

Phenotypes can be robust and evolvable if mutations have non-local effects on sequence constraints

Marcel Weiß^{1,2*} and Sebastian E. Ahnert^{1,2}

¹*Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, Cambridge, UK*

²*Sainsbury Laboratory, University of Cambridge, Cambridge, UK*

(Dated: December 5, 2017)

The mapping between biological genotypes and phenotypes plays an important role in evolution, and understanding the properties of this mapping is crucial to determine the outcome of evolutionary processes. One of the most striking properties observed in several genotype-phenotype (GP) maps is the positive correlation between the robustness and evolvability of phenotypes. This implies that a phenotype can be strongly robust against mutations and at the same time evolvable to a diverse range of alternative phenotypes. Here, we examine the causes for this positive correlation by introducing two analytically tractable GP map models that follow the principles of real biological GP maps. The first model is based on gene-like GP maps, reflecting the way in which genetic sequences are organised into protein-coding genes, and the second one is based on the GP map of RNA secondary structure. For both models, we find that a positive correlation between phenotype robustness and evolvability only emerges if mutations at one sequence position can have non-local effects on the sequence constraints at another position. This highlights that non-local effects of mutations are closely related to the coexistence of robustness and evolvability in phenotypes, and are likely to be an important feature of many biological GP maps.

I. INTRODUCTION

Biological examples of genotype-phenotype (GP) maps are extremely wide-ranging. For example, a GP map can refer to the mapping between DNA sequences and their encoded proteins or between RNA sequences and their folded spatial structures. Even GP maps of networks have been considered, like that of gene regulatory [1] or metabolic networks [2]. The properties of GP maps have been shown to strongly influence evolutionary outcomes [3, 4], particularly in systems without direct interactions between genotypes, or epigenetic effects. In general, studying the properties of real biological GP maps is a challenging task due to the extraordinary complexity and size of these maps, which means that only small parts can be captured experimentally. Progress has been made by focusing on abstract or computational models of GP maps, which can be studied exhaustively. One of them is the map between RNA sequences and their secondary structure, a tractable abstraction of the final spatial structure. There exist computational models [5, 6] that can predict the secondary structure of small RNA sequences, thereby allowing a full coverage of the GP map. Another example is the map between amino acid sequences and their folded tertiary structure predicted by the HP model [7, 8], a lattice model that approximates amino acids as being either hydrophobic (h) or polar (p). An even more abstract model is the so-called Polyomino model [9, 10], for which the map between different tile sets and the resulting assembled structures has been studied [11] as a representative of the formation of protein complexes (quaternary structure) by individual proteins.

The analysis of these GP maps [11–13] has shown that they share several properties which are likely to be universal for all GP maps. All maps show redundancy, meaning that there are much more genotypes than phenotypes and that different genotypes can map to the same phenotype. In addition, the maps show a skewed distribution for the number of genotypes per phenotype, also referred to as phenotype bias, meaning that a few phenotypes are mapped by a large number of genotypes, while many of the phenotypes are mapped by just a few genotypes. A further property shared by the maps is so-called shape space covering, which means that from any genotype of the GP map, a large fraction of all phenotypes can be accessed through just a few mutation steps. This property has been first described for RNA secondary structures by Fontana et al. [14].

A number of studies have discussed the concepts of robustness and evolvability in the context of GP maps. For the RNA secondary structure [11, 15] and Polyomino [11] GP maps, it has been found that robustness and evolvability are negatively correlated for genotypes, whereas they are positively correlated for phenotypes. This positive correlation has also been found recently for the GP map of transcription factor binding sites [16], which can be both robust and evolvable. In addition, several studies [11, 13, 17] have shown that the robustness of a phenotype scales logarithmically with the number of genotypes mapping to that phenotype.

Quantitative definitions of robustness and evolvability have been put forward by Wagner in the context of the RNA secondary structure GP map [15]. For a genotype, the robustness is defined as the fraction of mutations that do not change the phenotype of that genotype, and the evolvability as the number of distinct alternative phenotypes that are accessible through a single mutation from that genotype. For a phenotype, the robustness is

* mw636@cam.ac.uk

defined as the average over the robustness values of all genotypes mapping to that phenotype, and the evolvability as the number of distinct alternative phenotypes that are accessible through a single mutation from any of the genotypes mapping to that phenotype. The negative correlation found for genotypes is intuitive as a genotype cannot be both robust and evolvable. The positive correlation found for phenotypes is less intuitive, but crucial for evolution as it is an advantage if a phenotype on the one hand is robust against mutations but on the other hand also open to a large number of alternative phenotypes.

The space of all genotypes can be represented as a network in which each node corresponds to a genotype and each link to a single mutation step such that genotypes differing by a single mutation are connected. Since Maynard Smith and Eigen, this representation is referred to as the protein space [18] or sequence space [19]. The subset of genotypes mapping to a same phenotype is often referred to as a neutral network since mutations between genotypes in this network do not change the phenotype. Wagner [15] states two causes for a phenotype to have a high robustness and evolvability. Firstly, a large connected neutral network as this implies a high connectivity among genotypes with same phenotype and thus a high robustness. Secondly, the fact that different genotypes sampled from a neutral network have different alternative phenotypes in their neighbourhood as this implies a high evolvability if a large neutral network spans the space of genotypes.

Recently, an analytically tractable GP map model has been introduced that exhibits all the introduced properties found for GP maps [20]. This so-called Fibonacci model defines the set of genotypes as all possible binary sequences of fixed length. The phenotype of a genotype is defined as the part of the genotype sequence up to and including the first stop codon, which is a pre-defined sequence - in this case two consecutive ones, i.e. ‘11’. This divides a genotype sequence into a constrained and unconstrained part as mutations of the sequence positions preceding the stop codon change the phenotype, whereas mutations of the sequence positions following the stop codon have no effect on the phenotype. The authors state that it is this organisation into constrained and unconstrained parts that gives rise to the observed properties. In particular, the authors show that the positive correlation between phenotype robustness and evolvability results from the mutability of the stop codon as this can shift the boundary between the constrained and unconstrained part.

In this article, we further examine the possible causes for the positive correlation between phenotype robustness and evolvability. We do this by introducing two analytically tractable models. The first model follows the principles of gene-like GP maps and can be seen as a generalisation of the Fibonacci model. The second model follows the principles of sequence-to-structure GP maps like that of RNA secondary structure. By using reference

models in both cases we show that the positive correlation between phenotype robustness and evolvability depends on the possibility that a mutation in one sequence position changes the sequence constraints at another position - or more concisely, that mutations have non-local effects on sequence constraints.

II. GENE-LIKE GP MAP MODEL

The gene-like GP map model is based on the organisation of genetic sequences into protein-coding genes (genotypes) that are transcribed and translated into proteins (a form of phenotype).

The model works with an alphabet consisting of $k > 2$ letters, of which we choose one to be the stop codon, leaving $k - 1$ neutral letters. As the genotypes, we define all possible sequences of length L built from the letters of the alphabet, giving k^L genotypes in total. In analogy to the Fibonacci model [20], we define the phenotype of a genotype as the part of the genotype sequence up to and including the first stop codon (reading from left to right). We also refer to this part of the genotype sequence as the coding sequence. All genotypes that do not include a stop codon are assigned to a single so-called ‘undefined’ phenotype.

Within the genotype sequences, all letters are mutable, allowing changes between phenotypes. In order to determine the effect of a mutable stop codon, we consider a gene-like reference model that is identical to the gene-like model but that considers the first stop codon in a genotype sequence as rigid and allows no mutations of it. Both models are summarised in Figure 1.

In the following, we will analyse and compare key properties of the GP maps of both models. Figure 2 shows plots of these properties determined by an exhaustive enumeration of the GP maps.

A. Genotype robustness and evolvability

Using the definitions introduced by Wagner, genotype robustness and evolvability slightly differ for the gene-like model and its reference model.

