Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Using synthetic biology to study gene regulatory evolution

Justin Crocker[1] and Garth R Ilsley[2]

Transcriptional enhancers specify the precise time, level, and location of gene expression. Disentangling and characterizing the components of enhancer activity in multicellular eukaryotic development has proven challenging because enhancers contain activator and repressor binding sites for multiple factors that each exert nuanced, context-dependent control of enhancer activity. Recent advances in synthetic biology provide an almost unlimited ability to create and modify regulatory elements and networks, offering unprecedented power to study gene regulation. Here we review several studies demonstrating the utility of synthetic biology for studying enhancer function during development and evolution. These studies clearly show that synthetic biology can provide a way to reverse-engineer and reengineer transcriptional regulation in animal genomes with enormous potential for understanding evolution.

Addresses
[1] European Molecular Biology Laboratory, D-69117 Heidelberg, Germany
[2] Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan

Corresponding author: Crocker, Justin (justin.crocker@embl.de)

Current Opinion in Genetics & Development 2017, **47**:91–101

This review comes from a themed issue on **Evolutionary genetics**

Edited by **Eric S Haag** and **David L Stern**

For a complete overview see the Issue and the Editorial

Available online 29th September 2017

http://dx.doi.org/10.1016/j.gde.2017.09.001

What determines patterns of gene expression in space and time?

Morphological evolution often involves many mutations of small effect at the single nucleotide substitution level — not just one mutation of large effect [1] — consistent with Charles Darwin's favored view of a gradual evolutionary process. However, assaying changes of small effect is a challenge. In the words of the population geneticist John Gillespie regarding protein evolution: 'Perhaps we are investigating phenomena that are below the resolving power of our current techniques: phenomena large enough to dominate genetic drift, yet refractory to direct experimental investigations' [2]. This observation is no less relevant when applied to the regulatory regions of the genome, the transcriptional enhancers, that specify the precise time, level, and location of gene expression. Enhancers contain activator and repressor binding sites for multiple transcription factors and each transcription factor exerts a nuanced, context-dependent effect on enhancer activity [3–8]. Recent findings have shown that the output of the enhancer — the expression of the associated gene — can be adjusted gradually by modifying the transcription factors or their binding sites [5••,9••,10] and that individual transcription factor binding sites are often low-affinity [4,11,12••], with limited individual effect on gene expression when altered [4].

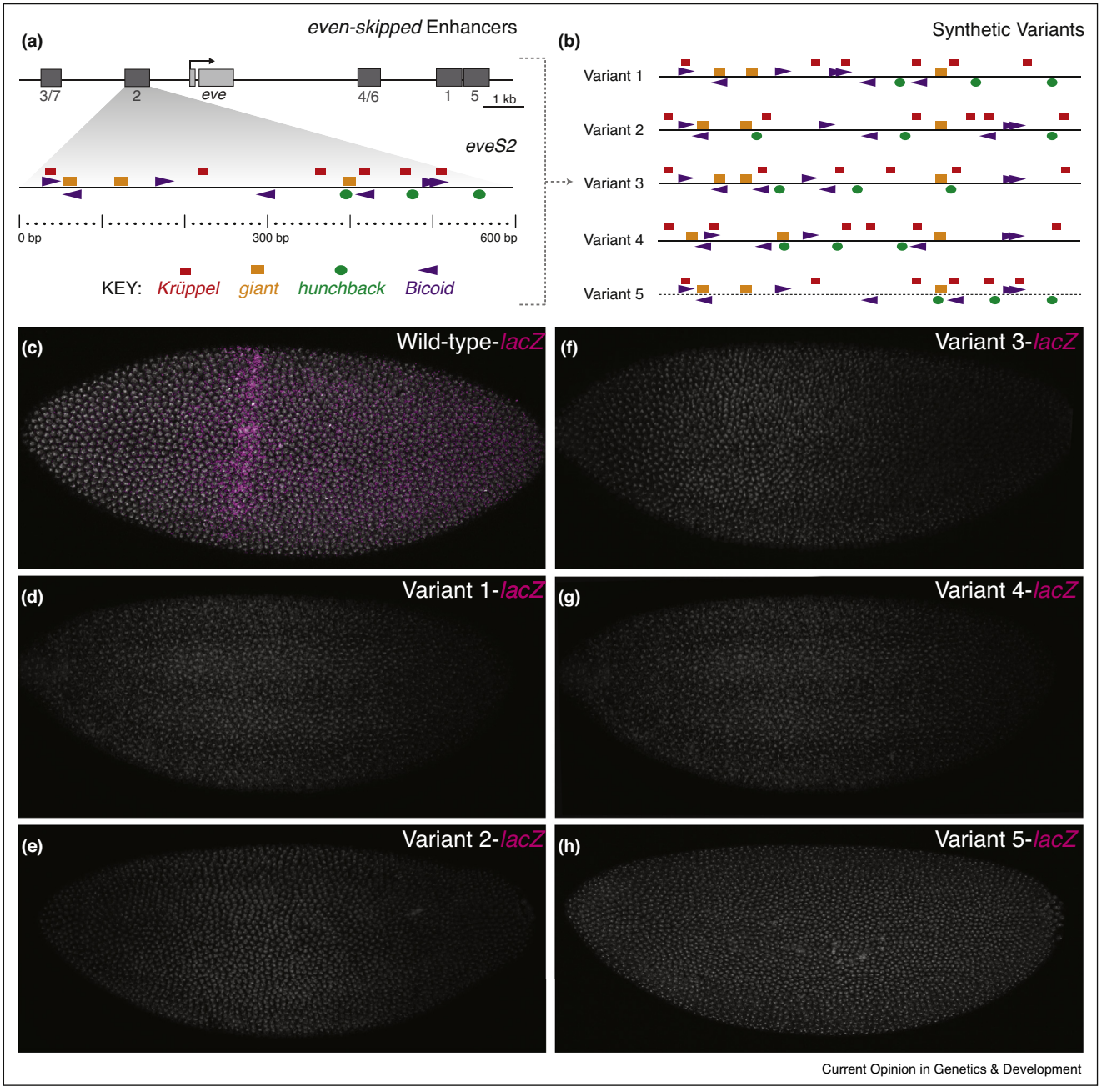
However, despite decades of research, a general and quantitative understanding of enhancer function has eluded discovery [7,8]. An example of this is the enhancer that drives expression of the second stripe of the gene *even-skipped (eveS2)*. Mechanisms for the precise regulation of *eveS2* have been explored by a range of methods including biochemistry, genetics, evolutionary genetics, live imaging, and computational modelling [13–19]. However, attempts to recreate a functional *eveS2* by building 'synthetic' enhancers with known binding sites have led to inconclusive results [20] (Figure 1) — indicating that we lack an understanding of the necessary components for enhancer function, and how these components function collectively.

Recently, there has been a rapid increase in synthetic biology with a focus on the construction of designed genetic systems. Synthetic biology is 'alternative chemistries, artificial cells, self-replicating macromolecules, *in silico* life forms, genetic circuits' [21]. Much of this work is focused on industrial applications, such as the construction of drug precursors [22], information storage [23], or diagnostics devices [24]. In this review, we explore how synthetic biology provides a powerful toolset to understand gene regulation and evolution, allowing biologists to construct tractable systems to test models *in vivo*.

