The long and winding road to understanding organismal construction: Reply to comments on "From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics"

Susanna Manrubia[a,b,1,2], José A. Cuesta[b,c,d,e,1], Jacobo Aguirre[b,f], Sebastian E. Ahnert[g,h], Lee Altenberg[i], Alejandro V. Cano[j,k], Pablo Catalán[b,c], Ramon Diaz-Uriarte[l,m], Santiago F. Elena[n,o], Juan Antonio García-Martín[p], Paulien Hogeweg[q], Bhavin S. Khatri[r,s], Joachim Krug[t], Ard A. Louis[u], Nora S. Martin[g,h], Joshua L. Payne[j,k], Matthew J. Tarnowski[v], Marcel Weiß[g,h]

[a] *Department of Systems Biology, Centro Nacional de Biotecnología (CSIC), Madrid, Spain*
[b] *Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain*
[c] *Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés, Spain*
[d] *Instituto de Biocomputación y Física de Sistemas Complejos (BiFi), Universidad de Zaragoza, Spain*
[e] *UC3M-Santander Big Data Institute (IBiDat), Getafe, Madrid, Spain*
[f] *Centro de Astrobiología, CSIC-INTA, ctra. de Ajalvir km 4, 28850 Torrejón de Ardoz, Madrid, Spain*
[g] *Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, Cambridge, UK*
[h] *Sainsbury Laboratory, University of Cambridge, Cambridge, UK*
[i] *University of Hawai'i at Manoa, Hawai, US*
[j] *Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland*
[k] *Swiss Institute of Bioinformatics, Lausanne, Switzerland*
[l] *Department of Biochemistry, Universidad Autónoma de Madrid, Madrid, Spain*
[m] *Instituto de Investigaciones Biomédicas "Alberto Sols" (UAM-CSIC), Madrid, Spain*
[n] *Instituto de Biología Integrativa de Sistemas, I^{2}SysBio (CSIC-UV), València, Spain*
[o] *The Santa Fe Institute, Santa Fe, New Mexico, USA*
[p] *Centro Nacional de Biotecnología (CSIC), Madrid, Spain*
[q] *Theoretical Biology and Bioinformatics Group, Utrecht University, The Netherlands*
[r] *The Francis Crick Institute, London, UK*
[s] *Department of Life Sciences, Imperial College London, London, UK*
[t] *Institute for Biological Physics, University of Cologne, Köln, Germany*
[u] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, UK*
[v] *School of Biological Sciences, University of Bristol, Bristol, UK*

[1] Organizing and leading authors; contributed equally to this work
[2] Corresponding author; smanrubia@cnb.csic.es

The series of comments [1–4] on our review on the map between genotypes, phenotypes, and eventually organisms [5] are gratifying in that they are exactly the discussion we hoped our review would engender. As a snapshot of the state of the field at a moment in time, it is missing the new contributions that have already come after it —only partly amended through this reply—, but we hope it may provide a citation portal to find such developments forward in time.

Understanding how a viable organism unfolds from the engagement of genomic sequence information in a suitable, but often variable, environment is a daunting challenge, and one that we will not see solved in the near future. The different aspects that need to be considered in the process are too many to be enumerated; even if an exhaustive list could be provided, the quantitative effect of each variable in the construction of the functional organism —an effect that is expected to vary with the specifics of each unfolding—, would remain largely unknown. In our review [5] we highlighted some aspects of the problem, focusing on commonalities that emerge from studies of (mainly) genotype-phenotype (GP) maps and pointing out unavoidable limitations of broader analyses, such as the impossibility to explore the whole of genotype spaces, the inherent stochasticity of the evolutionary process (which calls for statistical-mechanical approaches and limits its long-term predictability), or important epistemological difficulties to characterize a complete genotype-to-organism map.

In four comments to our review [1–4], further aspects of the problem are pointed out, and the discussion is enriched by different viewpoints and the addition of related open questions. Certainly, GP models have been instrumental to clarify important evolutionary features, such as the compatibility between high robustness and evolvability [2]. The existence of universal or quasi-universal properties of GP maps, regardless of the model under study, suggests that only major features are essential in biologically acceptable GP maps [6]: these might be summarized in the ability to navigate genotype spaces without losing functionality, which is a condition for innovation to be possible. All GP maps of interest fulfill this constraint, even if they differ in the details [3] and even when some of them yield a minute fraction of genotypes mapping onto functional phenotypes [4].

