A theoretical approach for quantifying the impact of changes in effective population size and expression level on the rate of coding sequence evolution.

T. Latrille[1,2], N. Lartillot[1]

[1]Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Évolutive UMR 5558, F-69622 Villeurbanne, France.

[2]École Normale Supérieure de Lyon, Université de Lyon, Université Lyon 1, Lyon, France

thibault.latrille@ens-lyon.org

January 13, 2021

Abstract

Molecular sequences are shaped by selection, where the strength of selection relative to drift is determined by effective population size (N_e). Populations with high N_e are expected to undergo stronger purifying selection, and consequently to show a lower substitution rate for selected mutations relative to the substitution rate for neutral mutations (ω). However, computational models based on biophysics of protein stability have suggested that ω can also be independent of N_e, a result proven under general conditions. Together, the response of ω to changes in N_e depends on the specific mapping from sequence to fitness. Importantly, an increase in protein expression level has been found empirically to result in decrease of ω, an observation predicted by theoretical models assuming selection for protein stability. Here, we derive a theoretical approximation for the response of ω to changes in N_e and expression level, under an explicit genotype-phenotype-fitness map. The method is generally valid for additive traits and log-concave fitness functions. We applied these results to protein undergoing selection for their conformational stability and corroborate out findings with simulations under more complex models. We predict a weak response of ω to changes in either N_e or expression level, which are interchangeable. Based on empirical data, we propose that fitness based on the conformational stability may not be a sufficient mechanism to explain the empirically observed variation in ω across species. Other aspects of protein biophysics might be explored, such as protein-protein interactions, which can lead to a stronger response of ω to changes in N_e.

Keywords protein stability · substitution rate · population-genetics · drift · expression level.

1 Introduction

Molecular sequences differ across species due to the particular history of nucleotide substitutions along their respective lineages. These substitutions in turn are the result of the interplay between evolutionary forces such as mutation and selection, whose relative forces are determined by the amount of random genetic drift. These forces have effects at different levels: mutations are carried by molecular sequences, selection is mediated at the level of individuals, while random genetic drift is a population sampling effect. Yet, they jointly contribute to the long-term molecular evolutionary process. Thus, the challenge of the study of molecular evolution is to tease out their respective contributions, based on comparative analyses.

One main aspect of this challenge is to correctly evaluate the role of random drift in the long term evolutionary process. Population genetics theory implies that the strength of drift, due to the stochastic sampling of mutations, is less pronounced in lineages with large effective population size (N_e), and as a consequence, the purification by selection of weakly deleterious mutations is more effective in large populations. This fundamental idea is at the core of the nearly-neutral theory of evolution. This theory posits that a substantial fraction of mutations are deleterious or weakly deleterious, and as a result, predicts that the substitution rate (relative to the neutral expectation), called ω, decreases along lineages with higher N_e (Ohta, 1972, 1992).

This prediction has been more quantitatively examined under the assumption that the selective effects of mutations are drawn from a fixed distribution of fitness effects (DFE) (Kimura, 1979; Welch *et al.*, 2008). Assuming a gamma distribution for the DFE, a key result obtained in this context is an approximate allometric scaling of ω as a function of N_e (i.e. $\omega \sim N_e^{-k}$), where k is the shape parameter of the DFE. In practice, DFEs are strongly leptokurtic, which thus predicts a weak negative relation between ω and N_e.

The study of protein-coding sequences evolution fostered another modelling approach, based on genotype-fitness maps instead of distribution of fitness effects. In this alternative approach, the selective effect of a mutation depends on the fitness of both the source and the target amino acids involved in the mutation event (Halpern and Bruno, 1998; Rodrigue *et al.*, 2010; Tamuri and Goldstein, 2012). Even though this modelling approach differs substantially from the one assuming a fixed DFE, it also predicts a negative correlation between ω and N_e, at least when the process is at equilibrium (Spielman and Wilke, 2015; Dos Reis, 2015).

Conversely, one striking theoretical result was the proof that ω is in fact predicted to be independent of N_e under relatively general circumstances, namely, whenever (i) the fitness is a log-concave function of a phenotype and (ii) the phenotype itself is equimutable. Equimutability states that the distribution of phenotypic changes due to mutation is independent of the current phenotype of individuals (Cherry, 1998). This general theoretical argument has been invoked in the context of *in silico* experiments of protein sequence evolution, assuming that proteins are under selection for their thermodynamic stability, with fitness being proportional to the folding probability of the protein (Goldstein, 2013). Thermodynamic stability is itself

computed using a 3D structural model of the protein. These computational experiments have led to the observation that ω is essentially independent of N_e. An explanation proposed for this result is that the distribution of changes in free energy of folding ($\Delta\Delta G$) due to mutations is approximately independent of the current free energy (ΔG), thus making the free energy of folding essentially equimutable.

However, the equimutability assumption is a relatively strong one, which also conflicts with combinatorial considerations about the relation between sequence and phenotype (Serohijos *et al.*, 2012). For example, if a protein sequence is already maximally stable, only destabilizing (or neutral) mutations can occur. More generally, assuming that the stability of a protein sequence reflects an underlying fraction of positions having already accepted destabilizing amino acids, then the probability of destabilizing mutational events is in turn expected to directly depend on the current stability of the protein.

Altogether, depending on the theoretical model mapping sequence to fitness, ω can be either independent or negatively correlated to N_e, or even positively if considering adaptive evolution and environmental changes (Lanfear *et al.*, 2014).

Empirically, variation in ω between lineages has been inferred using phylogenetic codon models applied to empirical sequences (Yang and Nielsen, 1998; Zhang and Nielsen, 2005). Confronting branch-specific ω estimates to life-history traits such as body mass or generation time uncovered a positive correlation (Popadin *et al.*, 2007; Nikolaev *et al.*, 2007). Subsequently, integrative inference methods combining molecular sequences and life-history traits have also found that ω correlates positively with traits such as longevity and body mass (Lartillot and Poujol, 2011; Figuet *et al.*, 2017). Since lineages with a large body size and extended longevity typically correspond to species with low N_e (Romiguier *et al.*, 2014), these empirical correlations suggest a negative correlation between ω and N_e, thus confirming the theoretical prediction of the nearly-neutral theory of evolution. However, the universality and robustness of the correlation between ω and life-history traits is still debated. Results have not been entirely consistent across independent studies. The correlation was found to be either not statistically significant (Lartillot and Delsuc, 2012), or even in the opposite direction depending on the specific clade under study or the potential biases taken into account (Lanfear *et al.*, 2010; Nabholz *et al.*, 2013; Lanfear *et al.*, 2014; Figuet *et al.*, 2016).

If empirical evidence for a negative correlation of ω with N_e is still not totally convincing, another empirical correlation is known to be much more robust. Indeed, expression level or protein abundance is one of the best predictors of ω, with highly expressed proteins typically having lower ω values, a correlation clearly significant although relatively weak (Duret and Mouchiroud, 2000; Rocha and Danchin, 2004; Drummond *et al.*, 2005; Zhang and Yang, 2015; Song *et al.*, 2017). Theoretical models, also based on protein stability, have been invoked to explain this negative correlation between ω and expression level (Wilke and Drummond, 2006; Drummond and Wilke, 2008). According to this argument, selection against protein misfolding due to toxicity, which is stronger for more abundant proteins, induces abundant proteins to evolve toward greater stability, resulting in a more constrained and more slowly evolving protein coding sequence (Serohijos *et al.*, 2012).

The possibility that expression level and N_e might play similar roles in the evolution of proteins has already been noticed. More precisely, under models of selection against protein misfolding, the free energy of

folding ΔG is predicted to vary similarly along a gradient of either N_e or expression level (Serohijos *et al.*, 2013). As a corollary, under strict equimutability of ΔG, these computational models imply that ω should also be independent of expression level (Serohijos *et al.*, 2012), akin to what is predicted with regards to changes in N_e.

Altogether, both theoretical results and empirical analyses are not yet conclusive about the question of how ω depends on N_e and expression level. In particular, the theoretical response of ω to changes in both N_e and expression level has not been quantified and, most importantly, has not been related to the specific map between genotype, phenotype and fitness. Such an analytical development would be useful to more decisively confront the theoretical predictions relating ω to both N_e and expression level to empirical data. Ultimately, relating proteins structural parameters to the response of ω would help to bridge the gap between protein thermodynamics on one side and comparative genomics on the other side.

Lastly, the theoretical results discussed so far are valid only at the mutation-selection-drift balance. In a non-equilibrium regime, however, and at least under a model assuming a site-independent genotype-fitness map, an increase in N_e first leads to an increase in ω caused by adaptive substitutions, and subsequently a decrease in ω due to stronger purifying selection in the long term (Jones *et al.*, 2016). Studying only equilibrium properties can thus be misleading. For this reason, the dynamic response of ω to changes in N_e must also be addressed, quantified, and its connection with the underlying selective landscape better characterized. Dynamic properties of ω to changes in N_e are of theoretical interest but are also empirically relevant, such that, if overlooked they could thwart the relation between theoretical expectations and empirical estimates.