Large-scale mutagenesis of developmental enhancers

A challenge in studying regulatory elements is explaining *why* natural architectures have evolved to be in their current state. For example, it took enormous effort to decipher the evolved regulatory changes in a robust enhancer of the *shavenbaby* gene; to discover that gain

Figure 1



Synthetic *eveS2* enhancers do not drive appreciable expression in embryos. **(a)** Schematic of the *even-skipped* locus, indicating early embryonic *cis*-regulatory stripe enhancers in grey boxes, and the eveS2 minimal enhancer, indicating binding sites for known TFs. **(b)** Different types of synthetic variants of the eveS2 element with either scrambled motifs, variant 1–4, or fixed motif distances with scrambled intervening sequences (dotted line). **(c)–(h)** Stage 5 embryos stained for β-Gal RNA carrying the indicated enhancers. In no cases do any of the synthetic enhances drive expression comparable to the wild-type enhancer (c). See also Vincent *et al*. [20].

of a repressor binding site overcame robustness encoded by multiple activator binding sites [25]. Such a lack of clarity is because, at present, we have a limited understanding of how natural genetic variants influence any given enhancer.

Synthetic saturation mutagenesis provides a powerful means to assay the transcriptional activities of thousands of regulatory elements in a single experiment — providing an *in vivo* assay of mutation effects. The principle of this technology was first applied towards high-resolution analysis of bacteriophage and mammalian promoters by

synthetic saturation mutagenesis *in vitro*, assaying all possible point mutations and small insertions/deletions of these promoters in parallel [26]. Similar assays have been used to dissect a number of promoters and enhancers [27–30]. Together, the results of these studies suggest that individual mutations often have modest effects on enhancer activity. Furthermore, low-affinity, non-canonical motifs [4,31] play a part, emphasizing the importance of more fine-grained experimental characterization.

Application of synthetic saturation mutagenesis to developmental systems will uncover what mutations augment the timing, level, or location of gene expression. However, to date most of these high-throughput assays rely on transient expression using episomal vectors. These assays were performed in the absence of native chromatin context and they are limited to cell types responsive to transfection. Recent experiments have used a lentiviral-based method in mammalian cells, allowing assays in a more biologically relevant context [32•]. In the future, 'lab on a chip' experiments [33,34] will facilitate such high-throughput methodology to be applied to classic model systems using stable transgenic constructs, providing a comprehensive understanding of enhancer function in a chromatin context during animal development.

The construction of synthetic versions of natural enhancers allows defined samples of variants to be examined, whether targeted or uniform. Importantly, such an approach to studying regulatory variation allows both the examination of 'evolutionarily relevant mutations' [35], those found at appreciable frequency in natural populations, as well as those that are rare and/or deleterious. For example, in a *Saccharomyces cerevisiae* study, the effects of polymorphism segregating in the *TDH3* promoter among 85 strains were compared to 235 defined, synthetic mutations in the same promoter [36••], demonstrating that selection on gene expression noise has a greater effect on sequence variation than selection on mean expression level. This fits well with results from engineered circuits in *B. subtilis*, which revealed that 'noisy' circuits provide broad physiological response ranges [37], demonstrating the utility of synthetic approaches. We anticipate that such research will be central to identifying and classifying the individual mutations responsible for variation.

Designed synthetic enhancers and high-throughput screens

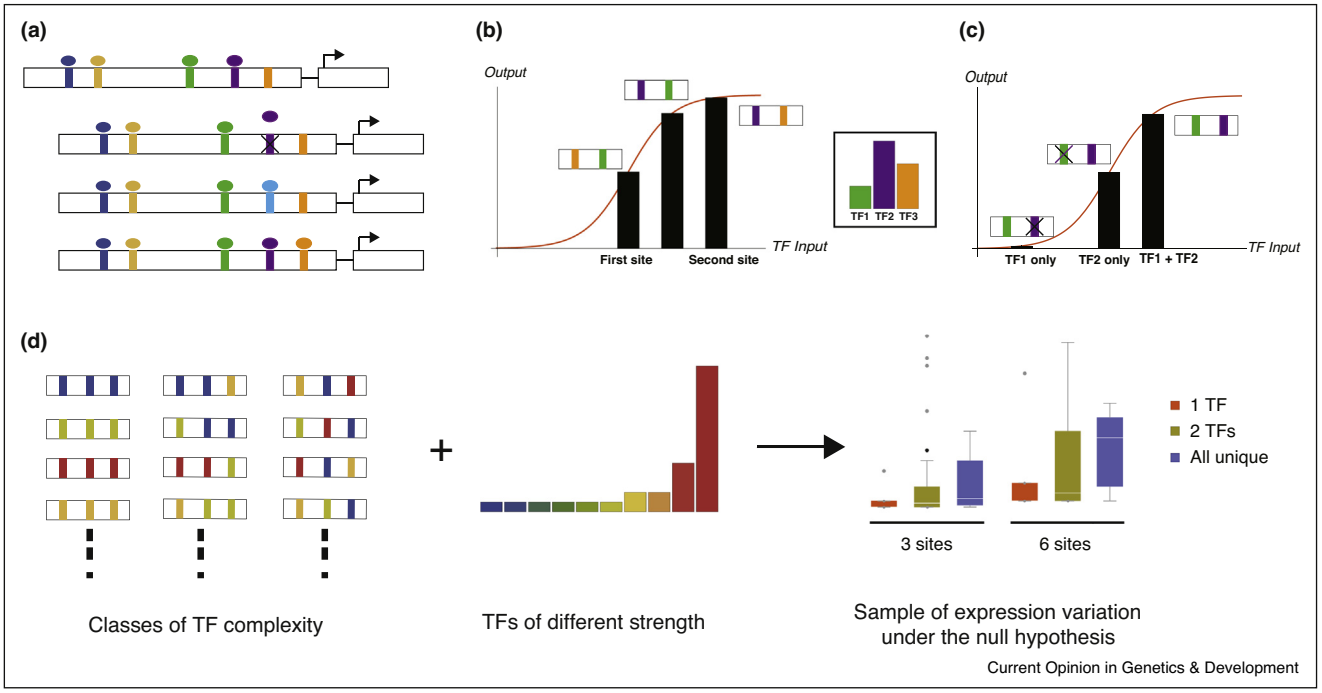
Complex libraries of designed synthetic regulatory elements can be used to explore enhancer structure and its effect on activity [38]. Using synthetic enhancers containing different patterns of twelve liver-specific transcription factor binding sites, Smith *et al.*, suggest that there is flexibility in binding site order [39], and that heterotypic enhancers, composed of different binding site motifs, are expressed more highly than their homotypic analogs. By contrast, in yeast, a library of 6500 synthesized sequences was analyzed, and it was found that transcription factor binding site number, location, and affinity were important for the activity. In *Drosophila*, Erceg and colleagues engineered 63 synthetic enhancers to assess the relationship between variation in the content and spacing of motifs within enhancers [40]. In over half the cases, elements containing only one or two types of transcription factor binding motifs were capable of driving specific patterns during development. Different motif organizations provided different degrees of robustness to enhancer activity, ranging from binary ON/OFF responses to subtle effects in levels and probabilities of expression. Similarly, a high-throughput screen in *Ciona* embryos identified synthetic notochord enhancers that are activated by a combination of two transcription factors [41]. Manipulation of these enhancers elucidated a 'regulatory code' for notochord-specific expression, whereby the optimal spacing of motifs compensates for low-affinity binding sites. Grossman and colleagues used a synthetic biology approach, paired with *in vivo* binding assays, to systematically dissect the contribution of genomic regulatory elements with PPARγ motifs [42•]. They found that different pairs of motifs followed different interaction rules, whether additive, subadditive or superadditive, with some pairs spatially constrained and others having more flexibility.

