3.2 On the importance of being Constant: a population genetics note about the influence of high order complementary epistasis

In 1895, Oscar Wilde wrote The Importance of Being Earnest, ridiculing Victorian conventions. Far from these considerations, it was shown that the biological relevance of being earnestness is very dependent on the sex of an organism and the natural mating system of the species to which it belongs (Wade et al., 1980; Darwin, 2008; Shuster, 2009; Tilquin et al., 2016). The same holds when one substitutes this first name for its French counterpart "Constant". Since Biology is about being adapted, its fitness may rely on constantly changing its phenotype when the environment varies, or, on the contrary, may require one to target one or few (in the case of balancing selection) constant optimal phenotypic values or ratios. For illustrative purpose, let us slip ourself in the shoes of Alice pilgriming in her Wonderland. When she suddenly grows up, her whole body extends, not only one specific part (Carroll, 1866). Were it not to be the case, this body would appear completely unfit due to simple physical properties such as the Fundamental principle of the dynamics of rotation regarding joints. Such conservation of scaling and relative proportions has been extensively studied for one century since the seminal work of D'arcy Thompson (Bonner et al., 1992) through the concept of allometry (Gould, 1966; Cheverud, 1982; Damuth, 2001; Dill et al., 2011), shedding light on these scaling laws (West et al., 2005), their range of variation and the reasons behind both their variability and similarity (Pélabon et al., 2014). We have shown previously that such constancy ratios should also been maintained along metabolic pathways since the final flux depends on the efficiency of all proteins involved in it. Consequently, it seems relevant to evaluate how this need for regularity impacts population genetics. This is what we propose to address in the following pages.

3.2.1 An introduction on epistasis, complexity and the need for complementarity

About complex genetic interactions

"*On rencontre sa destinée souvent par les chemins qu'on prend pour l'éviter*".

Jean de la Fontaine

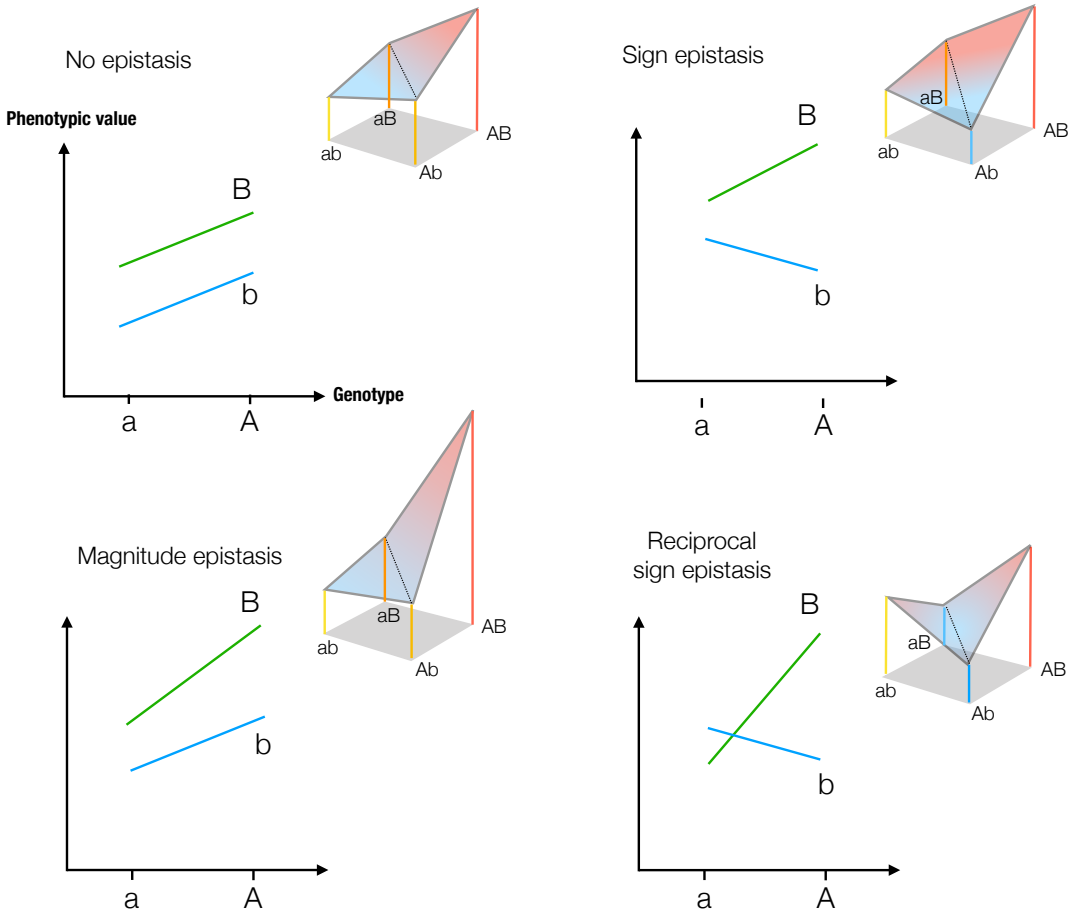

Figure 3.35: Main types of two-locus epistasis. First, no epistasis corresponds to the simple additivity model where loci (a/A or b/B) contribute the same amount to the trait no matter the value of a second locus involved in the trait. Magnitude epistasis depicts the case where synergistic or antagonistic effects come into play so that the effect of the combination is respectively increased or decreased compared to the expected additivity. With sign epistasis, the sign of the contribution brought by the second locus changes with the value of the first locus, which narrows the path towards the highest fitness. Reciprocal sign epistasis coincides with an extreme case of sign epistasis where both intermediate mutants (aB or Ab) are deleterious, giving birth to a fitness valley separating two local peaks.

Epistasis (see Figure 3.35) is a ubiquitous phenomenon in which the effect of a mutation differs depending on the genetic background in which it occurs (Bateson, 1909; Phillips, 2008; Domingo et al., 2019). As Weinreich et al. (2013) pointed out, it is a measure of our “surprise" insofar as we *a priori* expect mutational effects to be additive (Phillips, 2008; Domingo et al., 2019). Epistasis has long been known to occur between pairwise mutations where the combined effect between them results in a phenotype or a fitness not being the sum of that they would have in isolation from one another (Bateson, 1909). If Adaptation can sometimes still occur through few genetic changes (Orr, 2005), such epistatic interactions in general comes with several consequences. Forasmuch as it influences genetic architecture (Hermisson et al., 2003; Domingo et al., 2019; Sella et al.,

2019), it often narrows the path towards adaptive phenotypes (Poelwijk et al., 2007) and may even create fitness valleys under certain circumstances called reciprocal sign epistasis (Weinreich et al., 2005b; Poelwijk et al., 2011) - see Figure 3.35 for visual explanations.

Because of the process of genetic drift (Wright, 1930; Kimura, 1958; Ohta, 1992; Sella, 2009) that entails a mutational load (Haldane, 1937; Muller, 1950; Agrawal et al., 2012), lower effective population sizes (as was shown again for enzyme efficiencies in our work - see 3.1.1) are, on average, further from the optimum phenotype. This is notably the case when a species experiences a bottleneck (Wright, 1931; Nei et al., 1975) during which part of the genetic variability is lost even if adaptive. It was thus supposed that small effective population sizes could help escape from a local fitness peak by facilitating fixation of intermediate deleterious mutations (Wright, 1930; Wright, 1932) ; therefore, subdivision in small populations should be better at finding the highest peak though they are less efficient to climb up towards this peak. In a sense, as Jean de La Fontaine once put it, one may often meet his (Adaptive) Destiny on the (Neutral) route he took to escape from It. However, introducing polymorphism and recombination disproved this conclusion (Weinreich et al., 2005a) since combined mutations can either be found through stochastic tunneling (Iwasa et al., 2004b) – whereby a population jump to the fixation of an adaptive double mutation thanks to the transient segregation of a first deleterious mutation – or be brought together after having emerged in different lineages, and it was later shown that considering the width – *i.e.* the number of loci making up the valley – of valleys would even reverse Wright's initial intuition with higher populations more prone to cross large fitness valleys, especially if these valleys are much alike plateaus (Weissman et al., 2009).

Known as high(er)-order epistasis[2], this phenomenon involving many interacting loci usually gives rise to rugged fitness landscapes (Kauffman et al., 1987; Kauffman et al., 1989; Weinreich et al., 2013) where peaks and valleys quickly follow each other in both the phenotype and the genotype spaces, as can be observed in Figure 3.38-C. Based on the NK-model (see Box 3 to get to the nitty gritty of the model), Kauffman et al. (1987) showed that such high order epistasis – assimilated to a kind of complexity – comes at a great cost for fitness since more ruggedness in turn increases the probability of being trapped at a local optimum (Kauffman et al., 1987; Geard et al., 2002), a finding that was further confirmed when accounting for the possibility of neutral mutations, through NKq and NKp models (Barnett, 1998; Newman et al., 1998; Geard et al.,

[2]see Illustrated Glossary of complex genetic interactions in Box 2

2002). Under certain assumptions, this model makes possible to estimate analytically the average expected mutation-selection-drift balance (Weinberger, 1991) and its mutational load counterpart, which can thus be contrasted to other such predictive frameworks.

