locus. Initially, the population carries the wild-type genotype 00 with fitness 1. The single mutants 10 and 01 have fitness $1 + s_1$, $1 + s_2$ respectively, and the double mutant 11 is strongly deleterious compared to the single mutants. The rates of mutation per reproduction are μ_1 from 00 to 10 and μ_2 from 00 to 01. So there are two peaks in this fitness landscape, namely the single mutants 10 and 01, but the rate at which mutants on these peaks are produced is not necessarily the same.

Yampolski and Stoltzfus focused on the parameter regime where the higher fitness peak is less accessible ($\mu_1 > \mu_2$ but $s_1 < s_2$). At low mutation rates ($N\mu_1 \ll 1$), the fate of each mutant is decided without interference by other mutants, and the dynamics can be analyzed in the picture of evolution we have used throughout the thesis: Mutations create novel variants and selection decides whether this novelty becomes fixed in the population. Yampolski and Stoltzfus argued that in competition of two 'origin-fixation' processes, the ratio r of the outcomes is given by

$$\begin{aligned} r &\equiv \frac{\Pr(01 \text{ fixed})}{\Pr(10 \text{ fixed})} \\ &= \frac{\text{rate of origin-fixation 01}}{\text{rate of origin-fixation 10}} \\ &= \frac{\mu_2 P_1(s_2)}{\mu_1 P_1(s_1)} \\ &\approx \frac{\mu_2 s_2}{\mu_1 s_1} \end{aligned} \tag{7.5}$$

where $P_1(s) \approx 2s$ (cf. Eq. (2.15)) is the probability that a single mutant with selective advantage s goes to fixation, and the approximation is valid when $Ns \gg 1$.

The study by Yampolski and Stoltzfus is framed in the traditional allele-centered language of population genetics and ignores neutral spaces. Comparing their prediction to our microscopic approach allows us to see how neutral spaces affect adaptation by impacting both the origin and fixation of beneficial mutants. In our notation, we can replace the mu-

tation rates μ_i by the PPRs ν_{p_i}, which by definition (Eq. (4.1)) are just the probability per reproduction event that p_i is produced.

Even without a complete analysis, we can see that both the static correlations between genotypes and the dynamic fluctuations of reproduction may affect the origin-fixation argument. We will consider only the limit of monomorphic populations ($NL\mu \ll 1$) so that we may ignore the possibility of interference between mutants. Let $n_{p_1,g}$ and $n_{p_2,g}$ be the number of p_1- and p_2-genotypes accessible from current genotype g. As long as the population stays at g, the PPRs are simply $\nu_{p_i,g} = \mu n_{p_i,g}/(K-1)$ (cf. Section 6.2.1), and Eq. (7.5) remains valid if we replace μ_i by $\nu_{p_i,g}$. It is important to realize that the ratio of the $n_{p_i,g}$ is limited by the size of the mutational neighbourhood $(K-1)L$, so that the ratio $n_{p_2,g}/n_{p_1,g}$ can only be large when L is large. For our choice of phenotypes p_1 and p_2, the source neutral space contains 63 genotypes from which both p_1 and p_2 can be accessed. While these genotypes all have only a single p_2-neighbour, 47 have two or three neighbours mapping into p_1.

A more important consequence of strong phenotypic bias is that p_1 may occur in the neighbourhood of many more genotypes than p_2 does. Indeed, this is what we find for our choice of target phenotypes (Fig. 7.2 and Table 7.1): p_1 is accessible from about 77% of all genotypes in the source neutral space, while only about 4% of genotypes allow mutations to p_2. As a consequence, p_1 will often be produced by the population much earlier than p_2, and hence p_1 can go to fixation before p_2 is discovered. If the population is located at a genotype g such that p_1 is accessible while p_2 is not (ie., $n_{p_1,g} > 0$ but $n_{p_2,g} = 0$), then p_1 may be produced repeatedly before a single p_2-mutant arises: There may be a burst of p_1-mutants, as we discussed in the previous chapter. Consequently, the probability that p_1 is fixed from g is not the single-mutant fixation probability $P_1(s_1)$, but rather the conditional fixation probability $\Phi(s_1)$ (Eq. (6.6)); in large populations ($N \gg L$), $\Phi(s_1)$ will be much larger than $P_1(s_1)$ and hence it is quite likely that p_1 is fixed. Of course, the probability of fixing p_1 before discovering p_2 will further increase if the population moves through many genotypes like g before it encounters a portal genotype of p_2 for the first time.

In order to understand how much the localization on genotypes that do not allow the production of both phenotypes has an impact of evolutionary outcomes, we can use the

p_i	Structure	ρ_{p_i}	c_{p_iq}	g_{p_iq}
1	..(((...))).	0.41	0.11	0.7692
2	.((.(....)))	0.20	0.0024	0.0440

Table 7.1: Properties of the target phenotypes. The table lists properties for the two target phenotypes p_1 and p_2 (shown in Fig. 7.2 in grey and red, respectively). In addition to their secondary structures, their phenotypic mutational robustness ρ_{p_i}, their local frequency ϕ_{p_iq} and the fraction of portal genotypes g_{p_iq} in the source neutral space are given.

continuity approximation: By ignoring the discreteness of the source neutral space, we can use the mean-field PPRs to obtain a 'renormalized' expression for the ratio r of fixation probabilities, simply replacing μ_i by $\nu_{p_i\kappa} \propto c_{p_i\kappa}$ in Eq. (7.5):

$$r_{cont.} = \frac{c_{p_2\kappa} s_2}{c_{p_1\kappa} s_1} \tag{7.6}$$

Below, we shall measure not r but the probability P_2 that the population goes to fixation on the peak of lower accessibility (01), which is simply given by

$$P_2 = \frac{r_{cont.}}{1 + r_{cont.}} \tag{7.7}$$

Fig. 7.9 shows the outcomes of simulations with $N = 1000$ for different values of s_1 and s_2 over a wide range of mutation rates, corresponding to varying strengths of genetic drift; we considered each phenotype to be fixed when its frequency rose to 50% in the population. For an easier interpretation of the results, we have split the probability P_2 of ultimately fixing the less accessible, but more adaptive phenotype p_2 into the probability Δ_2 that it is discovered at all, and the probability Z_2 that it is fixed after discovery. Note that Z_2 is not exactly the same as the conditional fixation probability Φ, because Z_2 can only be calculated when the first p_2-mutant *has been* produced, while Φ assumes only that p_2-mutants *could be* produced.