Interpreting synthetic enhancer results

Layers of control and exceptions to the rule are typical in biology, but it remains an open question as to whether enhancers are built with a flexible architecture that can evolve and change easily or whether complex organization is intrinsic to their function [43].

Regulatory control in developmental biology has usually been explored using genetic tests, reporter assays or studies of DNA binding [44] where the experimental results are by design typically binary, revealing regulatory links that turn expression ON or OFF depending on the context. These results can be combined into complex regulatory networks [45], which have value in summarizing and interpreting discrete experimental data. However, without context the resulting formalism can encourage the view that gene regulation is brittle and that it evolves in steps. By contrast, construction of synthetic regulatory platforms is a powerful alternative to test models of enhancer function, with results that are more fine-grained and quantitative, hence requiring an alternative modelling approach. Various models of gene expression are available [46–50], but they are often complex with many parameters [48,51,52]. This range of choice does not help address a fundamental problem, which is that explaining enhancer function without reference to a suitable null hypothesis leads to an acceptance of models that are more complex than necessary and interpretations

Figure 2



Synthetic enhancer results evaluated against a null model assuming flexible enhancer architecture with collaborative TF binding, represented by a logistic model. **(a)** Different types of synthetic enhancer variants of a base enhancer can be evaluated using the null model, from top to bottom: deleting binding sites, modifying a site to target a specific TF or engineering a TF to target a novel site. **(b)** TF context under the null hypothesis. A new TF (orange) is engineered to replace the existing TF binding sites of a native enhancer (green and purple). The relative strengths of the different TFs are shown in the bar chart. The logistic model predicts that the orange TF will reduce expression when targeting the first site and will increase expression at the second site, thus showing that context-dependent activity can arise without localized cooperativity or enhancer structure. **(c)** Synergy under the null hypothesis. TF2 (purple) binds and activates more strongly than TF1 (green). Assuming no physical cooperativity, TF1 has little effect on its own, TF2 can increase expression, but both are required for full expression, demonstrating synergy. **(d)** Heterotypic binding sites can produce stronger expression than homotypic binding under the null hypothesis. Synthetic enhancer constructs can be categorized by level of complexity depending on how many TFs bind (left). If the TFs have different strengths of binding and activity (center), it is possible for heterotypic binding to produce stronger expression on average (right) under the null hypothesis, without requiring any special cooperativity or enhancer grammar.

that overstate the case of exceptions and enhancer complexity. This is illustrated with a few examples (Figure 2) using arguably the simplest predictive model as the null hypothesis, the logistic model. The base logistic model assumes independent binding of transcription factors, or perhaps more accurately, *collaborative binding*, where binding is dependent on cooperativity with all factors and the promoter, rather than with specific neighbors [53,54,55••]. Collaborative binding has important evolutionary implications in that transcription factor binding sites can be lost or gained gradually. In other words, enhancer structure is not brittle. Yet, despite this simplicity, the associated model explains how a transcription factor targeted to different sites can act to increase or decrease expression depending on context (Figure 2b) and that synergistic effects can result from collaborative binding and do not require pairwise interactions between transcription factors (Figure 2c).

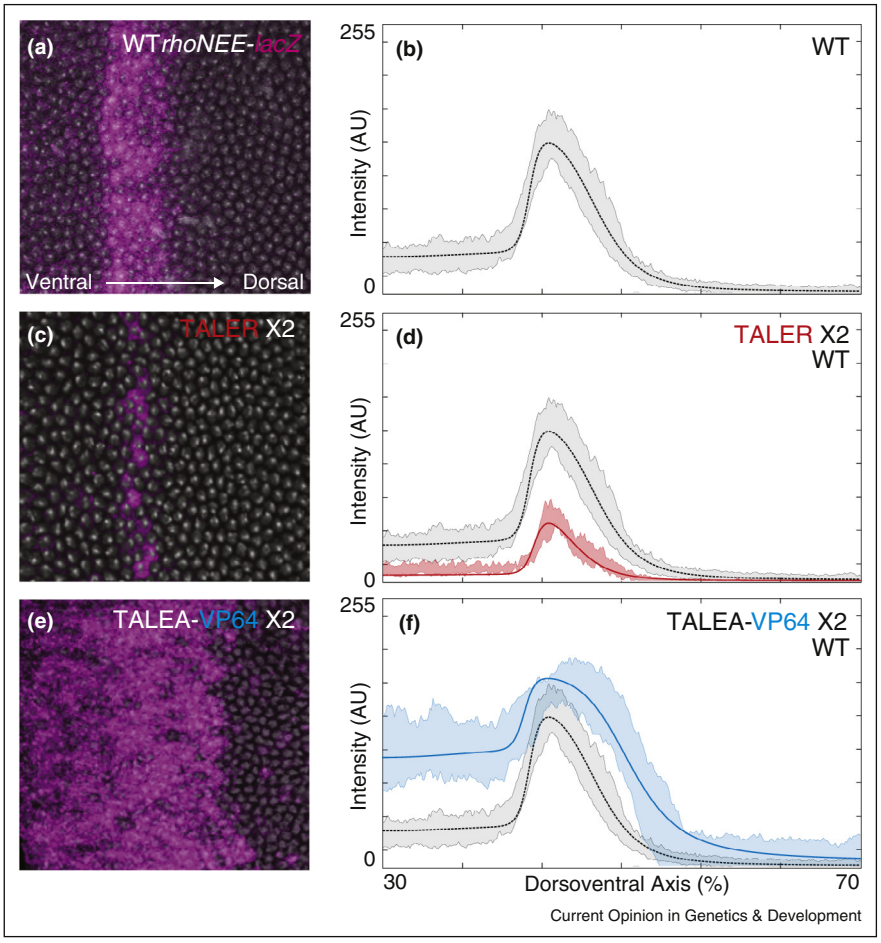
A further example is synthetic enhancers with heterotypic binding sites producing stronger expression on average than those with homotypic binding sites, which can be taken as evidence that increased binding complexity leads to stronger expression [39]. However, this result is also compatible with the null hypothesis of flexible architecture — if transcription factors have suitably different strengths (Figure 2d).

That said, the purpose of these observations is not to rule out the importance of pairwise cooperativity or enhancer grammar in specific contexts, but rather to show the value of the logistic model as an appropriate null hypothesis for interpreting synthetic enhancer results.