Box 1. A definition of high order epistasis

High order epistasis is the general case of epistasis where the interaction involves a great number of interacting parts owing to processes detailed in Box 2. It features the effect of the genomic background on lower order interactions (see Figure 3.37); noticeably, it was shown that the magnitude of epistasis sometimes seems to decline with its order (Ferretti et al., 2016; Weinreich et al., 2018), which means that high order interactions may be overlooked in certain cases.

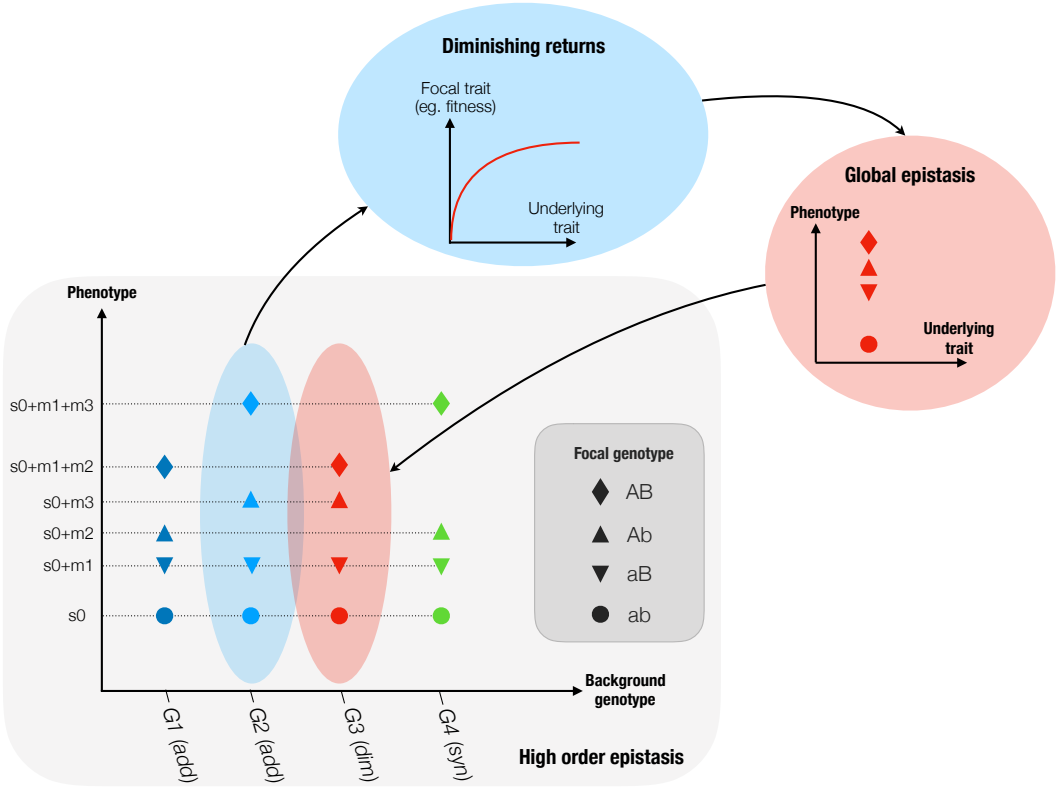

Figure 3.36: In **High order epistasis**, the genetic background (denoted G) modifies how loci combine to produce a phenotype (*eg.* $G1$ *vs* $G2$, with a lower level interaction involving two loci in the picture). Synergistic epistasis ('syn', occuring in $G4$) means that mutations produce a higher effect on the phenotype than the expected additive one. Diminishing-returns epistasis ('dim') stands for any mutational interaction in which the effect of combined advantageous mutations is less than the sum of their isolated effect. As a corollary, **global epistasis** is a special case of diminishing returns epistasis ('dim', occuring in $G3$) and occurs when several loci combine additively ('add') on an underlying - possibly unobserved - trait that influences an observed trait or phenotype (including fitness) through a diminishing returns relationship.

Box 2. Illustrated glossary on the processes behind high order epistasis

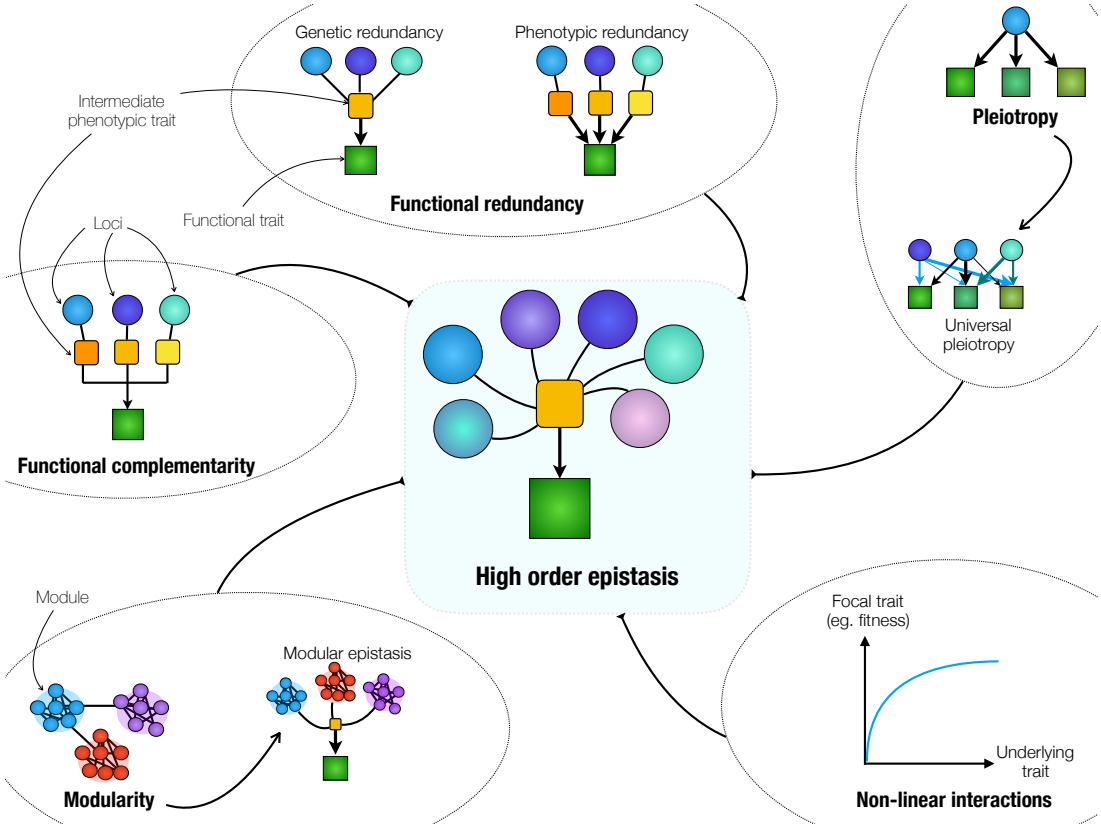

Figure 3.37: Summary of the genetic interactions involved in the emergence of high order epistasis. **Functional complementarity** results from the contribution of different intermediate phenotypic traits that each need be present in order to produce a functional phenotype: when different loci are involved in the intermediate traits, this gives rise to complementary epistasis (see also Figure 3.40). This is true for example when a specific color results from the involvement of several pigments. **Functional redundancy** describes the existence of multiple ways - often involving sub-phenotypic traits (***Phenotypic redundancy***) - to optimize a phenotypic trait (eg. concentration, affinity and catalytic rates for enzyme efficiency). ***Genetic redundancy*** is also possible, with distinct loci involved in the same phenotypic trait (*eg.* loci contributing to protein stability, or different genes involved in the same function) and may follow from duplication and sub-functionalization, for instance. **Pleiotropy** represents a process in which a specific locus (*eg.* single residue) influences several traits at once. Owing to the numerous interactions occuring in gene newtorks, many loci are assumed to influence many traits at once: this idea has been coined ***Universal pleiotropy***. Loci - or underlying phenotypic traits - may combine through **Non-linear interactions**, akin the diminshing returns law shown here, to produce a focal trait. **Modularity** defines the existence of relatively independent modules (*eg.* the arm and the leg) - in that they contribute independently to the phenotype - that are comprised of tightly interacting parts. Combined with epistasis, it extends the latter concept into ***Modular epistasis*** where each functional units (modules) display purely buffering or aggravating interactions.

