PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Review

Cite this article: Sambamoorthy G, Sinha H, Raman K. 2019 Evolutionary design principles in metabolism. *Proc. R. Soc. B* **286**: 20190098. http://dx.doi.org/10.1098/rspb.2019.0098

Received: 14 January 2019
Accepted: 14 February 2019

Subject Category:
Evolution

Subject Areas:
systems biology, evolution

Keywords:
adaptation, metabolic networks, robustness, systems biology, design principles

Authors for correspondence:
Himanshu Sinha
e-mail: sinha@iitm.ac.in
Karthik Raman
e-mail: kraman@iitm.ac.in

Evolutionary design principles in metabolism

Gayathri Sambamoorthy[1,2,3], Himanshu Sinha[1,2,3] and Karthik Raman[1,2,3]

[1]Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, [2]Initiative for Biological Systems Engineering (IBSE), and [3]Robert Bosch Centre for Data Science and Artificial Intelligence (RBCDSAI), Indian Institute of Technology Madras, Chennai 600036, India

HS, 0000-0001-7031-0491; KR, 0000-0002-9311-7093

Microorganisms are ubiquitous and adapt to various dynamic environments to sustain growth. These adaptations accumulate, generating new traits forming the basis of evolution. Organisms adapt at various levels, such as gene regulation, signalling, protein–protein interactions and metabolism. Of these, metabolism forms the integral core of an organism for maintaining the growth and function of a cell. Therefore, studying adaptations in metabolic networks is crucial to understand the emergence of novel metabolic capabilities. Metabolic networks, composed of enzyme-catalysed reactions, exhibit certain repeating paradigms or *design principles* that arise out of different selection pressures. In this review, we discuss the design principles that are known to exist in metabolic networks, such as functional redundancy, modularity, flux coupling and exaptations. We elaborate on the studies that have helped gain insights highlighting the interplay of these design principles and adaptation. Further, we discuss how evolution plays a role in exploiting such paradigms to enhance the robustness of organisms. Looking forward, we predict that with the availability of ever-increasing numbers of bacterial, archaeal and eukaryotic genomic sequences, novel design principles will be identified, expanding our understanding of these paradigms shaped by varied evolutionary processes.

1. Introduction

Microorganisms are ubiquitous and survive in dynamic environments where the level and nature of available nutrients change, for instance, the sources of carbon, nitrogen, phosphorous and the levels of oxygen. Sessile organisms are posed with a challenge to survive *in situ* and adapt in varying environments. Organisms adapt and acquire novel abilities predominantly by processes of gain-of-function and loss-of-function mutations, as well as horizontal gene transfer [1]. Adaptive evolution studies of *E. coli* suggest that some mutations result in gain-of-functionality, such as the ability to grow on a non-native carbon source [2]. Adaptation can also progress by loss-of-function mutations, where a change in the regulatory circuitry can rewire metabolism, resulting in adaptive fitness benefits [3]. Furthermore, bacteria adapt to changing environments predominantly by horizontal gene transfer, acquiring new genes, some of which establish novel functions [4]. Through these adaptive and selective processes, microbes have acquired a wide range of abilities—to grow in extreme physical conditions such as high and low temperatures, osmotic stress, low pH, high salt concentrations, high pressure and nutrient environments that are rich in sulfur, methane and iron [5–9]. Numerous examples of adaptations by organisms exist in nature, of which some are remarkable, e.g. *Sphingobium chlorophenolicum*, an organism that evolved to grow on pentachlorophenol, an anthropogenic pesticide, as a carbon source [10–12].

Adaptations in an organism occur at various levels, such as gene regulation, signalling, protein–protein interactions and metabolism. Of these, metabolism is one of the most complex cellular processes, and perhaps most vital for

THE ROYAL SOCIETY PUBLISHING

Table 1. Design principles in other networks. (The design principles discussed in the context of metabolism have been observed in other types of networks, such as gene regulatory and signalling networks. However, note that flux coupling, by definition, is applicable only to metabolic networks.)

design principle	description	reference
modularity	modules in gene expression networks for various cancer types	[19]
	modularity of feed-forward loops	[20]
	modularity and robustness in signalling networks	[21]
	modularity for establishing synthetic tools	[22]
	modularity in bacterial protein networks	[23]
synthetic lethality	synthetic genetic array analysis of *Escherichia coli* and yeast	[24,25]
	synthetic lethality analysis to target *Mycobacterium tuberculosis*	[26]
	synthetic lethal interaction of *PLK1* inhibitors and microtubule-destabilizing drug	[27]
	drug target identification in *Staphylococcus aureus* using synthetic lethality analysis	[28]
	synthetic lethality for targeting cancer	[29]
	understanding the redundant roles of teichoic acid polymers in regulating cell division, based on synthetic lethality	[30]
exaptation	exaptation of transposable elements—derived sequences	[31]
	functional exaptation of transposable elements in higher plants	[32]
	exaptation of the scnRNA pathway for regulating cellular genes	[33]
	understanding the mechanisms mediating the substrate preference of LuxI protein family through exaptation	[34]

maintaining several physiological functions. The adaptations and evolutionary changes that occur in metabolic processes are crucial, as these processes produce essential metabolites and small molecules that promote growth and manage adverse conditions such as temperature, oxidative and osmotic stresses [13,14]. Metabolism is composed of a complex network of numerous gene-encoded enzyme-catalysed reactions that use nutrients from external and internal environments to synthesize compounds necessary for the growth of organisms in a given environmental condition. Thus, a cell's *metabolic network* comprises all reactions across various biosynthetic and degradation pathways hitherto known to exist in an organism. The architecture and activity of this network are greatly influenced by the nutrients available for assimilation from the environment. A complex network of gene regulation underlies the switching on and off of various reactions, under different environmental conditions. These reactions, which are components of various metabolic pathways, are *active* in different scenarios. Therefore, building on a given *base* metabolic network, multiple architectures or *topologies* for the metabolic networks can exist. As these multiple topologies are interdependent, they need to be studied together to get a comprehensive understanding of the entire metabolic network.