For the gene-like model, consider a genotype with coding sequence length l , its robustness $r_g(l)$ is given by:

$$r_g(l) = \frac{(k-1)(L-l)}{(k-1)L} = 1 - \frac{l}{L} \quad (1)$$

as there are $(k-1)L$ total possible one-point mutations of which the $(k-1)(L-l)$ ones of the letters following the first stop codon do not change the phenotype. The gene-like reference model allows no mutations of the first stop codon. Thus, there are only $(k-1)(L-1)$ possible one-point mutations in total and the robustness of a genotype is given by:

$$r_g^{\text{ref}}(l) = \frac{(k-1)(L-l)}{(k-1)(L-1)} = \frac{L}{L-1} - \frac{l}{L-1} \quad (2)$$

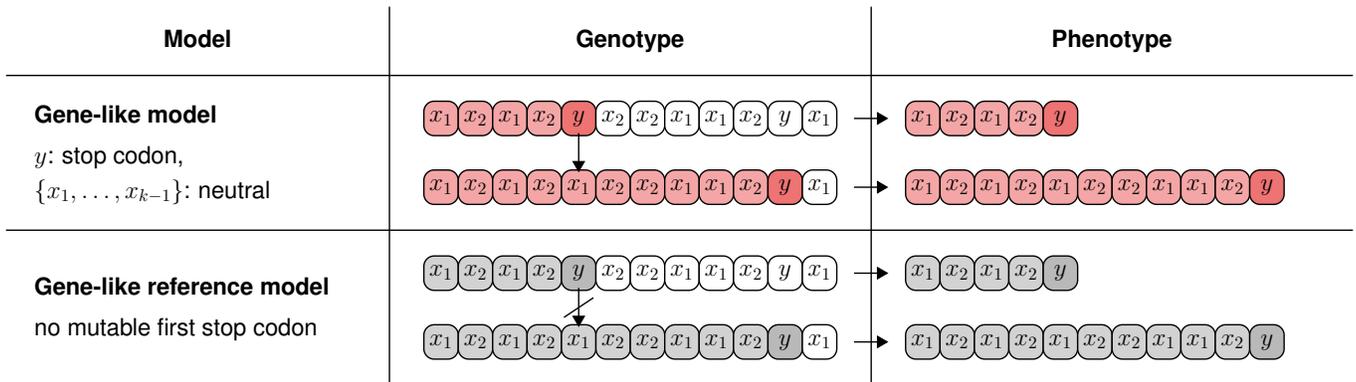


FIG. 1. Summary of the gene-like GP map model and its reference model together with examples of genotypes and their phenotypes. For both models, the same genotypes are shown. These two genotypes differ by the letter of the fifth site, which is a stop codon in the first genotype. The phenotype of a genotype is defined as the genotype sequence up to and including the first stop codon. In the gene-like model, this stop codon is mutable, allowing a transition from the first to the second genotype. In contrast, in the reference model, the first stop codon is not mutable.

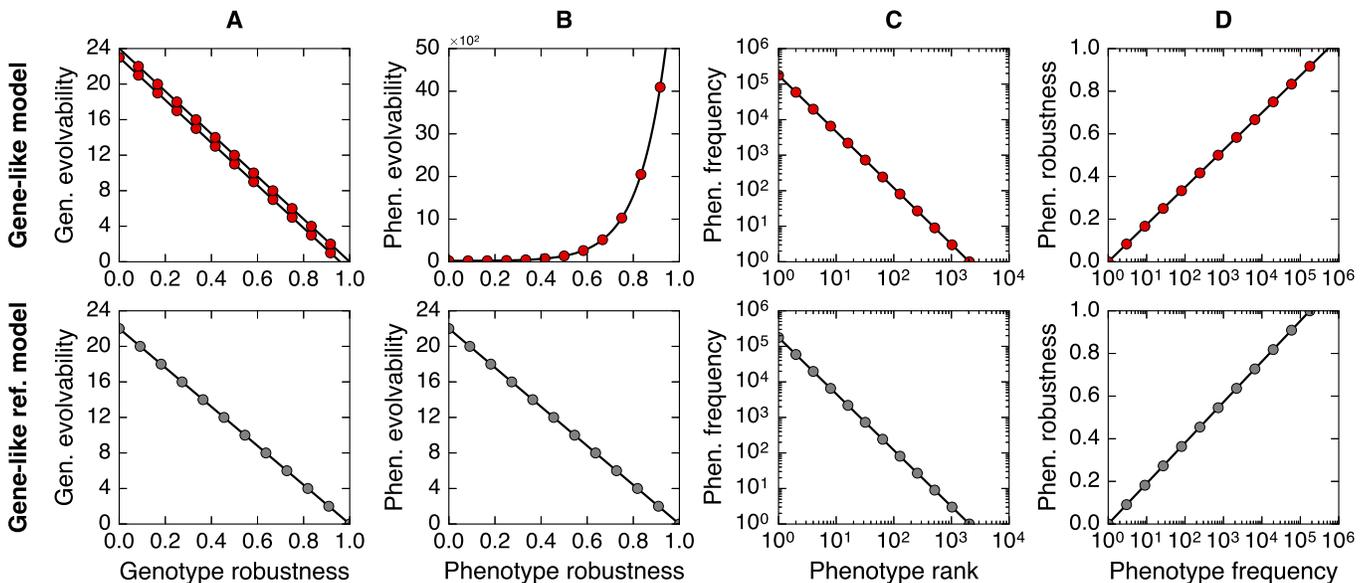


FIG. 2. Properties of the gene-like GP map model (TOP) and the gene-like reference model (BOTTOM), determined by an exhaustive enumeration of the GP maps. In all cases, an alphabet size of $k = 3$ and a genotype sequence length of $L = 12$ is chosen. Lines show analytical solutions. Both models qualitatively agree for all properties except for the correlation between phenotype evolvability and robustness (B), which is positive for the gene-like model (B, TOP), and negative for the reference model (B, BOTTOM). This highlights the importance of the mutable stop codon (and the resulting non-local effects of mutations) for the coexistence of robustness and evolvability in phenotypes. Note: The plots do not show the properties of the undefined phenotype (B-D) and the properties of its corresponding genotypes (A).

The robustness of genotypes with an undefined phenotype is not covered by these equations. Their robustness is in both cases given by $r_{g,\text{undef}} = r_{g,\text{undef}}^{\text{ref}} = (k-2)/(k-1)$.

For both models, all $(k-1)(l-1)$ one-point mutations of the letters preceding the first stop codon lead to distinct alternative phenotypes. In addition, only for the gene-like model, all $k-1$ one-point mutations of the stop codon either also lead to distinct alternative phenotypes if there is a second stop codon in the genotype sequence

or they all lead to the single undefined phenotype if there is no second stop codon. Hence, for the gene-like model, the evolvability $e_g(l)$ of a genotype with coding sequence length l is given by:

$$e_g(l) = (k-1)(l-1) + \begin{cases} (k-1) & \text{if 2nd stop} \\ 1 & \text{if no 2nd stop} \end{cases} \quad (3)$$

In contrast, for the gene-like reference model:

$$e_g^{\text{ref}}(l) = (k-1)(l-1) \quad (4)$$

Genotypes with undefined phenotype have in both models an evolvability of $e_{g,\text{undef}} = e_{g,\text{undef}}^{\text{ref}} = L$.

Expressing the genotype evolvability as a function of the robustness gives for the gene-like model:

$$e_g(r_g) = (k-1)(L - Lr_g - 1) + \begin{cases} (k-1) & \text{if 2nd stop} \\ 1 & \text{if no 2nd stop} \end{cases} \quad (5)$$

The distinction between the cases of a second or no second stop codon in the genotype sequence explains the two separate lines found in Figure 2(A,TOP). The corresponding expression for the gene-like reference model is:

$$e_g^{\text{ref}}(r_g) = (k-1)(L - (L-1)r_g - 1) \quad (6)$$

For both models, there is a negative correlation between the genotype evolvability and robustness in agreement with the results found for the Fibonacci model [20] and other GP maps [11, 15].