Indeed, the influence of complexity on evolutionary trajectories and outcomes has in parallel thoroughly been addressed through the lens of Fisher's geometrical model (Fisher, 1930; Orr, 1998; Orr, 2000; Poon et al., 2000; Martin et al., 2006; Tenaillon et al., 2007; Sella, 2009; Tenaillon, 2014). Within this framework, epistasis builds up from the mathematical underpinnings (Hartl et al., 1996; Orr, 2000; Tenaillon, 2014; Hwang et al., 2017) and the assumption of both non-linearities in the fitness landscapes(Hwang et al., 2017) and pleiotropy preexistence (Tenaillon, 2014) - see Box 4 for a description of Fisher's geometrical model and (Stearns, 2010) for a review on pleiotropy.Besides enabling to study the distribution of fitness effects (Martin et al., 2006; Lourenço et al., 2011), it is in principle possible through this framework to approach how complexity emerges and evolves (Orr, 1998; Martin et al., 2007; Gros et al., 2009; Le Nagard et al., 2011), what it really portrays[3], and even to try and derive that of an organism from the strength of drift experienced by an organism (when adopting a top-down approach) (Tenaillon et al., 2007). This is because epistasis and pleiotropy are intrinsic features to the multidimensional formulation of the model (Tenaillon, 2014) such that complexity needs not be defined in terms of the unknown explicit genotypic-phenotypic relationships, but can on the contrary be understood as the factor limiting the strength of selection when adopting a top-down approach (Le Nagard et al., 2011). Interestingly, Hwang et al. (2017) have recently disputed how we interpret the ins and outs of the model by shown, demonstrating that the amount of sign epistasis is contingent on the position in the landscape (see (C) in 3.39 to get a sense of why it is so); as well as it decreases with the predefined phenotypic dimensionality n. Yet more significant is their demonstration that fitness landscape complexity does not, in general, correlate with the n, which sheds light on the fact that complexity may take different irreducible and non-fungible forms.

Though fascinating, both these frameworks lack a mechanistic basis as highlighted recently by Martin (2014) and Yi et al. (2019), even if Fisher's geometrical model was specifically revived for this purpose by Hartl et al. (1996) and subsequently led to several insights about the mutational load and the expected distribution of mutation fitness effects (Poon et al., 2000; Martin et al., 2006; Martin et al., 2007). If we are to succesfully achieve a functional synthesis (Dean et al., 2007) predicated on an integrative approach (Gudelj et al., 2010), such approaches deserve, at least, to be informed by the very biological processes giving rise to genetic interactions. Otherwise, disentangling causes from effects in Evolution will remain to the state of wishful thinking.

[3]In a broad sense, it can cover independent phenotypic traits, epistasis, pleiotropy among other (and more organismic) features.

For instance, it has been shown that the distribution of fitness effects can be captured using such models (Martin et al., 2006; Huber et al., 2017), but how alternate and potentially more parsimonious explanations could account for similar patterns is largely unknown (Lourenço et al., 2011), a case that can also be made for fitness landscapes (Blanquart et al., 2016). In the same vein, determining how features such as modularity (Wagner et al., 2007; Segrè et al., 2005; Hartwell et al., 1999) or niche construction (Bajić et al., 2018) changes the landscape and whether it is a specific kind of intrinsic interactions or the evolutionary product that these interactions favour or even made necessary requires to understand how they all dynamically behave together, as was argued for molecular networks (Alexander et al., 2009).

This is especially significant - and, noticeably, a major challenge for biological understanding (Young et al., 2019) - inasmuch as the genotype-phenotype-fitness map stems from the intertwining between interacting genes - susceptible to evolve (Gros et al., 2009) during the course of Adaptation - with their basic and emerging physiochemical properties (Bershtein et al., 2017; Bergelson et al., 2021). Lately, shy first steps to fill in this theoretical gap have adopted the prolific paradigm of statistical physics - also used in (Sella et al., 2005) to determine genotype distribution at mutation-selection-drift balance - in an interesting attempt to derive the isotropic instance of Fisher's geometrical model from first principles (Martin, 2014) that draws inspiration from results of systems biology such as those of the FBA (Flux Balance Analysis) (Orth et al., 2010). However, FBA should not be the most appropriate framework to study their past Evolution for several reasons. First, FBA relies on an assumption of optimality to solve the systems of equations describing the process as well as on the existence of fixed molecule contents; moreover, such an assumption comes with a corollary that yield should be optimize rather than nutrient consumption, which is not always relevant from an eco-evolutionary standpoint [4]. Perhaps more importantly, because this framework does not deal with real mechanisms but only describes a complex system that has already evolved, it does not seem ideal to capture how genetic interactions evolved in the first place. On another hand, it seems nonetheless very well suited to tackle how these existing systems may react to changes and to give clues about their evolution from a given, already very complex point.

[4] See for example Schuster et al. (2008) .

Box 3. A brief note on Kaufman's NK model

NK models are a class of models which were initially built to understand the influence of complex epistasis interactions on the shape of the fitness landscape and the ability to adapt (Kauffman et al., 1987) but also eventually lead to progress in computer science and strategy sciences. Within this framework, an organism's fitness is simply the mean of all its components fitness. Its core assumptions are also very simple since only two parameters originally influence the fitness landscape: N denotes the number of components that contribute to fitness while K stands for the number of components influencing the fitness $c_i(i; K\text{other elements of the set of components})$ of each component (Csaszar, 2018). Each of these entities can exist into two forms ('0' and '1'). For example, let us say that the fitness of a fictional organism depends only on $N = 2$ traits which would be its brain size (small or large) and its body size (small or large). If $K = 1$, c_i is a each trait is independent such that if a large brain and a large body are adaptive, there is no epistasis and the fitness is just the average of that conferred by each of these features. However, if $K = 2$, the fitness component brought by both of them depends on the other trait value, which seems more realistic, as a large shape with a small brain to use it should be deleterious. For 2 dimensions, fitness can be represented through the four nodes of a square. Increasing N means that they map onto a N-dimensional hypercube, where the number of fitness maximum depends on K. This framework has largely been used to explain that epistasis complexity yields rugged fitness landscape as depicted in (C) of Figure 3.38. Although the framework is fascinating and has been very fruitful on several key issues, the latter analogy is not to be taken too literally, if taken at all, because it is not possible to assess whether the exploration of the N-dimensional rugged landscape is similar to the same process in an apparent 2 dimensional counterpart.

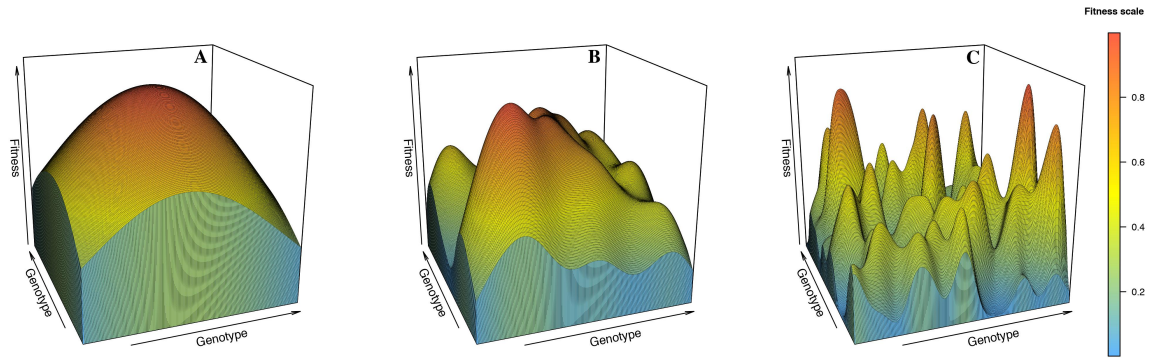

Figure 3.38: Illustration of the influence of complexity in NK models. When complexity is low ($K \approx 1$), the fitness landscape is smooth (A). Increasing K (toward the right) goes along with an increase in ruggedness of the landscape, which suggests that organisms can be trapped at local optimum - for instance in (C).

Box 4. A brief description of Fisher's geometrical model

In Fisher's geometrical model, a phenotype is described by a set of t idealized traits, each being independent from one another so that they are to be optimized specifically (Fisher, 1930) (see A below, in Figure 3.39). These traits can consequently be depicted as independent axes in the euclidean space of the relevant dimensionality, such that fitness isoclines are represented by hyperspheres of dimension t (Tenaillon, 2014). Phenotypes are subject to stabilizing selection towards an optimum: the more an organism approaches the optimum, the lower the selective pressure is. Fisher (1930) used this model to set the stage for quantitative genetics by showing that it is more efficient to rely on a large amount of loci with small effects than the opposite as small mutational hyperspheres display very few bias towards deleterious mutations on the contrary to large ones (see B below). (Global) epistasis is also an intrinsic feature of the framework (see C below for details).