Nitash and Adami [4] call attention to the relevance of measuring the information content of sequences to complement analyses of genotype spaces structure. This is closely related to the concept of sequence entropy —the logarithm of the number of sequences yielding a phenotype—, which we describe in detail using

a statistical mechanical framework of phenotypic evolution in the weak mutation regime (section 5.5 in [5]). As an example, Nitash and Adami investigate the set of the smallest self-replicators in the digital life system Avida [7] and show that an information-theoretic analysis reveals details about the robustness of replicators and its relationship with the sign of epistasis in their genotypes.

McCandlish [3] argues that the biophysical characteristics of specific systems play a role in evolutionary dynamics. The formation of GU pairs in RNA permits transitions between GC, GU and AU pairings, but changes from GC to CG, for instance, require rare double mutations. Therefore, neutral sets for RNA secondary structure consist *a priori* of many disconnected components depending on the choice for each base pair [3]. A question in this respect, however, is whether such disconnected components are equally abundant and, more importantly, whether disconnection implies lack of evolvability. Firstly, the components of the neutral network for a given RNA secondary structure differ in size when folding energy is taken into account [8, 9]. Indirect studies suggest that the largest among such components could become progressively dominant as the length of RNA sequences increases (see [6] and references therein). Second, fragmentation due to lack of percolation only holds for small components, a relevant observation when it comes to ensuring navigability of genotype spaces.

Another property of the RNA sequence-to-secondary structure map is that single mutations can cause large rearrangements, leading to punctuated patterns of adaptation [10]. Still, phenotypic stasis followed by sudden changes is not exclusive of adaptive changes in RNA secondary structure, since it has been empirically observed, e.g., in antigenic changes in viral populations [11]. It can be argued that any GP map that breaks the genotype space in a set of disjoint genotype networks for different phenotypes will by definition display sudden phenotypic transitions [12]. Actually, the fixation of a fitter phenotype in a population is in general an exponentially fast process compared with the search for new phenotypes, and the comparison between these two broadly different time scales can be described as punctuated equilibrium-like: sudden major changes interrupt prolonged stasis, over the relevant time scales, regardless of the mutational mechanism considered and the definition of phenotype.

As argued by McCandlish [3], the biophysical characteristics of particular evolutionary systems are important and do have an explanatory power. Also, system-specific and universal properties of GP maps can simultaneously hold: phenotypic bias is a universal property, but which phenotypes have high phenotypic frequencies is set by the biophysics of the system (e.g. number of stacks [13] and base pairs [6] in RNA, contact traces [14] in proteins). System-specific fea-

tures also yield important clues to understand how a variety of GP maps achieve robustness and evolvability. In two-letter alphabets, for example, neutral networks are small and disconnected [15], seriously compromising navigability. This situation is analogous to the phenomenon of quenched disorder at large population sizes occurring in the weak mutation regime [16]. This is also the case of systems with highly compressed information, as the smallest self-replicators in Avida [4]. However, navigability (and evolvability) can be restored if information becomes redundant in the genotype, if additional levels are added on top of the basic sequence-structure map [17], or if the population size is reduced allowing genetic drift to traverse local fitness valleys [16].

While the vast majority of RNA sequences stably fold into a minimum-free-energy secondary structure, functional phenotypes are extremely rare in other GP maps, as toyLIFE [18], regulatory gene networks [19] or metabolic reaction networks [20]. These systems display sufficient redundancy, but navigability relies on the existence of non-trivial correlations between genotypes: meaningful function clusters in the space of possible genotypes, and can be maintained while alternative phenotypes are explored—often through a variety of mutational mechanisms. When it comes to adaptation, nature uses multiple other tricks that are rarely considered in GP map studies, such as molecular mimicry [21], protein moonlighting [22] or enzyme promiscuity [23]. Inclusion of these forms of phenotypic redundancy and functional flexibility explicitly turn the GP map into a many-to-many relationship, and may significantly modify its topology. Our knowledge on the large-scale structure induced by generic GP maps is very limited, and future progress critically depends on our ability to extract model-independent features relevant in evolutionary and adaptive dynamics. In this respect, synthetic systems, such as Avida [7] or Dawkins' biomorphs [24], can provide important clues on the nature of universal features in evolving systems, and serve as examples of alternative solutions that differ in the details but coincide in essential mechanisms.

