Trends in
Ecology & Evolution

CellPress

Review

# Beyond modular enhancers: new questions in *cis*-regulatory evolution

Jeanne M.C. McDonald[1,*] and Robert D. Reed[1]

**Our understanding of how *cis*-regulatory elements work has advanced rapidly, outpacing our evolutionary models. In this review, we consider the implications of new mechanistic findings for evolutionary developmental biology. We focus on three different debates: whether evolutionary innovation occurs more often via the modification of old *cis*-regulatory elements or the emergence of new ones; the extent to which individual elements are specific and autonomous or multifunctional and interdependent; and how the robustness of *cis*-regulatory architectures influences the rate of trait evolution. These discussions lead us to propose new questions for the evo-devo of *cis*-regulation.**

## Highlights

Many studies challenge the view that gene expression is primarily controlled by rapidly evolving, highly specific, and independently acting enhancers.

Functional conservation of *cis*-regulatory elements, and the frequency of their co-option, may be underestimated because *cis*-regulatory elements can diverge considerably in sequence while binding the same transcription factors and performing the same developmental roles across species.

*cis*-Regulatory elements are often involved in regulating gene expression in the development of multiple traits and can be highly interdependent upon each other.

Some *cis*-regulatory elements are more robust to mutation compared with others, and the fragility of the *cis*-regulatory architecture of a trait could predict its rate of evolution.

New molecular insights prompt us to propose new questions for evolutionary developmental biology.

## *cis*-Regulatory evolution: new models, new questions

One of the aims of evolutionary developmental biology is to understand the genetic mechanisms that underlie the origin and evolution of traits [1]. Change in gene regulation is known to be an important driver of trait diversification. The evolution of **enhancers** (see Glossary) in particular has been proposed as a primary mechanism of trait evolution [1,2] because individual elements are thought to have highly compartmentalized, trait-specific functions [3–7]. In this review, we consider recent advances in our understanding of ***cis*-regulatory element** structure and function to expand and add nuance to models of regulatory evolution [8–10].

## Modification versus *de novo* evolution of *cis*-regulatory elements

There is a relatively limited repertoire of genes in eukaryotes that evolution continuously modifies and redeploys through **co-option** to generate the diverse forms we see in nature. In contrast to these gene sequences, most noncoding DNA is highly diverged across species [11,12]. Known or predicted *cis*-regulatory elements are more dissimilar in sequence when compared with genes [13–16] (Box 1), and genomic sites where **transcription factors** bind are often not shared across species* [19–22]. These observations indicate that *cis*-regulatory elements evolve more rapidly than genes and are rapidly gained and lost over time. At first glance, the regulatory genome appears to be constantly rewritten by evolution. Based on these data alone, then, we would expect that *de novo* evolution of regulatory elements contributes more to trait evolution than does modification of ancestral elements (Table 1). New data challenge this assumption, however.

Recent comparative studies show that many elements are more conserved across taxa than previously appreciated [45–47]. Some deeply conserved sequences are even shared between vertebrates and invertebrates, including humans and acorn worms [45], and others have been

[1]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA

*Correspondence:
jmm836@cornell.edu (J.M.C. McDonald).

* Several studies of closely related *Drosophila* species find higher conservation of binding sites and that binding diverges with evolutionary time [17,18]. The degree of conservation of binding sites could also be influenced by the developmental stage or tissue compared and/or the analytical approach chosen to estimate conservation.

found across all flowering plants [48]. Do these deeply conserved elements contribute to ongoing evolution, or are they too functionally constrained? A few case studies suggest that some evolutionary novelties do, in fact, appear to be products of mutations in deeply conserved elements [49–51]. For example, the evolution of a spot on the wing of the fly *Drosophila biarmipes* has been linked to changes in a conserved enhancer for the gene *yellow* [49]. This study is compelling because it suggests that trait evolution can be linked to functional evolution of ancestral regulatory elements. However, we still ultimately lack a critical mass of case studies[†] to be able to understand how modification of ancestral elements drives variation and adaptation of traits, and we see this as an important area of future research.

## *cis*-Regulatory ancestry: the problem of covert homology

While a subset of *cis*-regulatory elements are deeply conserved in sequence, many more may be functionally conserved even though their sequences have significantly diverged. Several recent studies suggest that *cis*-regulatory elements that are so divergent as to be unalignable between species are nonetheless orthologous [10,55–57]. *cis*-Regulatory elements can be functionally conserved without preserving exact DNA sequence or organization. These so-called '**covert homologs**' occur in the same genomic locations, bind the same transcription factors, and regulate target genes in the same manner across species [10,56–58], but their sequences are unalignable. Covert homology could be common, but the scale of the phenomenon remains unclear, and there is still no test to convincingly rule out convergent evolution as an explanation for some of the similarity between elements [59]. In any case, covert homology presents the possibility that the extent of *de novo cis*-regulatory element evolution has been overestimated, perhaps significantly.

## How do new expression domains originate?

New expression patterns can evolve via modification to existing elements or, alternatively, new elements can emerge to drive new domains via spontaneous mutation, duplication of existing elements, or **transposable element** insertion (Table 1). On the one hand, existing elements already have transcription factor binding sites that can be co-opted to drive new expression domains, and some researchers have argued that modification of existing elements is the primary mechanism of novel expression pattern evolution [6,49,52]. On the other hand, *cis*-regulatory element duplication and transposable element insertion could also generate new expression domains via reuse of existing transcription factor binding sites. By copying these binding sites into new elements, these mechanisms avoid the potential deleterious effects of modifying existing elements [60–62]. The origins of novel traits and adaptive radiations have been linked to duplications of genes and gene clusters, and there are a few studies linking phenotypic novelties specifically to the duplication of *cis*-regulatory elements [61,63,64]. For example, in the domestic chicken *Gallus domesticus,* tandem duplication of a noncoding region is associated with evolutionary changes in comb morphology and novel expression of a nearby gene in comb tissue [64]. There are also case studies implicating transposable element-derived regulatory sequences in the origins of novel and adaptive traits [27,65,66]. Transposable elements and *cis*-regulatory element duplication could both be productive sources of potentially adaptive *cis*-regulatory variation.