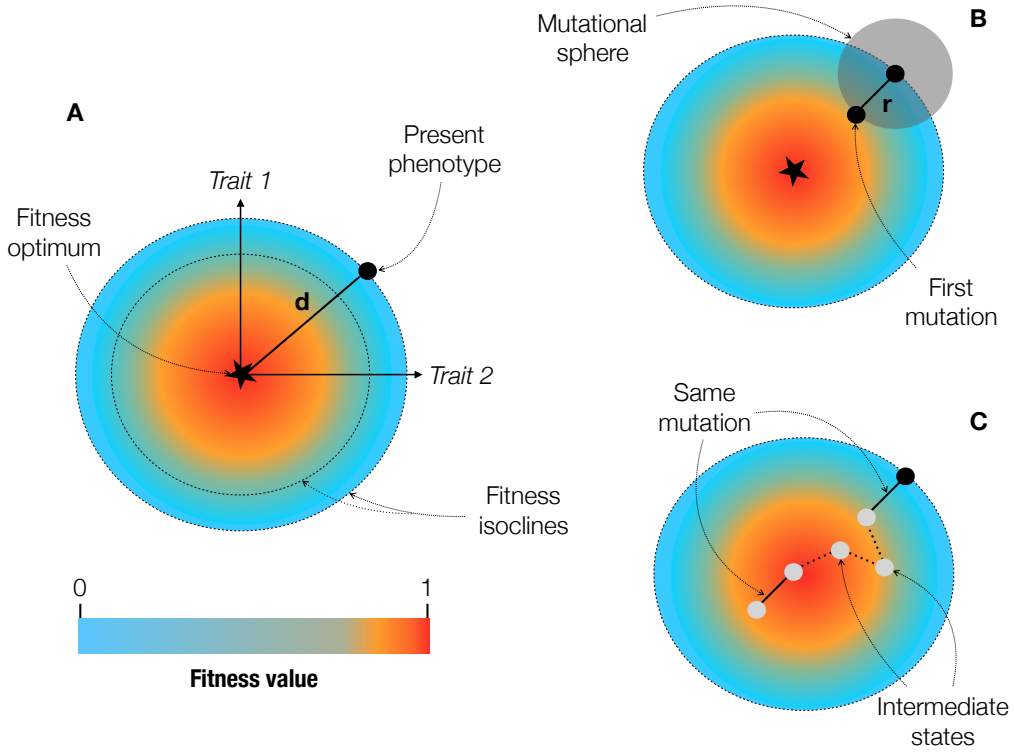

Figure 3.39: Illustrating the framework with the two dimensional isotropic Fisher's geometrical model. The 2D fitness landscape can be described by a circle whose center represents the optimum under stabilizing selection - denoted by the star - and is surrounded by concentric fitness isoclines. This is because fitness decreases identically along any axis in the isotropic instance of the model. As shown in (B), mutations are also considered isotropic and can be represented by a sphere of radius r: the larger this sphere, the larger the disequilibrium between advantageous mutations (grey blue area in (B) where the two spheres overlap) and deleterious ones (grey area in (B) where mutations pull organisms further from the optimum). (C) shows that mutations with exactly similar effects on a trait may have very different impacts on fitness depending on the position in the phenotypic space : epistasis is thereby built from the core assumptions of the model (Hwang et al., 2017).

Putting forward the concept of high order complementary epistasis

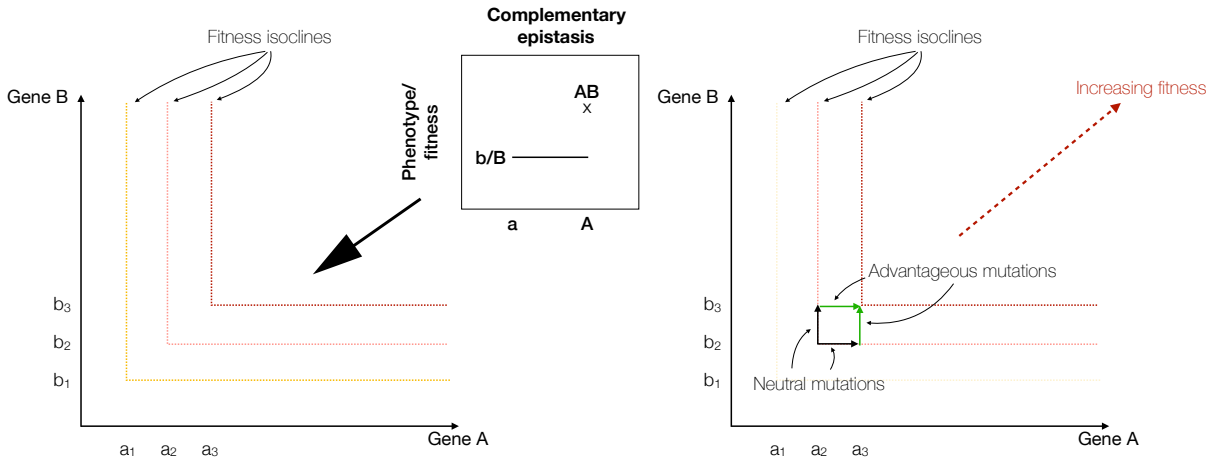

Figure 3.40: Description of complementary epistasis: on the left panel, its effect in the classical a/A-b/B case is shown in the short window, with the need for both A and B mutations to experience any gain of fitness, while the larger plot represents how phenotype/fitness isoclines can be mapped in the genotype space when more mutations are considered. On the right panel is also shown the mutational path that allows to gain extra fitness: first, a potentially advantageous but actually perfectly neutral mutation needs - at least - to exist on either one of the locus before the advantageous mutation can occur and give an actual extra push of fitness. This phenomenon can involve higher order interactions leading to the need for the segregation of numerous neutral mutations, which may besides be subject to mutational biases.

Be that as it may, it has not yet been determined how the combination of complementary epistasis (Crow et al., 1970; Sackton et al., 2016) – where a phenotypic trait can only be competitive if each and any of its underlying loci are (see Figure 3.40 for more details) – with global epistasis (Otwinowski et al., 2018) could specifically change population genetics predictions arising from the Neutral Theory of Evolution and its extensions (Kimura, 1968; Ohta, 1973; Ohta, 1992) while this process seems to be to be strongly supported by mechanistic underpinnings (Kacser et al., 1973; Hartl et al., 1985; Yi et al., 2019; Taverna et al., 2002; Bloom et al., 2005; Labourel et al., 2021) and Crutchfield et al. (2000)'s work suggested that neutral diffusion in the genotype space can guide the evolution towards higher complexity[5]. This is what we propose to do in what follows; to justify the interest of such an investigation and its consequences, we start with the presentation of a toy model based on simplistic assumptions from global epistasis. We then put forward an analytical treatment dealing with a simple instance of the process. Finally, we show how this

[5]In this line of work, this neutral diffusion drives the finding of a higher complexity attraction subbasin, which gives rise to (quick) epochal evolution followed by stasis, echoing the pioneering and controversial work on punctuated equilibrium by Gould and Eldredge (Eldredge et al., 1971; Eldredge et al., 1997).

framework may have a profound impact on our understanding of Evolution at different levels of biological organization and put forward the key components that a more complete instance of the model should include.

3.2.2 A first approach

A simplistic prediction of classical population genetics

We expect that complementary epistasis should influence both the speed of adaptation, as was observed for evolutionary escape (Weinreich et al., 2005a; Weissman et al., 2009) and the mutational load (see introduction for previous approaches on this idea), and focus here on this latter phenomenon. More particularly, we assume that fitness should decrease with the number of units/loci involved and that the mutational bias should play a major part in hampering Adaptation because it influences the preexistence of potential complementary beneficial mutations (at loci which are drifting because they differ from the worst one). Nevertheless, it should be possible to derive a rough prediction from the underlying premises of such a phenomenon.

Based on results of the neutral theory of Evolution (Kimura, 1962; Ohta, 1973), we indeed know that Natural Selection cannot screen mutations whose selective effect $|s|$ is below $1/N_e$ for haploid populations. If fitness is limited by a maximum value, this means that deleterious mutations with effects $|s| \approx 1/N_e$ evolve through genetic drift and that fitness at the mutation-selection-drift balance establishes around $\prec f_1^* \succ \approx 1 - 1/N_e$ when mutations are mostly deleterious and one locus is considered (Kimura, 1958). Let us say that an organism starts with the maximum possible fitness. Through drift on the first gene, f should decrease on average to approximately $1 - 1/N_e$. Drift on the second gene should again push fitness downwards since the maximum fitness relatively to which drift occurs is now set to $1 - 1/N_e$, such that $\prec f_2^* \succ \approx (1 - 1/N_e) \times (1 - 1/N_e) = (1 - 1/N_e)^2$. Therefore, considering n complementary genes yield the following prediction $\prec f_n^* \succ \approx (1 - 1/N_e)^n$. If $N_e >> n$, this can be summarized by $\prec f_n^* \succ \approx 1 - L^*$ using the first order Taylor expansion, with $L^* = n/N_e$ denoting the mutational load. In that case, this rough estimate is comparable in size to that from the Fisher's geometrical model in a N-dimensional space – namely $L^* \approx n/(n + 2N_e)$ – (Hartl et al., 1996; Poon et al., 2000; Sella, 2009) that considers the influence of complexity on the evolutionary balance.