B. Phenotype robustness and evolvability

Consider a phenotype with coding sequence length l . For both models, all genotypes mapping to that phenotype have the same robustness. Hence, by definition, the robustness $r_p(l)$ of that phenotype is identical to the corresponding genotype robustness:

$$r_p(l) = r_g(l) = 1 - \frac{l}{L} \quad (7)$$

$$r_p^{\text{ref}}(l) = r_g^{\text{ref}}(l) = \frac{L}{L-1} - \frac{l}{L-1} \quad (8)$$

The same holds for the undefined phenotype, its robustness is given by $r_{p,\text{undef}} = r_{g,\text{undef}} = r_{p,\text{undef}}^{\text{ref}} = r_{g,\text{undef}}^{\text{ref}} = (k-2)/(k-1)$.

In contrast, there is a fundamental difference in the phenotype evolvability for both models. For the gene-like reference model, from each genotype mapping to a particular phenotype, the same set of distinct alternative phenotypes is accessible through one-point mutations of the letters preceding the first stop codon. For this reason, the evolvability $e_p^{\text{ref}}(l)$ of a phenotype with coding sequence length l is identical to the corresponding genotype evolvability:

$$e_p^{\text{ref}}(l) = e_g^{\text{ref}}(l) = (k-1)(l-1) \quad (9)$$

This also holds for the gene-like model. However, due to the mutable first stop codon, the undefined phenotype or another $k-1$ alternative phenotypes can be accessed from any genotype through one-point mutations of the first stop codon, depending on whether there is a second stop codon. Since the genotypes mapping to a phenotype differ in the non-coding sequences following the first stop codon, this adds an additional set of distinct alternative phenotypes that are accessible from that phenotype.

Their number is given by one (undefined phenotype) plus $k-1$ times the number of different letter sequences ending with a stop codon of length up to $L-l$, i.e. the number of distinct coding sequences placed in the current non-coding parts of the genotypes mapping to the phenotype. Thus, for the gene-like model, the evolvability of a phenotype with coding sequence l is given by:

$$e_p(l) = (k-1)(l-1) + 1 + (k-1) \sum_{x=1}^{L-l} (k-1)^{x-1} \\ = 1 + (k-1) \left(l-1 + \frac{(k-1)^{L-l} - 1}{k-2} \right) \quad (10)$$

For both models, the evolvability of the undefined phenotype is given by $e_{p,\text{undef}} = e_{p,\text{undef}}^{\text{ref}} = \sum_{x=1}^L (k-1)^{x-1} = ((k-1)^L - 1)/(k-2)$.

Expressed as a function of the phenotype robustness, the phenotype evolvability for the gene-like model reads:

$$e_p(r_p) = 1 + (k-1) \left(L - Lr_p - 1 + \frac{(k-1)^{Lr_p} - 1}{k-2} \right) \quad (11)$$

Its derivative with respect to the robustness is:

$$\frac{de_p(r_p)}{dr_p} = (k-1)L \left(-1 + \frac{\ln(k-1)(k-1)^{Lr_p}}{k-2} \right) \quad (12)$$

This expression is larger than zero, i.e. there is a positive correlation between the phenotype evolvability and robustness, if $r_p > r_p^c$ with:

$$r_p^c = \frac{1}{L} \log_{k-1} \left(\frac{k-2}{\ln(k-1)} \right) \quad (13)$$

As it is $r_p^c < \frac{1}{L}$, the positive correlation holds for all robustness values except those corresponding to phenotypes with coding sequence length $l = L$ as $r_p(l = L) = 0$. In contrast, for the gene-like reference model, expressing the phenotype evolvability as a function of the robustness gives:

$$e_p^{\text{ref}}(r_p) = (k-1)(L - (L-1)r_p - 1) \quad (14)$$

This expression is identical to the corresponding one derived for the genotype quantities and so there is also a negative correlation between the phenotype evolvability and robustness for all robustness values.

The positive correlation found for the gene-like model agrees with the results found for the Fibonacci model [20] and qualitatively with those found for other GP maps [11, 15]. The fundamental difference between the results for the gene-like model and its reference model confirms what has been found for the Fibonacci model [20], namely the importance of a mutable stop codon, which contributes the term in the expression for the evolvability that leads to the positive correlation. A mutable stop codon allows a part of the genotype sequence to become coding that was not coding before. As genotypes mapping to the same phenotype have different non-coding sequences following the first stop codon, this can cause that

the sets of accessible alternative phenotypes from these genotypes differ, which is not the case for the gene-like reference model. In consequence, the smaller the coding sequence of a phenotype and thus the higher its robustness, the larger the number of genotypes with different non-coding sequences that map to it and thus the more significant the increase in evolvability.

C. Phenotype frequency, rank and robustness

Despite the stark difference in the correlation between phenotype robustness and evolvability, both gene-like models agree in several other properties mentioned in the introduction.

One of them is the skewed distribution of the number of genotypes per phenotype, which is commonly displayed as a phenotype frequency-rank plot. By the frequency of a phenotype, we simply refer to the number of genotypes mapping to it. As both models are based on the same assignment of genotypes to phenotypes, the frequency $f(l)$ of a phenotype with coding sequence l is in both cases identically given by:

$$f(l) = f^{\text{ref}}(l) = k^{L-l} \quad (15)$$

where we use the fact that all genotypes that map to a particular phenotype comprise the same coding sequence of length l but differ in the $L-l$ sites following the first stop codon, which can be occupied by any letter of the alphabet. Phenotypes with same coding sequence length have the same frequency. Thus, the number of phenotypes $C(l)$ with a given frequency $f(l)$ is for both models identical to the number of phenotypes (coding sequences) of length l :

$$C(l) = C^{\text{ref}}(l) = (k-1)^{l-1} \quad (16)$$

where we use the fact that the $l-1$ sites preceding the stop codon in a coding sequence of length l can be occupied by any of the $k-1$ neutral letters. By the rank of a phenotype, we refer to a ranking of the phenotypes according to their frequency such that the phenotype with highest frequency has rank one. As the frequency decreases with the coding sequence length, it follows for the rank $r(l)$ of a phenotype with coding sequence length l :

$$r(l) = r^{\text{ref}}(l) = \sum_{x=1}^{l-1} C(x) + 1 = \frac{(k-1)^{l-1} - 1}{k-2} + 1 \quad (17)$$

For simplicity, we neglect the undefined phenotype which has a frequency of $f_{\text{undef}} = (k-1)^L$. Expressed as a function of the rank, the frequency reads:

$$f(r) = f^{\text{ref}}(r) = k^{L-1-\log_{k-1}(1+(r-1)(k-2))} \quad (18)$$

For large rank values, i.e. $r \gg 1$, this function behaves as a power law:

$$f(r) = f^{\text{ref}}(r) \approx k^{L-1-\log_{k-1}(r(k-2))} = ar^{-\alpha} \quad (19)$$

with $a = k^{L-1-\log_{k-1}(k-2)}$ and exponent $\alpha = \log_{k-1}(k)$. This highly skewed distribution is in accordance to what has been found for other GP maps [11, 15]. In particular, the power law behaviour agrees with the results found for the Fibonacci model [20] and for model 1 in [21].

Both gene-like models also agree in the logarithmic scaling of the phenotype robustness with the frequency:

$$r_p(f) = 1 - \frac{L - \log_k(f)}{L} = \frac{\log_k(f)}{L} \quad (20)$$

$$r_p^{\text{ref}}(f) = \frac{L}{L-1} - \frac{L - \log_k(f)}{L-1} = \frac{\log_k(f)}{L-1} \quad (21)$$

The minor difference results from the different expressions for the phenotype robustness due to the consideration of a mutable first stop codon or not. Again, this logarithmic scaling is in agreement to what has been found for the Fibonacci model [20] and other GP maps [11, 13, 17].

