

Review

# From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics

Susanna Manrubia<sup>a,b,\*,1</sup>, José A. Cuesta<sup>b,c,d,e,1</sup>, Jacobo Aguirre<sup>b,f</sup>, Sebastian E. Ahnert<sup>g,h</sup>,  
Lee Altenberg<sup>k</sup>, Alejandro V. Cano<sup>l,m</sup>, Pablo Catalán<sup>b,c</sup>, Ramon Diaz-Uriarte<sup>n,o</sup>,  
Santiago F. Elena<sup>p,q</sup>, Juan Antonio García-Martín<sup>r</sup>, Paulien Hogeweg<sup>s</sup>, Bhavin  
S. Khatri<sup>t,u</sup>, Joachim Krug<sup>v</sup>, Ard A. Louis<sup>w</sup>, Nora S. Martin<sup>i,j</sup>, Joshua L. Payne<sup>l,m</sup>,  
Matthew J. Tarnowski<sup>x</sup>, Marcel Weiß<sup>i,j</sup>

<sup>a</sup> Department of Systems Biology, Centro Nacional de Biotecnología (CSIC), Madrid, Spain

<sup>b</sup> Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain

<sup>c</sup> Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés, Spain

<sup>d</sup> Instituto de Biocomputación y Física de Sistemas Complejos (BiFi), Universidad de Zaragoza, Spain

<sup>e</sup> UC3M-Santander Big Data Institute (IBiDat), Getafe, Madrid, Spain

<sup>f</sup> Centro de Astrobiología, CSIC-INTA, ctra. de Ajalvir km 4, 28850 Torrejón de Ardoz, Madrid, Spain

<sup>g</sup> Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge CB3 0AS, UK

<sup>h</sup> The Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, UK

<sup>i</sup> Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, Cambridge, UK

<sup>j</sup> Sainsbury Laboratory, University of Cambridge, Cambridge, UK

<sup>k</sup> University of Hawai'i at Manoa, HI, USA

<sup>l</sup> Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

<sup>m</sup> Swiss Institute of Bioinformatics, Lausanne, Switzerland

<sup>n</sup> Department of Biochemistry, Universidad Autónoma de Madrid, Madrid, Spain

<sup>o</sup> Instituto de Investigaciones Biomédicas “Alberto Sols” (UAM-CSIC), Madrid, Spain

<sup>p</sup> Instituto de Biología Integrativa de Sistemas, I<sup>2</sup>SysBio (CSIC-UV), València, Spain

<sup>q</sup> The Santa Fe Institute, Santa Fe, NM, USA

<sup>r</sup> Bioinformatics for Genomics and Proteomics, Centro Nacional de Biotecnología (CSIC), Madrid, Spain

<sup>s</sup> Theoretical Biology and Bioinformatics Group, Utrecht University, the Netherlands

<sup>t</sup> The Francis Crick Institute, London, UK

<sup>u</sup> Department of Life Sciences, Imperial College London, London, UK

<sup>v</sup> Institute for Biological Physics, University of Cologne, Köln, Germany

<sup>w</sup> Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, UK

<sup>x</sup> School of Biological Sciences, University of Bristol, Bristol, UK

Received 24 November 2020; accepted 1 March 2021

Available online 21 May 2021

Communicated by E. Shakhnovich

\* Corresponding author at: Department of Systems Biology, Centro Nacional de Biotecnología (CSIC), Madrid, Spain.  
E-mail address: [smanrubia@cnb.csic.es](mailto:smanrubia@cnb.csic.es) (S. Manrubia).

<sup>1</sup> Organizing and leading authors; contributed equally to this work.

---

## Abstract

Understanding how genotypes map onto phenotypes, fitness, and eventually organisms is arguably the next major missing piece in a fully predictive theory of evolution. We refer to this generally as the problem of the genotype-phenotype map. Though we are still far from achieving a complete picture of these relationships, our current understanding of simpler questions, such as the structure induced in the space of genotypes by sequences mapped to molecular structures, has revealed important facts that deeply affect the dynamical description of evolutionary processes. Empirical evidence supporting the fundamental relevance of features such as phenotypic bias is mounting as well, while the synthesis of conceptual and experimental progress leads to questioning current assumptions on the nature of evolutionary dynamics—cancer progression models or synthetic biology approaches being notable examples. This work delves with a critical and constructive attitude into our current knowledge of how genotypes map onto molecular phenotypes and organismal functions, and discusses theoretical and empirical avenues to broaden and improve this comprehension. As a final goal, this community should aim at deriving an updated picture of evolutionary processes soundly relying on the structural properties of genotype spaces, as revealed by modern techniques of molecular and functional analysis.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Molecular evolution; Genotype-phenotype map; Genotype network; Fitness landscape; Phenotypic bias; Experimental evolution

---

## 1. Introduction

How genetic variation contributes to phenotypic variation is an essential question that must be answered to understand the evolutionary process. The experimental characterisation of the genotype-phenotype (GP) relationship is a formidable theoretical and experimental challenge, but also an expensive task which suffers from severe practical limitations. Computational approaches have been recurrently used to make predictions of phenotypes from genotypes and to uncover the statistical features of that relationship. Advances notwithstanding, an apparently insurmountable problem remains: the astronomically large size of the space of genotypes. The space of possible phenotypic change and the probabilities of such change are directly determined by the architecture of the GP map; to quantify this map will improve our understanding of how the space of phenotypes is explored and answer important questions about the probability of evolutionary rescue or innovation under endogenous or exogenous changes.

Progress in our understanding of GP maps at various levels is of relevance for different scientific communities with interests that range from evolutionary theory to molecular design through genomic bases of disease aetiology. Understanding of how RNA, DNA or amino acid sequences map onto molecular function could be of great importance for more fundamental approaches in synthetic biology, biotechnology, and systems chemistry. In a broader ecological context, the way in which generic properties of the GP map shape adaptation have rarely been explored. As of today, the overarching question of whether organismal phenotypes can be predicted from microscopic properties of genotype spaces remains open.

In this review, we discuss the state-of-the-art of genotype-to-organism research and future research avenues in the field. The review is structured into four major parts. The first part is constituted by this introduction and Section 2, which puts in perspective how relevant the generation of variation is in the evolutionary process, and introduces important biases arising from the inherent structure of genotype spaces.

The second part comprises sections 3 to 6, where we discuss conceptual approaches to the static properties of GP maps and their dynamical consequences, as well as the evolution of GP maps themselves. The field is broad and several aspects have been addressed in previous reviews, so we only briefly summarise topics dealt with elsewhere. Therefore, we will succinctly present computational GP maps and only recapitulate, taking an integrative and explanatory viewpoint, the topological properties of the space of genotypes [1–6]. Section 3 constitutes a synthetic overview of GP map models, including paradigmatic examples such as RNA folding, more recent multi-level models such as toyLIFE, and a summary of artificial life examples. Readers familiar with those models can safely skip that section. Those models endow genotype spaces with topological properties that are briefly reviewed in the introduction of Section 4, which is mostly devoted to discussing possible roots for generic properties of a broad class of GP maps. Attention is subsequently devoted to population dynamics on genotype spaces, which has been a less explored topic. Section 5 describes transient and equilibrium dynamical features of evolutionary processes. First, it delves into the effects of recombination and mutation bias, and on phenotypic transitions caused by the hierarchical, networked

structure of genotype spaces. Then, a mean-field description that incorporates the essentials of GP map topology to clarify major dynamical features is discussed. The section finishes with a derivation of equilibrium properties in the context of statistical mechanics and some applied examples. Section 6 discusses the evolution of GP maps themselves by means of two illustrative examples: a scenario where a multifunctional quasispecies emerges and a model of virtual cells incorporating the evolution of genome size.

The third part, sections 7 and 8 is devoted to empirical GP maps and to biological applications, and mostly presents topics under development. Section 7 examines most recent achievements regarding the experimental characterisation of GP and genotype-to-function maps in molecules and simple organisms, and the different possibilities that current and future techniques might allow. It includes a formal discussion on how phenotypes can be inferred from genotypic data and fitness assays, and a discussion of the intimate relationship between fitness landscapes and GP maps. Section 8 exemplifies how concepts and techniques originating in quantitative studies of the GP map can enlighten useful approaches to diseases with a genetic component.

The fourth and last part presents a mostly self-contained overview of open questions and difficulties that the field faces, as well as some possible avenues for further progress, in Section 9. The paper closes with an outlook in Section 10 where we reflect on the feasibility of characterising the genotype-to-organism map, and on plausible epistemological difficulties to comprehend the organisation and complexity of full organisms.

## 2. GP maps and the importance of variation

Darwinian evolution requires heritable phenotypic variation, upon which natural selection acts. Much of traditional evolutionary theory has focused on the role of natural selection, while the study of variation has been much less developed. There are a number of reasons for this difference.

Firstly, there is an influential tradition, stemming from the early days of the modern synthesis, that any meaningful change over evolutionary time is ultimately caused by natural selection. One argument in favour of this thesis comes from the simple observation that a heritable phenotype with higher fitness will, over the generations, exponentially out-compete other phenotypes with lower fitness in the same population. Thus, differences in the rate at which mutations arrive will be swamped by the effect of fitness differences (there are much more sophisticated versions of this argument). Another argument, which is often more implicitly than explicitly made, is that a large part of variation is *isotropic*—in other words, it is not biased in one direction or another. Stephen J. Gould, who was critical of this viewpoint, expresses it as follows: “*variation becomes raw material only, an isotropic sphere of potential about the modal form of a species . . . [only] natural selection . . . can manufacture substantial, directional change*” [7]. Whether evolutionary trends must primarily be explained by natural selection, or whether anisotropic (biased) variation also plays a key role, is a complex question. While the arguments have moved on considerably since the critique of Gould, especially with the rise of evo-devo [8], they are far from being settled [9,10]. Ever since the modern synthesis, directed variation has been deemed anathema because it evokes the Lamarckian view of variation to facilitate adaptation. However, as the analysis of GP maps reveals, these maps are a major source of anisotropic variation, even if this variation is not necessarily biased in the most beneficial way for the organism.

The second reason why our understanding of variation is relatively underdeveloped is that working out the exact role played by the arrival of variation in evolutionary history is difficult because in nature we typically only observe the final outcomes of an evolutionary process. It is hard to know what variation may have arisen in the past, but not fixed, or what variation could have potentially arisen, but did not. For example, even when all potential variation is isotropic, the non-lethal variation may well be anisotropic, depending on the environment.

In this context, the study of GP maps is critical, because they provide access to the way that changes in genotypes, brought on by various kinds of mutations, are translated into phenotypic variation for the biological system that the map describes. They allow us to ask important *counterfactual* questions, such as what is the full spectrum of variation that could potentially arise? Working out how variation affects evolutionary outcomes depends on an understanding of such counterfactuals.

A final issue for understanding variation comes from the unfathomable vastness of genotype spaces, whose size grows exponentially with genome length, rapidly leading to hyperastronomical numbers of possibilities [11]. If these spaces are so unimaginably vast, then it might seem natural to conclude, as many have done, that the variation that appears in evolutionary history is largely contingent upon accidents of history, and unlikely to be repeated (see Ref. [11] for a discussion).

This problem of hyperastronomically large spaces means that only relatively simple GP maps allow global questions about the full spectrum of possible variation to be addressed. Nevertheless, important progress in this direction has been made through the use of GP maps that can be computationally explored and, more recently, through the development of quantitative approaches to shared generic properties. Among the latter, one of the most striking properties is a strong *bias* in the number of genotypes mapping to a phenotype [4,12]. This begs the question: Can this bias, which often extends over many orders of magnitude, affect evolutionary outcomes? Indeed, phenotypic bias, among other non-trivial properties of GP maps, does severely affect not only our understanding of how variation arises through random mutations, but also any accurate representation—be it metaphorical or formal—of evolutionary dynamics at large.

### 3. Models of the GP map

Maynard Smith introduced the notion of a mapping from a genetic space to a molecular structure—and with it the idea of a network linking viable genotypes—as a resolution of an evolutionary paradox pointed out by Salisbury [13]. In brief, Salisbury noted [14] that the number of possible amino acid sequences exceeds by many orders of magnitude the number of proteins that ever existed on Earth since the origin of life, and concluded from this fact that functionally effective proteins have a vanishingly small chance of arising by mutation. As a way out of this dilemma, Maynard Smith suggested that the existence of networks of functional proteins is essential to navigate the space of genotypes to produce a sequence of adaptive improvements and to explore new regions that, eventually, secure evolutionary innovation [15]. Formally, the space of genotypes can be defined as a network where nodes represent genotypes, with any two nodes linked if they are mutually accessible through a single point mutation [16]. A *neutral network* is therefore an ensemble of connected genotypes with the same fitness, including those with identical phenotypes. The empirical existence of such networks and their role in providing access to new phenotypes [17] was unequivocally demonstrated [18,19] four decades after Maynard Smith's conjecture.

Many studies have aimed at probing the statistical structure of the GP relationship, thus relying on the computational exploration of GP maps. Models of RNA secondary structure [16,20], protein secondary structures [21,22], gene regulatory networks [3,23], metabolic networks [24,25], protein complexes [26,27], artificial life [28], or multilevel maps such as a toyLIFE, which includes protein structure, regulatory, and metabolic networks [29,30], have been explored through the years. Computational frameworks often rely on building complete GP maps from exhaustive enumeration of genotypes (or sparse GP maps from large samples) in models with simple genotype-to-phenotype rules as the ones above. To study global properties of a GP map, such as phenotype frequencies, a large number of genotype-phenotype pairs have to be evaluated. With notable exceptions [31–33], some of which will be discussed in section 7 of this paper, the exhaustive study of GP maps represents an enormous challenge that has been restricted to systems where the phenotype can be found computationally from the genotypic information.

For the sake of simplicity, most GP computational maps assign a unique phenotype to each genotype, in a many-to-one representation. Some maps also take into account environmental factors such as temperature, which modify GP mapping rules and, therefore, include phenotypic plasticity in a streamlined fashion [34]. Other implementations also consider phenotypic promiscuity [35,36], that is, the possibility that each sequence maps to more than one phenotype under fixed environmental variables. However, many-to-many GP maps entail an exponentially increasing cost in computation time, so they have been rarely explored in depth (for exceptions see [24,34,37–43]).

Creating complete computational frameworks for GP models is a challenge—building complete GP maps for sequences as long as functional molecules in realistic environments is beyond our current computational power. Nevertheless, progress has been steady and significant. For example, and despite the freedom inherent to any definition of phenotype, many generalities have emerged from studying these models, and theoretical arguments to explain some of them have been developed [44–46]. These studies have led to a relatively sound understanding of the conditions that are behind different phenotype abundances, its relationship with robustness, and the topology of neutral networks [12,47,48].

In this section, we begin by briefly summarising a variety of GP maps that have been computationally studied to date. Some attention is devoted to RNA, a model for which we examine in perspective some of the important lessons learnt and discuss possible future contributions to GP map research. There is a substantial body of literature available, including comprehensive reviews [49] that we do not even attempt to summarise here. We finish this part discussing the GP maps of artificial life systems.

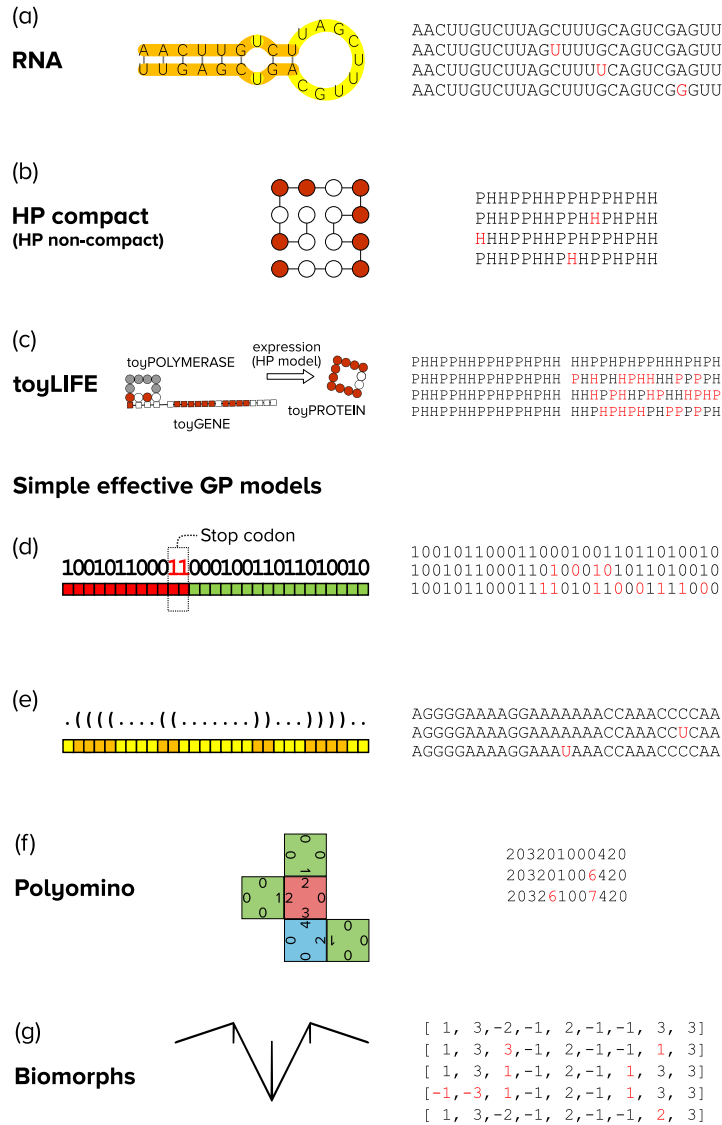


Fig. 1. Some examples of simple GP maps. For each model, and from left to right, we depict an example phenotype and some of the genotypes in its neutral network (mutations that do not change the phenotype are highlighted in red). (a) RNA sequence-to-structure is the paradigmatic GP map. Mutations that conserve the secondary structure appear in loops with a higher likelihood than in stems. (b) The HP model, both in compact or non-compact realisations, has been studied as a model for protein folding. (c) toyLIFE is a minimal model with several levels [29]. Sequences of the HP type are read and translated to proteins that interact through analogous HP rules to break down metabolites. (d) Fibonacci's model [44] relies on the separation between constrained and unconstrained positions in sequences to derive some formal properties of simple GP maps. (e) A generalisation of the idea of position-dependent constraints [45] provides a formal understanding of the ubiquitous lognormal distribution for neutral set sizes. (f) A polyomino model used to capture the essentials of quaternary protein structure [27]. (g) Dawkins' biomorphs are defined by genotypes with few parameters that define the generative rules of the structure [53]. Figure modified from Ref. [5].

While we focus here on sequence-to-structure and sequence-to-function maps, it is important to highlight that genotype-phenotype and genotype-fitness maps have also been studied in the context of development [50–52].

### 3.1. One-level GP models

Over the past three decades, the GP maps of several simple biological model systems have been studied in great detail. Fig. 1 summarises the essentials of some of the GP maps we will be discussing. Two classical examples are RNA secondary structure [16,20] and the HP model of protein folding [21,54]. The HP model represents proteins

on a regular lattice as self-avoiding chains of hydrophobic (H) or polar (P) beads. In its compact version the chains are forced to fold into rectangular configurations that leave no empty sites, while in the non-compact version all possible self-avoiding walks in the lattice are considered. The phenotype is defined as the minimum energy of a given configuration calculated from a contact potential between neighbouring (but not in the backbone) beads. Because RNA and HP models are relatively tractable, properties such as the distribution of the number of genotypes per phenotype [11,55,56], the phenotypic robustness and evolvability [57,58] (see Box 3.1) or the topological structure of neutral networks [48] could be systematically studied and compared [59].

Given the pivotal role proteins play in cellular processes, the protein sequence-to-structure map, of which the HP model constitutes the simplest realisation, is of great general interest [60–62]. The protein sequence-to-structure map has been also studied using more realistic, multi-parametric contact potentials [63–66] and coarse-grained models at different levels, such as the Polyomino model [26,27] for protein complexes. Some inferences about local and global properties of the protein sequence-to-structure-to-function GP map have also been made from experimental data [67–69], and estimates of neutral set sizes (NSSs) have been obtained from structural data [70].

Breakthroughs in the computational prediction of protein structure from amino acid sequence have recently been achieved by deep learning with artificial neural networks. The AlphaFold 2 system by DeepMind [71] outperformed 100 other teams in the 2020 Critical Assessment of Structure Prediction challenge (CASP13), with prediction accuracy rivalling that of experimental structure determination. Enormous resources are required for each evaluation, however, taking days of computational time for a single protein sequence. If subsequent development can maintain the accuracy while allowing exponential speedups, these computational systems should be able to open up entirely new investigations of the GP map for proteins.

A number of models work at levels above sequences. Simple gene regulatory networks act as effective genotypes in models that map them onto phenotypes defined as the steady-state gene expression pattern [75,76]. A metabolic genotype is defined as all chemical reactions an organism can catalyse via enzymes encoded in its genome; the phenotype is defined as viability in minimal chemical environments that differ in their sole carbon sources [77,78]. Those two models share the property that most genotypes do not map to any functional phenotype—it has been put forward that such a restrictive relationship may stem from a minimisation of the cost incurred by maintaining a complex functional network [79]. However, genotype spaces where function is sparse still contain large neutral networks that percolate that space and guarantee phenotypic innovation without loss of function [24,76,77].

There are compact [30] and non-compact [58] versions of the HP model with an overwhelming majority of non-functional genotypes where neutral networks are very small and mostly disconnected; therefore, innovation is severely hindered, if not plainly impossible, in those one-level maps. However, that lack of navigability turns out to be irrelevant if additional, higher levels, are taken into account.

### 3.2. Multi-level GP models

Most computational GP maps studied to date, including those discussed in the previous section, only include one level (or scale) of description, mapping genotypes of different kinds to their corresponding phenotypes (see, however, [80]). But even the simplest organisms include more than one level: RNAs and proteins will perform enzymatic and regulatory reactions that will in turn affect the availability of other molecules inside and outside the cell. If the study of one-level GP maps has led to great changes in our understanding of evolutionary theory, it stands to reason that studying multi-level GP maps will yield at least equally important insights.

It has been shown that multilevel models endowed with biophysically realistic interaction rules lead to the emergence of complex fitness landscapes that permit multiple, equally successful, evolutionary pathways [81,82] or the growth of organismal population size when protein-based, functional genotypes are discovered through evolution [83]. Recent proposals for multilevel models are the model of RNA-based virtual cells discussed in Section 6.2, a model of developmental spatial patterning [84,85] (see Section 5.5), and toyLIFE [29,30]. toyLIFE is a multi-level model that includes genes, proteins and metabolites, as well as their regulatory and metabolic interactions. toyGenes consist of binary sequences (the genotype) that are first mapped to HP-like proteins. None of these proteins can be obtained from any other through single-point mutations. Proteins interact between themselves, with the genome, and with metabolites. The phenotype is defined by the set of metabolites that a given sequence is able to catabolise. In its three-gene version, phenotype is mostly defined through the first two genes, which admit very few mutations, while

**Box 3.1a****Definitions**

*Function* Function is a contentious term [72,73] that is used to mean many things. In this review we are mostly referring to properties of proteins, such as stability, catalytic activity, and binding affinity.

*Genetic correlations* A GP map has this property if two sequences differing at a single site are more likely to generate the same phenotype than two arbitrary sequences [12].

*Genotype network* A set of mutually connected genotypes that have the same phenotype. This term is usually employed as a synonym of *neutral network*, although in some context a genotype network needs not be neutral—for instance, in the case of GP maps with both a categorical phenotype (e.g. molecular structure) as well as a quantitative fitness (e.g. thermodynamic stability of the structure).

*Genotypic evolvability* Total number of distinct alternative phenotypes that can be reached through point mutations from a single genotype [57].

*Genotypic robustness* Number of point mutations that do not change the phenotype of a specific genotype. It is analogous to the neutrality of a genotype.

*Navigability* Ability to navigate throughout genotype space via neutral mutations.

*Neutral network* A set of mutationally connected genotypes that have the same fitness, including those that have the same phenotype. Often, it refers to the largest connected component of a neutral set.

*Neutral set* A set of genotypes which have the same fitness, including those that have the same phenotype. The *neutral set size* is therefore the number of genotypes that map to a given phenotype.

*Organism* Any individual entity that embodies the properties of life, like a cell, an animal, or a plant. It is a synonym for “life form”. By extension, it also applies to artificial life forms.

*Phenotype* A property which is encoded in the genotype and is biologically relevant, for example a molecular structure. Though abstract, this broad definition allows a variety of models to be treated with the same terminology.

*Phenotypic robustness* Average genotypic robustness of all genotypes in a neutral network [57].

*Phenotypic evolvability* Total number of distinct alternative phenotypes that can be reached through point mutations from a phenotype’s neutral network [57].

*Plasticity* Quality of a genotype leading to the production of more than one phenotype depending on the environment [42].

*Promiscuity* Quality of a genotype leading to the production of more than one phenotype in the same environment.

*Quasispecies* Population structure with a large number of variant genomes related by mutations. Quasispecies typically arise under high mutation rates as possible mutants change in relative frequency as replication and selection proceed [74].

*Shape-space-covering* A GP map has the shape space covering property if, given a phenotype, only a small radius around a sequence encoding that phenotype needs to be explored in order to find the most common phenotypes [16].

*Versatility* A quantitative measure of the rescaled robustness of a specific sequence position [46].

**Box 3.1b**

## Acronyms

CPMs Cancer progression models  
 DAG Directed acyclic graph  
 FACS Fluorescence-activated cell sorting  
 FPGA Field-programmable gate array  
 GP Genotype-to-phenotype  
 MAVEs Multiplexed assays for variant effects  
 MFE Minimum free energy  
 MPRAs Massively parallel reporter assays  
 NSS Neutral set size  
 OLS Oligo(nucleotide) library synthesis  
 SCRaMbLE Synthetic Chromosome Recombination and Modification by LoxP-mediated Evolution

the third gene is essentially free to mutate, thus restoring evolvability to the system. Additionally, the existence of promiscuous sequences further enhances navigability when environmental factors such as temperature are considered [86]. Promiscuity was recognised long ago as a key property in adaptive processes [35] that, as of yet, has not been explored in most GP maps.

One of the most interesting results to come out of an early exploration of toyLIFE's metabolic GP map is that adding levels of complexity to a phenotypic definition actually increases robustness [30]: proteins can change and become non-functional, and regulatory functions can be altered, while the overall metabolic function remains constant. This suggests that the potential for cells to evolve toward new evolutionary challenges has been significantly underestimated in the past.

### 3.3. RNA

RNA is the most paradigmatic model for studying GP relationships and constructing GP maps [16,46,48,87–95]. Two major breakthroughs behind its popularity were the development of empirically based energy models—of which the most widespread is the Turner nearest neighbour energy model [96]—, and two fast dynamic programming algorithms to determine the minimum free energy (MFE) secondary structure [97] and to compute the partition function [98] of a sequence. In general, a sequence can fold into a number of secondary structures and the energy models and dynamic programming algorithms have made it possible to select low-energy structures [99], quantify their free energies [100] and use this to define a GP map in several ways: one GP map definition considers a single structure per sequence, usually the minimum-free-energy structure [16]. This will lead to a many-to-one GP map, where each sequence maps to a single structure, but each structure can be generated by a number of different sequences. An alternative definition allows several low-free-energy structures per sequence, which leads to a more complex many-to-many relationship. Together, these different studies defined a range of formal measures to quantify some of the key features of GP relationships, such as plasticity, evolvability, robustness and modularity [37]. The results obtained with RNA through the years have served as inspiration and guide to our intuition when faced with other GP maps.

#### 3.3.1. Phenotypic bias in RNA

We will start by reviewing results from the commonly studied many-to-one GP map, where the focus is solely on the predicted minimum-free-energy structure of each sequence. The largest exhaustive enumeration performed for RNA sequences, of length  $L = 20$ , yielded 10 orders of magnitude difference in the number of genotypes mapping from the most rare to the most frequent secondary structure phenotypes [94]. Approximate calculations of NSSs for longer sequences [46,95] show that this variance grows rapidly with increasing length. For example, for  $L = 100$  this difference is expected to be over 50 orders of magnitude: these maps are extremely biased. In an important study [92] the NSSs for longer length RNA were calculated using a sampling technique. When comparing to structures in the fRNAdb database for functional non-coding RNA (ncRNA) [101], they found, for systems of lengths  $L = 30$  to  $L = 50$ , that the natural RNA secondary structures were typically among those with larger NSS. These results suggested that the strong bias in the GP map was reflected in the secondary structures found in nature.