Adaptations can facilitate new capabilities such as the ability to grow in novel environments, via changes in the architecture of the metabolic network. These changes typically increase the fitness and viability of an organism in new environments and enable it to evolve. Across the metabolic networks of diverse organisms, a remarkable variation is observed; this diversity represents a 'universe' of possible metabolic reactions, some of which an organism can acquire for adaptation [4]. However, this ability of an organism to acquire new metabolic reactions is limited by the extent of microbiome diversity in its environmental niche and its inherent metabolic capabilities. As an organism evolves in multiple environments, mutations tend to accumulate, unless genetic interactions between these mutations are detrimental and selected against. Even in such a scenario, where certain genetically interacting mutations are mildly detrimental, many organisms tend to accumulate these combinations, if they have a fitness advantage in some other environmental conditions [15–18].

In this review, we illustrate how a few repeating paradigms, or *design principles*, namely functional redundancy, flux coupling, modularity and exaptation, have emerged in diverse organisms enabling them to adapt and survive in a multitude of environments. Although design principles have previously been elucidated in other types of networks as well (table 1), in this review, we focus specifically on metabolic networks. It is important to note that these design principles are not strategies preferred by the organism *per se* but are essentially *post facto* observations of paradigms that have emerged as a consequence of the evolutionary processes that operate in different environments. We discuss several *in silico* and empirical studies that illustrate these recurring principles in metabolic networks, and summarize the progress that has been made towards understanding various underlying aspects of metabolic network structure and regulation.

2. Design principles in metabolism

(a) Functional redundancy

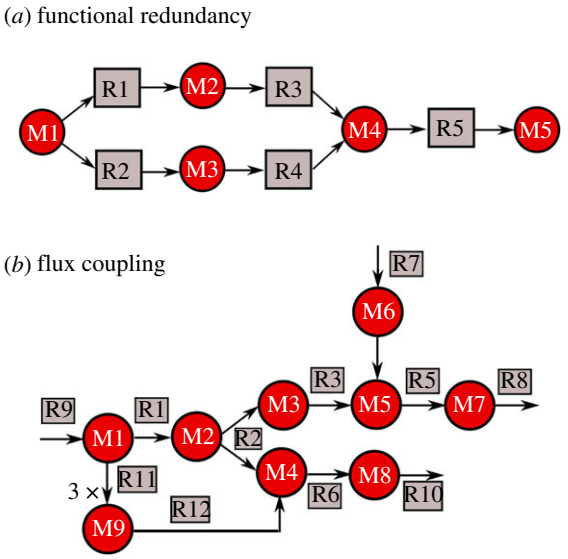
Functional redundancy is a defining feature of all metabolic networks. Functional redundancy is characterized by the ability of an organism to use alternative fluxes (reactions, their corresponding enzymes or genes) in a metabolic network in a given environment, which can completely or partially compensate for the loss of the other [35]. Through functional redundancy of key reactions or enzymes,

metabolic networks have evolved to reduce the chances of disruption of core metabolic pathways. A highly powerful approach, particularly suited to study and understand functional redundancy in metabolic networks, is the analysis of synthetic lethal (SL) combinations or synthetic lethality. An SL pair in an organism refers to a pair of genes or reactions, which when deleted simultaneously, lead to a lethal phenotype [36], but the organism can sustain growth when even one of the genes/reactions in the pair is present. An SL, therefore, is the extreme case of a genetic interaction, where the combined effect of two genetic mutations differs markedly from those of their individual effects [37]. Figure 1*a* depicts a schematic metabolic network where metabolite M5 is essential for the survival of the organism. Here, the reactions {R1, R2} form a synthetic lethal pair; only upon deletion of both these reactions, the metabolite M5 cannot be produced. Thus, the pathways {M1 → R1 → M2 → R3 → M4} and {M1 → R2 → M3 → R4 → M4} are redundant pathways, ultimately producing the same metabolite.

Functional redundancy is usually achieved when genes, pathways, chromosomes or genomes get duplicated. The duplicated set of metabolic genes and reactions often leads to conservation of metabolic function; alternatively, these duplicated genes may diversify, to *innovate* novel functions. Gene duplicates play a major role in the maintenance of metabolic functions; indeed, they are known to occur more frequently in metabolic proteins than non-metabolic proteins [38]. Genes involved in central metabolism, catabolic pathways and reactions that carry a high metabolic flux also tend to have a higher rate of duplication compared to genes in anabolic pathways [38,39]. Some of these duplicate genes acquire specialized functions and possess not just a single essential function but an array of overlapping functions [40]. The selection pressure that retains such duplicates is still not clearly understood [40]. However, some studies suggest that redundant genes are preserved by evolution; analysis of gene expression data in yeast and mammals shows that the expression levels of duplicated genes decrease, following the duplication event [41]. Owing to the reduced selection pressure on genes with lowered expression, the retention of duplicate genes is facilitated [41]. Apart from studying SLs in metabolic reaction networks *in silico*, SLs have also been studied *genetically*, by deleting pairs of genes and testing for growth in a given environmental condition [25]. Extensive studies in yeast have shown that understanding gene–gene interactions gives insight into the structure and function of the genetic networks and also the SL interactions that exist [25,42]. A study on yeast genetic interactions suggests that at least 23% of the synthetic lethal genetic interactions are conserved between *Schizosaccharomycetes pombe* and *Saccharomycetes cerevisiae*. Although the two organisms are separated by hundreds of million years of evolution, they possess conserved synthetic interactions and essential genes, suggesting that evolutionary pressures play a significant role in retaining such interactions, even in eukaryotes [43].

