Author for correspondence:
Justin Crocker
e-mail: justin.crocker@embl.de

A synthetic synthesis to explore animal evolution and development

Mindy Liu Perkins, Lautaro Gandara and Justin Crocker

Developmental Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany

(iD) MLP, 0000-0001-7839-5055; LG, 0000-0003-3695-6592; JC, 0000-0002-5113-0476

Identifying the general principles by which genotypes are converted into phenotypes remains a challenge in the post-genomic era. We still lack a predictive understanding of how genes shape interactions among cells and tissues in response to signalling and environmental cues, and hence how regulatory networks generate the phenotypic variation required for adaptive evolution. Here, we discuss how techniques borrowed from synthetic biology may facilitate a systematic exploration of evolvability across biological scales. Synthetic approaches permit controlled manipulation of both endogenous and fully engineered systems, providing a flexible platform for investigating causal mechanisms *in vivo*. Combining synthetic approaches with multi-level phenotyping (phenomics) will supply a detailed, quantitative characterization of how internal and external stimuli shape the morphology and behaviour of living organisms. We advocate integrating high-throughput experimental data with mathematical and computational techniques from a variety of disciplines in order to pursue a comprehensive theory of evolution.

This article is part of the theme issue 'Genetic basis of adaptation and speciation: from loci to causative mutations'.

> Life is infinitely stranger than anything which the mind of man could invent.
> – Sherlock Holmes, 'A Case of Identity' in *The Adventures of Sherlock Holmes* (1891)

1. Introduction

Over a century of advances in molecular techniques have gifted evolutionary biologists greater access to the source of heritable variation than Mendel or Darwin could have imagined (table 1). We are only now beginning to understand what processes and principles convert an enormous diversity of genotypes into an even greater diversity of phenotypic forms. In the search for general principles of animal evolution, researchers have found that evolution is driven both by mutations to proteins and mutations to regulatory regions. Although it has been proposed that mutating *cis*-regulatory regions may avoid the pleiotropic effects of mutating proteins [7], recent studies indicate that modifying transcriptional enhancers may also introduce extensive pleiotropy [8]. Indeed, evolutionary innovations in regulatory regions have been found to affect every level from morphology to physiology to behaviour [6] even when target proteins are functionally conserved [9–14]. These observations highlight that a comprehensive evolutionary theory necessitates functional understanding of how regulatory regions orchestrate the activities of cells, tissues, and ultimately whole organisms.

We face a number of fundamental challenges to elucidating phenotypic evolution through regulatory mutations. On a very basic level, it is still difficult to predict gene expression levels from an arbitrary promoter by an enhancer of known sequence. Recent progress in mathematical and mechanistic modelling has introduced general thermodynamic frameworks for the interactions of regulatory factors with enhancers [15], and reliable quantitative predictions have

Table 1. Defined terms.

synthetic biology	Creating biological systems in order to establish control over cellular behaviours (modified from Bashor & Collins [1]).
synthetic approach	Introducing an artificial or constructed element into a biological context; e.g. introducing mutations/duplications/indels, genetically engineered constructs (including circuits) (see also Garcia *et al.* [2]).
genotype-to-phenotype map	Sum of ways by which genotypic information influences the phenotype of an organism (adapted from Houle [3]).
phenomics	Acquisition of high-dimensional phenotypic data on an organism-wide scale. While genomic methods can aspire to survey genetic information comprehensively, the vast information content of phenotypes prevents their exhaustive characterization. Phenomics, instead, relies on prioritizing what to measure (adapted from Houle [3]).
enhancer	A contiguous DNA segment capable of boosting transcription from the promoter of a target gene, which could be located thousands of base pairs away. Enhancers can be found upstream or downstream to their target promoter and can even be located within transcriptional units.
regulatory region	DNA sequence that alters the expression of target genes. In this review, we will use the term primarily to refer to enhancers, promoters, silencers and insulators, which function through the binding of transcription factors or other regulatory molecules to DNA.
gene regulatory network	Set of transcription factors and signalling molecules that interact with each other and with DNA to regulate the expression of a set of genes (some of which may encode the transcription factors themselves).
complexity	Has many precise mathematical definitions in different contexts and fields, related in some way to the ease or difficulty of describing a given structure. For our purposes, the 'complexity' of a system scales with the number of potential behaviours that the system could demonstrate; e.g. for networks, 'complexity' roughly scales with size (number of elements and interactions between them).
modularity	The degree to which a network can be divided into independent subnetworks responsible for executing particular functions.
robustness	The degree to which a system is sensitive to perturbation or variation in architecture, environment, noise, parameters, etc. [4]. It can be precisely defined depending upon the performance measure of interest [5], though its usage in biological literature is not standardized, partly because it is rarely quantified.
evolvability	The capacity of a population to produce the heritable phenotypic variation of a kind that is not unconditionally deleterious (adapted from Masel & Trotter [6]).

been produced for specific cases in developing organisms (e.g. [16,17]). However, we have yet to generate models approaching the accuracy and generalizability of comparable theories for prokaryotic gene expression, which is itself still an active area of research beyond certain controlled contexts [18].

Additionally, we have only a limited empirical understanding of the possible paths for regulatory evolution. Most evidence derives from either observational measures of standing variation or biased perturbations of genomic regions that are already known to bind transcription factors strongly. This methodological skew reflects a general trend of relying more on genome-scale experimental data than would be typical for focused experimental studies [19]. Even when key regulatory regions are positively identified, their genetic variants may contribute very little to measurable phenotypic differences [20,21] or trigger cascading effects that result in pleiotropy [22–24]. Both of these phenomena emerge in part through complex interactions across organizational scales from cells to tissues to organs, each of which is subject to selection pressures [25]. Thus, cultivating a causal understanding of evolution will entail transferring insights from one level of biological organization to another with predictive power and, ideally, interpretability.

Here, we discuss gene regulation as a motivating example of how synthetic approaches can uncover evolutionary principles across biological scales. Many of these approaches are borrowed from synthetic biology, an interdisciplinary field that emphasizes rationally designing organisms to execute predefined functions, predicated on a practical understanding of the genetic and biophysical causes behind phenotypic features. Recent advances in synthetic technologies and theories have targeted throughput and control across multiple levels of biophysical and functional organization [1,26]. We outline how these developments are furnishing experimental evolution with the very toolkit it needs to chart a quantitative course from genotype to phenotype (figure 1).

2. Synthetic approaches to manipulating regulatory regions

Studies of transcriptional regulation often focus on developmental enhancers, which specify the precise time, level and location of gene expression during embryonic growth [27] and therefore act as major contributors to phenotypic evolution [27,28]. We now have the ability to map millions of putative enhancer regions across genomes [29], but defining which of these sites are functional remains a challenge. Inversely, we cannot yet reconstitute even the most well-characterized developmental enhancers from synthetic binding sites [30,31]. Deriving a causal relationship between regulatory sequences and phenotypes will require a better
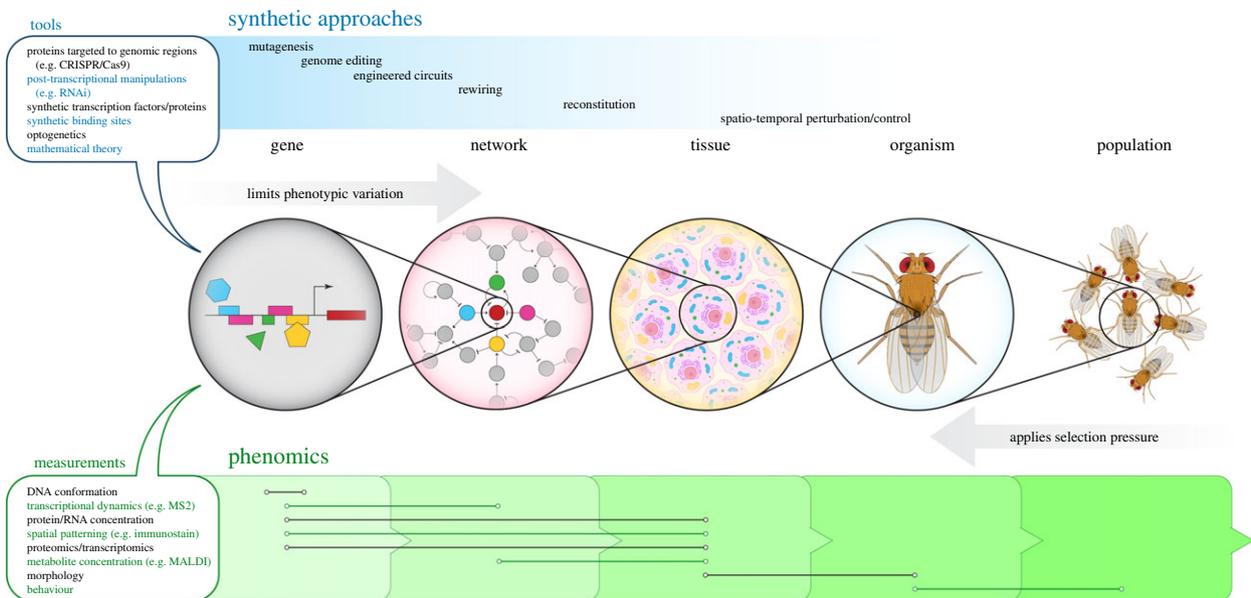
Figure 1. Synthetic approaches coupled with multi-level phenotypic measurements (phenomics) can characterize interactions across biological scales, which shape the possible and actual behaviours of living systems. Lower levels of organization tend to limit phenotypic variation at higher levels, while functional needs at higher levels may impose selection pressures on lower levels. Both processes shape evolvability.

understanding of the molecular mechanisms driving precise patterns of gene expression, as well as the consequent roles of cryptic variation and 'robustness' [32,33] on enhancer evolution. We propose that synthetic approaches, which involve introducing artificial elements into a biological context, enable experimental exploration of possibilities unsampled by natural evolutionary processes, while simultaneously granting the high levels of precision and control that are necessary to resolve causal factors across scales.