III. RNA-LIKE GP MAP MODEL

In contrast to the gene-like GP map model, which considers concrete letter sequences as the phenotypes, the RNA-like model defines phenotypes as the structures formed by bonds between sequence positions of the genotypes, similar to the RNA secondary structure GP map.

For the model, we again consider an alphabet consisting of $k > 2$ letters and allocate a special role to one of the letters. Here, we choose one letter to be the ‘coding’ letter, while the other $k-1$ letters are labelled as ‘non-coding’. As for the gene-like model, we define the genotypes as all possible sequences of length L built from the letters of this alphabet. However, as the phenotype of a genotype, we now define the structure that results by linking pairs of furthest-separated coding letters within the genotype sequence, proceeding from the outside to the inside to ensure a planar structure, i.e. such that there are no overlaps between links. Note that the positions of the linked coding letters suffice as a unique representation of the phenotype. Genotypes that do not form any links are again assigned to a single undefined phenotype.

In order to analyse the effect of the links, we again consider a reference model. It works with the same genotypes as the RNA-like model, but defines the phenotypes simply by the position of the coding letters within the genotype sequences. Both models are summarised in Figure 3.

As before, we will analyse and compare key properties of the GP maps of both models in the following. Figure 4 shows plots of these properties determined by an exhaustive enumeration of the GP maps.

A. Genotype robustness and evolvability

For the RNA-like model, the robustness of a genotype depends on the number of coding letters within the geno-

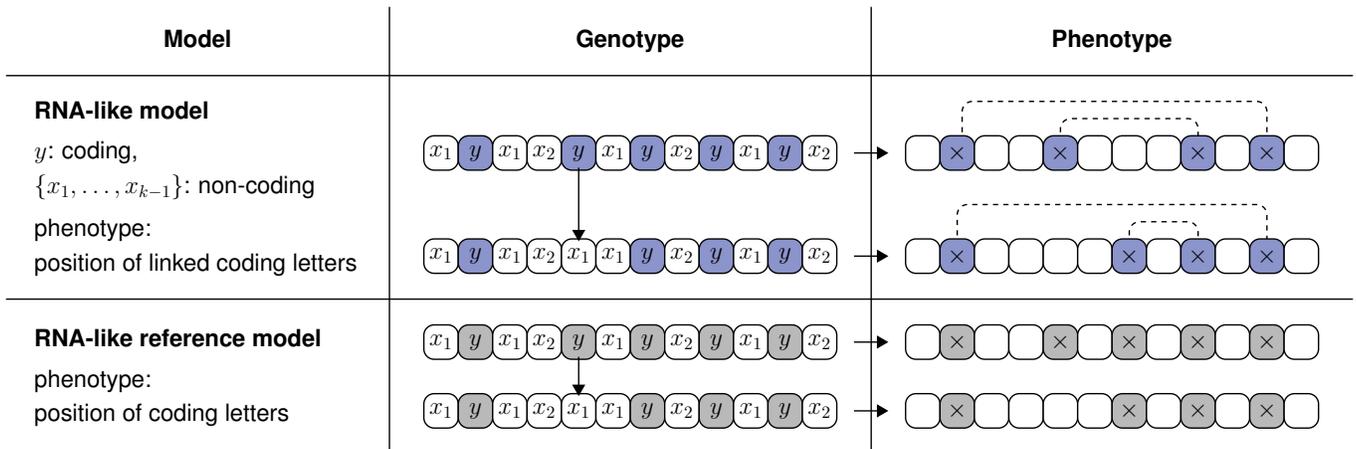


FIG. 3. Summary of the RNA-like GP map model and its reference model together with examples of genotypes and their phenotypes. For both models, the same genotypes are shown, which differ by the letter of the fifth site. For the RNA-like model, the phenotype of a genotype is defined as the structure that results by linking pairs of furthest-separated coding letters within the genotype sequence, proceeding from the outside to the inside. This is equivalent to defining the phenotype by the position of the linked coding letters. For the reference model, the phenotype of a genotype is simply defined by the position of the coding letters.

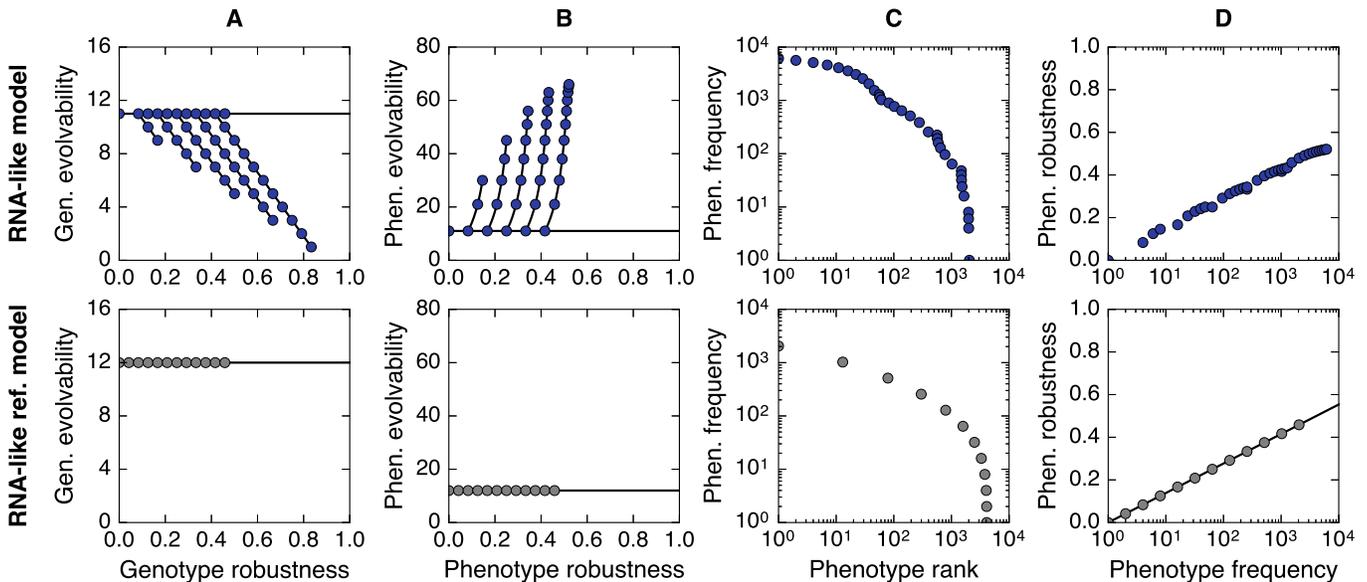


FIG. 4. Properties of the RNA-like GP map model (TOP) and the RNA-like reference model (BOTTOM), determined by an exhaustive enumeration of the GP maps. In all cases, an alphabet size of $k = 3$ and a genotype sequence length of $L = 12$ is chosen. Lines show analytical solutions. While both models qualitatively agree in the phenotype frequency-rank plot (C) and phenotype robustness-frequency plot (D), the robustness and evolvability for both genotypes and phenotypes exhibit very different relationships in the two models (A and B). For the reference model, genotype as well as phenotype robustness and evolvability are uncorrelated, whereas for the RNA-like model, negative correlations emerge for the genotype quantities (A, TOP) and positive ones for the phenotype quantities (B, TOP). This highlights the importance of non-local effects of mutations for the coexistence of robustness and evolvability in phenotypes. Note: The plots do not show the properties of the undefined phenotype (B-D) and the properties of its corresponding genotypes (A).

type sequence, which we denote by i . Consider a genotype with i coding letters. Each of the $L - i$ non-coding letters within the sequence can be mutated to any of the other non-coding letters without changing the phenotype, giving $(k - 2)(L - i)$ neutral one-point mutations.