Testing models of synthetic enhancer activity

To test models of enhancer activity in a developmental context, we have employed engineered transcription-activator-like proteins (TALEs) fused to activators or repressors [56] (Figure 3), TALEAs and TALERs,

Figure 3



Logistic model predicts the quantitative control of the *rhomboid (rho)* enhancer. **(a, c, e)** Stage 5 embryos stained for *β-Gal* RNA carrying the indicated TALE-*VP64* activator (TALEA) or TALE-*hairy* repressor (TALER) constructs and *rhomboid* enhancers. Adapted from Crocker *et al*. [5••]. **(b, d, f)** Profiles of average expression levels for the embryonic genotypes shown in panels (a, c, e) ($n = 10$ for each genotype). In all plots, the dashed black line denotes the wild-type embryo predictions, with red TALER-*hairy* or TALEA-*VP64* predictions, respectively. Shaded areas indicate one standard deviation of experimental embryonic data. AU indicates Arbitrary Units of fluorescence intensity. The model predicts increased expression in the dorsal and ventral regions with the addition of TALEAs (blue dashed lines).

respectively, to target novel transcription factor activity to enhancers [5••]. We employed 'sequence-free' models, which abstract enhancers as simple machines that measure transcriptional activation and repression as inputs, and produce transcriptional outputs as mathematical functions [47,57–61]. Regulatory input was found to combine linearly with existing inputs and a sequence-free approach accurately modelled each enhancer's transcriptional output, providing a method to quantitatively control enhancers *in vivo*. That enhancers can function as simple input/output devices, where similar transcription factors can substitute for each other is supported by Stampfel and colleagues, who explored the regulatory contributions of transcription factors and cofactors in various combinations by recruiting GAL4 DNA-binding-domain fusions of many *Drosophila* transcription factors and cofactors to different enhancers [62•]. Similarly, Khalil and colleagues used artificial zinc-finger transcription factors to create synthetic transcription factors and used these to wire synthetic transcriptional circuits in yeast [63]. They engineered tunable transcriptional outputs by adjusting key properties; including specificity, affinity, syntax, and protein–protein interactions. In *Drosophila*, a well-defined set of transcriptional modules was used to test a fractional occupancy-based model, which explained the effect of repressors on endogenous activators [50]. Thus, construction of synthetic regulatory platforms is a powerful way to test simple models of enhancer function.

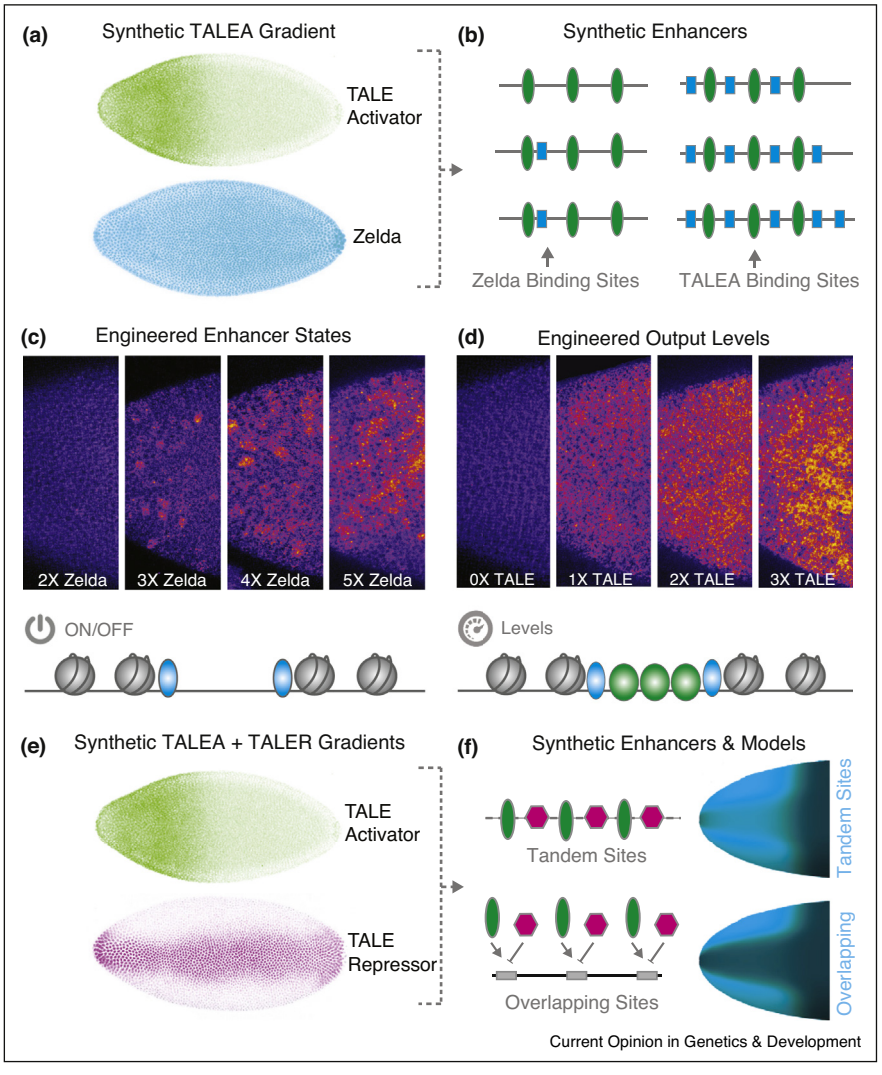
Together, these results have broad evolutionary implications as the activities of individual, and cooperative pairs of transcription factor binding sites are combined in a linear manner to produce a sigmoidal output. Combined with the wide spectrum of individual binding sites

that can be bound by individual transcription factors [4,64–66], this means that the same enhancer activity can be encoded by a vast number of different sequences. This is consistent with a large body of data demonstrating the rapid evolution, and variable architectures, of transcriptional enhancers with conserved functions [19,53,60,67–72].

Building synthetic transcriptional regulatory networks

Transcriptional networks contain the information required to confer robust positional information within a developing embryo. However, these networks have remained recalcitrant to modelling, as every parameter, such as diffusion and binding coefficients, are usually fitted values, and not derived from *in vivo* data. Fully synthetic networks allow the exploration of network architecture with fewer unknowns. In mammalian cell-

Figure 4



A fully synthetic transcriptional platform in *Drosophila* consisting of engineered transcription factor gradients and artificial enhancers. **(a)** Schematic representation of the approach used to build a TALEA gradient using the *hunchback* (*hb*) promoter (adapted from [9••]). **(b)** Schematic of synthetic enhancers built to detect the TALEA gradient. **(c)** Stage 5 embryos stained for *lacZ* expression from synthetic enhancers with the indicated number of Zelda binding sites (adapted from [9••]). The number of Zelda motifs changes the probability of being ON or OFF — or the 'state' of expression. **(d)** Stage 5 embryos stained for *lacZ* expression from synthetic enhancers with the indicated number of TALEA binding sites. The number of TALEA motifs changes the levels of expression. **(e)** Expression patterns of opposing TALEA and TALER gradients. **(f)** Schematic of synthetic enhancers used to test the effect of tandem versus overlapping activator and repressor binding sites. TF occupancy models of embryos with the indicated TALEA and TALER arrangement accurately predict that binding site arrangement provides precision to enhancer expression; see also Crocker *et al*. 2017 [9••].

systems there is an expanding number of studies exploring synthetic gene circuits [73]. Many synthetic networks have focused on unique functions of a particular circuit, for example complex dynamics [74–77] or information processing [78–85]. Other efforts use gene circuits that interface with endogenous inputs from the cell or environment to drive a desired response [24,86•,87].

This framework can be applied to test general principles of development [88]. For example, Crocker and colleagues constructed a simple synthetic enhancer platform in *Drosophila* embryos that responds to an engineered morphogen gradient (Figure 4) [9••]. They observed that while levels of expression are controlled by activator binding site number, the ON or OFF 'state' of enhancer activity was controlled by the number of Zelda binding sites — a so-called 'pioneer factor'. Furthermore, they show how overlapping binding sites can create well-defined expression boundaries during development. Similarly, Bintu *et al.* monitored a transcriptional reporter gene carried on a synthetic human chromosome and found that gene silencing was all-or-none, and that the duration of recruitment of the chromatin regulators determined the fraction of cells that were silenced [89]. Thus, distinct epigenetic modifiers can produce different types of repression and epigenetic memory. The use of 'epigenome editing' [90–93], combined with synthetic enhancer platforms will allow deeper investigation into how different protein domains contribute to enhancer activity than is possible using native enhancers.