GP maps are limited to specific aspects of the development process, unavoidably leaving a large gap between the molecular phenotypes in many models and the properties of whole cells or organisms [25]. The next crucial level to integrate in the overall description is the map from genotype to fitness (GF). Though the GF map does not always require an explicit definition of phenotype, the connection between phenotype and fitness, when possible, becomes essential for predictions of evolution by natural selection, as de Visser points out [1]. The connection between genotype and fitness has been explored in our review [5], aware as we are of the non-trivial role played by phenotypic bias and GP maps in

evolutionary dynamics (e.g. in speciation [26]). Basic organismal body plans, for instance, are conserved over many millions of years through stabilizing selection, and yet the underlying gene regulation and genetic makeup between species can dramatically diverge —this is known as developmental system drift. Similarly, as also discussed in [5], basic protein folds are under strong stabilising selection and highly conserved: simple models of GP maps suffice to explain the observation of the marginal stability of proteins [25, 27].

This nonetheless, various other possibilities have to be considered to complete the GF relationship, unavoidably complicating the description. Fitness is difficult to quantify due to its dependence on exogenous and endogenous variation. Fitness can be only defined in a specific environment, since the adaptive value of a phenotype (or a genotype) depends on the conditions under which it has to function: experiments with viral populations nicely illustrate this dependence by showing how the value of mutations depends on the infected host [28]. Actually, a strict separation between phenotype and fitness is in general not possible. The ability of a genotype to produce different phenotypes when exposed to different environments (phenotypic plasticity) causes an intimate, environment-mediated relationship between genotype, phenotype, and fitness: RNA sequences, for example, yield a simple case because they fold into different secondary structures depending on temperature [29], pH or the molecular context. The scenario becomes more complex when we consider that organisms themselves interact with the environment in non-trivial ways. Organism-environment interactions alter the fitness landscape through short-term (e.g. metabolic [30]) and long-term (as in an extended phenotype [31] and niche construction) modifications of expressed phenotypes that in turn affect both fitness and the environmental conditions [25]. Such feedback loops between phenotype and environment might lead to non-commutativity of mutations [32] (the fitness landscape in the neighborhood of a given genotype depends on the evolutionary history [33]), turning evolutionary pathways highly dependent on contingent events.

Can the (quasi-)universal features of GP maps be extrapolated to GF maps? This is a relevant question raised by de Visser [1], who suggests future research avenues: the development of fitness models of phenotypes (possibly including feedback loops from the metabolic activity of the organism, to begin with) and the exploration of how GP maps change under environmental variation. Efforts in these two directions exist but have not yet been conceptually integrated to the degree that multiple GP models have. Some instances of feedback between the molecular environment and the unfolded (functional) genotypes have been addressed in our review [5] when we discuss GP maps as evolving objects. Such

an evolution is the result of self-organization in evolving populations, where relevant biological functions themselves (equivalent to a GF map) emerge from intra-specific interactions in an RNA molecular quasispecies [34, 35]. Few other models have addressed the evolution of the GP map in a complex, cellular-like environment under a fixed fitness function [36]. Adaptation in variable fitness landscapes has been investigated through the introduction of seascapes [37], for example, though the inability of static fitness landscapes to capture environmental changes already worried Wright himself [38]. Despite its relevance, this latter question has received limited attention so far.

Recent work [39] on genotype to phenotype to fitness maps has demonstrated that the nonlinear relationships between these different levels can interact in surprising ways, reducing, for example, the amount of reciprocal sign epistasis, that normally frustrates adaptive dynamics. For transcription factor landscapes, these interactions lead to an important enhancement of the likelihood that low- and intermediate-affinity binding sites fix in a population, over and above the *arrival of the frequent* effect [40] that enhances the likelihood that such phenotypes appear as potential variation in the first place.

