news & views

FITNESS LANDSCAPES

Life finds a way

Fitness landscapes were described almost a century ago as smooth surfaces with peaks and valleys that are difficult to navigate. Now, more realistic high-dimensional genotype–phenotype maps show that fitness maxima can be reached from almost any other phenotype while avoiding fitness valleys, which are very rare.

Jacobo Aguirre

A fitness landscape is a well-known simplified visualization of the relationship between genotypes and their fitness, or reproductive success. In evolutionary geneticist Sewall Wright's original conception[1], similar genotypes (or organisms) are close to each other on the continuous surface of the landscape, and height indicates the corresponding fitness. In consequence, organisms are constrained to move along ridges if they exist, or risk extinction by crossing valleys of low fitness, in order to reach fitness peaks (Fig. 1a). Even though Wright was aware of some limitations of his metaphor, its simplicity and beauty helped it to spread rapidly through hundreds of scientific articles and books. Over the years, others, such as Sergey Gavrilets[2]—who introduced the concept of 'holey adaptive landscapes'—have pointed out that this approach was too simplistic and could even be misleading given the multidimensionality of sequence space. This means that how genotypes map onto phenotypes and then onto organismal functions remains an open question in evolutionary biology[3].

Writing in *Nature Ecology and Evolution*, Greenbury et al[4]. show how common properties of high-dimensional genotype–phenotype maps influence the evolutionary navigability of fitness landscapes, making use of biologically realistic models and experimental databases. They show that fitness maxima can be reached from almost any other phenotype without passing through a fitness valley, confirming that fitness landscapes are navigable (Fig. 1b). As a consequence of their work, the very concept of fitness valleys becomes blurred, as these turn out to be rare and are no longer the frontiers between well-defined peaks.

The authors studied the navigability of three widely studied model genotype–phenotype maps: RNA secondary structures, protein tertiary structures and protein complexes. They defined the accessible paths of point mutations between a low-fitness phenotype and a high-fitness phenotype as those with monotonically increasing

fitness. The navigability was then measured as the average probability of finding an accessible path between the different phenotypes in the fitness landscape. Their work convincingly shows that the networked and multi-dimensional nature of these model fitness landscapes presents a strong positive impact of dimensionality and redundancy (the size of neutral sets, that is, of sets of genotypes that share the same phenotype) on evolutionary navigability. Furthermore, they verified the results with sample real RNA secondary structures from the functional RNA database (fRNAdb), suggesting that accessible paths are likely to be exploited in the course of biological evolution. Interestingly, most of the numerical work made use of biologically unrealistic random fitness landscapes, in which similar phenotypes did not in general show similar fitness. Even in this strongly noisy environment, navigability was present. When the authors used more realistic landscapes, that incorporated fitness correlations based on phenotypic similarity and in which the concept of accessible paths was loosened to accept small losses of fitness in intermediate steps, navigability was, as expected, strongly enhanced.

Extensive literature has focused on understanding genotype–phenotype maps, resulting in a constant updating of the seminal interpretation of fitness landscapes[3]. For instance, Franke et al[5]. studied the dependence of the evolutionary accessibility of fitness landscapes on their dimensionality by using several landscape models with tuneable ruggedness, and compared the results with an empirical landscape. More recently, Catalán et al. introduced adaptive multiscapes[6] to address the challenge of visualizing realistic fitness landscapes. This approach overcame some of the deficiencies of Wright's concept with a network-based landscape formed by interconnected layers, each of which contains the phenotypes viable in a given environment. The novel contribution of Greenbury et al. to these previous approaches is a systematic quantification of the navigability of realistic
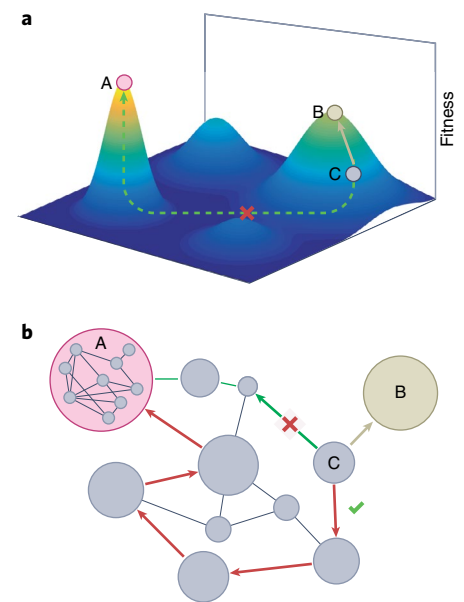