In this context, the aim of the present study is to characterize the dynamics and equilibrium response of ω to changes in N_e and expression level, and to relate this response to structural parameters of the model. To this effect, we develop a general mathematical approach to derive a quantitative approximation of the response of ω to changes in N_e and expression level, in the context of a given genotype-phenotype-fitness map, as depicted in figure 1. In the light of previously published empirical estimates from protein thermodynamics and comparative genomics, we discuss the articulation between empirical data and our mechanistic model. We also discuss some of the alternative biophysical mechanisms that could determine the selective landscape on protein-coding sequences, and how they would modulate the response of ω to changes in N_e and expression level.

2 Results

2.1 Models of evolution

The results that are presented below are valid for a general category of models of sequence evolution, based on an additive trait x, such that the coding positions of the sequence contribute additively to the trait. The trait is under directional selection specified by a decreasing and log-concave fitness function $W(x)$. As a specific example, we more specifically consider a model of protein evolution under the constraint of thermodynamic stability, as depicted in the left panel of figure 1. This model is inspired from previous work (Williams *et al.*,

2006; Goldstein, 2011; Pollock *et al.*, 2012), except that we make several simplifying assumptions, allowing us to derive analytical equations.

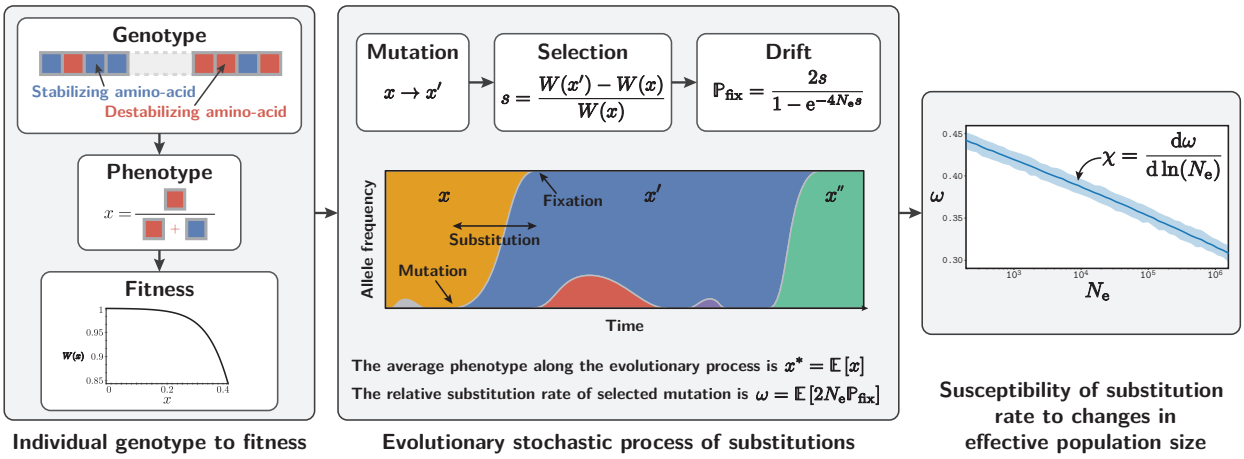

Figure 1: Outline of the theoretical results. The genotype to fitness relationship is depicted in the left panel. The phenotype (x) is a real-valued function of the genotype (i.e. the amino-acid sequence), and is defined in our model as the fraction of destabilizing amino acids in the sequence. Fitness is a decreasing log-concave function of the phenotype, depending on structural parameters of the model. Once the relation from genotype to fitness is defined, the substitution process proceeds as presented in the middle panel. For a given effective population size N_e, the evolutionary process results in an average value of the phenotype x^* and an average substitution rate (relative to the neutral rate) ω. Averaging over time is equivalent to determining the statistical equilibrium, by ergodicity of the stochastic process. The slope of the scaling of the equilibrium ω as a function of log-N_e defines the susceptibility χ, which is a function of the structural parameters defined by the phenotype-fitness map.

In the original biophysical model, protein stability is determined by the difference in free energy between the folded and unfolded conformations, called ΔG and measured in kcal/mol. Technically, free energy is computed based on the 3D conformation of the protein and using statistical potentials. As a result, the stabilizing or destabilizing effect of an amino acid at a particular site depends on amino acids present in the vicinity in 3D conformation, thus implementing what has been called specific epistasis (Starr and Thornton, 2016).

Here, we approximate this model such that the (de-)stabilizing effect at a particular site, such as measured by the $\Delta\Delta G$ of the mutation, does not depend on other neighbouring residues, thus disregarding specific epistasis (Dasmeh *et al.*, 2014). Instead, each site contributes independently and additively to ΔG. In addition, we assume that, for each site of the sequence, only one amino acid is stabilizing the protein. All 19 other amino acids are equally destabilizing. Each site bearing a destabilizing amino acid contributes an excess of $\Delta\Delta G > 0$ (in kcal/mol) to the total ΔG. The smallest achievable value of ΔG, obtained when all amino acids of the sequence are stabilizing, is noted $\Delta G_{\min} < 0$. In this model, the most succinct phenotype of a given genotype (i.e. sequence) is just the proportion of destabilizing amino acids in the sequence, defined

as $0 \leq x \leq 1$. Thus ΔG is a linear function of x:

$$\Delta G(x) = \Delta G_{\min} + n\Delta\Delta G x, \tag{1}$$

where n is the number of sites in the sequence.

For a given ΔG, thermodynamic equations allows one to derive the proportion of protein molecules that are in the native (folded) conformation in the cytoplasm. This fraction is assumed to be a proxy for fitness, motivated in part by the fact that a protein must be folded to perform its function. A slightly different model will be considered below, in order to take into account protein expression level (see section 2.3).

Analytically, the fitness function is given by the Fermi Dirac distribution and is typically close to 1, leading to a first-order approximation (Goldstein, 2011):

$$W(x) = \frac{1}{1 + e^{\beta(\Delta G_{\min} + n\Delta\Delta G x)}}, \tag{2}$$

$$\Rightarrow W(x) \simeq 1 - e^{\beta(\Delta G_{\min} + n\Delta\Delta G x)}, \tag{3}$$

$$\Rightarrow f(x) = \ln(W(x)) \simeq e^{\beta(\Delta G_{\min} + n\Delta\Delta G x)}, \tag{4}$$

where W is the Wrightian fitness for a given phenotype and f is the Malthusian fitness (or log-fitness). Here, $\Delta G_{\min}$ and $\Delta\Delta G$ are defined as above, and the parameter β is 1.686 mol/kcal at $25°C$ (or 298.2K).

Of note, even though the phenotypic effect of a mutation at a given site does not depend on the amino-acids that are present at other sites (i.e. the trait is additive), the fitness effect of a mutation still depends on other sites (i.e. the log-fitness is not additive). As a result, the molecular evolutionary process is site-interdependent, a property referred to as non-specific epistasis (Starr and Thornton, 2016; Dasmeh and Serohijos, 2018).

2.2 Response of ω to changes in N_e. Analytical approximation

For a given effective population size N_e, the evolutionary process reaches an equilibrium (figure 1, middle panel). This substitution rate at this equilibrium, normalized by the substitution rate of neutral of mutations to discard the influence of the underlying mutation rate, is denoted ω. This relative rate can also be interpreted as the mean fixation probability of mutations scaled by the fixation probability of neutral alleles $p = 1/2N_e$, the mean being weighted by the probability of occurrence of mutations in the population. As a result, an $\omega < 1$ indicates that mutations are negatively selected on average, and ω decreases with increasing strength of purifying selection.

In this section we present an analytical approximate solution for the response of ω after a change in N_e (in log space), as depicted in the right panel of figure 1. We call this response the susceptibility of ω to changes in N_e, and denote it as χ:

$$\chi = \frac{\mathrm{d}\omega}{\mathrm{d}\ln(N_e)} \tag{5}$$

Deriving χ is done in two steps. First, we determine the mean phenotype at equilibrium, when evolutionary forces of mutation, selection and genetic drift compensate each other. Subsequently, differential calculus is used to compute the response of the equilibrium phenotype to a change in N_e, which allows us to ultimately derive an equation for χ. The main results of our derivation are given both in the general case of any

(log-concave) phenotype-fitness map, and in the specific case of the biophysical model introduced above. A more detailed derivation is available in the supplementary materials.