We start by exploring how synthetic approaches may clarify regulatory logic. Mutational scanning provides a powerful means to assay the activities of regulatory elements and to explore evolutionary potential [34–38]. These studies involve screening an unbiased assay of genetic variants and can therefore identify evolutionarily relevant sequences that are not surfaced by genomic or designed-sequence approaches based on previously identified binding sites [39]. In particular, mutational scanning experiments have uncovered non-canonical and low-affinity binding sites [39–41] that appear in appreciable frequencies in natural populations [42], as well as rare or deleterious mutations that are difficult to observe *in vivo* [43]. A recent study in *Drosophila melanogaster* used an automated robotics pipeline to survey an unbiased mutation library for a developmental enhancer, finding that almost all mutations altered gene expression and that the location and levels were highly interdependent [44]. The results are consistent with mammalian studies including both developmental enhancers [45] and promoters [46], indicating that widespread pleiotropic effects may constrain the evolvability of developmental enhancers [47,48]. In the future, mutational scanning could be used to define the essential features of enhancers and inform predictive models of enhancer function.

Most studies such as the one just described rely on reporter constructs to evidence phenotypic responses to genotypic changes. While reporter gene-based assays can provide information on the activity of individual regulatory regions, there are several potential caveats to generalizing the conclusions. First, removing an enhancer from its native location and placing

it in a novel chromatin environment may facilitate the binding of transcription factors to sites that are not used in their native context [49]. Second, many reporter constructs are based on heterologous promoters, which may augment gene expression relative to the endogenous promoter [50] or behave differently from the native locus if the distance (in bp) from promoter to enhancer is not retained. Third, what is identified as the enhancer (and therefore included in the reporter construct) may miss essential components regulating the target promoter, including binding sites adjacent to the native enhancer [51–53] or additional enhancers located elsewhere in the genome [54–58], which can buffer the phenotypic effects observed from mutations to the construct. Finally, in some cases, it is not even possible to isolate individual regulatory regions that drive reporter gene expression [59,60], emphasizing the need to explore beyond minimal, 'modular' enhancers.

Classical reporter-gene assays can be complemented by a number of methods and techniques from synthetic biology for interrogating native loci. Unlike screens performed in transgenes, assays that modify endogenous sequences can be used to directly test the physiological effects of different genetic variants. This capability is crucial given that changes in gene expression levels alone are not necessarily correlated with changes in evolutionary fitness [61]. New technological advances will facilitate building mutagenesis libraries in endogenous loci similar to the ones that rely on reporter-gene assays. Genome-editing tools, including the CRISPR-associated protein 9 (CRISPR-Cas9) system, have been used for high-throughput loss-of-function screens of regulatory elements [62–68]. CRISPR-Cas nucleases introduce insertion and deletion (indel) mutations of variable length, increasing the scope of genetic variations that can be tested relative to conventional mutagenesis techniques. CRISPR-Cas9 and other DNA-binding proteins such as TALE-effectors can also be fused to activator and repressor domains and then targeted to specific enhancers in order to modulate gene expression *in vivo* [69–73]. Activator and repressor domains may remodel chromatin structure across hundreds of nucleotides [71],

introducing larger perturbations than CRISPR-Cas nucleases and thereby improving enhancer detection.

Coupling synthetically perturbed enhancers with live imaging holds enormous potential to shed light on how enhancers establish precise expression patterns and confer phenotypic robustness. Recent technological advances have enabled real-time imaging of transcription by tagging mRNA with multiple repeats of a stem-loop sequence that is recognized by a binding protein fused with a fluorescent protein. The resulting 'transcription spots' are then visualized using standard live microscopy techniques [74], revealing the precise timing and spatial location of gene expression. For example, the *even-skipped stripe two* (eveS2) enhancer, one of the most well-characterized developmental enhancers [75], is active across a short window of time in the early embryo [76] and drives highly dynamic spatial patterns of expression [77]. Simultaneous imaging of transcription factors and regulatory targets [78] will be particularly instructive for probing the regulatory logic of individual loci as well as interactions among transcription factors.

We have presented here just a few of the available techniques for modifying regulatory regions and measuring their effects. Future efforts in technological development should focus on comprehensive approaches to interrogating regulatory elements in their native context. Analysis of the resulting data will be facilitated by our enhanced capacity to detect genetic variants with small effect sizes, as well as continued improvements to the sensitivity and throughput of live-imaging methods. With a combination of these approaches, we may begin to tease out the precise relationships between regulatory sequences, transcription factors and the expression patterns of target genes—the first step toward understanding the potential of genetic variants, realized or unrealized, to influence downstream phenotypes.

3. Form and function in gene regulatory networks

Synthetic approaches suitable for individual regulatory regions naturally extend to the study of gene regulatory networks (GRNs), or functional groupings of genes that produce phenotypes from genotypes by modulating expression levels in response to internal and external signals. Even small GRNs can execute involved processes such as disturbance rejection [79] and fold-change detection [80] over timescales much longer than the few seconds required for transcription initiation [16]. Although complex functions require complex networks, simple functions do not require simple networks. Rather, the same network structure may perform multiple functions depending on context or parameters, and the same function may be performed by multiple networks. These observations underlie nearly every aspect of network evolution, from the emergence of complexity through non-adaptive mechanisms [81,82] to the prevalence of 'network drift' within and between related species [83–86]. Classic challenges to the evolutionary study of GRNs involve their identification, level of conservation and ease of emergence or 'rewiring' based on mechanisms of binding site or enhancer turnover [12,13]. The synthetic biological emphasis on mathematical 'design principles' provides a fresh perspective from which to consider these challenges.

At the level of networks, understanding causal relationships between genotypes and phenotypes is complicated in a number of ways. If a regulated gene is itself a transcription factor, then downstream genes may constrain the minimum or maximum expression of the gene [87]. Compounded constraints for genes involved in multiple pathways may result in the emergence of non-adaptive features [88]. Even genes that are not directly interacting can become coupled through biophysical processes. For example, crosstalk, in which regulatory factors bind nonspecific or multiple targets, introduces a tradeoff between site specificity and the number of regulatory targets [89] as well as pressure to reduce active regulation [90], among other theoretical implications [91–93]. Given the high potential for coupling among GRNs, it is perhaps little surprise that researchers have proposed an 'omnigenic model' in which essentially all genes with regulatory variants contribute to complex traits [24].

Synthetic biologists have developed a toolkit sufficient for constructing entirely synthetic GRNs from engineered enhancers, promoters and transcription factors in bacteria and eukaryotes [94]. Considerable conceptual and experimental effort has aimed at understanding the reliability and robustness of these circuits, with a recent focus on predicting [95–99] and mitigating [100,101] the impact of crosstalk and resource sharing [1,102,103]. The tools and theories developed to design GRNs are grounded in approaches from engineering and physics, and hold enormous potential for helping evolutionary biologists to make sense of network-level changes over time.

For example, networks are subject to hard limits on their functionality that shape, directly or indirectly, the optima of the fitness landscape. Mathematical theory has elucidated fundamental limits to such functions as noise suppression [104], fold-change detection [105] and chemical sensing [106]; uncovered insurmountable tradeoffs among aspects of GRN performance [107–109] that may depend on system size [110]; and identified schemes that optimally achieve desired behaviours [111,112] under relevant constraints [113]. How closely networks operate to their theoretical optima—and how closely they can be brought to approach these optima through experimental evolution or synthetic recapitulation—may inform our understanding of selection pressures, adaptive responses and the balance between necessity and sufficiency.

Similarly, both evolutionary and synthetic biologists have investigated the degree to which GRNs are modular [114]—the former to determine whether decentralized control confers adaptive advantage [82,115–119], the latter to simplify the analysis and design of sophisticated circuitry [1,120], and both with the hope of improving human interpretability of complex structures. In principle, modules can be any size, which has historically biased research toward small, prevalent architectures or 'motifs' that are relatively simple to model, build and interpret [121–123]. Regardless of how a particular module is identified, its evolvability is strongly constrained by the pattern of regulatory interactions (topology) [124], especially when an original function must be preserved even as novel phenotypes arise [125–128]. Experiments with synthetic rewiring in bacteria have shown that topological differences can cause modules with the same average behaviour to diverge in their responses to stochastic fluctuations in chemical concentrations [129]. Moreover, the *strength* of regulatory interactions typically determines module behaviour, such that modules may vary

in their robustness to enhancer mutations or noisy gene expression depending upon their quantitative biophysical properties [5,128,130–132].

The degree to which modules are (i) separable from the context of the full GRN or whole cell and (ii) sensitive to variation in topology or biophysical parameters begs a host of evolutionary questions. Do less parameter-sensitive modules exhibit greater standing variation? Are parameter-sensitive modules more likely to be evolutionary co-opted? How do the relative sensitivities of overlapping modules determine evolvability [133–136]? To what extent does the parameter sensitivity of a module constrain the remainder of the network to maintain its own narrow operating range? Addressing these questions would help clarify when fitness benefits derive from 'internal' improvements in the function of a module or 'external' improvements in the behaviour with respect to other modules [137].

A straightforward way to proceed involves using the techniques discussed in the previous section to simultaneously disturb multiple regulatory regions belonging to the same GRN. It may be difficult, however, to infer causation from experiments conducted on complex and incompletely characterized endogenous systems. A complementary approach is to explore evolutionary principles in simple synthetic networks whose constituent parts can be individually measured and manipulated, and whose functions are understood *a priori*. For example, it has been shown that the set of regulatory interactions among three genes in a synthetic circuit biases the phenotypic variation observed after random mutation to the regulatory regions, even if the initial phenotypes are identical [138]. Similar techniques could be used to quantify dynamical behaviours and probe interconnections of modules, yielding invaluable predictive insights into the network-level effects of enhancer variation. The degree to which synthetic networks can be made orthogonal to endogenous pathways will present challenges to 'parallel evolution' experiments, but could also be deliberately employed to investigate the evolutionary implications of resource sharing, dosage change or crosstalk, which have been proposed to act as selective pressures against complexity [92].