A toy evolutionary model

In this toy model, we consider a fitness landscape subject to saturation due to diminishing returns epistasis (Tokuriki et al., 2012; Kaltenbach et al., 2014) as has widely been documented for proteins in the case of stability (Taverna et al., 2002; Bloom et al., 2005; Bloom et al., 2006; Kaltenbach et al., 2014) and catalytic efficiency (Dykhuizen et al., 1987; Hartl et al., 1985; Yi et al., 2019; Labourel et al., 2021). It has also been recently shown that global diminishing returns epistasis (Kryazhimskiy et al., 2014; Bahcall, 2014; Otwinowski et al., 2018) should arise for a complex trait as a by-product of the distribution of fitness effects (Reddy et al., 2020). Such a fitness landscape can typically be described through a sigmoid function whose shape is given by the following equation (similar to that of Michaelis Menten):

$$f(x) = \frac{x}{x + K_X} \tag{3.23}$$

The evolutionary process is then simulated using the probability of fixation (McCandlish et al., 2014), which is a classical result in population genetics (Haldane, 1927; Kimura, 1962; Wright, 1931). Within this simplified framework, neither clonal interference nor double/multiple mutants are considered, meaning that the fixation process concerns only one mutation (on one gene) at a time through a pairwise competition. Under this assumption, the probability that a mutation occurring in a haploid population is eventually fixed is given by:

$$P_{\text{fix}}(s, N_e) = \frac{1 - e^{-2s}}{1 - e^{-2N_e s}} (\approx \frac{2s}{1 - e^{-2N_e s}}, \text{ when } s << 1) \tag{3.24}$$

Let us say that $\mathbf{X} = (X_1, ..., X_n)$ is the vector representing a state of the pool of n complementary gene[6] where X_i denotes the phenotypic value of the i^{eth} gene, and that $s'_m = \frac{f(X'_m) - f(X_m)}{f(X_m)}$ denotes the potential selective value of a mutation X'_m with fitness $f(X'_m)$ occurring on the gene m whose current value in the population is X_m (and fitness $f(X_m)$). The fitness function detailed above therefore determines the maximum fitness a gene can potentially induce (*e.g.* the maximum catalytic flux an enzyme may be able to sustain without incurring costs). The actual selective value of a mutation, and its evolutionary fate, depends on the whole genetic background $\mathbf{X}$ within which it occurs owing to the specific process of complementary epistasis that changes the selective

[6]For convenience, we only mention genes, but as stated previously, it may also apply to loci or organismic units, for instance.

effect this mutation provides to its carrier. Two cases have to be distinguished, as we specified below and on Figure 3.41.

First, if the mutation affects the less efficient gene of a pool of complementary genes – *i.e.* $X_m = \min_{i \in S_g}(\mathbf{X})$, such that $f_{\mathbf{X}} = f(X_m)$, with S_g the set of genes involved in the phenotypic set $\mathbf{X}$ – the probability of fixation of the mutation X'_m rises only up to the threshold where it is no longer the worst gene among the pool. It yields:

$$P_{\text{fix},X'_m} = \begin{cases} P_{\text{fix}}(s'_m, N_e), \text{ when } s'_m \leq \Delta s_{\{m,max\}} \\ P_{\text{fix}}(\Delta s_{\{m,max\}}, N_e), \text{ otherwise,} \end{cases} \tag{3.25}$$

where $\Delta s_{\{m,max\}} = \min_{\substack{i \in S_g \\ i \neq m}} \frac{f(X_i)}{f(X_m)} - 1$ (with $\Delta s_{\{m,max\}} \geq 0$).

Conversely, if the mutation influences the phenotypic value of any other gene – *i.e.* $X_m > \min_{i \in S_g}(\mathbf{X})$, with S_g the set of genes involved in the phenotype $\mathbf{X}$ – its probability of fixation is that of a perfectly neutral mutation as long as the phenotypic value for this gene remains above the minimum of the set while it is that of a disadvantageous mutation – relatively to this threshold – when it falls under it, such that:

$$P_{\text{fix},X'_m} = \begin{cases} 1/N_e, \text{ when } s'_m \geq \Delta s_{\{m,min\}} \\ P_{\text{fix}}(s_{\{act,m'\}}, N_e), \text{ otherwise,} \end{cases} \tag{3.26}$$

where $\Delta s_{\{m,min\}} = \min_{i \in S_g} \frac{f(X_i)}{f(X_m)} - 1$ (with $\Delta s_{\{m,min\}} \leq 0$) and $s_{\{act,m'\}} = \frac{f(X'_m)}{\min_{i \in S_g} f(X_i)} - 1$.

We considered the influence of different distribution of fitness effects by drawing phenotypic values X'_m according to the following equation :

$$\log_{10}(X'_m) \sim \mathcal{N}(\log_{10}(X_m) - b, \sigma_X), \tag{3.27}$$

where b represents the intensity of the bias towards deleterious mutations. To comply with estimates on biological phenotypic traits, be they catalytic constants (Carlin et al., 2016) or gene expression (Metzger et al., 2016; Hodgins-Davis et al., 2019), the distribution of mutational phenotypic effects is modelled through a Gaussian distribution whose mean depends on the present value of the trait X_m at the loci m and a variance σ_X^2 (and the aformentionned mutational bias).

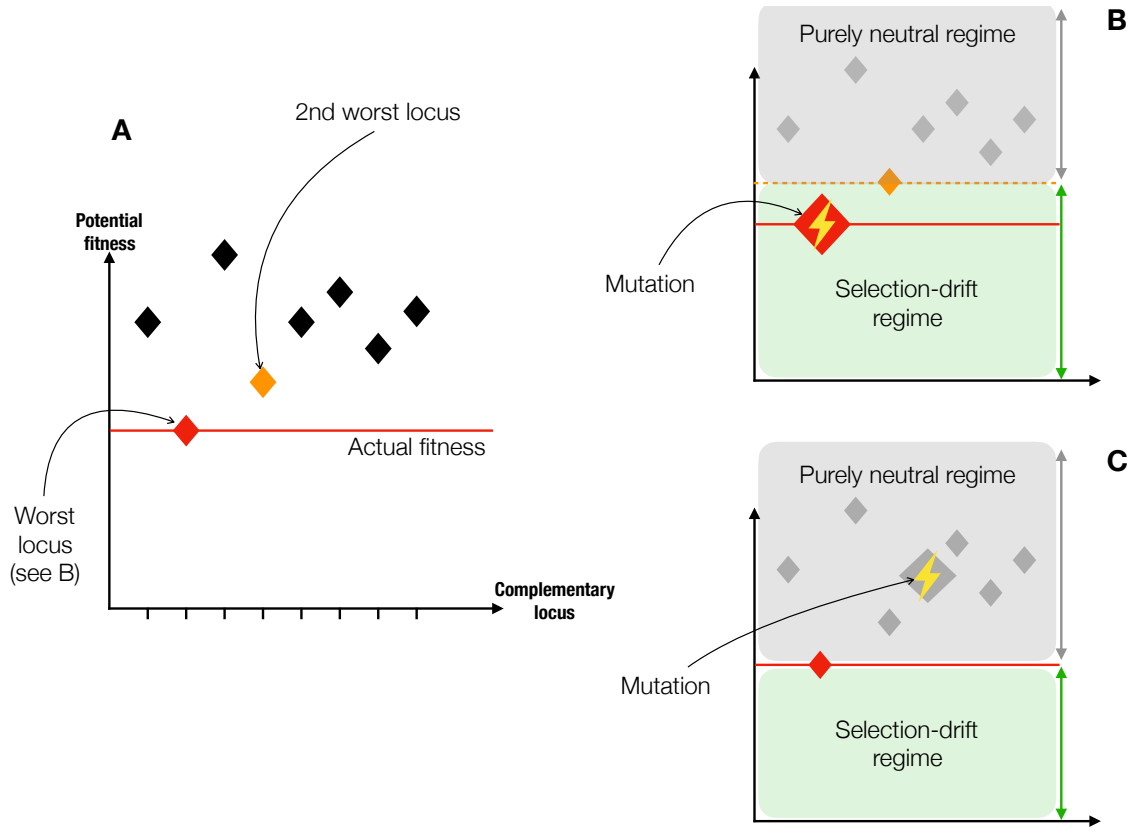

Figure 3.41: Explanations on how the probability of fixation depends on the locus affected by a mutation: as shown in A, the actual fitness is determined by the potential fitness of the worst locus. In accordance with that, two cases are to be distinguished : if the mutation hits the worst one (B), it evolves under selection-drift regime (with P_{fix} corresponding to the single locus model expectation) up until its potential fitness reaches the 2^{nd} worst one of the set of complementary loci, where any extra improvement would have the same selective value as that provided by the difference between the two worst loci; on the other hand, if the mutation hits one of the other (then the worst one) loci, it evolves under pure neutrality ($P_{fix} = 1/N_e$ for haploid Wright-Fisher process) down to the actual fitness because it does not affect this latter, and it is only below this threshold that it can be counter-selected according to the drift-selection regime.

We examined different situations ranging from cases where no mutational bias exists ($b = 0$) to some with high ones ($|b| = \lambda X_m$, where $\lambda >> 0$). Both because it seems more realistic (Carlin et al., 2016) and because the optimization process would otherwise be far longer, mutations are drawn for the log_{10} value of the trait in such a way that those affecting higher phenotypic values have a proportionally higher variability and are more biased (when the bias is not null).