For a given genotype, mutations can have various effects: they can increase or decrease the proportion of destabilizing amino acids, or do nothing if the mutation is between two destabilizing amino acids. To derive the probabilities of such events to occur, we also make the simplifying assumption that all transitions between amino acids are equiprobable. Altogether, any mutation in the sequence can then have a phenotypic effect of 0 or $\delta x = 1/n$, with probabilities of transitions equal to:

$$\begin{cases} \delta x & \text{with probability } 1-x, \\ 0 & \text{with probability } \frac{18x}{19}, \\ -\delta x & \text{with probability } \frac{x}{19}. \end{cases} \tag{6}$$

In the extreme case of an optimal phenotype ($x = 0$), only destabilizing mutations are proposed. Moreover, the probability to propose a stabilizing mutation (effect $-\delta x$), or a neutral mutation (effect 0), is proportional to x. Conversely, the probability to propose a destabilizing mutation is equal to $(1 - x)$. As a result, the mutation bias is proportional to $(1 - x)/x$. This mutation bias fundamentally reflects a combinatorial effect, due to the number of mutational opportunities available in either direction.

Second, we need to determine the strength of selection acting on mutations. Destabilizing mutations are selected against with a negative selection coefficient which can be approximated by:

$$s \simeq \frac{1}{n}\frac{\partial f(x)}{\partial x} \tag{7}$$

$$\Rightarrow s \simeq -\beta \Delta\Delta G e^{\beta(\Delta G_{\min} + n\Delta\Delta G x)}, \tag{8}$$

where $f = \ln(W)$ is the log-fitness (or Malthusian fitness). Conversely, stabilizing mutations will be under positive selection with opposite sign but same absolute value. It is important to realize that the selective effect is dependent on x. Furthermore, because the fitness function is log-concave, the absolute value of s increases with x.

Based on these expressions for the mutational and selective pressures, one can then study the trajectory followed by the evolutionary process. Starting from an optimal sequence, mostly destabilizing mutations will occur, some of which may reach fixation and accumulate until selection coefficients against new deleterious mutations is too strong, at which point the protein will reach a point of equilibrium called marginal stability (Taverna and Goldstein, 2002; Bloom *et al.*, 2007). Most importantly, the probability of fixation of mutations is affected by genetic drift, and thus depends on the effective population size (N_e). At the equilibrium between mutation, selection and drift, the process fluctuates through the occurrence of advantageous and deleterious substitutions compensating each other. This equilibrium can be determined by expressing the constraint that the selection coefficient of substitutions is expected to be null on average (Goldstein, 2013). Formally, and after simplification, the equilibrium phenotype denoted x^* is given in the general case by:

$$\ln\left(\frac{1-x^*}{x^*}\right) + \ln(19) \simeq -\frac{4N_e}{n}\frac{\partial f(x^*)}{\partial x^*} \tag{9}$$

$$\Rightarrow \ln\left(\frac{1-x^*}{x^*}\right) + \ln(19) \simeq 4N_e\beta\Delta\Delta G e^{\beta(\Delta G_{\min} + n\Delta\Delta G x^*)}, \tag{10}$$

in the more specific case of the biophysical model. This equation essentially expresses the mutation-selection equilibrium: the left-hand side of the equation is the log of the mutation bias at x, while the right-hand side is simply $4N_e s$, the scaled selection coefficient.

This equation cannot be solved explicitly for x^*, but a qualitative intuition on the consequences of change in N_e to the equilibrium phenotype x^* is given in figure 2. Intuitively, an increase in N_e results in a more optimal phenotype, closer to 0. The mutation bias (left-hand side of equation 10) decreases with x while the strength of selection (right-hand side of equation 10) increases with x, and the equilibrium phenotype is obtained at their intersection. An increase in N_e leads to shifting the selective response upward, which then results in a leftward shift of the equilibrium phenotype (i.e. closer to 0). The leftward shift is smaller for selective strengths characterized by a steeper curve, resulting in qualitatively weaker susceptibility of the equilibrium phenotype to changes in N_e

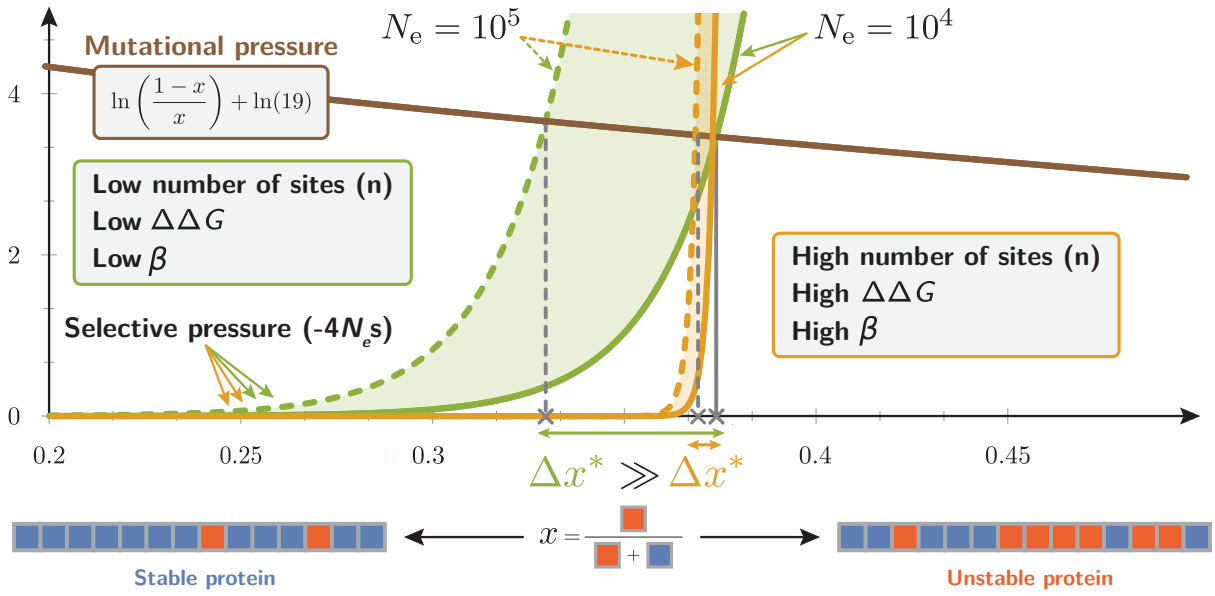

Figure 2: Response of the equilibrium phenotype after a change in N_e. The equilibrium phenotype x^* is obtained when the selective pressure equals to the mutational pressure (equation 10). The selective pressure (right-hand side of eq. 10) increases exponentially with x where $\beta n \Delta\Delta G$ is the exponential growth rate (yellow and green curves). When $\beta n \Delta\Delta G$ is large, increasing N_e by an order of magnitude (yellow dotted curves) very moderately impacts the equilibrium phenotype (small Δx^*). In contrast, for small $\beta n \Delta\Delta G$ (green curves), the equilibrium phenotype is more strongly impacted by a change in N_e (large Δx^*). Finally, response of x^* to changes in N_e reflects the response of ω since both are approximately equal (equation 11).

The results obtained thus far only relate the equilibrium phenotype (x^*) to N_e. To capture how ω varies with N_e, we also need to obtain an expression for ω as a function of x^*. At equilibrium we can derive (supplementary materials) the expected substitution rate of mutations, and thus ω, which simply approximates to:

$$\omega \simeq x^* \tag{11}$$

This simple approximation is due to the fact that the substitutions between two destabilizing amino acids (which are neutral) compose the largest proportion of proposed mutations having a substantial probability of fixation (equation 6). In contrast, stabilizing mutations are rare, while destabilizing mutations have a low probability of fixation. Since there is a fraction x^* of sites already occupied by a destabilizing amino-acid, these neutral substitutions occur at rate x^*.

Combined together, these analytical approximations yield the susceptibility (equation 5) of ω to a change in N_e:

$$\chi = \frac{d\omega}{d\ln(N_e)} \simeq -\frac{\frac{\partial f(x^*)}{\partial x^*}}{\frac{n}{4N_e}\frac{\partial \ln[(1-x^*)/x^*]}{\partial x^*} + \frac{\partial^2 f(x^*)}{\partial x^{*2}}}. \tag{12}$$

The two terms of the denominator correspond to the derivative of the mutational bias and the scaled selection coefficient, respectively. However, the mutational bias decreases weakly with x (blue curve on figure 2) while the strength of selection increases sharply with x (red and green curves). As a consequence, the derivative of the mutational bias is much lower than the derivative of the selection coefficient around the equilibrium point (i.e. the phenotype is *nearly* equimutable). The second term can therefore be ignored, which leads to a very compact equation for susceptibility χ in the general case:

$$\chi \simeq -\frac{\frac{\partial f(x^*)}{\partial x^*}}{\frac{\partial^2 f(x^*)}{\partial x^{*2}}} \tag{13}$$

The susceptibility is thus equal to the inverse of the relative curvature, i.e. the ratio of the second to the first derivatives, of the log-fitness function, taken at the equilibrium phenotype. Of note, this susceptibility is strictly negative for decreasing log-concave fitness functions, asserting that ω is a decreasing function of N_e. In addition, the susceptibility itself is low in absolute value (i.e. ω responds more weakly) for strongly concave log-fitness functions. This equation quantitatively captures the intuition developed in figure 2, namely that the response of ω is very weak if the selection curve is very steep around the equilibrium set point (red curve compared to green curve).