Truly leveraging insights from simple synthetic modules will require theoretical and experimental work to predict the behaviour of full networks from constituent components under interconnection [139]. Such a programme necessitates 'parts characterization' of single loci or entire modules, in line with historical engineering practices and repeated calls for standardization across synthetic biology [1,26]. Most existing research has been carried out in microbes (e.g. [18]), but future efforts should also focus on multi-cellular sexually reproducing eukaryotes in particular, as to the best of our knowledge very little research has dealt with the implications of enhancer heterozygosity for network function [140], the interaction of population size with network evolution [115,141], or the role of 'multiple inheritances' in generating variation in network architecture and thus phenotype [137,142]. Whether disrupting endogenous networks or constructing synthetic ones, experiments linking *cis*-regulatory variation to the output of GRNs will require new techniques and tools capable of manipulating and monitoring multiple genes simultaneously. Such detailed study of regulatory regions in context will be essential for understanding biophysical and evolutionary constraints, and therefore for predicting phenotypic consequences on larger organizational scales.

4. Leveraging phenomics to explore animal evolution and development

Historically, most efforts to study adaptation have involved providing lists of genetic variants that influence the phenotype of interest. Large-scale approaches such as genome-wide association studies identify regions correlated with the observed phenotypic variation, but often fall short of functional validation [21,143]. This issue, commonly referred to as *the missing heritability problem* [143], is often assumed to reflect the wide distribution of genes and regulatory regions along the genome [24] that limits the explanatory potential of genomic approaches. The problem could also arise, however, from the failure of genomic approaches to consider contextual factors, such as environmental stimuli or feedback loops between phenotypic levels, that also influence organismal behaviour. Therefore, a complete genotype-to-phenotype map will require not only a comprehensive account of genetic diversity, but also (i) a high-dimensional, multi-scale depiction of the phenotypic space (phenomics), (ii) a standardized set of techniques borrowed from synthetic biology for targeted manipulation of the system of study and (iii) a systematic exploration of environmental conditions, closer to native ecological settings where food shortages or extreme temperatures are plausible scenarios.

Phenomic characterization is becoming more feasible as high-throughput phenotyping technologies become more available, ranging from transcriptomics and epigenomics to proteomics and metabolomics. Techniques that can link single-cell data to tissue architecture will be particularly useful, as bulk datasets tend to hide effects that are restricted to specific cell types or regions but that could deeply affect organism-level, fitness-related phenotypes such as mating behaviours or survivability. Recently, single-cell RNA-seq has been performed on relatively large samples from various species [144–146] including humans [147,148], revealing significant levels of cell-to-cell gene expression variability. New advances in robotics and automated microscopy techniques have also increased the feasibility of high-throughput screenings based on *in situ* hybridization or immunostaining methods [149]. Moving a step further from the analysis of gene expression patterns, technological developments now allow large-scale monitoring of more complex phenotypic features; for example, MALDI-imaging mass spectrometry can produce metabolomics data with high spatial resolution [150–153]. Even the behaviour of whole organisms can be analysed using automated video tracking of confined [154] or free-living [155] animals. Together, these high-throughput approaches permit assessing phenotypes in a standardized and replicable manner.

Though it is often assumed that causality runs from genotype to phenotype, phenomics characterization could reveal causal links between *phenotypic* layers (e.g. increased risk of lung cancer or type II diabetes due to tobacco addiction or obesity, respectively), thereby overriding the need for genetics to act as the preferential level of causal explanation [3]. New synthetic approaches are starting to allow the direct manipulation of specific phenotypic levels while exerting a minimal perturbation on the global biological system, turning the search for causality beyond genetics into a more practicable endeavour. Optogenetics, the experimental control of protein function mediated by light, enables modulating cellular pathways with high spatio-temporal resolution in a quantitative manner [156]. Thus, it is now possible to trigger cell

polarization [157], alter neuronal physiology in real time [158] or induce tissue morphogenesis [159] with precisely targeted techniques in intact organisms. These features include the possibility to persist in an out-of-equilibrium state, the engagement in cycles of growth and division, and, with particular relevance to the present discussion, the potential to evolve. Thus, these synthetic cellular systems could be used to explore cell-level mechanisms that might be acting as evolutionary constraints.

A comprehensive understanding of genotype-to-phenotype mapping will also necessitate a mechanistic description of how regulatory networks integrate the enormous range of environmental stimuli that shape phenotypic outputs. The fact that environmental inputs can modify developmental programmes and lead to altered phenotypes, a process called phenotypic plasticity, has been deeply studied by ecologists and geneticists [160,161]. Developmental biology, however, has historically paid less attention to this concept [162], focusing instead on the reverse phenomenon of canalization: how developmental programmes buffer environmental (and genetic) variation to reproducibly generate phenotypes. Hence, several mechanisms have been identified that allow developmental programmes to reject environmental perturbations [163], while explanations for how these programmes might instead *leverage* environmental cues are still scarce [164].

An underlying assumption of experimental biology is that the laboratory environment allows researchers to control external conditions and thus standardize the effect of the environment on the system under study. This approach, however, does not minimize the environmental dependency of the analysed phenomenon; it just provides a set of standard conditions that are used to compare different experiments from which more general rules and principles are then derived. Thus, many conclusions obtained from laboratory experiments do not hold when they are tested in natural environments (reviewed in [165]). If the biological processes under scrutiny are the mechanisms through which genotypes lead to phenotypes, which, as mentioned above, are known to be strongly dependent on environmental factors [160,161], restricting the external conditions to the laboratory environment seriously impairs our understanding of the phenomenon.

For these reasons, we advocate a deeper exploration of external conditions when studying genotype-to-phenotype mapping, beyond the optimal laboratory environment and closer to the actual stimuli that the studied organisms can find in their natural environments. Focusing on systems in their natural context with functional biology is gaining traction in the evolutionary community [166]. We propose that pairing such approaches with systematic manipulations imported from synthetic biology would provide new insights into evolutionary developmental biology. For example, unbiased characterization of genomic response to external signals would involve testing libraries of synthetic circuits under an extensive variety of tightly controlled environmental conditions that emulate realistic settings for the host organism. Far from being limited to controlled laboratory experiments, these studies could be carried out on multi-scale platforms such as Ecotrons [167,168].

Environmental effects on phenotype could also be systematically and selectively interrogated through synthetic interventions that trigger, disable or modify signalling pathways

naturally linked to selective stimuli, such as nutrient-sensitive pathways [169], oxygen-sensing mechanisms [170] or light-controlled circadian signalling [171]. This approach is not often coupled with extensive phenotypic quantification, as the target pathway is typically identified for its ability to modulate a particular behaviour of interest, rather than for 'pleiotropic', knock-on, or incremental effects across the whole organism. By combining such methods with phenomics, a deeper understanding of how specific external factors propagate across phenotypic levels could be obtained.

Phenomics approaches will produce huge datasets that vary significantly in content as well as the method and framework of acquisition. How to integrate information and extract meaningful conclusions from diverse, high-dimensional datasets may become one of the most urgent biological challenges in the near future (reviewed in [172,173]). Addressing this challenge will require (i) developing standard methods for acquiring and processing different types of data, which will enable future automation; (ii) stimulating data sharing and data reuse in order to avoid duplication of efforts; and (iii) developing next-generation and community-oriented data platforms to facilitate data accessibility and standardization of processing methods [174]. Only by overcoming these limitations will it be possible to draw biological conclusions from phenomic measurements.

Collectively, the technologies and methods discussed here would allow researchers to empirically test the principles of genotype-to-phenotype mapping through a wide variety of synthetic approaches coupled with phenomic characterizations and controlled exposure to a range of realistic selective stimuli. The proposed experimental pipeline could identify and modulate contributing factors across biological scales and phenotypic levels up through environmental interactions, thereby providing a framework for the causal understanding of adaptive responses. Furthermore, the ability to measure multiple phenotypic levels in individual organisms will permit correlating stochasticity or variability across organizational scales, illuminating an otherwise elusive means by which phenotypic variation may be generated [175–177]. Systematic applications of the suggested techniques could ultimately help us distinguish between contingent limitations of known regulatory mechanisms and real biophysical constraints, which will be essential to provide evolutionary biology with predictive power.

5. Uniting synthetic and evolutionary biology: metabolism as a case study

In the previous sections, we have outlined synthetic approaches to perturb individual regulatory regions and GRNs, as well as emerging technologies for manipulating and characterizing phenotypes at the level of cells, tissues or organisms. Here, we demonstrate the joint potential of these approaches using metabolism as a case study. Metabolic networks—which have been the subject of intense study by both evolutionary and synthetic biologists—are more complex than GRNs, as they are not only regulated by transcription factors, but also respond to changes in the concentrations of metabolites that are themselves part of the network. This additional regulatory layer is essential for quick responses to a constantly changing extracellular environment [178,179]. Thus, metabolic processes are

particularly suitable for examining genotype-to-phenotype mapping, as they couple regulatory networks that are hard-wired in the genome (i.e. the levels and spatial distribution of enzymes or metabolic regulators) with external factors that are essential for the survival of the organism, such as nutrient availability. Though complex, the topology and dynamics of metabolic pathways are well characterized at the molecular and mathematical levels [180], permitting computational modelling and experimental manipulation to a higher degree than most biological systems.

Explaining the mechanisms that could originate such intricate networks constitutes a major question for evolutionary biologists. Although metabolic engineering has been mostly driven by industrial motivations (specifically the high cost of chemical synthesis), the quest to optimize bacteria or yeast strains to produce commercially viable compounds has also provided insight into the biophysical constraints, and thus the evolutionary potential, of metabolic networks. Metabolic engineering ranges from single genetic alterations that enhance an endogenous metabolic route [181,182] to the construction of entirely synthetic circuits that lead to compounds not produced by the wild-type strain [183]. However, the complex intertwinement of metabolic networks often leads to unexpected or undesired outcomes, limiting the scope of genetic engineering efforts to manipulate metabolism [184]. Instead, yield optimization is achieved by the iterative selection of the desired phenotypes. The success of selective breeding is subject to evolutionary constraints affecting the yield of biosynthetic processes, including the level of genomic plasticity that a particular prokaryote 'species' can tolerate [185,186] and the robustness of metabolic flux [187–189]. Hence, metabolic engineering combined with selective breeding (directed evolution) appears to be a well-suited approach to study evolutionary mechanisms, as it provides an experimental framework to explore how altered genotypes affect fitness-related phenotypes, to identify bottlenecks and possible ways to overcome them, and to quantitatively address system-level features such as complexity and optimality.