There is, however, experimental evidence suggesting that the appearance of *de novo* regulatory elements could play an even greater role in the emergence of new gene functions. Several recent mutational screens suggest that novel regulatory elements are more likely to contribute to novel spatial domains of gene expression than are mutations of ancestral elements. For example,

[†] While we lack a critical mass to say anything conclusively, a few case studies provide a good foundation for our understanding of *cis*-regulatory element modification and trait origin [52,53]. However, there are not many studies that conclusively demonstrate co-option of conserved elements and rule out the possibility of variation in upstream regulatory proteins as the sole explanation for differences in enhancer function across species. Please see [54] for an explanation of *cis* versus *trans* evolution and how to distinguish between them.

## Glossary

**Chromatin:** complex of DNA and proteins that condense it in the eukaryotic cell. Chromatin accessibility refers to whether transcription factors and polymerases are able to access a region of DNA to regulate transcription, or whether the DNA is compacted and inaccessible.
***cis*-Regulatory element:** unit of noncoding DNA that regulates gene expression.
**Co-option:** evolutionary repurposing of an existing part of an organism, such as a *cis*-regulatory element, gene, or morphological structure, for a new function.
**Covert homolog:** *cis*-regulatory element that retains its ancestral function without conservation of sequence. These homologs are often identified by conservation of transcription factor-binding sites, genomic location, and function.
**Enhancer:** regulatory sequence that increases gene transcription.
**Insulator:** regulatory sequence that blocks enhancing and/or silencing effects from nearby DNA.
**Interdependence:** condition that the proper function of an element requires the presence of, or interaction with, at least one other element.
**Modularity:** ability of a *cis*-regulatory element to independently control a specific domain of gene expression.
**Necessity:** condition that a unit of DNA is required for the normal development of a trait.
**Pleiotropy:** condition of a single unit of DNA affecting multiple distinct traits.
**Promoter:** regulatory sequence where gene transcription begins.
**Robustness:** ability of phenotypes to be unaffected by environmental and/or genetic perturbations.
**Silencer:** regulatory sequence that reduces gene transcription.
**Sufficiency:** condition that a unit of DNA can induce a phenotype independent of other elements.
**Transcription factor:** protein that regulates gene expression by binding to DNA. Transcription factors can both activate and repress transcription of DNA into RNA.
**Transposable element:** unit of DNA that is able to move to a different position in the genome, often referred to colloquially as 'jumping gene'. Transposable elements account for a large fraction of most eukaryotic genomes.

## Box 1. What are *cis*-regulatory elements and how are they identified?

*cis*-Regulatory elements are noncoding units of DNA that regulate gene transcription. They are often categorized into four different classes: promoters, enhancers, silencers, and insulators. Promoters and enhancers activate transcription, while silencers repress it. Insulators regulate genome organization, both promoting and inhibiting gene transcription by restricting the activity of enhancers and silencers to specific genomic regions. *cis*-Regulatory elements are often classified as only one of these types, but there is growing evidence that many of them can fit multiple categories: an enhancer can act as a silencer for the same gene in different developmental contexts [123–125]; the promoter of one gene can act as an enhancer for another gene [126,127]; and a promoter for one gene can block nearby enhancers from activating neighboring genes, acting as a type of insulator [128–130].

*cis*-Regulatory elements are identified using a combination of sequencing and functional experiments (Figure I). First, different types of sequencing, such as assay for transposase-accessible chromatin sequencing (ATAC-seq) and chromatin immunoprecipitation sequencing (ChIP-seq), can identify potential *cis*-regulatory elements (Figure IA). While *cis*-regulatory elements are sometimes defined as much larger regions encompassing multiple distinct accessible regions (e.g., 'promoter' is sometimes used to refer to the whole upstream region of a gene), here we define a *cis*-regulatory element as a distinct region of accessible DNA delimited by changes in chromatin accessibility or by histone marks. Deleting candidate elements tests for the necessity of those elements in trait development (Figure IB). Assays for changes in gene expression using *in situ* hybridization (Figure IB) or changes in phenotype can confirm regulatory activity and phenotypic effects, respectively. In addition, reporter construct experiments determine the sufficiency of individual elements to affect gene expression by placing them upstream of a visually detectable reporter protein (Figure IC). For more detailed reviews of these methods and their many variations, see [131–133].