SLs in metabolic networks manifest either as redundant genes, reactions or pathways. Functionally redundant reactions and pathways identified from SL analyses play an important role in conferring robustness against single mutations. Deletion of an SL pair abrogates growth (figure 1*a*), owing to the deleterious effect on one or more key metabolic functions, and gives insights into the functionally redundant reactions or pathways [35]. Thus, the identification of SL reaction pairs is a powerful technique to identify alternative pathways that can carry out similar functions. In the majority of cases, SL pairs occur in alternative pathways where one reaction is active in the wild-type and the other is inactive [35]. The deletion of inactive pathways does not have phenotypic effects; upon deletion of the active reaction, fluxes are re-routed through the alternative pathway, enabling the organism to grow [35,44]. Systems-level analyses of the metabolic networks of *Escherichia coli* and yeast found that most of the redundant reactions are preserved, contributing directly towards increasing fitness. Although these redundancies can be regarded as adaptations against deleterious mutations, the evolutionary mechanisms that retain such overlapping pathways are still not clear [45]. Despite the fact that multiple alternative pathways exist, evolution is believed to select those that are known to increase fitness. A classic example is that of alternative pathways for glycolysis and gluconeogenesis [46]. Although multiple pathways exist, the other pathways carry lower flux when compared with the cardinal pathway in a given physiological state. However, the alternative pathways may supersede the cardinal pathway under other physiological conditions.

A systems-level flux balance analysis of yeast in a wide range of environments suggests that genetic interactions, particularly synthetic lethals, have a strong dependence on the existing environmental conditions [47]. A pair of synthetic lethal genes is environment-dependent, i.e. they are not redundant in all environments, and therefore, their ability to compensate varies considerably. Thus, environmental constraints play a key role in restricting the alternative pathways that are open to selection.



examples: fully coupled: R1-R2, R2-R3, R4-R7; partially coupled: R6-R9, R9-R10; directionally coupled: R1-R5, R2-R6, R3-R8; uncoupled: R1-R7, R2-R11, R3-R12;

Figure 1. Functional redundancy and flux coupling. (*a*) The figure shows a sample metabolic network to illustrate redundancy. Circles denote metabolites (M1, M2, etc.), rectangles represent reactions (R1, R2, etc.). The reactions pairs {R1, R2}, {R1, R4}, {R2, R3}, {R3, R4} form synthetic lethal pairs and the pathways {M1 → R1 → M2 → R3 → M4} and {M1 → R2 → M3 → R4 → M4} are redundant for the production of the metabolite M4. (*b*) A sample metabolic network where circles denote metabolites (M1, M2, etc.) and rectangles represent reactions. All the stoichiometric coefficients in the network are 1 except reaction R11, where 1 mol of M1 is converted to 3 mol of M9. The three types of flux coupling are illustrated.

Microorganisms have thus evolved to achieve redundancy in key reactions and pathways that are vital for the growth of the organism; the redundancy enables buffering, and rescue from single mutations. For instance, following gene duplication, the organism might have retained both genes although one of them would have since evolved to perform a different function still preserving the native function. Furthermore, as a result of adaptation to a new environment, it is possible that reactions which perform different functions adapt to compensate for one another [48].

(b) Flux coupling

The fluxes of reactions in metabolic networks are seldom independent; they are often correlated across different metabolic pathways. A comprehensive understanding of such correlations or dependencies between reactions is achieved through flux coupling analysis (FCA) [49]. FCA reveals the functional dependencies between reactions and the coordination of regulation across a metabolic network. FCA also provides insights into gene essentiality, gene regulation, network evolution and metabolic network hierarchy [50,51].

Reactions can be coupled in three different ways: fully, partially or directionally (figure 1*b*). In fully coupled reactions, a flux in one reaction implies a fixed flux in the other reaction, such that their fluxes are proportional to each other (reactions {R1, R2} and {R2, R3} in figure 1*b*). In partially coupled reactions, the flux of one reaction implies flux in the other reaction but their fluxes are not proportional, i.e. either one of the reactions can have a variable flux (reactions {R6, R9} and {R9, R10} in figure 1*b*). In both the above categories, if one reaction carries a flux, it is essential that the other reaction carries a flux as well. In the case of directionally coupled reactions, when the first reaction carries a non-zero flux, the second reaction also has a non-zero flux; however, the converse does not hold, thus illustrating an asymmetric (i.e. directional) dependence between the reactions (reactions {R1, R5} and {R2, R6} in figure 1*b*) [49]. The reactions that do not fall into any of the above categories are said to be uncoupled (reactions {R1, R7} and {R2, R11} in figure 1*b*). Efficient methods have been developed for identifying coupled reactions, notably flux coupling finder (FCF) [49], feasibility-based flux coupling analysis (FFCA) [52] and F2C2 [50].

Flux coupling analyses have shown that genes, whose reaction fluxes are fully coupled, are co-expressed and co-regulated; further, many such genes also share transcriptional regulators or reside in the same operon [53]. Flux coupling relationships have been used to construct a hierarchical flux coupling graph for *E. coli* metabolism, illustrating that the reactions essential for a wide range of environments are conserved during evolution [51]. This coupling graph also established that newly acquired reactions (e.g. via horizontal gene transfer), which facilitate adaptation to specific environments, get placed at the top of the hierarchy [51]. Identification of fully and directionally coupled reactions in *E. coli* and yeast reflects that a large fraction of genes catalysing such reactions are co-expressed and are regulated by the same transcription factor [53,54]. The nature of flux coupling gives a notion of the extent of genes being co-regulated [53]. Directionally coupled reactions give insight into the asymmetric relationship between proteins [55]. Such relationships are consistent with asymmetry in gene expression patterns [55].