It is possible to extend these systems to build more sophisticated synthetic regulatory systems that could be engineered to test the roles of specific features of regulatory architecture during development. Such networks could be used to test directly the roles of canonical 'network motifs' during development [94], such as the role of feed-forward loops, biological noise, oscillations, systems drift, and developmental precision [11,93,95–97]. Synthetic networks could incorperate highly modular construction [98,99], synthetic receptors [100,101] and engineered cell-contractility [102] for the transduction of transcriptional inputs into diverse outputs, providing tests of models for transcription and development [103–106].

Adaptive evolution in animals is a complex optimization process that is highly combinatorial [107]. This complexity is due in part because selection acts over multiple developmental scales — cells, fields of cells, tissues, organs — across different gene regulatory networks, and across different developmental stages [108]. Therefore, it can be difficult to infer past selection events from genomic data. For example, comparative studies of distantly related species and genetic analysis of closely related species indicate that many equivalent characters between taxa differ in their gene regulatory networks, either due to convergent evolution or as a result of developmental system drift [109–111]. Additionally, we know little about the 'ruggedness' of biological fitness landscapes [112], yet the architecture of regulatory networks may dramatically shape evolutionary trajectories [109]. Understanding how networks evolve requires knowledge of the molecular components, their actions, and the dynamics of these interactions.

The ability to build synthetic networks in developmental systems allows the exploration of network architecture with fewer unknowns. Thus, like synthetic enhancer platforms [9••], the construction of synthetic regulatory networks will be a powerful way to test simple models of network function. Such systems will provide many variants, each of which can be clearly defined, freeing evolutionary biologists from both the limited biological data that may exist in natural populations as well as the complexities inherent in densely connected networks [113,114]. The dissection of composite traits into much smaller, discrete problems has proven to be a very successful research program [1,25,115], even though most 'traits' of current evolutionary interest are composites of multiple developmental problems [35]. Understanding how variation in network structure relates to variations in traits will clarify the extent and importance of network modularity in evolution. We anticipate that synthetic networks will provide direct and tractable tests of how gene regulatory networks function, providing new material for the study of variation [116].

Conclusion

The goal of developmental biology is to understand how a single cell develops into a multicellular animal. This complex process requires that cells divide, differentiate, and form precise patterns. Over the past century, we have learned an enormous amount about this process — deciphering the molecules and pathways essential for animal development. Although we have identified many of the constitutive components of development, understanding how networks of these molecules interact to build multicellular animals has proven challenging. Developmental biology can now exploit genome engineering technologies, high-throughput robotics, live-imaging, and single-molecule imaging techniques in live embryos [56,117–120]. Such assays will allow us to measure the number of transcription factor molecules in cells, molecular interactions *in vivo*, and the rates of transcription and translation of genes. Along with these experimental goals, improved models that simplify and abstract aspects of the system will help provide an intuitive understanding of its operation. Together, synthetic biology will allow deeper investigation into the process of development creating a cross-cutting approach to study gene regulatory evolution.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Frankel N, Erezyilmaz DF, McGregor AP, Wang S, Payre F, Stern DL: **Morphological evolution caused by many subtle-effect substitutions in regulatory DNA**. *Nature* 2011, **474**:598-603.

2. Gillespie JH: *The Causes of Molecular Evolution*. Oxford University Press; 1994.

3. Crocker J, Tamori Y, Erives A: **Evolution acts on enhancer organization to fine-tune gradient threshold readouts**. *PLoS Biol* 2008, **6**:e263.

4. Crocker J, Abe N, Rinaldi L, McGregor AP, Frankel N, Wang S, Alsawadi A, Valenti P, Plaza S, Payre F *et al.*: **Low affinity binding site clusters confer hox specificity and regulatory robustness. [Internet]**. *Cell* 2015, **160**:191-203.

5. •• Crocker J, Ilsley GR, Stern DL: **Quantitatively predictable control of Drosophila transcriptional enhancers in vivo with engineered transcription factors. [Internet]**. *Nat Genet* 2016, **48**:292-298.

By combining quantitative models of enhancer function with manipulations using engineered transcription factors, the authors show that enhancer function can be controlled in a quantitatively predictable manner.

6. Spitz F, Furlong EEM: **Transcription factors: from enhancer binding to developmental control. [Internet]**. *Nat Rev Genet* 2012, **13**:613-626.

7. Levo M, Segal E: **In pursuit of design principles of regulatory sequences. [Internet]**. *Nat Rev Genet* 2014, **15**:453-468.

8. Shlyueva D, Stampfel G, Stark A: **Transcriptional enhancers: from properties to genome-wide predictions. [Internet]**. *Nat Rev Genet* 2014, **15**:272-286.

9. •• Crocker J, Tsai A, Stern DL: **A fully synthetic transcriptional platform for a multicellular eukaryote. [Internet]**. *Cell Rep* 2017, **18**:287-296.

By combining engineered transcription factor gradients with synthetic enhancers, the authors show how synthetic networks can be constructed in developmental systems.

10. Barrera LA, Vedenko A, Kurland JV, Rogers JM, Gisselbrecht SS, Rossin EJ, Woodard J, Mariani L, Kock KH, Inukai S *et al.*: **Survey of variation in human transcription factors reveals prevalent DNA binding changes. [Internet]**. *Science (80-.)* 2016, **351**:1450-1454.

11. Lorberbaum DS, Ramos AI, Peterson KA, Carpenter BS, Parker DS, De S, Hillers LE, Blake VM, Nishi Y, McFarlane MR *et al.*: **An ancient yet flexible *cis*-regulatory architecture allows localized Hedgehog tuning by *patched/Ptch1*. [Internet]**. *Elife* 2016, **5**:599-604.

12. •• Farley EK, Olson KM, Zhang W, Brandt AJ, Rokhsar DS, Levine MS: **Suboptimization of developmental enhancers. [Internet]**. *Science (80-.)* 2015, **350**:325-328.

A high-throughput screen of millions of variants of the Otx-a enhancer in Ciona revealed that suboptimal binding sites are important for proper enhancer specificity and that subtle changes in binding site spacing can affect activity.

13. Small S, Warrior R, Kraut R, Levine M, Hoey T, Warrior R, Levine M: **Transcriptional regulation of a pair-rule stripe in Drosophila. [Internet]**. *Genes Dev* 1991, **5**:827-839.

14. Arnosti DN, Barolo S, Levine M, Small S: **The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. [Internet]**. *Development* 1996, **122**:205-214.

15. Ludwig MZ, Kreitman M: **Evolutionary dynamics of the enhancer region of even-skipped in Drosophila. [Internet]**. *Mol Biol Evol* 1995, **12**:1002-1011.

16. Frasch M, Warrior R, Tugwood J, Levine M: **Molecular analysis of even-skipped mutants in Drosophila development. [Internet]**. *Genes Dev* 1988, **2**:1824-1838.