In Fig. 7.9a we find that Δ_2 is independent of s_2; of course, this is what we expect from a blind evolutionary search. The decrease of Δ_2 with s_1 is equally intuitive: When p_1 is more fit, it is more likely to be fixed before the much rarer p_2 is found. And it is not surprising that Δ_2 also decreases with μ: The discovery time scales with $1/\mu$ (if not stronger,

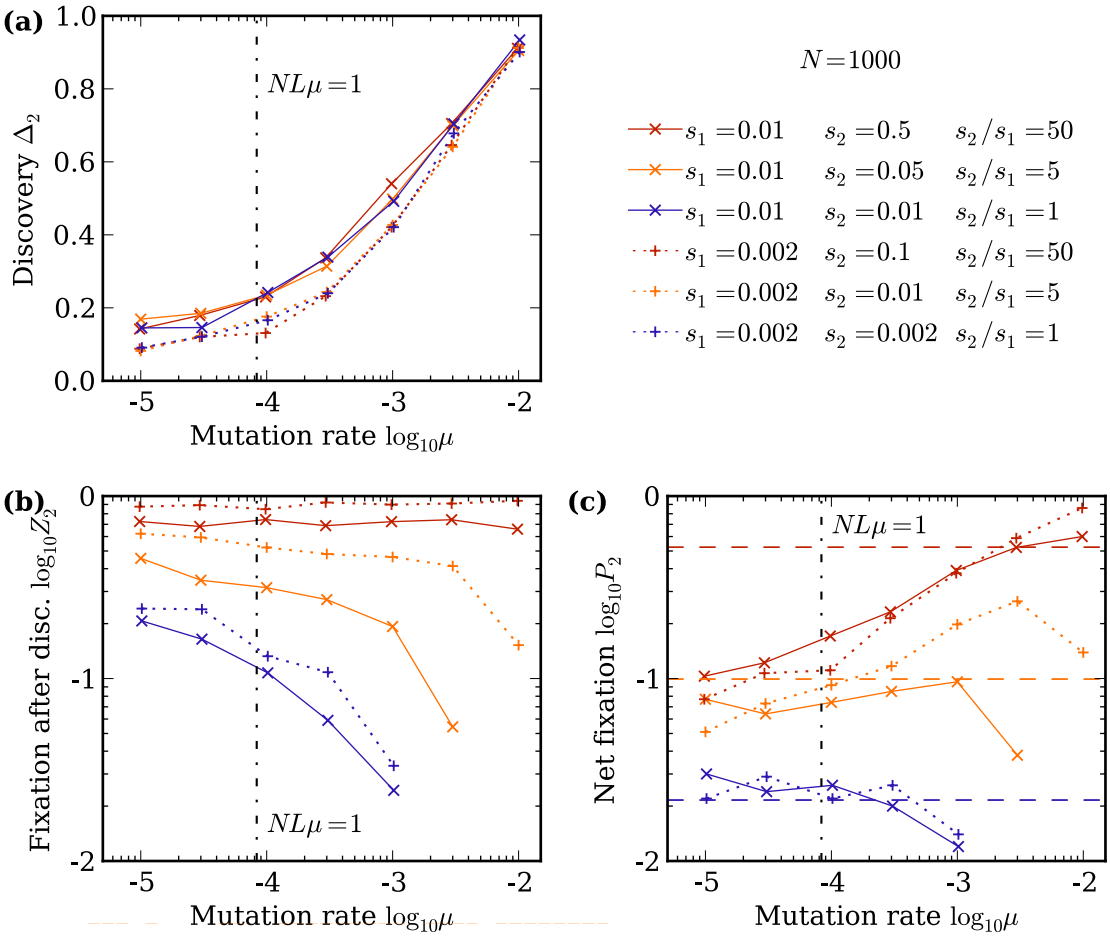

Figure 7.9: Phenotypic bias and neutral space structure can steer populations towards sub-optimal phenotypes. **(a)** shows the probability Δ that p_2 is discovered before p_1 is fixed. **(b)** displays the probability Z_2 that p_2 is fixed after discovery. **(c)** shows the product $P_2 = \Delta_2 Z_2$, which is the total probability that the population adapted to p_2; the dashed lines correspond to Eq. (7.7). For all panels, the colour encodes the selective bias s_2/s_1, while the markers indicate the adaptive advantage s_1 of the phenotype p_1 favoured by accessibility. Either phenotype was considered fixed if it had a frequency of 50% in the population. For each mutation rate, the results were obtained from 1000 simulations with a population of size $N = 1000$. The dash-dotted vertical line at $NL\mu = 1$ roughly separates the monomorphic and polymorphic regimes.

cf. Fig. 7.5), and hence the time between the discovery of p_1 and p_2 also scales like $1/\mu$. Thus as μ becomes small, the time to fix p_1 before discovering p_2 increases, thus reducing the probability that p_2 is found.

Fig. 7.9b shows the probability Z_2 that p_2 is fixed after its first discovery. At low mutation rates, we find that Z_2 increases with s_2, which again is what we should expect: Fitter phenotypes are more likely to be fixed. On the other hand Z_2 decreases as we make s_1 large, albeit an increase in s_1 also increases s_2 (since we keep the ratio s_2/s_1 fixed). We can understand this result in terms of the competition of p_2- and p_1-individuals, as p_1 can also often be produced while p_2 is accessible, quite likely even at a higher rate. So the decrease in fixation probability of the p_1-individuals increases the chance of producing the p_2-mutant that is destined to be fixed. It is interesting to note that Z_2 vanishes at high mutation rates, unless s_2 is very large. This is an instance of the survival of the flattest [140] (see also Section 2.3.1): Because of its lower robustness, p_2 cannot be fixed, even though it is discovered with high probability.

Finally, Fig. 7.9c displays the net fixation probability P_2 of p_2, given by the product $P_2 = \Delta_2 Z_2$. Depending on the ratio s_2/s_1, we observe qualitatively different behaviour. If the fitness advantage of p_2 is large ($s_2/s_1 \approx c_1/c_2$, shown in red), fixation is limited by discovery, and hence the net fixation probability decreases as genetic drift becomes stronger (smaller μ). When the selective benefit of p_2 is not too large (orange lines), the importance of robustness leads to the emergence of an intermediate mutation rate that maximizes the probability of fixing the fitter, but rarer phenotype. Finally, when p_2 has no fitness advantage over p_1 (blue lines), it is generally unlikely that the population fixes on p_2, and at high mutation rates, we again observe that selection for robustness prevents the fixation of p_2 completely. It is worth noting that the continuity approximation treatment of the origin-fixation process (Eq. (7.6)) only gives a reasonable prediction when $s_1 = s_2$ (and low μ, for which the result has been derived): The internal, discrete structure of neutral spaces has non-trivial consequences for adaptive evolution which should not be ignored.

In summary, our study of fixation in a two-peaked fitness landscape highlights the importance to account for neutral spaces when thinking about evolutionary dynamics: At

high mutation rates, mutational robustness (a direct consequence of the existence of neutral spaces) becomes an important factor to determine evolutionary outcomes; and when mutation rates are low and genetic drift is strong, the slow speed of neutral exploration can prevent the discovery of rare, highly beneficial phenotypes and instead facilitate the fixation of easily accessible, yet sub-optimal phenotypes. In either case, we note that the effect of neutral spaces is to favour frequent (hence robust and more accessible) phenotypes over fit ones.

7.5 Summary and discussion

In this chapter, we studied simulations of evolutionary dynamics using RNA as an example of a GP map with strong genotypic correlations. Such correlations violate the mean-field assumption that we used in the previous chapters to construct our theoretical description of evolution on neutral spaces.

Chapter 3 showed that neutral spaces in RNA have non-trivial macroscopic structure: Neutral networks are not fully connected, but contain a relatively small number of large components, which may be accompanied by many much smaller components. Here, we zoomed in on individual components; visual inspection (cf. Fig. 7.2) suggests that even the individual components may be internally structured. More importantly for the introduction of variation, we found that the portal genotypes to a particular phenotype may be clustered within a neutral space. We argued that such clustering has two effects: First, on average it delays the discovery of the respective phenotype. Second, once discovered, the phenotype will be produced repeatedly, since a neutral fixation is likely to stay within the cluster, thus enhancing the short-term correlations in the phenotypes of non-neutral mutants.