If i is odd, the innermost coding letter will stay unlinked in the phenotype. This letter can be mutated to any of the non-coding letters without changing the phenotype, giving further $k - 1$ neutral one-point-mutations. If i is even, the robustness of the genotype also depends on the

number of non-coding letters between the innermost pair of coding letters, which we denote by j . Each of these j non-coding letters can also be mutated to the coding letter without changing the phenotype, because this additional coding letter will stay unlinked. This gives further j neutral one-point mutations. Thus, it follows for the genotype robustness:

$$r_g(i, j) = \begin{cases} \frac{(k-2)(L-i)+(k-1)}{(k-1)L} & \text{if } i \text{ odd} \\ \frac{(k-2)(L-i)+j}{(k-1)L} & \text{if } i \text{ even} \end{cases} \\ = \begin{cases} \frac{k-2}{k-1} - \frac{(k-2)i}{(k-1)L} + \frac{1}{L} & \text{if } i \text{ odd} \\ \frac{k-2}{k-1} - \frac{(k-2)i}{(k-1)L} + \frac{j}{(k-1)L} & \text{if } i \text{ even} \end{cases} \quad (22)$$

Genotypes with no or one coding letter lead to the undefined phenotype, their robustness is given by $r_{g,\text{undef}} = 1$ and $r_{g,\text{undef}} = ((k-2)(L-1) + (k-1))/((k-1)L)$, respectively.

For the RNA-like reference model, the robustness of a genotype solely depends on the number of coding letters i and no distinction between cases is necessary. As for the RNA-like model, the phenotype does not change if one of the non-coding letters within the sequence is mutated to any of the other non-coding letters. However, all mutations from a non-coding letter to the coding one or vice versa change the phenotype. Hence, the robustness of a genotype with i coding letters is given by:

$$r_g^{\text{ref}}(i) = \frac{(k-2)(L-i)}{(k-1)L} = \frac{k-2}{k-1} - \frac{(k-2)i}{(k-1)L} \quad (23)$$

The robustness of genotypes mapping to the undefined phenotype, i.e. those with no coding letter, is given by $r_{g,\text{undef}}^{\text{ref}} = (k-2)/(k-1)$.

For the RNA-like model, we also need to consider even and odd i separately when calculating the genotype evolvability. If i is odd, mutating any of the coding letters within the sequence to a non-coding one or vice versa changes the link structure and so the phenotype. The only exception is the one currently unlinked coding letter, as mutating this letter to a non-coding one has no effect on the phenotype. Thus, in total, $L-1$ distinct alternative phenotypes are accessible through one-point mutations. If i is even, again all mutations from a coding letter to a non-coding one or vice versa change the phenotype, but now apart from the j non-coding letters between the innermost pair of coding letters because mutating one of them to the coding letter has no effect on the phenotype. In addition, mutating one of the two innermost coding letters to a non-coding letter has the same effect, it will lead to a removal of the innermost link. Hence, in total, $L-j-1$ distinct alternative phenotypes are accessible through one-point mutations. Bringing everything together, the genotype evolvability is given by:

$$e_g(i, j) = \begin{cases} L-1 & \text{if } i \text{ odd} \\ L-j-1 & \text{if } i \text{ even} \end{cases} \quad (24)$$

Genotypes with undefined phenotype, i.e. those with no or one coding letter, have an evolvability of $e_{g,\text{undef}} = 0$

and $e_{g,\text{undef}} = L-1$, respectively. In general, as both the genotype robustness and evolvability depend on i and j , it is not possible to express the evolvability as a function solely of the robustness. Nevertheless, a relationship between both quantities can be determined. As can be seen from (24), for all odd values of i , the evolvability is constant and uncorrelated from the robustness. In Figure 4(A,TOP), this is marked by a horizontal line, on which every second data point corresponds to an odd value of i . By contrast, for each fixed even value of i , there is a negative correlation between evolvability and robustness as the robustness increases with j (see (22)), while the evolvability decreases with j (see (24)), i.e. the larger the gap between the innermost pair of coding letters, the more robust and less evolvable a genotype. These negative correlations are marked by the lines with negative slope in Figure 4(A,TOP). Each of these lines of data points corresponds to a fixed even value of $i \in \{2, 4, \dots, 2\lfloor L/2 \rfloor\}$ and the respective values of $j \in \{0, 1, \dots, L-i\}$. As the robustness increases with decreasing i , the lines correspond to decreasing values of i from left to right. In addition, since the number of allowed values of j increases with decreasing i , lower evolvability values are reached from left to right, i.e. with increasing robustness. In the limit of large alphabet sizes, $k \rightarrow \infty$, the genotype robustness loses its dependence on j (see (22)) and the lines of data points with negative slope would converge to vertical lines. In consequence, the negative correlation for each fixed even value of i vanishes. But it still holds that lower evolvability values are reached with increasing robustness.

In contrast, for the RNA-like reference model, all genotypes have the same evolvability independent of the number of coding letters i :

$$e_g^{\text{ref}}(i) = L \quad (25)$$

as all mutations from a coding letter to a non-coding one or vice versa change the position of the coding sites and so the phenotype, giving L distinct alternative phenotypes accessible through one-point mutations. This also holds for the genotypes belonging to the undefined phenotype, which have an evolvability of $e_{g,\text{undef}}^{\text{ref}} = L$. There is no correlation between the genotype evolvability and robustness in the RNA-like reference model.

This significant difference between the RNA-like model and its reference model shows that for the RNA-like GP map model the consideration of links between sequence positions of the genotypes seems to be crucial for the emergence of negative correlations between the genotype robustness and evolvability as it gives genotypes the possibility to be both more robust and less evolvable. The reason why we do not find a negative correlation for the RNA-like reference model but for the gene-like reference model is that for the RNA-like reference model only the position of the coding letters matters, whereas for the gene-like reference model the concrete letters in a coding sequence matter for the phenotype definition. This means that both the robustness and evolvability of geno-

types depend on the length of the coding sequence, resulting in a negative correlation between the two.

B. Phenotype robustness and evolvability

For the RNA-like model, the robustness of a phenotype depends on two variables, namely the number of linked sites and the number of unlinked sites between the innermost pair of linked sites, which we denote by i and j , in analogy to the characterisation of genotypes. A phenotype with i linked sites (even number) and j sites between the innermost linked pair either results from genotypes with i coding letters (even number) that are placed at the same positions as the linked sites, there are in total $(k-1)^{L-i}$ of such genotypes, or from genotypes with $i+1$ coding letters (odd number) that are also placed at the same positions as the linked sites except for the innermost coding letter that can be placed at any of the j sites between the two second-innermost coding letters, there are in total $j(k-1)^{L-i-1}$ of such genotypes. For this reason, the phenotype robustness is given by averaging over the genotype robustness values of these two genotype kinds, weighted by the actual number of the genotypes:

$$\begin{aligned} r_p(i, j) &= \frac{(k-1)^{L-i} r_g(i, j) + j(k-1)^{L-i-1} r_g(i+1, j)}{(k-1)^{L-i} + j(k-1)^{L-i-1}} \\ &= \frac{k-2}{k-1} - \frac{(k-2)i}{(k-1)L} + \frac{1}{(k-1)L} \cdot \frac{kj}{k-1+j} \end{aligned} \quad (26)$$

In a similar way, the robustness of the undefined phenotype, originating from genotypes with no or one coding letter, is given by:

$$\begin{aligned} r_{p, \text{undef}} &= \frac{(k-1)^L(1) + L(k-1)^{L-1} \left(\frac{(k-2)(L-1)+(k-1)}{(k-1)L} \right)}{(k-1)^L + L(k-1)^{L-1}} \\ &= \frac{k-1 + L \left(\frac{(k-2)(L-1)+(k-1)}{(k-1)L} \right)}{k-1 + L} \end{aligned} \quad (27)$$

For the RNA-like reference model, all genotypes belonging to a phenotype with i coding sites have the same robustness and so the phenotype robustness is identical to the corresponding genotype robustness:

$$r_p^{\text{ref}}(i) = r_g^{\text{ref}}(i) = \frac{k-2}{k-1} - \frac{(k-2)i}{(k-1)L} \quad (28)$$

The same holds for the robustness of the undefined phenotype, i.e. $r_{p, \text{undef}}^{\text{ref}} = r_{g, \text{undef}}^{\text{ref}} = (k-2)/(k-1)$.