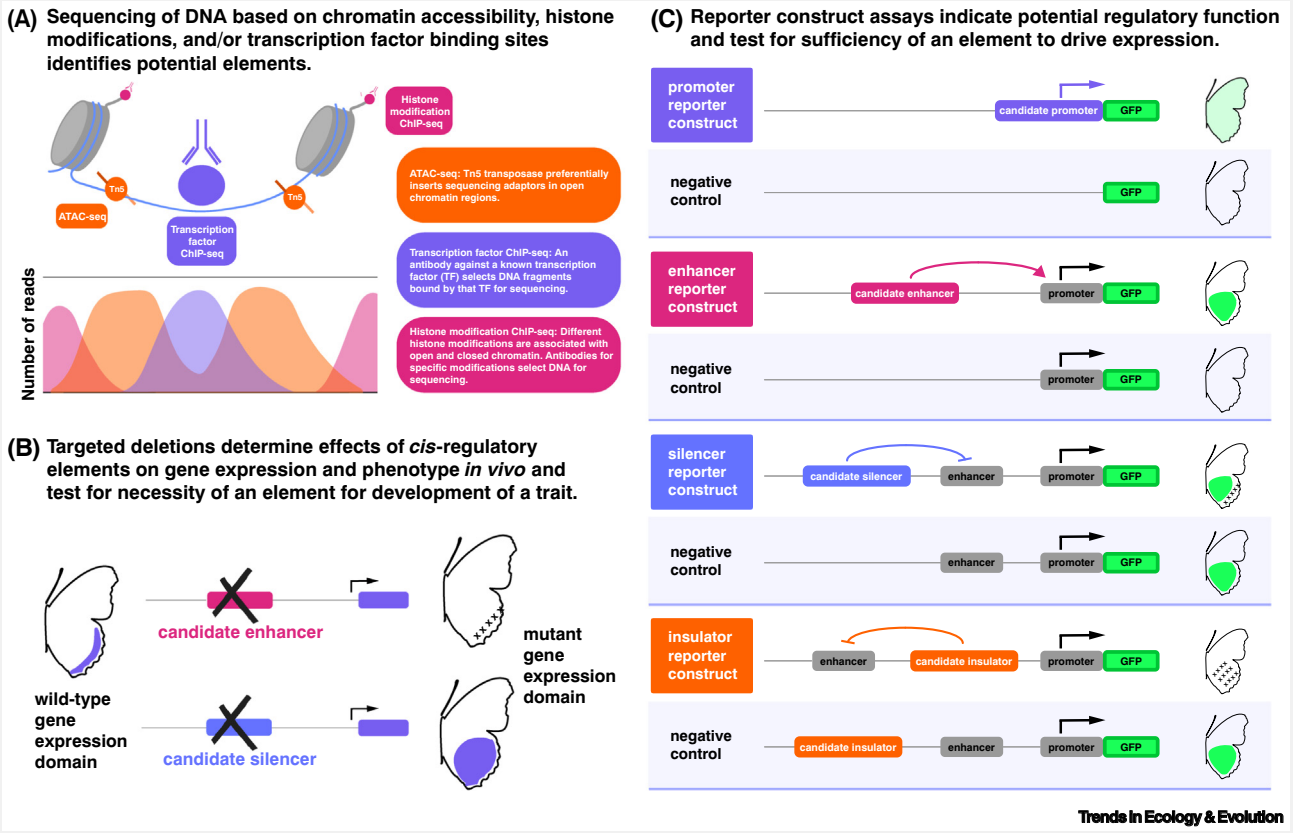


Figure I. Methods of characterizing *cis*-regulatory elements. Abbreviations: ATAC-seq, assay for transposase-accessible chromatin sequencing, ChIP-seq, chromatin immunoprecipitation sequencing.

Galupa *et al.* found that most point mutations of four well-characterized embryonic enhancers in *Drosophila melanogaster* did not change which spatial domain or developmental stage the gene was expressed in [67]. This constraint has also been described in **promoters**: mutations in promoter sequence more often alter the level of transcription than the spatial pattern of gene expression [68]. By contrast, Galupa *et al.* found that most synthetic random sequences were able to

Table 1. Mechanisms of *cis*-regulatory element *de novo* emergence and modification[a]

| | | |
|---|---|---|
| **Mechanisms of *cis*-regulatory element emergence** | | |
| Spontaneous mutations | | [23,24] |
| *cis*-Regulatory element duplication | | [25,26] |
| Transposable element insertion | | [27–29] |
| **Mechanisms of *cis*-regulatory element modification** | | |
| Changes to transcription factor binding sites of an element | Gain | [30,31] |
| | Loss | [30,32,33] |
| | Affinity | [9,34,35] |
| | Arrangement (spacing, order, orientation) | [36–38] |
| Changes to neighboring sequences | Positional effects | [39,40] |
| | Direct and indirect interactions with other elements | [41,42] |
| | Changes to TAD boundaries | [43,44] |

[a]References include studies of the mechanism affecting gene expression and evolutionary case studies. Only a few key studies are included due to citation limits.

drive expression of a reporter gene across multiple stages of development and across many different tissues [67]. These results are consistent with other studies that show a significant proportion of random sequences can be sufficient to drive reporter gene expression, and that many more are only one mutation away from driving expression [69–72].

The results described in the preceding text suggest that the potential for novel variation scales inversely with the age of an enhancer. We do not know, however, how generalizable this is across *cis*-regulatory elements. For instance, we might also hypothesize that less developmentally essential enhancers could have more latitude to generate variation than highly essential enhancers. Indeed, there are observational data to suggest that some *cis*-regulatory elements are more constrained than others: enhancers that are deeply conserved in sequence are often involved in embryonic development [73]. More studies of different types of elements that regulate different types of genes in different systems can help determine whether there are predictable patterns underlying the genesis of novel variation.

## Expansion of *cis*-regulatory material can expand opportunities for evolution

Quite a bit of correlative evidence suggests that the birth and expansion of *cis*-regulatory elements plays a large role in phenotypic evolution. The more noncoding DNA surrounds a gene, the more developmental roles it has [74]. The longer an enhancer is, the more cell types it is active in [75]. The more noncoding DNA around the neuronal genes of a taxon, the more complex its nervous system [76,77]. This all suggests that old genes gain new roles by adding either new *cis*-regulatory elements or new binding sites to ancestral elements.

Despite these general patterns, we still expect co-option of ancestral elements to be involved in phenotypic innovation, even when those innovations are associated with the addition of novel *cis*-regulatory elements. This is because new and old elements for the same gene cannot be assumed to be acting independently of each other. *cis*-Regulatory elements interact in different combinations to form transient, dynamic 'hubs' of elements that facilitate transcription [78,79]. As we discuss in the next section, this **interdependence** among elements is likely more common than previously thought. Artificially placing enhancers from different genes next to each

other can generate novel expression domains that neither enhancer could individually produce [80]. In a similar way, new elements could interact with old elements to expand the repertoire of expression domains of a gene, co-opting both the gene and its ancestral *cis*-regulatory elements.