Furthermore, FCA helps in identifying reactions that directly have an effect on the phenotype of the organism, e.g. the reactions that are directly coupled with growth. These reactions, which are essential for cell growth, remain conserved during evolution [51]. Adaptation of organisms to new environments results in the addition of new reactions; as evolution progresses, some of these are retained as they are coupled with essential reactions. A flux coupling study of 23 organisms from different kingdoms of life identified the smallest number of reactions, the *driver set*, which can control the entire metabolic network of an organism. These reactions are involved in complex transcriptional and post-transcriptional regulation in the metabolic network of *E. coli*; thus, the regulation of the cell depends on these driver reactions for efficient control of the entire metabolic network. Such driver reactions can also serve as targets for optimization of cellular metabolite production and are thus useful for metabolic engineering applications. The study also identified driver reactions in human cancer cells, which are known to play a major role in malignancy. Consequently, these reactions can serve as potential drug targets [54].

In summary, FCA is crucial to identifying essential genes and reactions in metabolic networks, and can provide information about the co-occurrence of genes, co-expression of genes and the evolution of metabolic networks. Notably, this is possible without even the knowledge of the genes themselves. FCA can aid in the identification of drug targets for pathogenic organisms [54]. FCA also has important implications for metabolic engineering applications because it helps in identifying the few key reactions that may alter the flow of flux through the entire metabolic network. Overall, the analysis of coupling of reaction fluxes in a metabolic network can reveal functional dependencies and shed light on the coordination of regulation across the network.

(c) Modularity

Biological networks are highly modular in nature, which means that they can be decomposed into subunits, or *modules*, which can function relatively independently and are highly interconnected [56] (figure 2*a*). Nodes (genes or reactions) within a module are more densely connected when compared with nodes across modules. Biological networks at various levels exhibit such differentiable entities organized into modules. Many studies have established the existence of modularity in biological networks [57–59]. Modularity allows segregation of functions, and thus serves as an effective mechanism to insulate large complex systems from the effects of local damages, resulting in robustness [60,61].

Metabolic networks exhibit modularity with a large number of reactions organized into modules; the majority of reactions in a module are associated with the same biochemical pathway [62,63]. Further, the establishment of modules facilitates adaptation to new environments [62]. A number of studies have been carried out to understand the origin of such distinctive structures in metabolic networks. Studies have shown that variability in environments plays a major role in the formation of modules. For instance, the evolution of networks into modules was found to be more rapid in networks that are exposed to temporally varying or alternating environments [64,65]. Analysis of several bacterial species differing in their habitat and environmental

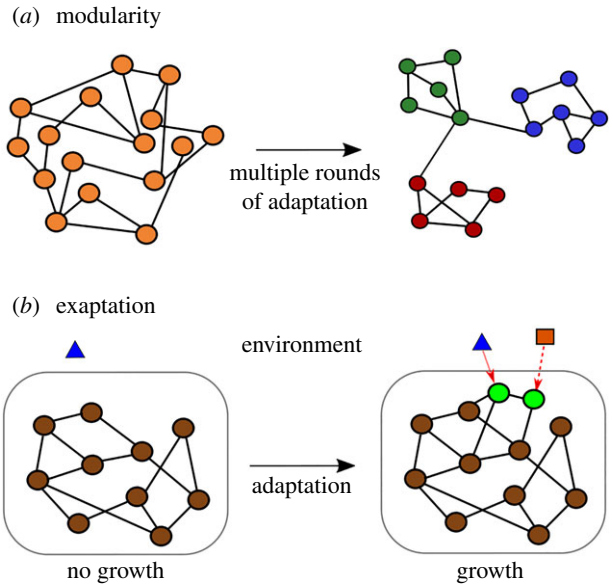


Figure 2. Modularity and exaptation. Schematic drawing of metabolic networks, demonstrating the concepts of modularity and exaptation. (*a*) Metabolic network forming modules with more connections within modules and fewer connections across modules. The original network is less modular (more inter-connectivity, than intra-connectivity), compared to the one on the right, where tightly connected modules (indicated by different colours) exist, with relatively few interconnections. (*b*) The metabolic network (where circles denote the reactions present) on the left cannot grow in an environment (blue triangle). Following adaptation, where it has acquired reactions (highlighted in green), the network exhibits growth in the environment. However, the highlighted reactions turn out to be an exaptation, when exposed to another environment (brown square), where the pre-existing reactions (highlighted green) enable growth in this new environment as well. In both panels, circles denote reactions and lines connect reactions that share a metabolite. (Online version in colour.)

variability could conclusively explain that more variable environments promote the evolution of highly modular metabolic networks, compared with constant environments [65]. The metabolic networks that are more versatile, i.e. that can grow in many different environments, have been shown to be highly modular [62]. Taken together, these studies highlight that adaptation plays a major role in the formation of such modular architectures [56,61].

Modularity is influenced by the relationship of horizontal gene transfer (HGT), habitats and environments [66]. Studies have suggested that HGT plays a major role in the emergence of modularity in metabolic networks [4,65,67]. A study on the relationship of the proportion of horizontally transferred genes and modularity for 94 bacteria revealed a strong positive correlation [65]. Other simulations showed that a combined selection pressure to maximize performance and simultaneously minimize the cost of network connectivity can enhance modularity even under constant environments [68].

As networks are subjected to varying environments, modules composed of reactions that operate in different environments emerge, thus enabling adaptation in new environments. Thus, evolution appears to shape metabolic networks to form segregated modules corresponding to specific defined functions. Modularity establishes isolation of functions and facilitates co-option, thus giving rise to evolutionary innovations [69].