17. Jiang P, Ludwig MZ, Kreitman M, Reinitz J: **Natural variation of the expression pattern of the segmentation gene even-skipped in melanogaster. [Internet]**. *Dev Biol* 2015, **405**:173-181.

18. Ludwig MZ, Patel NH, Kreitman M: **Functional analysis of eve stripe 2 enhancer evolution in Drosophila: rules governing conservation and change**. *Development* 1998, **125**:949-958.

19. Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB: **Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. [Internet]**. *PLoS Genet* 2008, **4**:e1000106.

20. Vincent BJ, Estrada J, DePace AH: **The appeasement of Doug: a synthetic approach to enhancer biology. [Internet]**. *Integr Biol* 2016 http://dx.doi.org/10.1039/C5IB00321K.

21. Bayer TS: **Using synthetic biology to understand the evolution of gene expression. [Internet]**. *Curr Biol* 2010, **20**:R772-R779.

22. DeLoache WC, Russ ZN, Narcross L, Gonzales AM, Martin VJJ, Dueber JE: **An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. [Internet]**. *Nat Chem Biol* 2015, **11**:465-471.

23. Church GM, Gao Y, Kosuri S: **Next-generation digital information storage in DNA. [Internet]**. *Science (80-.)* 2012, **337** 1628-1628.

24. Slomovic S, Pardee K, Collins JJ: **Synthetic biology devices for in vitro and in vivo diagnostics. [Internet]**. *Proc Natl Acad Sci U S A* 2015, **112**:14429-14435.

25. Preger-Ben Noon E, Davis FP, Stern DL: **Evolved repression overcomes enhancer robustness. [Internet]**. *Dev Cell* 2016, **39**:572-584.

26. Patwardhan RP, Lee C, Litvin O, Young DL, Pe'er D, Shendure J: **High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. [Internet]**. *Nat Biotechnol* 2009, **27**:1173-1175.

27. Kinney JB, Murugan A, Callan CG, Cox EC: **Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. [Internet]**. *Proc Natl Acad Sci U S A* 2010, **107**:9158-9163.

28. Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee S-I, Cooper GM *et al.*: **Massively parallel functional dissection of mammalian enhancers in vivo. [Internet]**. *Nat Biotechnol* 2012, **30**:265-270.

29. Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, Feizi S, Gnirke A, Callan CG, Kinney JB *et al.*: **Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. [Internet]**. *Nat Biotechnol* 2012, **30**:271-277.

30. Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA: **Complex effects of nucleotide variants in a mammalian cis-regulatory element. [Internet]**. *Proc Natl Acad Sci* 2012, **109**:19498-19503.

31. Ramos AI, Barolo S, PTRS B: **Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution**. *Philos Trans R Soc L B Biol Sci* 2013, **368**:20130018.

32. • Maricque BB, Dougherty JD, Cohen BA: **A genome-integrated massively parallel reporter assay reveals DNA sequence determinants of *cis*-regulatory activity in neural cells. [Internet]**. *Nucleic Acids Res* 2016, **9**.

The authors develop a novel lentiviral-based, massively parallel reporter gene assay, to study the function of genome-integrated regulatory elements in any mammalian cell type.

33. Levario TJ, Zhao C, Rouse T, Shvartsman SY, Lu H: **An integrated platform for large-scale data collection and precise**

perturbation of live Drosophila embryos. [Internet]. *Sci Rep* 2016, **6**:21366.

34. Cornaglia M, Mouchiroud L, Marette A, Narasimhan S, Lehnert T, Jovaisaite V, Auwerx J, Gijs MAM: **An automated microfluidic platform for *C. elegans* embryo arraying, phenotyping, and long-term live imaging. [Internet]**. *Sci Rep* 2015, **5**:10192.

35. Stern DL: **Evolutionary developmental biology and the problem of variation. [Internet]**. *Evolution* 2000, **54**:1079-1091.

36. •• Metzger BPH, Yuan DC, Gruber JD, Duveau F, Wittkopp PJ: **Selection on noise constrains variation in a eukaryotic promoter. [Internet]**. *Nature* 2015, **521**:344-347.
This study shows how systematically examining the effects of promoter mutations can inform our understanding of evolutionary mechanisms.

37. Cağatay T, Turcotte M, Elowitz MB, Garcia-Ojalvo J, Süel GM, Kobayashi M, Okamura H, Walt DR, Collins JJ: **Architecture-dependent noise discriminates functionally analogous differentiation circuits. [Internet]**. *Cell* 2009, **139**:512-522.

38. Amit R, Garcia HGG, Phillips R, Fraser SEE: **Building enhancers from the ground up: a synthetic biology approach. [Internet]**. *Cell* 2011, **146**:105-118.

39. Smith RP, Taher L, Patwardhan RP, Kim MJ, Inoue F, Shendure J, Ovcharenko I, Ahituv N: **Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. [Internet]**. *Nat Genet* 2013, **45**:1021-1028.

40. Erceg J, Saunders TE, Girardot C, Devos DP, Hufnagel L, Furlong EEM: **Subtle changes in motif positioning cause tissue-specific effects on robustness of an enhancer's activity. [Internet]**. *PLoS Genet* 2014, **10**:e1004060.

41. Farley EK, Olson KM, Zhang W, Rokhsar DS, Levine MS: **Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. [Internet]**. *Proc Natl Acad Sci* 2016, **113**:6508-6513.

42. • Grossman SR, Zhang X, Wang L, Engreitz J, Melnikov A, Rogov P, Tewhey R, Isakova A, Deplancke B, Bernstein BE *et al.*: **Systematic dissection of genomic features determining transcription factor binding and enhancer function. [Internet]**. *Proc Natl Acad Sci* 2017, **114**:E1291-E1300.
Using reporter assays of tens of thousands of natural and synthetic enhancers, this study shows that PPARγ binding is largely governed by binding site affinity and chromatin accessibility, whereas enhancer activity depends on dozens of transcription factors that bind in the immediate vicinity.

43. Arnosti DN, Kulkarni MM: **Transcriptional enhancers: intelligent enhanceosomes or flexible billboards? [Internet]**. *J Cell Biochem* 2005, **94**:890-898.

44. Glassford WJ, Johnson WC, Dall NR, Smith SJ, Liu Y, Boll W, Noll M, Rebeiz M: **Co-option of an ancestral Hox-regulated network underlies a recently evolved morphological novelty. [Internet]**. *Dev Cell* 2015, **34**:520-531.

45. Davidson EH: *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*. Academic Press; 2006.

46. Turing AM: **The chemical basis of morphogenesis. [Internet]**. *Bull Math Biol* 1952, **237**:37-72.

47. Jaeger J, Surkova S, Blagov M, Janssens H, Kosman D, Kozlov KN, Manu, Myasnikova E, Vanario-Alonso CE, Samsonova M *et al.*: **Dynamic control of positional information in the early Drosophila embryo. [Internet]**. *Nature* 2004, **430**:368-371.

48. Janssens H, Hou S, Jaeger J, Kim A-R, Myasnikova E, Sharp D, Reinitz J: **Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster* even skipped gene. [Internet]**. *Nat Genet* 2006, **38**:1159-1165.

49. Kim A-R, Martinez C, Ionides J, Ramos AF, Ludwig MZ, Ogawa N, Sharp DH, Reinitz J: **Rearrangements of 2.5 kilobases of noncoding DNA from the Drosophila even-skipped locus define predictive rules of genomic cis-regulatory logic. [Internet]**. *PLoS Genet* 2013, **9**:e1003243.