## Evolution and modularity: when do we expect autonomy versus interdependence of *cis*-regulatory elements?

**Modularity** of *cis*-regulatory elements, specifically enhancers, has long been a defining paradigm in evolutionary developmental biology [1]. While 'modularity' has many definitions, here we refer to the common definition that describes individual enhancers as being independently responsible for highly specific expression domains [1,81–83] (Box 2). Mutation in one modular enhancer should only affect one modular trait. Enhancer modularity has been suggested to explain why morphological evolution occurs more often in *cis* than in *trans,* and that systems composed of modular enhancers could be more evolvable due to the reduction in **pleiotropy** [84].

### The modularity paradigm emerged from reporter construct studies

The modularity paradigm derived from, and is supported by, a long history of reporter construct experiments, which demonstrate that many enhancers are individually sufficient to drive gene expression in specific domains [85–88] (Box 1). By contrast, enhancer deletion experiments have thus far largely failed to show correspondingly specific effects on development [89] – many deletions affect multiple traits [50,90,91]. These deletion studies remain limited in number, but they suggest that both interdependence and pleiotropy are both common attributes of *cis*-regulatory architectures (Box 2). How do we reconcile these apparently conflicting observations of *cis*-regulatory element specificity? Some possibilities include:

(i) Secondary effects of deletions on interactions between elements: individual elements might encode tissue-specific information and yet be functionally pleiotropic due to physical interactions with other elements or by acting as necessary spacers between elements [79,92,93]. For example, deletion of one element may affect the proper function of other elements by disrupting the 3D structure of interacting elements that controls gene expression;

(ii) Initiation versus maintenance of gene expression: deletion experiments could be targeting maintenance elements, which are elements that cause an expression state (i.e., activated or repressed) to persist, such as by binding Trithorax group and Polycomb group proteins [94–96]. Deletion of these elements would disrupt normal expression across multiple developmental contexts. Meanwhile, elements that initiate the expression state might still do so in a context-specific and autonomous manner [97];

(iii) Ascertainment bias: experimental approaches coupled with peculiarities of particular model systems may have caused us to favor one model over the other. For example, the modularity model relies primarily on reporter construct data, which are problematic because (1) they only test for **sufficiency** not **necessity** (i.e., they demonstrate only the individual potential of a *cis*-

### Box 2. How do we characterize *cis*-regulatory element modularity?

Modular enhancers are often defined by functional autonomy, where they are individually sufficient to drive expression in only one specific spatial domain. Both independence from other elements and context specificity can be plotted on spectra as in Figure I, such that individual elements have different degrees of modularity in both structure (dependence on interactions with other elements) and function (number of different expression domains regulated during development). In this way, *cis*-regulatory elements can vary in their degree of modularity with respect to both their specificity to individual contexts and their autonomy to drive expression without interacting with other elements. These need not vary together: some elements could be highly autonomous but nonspecific, whereas others could be specific to only one context but not sufficient to drive that expression domain without other elements (Figure I).

To understand how the degree of modularity varies across different *cis*-regulatory elements and how this influences evolution, both high-throughput screens, such as massively parallel reporter assays and mutation libraries, and detailed case studies of the *cis*-regulatory architectures of individual genes are necessary.