Fig. 1 | Evolution of fitness landscapes from Wright to the present day. In the low-dimensional traditional fitness landscape plotted in **a**, a genotype of phenotype C can reach fitness local maximum B through monotonically increasing fitness mutations, but to reach fitness global maximum A it must cross a deep valley, with risk of extinction. In the more realistic phenotype network described in Greenbury et al[4]. and shown in **b**, nodes represent the neutral network of genotypes with a single phenotype, node sizes are the associated fitness, and two nodes are connected if at least one genotype of the first node is one mutation away from a genotype of the second node. A genotype of phenotype C can now avoid fitness valleys and reach both fitness maxima A and B through mutations of increasing fitness. This is a simple sketch of how the multidimensional, networked and extremely complex nature of realistic fitness landscapes makes them evolutionarily navigable.

evolutionary landscapes, including a strong effort to deal with the inherent limitations of numerical simulations. Their results reach a level of generality that will allow application

to a wide variety of both genotype–phenotype models and real systems.

Several critical challenges remain for achieving a full comprehension of genotype–phenotype complexity. One is the unimaginably large sequence spaces associated with even the smallest relevant biomolecules. For example, the shortest transfer RNA is 75 nucleotides long, which means that its genotype network size is 4^{75} (~10^{45}), and the same applies to the sequence space associated with short proteins. We will probably never be able to numerically encompass such large datasets. For this reason, we need a formal theory that is compatible with these large sets, but also methods to numerically obtain statistically significant fitness landscapes from limited data samples.

A second major task is the creation of more realistic genotype–phenotype maps, benefiting from recent developments in network science, big data and machine learning. Most current RNA genotype–phenotype maps deal with evolutionary processes in which the minimum-free-energy secondary structure is the only phenotype associated with a sequence, and single mutations are the unique source of variability. However, RNA genotypes are known to fold into different structures that could be integrated in a many-to-many map[7], while multiple mutations, insertions and deletions of nucleotides[8] and even sequence recombination[9] should also be taken into account, as these are typical in real biological systems.

In summary, the next generation of scientists will face the task of working with fitness landscapes whose nature is far more complex than they were taught. Fortunately, those who enter there need not abandon all hope, because these multidimensional and networked entities, although unconceivably large and astonishingly interwoven, are navigable. ❐

Jacobo Aguirre [iD][1,2][✉]

¹*Centro de Astrobiología (CAB), CSIC-INTA, Carretera de Ajalvir, Torrejón de Ardoz, Madrid, Spain.* ²*Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain.*
✉*e-mail:* jaguirre@cab.inta-csic.es

References
1. Wright, S. *Proceedings of the XI International Congress of Genetics* **8**, 209–222 (1932).
2. Gavrilets, S. *Trends Ecol. Evol.* **12**, 307–312 (1997).
3. Manrubia, S. et al. *Phys. Life Rev.* **38**, 55–106 (2021).
4. Greenbury, S. F., Louis, A. A. & Ahnert, S. E. *Nat. Ecol. Evol.* https://doi.org/10.1038/s41559-022-01867-z (2022).
5. Franke, J. & Klözer, A. *PLOS Comput. Biol.* **7**, e1002134 (2011).
6. Catalán, P., Arias, C. F., Cuesta, J. A. & Manrubia, S. *Biol. Direct* **12**, 1–15 (2017).
7. Aguirre, J., Catalán, P., Cuesta, J. A. & Manrubia, S. *Open Biol.* **8**, 180069 (2018).
8. Martin, N. S. & Ahnert, S. E. *J. R. Soc. Interface* **18**, 20210380 (2021).
9. Klug, A., Park, S.-C. & Krug, J. *PLOS Comput. Biol.* **15**, e1006884 (2019).

Competing interests
The author declares no competing interests.