Note that there exists a broad scientific literature on the distribution of fitness effects of mutations (Keightley et al., 2007; Orr, 2003; Gillespie, 1984) and some previous theoretical arguments for it (see for instance Martin et al., 2006 and Rice et al., 2015), but we are here interested in the

making of these effects from underlying causes and cannot, as a consequence, use distributions that result from evolved phenotype-fitness maps. We shall later discuss this point as a direct perspective of the present project.

Simulation results

As an attempt to evaluate the relevance of studying this phenomenon, we tested as a premise the influence of complementary epistasis on the mutational load. To do so, we set $K_X = 1$ in equation (3.23). Three mutational biases were considered $(0;-\sigma_X;-2\sigma_X)$ ranging from no bias to a high one through which only 1 mutations out of 40, approximately, is advantageous. We also studied two different levels of phenotypic variability among mutants for the log_{10} of X'_m ($\sigma_X = 0.25$ and $\sigma_X = 0.5$), with the highest one producing advantageous mutations with greater selective effects while decreasing the relative pool of slightly deleterious ones. These sets of parameters are in line with the aforementioned estimates for some distribution of phenotypic effects of mutations (Carlin et al., 2016; Metzger et al., 2016). The initial value of each complementary phenotypic trait was set to $X_i = X_0$ for any gene, with $X_0 = 1 - 1/N_e$ to start from the null hypothesis of nearly-neutral Evolution under high deleterious bias (but only one module). This starting point thus coincides with the classical Wright and Kimura's expectation and should outstrip the final value as any extra epistasis comes at a cost. Finally, the probability of fixation was computed through either equation (3.25) or equation (3.26) depending on the locus which was affected and the simulations were ran for an average of 10^3 mutation events per gene per ideal Wright-Fisher individual (*eg.* if $N_e = 10^2$, 10^5 mutations are drawn per gene).

We simulated the evolutionary process for $N_e = 10$ (see Appendix section at the end of this subsection) and $N_e = 100$ and present the results obtained for the mutational load in this latter case. In line with expectations, we observed that the number of complementary genes (denoted under the generic term of modules in the figure) severely impairs the strength of Natural Selection, with a decrease of the order of $n_{mod} \times 1/N_e$ when mutational biases are considered. However, this is not the case when no bias is considered with a decrease being far more limited that does not jeopardise the optimisation process albeit marginally - see (A) and (B) on Figure (3.42). Because deviations from the expected steady-state can accumulate, the balance is also more sensitive to the variability of mutational effects, with lower variability coming with a predictable decreased fitness - due to an increased amount of slightly deleterious mutations and a decreased amount

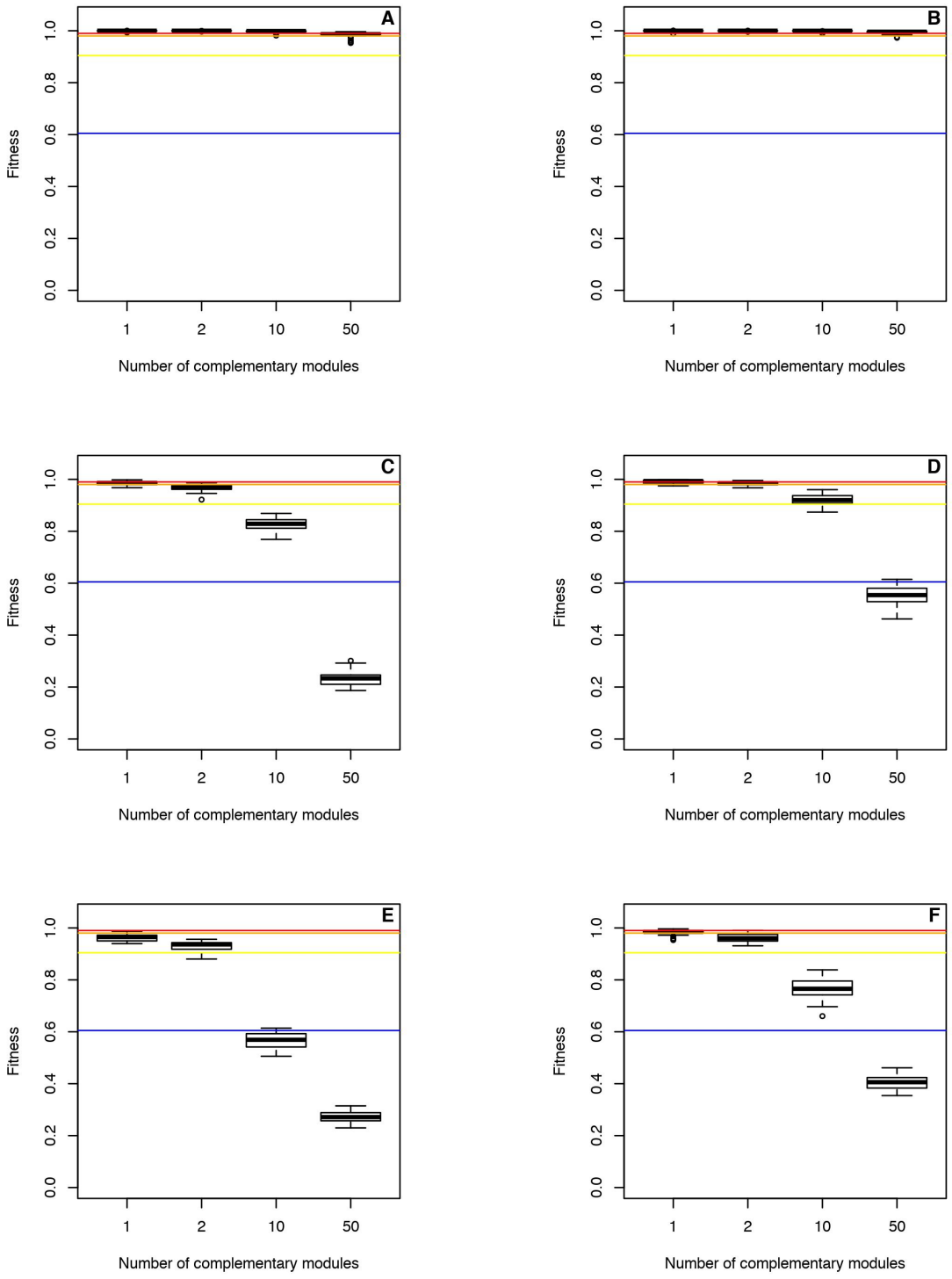

Figure 3.42: Simulation outcomes for the fitness at mutation-selection-drift balance with $N_e = 10^2$ under the interplay of complementary epistasis and diminishing returns epistasis (through the fitness landscape of traits). The number of complementary modules (*eg.* genes) varies from 1 to 50. Each line corresponds to a level of mutational bias: null in (A) and (B), low in (C) and (D), high in (E) and (F) - see text for details; each column represents a level of mutational variability, with (A), (C) and (E) having low variability while (B), (D) and (F) display a moderately high variability.

of largely advantageous mutations - for the same level of mutational bias (compare (C) and (E) with (D) and (F) on Figure (3.42)). Owing to the very low effective population sizes, drift overwhelms Natural Selection and pushes the balance towards very low values, a process already documented in models based on universal pleiotropy such as (Hartl et al., 1996) and (Poon et al., 2000). But we show here using a toy model that such complexity-selection trade-off may be readily observed without any preexisting pleiotropic relationship (and without accounting for the decrease of effective population size usually associated with complexity), which lends credit to the idea of an intrinsic cost to complexity.

3.2.3 Towards an analytical treatment ?

In order to grasp the effect of high order complementary epistasis, it is of great interest to derive results in terms of phenotypes and fitness distributions, rather than only the mutation load it entails. Based on insights by (Sella et al., 2005), who shown that assuming weak mutations - such as in Moran or Wright-Fisher model - it is possible to determine the distribution of fitness at evolutionary steady-state as a combination between the neutral distribution of fitness and the effect of the power of selection N_e using the formula:

$$\rho^*(f) = \frac{f^v.\rho_g(f)}{\int_0^1 f^v.\rho_g(f)\,\mathrm{d}f}, \tag{3.28}$$

where v denotes the temperature analogy and can be estimated for both Moran and Wright-Fisher models (*eg.* haploid Wright-Fisher model $v = 2(N_e - 1)$), f represents any possible value of fitness for a genotype, and $\rho_g(f)$ is the neutral distribution of fitnesses in the genotype space. In this formula, one assumes symmetrical mutations ($\mu_{ij} = \mu ji$), but it is possible to relax this assumption (Sella et al., 2005).