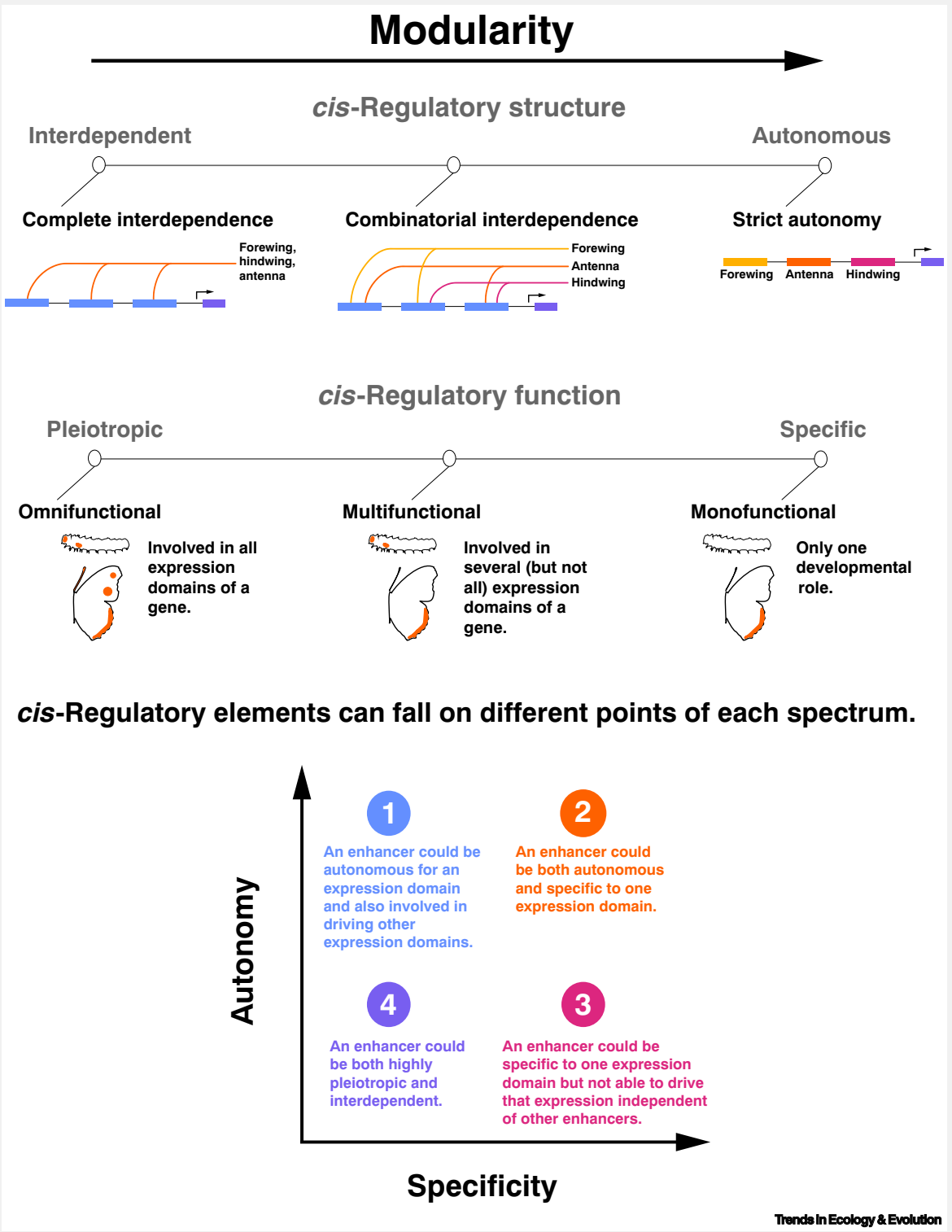


Figure I. The modularity of a *cis*-regulatory element is measured by its autonomy and specificity.

regulatory element to regulate gene expression, not its necessity for organismal development) (Box 1); (2) they cannot easily detect negative effects on gene regulation (i.e., **silencer** elements); and (3) investigators rarely look at multiple time points or tissues when they are focused on specific traits, thus pleiotropic effects may go undetected (Box 2); and

(iv) Inconsistent granularity: There has been inconsistent granularity in *cis*-regulatory assays across case studies, whereby many reporter and deletion experiments focus on long regions that may contain multiple distinct *cis*-regulatory elements. In the context of reporters, interaction between multiple adjacent enhancers and/or silencers in a long construct could produce refined, modular-appearing expression patterns. Conversely, induced deletions spanning multiple elements could produce more pleiotropic phenotypes, where multiple traits are affected [63].

We do not know enough to say how important any of the aforementioned effects are in terms of informing generalities; nevertheless, we cannot rule out enhancer pleiotropy. There are cases of single transcription factor binding sites, and even single nucleotides, affecting the development of multiple traits, so not all pleiotropic enhancers can be easily divided into modular subelements [98,99], and highly granular, systematic studies in *Drosophila* have shown that enhancer pleiotropy is prevalent [8,9,100,101].

### Evolutionary consequences of *cis*-regulatory element pleiotropy

If most enhancers are not as rigidly trait-specific or autonomous as previously thought, we must then consider the evolutionary consequences of *cis*-regulatory pleiotropy. We illustrate the *cis*-regulatory architecture of a gene using two related spectra of modularity: the level of interdependence among elements and the level of pleiotropy of individual elements (Box 2). We propose that both of these attributes influence the evolvability of a gene's expression. For example, the more pleiotropic an element is, the more likely it is that a mutation will have deleterious effects. Conversely, functionally autonomous elements with highly specific effects will be less evolutionarily constrained.

Beyond simply being constrained, however, highly interdependent pleiotropic architectures could be pre-adapted to achieve a higher level of precision in timing and spatial activation than simpler systems [102,103]. A combinatorial system of interdependent elements could also realize more 'higher-order' patterns than a simple system of discrete modular elements. That is, there could be a greater diversity of nuanced gene expression patterns that emerge from groupings of interacting elements, compared with what could be generated by an array of monofunctional, independently-acting elements. More complex interactive structures could expand the capacity of a gene's *cis*-regulatory region to generate new patterns of expression.

### Do enhancers evolve to become more or less specialized over time?

Traditional models of modular enhancer evolution usually assume an additive process, in which new expression domains are added via the *de novo* evolution of new autonomous elements [82]. However, to our knowledge, there is little evidence that *de novo* enhancers emerge specialized and sufficient to activate expression in discrete, highly specific spatial domains. One might instead predict that enhancers initially drive expression quite broadly, like random sequences often do when inserted into reporter constructs [67]. This prediction may not apply to some developmental genes, for which broad expression is likely deleterious. Nevertheless, across all enhancers of all ages, most are active in multiple contexts [8,75].

What selective processes could drive these broad-acting or pleiotropic enhancers to become more context-specific? Selection could favor an enhancer allele because of an advantageous

effect on one trait despite weak deleterious effects on another trait, resulting in antagonistic pleiotropy. One way of resolving this antagonism would involve degeneration of the enhancer's effect on the negatively affected trait, thus leading to increased specificity. Context-specific *cis*-regulatory elements could also evolve as a result of selection for **robustness**. For example, redundant shadow enhancers might be selected for under environmental perturbation because they confer higher developmental robustness [104]. In more stable environments, these enhancers would be released from constraints and could become specialized for separate developmental roles depending on the population size and rates of binding site gain and loss through a process similar to the subfunctionalization of duplicate genes [61]. Neutral evolution alone could also produce context-specific elements. Many genes are regulated by *cis*-regulatory elements that have been duplicated, and these duplicate elements could similarly diverge to become specialized [83,105,106].

It remains unknown whether there is a trend for existing *cis*-regulatory elements to become more or less multifunctional over time. Are regulatory elements like most genes, repeatedly undergoing cooption to progressively accumulate new functions over time? Or do they originate as relatively nonspecific, and become progressively more specialized? Or perhaps there is a more nuanced, yet predictable, arc in the typical lifespan of a *cis*-regulatory element (e.g., broad → specialized → pleiotropic). With new methods now available in comparative systems, we are poised to address these kinds of questions.

## Regulatory fragility, robustness, and the tempo of evolution

The discovery that deleting one enhancer can affect multiple phenotypes, sometimes drastically, challenges more than one paradigm in regulatory genomics. It should not be so trivial to experimentally crack nature's molds if *cis*-regulatory architecture is robust. Deletions as small as 18 base pairs of noncoding DNA have significant effects on the pigmentation of a butterfly, for example [107]. On the other hand, deletions of enhancers of 1 kb or more have no noticeable effect on mouse development [108]. Clearly, some *cis*-regulatory architectures are more robust to mutation than others. We propose that this variation in robustness could correspond to variation in evolvability.