A comprehensive understanding of metabolism may be even more relevant to studying evolution in multi-cellular than unicellular organisms, given that metabolic processes, in addition to managing the bioenergetic requirements of cells, can directly alter signalling pathways and developmental programmes (reviewed in [190,191]). For example, the intracellular concentration of specific metabolites has been shown to regulate key ontogenetic processes such as zygotic genome activation [192] or the development of the presomitic mesoderm in vertebrates [193,194]. Thus, environmentally modulated metabolic states could impact a wide range of fitness-related phenotypes through pathways at all levels from embryonic development to organismal behaviour.

In sum, as technological advances increase our throughput to measure metabolism [3,195], we anticipate that synthetic perturbations to metabolic pathways could be used to explore functional evolution across animal populations under realistic environmental conditions. The combined use of synthetic approaches and phenomics offers an unprecedented opportunity to investigate how metabolic networks integrate genetics and external inputs into coherent phenotypic outputs, shedding light on the basic principles and mechanisms that underlie adaptive responses to environmental challenges.

6. Synthesis at scale: prospectus in synthetic evolutionary biology

Deep comprehension of the mechanisms behind genotype-to-phenotype mapping is essential for predicting adaptive evolution. Charting the map will require a concerted effort to blend mathematical theory and computational modelling with large-scale experiments yielding high-dimensional quantitative data. Synthetic approaches are a particularly promising avenue for empirically testing ideas inferred from genomic and morphological studies of extinct and extant lineages. The ability to manipulate individual genetic sequences has already produced insights into the origins of pleiotropy at the regulatory and network levels. Advancements in high-throughput methods for synthetic perturbation and phenomic measurement promise to elucidate macroscale morphological and behavioural changes relevant to natural selection. The sizable resource investment required by this 'synthetic synthesis' will be justified to produce new, qualitatively diverse data sufficient to identify not only causal factors behind specific phenotypes but also general rules for how perturbations to genotype are propagated across phenotypes. In this way, we might finally begin to develop a detailed understanding of the multi-scale interactions and constraints that shape the evolvability of biological systems.

Despite recent successes mapping targeted genomic perturbations to phenotypic outcomes, we have yet to harness the full potential of experimental evolution combined with synthetic approaches. For practical reasons, most existing efforts have focused on microbial systems, including a study of almost 600 *E. coli* strains that found network linkages synthetically introduced into endogenous GRNs tend to enhance adaptability to selection pressures [196]. Yet many more ongoing studies in microbial evolution could still benefit from synthetic augmentation. For example, only two of 58 possible cross-feeding interactions in *E. coli* have been observed to evolve under experimental conditions, despite the fact that the metabolic rewiring required to generate these two interactions is theoretically no less complex than to generate the other outcomes [197]. The ability to synthetically rewire elements of the metabolic network for one or both strains offers a controlled way to compare the performance of various cross-feeding pairs and systematically characterize environmental constraints, as well as to investigate the extent to which evolutionary innovations in one lineage trigger speciation in cohabiting lineages.

Synthetic approaches appropriate for microbial studies provide direct inroads into controlled experiments involving multi-cellular organisms. Gut microbiota are one avenue of particular interest, as they are the subject of significant medical attention as well as a growing body of evolutionary work [198,199]. One study in *D. melanogaster* proposed that microbiota contribute directly to speciation in host organisms [200], but subsequent small-scale experiments have failed to consistently replicate the results or indicate a clear mechanistic connection [201]. The combination of libraries of bacteria harbouring synthetic metabolisms with simulated ecological conditions could clarify the real extent and ecological relevance of this phenomenon.

As technological and theoretical progress continues, it will be crucial to extend the application of synthetic approaches to experimental evolution in multi-cellular organisms. A straightforward first step involves using mutational screens in transgenes or endogenous loci to generate sexually

reproducing populations with varying degrees of genetic variation, in order to systematically quantitate how heterozygosity and genetic diversity influence evolutionary fitness. Given several starting populations homozygous at the allele(s) of interest, it should be possible to quantify the effects of heterozygosity in all possible pairwise couplings before moving to mixed populations of three or more lineages combined. This sequential, 'tiered' approach could be especially important for teasing out combinatorial effects, such as the recent observation that larval phenotypes in annelids result from the interplay of three genomes: maternal, paternal and zygotic [142]. Mixed populations of alleles could then be experimentally evolved under various selection pressures to link the degree of genetic diversity in a population to the dynamics of evolutionary adaptation.

Simultaneously, synthetic approaches should be leveraged to deepen our understanding of how GRNs evolve, which will be crucial for interrogating emergent features of complex multi-cellular organisms. For example, the flexible relationship between network structure and function affords greater potential for phenotypic variation than has been described in natural systems. While individual modules of fixed topology may be constrained in their evolvability, it appears that larger or more complex networks may be less constrained, due to the presence of redundant or 'buffering' mechanisms as well as the combinatorial increase in possible compensatory mechanisms. Future experiments with fully synthetic networks will permit more detailed research into the evolutionary potential of GRNs that interact to varying degrees with endogenous components at different times during the life cycle of an organism. Key to such studies is the ability to construct topologies unencumbered by evolutionary lineage and to sample a larger range of parameter space than could be expected to persist in a naturally evolving population. Eventually, the general principles derived from such experiments could be used to build circuits that directly influence fitness, in order to systematically investigate the evolvability of biological networks under controlled selection regimes.

Finally, multi-cellular organisms present unique challenges and opportunities for evolutionary study that should be both considered and capitalized upon. Notably, animals and plants possess a plethora of cell types and tissues that differ in behaviour almost exclusively due to differences in gene regulation [202]. Phenomic approaches in experimental evolution will be particularly useful to quantify changes in multiple tissues across many individuals in a population and therefore to trace the degree to which cells that share genetic material evolve independently. Here, too, synthetic reconstitution of cell-type-specific pathways might help to decouple intrinsic constraints on network function from constraints that arise internally due to shared usage of protein and regulatory apparatuses across different cell types within the same organism. In the future, more sophisticated synthetic manipulations to developmental programmes could shape the morphology of species and hence their environmental interactions, in order to explore macroscale selection pressures including physical or (eu)social considerations that may be irrelevant to microscopic species. Improvements to Ecotrons or other simplified ecosystems will further allow the methods proposed in this article to extend to community-level studies of several multi-cellular species evolving in tandem.

Though their motivations differ, evolutionary and synthetic biologists are both concerned with the mechanisms that decode genetic variation into phenotypic variation. Inspired by this convergence, we have illustrated here just a few of the ways in which synthetic approaches may shed light on outstanding problems in the evolution of genetic regulation by transcription factors. The framework we propose applies equally well to synthetic interventions modulating protein function [203–205], genome topology [206], cell-to-cell communication [207,208], post-transcriptional regulation [209,210] and neural behaviour [211,212], with increasing versatility as researchers continue to develop techniques for genetic engineering at scale. These studies will provide insight into basic evolutionary principles that will in turn enhance our ability to implement rational designs for living systems. Thus might we close the feedback loop between selected and synthesized, furnishing the variation of perspective from which human ingenuity has been, and is being, evolved.

royalsocietypublishing.org/journal/rstb Phil. Trans. R. Soc. B 377. 20200517

8

References

1. Bashor CJ, Collins JJ. 2018 Understanding biological regulation through synthetic biology. *Annu. Rev. Biophys.* **47**, 399–423. (doi:10.1146/annurev-biophys-070816-033903)

2. Garcia HG, Brewster RC, Phillips R. 2016 Using synthetic biology to make cells tomorrow's test tubes. *Integr. Biol.* **8**, 431–450. (doi:10.1039/c6ib00006a)

3. Houle D, Govindaraju DR, Omholt S. 2010 Phenomics: the next challenge. *Nat. Rev. Genet.* **11**, 855–866. (doi:10.1038/nrg2897)

4. Lander AD. 2007 Morpheus unbound: reimagining the morphogen gradient. *Cell* **128**, 245–256. (doi:10.1016/j.cell.2007.01.004)

5. Khammash M. 2016 An engineering viewpoint on biological robustness. *BMC Biol.* **14**, 22. (doi:10.1186/s12915-016-0241-x)

6. Masel J, Trotter MV. 2010 Robustness and evolvability. *Trends Genet.* **26**, 406–414. (doi:10.1016/j.tig.2010.06.002)

7. Wray GA. 2007 The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* **8**, 206–216. (doi:10.1038/nrg2063)

8. Sabarís G, Laiker I, Preger-Ben Noon E, Frankel N. 2019 Actors with multiple roles: pleiotropic enhancers and the paradigm of enhancer modularity. *Trends Genet.* **35**, 423–433. (doi:10.1016/j.tig.2019.03.006)

9. King MC, Wilson AC. 1975 Evolution at two levels in humans and chimpanzees. *Science*

188, 107–116. (doi:10.1126/science.1090005)

10. Prud'homme B, Gompel N, Carroll SB. 2007 Emerging principles of regulatory evolution. *Proc. Natl Acad. Sci. USA* **104**, 8605–8612. (doi:10.1073/pnas.0700488104)

11. Carroll SB. 2008 Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36. (doi:10.1016/j.cell.2008.06.030)

12. Halfon MS. 2017 Perspectives on gene regulatory network evolution. *Trends Genet.* **33**, 436–447. (doi:10.1016/j.tig.2017.04.005)

13. Rebeiz M, Patel NH, Hinman VF. 2015 Unraveling the tangled skein: the evolution of transcriptional regulatory networks in development. *Annu. Rev. Genomics Hum. Genet.* **16**, 103–131. (doi:10.1146/annurev-genom-091212-153423)

14. Stern DL, Orgogozo V. 2008 The loci of evolution: how predictable is genetic evolution? *Evolution* **62**, 2155–2177. (doi:10.1111/j.1558-5646.2008.00450.x)

15. He X, Samee MAH, Blatti C, Sinha S. 2010 Thermodynamics-based models of transcriptional regulation by enhancers: the roles of synergistic activation, cooperative binding and short-range repression. *PLoS Comput. Biol.* **6**, e1000935. (doi:10.1371/journal.pcbi.1000935)

16. Lammers NC, Kim YJ, Zhao J, Garcia HG. 2020 A matter of time: using dynamics and theory to uncover mechanisms of transcriptional bursting. *Curr. Opin. Cell Biol.* **67**, 147–157. (doi:10.1016/j.ceb.2020.08.001)

