Chapter 18
How Evolutionary Systems Biology Will Help Understand Adaptive Landscapes and Distributions of Mutational Effects

Laurence Loewe

Abstract Population genetics and ecology have been modeling biological systems quantitatively for over 8 decades and their results have contributed greatly to our understanding of the natural world and its evolution. Theories in these areas necessarily had to focus on comparisons of the contribution of different individuals to changes in the bigger picture at the expense of ignoring much of the complexity that exists inside individuals. Current systems biology provides new insights into this complexity within organisms. Here I review developments in evolutionary systems biology that have the potential to lead to a more unified approach that integrates contributions from current systems biology and population genetics. Central integrative concepts in this approach are the adaptive landscape and distributions of mutational effects. Both capture our understanding of the fitness of individuals and how it can change. Fitness is frequently used in population genetics to summarize key properties of individuals. Such properties emerge from the complexity of molecular processes within individuals, often in interaction with the environment. The general principles of this approach are reviewed here. This work can open up new avenues for computing critical quantities for models of long-term evolution, including epistasis, the distribution of deleterious mutational effects, and the frequency of adaptive mutations.

1 Introduction

Fitness is a central concept in modern biology. It owes its influence to the unifying power of evolution. The fundamental definition of fitness is simple: the fitness of a genotype predicts how successful it will be at contributing offspring to the

L. Loewe (✉)
Laboratory of Genetics and Wisconsin Institute for Discovery, University of Wisconsin-Madison, 330 North Orchard St., Madison, WI 53715, USA
e-mail: loewe@wisc.edu

O.S. Soyer (ed.), *Evolutionary Systems Biology*, Advances in Experimental Medicine and Biology 751, DOI 10.1007/978-1-4614-3567-9_18,
© Springer Science+Business Media, LLC 2012

next generation. There are a number of different mathematical formalisms that can be used to describe this concept, most of which use simplifications that are appropriate for specific biological systems [1–3]. Ultimately, only two traits contribute to fitness throughout the broad diversity observed in biology: survival and fecundity. The importance of these two contributions throughout the life of individuals varies greatly for different systems, as do the trade-offs between them. Investigating the many individual traits that combine to define these two high-level composite traits leads to even more complexity and diversity. Different fields that study evolution stop at different points in their quest to drill down into deeper details. For example, many studies in population genetics assume that investigating differences in survival between discrete generations is sufficient, whereas studies in life-history evolution use a much more fine-grained approach [3]. Regardless of the level of detail used in an analysis, the fitness of individuals always exists independently of our models. Fitness is an important factor that governs evolution, whether we can compute it or not. Our poor track record in predicting fitness occasionally feeds the misunderstanding that the concept of fitness is part of a tautology ("The fittest survive. Who are the fittest? Those that survive"). The tautology falls apart in the relevant models of population genetics that predict evolutionary outcomes based on measured fitness values. While the fitness of individuals is difficult to predict, it can be measured more readily in the wild and its heritability is substantial [4–6]. Heritability implies that changes in the genes that contribute to fitness will lead to changes in survival or fecundity for individuals that carry those genes. This is equivalent to saying that DNA mutations can have a substantial effect on fitness.

The link between mutations and their ultimate effects on fitness can be less than obvious. Examples for obvious links are Mendelian diseases that are characterized by few mutations of large effects [7]. On many other occasions the link is less obvious, especially for the many mutational effects on fitness that are very small [8–10]. The effects of these mutations can be estimated from DNA sequence diversity data with the help of population genetics models [10–15]. However, these models depend on various assumptions that can be difficult to check independently, such as ancient effective population sizes or the condition of having reached mutation-selection-drift balance after recent demographic events. Sometimes theoretical insights can lead to important conceptual results that prove helpful in predicting experimental outcomes as in studies of the distribution of advantageous mutations [16, 17]. Generally, properties of harmful DNA changes are much easier to infer than those of advantageous mutations, due to the relative abundance of the former. Thus, an independent approach for inferring small mutational effects would be helpful.