To understand the consequences of *cis*-regulatory robustness, we should first consider the mechanisms by which it can be achieved. Robustness to mutational or environmental perturbation can be conferred by functional redundancy both within and between *cis*-regulatory elements. This redundancy is typically manifested by clustering repeats of certain transcription factor binding sites [104,108–115]. *cis*-Regulatory elements can also tolerate many single nucleotide substitutions because most transcription factors can tolerate a certain amount of flexibility in binding site sequence motifs, and *cis*-regulatory elements can sometimes maintain the same function despite changes in the arrangement of binding sites [37,116]. Some enhancers are even robust to insertion of **insulators** and are still able to interact with their target gene's promoter to drive expression [117]. Most of these studies of robustness have been conducted only within the past decade, and we do not yet have a good sense of how generalizable they are or whether these mechanisms contribute to the developmental buffering of some classes of traits more than others.

We can make some predictions regarding the robustness and fragility of traits based on the emerging case studies we do have, however. Surprisingly, deletion of highly conserved elements often has no detectable effect on development (e.g., vertebrate embryogenesis), while deletion of other elements that are diverged in sequence across species can have pronounced morphological effects (e.g., butterfly wing patterns) [50,108]. These results are consistent with the

expectation that networks under stabilizing selection for a long period of time should become less sensitive to mutation, unlike those that are evolving rapidly under strong positive selection [118]. Therefore, we predict that traits with more robust *cis*-regulatory architectures are more constrained in their evolution because genetic variation will generate less phenotypic variation that can be selected upon, while traits that are evolving rapidly are likely to have more easily mutable regulatory architectures with *cis*-regulatory elements that are more sensitive to mutation [119].

### Can robustness beget novelty?

A simple negative relationship between robustness and rate of trait evolution makes sense, but it may also appear to contradict our previous suggestion that redundant enhancers serve as a source of innovation. We propose, however, that this robustness-to-novelty effect could be an important mode of *cis*-regulatory evolution, since it simultaneously buffers an ancestral trait while providing the raw material for variations on a regulatory theme or even completely novel expression domains. We also note that robust traits could accumulate cryptic genetic variation that could later be selected upon [120]. Thus, robust regulatory architectures could also provide 'safe' vessels in which regulatory variation can accumulate, and from which novelty can spring.

## Concluding remarks

We conclude with more questions than answers, in part because many recent mechanistic studies show that *cis*-regulatory elements are much more complex than previously appreciated. *cis*-Regulatory elements interact with each other in dynamic assemblages and often have multiple functional roles and expression domains, rather than single specific ones. Homology cannot be reliably inferred from sequence similarity; covertly homologous elements may be found in abundance. We have few data-informed models of the general evolutionary trajectories of enhancers, and whether they tend to become more or less autonomous over time. Some highly conserved elements are redundant, while others are sensitive to small mutations. These findings raise many new questions for evolutionary biologists (see Outstanding questions).

These new questions will need to be addressed through the synthesis of detailed investigations of *cis*-regulatory evolution across populations and species. Specifically, a combination of methods, including comparative **chromatin** annotations [e.g., assay for transposase-accessible chromatin sequencing (ATAC-seq) and chromatin immunoprecipitation sequencing (ChIP-seq)], targeted deletions, and reporter constructs (Box 1), will be necessary to measure and detect patterns in the relative degree of modularity of different elements. Since reporter construct experiments are easier than deletions in some systems, and the converse is true in others, there have been different estimates of the prevalence of *cis*-regulatory pleiotropy from different studies. Consistently pairing both approaches, to assess both necessity and sufficiency in the same experimental system, could help resolve many outstanding questions. Furthermore, the function of a *cis*-regulatory element can elude detection by deletion experiments if it is redundant. Therefore, accurately describing the function and robustness of a *cis*-regulatory architecture will sometimes require systematic perturbation of many regions within a large locus. Although conducting this level of detailed investigation across species and populations will be labor intensive, it will also shed new light on *cis*-regulatory evolution that neither comparative sequencing studies nor single-species mechanistic studies could do alone.

By focusing this review specifically on *cis*-regulatory elements, we could not cover the full breadth of regulatory evolution. Alternative levels of regulation (e.g., splice variation, transcript untranslated regions, etc.) have been comparatively neglected in evolutionary biology [121], although that is beginning to change [122]. In general, further integration of molecular genetics and

### Outstanding questions

To what extent do new phenotypes evolve via changes to ancestral regulatory elements versus the birth of new elements?

Are certain types of *cis*-regulatory element more likely to be co-opted for new roles, rather than constrained to specific, specialized roles?

Over time, do *cis*-regulatory elements tend to evolve toward specificity or multifunctionality?

How does the relative robustness of the *cis*-regulatory architecture of a trait constrain or facilitate phenotypic evolution?

evolutionary biology will benefit each field. Just as studies of specific mechanisms inform our predictions for *cis*-regulatory evolution here, studies of variation in development will also inform our understanding of how *cis*-regulatory elements work.

## Acknowledgments

The authors would like to thank three anonymous reviewers, Leslie Babonis, Ben McCormick, and Kevin McDonald for feedback on the manuscript. This work was supported by NSF GRFP DGE-1650441 to J.M.C.M. and NSF grants IOS-2128164 and DEB-2242865 to R.D.R.

## Declaration of interests

The authors have no interests to declare.