(d) Exaptation

Pre-adaptations, or *exaptations*, are defined as traits that are initially naturally selected for a specific role, but later co-opted for a different purpose [70] (figure 2*b*). While adaptation involves the addition of new features that arise out of natural selection, exaptations involve co-option of an already existing feature for some other function. Exaptations thus arise *non-adaptively*, yet play an important role in the establishment of novel biological functions.

Metabolic networks selectively evolved on one carbon source were found to exapt on different carbon sources, illustrating that carbon source use can arise non-adaptively, and this paradigm serves as an important feature in the evolution of metabolic networks [71]. As illustrated in figure 2*b*, an organism adapts to an environment by acquiring certain reactions (green circles), which at a later stage become essential for survival in a different environment. Therefore, these reactions are said to have exapted for growth on the carbon source (figure 2*b*). A large number of metabolic networks evolved *in silico* to be viable on one particular carbon source were found to be viable on multiple other carbon sources that were not selected for. This illustrates the fact that adaptation to one carbon source imparts the network added advantage, namely growth on additional carbon sources. Metabolic networks that had a greater yield on complex carbon sources were found to be viable on a larger number of other carbon sources. Also, the chemical nature of the carbon sources was shown to influence the effect of metabolic network exaptation on a new source [72].

Complex innovations have also been shown to arise from stepwise additions of new reactions to a metabolic network [73]. The addition of even a single reaction was at times sufficient to pre-adapt and help form a complex trait later, in a new environment. A recent study on the history of bi-functionality for sugar isomerase HisA, a part of the histidine biosynthesis pathway [74], showed that the ancestral HisA, apart from histidine biosynthesis, can also catalyse the isomerization of a substrate usually catalysed by TrpF. This bi-functionality is estimated to have been conserved for two billion years without any selection pressure. Actinobacteria have lost the TrpF activity, yet possess the bi-functional HisA homologue, which performs the functions of both TrpF and HisA. This empirically exemplifies the evolution of novel functions through exaptation.

Exaptations are particularly interesting in metabolism because the addition of one or a few reactions to a metabolic network often enables survival on new carbon sources. Thus, exaptations can provide novel capabilities to metabolic networks and may play an important role in facilitating adaptation, highlighting evolutionary innovations.

3. Outlook

Metabolic networks are shaped through a variety of complex evolutionary processes, depending on the nature of the environments organisms are exposed to, and the consequent selection pressures. Despite the variety of mechanisms and selection pressures, all metabolic networks display distinctive features or design principles, which serve to increase the survivability of an organism, or its robustness. An understanding of such organizational principles can also point towards the evolutionary processes shaping various cellular networks.

Design principles such as functional redundancy and flux coupling illustrate fundamental constraints underlying metabolic networks. These constraints become important as one seeks to manipulate the flow of flux through various pathways towards applications in metabolic engineering or combinatorial targets for therapeutics [29,54,75,76]. Knowledge on the development of modules and exaptation in metabolic networks paves the way to obtain a better picture on the adaptation of organisms to new environments and thereby evolution.

Despite substantial advances made by the numerous studies we have discussed herein, the study of design principles in biological networks is still very nascent, with several open questions. For instance, while it is known that gene duplication leads to redundant functions, does it also facilitate flux coupling? Given that we have quite a good understanding of duplicated genes, can we predict functional redundancy and flux coupling? Further, are some design principles more likely to occur in a given evolutionary context, or under some specific evolutionary pressure?

While it is known that alternating selection pressures enhance the modularity of networks [77], it is harder to pin down the exact origins of some of the other design principles. Of the design principles we have enumerated, exaptation appears to be the most challenging to study, given that exaptations appear to arise serendipitously, repurposing existing machinery for a new function. A limited number of studies of exaptation have examined only a restrictive set of possible new functions, e.g. the ability to grow on a different carbon source. Identifying additional examples of exaptations will probably further our understanding on the emergence of possible exapted genes and reactions, which may help identify the patterns and nature of exaptation. Although we have chosen to restrict ourselves to metabolic networks in this review, interesting questions emerge, as we consider the large-scale organization of biological networks, where the metabolic networks are embedded, e.g. modules, that span across networks, *viz*. metabolic, regulatory and signalling networks.

Finally, it is important to note that the design principles outlined here are by no means exhaustive. Notwithstanding the number of studies discussed here, it is estimated that less than 35% of its entire metabolic network has been discovered, even for very well-studied organisms such as *E. coli* [78]. Given that we currently know but a small part of the universe of possible metabolic reactions (KEGG: http://www.kegg.jp/) [79], it is likely that other complex mechanisms and design principles may exist, especially in under-studied and novel organisms. Consequently, it will be exciting to study how metabolic networks from very different organisms behave: do archaeal metabolic networks share the design principles outlined? Alternatively, in the case of extremophiles, which may have fundamentally different metabolic network architectures, do evolutionary forces that determine these design principles act analogously in extreme environments? Another interesting dimension of research would be to explore the correlation between genome (or metabolic network) size or complexity and various design principles. Can smaller metabolic networks still demonstrate functional redundancy or carbon source exaptations, while more complex networks that have more constraints demonstrate different design principles? With an increasing number of genomes being sequenced, the studies of metabolic networks and their design principles will further unravel evolutionary adaptation and constraints in these networks, as organisms adapt to varying and divergent environments.

Data accessibility. This article has no additional data.

Authors' contributions. All authors contributed significantly to the writing and editing of the manuscript.

Competing interests. The authors declare to have no competing interests.

Funding. We received no funding for this study.