17. Park J *et al.* 2019 Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. *eLife* **8**, e41266. (doi:10.7554/eLife.41266)

18. Phillips R, Belliveau NM, Chure G, Garcia HG, Razo-Mejia M, Scholes C. 2019 Figure 1 theory meets figure 2 experiments in the study of gene expression. *Annu. Rev. Biophys.* **48**, 121–163. (doi:10.1146/annurev-biophys-052118-115525)

19. Halfon MS. 2019 Studying transcriptional enhancers: the founder fallacy, validation creep, and other biases. *Trends Genet.* **35**, 93–103. (doi:10.1016/j.tig.2018.11.004)

20. Frankel NS, Erezyilmaz DF, McGregor AP, Wang S, Payre F, Stern DL. 2011 Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* **474**, 598–603. (doi:10.1038/nature10200)

21. Rockman MV. 2012 The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**, 1–7. (doi:10.1111/j.1558-5646.2011.01486.x)

22. Paaby AB, Rockman MV. 2013 The many faces of pleiotropy. *Trends Genet.* **29**, 66–73. (doi:10.1016/j.tig.2012.10.010)

23. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. 2013 Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495. (doi:10.1038/nrg3461)

24. Boyle EA, Li YI, Pritchard JK. 2017 An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169**, 1177–1186. (doi:10.1016/j.cell.2017.05.038)

25. Davidson EH, Erwin DH. 2006 Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796–800. (doi:10.1126/science.1113832)

26. Currin A, Parker S, Robinson CJ, Takano E, Scrutton NS, Breitling R. 2021 The evolving art of creating genetic diversity: from directed evolution to synthetic biology. *Biotechnol. Adv.* **50**, 107762. (doi:10.1016/j.biotechadv.2021.107762)

27. Long HK, Prescott SL, Wysocka J. 2016 Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170–1187. (doi:10.1016/j.cell.2016.09.018)

28. Wittkopp PJ, Kalay G. 2012 *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69. (doi:10.1038/nrg3095)

29. Abascal F *et al.* 2020 Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710. (doi:10.1038/s41586-020-2493-4)

30. Vincent BJ, Estrada J, DePace AH. 2016 The appeasement of Doug: a synthetic approach to enhancer biology. *Integr. Biol.* **8**, 475–484. (doi:10.1039/c5ib00321k)

31. Crocker J, Ilsley GR. 2017 Using synthetic biology to study gene regulatory evolution. *Curr. Opin. Genet. Dev.* **47**, 91–101. (doi:10.1016/j.gde.2017.09.001)

32. Félix MA, Barkoulas M. 2015 Pervasive robustness in biological systems. *Nat. Rev. Genet.* **16**, 483–496. (doi:10.1038/nrg3949)

33. Payne JL, Wagner A. 2019 The causes of evolvability and their evolution. *Nat. Rev. Genet.* **20**, 24–38. (doi:10.1038/s41576-018-0069-z)

34. Weingarten-Gabbay S, Nir R, Lubliner S, Sharon E, Kalma Y, Weinberger A, Segal E. 2019 Systematic interrogation of human promoters. *Genome Res.* **29**, 171–183. (doi:10.1101/gr.236075.118)

35. Mogno I, Kwasnieski JC, Cohen BA. 2013 Massively parallel synthetic promoter assays reveal the in vivo effects of binding site variants. *Genome Res.* **23**, 1908–1915. (doi:10.1101/gr.157891.113)

36. Kircher M, Xiong C, Martin B, Schubach M, Inoue F, Bell RJA, Costello JF, Shendure J, Ahituv N. 2019 Saturation mutagenesis of twenty disease-associated regulatory elements at single base-pair resolution. *Nat. Commun.* **10**, 1–5. (doi:10.1038/s41467-019-11526-w)

37. Patwardhan RP *et al.* 2012 Massively parallel functional dissection of mammalian enhancers in vivo. *Nat. Biotechnol.* **30**, 265–270. (doi:10.1038/nbt.2136)

38. Patwardhan RP, Lee C, Litvin O, Young DL, Pe'Er D, Shendure J. 2009 High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat. Biotechnol.* **27**, 1173–1175. (doi:10.1038/nbt.1589)

39. de Boer CG, Vaishnav ED, Sadeh R, Abeyta EL, Friedman N, Regev A. 2020 Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. *Nat. Biotechnol.* **38**, 56–65. (doi:10.1038/s41587-019-0315-8)

40. Farley EK, Olson KM, Zhang W, Brandt AJ, Rokhsar DS, Levine MS. 2015 Suboptimization of developmental enhancers. *Science* **350**, 325–328. (doi:10.1126/science.aac6948)

41. Crocker J, Preger-Ben Noon E, Stern DL. 2016 The soft touch: low-affinity transcription factor binding sites in development and evolution. *Curr. Top. Dev. Biol.* **117**, 455–469. (doi:10.1016/bs.ctdb.2015.11.018)

42. Maurano MT *et al.* 2012 Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195. (doi:10.1126/science.1222794)

43. Visel A, Rubin EM, Pennacchio LA. 2009 Genomic views of distant-acting enhancers. *Nature* **461**, 199–205. (doi:10.1038/nature08451)

44. Fuqua T *et al.* 2020 Dense and pleiotropic regulatory information in a developmental enhancer. *Nature* **587**, 235–239. (doi:10.1038/s41586-020-2816-5)

45. Kvon EZ *et al.* 2020 Comprehensive in vivo interrogation reveals phenotypic impact of human enhancer variants. *Cell* **180**, 1262–1271. (doi:10.1016/j.cell.2020.02.031)

46. Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA. 2012 Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc. Natl Acad. Sci. USA* **109**, 19 498–19 503. (doi:10.1073/pnas.1210678109)

47. Uller T, Moczek AP, Watson RA, Brakefield PM, Laland KN. 2018 Developmental bias and evolution: a regulatory network perspective. *Genetics* **209**, 949–966. (doi:10.1534/genetics.118.300995)

48. Blomberg SP, Garland T. 2002 Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J. Evol. Biol.* **15**, 899–910. (doi:10.1046/j.1420-9101.2002.00472.x)

49. Dey SS, Foley JE, Limsirichai P, Schaffer DV, Arkin AP. 2015 Orthogonal control of expression mean and variance by epigenetic features at different genomic loci. *Mol. Syst. Biol.* **11**, 806. (doi:10.15252/msb.20145704)

50. Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, Stark A. 2015 Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556–559. (doi:10.1038/nature13994)

51. Kim AR, Martinez C, Ionides J, Ramos AF, Ludwig MZ, Ogawa N, Sharp DH, Reinitz J. 2013 Rearrangements of 2.5 kilobases of noncoding DNA from the *Drosophila even-skipped* locus define predictive rules of genomic *cis*-regulatory logic. *PLoS Genet.* **9**, e1003243. (doi:10.1371/journal.pgen.1003243)

52. Crocker J, Stern DL. 2017 Functional regulatory evolution outside of the minimal *even-skipped* stripe 2 enhancer. *Development* **144**, 3095–3101. (doi:10.1242/dev.149427)

53. Lopez-Rivera F, Foster Rhoades OK, Vincent BJ, Pym ECG, Bragdon MDJ, Estrada J, DePace AH, Wunderlich Z. 2020 A mutation in the *Drosophila melanogaster eve* stripe 2 minimal enhancer is buffered by flanking

sequences. *Genes Genomes Genet.* **10**, 4473–4482. (doi:10.1534/g3.120.401777)

54. Hong JW, Hendrix DA, Levine MS. 2008 Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314. (doi:10.1126/science.1160631)

55. Perry MW, Boettiger AN, Bothma JP, Levine M. 2010 Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20**, 1562–1567. (doi:10.1016/j.cub.2010.07.043)

56. Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. 2010 Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493. (doi:10.1038/nature09158)

57. Tsai A, Alves MRP, Crocker J. 2019 Multi-enhancer transcriptional hubs confer phenotypic robustness. *eLife* **8**, e45325. (doi:10.7554/eLife.45325)

58. Delker RK, Ranade V, Loker R, Voutev R, Mann RS. 2019 Low affinity binding sites in an activating CRM mediate negative autoregulation of the *Drosophila* Hox gene *Ultrabithorax*. *PLoS Genet.* **15**, e1008444. (doi:10.1371/journal.pgen.1008444)

59. Davis GK, Srinivasan DG, Wittkopp PJ, Stern DL. 2007 The function and regulation of *Ultrabithorax* in the legs of *Drosophila melanogaster*. *Dev. Biol.* **308**, 621–631. (doi:10.1016/j.ydbio.2007.06.002)

60. Klingler M, Soong J, Butler B, Gergen JP. 1996 Disperse versus compact elements for the regulation of runt stripes in *Drosophila*. *Dev. Biol.* **177**, 73–84. (doi:10.1006/dbio.1996.0146)

61. Duveau F, Toubiana W, Wittkopp PJ. 2017 Fitness effects of *cis*-regulatory variants in the *Saccharomyces cerevisiae* TDH3 promoter. *Mol. Biol. Evol.* **34**, 2908–2912. (doi:10.1093/molbev/msx224)

62. Canver MC *et al.* 2015 *BCL11A* enhancer dissection by Cas9-mediated *in situ* saturating mutagenesis. *Nature* **527**, 192–197. (doi:10.1038/nature15521)

63. Vierstra J *et al.* 2015 Functional footprinting of regulatory DNA. *Nat. Methods* **12**, 927–930. (doi:10.1038/nmeth.3554)

64. Fulco CP *et al.* 2016 Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**, 769–773. (doi:10.1126/science.aag2445)

65. Rajagopal N *et al.* 2016 High-throughput mapping of regulatory DNA. *Nat. Biotechnol.* **34**, 167–174. (doi:10.1038/nbt.3468)

66. Korkmaz G, Lopes R, Ugalde AP, Nevedomskaya E, Han R, Myacheva K, Zwart W, Elkon R, Agami R. 2016 Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.* **34**, 192–198. (doi:10.1038/nbt.3450)

67. Diao Y *et al.* 2016 A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. *Genome Res.* **26**, 397–405. (doi:10.1101/gr.197152.115)

68. Sanjana NE, Wright J, Zheng K, Shalem O, Fontanillas P, Joung J, Cheng C, Regev A, Zhang F. 2016 High-resolution interrogation of functional elements in the noncoding genome. *Science* **353**, 1545–1549. (doi:10.1126/science.aaf7613)