50. Fakhouri WD, Ay A, Sayal R, Dresch J, Dayringer E, Arnosti DN: **Deciphering a transcriptional regulatory code: modeling short-range repression in the Drosophila embryo. [Internet]**. *Mol Syst Biol* 2010, **6**:341.

51. Ay A, Arnosti DN: **Mathematical modeling of gene expression: a guide for the perplexed biologist. [Internet]**. *Crit Rev Biochem Mol Biol* 2011, **46**:137-151.

52. Samee MAH, Sinha S: **Quantitative modeling of a gene's expression from its intergenic sequence. [Internet]**. *PLoS Comput Biol* 2014, **10**:e1003467.

53. Junion G, Spivakov M, Girardot C, Braun M, Gustafson EH, Birney E, Furlong EEM: **A transcription factor collective defines cardiac cell fate and reflects lineage history. [Internet]**. *Cell* 2012, **148**:473-486.

54. Reiter F, Wienerroither S, Stark A: **Combinatorial function of transcription factors and cofactors. [Internet]**. *Curr Opin Genet Dev* 2017, **43**:73-81.

55. •• Deplancke B, Alpern D, Gardeux V: **The genetics of transcription factor DNA binding variation. [Internet]**. *Cell* 2016, **166**:538-554.
This review discusses recent findings that the molecular mechanisms underlying TF-DNA binding variation are more complex than originally anticipated and reviews the proposed mechanisms for local, proximal, or distal genetic variation-mediated changes in TF-DNA binding.

56. Crocker J, Stern DL: **TALE-mediated modulation of transcriptional enhancers in vivo. [Internet]**. *Nat Methods* 2013, **10**:762-767.

57. Glass L, Kauffman SA: **The logical analysis of continuous, non-linear biochemical control networks. [Internet]**. *J Theor Biol* 1973, **39**:103-129.

58. Reinitz J, Mjolsness E, Sharp DH: **Model for cooperative control of positional information in Drosophila by bicoid and maternal hunchback. [Internet]**. *J Exp Zool* 1995, **271**:47-56.

59. Reinitz J, Sharp DH: **Mechanism of eve stripe formation. [Internet]**. *Mech Dev* 1995, **49**:133-158.

60. Ilsley GR, Fisher J, Apweiler R, Depace AH, Luscombe NM: **Cellular resolution models for even skipped regulation in the entire Drosophila embryo. [Internet]**. *Elife* 2013, **2**:e00522.

61. Staller MV, Vincent BJ, Bragdon MDJ, Lydiard-Martin T, Wunderlich Z, Estrada J, DePace AH: **Shadow enhancers enable Hunchback bifunctionality in the Drosophila embryo. [Internet]**. *Proc Natl Acad Sci U S A* 2015, **112**:785-790.

62. • Stampfel G, Kazmar T, Frank O, Wienerroither S, Reiter F, Stark A: **Transcriptional regulators form diverse groups with context-dependent regulatory functions. [Internet]**. *Nature* 2015, **528**:147-151.
The regulatory contributions of TFs and cofactors in various combinations was examined by recruiting GAL4 DNA-binding-domain fusions of hundreds of Drosophila TFs and cofactors to different enhancers.

63. Khalil AS, Lu TK, Bashor CJ, Ramirez CL, Pyenson NC, Joung JK, Collins JJ, Ye Z, Park IH, Daley GQ *et al.*: **A synthetic biology framework for programming eukaryotic transcription functions. [Internet]**. *Cell* 2012, **150**:647-658.

64. Wong D, Teixeira A, Oikonomopoulos S, Humburg P, Lone IN, Saliba D, Siggers T, Bulyk M, Angelov D, Dimitrov S *et al.*: **Extensive characterization of NF-κB binding uncovers non-canonical motifs and advances the interpretation of genetic functional traits. [Internet]**. *Genome Biol* 2011, **12**:R70.

65. Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X *et al.*: **Diversity and complexity in DNA recognition by transcription factors. [Internet]**. *Science* 2009, **324**:1720-1723.

66. Zuo Z, Stormo GD: **High-resolution specificity from DNA sequencing highlights alternative modes of Lac repressor binding. [Internet]**. *Genetics* 2014, **198**:1329-1343.

67. Rastegar S, Hess I, Dickmeis T, Nicod JC, Ertzer R, Hadzhiev Y, Thies W-G, Scherer G, Strähle U: **The words of the regulatory code are arranged in a variable manner in highly conserved enhancers. [Internet]**. *Dev Biol* 2008, **318**:366-377.

68. Jin H, Stojnic R, Adryan B, Ozdemir A, Stathopoulos A, Frasch M: **Genome-wide screens for in vivo Tinman binding sites identify**

cardiac enhancers with diverse functional architectures. [Internet]. *PLoS Genet* 2013, **9**:e1003195.

69. Brown CD, Johnson DS, Sidow A: **Functional architecture and evolution of transcriptional elements that drive gene coexpression**. *Science (80-.)* 2007, **317**:1557-1560.

70. Menoret D, Santolini M, Fernandes I, Spokony R, Zanet J, Gonzalez I, Latapie Y, Ferrer P, Rouault H, White KP *et al.*: **Genome-wide analyses of Shavenbaby target genes reveals distinct features of enhancer organization. [Internet]**. *Genome Biol* 2013, **14**:R86.

71. Lusk RW, Eisen MB: **Evolutionary mirages: selection on binding site composition creates the illusion of conserved grammars in Drosophila enhancers. [Internet]**. *PLoS Genet* 2010, **6**: e1000829.

72. Liberman LM, Stathopoulos A: **Design flexibility in cis-regulatory control of gene expression: synthetic and comparative evidence. [Internet]**. *Dev Biol* 2009, **327**:578-589.

73. Ausländer S, Fussenegger M: **Engineering gene circuits for mammalian cell-based applications. [Internet]**. *Cold Spring Harb Perspect Biol* 2016, **8**:a023895.

74. Elowitz MB, Leibler S: **A synthetic oscillatory network of transcriptional regulators. [Internet]**. *Nature* 2000, **403**:335-338.

75. Gardner TS, Cantor CR, Collins JJ: **Construction of a genetic toggle switch in *Escherichia coli*. [Internet]**. *Nature* 2000, **403**:339-342.

76. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J: **A fast, robust and tunable synthetic gene oscillator. [Internet]**. *Nature* 2008, **456**:516-519.

77. Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M: **A tunable synthetic mammalian oscillator. [Internet]**. *Nature* 2009, **457**:309-312.

78. Ausländer S, Ausländer D, Müller M, Wieland M, Fussenegger M: **Programmable single-cell mammalian biocomputers. [Internet]**. *Nature* 2012, **487**:123-127.

79. Benenson Y: **Biomolecular computing systems: principles, progress and potential. [Internet]**. *Nat Rev Genet* 2012, **13**:455-468.

80. Deans TL, Cantor CR, Collins JJ: **A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. [Internet]**. *Cell* 2007, **130**:363-372.

81. Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ: **Synthetic gene networks that count. [Internet]**. *Science* 2009, **324**:1199-1202.

82. Green AA, Silver PA, Collins JJ, Yin P: **Toehold switches: de-novo-designed regulators of gene expression. [Internet]**. *Cell* 2014, **159**:925-939.

83. Park S-H, Zarrinpar A, Lim WA: **Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. [Internet]**. *Science (80-.)* 2003:299.

84. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y: **A universal RNAi-based logic evaluator that operates in mammalian cells. [Internet]**. *Nat Biotechnol* 2007, **25**:795-801.

85. Tamsir A, Tabor JJ, Voigt CA: **Robust multicellular computing using genetically encoded NOR gates and chemical "wires". [Internet]**. *Nature* 2011, **469**:212-215.

86. • Angelici B, Mailand E, Haefliger B, Benenson Y: **Synthetic biology platform for sensing and integrating endogenous transcriptional inputs in mammalian cells. [Internet]**. *Cell Rep* 2016, **16**:2525-2537.
The authors establish a framework for developing synthetic gene networks that interface with cellular processes through transcriptional regulators.

87. Ausländer D, Ausländer S, Charpin-El Hamri G, Sedlmayer F, Müller M, Frey O, Hierlemann A, Stelling J, Fussenegger M, Newgard CB *et al.*: **A synthetic multifunctional mammalian pH sensor and CO_2 transgene-control device. [Internet]**. *Mol Cell* 2014, **55**:397-408.

88. Davies J: **Using synthetic biology to explore principles of development. [Internet]**. *Development* 2017:144.

89. Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M, Elowitz MB: **Dynamics of epigenetic regulation at the single-cell level. [Internet]**. *Science (80-.)* 2016:351.

90. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR: **Dynamics and memory of heterochromatin in living cells. [Internet]**. *Cell* 2012, **149**:1447-1460.

91. Keung AJJ, Bashor CJJ, Kiriakov S, Collins JJJ, Khalil ASS: **Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. [Internet]**. *Cell* 2014, **158**:110-120.

92. Ragunathan K, Jih G, Moazed D: **Epigenetic inheritance uncoupled from sequence-specific recruitment. [Internet]**. *Science (80-.)* 2015, **348** 1258699-1258699.

93. Tkačik G, Dubuis JO, Petkova MD, Gregor T: **Positional information, positional error, and readout precision in morphogenesis: a mathematical framework. [Internet]**. *Genetics* 2015:199.

94. Alon U: **Network motifs: theory and experimental approaches. [Internet]**. *Nat Rev Genet* 2007, **8**:450-461.

95. Rowland MA, Abdelzaher A, Ghosh P, Mayo ML: **Crosstalk and the dynamical modularity of feed-forward loops in transcriptional regulatory networks. [Internet]**. *Biophys J* 2017, **112**:1539-1550.

96. Suderman R, Bachman JA, Smith A, Sorger PK, Deeds EJ: **Fundamental trade-offs between information flow in single cells and cellular populations. [Internet]**. *Proc Natl Acad Sci* 2017, **114**:5755-5760.

97. Wotton KR, Jiménez-Guri E, Crombach A, Janssens H, Alcaine-Colet A, Lemke S, Schmidt-Ott U, Jaeger J: **Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. [Internet]**. *Elife* 2015:4.

98. Li Y, Jiang Y, Chen H, Liao W, Li Z, Weiss R, Xie Z: **Modular construction of mammalian gene circuits using TALE transcriptional repressors. [Internet]**. *Nat Chem Biol* 2015, **11**:207-213.

99. Weinberg BH, Pham NTH, Caraballo LD, Lozanoski T, Engel A, Bhatia S, Wong WW: **Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. [Internet]**. *Nat Biotechnol* 2017, **35**:453-462.

100. Morsut L, Roybal KT, Xiong X, Gordley RM, Coyle SM, Thomson M, Lim WA: **Engineering customized cell sensing and response behaviors using synthetic notch receptors. [Internet]**. *Cell* 2016, **164**:780-791.

101. He L, Huang J, Perrimon N: **Development of an optimized synthetic Notch receptor as an in vivo cell–cell contact sensor. [Internet]**. *Proc Natl Acad Sci U S A* 2017, **114**:5467-5472.

102. Guglielmi G, Barry JD, Huber W, De Renzis S: **An optogenetic method to modulate cell contractility during tissue morphogenesis. [Internet]**. *Dev Cell* 2015, **35**:646-660.

103. Cohen M, Page KM, Perez-Carrasco R, Barnes CP, Briscoe J: **A theoretical framework for the regulation of Shh morphogen-controlled gene expression. [Internet]**. *Development* 2014:141.

104. Scholes C, DePace AH, Sánchez Á: **Combinatorial gene regulation through kinetic control of the transcription cycle. [Internet]**. *Cell Syst* 2017, **4**:97-108 e9.

105. Estrada J, Wong F, DePace A, Gunawardena J: **Information integration and energy expenditure in gene regulation. [Internet]**. *Cell* 2016, **166**:234-244.

106. Cotterell J, Robert-Moreno A, Sharpe J: **A local, self-organizing reaction-diffusion model can explain somite patterning in embryos. [Internet]**. *Cell Syst* 2015, **1**:257-269.

107. Kauffman S, Levin S: **Towards a general theory of adaptive walks on rugged landscapes. [Internet]**. *J Theor Biol* 1987, **128**:11-45.

108. Davidson EH, Erwin DH: **Gene regulatory networks and the evolution of animal body plans. [Internet]**. *Science* 2006, **311**:796-800.

109. Stern DL, Orgogozo V: **The loci of evolution: how predictable is genetic evolution?. [Internet]**. *Evolution* 2008, **62**:2155-2177.

110. Stern DL: **The genetic causes of convergent evolution. [Internet]**. *Nat Rev Genet* 2013, **14**:751-764.

111. True JR, Haag ES: **Developmental system drift and flexibility in evolutionary trajectories. [Internet]**. *Evol Dev* 2001, **3**:109-119.

112. Orr HA: **Fitness and its role in evolutionary genetics. [Internet]**. *Nat Rev Genet* 2009, **10**:531-539.

113. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I *et al.*: **Transcriptional regulatory networks in *Saccharomyces cerevisiae*. [Internet]**. *Science (80-.)* 2002:298.

114. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG *et al.*: **Core transcriptional regulatory circuitry in human embryonic stem cells. [Internet]**. *Cell* 2005, **122**:947-956.

115. Prud'homme B, Gompel N, Carroll SB: **Emerging principles of regulatory evolution. [Internet]**. *Proc Natl Acad Sci U S A* 2007, **104(Suppl)**:8605-8612.

116. Bateson W: *Materials for the Study of Variation: Treated With Especial Regard to Discontinuity in the Origin of Species*. Macmillian and Co.; 1894.

117. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M: **Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. [Internet]**. *Proc Natl Acad Sci U S A* 2014, **111**:10598-10603.

118. Cho W-K, Jayanth N, English BP, Inoue T, Andrews JO, Conway W, Grimm JB, Spille J-H, Lavis LD, Lionnet T *et al.*: **RNA Polymerase II cluster dynamics predict mRNA output in living cells**. *Elife* 2016, **5**:681-686.

119. Grimm JB, English BP, Chen J, Slaughter JP, Zhang Z, Revyakin A, Patel R, Macklin JJ, Normanno D, Singer RH *et al.*: **A general method to improve fluorophores for live-cell and single-molecule microscopy. [Internet]**. *Nat Methods* 2015, **12**:244-250.

120. Garcia HGG, Tikhonov M, Lin A, Gregor T: **Quantitative imaging of transcription in living drosophila embryos links polymerase activity to patterning. [Internet]**. *Curr Biol* 2013, **23**:2140-2145.