More generally, it has recently been argued [41] that explicitly including the phenotype as an intermediate step between genotype and fitness, and therefore implicitly including a number of key organizational properties of the mapping from genotypes to phenotypes – such as large neutral networks, neutral correlations (or high mutational robustness) that facilitate neutral exploration and high dimensions—greatly increases the number of accessible paths with monotonically increasing fitness or navigability, in associated fitness landscapes, even under a worst-case scenario of random fitness assignment to the phenotypes.

Such enhanced navigability may help explain the remarkably tight correlation between the frequency with which RNA secondary structures are found in nature, and the frequency with which they arise as potential variation [42]. Of course at the level of individual phenotypes, fitness matters greatly, but at the level of distributions (such as the probability of obtaining a certain coarse-grained shape) the fitness effects apparently wash out in this system.

A similar strong sculpting by the GP map has been observed for the distribution of protein cluster topologies from the Protein Data bank, providing a non-adaptive explanation for the strong preference for symmetry observed in nature [43]. Of course fitness effects play an important role in all these systems, but nature can only fix those phenotypes that appear in the first place.

Arguments based on the coding theorem from algorithmic information theory [44] suggest that, under some mild assumptions, GP maps should be gener-

ically biased with an exponential drop in the size of neutral sets with linear increases in the descriptional (Kolmogorov) complexity of the phenotype [43]. One of the really interesting open questions is whether the strong bias towards simple phenotypes leads to phenotypes that are also favored by natural selection. One potential property to investigate is higher-than-average robustness exhibited by these preferred phenotypes. This property will in turn, make it easier to accumulate other functions, and so may facilitate the emergence of modularity, making evolution more evolvable.

In a scenario that seems to become increasingly complex, it is important to come back to the relevant question: Do we need to know all details of GF maps to predict (even if in a probabilistic way) the evolutionary dynamics of populations? Manhart and Bonhoeffer [2] provide an important clue by noting that recent investigation of some phenotype-fitness maps reveals a low dimensionality (of order 10) of fitness-relevant phenotypes [45, 46], supporting the hypothesis advanced in [47] that phenotypic variation falls within low-dimensional spaces. This is an encouraging possibility, though the point is whether the properties of phenotype to fitness maps can be generalized across environmental conditions, erasing possible idiosyncratic dependencies of fitness on environments and, eventually, rendering fitness landscapes a useful tool to predict evolutionary dynamics [2].

References

[1] J. A. G. de Visser, Genotype-phenotype maps and the predictability of evolution: Comment on "From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics" by Susanna Manrubia et al., Physics of Life Reviews 39 (2021) 79–81. doi:https://doi.org/10.1016/j.plrev.2021.08.003.

[2] M. Manhart, S. Bonhoeffer, The search for universality in evolutionary landscapes: Comment on "From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics" by Susanna Manrubia, José A. Cuesta, et al., Physics of Life Reviews 39 (2021) 76–78. doi:https://doi.org/10.1016/j.plrev.2021.08.004.

[3] D. M. McCandlish, System-specificity of genotype-phenotype map structure: Comment on "From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics" by Su-

sanna Manrubia et al., Physics of Life Reviews 39 (2021) 73–75. doi:https://doi.org/10.1016/j.plrev.2021.08.005.

[4] Nitash C G, C. Adami, Information-theoretic characterization of the complete genotype-phenotype map of a complex pre-biotic world: Comment on "From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics" by Susanna Manrubia et al., Physics of Life Reviews 38 (2021) 111–114. doi:https://doi.org/10.1016/j.plrev.2021.06.001.

[5] S. Manrubia, J. A. Cuesta, J. Aguirre, S. E. Ahnert, L. Altenberg, A. V. Cano, P. Catalán, R. Diaz-Uriarte, S. F. Elena, J. A. García-Martín, P. Hogeweg, B. S. Khatri, J. Krug, A. A. Louis, N. S. Martin, J. L. Payne, M. J. Tarnowski, M. Weiß, From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics, Physics of Life Reviews 38 (2021) 55–106. doi:https://doi.org/10.1016/j.plrev.2021.03.004.