For the RNA-like model, the evolvability of a phenotype with i linked sites and j sites between the innermost pair of linked sites is given by:

$$e_p(i, j) = e_g(i, j) + j e_g(i+1, j) - \sum_{x=0}^{j-1} x \quad (29)$$

$$= (L-1) + \left(L - \frac{3}{2} \right) j - \frac{1}{2} j^2 \quad (30)$$

From all genotypes with i coding letters (even number) mapping to that phenotype, the sets of accessible alternative phenotypes are the same. Hence, they just contribute one time their genotype evolvability to the phenotype evolvability (first term in (29)). From all genotypes with $i+1$ coding letters (odd number) mapping to that phenotype, the sets of accessible alternative phenotypes vary across the genotypes, but they are the same for genotypes which have the innermost coding letter positioned at the same site. As there are j allowed sites for the innermost coding letter, these genotypes, in principle, could contribute j times their genotype evolvability to the phenotype evolvability (second term in (29)). The sets of accessible alternative phenotypes from genotypes with i and those with $i+1$ coding letters are distinct. However, not yet all considered accessible alternative phenotypes from genotypes with $i+1$ coding letters are distinct and the second term needs to be corrected by the number of multiple counted phenotypes (third term in (29)). The form of this correction term can be explained in the following way. First, consider the set of distinct alternative phenotypes accessible from the genotypes for which the innermost coding letter is positioned at the ‘first’ site between the second-innermost pair of coding letters. Then, consider the genotypes for which this innermost coding letter is positioned at the ‘second’ site. Through a mutation of the non-coding letter at the ‘first’ site, an alternative phenotype occurs by linking the new pair of coding letters at the ‘first’ and ‘second’ site. However, this particular phenotype has already been accounted for by the genotypes for which the innermost coding letter is at the ‘first’ site (there accessible through a mutation of the non-coding letter at the ‘second’ site). In a similar way, for genotypes, for which the unlinked coding letter is at the ‘third’ site, two accessible alternative phenotypes have already been accounted for, and so on. Summing up gives the third term. The evolvability of the undefined phenotype is given by $e_{p, \text{undef}} = 0 + L(L-1) - \frac{1}{2}(L-1)L = \frac{1}{2}L(L-1)$. The phenotype evolvability only depends on j . But since the phenotype robustness depends on i and j , the phenotype evolvability cannot be expressed as a function of the robustness alone. However, a correlation between both quantities can be determined for fixed values of $i \in \{2, 4, \dots, 2\lfloor L/2 \rfloor\}$. As can be derived from (26), the phenotype robustness increases with j for each fixed value of i . Also the phenotype evolvability increases with j for each fixed value of i since it holds:

$$\frac{de_p(i, j)}{dj} = L - \frac{3}{2} - j > 0 \quad (31)$$

as $j \in \{0, 1, \dots, L-i\}$ and so $\max(j) = L-2$. In consequence, for each fixed value of i , there is a positive correlation between the phenotype evolvability and robustness. In Figure 4(B, TOP), these positive correlations are marked by the lines with positive slope. Each of these lines of data points corresponds to a fixed value of $i \in \{2, 4, \dots, 2\lfloor L/2 \rfloor\}$ and the respective values of

$j \in \{0, 1, \dots, L - i\}$. Similar to what has been found for the genotype quantities, the phenotype robustness increases with decreasing i and so the lines correspond to decreasing values of i from left to right. As a smaller i implies a larger value range of j , higher evolvability values are reached from left to right, i.e. with increasing robustness. In the limit of large alphabet sizes, $k \rightarrow \infty$, the phenotype robustness loses its dependence on j (see (26)) and the lines of data points with positive slope would converge to vertical lines. In consequence, the positive correlation for each fixed value of i vanishes. But it still holds that higher evolvability values are reached with increasing robustness.

In contrast, for the RNA-like reference model, from each genotype belonging to a particular phenotype with i coding sites, the same set of distinct alternative phenotypes is accessible and the phenotype evolvability is identical to the genotype evolvability:

$$e_p^{\text{ref}}(i) = e_g^{\text{ref}}(i) = L \quad (32)$$

The same holds for the undefined phenotype, its evolvability is given by $e_{p,\text{undef}}^{\text{ref}} = e_{g,\text{undef}}^{\text{ref}} = L$. Hence, there is no correlation between the phenotype evolvability and robustness in the RNA-like reference model.

This significant difference between the RNA-like model and its reference model shows that the positive correlations between the phenotype robustness and evolvability depend on relationships between sequence positions of the genotypes. If a mutation occurs at one site of the genotype sequence it can affect the role of a letter at a different site. This becomes important when the innermost coding letter is unlinked in the case of an odd number of coding letters. Mutating another letter from non-coding to coding or vice versa can ‘switch on’ this innermost coding letter and link it such that it has an effect on the phenotype. As this innermost coding letter can be positioned at different sites within the genotype sequences of genotypes mapping to the same phenotype, this can cause that the sets of accessible alternative phenotypes from genotypes with same phenotype differ. The larger the gap between the innermost linked pair in a phenotype and thus the higher its robustness, the more significant this effect and the higher the increase in evolvability.

C. Phenotype frequency, rank and robustness

Similar to what has been found for the gene-like model and its reference, the RNA-like model and its reference qualitatively agree in other properties of the GP maps like the shape of the phenotype frequency-rank plot and the logarithmic scaling of the phenotype robustness with the frequency. Here we leave it with a qualitative description of these results shown in Figures 4(C) and (D), a detailed analytical description can be found in the appendix A.

Instead of a pure power law behaviour, the phenotype frequency-rank plots for both RNA-like models show a

strong cut-off for large rank values (see Figures 4(C)). This shape agrees with the one found for the frequency-rank plot of RNA secondary structure of same length [11], thereby highlighting the agreement of our simple model with real sequence-to-structure GP maps. The reason for this shape can be explained in the following way. For the RNA-like reference model, the phenotype frequency only depends on the number of coding sites within the phenotype sequence, and decreases exponentially with it. Thus, the number of phenotypes with a given frequency is identical to the number of phenotypes with the corresponding number of coding sites, which, in turn, is identical to the number of different ways to place these coding sites across the phenotype sequence of fixed length (mathematically described by a binomial coefficient). For a large number of coding sites, corresponding to a low phenotype frequency, the number of different ways to place the coding sites is highly limited and so the number of phenotypes with low frequency, thereby explaining the observed cut-off in the frequency-rank plot. For the RNA-like model, the frequency of a phenotype not only depends on the number of linked sites, but also on the size of the gap between the innermost linked pair. For this reason, a much higher number of data points appears in the frequency-rank plot (see Figure 4(C,TOP)). Both frequency-rank plots show a similar overall shape because for the RNA-like model, the phenotype frequency as well as the phenotype rank depend in a quite similar way on the number of linked sites as both quantities depend on the number of coding sites in the RNA-like reference model.

The exponential scaling of the phenotype frequency with the number of linked sites (RNA-like model) or the number of coding sites (RNA-like reference model) also causes the logarithmic scaling of the phenotype robustness with the frequency observed for both models (see Figures 4(D)).