69. Crocker J, Stern DL. 2013 TALE-mediated modulation of transcriptional enhancers *in vivo*. *Nat. Methods* **10**, 762–767. (doi:10.1038/nmeth.2543)

70. Fulco CP *et al.* 2019 Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669. (doi:10.1038/s41588-019-0538-0)

71. Simeonov DR *et al.* 2017 Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature* **549**, 111–115. (doi:10.1038/nature23875)

72. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA. 2015 Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517. (doi:10.1038/nbt.3199)

73. Ewen-Campen B *et al.* 2017 Optimized strategy for in vivo Cas9-activation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **114**, 9409–9414. (doi:10.1073/pnas.1707635114)

74. Garcia HG, Tikhonov M, Lin A, Gregor T. 2013 Quantitative imaging of transcription in living *Drosophila* embryos links polymerase activity to patterning. *Curr. Biol.* **23**, 2140–2145. (doi:10.1016/j.cub.2013.08.054)

75. Small S, Blair A, Levine M. 1992 Regulation of even-skipped stripe 2 in the *Drosophila* embryo. *EMBO J.* **11**, 4047–4057. (doi:10.1002/j.1460-2075.1992.tb05498.x)

76. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M. 2014 Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos. *Proc. Natl Acad. Sci. USA* **111**, 10 598–10 603. (doi:10.1073/pnas.1410022111)

77. Briscoe J, Small S. 2015 Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996–4009. (doi:10.1242/dev.129452)

78. Bothma JP, Norstad MR, Alamos S, Garcia HG. 2018 LlamaTags: a versatile tool to image transcription factor dynamics in live embryos. *Cell* **173**, 1810–1822.e16. (doi:10.1016/j.cell.2018.03.069)

79. Shinar G, Feinberg M. 2010 Structural sources of robustness in biochemical reaction networks. *Science* **327**, 1389–1391. (doi:10.1126/science.1183372)

80. Goentoro L, Shoval O, Kirschner MW, Alon U. 2009 The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol. Cell* **36**, 894–899. (doi:10.1016/j.molcel.2009.11.018)

81. Soyer OS, Bonhoeffer S. 2006 Evolution of complexity in signaling pathways. *Proc. Natl Acad. Sci. USA* **103**, 16 337–16 342. (doi:10.1073/pnas.0604449103)

82. Lynch M. 2007 The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc. Natl Acad. Sci. USA* **104**(Suppl. 1), 8597–8604. (doi:10.1073/pnas.0702207104)

83. True JR, Haag ES. 2001 Developmental system drift and flexibility in evolutionary trajectories. *Evol. Dev.* **3**, 109–119. (doi:10.1046/j.1525-142x.2001.003002109.x)

84. Wotton KR, Jiménez-Guri E, Crombach A, Janssens H, Alcaine-Colet A, Lemke S, Schmidt-Ott U, Jaeger J. 2015 Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. *eLife* **4**, e04785. (doi:10.7554/eLife.04785)

85. Suryamohan K, Hanson C, Andrews E, Sinha S, Scheel MD, Halfon MS. 2016 Redeployment of a conserved gene regulatory network during *Aedes aegypti* development. *Dev. Biol.* **416**, 402–413. (doi:10.1016/j.ydbio.2016.06.031)

86. Ewe CK, Torres Cleuren YN, Rothman JH. 2020 Evolution and developmental system drift in the endoderm gene regulatory network of *Caenorhabditis* and other nematodes. *Front. Cell Dev. Biol.* **8**, 170. (doi:10.3389/fcell.2020.00170)

87. Chatterjee S, Kapoor A, Akiyama JA, Auer DR, Lee D, Gabriel S, Berrios C, Pennacchio LA, Chakravarti A. 2016 Enhancer variants synergistically drive dysfunction of a gene regulatory network in hirschsprung disease. *Cell* **167**, 355–368.e10. (doi:10.1016/j.cell.2016.09.005)

88. Dickinson WJ. 1988 On the architecture of regulatory systems: evolutionary insights and implications. *Bioessays* **8**, 204–208. (doi:10.1002/bies.950080608)

89. Sengupta AM, Djordjevic M, Shraiman BI. 2002 Specificity and robustness in transcription control networks. *Proc. Natl Acad. Sci. USA* **99**, 2072–2077. (doi:10.1073/pnas.022388499)

90. Grah R, Friedlander T. 2020 The relation between crosstalk and gene regulation form revisited. *PLoS Comput. Biol.* **16**, e1007642. (doi:10.1371/journal.pcbi.1007642)

91. Soltani M, Bokes P, Fox Z, Singh A. 2015 Nonspecific transcription factor binding can reduce noise in the expression of downstream proteins. *Phys. Biol.* **12**, 055002. (doi:10.1088/1478-3975/12/5/055002)

92. Friedlander T, Prizak R, Guet CC, Barton NH, Tkačik G. 2016 Intrinsic limits to gene regulation by global crosstalk. *Nat. Commun.* **7**, 12307. (doi:10.1038/ncomms12307)

93. Das D, Dey S, Brewster RC, Choubey S. 2017 Effect of transcription factor resource sharing on gene expression noise. *PLoS Comput. Biol.* **13**, e1005491. (doi:10.1371/journal.pcbi.1005491)

94. Engstrom MD, Pfleger BF. 2017 Transcription control engineering and applications in synthetic biology. *Synth. Syst. Biotechnol.* **2**, 176–191. (doi:10.1016/j.synbio.2017.09.003)

95. Gyorgy A, Del Vecchio D. 2014 Modular composition of gene transcription networks. *PLoS Comput. Biol.* **10**, e1003486. (doi:10.1371/journal.pcbi.1003486)

96. Landman J, Brewster RC, Weinert FM, Phillips R, Kegel WK. 2017 Self-consistent theory of transcriptional control in complex regulatory architectures. *PLoS ONE* **12**, e0179235. (doi:10.1371/journal.pone.0179235)

97. Wei L, Yuan Y, Hu T, Li S, Cheng T, Lei J, Xie Z, Zhang MQ, Wang X. 2019 Regulation by competition: a hidden layer of gene regulatory network. *Quant. Biol.* **7**, 110–121. (doi:10.1007/s40484-018-0162-5)

98. Weinert FM, Brewster RC, Rydenfelt M, Phillips R, Kegel WK. 2014 Scaling of gene expression with transcription-factor fugacity. *Phys. Rev. Lett.* **113**, 258101. (doi:10.1103/PhysRevLett.113.258101)

99. Qian Y, Del Vecchio D. 2021 Robustness of networked systems to unintended interactions with application to engineered genetic circuits. *IEEE Trans. Control Network Syst.* **8**, 1705–1716. (doi:10.1109/TCNS.2021.3078144)

100. Ceroni F *et al.* 2018 Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387–393. (doi:10.1038/nmeth.4635)

101. Frei T, Cella F, Tedeschi F, Gutiérrez J, Stan GB, Khammash M, Siciliano V. 2020 Characterization and mitigation of gene expression burden in mammalian cells. *Nat. Commun.* **11**, 4641. (doi:10.1038/s41467-020-18392-x)

102. Boo A, Ellis T, Stan GB. 2019 Host-aware synthetic biology. *Curr. Opin. Syst. Biol.* **14**, 66–72. (doi:10.1016/j.coisb.2019.03.001)

103. McBride CD, Grunberg TW, Del Vecchio D. 2021 Design of genetic circuits that are robust to resource competition. *Curr. Opin. Syst. Biol.* **28**, 100357. (doi:10.1016/j.coisb.2021.100357)

104. Lestas I, Vinnicombe G, Paulsson J. 2010 Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**, 174–178. (doi:10.1038/nature09333)

105. Skataric M, Nikolaev EV, Sontag ED. 2015 Fundamental limitation of the instantaneous approximation in fold-change detection models. *IET Syst. Biol.* **9**, 1–15. (doi:10.1049/iet-syb.2014.0006)

106. ten Wolde PR, Becker NB, Ouldridge TE, Mugler A. 2016 Fundamental limits to cellular sensing. *J. Stat. Phys.* **162**, 1395–1424. (doi:10.1007/s10955-015-1440-5)

107. Chandra FA, Buzi G, Doyle JC. 2011 Glycolytic oscillations and limits on robust efficiency. *Science* **333**, 187–192. (doi:10.1126/science.1200705)

108. Olsman N, Baetica AA, Xiao F, Leong YP, Murray RM, Doyle JC. 2019 Hard limits and performance tradeoffs in a class of antithetic integral feedback networks. *Cell Syst.* **9**, 49–63.e16. (doi:10.1016/j.cels.2019.06.001)

109. Siami M, Motee N, Buzi G, Bamieh B, Khammash M, Doyle JC. 2020 Fundamental limits and tradeoffs in autocatalytic pathways. *IEEE Trans. Automat. Contr.* **65**, 733–740. (doi:10.1109/TAC.2019.2921671)

110. Suderman R, Bachman JA, Smith A, Sorger PK, Deeds EJ. 2017 Fundamental trade-offs between information flow in single cells and cellular populations. *Proc. Natl Acad. Sci. USA* **114**, 5755–5760. (doi:10.1073/pnas.1615660114)

111. Tkačik G, Walczak AM. 2011 Information transmission in genetic regulatory networks: a review. *J. Phys. Condens. Matter* **23**, 153102. (doi:10.1088/0953-8984/23/15/153102)

112. Aoki SK, Lillacci G, Gupta A, Baumschlager A, Schweingruber D, Khammash M. 2019 A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* **570**, 533–537. (doi:10.1038/s41586-019-1321-1)

113. Malaguti G, Ten Wolde PR. 2021 Theory for the optimal detection of time-varying signals in cellular sensing systems. *eLife* **10**, e62574. (doi:10.7554/eLife.62574)

114. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999 From molecular to modular cell biology. *Nature* **402**, C47–C52. (doi:10.1038/35011540)

115. Force A, Cresko WA, Pickett FB, Proulx SR, Amemiya C, Lynch M. 2005 The origin of subfunctions and modular gene regulation. *Genetics* **170**, 433–446. (doi:10.1534/genetics.104.027607)

116. 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)