Our simulations confirmed both of these expectations: Correlations among genotypes and their neighbour phenotypes generally slow down the introduction of novel phenotypes, and aid the fixation of beneficial phenotypes once discovered. This delay in the discovery of alternative phenotypes also affects the population evolvability. Our simulations underline the claim of Chapter 3 that the spectrum of phenotypes produced by an evolving population

is not only sensitive to the current phenotype of the population, but also to the populations location in genotype space: Populations on different NCs produce distinct phenotypes.

While quantitative predictions of the mean-field theory are not very accurate in RNA, the qualitative features of the mean-field theory are preserved: Rare phenotypes are typically found later than frequent ones; population evolvability increases only slowly with time in monomorphic populations; genetic drift leads to correlations among the phenotypes of non-neutral mutants over times given by (at least) the inverse of the neutral fixation rate. All these results underline the importance of neutral spaces for evolutionary dynamics: While we may not (yet) be able to predict the outcome of evolutionary processes accurately, knowledge of neutral spaces is clearly indispensable if we want to make sense of evolution.

Finally, we studied the impact on phenotypic bias on adaptive evolution, combining the necessity to discover an advantageous phenotype with the requirement of fixing it in the population. In a fitness landscape with two peaks, we showed that the mutational accessibility of a peak is often more important than its fitness in determining which peak the population evolves to. By comparing our setting to the population genetics approach of Yampolski and Stoltzfus [146], in which neutral spaces were ignored, we were able to see how phenotypic bias and the discrete structure of neutral spaces affects the course of evolution: At low mutation rates, the discovery of a frequent, but only mildly beneficial phenotype often leads to its fixation (cf. Section 6.3); rarer but potentially much more adaptive phenotypes are not discovered and hence cannot be fixed. Because of phenotypic bias, the variation produced by an evolving population may not be sufficient to adapt to rare phenotypes, independent of their fitness. At high mutation rates, populations produce more variation and often also discover rare phenotypes; but in this regime, the limited robustness of rare phenotypes impedes their fixation, and the population still adapts to the less fit, but more robust phenotypes. Overall, these results support and extend the 'ascent of the abundant' hypothesis [24]: The reason we observe a particular phenotype in nature need not be its global optimality, but rather the fact that it simultaneously optimizes mutational accessibility and reproductive success.

Chapter 8

Conclusion

8.1 Summary

The central theme of this thesis has been the question how neutral spaces impact evolutionary dynamics, both by enabling and constraining the introduction of phenotypic variation and by affecting adaptation to beneficial phenotypes.

Neutral spaces, or neutral networks [115], arise when the genotype-phenotype (GP) map [4], which determines how mutations in heritable information affect the physical properties of organisms, is many-to-one so that many genotypes give rise to the same phenotype. When these genotypes are connected by small mutational steps, they form (connected) neutral spaces which allow populations to sample a large variety of genotypes without suffering from reductions in fitness [126].

Chapter 3 laid the foundations for our dynamic models by investigating the static structure of neutral spaces, using the well-studied GP map of RNA secondary structure folding [23, 42, 64] as our model system. RNA illustrated the main characteristics of neutral spaces that motivated our further study. First, large connected neutral spaces exist; this result itself is non-trivial and would not typically occur if GP maps were associating genotypes to phenotypes randomly. Second, RNA showed strong phenotypic bias, which has also been observed in a variety of other biological systems [13, 14, 32, 40, 67, 82, 104, 106]; the sizes of neutral spaces corresponding to different phenotypes vary in their size (that is, the number

of genotypes in the space) by many orders of magnitude; our results suggests that this bias becomes stronger exponentially as we go from the short sequences whose neutral spaces we could study exhaustively (sequence lengths L in the range from 12 to 20) to biologically more interesting lengths (e.g. $L = 76$ for tRNA [98]).

Neutral spaces in RNA show a high degree of robustness: Genotypes mapping into the same phenotype are often close to each other. Nonetheless, we could show explicitly that the exploration of a neutral space can bring the population close to a variety of phenotypes that is larger than the number of phenotypes accessible from a single genotype, confirming earlier sampling results by Wagner [127]. We also showed that phenotypic bias may limit the practical usefulness of neutral exploration over short times [24], but the static considerations alone were not sufficient to decide whether robustness ultimately hinders or promotes the capacity of a population to introduce novel phenotypes, which is also called evolvability [4, 6, 28, 99, 127].

In an attempt to resolve this question, Chapter 4 introduced a model of neutral exploration dynamics which included an explicit GP map into the classic Wright-Fisher model [41, 143]. The conceptual advance of this approach was to distinguish between the introduction of genotypic variation through mutations and the outcome of mutations at the level phenotypes. Two types mean-field approximations allowed us to obtain analytic insight. First, the continuity approximation completely ignored neutral exploration and lead to intuitive results for the importance of phenotypic bias: Rare phenotypes are much harder to discover than frequent ones. Second, in the simplest case of neutral exploration dynamics (the random walk over a neutral space performed by a monomorphic population), we found that a mean-field treatment of the GP map leads to the same signature of phenotypic bias; the calculation showed that neutral exploration is slow and delays the discovery of rare phenotypes even further than phenotypic bias alone.

In addition, our approximations lead to three interesting results: First, we found that the ratio of genome length to population size is an important measure of the importance of fluctuations under genetic drift. Second, we could argue that intermediate strengths of genetic drift can be handled in terms of a single fitting parameter. Third, we were able

to give an intuitive explanation for the non-monotonic dependence of the discovery time of an alternative phenotype on the mutational robustness of the population, which was first derived by Draghi et al. [33] through quite involved calculations: Greater robustness allows for faster neutral exploration, but at the same time reduces the probability of non-neutral mutations that produce novel phenotypes.

In Chapter 5, we proposed a dynamic measure of evolvability, namely the population evolvability that counts the number of phenotypes produced by the population over time. After relating population evolvability to the static evolvability measures of Cowperthwaite et al. [24] and Wagner [127], which emerged at short and long times respectively, we studied the impact of phenotypic bias on evolvability. In the sequential approximation, which assumes that phenotypes are discovered in the order of their frequency, we showed that exponential phenotypic bias leads to strongly diminishing returns of neutral exploration: The number of phenotypes discovered increases only logarithmically with search time. This result underlines that the apparent tension of robustness and evolvability may not be decidable in principle, but that their relationship may depend significantly on the structure and inter-connections of neutral spaces and on the time-scales under consideration.

The diminishing returns of neutral exploration are caused by the production of 'secondary' mutants, which carry phenotypes that were already discovered. In Chapter 6, we studied the implications of the mutants for the fixation of beneficial phenotypes. First, we showed that in the limit of large monomorphic populations, the discreteness of neutral spaces and the slow speed of neutral exploration lead to bursts of mutants with the same phenotype produced in relatively quick succession. Second, we showed that these bursts can significantly increase the probability of fixation of a beneficial phenotype once it becomes accessible through mutations. We also showed that neutral spaces (through the mutational robustness they impart on evolving populations) reduce the probability of adaptive transitions, because populations can move away from beneficial phenotypes into other regions of the neutral space.