Current systems biology might offer a different perspective. While the fitness of individuals can be seen as the starting point for quantitative predictions in ecology and evolution, fitness might also be seen as the highest goal for quantitative predictions in current systems biology. Current systems biology usually focuses on a subset of the large number of complex processes that occur concurrently inside of individuals [18, 19]. Substantial efforts have been dedicated to developing quantitative models at appropriate levels of abstraction. These models are not

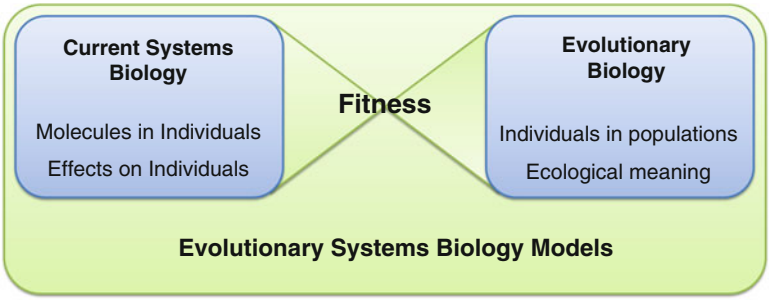

Fig. 18.1 The central role of fitness as a bridge between models in current systems biology and evolutionary biology. Evolutionary systems biology models are characterized by their drive to integrate results from different domains (modified from [21])

usually connected explicitly to traits that are important for fitness. However, it is very likely that at least some of the properties predicted by current systems biology models will be important for either survival or fecundity. Finding and measuring these properties might at times be difficult and can require substantial biological intuition for their identification and substantial computational skill for their quantification. However, if such properties indeed capture some aspect of importance to fitness, then they will be correlated with fitness. This correlation will be linear for mutational effects that are small enough (for exactly the same reasons that allow arbitrary smooth functions to be approximated by straight lines over short enough intervals). Since these fitness correlates are not fitness itself and thus not subject to the optimizing trade-offs that usually affect fitness in nature, future laboratory work in artificial environments will likely be able to observe fitness correlate values that are larger than those of the wild type. If this is combined with an appropriate model of the trade-offs that comprise fitness, then harmful and beneficial changes to fitness might be computed with equal ease.

Computing fitness correlates and their resulting adaptive landscapes have been at the core of a series of proposals that develop a mechanistic understanding of evolutionary systems biology [20–22]. In this framework the concept of fitness is used to facilitate a separation of concerns between different domains and to mediate important interactions that build bridges between current systems biology and evolutionary biology (Fig. 18.1).

The remainder of this chapter reviews various difficulties that arise when computing fitness and potential ways for addressing them.

2 Levels of the Adaptive Landscape

The adaptive landscape has a substantial intuitive appeal for explaining key components of the evolutionary process, such as the observation that populations will usually evolve towards the nearest local optimum. In such landscapes the position

in a two-dimensional *plane* of genotypes determines the one-dimensional *height* that indicates fitness. This structure captures a true causality, if we assume that the environment is constant and stochastic effects are absent. Adaptive landscapes have been helpful in understanding speciation processes [23]. However, most genotypes require more than two dimensions for an adequate description, and it is not clear how a high-level parameter such as fitness can be computed from them. Such difficulties have led to the criticism that adaptive landscapes are not as useful for research as they seem.

Indeed, the difficulties of mapping genotypes to phenotypes and ultimately to fitness are enormous, as they require mastering a substantial number of problems viewed as grand challenges in modern biology. The arrival of current systems biology has substantially improved the conceptual situation, as it is developing approaches for simulating ensembles of many parts that could previously only be investigated individually. This has enabled the definition of "levels" of the adaptive landscape [21], based on

- An abstraction and
- A chain of causality.

The abstraction replaces the causality captured in adaptive landscapes with a function:

$$\text{height} = \text{function}(\text{position in plane}), \tag{18.1}$$

where *height* is a "high-level" parameter that is conceptually closer to fitness (higher = later in the chain of causality), *position in plane* is a "low-level" parameter that is conceptually closer to a DNA-sequence genotype (lower = earlier in the chain of causality), and both parameters have as many dimensions as necessary to describe the relevant data with reasonable precision. Based on this abstraction, different types of biological research results that are relevant for adaptive landscapes can be ordered into a chain of causality, where the output of a lower-level function serves as the input for a higher-level function. For three levels, we have formally:

$$\begin{aligned} h_1 &= \text{AdaptL1}\,(p_1) \\ h_2 &= \text{AdaptL2}\,(p_2 = h_1) \\ h_3 &= \text{AdaptL3}\,(p_3 = h_2) \end{aligned}$$

equivalent to

$$h_3 = \text{AdaptL3}\,(\text{AdaptL2}\,(\text{AdaptL1}\,(p_1))), \tag{18.2}$$

where $p_1, \ldots, p_3$ denote *positions in planes*, $h_1, \ldots, h_3$ denote *heights*, AdaptL1,..., AdaptL3 denote functions that encapsulate the causality of the corresponding level of the adaptive landscape and numbers indicate the corresponding level. Please note that AdaptL might produce probability distributions as output in order to model stochastic systems. Building on these concepts we can define levels of adaptive landscapes as given in Table 18.1.