In the specific case of the biophysical model, the susceptibility (χ) further simplifies to:

$$\chi \simeq -\frac{1}{\beta n \Delta\Delta G}, \tag{14}$$

meaning that ω is linearly decreasing with N_e (in log scale) since χ is independent of x^*, or, in other words, that the exact value of the equilibrium phenotype has no impact on the slope. Moreover, only the compound parameter $\beta\Delta\Delta Gn$ has an impact on the slope of the linear relationship. Thus, in particular, the slope of the linear relationship between ω and N_e is affected by $\Delta\Delta G$ but not by $\Delta G_{\min}$. Of note, empirically, only relative values of N_e (up to a multiplicative constant) are required to obtain an estimate of χ.

2.3 Response of ω to changes in protein expression level

Effective population size is not the sole predictor of ω, and expression level (or protein abundance) is also negatively correlated to ω. However, our previous model, which assumes that fitness is proportional to the folded fraction, and is thus independent of protein abundance, does not express the fact that selection is typically stronger for proteins characterized by higher levels of expression. An alternative biophysical model is to assume that each misfolded protein molecule has the same relative effect on fitness, caused by its toxicity

for the cell (Drummond *et al.*, 2005; Wilke and Drummond, 2006; Drummond and Wilke, 2008; Serohijos *et al.*, 2012).

Our general derivation can be directly applied to this case. For a given protein with expression level y and a cost A representing the selective cost per misfolded molecule (positive constant), the fitness and selection coefficient can be defined as follows:

$$f(x) \simeq -Aye^{\beta(\Delta G_{\min} + n\Delta\Delta G x)} \tag{15}$$

$$\Rightarrow s \simeq -\beta\Delta\Delta G Aye^{\beta(\Delta G_{\min} + \Delta\Delta G n x)}. \tag{16}$$

Under this model, the total selective cost of a destabilizing mutation is now directly proportional to the total amount of misfolded proteins. This fitness function leads to the following expression for the mutation-selection-drift equilibrium:

$$\ln\left(\frac{1-x^*}{x^*}\right) + \ln(19) = 4N_e y A\beta\Delta\Delta G n e^{\beta(\Delta G_{\min} + \Delta\Delta G n x^*)}. \tag{17}$$

Importantly, in this equation, N_e and y are confounded factors appearing only as a product. This means that increasing either N_e or y leads to same change in equilibrium phenotype, and hence the same change in ω. In other words, the susceptibility of the response to changes in either N_e or expression level is the same:

$$\chi = \frac{\mathrm{d}\omega^*}{\mathrm{d}\ln(N_e)} = \frac{\mathrm{d}\omega^*}{\mathrm{d}\ln(y)} \simeq -\frac{1}{\beta n \Delta\Delta G}. \tag{18}$$

A similar result can be obtained under other models relating phenotype to fitness, for example if the selective cost is due to translational errors (supplementary materials). Alternatively if the protein is assumed to be regulated such as to reach a specific level of functional protein abundance under a general cost-benefit argument (Cherry, 2010; Gout *et al.*, 2010), a multiplicative factor depending solely on the expression level is prefixed (supplementary materials). Altogether, we theoretically obtain the same linear decrease of ω with regards to either effective population size or expression level (in log space) under a broad variety of hypotheses.

2.4 Simulation experiments

Our theoretical derivation of the susceptibility of ω to changes in N_e (and expression level) is based on several simplifying assumptions about the evolutionary model and makes multiple approximations. In order to test the robustness of our main result, we therefore conducted systematic simulation experiments, relaxing several of these assumptions. In each case, simulations were conducted under a broad range of values of N_e, monitoring the average ω observed at equilibrium and plotting the scaling of these measured equilibrium ω as a function of N_e.

Specifically, with respect to mutations, our derivation assumes that all amino-acid transitions are equiprobable, or in other words, the complexity of the genetic code is not taken into account. Simulating evolution of DNA sequence and invoking a matrix of mutation rates between nucleotide allows us to test the robustness of our results to this assumption. Furthermore, with regard to the phenotypic effects of amino-acid changes, in our derivation, we assumed that all destabilizing amino acids have an identical impact on protein stability. In reality, one would expect conservative amino-acid replacements to be less destabilizing than radical changes.

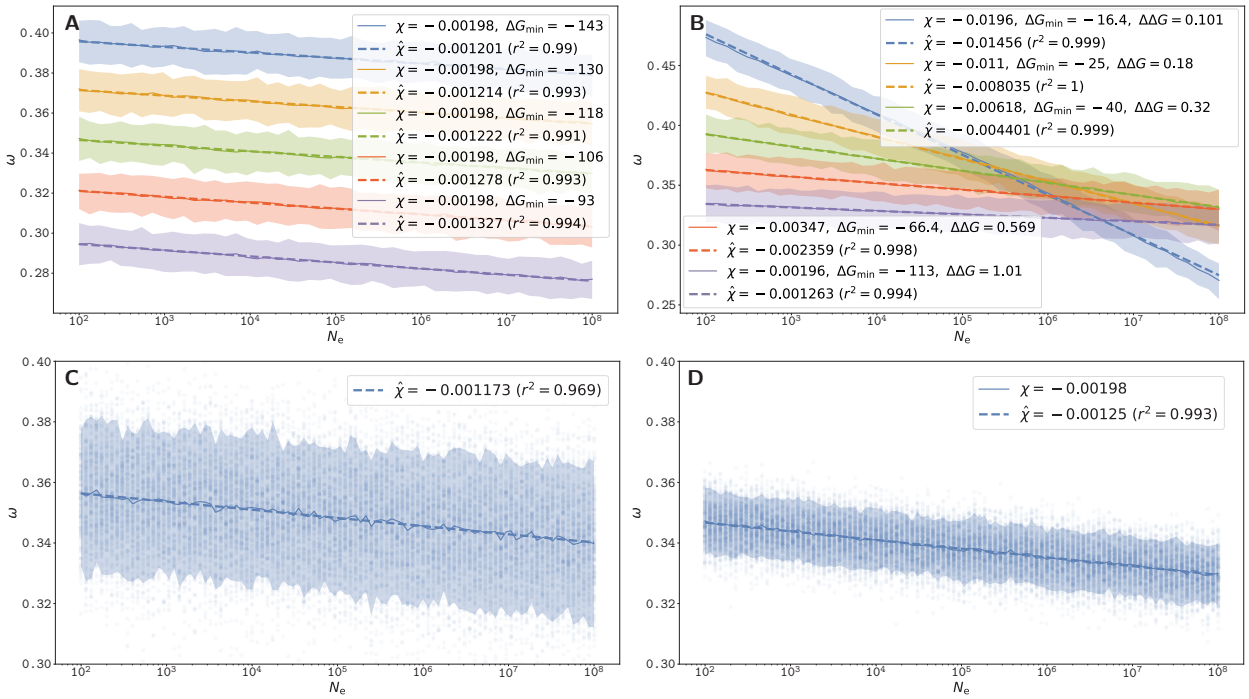

Figure 3: Scaling of equilibrium ω as a function of log-Ne, under the additive phenotype model using Grantham distances (A,B,D) or the explicit biophysical model using a statistical potential (C), with $n = 300$ and $\beta = 1.686$. 200 replicates per N_e value are shown (dots). Solid lines are average over replicates, and shaded areas are 90% confidence interval. The slope (or susceptibility ($\hat{\chi}$), is estimated by linear regression (dashed lines). (A): ΔG_{min} are given in the legend, and $\Delta\Delta G = 1$. Decreasing ΔG_{min} (to more negative values) increases ω but does not impact the slope. (B): $\Delta\Delta G$ is increased and ΔG_{min} is changed accordingly such that the equilibrium value x^* is kept constant, by solving numerically equation 10. The estimated susceptibility ($\hat{\chi}$) decreases proportionally to the inverse of $\Delta\Delta G$, as predicted by our theoretical model. (C): Stability of the folded native state is computed using 3D structural conformations and pairwise contact potentials. (D): Additive model with $\Delta G_{min} = -118$ kcal/mol and $\Delta\Delta G = 1$ kcal/mol matches structural model shown in C (although with less variance).

This assumption is relaxed in our simulation, such that destabilizing mutations in each position are now proportional to the Grantham distance (Grantham, 1974) between the optimal amino acid in this position and the amino acid proposed by the non-synonymous mutation. Finally, our derivation assumes that the number of sites in the sequence (n) is large, such that the selection coefficient is well approximated by the fitness derivative (equation 7). The robustness of this approximation was tested by conducting simulations with finite sequences of realistic length ($n = 300$ coding positions).