We repeatedly used a random GP map to simulate evolutionary dynamics and test our predictions. This mapping can lead to large neutral spaces of high mutational robustness

and allowed us to explicitly enforce strong phenotypic bias; further, there are no correlations among the phenotypes of adjacent genotypes except for frozen fluctuations in the landscape, and thus the random GP map fulfilled our mean-field assumptions which can be viewed as the requirement of homogeneous neutral spaces. Our simulations under the random GP map indeed confirmed our theoretical predictions: First, phenotypic bias reduces the variety of phenotypes a population can produce quickly. Second, neutral exploration is typically slow; the impact of intermediate strengths of genetic drift can be captured by a single fitting parameter. Third, bursts of similar mutants occur in monomorphic populations and can make the fixation of accessible, beneficial phenotypes highly likely.

Chapter 7 brought together the different results of the previous chapters by studying the dynamics of neutral exploration on RNA neutral spaces. While the quantitative accuracy of the mean-field calculations is reduced by the heterogeneity of RNA neutral spaces, the qualitative features remain valid: The static correlations in the spaces, which we discussed in Chapter 3, on average lead to an additional delay in the discovery of alternative phenotypes via neutral exploration, and to more pronounced bursts.

Finally, Section 7.4 brought together our study of neutral exploration and of fixation: We considered the fate of a population evolving on a fitness landscape of two disconnected peaks, one corresponding to a frequent, but only slightly beneficial phenotype while the phenotype of the other peak is less frequent but more fit. This setting allowed us to compare the relative importance of accessibility and fitness for the eventual fixation of a phenotype. By comparing our simulations to calculations which ignored neutral spaces [146], we showed that the slow speed of neutral exploration, which is typically required to discover rare phenotypes, makes the fixation of rare, yet highly adaptive phenotypes quite unlikely, because less beneficial but more easily produced phenotypes are often fixed much faster.

Our emphasis on the importance of mutations to generate variation allows us to rationalize several recent experimental results. For example, Burch and Chao showed that the course of evolution of an RNA virus is sensitive to the population's initial location in genotype space [16]. This finding not only underlines the heterogeneity of genotype space, but also shows that even for viruses (which have relatively high mutation rates [16, 34]), not all

alternative phenotypes are immediately accessible to the population.

The slow discovery of rare phenotypes, exacerbated by the necessity of neutral exploration (cf. Section 4.2.5), suggests that during the course of evolution on earth, probably not all possible phenotypes have been produced. A study by Povolotskaya and Kondrashov [103] underlines this idea: Due to the intricate structure of realistic neutral spaces (such as we have seen in Chapters 3 and 7), evolving populations continue to venture into unexplored regions of genotype space.

Another important result comes from the work of Hayden et al. [63], who showed that pre-existing genetic variation in a population allows for faster adaptation upon an environmental change. Hayden et al. showed that the beneficial mutants were already present in the population before the environmental shift, but only in the new environment did these mutants have a fitness benefit; in their experiment, the environmental shift changed the GP map, while in our theoretical setting, we only consider changes in the phenotype-fitness map. However, the authors also showed that in a population without pre-existing variation, the response to the environmental shift was significantly slowed down, highlighting that the introduction of variation can be the limiting factor of evolutionary progress.

We stress that the line of argument presented in this thesis is not meant to deny the role of natural selection in evolution, but rather to stress that other processes can be at least equally important (Lynch [84] gives an excellent discussion of this point). In particular, phenotypic bias affects the introduction of selectable variation [146] and provides an alternative hypothesis for why particular phenotypes are observed in nature. It is interesting to note that Jörg et al. found that the secondary structures of several RNA enzymes observed in nature appear to be associated with large neutral spaces [68]. Of course, this result does not imply that selection has played no role in the evolution of these enzymes; as we have seen in Fig. 3.5, the most abundant secondary structures appear to be quite regular and it has been speculated that this regularity allows for the construction of more efficient enzymes [42].

In summary, the results presented in this thesis underline that evolution is a highly complex process. While the role of neutral spaces in enabling neutral exploration has been recognized previously, the developments in this thesis also highlight the complementary

importance of neutral spaces to determine the accessibility of phenotypes. Overall, our results suggest that statistical predictions about evolutionary dynamics can be made from first principles, but in order to do so, knowledge about neutral spaces and their internal and external connections is indispensable.

8.2 Discussion and outlook

Fitness is one of the fundamental concepts for theoretical approaches to evolution. As we have seen in Chapter 2, a fruitful definition of fitness can be given in terms of the (expected) number of offspring of an individual. However, measuring fitness accurately is notoriously difficult, not least because the number of offspring is often hard to obtain experimentally. Instead, we have focused on phenotypes, which may are often easier to observe in nature, and we have postulated that there is an (effectively) unique mapping from phenotypes to fitness. But when are two phenotypes different? In RNA secondary structures, it appears that the answer is straightforward: Two structures are different if their base-pair arrangement is not identical. However, this is a rather coarse-grained description which ignores thermodynamic details such as the free energy and stability of the native fold [42]. Depending on the context in which an RNA molecule performs a function, these criteria (and others, such as the three-dimensional structure of the folded molecule or sequence dependencies in binding sites) may well be important. Overall, determining which aspects of a phenotype are relevant determinants of fitness is a complex problem, whose solution is likely to depend on the specific biological system under study. Elucidating the GP map and its role in the evolution of more complex systems is an important direction for future work.

Another interesting direction for future work is to study more involved dynamic scenarios. For example, we have always assumed a maximum degree of competition between individuals in the population: In principle, a single individual's offspring can form the entire population in the next generation. This assumption may be valid for well-mixed populations, but may fail when populations are spatially structured and migration between different sub-populations may limit genetic (and phenotypic) exchange in the total population

[92]. Further, the impact of neutral spaces and phenotypic bias in the evolution of sexually producing organisms is an interesting question. Section (3.4) scratched the surface of the complexities that arise in sexual organisms: Because reproduction brings together different sources of genetic material, the genotypic variation produced in sexual organisms is typically greater than in asexual ones. Whether this increased capacity for variation can simply be captured in terms of our mean-field description of phenotypic bias is an important open question.

Overall, our mean-field approach has been quite successful in describing the evolutionary dynamics on homogeneous neutral spaces. However, we have seen that neutral spaces arising from more biologically realistic GP maps can show strong heterogeneity across several scales. While it appears that our mean-field approach still gives a good first-order intuition, it is an evident question whether we can be more accurate. One possible way to make progress is try to divide a complex neutral space into homogeneous parts. Of course, this is a hard problem, and it is not clear a priori that this approach can be successful. Yet an interesting strategy has been put forward in a recent paper by Aguirre et al. [2], who used methods of community detection [45, 90, 102] on neutral spaces. Preliminary results suggest that this may be a useful avenue for further study.

Ultimately, if we wish to predict evolutionary dynamics, we need to understand evolutionary pathways across several phenotypes and neutral spaces. The success of the mean-field approach suggests that there may be useful ways of coarse-graining the description of neutral spaces in terms of their inter-connections. Given the insights we could gain into the dynamics of neutral exploration, extending our model to long-term evolution is another promising endeavour.

Bibliography

[1] http://judy.sourceforge.net.

[2] J. Aguirre, J. M. Buldù, M. Stich, and S. C. Manrubia. Topological structure of the space of phenotypes: The case of RNA neutral networks. *PLoS ONE*, 6(10):e26324, 2011.

[3] S. E. Ahnert, I. G. Johnston, T. M. A. Fink, J. P. K. Doye, and A. A. Louis. Self-assembly, modularity, and physical complexity. *Phys. Rev. E*, 82(2):026117, 2010.

[4] P. Alberch. From genes to phenotype: dynamical systems and evolvability. *Genetica*, 84(1):5–11, 1991. ISSN 0016-6707.