IV. DISCUSSION AND CONCLUSION

We introduced two simple and analytically tractable GP map models, a gene-like model that follows the principles of gene expression and a RNA-like model in analogy to sequence-to-structure GP maps like that of RNA secondary structure. Both models reproduce several of the properties found for other GP maps, like a skewed distribution of the number of genotypes per phenotype, a negative correlation between genotype robustness and evolvability, a positive correlation between phenotype robustness and evolvability and a logarithmic scaling of the phenotype robustness with the frequency. By introducing a reference model for each of the models we were able to isolate the characteristics of the models that are responsible for the positive correlation between phenotype robustness and evolvability. For the gene-like model, this is the mutable first stop codon, and for the RNA-like model, this is the existence of links between sites. At

first sight, both of these characteristics may seem very different, but they have a similar effect. Both mean that the mutation of a letter at one site of the sequence can affect the role of another letter elsewhere in the sequence, possibly a great distance away. In the gene-like model, a mutation of the first stop codon in a sequence can ‘switch on’ another stop codon in the sequence, and in the RNA-like model, a mutation from a non-coding to the coding letter or vice versa can ‘switch on’ (‘switch off’) another coding letter that is currently unlinked (linked). In general, these characteristics can allow genotypes that map to the same phenotype to access different sets of alternative phenotypes through mutations. This is compatible with Wagner’s findings [15] that the sets of accessible alternative phenotypes from genotypes sampled from the same neutral network mostly differ, and his suggestion that this is one of the causes for the positive correlation between phenotype robustness and evolvability. For the RNA-like model, the dependence between sequence positions also leads to the negative correlations between genotype robustness and evolvability, which are not found in the RNA-like reference model.

The identified non-local mutation effects resemble the widely-studied phenomenon of epistasis, but also differ from the conventional notion of epistasis in that a mutation in one location is not necessarily linked to a mutation in another specific sequence position, but instead can change sequence constraints at a range of different sequence positions elsewhere. The mutation of a stop codon for instance would normally not be regarded as an epistatic effect.

For the Fibonacci GP map model [20], it has been proposed that it is the organisation of the sequences into constrained and unconstrained parts that determines fundamental properties of GP maps. This can be further supported by our introduced models. For both the gene-like and RNA-like model, we have found that the respective main and reference models agree in properties like the shape of the frequency-rank plot and the logarithmic scaling of the phenotype robustness with the frequency. This shows that for these properties, the additional features of a mutable stop codon or links are not relevant and they are likely to be solely determined by the organisation of the sequences into constrained and unconstrained parts. For the gene-like and gene-like reference model, it is the first stop codon that divides a genotype sequence into a constrained and an unconstrained part. For the RNA-like model, the linked sites can be seen as constrained and the unlinked sites as largely unconstrained. This is because we restrict ourselves to an alphabet with only one coding letter but at least two non-coding letters such that the letter of an unlinked site can be mutated without changing the phenotype, while it always changes if the letter of a linked site is mutated. In a similar way, this holds for the RNA-like reference model, for which the coding sites can be seen as constrained and the non-coding sites as largely unconstrained.

In [21], it has been investigated how the organisation

of sequences into constrained and unconstrained parts affects the form of the neutral network size distribution. This distribution ranges from a pure power law distribution if the constrained and unconstrained sites occupy fixed positions at each end of the sequence, respectively, to a log-normal distribution if the constrained sites can occupy any position in the sequence. We did not explicitly consider the distribution of the neutral network sizes, but the phenotype frequency-rank function, which is related to it. Here we find a pure power law behaviour for the gene-like model, for which the sequences are just split into one constrained and unconstrained part, and a strong deviation from a power law behaviour for the RNA-like model, for which the constrained (linked) sites can occupy any position in the sequence. Both is in accordance to the findings in [21].

To summarise, we introduce two GP map models that exhibit the same properties observed in previous analytical and computational GP map models, but that allow us to examine the specific requirements for a positive correlation between phenotype robustness and evolvability, one of the most striking properties of GP maps. One might argue that these models are very abstract, and far removed from real biological GP maps. We have considered a number of more advanced models, such as, in the gene-like case, models that include a start codon or even multiple start and stop codons. For the RNA-like case, we also considered models that only allow links between specific letter pairs. Apart from losing the analytically tractability, these alternative, more complex models did not give any significant new insights as all of the observed GP map properties can be observed in the simpler models we have introduced here. It therefore seems that non-local effects of mutations on sequence constraints are indeed likely to be an important feature in biological GP maps.

Author contributions

MW and SEA designed the analysis, MW carried out the analysis, MW and SEA wrote the manuscript.

Competing interests

We declare we have no competing interests.

Funding

MW was supported by the EPSRC and the Gatsby Charitable Foundation. SEA was supported by the Royal Society and the Gatsby Charitable Foundation.

Appendix A: Further analytical description of the RNA-like model and its reference model

1. Phenotype frequency

For the RNA-like model, the frequency of a phenotype depends on the number of linked sites, denoted by i , and the number of unlinked sites between the innermost pair of linked sites, denoted by j . As mentioned in the article, a phenotype with i linked sites (even number) and j sites between the innermost linked pair either results from genotypes with i coding letters (even number) that are placed at the same positions as the linked sites, there are in total $(k-1)^{L-i}$ of such genotypes, or from genotypes with $i+1$ coding letters (odd number) that are also placed at the same positions as the linked sites except for the innermost coding letter that can be placed at any of the j sites between the two second-innermost coding letters, there are in total $j(k-1)^{L-i-1}$ of such genotypes. Thus, the phenotype frequency is given by:

$$f(i, j) = (k-1)^{L-i} + j(k-1)^{L-i-1} \quad (\text{A1})$$

The frequency of the undefined phenotype, which results from genotypes with no or only one coding letter, is given by $f_{\text{undef}} = (k-1)^L + L(k-1)^{L-1}$.

For the RNA-like reference model, the frequency of a phenotype solely depends on the number of coding sites, denoted by i . A phenotype with i coding sites simply results from all genotypes with i coding letters placed at these sites and so its frequency is given by:

$$f^{\text{ref}}(i) = (k-1)^{L-i} \quad (\text{A2})$$

The frequency of the undefined phenotype is given by $f_{\text{undef}}^{\text{ref}} = (k-1)^L$.

2. Phenotype frequency versus rank

For the RNA-like model, in general, the total number of phenotypes with a given frequency $f(i, j)$ is not simply identical to the number of phenotypes with i linked sites and j sites between the innermost pair of linked sites as there might be phenotypes characterised by different values $(i', j') \neq (i, j)$ that also have the same frequency $f(i', j') = f(i, j)$. In consequence, it is not possible to find a simple analytical expression for the total number of phenotypes with a given frequency $f(i, j)$, which would be useful to calculate the phenotype rank. However, it is at least possible to find an analytical expression for the number of phenotypes characterised by the same values of i and j , which definitely have the same frequency $f(i, j)$.

We denote this number by $\tilde{C}(i, j)$:

$$\begin{aligned} \tilde{C}(i, j) &= \prod_{x=1}^{L-i-j} ((i+j+1) - (j+1) + (x-1)) \\ &\quad \times \frac{1}{(L-i-j)!} \\ &= \prod_{x=1}^{L-i-j} \frac{i-1+x}{x} = \binom{i-1+(L-i-j)}{L-i-j} \\ &= \binom{L-1-j}{L-i-j} = \binom{L-1-j}{i-1} \end{aligned} \quad (\text{A3})$$

It is identical to the number of ways to arrange i linked sites across a sequence of length L under the constraint that the innermost linked pair is separated by j sites. In order to derive this expression (first equality), we make use of an analogy. Imagine a ‘test sequence’ consisting of i linked sites and j unlinked sites representing the innermost gap, i.e. $i+j$ sites in total. Now, by analogy, the number of ways to arrange i linked sites across a sequence of length L under the constraint that the innermost linked pair is separated by j sites is given by the number of ways to fill up the ‘test sequence’ with $L-(i+j)$ further unlinked sites such that a sequence of length L results while i and j remain unchanged. For the first additional site, there are $(i+j+1) - (j+1)$ allowed slots available at the sequence. In total, there are $(i+j+1)$ slots available, but adding to the $(j+1)$ slots in the innermost gap would change j . For the second additional site, there are $(i+j+1) - (j+1) + 1$ allowed slots, and so on. Thus, the total number of ways is given by the product of these factors over the $L-(i+j)$ unlinked sites that have to be added, divided by $(L-(i+j))!$ as the concrete order of these added sites does not matter.