These simulation experiments demonstrate, first, that the relation between ω and log-N_e is indeed linear, at least in the range explored here, and that the slope of the linear regression matches the expected theoretical value (figure 3A). Secondly, we observe that the parameter ΔG_{min} has virtually no effect on the slope of the linear regression, as also expected theoretically (figure 3B). Instead, decreasing ΔG_{min} (to more negative

values) merely results in an overall increase in ω over the whole range of N_e (i.e. has an impact on the intercept, not on the slope of the relation). This is due to the fact that decreasing $\Delta G_{\min}$ shifts the equilibrium to higher x^*, since more destabilizing sites can then reach fixation before reaching the point of marginal stability.

Finally, we relaxed our assumption that each site of the sequence contributes independently to ΔG, by taking into account the 3D structure of protein and using a statistical potential to estimate ΔG (supplementary materials). We implemented the original model considered in Williams *et al.* (2006), Goldstein (2011) and Pollock *et al.* (2012), in which the free energy is computed based on the 3D conformation using pairwise contact potential energies between neighbouring amino-acid residues (Miyazawa and Jernigan, 1985). The original works showed that under this model, ω is approximately independent of N_e (Goldstein, 2013). Using extensive simulations in order to obtain sufficient resolution, we observe that ω is in fact weakly dependent on N_e, being again approximately linear with log-N_e (figure 3C). Moreover, the observed slope ($\widehat{\chi} = -0.00117$) matches the slope obtained under the model of additive ΔG ($\widehat{\chi} = -0.00125$, figure 3D), considering an empirical $\Delta\Delta G = 1.0$ kcal/mol for destabilizing mutations and $n = 300$. In this experiment (figure 3D), $\Delta G_{\min}$ was set to -118 kcal/mol, which is the ΔG of the optimal (maximally stable) sequence of 300 sites (Goldstein, 2011).

2.5 Time to relaxation

Although the equilibrium value of ω after changes in N_e is an important feature of the ω-N_e relationship, another characteristic that is scarcely studied is the dynamic aspect (Jones *et al.*, 2016), particularly the relaxation time to reach the new equilibrium ω. We observed in our simulations that the determining factor of the relaxation time is the number of sites n (figure 4A), such that the return to equilibrium is faster for longer sequences. This observation matches the theoretical prediction that more mutational opportunities are available for longer sequences, driving the trait close to equilibrium at a faster rate.

It may be useful to compare the relaxation pattern observed here with the predictions under two alternative models of sequence evolution, representing two extreme scenarios. On one hand, having fitness modelled at the level of sites, such as contemplated by many phylogenetic mutation-selection models (Halpern and Bruno, 1998; Rodrigue *et al.*, 2010; Tamuri and Goldstein, 2012), leads to a situation where every site has to adapt on its own to the new change in N_e. The relaxation time is then very long, on the order of the inverse of the per-site substitution rate. On the other hand, assuming a fixed distribution of fitness effect (DFE) as in Welch *et al.* (2008), the response of ω is instantaneous (figure 4B). Our model is effectively in between these two extreme scenarios.

Another characteristic observed in these non-equilibrium experiments is the discontinuity of ω after a change in N_e. Most importantly, both an increase and decrease in N_e lead to a discontinuity (figure 4A & 4B). These non-equilibrium behaviors can both be explained mechanistically. Under low N_e, the phenotype is far away from the optimal phenotype because the efficacy of selection is weaker. A sudden increase in N_e results first in a short traction toward a more optimal phenotype, which results in a suddenly higher ω, caused by a transient adaptation of the protein toward a higher stability. Conversely, under high N_e the phenotype is closer to optimal and the purification of deleterious mutations is stronger. The reaction to a decrease in N_e is

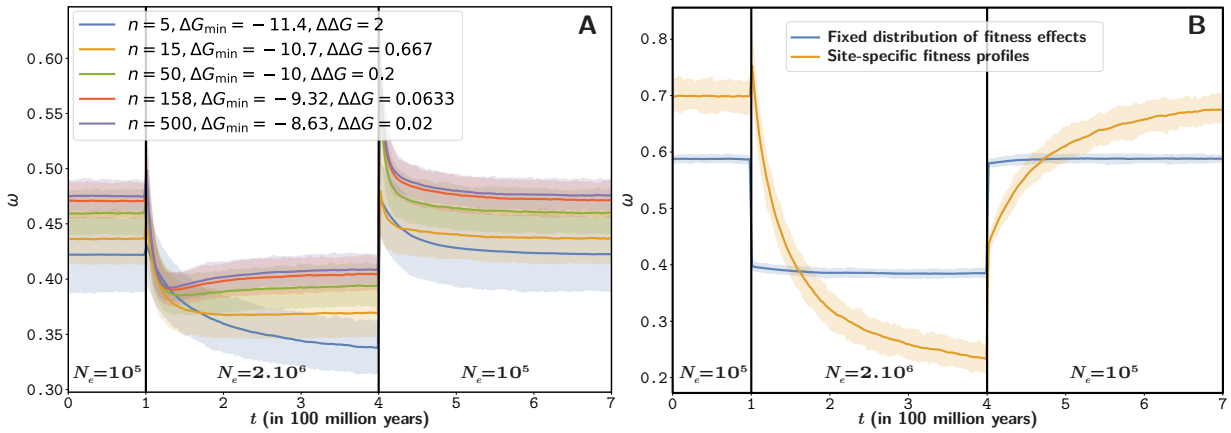

Figure 4: Relaxation of ω after a change in N_e Solid line corresponds to the average over 1000 replicates and the shaded area corresponds to the 90% interval among replicates. The mutation rate (μ) is $1e{-}8$ per year per site, and the total evolutionary period is 700 million years. (A): $\beta = 1.686$ for all simulations. The DNA sequence of 500 sites is divided into exons of equal size. However the number of sites per exon changes between simulations from $n = 5$ to $n = 500$. Moreover, $\Delta\Delta G$ is changed according to the exon size such that $n\Delta\Delta G$ (and as a result, the susceptibility) are kept constant, and $\Delta G_{\min}$ is changed accordingly such that the equilibrium value x^* is kept constant, by solving numerically equation 10. Thus, regardless of exon size, x^* and χ are kept constant and thus the observed effect is due to the number of sites in the exon. We observe that increasing the number of sites leads to a reduced time to reach the new equilibrium. (B): In the context of a time-independent fitness landscape (yellow curve), where each amino acid has different fitness (site-specific profiles), the time taken to reach the new equilibrium value of ω after a change in N_e is long. In the context of a fixed distribution of fitness effects (blue curve), the relaxation time is non-existent and the new equilibrium value of ω is reached instantaneously.

a relaxation of the purification and thus an ω closer to the neutral case, which results into higher ω until reaching the point of marginal stability. To note, an increase in N_e can theoretically and possibly lead to an ω that is temporarily greater than 1 due to adaptive evolution (Jones *et al.*, 2016), while a decrease in N_e always imply an $\omega < 1$, as it gives at most a neutral regime of relaxed selection.

3 Discussion

We provide a compact analytical result for the equilibrium response (which, by analogy with thermodynamics, we call the susceptibility) of ω to changes in N_e, and we relate this response to the parameterization of the genotype-phenotype-fitness map. An application to a model of selection against protein misfolding shows that the response of ω to variation in N_e (in log space) is linear, with a negative slope. Furthermore, this application demonstrates that effective population size and protein expression level are interchangeable with respect to their impact on the response of ω. Our compact theoretical results, which were obtained by making several simplifying assumptions, are supported by more complex simulations of protein evolution relaxing

these assumptions. In particular, our theoretical predictions are verified under a numerical model of protein evolution in which the free energy is computed based on the 3D structure.

Overall, the susceptibility (χ) is a function of the structural parameters of the protein and takes a very simple analytical form, being inversely proportional to the product of three terms: the sequence size, the inverse temperature (β), and the average change in conformational energy of destabilizing mutations ($\Delta\Delta G$). Quantitatively, this product can be several orders of magnitude greater than 1 in practice, such that the susceptibility of ω, which is its inverse, is typically small. Previous studies using this model presented an apparent lack of response of ω to changes in N_e (Goldstein, 2013). We refine this result, by observing that there is in fact a very subtle and weak relation, which requires extensive computation to be detected, but which is well predicted by our theoretical derivation. Based on empirical estimates of the structural parameters $\beta = 1.686$, $n = 300$ sites and $\Delta\Delta G = 1.0$ kcal/mol for destabilizing mutations (Zeldovich *et al.*, 2007), the estimated susceptibility is $\hat{\chi} \simeq -0.002$. In other words, for a relative increase in N_e or expression level of 6 orders of magnitude, a factor approximately equal to 0.01 is subtracted from ω, a subtle relationship that requires laborious effort to be detected in simulated data.