Another interesting set of studies compared structural features (e.g. distributions of stack and loop sizes) of natural secondary structures and those obtained when randomly sampling over sequences. They found that many are quite similar [87], and that natural and random RNA share strong similarities in the sequence nucleotide composition of secondary structure motifs such as stems, loops, and bulges [90]. Why should random sampling over sequences generate distributions that are so similar to natural RNA, where natural selection would normally be thought to play an important role?

The study of much larger datasets of natural RNA from the fRNAdb database—and for lengths ranging from  $L = 20$  to  $L = 126$ —demonstrated that the distributions of various structural features, and also properties such as the genotypic robustness, are very close to those obtained by random sampling over genotypes [95]. Furthermore, the distribution of NSS for natural RNA was found to closely follow the NSS distribution that arises upon random sampling of phenotypes. If one were to simply randomly sample over phenotypes, very significant differences with random genotype sampling (and natural RNA) would be found. By working out these counterfactuals it was therefore possible to demonstrate that the way in which variation arises through a GP map is dramatically different from the naive expectation that all potential variation is equally likely.

The close agreement of the distributions found in nature and those found by random sampling of genotypes via the GP map is very surprising given that natural selection is expected to be an important factor in the process that allows a particular functional RNA to fix in a population. The fact that its effect is not really visible for the properties above, at least when compared to a null model of random sampling genotypes, would appear to be strong evidence for the importance of anisotropic variation in determining evolutionary outcomes. However, before this conclusion can be drawn, it is important to remember that evolution does not proceed by random sampling of genotypes. Instead, it typically starts with a particular genotype and phenotype, and alters it via mutations that in turn generate new phenotypes that are either fixed or disappear over the generations in evolving populations. Given the hyper-astronomically large size of these spaces, it is not clear that such a local search should be at all similar to the results of random sampling of genotypes, which is a global property that does not depend on the starting point in genotype space.

Still, a counterexample of natural RNA where selection seems to have played a visible effect is that of viroids. Viroids are small, non-coding, circular RNA molecules that infect plants [102]. Viroids have compact secondary structures that constrain their evolution [103] and whose preservation seems essential to avoid degradation and inactivation [104], and to minimise the effect of deleterious mutations [105,106]. Viroids bear a number of paired nucleotides well above random expectations [107], such that the estimated NSSs of typical viroids are significantly below those of random sequences. For example, a typical structure for a circular RNA of length 399 has an average of 230 paired nucleotides and about  $10^{91}$  compatible sequences. However, the largest known viroid is *Chrysanthemum chlorotic mottle viroid*, which matches that length, but has 280 paired nucleotides and an NSS of about  $10^{72}$  genotypes [108].

### 3.3.2. Promiscuity in RNA

Beyond the many-to-one GP map, many-to-many GP maps that take into account the MFE structure and suboptimal structures in the Boltzmann ensemble have been studied [34,37,42]. Suboptimal structures can be included according to several criteria: either all structures which fall within a fixed free energy range from the MFE structure [34,37] are considered or only structures which have the same free energy as the MFE structure up to the energy resolution of the computational model [42]. First, a link was found between the suboptimal phenotypes of a sequence in the many-to-many GP map and the phenotypes in the mutational neighbourhood of the same sequence in the corresponding many-to-one GP map [37]. Secondly, genotypes with low promiscuity were shown to have MFE structures with higher modularity [37]. Finally, it was found that evolving populations encounter a higher number of phenotypes if suboptimal phenotypes are included [34]. Altogether, these observations point to the important adaptive role of molecular promiscuity by supplying alternative phenotypes in the absence of mutations, and so redefining the fitness landscape [5].

### 3.3.3. Hints from RNA inverse folding algorithms

The characterisation of functional phenotypes by designing sequences that fold into a given RNA secondary structure has been much less explored than the direct fold of given sequences. Finding sequences that yield a particular secondary structure is known as the RNA *inverse folding* problem. This is an NP-complete problem even for the MFE structure [109], hence a very demanding computational task. As a consequence, most approaches are based on local search algorithms [110]. Actually, RNA inverse folding algorithms are mostly intended for synthetic design,

though they have occasionally been used to investigate GP relationships [57,111]. However, their use is controversial due to the intrinsic bias of the underlying local search algorithms [112], which are not complete by definition and therefore produce biased samples under multiple runs. This caveat notwithstanding, there are some inverse folding methodologies that appear more suitable for this purpose.

The first method is a *soft inverse folding* approach which implements a dynamic programming algorithm to compute the RNA *dual partition function* [113]. This partition function is defined as the sum of Boltzmann factors  $\sum_{\sigma} \exp(-E(\sigma, \Sigma)/T)$ , where  $E(\sigma, \Sigma)$  is the energy of the RNA nucleotide sequence  $\sigma$  compatible with a target structure  $\Sigma$ , and  $T$  the absolute temperature (in units of energy). An energy weighted sampling from the low energy ensemble of sequences that are compatible with the given secondary structure is performed to calculate this partition function. While this approach is not particularly practical for synthetic design, it provides insights into molecular evolution.

This theoretical abstraction and the measures derived from it, such as the *expected dual energy*, can provide useful information about general properties of the phenotypes without exploring the whole genotype space. Computational analyses based on the nearest neighbour energy model over all the RNA sequences in the Rfam database [114] indicate that natural RNAs fold into secondary structures with energy higher than expected for sequences with the same length and GC content. Possible explanations for this observation are either that functional RNAs are not under evolutionary pressure to be highly thermodynamically stable or that sequence requirements prevent reaching minimum folding energies. On the other hand, experimental studies confirm that even random sequences frequently acquire compact folds similar to those of natural RNAs. Empirical observations further indicate that natural selection could be a determinant factor to achieve unique, stable tertiary folds—i.e. without major competing phenotypes—under natural conditions [88]. Besides, the controlled bias in this sampling methodology provides a delimited context to evaluate the properties that characterise a functional RNA with respect to sequences with similar structure. Simulations using this approach indicate that bacterial ncRNAs are more plastic and less robust than other sequences with similar structure [113].

Although the samples returned by this algorithm are representative of the low energy ensemble of sequences of the given structure, the MFE structure of individual sequences is not necessarily the target structure. However, the proportion of alternative MFE structures of the sampled sequences is the distribution of *competing phenotypes* in the low energy ensemble of the target structure, which can in turn be interpreted as an estimate of the structures that are likely to coexist with that phenotype in a many-to-many GP map.

Similar algorithms for computing and sampling from the RNA *dual partition function* with additional constraints have been developed and used to determine the neutral path between sequences in the same phenotype [115].

The second methodology is complete inverse folding based on constraint programming [116]. The constraint programming paradigm avoids exploring the whole sequence space when structural, sequence or environmental restrictions are included. These restrictions comprise, among many others, GC content, sequence motifs, multiple local and global structures and folding temperatures. Rather than slowing down the search, each constraint increases the speed of this algorithm. This algorithm can potentially retrieve all sequences that meet the requirements or conclude that no solution exists. In practice, the running time depends on the sequence space defined by the given constraints. These features make it appropriate for the study of genotype-phenotype-function relationships of moderately small functional RNAs with known moieties, or of regulatory RNA elements like riboswitches and thermoswitches.

Some examples of the performance of complete inverse folding based on constraint programming are the computationally-based suggestion that the conserved GUH (no G) motif in the hammerhead ribozyme type III cleavage site of *Peach latent mosaic viroid* is due to structural, rather than functional, requirements [117], or that natural thermoswitches do not seem to be optimised to maximise the probability difference between the active and inactive structures at the corresponding folding temperatures [118].

### 3.4. Artificial life

Evolutionary processes have not only been studied in biology, but also in man-made systems. Some models were designed to simulate biological evolution computationally and mimic biological properties. A widely used example is the digital model of a biological organism called Avida [28]. Avida organisms are pieces of code which can self-replicate and evolve towards optimal usage of computational resources. Richard Dawkins introduced a different form of artificial life to study evolution: biomorphs [53] are two-dimensional stick figures produced recursively from a

genotype, which consists of nine integer numbers. These biomorphs resemble abstract animal or plant shapes. Lindenmayer systems are another famous recursive model which can produce plant-like figures [119,120]. These model systems are abstractions of biological organisms, but they all imitate properties of biological systems: the recursive branching rules in Lindenmayer's systems and later in Dawkins' biomorphs were inspired by plant development, whereas Avida digital organisms have a metabolism and compete, just like bacteria [28,53,120]. However, evolutionary principles have been applied even more generally: the study of programmable electronic hardware has been addressed using the GP framework [121]. Circuit configurations were treated as genotypes and the function which a circuit computes as the corresponding phenotype.

Here we will focus on results for four artificial life models: Avida organisms [122], biomorphs [53,123], the 2PDOL model [124,125], which is based on Lindenmayer's systems, and FPGAs [121], a type of programmable electronic circuits. These studies have focused on different properties, which makes a direct and quantitative comparison difficult. However, similarities between these artificial life GP maps and molecular sequence-to-structure GP maps exist [121,122,124,125]: first, in three of these four systems the number of genotypes mapping to a given phenotype was estimated and found to vary significantly between phenotypes [121–123]. For the fourth model a related quantity, the neutral set diameter, was also found to differ between phenotypes [125]. Such a heterogeneity, or phenotypic bias, in the distribution of genotypes over phenotypes has long been observed in molecular structure GP maps [16,54]. Second, a high degree of genotypic robustness was observed, which enables the formation of neutral networks [121–123,125]. This property was also first found in molecular structure GP maps [21] and is referred to as genetic correlations [12]. A third shared property follows from the vastly different NSS: the probability of transitioning from a larger to a chosen smaller neutral set by point mutations is much smaller than that in the reverse direction. This asymmetry is known from molecular structure GP maps [126] and has been confirmed for two of the artificial life GP maps: Avida [122] and the 2PDOL model [124].

In addition to these shared properties, there are points in which the various artificial life systems differ. In Avida, a high fraction of genotypes is considered inviable because the organisms are unable to reproduce [122], whereas in the biomorphs system all genotypes produce well-defined drawings and all stick figures are viable until an external decision is made about the fitness of specific shapes. In molecular GP maps the fraction of viable genotypes also depends on the system: in studies of model proteins, a large fraction of genotypes does not fold into a unique structure and is considered unstable, whereas for RNA secondary structure a minimum free energy structure is found for most sequences [59]. Further comparisons could be made once quantities defined for GP maps, such as phenotypic robustness and evolvability, NSS, and mean-field mutation probabilities, are evaluated consistently for all of these and further artificial life models. Commonalities between artificial life and molecular structure GP maps dominate the picture at present, but future research may also identify differences between these two groups of models.

#### 4. The universal topology of genotype spaces

Some of the results highlighted in the former section hint at the possibility that any sensible GP map (and, by extension, artificial life system) is characterised by a generic set of structural properties that appear repeatedly, with small quantitative variations, regardless the specifics of each map. Extensive research performed in recent years has confirmed this possibility to an unexpected degree.

Some of the commonalities documented are navigability, as reflected in the ubiquitous existence of large neutral networks for common phenotypes that span the whole space of genotypes, a negative correlation between genotypic evolvability and genotypic robustness, a positive correlation between phenotypic evolvability and phenotypic robustness, a linear growth of phenotypic robustness with the logarithm of the NSS, or a near lognormal distribution of the latter. There are recent and comprehensive reviews of the properties measured and shared by different GP maps [1–6]. In the following sections, we discuss new views on the plausible roots of this seemingly universal class of GP maps.

##### 4.1. Possible roots of universality in GP maps

The question obviously arises: Why are structural properties of GP maps unaltered by the details of the mapping? Part of the answer must lie in the topology of the very high dimensional spaces governing the relationship between genotypes and phenotypes. Our intuitions often fail us here because these spaces are highly interconnected. Although their volumes grow exponentially with sequence length, distances are linear. For example, if one could make every

RNA of length  $L = 79$ , the molecules would weigh more than the Earth [11]. Yet none of those strands is more than 79 point mutations away from any other.

One way this interconnection manifests itself is through the property of shape-space covering, a term first introduced for GP maps in the RNA context [16], and borrowed from its original use in immunology [127]. It captures the fact that many phenotypes are only a handful of mutations away from one another. While this property has been best studied in the secondary structure RNA GP map, it has also been shown to be present in the HP model [59,128], toyLIFE [30], the polyominoes [27], and a model of gene expression [84] (where it is described as ergodicity of phenotypic exploration). Shape-space covering suggests that no matter where you start, many other phenotypes are in principle close by in terms of Hamming distance. In the cases above, this holds even if the search begins in an arbitrary genotype. In GP maps where function is sparse in genotype space [24,76,77], phenotypes are still close to each other, but links are established through a limited number of genotypes that might take a long time to find through random walks on the neutral network.

#### 4.2. Constrained and unconstrained sequence positions. Formalising neutrality and evolvability

The intuitions above have received quantitative support from analytically tractable, streamlined GP maps which aim to capture the essentials of generic GP map features. These models, the results attained and the clues they provide are summarised in this section, which might appear slightly technical to the non-familiar reader but clarifies possible constructive principles of evolutionarily apt GP maps.

Highly simplified, abstract GP maps can reproduce many of the generic properties discussed [44,129]. These simplified maps hint at two major possible causes underlying structural universality: (i) the partition of sequence regions into constrained and unconstrained parts and (ii) non-local interdependence of sequence positions with regard to their constraints (as sketched in Fig. 1 (d, e)). Let us illustrate how to derive general results with a simple example. Consider a sequence of length  $L$  whose first  $\ell$  positions are fully constrained (changing any of those positions amounts to changing the phenotype) and the remaining  $L - \ell$  positions are neutral (changes do not affect phenotype). If every position in the sequence admits  $k$  possible values ( $k = 2$  for a binary alphabet, formed for example by symbols  $\{0, 1\}$  and  $k = 4$  in quaternary alphabets, as  $\{A, C, G, U\}$ ), then there are  $k^\ell$  different phenotypes for every value of  $\ell$ , each of size  $k^{L-\ell}$ . Using a rank-ordering of the sizes of phenotypes and some simple algebra, it is easy to conclude that the probability  $p(S)$  that a phenotype has size  $S$  is  $p(S) \propto S^{-\alpha}$ , with  $\alpha = 2$  [45]. If the restriction of a position being fully constrained or neutral is relaxed, different values of the exponent  $\alpha$  can be obtained. The exponent also changes if a stop codon (equivalent to considering an  $\ell$ -dependent amount of lethal mutations) is introduced [44].

Interestingly, the shape of the distribution  $p(S)$  changes to a lognormal function if, in the examples above, constrained sites can be arbitrarily distributed along the sequence. In general, the positions of a sequence are neither constrained nor neutral, but versatile in varying degrees. Let us define the *versatility*  $v_i$  of position  $i$  in a sequence as the average number of alphabet letters at that site that do not modify the phenotype. This extends the ideas above and provides a simple estimation of neutral set size  $S$ , as  $S = v_1 v_2 v_3 \dots v_L$ . This estimated value has been shown to be a very good approximation to the NSS in several GP models such as RNA, HP and toyLIFE [46] (Fig. 2a). In all those cases and several others, the distribution  $p(S)$  is compatible with a lognormal, which can be analytically derived under very generic assumptions in the case of RNA [45] (Fig. 2b). Moreover, the results suggest that this approximation can be extrapolated to larger sizes. Additional properties, such as genotypic and phenotypic robustness, can be analytically obtained in such effective models [44], which constitute a sound first step towards deriving a formal theory of genotype spaces and their universal properties.

Generic biological sequences display the characteristics above in almost every biological context: exons and introns correspond to constrained and unconstrained regions, as do genes and noncoding intergenic sequences. Start and stop codons as well as interactions between transcription factors and their targets are examples of the interdependence of one sequence region on the constraint of another. As a result it is likely that the same GP map properties we observe in abstract model systems also hold for much more complex and biologically realistic phenotypes. The challenge in these more complex GP maps, however, is the vast size of the genotype space. A protein of 300 residues has a sequence space of size  $20^{300}$ . Approaches that can estimate the structural properties of a GP map from relatively small samples are therefore essential. Knowledge of these properties is not just interesting for the study of GP maps, but also has potentially useful applications [95]. Being able to measure properties such as the phenotypic robustness, evolvability, and neutral network size of phenotypes in more complex GP maps would therefore provide a powerful methodological

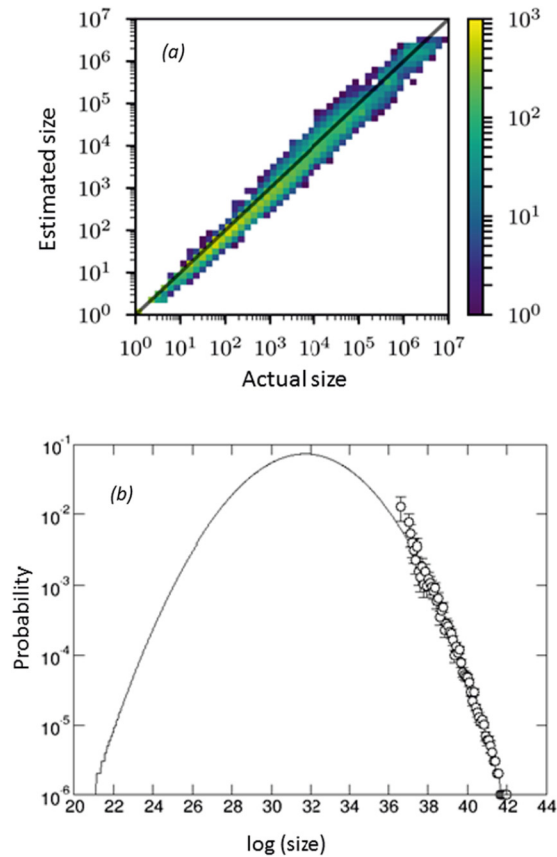


Fig. 2. Predicted and measured properties of RNA phenotypes. (a) Log-log-log histogram of the estimated abundance vs. actual abundance of four-letter RNA of length  $L = 16$  phenotypes using versatility as defined in the main text [46]; (b) NSSs estimated using [92] natural RNAs of length  $L = 100$  obtained from the ncRNA database [95]. Random evolutionary search is highly skewed towards the largest phenotypes, as evidenced by the predicted shape of the full, lognormal distribution (solid curve): phenotypes of small and typical sizes are not found in nature.

tool for the prediction of evolutionary pathways. The division of sequences into constrained and unconstrained regions is also likely to make prediction of structural GP map properties from local samples easier. This is because a division of sequences into constrained and unconstrained regions implies that many sequence positions are largely independent of each other with regard to their phenotypic effect. While important interdependencies remain, which particularly affect evolvability, the fact that interdependent sequence positions are likely to constitute a relatively small fraction of the total sequence means that a sampling approach is feasible for the purpose of estimating neutral network sizes and phenotypic robustness.

## 5. Evolutionary dynamics on genotype spaces

In the previous sections we have discussed the static properties of genotype spaces, their plausible universality and some basic principles that may underlie their topology. Such findings are relevant by themselves, but a further aim is to uncover the consequences of genotype space architecture in evolutionary dynamics. Evolution can be pictured as the navigation on the space of all possible genotypes [13], and GP maps describe the way different phenotypes are organised in such a space [130]. This organisation and the intrinsic structure of GP maps affects, among others, the ability to find genotypes and phenotypes in evolutionary searches [91,94], as well as the rate of adaptation [131,132].

Early studies of dynamics on neutral networks quantified the trend of populations to maximise genotypic robustness by demonstrating that mutation-selection equilibrium is solely determined by the network topology [133]. Still, the time to reach equilibrium is an inverse function of the mutation rate [47]. Neutral networks in GP maps, as well as in a few instances where this property could be quantified, are assortative [48]: the neutrality of genotypes one mutation

away from each other is positively correlated. As a result, the dynamics is naturally canalised towards maximally connected regions [37], resulting in an acceleration in the rate of accumulation of neutral mutations with time [132].

In more recent analyses, attention has turned towards the effect of phenotypic bias in adaptation, as we have already discussed by means of enlightening studies with RNA. The question has been also investigated using a modified version of toyLIFE to model pattern-formation in regulatory networks [86,134] aimed at finding out how evolution chooses between two *a priori* equally fit phenotypes. It turns out that evolutionary dynamics at the phenotypic level cannot be well described by a Markovian process between phenotypes [132], because of the nontrivial topology of each phenotype's neutral network [5]. As a matter of fact, the escape time from one phenotype does not follow an exponential distribution, as most evolutionary models assume. This is one instance of the so-called phenotypic entrapment [132], in which the trend of populations to become trapped in increasingly robust regions of a phenotype neutral network results in a long-tailed distribution of escape times: either the population escapes very fast, or takes a very long time to do it.

Accounts of evolution on neutral networks driven by point mutations and the corresponding mathematical formalism can be found elsewhere [5,135,136], though some essentials will be also described here. In this section we will mainly discuss the effects of a largely disregarded but essential evolutionary mechanism (recombination) and how mutational bias affects isotropic searches.

We continue with evolutionary dynamics on genotype and phenotype networks defined by point mutations, where if we make an ergodic assumption that all typical phenotypes are locally accessible we are led in a natural manner to the formulation of the statistical mechanics of phenotypic evolution. We close by discussing a number of applications where these ergodic assumptions are most appropriate.

The approaches in this section differ in the formalism used (complex networks at large, mean-field effective models and statistical mechanics) but all converge in the main emerging lesson: the size of a phenotype plays a role in evolution comparable to that of fitness. Quantification of their relative weight through formal approaches might eventually settle the false dichotomy between neutralism and adaptationism.

### 5.1. Robustness and recombination

Genotypic robustness is a property of the GP map that quantifies to what extent functional genotypes can be maintained in the presence of random mutations [89,137–139]. Specifically, consider a genotype encoded by a sequence  $\sigma$  of length  $L$  that admits a total of  $(k-1)L$  single point mutations (recall that  $k$  is the size of the alphabet). Genotypes are classified to be either viable (functional) or lethal (non-functional). Then the genotypic robustness  $r_\sigma$  of a viable genotype is defined as the fraction of mutations that maintain viability [140],  $r_\sigma = n_\sigma / (k-1)L$ , where  $n_\sigma$  is the number of viable mutational neighbours. The population-averaged robustness is correspondingly defined as

$$r = \sum_{\sigma \in V} r_\sigma v_\sigma^*, \quad (1)$$

where  $V$  denotes the set of viable genotypes and  $v_\sigma^*$  is the stationary frequency of genotype  $\sigma$ . Two limiting cases are of particular interest. If the product of population size  $N$  and mutation rate  $U$  per individual and generation is small,  $NU \ll 1$ , the population is monomorphic and performs a random walk on the network of viable states. The stationary frequency distribution is then uniform and (1) reduces to

$$r_0 = |V|^{-1} \sum_{\sigma \in V} r_\sigma, \quad (2)$$

where  $|V|$  is the number of viable genotypes. On the other hand, when  $NU \gg 1$ , the stationary frequency distribution is determined by mutation-selection balance and can be shown to be given by the leading eigenvector of the adjacency matrix of the network of viable genotypes [128,133], see also Section 5.3. The population robustness  $r$  is related to the corresponding eigenvalue and exceeds the uniform robustness  $r_0$  whenever the network is inhomogeneous. This implies that selection in large populations increases robustness by focusing the population in highly connected regions of the network.

Numerical studies of recombining populations on various types of genotype networks have indicated that recombination enhances the focusing effect of selection and thus substantially increases genotypic robustness [141–146].

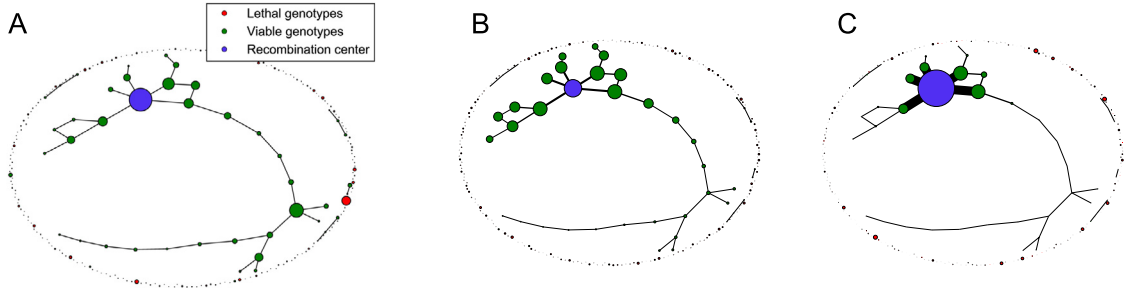


Fig. 3. Genotype network generated by assigning viable genotypes at random with probability  $p = 0.2$  to binary sequences of length  $L = 8$ . The largest connected component of viable genotypes is shown in the centre of each panel, and smaller components and isolated nodes are arranged in a ring surrounding the central component. (a) Network structure visualised by the recombination weight  $\omega_\sigma$ . Node areas are proportional to  $\omega_\sigma^6$  and the recombination centre is marked in purple. (b) Stationary frequency distribution of a non-recombining population. Node areas are proportional to the stationary frequency  $v_\sigma^*$  of the respective genotype, and the edge width between neighbouring genotypes  $\sigma, \tau$  is proportional to  $\max[v_\sigma^*, v_\tau^*]$ . (c) Same as (b) for a recombining population. Note that the population is much more strongly concentrated on the recombination centre than in panel (b). In panels (b) and (c) the mutation rate per site is  $\mu = U/L = 0.001$ . Courtesy of Alexander Klug.

Recently, a systematic and largely analytic investigation of the relationship between recombination and genotypic robustness within the framework of deterministic mutation-selection-recombination models has been presented [147].

As a simple but informative example, consider the space of binary sequences  $\{0, 1\}^L$  endowed with a ‘mesa’ landscape where genotypes carrying up to  $\eta$  1’s are viable and all others are lethal [148]. The genotypes on the brink of the mesa carry exactly  $\eta$  mutations and have robustness  $r_\sigma = \eta/L$ , whereas all others have robustness  $r_\sigma = 1$ . Combinatorial considerations show that the uniform robustness  $r_0 \approx 2\eta/L$  for large  $L$  and  $\eta < L/2$ , reflecting the fact that a large fraction of genotypes is located at the brink for purely entropic reasons. The maximal robustness that can be achieved through selection alone is [148]

$$r \approx 2\sqrt{\frac{\eta}{L} \left(1 - \frac{\eta}{L}\right)} \text{ for } \eta < L/2, \quad (3)$$

which exceeds  $r_0$  but is small compared to unity when  $\eta \ll L$ . Thus selection only partly counteracts the entropic outward pressure and as a consequence a large part of the population is still located near the brink under mutation-selection balance. By contrast, in the presence of recombination  $r \rightarrow 1$  for small mutation rates, because the contracting property of recombination efficiently transfers the population to the interior of the mesa where all genotypes are surrounded by viable mutants.

Simulations on different types of random genotype networks show that the massive enhancement of robustness found for the mesa landscape is generic, and typically a recombination rate on the order of the mutation rate suffices to achieve this effect. It is not obvious that the focusing of the population towards the centre of its genotypic range by recombination should generally increase robustness in this case, because viable and lethal genotypes are randomly interspersed in the network. To rationalise the observed increase in robustness it is useful to quantify the likelihood of a genotype  $\sigma$  to be created by recombination through its *recombination weight*  $\omega_\sigma$  defined by

$$\omega_\sigma = \frac{1}{|G|} \sum_{\kappa \in V, \tau \in V} R_{\sigma|\kappa\tau}. \quad (4)$$

Here  $R_{\sigma|\kappa\tau}$  denotes the probability that  $\sigma$  is generated by crossover from  $\kappa$  and  $\tau$  and  $|G|$  is the total number of genotypes. The normalisation ensures that  $\omega_\sigma \in [0, 1]$ , and the recombination weights sum to  $\sum_\sigma \omega_\sigma = |V|^2/|G|$ . The genotype that maximises  $\omega_\sigma$  is called the *recombination centre* of the network and provides a good predictor for the point of concentration of the recombining population in the limit  $U \rightarrow 0$  (see Fig. 3 for an example). Moreover, for two classes of random, percolation-type genotype networks and one empirical fitness landscape, the recombination weight  $\omega_\sigma$  was found to be positively correlated with the genotypic robustness  $r_\sigma$ .