[5] U. Alon. *An introduction to systems biology: design principles of biological circuits*, volume 10. CRC press, 2007.

[6] L. Altenberg. Genome growth and the evolution of the genotype-phenotype map. *Evolution and biocomputation*, 899:205–259, 1995.

[7] A. Babajide, I. L. Hofacker, M. J. Sippl, and P. 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, 1997.

[8] J. Bard. A systems biology view of evolutionary genetics. *BioEssays*, 32(7):559–563, 2010.

[9] W. Bateson, E. R. Saunders, R. C. Punnett, and C. C. Hurst. Experimental studies in the physiology of heredity. *Reports to the Evolution Committee of the Royal Society*, 2:1–55, 1905.

[10] R. E. Beardmore, I. Gudelj, D. A. Lipson, and L. D. Hurst. Metabolic trade-offs and the maintenance of the fittest and the flattest. *Nature*, 472(7343):342–346, 2011.

[11] D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, and H. R. e. a. Bignell. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218): 53–59, 2008.

[12] Boldhaus, G. and Klemm, K. Regulatory networks and connected components of the neutral space. *Eur. Phys. J. B*, 77(2):233–237, 2010.

[13] E. Borenstein and D. C. Krakauer. An end to endless forms: Epistasis, phenotype distribution bias, and nonuniform evolution. *PLoS Comp. Biol.*, 4(10):e1000202, 2008.

[14] E. Bornberg-Bauer. How are model protein structures distributed in sequence space? *Biophys. J.*, 73(5):2393–2403, 1997.

[15] E. Bornberg-Bauer and H. S. Chan. Modeling evolutionary landscapes: Mutational stability, topology, and superfunnels in sequence space. *Proc. Natl. Acad. Sci. USA*, 96(19):10689 –10694, 1999.

[16] C. L. Burch and L. Chao. Evolvability of an RNA virus is determined by its mutational neighbourhood. *Nature*, 406(6796):625–628, 2000.

[17] B. Charlesworth. Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.*, 10(3):195–205, 2009.

[18] H. Check et al. Human genome at ten: Life is complicated. *Nature*, 464(7289):664, 2010.

[19] S. Ciliberti, O. C. Martin, and A. Wagner. Innovation and robustness in complex regulatory gene networks. *Proc. Natl. Acad. Sci. USA*, 104(34):13591–13596, 2007.

[20] S. Ciliberti, O. C. Martin, and A. Wagner. Robustness Can Evolve Gradually in Complex Regulatory Gene Networks with Varying Topology. *PLoS Comp. Biol.*, 3(2): e15, 2007.

[21] M. Conrad. The geometry of evolution. *Biosystems*, 24(1):61–81, 1990.

[22] M. Conrad and W. Ebeling. M.V. Volkenstein, evolutionary thinking and the structure of fitness landscapes. *BioSystems*, 27(3):125–128, 1992.

[23] M. C. Cowperthwaite and L. A. Meyers. How mutational networks shape evolution: lessons from RNA models. *Annu. Rev. Ecol. Evol. Syst.*, 38:203–230, 2007.

[24] M. C. Cowperthwaite, E. P. Economo, W. R. Harcombe, E. L. Miller, and L. A. Meyers. The Ascent of the Abundant: How Mutational Networks Constrain Evolution. *PLoS Comp. Biol.*, 4(7):e1000110, 2008.

[25] J. F. Crow and M. Kimura. Evolution in sexual and asexual populations. *American Naturalist*, 49(909):439–450, 1965.

[26] B. Daniels, Y. Chen, J. Sethna, R. Gutenkunst, and C. Myers. Sloppiness, robustness, and evolvability in systems biology. *Curr. Op. Biotech.*, 19(4):389–395, 2008.

[27] C. R. Darwin. *On the Origin of Species*. Murray, 1859.

[28] R. Dawkins. *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company, 1986.

[29] H. de Vries. *Species and Varieties, Their Origin by Mutation*. The Open Court Publishing Company, 1904.

[30] M. A. DePristo, D. M. Weinreich, and D. L. Hartl. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat. Rev. Genet.*, 6(9):678–687, 2005. ISSN 1471-0056.

[31] B. Derrida and L. Peliti. Evolution in a flat fitness landscape. *Bull. Math. Biol.*, 53 (3):355–382, 1991. ISSN 0092-8240.

[32] K. Dingle and A. A. Louis. in preparation.

[33] J. A. Draghi, T. L. Parsons, G. P. Wagner, and J. B. Plotkin. Mutational robustness can facilitate adaptation. *Nature*, 463(7279):353–355, 2010.

[34] J. W. Drake, B. Charlesworth, D. Charlesworth, and J. F. Crow. Rates of spontaneous mutation. *Genetics*, 148(4):1667–1686, 1998.

[35] M. Eigen. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523, 1971.

[36] M. Eigen and P. Schuster. A principle of natural self-organization. *Naturwissenschaften*, 64(11):541–565, 1977.

[37] M. Eigen and P. Schuster. The hypercycle. *Naturwissenschaften*, 65(1):7–41, 1978.

[38] M. Eigen and P. Schuster. *The Hypercycle: A principle of self-organization.* Springer, Berlin and Heidelberg, Germany, 1979.

[39] E. H. Ekland, J. W. Szostak, and D. P. Bartel. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science*, 269(5222):364–370, 1995.

[40] E. Ferrada and A. Wagner. Evolutionary innovations and the organization of protein functions in genotype space. *PLoS ONE*, 5(11):e14172, 2010.

[41] R. A. Fisher. *The genetical theory of natural selection.* Clarendon Press, 1930.

[42] W. Fontana. Modeling 'evo-devo' with RNA. *BioEssays*, 24(12):1164–1177, 2002.

[43] W. Fontana, D. A. M. Konings, P. F. Stadler, and P. Schuster. Statistics of RNA secondary structures. *Biopolymers*, 33(9):1389–1404, 1993.

[44] R. Forster, C. Adami, and C. Wilke. Selection for mutational robustness in finite populations. *J. Theor. Biol.*, 243(2):181–190, 2006.

[45] S. Fortunato. Community detection in graphs. *Phys. Rep.*, 486(3-5):75–174, 2010.

[46] H. Frauenfelder and D. T. Leeson. The energy landscape in non-biological and biological molecules. *Nature Structural & Molecular Biology*, 5(9):757–759, 1998.

[47] S. J. Freeland and L. D. Hurst. The genetic code is one in a million. *J. Mol. Evol.*, 47 (3):238–248, 1998. ISSN 0022-2844.

[48] T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991.

[49] A. Gallagher. *Evolvability: a Formal Approach.* PhD thesis, University of Oxford, 2009.

[50] H. Gao, J. M. Granka, and M. W. Feldman. On the classification of epistatic interactions. *Genetics*, 184(3):827–837, 2010.

[51] S. Gavrilets. *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, NJ, 2004.

[52] S. Gavrilets and J. Gravner. Percolation on the fitness hypercube and the evolution of reproductive isolation. *J. Theor. Biol.*, 184(1):51–64, JAN 7 1997.

[53] J. H. Gillespie. Molecular evolution over the mutational landscape. *Evolution*, 38(5): 1116–1129, 1984.

[54] J. H. Gillespie. *The causes of molecular evolution*. Oxford University Press, Oxford, UK, 1994.

[55] N. Goldenfeld and C. Woese. Life is physics: Evolution as a collective phenomenon far from equilibrium. *Ann. Rev. Cond. Matt. Phys.*, 2(1):375–399, 2011.