## References

1. Carroll, S.B. (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134, 25–36
2. Wittkopp, P.J. and Kalay, G. (2011) Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13, 59–69
3. McGregor, A.P. *et al.* (2007) Morphological evolution through multiple cis-regulatory mutations at a single gene. *Nature* 448, 587–590
4. Reed, R.D. *et al.* (2011) Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333, 1137–1141
5. Chan, Y.F. *et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327, 302–305
6. Prud'homme, B. *et al.* (2007) Emerging principles of regulatory evolution. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8605–8612
7. Prud'homme, B. *et al.* (2006) Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* 440, 1050–1053
8. Sabaris, G. *et al.* (2019) Actors with multiple roles: pleiotropic enhancers and the paradigm of enhancer modularity. *Trends Genet.* 35, 423–433
9. Fuqua, T. *et al.* (2020) Dense and pleiotropic regulatory information in a developmental enhancer. *Nature* 587, 235–239
10. Wong, E.S. *et al.* (2020) Deep conservation of the enhancer regulatory code in animals. *Science* 370, eaax8137
11. Jareborg, N. *et al.* (1999) Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. *Genome Res.* 9, 815–824
12. Shabalina, S.A. and Kondrashov, A.S. (1999) Pattern of selective constraint in *C. elegans* and *C. briggsae* genomes. *Genet. Res.* 74, 23–30
13. Andrews, G. *et al.* (2023) Mammalian evolution of human cis-regulatory elements and transcription factor binding sites. *Science* 380, eabn7930
14. Bernstein, B.E. *et al.* (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169–181
15. Yue, F. *et al.* (2014) A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515, 355–364
16. Yousaf, A. *et al.* (2015) The evolution of bony vertebrate enhancers at odds with their coding sequence landscape. *Genome Biol. Evol.* 7, 2333–2343
17. Bradley, R.K. *et al.* (2010) Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related *Drosophila* species. *PLoS Biol.* 8, e1000343
18. Paris, M. *et al.* (2013) Extensive divergence of transcription factor binding in *Drosophila* embryos with highly conserved gene expression. *PLoS Genet.* 9, e1003748
19. Odom, D.T. *et al.* (2007) Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat. Genet.* 39, 730–732
20. Borneman, A.R. *et al.* (2007) Divergence of transcription factor binding sites across related yeast species. *Science* 317, 815–819
21. Schmidt, D. *et al.* (2010) Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* 328, 1036–1040
22. Prasad, N. *et al.* (2016) A comparative genomic analysis of targets of Hox protein Ultrabithorax amongst distant insect species. *Sci. Rep.* 6, 27885
23. Chen, L. *et al.* (2016) A de novo silencer causes elimination of MITF-M expression and profound hearing loss in pigs. *BMC Biol.* 14, 52
24. Eichenlaub, M.P. and Ettwiller, L. (2011) De novo genesis of enhancers in vertebrates. *PLoS Biol.* 9, e1001188
25. Will, A.J. *et al.* (2017) Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nat. Genet.* 49, 1539–1545
26. Goode, D.K. *et al.* (2011) Minor change, major difference: divergent functions of highly conserved cis-regulatory elements subsequent to whole genome duplication events. *Development* 138, 879–884
27. Ye, M. *et al.* (2020) Specific subfamilies of transposable elements contribute to different domains of T lymphocyte enhancers. *Proc. Natl. Acad. Sci. U. S. A.* 117, 7905–7916
28. Nishihara, H. *et al.* (2016) Coordinately co-opted multiple transposable elements constitute an enhancer for wnt5a expression in the mammalian secondary palate. *PLoS Genet.* 12, e1006380
29. Emera, D. *et al.* (2012) Convergent evolution of endometrial prolactin expression in primates, mice, and elephants through the independent recruitment of transposable elements. *Mol. Biol. Evol.* 29, 239–247
30. Preger-Ben Noon, E. *et al.* (2016) Evolved repression overcomes enhancer robustness. *Dev. Cell* 39, 572–584
31. Ling, L. *et al.* (2023) Increased chromatin accessibility promotes the evolution of a transcriptional silencer in *Drosophila*. *Sci. Adv.* 9, eade6529
32. Jeong, S. *et al.* (2006) Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* 125, 1387–1399
33. Kvon, E.Z. *et al.* (2016) Progressive loss of function in a limb enhancer during snake evolution. *Cell* 167, 633–642
34. Ramos, A.I. and Barolo, S. (2013) Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20130018
35. Lim, F. *et al.* (2022) Affinity-optimizing enhancer variants disrupt development. *Nature* 626, 151–159
36. Yie, J. *et al.* (1999) Mechanism by which the IFN-β enhanceosome activates transcription. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13108–13113
37. Shen, Z. *et al.* (2022) Systematic analysis of naturally occurring insertions and deletions that alter transcription factor spacing identifies tolerant and sensitive transcription factor pairs. *eLife* 11, e70878
38. Georgakopoulos-Soares, I. *et al.* (2023) Transcription factor binding site orientation and order are major drivers of gene regulatory activity. *Nat. Commun.* 14, 2333