If this correlation were a generic feature of GP maps, it would constitute a mechanistic explanation for how recombination acts to enhance genotypic robustness. Future work should therefore elucidate the conditions on the topology of the genotype network required for such a correlation to be present. It is not difficult to construct counterexamples where the recombination centre has low robustness, e.g., by placing a hole of lethal genotypes at the centre of a mesa

landscape. Only the investigation of specific, biophysically motivated GP maps such as RNA secondary structures or lattice proteins will clarify whether or not such instances are statistically relevant. More broadly, it appears that a common perspective on recombination, robustness and evolvability [89,138,139] may help to develop and test novel hypotheses about the evolutionary origins of these important biological phenomena.

## 5.2. Mutation bias

Some regions of genotype space exhibit biases in the mutations they contain. For instance, GC-rich regions have more  $G \leftrightarrow C$  transversion (purine-to-pyrimidine or pyrimidine-to-purine) mutations than transitions (pyrimidine-to-pyrimidine or purine-to-purine mutations). This may interact with biases in the generation of genetic variation, because some mutations occur more frequently than others. For instance, the rate of  $A \leftrightarrow G$  transitions exceeds the rate of  $T \leftrightarrow C$  transitions in transcribed human genes, whereas there is no significant difference in non-transcribed regions [149]. Furthermore, CpG dinucleotides—regions of DNA where C follows G—are considered “hot spots” for  $G \rightarrow A$  and  $C \rightarrow T$  transition mutations [150]. Other forms of mutation bias such as deletion bias and strand-specific bias have been reported in bacterial genomes [151,152].

Under certain population genetics conditions, mutation bias can be an orienting factor in adaptive evolution [153,154], and several experimental evolution studies indicate that mutation bias can influence trajectories of adaptive protein evolution [155,156]. It is possible to get a better understanding of how such mutation biases affect the outcomes and mutational trajectories of adaptive evolution by studying their impact on the navigability of GP maps.

Instead of the classic depiction of a GP map in which all the possible mutations are equally likely to occur, one could consider regions of the genotype space being differentially prone to distinct kinds of mutations. Ultimately this would affect the probability of traversing different edges in the genotype network and, therefore, its navigability. In this context, a *mutation bias weight* could be formally defined and introduced into a more general formulation of genotype networks, by biasing the accessibility of different genotypes. Understanding the potential evolutionary implications of mutation biased GP maps could provide us with valuable information about the nature of the systems they represent. For example, if a bias towards certain kinds of mutations enhances the ability to find the adaptive peaks of a certain GP map, a testable prediction could be that adaptive genotypes are more likely to evolve in regions of the genome that are prone to that particular kind of mutation.

Moreover, integrating mutation bias into the study of GP maps can change properties such as robustness and evolvability [157,158]. Both robustness and evolvability are based on the structure of genotypic neighbourhoods, and this structure can change if mutation bias is considered. For instance, a genotype might seem highly robust when most of its neighbours in the genotype space map onto the same phenotype. However, if there is a sufficiently high mutation bias towards mutations that do not preserve that phenotype, robustness would be diminished. The same principle can apply to evolvability.

## 5.3. Phenotypic transitions as competitions between networks

Formal studies of the way the structure and navigability of GP maps affects evolutionary dynamics can provide insights into the mechanisms underlying adaptive evolution, robustness and the emergence of phenotypic innovations. In the previous two sections, it has been shown that links between genotypes in a genotype network are weighted: microscopic mechanisms such as recombination and mutation bias modify the likelihood of transitions between pairs of genotypes. Constant link weights of a generic transition matrix  $\mathbf{M}$  correctly describe mutation bias, but cannot account for the effects of recombination, since in the latter case they depend on the abundances of each genotype  $v_\sigma$  in a nonlinear way, and in general are a time-dependent quantity. The simultaneous consideration of point mutations and recombination in a network framework remains as a topic for future studies.

In the following, we summarise a mutation-selection evolutionary process on a network of genotypes subject only to point mutations using tools from complex network theory. Consider a vector  $\vec{n}(t)$  whose components are the population of individuals at each node at time  $t$  (upon normalisation, each component  $n_\sigma(t)$  is the frequency of the genotype  $v_\sigma(t)$ ). Then,

$$\vec{n}(t+1) = \mathbf{M}\vec{n}(t) \quad (5)$$

**Box 5.3**

## Genotype spaces as networks of networks

Populations evolve in steadily changing environments where the impact of internal and external perturbations can rarely be considered in full. Often, nonlinear responses to small external changes hinder predictability, as weak perturbations might trigger critical transitions that strongly influence the fate of whole ecosystems [159,160]. Complex network theory and the tools associated to it offer a powerful framework to tackle this type of dynamical systems, since a multitude of natural systems can be modelled as nodes (agents) connected by links (interactions).

While network science has largely focused on single networks, in the last decade the study of dynamical properties on networks of networks or, in a more general way, on multilayer networks [161], has attracted wide attention [162,163]. One important motivation has been the finding that robustness, synchronisation or cooperation lead to different behaviour when studied in isolated or in interconnected networks [164–167]. However, the main reason for this change of perspective has been to realise that many natural systems, beyond displaying a network-like organisation, are also made of interacting and competing networks at very different scales, from the molecular level to supranational organisations [168].

The extent to which network science can foster our knowledge and comprehension of the evolution and adaptation of heterogeneous populations in an ever changing biosphere is a relevant open question. In particular, the *theory of competing networks* can be used to analyse the evolutionary dynamics of populations in a space of genotypes that can be regarded as a network of networks [169]. From this viewpoint, population evolution is described as a competition for resources of a certain kind, where the competitors are whole networks instead of independent nodes [170].

represents the dynamics of the population, where  $\mathbf{M}$  is a transition matrix with information on the fitness of each genotype, on the mutation and replication process, and on the weighted topology of the network.  $\vec{n}(t)$  describes the distribution, at each time  $t$ , of the population of sequences on the space of genotypes. As already stated, mutation-selection equilibrium is independent of the initial state and given by the eigenvector  $\vec{u}_1$  associated to the largest eigenvalue  $\lambda_1$  of  $\mathbf{M}$ . Furthermore,  $\lambda_1$  yields the growth rate of the population at equilibrium, and  $\vec{u}_1$  is also a measure (known as eigenvector centrality) of the topological importance of a node in a network [171].

In the context of the theory of competing networks, any dynamics that takes place on networks interconnected through a limited number of links (networks of networks), can be often characterised as a competition where the contenders are whole networks, and where eigenvector centrality represents the resource that the agents compete for (see Box 5.3). The final outcome of such a struggle for centrality strongly depends on the internal structure of the competing networks and on the links connecting them [170].

On the other hand, it has been shown [172] that even when environmental perturbations are weak, populations may suffer critical transitions in their genomic composition when the fraction of lethal mutations (i.e. of zero-fitness genotypes) is sufficiently high—of the order of that observed in natural populations [173]. A recent analysis of these results suggested that the space of genotypes can be regarded as a network of networks in “competition” to attract population [169], and that knowledge of the topology of the space of genotypes entails a certain predictive capability of the future evolutionary dynamics of the population under study. In fitness landscapes with a large fraction of lethal genotypes (as it could be the case of the non-compact HP model, GP maps for gene regulatory networks, or models for metabolism), the space of genotypes is formed by many subnetworks connected through narrow adaptive pathways. This topology induces drastic transitions of population from one subnetwork to another, occasionally causing the extinction of the population. The key topological element underlying sudden genomic shifts is the high heterogeneity in the network describing and linking viable genotypes. This topology can arise under a significant fraction of lethal mutations (or non-viable genotypes), but the same phenomenon is observed in rugged fitness landscapes.

### 5.3.1. An empirical test of the theory: transition forecast

It is highly likely that large molecular populations able to evolve fast, such as RNA viruses, can provide an empirical test of this predicted critical behaviour. The enormous advances of high-throughput sequencing allow for a very precise description of the populations at the molecular level, and in particular of the abundances of the coexisting genotypes. This information might be used to build the space of sequences associated to a population that evolves in

a changing environment, and thus a proxy of the network of genotypes where the population evolves. Applying the theory of competing networks it is conceivable that the eigenvalues of the different subnetworks and the centrality of the connector nodes would provide valuable information on how environmental variability affects the sharpness of the transitions and on the chances that the population could survive. The combination of tools from complex networks theory and the last decades' research on state shifts in the biosphere [174,175] might eventually lead to a prediction of the time left until the transition occurs. This prediction is important because, once a tipping point takes place, it becomes very difficult, if not impossible, to return to the previous state. At present, a wide variety of early warning signals for state shifts has already been characterised, but none of them yields precise information about the time left before the tipping point is reached [176,177]. However, calculations of the minimal distance between the first and second eigenvalues associated to the transition matrix  $\mathbf{M}$  could be used to obtain a first estimation of the time to the transition [172]. In an evolving population, the relative abundances of the different genotypes could be used as an approximation of the eigenvector  $\tilde{u}_1$ ; a measure of the growth rate of the population at equilibrium could yield the largest eigenvalue  $\lambda_1$ , and  $\lambda_2$  might be estimated by quantifying how resilient the population is to external perturbations [178]. A sufficiently precise measurement of these quantities would represent a very fruitful connection between actual evolving populations and a dynamical description of possible sudden evolutionary transitions. On a related note, regarding the space of genotypes as a network of networks entails a more coarse-grained, effective model where each genotype network can be considered as a single node, and where the dynamics can be simplified to account only for changes in the phenotype. Links in this higher-level description would have a weight proportional to the within- and between-phenotypes links. At odds with the description at the genotype level though, transitions between phenotypes are no longer symmetrical [91,179], nor is the dynamics describing these transitions Markovian any more [132,180].

#### 5.4. A mean-field description of phenotype networks

The qualitative properties of a high-dimensional evolutionary search are inherent to navigable GP maps and very likely responsible for some of the generic features described in Section 4. Despite all caveats that the complex dynamics at the genotype level may raise due to its non-Markovian nature [132,180], the high dimensionality of genotype spaces helps us understand why a simple mean field model [94], which averages over much of the local structure of a neutral set, succeeds in capturing some of those generic, dynamical properties. The model works with  $\phi_{\xi\chi}$ , the probability that a point mutation for genotypes that map to phenotype  $\xi$  generates a genotype for phenotype  $\chi$ , averaged over all genotypes that generate  $\xi$ . By measuring the  $\phi_{\xi\chi}$ , a weighted network between all the phenotypes can be defined, with  $\phi_{\xi\chi}$  as the weights. This allows for a much simpler dynamics that ignores the individual genotypes, and so analytic results can be derived for many properties in dynamical regimes ranging from the monomorphic to the fully polymorphic limits. Interestingly, for RNA, as well as for a number of other GP maps [12], it was found to a good first approximation that if  $\xi \neq \chi$  then

$$\phi_{\xi\chi} \approx f_{\chi}, \quad (6)$$

where  $f_{\chi}$  is the global frequency of phenotype  $\chi$ , i.e., the fraction of genotypes that map to  $\chi$ . Since the  $f_{\chi}$  range over many orders of magnitude, so do the  $\phi_{\xi\chi}$ . In contrast to the case where  $\xi = \chi$ , the robustness of phenotype  $\xi$  is  $\phi_{\xi\xi} \propto \log(f_{\xi})$ , and so varies much less with NSS. This property of the robustness is critical for neutral exploration. The mean field model predicts that for many different starting phenotypes  $\xi$ , the probability that a different phenotype  $\chi$  will appear as potential variation will scale as  $f_{\chi}$ .

For several GP maps, this simplified model does an excellent job at predicting the rates at which variation arises in full GP map simulations. Since NSS, or equivalently  $f_{\chi}$ , varies over many orders of magnitude, this argument predicts that, to first order, the rate at which variation arises will also vary over many orders of magnitude. Therefore, even though the set of physically possible variations may be very large, only a tiny fraction of the most frequent phenotypes will ever be presented to natural selection. This *arrival of the frequent* effect [94] is therefore very strong. Fundamentally it is a non-steady-state effect, since the longer an evolutionary run proceeds, the more the potential variation with lower  $f_{\chi}$  becomes likely to appear. The arrival of the frequent differs from the *survival of the flattest*, [181] which describes the situation where a fitness peak with lower fitness can nevertheless dominate over a higher fitness peak with a lower NSS. The latter effect can be analysed in a steady-state framework, whereas the former effect cannot.

Let us return in this context to the question of why so many structural features, as well as the genotypic robustness of RNA secondary structures, are so accurately predicted by a null model that ignores selection entirely. The arguments above suggest that even in the more complex situation of RNA evolution in nature, variation will nevertheless to a good first approximation arise with a probability proportional to its NSS. Since this rate varies by so many orders of magnitude, this arrival of the frequent effect determines what natural selection can work with, and so tends to dominate over local fitness effects. Rare phenotypes will almost have no bearing on evolutionary dynamics: they will hardly be found by a population searching for an adaptive solution and, if they are found, they will be quickly lost due to mutations. This is akin to an entropic effect in statistical physics: dynamics tend to favour macrostates with a larger set of microstates.

In other words, natural selection can only act on variation that has been pre-sculpted by the GP map. For the case of RNA described in Section 3.3.1, it appears that it mainly works by further refining parts of the sequence. This picture of the primacy of variation stands in sharp contrast to more traditional arguments about the importance of natural selection as an ultimate explanation of any evolutionary trend. It also raises many open questions. Are there other GP maps for which we can see such dramatic effects in nature? There are certainly conditions where this primacy of variation is incorrect. But how, and when does this GP map based picture of pre-sculpted variation break down? The exceptional case of viroids, discussed in Section 3.3 might be one such example, and provide clues to seek answers to the latter question.

### 5.5. Equilibrium properties and statistical mechanical analogies in the weak mutation regime

The broad question of optimisation in evolution, such as the existence of a Lyapunov function, describing a general dynamics and approach to equilibrium was first addressed by Iwasa [182] in his definition of “free fitness”, in analogy to the free energy of statistical mechanics, and then later rediscovered for the particular case of the weak mutation regime [158] and in the context of the evolution of quantitative traits [183,184]. The key insight, is that, at finite population size, not fitness itself but a combination of fitness and Shannon entropy (weighted by  $1/N_e$ , where  $N_e$  is the effective population size) is optimised over the evolutionary degrees of freedom of interest. This perception is consistent with the mean-field description reviewed in the previous section, where it has been shown that phenotypic bias is at least as relevant as phenotype fitness in evolutionary dynamics.

From a statistical mechanics viewpoint, and in the weak mutation regime ( $N_e U \ll 1$ ,  $N_e U \ln(N_e s) \ll 1$ , where  $s$  is the gain in fitness brought about by a mutation), populations are approximately monomorphic and the degrees of freedom of interest are the different alleles, codons or genotypes, which are fixed, or not, in the population; evolution can be described by a Markov process, where populations effectively jump sequentially between adjacent genotypes by a substitutional process [153], where in equilibrium (assuming uniform mutation and the fixation probability given by the diffusion approximation) the probability of occupation is given by the Boltzmann distribution

$$p_\sigma = e^{2N_e F_\sigma} / Z, \quad (7)$$

where  $F_\sigma$  is the fitness of genotype, allele, or codon  $\sigma$ , and  $Z$  is the partition function. It is clear that  $N_e$  plays the role of inverse temperature, such that fitness dominates at large population sizes (low temperature) and genetic drift for small population sizes (high temperature). Many of the calculational tools of statistical mechanics and generating functions then carry over to evolutionary problems [184] under usual ergodic assumptions.

The statistical mechanical analogy finds particular use in understanding the evolution of phenotypes arising from GP maps. Here, selection acts on phenotypes, but mutation and variation arise at the level of genotypes. Keeping in mind the many-to-one nature of most GP maps and phenotypic bias, the Boltzmann distribution of genotypes can be recast in terms of a Boltzmann distribution of phenotypes [84,185]; as each genotype mapping to a given phenotype must by definition have the same fitness, the probability of each phenotype is the Boltzmann factor ( $e^{2N_e F(\xi)}$ ) weighted by the degeneracy of each phenotype  $\Omega(\xi) = k^L f_\xi$ , giving

$$p(\xi) = e^{2N_e \Phi(\xi)} / Z, \quad (8)$$

where

$$\Phi(\xi) = F(\xi) + \frac{S(\xi)}{2N_e} \quad (9)$$

is the free fitness of phenotypes,  $S(\xi) = \ln(\Omega(\xi))$  being the Boltzmann or *sequence entropy* of phenotypes. We see that for small populations phenotypes with larger sequence entropy are favoured by genetic drift in evolution.

### 5.5.1. Statistical mechanics of the evolution of transcription factor-DNA binding

The ideas above first formally found application in simple biophysical models of transcription factor DNA binding [186–188], where the degeneracy of binding affinities can be exactly quantified under simplifying assumptions. It is typically found that for a given transcription factor, the amino acids at the binding interface tend to have a strong preference to bind a single nucleotide; it is then mismatched nucleotides that control the binding affinity, as these are strongly destabilising, since hydrogen bonding is disrupted at the interface, as well as the lost hydrogen bonds with water molecules. A simple model of transcription factor binding, assumes binding between protein and DNA can be reduced to either a quaternary or binary alphabet, where the binding energy  $E$  is proportional to the number of mismatches, or Hamming distance,  $h$ ,  $E = \epsilon h$ . The degeneracy function is then related to the binomial coefficient

$$\Omega(h) \propto \binom{L}{h} (k-1)^h. \quad (10)$$

This simple combinatorial argument shows that there is a huge degeneracy, or phenotypic, bias towards poor binding in this genotype-phenotype map. This methodology has been used to infer the effective genome-wide fitness landscape for transcription factor DNA binding in *Escherichia coli* and yeast [187–189], suggesting that on average binding is under stabilising selection, with monotonically decreasing fitness with decreasing binding affinity.

This simple model of transcription factor DNA binding suggests that smaller populations bear a significantly greater drift load under stabilising selection, than would be predicted if we assumed evolution based on phenotypes only [185,190]; while selection pushes populations to larger binding affinities, there is an opposing sequence entropic pressure for poorer binding. In equilibrium, these opposing tendencies are balanced, and it is the free fitness that is maximised, not fitness. This effect of sequence entropy on drift load is significantly greater than would be expected for a trait under stabilising selection, which ignores any degeneracy (see Box 5.5.1).

### 5.5.2. Evolution of genotypic divergence and reproductive isolation

One consequence of this significantly larger drift load is that in (allopatrically) diverging populations this gives rise to the prediction that reproductive isolation arises more quickly due to common ancestors already having more maladapted transcription factor-binding site pairs on average [185,190,191] (Fig. 4); if the common ancestor has a binding affinity closer to being deleterious (but kept in check by stabilising selection), then in hybrids Dobzhansky-Muller incompatibilities [192–194], which are incompatible combinations of transcription factors and DNA binding sites, arise more quickly after divergence. In particular, this mechanism is broadly consistent with trends seen in field-data [195–197] and diversification rates in phylogenetic trees [198–200], and so gives a robust explanation of how stabilising selection can give rise to this population size effect in speciation, without requiring passing through fitness valleys as do models based on the founder effect [201–204]. This model also predicts that those transcription factor-DNA binding site pairs, which are under weaker selection across a genome, would for the same reason give rise to a greater contribution to reproductive isolation, as the balance between selection and sequence entropy would be shifted to give common ancestors with weaker binding on average [190].

### 5.5.3. Marginal stability of compact proteins

Equilibrium statistical mechanical ideas also have the potential to explain the observed marginal stability of compact proteins. Various databases show that proteins have stabilities (measured through free-energy differences) of order  $\Delta G \sim -10$  kcal/mol, which is only a few hydrogen bonds [205], when their potential maximum stability could be orders of magnitude greater. Although adaptive explanations have been suggested related to the necessity of protein flexibility [206], a more straightforward explanation is offered in light of the free fitness of phenotypes; for a given chain length there are many more sequences that give poor protein stability and so this sequence entropy pressure balances the tendency of natural selection to choose proteins of higher stability. Simulations and theory of the evolution of protein folding [205,207–210] reproduce this marginal stability, with a particular property that the distribution of  $\Delta\Delta G$ , the change caused by mutations in the stability of a protein, is roughly independent of  $\Delta G$ , the stability of the protein. The marginal stability of proteins shows an interesting behaviour as function of population size  $N_e$ ; as we

**Box 5.5.1**

## Quadratic free fitness landscape

We examine a quadratic free fitness landscape, which is a simple description for the statistical mechanics of transcription factor DNA binding. For moderately large  $L$ , we can make the standard Gaussian approximation of the binomial distribution in the degeneracy function (10), such that the sequence entropy function (up to a constant) is approximately quadratic

$$S(h) \approx \frac{k}{2\langle h \rangle} (h - \langle h \rangle)^2, \quad (11)$$

where  $\langle h \rangle = (k - 1)L/k$  is the mean of the binomial distribution ( $\langle h \rangle = L/2$  for binary, and  $\langle h \rangle = 3L/4$  for quaternary alphabets). Further, if we assume that the fitness landscape of binding affinities is quadratic of the form  $F(h) = -\frac{1}{2}\kappa_F h^2$ , whose maximum is for the best binders ( $h = 0$ ), then from Eq. (9) the free fitness of binding energies/Hamming distance  $h$  is then itself quadratic with new curvature  $\kappa = \kappa_F + \frac{k}{2\langle h \rangle N_e}$  and with shifted maximum  $h^* = \frac{k}{2N_e \kappa}$ :

$$\Phi(h) = \frac{1}{2}\kappa (h - h^*)^2. \quad (12)$$

This new maximum is shifted to poorer binding affinities and represents the balance between selection and sequence entropy:

$$\frac{d\Phi}{dh} = \frac{dF}{dh} + \frac{1}{2N_e} \frac{dS}{dh} = 0.$$

It is instructive that the drift-load for this simple GP map varies as  $D \sim N_e^{-1}$ , a far stronger dependence on population size than if we considered evolution on only a phenotypic landscape  $F(r)$ , which would vary as  $D \sim N_e^{-1/2}$ ; this significant difference arises as the sequence entropic pressure causes the peak of the phenotypic distribution to shift, whilst ignoring this would simply give rise to a broadening of the distribution.

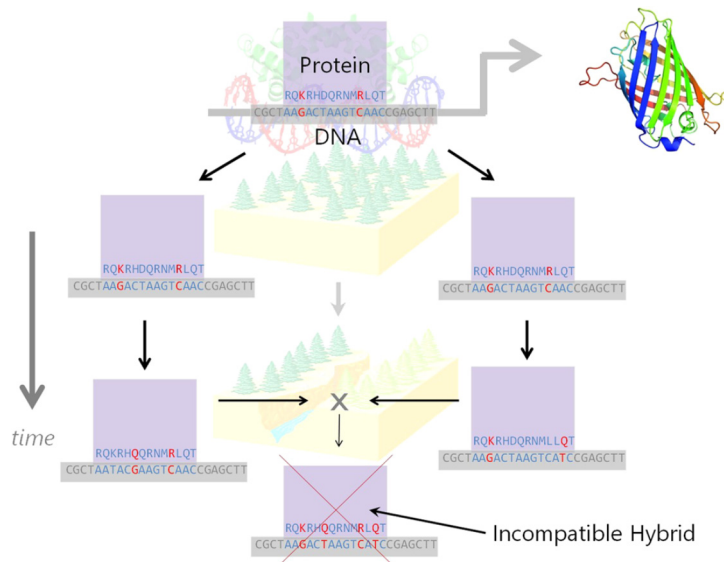


Fig. 4. Divergence for a simple GP map of transcription factor DNA binding. After a geographic split a once unique species evolves into two independent ones. Within simulations, the fitness of various hybrid combinations of loci from each lineage can be calculated and the number of inviable combinations (incompatibilities) recorded. Numerical results show that incompatibilities arise more quickly in smaller populations [190].

might expect, as the population size is increased, selection dominates genetic drift, and simulations [209,211] show that the average stability

$$\langle \Delta G \rangle \sim -k_B T \ln(N_e) \quad (13)$$

in the weak mutation limit ( $\mu N_e \ll 1$ ). This result can be obtained from simple scaling arguments [211], and it has been theoretically shown that, under global selection against misfolding, a broader scaling relationship between protein folding stability, protein cellular abundance, and effective population size holds [212]. Also, Eq. (13) can be shown to arise from similar reasoning to that in Box 5.5.1, as a result of a balance between selection and sequence entropy enhanced genetic drift at a given population size  $N_e$  [213].

#### 5.5.4. Free fitness, universality and developmental system drift

The question naturally arises: how universal or general is this effect in the genetic divergence of populations under stabilising selection? In developmental biology it is commonly found that closely related species that have similar organismal level phenotypes, such as body plans, nonetheless have diverged in the regulatory networks that control this patterning [214–216]. This cryptic variation is known as “developmental system drift” [217,218] and is a potential source of hybrid incompatibilities and ultimately reproductive isolation as previously explored by using simple gene regulatory networks [219,220]. However, in analogy to transcription factor DNA binding, if the GP map of developmental patterning has large degeneracy or phenotypic biases, we may also expect to find a rapid increase in the rate that hybrid incompatibilities arise as the population size decreases.

To explore this, a previously studied multi-level GP map for developmental spatial patterning [84] was used, in which gene regulation is manifested by multiple transcription factor DNA binding interactions, each described by a Hamming distance model as described above. Importantly, the GP map has the property that stabilising selection acts to maintain a body patterning phenotype, but the underlying genotypic degrees of freedom and molecular (binding energy) phenotypes can drift. In addition, the GP map has the essential property that allows an equilibrium analysis [84] that it is ergodic for small population sizes, which is in itself surprising given its state space is many orders of magnitude greater than can be explored in any realistic or relevant evolutionary timescale [221]. This ergodic property is closely related to the idea of space-shape covering, where the property of high dimensional maps means many phenotypes are potentially accessible from each other by only a few mutations. The main result is that in this more complex GP map, reproductive isolation also arises more quickly for small populations [85], which is related to the strong phenotypic bias. In addition, analogous to transcription factor DNA binding, it is also found that the molecular binding energy phenotypes—that underlie the organismal level patterning phenotype—which are under weakest selection, are most likely to give rise to the earliest hybrid incompatibilities.

Altogether, these results point to a universal picture to understand divergence between populations and the role of population size for strongly conserved traits; high-fitness phenotypes tend to be also highly specified, which means in converse low fitness phenotypes will have a large relative degeneracy or phenotypic bias. This means that the balance between fitness and sequence entropy, embodied by the maximum of free fitness, will be a strong feature of the equilibrium probability distribution of strongly conserved phenotypes, which are under stabilising selection. For simple biophysical traits like transcription factor DNA binding or protein stability, it is clear this is true since there will always only be a few sequences that give maximum affinity or stability, however, this has been found to be true even in a more complex GP map for developmental system drift. It is likely that in some way the sequence entropy constraints of transcription factor DNA binding propagate up in determining the sequence entropy of the organismal level patterning phenotype. The open questions are: how universal is this phenomenon?, will far more complex GP maps also show this behaviour?, will such maps maintain their property of ergodicity?, and is there a broad theoretical framework that can address this question without the more complex and computationally intensive simulations needed to address the former? Beyond an equilibrium analysis, there is the open question of dynamics and adaptation in GP maps [94,132,185,222], as well as extending this formalism to the strong mutation regime, yet still at finite population size (as compared to the infinite population size, quasispecies regime [182,183,222,223]).