[56] W. Grüner, R. Giegerich, D. Strothmann, C. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster. Analysis of RNA sequence structure maps by exhaustive enumeration i. neutral networks. *Monatsh. Chemie*, 127(4):355–374, 1996.

[57] R. Gutenkunst, F. Casey, J. Waterfall, C. Myers, and J. Sethna. Extracting falsifiable predictions from sloppy models. *Annals of the New York Academy of Sciences*, 1115 (1):203–211, 2007.

[58] R. Gutenkunst, J. Waterfall, F. Casey, K. Brown, C. Myers, and J. Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comp. Biol.*, 3(10): e189, 2007.

[59] D. Haig and L. D. Hurst. A quantitative measure of error minimization in the genetic code. *J. Mol. Evol.*, 33(5):412–417, 1991. ISSN 0022-2844.

[60] J. B. S. Haldane. *The causes of evolution*. Longmans, Green and co., London and New York, 1932.

[61] R. W. Hamming. Error detecting and error correcting codes. *Bell System technical journal*, 29(2):147–160, 1950.

[62] D. L. Hartl and A. G. Clark. *Principles of population genetics*. Sinauer Associates; Sunderland, MA, 1997.

[63] E. J. Hayden, E. Ferrada, and A. Wagner. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature*, 474(7349):92–95, 2011.

[64] I. L. Hofacker, W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatsh. Chemie*, 125(2): 167–188, 1994.

[65] M. A. Huynen. Exploring phenotype space through neutral evolution. *J. Mol. Evol.*, 43(3):165–169, 1996. ISSN 0022-2844.

[66] S. Itzkovitz and U. Alon. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Research*, 17(4):405–412, 2007.

[67] I. G. Johnston, S. E. Ahnert, J. P. K. Doye, and A. A. Louis. Evolutionary dynamics in a simple model of self-assembly. *Phys. Rev. E*, 83(6):066105, 2011.

[68] T. Jörg, O. Martin, and A. Wagner. Neutral network sizes of biological RNA molecules can be computed and are not atypically small. *BMC Bioinf.*, 9(1):464, 2008.

[69] S. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.*, 128(1):11–45, 1987.

[70] M. Kellis, B. W. Birren, and E. S. Lander. Proof and evolutionary analysis of ancient genome duplication in the yeast saccharomyces cerevisiae. *Nature*, 428(6983):617–624, 2004.

[71] M. Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713, 1962.

[72] M. Kimura. Diffusion models in population genetics. *J. Appl. Prob.*, 1(2):177–232, 1964.

[73] M. Kimura. Evolutionary rate at the molecular level. *Nature*, 217(5129):624, 1968.

[74] M. Kimura. *The Neutral Theory of Molecular Evolution*. Cambrige University Press, Cambridge, UK, 1985.

[75] M. Kimura. The role of compensatory neutral mutations in molecular evolution. *J. Genetics*, 64(1):7–19, 1985. ISSN 0022-1333.

[76] M. Kirschner and J. Gerhart. Evolvability. *Proc. Nat. Acad. Sci. USA*, 95(15):8420–8427, 1998.

[77] E. Koonin, K. Makarova, and L. Aravind. Horizontal gene transfer in prokaryotes: Quantification and classification 1. *Ann. Rev. Microbiol.*, 55(1):709–742, 2001.

[78] E. Lander, L. Linton, B. Birren, C. Nusbaum, M. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, et al. Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921, 2001.

[79] M. Lässig. From biophysics to evolutionary genetics: statistical aspects of gene regulation. *BMC Bioinf.*, 8(Suppl 6):S7, 2007.

[80] K. Lau, S. Ganguli, and C. Tang. Function constrains network architecture and dynamics: A case study on the yeast cell cycle boolean network. *Phys. Rev. E*, 75(5): 051907, 2007.

[81] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang. The yeast cell-cycle network is robustly designed. *Proc. Nat. Acad. Sci. USA*, 101(14):4781, 2004.

[82] H. Li, R. Helling, C. Tang, and N. Wingreen. Emergence of preferred structures in a simple model of protein folding. *Science*, 273(5275):666 –669, 1996.

[83] D. J. Lipman and W. J. Wilbur. Modelling Neutral and Selective Evolution of Protein Folding. *Proc. Roy. Soc. B*, 245(1312):7 –11, 1991.

[84] M. Lynch. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc. Nat. Acad. Sci. USA*, 104(Suppl 1):8597, 2007.

[85] J. Maynard Smith. Natural Selection and the Concept of a Protein Space. *Nature*, 225:563–564, 1970.

[86] G. Mendel. Experiments in plant hybridization. *Proceedings of the Natural History Society of Brünn*, 4:281–321, 1866.

[87] P. A. P. Moran. *The statistical processes of evolutionary theory.* Oxford University Press, Oxford, UK, 1962.

[88] V. Mustonen, J. Kinney, C. Callan, and M. Lässig. Energy-dependent fitness: A quantitative model for the evolution of yeast transcription factor binding sites. *Proc. Nat. Acad. Sci. USA*, 105(34):12376, 2008.

[89] M. Nei. *Molecular evolutionary genetics.* Columbia University Press, New York, NY, 1987.

[90] M. Newman. *Networks: an introduction.* Oxford University Press, Oxford, UK, 2010.

[91] M. Nirenberg and P. Leder. RNA codewords and protein synthesis. *Science*, 145(3639): 1399–1407, 1964.

[92] M. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life.* Harvard University Press, Cambridge, MA, 2006.

[93] T. Ohta. Slightly deleterious mutant substitutions in evolution. *Nature*, 246(5428): 96–98, 1973.

[94] H. Orr. The population genetics of adaptation: the adaptation of DNA sequences. *Evolution*, 56(7):1317–1330, 2002.

[95] H. Orr. The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.*, 6(2): 119–127, 2005.

[96] Z. Patwa and L. M. Wahl. The fixation probability of beneficial mutations. *Roy. Soc. Interface*, 5(28):1279–1289, 2008.

[97] P. C. Phillips. Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.*, 9(11):855–867, 2008.

[98] R. Phillips, J. Kondev, and J. Theriot. *Physical Biology of the Cell.* Garland Science, New York, NY, 2008.

[99] M. Pigliucci. Opinion - Is evolvability evolvable? *Nat. Rev. Genet.*, 9(1):75–82, 2008.

[100] M. Pigliucci. Sewall Wright's adaptive landscapes: 1932 vs. 1988. *Biology & Philosophy*, 23(5):591–603, 2008.

[101] M. Pigliucci. Genotype–phenotype mapping and the end of the 'genes as blueprint' metaphor. *Phil. Trans. Roy. Soc. B*, 365(1540):557–566, 2010.

[102] M. A. Porter, J. P. Onnela, and P. J. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009.

[103] I. S. Povolotskaya and F. A. Kondrashov. Sequence space and the ongoing expansion of the protein universe. *Nature*, 465(7300):922–926, 2010.

[104] S. Psujek and R. D. Beer. Developmental bias in evolution: evolutionary accessibility of phenotypes in a model evo-devo system. *Evolution & Development*, 10(3):375–390, 2008.

[105] K. Raman and A. Wagner. The evolvability of programmable hardware. *Roy. Soc. Interface*, 8(55):269–281, 2011.

[106] K. Raman and A. Wagner. Evolvability and robustness in a complex signalling circuit. *Mol. BioSyst.*, 7(4):1081–1092, 2011.