While definitions of hierarchies as in Table 18.1 will always be arbitrary to some degree, a number of deliberate choices were made in this case to facilitate

Table 18.1 Levels of the adaptive landscape [21]

Level	Plane (given)	Discipline determining	Height (desired)
7	Fitness of individuals in a population	Simple statistics	Mean fitness of a population
6	Observable fitness correlates	Life-history theory, trade-off analyses	Fitness of individuals
5	Computable candidate fitness correlates	Complex experiments	Observable fitness correlates
4	Computable emergent properties	EvoSysBio	Computable candidate fitness correlates
3	Molecular functions	Current systems biology simulations	Computable emergent properties
2	Molecular structures	Structural biology	Molecular functions
1	DNA sequences	Structural biology	Molecular structures

the integration of existing evidence into a bigger coherent picture. For example, one might argue that systems biology simulations at "Level Three" should be described by several levels representing processes that operate only inside cells, inside tissues, inside organs and finally inside individuals. This can still be done by defining "sub-levels" or by adding more levels to the chain of causality, if deemed helpful. However, the biggest challenge will be to define the computational models that span all these hierarchical levels of organization. To avoid overcomplicating this hierarchy or introducing an inflation of various potentially conflicting level definitions, all these details were summarized into "Level Three", leaving decisions about what to model to the choice of abstractions in current systems biology. Thus, some models can choose to track the details of transcription and translation, whereas others can operate at a coarser grained level without affecting the definition of levels in Table 18.1.

Linear Fitness Correlate Hypothesis. One might also argue that level 5 does not reflect any objective hierarchical reality in biology, but rather is merely an artifact of the way we observe biology (i.e. through computation or through experiments). However, the biggest challenge here is to arrive at the point where this level is indeed superfluous. This will be the case, when the stack of computational predictions from all previous levels can be demonstrated to match equivalent experimental observations of fitness correlates. The challenge in this is that the same relative change in the value of a fitness correlate needs to be determined independently in two different ways: Mutants with known genotypes must exist for which the fitness correlates in question can be measured experimentally in vivo. Independently of that, the genotypic information of these mutants has to be used to predict the same fitness correlates in silico. One can only claim to have mastered the system if the relative changes of fitness correlates between substantial numbers of mutants are the same in predictions and observations. If this is true and not an artifact of stochasticity or measurement noise in the system, then one can accept the "Linear Fitness Correlate Hypothesis" for that system [21]. It derives its name from the

expectation that predicted and observed fitness correlates for all mutants in a given system should be on a straight line if the system has indeed been truly understood.

Linking the adaptive landscape to current biological research. The various levels presented in Table 18.1 are noteworthy in the following respect. In combination they completely map genotypes to fitness and for each level studies can be found that demonstrate in some biological system how this level can be mastered in principle, either through experimental observation or through computational prediction. However, what is missing are nontrivial study systems that go all the way from DNA sequences to fitness. Work predicting growth of biomass in bacteria by flux-balance-analysis [24,25] is one of the several promising examples (for more examples, see below, other chapters in this book and [21]). New developments in contemporary biology have brought us to the point where further progress towards predicting fitness correlates is rapidly coming closer for many interesting systems. However, many exciting methods are not general and work well only for a limited number of systems. Thus, much more work will have to be done to broaden the applicability of these methods in systems biology and structural biology if predicting adaptive landscapes is to become a routine activity.

3 Distributions of Mutational Effects Can Visualize Adaptive Landscapes

Besides their poor computability, adaptive landscapes have another feature that makes them difficult to use: their high dimensionality. One can debate whether dimensions of the genotype *plane* should be a relatively small number of quantitative traits or the much larger number of genes in an organism or the even larger number of nucleotides in the genome. In either case the number of resulting dimensions of the adaptive landscape is much larger than humans can usually handle cognitively. This makes adaptive landscapes fundamentally difficult to visualize and renders the many popular images of such landscapes useless at best and misleading at worst. Thus, new innovative ways are needed for visualizing the high-dimensional results that come from adaptive landscape analyses.