References

1. Wagner A. 2012 Metabolic networks and their evolution. In *Evolutionary Systems Biology/Advances in experimental medicine and biology* (ed. OS Soyer), vol. 751, pp. 29–52. Basingstoke, UK: Springer Nature.
2. Lee DH, Palsson BO. 2010 Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a nonnative carbon source, L-l,2-propanediol. *Appl. Environ. Microbiol.* **76**, 6327. (doi:10.1128/AEM.00373-10)
3. D'Souza G *et al.* 2016 Experimental evolution of metabolic dependency in bacteria. *PLoS Genet.* **12**, e1006364. (doi:10.1371/journal.pgen.1006364)
4. Pál C, Papp B, Lercher MJ. 2005 Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37**, 1372–1375. (doi:10.1038/ng1686)
5. Beales N. 2004 Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: a review. *Compr. Rev. Food Sci. Food Saf.* **3**, 1–20. (doi:10.1111/j.1541-4337.2004.tb00057.x)
6. Morgan-Kiss RM, Priscu JC, Pocock T, Gudynaite-Savitch L, Huner NPA. 2006 Adaptation and acclimation of photosynthetic microorganisms to permanently cold environments. *Microbiol. Mol. Biol. Rev.* **70**, 222–252. (doi:10.1128/MMBR.70.1.222-252.2006)
7. Morozkina EV, Slutskaia ES, Fedorova TV, Tugaĭ TI, Golubeva LI, Koroleva OV. 2010 Extremophilic microorganisms: biochemical adaptation and biotechnological application (review). *Prikl. Biokhim. Mikrobiol.* **46**, 5–20. (doi:10.1134/S0003683810010011)
8. Henrique RP. 2013 Extremophiles and extreme environments. *Life* **3**, 482–485. (doi:10.3390/life3030482)
9. Tolner B, Poolman B, Konings WN. 1997 Adaption of microorganisms and their transport systems to high temperatures. *Comp. Biochem. Physiol.* **118**, 423–428.
10. Cai M, Xun L. 2002 Organization and regulation of pentachlorophenol-degrading genes in *Sphingobium chlorophenolicum* ATCC 39723. *J. Bacteriol.* **184**, 4672–4680. (doi:10.1128/JB.184.17.4672-4680.2002)
11. Rehmann L, Daugulis AJ. 2008 Enhancement of PCB degradation by *Burkholderia xenovorans* LB400 in biphasic systems by manipulating culture conditions. *Biotechnol. Bioeng.* **99**, 521–528. (doi:10.1002/bit.21610)
12. Copley SD. 2010 NIH public access. *Dev. Biol.* **5**, 559–566. (doi:10.1038/nchembio.197.Evolution)
13. Jain M. 2013 Emerging role of metabolic pathways in abiotic stress tolerance. *J. Plant Biochem. Physiol.* **1**, 41–61. (doi:10.4172/2329-9029.1000108)
14. Parrou JL, Teste MA, François J. 1997 Effects of various types of stress on the metabolism of reserve carbohydrates in *Saccharomyces cerevisiae*: genetic evidence for a stress-induced recycling of glycogen and trehalose. *Microbiology* **143**, 1891–1900. (doi:10.1099/00221287-143-6-1891)
15. Rigato E, Fusco G. 2016 Enhancing effect of phenotype mutational robustness on adaptation in

Escherichia coli. *J. Exp. Zool. Part B Mol. Dev. Evol.* **326**, 31–37. (doi:10.1002/jez.b.22662)