Now, the rank of a particular phenotype is given by the sum of $\tilde{C}(i, j)$ over all values of i and j that correspond to phenotype frequencies that are higher than the one of this considered phenotype. However, in general, there exist no simple analytical expression for the phenotype rank since the phenotype frequency does not decrease in a simple sortable order with the values of i and j and since there may be phenotypes characterised by different values of i and j but same frequency, as mentioned before.

For the RNA-like reference model, the phenotype frequency solely depends on the number of coding sites i . Thus, the number of phenotypes with a given frequency $f(i)$ is identical to the number of phenotypes with i coding sites, which is given by a binominal coefficient:

$$C^{\text{ref}}(i) = \binom{L}{i} \quad (\text{A4})$$

As the frequency decreases with increasing i , it follows for the rank of a phenotype with i coding sites:

$$r^{\text{ref}}(i) = \sum_{x=1}^{i-1} C^{\text{ref}}(x) + 1 = \sum_{x=1}^{i-1} \binom{L}{x} + 1 \quad (\text{A5})$$

For simplicity, we neglect the undefined phenotype. Due to the sum over binomial coefficients, there exists no simplified expression for this equation and so no analytical expression for the frequency as a function of the rank. However, it is possible to determine the value pairs $(r(i), f(i))$ for all values of $i \in \{1, 2, \dots, L\}$, which agree with the data points shown in Figure 4(C,BOTTOM). As mentioned in the article, the observed cut-off in the frequency-rank plot for large rank values results from the fact that the frequency decreases exponentially with i for all values of i , while the rank only marginally increases with i for large values of i .

For the RNA-like model, as can be seen from the formulas, the phenotype frequency as well as the number of phenotypes characterised by the same values of i and j depend in a quite similar way on i as the corresponding quantities do for the RNA-like reference model. In addition, for the phenotype frequency, this dependence on i dominates over the additional dependence on j . Thus, the frequency-rank plot (see Figure 4(C,TOP)) shows a similar overall shape as for the RNA-like reference model. Because of the additional dependence on j , the data points appear in a group-like pattern, whereby each group roughly corresponds to one fixed value of i and the respective values of j .

3. Phenotype robustness versus frequency

For the RNA-like model, it is not possible to express the phenotype robustness as a function solely of the phenotype frequency because both the robustness and frequency depend on i and j . However, it is possible to determine a dependence between the robustness and frequency by assuming either i or j to be fixed. Solving the phenotype frequency (A1) for j :

$$j(f, i) = \frac{f}{(k-1)L^{i-1}} - (k-1) \quad (\text{A6})$$

and inserting it into the expression for the phenotype robustness (26), gives the robustness as a function of the frequency and i :

$$r_p(f, i) = \frac{k-2}{k-1} - \frac{(k-2)i}{(k-1)L} + \frac{k}{(k-1)L} - \frac{k(k-1)L^{i-1}}{L} \cdot \frac{1}{f} \quad (\text{A7})$$

In consequence, for fixed i , the robustness scales with the frequency according to $r_p(f) = \alpha - \beta/f$ with $\alpha = \text{constant}$ and $\beta = \text{constant}$. In contrast, solving the phenotype frequency (A1) for i :

$$i(f, j) = -\log_{k-1}(f) + \log_{k-1}(k-1+j) + (L-1) \quad (\text{A8})$$

and inserting it into (26), gives the robustness as a function of the frequency and j :

$$r_p(f, j) = \frac{(k-2)\log_{k-1}(f)}{(k-1)L} - \frac{(k-2)\log_{k-1}(k-1+j)}{(k-1)L} + \frac{k-2}{(k-1)L} + \frac{1}{(k-1)L} \cdot \frac{kj}{k-1+j} \quad (\text{A9})$$

Hence, for fixed j , the robustness scales with the logarithm of the frequency according to $r_p(f) = \alpha' \log_{k-1}(f) + \beta'$ with $\alpha' = \text{constant}$ and $\beta' = \text{constant}$.

For the phenotype robustness as well as the phenotype frequency, the dependence on i dominates over the dependence on j (see original expressions for both quantities: (26) and (A1)). Thus, when expressed as a function of each other, overall, the second derived scaling of the phenotype robustness with the logarithm of the frequency will dominate. This is reflected by the plot in Figure 4(D,TOP), which shows a group-like pattern of data points. Each group of data points roughly corresponds to a fixed value of $i \in \{2, 4, \dots, 2\lfloor L/2 \rfloor\}$ and the respective values of $j \in \{0, 1, \dots, L-i\}$. Overall, the groups of data points lie on a line in the semi-log plot, i.e. behave according to (A9) for fixed j , while the single points within each group behave according to (A7) for fixed i .

For the RNA-like reference model, it is possible to express the robustness as a function solely of the frequency:

$$r_p^{\text{ref}}(f) = \frac{(k-2)\log_{k-1}(f)}{(k-1)L} \quad (\text{A10})$$

The robustness scales linearly with the logarithm of the frequency, the same overall behaviour as found for the RNA-like model.

[1] S. Ciliberti, O. C. Martin, and A. Wagner, Proceedings of the National Academy of Sciences **104**, 13591 (2007).
 [2] J. F. Matias Rodrigues and A. Wagner, PLOS Computational Biology **5**, 1 (2009).
 [3] S. Schaper and A. A. Louis, PLOS ONE **9**, 1 (2014).
 [4] K. Dingle, S. Schaper, and A. A. Louis, Interface Focus **5** (2015), 10.1098/rsfs.2015.0053.

[5] P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker, Proceedings of the Royal Society of London B: Biological Sciences **255**, 279 (1994).
 [6] I. L. Hofacker, Nucleic Acids Research **31**, 3429 (2003).
 [7] K. A. Dill, Biochemistry **24**, 1501 (1985).
 [8] K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989).

- [9] S. E. Ahnert, I. G. Johnston, T. M. A. Fink, J. P. K. Doye, and A. A. Louis, *Phys. Rev. E* **82**, 026117 (2010).
- [10] I. G. Johnston, S. E. Ahnert, J. P. K. Doye, and A. A. Louis, *Phys. Rev. E* **83**, 066105 (2011).
- [11] S. F. Greenbury, I. G. Johnston, A. A. Louis, and S. E. Ahnert, *Journal of The Royal Society Interface* **11** (2014), 10.1098/rsif.2014.0249.
- [12] E. Ferrada and A. Wagner, *Biophysical Journal* **102**, 1916 (2012).
- [13] S. F. Greenbury, S. Schaper, S. E. Ahnert, and A. A. Louis, *PLOS Computational Biology* **12**, 1 (2016).
- [14] W. Fontana, D. A. M. Konings, P. F. Stadler, and P. Schuster, *Biopolymers* **33**, 1389 (1993).
- [15] A. Wagner, *Proceedings of the Royal Society of London B: Biological Sciences* **275**, 91 (2008).
- [16] J. L. Payne and A. Wagner, *Science* **343**, 875 (2014).
- [17] J. Aguirre, J. M. Buldú, M. Stich, and S. C. Manrubia, *PLOS ONE* **6**, 1 (2011).
- [18] J. Maynard Smith, *Nature* **225**, 563 (1970).
- [19] M. Eigen, R. Winkler-Oswatitsch, and A. Dress, *Proceedings of the National Academy of Sciences of the United States of America* **85**, 5913 (1988).
- [20] S. F. Greenbury and S. E. Ahnert, *Journal of The Royal Society Interface* **12** (2015), 10.1098/rsif.2015.0724.
- [21] S. Manrubia and J. A. Cuesta, *Journal of The Royal Society Interface* **14** (2017), 10.1098/rsif.2016.0976.