One approach may come from an unexpected corner and has been used for a long time in population genetics models: distributions of mutational effects (DMEs). DMEs can be thought of as local excerpts of the adaptive landscape. They show all points of the adaptive landscape that are relevant to the next mutational step, given a specified starting point. DMEs can be compiled in the following way:

1. Select a wild-type reference genotype as a starting point.
2. Determine the fitness for the starting point.
3. Add a mutation by randomly changing the genotype.
4. Determine changes in fitness for the new genotype with the mutation.
5. Go to (3) until the DME has been sampled with sufficient accuracy.

A few comments may be in order here. The choice of the wild-type reference point will usually be guided by the availability of a computational model that has been built to resemble a naturally existing system. The choice of what constitutes the equivalent of a "genotype" and the equivalent of "fitness" can be very model dependent. As highlighted in Table 18.1, there are many disciplines that provide datasets that predict the height from a point in the plane of a level of the adaptive landscape. For example, computational studies of RNA folding can be used to study aspects of adaptive landscapes by predicting patterns of thermodynamic stability in RNA [26]. In this case the plane is defined by RNA sequences and the height by the thermodynamic stability of their secondary structures. In another example, computational studies of genotype networks indicating the presence or absence of particular genes can be used to investigate aspects of adaptive landscapes by predicting biomass production [25]. Here the plane is defined by a combination of chemical reactions present in the cell, whereas the height is given by the rate of biomass production in a defined environment. The DMEs resulting from such an analysis are obviously limited to statements about the corresponding level of the adaptive landscape and the system studied. Further work will be necessary to determine how general such conclusions can be.

The nature of "mutations" in all scenarios is that of "steps in the plane of the landscape." What those steps mean and how step sizes are distributed in various dimensions of the plain depends on system details as indicated above. Ultimately, "mutations" are DNA changes that are weighted by the frequency of their respective occurrence. For higher levels, "mutations" could also be substituted by changes in biochemical reaction rates or changes in quantitative traits, again weighted by their relative rate of occurrence.

Likewise, the nature of "fitness" will vary with levels. On lower levels it will mostly quantify the difference between a given system and an "optimal" system. Optimality may be assessed by biological considerations of the nature of the system and its context and will have to be judged by how well these considerations reflect reality. If this is too difficult, the following simplified substitute might work for some systems:

1. Take any existing natural system as a template;
2. Assume that natural selection has optimized it sufficiently; and
3. Compare any mutated system to the template by using measures that are considered to be relevant to optimality. While both approaches cannot guarantee relevant results, they provide ample opportunities for improving our understanding of the systems investigated.

DMEs are ultimately histograms that quantify how many mutations will have an effect of a given size. In order to observe them properly, sufficiently many different mutations need to be sampled. How many will depend on the number of potential mutations possible, the stochastic or deterministic nature of the models that compute effects, and the questions driving the analysis. Generally, most analyses will need large sample sizes to compute relevant DMEs.

There are a large number of different types of changes that can be applied to biological systems in order to investigate mutations. There are also a large number of different types of emergent properties that might be important for fitness. This is true for each single level of the adaptive landscape, but even more so, when different levels are considered. The resulting explosion of opportunities for analysis leads to a proliferation of DMEs that can be very confusing. To facilitate comparisons, a nomenclature was designed that helps directing the focus of a discussion to particular DMEs [22]. The core of this nomenclature centers on the three abbreviations in "DME," each of which stands for one of the following concepts and is accompanied by indices:

- *Distribution sign*. Generally, DMEs include increases and decreases in the possible changes that they describe. Sometimes it is helpful to focus on one and exclude the other, e.g. to avoid interferences from deleterious mutations, when discussing fitness increases. In this nomenclature, adding "D" or "I" can be used to indicate a focus on **d**ecreases or **i**ncreases, respectively. Using superscripts ^{D} or ^{I} indicate a reference to respective decreases or increases in the high-level effect ("the height"). Subscripts $_{D}$ or $_{I}$ describe corresponding low-level changes ("the plane") if these are characterized by parameters and not by DNA sequence states. By convention, omission is equivalent to specifying "$_{DI}$" or "$_{DI}$". For example, "$\mathrm{D}^{\mathrm{I}}_{\mathrm{D}}\mathrm{M}_r\mathrm{E}^{F}$" indicates a distribution of effects that exclusively increase fitness F, but only by decreasing the biochemical rate r.
- *Mutations.* Ultimately, mutations are DNA changes. To denote this, by convention, no further indices are added. Unfortunately our capacity to predict biochemical and higher-level functions from DNA sequences are very limited at the moment. Therefore, some studies might want to introduce effects of mutations by merely changing a biochemical reaction rate parameter directly. In this nomenclature this is done by adding the rate as a subscript after the "M" in "DME." For example, to contrast the different DMEs that result from changes in the biochemical rates r and s, one writes "$\mathrm{DM}_r\mathrm{E}$" and "$\mathrm{DM}_s\mathrm{E}$", respectively.
- *Effects*. The ultimate mutational effects are effects on fitness. Since these are very difficult to compute, many DME analyses will use various fitness correlates to measure effects. In the absence of meaningfully defined fitness correlates, other properties of interest may be used. To indicate that effects on property x are analyzed in a DME, the nomenclature indicates the property as a superscript after the "E," resulting in "DME^{x}." A superscript is used to highlight that effects focus on analyses of high-level properties of a system. For example, effects on a general measure of fitness F might be referred to as "DME^{F}." By convention, the superscript may be omitted if it refers to a general measure of fitness that does not need to be characterized in more detail.

For example, using this nomenclature it is possible to focus on

'*high-level increases caused by low-level decreases in a*
d*istribution of*
m*utational changes in biochemical reaction rate* r *with*
e*ffects on high-level property* p'

by simply referring to

"$\mathrm{D}^{\mathrm{I}}_{\mathrm{D}}\mathrm{M}_r\mathrm{E}^p$".

In this nomenclature a comprehensive description of the simple abbreviation "DME" would be "high-level increases and decreases caused by any low-level changes in a **d**istribution of **m**utational changes in DNA that have **e**ffects on fitness." This nomenclature is used in [22] to discuss properties of a simple model of a circadian clock.

Many different DMEs will need to be analyzed in order to obtain a reasonably comprehensive picture of the adaptive landscape. A concise nomenclature facilitates the necessary discussions. This requires the definition of a broad range of diverse candidate fitness correlates.

4 Example: Candidate Fitness Correlates (CFCs) in Circadian Clocks

To provide an example for how links could be defined between a systems biology model and fitness, we will briefly discuss a candidate fitness correlate in a simple circadian clock. Circadian clocks are of immense importance for a wide range of biological activities, many of them relevant to fitness, such as growth in plants [27]. Thus, the quality of a circadian clock has an impact on fitness. How can this insight be used to construct a candidate fitness correlate that facilitates the evolutionary systems biology analyses discussed above? To see this, we need to take a step back.

The purpose of a circadian clock is to internally represent an external change in the environment and thus enable switching important genes on or off. It is important to accurately inform about the external state, as any misinformation is likely to have severe consequences for the capability of an organism to exploit opportunities for growth or survival.

This insight can be formalized. A system with external *day-night* oscillations and internal *clock-is-on-or-off* oscillations will lead to four principal states as detailed in Fig. 18.2. Using timeseries of molecule counts and changes in external state as observed in a simulation allows us to compute the fraction of total time the clock spends in each of the four different states. Depending on whether a high molecular count is activating or repressing and whether it affects mostly genes used during day or during night, we can define cyclical (F_c) and anticyclical (F_a) CFCs:

$$\begin{aligned} \mathrm{CFC1} &: F_c = (T_{1\mathrm{D}} + T_{0\mathrm{N}})/T_{\mathrm{total}} \\ \mathrm{CFC2} &: F_a = (T_{0\mathrm{D}} + T_{1\mathrm{N}})/T_{\mathrm{total}}, \end{aligned} \tag{18.3}$$

where $T_{1\mathrm{D}}$, $T_{0\mathrm{N}}$, $T_{0\mathrm{D}}$, and $T_{1\mathrm{N}}$ sum over all time when the system is "On" during "Day," "Off" during "Night," "Off" during "Day," and "On" during "Night," respectively, and T_{total} is the total time. CFC2 has been computed by dedicated code