16. McGuigan K, Sgrò CM. 2009 Evolutionary consequences of cryptic genetic variation. *Trends Ecol. Evol.* **24**, 305–311. (doi:10.1016/j.tree.2009.02.001)
17. Masel J. 2006 Cryptic genetic variation is enriched for potential adaptations. *Genetics* **172**, 1985–1991. (doi:10.1534/genetics.105.051649)
18. Gibson G, Dworkin I. 2004 Uncovering cryptic genetic variation. *Nat. Rev. Genet.* **5**, 681–690. (doi:10.1038/nrg1426)
19. Yang Y, Han L, Yuan Y, Li J, Hei N, Liang H. 2014 Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. *Nat. Commun.* **5**, 3231. (doi:10.1038/ncomms4231)
20. Rowland MA, Abdelzaher A, Ghosh P, Mayo ML. 2017 Crosstalk and the dynamical modularity of feed-forward loops in transcriptional regulatory networks. *Biophys. J.* **112**, 1539–1550. (doi:10.1016/j.bpj.2017.02.044)
21. Tran T-D, Kwon Y-K. 2013 The relationship between modularity and robustness in signalling networks. *J. R. Soc. Interface* **10**, 20130771. (doi:10.1098/rsif.2013.0771)
22. Gordley RM, Bugaj LJ, Lim WA. 2016 Modular engineering of cellular signaling proteins and networks. *Curr. Opin. Struct. Biol.* **39**, 106–114. (doi:10.1016/j.sbi.2016.06.012)
23. Typas A, Sourjik V. 2015 Bacterial protein networks: properties and functions. *Nat. Rev. Micro* **13**, 559–572. (doi:10.1038/nrmicro3508)
24. Butland G *et al.* 2008 eSGA: *E. coli* synthetic genetic array analysis. *Nat. Methods* **5**, 789–795. (doi:10.1038/nmeth.1239)
25. Tong AHY. 2001 Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368. (doi:10.1126/science.1065810)
26. Kalia NP *et al.* 2017 Exploiting the synthetic lethality between terminal respiratory oxidases to kill *Mycobacterium tuberculosis* and clear host infection. *Proc. Natl Acad. Sci. USA* **114**, 7426–7431. (doi:10.1073/pnas.1706139114)
27. Hugle M, Belz K, Fulda S. 2015 Identification of synthetic lethality of *PLK1* inhibition and microtubule-destabilizing drugs. *Cell Death Differ.* **22**, 1–11. (doi:10.1038/cdd.2015.59)
28. Pasquina L *et al.* 2016 A synthetic lethal approach for compound and target identification in *Staphylococcus aureus*. *Nat. Chem. Biol.* **12**, 40–45. (doi:10.1038/nchembio.1967)
29. McLornan DP, List A, Mufti GJ. 2014 Applying synthetic lethality for the selective targeting of cancer. *N. Engl. J. Med.* **371**, 1725–1735. (doi:10.1056/NEJMra1407390)
30. Santa MJP, Sadaka A, Moussa SH, Brown S, Zhang YJ, Rubin EJ, Gilmore MS, Walker S. 2014 Compound-gene interaction mapping reveals distinct roles for *Staphylococcus aureus* teichoic acids. *Proc. Natl Acad. Sci. USA* **111**, 12 510–12 515. (doi:10.1073/pnas.1404099111)
31. De Souza FSJ, Franchini LF, Rubinstein M. 2013 Exaptation of transposable elements into novel Cis-regulatory elements: is the evidence always strong? *Mol. Biol. Evol.* **30**, 1239–1251. (doi:10.1093/molbev/mst045)
32. Cui X, Cao X. 2014 Epigenetic regulation and functional exaptation of transposable elements in higher plants. *Curr. Opin. Plant Biol.* **21**, 83–88. (doi:10.1016/j.pbi.2014.07.001)
33. Singh DP *et al.* 2014 Genome-defence small RNAs exapted for epigenetic mating-type inheritance. *Nature* **509**, 447–452. (doi:10.1038/nature13318)
34. Christensen QH, Brecht RM, Dudekula D, Greenberg EP, Nagarajan R. 2014 Evolution of acyl-substrate recognition by a family of acyl-homoserine lactone synthases. *PLoS ONE* **9,** e0112464. (doi:10.1371/journal.pone.0112464)
35. Güell O, Sagués F, Serrano MÁ. 2014 Essential plasticity and redundancy of metabolism unveiled by synthetic lethality analysis. *PLoS Comput. Biol.* **10,** e1003637. (doi:10.1371/journal.pcbi.1003637)
36. Hartman JL, Garvik B, Hartwell L. 2001 Principles for the buffering of genetic variation. *Science* **291**, 1001–1004. (doi:10.1126/science.1056072)
37. Mani R, St. Onge RP, Hartman JL, Giaever G, Roth FP. 2008 Defining genetic interaction. *Proc. Natl Acad. Sci. USA* **105**, 3461–3466. (doi:10.1073/pnas.0712255105)
38. Marland E, Prachumwat A, Maltsev N, Gu Z, Li WH. 2004 Higher gene duplicabilities for metabolic proteins than for nonmetabolic proteins in yeast and *E. coli*. *J. Mol. Evol.* **59**, 806–814. (doi:10.1007/s00239-004-0068-x)
39. Vitkup D, Kharchenko P, Wagner A. 2006 Influence of metabolic network structure and function on enzyme evolution. *Genome Biol.* **7**, R39. (doi:10.1186/gb-2006-7-5-r39)
40. Kuepfer L, Sauer U, Blank LM. 2005 Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome Res.* **15**, 1421–1430. (doi:10.1101/gr.3992505)
41. Qian W, Liao BY, Chang AYF, Zhang J. 2010 Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet.* **26**, 425–430. (doi:10.1016/j.tig.2010.07.002)
42. Boone C, Bussey H, Andrews BJ. 2007 Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449. (doi:10.1038/nrg2085)
43. Dixon SJ, Andrews BJ, Boone C. 2009 Exploring the conservation of synthetic lethal genetic interaction networks. *Commun. Integr. Biol.* **2**, 78–81. (doi:10.4161/cib.7501)
44. Ghim CM, Goh K, Kahng B. 2005 Lethality and synthetic lethality in the genome-wide metabolic network of *Escherichia coli*. *J. Theor. Biol.* **237**, 401–411. (doi:10.1016/j.jtbi.2005.04.025)
45. Papp B, Teusink B, Notebaart RA. 2009 A critical view of metabolic network adaptations. *HFSP J.* **3**, 24–35. (doi:10.2976/1.3020599)
46. Court SJ, Waclaw B, Allen RJ. 2015 Lower glycolysis carries a higher flux than any biochemically possible alternative. *Nat. Commun.* **6**, 8427. (doi:10.1038/ncomms9427)
47. Harrison R, Papp B, Pál C, Oliver SG, Delneri D. 2007 Plasticity of genetic interactions in metabolic networks of yeast. *Proc. Natl Acad. Sci. USA* **104**, 2307–2312. (doi:10.1073/pnas.0607153104)
48. Sambamoorthy G, Raman K. 2018 Understanding the evolution of functional redundancy in metabolic networks. *Bioinformatics* **34**, i981–i987. (doi:10.1093/bioinformatics/bty604)
49. Burgard AP, Nikolaev EV, Schilling CH, Maranas CD. 2004 Flux coupling analysis of genome-scale metabolic network. *Genome Res.* **14**, 301–312. (doi:10.1101/gr.1926504.)
50. Larhlimi A, David L, Selbig J, Bockmayr A. 2012 F2C2: a fast tool for the computation of flux coupling in genome-scale metabolic networks. *BMC Bioinf.* **13**, 57. (doi:10.1186/1471-2105-13-57)
51. Hosseini Z, Marashi SA. 2015 Hierarchical organization of fluxes in *Escherichia coli* metabolic network: using flux coupling analysis for understanding the physiological properties of metabolic genes. *Gene* **561**, 199–208. (doi:10.1016/j.gene.2015.02.032)
52. David L, Marashi S-A, Larhlimi A, Mieth B, Bockmayr A. 2011 FFCA: a feasibility-based method for flux coupling analysis of metabolic networks. *BMC Bioinf.* **12**, 236. (doi:10.1186/1471-2105-12-236)
53. Notebaart RA, Teusink B, Siezen RJ, Papp B. 2008 Co-regulation of metabolic genes is better explained by flux coupling than by network distance. *PLoS Comput. Biol.* **4**, 157–163. (doi:10.1371/journal.pcbi.0040026)
54. Basler G, Nikoloski Z, Larhlimi A, Barabasi AL, Liu YY. 2016 Control of fluxes in metabolic networks. *Genome Res.* **26**, 956–968. (doi:10.1101/gr.202648.115)
55. Notebaart RA, Kensche PR, Huynen MA, Dutilh BE. 2009 Asymmetric relationships between proteins shape genome evolution. *Genome Biol.* **10**, R19. (doi:10.1186/gb-2009-10-2-r19)
56. Wagner GP, Pavlicev M, Cheverud JM. 2007 The road to modularity. *Nat. Rev. Genet.* **8**, 921–931. (doi:10.1038/nrg2267)
57. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999 From molecular to modular cell biology. *Nature* **402**, C47–C52. (doi:10.1038/35011540)
58. Wagner GP, Altenberg L. 1996 Perspective: complex adaptations and the evolution of evolvability Gunter. *Evolution* **50**, 967–976. (doi:10.1017/CBO9781107415324.004)
59. Kashtan N, Alon U. 2005 Spontaneous evolution of modularity and network motifs. *Proc. Natl Acad. Sci. USA* **102**, 13 773–13 778. (doi:10.1073/pnas.0503610102)
60. Hiroaki K. 2004 Biological robustness. *Nat. Rev. Genet.* **5**, 826–837. (doi:10.1007/978-3-7643-7567-6_10)
61. Hintze A, Adami C. 2008 Evolution of complex modular biological networks. *PLoS Comput. Biol.* **4**, e40023. (doi:10.1371/journal.pcbi.0040023)
62. Samal A *et al.* 2011 Environmental versatility promotes modularity in genome-scale metabolic