## 6. GP maps as evolving objects

Important new insights on quantitative features of adaptation have been obtained by studying evolutionary processes with realistic, highly nonlinear GP maps, as presented in previous sections. However, the concept of a

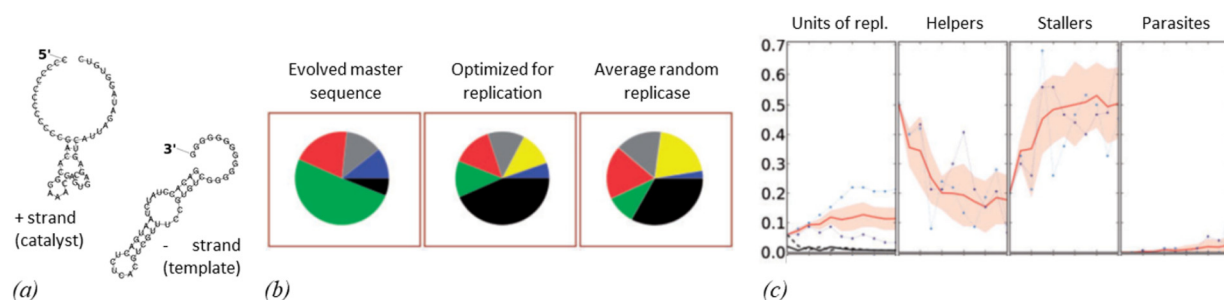


Fig. 5. The region of the genotype space selected by a population within an RNA world model is highly special as compared to controls. (a) Secondary structure of a replicase (+ strand) and its - strand which is optimized for maximum replication rate. (b) Functional classes in the 1-mutation-away mutational neighbourhood. Black: replicases; yellow: parasites; green: helpers; red: stallers; gray: junk; blue: unclassified. From left to right, pie charts correspond to a purely evolved replicase, a replicase optimized for maximum replication rate, and an average random replicase, respectively. Strong reduction of replicases and parasites and strong over-representation of helpers convey robustness to high mutation rates. (c) 1- to 10-mutations-away neighbourhoods ( $x$ -axis of each functional type) of the evolved GP map: at larger mutational (and therewith spatial) distances, frequencies of helper mutants decline drastically, whereas the frequency of stallers increases, thereby preferentially helping the ancestor and stalling others, including parasites.

predefined GP map on which evolutionary processes occur is not realistic. Not only should we expect the phenotype-to-fitness relation to vary due to environmental fluctuations—changing the fitness landscape into a seascape [224]—, but it is the GP map itself that is subject to evolution. Indeed, one might argue that in the long term what occurs is the evolution of the GP map, rather than a simple adaptation on a sort of preexisting genotype space.

By evolution of the GP map we mean two things: first, that the assignment of phenotypes to genotypes is a dynamic process that depends on context. As a consequence, the same genotype can present very different phenotypes during the course of evolution. And, second, that the dimensionality of the map changes during the course of evolution [83]. Indeed, duplications, deletions or large-scale chromosomal rearrangements, among others, are very frequent and often related to the acquisition of new or different phenotypic features [225].

In the following we will explore two computational models where the GP map itself is allowed to evolve, each demonstrating one of the features of GP map evolution mentioned above. We focus our discussion on the evolution of mutational neighbourhood, that determines which phenotypes are accessible from an evolved (evolving) genotype. As we will see, evolved populations in these two models fine-tune their mutational neighbourhoods so that adaptive phenotypes arise more frequently as a result of mutation. It appears that the explicit consideration of the mutational neighbourhood determined by the evolution of the GP map is essential for understanding not only long and short term evolution, but also the functioning of present day organisms.

### 6.1. Evolution of a multifunctional quasispecies in an RNA world model

The RNA World model [226] envisages a plausible scenario for the origin and early evolution of life. Understanding how the RNA World could have arisen involves explaining how diverse molecular function might emerge in the absence of faithful replication. Interestingly, it has been suggested that phenotypic bias could have played a main role in solving this problem [227]. Evolution and selection become possible only once the replication machinery is in place. In perspective, two alternative approaches have been extensively used for studying evolution of the RNA world: those that study the evolution on the RNA-sequence-to-secondary-structure GP map and those studying the impact of spatial pattern formation on what is selected. While the former class of models study the RNA world using the GP map with predefined fitness criteria, the latter explores the eco-evolutionary dynamics of replicator interactions without a predefined fitness. These two approaches have been combined [228,229] in a case study of the evolution of the qualitative, emergent functional properties of mutational neighbourhoods.

In this model, RNA sequences are embedded in a 2D grid and interact with their closest neighbours by complementary base pairing, forming complexes. If one of the molecules  $X$  folds into a structure with predefined motifs and binds to a molecule  $Y$ , replication can occur and the complementary strand of the molecule  $Y$  is formed. No fitness is explicitly defined, and therefore it arises as an emergent property of the population. Because of the spatial embedding, the interactions that occur are shaped by emergent spatial structures. Such emergent spatial structures constitute a new

level of selection and deeply affect the evolutionary outcome of replicators (as it has been shown in previous examples related to the RNA world [230,231]).

At all mutation rates studied, replicases rapidly evolve symmetry breaking between the complementary RNA strands, with one strand having replicase functionality and the complementary strand evolving an optimal template function—i.e. optimally binding the replicase. This symmetry breaking is also seen in non-GP-map-based toy models [232,233]. The stationary phenotypic composition of the population, however, strongly depends on mutation rate. At high mutation rates, only one, highly polymorphic quasispecies of replicases exists, whereas at lower mutation rates multiple quasispecies coexist. At intermediate mutation rates, there is coexistence between replicases and parasites, RNA molecules that act as templates for the replicases but which have no catalytic function themselves. At lower mutation rates, two different replicase-parasite communities coexist. Finally, at the lowest mutation rates these communities compete with each other sometimes going to extinction.

Let us focus now on the mutational neighbourhood of the replicases that evolve at the highest sustainable mutation rates. The functional composition (see Box 6.1) of the mutational neighbourhood of such evolved replicases is compared to two controls, i.e. a replicase that has been optimised for its replication rate, and randomly sampled replicases (Fig. 5b). In the mutational neighbourhood of the evolved replicases, replicases are scarce, parasites are missing, helpers are over-represented, and non-viable stallers are above average. In contrast, the controls have many replicases and parasites, and much fewer helpers (Fig. 5b).

The advantages of the multifunctional organisation of the mutational neighbourhood can be understood as follows. Non-viable mutants tend to be spatially close to their ancestor. Thus the helpers, in the close mutational neighbourhood, tend to help their ancestor and siblings rather than others. The non-viable helpers are essential for survival: if they are eliminated, the whole system goes extinct. This advantage of helpers is true only because there are no parasites in the mutational—and therefore spatial—neighbourhood. In contrast, stallers are detrimental for the system, but less so for their ancestor, for whose survival they are essential. This is because there are fewer stallers in the close neighbourhood than farther away (Fig. 5b and c) and they therefore hinder others more than the ancestor. In particular, they stall parasites if they emerge farther away. Indeed, if stallers are killed, parasites invade the system forming the two-species system characteristic of lower mutation rates.

In this scenario, functions were not pre-conceived, but emerged. Because the implemented GP map is actually the classical RNA GP map, and only point mutations are considered, the evolutionary dynamics could be seen as evolution *on* this fixed GP map. However, the GP map described in terms of these functions, and their structural implementation, fits better in the conceptualisation of evolution *of* the GP map, where some phenotypes (and thus functions, like helpers or stallers) evolve not as separate lineages, but as mutants in the evolved mutational neighbourhood of a replicase.

## 6.2. Evolution of genome size and evolvability in virtual cells

Now we explore *virtual cells* (Fig. 6a), a second model where we allow the dimensionality of the GP map to change while fixing a fitness function [234–236]. The system consists of a genome with genes coding for enzymes, pumps and transcription factors, as well as transcription factor binding sites. The transcribed genes form a simple metabolic network, which pumps in resources and transforms them into energy and building blocks. The external resource fluctuates, and fitness is homeostasis: the energy carrier and internal resource have to be close to a preset value. Average homeostasis over a cell's lifetime determines its fitness at replication. Mutations include changes in parameter values as well as gene duplications, deletions and large chromosomal rearrangements. Thus, genome size is variable.

Fig. 6b summarises the dynamics of evolved virtual cells at different stages. Early in evolution, genome size expands dramatically and immediately declines sharply. Although this transient is not generic, it occurs in those evolutionary runs which later reach high fitness. Interestingly, genome expansion does not entail an immediate fitness benefit, since there is no difference, in this time frame, between runs in which the common ancestor does and does not expand its genome. Subsequently, gene loss dominates the evolutionary dynamics most of the time, and often conveys increases in fitness.

The mutational neighbourhood, represented here as the fraction of mutants with decreasing fitness, has a characteristic “U-shape” (Fig. 6c), which becomes more pronounced during evolution. The fraction of neutral mutants remains the same despite an overall fitness increase, whereas the fraction of lethal mutations increases and the fraction of slightly deleterious mutations decreases.

**Box 6.1**

## Emergent functional classes in an RNA world model

In the RNA world model here described [228,229], molecular phenotypes were determined for pairs of complementary sequences (+ and – strands), based on specific structure motifs. Under evolution, the following phenotypes emerge:

**Self-replicases** can replicate both other molecules as well as themselves.

**Parasites** are RNA sequences that only work as templates and have no replicase ability.

**Helpers** can replicate other molecules but cannot be replicated.

**Stallers** can engage molecules in complexes, but can neither replicate them nor be replicated.

**Junk** cannot form complexes and are therefore mostly inert.

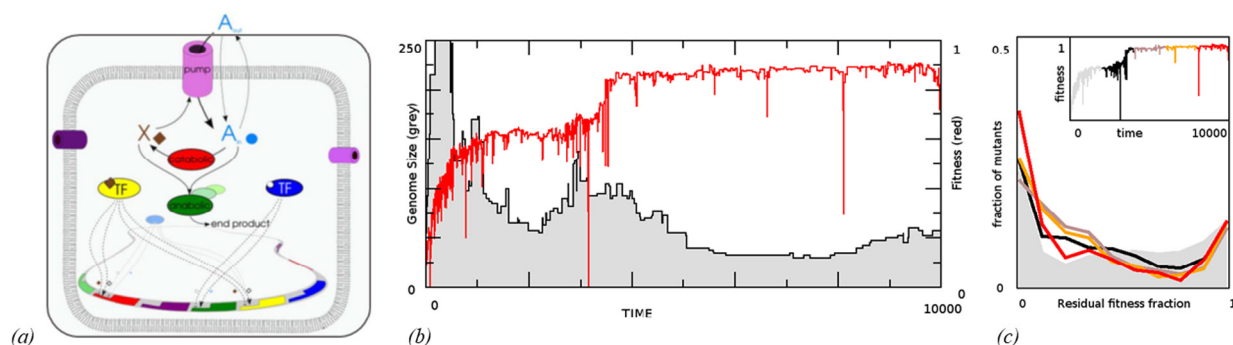


Fig. 6. Virtual cell and evolutionary dynamics. (a) Scheme of a virtual cell. (b) Common ancestor through time. Red line is the average fitness in three standard environments; shaded grey area depicts genome size, showing initial genome inflation followed by streamlining (gene loss). (c) Mutational neighbourhood of the ancestor in various time periods, colour coded according to inset: during evolution the mutational neighbourhood changes from the initial fitness distribution of the shaded grey area to a more pronounced “U-shape”, with peaks at neutral mutations (right side) and strongly deleterious mutations (left side).

Through this process, populations become highly evolvable. After a drastic environmental change (here implemented as a change in basic parameters) it sometimes takes only a few minor mutations to recover from nearly zero fitness to a value comparable to that previous to the environmental change; in other cases, a relatively fast recovery of fitness is mediated by genome expansion. After repeated environmental switches, evolvability through few mutations becomes common. Such fast evolvability turns out to be easier to evolve than regulatory mechanisms to adapt to changing environments [236].

These results are consistent with experimental reports. Phylogenetic reconstructions of long-term evolution show surprisingly large genome sizes of common ancestors (LUCA and LECA) and evolutionary dynamics dominated by gene loss [237]. The U-shape mutational neighbourhood has been observed in yeast [238] and viruses [239]. Lastly, fast adaptation to environmental changes mediated by few mutations or by genome expansion are well documented in many evolutionary experiments, for instance in yeast [240]. Antibiotic production in *Streptomyces* is done by highly unfit mutants [241], an evolutionary signature resembling the multifunctional quasispecies described in the RNA example. It is remarkable that all these surprising evolutionary signatures emerge in a minimal cell model, suggesting that they are generic features of Darwinian evolution, if genome organisation and the GP map are allowed to evolve.

## 7. Empirical genotype-to-phenotype and genotype-to-function maps

Technological advances are facilitating the experimental characterization of GP maps at ever-increasing resolution and scale [242,243]. The phenotypes of such maps include the activity or binding specificity of macromolecules such as RNA and proteins [67,244–246], the exonic composition of transcripts [247], the spatiotemporal gene expression pattern of regulatory circuits [248,249], as well as the function and flux of metabolic pathways [250]. In some cases, it is even possible to measure organismal fitness *en masse* [251–254].

When combined with a mapping from phenotype to fitness, biophysical GP maps provide a principled approach to constructing a fitness landscape over the space of genotypes. In situations where an empirical genotype-fitness map is available but a mechanistic understanding of its structure is lacking, one may instead try to infer the hidden phenotypic level from the genotype-fitness data. Ideally the inferred phenotypes can be interpreted biologically, but even when this is not the case, the introduction of an intermediate phenotypic layer helps to organise the high-dimensional genotypic data set and to reduce its complexity.

In this section, we begin with GP maps that have been empirically characterised. Sometimes, the quantity that is experimentally accessible is fitness, and not phenotype. We discuss how empirical data of that kind can be used to infer the structure of fitness landscapes and some properties of the phenotypic level. Then, we delve into the characterisation of GP and genotype-to-function maps in virus populations, discussing as well the implications of those maps in evolutionary dynamics under constant and variable environments. Next we address the inference of intermediate phenotypes from genotype-fitness data and, finally, we discuss approaches to the experimental characterisation of GP maps that may be relevant to synthetic biologists.

### 7.1. Empirical GP maps

There are three main approaches for constructing empirical GP maps [242,243]: (i) a combinatorially complete map is constructed using all possible combinations of a small set of mutations, such as those that occurred along an adaptive trajectory in a laboratory evolution experiment or in natural history [255]; (ii) a deep mutational scan assays the phenotypes of all single mutants, as well as many double- and triple-mutants of a single wild-type genotype [256]; and (iii) an exhaustively-enumerated map is constructed from all possible genotypes—something which is only possible for very small genotype spaces [32,33,257]. In some cases, such as with antibody repertoires [258,259] or viral populations [260,261], a fourth method of construction is possible. Specifically, one can directly construct a small portion of an empirical GP map by collecting a large number of genotypes with a particular phenotype from nature (e.g., the ability of an antibody to bind an antigen). Below, we describe a recent example from each of the three main categories, highlighting the biological insights gained from the construction and analysis of such maps.

#### 7.1.1. A combinatorially complete map

Alternative splicing is a key step of post-transcriptional gene regulation, and exonic mutations that affect splicing are commonly implicated in disease [262]. All possible combinations of mutations that occurred in the evolution of exon 6 in the human *FAS* gene since the last common ancestor of humans and lemurs have been analysed [263]. A total of 3,072 genotypes were assayed for the percentage of transcript isoforms that included the exon. This phenotype of “percentage spliced-in” varied from 0% to 100% among the 3,072 genotypes, indicating that in combination, these mutations are capable of producing the full range of exon inclusion levels. Importantly, the phenotypic change induced by a mutation depended non-monotonically upon the phenotype of the genotype in which the mutation was introduced, such that mutations to genotypes near the full-exclusion or full-inclusion phenotypic bounds had the smallest effects, whereas mutations to genotypes with intermediate inclusion levels had the largest effects. The resulting biological insight is that the evolution of an alternative exon from a constitutive exon will require several mutations, because mutation effect sizes are smallest when the exon is near full inclusion. This observation led to the mathematical derivation of a scaling law that applies to this and possibly other GP maps, and that may aid in the development of drugs aimed at targeting splicing for therapeutic benefit, by helping to predict drug-sensitive splicing events.

Another example of a combinatorially complete map will be explored more fully in Section 7.2.

#### 7.1.2. A deep mutational scan assay

Amino acid metabolism is fundamental to life, and is driven by complex metabolic and regulatory pathways. A deep mutational scan of nineteen genes involved in four pathways that affect lysine flux in *E. coli* was performed [250]. The resulting GP map consisted of 16,300 genotypes, each of which was assayed for its resistance to a lysine analogue that induces protein misfolding and reduces cell growth. The phenotype was therefore grown in the presence of the analogue. Several resistance-conferring mutations were identified, including mutations in transporters, regulators, and biosynthetic genes. For example, such mutations were often observed in a lysine transporter called LysP. These were relatively evenly distributed across the gene, suggesting that loss-of-function mutations were a common evolutionary path toward abrogated transport of the lysine analogue. More generally, this study represents a proof-of-concept that

deep mutational scanning experiments can be scaled up from individual macromolecules to regulatory and metabolic pathways.

### 7.1.3. An exhaustively enumerated map

Binding of regulatory proteins to DNA and RNA molecules are central to transcriptional and post-transcriptional gene regulation, respectively. The robustness and evolvability of these two layers of gene regulation has been studied via a comparative analysis of two empirical GP maps [264]. At the transcriptional level, interactions between DNA and transcription factors were considered, where a genotype was a short DNA sequence (a transcription factor binding site) whose phenotype was its molecular capacity to bind a transcription factor. At the post-transcriptional level, interactions between RNA and RNA binding proteins were analysed, where a genotype was a short RNA sequence (an RNA binding protein binding site) whose phenotype was the capacity to bind an RNA-binding protein. Though robustness at both layers of gene regulation was comparable, there were marked differences in evolvability, which were suggestive of qualitatively different architectural features in the two GP maps. Specifically, the genotype networks of binding sites for RNA binding proteins were separated by more mutations than the genotype networks of binding sites for transcription factors, rendering mutations to the binding sites of RNA binding proteins less likely to bring forth phenotypic variation than mutations to the binding sites of transcription factors. These observations are consistent with the rapid turnover of transcription factor binding sites among closely related species, as well as with the relatively high conservation levels of binding sites for RNA binding proteins. This comparative analysis may therefore help to explain why transcriptional regulation is more commonly implicated in evolutionary adaptations and innovations than post-transcriptional regulation mediated by RNA binding proteins.

## 7.2. Empirical fitness landscapes and adaptive dynamics of viral populations

The topography of fitness landscapes is key to understand evolutionary dynamics, and recent studies have focused on epistasis as a measure of landscape ruggedness (see Box 7.2). Two different experimental approaches have been taken to characterise the ruggedness of fitness landscapes through epistasis: a first, simpler approach is to analyse the epistasis among random pairs of mutations [239,265–267], while a more exhaustive approach relies on reconstructing a combinatorial fitness landscape that includes all possible combinations among a set of  $m$  mutations [242]. Usually, these  $m$  mutations have been observed during experimental evolution and adaptation of populations to novel environments. Such empirical landscapes have been characterised for bacteria [255,268–272], protozoa [155], fungi [273,274], and human immunodeficiency virus type-1 (HIV-1) [260,275–277].

In what follows, we review work focusing on the topography of an RNA virus fitness landscape. We begin with an investigation of how prevalent different epistasis types are (see Box 7.2), and then continue with the influence of landscape topography on the evolutionary potential of a virus population. Finally, we discuss the relevance of the environment on viral evolution, through analyses of landscapes on different host species.

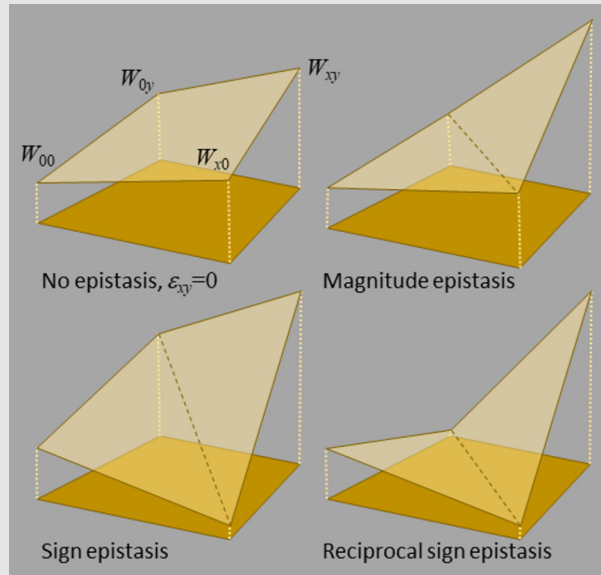
### 7.2.1. Description of epistasis among random pairs of mutations

The analysis of the effects of mutations on fitness provides information about the degree of ruggedness of the landscape at a coarse-grained level. In a study with *Tobacco etch potyvirus* (TEV) [283], 20 single nucleotide substitution mutations randomly scattered along the RNA genome of the virus were analysed. These mutations were deleterious when evaluated in *Nicotiana tabacum*, its natural host, through competition experiments against a reference TEV strain [284]. Those single mutations were randomly combined to yield 53 double mutants, whose fitness was measured also in *N. tabacum*. Twenty combinations rendered  $\epsilon_{xy}$  values significantly deviating from the null expectation, 11 of which were positive and 9 negative (see Box 7.2). Interestingly, these nine cases were all examples of synthetic lethality, that is, single mutations were deleterious but viable, but in combination became lethal. This represents an extreme case of negative epistasis. Previous studies with other RNA viruses obtained comparable epistatic interactions in type and sign [239,266,285].

How can we explain positive epistasis in the small and compact genomes of RNA viruses? Given the lack of genetic and functional redundancy and, in many cases, overlapping genes and multifunctional proteins, a small number of mutations can produce a strong deleterious effect. But, as mutations accumulate, they affect the same function with increasing probability and thus, their marginal contribution to fitness diminishes. Hence, the observed fitness is above the expected multiplicative value. In other words, epistasis is positive.

**Box 7.2**

## Epistasis and fitness landscapes



Epistasis means that the phenotypic effects of a mutation depend on the genetic background (genetic sequence) in which it occurs [269]. Whereas the concept applies principally to any phenotypic trait, in the evolutionary context epistasis for fitness is of primary importance, and we will focus on this case in what follows. The degree of epistasis,  $\epsilon_{xy}$ , between a pair of mutations  $x$  and  $y$  can be estimated as  $\epsilon_{xy} = W_{00}W_{xy} - W_{x0}W_{0y}$ , where  $W_{00}$  is the fitness of the non-mutated genotype,  $W_{xy}$  the experimentally determined fitness of the double mutant and  $W_{x0}$  and  $W_{0y}$  are the measured fitness of each single mutant. Under a multiplicative fitness effect model,  $W_{x0}W_{0y}/W_{00}$  represents the expected fitness value of the double mutant and, therefore,  $\epsilon_{xy}$  represents the deviation from this null hypothesis. The sign of  $\epsilon_{xy}$  corresponds to the sign of epistasis.

Magnitude epistasis causes deviations from the multiplicative model, but the landscape remains monotonic; sign epistasis means that the fitness sign of at least one of the mutations in the pair changes in presence of the other mutation; reciprocal sign epistasis occurs when both mutations change the sign of their fitness effect when combined, so both potential adaptive pathways connecting the nonmutated ancestor with the double mutant necessarily must cross a valley. Epistasis thus determines the ruggedness of a fitness landscape [278–280] and therefore the accessibility of adaptive pathways [281]. If there is either magnitude epistasis or no epistasis at all, fitness landscapes are smooth and single-peaked, and evolving populations can reach the global maximum. In the case of sign epistasis, only a fraction of the total paths to the optimum are accessible. Reciprocal sign epistasis is a necessary but not sufficient condition for rugged landscapes with multiple local optima [280], a situation where an evolving population might get stuck into suboptimal peaks. Most studies on epistasis have focused on pairwise epistasis, ignoring interactions among more than two mutations. However, higher-order epistasis appears in almost every published combinatorial fitness landscape [282], so the topographical features of fitness landscapes seem to depend on all orders of epistasis.

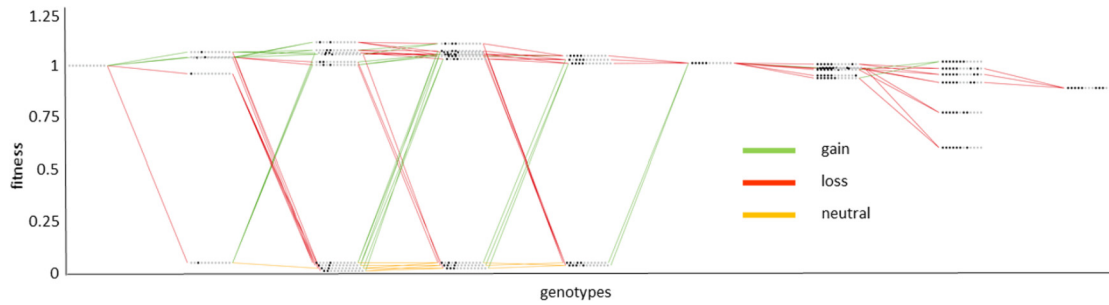


Fig. 7. Snapshot of an empirical fitness landscape constructed with combinations of mutations observed during experimental adaptation of tobacco etch potyvirus (TEV) to its new experimental host *Arabidopsis thaliana*. Each string of dots represents a genotype. Black dots represent a mutation in the corresponding locus, while grey dots correspond to the wild-type allele at that locus. Green lines stand for mutations with beneficial effect, red lines for deleterious mutations and orange lines for neutral mutations in the corresponding genetic background. Lines link genotypes which are one mutation away. The global optimum for this landscape corresponds to the 01001000 genotype. Data has been processed with MAGELLAN, which qualitatively orders genotypes along the x-axis according to the number of mutations [286].

### 7.2.2. Description of a combinatorial landscape and higher-order epistasis

TEV was evolved in a novel host, *Arabidopsis thaliana*, until it achieved high fitness [287]. The consensus genome of this adapted strain had only six mutations, three of which were nonsynonymous. The fitness effect of five of these mutations (the sixth one had to be discarded) was individually evaluated: two were significantly beneficial (one synonymous and one nonsynonymous), one was neutral (nonsynonymous), one deleterious (synonymous), and one lethal (nonsynonymous) [287]. All  $2^5 = 32$  possible genotypes that result from combining the observed five mutations were created in order to generate a complete five-sites landscape (Fig. 7), with abundant epistasis. Thus, the obtained landscape was rugged and without neutrality.

The pervasiveness of higher-order epistatic interactions in all empirically characterised combinatorial landscapes [282] prompted its study in the small TEV combinatorial landscape [288]. Using the Walsh-transform method [282, 289], higher-order epistatic interactions were found to be as important as pairwise interactions to fully understand the topological properties of adaptive landscapes.

Interestingly, and despite previous reports claiming that pervasive epistasis results in predictable evolutionary dynamics [242], repeated evolutionary experiments starting from different genotypes of the TEV virus resulted in different evolutionary endpoints, and populations were able to escape local optima, moving efficiently in this highly rugged landscape, with new mutations appearing in the course of evolution [290]. This result suggests that evolutionary predictions based on extrapolations from non-exhaustive fitness landscapes have to be taken with care, as evolving populations are often able to find new, previously undescribed mutations that introduce new evolutionary dimensions.

### 7.2.3. Effect of host species on the topography

Another quite common observation in evolutionary experiments with RNA viruses, as well as in natural populations, is the existence of pleiotropic fitness costs across different hosts [291]—beneficial mutational effects in one host may become deleterious in an alternative host. These negative fitness effects limit the host range of viruses to closely related species that share most of the molecular targets needed for the virus to complete its infectious cycle.

The concept of pleiotropy can be explored in terms of changes in the topography of fitness landscapes across hosts. The fitness of TEV single and double mutants was measured in four different susceptible hosts that differed in their degree of genetic relatedness [267]: the natural host *N. tabacum*, *Datura stramonium* (in the same botanical family, Solanaceae), *Helianthus annuus* (an Asteraceae phylogenetically related to the Solanaceae—both are Asterids), and in *Spinacea oleraceae* (an Amaranthaceae). Both the sign and the magnitude of epistasis changed across hosts: epistasis was positive ( $\epsilon_{xy} > 0$ ) only in the natural host, and it diminished as the host species relatedness to *N. tabacum* decreased.

The topography of the combinatorial fitness landscape was more rugged in *N. tabacum* when evaluated with the TEV strain adapted to *A. thaliana* [292]. Though the global optimum was the same in both landscapes, it was less accessible in *N. tabacum* given the greater magnitude of reciprocal sign epistasis in its vicinity.

Altogether, these results suggest that the topography of the adaptive fitness landscape for an RNA virus is strongly dependent on the environment (host species), though some general properties, such as the existence of lethal geno-

types, minimal or null neutrality and high ruggedness, remain. In this light, novel frameworks that explicitly account for environmental changes on the properties of landscapes, such as seascapes or adaptive multiscales [293], should yield a better picture to think about these experiments and even provide some predictive power.