[6] J. A. Garcia-Martin, P. Catalán, J. A. Cuesta, S. Manrubia, Statistical theory of phenotype abundance distributions: A test through exact enumeration of genotype spaces, Europhys. Lett. 123 (2018) 28001.

[7] C. Adami, C. Brown, Evolutionary Learning in the 2D Artificial Life System "Avida", in: Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems, The MIT Press, 1994. doi:10.7551/mitpress/1428.003.0049.

[8] S. F. Greenbury, S. Schaper, S. E. Ahnert, A. A. Louis, Genetic correlations greatly increase mutational robustness and can both reduce and enhance evolvability, PLoS Comput. Biol. 12 (2016) e1004773.

[9] W. Grüner, R. Giegerich, D. Strothmann, C. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, P. Schuster, Analysis of RNA sequence structure maps by exhaustive enumeration II. Structures of neutral networks and shape space covering, Monatsh. Chem. 127 (1996) 375–389.

[10] M. A. Huynen, P. F. Stadler, W. Fontana, Smoothness within ruggedness: The role of neutrality in adaptation, Proc. Natl. Acad. Sci. USA 93 (1996) 397–401.

[11] K. Koelle, S. Cobey, B. Grenfell, M. Pascual, Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans, Science 314 (2006) 1898–1903.

[12] J. Aguirre, P. Catalán, J. A. Cuesta, S. Manrubia, On the networked architecture of genotype spaces and its critical effects on molecular evolution, Open Biol. 8 (2018) 180069.

[13] K. Dingle, S. Schaper, A. A. Louis, The structure of the genotype-phenotype map strongly constrains the evolution of non-coding RNA, Interface Focus 5 (2015) 20150053.

[14] J. L. England, E. I. Shakhnovich, Structural determinant of protein designability, Phys. Rev. Lett. 90 (2003) 218101.

[15] C. Holzgräfe, A. Irbäck, C. Troein, Mutation-induced fold switching among lattice proteins, J. Chem. Phys. 135 (2011) 195101.

[16] B. S. Khatri, T. C. B. McLeish, R. P. Sear, Statistical mechanics of convergent evolution in spatial patterning, Proc. Natl. Acad. Sci. USA 106 (2009) 9564–9569.

[17] P. Catalán, A. Wagner, S. Manrubia, J. A. Cuesta, Adding levels of complexity enhances robustness and evolvability in a multi-level genotype-phenotype map, J. R. Soc. Interface 15 (2018) 20170516.

[18] C. F. Arias, P. Catalán, S. Manrubia, J. A. Cuesta, toyLIFE: a computational framework to study the multi-level organisation of the genotype-phenotype map, Sci. Rep. 4 (2014) 7549.

[19] S. Ciliberti, O. C. Martin, A. Wagner, Innovation and robustness in complex regulatory gene networks, Proc. Natl. Acad. Sci. USA 104 (2007) 13595–13596.

[20] J. F. Matias Rodrigues, A. Wagner, Evolutionary plasticity and innovations in complex metabolic reaction networks, PLoS Comput. Biol. 5 (2009) e1000613.

[21] K. W. Wucherpfennig, Structural basis of molecular mimicry, Journal of Autoimmunity 16 (2001) 293–302.

[22] J. Piatigorsky, Gene sharing and evolution: the diversity of protein functions, Harvard University Press Cambridge MA:, 2007.

[23] O. Khersonsky, D. S. Tawfik, Enzyme promiscuity: a mechanistic and evolutionary perspective, Annu. Rev. Biochem. 79 (2010) 471–505.

[24] R. Dawkins, The evolution of evolvability, in: S. Kumar, P. J. Bentley (Eds.), On Growth, Form and Computers, Academic Press, London, 2003, pp. 239–255.

[25] S. Bershtein, A. W. Serohijos, E. I. Shakhnovich, Bridging the physical scales in evolutionary biology: From protein sequence space to fitness of organisms and populations, Curr. Opin. Struc. Biol. 42 (2017) 31–40.

[26] B. S. Khatri, R. A. Goldstein, Biophysics and population size constrains speciation in an evolutionary model of developmental system drift, PLoS Comput. Biol. 15 (2019) 1–24.