3.1 Adequacy to empirical data

Empirically, variation in ω along the branches of phylogenetic trees has been inferred and correlated to proxies of N_e, such as body size or other life-history traits. These analyses showed mitigated support for a negative relation between ω and N_e (Lanfear *et al.*, 2014). More recently, phylogenetic integrative methods refined the estimate of covariation between ω and N_e along lineages by leveraging polymorphism data (Brevet and Lartillot, 2019). This approach gives an estimate of $\hat{\chi} \simeq 0.02$ in primates (supplementary materials) at least one order of magnitude greater than the quantitative estimate obtained above from the biophysical model. More empirical data across different clades would be required to robustly consolidate such empirical estimates, but as of yet, these results are challenging the idea of a very weak response.

The relation between ω and expression level provides an independent, and potentially more robust, source of empirical observation. Our theoretical results suggest that, under relatively general conditions, the response of ω to expression level should be of the same magnitude than the response to N_e. Empirically, the protein expression level is one of the best predictors of ω and the empirical estimation of χ in fungi, archaea and bacteria varies in the range $[-0.046; -0.021]$ (supplementary materials) extracted from Zhang and Yang (2015). Estimation in animals and plants gives somewhat lower estimates, in the range of $[-0.026; -0.004]$, although still higher (in absolute value) than -0.002.

Additionally, another empirical observation is the negative relation between the mean destabilizing effect of mutations (mean $\Delta\Delta G$) and the ΔG of the protein. Such a relation is empirically observed in Serohijos *et al.* (2012), where the slope of the linear regression is -0.13 ($r^2 = 0.04$). The slope of the linear correlation observed in our simulations is weaker, with an observed slope of -0.01 ($r^2 = 0.29$) under the 3D biophysical model, and -0.003 ($r^2 = 0.33$) under the model of additive phenotype parameterized by $\Delta\Delta G = 1$ and $n = 300$ (supplementary materials). This observation also sheds light on the correlation between ω and N_e in empirical data and in our model. Indeed, equimutability, or namely that the distribution of $\Delta\Delta G$ of

mutations is independent of ΔG is a necessary condition to observe independence between ω and N_e (Cherry, 1998). In our model, the average $\Delta\Delta G$ of mutations at equilibrium depends on ΔG due to combinatorial considerations, but this dependence is weaker than empirically observed, which also translates into a weaker susceptibility of ω to changes in N_e or expression level than empirically observed.

Thus, overall, the response of ω to either N_e or expression level predicted by the biophysical model considered above seems lower than what is empirically observed. There are several possible explanations for this discrepancy. First, the biophysical model might be valid, but the numerical estimates used for n or $\Delta\Delta G$ could be inadequate. A $\Delta\Delta G$ of 1.0 kcal/mol for destabilizing mutations seems to correspond to empirical estimates (Zeldovich *et al.*, 2007). On the other hand, the effective number of positions implicated in the trait might be smaller than the total number of residues in the protein. In our model, all positions in the protein can in principle compensate for the destabilizing effect of a mutation at a particular position. In practice, the number of sites susceptible to compensate each other is probably smaller, resulting in a stronger departure from equimutability.

Alternatively, the biophysical model considered here might be too restrictive. Recent empirical studies have provided evidence against the hypothesis that the rate of sequence evolution is driven solely by the toxicity effect of unfolded proteins (Plata and Vitkup, 2017; Razban, 2019; Biesiadecka *et al.*, 2020). Notably, the response of ω to changes in expression level has also been found theoretically to arise as a consequence of protein-protein interactions, where protein may either be in free form or engaged in non-specific interactions (Yang *et al.*, 2012; Zhang *et al.*, 2013). In non-specific interactions at the protein surface, stabilizing amino acids are hydrophilic and destabilizing amino acids are hydrophobic, sticking to hydrophobic residues at the surface of other proteins (Dixit and Maslov, 2013; Manhart and Morozov, 2015).

Our theoretical results can be applied more broadly to protein-protein interactions using a mean-field argument (supplementary materials). Fitting this model with empirical structural estimates (Janin, 1995; Zhang *et al.*, 2008), we obtain a susceptibility of $\chi \simeq -0.2$ thus a much stronger response than under the model based on conformational stability. This much stronger response is due to fewer sites in the protein being involved in protein-protein interaction than for conformational stability, in addition to a lower free energy engaged in contact between residues.

Altogether, fitness based on protein stability is a compelling model of molecular evolution, but may not be a sufficiently comprehensive model to explain the amplitude of variation of ω empirically observed along a gradient of either effective population size or protein expression level. The net response of ω to changes in N_e or expression level could have several biophysical causes, which in the end would imply a weak but still empirically measurable response.

3.2 The statistical mechanics of molecular evolution

This study describes the signature imprinted on DNA sequences by an evolutionary process by merging equations from population genetics and from structural physicochemical first principles. More generally, it outlines a general approach for deriving quantitative predictions about the observable macroscopic properties of the molecular evolutionary process based on an underlying microscopic model of the detailed relation

between sequence, phenotype and fitness. In this respect, it borrows from statistical mechanics, attempting approximations to derive analytically tractable results (Sella and Hirsh, 2005; Mustonen and Lässig, 2009; Bastolla *et al.*, 2012, 2017) The robustness of results can be assessed by computational implementations and simulations. Computational models offer a means to test the validity and robustness, while mathematical models offer an intuitive mechanistic mental analogy.

Ultimately, the approach could be generalized to other aspects of the evolutionary process. Beyond ω, other macroscopic observables could be of interest, for example site entropy, i.e. the effective number of observed amino acids per site at equilibrium (Goldstein and Pollock, 2016; Jimenez *et al.*, 2018; Jiang *et al.*, 2018), or the nucleotide or amino-acid composition. In addition to N_{e}, other evolutionary forces could also be considered, for instance the mutational bias or GC-biased gene conversion. The susceptibility of the macroscopic observables to changes in the strength of these underlying forces could then more generally be investigated. As such, the framework outlined here could foster a better understanding of observable signatures of the long-term evolutionary process emerging from ecological parameters and molecular physico-chemical first principles, by carefully teasing out the combined effects of mutation, selection and drift.

4 Materials & Methods

Protein sequence evolution is simulated under an origin-fixation model (McCandlish and Stoltzfus, 2014), i.e. the whole population is considered monomorphic and only the succession of fixation events is modeled. Given the currently fixed sequence $\mathbb{S}$, we define $\mathcal{M}(\mathbb{S})$ as the set of all possible mutant that are one nucleotide away from $\mathbb{S}$. Non-sense mutants are not considered. For a protein of n amino-acid sites, $|\mathcal{M}(\mathbb{S})| \leq 9n$, since each codon has a maximum of 9 possible nearest neighbors that are not stop codons. For each mutant sequence $\mathbb{S}' \in \mathcal{M}(\mathbb{S})$, we compute its fitness and subsequently the selection coefficient of the mutant:

$$s(\mathbb{S}, \mathbb{S}') = \frac{W(\mathbb{S}') - W(\mathbb{S})}{W(\mathbb{S})}, \tag{19}$$

$$\Rightarrow s(\mathbb{S}, \mathbb{S}') \simeq f(\mathbb{S}') - f(\mathbb{S}'), \tag{20}$$

where W is the Wrightian fitness for a given phenotype and f is the Malthusian fitness (or log-fitness).

The waiting time before the next mutant invading the population, and the specific mutation involved in this event, are chosen using Gillespie's algorithm (Gillespie, 1977), according to the rates of substitution between $\mathbb{S}$ and each $\mathbb{S}' \in \mathcal{M}(\mathbb{S})$, which are given by:

$$Q_{\mathbb{S},\mathbb{S}'} = \mu_{\mathbb{S},\mathbb{S}'} \frac{4N_{\mathrm{e}} s(\mathbb{S}, \mathbb{S}')}{1 - \mathrm{e}^{-4N_{\mathrm{e}} s(\mathbb{S},\mathbb{S}')}}, \tag{21}$$

where $\mu_{\mathbb{S},\mathbb{S}'}$ is the mutation rate between $\mathbb{S}$ and $\mathbb{S}'$, determined by the underlying $4x4$ nucleotide mutation rate matrix, and $Q_{\mathbb{S},\mathbb{S}'} = \mu_{\mathbb{S},\mathbb{S}'}$ in the case of synonymous substitutions. Various optimizations are implemented to reduce the computation time of mutant fitness. The simulation starts with a burn-in period to reach mutation-selection-drift equilibrium.

4.1 Models of the fitness function

Under the additive model for the free energy, the difference in free energy between folded and unfolded state is assumed to be given by:

$$\Delta G\left(\mathbb{S}\right) = \Delta G_{\min} + n\Delta\Delta G * x\left(\mathbb{S}\right),$$

where $0 \leq x\left(\mathbb{S}\right) \leq 1$ is the distance of $\mathbb{S}$ to the optimal sequence (i.e. the fraction of sites occupied by a destabilizing amino-acid). For each site of the sequence, the optimal amino acids are chosen randomly at initialization, and the distance between the current amino acid and the optimal is scaled by the Grantham amino-acid distance (Grantham, 1974). The Wrightian fitness is defined as the probability of our protein to be in the folded state, given by the Fermi-Dirac distribution:

$$W(\mathbb{S}) = \frac{e^{-\beta\Delta G(\mathbb{S})}}{1 + e^{-\beta\Delta G(\mathbb{S})}} = \frac{1}{1 + e^{\beta\Delta G(\mathbb{S})}}, \tag{22}$$

where β is the inverse of the temperature ($\beta = 1/kT$).