### 7.3. Inferring phenotypes from genotype-fitness maps

Massively parallel empirical studies that examine a large ensemble of genotypes often yield information on their biological activity, or their overall fitness, while the identification of the phenotypes involved becomes difficult. In such cases it is possible to infer a phenotypic level, ideally endowed with a biological meaning, from data analogous to that of the previous sections. The key assumption underlying these formal approaches is that the mutational effects on the unobserved phenotypic traits are additive, such that any epistatic interaction for fitness arises from the nonlinearity of the phenotype-fitness map [294]. In the simplest case of a one-dimensional trait that maps monotonically to fitness, the trait variable has been referred to as the *fitness potential* [295,296] and the nonlinearity of the phenotype-fitness map as *global epistasis* [297]. Because a monotonic phenotype-fitness map preserves the rank ordering of genotypes with respect to fitness, it can account for magnitude epistasis, but not for sign epistasis [279].

Sign epistasis can however arise from non-monotonic one-dimensional phenotype-fitness maps. Consider a fitness function  $f(x)$  where the wild type trait value is located at  $x = 0$  and grows toward a single phenotypic optimum at  $x_{\text{opt}} > 0$ . A mutation that increases the trait value by an amount  $\Delta x < x_{\text{opt}}$  is then beneficial on the wild type background but deleterious on a background with trait value  $x \geq x_{\text{opt}}$  that overshoots the optimum. In an experimental study of the ssDNA bacteriophage ID11, it was found that this scenario explains the pairwise epistatic interactions between 9 individually beneficial mutations rather well [298]. In this case, the phenotype-fitness map was taken to be a gamma function with 4 parameters, and the unknown phenotype was parametrised by the 9 single mutational effects. The joint inference of the phenotypic effects and the phenotype-fitness map thus required 13 parameters to be estimated from the fitness values of 9 single and 18 double mutants.

The range of epistatic interaction patterns that can be generated from a one-dimensional phenotypic trait subject to a single-peaked phenotype-fitness map is obviously limited. In particular, any evolutionary trajectory composed of mutations that are individually beneficial on the wild-type background can display at most one fitness maximum. This criterion was used in a recent study of the combined resistance effects of synonymous mutations in the antibiotic resistance enzyme TEM-1  $\beta$ -lactamase challenged by cefotaxime to conclude that the phenotype underlying these effects is most likely multidimensional [299]. Multidimensional phenotypes allow for more versatile interaction structures but also require more parameters to be inferred from data. In a study of non-synonymous resistance mutations in TEM-1, a two-dimensional phenotype combined with a sigmoidal phenotype-fitness map was found to provide a good description of the measured resistance values [300]. In this case one of the phenotypes was taken to be protein stability, which was determined computationally, whereas the second phenotype was inferred along the lines of the experiment with the ID11 bacteriophage above [298]. A similar approach has been applied to the fitness landscape of a norovirus escaping a neutralising antibody, where the folding stability and binding affinity of the capsid protein were mapped to the probability of infection [254]. Importantly, in two-dimensional phenotype-fitness maps sign epistasis can emerge even in the absence of a phenotypic optimum [300,301].

The models described so far can be viewed as variants of the *geometric model* devised by Ronald Fisher to argue that the adaptation of complex phenotypes must proceed in small steps [302–304]. Originally, Fisher's geometric model (FGM) did not include the assumption of additive phenotypes, which was introduced later in a study of pairwise epistasis between mutations in *Escherichia coli* and vesicular stomatitis virus [305]. In its modern formulation, the model is based on a set of  $d$  real-valued traits forming a vector  $\vec{x} = (x_1, x_2, \dots, x_d)$  in the  $d$ -dimensional Euclidean space  $\mathbb{R}^d$  and a nonlinear phenotype-fitness function  $f(\vec{x})$  with a single optimum that is conventionally located at the origin  $\vec{x} = 0$  (Fig. 8). The genotype is described by a sequence  $\tau = (\tau_1, \tau_2, \dots, \tau_L)$  of  $L$  symbols  $\tau_i$  drawn from the allele set  $\{0, 1, \dots, k-1\}$ , where  $\tau_i = 0$  denotes the wild type allele and in most cases a binary alphabet with  $k = 2$  has been considered. The additive GP map takes the form [306]

$$\vec{x}(\tau) = \vec{x}_0 + \sum_{a=1}^{k-1} \sum_{i=1}^L \delta_{\tau_i, a} \vec{v}_{i, a}, \quad (14)$$

where  $\vec{x}_0$  is the wild type phenotype and the vector  $\vec{v}_{i, a} \in \mathbb{R}^d$  describes the phenotypic effect of the mutation  $0 \rightarrow a$  at the  $i$ 'th genetic locus. The genotype-fitness map is then obtained as  $F(\tau) = f[\vec{x}(\tau)]$ . In applications of FGM to

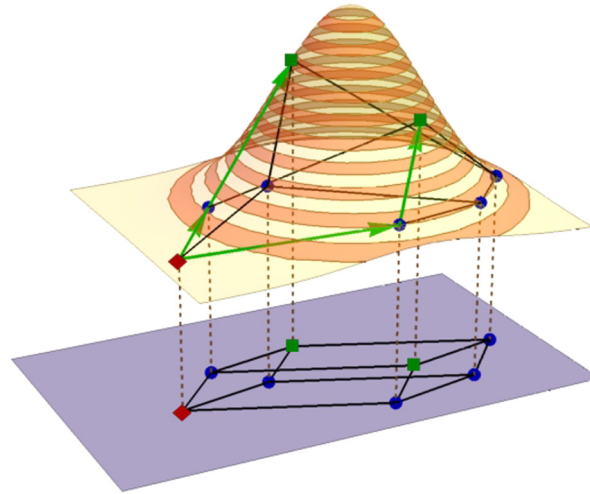


Fig. 8. Illustration of Fisher's geometric model for a two-dimensional phenotype and a single-peaked phenotype-fitness map. Three phenotypic mutations originating from the wild type (marked in red) combine additively, giving rise to a distorted three-dimensional cube in the phenotypic plane. As a consequence of the nonlinear mapping to fitness, two of the double mutants (marked in green) become local fitness maxima in the induced genotypic landscape. Courtesy of Sungmin Hwang.

experimental data, the mutational effects  $\vec{v}_{i,a}$  are usually treated as random vectors drawn from a multivariate Gaussian distribution. Rather than inferring specific phenotypes, such analyses yield gross statistical features of the phenotypic landscape, such as the number of phenotypic traits  $d$ , the distance of the wild type phenotype from the optimum  $|\vec{x}_0|$ , and the variance of phenotypic mutational effects [305,307–309].

Current high-throughput sequencing methods are capable of measuring fitness and other functional phenotypes for hundreds of thousands of genotype sequences in a single experiment, and methods based on the inference of unobserved additive traits provide an important tool for organising and interpreting the resulting data sets. A recent large-scale analysis of the fitness landscape of a segment of the *His3* gene in yeast built out of amino acid substitutions from extant species made use of a deep learning approach to infer the additive phenotype and its nonlinear mapping to fitness [310]. Remarkably, large parts of the data were well described by a one-dimensional fitness potential combined with a sigmoidal phenotype-fitness map, suggesting that much of the observed complexity of epistatic interactions could potentially be explained in terms of thermodynamic considerations [297]. Combining such data-driven inference methods with biophysical modelling and functional information appears to be a promising route towards a deeper understanding of the relation between genotype, phenotype and fitness on the molecular level [311].

#### 7.4. Synthetic biology approaches to characterising GP maps

A major goal in the field of synthetic biology is the re-purposing of biological components and systems to create living cells with new, designed functionalities. So what is the link between synthetic biology and GP maps? Faced with the challenge of understanding the function of biological parts and using this insight to rationally engineer cells, synthetic biologists frequently assemble large numbers of genetic designs (genotypes) and measure key aspects of the resultant cellular phenotypes. In doing so, novel methods for characterising GP maps have been developed (Fig. 9). Key to many of these are two capabilities. First, it is necessary to be able to construct large numbers of diverse genotypes (referred to as *libraries*) in a structured way. For example, assembling many genetic circuits simultaneously, each one containing a different combination of functional DNA parts (e.g. protein coding genes or regulatory elements like promoters, ribosome binding sites and terminators) [312–314]. Second, it should be possible to test these designs *en masse*. To support both requirements, high-throughput, pooled DNA assembly and sequencing methods have been developed to measure the phenotype of every genetic circuit design (genotype) across huge libraries, effectively creating a detailed GP map.

Genotype libraries can be constructed in many ways, each with their own advantages and pitfalls. Perhaps the simplest and most transferable protocol involves pooled synthesis of a large library of pre-defined DNA parts, insertion

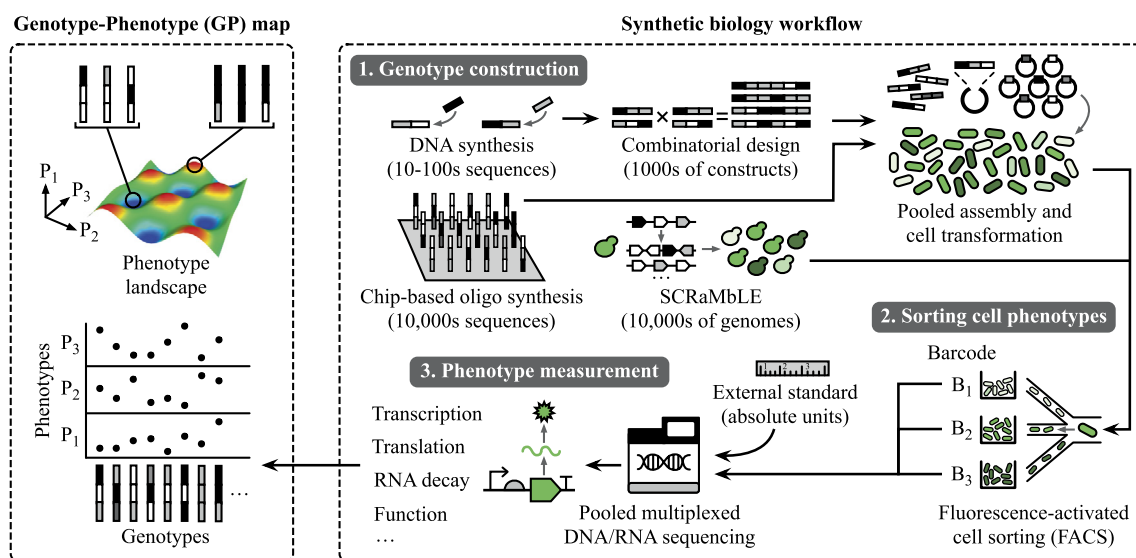


Fig. 9. Synthetic biology methodologies can support the construction of detailed GP maps. Diverse sets of genotypes in cells can be generated using combinatorial DNA assembly [315–317], chip-based DNA synthesis [318], or systems to induce structural DNA rearrangements, e.g. Synthetic Chromosome Recombination and Modification by LoxP-mediated Evolution (SCRaMbLE) [319]. Pooled libraries of cells can then be sorted into physically separated groups based on a parameter of interest, e.g. fluorescence of cells using fluorescence-activated cell sorting (FACS). Barcoded sequencing libraries can be generated from cells in each group and deep-sequencing of DNA/RNA performed to measure a wide range of phenotypic properties [313,320–325]. The inclusion of external standards during the sequencing allows for the conversion of relative phenotypic measures into absolute units that are comparable across contexts [326]. Genetic diagrams drawn using Synthetic Biology Open Language Visual notation [327]. Image courtesy of Thomas E. Gorochoowski.

of each part into a circular plasmid backbone that enables self-replication in cells, and transformation of the resulting plasmid library in the host cell of interest. This method can be used in combination with oligo(nucleotide) library synthesis (OLS) [318] for generation of the DNA part library. Whilst limitations include genotype length (up to 200 nucleotides) and accuracy (error rate of 1 in 200 nucleotides), accessible genotype libraries allow access to regions of genotype space distant from one another, with the latest OLS derived study characterising 244,000 sequences simultaneously [320]. Other approaches for constructing libraries of genotypes include multiplexed DNA assembly [315,317,328], and site-specific incorporation of random genetic diversity [329–332]. The latter approach was recently used to characterise millions of promoter variants [333].

Multiplexed measurement of many different phenotypes of the constructed genotype library is possible [320,326], though it must be ensured that each genotype contains a unique nucleotide-encoded barcode, to enable sequencing reads to be matched to the correct genotype [334]. Sequenceable phenotypes such as DNA or RNA abundance [320, 323,324,330] can be studied directly, in absolute units [321,322]. Non-sequenceable phenotypes can be measured too, by sorting phenotypes into groups and then appending a unique barcode sequence to genotypes in each group (Fig. 9). In this way, genotypes are mapped to phenotype categories. A detailed framework for design of pooled sequencing experiments (Multiplexed Assays for Variant Effects, MAVEs or Massively Parallel Reporter Assays, MPRA) is available [313].

Much of the focus to date has been on using these methods to characterise genetic part function, measuring the behaviour of parts taken from distantly related species or designed *in silico*. However characterisation of mutationally-connected genotype networks elucidates structural properties of GP maps for phenotypes which have not previously been characterised empirically at such scale. The resulting data can enable the construction of new *in silico* models for predicting phenotypes from genotypes [264,335] (section 7.1.3), a goal common to both synthetic biology and GP map researchers. Indeed, the synergy goes both ways: GP map studies highlight important principles which are only beginning to be considered by synthetic biologists during genetic circuit design. A clear example is genotypic robustness to mutations [264], which may prove important for genetic circuit longevity [336,337]. New science and technology brings new questions: ecological implications of microorganisms containing mutationally robust synthetic sequences have yet to be considered.

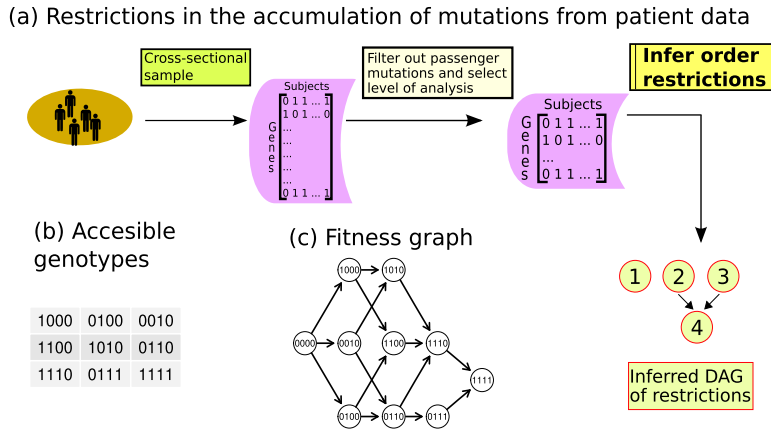


Fig. 10. Cancer progression models. (a) Main steps in the analysis of patient data. On the right, the DAG of restrictions shows genes in the nodes; an arrow from gene  $i$  to gene  $j$  indicates that a mutation in gene  $i$  must occur before a mutation in gene  $j$  can occur and, thus, indicates a direct dependency of a mutation in gene  $j$  on a mutation in gene  $i$ . The absence of an arrow between two genes means that there are no direct dependencies between the two genes. According to this DAG a mutation in the fourth gene can only be observed if both the second and third genes are mutated, but mutations in the first, second, and third gene do not have any dependencies among themselves. (b) Genotypes that fulfil the restrictions encoded in the DAG of restrictions: these are the accessible genotypes under the DAG. Genotypes are shown as sequences of 0 s and 1 s, where “1100” means a genotype with the first and second genes mutated. (c) Fitness graph or graph of mutational paths between accessible genotypes; nodes are genotypes (not genes) and arrows point toward mutational neighbours of higher fitness (thus, two genotypes connected by an arrow differ in one mutation that increases fitness [242,342]). Under CPMs, each new driver mutation with its dependencies satisfied increases fitness; therefore, all accessible genotypes that differ by exactly one mutation are connected in the fitness graph and the genotype with all driver genes mutated is the single fitness maximum. The fitness graph shows all the paths of tumor progression that start from the “0000” genotype and end in the fitness maximum. Figures modified from [343,344].

Expansion of this approach to study GP maps for different genotypes and phenotypes lies ahead. Innovations in nucleic acid sequencing are beginning to open up high-throughput characterisation of new types of phenotype in detail without sorting, such as epigenetic signatures or protein concentrations [325,338,339]. The advent of long read sequencing [340] beckons high-throughput characterisation of GP maps for whole-cell genotypes. This is becoming possible with methods for high-throughput genome modification [341], such as SCRaMbLE which uses recombination for *in vivo* combinatorial genomic rearrangement [319].

Pooled DNA assembly and sequencing is by no means the final solution for synthetic biologists or GP map researchers: crucially, it is limited by the number and length of assembleable genotypes and to phenotypes that can be inferred from sequencing data or for which high-throughput sorting methods (e.g. FACS) exist. Nonetheless, this approach offers a significant increase in the size of GP maps that can be studied empirically and highlights the potential for mutually beneficial collaboration across these two emerging areas of biological research.

## 8. Consequences of GP maps for models of tumour evolution and cancer progression

Epistatic interactions between genetic alterations can constrain the order of accumulation of mutations during cancer progression (e.g. in colorectal cancer, mutations in the *APC* gene are an early event that generally precedes mutations in the *KRAS* gene [345]). Cancer progression models (CPMs) have been developed to try to identify these restrictions during tumour progression using cross-sectional mutation data [346,347]. CPMs take as input a cross-sectional sample from a population of cancer patients: each individual or patient provides a single observation, the cancer genotype in that patient. Thus, the input for CPMs is a matrix of individuals or patients by alteration events, where each entry in the matrix is binary coded as mutated/not-mutated or altered/not-altered (Fig. 10). The output from CPMs are directed acyclic graphs (DAGs) that encode the restrictions inferred (which are in fact sign epistasis relationships [43,342]). In these DAGs, an edge between nodes  $i$  and  $j$  is to be interpreted as a direct dependence of an alteration of event  $j$  on an alteration of event  $i$ ;  $j$  should never be observed altered unless  $i$  is also altered. CPMs regard different patients as replicate evolutionary experiments, assume that the cancer cells in all patients are under the same genetic constraints [345–347], and ignore back mutations in the alteration of driver events. Thus, CPMs implicitly

encodes all the possible mutational paths or trajectories of tumour progression (Fig. 10) [343], and some methods (e.g., CBN) provide estimates of the probabilities of the different paths of tumour progression [343,348]. As in other domains, such as predicting antibiotic resistance, even small increases in our capacity to predict disease progression would be valuable for diagnostic, prognostic, and treatment purposes [349]; this renders CPMs a potentially useful tool in precision medicine. (Note that the focus here is on predicting mutational paths, but see also Section 5.3.1 for transition forecasts, a different objective and approach when predicting evolution using GP maps.)

Several CPM methods have been developed, including oncogenetic trees (OT) [350,351], conjunctive Bayesian networks (CBN) [345,352,353], and CAncer PRogression Inference (CAPRI) [354,355]. The different methods differ in their model fitting procedures and in the types of restrictions they can represent. For example, OT can only return trees, where a mutation in a given gene has a direct dependence on only one other gene mutation; this is in contrast to CAPRI and CBN, where a mutation in a gene can depend on mutations in two or more different genes, and thus CAPRI and CBN return as output DAGs where some nodes can have multiple parents, as shown in Fig. 10. All CPMs focus on “driver alterations”, i.e. those believed to actually drive, through selection, cancer progression (in contrast to so-called passenger mutations or hitchhikers). The types of alterations studied in CPMs range from changes in genes and pathways to gains and losses of chromosomal regions [345,351,355].

CPMs model sign epistasis, but they cannot model reciprocal sign epistasis (see section 6) [43], and thus CPMs effectively consider fitness landscapes with a single global peak. As a consequence, predictions of tumour progression, compared to the true paths of tumour progression, are very poor under multi-peaked fitness landscapes [343]. Remarkably, even in the latter scenario, CPMs could be used to estimate an upper bound to the true evolutionary unpredictability [343]; the analysis of twenty-two cancer data sets shows many of them to have low unpredictability.

CPMs do not force us to try to infer restrictions in the order of accumulation of alterations at any particular level or layer, in so far as the alterations examined can be regarded as heritable alterations with no back mutations. Thus, we could use the layers or levels of analysis (see also section 3.2) that are more relevant (e.g., metabolic pathways) or more likely to satisfy assumptions. Germane to this task are phenotypic bias (see section 3.3.1) and the effect on evolvability of many-to-many GP maps, relevant in the context of cancer [6,356]. These results suggest examining which is the most appropriate layer of analysis when using CPMs, which might not be gene alterations, but a layer closer to a “heritable phenotype”. On the one hand, layers other than genes could allow us to maximise predictive ability (related to ideas on how to choose the relevant phenotypic dimensions [357]). On the other hand, at other layers of analysis CPMs’ assumptions might be more likely to be satisfied—in particular, the lack of reciprocal sign epistasis and local fitness maxima, as well as the absence of disjunctive (OR) relationships in dependencies between alterations (when a mutation in a gene can happen if a mutation in at least one of its parents has occurred; in Fig. 10, under an OR model, a mutation in gene 4 would need one of genes 2 or 3 to be mutated, but not both) [43,343].

CPMs assume Markovian evolution. However, non-Markovian dynamics on neutral networks [132] (see also section 5) raises issues about choosing the layer of analysis for CPMs. For example, it seems unlikely that we could detect the existence of non-Markovian evolution reliably from the cross-sectional data used by CPMs. Additionally, non-Markovian evolution might be strong enough to cancel out possible benefits of working at other layers, and it could even be having an effect at the usual gene level of analysis where we label genes as altered/not-altered (mutated/not-mutated), because there is a many-to-one mapping between mutations in individual DNA bases and “altered” gene status.

The effects that environmental changes might have on evolutionary dynamics, given the dependence of epistatic relationships and fitness landscapes on the environment [6,169,243,267,292] (see also sections 6 and 7.2), could be particularly relevant for the use of CPMs if, as posited by the “adaptive oncogenesis” hypothesis [358], a key contribution to the relationship between age and cancer is the change in tissue fitness landscape with age (briefly, under the fitness landscapes of youth most mutants would have low fitness, unlike in the landscapes at older age). At a minimum, stratification of data sets by age would be warranted.

The sheer size of genotype and phenotype spaces is a potential matter of concern for CPMs, since the latter can only analyse a limited number of events. High-dimensional fitness landscapes might show increased mutational accessibility [359] and thus show both increased evolvability [243] and decreased evolutionary predictability. From the point of view of predicting tumour evolution with CPMs, robustness to alterations in the features examined by the CPM would of course be a hurdle; but then, hopefully, these features would not have been regarded as “drivers”. However, it should be mentioned here that “passenger” mutations in cancer, traditionally considered neutral, might actually reduce fitness of cancer cells and prevent tumour progression [360]; this raises the question of how to incor-

porate this lack of robustness in CPMs and, more generally, the extent of robustness and fitness landscape navigability in the cancer genome. The existence of a large pool of mildly deleterious passengers can also have consequences for procedures, such as CPMs, that analyse only a small subspace of the GP map.

Of note, CPMs are often used in cancer progression scenarios where aneuploidies and karyotypic changes are common [361]. This becomes an example of evolution of the GP map [234–236], a question deeply related to the proper comparison between GP maps of variable sequence length (see section 6). Choosing the right layer of analysis might again alleviate this problem, at least from the point of view of using CPMs to predict tumour evolution. Possible applicability of concepts emerging from evolving GP maps to the cancer genome is intriguing, especially given the possible costs of chromosomal instability and aneuploidy in cancer [360], with the caveat that cancer constitutes a short-term evolution experiment that starts from cells with a long evolutionary history and that dies with its host [362]. Finally, the extent to which neutrality and phenotypic bias (see sections 3.3.1 and 5) affect CPMs remain as open questions, since CPMs are predicated on the idea that natural selection is what matters for the features studied.

## 9. Summary and short-term perspectives

Exhaustive enumerations of genotype spaces are only feasible for short sequence lengths. These enumerations may be sufficient in specific empirical cases, as to study transcription factor binding sites (see Sections 5.5 and 7.1.3) or to build, in the near future, the first complete RNA GP maps incorporating experimentally measured fitness (SELEX experiments of small synthetic aptamers exploring the whole sequence space of length 24, such as those of Ref. [33], are already available). However, the number of possible genotypes for most biologically relevant sequence lengths is out of reach and, in the vast majority of cases, will always be: The estimated number of particles in the universe is of order  $10^{80}$ , a quantity comparable to the number of RNA sequences of length  $L = 133$  (the shortest known viroid has length  $L = 246$ ) or to that of proteins with 62 amino acids (the class of “small proteins” refers to those with fewer than 100 amino acids).

On the other hand, complete GP maps using RNA folding, the HP model, and toyLIFE [46], or using transcription factor binding [186,190], have proven to be very valuable resources for unveiling and testing some general properties of GP relationships, which seem to be common to several models. Further efforts toward theoretical developments that allow extrapolations to arbitrarily large genotype sizes, as well as approaches targeting higher levels of abstraction to study GP relationships without exploring the whole genotype space [113], appear as two main avenues to complement computational studies. Though the specifics of folding algorithms do not seem to affect the statistical properties of GP maps, we cannot forget that the predictive abilities of those algorithms depend on the accuracy of the energy model and its parameters, which in the case of RNA or proteins are extrapolated from experimental measurements obtained under very specific conditions. Therefore, any improvement on this aspect will have a huge impact, not only in the accuracy of RNA, proteins, and possibly other GP maps, but in every related research field concerned with functional prediction. Computational analyses might also benefit from approaches that do not demand an exhaustive enumeration, but are tailored to test theoretical predictions, for example. One of them might be computations of the dual partition function of multiple RNA structures. Also, complete inverse folding methodologies can be used to develop computational frameworks for the study of genotype-phenotype-function relationships. Current algorithms can potentially build partial GP maps focused on phenotypes of interest, in which experimental data available can be fit, thus providing an appropriate context to make predictions and guide further experiments. New tools able to produce reliable estimates of structural properties, like neutral set size, robustness or evolvability, should be ideally independent of the GP map, as well as experimentally compatible, i.e. they should allow predictions from small samples of genotypes. In this context, a sampling method that produces a genotype sample that optimally represents the phenotype of interest would be an important advance. There was some progress towards this goal in the form of a computational tool [92] which produces estimates of the size and robustness of RNA secondary structure phenotypes—but is in principle transferable to other GP maps—or through the estimation of the versatility of genotypes (see Section 4.2), but ample space for improvement remains. Also, and while these tools can predict neutral set size and robustness, no approaches exist yet for estimating phenotype evolvability or phenotype-phenotype correlations, which would yield the framework required to understand the evolution of evolvability or the deeply related concept of selection of the mutational neighbourhood.

We are only beginning to understand how the structure of GP maps depends upon environmental conditions [363–366]. We mostly ignore how the structure of a GP map changes with the dimensionality of genotype space,

a topic that, beyond simple evolvable cells [236] or toyLIFE [29], could potentially be explored using artificial genetic codes [367,368] or expanded nucleotide alphabets [369]. Finally, no matter the technological advance, the hyper-astronomical size of genotype space precludes the experimental construction of exhaustively-enumerated GP maps for large macromolecules, gene regulatory circuits, and metabolic pathways [11]. This inconvenient fact necessitates the development of methods that can reliably infer the structure of a GP map from a relatively small sample of the map [370,371]. Besides analytical approaches based on generic properties of GP maps that allow inferences of their large-scale structure (see section 4), advances in deep learning are already offering promising solutions to this key problem [372].

### 9.1. Towards an improved understanding of GP map architecture: Is it universal?

As discussed in section 4, notable similarities exist amongst GP map properties giving rise to the notion of “universal” [4,27] properties of GP maps, such as genetic correlations and phenotypic bias. Phenotypic bias, genetic correlations and evolvability are discussed in most studies of GP maps, but other properties, such as the assortativity of neutral networks [48], have only been analysed for some models. These topological properties could either provide a way of distinguishing between sequence-to-structure and artificial life GP maps or they could also be “universal” across a variety of models. At present, the universality of the structure induced in genotype spaces by evolutionarily sensible GP maps is a conjecture that those analyses, among others, could help to prove or disprove. Behind this conjecture, there is the main question of which fundamental mechanisms are responsible for the potentially universal features. As discussed in Section 4.1, spaces of high dimensionality that facilitate interconnections between genotypes and phenotypes seem to be a must.