117. Hintze A, Adami C. 2008 Evolution of complex modular biological networks. *PLoS Comput. Biol.* **4**, e23. (doi:10.1371/journal.pcbi.0040023)

118. Davidson EH. 2010 Emerging properties of animal gene regulatory networks. *Nature* **468**, 911–920. (doi:10.1038/nature09645)

119. West-Eberhard MJ. 2019 Modularity as a universal emergent property of biological traits. *J. Exp. Zool. B* **332**, 356–364. (doi:10.1002/jez.b.22913)

120. Purnick PEM, Weiss R. 2009 The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* **10**, 410–422. (doi:10.1038/nrm2698)

121. Alon U. 2007 Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461. (doi:10.1038/nrg2102)

122. Konagurthu AS, Lesk AM. 2008 On the origin of distribution patterns of motifs in biological networks. *BMC Syst. Biol.* **2**, 73. (doi:10.1186/1752-0509-2-73)

123. Jiménez A, Cotterell J, Munteanu A, Sharpe J. 2017 A spectrum of modularity in multi-functional gene circuits. *Mol. Syst. Biol.* **13**, 925. (doi:10.15252/msb.20167347)

124. Matias Rodrigues JF, Wagner A. 2009 Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLoS Comput. Biol.* **5**, e1000613. (doi:10.1371/journal.pcbi.1000613)

125. MacCarthy T, Seymour R, Pomiankowski A. 2003 The evolutionary potential of the *Drosophila* sex determination gene network. *J. Theor. Biol.* **225**, 461–468. (doi:10.1016/S0022-5193(03)00282-0)

126. Torres-Sosa C, Huang S, Aldana M. 2012 Criticality is an emergent property of genetic networks that exhibit evolvability. *PLoS Comput. Biol.* **8**, e1002669. (doi:10.1371/journal.pcbi.1002669)

127. Jiménez A, Cotterell J, Munteanu A, Sharpe J. 2015 Dynamics of gene circuits shapes evolvability. *Proc. Natl Acad. Sci. USA* **112**, 2103–2108. (doi:10.1073/pnas.1411065112)

128. Payne JL, Wagner A. 2015 Function does not follow form in gene regulatory circuits. *Sci. Rep.* **5**, 13015. (doi:10.1038/srep13015)

129. Cağatay T, Turcotte M, Elowitz MB, Garcia-Ojalvo J, Süel GM. 2009 Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* **139**, 512–522. (doi:10.1016/j.cell.2009.07.046)

130. Ingram PJ, Stumpf MPH, Stark J. 2006 Network motifs: structure does not determine function. *BMC Genomics* **7**, 108. (doi:10.1186/1471-2164-7-108)

131. Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U. 2006 Plasticity of the *cis*-regulatory input function of a gene. *PLoS Biol.* **4**, e45. (doi:10.1371/journal.pbio.0040045)

132. Siegal ML, Promislow DEL, Bergman A. 2007 Functional and evolutionary inference in gene networks: does topology matter? *Genetica* **129**, 83–103. (doi:10.1007/s10709-006-0035-0)

133. Von Dassow G, Odell GM. 2002 Design and constraints of the *Drosophila* segment polarity module: robust spatial patterning emerges from intertwined cell state switches. *J. Exp. Zool.* **294**, 179–215. (doi:10.1002/jez.10144)

134. Ishihara S, Fujimoto K, Shibata T. 2005 Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. *Genes Cells* **10**, 1025–1038. (doi:10.1111/j.1365-2443.2005.00897.x)

135. 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)

136. Verd B, Monk NA, Jaeger J. 2019 Modularity, criticality, and evolvability of a developmental gene regulatory network. *eLife* **8**, e42832. (doi:10.7554/eLife.42832)

137. Papale F, Saget J, Bapteste É. 2020 Networks consolidate the core concepts of evolution by natural selection. *Trends Microbiol.* **28**, 254–265. (doi:10.1016/j.tim.2019.11.006)

138. Schaerli Y, Jiménez A, Duarte JM, Mihajlovic L, Renggli J, Isalan M, Sharpe J, Wagner A. 2018 Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. *Mol. Syst. Biol.* **14**, e8102. (doi:10.15252/msb.20178102)

139. Grunberg TW, Del Vecchio D. 2020 Modular analysis and design of biological circuits. *Curr. Opin. Biotechnol.* **63**, 41–47. (doi:10.1016/j.copbio.2019.11.015)

140. Dickinson WJ. 1979 Complex *cis*-acting regulatory genes demonstrated in *Drosophila* hybrids. *Dev. Genet.* **1**, 229–240. (doi:10.1002/dvg.1020010305)

141. Lynch M. 2007 The evolution of genetic networks by non-adaptive processes. *Nat. Rev. Genet.* **8**, 803–813. (doi:10.1038/nrg2192)

142. Zakas C, Rockman MV. 2021 Baby makes three: maternal, paternal, and zygotic genetic effects shape larval phenotypic evolution. *Evolution* **75**, 1607–1618. (doi:10.1111/evo.14244)

143. Pallares LF. 2019 Searching for solutions to the missing heritability problem. *eLife* **8**, e53018. (doi:10.7554/eLife.53018)

144. Achim K, Pettit JB, Saraiva LR, Gavriouchkina D, Larsson T, Arendt D, Marioni JC. 2015 High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat. Biotechnol.* **33**, 503–509. (doi:10.1038/nbt.3209)

145. Brunet Avalos C, Maier GL, Bruggmann R, Sprecher SG. 2019 Single cell transcriptome atlas of the *Drosophila* larval brain. *eLife* **8**, e50354. (doi:10.7554/eLife.50354)

146. Lorenzo R, Onizuka M, Defrance M, Laurent P. 2020 Combining single-cell RNA-sequencing with a

molecular atlas unveils new markers for *Caenorhabditis elegans* neuron classes. *Nucleic Acids Res.* **48**, 7119–7134. (doi:10.1093/nar/gkaa486)

147. Zhong S *et al*. 2018 A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528. (doi:10.1038/nature25980)

148. Yan L *et al*. 2013 Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139. (doi:10.1038/nsmb.2660)

149. Fuqua T, Jordan J, Halavatyi A, Tischer C, Richter K, Crocker J. 2021 An open-source semi-automated robotics pipeline for embryo immunohistochemistry. *Sci. Rep.* **11**, 10314. (doi:10.1038/s41598-021-89676-5)

150. Walch A, Rauser S, Deininger SO, Höfler H. 2008 MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochem. Cell Biol.* **130**, 421–434. (doi:10.1007/s00418-008-0469-9)

151. Neumann EK, Migas LG, Allen JL, Caprioli RM, Van de Plas R, Spraggins JM. 2020 Spatial metabolomics of the human kidney using MALDI trapped ion mobility imaging mass spectrometry. *Anal. Chem.* **92**, 13 084–13 091. (doi:10.1021/acs.analchem.0c02051)

152. Dilillo M *et al*. 2017 Ultra-high mass resolution MALDI imaging mass spectrometry of proteins and metabolites in a mouse model of glioblastoma. *Sci. Rep.* **7**, 603. (doi:10.1038/s41598-017-00703-w)

153. Rudd D, Ronci M, Johnston MR, Guinan T, Voelcker NH, Benkendorff K. 2015 Mass spectrometry imaging reveals new biological roles for choline esters and Tyrian purple precursors in muricid molluscs. *Sci. Rep.* **5**, 13408. (doi:10.1038/srep13408)

154. Berman GJ, Choi DM, Bialek W, Shaevitz JW. 2014 Mapping the stereotyped behaviour of freely moving fruit flies. *J. R. Soc. Interface* **11**, 20140672. (doi:10.1098/rsif.2014.0672)

155. Rodríguez-Muñoz R, Bretman A, Slate J, Walling CA, Tregenza T. 2010 Natural and sexual selection in a wild insect population. *Science* **328**, 1269–1272. (doi:10.1126/science.1188102)

156. Kwon E, Heo WD. 2020 Optogenetic tools for dissecting complex intracellular signaling pathways. *Biochem. Biophys. Res. Commun.* **527**, 331–336. (doi:10.1016/j.bbrc.2019.12.132)

157. Wang X, He L, Wu YI, Hahn KM, Montell DJ. 2010 Light-mediated activation reveals a key role for Rac in collective guidance of cell movement *in vivo*. *Nat. Cell Biol.* **12**, 591–597. (doi:10.1038/ncb2061)

158. Carter ME, de Lecea L. 2011 Optogenetic investigation of neural circuits *in vivo*. *Trends Mol. Med.* **17**, 197–206. (doi:10.1016/j.molmed.2010.12.005)

159. Izquierdo E, Quinkler T, De Renzis S. 2018 Guided morphogenesis through optogenetic activation of Rho signalling during early *Drosophila* embryogenesis. *Nat. Commun.* **9**, 2366. (doi:10.1038/s41467-018-04754-z)

160. Favé MJ *et al*. 2018 Gene-by-environment interactions in urban populations modulate risk

phenotypes. *Nat. Commun.* **9**, 827. (doi:10.1038/s41467-018-03202-2)

161. Li X, Guo T, Mu Q, Li X, Yu J. 2018 Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proc. Natl Acad. Sci. USA* **115**, 6679–6684. (doi:10.1073/pnas.1718326115)

162. Erwin DH. 2017 Developmental push or environmental pull? The causes of macroevolutionary dynamics. *Hist. Phil. Life Sci.* **39**, 36. (doi:10.1007/s40656-017-0163-0)

163. Gursky VV, Surkova SY, Samsonova MG. 2012 Mechanisms of developmental robustness. *Biosystems* **109**, 329–335. (doi:10.1016/j.biosystems.2012.05.013)

164. Gilbert SF, Bosch TCG, Ledón-Rettig C. 2015 Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat. Rev. Genet.* **16**, 611–622. (doi:10.1038/nrg3982)

165. Calisi RM, Bentley GE. 2009 Lab and field experiments: are they the same animal? *Horm. Behav.* **56**, 1–10. (doi:10.1016/j.yhbeh.2009.02.010)

166. Bergelson J, Kreitman M, Petrov DA, Sanchez A, Tikhonov M. 2021 Functional biology in its natural context: a search for emergent simplicity. *eLife* **10**, e67646. (doi:10.7554/eLife.67646)