[107] C. Reidys, P. Stadler, and P. Schuster. Generic properties of combinatory maps: Neutral networks of RNA secondary structures. *Bull. Math. Biol.*, 59(2):339–397, 1997.

[108] C. M. Reidys. Random induced subgraphs of generalized n-cubes. *Adv. Appl. Math.*, 19(3):360–377, 1997.

[109] C. M. Reidys. Large components in random induced subgraphs of n-cubes. *Discrete Mathematics*, 309(10):3113–3124, 2009.

[110] C. M. Reidys and P. F. Stadler. Combinatorial landscapes. *SIAM Review*, 44(1):3–54, 2002.

[111] M. Rendel. Adaptive evolutionary walks require neutral intermediates in RNA fitness landscapes. *Theor. Pop. Biol.*, 79(1):12–18, 2011.

[112] J. F. M. Rodrigues and A. Wagner. Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLoS Comp. Biol.*, 5(12):e1000613, 2009.

[113] A. Samal, J. M. Rodrigues, J. Jost, O. Martin, and A. Wagner. Genotype networks in metabolic reaction spaces. *BMC Sys. Bio.*, 4(1):30, 2010. ISSN 1752-0509.

[114] S. Schaper, I. G. Johnston, and A. A. Louis. Epistasis can lead to fragmented neutral spaces and contingency in evolution. *Proc. Roy. Soc. B*, 279(1734):1777–1783, 2012.

[115] P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker. From sequences to shapes and back: A case study in RNA secondary structures. *Proc. Roy. Soc. B*, 255(1344): 279–284, 1994.

[116] D. Sherrington. Landscape paradigms in physics and biology: introduction and overview. *Physica D*, 107(2-4):117–121, 1997.

[117] D. Sherrington. Physics and complexity. *Phil. Trans. Roy. Soc. A*, 368(1914):1175–1189, 2010.

[118] Sumedha, O. C. Martin, and L. Peliti. Population size effects in evolutionary dynamics on neutral networks and toy landscapes. *J. Stat. Mech.: Theory and Experiment*, 2007 (05):P05011, 2007. ISSN 1742-5468.

[119] D. Tautz. Evolution of transcriptional regulation. *Curr. Op. Genet. & Devel.*, 10(5): 575–579, 2000.

[120] M. Transtrum, B. Machta, and J. Sethna. Geometry of nonlinear least squares with applications to sloppy models and optimization. *Phys. Rev. E*, 83(3):036701, 2011.

[121] E. van Nimwegen and J. Crutchfield. Metastable evolutionary dynamics: Crossing fitness barriers or escaping via neutral paths? *Bull. Math. Biol.*, 62(5):799–848, 2000. ISSN 0092-8240.

[122] E. Van Nimwegen, J. Crutchfield, and M. Mitchell. Finite populations induce metastability in evolutionary search. *Phys. Lett. A*, 229(3):144–150, 1997.

[123] E. van Nimwegen, J. P. Crutchfield, and M. Huynen. Neutral Evolution of Mutational Robustness. *Proc. Nat. Acad. Sci. USA*, 96(17):9716–9720, 1999.

[124] K. Vetsigian, C. Woese, and N. Goldenfeld. Collective evolution and the genetic code. *Proc. Nat. Acad. Sci. USA*, 103(28):10696–10701, 2006.

[125] A. Wagner. *Robustness and Evolvability in Living Systems*. Princeton University Press, Princeton, NJ, 2005.

[126] A. Wagner. Neutralism and selectionism: a network-based reconciliation. *Nat. Rev. Genet.*, 9(12):965–974, 2008.

[127] A. Wagner. Robustness and evolvability: a paradox resolved. *Proc. Roy. Soc. B*, 275 (1630):91–100, 2008.

[128] A. Wagner. *The origins of evolutionary innovations*. Oxford University Press, Oxford, UK, 2011.

[129] A. Wagner. Genotype networks shed light on evolutionary constraints. *Trends in ecology & evolution*, 26(11):577–584, 2011.

[130] A. Wagner. The molecular origins of evolutionary innovations. *Trends in Genetics*, 27 (10):397–410, 2011.

[131] G. P. Wagner and L. Altenberg. Perspective: Complex adaptations and the evolution of evolvability. *Evolution*, 50(3):967–976, 1996. ISSN 00143820.

[132] J. Waterfall, F. Casey, R. Gutenkunst, K. Brown, C. Myers, P. Brouwer, V. Elser, and J. Sethna. Sloppy-model universality class and the vandermonde matrix. *Phys. Rev. Lett.*, 97(15):150601, 2006.

[133] M. S. Waterman. Secondary structure of single-stranded nucleic acids. *Adv. math. suppl. studies*, 1:167–212, 1978.

[134] J. D. Watson and F. H. C. Crick. A Structure for Deoxyribose Nucleic Acid. *Nature*, 171:737–738, 1953.

[135] D. M. Weinreich, R. A. Watson, and L. Chao. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*, 59(6):1165–1174, 2005. ISSN 1558-5646.

[136] D. M. Weinreich, N. F. Delaney, M. A. DePristo, and D. L. Hartl. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, 312(5770):111–114, 2006.

[137] D. A. Wheeler, M. Srinivasan, M. Egholm, Y. Shen, L. Chen, A. McGuire, W. He, Y. J. Chen, V. Makhijani, and G. T. e. a. Roth. The complete genome of an individual by massively parallel DNA sequencing. *Nature*, 452(7189):872–876, 2008.

[138] C. Wilke. Quasispecies theory in the context of population genetics. *BMC Evol. Biol.*, 5(1):44, 2005.

[139] C. Wilke and C. Adami. Evolution of mutational robustness. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 522(1-2):3–11, 2003.

[140] C. O. Wilke, J. L. Wang, C. Ofria, R. E. Lenski, and C. Adami. Evolution of digital organisms at high mutation rates leads to survival of the attest. *Nature*, 412(6844): 331–333, 2001.

[141] C. Woese. *The Genetic Code: The Molecular Basis for Genetic Expression*. Harper & Row, New York, NY, 1967.

[142] C. R. Woese, D. H. Dugre, S. A. Dugre, M. Kondo, and W. C. Saxinger. On the fundamental nature and evolution of the genetic code. *Cold Spring Harbor Symposia on Quantitative Biology*, 31:723 –736, 1966.

[143] S. Wright. Evolution in mendelian populations. *Genetics*, 16(2):97, 1931.

[144] S. Wright. The shifting balance theory and macroevolution. *Ann. Rev. Genet.*, 16(1): 1–20, 1982.

[145] S. G. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. 6th Int. Cong. Genet.*, 1:356–366, 1932.

[146] L. Y. Yampolsky and A. Stoltzfus. Bias in the introduction of variation as an orienting factor in evolution. *Evolution & Development*, 3(2):73–83, 2001.

Appendix A

Neutral diversity in the absence of selection

The material presented here provides an extension of the work of Derrida and Peliti [31], who studied the dynamics of the Wright-Fisher model in a flat fitness landscape. They mapped this system onto a spin-glass model, in which a binary alphabet ($K = 2$) is natural. Here, we consider the case of general K and provide a derivation of Equation (2.31).

As a first step towards understanding the evolution and maintenance of neutral diversity, we consider a random walk on the genotype hypercube. Specifically, we are interested in the number of mutations d that the walker acquires with respect to its starting position during t generations, given that a mutation at a particular locus occurs with probability μ per generation.