More specific explanations for the striking similarities detected among dissimilar GP maps have come from simple analytic models [44,45] (see section 4.2). These models differ, but qualitatively they are all based on the fact that, depending on the phenotype, a part of the genotype is more constrained than the rest, for example to enable base pairing in RNA [45]. Interestingly, such a model can also be constructed to predict GP map properties in Richard Dawkins’ biomorphs [123]. The fact that such widely different GP maps can be understood with similar models, supports the hypothesis that sequence constraints are an important cause for the observed similarities between GP maps. A question for future research is the extent to which these sequence constraints generalise to other biological or artificial life GP maps. Are there any counter-examples? And do the kind of assumptions about sequence constraints in the analytic models always hold or can we observe GP maps with similar properties which cannot be modelled based on sequence constraints?

In the context of mathematical models of GP maps, it would also be desirable to further develop the existing models to explain more complex and biologically relevant situations, and to find out whether generic structural properties of genotype spaces are maintained under those circumstances. More realistic models should include mutations other than point mutations, such as deletions, duplications or insertions (there are just a few examples where the genome size is variable, among them that described in Section 6.2), and recombination (see Section 5.1). Extension to many-to-many GP maps by allowing multiple and semi-optimal phenotypes for a genotype—as it is the case for RNA sequences, for which there can exist multiple secondary structures with quite similar free energies—seems essential to fully understand adaptability [39,40]. Models such as toyLIFE and virtual cells should be further studied if we want to explore issues relevant to synthetic biology, among others. A very relevant question for the synthetic biology community has been to design gene regulatory circuits that are mutationally robust [373]. Results with toyLIFE show that genotypic robustness is a function not only of the individual components, but also of the complete network, which could be designed to be robust even if the individual components are not [30]. Actually, these extended models would have to come along with redefinitions of structural properties like neutral set size, robustness and evolvability.

### 9.2. Evolution on and of genotype spaces

Since the eventual aim of GP map studies is to understand evolutionary processes, a key question is how each of the universal GP map properties—should they exist—affect evolutionary outcomes. Phenotypic bias, for example, implies that only very abundant phenotypes will be visited when adapting to a new evolutionary challenge. This implies that the evolutionary search is constrained to look in the space of very abundant solutions. This constraint might lead to a limitation in the number of possible phenotypes attainable through evolution: it has been put forward, and

supported with simple developmental models, that the small fraction of phenotypes visible to evolution are highly clustered in morphospace and that the most frequent phenotypes are the most similar [374]—recalling the relevance of phenotype-phenotype correlations. Since evolutionary search is a consequence of the stochastic nature of the evolutionary dynamics, and is not dependent on the particulars of the GP map, there is no reason why this phenomenon should not be observed in real GP maps.

It has been also shown that transition times between phenotypes depend very strongly on how they are connected in genotype space, and there is also a strong indication that the genotypic robustness in a neutral network plays a role [375]: transition times between phenotypes depend strongly on how accessible a given phenotype is from the most robust genotypes. Because evolution naturally tends to visit the most robust genotypes [133], their connections to other phenotypes may be more relevant than those of less robust genotypes. The natural question to address is if there is a mutational bias in evolution towards phenotypes that connect to robust genotypes. Though computational GP maps have been the primary tool to explore this question in depth, some experimental work on this topic has been carried out as well. Indeed, *Pseudomonas aeruginosa* preferentially chooses three particular mutational pathways to evolve an adaptive phenotype under certain conditions [376]. When these pathways are repressed (through gene knockouts), the bacteria are able to evolve the same phenotype, but using new mutations. Actually, those mutations were available in the original population, but the probability of fixing them is very small compared to the three preferred pathways. This work gives empirical support to the relevance of mutational neighbourhoods for evolution (see Section 6.1), and highlights once more the need for further computational and experimental investigations of this topic.

Our knowledge of how the GP map properties individually affect evolutionary outcomes is still incomplete. For example, phenotypic bias is known to affect evolutionary outcomes due to at least three mechanisms: the ‘survival of the flattest’ [181], the ‘arrival of the frequent’ [94], and its effect on the free fitness of phenotypes in monomorphic regime [84,158,182,190]. Despite this progress, it may still be difficult to estimate for more complex cases than the scenario studied by Schaper and Louis, how strong the ‘arrival of the frequent’ effect will be and whether phenotypic frequency or phenotypic fitness are likely to determine evolutionary outcomes. Ultimately, this knowledge will help us answer the bigger questions of whether and how we can use GP maps to predict short- and long-term trends in evolution [377–379].

The application of the tools of network science to the previous context, and to evolutionary dynamics of heterogeneous populations at large, opens a promising avenue that, as of today, faces however some limitations and difficulties. First, and although the hyperastronomical sizes of genotype networks seem an insurmountable obstacle, the theory of competing networks shows that, in genotype spaces where function is relatively sparse, only the much smaller local subnetworks are relevant to analyse the evolution of populations—while the rest of the huge network of networks is in practice negligible [169]. A different promising avenue is the generic construction of a phenotype network that can be computationally—and likely analytically—tackled [91,94,132]. Second, most theoretical work has been developed using models that only consider point mutations. The introduction of different mutational mechanisms, as discussed in the previous sections, would drastically transform the topology and spectral properties of genotype networks. However, once the new network defined through those rules is known, the analysis proceeds following standard procedures. When the GP map is many-to-many, either due to environmental changes or to phenotypic promiscuity, more complex configurations such as multi-layer networks should be introduced to properly describe the evolution of the system [5,293]. Recombination cannot be easily cast in this network framework, which is unable to describe the process in detail [141,273,380]. Different approaches can be used in this case though (see Section 5.1), and hopefully combined to eventually yield a unified formal description of dynamics under a variety of microscopic processes generating diversity.

## 10. Outlook: On the feasibility of a complete genotype-to-organism map

Systems such as RNA folding, protein secondary structure, and transcription factor binding are attractive models for understanding the GP map because it is possible to compute the map from physical first-principles. But these processes are only the first steps in the long chain of interactions whose end result is organismal function, structure, viability, and reproduction [311]. At these higher levels of integration, it is the integration itself that in large measure determines the GP map. To address the question of whether evolution *of* the GP map or evolution *on* the GP map is the appropriate framework, it may be helpful to use a distinction [381] of two different properties of the GP map: *generative* properties—how the genotype is actually used to produce the phenotype—and *variational* properties—the

way that changes in the genotype map to changes in the phenotype. More recently, this distinction has been called “formative” and “differential” properties, respectively [382].

Unravelling the generative properties of the GP map is the main agenda of molecular and developmental biology. The variational properties ultimately derive from the generative properties, so the question is whether anything systematic about either can be predicted from evolutionary theory. Tremendous resources have been dedicated to molecular and cellular biology with the promise that, by identifying all the parts and interactions involved in a biological phenomenon, it could be understood, controlled, and even synthesised. The fruition of this promise has been realised in many cases, as attested to by the advent of successful treatments for many diseases. This was the justification of the Human Genome Project, with the hope that once all of the human DNA sequence was known, the genetic basis of diseases and organismal functions would be attainable. But a surprise from the Human Genome Project was the “missing heritability” that emerged from genome-wide association studies (GWAS). Analysis of DNA sequences could identify only a small fraction of the genes responsible for human phenotypes known from family studies to have high heritability.

Currently there are contradictory findings about the GP map at the whole organism level. On the one hand are studies which find that organisms exhibit a modular structure over large classes of phenotypic variables [383] to the point where modularity is often stated as an accepted fact [384]. On the other hand are studies which find that almost every gene affects many characters (universal pleiotropy) and almost every character is affected by many genes, summarised as the *omnigenic model* of the GP map [385]. The omnigenic model proposes that while there may be “core genes” contributing to any given phenotype, the network of gene-interactions has a “small world” topology, a property that leads to broad pleiotropy and polygeny in the GP map.

It may prove helpful that there is another field that is also trying to understand how complex functional behaviours emerge out of the interaction of thousands or millions of simple parts—the artificial neural network community. Artificial neural networks (ANNs) have now been created whose behaviours rival or exceed certain human cognitive capabilities. Because of the recent achievements of ANNs, the field is currently in an explosive state of development. The achievement of the engineers outpaced the understanding of the theoreticians as to why deep learning networks perform so well, and the theoreticians are working to catch up. As of now, there are several observations about ANNs that may be instructive to those making computational models of the GP map.

The multilayered “deep neural networks” (DNNs), which have proven to be the most successful ANNs, are defined by a collection of thousands or millions of algorithmically learned numbers. The numbers specify the weights of connections between nodes, and each node sums its inputs from other nodes, and then outputs a function of this sum as inputs to other nodes. What is most notable about DNN engineering is that there is very little interest in the specific values of the numbers, and no way to understand how the specific numbers generate the network behaviour. While there has been some success at *interpreting* DNNs—where one identifies what feature of an input causes a particular neuron in the network to activate—there is currently little understanding of how the all weights connecting the neurons produce this behaviour. The main focus has been on the *processes* that generate the numbers, and this is where theoreticians are attempting to generate understanding. The most successful process for training the weights is based on their variational properties: how changes in the weights change the error between the network’s actual behaviour and its desired behaviour. The methods of back-propagation and stochastic gradient descent change the weights until there is little or no error on a set of training examples applied to the network [386]. The variational properties of greatest interest are how the network behaves on novel inputs, and how changes in the inputs map to changes in the network behaviour.

A similar situation may hold in complex organisms. Without understanding or even knowing how the thousands of organismal components are generating phenotypes on the whole-organism level, we may nevertheless be able to understand its variational properties based on evolutionary processes. Here we briefly list a few principal processes that are understood to shape the variational properties of genotypes.

### 10.1. The evolution of re-evolvability under varying selection

In a number of different GP map models, evolution under recurring variation in natural selection moves the genome to places in genotype space where fewer and fewer mutations are needed to re-evolve previous adaptations when the old environment returns. This has been observed for a model of two neutral networks [387] (also called the evolution of “genetic potential” [388]), networks of logic gates (the variation is called “modularly varying goals” [389]), gene

regulatory networks and the virtual cells discussed above (just called the “evolution of evolvability” [236,390,391]). However, not all GP maps support this phenomenon [389]. Exactly what properties a GP map must possess to allow the evolution of re-evolvability remains an open problem.

### 10.2. Constructional selection

Genes not only provide material for the generation of the phenotype, but provide degrees of freedom for varying the phenotype. A gene duplication or *de novo* gene origin thus differs from a point mutation in that it increases the degrees of freedom of the GP map, and thus adds new variational properties to the genome. Gene duplications and deletions are frequent events in eukaryotic reproduction. Any variational property of a gene that is associated with the gene being retained in the genome can thus become enriched in the GP map [381]. The likelihood of a duplicate copy of a gene to be retained by evolution has been called its “gene duplicability” [392]. The identification of gene properties that are associated with gene duplicability is an active area of research. Some of the properties identified include:

- peripheral versus central position in protein-protein interaction networks [393];
- high levels of gene expression [394];
- high rates of sequence evolution before duplication [395];
- ordered, versus intrinsically disordered proteins [396];
- signalling, transport, and metabolism functions increase gene duplicability, while involvement in genome stability and organelle function reduces it, for whole genome duplications in plants [397].

While the causes of differential gene duplicability have been subject of a great deal of investigation, its *consequences* for organismal evolvability have received limited thought. Quantitative models for how differences in gene duplicability can shape the variational properties of the entire genome [381] have been applied to examples of evolutionary computation [398–400] under the rubric “constructional selection”. One can conceive of the genome as a population of genes, and differences in gene duplicability as fitness differences, not on the organismal level, but on this level of genome-as-population. Constructional selection results in the enrichment of the genome in genes that have a higher likelihood of being retained when copies of them are created. These are gene copies that evolve to where deletion or inactivation becomes deleterious to the organism. This occurs for genes more likely to subfunctionalise, or escape adaptive conflict, or neofunctionalise. It provides a ubiquitous mechanism for the evolution of evolvability.

### 10.3. Entropic evolutionary forces

The GP map is mostly cast as a many-to-one map because there may be multiple genotypes that result in the same phenotype, due to low-level properties such as synonymous codons, but also due to multiple ways that ligand-receptor bonds may be achieved, and multiple ways that the same gene regulatory interactions may be encoded. This degeneracy of the GP map [401] creates the possibility of evolution along neutral networks of mutationally-connected genotypes with the same fitness. The randomness of evolution along neutral networks brings forth statistical mechanical forces of entropy increase. This entropic behaviour has been described as “biology’s first law” [402].

Entropic phenomena that result from evolution along neutral networks include:

**Subfunctionalization.** If different functions in a gene are modular enough so that they can be individually disabled without affecting each other, then the process of gene duplication and complementary loss of functions effectively spreads the functions among multiple genes [403–405]. Since there are many more ways to spread the functions apart than to keep them in one gene, there is an entropic force in the direction of separating separable functions.

**Constructive neutral evolution.** Stoltzfus [404] introduced the general concept of entropic processes that produce greater genetic complexity to traits simply because there happen to be more complex ways to generate a trait than

there are simple ways. If there are neutral mutational pathways between alternate means of generating traits, then the more numerous class will come to dominate.

**Non-optimal phenotypes.** As explored in Section 5.5, in the context of the weak mutation monomorphic regime, there is an exact analogy to statistical mechanics, embodied in a quantity called free fitness [182,185], which is the sum the fitness of phenotypes and the sequence entropy (log degeneracy) weighted by the analogue of temperature, the inverse of the population size. This means that for small populations, evolution gives rise to non-optimal phenotypes that balance fitness and entropy, or free fitness.

**Developmental systems drift.** Primary sequences may diverge between species even while the same developmental outcomes are maintained [217]. Within the free fitness framework traits this has been explored (Section 5.5) under stabilising selection, and for small populations it is predicted that the effect of sequence entropy is that populations develop isolation more quickly [85].

The great variation in genome sizes over different taxa and even within closely related taxa suggests that the quantity of DNA maintained in the genome may function as a quantitative trait subject to species-specific natural selection. Just as in physical systems, where entropic forces can be counteracted by energy potentials, natural selection on genomic complexity as a quantitative trait may counteract the entropic tendencies in constructive neutral evolution. Such dynamics may be at work in the genomic streamlining discussed in Section 7. As is seen with the infinitesimal model of quantitative genetics, even though any individual streamlining event or an individual complexification event may have unobservable effects on fitness, the aggregate forces of entropic complexification and quantitative selection on genome size may statistically push the genome toward a balance, the character of which depends on the species-specific costs of genome maintenance.

#### 10.4. Omnigenic integration

When adaptations are produced by large scale interactions of organismal components, the GP map can be expected to be highly polygenic. Complex interactions of many components make the individual components also highly pleiotropic. Any given genetic change may beneficially affect certain traits while being detrimental to others. When their net effect is beneficial, then they are selected, but the deleterious effects they produce on certain traits creates the opportunity for other genetic variation to compensate for these effects. The GP map then becomes a patchwork of compensatory effects. In the limit of small effects, this patchwork becomes Fisher's *infinitesimal model*, [406] in which pleiotropy and polygeny are continuous and ubiquitous and there is little structuring of the GP map.

#### 10.5. Selection for mixability

Natural selection in sexual organisms with genetic recombination favours alleles that have high average fitness among all the different genotypes in which they appear in the population. An allele which might produce a highly adaptive phenotype when combined with just the right alleles at the same or other loci faces the breakup of such an advantageous combination due to segregation and recombination. Alleles which produce a reliable fitness advantage regardless of the genetic variation they are recombined with—a property called “mixability”—have a selective advantage [407]. The aggregate consequence of selection for mixability is toward greater modularity in the production of phenotypes: alleles individually produce the adaptive advantage without reliance on particular states of alleles at other loci. It suggests a process that counteracts the “omnigenic” model of complete genomic integration. The consequences of selection for mixability on the GP map have only begun to be elucidated [408].

#### 10.6. Epistatic smoothing of the fitness landscape

Conrad noted that a mutation which smoothed the fitness landscape for other loci would enhance their chance of producing advantageous mutations, and hitchhike along with such mutations, thus providing a constant force toward reducing reciprocal sign epistasis. This is the earliest mechanism proposed for the evolution of evolvability [409,410], and has yet to be fully investigated theoretically.

### 10.7. Summary

We have identified a patchwork of processes that in principle are able to shape the variational properties of the GP map for phenotypes at the level of whole organisms, where complex integration leaves us unable to derive the properties from physical first-principles. This is an area in which evolutionary theory needs much greater development. At levels of complexity at which detailed reductionist modelling is currently impossible, we have surveyed efforts to date that attempt to analyse how evolutionary processes shape the GP map. The body of results described, while not a fully fleshed-out theory, is perhaps sufficient to demonstrate that this process-based approach can inform a research program for the GP map at the whole organism level.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

All authors are indebted to the Centre Européen de Calcul Atomique et Moléculaire (CECAM) for supporting the organization of the workshop “From genotypes to function. Challenges in the computation of realistic genotype-phenotype maps”, which took place in Zaragoza (March 13th to March 15th, 2019) and triggered the production of this work. These are additional sources of financial support of the authors:

SM: grant FIS2017-89773-P (MINECO/FEDER, EU); “Severo Ochoa” Centers of Excellence to CNB, SEV 2017-0712.

JAC: grants FIS2015-64349-P (MINECO/FEDER, EU) and PGC2018-098186-B-I00 (MICINN/FEDER, EU).

JA: grant FIS2017-89773-P (MINECO/FEDER, EU); “Severo Ochoa” Centers of Excellence to CNB, SEV 2017-0712.

SEA: the Gatsby Charitable Foundation.

LA: Foundational Questions Institute (FQXi) and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation, for FQXi Grant number FQXi-RFP-IPW-1913, Stanford Center for Computational, Evolutionary and Human Genomics and the Morrison Institute for Population and Resources Studies, Stanford University, the 2015 Information Processing in Cells and Tissues Conference, and the Mathematical Biosciences Institute at The Ohio State University, for its support through National Science Foundation Award #DMS 0931642.

PC: Ramón Areces Postdoctoral Fellowship.

RDU: grant BFU2015-67302-R (MINECO/FEDER, EU).

SFE: grants BFU2015-65037-P (MCIU-FEDER) and PROMETEOII/2014/012 (Generalitat Valenciana).

JK: DFG within CRC1310 “Predictability in Evolution”.

NSM: Gates Cambridge Scholarship; Winton Programme for the Physics of Sustainability.

JLP: Swiss National Science Foundation, grant PP00P3\_170604.

MJT: grants EP/L016494/1 (EPSRC/BBSRC Centre for Doctoral Training in Synthetic Biology) and BB/L01386X/1 (BBSRC/EPSRC Synthetic Biology Research Centre, BrisSynBio).

MW: the EPSRC and the Gatsby Charitable Foundation.

### References

- [1] Reidys C, Stadler PF, Schuster P. Generic properties of combinatorial maps: neutral networks of RNA secondary structures. *Bull Math Biol* 1997;59:339–97.
- [2] Stadler PF, Stadler BMR. Genotype-phenotype maps. *Biol Theor* 2006;1:268–79.
- [3] Wagner A. The origins of evolutionary innovations. Oxford University Press; 2011.
- [4] Ahnert SE. Structural properties of genotype-phenotype maps. *J R Soc Interface* 2017;14:20170275.
- [5] Aguirre J, Catalán P, Cuesta JA, Manrubia S. On the networked architecture of genotype spaces and its critical effects on molecular evolution. *Open Biol* 2018;8:180069.
- [6] Nichol D, Robertson-Tessi M, Anderson ARA, Jeavons P. Model genotype-phenotype mappings and the algorithmic structure of evolution. *J R Soc Interface* 2019;16:20190332. <https://doi.org/10.1098/rsif.2019.0332>.
- [7] Gould SJ. The structure of evolutionary theory. Harvard University Press; 2002.

- [8] Love AC, editor. Conceptual change in biology: scientific and philosophical perspectives on evolution and development. Dordrecht: Springer; 2015.
- [9] Laland K, Uller T, Feldman M, Sterelny K, Müller GB, Moczek A, et al. Does evolutionary theory need a rethink? *Nature* 2014;514:161–4.
- [10] Stoltzfus A. Understanding bias in the introduction of variation as an evolutionary cause. In: Uller T, Laland KN, editors. *Evolutionary causation: biological and philosophical reflections*. Cambridge, MA: MIT Press; 2019.
- [11] Louis AA. Contingency, convergence and hyper-astronomical numbers in biological evolution. *Stud Hist Philos Sci C* 2016;58:107–16.
- [12] Greenbury SF, Schaper S, Ahnert SE, Louis AA. Genetic correlations greatly increase mutational robustness and can both reduce and enhance evolvability. *PLoS Comput Biol* 2016;12:e1004773.
- [13] Maynard Smith J. Natural selection and the concept of a protein space. *Nature* 1970;225:563–4.
- [14] Salisburry FB. Natural selection and the complexity of the gene. *Nature* 1969;224:342–3.
- [15] Huynen MA. Exploring phenotype space through neutral evolution. *J Mol Evol* 1996;43:165–9.
- [16] Schuster P, Fontana W, Stadler PF, Hofacker IL. From sequences to shapes and back: a case study in RNA secondary structures. *Proc R Soc Lond B* 1994;255:279–84.
- [17] Fontana W, Schuster P. Continuity in evolution: on the nature of transitions. *Science* 1998;280:1451–5.
- [18] Koelle K, Cobey S, Grenfell B, Pascual M. Epochal evolution shapes the phylodynamics of interpanemic influenza A (H3N2) in humans. *Science* 2006;314:1898–903.
- [19] Schultes EA, Bartel DP. One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science* 2000;289:448–52.
- [20] Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P. Fast folding and comparison of RNA secondary structures. *Monatsh Chem* 1994;125:167–88.
- [21] Lipman DJ, Wilbur WJ. Modelling neutral and selective evolution of protein folding. *Proc R Soc Lond* 1991;245:7–11.
- [22] Irback A, Troein C. Enumerating designing sequences in the HP model. *J Biol Phys* 2002;28:1–15.
- [23] Payne JL, Moore JH, Wagner A. Robustness, evolvability, and the logic of genetic regulation. *Artif Life* 2014;20:111–26.
- [24] Barve A, Wagner A. A latent capacity for evolutionary innovation through exaptation in metabolic systems. *Nature* 2013;500:203–6.
- [25] Hosseini S-R, Barve A, Wagner A. Exhaustive analysis of a genotype space comprising  $10^{15}$  central carbon metabolisms reveals an organization conducive to metabolic innovation. *PLoS Comput Biol* 2015;11:e1004329.
- [26] Johnston IG, Ahnert SE, Doye JP, Louis AA. Evolutionary dynamics in a simple model of self-assembly. *Phys Rev E* 2011;83:066105.
- [27] Greenbury SF, Johnston IG, Louis AA, Ahnert SE. A tractable genotype-phenotype map for the self-assembly of protein quaternary structure. *J R Soc Interface* 2014;11:20140249.
- [28] Ofria C, Wilke CO. Avida: a software platform for research in computational evolutionary biology. *Artif Life* 2004;10:191–229.
- [29] Arias CF, Catalán P, Manrubia S, Cuesta JA. toyLIFE: a computational framework to study the multi-level organisation of the genotype-phenotype map. *Sci Rep* 2014;4:7549.
- [30] Catalán P, Wagner A, Manrubia S, Cuesta JA. Adding levels of complexity enhances robustness and evolvability in a multi-level genotype-phenotype map. *J R Soc Interface* 2018;15:20170516.
- [31] Aguilar-Rodríguez J, Payne JL, Wagner A. A thousand empirical adaptive landscapes and their navigability. *Nat Ecol Evol* 2017;1:45.
- [32] Rowe W, Platt M, Wedge DC, Day PJ, Kell DB, Knowles J. Analysis of a complete DNA–protein affinity landscape. *J R Soc Interface* 2009;7:397–408.
- [33] Jiménez JJ, Xulvi-Brunet R, Campbell GW, Turk-MacLeod R, Chen IA. Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proc Natl Acad Sci USA* 2013;110:14984–9.
- [34] Wagner A. Mutational robustness accelerates the origin of novel RNA phenotypes through phenotypic plasticity. *Biophys J* 2014;106:955–65.
- [35] Jensen RA. Enzyme recruitment in evolution of new function. *Annu Rev Microbiol* 1976;30:409–25.
- [36] Aharoni A, Gaidukov L, Khersonsky O, Gould SM, Roodveldt C, Tawfik DS. The ‘evolvability’ of promiscuous protein functions. *Nat Genet* 2005;37:73.
- [37] Ance LW, Fontana W. Plasticity, evolvability, and modularity in RNA. *J Exp Zool* 2000;288:242–83.
- [38] Espinosa-Soto C, Martin OC, Wagner A. Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits. *BMC Evol Biol* 2011;11:5.
- [39] de Boer FK, Hogeweg P. Less can be more: RNA-adapters may enhance coding capacity of replicators. *PLoS ONE* 2012;7:e29952.
- [40] de Boer FK, Hogeweg P. Mutation rates and the evolution of multiple coding in RNA-based protocells. *J Mol Evol* 2014;79:193–203.
- [41] Payne JL, Wagner A. Latent phenotypes pervade gene regulatory circuits. *BMC Syst Biol* 2014;8:64.
- [42] Rezazadegan R, Reidys C. Degeneracy and genetic assimilation in RNA evolution. *BMC Bioinform* 2018;19:543.
- [43] Diaz-Urriarte R. Cancer progression models and fitness landscapes: a many-to-many relationship. *Bioinformatics* 2018;34:836–44.
- [44] Greenbury S, Ahnert S. The organization of biological sequences into constrained and unconstrained parts determines fundamental properties of genotype–phenotype maps. *J R Soc Interface* 2015;12:20150724.
- [45] Manrubia S, Cuesta JA. Distribution of genotype network sizes in sequence-to-structure genotype-phenotype maps. *J R Soc Interface* 2017;14:20160976.
- [46] Garcia-Martin JA, Catalán P, Cuesta JA, Manrubia S. Phenotype size distributions in exact enumerations of genotype spaces. *Europhys Lett* 2018;123:28001.
- [47] Aguirre J, Buldú JM, Manrubia SC. Evolutionary dynamics on networks of selectively neutral genotypes: effects of topology and sequence stability. *Phys Rev E* 2009;80:066112.
- [48] Aguirre J, Buldú JM, Stich M, Manrubia SC. Topological structure of the space of phenotypes: the case of RNA neutral networks. *PLoS ONE* 2011;6:e26324.
- [49] Schuster P. Prediction of RNA secondary structures: from theory to models and real molecules. *Rep Prog Phys* 2006;69:1419–77.
- [50] Salazar-Ciudad I, Jernvall J. A computational model of teeth and the developmental origins of morphological variation. *Nature* 2010;464:583–6.