For simulations under a 3D model of protein conformations, we adapted the model developed in Goldstein and Pollock (2017) to our C++ simulator (see supplementary materials).

For simulations under a site-independent fitness landscape, with site-specific fitness profiles, the protein log-fitness is computed as the sum of amino-acid log-fitness coefficients along the sequence. In this model, each codon site i has its own fitness profile, denoted $\boldsymbol{\phi}^{(i)} = \{\phi_{\mathrm{a}}^{(i)},\ 1 \leq \mathrm{a} \leq 20\}$, a vector of 20 amino-acid scaled (Wrightian) fitness coefficients. Since $\mathbb{S}[i]$ is the codon at site i, the encoded amino acid is $\mathcal{A}\left(\mathbb{S}[i]\right)$, hence the fitness at site i is $\boldsymbol{\phi}^{(i)}_{\mathcal{A}(\mathbb{S}[i])}$. Altogether, the selection coefficient of the mutant $\mathbb{S}'$ is:

$$s\left(\mathbb{S}, \mathbb{S}'\right) = \sum_{i=1}^{n} \ln\left(\frac{\boldsymbol{\phi}^{(i)}_{\mathcal{A}(\mathbb{S}'[i])}}{\boldsymbol{\phi}^{(i)}_{\mathcal{A}(\mathbb{S}[i])}}\right), \tag{23}$$

The fitness vectors $\boldsymbol{\phi}^{(i)}$ used in this study are extracted from Bloom (2017). They were experimentally determined by deep mutational scanning.

For simulations assuming a fixed distribution of fitness effects (DFE), the selection coefficient of the mutant $\mathbb{S}'$ is gamma distributed (shape $k > 0$):

$$-s\left(\mathbb{S}, \mathbb{S}'\right) \sim \mathrm{Gamma}\left(|\bar{s}|, k\right) \tag{24}$$

4.2 Computing ω along the simulation

From the set of mutants $\mathcal{M}\left(\mathbb{S}\right)$ that are one nucleotide away from $\mathbb{S}$, we define the subsets $\mathcal{N}\left(\mathbb{S}\right)$ of non-synonymous and synonymous mutants ($\mathcal{N}\left(\mathbb{S}\right) \subseteq \mathcal{M}\left(\mathbb{S}\right)$). The ratio of non-synonymous over synonymous substitution rates, given the sequence $\mathbb{S}$ at time t is defined as (Spielman and Wilke, 2015; Dos Reis, 2015; Jones *et al.*, 2016):

$$\omega(t) = \frac{\sum\limits_{\mathbb{S}' \in \mathcal{N}(\mathbb{S})} \mu_{\mathbb{S},\mathbb{S}'} \dfrac{4N_{\mathrm{e}} s\left(\mathbb{S}, \mathbb{S}'\right)}{1 - \mathrm{e}^{-4N_{\mathrm{e}}\left(\mathbb{S},\mathbb{S}'\right)}}}{\sum\limits_{\mathbb{S}' \in \mathcal{N}(\mathbb{S})} \mu_{\mathbb{S},\mathbb{S}'}} \tag{25}$$

Averaged over all branches of the tree, the average ω is:

$$\omega = \langle \omega(t) \rangle \,, \tag{26}$$

$$= \int_t \omega(t) \mathrm{d}t, \tag{27}$$

where the integral is taken over all branches of the tree, while the integrand $\omega(t)$ is a piece-wise function changing after every point substitution event.

5 Reproducibility - Supplementary Materials

The simulators written in `C++` are publicly available under MIT license at https://github.com/ThibaultLatrille/SimuEvol. The mathematical developments under the general case of an arbitrary additive trait and an arbitrary log-concave fitness function, and the derived susceptibility under various fitness functions, as well as supplementary figures, are available in supplementary materials. The scripts and instructions necessary to reproduce this study are available at https://github.com/ThibaultLatrille/GenotypePhenotypeFitness.

6 Author contributions

TL gathered and formatted the data, developed the new models in `SimuEvol` and conducted all analyses, in the context of a PhD work (Ecole Normale Superieure de Lyon). TL and NL both contributed to the writing of the manuscript.

7 Acknowledgements

We wish to thank Julien Joseph for whiteboard mathematical sessions. We gratefully acknowledge the help of Nicolas Rodrigue and Laurent Duret for their input on this work and their comments on the manuscript. This work was performed using the computing facilities of the CC LBBE/PRABI. Funding: French National Research Agency, Grant ANR-15-CE12-0010-01 / DASIRE.

References

Bastolla, U., Bruscolini, P., and Velasco, J. L. 2012. Sequence determinants of protein folding rates: Positive correlation between contact energy and contact range indicates selection for fast folding. *Proteins: Structure, Function and Bioinformatics*, 80(9): 2287–2304.

Bastolla, U., Dehouck, Y., and Echave, J. 2017. What evolution tells us about protein physics, and protein physics tells us about evolution.

Biesiadecka, M. K., Sliwa, P., Tomala, K., and Korona, R. 2020. An Overexpression Experiment Does Not Support the Hypothesis That Avoidance of Toxicity Determines the Rate of Protein Evolution. *Genome Biology and Evolution*, 12(5): 589–596.

Bloom, J. D. 2017. Identification of positive selection in genes is greatly improved by using experimentally informed site-specific models. *Biology Direct*, 12(1): 1.

Bloom, J. D., Raval, A., and Wilke, C. O. 2007. Thermodynamics of Neutral Protein Evolution. *Genetics*, 175(1): 255 LP – 266.

Brevet, M. and Lartillot, N. 2019. Reconstructing the history of variation in effective population size along phylogenies. *bioRxiv*, page 793059.

Brush, S. G. 1967. History of the Lenz-Ising model. *Reviews of Modern Physics*, 39(4): 883–893.

Cherry, J. L. 1998. Should We Expect Substitution Rate to Depend on Population Size? *Genetics*, 150(2).

Cherry, J. L. 2010. Expression Level, Evolutionary Rate, and the Cost of Expression. *Genome Biology and Evolution*, 2: 757–769.

Dasmeh, P. and Serohijos, A. W. R. 2018. Estimating the contribution of folding stability to nonspecific epistasis in protein evolution. *Proteins: Structure, Function, and Bioinformatics*, 86(12): 1242–1250.

Dasmeh, P., Serohijos, A. W., Kepp, K. P., and Shakhnovich, E. I. 2014. The Influence of Selection for Protein Stability on dN/dS Estimations. *Genome Biology and Evolution*, 6(10): 2956–2967.

Dixit, P. D. and Maslov, S. 2013. Evolutionary Capacitance and Control of Protein Stability in Protein-Protein Interaction Networks. *PLoS Computational Biology*, 9(4).

Dos Reis, M. 2015. How to calculate the non-synonymous to synonymous rate ratio of protein-coding genes under the fisher-wright mutation-selection framework. *Biology Letters*, 11(4).

Drummond, D. A. and Wilke, C. O. 2008. Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution. *Cell*, 134(2): 341–352.

Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., and Arnold, F. H. 2005. Why highly expressed proteins evolve slowly. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40): 14338–14343.

Duret, L. and Mouchiroud, D. 2000. Determinants of Substitution Rates in Mammalian Genes: Expression Pattern Affects Selection Intensity but Not Mutation Rate. *Molecular Biology and Evolution*, 17(1): 68–070.

Figuet, E., Nabholz, B., Bonneau, M., Mas Carrio, E., Nadachowska-Brzyska, K., Ellegren, H., and Galtier, N. 2016. Life History Traits, Protein Evolution, and the Nearly Neutral Theory in Amniotes. *Molecular Biology and Evolution*, 33(6): 1517–1527.

Figuet, E., Ballenghien, M., Lartillot, N., and Galtier, N. 2017. Reconstruction of body mass evolution in the Cetartiodactyla and mammals using phylogenomic data. *bioRxiv*, pages 139147, ver. 3 peer–reviewed and recommended by PC.

Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25): 2340–2361.

Goldstein, R. A. 2011. The evolution and evolutionary consequences of marginal thermostability in proteins. *Proteins: Structure, Function and Bioinformatics*, 79(5): 1396–1407.

Goldstein, R. A. 2013. Population Size Dependence of Fitness Effect Distribution and Substitution Rate Probed by Biophysical Model of Protein Thermostability. *Genome Biology and Evolution*, 5(9): 1584–1593.