[27] D. M. Taverna, R. A. Goldstein, Why are proteins marginally stable?, Proteins 46 (2002) 105–109.

[28] S. Bedhomme, J. Hillung, S. F. Elena, Emerging viruses: why they are not jacks of all trades?, Curr. Opin. Virol. 10 (2015) 1–6.

[29] P. Catalán, C. F. Arias, J. A. Cuesta, S. Manrubia, Adaptive multiscapes: An up-to-date metaphor to visualize molecular adaptation, Biol. Direct 12 (2017) 7.

[30] F. Pinheiro, O. Warsi, D. Andersson, M. Lässig, Metabolic fitness landscapes predict the evolution of antibiotic resistance, Nature Ecology & Evolution 5 (2021) 1–11. doi:10.1038/s41559-021-01397-0.

[31] R. Dawkins, The extended phenotype, Oxford University Press, Oxford, UK, 1999.

[32] D. Bajić, J. C. C. Vila, Z. D. Blount, A. Sánchez, On the deformability of an empirical fitness landscape by microbial evolution, Proceedings of the National Academy of Sciences 115 (2018) 11286–11291. doi:10.1073/pnas.1808485115.

[33] A. Agarwala, D. S. Fisher, Adaptive walks on high-dimensional fitness landscapes and seascapes with distance-dependent statistics, Theoretical Population Biology 130 (2019) 13–49. doi:https://doi.org/10.1016/j.tpb.2019.09.011.

[34] N. Takeuchi, P. Hogeweg, Evolution of complexity in rna-like replicator systems, Biol. Direct 3 (2008) 11.

[35] E. S. Colizzi, P. Hogeweg, Evolution of functional diversification within quasispecies, Genome Biol. Evol. 6 (2014) 1990–2007.

[36] T. D. Cuypers, J. P. Rutten, P. Hogeweg, Evolution of evolvability and phenotypic plasticity in virtual cells, BMC Evol. Biol. 17 (2017) 60.

[37] V. Mustonen, M. Lässig, From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation, Trends Genet. 25 (2009) 111–119.

[38] S. Wright, The roles of mutation, inbreeding, crossbreeding and selection in evolution, Proc. 6^{th} Int. Congr. Genet. 1 (1932) 356–366.

[39] M. Srivastava, J. L. Payne, The transformability of genotype-phenotype landscapes, bioRxiv (2022).

[40] S. Schaper, A. A. Louis, The arrival of the frequent: How bias in genotype-phenotype maps can steer populations to local optima, PLoS ONE 9 (2014) e86635.

[41] S. F. Greenbury, A. A. Louis, S. E. Ahnert, The structure of genotype-phenotype maps makes fitness landscapes navigable, bioRxiv (2021).

[42] K. Dingle, F. Ghaddar, P. Šulc, A. A. Louis, Phenotype bias determines how natural rna structures occupy the morphospace of all possible shapes, Molecular biology and evolution 39 (2022) msab280.

[43] I. G. Johnston, K. Dingle, S. F. Greenbury, C. Q. Camargo, J. P. Doye, S. E. Ahnert, A. A. Louis, Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution, Proceedings of the National Academy of Sciences 119 (2022) e2113883119.

[44] K. Dingle, C. Q. Camargo, A. A. Louis, Input–output maps are strongly biased towards simple outputs, Nature communications 9 (2018) 1–7.

[45] T. P. Wytock, A. E. Motter, Predicting growth rate from gene expression, Proceedings of the National Academy of Sciences 116 (2019) 367–372. doi:10.1073/pnas.1808080116.

[46] T. Maeda, J. Iwasawa, H. Kotani, N. Sakata, M. Kawada, T. Horinouchi, A. Sakai, K. Tanabe, C. Furusawa, High-throughput laboratory evolution reveals evolutionary constraints in escherichia coli, Nature Communications 11 (2020) 5970. doi:10.1038/s41467-020-19713-w.

[47] L. Altenberg, Modularity in evolution: Some low-level questions, in: W. Callebaut, D. Rasskin-Gutman, H. A. Simon (Eds.), Modularity: Understanding the Development and Evolution of Complex Natural Systems, MIT Press, Cambridge, MA, 2005.