- [51] Salazar-Ciudad I, Marín-Riera M. Adaptive dynamics under development-based genotype–phenotype maps. *Nature* 2013;497:361–4.
- [52] Hagolani PF, Zimm R, Vroomans R, Salazar-Ciudad I. On the evolution and development of morphological complexity: a view from gene regulatory networks. *PLoS Comput Biol* 2021;17:e1008570.
- [53] Dawkins R. The evolution of evolvability. In: Kumar S, Bentley PJ, editors. *On growth, form and computers*. London: Academic Press; 2003. p. 239–55.
- [54] Li H, Helling R, Tang C, Wingreen N. Emergence of preferred structures in a simple model of protein folding. *Science* 1996;273:666–9.
- [55] Stich M, Briones C, Manrubia SC. On the structural repertoire of pools of short, random RNA sequences. *J Theor Biol* 2008;252:750–63.
- [56] Shahrezaei V, Hamedani N, Ejtehadi MR. Protein ground state candidates in a simple model: an enumeration study. *Phys Rev E* 1999;60:4629–36.
- [57] Wagner A. Robustness and evolvability: a paradox resolved. *Proc R Soc Lond B* 2007;275:91–100.
- [58] Holzgräfe C, Irbäck A, Troein C. Mutation-induced fold switching among lattice proteins. *J Chem Phys* 2011;135:195101.
- [59] Ferrada E, Wagner A. A comparison of genotype-phenotype maps for RNA and proteins. *Biophys J* 2012;102:1916–25.
- [60] Shakhnovich E. Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. *Chem Rev* 2006;106:1559–88.
- [61] Ciliberti S, Martin OC, Wagner A. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput Biol* 2007;3:e15.
- [62] Chen Y, Ding F, Nie H, Serohijos AW, Sharma S, Wilcox KC, et al. Protein folding: then and now. *Arch Biochem Biophys* 2008;469:4–19.
- [63] Mirny LA, Shakhnovich EI. How to derive a protein folding potential? A new approach to an old problem. *J Mol Biol* 1996;264:1164–79.
- [64] Buchler NEG, Goldstein RA. Effect of alphabet size and foldability requirements on protein structure designability. *Proteins* 1999;34:113–24.
- [65] Li H, Tang C, Wingreen NS. Designability of protein structures: a lattice-model study using the Miyazawa-Jernigan matrix. *Proteins* 2002;49:403–12.
- [66] Bastolla U, Porto M, Roman HE, Vendruscolo M. Connectivity of neutral networks, overdispersion, and structural conservation in protein evolution. *J Mol Biol* 2003;56:243–54.
- [67] Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, et al. Local fitness landscape of the green fluorescent protein. *Nature* 2016;533:397–401.
- [68] Ferrada E, Wagner A. Protein robustness promotes evolutionary innovations on large evolutionary time-scales. *Proc R Soc Lond B* 2008;275:1595–602.
- [69] Ferrada E, Wagner A. Evolutionary innovations and the organization of protein functions in genotype space. *PLoS ONE* 2010;5:e14172.
- [70] England JL, Shakhnovich EI. Structural determinant of protein designability. *Phys Rev Lett* 2003;90:218101.
- [71] Callaway E. ‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures. *Nature* 2020;588:203–4.
- [72] Graur D, Zheng Y, Price N, Azevedo RBR, Zufall RA, Elhaik E. On the immortality of television sets: “function” in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol Evol* 2013;5:578–90.
- [73] Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, et al. Defining functional DNA elements in the human genome. *Proc Natl Acad Sci USA* 2014;111:6131–8.
- [74] Domingo E, Perales C. Viral quasispecies. *PLoS Genet* 2019;15:e1008271.
- [75] Wagner A. How the global structure of protein interaction networks evolves. *Proc R Soc Lond B* 2003;270:457–66.
- [76] Ciliberti S, Martin OC, Wagner A. Innovation and robustness in complex regulatory gene networks. *Proc Natl Acad Sci USA* 2007;104:13595–6.
- [77] Matias Rodrigues JF, Wagner A. Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLoS Comput Biol* 2009;5:e1000613.
- [78] Samal A, Rodrigues JFM, Jost J, Martin OC, Wagner A. Genotype networks in metabolic reaction spaces. *BMC Syst Biol* 2010;4:30.
- [79] Leclerc RD. Survival of the sparsest: robust gene networks are parsimonious. *Mol Syst Biol* 2008;4:213.
- [80] Serohijos AWR, Shakhnovich EI. Merging molecular mechanisms and evolution: theory and computation at the interface of biophysics and evolutionary population genetics. *Curr Opin Struct Biol* 2014;26:84–91.
- [81] Heo M, Kang L, Shakhnovich EI. Emergence of species in evolutionary “simulated annealing”. *Proc Natl Acad Sci USA* 2008;106:1869–74.
- [82] Heo M, Maslov S, Shakhnovich E. Topology of protein interaction network shapes protein abundances and strengths of their functional and nonspecific interactions. *Proc Natl Acad Sci USA* 2011;108:4258–63.
- [83] Zeldovich KB, Chen P, Shakhnovich BE, Shakhnovich EI. A first-principles model of early evolution: emergence of gene families, species, and preferred protein folds. *PLoS Comput Biol* 2007;3:e139.
- [84] Khatri BS, McLeish TCB, Sear RP. Statistical mechanics of convergent evolution in spatial patterning. *Proc Natl Acad Sci USA* 2009;106:9564–9.
- [85] Khatri BS, Goldstein RA. Biophysics and population size constrains speciation in an evolutionary model of developmental system drift. *PLoS Comput Biol* 2019;15:1–24.
- [86] Catalán P. Models in molecular evolution: the case of toyLIFE. Ph.D. thesis. Universidad Carlos III de Madrid; 2017.
- [87] Fontana W, Konings DA, Stadler PF, Schuster P. Statistics of RNA secondary structures. *Biopolymers* 1993;33:1389–404.
- [88] Schultes EA, Spasic A, Mohanty U, Bartel DP. Compact and ordered collapse of randomly generated RNA sequences. *Nat Struct Mol Biol* 2005;12:1130–6.
- [89] Wagner A. Robustness and evolvability in living systems. Princeton, NJ: Princeton University Press; 2005.
- [90] Smit S, Yarus M, Knight R. Natural selection is not required to explain universal compositional patterns in rRNA secondary structure categories. *RNA* 2006;12:1–14.
- [91] Cowperthwaite MC, Economo EP, Harcombe WR, Miller EL, Meyers LA. The ascent of the abundant: how mutational networks constrain evolution. *PLoS Comput Biol* 2008;4:e1000110.
- [92] Jörg T, Martin OC, Wagner A. Neutral network sizes of biological RNA molecules can be computed and are not atypically small. *BMC Bioinform* 2008;9:464.

- [93] Stich M, Manrubia SC. Motif frequency and evolutionary search times in RNA populations. *J Theor Biol* 2011;280:117–26.
- [94] Schaper S, Louis AA. The arrival of the frequent: how bias in genotype-phenotype maps can steer populations to local optima. *PLoS ONE* 2014;9:e86635.
- [95] Dingle K, Schaper S, Louis AA. The structure of the genotype-phenotype map strongly constrains the evolution of non-coding RNA. *Interface Focus* 2015;5:20150053.
- [96] Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 1999;288:911–40.
- [97] Zuker M, Stiegler P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res* 1981;9:133–48.
- [98] McCaskill JS. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers* 1990;29:1105–19.
- [99] Wuchty S, Fontana W, Hofacker IL, Schuster P. Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* 1999;49:145–65.
- [100] Lorenz R, Bernhart SH, Höner zu Siederdisen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA package 2.0. *Algorithms Mol Biol* 2011;6:26.
- [101] Kin T, Yamada K, Terai G, Okida H, Yoshinari Y, Ono Y, et al. fRNAdb: a platform for mining/annotating functional RNA candidates from non-coding RNA sequences. *Nucleic Acids Res* 2007;35:D145–8.
- [102] Diener TO. Potato spindle tuber “virus”, IV: a replicating, low molecular weight RNA. *Virology* 1971;45:411–28.
- [103] Elena SF, Gómez G, Daròs J-A. Evolutionary constraints to viroid evolution. *Viruses* 2009;1:241–54.
- [104] Serio FD, Navarro B, Flores R. Origin and evolution of viroids. In: Hadidi A, Flores R, Randles J, Palukaitis P, editors. *Viroids and satellites*. Academic Press; 2017. p. 125–34.
- [105] Sanjuán R, Forment J, Elena SF. In silico predicted robustness of viroids RNA secondary structures, I: the effect of single mutations. *Mol Biol Evol* 2006;23:1427–36.
- [106] Sanjuán R, Forment J, Elena SF. In silico predicted robustness of viroids RNA secondary structures, II: interactions between mutation pairs. *Mol Biol Evol* 2006;23:2123–30.
- [107] Cuesta JA, Manrubia S. Enumerating secondary structures and structural moieties for circular RNAs. *J Theor Biol* 2017;419:375–82.
- [108] Catalán P, Elena SF, Cuesta JA, Manrubia S. Parsimonious scenario for the emergence of viroid-like replicons de novo. *Viruses* 2019;11:425.
- [109] Schnall-Levin M, Chindelevitch L, Berger B. Inverting the viterbi algorithm: an abstract framework for structure design. In: Cohen WW, McCallum A, Roweis ST, editors. *Proceedings of the 25th international conference on machine learning*, vol. 307. New York: ACM; 2008. p. 904–11.
- [110] Churkin A, Retwitzer MD, Reinharz V, Ponty Y, Waldispühl J, Barash D. Design of RNAs: comparing programs for inverse RNA folding. *Brief Bioinform* 2017;19:350–8.
- [111] Borenstein E, Rupp E. Direct evolution of genetic robustness in microRNA. *Proc Natl Acad Sci USA* 2006;103:6593–8.
- [112] Szöllősi GJ, Derényi I. Congruent evolution of genetic and environmental robustness in micro-RNA. *Mol Biol Evol* 2009;26:867–74.
- [113] Garcia-Martin JA, Bayegan AH, Dotu I, Clote P. RNAfoldPF: software to compute the dual partition function with sample applications in molecular evolution theory. *BMC Bioinform* 2016;17:424.
- [114] Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, et al. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res* 2018;46:D335–42.
- [115] Barrett C, He Q, Huang FW, Reidys CM. An efficient dual sampling algorithm with Hamming distance filtration. *J Comput Biol* 2018;25:1179–92.
- [116] Garcia-Martin JA, Clote P, Dotu I. RNAiFOLD: a constraint programming algorithm for RNA inverse folding and molecular design. *J Bioinform Comput Biol* 2013;11:1350001.
- [117] Dotu I, Garcia-Martin JA, Slinger BL, Mechery V, Meyer MM, Clote P. Complete RNA inverse folding: computational design of functional hammerhead ribozymes. *Nucleic Acids Res* 2014;42:11752–62.
- [118] Garcia-Martin JA, Dotu I, Fernandez-Chamorro J, Lozano G, Ramajo J, Martinez-Salas E, et al. RNAiFold2T: constraint programming design of thermo-IRES switches. *Bioinformatics* 2016;32:i360–8.
- [119] Lindenmayer A. Mathematical models for cellular interactions in development, I: filaments with one-sided inputs. *J Theor Biol* 1968;18:280–99.
- [120] Lindenmayer A. Mathematical models for cellular interactions in development, II: simple and branching filaments with two-sided inputs. *J Theor Biol* 1968;18:300–15.
- [121] Raman K, Wagner A. The evolvability of programmable hardware. *J R Soc Interface* 2011;8:269–81.
- [122] Fortuna MA, Zaman L, Ofria C, Wagner A. The genotype-phenotype map of an evolving digital organism. *PLoS Comput Biol* 2017;13:e1005414.
- [123] Martin NS, Camargo CQ, Louis AA. *Untitled data*, 2020, unpublished.
- [124] Lehre PK, Haddow PC. Accessibility between neutral networks in indirect genotype-phenotype mappings. In: 2005 IEEE congress on evolutionary computation, vol. 1; 2005. p. 419–26.
- [125] Lehre PK, Haddow PC. Phenotypic complexity and local variations in neutral degree. *Biosystems* 2007;87:233–42.
- [126] Fontana W, Schuster P. Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. *J Theor Biol* 1998;194:491–515.
- [127] Perelson AS, Oster GF. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *J Theor Biol* 1979;81:645–70.
- [128] Bornberg-Bauer E, Chan HS. Modeling evolutionary landscapes: mutational stability, topology, and superfunnels in sequence space. *Proc Natl Acad Sci USA* 1999;96:10689–94.

- [129] Weiß M, Ahnert SE. Phenotypes can be robust and evolvable if mutations have non-local effects on sequence constraints. *J R Soc Interface* 2018;15:20170618.
- [130] Alberch P. From genes to phenotype: dynamical systems and evolvability. *Genetica* 1991;84:5–11.
- [131] Draghi JA, Parsons TL, Wagner GP, Plotkin JB. Mutational robustness can facilitate adaptation. *Nature* 2010;463:353–5.
- [132] Manrubia S, Cuesta JA. Evolution on genotype networks accelerates the ticking rate of the molecular clock. *J R Soc Interface* 2015;12:20141010.
- [133] van Nimwegen E, Crutchfield JP, Huynen M. Neutral evolution of mutational robustness. *Proc Natl Acad Sci USA* 1999;96:9716–20.
- [134] Catalán P, Manrubia S, Cuesta JA. Populations of genetic circuits are unable to find the fittest solution in a multilevel genotype-phenotype map. *J R Soc Interface* 2020;17:20190843.
- [135] Wilke CO. Adaptive evolution on neutral networks. *Bull Math Biol* 2001;63:715–30.
- [136] Reidys CM, Forst CV, Stadler PF. Replication and mutation on neutral networks. *Bull Math Biol* 2001;63:57–94.
- [137] de Visser JAG, Hermisson J, Wagner GP, Meyers LA, Bagheri-Chaichian H, Blanchard JL, et al. Perspective: evolution and detection of genetic robustness. *Evolution* 2003;57:1959–72.
- [138] Lenski RE, Barrick JE, Ofria C. Balancing robustness and evolvability. *PLoS Biol* 2006;4:e428.
- [139] Masel J, Trotter MV. Robustness and evolvability. *Trends Genet* 2010;26:406–14.
- [140] Chen P, Shakhnovich EI. Lethal mutagenesis in viruses and bacteria. *Genetics* 2009;183:639–50.
- [141] Azevedo RB, Lohaus R, Srinivasan S, Dang KK, Burch CL. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature* 2006;440:87.
- [142] Hu T, Banzhaf W, Moore JH. The effects of recombination on phenotypic exploration and robustness in evolution. *Artif Life* 2014;20:457–70.
- [143] Huynen MA, Hogeweg P. Pattern generation in molecular evolution: exploitation of the variation in RNA landscapes. *J Mol Evol* 1994;39:71–9.
- [144] Singhal S, Gomez SM, Burch CL. Recombination drives the evolution of mutational robustness. *Curr Opin Syst Biol* 2019;13:142–9.
- [145] Szöllösi GJ, Derényi I. The effect of recombination on the neutral evolution of genetic robustness. *Math Biosci* 2008;214:58–62.
- [146] Xia Y, Levitt M. Roles of mutation and recombination in the evolution of protein thermodynamics. *Proc Natl Acad Sci USA* 2002;99:10382–7.
- [147] Klug A, Park S-C, Krug J. Recombination and mutational robustness in neutral fitness landscapes. *PLoS Comput Biol* 2019;15:e1006884.
- [148] Wolff A, Krug J. Robustness and epistasis in mutation-selection models. *Phys Biol* 2009;6:036007.
- [149] Green P, Ewing B, Miller W, Thomas PJ, Program NCS, Green ED. Transcription-associated mutational asymmetry in mammalian evolution. *Nat Genet* 2003;33:514–7. <https://doi.org/10.1038/ng1103>.
- [150] Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. *Genetics* 2000;156:297–304.
- [151] Paul S, Million-Weaver S, Chattopadhyay S, Sokurenko E, Merrih H. Accelerated gene evolution through replication-transcription conflicts. *Nature* 2013;495:512–5.
- [152] Mira A, Ochman H. Gene location and bacterial sequence divergence. *Mol Biol Evol* 2002;19:1350–8.
- [153] McCandlish DM, Stoltzfus A. Modeling evolution using the probability of fixation: history and implications. *Q Rev Biol* 2014;89:225–52.
- [154] Stoltzfus A, McCandlish DM. Mutational biases influence parallel adaptation. *Mol Biol Evol* 2017;34:2163–72.
- [155] Lozovsky ER, Chookajorn T, Brown KM, Imwong M, Shaw PJ, Kamchonwongpaisan S, et al. Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proc Natl Acad Sci USA* 2009;106:12025–30.
- [156] Rokytá DR, Joyce P, Caudle SB, Wichman HA. An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nat Genet* 2005;37:441–4.
- [157] Cano AV, Payne JL. Mutation bias interacts with composition bias to influence adaptive evolution. *PLoS Comput Biol* 2020;16:e1008296.
- [158] Sella G, Hirsh AE. The application of statistical physics to evolutionary biology. *Proc Natl Acad Sci USA* 2005;102:9541–6.
- [159] May RM. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* 1977;269:471.
- [160] Scheffer M, Carpenter S, Foley JA, Folke C, Walker B. Catastrophic shifts in ecosystems. *Nature* 2001;413:591.
- [161] Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. *J Comput Netw* 2010;2:203–71.
- [162] Gao J, Buldyrev SV, Havlin S, Stanley HE. Robustness of a network of networks. *Phys Rev Lett* 2011;107:195701.
- [163] Quill Elizabeth. When networks network: once studied solo, systems display surprising behavior when they interact. *Sci News* 2012;182:18–25.
- [164] Buldyrev SV, Parshani R, Gerald P, Stanley EH, Havlin S. Catastrophic cascade of failures in interdependent networks. *Nature* 2010;464:1025–8.
- [165] Aguirre J, Sevilla-Escoboza R, Gutiérrez R, Papo D, Buldú JM. Synchronization of interconnected networks: the role of connector nodes. *Phys Rev Lett* 2014;112:248701.
- [166] Gómez-Gardeñes J, Reinares I, Arenas A, Floria LM. Evolution of cooperation in multiplex networks. *Sci Rep* 2012;2:620.
- [167] Iranzo J, Buldú JM, Aguirre J. Competition among networks highlights the power of the weak. *Nat Commun* 2016;7:13273.
- [168] Buldú J, Pablo-Martí F, Aguirre J. Taming out-of-equilibrium dynamics on interconnected networks. *Nat Commun* 2019;10:5314.
- [169] Yubero P, Manrubia S, Aguirre J. The space of genotypes is a network of networks: implications for evolutionary and extinction dynamics. *Sci Rep* 2017;7:13813.
- [170] Aguirre J, Papo D, Buldú JM. Successful strategies for competing networks. *Nat Phys* 2013;9:230–4.
- [171] Newman MEJ. *Networks: an introduction*. New York: Oxford University Press; 2010.
- [172] Aguirre J, Manrubia S. Tipping points and early warning signals in the genomic composition of populations induced by environmental changes. *Sci Rep* 2015;5:9664.
- [173] Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. *Nat Rev Genet* 2007;8:610–8.
- [174] Barnosky AD, Hadly EA, Bascompte J, Berlow EL, Brown JH, Fortelius M, et al. Approaching a state shift in Earth's biosphere. *Nature* 2012;486:52–8.

- [175] Brook BW, Ellis EC, Perring MP, Mackay AW, Blomqvist L. Does the terrestrial biosphere have planetary tipping points? *Trends Ecol Evol* 2013;28:396–401.
- [176] Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, et al. Early-warning signals for critical transitions. *Nature* 2009;461:53.
- [177] Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, et al. Anticipating critical transitions. *Science* 2012;338:344–8.
- [178] Dai L, Vorselen D, Korolev KS, Gore J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* 2012;336:1175–7.
- [179] Fontana W. Modelling ‘evo-devo’ with RNA. *BioEssays* 2002;24:1164–77.
- [180] Huynen MA, Stadler PF, Fontana W. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc Natl Acad Sci USA* 1996;93:397–401.
- [181] Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* 2001;412:331–3.
- [182] Iwasa Y. Free fitness that always increases in evolution. *J Theor Biol* 1988;135:265–81.
- [183] Barton NH, Coe JB. On the application of statistical physics to evolutionary biology. *J Theor Biol* 2009;259:317–24.
- [184] Barton N, de Vladar H. Statistical mechanics and the evolution of polygenic quantitative traits. *Genetics* 2009;181:997–1011.
- [185] Khatri BS, Goldstein RA. A coarse-grained biophysical model of sequence evolution and the population size dependence of the speciation rate. *J Theor Biol* 2015;378:56–64.
- [186] Berg J, Willmann S, Lässig M. Adaptive evolution of transcription factor binding sites. *BMC Evol Biol* 2004;4:42.
- [187] Mustonen V, Lässig M. Evolutionary population genetics of promoters: predicting binding sites and functional phylogenies. *Proc Natl Acad Sci USA* 2005;102:15936–41.
- [188] Mustonen V, Kinney J, Callan CG, Lässig M. Energy-dependent fitness: a quantitative model for the evolution of yeast transcription factor binding sites. *Proc Natl Acad Sci USA* 2008;105:12376–81.
- [189] Haldane A, Manhart M, Morozov AV. Biophysical fitness landscapes for transcription factor binding sites. *PLoS Comput Biol* 2014;10:e1003683.
- [190] Khatri BS, Goldstein RA. Simple biophysical model predicts faster accumulation of hybrid incompatibilities in small populations under stabilizing selection. *Genetics* 2015;201:1525–37.
- [191] Tulchinsky AY, Johnson NA, Watt WB, Porter AH. Hybrid incompatibility arises in a sequence-based bioenergetic model of transcription factor binding. *Genetics* 2014;198:1155–66.
- [192] Dobzhansky T. Studies on hybrid sterility, II: localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* 1936;21:113–35.
- [193] Muller H. Isolating mechanisms, evolution and temperature. *Biol Symp* 1942;6:71–125.
- [194] Bateson W. Heredity and variation in modern lights. In: Seward AC, editor. *Darwin and modern science: essays in commemoration of the centenary of the birth of Charles Darwin and of the fiftieth anniversary of the publication of the origin of species*. Cambridge library collection–Darwin, evolution and genetics. Cambridge University Press; 2009. p. 85–101.
- [195] Fitzpatrick BM. Rates of evolution of hybrid inviability in birds and mammals. *Evolution* 2004;58:1865–70.
- [196] Stelkens RB, Young KA, Seehausen O. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution* 2010;64:617–33.
- [197] Cooper A, Penny D. Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. *Science* 1997;275:1109–13.
- [198] Coyne JA, Orr HA. The evolutionary genetics of speciation. *Philos Trans R Soc B* 1998;353:287–305.
- [199] Barraclough TG, Nee S. Phylogenetics and speciation. *Trends Ecol Evol* 2001;16:391–9.
- [200] Nee S. Inferring speciation rates from phylogenies. *Evolution* 2001;55:661–8.
- [201] Lande R. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution* 1979;33:234–51.
- [202] Lande R. Expected time for random genetic drift of a population between stable phenotypic states. *Proc Natl Acad Sci USA* 1985;82:7641–5.
- [203] Barton NH, Charlesworth B. Genetic revolutions, founder effects, and speciation. *Annu Rev Ecol Syst* 1984;15:133–64.
- [204] Barton N, Rouhani S. The frequency of shifts between alternative equilibria. *J Theor Biol* 1987;125:397–418.
- [205] Zeldovich KB, Chen P, Shakhnovich EI. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci USA* 2007;104:16152–7.
- [206] Závodszy P, Kardos J, Svingor A, Petsko GA. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc Natl Acad Sci USA* 1998;95:7406–11.
- [207] Taverna DM, Goldstein RA. Why are proteins marginally stable? *Proteins* 2002;46:105–9.
- [208] Bloom JD, Raval A, Wilke CO. Thermodynamics of neutral protein evolution. *Genetics* 2007;175:255–66.
- [209] Goldstein RA. The evolution and evolutionary consequences of marginal thermostability in proteins. *Proteins* 2011;79:1396–407.
- [210] Serohijos AWR, Rimas Z, Shakhnovich EI. Protein biophysics explains why highly abundant proteins evolve slowly. *Cell Rep* 2012;2:249–56.
- [211] Wylie CS, Shakhnovich EI. A biophysical protein folding model accounts for most mutational fitness effects in viruses. *Proc Natl Acad Sci USA* 2011;99:16–21.
- [212] Serohijos AWR, Lee SYR, Shakhnovich EI. Highly abundant proteins favor more stable 3D structures in yeast. *Biophys J* 2013;104:L01–3.
- [213] Khatri BS. Free fitness and the marginal stability of proteins, 2019, unpublished.
- [214] Matute DR, Coyne JA. Intrinsic reproductive isolation between two sister species of *Drosophila*. *Evolution* 2010;64:903–20.
- [215] Verster AJ, Ramani AK, McKay SJ, Fraser AG. Comparative RNAi screens in *C. elegans* and *C. briggsae* reveal the impact of developmental system drift on gene function. *PLoS Genet* 2014;10:e1004077.
- [216] Wotton KR, Jiménez-Guri E, Crombach A, Janssens H, Alcaine-Colet A, Lemke S, et al. Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. *eLife* 2015;4:e04785.
- [217] True JR, Haag ES. Developmental system drift and flexibility in evolutionary trajectories. *Evol Dev* 2001;3:109–19.

- [218] Haag ES. The same but different: worms reveal the pervasiveness of developmental system drift. *PLoS Genet* 2014;10:e1004150.
- [219] Johnson NA, Porter AH. Rapid speciation via parallel, directional selection on regulatory genetic pathways. *J Theor Biol* 2000;205:527–42.
- [220] Johnson NA, Porter AH. Toward a new synthesis: population genetics and evolutionary developmental biology. *Genetica* 2001;112:45–58.
- [221] McLeish TCB. Are there ergodic limits to evolution? Ergodic exploration of genome space and convergence. *Interface Focus* 2015;5:20150041.
- [222] Nourmohammad A, Schifffels S, Lässig M. Evolution of molecular phenotypes under stabilizing selection. *J Stat Mech Theory Exp* 2013;2013:P01012.
- [223] Khatri BS. Survival of the frequent at finite population size and mutation rate: filling the gap between quasispecies and monomorphic regimes. *BioRxiv* 2018. <https://doi.org/10.1101/375147>.
- [224] Mustonen V, Lässig M. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet* 2009;25:111–9.
- [225] Kent WJ, Baertsch R, Hinrichs A, Miller W, Haussler D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proc Natl Acad Sci USA* 2003;100:11484–9.
- [226] Gilbert W. The RNA world. *Nature* 1986;319:618.
- [227] Briones C, Stich M, Manrubia SC. The dawn of the RNA world: toward functional complexity through ligation of random RNA oligomers. *RNA* 2009;15:743–9.
- [228] Takeuchi N, Hogeweg P. Evolution of complexity in RNA-like replicator systems. *Biol Direct* 2008;3:11.
- [229] Colizzi ES, Hogeweg P. Evolution of functional diversification within quasispecies. *Genome Biol Evol* 2014;6:1990–2007.
- [230] Boerlijst MC, Hogeweg P. Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. *Physica D* 1991;48:17–28.
- [231] Takeuchi N, Hogeweg P. Multilevel selection in models of prebiotic evolution ii: a direct comparison of compartmentalization and spatial self-organization. *PLoS Comput Biol* 2009;5:e1000542.
- [232] Takeuchi N, Hogeweg P, Kaneko K. The origin of a primordial genome through spontaneous symmetry breaking. *Nat Commun* 2017;8:250.
- [233] von der Dunk S, Colizzi E, Hogeweg P. Evolutionary conflict leads to innovation: symmetry breaking in a spatial model of RNA-like replicators. *Life* 2017;7:43.
- [234] Cuypers TD, Hogeweg P. Virtual genomes in flux: an interplay of neutrality and adaptability explains genome expansion and streamlining. *Genome Biol Evol* 2012;4:212–29.
- [235] Cuypers TD, Hogeweg P. A synergism between adaptive effects and evolvability drives whole genome duplication to fixation. *PLoS Comput Biol* 2014;10:e1003547.
- [236] Cuypers TD, Rutten JP, Hogeweg P. Evolution of evolvability and phenotypic plasticity in virtual cells. *BMC Evol Biol* 2017;17:60.
- [237] Koonin EV. The biological Big Bang model for the major transitions in evolution. *Biol Direct* 2007;2:21.
- [238] Wloch DM, Szafraniec K, Borts RH, Korona R. Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. *Genetics* 2001;159:441–52.
- [239] Sanjuán R, Moya A, Elena SF. The contribution of epistasis to the architecture of fitness in an RNA virus. *Proc Natl Acad Sci USA* 2004;101:15376–9.
- [240] Yona AH, Manor YS, Herbst RH, Romano GH, Mitchell A, Kupiec M, et al. Chromosomal duplication is a transient evolutionary solution to stress. *Proc Natl Acad Sci USA* 2012;109:21010–5.
- [241] Zhang Z, Du C, de Barsey F, Liem M, Liakopoulos A, van Wezel GP, et al. Antibiotic production in *Streptomyces* is organized by a division of labor through terminal genomic differentiation. *Sci Adv* 2020;6:eay5781.
- [242] De Visser JAGM, Krug J. Empirical fitness landscapes and the predictability of evolution. *Nat Rev Genet* 2014;15:480.
- [243] Payne JL, Wagner A. The causes of evolvability and their evolution. *Nat Rev Genet* 2019;20:24–38.
- [244] Olson CA, Wu NC, Sun R. A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain. *Curr Biol* 2014;24:2643–51.
- [245] Pitt JN, Ferré-D'Amaré AR. Rapid construction of empirical RNA fitness landscapes. *Science* 2010;330:376–9.
- [246] Diss G, Lehner B. The genetic landscape of a physical interaction. *eLife* 2018;7:e32472.
- [247] Julien P, Miñana B, Baeza-Centurion P, Valcárcel J, Lehner B. The complete local genotype-phenotype landscape for the alternative splicing of a human exon. *Nat Commun* 2016;7:11558.
- [248] Schaerli Y, Munteanu A, Gili M, Cotterell J, Sharpe J, Isalan M. A unified design space of synthetic stripe-forming networks. *Nat Commun* 2014;5:4905.
- [249] Schaerli Y, Jiménez A, Duarte JM, Mihajlovic L, Renggli J, Isalan M, et al. Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. *Mol Syst Biol* 2018;14:e8102.
- [250] Bassalo MC, Garst AD, Choudhury A, Grau WC, Oh EJ, Spindler E, et al. Deep scanning lysine metabolism in *Escherichia coli*. *Mol Syst Biol* 2018;14:e8371.
- [251] Li C, Qian W, Maclean CJ, Zhang J. The fitness landscape of a tRNA gene. *Science* 2016;352:837–40.
- [252] Puchta O, Cseke B, Czaja H, Tollervey D, Sanguinetti G, Kudla G. Network of epistatic interactions within a yeast snoRNA. *Science* 2016;352:840–4.
- [253] Venkataram S, Dunn B, Li Y, Agarwala A, Chang J, Ebel ER, et al. Development of a comprehensive genotype-to-fitness map of adaptation-driving mutations in yeast. *Mol Syst Biol* 2016;167:1585–96.
- [254] Rotem A, Serohijos AW, Chang CB, Wolfe JT, Fischer AE, Mehoke TS, et al. Evolution on the biophysical fitness landscape of an RNA virus. *Mol Biol Evol* 2018;35:2390–400.
- [255] Weinreich DM, Delaney NF, DePristo MA, Hartl DL. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 2006;312:111–4.
- [256] Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. *Nat Methods* 2014;11:801–7.