For example, if fitness is uniformly distributed between 0 and 1, meaning that $\rho_g(f) \sim \mathcal{U}(0,1)$, it is straightforward to show that the steady-state distribution of fitness is given by:

$$\rho^*(f) = \frac{f^v}{1/(v+1)} = (v+1)f^v \tag{3.29}$$

As a direct consequence, such a fitness distribution yields the following mutational load:

$$\begin{aligned} L^* = 1 - \prec f^* \succ &= 1 - (v+1)\int_0^1 f^{v+1}.\rho_g(f)\,\mathrm{d}f \\ &= 1 - \frac{v+1}{v+2} \\ &= \frac{1}{v+2} \\ &(= \frac{1}{2N_e} \text{for an haploid Wright-Fisher population}) \end{aligned}$$

Simultaneously, the fundamental theorem of order statistics (Casella et al., 2002) enables one to determine the distribution of the k^{ieth} order statistic. More particularly, the distribution followed by the minimum $X_{(1)} = \min\{X_1, X_2, ..., X_n\}$ of a n-tuple of random variables identically distributed obeys the following equation:

$$f_{X_{(1)}}(x) = n\ f(x)\ (1 - F(x))^{n-1}, \tag{3.30}$$

where $f(x)$ and $F(x)$ respectively represents the density and cumulative probability distribution of these variables.

It is therefore possible to determine the neutral distribution of fitness using known distributions under certain circumstances. A simple didactic case corresponds to the aforementioned uniform distribution, for which it is known that the first order statistic obeys a Beta distribution $\mathcal{B}(1, n)$ when n variables are drawn. The neutral distribution of fitness under complementary epistasis is therefore given by this distribution, which yields that the fitness distribution at evolutionary steady-state also follows a Beta distribution as:

$$\rho^*(f) = \frac{f^v.(1-f)^{n-1}}{\int_0^1 f^v.(1-f)^{n-1}\,\mathrm{d}f}, \tag{3.31}$$

Notice that the scaling factor made by the beta function vanishes as it is present on both sides of the fraction. It follows that $\rho^*(f) \sim \mathcal{B}(v+1, n)$, whose expectancy - which corresponds to fitness at mutation-selection-drift balance - is given by $\prec f^*_{\mathcal{U}_n} \succ = \frac{v+1}{v+1+n}$, which yields $\prec f^*_{hap,\mathcal{U}_n} \succ = \frac{2N_e-1}{2N_e-1+n}$ in an haploid population for a corresponding mutational load of $L^*_{hap,\mathcal{U}_n} = \frac{n}{n+2N_e-1}$. Remarkably, this burden is the exact same quantity than that derived from Fisher's N dimensional

geometric model, meaning that the latter coincides with a model of complementary epistasis where the neutral distribution of fitness is uniformly distributed.

Nonetheless, we know that such a distribution is not very realistic. Instead, the fitness distribution is more likely to obey a Beta-like distribution whose parameters may differ on a case-by-case basis. Beta-distributions do not enable analytic tractability, but Kumaraswamy (1980) put forward the beta-like Kumaraswamy distribution, which allows such calculations (Jones, 2009). This distribution is described by the following DFD:

$$f(x;a,b) = abx^{a-1}(1-x^a)^{b-1}\mathbb{1}_{[0,1]},$$

for which it is possible to write the first order statistic using Eq.(3.30) and to apply it to the fitness distribution:

$$\rho_g(f) = n \times abf^{a-1}(1-f^a)^{b-1}\mathbb{1}_{[0,1]}((1-f^a)^b))^{n-1}$$

Rearranging the terms to gather n and b ones, it is straightforward to show that $\rho_g(f) \sim \mathcal{K}(a,nb)$. Consequently, the distribution of fitness at evolutionary steady-state obeys the following law:

$$\rho^*(f) = \frac{a(nb)f^{v+a-1}(1-f^a)^{nb-1}\mathbb{1}_{[0,1]}}{\int_0^1 a(nb)f^{v+a-1}(1-f^a)^{nb-1}\,\mathrm{d}f} \tag{3.32}$$

Furthermore, another interesting property of this distribution is that its moments may be expressed analytically through a combination of beta functions and thus also through one of the well-known gamma distributions. It is very helpful because looking carefully at Sella et al. (2005)'s formula shows that the fitness at evolutionary steady-state simply reduces to the ratio between the $v+1^{eth}$ and the v^{eth} moments of the neutral fitness distribution. These considerations lead to the following expression:

$$\prec f^*_{\mathcal{K}_v(a,b)} \succ = \frac{nb.B(1+(v+1)/a, nb)}{nb.B(1+v/a, nb)} = \frac{\Gamma(1+(v+1)/a)}{\Gamma(1+v/a)} \times \frac{\Gamma(1+nb+v/a)}{\Gamma(1+nb+(v+1)/a)}, \tag{3.33}$$

where $\Gamma(x,y)$ denotes the gamma function[7].

When $a = 1$ (or more broadly $a \propto 1/k \mid k \in \mathbb{N}^*$), the gamma function can be expressed

[7]Tha gamma function generalizes the factorial concept to the set of complex numbers.

as a factorial, which simplifies the reduction of Eq.(3.33). Yet, the relevance of using Beta-like distributions for the neutral distribution of fitness is to determine the influence of the balance between deleterious and advantageous mutations on the mutational load, which critically depends on both the ratio between a and b, and their absolute values: more specifically, setting $a > 1$ allows one to explore the influence of bell-like curves, as can be shown by drawing the DFD for a sample of parameters set:

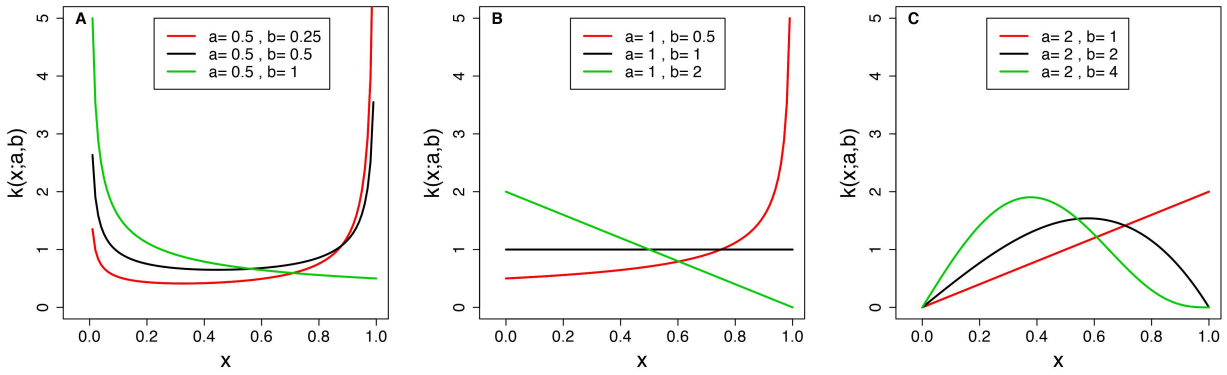

Figure 3.43: Subset of Kumaraswamy distributions: each panel corresponds to a given value of a: $a < 1$ in A produces distributions with an excess of extreme values as is often described about the distribution of fitness effects, $a = 1$ in B represents simpler distributions, while $a = 2$ in C encompasses bell-like distributions. The balance between low and high fitnesses is mainly determined by the ratio between a and b: the higher b, the largest deleterious variants.

From the property connecting the Gamma function to factorial numbers for integer numbers – $\Gamma(n+1) = n!$ – it is straightforward to derive the steady-state fitness and its corollary mutational load when a can be written under the form $1/k$ with $k \in \mathbb{N}^*$, which is given by:

$$\prec f^*_{\mathcal{K}_v(a,b)} \succ = \prod_{i=1}^{1/a} \frac{2N_e/a + i}{2N_e/a + nb + i} \tag{3.34}$$

In parallel, it is also possible to determine an analytical solution when $a = 2$ taking advantages of the well-known formula for half integers, which states that $\Gamma(m + 1/2) = \frac{(2m)!}{2^{2m}m!}\sqrt{\pi}$ if $m \in \mathbb{N}^*$. Assuming that nb is an integer, it is therefore possible to determine $\frac{\Gamma(N+1/2)}{\Gamma(N)}$ as well as $\frac{\Gamma(N+1/2+nb)}{\Gamma(N+nb)}$, which gives:

$$\prec f^*_{\mathcal{K}_v(2,b)} \succ = 2^{2nb} . \frac{(2N_e)!}{N_e!(N_e-1)!} . \frac{(N_e+nb)!(N_e+nb-1)!}{(2(N_e+nb))!} \tag{3.35}$$

Stirling's formula[8] (Dutka, 1991) states that when K is a large integer, the factorial matches the much simpler form $K! = \sqrt{2\pi K}(\frac{K}{e})^n$. Given that $\frac{(2K)!}{K!(K-1)!} = \frac{(2K)!K}{(K!)^2}$, substituting Stirling's formula in a term of the form found in Eq.(3.35) yields:

$$\frac{(2K)!}{K!(K-1)!} \sim 2^{2K}\sqrt{\frac{K}{\pi}}, \text{ for K} \in \mathbb{N}^*$$

Finally, it yields for $a = 2$ and $v = 2(N_e - 1)$:

$$\prec f^*_{\mathcal{K}_v(a,b)} \succ = \sqrt{\frac{N_e}{N_e + nb}} \tag{3.36}$$

Using first order approximation, it follows from these findings that when $nb << N_e$, the Poon-Otto's burden of complexity holds with $L^* \approx \frac{nb}{2N_e}$, which confirms that these confounding causes may be difficult to disentangle through data analysis alone. By contrast, if the number of interacting units is large, the mutational load can depart largely from this previous prediction, especially when nb approaches or exceeds the order of magnitude of N_e. Noticeably, this new formulation shows that parameters b (which can be seen as a mutational bias towards deleterious mutations) and n are equally contributing to the load, and their importance critically depends on the shape of the neutral fitness distribution. Focusing on the most documented distributions, those displaying an excess of extreme values - both low and high ones potentially-, correspond to cases where the mutational load is increased if $b > 1$, and decreased otherwise, for a given number of complementary units. But in contrast with Fisher's model, it seems plausible if not inevitable to assume that n outnumbers small or even intermediate N_e since each of the units can for instance represent an enzyme's or a protein's activity.