networks. *BMC Syst. Biol.* **5**, 135. (doi:10.1186/1752-0509-5-135)
63. Guimerà R, Nunes Amaral LA. 2005 Functional cartography of complex metabolic networks. *Nature* **433**, 895–900. (doi:10.1038/nature03288)
64. Parter M, Kashtan N, Alon U. 2007 Environmental variability and modularity of bacterial metabolic networks. *BMC Evol. Biol.* **7**, 169. (doi:10.1186/1471-2148-7-169)
65. Kreimer A, Borenstein E, Gophna U, Ruppin E. 2008 The evolution of modularity in bacterial metabolic networks. *Proc. Natl Acad. Sci. USA* **105**, 6976–6981. (doi:10.1073/pnas.0712149105)
66. Takemoto K. 2013 Does habitat variability really promote metabolic network modularity? *PLoS ONE* **8**, 2–10. (doi:10.1371/journal.pone.0061348)
67. Deem MW. 2013 Statistical mechanics of modularity and horizontal gene transfer. *Annu. Rev. Condens. Matter Phys.* **4**, 287–311. (doi:10.1146/annurev-conmatphys-030212-184316)
68. Clune J, Mouret J-B, Lipson H. 2013 The evolutionary origins of modularity. *Proc. R. Soc. B* **280**, 20122863. (doi:10.1098/rspb.2012.2863)
69. Espinosa-Soto C, Wagner A. 2010 Specialization can drive the evolution of modularity. *PLoS Comput. Biol.* **6**, e1000719. (doi:10.1371/journal.pcbi.1000719)
70. Gould SJ, Vrba ES. 1982 Exaptation: a missing term in the science of form. *Paleobiology* **8**, 4–15. (doi:10.1017/S0094837300004310)
71. Barve A, Wagner A. 2013 A latent capacity for evolutionary innovation through exaptation in metabolic systems. *Nature* **500**, 203–206. (doi:10.1038/nature12301)
72. Hosseini S-R, Wagner A. 2016 The potential for non-adaptive origins of evolutionary innovations in central carbon metabolism. *BMC Syst. Biol.* **10**, 97. (doi:10.1186/s12918-016-0343-7)
73. Szappanos B *et al.* 2016 Adaptive evolution of complex innovations through stepwise metabolic niche expansion. *Nat. Commun.* **7**, 11607. (doi:10.1038/ncomms11607)
74. Plach MG, Reisinger B, Sterner R, Merkl R. 2016 Long-term persistence of bi-functionality contributes to the robustness of microbial life through exaptation. *PLoS Genet.* **12**, e1005836. (doi:10.1371/journal.pgen.1005836)
75. Leung AWY, de Silva T, Bally MB, Lockwood WW. 2016 Synthetic lethality in lung cancer and translation to clinical therapies. *Mol. Cancer* **15**, 61. (doi:10.1186/s12943-016-0546-y)
76. Jerby-Arnon L *et al.* 2014 Predicting cancer-specific vulnerability via data-driven detection of synthetic lethality. *Cell* **158**, 1199–1209. (doi:10.1016/j.cell.2014.07.027)
77. Kashtan N, Noor E, Alon U. 2007 Varying environments can speed up evolution. *Proc. Natl Acad. Sci. USA* **104**, 13 711–13 716. (doi:10.1073/pnas.0611630104)
78. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BØ. 2011 A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. *Mol. Syst. Biol.* **7**, 535. (doi:10.1038/msb.2011.65)
79. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014 Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* **42**, 199–205. (doi:10.1093/nar/gkt1076)