When $d(t)$ is the Hamming distance of the genotype at t to the source genotype (ie, $d(t = 0) = 0$) and μ is small enough that double mutations can be ignored (ie. $L\mu \ll 1$), we have the following transition probabilities:

$$\begin{aligned} P(d-1 \to d) &= L\mu \frac{L-d+1}{L} \\ P(d \to d) &= (1 - L\mu) + L\mu \frac{d}{L} \frac{K-2}{K-1} \\ P(d+1 \to d) &= L\mu \frac{d+1}{L} \frac{1}{K-1} \end{aligned}$$

In words, $P(d-1 \to d)$ is the probability that a mutation happens ($L\mu$), and that this mutation happens at one of the $L-(d-1)$ loci which are currently still identical with the initial genotype. $P(d \to d)$ is the probability that either no mutation happens ($1 - L\mu$) or that the mutation changes one of the d already mutated loci into another allele that also differs from the initial genotype (note that in the case of binary sequences, $K = 2$, such mutations are not possible). Finally $P(d+1 \to d)$ is the probability of back mutations, which occur when one of the $d+1$ loci that has already mutated undergoes another mutation which takes it back to the original allele.

Let $\pi_d(t)$ be the probability that at time t, the walker is at Hamming distance d from the initial genotype. In a continuous time approximation, we thus have

$$\frac{d\pi_d}{dt} = \mu \left(\frac{d+1}{K-1} \pi_{d+1} + (L-d+1)\pi_{d-1} - \left(L - d\frac{K-2}{K-1} \right) \pi_d \right) \tag{A.1}$$

To solve this equation, we make the ansatz

$$\pi_d(t) = \binom{L}{d} (1-q)^d \, q^{L-d} \tag{A.2}$$

where $q(t)$ is the probability that a randomly chosen locus has not undergone a mutation by time t. By tedious, but straightforward calculation, we can show that

$$q(t) = \frac{1 + (K-1)\exp\left(-\frac{K}{K-1}\mu t\right)}{K} \tag{A.3}$$

Now we can translate these results for the random walk into the dynamics of an evolving population under the Wright-Fisher scheme (see Section 2.1.2). We are interested here in the Hamming distance of the genotypes of two randomly chosen individuals of a population of size N. To calculate this measure of diversity, we can trace the ancestry of the two individuals back in time to the point where they had a common ancestor, and hence the same genotype. Derrida and Peliti [31] showed how the time T back to the common ancestor is distributed: Looking back in time, at each generation every individual 'picks' its ancestor randomly from N potential parents. Hence the probability that two individuals pick the same parent is $1/N$ per generation, and so T follows a geometrical distribution with mean N. Moving to a continuous time approximation, the probability that two random individuals had their most recent common ancestor between time t and $t + dt$ is thus

$$P_a(t)dt = e^{-t/N} dt/N \tag{A.4}$$

After diverging from the common ancestor, both individuals performed indepent random walks on genotype space. So by combining Equations (A.2) and (A.4), we find that the probability that the genotypes of two random individuals in the population have Hamming distance d is simply

$$P^H(d) = \int_0^\infty \frac{dt}{N} \exp(-t/N)\pi_d(2t) \tag{A.5}$$

where the factor of 2 in the argument of π_d accounts for the fact the both individuals diverged from the common ancestor. While this integral can be carried out exactly, the calculation involves unwieldy special functions. Instead, we consider the limit of large L: Since $\pi_d(t)$ is a binomial distribution, in the large L limit the variance becomes small compared to the mean and we have approximately

$$\pi_d(t) \approx \delta(d - \bar{d}(t))$$

where $\delta(x)$ is the Dirac δ-function and

$$\bar{d}(t) = \frac{L(K-1)}{K}\left(1 - \exp\left(-\frac{K-1}{K}\mu t\right)\right)$$

is the expectation of d over the distribution $\pi_d(t)$.

For two functions $f(x), g(x)$ we have

$$\int_D dx f(x)\delta(g(x)) = \sum_{x_0} \frac{f(x_0)}{g'(x_0)}$$

where D is the domain of integration, $g' \equiv dg/dx$ and the sum goes over all the $x_0 \in D$ for which $g(x_0) = 0$. So we obtain

$$\begin{aligned} P^H(d) &= \int_0^\infty \frac{dt}{N} \exp(-t/N)\delta\left(d - \frac{L(K-1)}{K}\left(1 - \exp\left(-\frac{K-1}{K}\mu 2t\right)\right)\right) \\ &= \frac{\sigma}{d_\infty}\left(1 - \frac{d}{d_\infty}\right)^{\sigma-1} \end{aligned} \tag{A.6}$$

where $\sigma = (K-1)/(2K\mu N)$ and $d_\infty = L(K-1)/K$. Note that $P^H(d)$ can be complex when $d > d_\infty$ and $\sigma < 1$, and should be taken to vanish in this case.

Finally, we average over all pairs of individuals to obtain the average Hamming distance $\langle d \rangle$ of two random individuals in the population:

$$\begin{aligned} \langle d \rangle &= \sum_{d=0}^{d_\infty} dP^H(d) \\ &\approx d_\infty \frac{\int_0^1 dx \sigma x (1-x)^{\sigma-1}}{\int_0^1 \sigma(1-x)^{\sigma-1}} \\ &= \frac{d_\infty}{1+\sigma} \\ &= \frac{d_\infty}{1 + \frac{d_\infty}{2NL\mu}} \end{aligned} \tag{A.7}$$

where the second step used $x = d/d_\infty$ and replaced the summation by an integral. This concludes the derivation of Equation (2.31).

Appendix B

Algorithmic details

The first computational problem in characterizing the properties of neutral spaces is to find all their genotypes. The following procedure [56], can be used to map out a single neutral component (NC) of phenotype p, given a starting genotype g_0.

1. Initialize two empty lists U and V. Append g_0 into U.
2. Pick a genotype g from U.
3. For each neighbour g' of g, obtain its phenotype p' from the GP map. If $p' = p$ and g' is neither in U nor V, append g' to U.
4. Remove g from U and append g to V.
5. If U is not empty, return to step 2.
6. When U is empty, V contains a neutral component of phenotype p.

The function of the algorithm becomes clear when think of U as the set of 'unvisited' genotypes, and of V as the 'visited' genotypes. So the algorithm takes an unvisited genotype, visits all its neighbours, and if a new neutral neighbour is found, this neighbour is marked as unvisited. The algorithm terminates when all neutrally accessible genotypes have been visited.

By extending step 3 of the algorithm, we can also characterize the mutational neighbourhood of the neutral component. To this end, we simply keep track how often each alternative phenotype p' is encountered.

We note that this algorithm can be used whenever all mutants can be found from a single genotype alone. That is, the algorithm is not restricted to single point mutations, but can in principle also be applied for double mutations, insertions or deletions. The case of crossover (cf. Section 3.4) is more involved, simply because crossover requires two genotypes.

Regarding an efficient implementation, the main challenge is that at the outset, the size of the NC is unknown[1]. And since NCs can be vast (in Section 3.3.5, we deal with single NCs of more than 5×10^8 genotypes), it is important to have a data structure that is compact in memory and enables fast addition, removal and checking of elements. JudyArrays [1] proved to be an excellent tool to achieve these goals.

[1] If we know the total size of the neutral network from exhaustive enumeration, this size provides an upper bound.