3.2.4 Discussion : how other biological features impact these predictions ?

Because the aim of this new framework is to try and understand the influence of epistasis from physical and chemical first principles, further developments will definitely need include other relevant features that should influence evolutionary outcomes. First, it is known that mutations occurring on coding sequences, even being synonymous, are rarely exactly neutral, reflecting in particular the cost arising from codon usage bias (Ikemura, 1985; Galtier et al., 2018; LaBella et

[8] As often happens in the History of Mathematics, Stirling's formula, despite its baptismal name, was not formulated first by Stirling but by De Moivre.

al., 2019), a feature that definitely influences fitness landscapes (Fragata et al., 2018). In parallel, not all kind of phenotypic traits undergo deleterious mutational biases as simple as that presented up until now : this is especially true for levels of expression[9], which often result themselves from complex changes in gene networks features such as the level of trans-regulatory elements, their affinity and specificity, and a pleitropic set of relations with cis-regulatory elements involving the whole genome (Hodgkin, 1998; Chesmore et al., 2016; Boyle et al., 2017; Sella et al., 2019). For illustrative purpose, let us imagine that a given gene is mainly regulated by a transcriptional inhibitor: as transcription factors face a mutational bias towards lower affinities, this bias should enhance gene expression rather than decrease it. Last but not least, if biological systems may be thought to be under directional selection to maximize growth and biomolecules production, the intertwining of pathways and biological reactions is more likely to result in a slightly different picture - portrayed on Figure 3.44 - than that displayed by a simple positive complementary epistasis: proteins that are too efficient may at some point yield collateral damages either because they monopolize resources in vain or as they induce toxicity (*eg.* producing too many metabolites) (Lilja et al., 2017; Niehaus et al., 2020; Labourel et al., 2021). Consequently, it seems more realistic that each of the non-worst genes is under more or less relaxed stabilizing selection towards an optimum, coinciding with the worst gene - that can only change when the worst gene of the set improves to a higher value - which means that even when no mutational bias exists, most genes involved in complementary epistasis should be pulled towards the worst one, for complementary epistasis becomes negative in those cases. Likewise, it seems judicious to study whether the effect of such epistasis differs when it occurs within entities part of a linear system or between parallel entities (which could of course be made by lower-level entities themselves) like two pathways comprised of proteins, which would also be philosophically interesting for its shared similarity with electric or heat systems - as recently outlined by Yi et al. (2019) - where parallel and series circuit were shown to work differently a long while ago, by the then student Gustav Kirchhoff (Charbonneau, 2014).

Being directly based on functional insights, such a framework seems very relevant to start tackling the challenges set out by the joint evolution between basic functional epistasis – and more broadly, genetic interactions – and Adaptation from a population genetics theoretical perspective

[9] And organic shapes, albeit for different reasons.

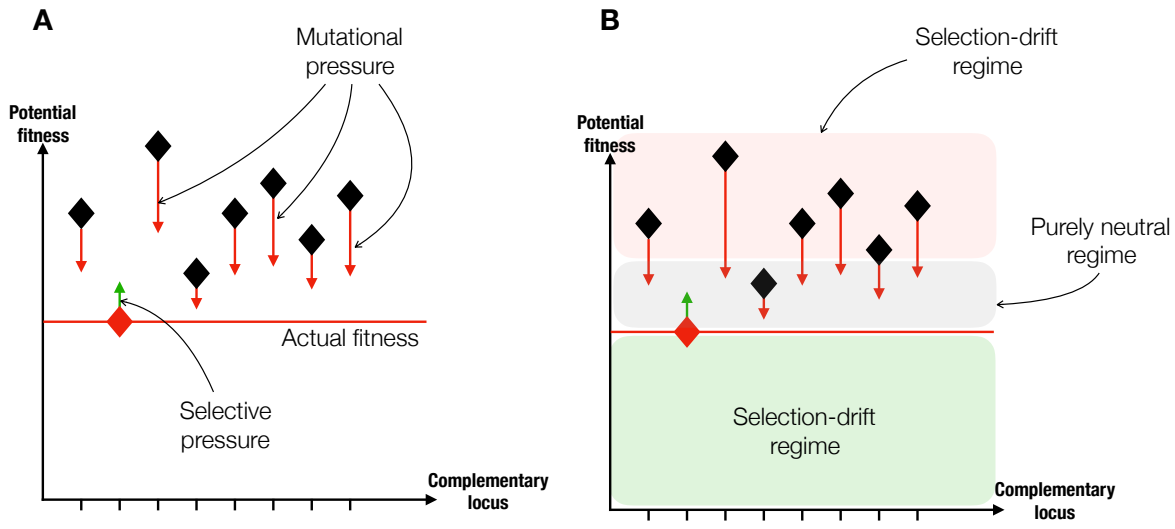

Figure 3.44: Owing to mutational biases, the mutational process exerts a pressure towards lower fitness phenotypes (A), which is not counterbalanced by Natural Selection above the actual fitness: this phenomenon may look like stabilizing selection since most mutations are going to stabilize the actual phenotype by pushing all loci towards the same lowest level coinciding with actual fitness. But in fact, it is even likely that mutations towards very high values of potential fitness can often prove to be deleterious rather than neutral (*eg.* an enzyme far more efficient than its neighbours in a pathway may produce a metabolite so quickly that it ends up reaching toxic cellular levels.). In turn, only a small range of adaptive phenotypes exists at any given time (B). Despite the trait being under directional selection, this latter process thence mimics stabilizing selection, although towards an optimal area rather than a precise point, whilst the mutational pressure further stabilizes the actual deleterious phenotype.

insofar as it may allow to draw very general and testable conclusions. In spite of its apparent specificity, it can indeed describe both intra-level and multilevel evolution: it is appropriate to describe the joint evolution of organs/appendages, as well as enzymes along a pathway, or organs/appendages and enzymes all together, provided that one knows how the genotype-phenoype-fitness map builds up. Many further developments are thereby possible.

One natural sequel would of course be to refine its components and account for the existence of compensatory mutations entailed by the multidimensional phenotypic redundancy of some biological features: higher enzyme concentrations can buffer lower kinetic enzymes - though it comes with the cost of a protein burden (Koch, 1983; Dill et al., 2011; Kafri et al., 2016); villi can theoretically relax the selective pressure acting on enzymes for the absorption of nutrients; a longer calf can compensate for a smaller thigh or vice-versa, etc. To expand further in this area, it would also be relevant to see what happens when genes undergoing true stabilizing selection come

into play, as they are also widespread[10](Sella et al., 2019), and whether the joint evolution with the fitness effects of mutations - whose distribution both impacts the course and the outcome of Evolution while inescapably being subject to It - could overturn expectations.

However, it would seem premature to further investigate these specific building blocks more than others while we know that there exists plenty other kind of genetic interactions (*eg.* biological modularity (Wagner, 1996)), that each deserves to be accounted for on a mechanistic basis: for instance, pleiotropy is, like epistasis and the distribution of fitness effects, both a result of Evolution and an intrinsic biological phenomenon (Wagner, 1996; Chesmore et al., 2016) depending on its underpinnings, which are numerous. This means that understanding its joint evolution with epistasis requires first to inform which part is intrinsic to biological systems[11], when and how it can be alleviated, and how it finally impacts the genotype-phenotype-fitness map before these systems are studied using a complete framework, which undeniably sketches a rather more distant objective.

Instead, it seems more appropriate to adopt a step-by-step approach where population genetics and first principle fitness landscapes are built in parallel, as was done in the past to understand the evolution of stability (Taverna et al., 2002; Bloom et al., 2004) and its evolutionary (Drummond et al., 2005; Bloom et al., 2006; Drummond et al., 2008; Tokuriki et al., 2009; Serohijos et al., 2012; Dasmeh et al., 2014; Echave et al., 2017) and functional consequences (Bloom et al., 2007; Geiler-Samerotte et al., 2011; Dasmeh et al., 2018), and to derive the models of genetic interactions and constraints from these latter ones rather than taking them for granted because they currently exist after billion years of evolutionary history.

[10]This is noticeably interesting for such fitness landscapes - including stabilizing selection - have been empirically documented in the case of drug resistance in microorganisms (Ford et al., 2020).

[11]In order to avoid introducing constraints which may be the result of Evolution, be it adaptive or not.