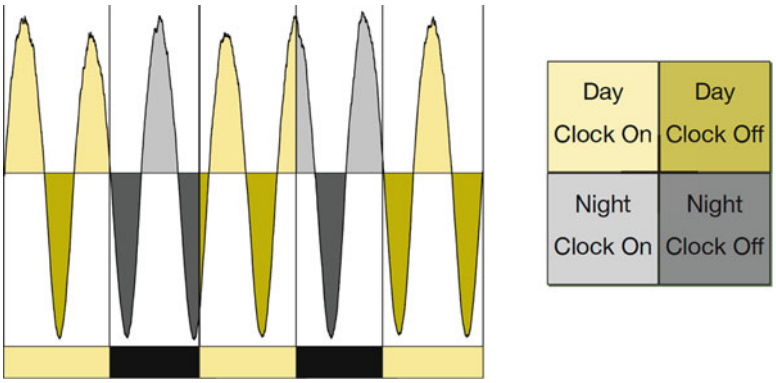

Fig. 18.2 Each point in time in a system with external and internal oscillations around thresholds can be assigned unambiguously to one of the four states given in the table on the *right*. The bars underneath the time course on the *left* depict the external cycles of day and night, whereas the oscillations around the threshold indicate the behavior of the internal circadian clock (with kind permission from [22])

that analyzed time series data from a simple circadian clock simulation to compute a $\mathrm{DM}_{k1}\mathrm{E}^{Fa}$ where k_1 is the effective rate of repressor accumulation in the nucleus [22]. Current work focuses on analyzing this and other CFCs in a more realistic model of the circadian clock of the green algae *Ostreococcus tauri* [28].

5 Perspectives

The general approach presented here can in principle harness the mechanistic understanding inherent in systems biology models in order to compute properties of importance in population genetics. Such properties include parameters that quantify DMEs [21, 22, 28, 29] and epistasis [30–33]. There is a notorious lack of information in population genetics about mutational effects and epistasis. Both are hard to measure, yet of crucial importance for many important theories. Results from current systems biology models will help improve our understanding of these evolutionary theories by providing better estimates of key parameters if analyzed by the approach presented here.

Current systems biology will equally benefit from a deeper understanding of evolutionary questions. A substantial number of important objects of study in current systems biology include entities that carry genomes and replicate over many generations. Such systems are most elegantly described by the formalisms that have been developed over decades in population genetics. For example, cancer cells replicate where they should not. They do this, because of mutations that provide

them with a selective advantage over cells that do not grow in the same tissue. Such a system is best described by a population genetics model, which can predict growth if properly parameterized.

The mechanistic approach for evolutionary systems biology presented here is not the only possible approach. Other approaches for evolutionary systems biology focus on comparing results using methods from systems biology and other fields to characterize phenotypic properties of diverse organisms connected by a phylogenetic tree [34, 35].

The specific contribution of mechanistic evolutionary systems biology is a deeper understanding of how various processes govern the long-term dynamics of a system under investigation. Thus, evolutionary systems biology has the potential to help us understand a broad range of scientific questions by bringing together the disciplines of systems biology and population genetics, both of which have developed independent and impressive bodies of theory for quantifying our understanding of life [21].

Acknowledgements I thank the editor and two anonymous reviewers for helpful comments on this manuscript and the Wisconsin Institute for Discovery at the University of Wisconsin-Madison for support.

References

1. Crow JF., Kimura M (1970, 2009) An introduction to population genetics theory. Blackburn Press, Caldwell
2. Charlesworth B (1994) Evolution in age-structured populations. Cambridge studies in mathematical biology, vol 13, 2nd edn. Cambridge University Press, Cambridge
3. Brommer JE (2000) The evolution of fitness in life-history theory. Biol Rev Camb Philos Soc 75(3):377–404
4. Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P (2001) The strength of phenotypic selection in natural populations. Am Nat 157(3):245–261
5. Endler JA (1986) Natural selection in the wild. Monographs in population biology, vol 21. Princeton University Press, NJ
6. Mitton JB (1997) Selection in natural populations. Oxford University Press, Oxford
7. Sunyaev S, Ramensky V, Koch I, Lathe W 3rd, Kondrashov AS, Bork P (2001) Prediction of deleterious human alleles. Hum Mol Genet 10(6):591–597
8. Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. Nat Rev Genet 8(8):610–618
9. Loewe L, Charlesworth B (2006) Inferring the distribution of mutational effects on fitness in Drosophila. Biol Lett 2(3):426–430
10. Loewe L, Charlesworth B, Bartolomé C, Nöel V (2006) Estimating selection on nonsynonymous mutations. Genetics 172:1079–1092
11. Eyre-Walker A, Woolfit M, Phelps T (2006) The distribution of fitness effects of new deleterious amino acid mutations in humans. Genetics 173(2):891–900
12. Keightley PD, Halligan DL (2011) Inference of site frequency spectra from high-throughput sequence data: quantification of selection on nonsynonymous and synonymous sites in humans. Genetics 188(4):931–940 doi:10.1534/genetics.111.128355