167. Granjou C, Walker J. 2016 Promises that matter: reconfiguring ecology in the ecotrons. *Sci. Technol. Stud.* **29**, 49–67. (doi:10.23987/sts.58844)

168. Sauze J *et al*. 2020 The European Ecotron of Montpellier: experimental platforms to study ecosystem response to climate change. In *EGU General Assembly Conf., Online, 4–8 May 2020*, pp. 8650. (doi:10.5194/egusphere-egu2020-8650)

169. González A, Hall MN, Lin SC, Hardie DG. 2020 AMPK and TOR: the Yin and Yang of cellular nutrient sensing and growth control. *Cell Metab.* **31**, 472–492. (doi:10.1016/j.cmet.2020.01.015)

170. Wilson JW, Shakir D, Batie M, Frost M, Rocha S. 2020 Oxygen-sensing mechanisms in cells. *FEBS J.* **287**, 3888–3906. (doi:10.1111/febs.15374)

171. Ko GYP. 2020 Circadian regulation in the retina: from molecules to network. *Eur. J. Neurosci.* **51**, 194–216. (doi:10.1111/ejn.14185)

172. Lapatas V, Stefanidakis M, Jimenez RC, Via A, Schneider MV. 2015 Data integration in biological research: an overview. *J. Biol. Res.* **22**, 9. (doi:10.1186/s40709-015-0032-5)

173. Thessen AE, Bogdan P, Patterson DJ, Casey TM, Hinojo-Hinojo C, de Lange O, Haendel MA. 2021 From reductionism to reintegration: solving society's most pressing problems requires building bridges between data types across the life sciences. *PLoS Biol.* **19**, e3001129. (doi:10.1371/journal.pbio.3001129)

174. Alexandrov T, Ovchinnikova K, Palmer A, Kovalev V. 2019 METASPACE: a community-populated knowledge base of spatial metabolomes in health and disease. *BioRxiv*. (doi:10.1101/539478)

175. Giri R, Papadopoulos DK, Posadas DM, Potluri HK, Tomancak P, Mani M, Carthew RW. 2020 Ordered patterning of the sensory system is susceptible to

stochastic features of gene expression. *eLife* **9**, e53638. (doi:10.7554/eLife.53638)

176. Sosnik J, Zheng L, Rackauckas CV, Digman M, Gratton E, Nie Q, Schilling TF. 2016 Noise modulation in retinoic acid signaling sharpens segmental boundaries of gene expression in the embryonic zebrafish hindbrain. *eLife* **5**, e14034. (doi:10.7554/eLife.14034)

177. Kiskowski M, Glimm T, Moreno N, Gamble T, Chiari Y. 2019 Isolating and quantifying the role of developmental noise in generating phenotypic variation. *PLoS Comput. Biol.* **15**, e1006943. (doi:10.1371/journal.pcbi.1006943)

178. Liu D, Zhang F. 2018 Metabolic feedback circuits provide rapid control of metabolite dynamics. *ACS Synth. Biol.* **7**, 347–356. (doi:10.1021/acssynbio.7b00342)

179. Voet D, Voet JG. 2011 *Biochemistry*, 4th edn, pp. 1088–1106. Hoboken, NJ: John Wiley & Sons, Inc.

180. 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)

181. Vrljic M, Sahm H, Eggeling L. 1996 A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. *Mol. Microbiol.* **22**, 815–826. (doi:10.1046/j.1365-2958.1996.01527.x)

182. Morbach S, Sahm H, Eggeling L. 1996 l-Isoleucine production with *Corynebacterium glutamicum*: further flux increase and limitation of export. *Appl. Environ. Microbiol.* **62**, 4345–4351. (doi:10.1128/aem.62.12.4345-4351.1996)

183. Szczebara FM *et al*. 2003 Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat. Biotechnol.* **21**, 143–149. (doi:10.1038/nbt775)

184. Petri R, Schmidt-Dannert C. 2004 Dealing with complexity: evolutionary engineering and genome shuffling. *Curr. Opin. Biotechnol.* **15**, 298–304. (doi:10.1016/j.copbio.2004.05.005)

185. Nakamura Y, Nishio Y, Ikeo K, Gojobori T. 2003 The genome stability in *Corynebacterium* species due to lack of the recombinational repair system. *Gene* **317**, 149–155. (doi:10.1016/S0378-1119(03)00653-X)

186. Aras RA, Kang J, Tschumi AI, Harasaki Y, Blaser MJ. 2003 Extensive repetitive DNA facilitates prokaryotic genome plasticity. *Proc. Natl Acad. Sci. USA* **100**, 13 579–13 584. (doi:10.1073/pnas.1735481100)

187. Stephanopoulos G, Vallino JJ. 1991 Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252**, 1675–1681. (doi:10.1126/science.1904627)

188. Almaas E, Kovács B, Vicsek T, Oltvai ZN, Barabási AL. 2004 Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* **427**, 839–843. (doi:10.1038/nature02289)

189. Aslan S, Noor E, Bar-Even A. 2017 Holistic bioengineering: rewiring central metabolism for enhanced bioproduction. *Biochem. J.* **474**, 3935–3950. (doi:10.1042/BCJ20170377)

190. Gándara L, Wappner P. 2018 Metabo-Devo: a metabolic perspective of development. *Mech. Dev.* **154**, 12–23. (doi:10.1016/j.mod.2018.02.004)

191. Miyazawa H, Aulehla A. 2018 Revisiting the role of metabolism during development. *Development* **145**, dev131110. (doi:10.1242/dev.131110)

192. Nagaraj R, Sharpley MS, Chi F, Braas D, Zhou Y, Kim R, Clark AT, Banerjee U. 2017 Nuclear localization of mitochondrial TCA cycle enzymes as a critical step in mammalian zygotic genome activation. *Cell* **168**, 210–223.e11. (doi:10.1016/j.cell.2016.12.026)

193. Bulusu V *et al.* 2017 Spatiotemporal analysis of a glycolytic activity gradient linked to mouse embryo mesoderm development. *Dev. Cell* **40**, 331–341.e4. (doi:10.1016/j.devcel.2017.01.015)

194. Oginuma M, Moncuquet P, Xiong F, Karoly E, Chal J, Guevorkian K, Pourquié O. 2017 A gradient of glycolytic activity coordinates FGF and Wnt signaling during elongation of the body axis in amniote embryos. *Dev. Cell* **40**, 342–353.e10. (doi:10.1016/j.devcel.2017.02.001)

195. de Raad M, Fischer CR, Northen TR. 2016 High-throughput platforms for metabolomics. *Curr. Opin. Chem. Biol.* **30**, 7–13. (doi:10.1016/j.cbpa.2015.10.012)

196. Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, Garriga-Canut M, Serrano L. 2008 Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452**, 840–845. (doi:10.1038/nature06847)

197. San Roman M, Wagner A. 2020 Acetate and glycerol are not uniquely suited for the evolution of cross-feeding in *E. coli*. *PLoS Comput. Biol.* **16**, e1008433. (doi:10.1371/journal.pcbi.1008433)

198. Moeller AH, Li Y, Mpoudi Ngole E, Ahuka-Mundeke S, Lonsdorf EV, Pusey AE, Peeters M, Hahn BH, Ochman H. 2014 Rapid changes in the gut microbiome during human evolution. *Proc. Natl Acad. Sci. USA* **111**, 16 431–16 435. (doi:10.1073/pnas.1419136111)

199. Garud NR, Good BH, Hallatschek O, Pollard KS. 2019 Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLoS Biol.* **17**, e3000102. (doi:10.1371/journal.pbio.3000102)

200. Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E. 2010 Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **107**, 20 051–20 056. (doi:10.1073/pnas.1009906107)

201. Obadia B, Keebaugh ES, Yamada R, Ludington WB, Ja WW. 2018 Diet influences host-microbiota associations in *Drosophila*. *Proc. Natl Acad. Sci. USA* **115**, E4547–E4548. (doi:10.1073/pnas.1804948115)

202. Arendt D *et al.* 2016 The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757. (doi:10.1038/nrg.2016.127)

203. Chowdhury R, Maranas CD. 2020 From directed evolution to computational enzyme engineering—a review. *AIChE J.* **66**, e16847. (doi:10.1002/aic.16847)

204. Yang KK, Wu Z, Arnold FH. 2019 Machine-learning-guided directed evolution for protein engineering. *Nat. Methods* **16**, 687–694. (doi:10.1038/s41592-019-0496-6)

205. Li JW, Zhang XY, Wu H, Bai YP. 2020 Transcription factor engineering for high-throughput strain evolution and organic acid bioproduction: a review. *Front. Bioeng. Biotechnol.* **8**, 98. (doi:10.3389/fbioe.2020.00098)

206. Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M, Elowitz MB. 2016 Dynamics of epigenetic regulation at the single-cell level. *Science* **351**, 720–724. (doi:10.1126/science.aab2956)

207. Toda S, Frankel NW, Lim WA. 2019 Engineering cell–cell communication networks: programming multicellular behaviors. *Curr. Opin. Chem. Biol.* **52**, 31–38. (doi:10.1016/j.cbpa.2019.04.020)

208. Perkins ML, Benzinger D, Arcak M, Khammash M. 2020 Cell-in-the-loop pattern formation with optogenetically emulated cell-to-cell signaling. *Nat. Commun.* **11**, 1355. (doi:10.1038/s41467-020-15166-3)

209. Makeyev EV, Maniatis T. 2008 Multilevel regulation of gene expression by microRNAs. *Science* **319**, 1789–1790. (doi:10.1126/science.1152326)

210. Bartoli V, Meaker GA, di Bernardo M, Gorochowski TE. 2020 Tunable genetic devices through simultaneous control of transcription and translation. *Nat. Commun.* **11**, 2095. (doi:10.1038/s41467-020-15653-7)

211. Kim CK, Adhikari A, Deisseroth K. 2017 Integration of optogenetics with complementary methodologies in systems neuroscience. *Nat. Rev. Neurosci.* **18**, 222–235. (doi:10.1038/nrn.2017.15)

212. Rabinowitch I. 2020 What would a synthetic connectome look like? *Phys. Life Rev.* **33**, 1–15. (doi:10.1016/j.plrev.2019.06.005)

13

royalsocietypublishing.org/journal/rstb *Phil. Trans. R. Soc. B* **377**: 20200517