- [257] Payne JL, Wagner A. The robustness and evolvability of transcription factor binding sites. *Science* 2014;343:875–7.
- [258] Adams RM, Kinney JB, Walczak AW, Mora T. Epistasis in a fitness landscape defined by antibody-antigen binding free energy. *Cell Syst* 2019;8:1–8.
- [259] Miho E, Roškar R, Greiff V, Reddy ST. Large-scale network analysis reveals the sequence space architecture of antibody repertoires. *Nat Commun* 2019;10:1321.
- [260] Hinkley T, Martins J, Chappey C, Haddad M, Stawiski E, Whitcomb JM, et al. A systems analysis of mutational effects in HIV-1 protease and reverse transcriptase. *Nat Genet* 2011;43:487–9.
- [261] Acevedo A, Brodsky L, Andino R. Mutational and fitness landscapes of an RNA virus revealed through population screening. *Nature* 2014;505:686–90.
- [262] Daguene E, Dujardin G, Valcárcel J. The pathogenicity of splicing effects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. *EMBO Rep* 2015;16:1640–55.
- [263] Baeza-Centurion P, Miñana B, Schmiedel JM, Valcárcel J, Lehner B. Combinatorial genetics reveals a scaling law for the effects of mutations on splicing. *Cell* 2019;176:1–15.
- [264] Payne JL, Khalid F, Wagner A. RNA-mediated gene regulation is less evolvable than transcriptional regulation. *Proc Natl Acad Sci USA* 2018;115:3481–90.
- [265] Elena SF, Lenski RE. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* 1997;390:395–8.
- [266] Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. Evidence for positive epistasis in hiv-1. *Science* 2004;306:1547–50.
- [267] Lalić J, Elena SF. Epistasis between mutations is host-dependent for an RNA virus. *Biol Lett* 2013;9:20120396.
- [268] Lunzer M, Miller SP, Felsheim R, Dean AM. The biochemical architecture of an ancient adaptive landscape. *Science* 2005;310:499–501.
- [269] Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 2007;445:383–6.
- [270] Dawid A, Kiviet DJ, Kogenaru M, de Vos M, Tans SJ. Multiple peaks and reciprocal sign epistasis in an empirically determined genotype-phenotype landscape. *Chaos* 2010;20:026105.
- [271] Chou H-H, Chiu H-C, Delaney NF, Segrè D, Marx CJ. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 2011;332:1190–2.
- [272] Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 2011;332:1193–6.
- [273] de Visser JA, Park S, Krug J. Exploring the effect of sex on empirical fitness landscapes. *Am Nat* 2009;174:S15–30.
- [274] Hall DW, Agan M, Pope SC. Fitness epistasis among 6 biosynthetic loci in the budding yeast *Saccharomyces cerevisiae*. *J Hered* 2010;101:S75–84.
- [275] da Silva J, Coetzer M, Nedellec R, Pastore C, Mosier DE. Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. *Genetics* 2010;185:293–303.
- [276] Kouyos RD, Leventhal GE, Hinkley T, Haddad M, Whitcomb JM, Petropoulos CJ, et al. Exploring the complexity of the HIV-1 fitness landscape. *PLoS Genet* 2012;8:e1002551.
- [277] da Silva J, Wyatt SK. Fitness valleys constrain hiv-1's adaptation to its secondary chemokine coreceptor. *J Evol Biol* 2014;27:604–15.
- [278] Wright S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In: *Proc. 6th int. Congr. genet.*, vol. 1; 1932. p. 356–66.
- [279] Weinreich DM, Watson RA, Chao L. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 2005;59:1165–74.
- [280] Poelwijk FJ, Tănase-Nicola S, Kiviet DJ, Tans SJ. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J Theor Biol* 2011;272:141–4.
- [281] Schaper S, Johnston IG, Louis AA. Epistasis can lead to fragmented neutral spaces and contingency in evolution. *Proc R Soc Lond B* 2011;279:1777–83.
- [282] Weinreich DM, Lan Y, Wylie CS, Heckendorn RB. Should evolutionary geneticists worry about higher-order epistasis? *Curr Opin Genet Dev* 2013;23:700–7.
- [283] Lalić J, Elena SF. Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity* 2012;109:71–7.
- [284] Carrasco P, de la Iglesia F, Elena SF. Distribution of fitness and virulence effects caused by single-nucleotide substitutions in tobacco etch virus. *J Virol* 2007;81:12979–84.
- [285] Sanjuán R. Quantifying antagonistic epistasis in a multifunctional RNA secondary structure of the Rous sarcoma virus. *J Gen Virol* 2006;87:1595–602.
- [286] Brouillet S, Annoni H, Ferretti L, Achaz G. Magellan: a tool to explore small fitness landscapes. *bioRxiv* 2015. <https://doi.org/10.1101/031583>. <https://www.biorxiv.org/content/early/2015/11/13/031583.full.pdf>.
- [287] Agudelo-Romero P, Carbonell P, Perez-Amador MA, Elena SF. Virus adaptation by manipulation of host's gene expression. *PLoS ONE* 2008;3:1–7.
- [288] Lalić J, Elena SF. The impact of high-order epistasis in the within-host fitness of a positive-sense plant RNA virus. *J Evol Biol* 2015;28:2236–47.
- [289] Frank RR, Poelwijk J, Krishna Vinod. The context-dependence of mutations: a linkage of formalisms. *PLoS Comput Biol* 2016;12:e1004771.
- [290] Cervera H, Lalić J, Elena SF. Efficient escape from local optima in a highly rugged fitness landscape by evolving RNA virus populations. *Proc R Soc Lond B* 2016;283:20160984.
- [291] Bedhomme S, Hillung J, Elena SF. Emerging viruses: why they are not jacks of all trades? *Curr Opin Virol* 2015;10:1–6.
- [292] Cervera H, Lalić J, Elena SF. Effect of host species on topography of the fitness landscape for a plant RNA virus. *J Virol* 2016;90:10160–9.
- [293] Catalán P, Arias CF, Cuesta JA, Manrubia S. Adaptive multiscales: an up-to-date metaphor to visualize molecular adaptation. *Biol Direct* 2017;12:7.
- [294] Domingo J, Baeza-Centurion P, Lehner B. The causes and consequences of genetic interactions (epistasis). *Annu Rev Genomics Hum Genet* 2019;20:433–60.

- [295] Kondrashov FA, Kondrashov AS. Multidimensional epistasis and the disadvantage of sex. *Proc Natl Acad Sci USA* 2001;98:12089–92.
- [296] Milkman R. Selection differentials and selection coefficients. *Genetics* 1978;88:391–403.
- [297] Otwinowski J, McCandlish DM, Plotkin JB. Inferring the shape of global epistasis. *Proc Natl Acad Sci USA* 2018;115:E7550–8.
- [298] Rokyta DR, Joyce P, Caudle SB, Miller C, Beisel CJ, Wichman HA. Epistasis between beneficial mutations and the phenotype-to-fitness map for a ssDNA virus. *PLoS Genet* 2011;7:e1002075.
- [299] Zwart MP, Schenk MF, Hwang S, Koopmanschap B, de Lange N, van de Pol L, et al. Unraveling the causes of adaptive benefits of synonymous mutations in TEM-1  $\beta$ -lactamase. *Heredity* 2018;121:406–21.
- [300] Schenk MF, Szendro IG, Salverda ML, Krug J, de Visser JAGM. Patterns of epistasis between beneficial mutations in an antibiotic resistance gene. *Mol Biol Evol* 2013;30:1779–87.
- [301] Manhart M, Morozov AV. Protein folding and binding can emerge as evolutionary spandrels through structural coupling. *Proc Natl Acad Sci USA* 2015;112:1797–802.
- [302] Blanquart F, Achaz G, Bataillon T, Tenaillon O. Properties of selected mutations and genotypic landscapes under Fisher's geometric model. *Evolution* 2014;68:3537–54.
- [303] Fisher RA. The genetical theory of natural selection. Oxford: Clarendon Press; 1930.
- [304] Tenaillon O. The utility of Fisher's geometric model in evolutionary genetics. *Annu Rev Ecol Syst* 2014;45:179–201.
- [305] Martin G, Elena SF, Lenormand T. Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat Genet* 2007;39:555–60.
- [306] Hwang S, Park S-C, Krug J. Genotypic complexity of Fisher's geometric model. *Genetics* 2017;206:1049–79.
- [307] Blanquart F, Bataillon T. Epistasis and the structure of fitness landscapes: are experimental fitness landscapes compatible with Fisher's geometric model? *Genetics* 2016;203:847–62.
- [308] Schoustra S, Hwang S, Krug J, de Visser JAGM. Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. *Proc R Soc Lond B* 2016;283:20161376.
- [309] Weinreich DM, Knies JL. Fisher's geometric model of adaptation meets the functional synthesis: data on pairwise epistasis for fitness yields insights into the shape and size of phenotype space. *Evolution* 2013;67:2957–72.
- [310] Pokusaeva VO, Usmanova DR, Putintseva EV, Espinar L, Sarkisyan KS, Mishin AS, et al. An experimental assay of the interactions of amino acids from orthologous sequences shaping a complex fitness landscape. *PLoS Genet* 2019;15:e1008079.
- [311] Bershtein S, Serohijos AW, Shakhnovich EI. Bridging the physical scales in evolutionary biology: from protein sequence space to fitness of organisms and populations. *Curr Opin Struct Biol* 2017;42:31–40.
- [312] Brophy JA, Voigt CA. Principles of genetic circuit design. *Nat Methods* 2014;11:508–20.
- [313] Gasperini M, Starita L, Shendure J. The power of multiplexed functional analysis of genetic variants. *Nat Protoc* 2016;11:1782–7.
- [314] Appleton E, Madsen C, Roehner N, Densmore D. Design automation in synthetic biology. *Cold Spring Harb Perspect Biol* 2017;9:a023978.
- [315] Woodruff L, Gorochowski TE, Roehner N, Mikkelsen T, Densmore D, Gordon DB, et al. Registry in a tube: multiplexed pools of retrievable parts for genetic design space exploration. *Nucleic Acids Res* 2017;45:1553–65.
- [316] Gorochowski TE, van den Berg E, Kerkman R, Roubos JA, Bovenberg RAL. Using synthetic biological parts and microbioreactors to explore the protein expression characteristics of *Escherichia coli*. *ACS Synth Biol* 2014;3:129–39.
- [317] Plesa C, Sidore AM, Lubock NB, Zhang D, Kosuri S. Multiplexed gene synthesis in emulsions for exploring protein functional landscapes. *Science* 2018;359:343–7.
- [318] Kosuri S, Church GM. Large-scale de novo DNA synthesis: technologies and applications. *Nat Methods* 2014;11:499–507.
- [319] Blount BA, Gowers G-OF, Ho JCH, Ledesma-Amaro R, Jovicevic D, McKiernan RM, et al. Rapid host strain improvement by in vivo rearrangement of a synthetic yeast chromosome. *Nat Commun* 2018;9:1932.
- [320] Cambray G, Guimaraes JC, Arkin AP. Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in *Escherichia coli*. *Nat Biotechnol* 2018;36:1005–15.
- [321] Gorochowski TE, Chelysheva I, Eriksen M, Nair P, Pedersen S, Ignatova Z. Absolute quantification of translational regulation and burden using combined sequencing approaches. *Mol Syst Biol* 2019;15:e8719.
- [322] Gorochowski TE, Espah Borujeni A, Park Y, Nielsen AAK, Zhang J, Der BS, et al. Genetic circuit characterization and debugging using RNA-seq. *Mol Syst Biol* 2017;13:952.
- [323] Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y, Arkin AP, et al. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc Natl Acad Sci USA* 2013;110:14024–9.
- [324] Johns NI, Gomes ALC, Yim SS, Yang A, Blazewski T, Smillie CS, et al. Metagenomic mining of regulatory elements enables programmable species-selective gene expression. *Nat Methods* 2018;15:323–9.
- [325] Liszczak G, Muir TW. Nucleic acid-barcoding technologies: converting DNA sequencing into a broad-spectrum molecular counter. *Angew Chem, Int Ed Engl* 2019;58:4144–62.
- [326] Gorochowski TE, Ellis T. Designing efficient translation. *Nat Biotechnol* 2018;26:934–5.
- [327] Der BS, Glassey E, Bartley BA, Enghuus C, Goodman DB, Gordon DB, et al. DNAplotlib: programmable visualization of genetic designs and associated data. *ACS Synth Biol* 2017;6:1115–9.
- [328] Hughes RA, Ellington AD. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb Perspect Biol* 2017;9:a023812.
- [329] Komura R, Aoki W, Motone K, Satomura A, Ueda M. High-throughput evaluation of T7 promoter variants using biased randomization and DNA barcoding. *PLoS ONE* 2018;13:e0196905.
- [330] Patwardhan RP, Lee C, Litvin O, Young DL, Pe'er D, Shendure J. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat Biotechnol* 2009;27:1173–5.
- [331] Cozens C, Pinheiro VB. Darwin assembly: fast, efficient, multi-site bespoke mutagenesis. *Nucleic Acids Res* 2018;46:e51.

- [332] Holmqvist E, Reimegård J, Wagner EGH. Massive functional mapping of a 5'-UTR by saturation mutagenesis, phenotypic sorting and deep sequencing. *Nucleic Acids Res* 2013;41:e122.
- [333] de Boer FK, Vaishnav ED, Sadeh R, Abeyta EL, Friedman N, Regev A. Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. *Nat Biotechnol* 2020;38:56–65.
- [334] Church GM, Kieffer-Higgins S. Multiplex DNA sequencing. *Science* 1988;240:185–8.
- [335] Cuperus JT, Groves B, Kuchina A, Rosenberg AB, Jojic N, Fields S, et al. Deep learning of the regulatory grammar of yeast 5' untranslated regions from 500,000 random sequences. *Genome Res* 2017;27:2015–24.
- [336] Bull JJ, Barrick JE. Arresting evolution. *Trends Genet* 2017;33:910–20.
- [337] Sleight SC, Sauro HM. Visualization of evolutionary stability dynamics and competitive fitness of *Escherichia coli* engineered with randomized multigene circuits. *ACS Synth Biol* 2013;2:519–28.
- [338] Yus E, Yang JS, Sogues A, Serrano L. A reporter system coupled with high-throughput sequencing unveils key bacterial transcription and translation determinants. *Nat Commun* 2017;8:368.
- [339] Sze JYY, Ivanov AP, Cass AEG, Edel JB. Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. *Nat Commun* 2017;8:1552.
- [340] van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. *Trends Genet* 2018;34:666–81.
- [341] Chari R, Church GM. Beyond editing to writing large genomes. *Nat Rev Genet* 2017;18:749–60.
- [342] Crona K, Greene D, Barlow M. The peaks and geometry of fitness landscapes. *J Theor Biol* 2013;317:1–10.
- [343] Diaz-Uriarte R, Vasallo C. Every which way? On predicting tumor evolution using cancer progression models. *PLoS Comput Biol* 2019;15:1–29.
- [344] Diaz-Uriarte R. Identifying restrictions in the order of accumulation of mutations during tumor progression: effects of passengers, evolutionary models, and sampling. *BMC Bioinform* 2015;16.
- [345] Gerstung M, Eriksson N, Lin J, Vogelstein B, Beerenwinkel N. The temporal order of genetic and pathway alterations in tumorigenesis. *PLoS ONE* 2011;6:1–9.
- [346] Beerenwinkel N, Schwarz RF, Gerstung M, Markowitz F. Cancer evolution: mathematical models and computational inference. *Syst Biol* 2014;64:e1–25.
- [347] Beerenwinkel N, Greenman CD, Lagergren J. Computational cancer biology: an evolutionary perspective. *PLoS Comput Biol* 2016;12:e1004717.
- [348] Hosseini S-R, Diaz-Uriarte R, Markowitz F, Beerenwinkel N. Estimating the predictability of cancer evolution. *Bioinformatics* 2019;35: i389–i397.
- [349] Toprak E, Veres A, Michel J-B, Chait R, Hartl DL, Kishony R. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat Genet* 2012;44:101–5.
- [350] Szabo A, Boucher KM. Oncogenetic trees. In: Tan W-Y, Hanin L, editors. *Handbook of cancer models with applications*. World Scientific; 2008. p. 1–24.
- [351] Desper R, Jiang F, Kallioniemi O-P, Moch H, Papadimitriou CH, Schäffer AA. Inferring tree models for oncogenesis from comparative genome hybridization data. *J Comput Biol* 1999;6:37–51.
- [352] Gerstung M, Baudis M, Moch H, Beerenwinkel N. Quantifying cancer progression with conjunctive Bayesian networks. *Bioinformatics* 2009;25:2809–15.
- [353] Montazeri H, Kuipers J, Kouyos R, Böni J, Yerly S, Klimkait T, et al. Large-scale inference of conjunctive Bayesian networks. *Bioinformatics* 2016;32: i727–i735.
- [354] Ramazzotti D, Caravagna G, Olde Loohuis L, Graudenzi A, Korsunsky I, Mauri G, et al. CAPRI: efficient inference of cancer progression models from cross-sectional data. *Bioinformatics* 2015;31:3016–26.
- [355] Caravagna G, Graudenzi A, Ramazzotti D, Sanz-Pamplona R, De Sano L, Mauri G, et al. Algorithmic methods to infer the evolutionary trajectories in cancer progression. *Proc Natl Acad Sci USA* 2016;113:E4025–34.
- [356] Frank SA, Rosner MR. Nonheritable cellular variability accelerates the evolutionary processes of cancer. *PLoS Biol* 2012;10:e1001296.
- [357] Altenberg L. Modularity in evolution: some low-level questions. In: Callebaut W, Rasskin-Gutman D, Simon HA, editors. *Modularity: understanding the development and evolution of complex natural systems*. Cambridge, MA: MIT Press; 2004.
- [358] DeGregori JD. *Adaptive oncogenesis: a new understanding of how cancer evolves inside US*. Cambridge, MA: Harvard University Press; 2018.
- [359] Gavrillets S. Evolution and speciation on holey adaptive landscapes. *Trends Ecol Evol* 1997;12:307–12.
- [360] McFarland CD, Yaglom JA, Wojtkowiak JW, Scott JG, Morse DL, Sherman MY, et al. The damaging effect of passenger mutations on cancer progression. *Cancer Res* 2017;77:4763–72.
- [361] Heng HH. The genomic landscape of cancers. In: Ujvari B, Roche B, Thomas F, editors. *Ecology and evolution of cancer*. London: Academic Press; 2017. p. 69–86.
- [362] Kokko H, Schindler S, Sprouffske K. Chapter 22—Searching for a cancer-proof organism: it's the journey that teaches you about the destination. In: Ujvari B, Roche B, Thomas F, editors. *Ecology and evolution of cancer*. London: Academic Press; 2017. p. 247–54.
- [363] De Vos MG, Dawid A, Sunderlikova V, Tans SJ. Breaking evolutionary constraint with a tradeoff ratchet. *Proc Natl Acad Sci USA* 2015;112:14906–11.
- [364] Steinberg B, Ostermeier M. Environmental changes bridge evolutionary valleys. *Sci Adv* 2016;2:e1500921.
- [365] Li C, Zhang J. Multi-environment fitness landscapes of a tRNA gene. *Nat Ecol Evol* 2018;2:1025–32.
- [366] Gorter FA, Aarts MGM, Zwaan BJ, de Visser JAGM. Local fitness landscapes predict yeast evolutionary dynamics in directionally changing environments. *Genetics* 2018;208:307–22.
- [367] Zhang Y, Ptacin JL, Fischer EC, Aerni HR, Caffaro CE, San Jose K, et al. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* 2017;551:644–7.

- [368] Fredens J, Wang K, de la Torre D, Funke LFH, Robertson WE, Christova Y, et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 2019;569:514–8.
- [369] Hoshika S, Leal NA, Kim M-J, Kim M-S, Karalkar NB, Kim H-J, et al. Hachimoji DNA and RNA: a genetic system with eight building blocks. *Science* 2019;363:884–7.
- [370] Otwinowski J, Plotkin J. Inferring fitness landscapes by regression produces biased estimates of epistasis. *Proc Natl Acad Sci USA* 2014;111:E2301–9.
- [371] du Plessis L, Leventhal GE, Bonhoeffer S. How good are statistical models at approximating complex fitness landscapes? *Mol Biol Evol* 2016;33:2454–68.
- [372] Riesselman AJ, Ingraham JB, Marks DS. Deep generative models of genetic variation capture the effects of mutations. *Nat Methods* 2018;15:816–22.
- [373] Chen B-S, Hsu C-Y, Liou J-J. Robust design of biological circuits: evolutionary systems biology approach. *J Biomed Biotechnol* 2011;2011:304236.
- [374] Borenstein E, Krakauer DC. An end to endless forms: epistasis, phenotype distribution bias, and nonuniform evolution. *PLoS Comput Biol* 2008;4:e1000202.
- [375] Hu T, Payne JL, Banzhaf W, Moore JH. Evolutionary dynamics on multiple scales: a quantitative analysis of the interplay between genotype, phenotype, and fitness in linear genetic programming. *Genet Program Evol Mach* 2012;13:305–37.
- [376] Lind PA, Farr AD, Rainey PB. Experimental evolution reveals hidden diversity in evolutionary pathways. *eLife* 2015;4:e07074.
- [377] Lässig M, Mustonen V, Walczak AM. Predicting evolution. *Nat Ecol Evol* 2017;1:0077.
- [378] Milocco L, Salazar-Ciudad I. Is evolution predictable? Quantitative genetics under complex genotype-phenotype maps. *Evolution* 2019;74:230–44.
- [379] Nosil P, Flaxman SM, Feder JL, Gompert Z. Increasing our ability to predict contemporary evolution. *Nat Commun* 2020;11:5529.
- [380] Paixão T, Bassler K, Azevedo RBR. Emergent speciation by multiple Dobzhansky-Muller incompatibilities. *bioRxiv* 2015. <https://doi.org/10.1101/008268>. <https://www.biorxiv.org/content/10.1101/008268v3>.
- [381] Altenberg L. Genome growth and the evolution of the genotype-phenotype map. In: Banzhaf W, Eeckman FH, editors. *Evolution and biocomputation: computational models of evolution*. Lecture notes in computer science, vol. 899. Berlin: Springer-Verlag; 1995. p. 205–59.
- [382] Orgogozo V, Morizot B, Martin A. The differential view of genotype-phenotype relationships. *Front Genet* 2015;6:179.
- [383] Wagner GP, Zhang J. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet* 2011;12:204–11.
- [384] Espinosa-Soto C. On the role of sparseness in the evolution of modularity in gene regulatory networks. *PLoS Comput Biol* 2018;14:e1006172.
- [385] Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017;169:1177–86.
- [386] Bottou L. Large-scale machine learning with stochastic gradient descent. In: Lechevallier Y, Saporta G, editors. *Proceedings of COMPSTAT 2010*. Berlin: Springer; 2010. p. 177–86.
- [387] Draghi J, Wagner GP. Evolution of evolvability in a developmental model. *Evolution* 2008;62:301–15.
- [388] Meyers LA, Ancel FD, Lachmann M. Evolution of genetic potential. *PLoS Comput Biol* 2005;1:e32.
- [389] Kashtan N, Alon U. Spontaneous evolution of modularity and network motifs. *Proc Natl Acad Sci USA* 2005;102:13773–8.
- [390] Crombach A, Hogeweg P. Chromosome rearrangements and the evolution of genome structuring and adaptability. *Mol Biol Evol* 2007;24:1130–9.
- [391] Crombach A, Hogeweg P. Evolution of evolvability in gene regulatory networks. *PLoS Comput Biol* 2008;4:e1000112.
- [392] Yang J, Lusk R, Li W-H. Organismal complexity, protein complexity, and gene duplicability. *Proc Natl Acad Sci USA* 2003;100:15661–5.
- [393] Chen C-Y, Ho A, Huang H-Y, Juan H-F, Huang H-C. Dissecting the human protein-protein interaction network via phylogenetic decomposition. *Sci Rep* 2014;4:7153.
- [394] Mattenberger F, Sabater-Muñoz B, Toft C, Sablok G, Fares MA. Expression properties exhibit correlated patterns with the fate of duplicated genes, their divergence, and transcriptional plasticity in *Saccharomycotina*. *DNA Res* 2017;24:559–70.
- [395] O'toole ÁN, Hurst LD, McLysaght A. Faster evolving primate genes are more likely to duplicate. *Mol Biol Evol* 2017;35:107–18.
- [396] Banerjee S, Feyertag F, Alvarez-Ponce D. Intrinsic protein disorder reduces small-scale gene duplicability. *DNA Res* 2017;24:435–44.
- [397] Li Z, Defoort J, Tasdighian S, Maere S, Van de Peer Y, De Smet R. Gene duplicability of core genes is highly consistent across all angiosperms. *Plant Cell* 2016;28:326–44.
- [398] Altenberg L. The evolution of evolvability in genetic programming. In: Kinnear KE, editor. *Advances in genetic programming*. Cambridge, MA: MIT Press; 1994. p. 47–74.
- [399] Altenberg L. Evolving better representations through selective genome growth. In: Fogel DB, editor. *Proceedings of the 1st IEEE conference on evolutionary computation*. Piscataway, NJ: IEEE; 1994. p. 182–7. Part 1 (of 2).
- [400] Altenberg L. Emergent phenomena in genetic programming. In: Sebald A, Fogel L, editors. *Evolutionary programming — proceedings of the third annual conference*. World Scientific Publishing. ISBN 9810218109, 1994. p. 233–41.
- [401] Whitacre J, Bender A. Degeneracy: a design principle for achieving robustness and evolvability. *J Theor Biol* 2010;263:143–53.
- [402] McShea DW, Brandon RN. *Biology's first law: the tendency for diversity and complexity to increase in evolutionary systems*. Chicago: University of Chicago Press; 2010.
- [403] Hughes AL. The evolution of functionally novel proteins after gene duplication. *Proc R Soc Lond B* 1994;256:119–24.
- [404] Stoltzfus A. On the possibility of constructive neutral evolution. *J Mol Evol* 1999;49:169–81.
- [405] Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 1999;151:1531–45.
- [406] Barton N, Etheridge A, Véber A. The infinitesimal model: definition, derivation, and implications. *Theor Popul Biol* 2017;118:50–73.
- [407] Livnat A, Papadimitriou C, Dushoff J, Feldman M. A mixability theory for the role of sex in evolution. *Proc Natl Acad Sci USA* 2008;105:19803–8.

- [408] Livnat A, Papadimitriou C, Pippenger N, Feldman M. Sex, mixability, and modularity. *Proc Natl Acad Sci USA* 2010;107:1452–7.
- [409] Conrad M. The importance of molecular hierarchy in information processing. *Towards Theor Biol* 1972;4:222–8.
- [410] Conrad M. Bootstrapping on the adaptive landscape. *BioSystems* 1979;11:167–82.