Goldstein, R. A. and Pollock, D. D. 2016. The tangled bank of amino acids. *Protein Science*, 25(7): 1354–1362.

Goldstein, R. A. and Pollock, D. D. 2017. Sequence entropy of folding and the absolute rate of amino acid substitutions. *Nature Ecology & Evolution*, 1(12): 1923–1930.

Gout, J.-F., Kahn, D., and Duret, L. 2010. The Relationship among Gene Expression, the Evolution of Gene Dosage, and the Rate of Protein Evolution. *PLoS Genetics*, 6(5): e1000944.

Grantham, R. 1974. Amino acid difference formula to help explain protein evolution. *Science*, 185(4154): 862–864.

Halpern, A. L. and Bruno, W. J. 1998. Evolutionary distances for protein-coding sequences: modeling site-specific residue frequencies. *Molecular biology and evolution*, 15(7): 910–917.

Janin, J. 1995. Protein-protein recognition.

Jiang, Q., Teufel, A. I., Jackson, E. L., and Wilke, C. O. 2018. Beyond thermodynamic constraints: Evolutionary sampling generates realistic protein sequence variation. *Genetics*, 208(4): 1387–1395.

Jimenez, M. J., Arenas, M., and Bastolla, U. 2018. Substitution rates predicted by stability-constrained models of protein evolution are not consistent with empirical data. *Molecular Biology and Evolution*, 35(3): 743–755.

Jones, C. T., Youssef, N., Susko, E., and Bielawski, J. P. 2016. Shifting Balance on a Static Mutation–Selection Landscape: A Novel Scenario of Positive Selection. *Molecular Biology and Evolution*, 34(2): msw237.

Kimura, M. 1979. Model of effectively neutral mutations in which selective constraint is incorporated. *Proceedings of the National Academy of Sciences of the United States of America*, 76(7): 3440–3444.

Lanfear, R., Ho, S. Y., Love, D., and Bromham, L. 2010. Mutation rate is linked to diversification in birds. *Proceedings of the National Academy of Sciences of the United States of America*, 107(47): 20423–20428.

Lanfear, R., Kokko, H., and Eyre-Walker, A. 2014. Population size and the rate of evolution.

Lartillot, N. and Delsuc, F. 2012. Joint reconstruction of divergence times and life-history evolution in placental mammals using a phylogenetic covariance model. *Evolution*, 66(6): 1773–1787.

Lartillot, N. and Poujol, R. 2011. A Phylogenetic Model for Investigating Correlated Evolution of Substitution Rates and Continuous Phenotypic Characters. *Molecular Biology and Evolution*, 28(1): 729–744.

Manhart, M. and Morozov, A. V. 2015. Protein folding and binding can emerge as evolutionary spandrels through structural coupling. *Proceedings of the National Academy of Sciences of the United States of America*, 112(6): 1797–1802.

McCandlish, D. M. and Stoltzfus, A. 2014. Modeling Evolution Using the Probability of Fixation: History and Implications. *The Quarterly Review of Biology*, 89(3): 225–252.

Miyazawa, S. and Jernigan, R. L. 1985. Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules*, 18(3): 534–552.

Mustonen, V. and Lässig, M. 2009. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends in genetics*, 25(3): 111–119.

Nabholz, B., Uwimana, N., and Lartillot, N. 2013. Reconstructing the Phylogenetic History of Long-Term Effective Population Size and Life-History Traits Using Patterns of Amino Acid Replacement in Mitochondrial Genomes of Mammals and Birds. *Genome Biology and Evolution*, 5(7): 1273–1290.

Nikolaev, S. I., Montoya-Burgos, J. I., Popadin, K., Parand, L., Margulies, E. H., Antonarakis, S. E., Bouffard, G. G., Idol, J. R., Maduro, V. V., Blakesley, R. W., Guan, X., Hansen, N. F., Maskeri, B., McDowell, J. C., Park, M., Thomas, P. J., and Young, A. C. 2007. Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements. *Proceedings of the National Academy of Sciences of the United States of America*, 104(51): 20443–20448.

Ohta, T. 1972. Population size and rate of evolution. *Journal of Molecular Evolution*, 1(4): 305–314.

Ohta, T. 1992. The nearly neutral theory of molecular evolution. *Annual Review of Ecology and Systematics*, 23(1992): 263–286.

Plata, G. and Vitkup, D. 2017. Protein Stability and Avoidance of Toxic Misfolding Do Not Explain the Sequence Constraints of Highly Expressed Proteins. *Molecular Biology and Evolution*, 35(3): 700–703.

Pollock, D. D., Thiltgen, G., and Goldstein, R. A. 2012. Amino acid coevolution induces an evolutionary Stokes shift. *Proceedings of the National Academy of Sciences of the United States of America*, 109(21): E1352–E1359.

Popadin, K., Polishchuk, L. V., Mamirova, L., Knorre, D., and Gunbin, K. 2007. Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. *Proceedings of the National Academy of Sciences of the United States of America*, 104(33): 13390–13395.

Razban, R. M. 2019. Protein Melting Temperature Cannot Fully Assess Whether Protein Folding Free Energy Underlies the Universal Abundance–Evolutionary Rate Correlation Seen in Proteins. *Molecular Biology and Evolution*, 36(9): 1955–1963.

Rocha, E. P. C. and Danchin, A. 2004. An Analysis of Determinants of Amino Acids Substitution Rates in Bacterial Proteins. *Molecular Biology and Evolution*, 21(1): 108–116.

Rodrigue, N., Philippe, H., and Lartillot, N. 2010. Mutation-selection models of coding sequence evolution with site-heterogeneous amino acid fitness profiles. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10): 4629–34.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J. M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A. A.-T., Weinert, L. A., Belkhir, K., Bierne, N., Glémin, S., and Galtier, N. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, 515(7526): 261–263.

Sella, G. and Hirsh, A. E. 2005. The application of statistical physics to evolutionary biology. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27): 9541–9546.

Serohijos, A. W., Rimas, Z., and Shakhnovich, E. I. 2012. Protein Biophysics Explains Why Highly Abundant Proteins Evolve Slowly. *Cell Reports*, 2(2): 249–256.

Serohijos, A. W., Lee, S. Y., and Shakhnovich, E. I. 2013. Highly abundant proteins favor more stable 3D structures in yeast. *Biophysical Journal*, 104(3): L1–L3.

Song, H., Gao, H., Liu, J., Tian, P., and Nan, Z. 2017. Comprehensive analysis of correlations among codon usage bias, gene expression, and substitution rate in Arachis duranensis and Arachis ipaënsis orthologs. *Scientific Reports*, 7(1): 1–12.

Spielman, S. J. and Wilke, C. O. 2015. The relationship between dN/dS and scaled selection coefficients. *Molecular biology and evolution*, 32(4): 1097–1108.

Starr, T. N. and Thornton, J. W. 2016. Epistasis in protein evolution.

Tamuri, A. U. and Goldstein, R. A. 2012. Estimating the distribution of selection coefficients from phylogenetic data using sitewise mutation-selection models. *Genetics*, 190(March): 1101–1115.

Taverna, D. M. and Goldstein, R. A. 2002. Why are proteins marginally stable? *Proteins: Structure, Function, and Bioinformatics*, 46(1): 105–109.

Welch, J. J., Eyre-Walker, A., and Waxman, D. 2008. Divergence and polymorphism under the nearly neutral theory of molecular evolution. *Journal of Molecular Evolution*, 67(4): 418–426.

Wilke, C. O. and Drummond, D. A. 2006. Population Genetics of Translational Robustness. *Genetics*, 173(1): 473 LP – 481.

Williams, P. D., Pollock, D. D., Blackburne, B. P., and Goldstein, R. A. 2006. Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Computational Biology*, 2(6): 0598–0605.

Yang, J. R., Liao, B. Y., Zhuang, S. M., and Zhang, J. 2012. Protein misinteraction avoidance causes highly expressed proteins to evolve slowly. *Proceedings of the National Academy of Sciences of the United States of America*, 109(14): E831–E840.

Yang, Z. and Nielsen, R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution*, 46(4): 409–418.

Zeldovich, K. B., Chen, P., and Shakhnovich, E. I. 2007. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41): 16152–16157.

570 Zhang, J. and Nielsen, R. 2005. Evaluation of an improved branch-site likelihood method for detecting
571 positive selection at the molecular level. *Molecular biology and evolution*, 22(12): 2472–2479.

572 Zhang, J. and Yang, J. R. 2015. Determinants of the rate of protein sequence evolution.

573 Zhang, J., Maslov, S., and Shakhnovich, E. I. 2008. Constraints imposed by non-functional protein–protein
574 interactions on gene expression and proteome size. *Molecular Systems Biology*, 4(1): 210.

575 Zhang, X., Perica, T., and Teichmann, S. A. 2013. Evolution of protein structures and interactions from the
576 perspective of residue contact networks.