13. Schneider A, Charlesworth B, Eyre-Walker A, Keightley PD (2011) A method for inferring the rate of occurrence and fitness effects of advantageous mutations. Genetics 189:1427–1437
14. Keightley PD, Eyre-Walker A (2007) Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. Genetics 177(4):2251–2261
15. Keightley PD, Eyre-Walker A (2010) What can we learn about the distribution of fitness effects of new mutations from DNA sequence data? Philos Trans Roy Soc Lond B Biol Sci 365(1544):1187–1193 doi:10.1098/rstb.2009.0266
16. Orr HA (2005) The genetic theory of adaptation: a brief history. Nat Rev Genet 6(2):119–127
17. Orr HA (2010) The population genetics of beneficial mutations. Philos Trans Roy Soc Lond B Biol Sci 365:1195–1201
18. Kitano H (2002) Computational systems biology. Nature 420(6912):206–210
19. Kitano H (2002) Systems biology: a brief overview. Science 295(5560):1662–1664
20. Loewe L (2007) Poster: An evolutionary framework for systems biology. In: 41th Population Genetics Group Meeting, Warwick, UK. http://www.populationgeneticsgroup.org/wpcontent/uploads/2010/01/PGGWarwickProgramme.pdf and http://evolution.ws/people/loewe/posters
21. Loewe L (2009) A framework for evolutionary systems biology. BMC Syst Biol 3:27
22. Loewe L, Hillston J (2008) The distribution of mutational effects on fitness in a simple circadian clock. Lect Notes Bioinform 5307:156–175
23. Gavrilets S (2004) Fitness landscapes and the origin of species, Monographs in population biology, Princeton University Press, NJ 41:476
24. Joyce AR, Palsson BO (2008) Predicting gene essentiality using genome-scale in silico models. Meth Mol Biol 416:433–457
25. Samal A, Matias Rodrigues JF, Jost J, Martin OC, Wagner A (2010) Genotype networks in metabolic reaction spaces. BMC Syst Biol 4:30
26. Cowperthwaite MC, Economo EP, Harcombe WR, Miller EL, Meyers LA (2008) The ascent of the abundant: how mutational networks constrain evolution. PLoS Comput Biol 4(7):e1000110
27. Dodd AN, Salathia N, Hall A, Kevei E, Toth R, Nagy F, Hibberd JM, Millar AJ, Webb AAR (2005) Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science 309(5734):630–633
28. Loewe L, Hillston J (2012) Evolutionary systems biology estimates of distributions of mutational effects in a circadian clock (in preparation)
29. MacLean RC, Perron GG, Gardner A (2010) Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in Pseudomonas aeruginosa. Genetics 186(4):1345–1354. doi:10.1534/genetics.110.123083
30. You L., Yin J (2002) Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage t7. Genetics 160(4):1273–1281
31. You L, Suthers PF, Yin J (2002) Effects of Escherichia coli physiology on growth of phage T7 in vivo and in silico. J Bacteriol 184(7):1888–1894
32. Endy D, You L, Yin J, Molineux IJ (2000) Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. Proc Natl Acad Sci USA 97(10):5375–5380
33. Segre D, Deluna A, Church GM, Kishony R (2005) Modular epistasis in yeast metabolism. Nat Genet 37(1):77–83
34. Medina M (2005) Genomes, phylogeny, and evolutionary systems biology. Proc Natl Acad Sci USA 102(1):6630–6635
35. Koonin EV, Wolf YI (2006) Evolutionary systems biology: links between gene evolution and function. Curr Opin Biotechnol 17(5):481–487