39. Bateman, J.R. and Johnson, J.E. (2022) Altering enhancer-promoter linear distance impacts promoter competition in cis and in trans. *Genetics* 222, iyac098
40. Hao, N. *et al.* (2019) Positive and negative control of enhancer-promoter interactions by other DNA loops generates specificity and tunability. *Cell Rep.* 26, 2419–2433
41. Choi, J. *et al.* (2021) Evidence for additive and synergistic action of mammalian enhancers during cell fate determination. *eLife* 10, e65381
42. Cho, S.W. *et al.* (2018) Promoter of lncRNA gene PVT1 is a tumor-suppressor DNA boundary element. *Cell* 173, 1398–1412
43. Krefting, J. *et al.* (2018) Evolutionary stability of topologically associating domains is associated with conserved gene regulation. *BMC Biol.* 16, 87
44. Willemin, A. *et al.* (2021) Induction of a chromatin boundary in vivo upon insertion of a TAD border. *PLoS Genet.* 17, e1009691
45. Simakov, O. *et al.* (2015) Hemichordate genomes and deuterostome origins. *Nature* 527, 459–465
46. Pennacchio, L.A. *et al.* (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444, 499–502
47. Wang, W. *et al.* (2007) Comparison of Pax1/9 locus reveals 500-Myr-old syntenic block and evolutionary conserved non-coding regions. *Mol. Biol. Evol.* 24, 784–791
48. Burgess, D. and Freeling, M. (2014) The most deeply conserved noncoding sequences in plants serve similar functions to those in vertebrates despite large differences in evolutionary rates. *Plant Cell* 26, 946–961
49. Gompel, N. *et al.* (2005) Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433, 481–487
50. Mazo-Vargas, A. *et al.* (2022) Deep cis-regulatory homology of the butterfly wing pattern ground plan. *Science* 378, 304–308
51. Rebeiz, M. *et al.* (2011) Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10036–10043
52. Glassford, W.J. *et al.* (2015) Co-option of an ancestral Hox-regulated network underlies a recently evolved morphological novelty. *Dev. Cell* 34, 520–531
53. Arnoult, L. *et al.* (2013) Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* 339, 1423–1426
54. Signor, S.A. and Nuzhdin, S.V. (2018) The evolution of gene expression in cis and trans. *Trends Genet.* 34, 532–544
55. Taher, L. *et al.* (2011) Genome-wide identification of conserved regulatory function in diverged sequences. *Genome Res.* 21, 1139–1149
56. Yang, S. *et al.* (2015) Functionally conserved enhancers with divergent sequences in distant vertebrates. *BMC Genomics* 16, 882
57. Khoueiry, P. *et al.* (2017) Uncoupling evolutionary changes in DNA sequence, transcription factor occupancy and enhancer activity. *eLife* 6, e28440
58. Weirauch, M.T. and Hughes, T.R. (2010) Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. *Trends Genet.* 26, 66–74
59. Lusk, R.W. and Eisen, M.B. (2010) Evolutionary mirages: selection on binding site composition creates the illusion of conserved grammars in *Drosophila* enhancers. *PLoS Genet.* 6, e1000829
60. Sundaram, V. and Wysocka, J. (2020) Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 375, 20190347
61. Wagner, G.P. (2018) *Homology, Genes, and Evolutionary Innovation*, Princeton University Press
62. Rebeiz, M. and Tsiantis, M. (2017) Enhancer evolution and the origins of morphological novelty. *Curr. Opin. Genet. Dev.* 45, 115–123
63. Murugesan, S.N. and Monteiro, A. (2022) Evolution of modular and pleiotropic enhancers. *J. Exp. Zool. B Mol. Dev. Evol.* 340, 105–115
64. Dorshorst, B. *et al.* (2015) A genomic duplication is associated with ectopic eomesodermin expression in the embryonic chicken comb and two duplex-comb phenotypes. *PLoS Genet.* 11, e1004947
65. Schrader, L. and Schmitz, J. (2019) The impact of transposable elements in adaptive evolution. *Mol. Ecol.* 28, 1537–1549
66. Indjeian, V.B. *et al.* (2016) Evolving new skeletal traits by cis-regulatory changes in bone morphogenetic proteins. *Cell* 164, 45–56
67. Galupa, R. *et al.* (2023) Enhancer architecture and chromatin accessibility constrain phenotypic space during *Drosophila* development. *Dev. Cell* 58, 51–62
68. Li, X.C. *et al.* (2023) Mutational scans reveal differential evolvability of *Drosophila* promoters and enhancers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 378, 20220054
69. Smith, R.P. *et al.* (2013) A compact, in vivo screen of all 6-mers reveals drivers of tissue-specific expression and guides synthetic regulatory element design. *Genome Biol.* 14, R72
70. Yona, A.H. *et al.* (2018) Random sequences rapidly evolve into de novo promoters. *Nat. Commun.* 9, 1530
71. Arnold, C.D. *et al.* (2014) Quantitative genome-wide enhancer activity maps for five Drosophila species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat. Genet.* 46, 685–692
72. de Boer, C.G. *et al.* (2020) Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. *Nat. Biotechnol.* 38, 56–65
73. Plessy, C. *et al.* (2005) Enhancer sequence conservation between vertebrates is favoured in developmental regulator genes. *Trends Genet.* 21, 207–210
74. Nelson, C.E. *et al.* (2004) The regulatory content of intergenic DNA shapes genome architecture. *Genome Biol.* 5, R25
75. Laiker, I. and Frankel, N. (2022) Pleiotropic enhancers are ubiquitous regulatory elements in the human genome. *Genome Biol. Evol.* 14, evac071
76. Taft, R.J. *et al.* (2007) The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* 29, 288–299
77. Closser, M. *et al.* (2022) An expansion of the non-coding genome and its regulatory potential underlies vertebrate neuronal diversity. *Neuron* 110, 70–85
78. Furlong, E.E.M. and Levine, M. (2018) Developmental enhancers and chromosome topology. *Science* 361, 1341–1345
79. Lim, B. and Levine, M.S. (2021) Enhancer-promoter communication: hubs or loops? *Curr. Opin. Genet. Dev.* 67, 5–9
80. Shermoen, A.W. *et al.* (1987) Developmental regulation by an enhancer from the Sgs-4 gene of *Drosophila*. *EMBO J.* 6, 207–214
81. Koshikawa, S. (2015) Enhancer modularity and the evolution of new traits. *Fly* 9, 155–159
82. Wilkins, A.S. (2002) *The Evolution of Developmental Pathways*, Oxford University Press
83. Lynch, M. (2007) *The Origins of Genome Architecture*, Oxford University Press
84. Stern, D.L. and Orgogozo, V. (2008) The loci of evolution: how predictable is genetic evolution? *Evolution* 62, 2155–2177
85. Rebeiz, M. *et al.* (2009) Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326, 1663–1667
86. Koshikawa, S. *et al.* (2015) Gain of cis-regulatory activities underlies novel domains of wingless gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 112, 7524–7529
87. Rice, G.R. *et al.* (2019) Modular tissue-specific regulation of doublesex underpins sexually dimorphic development in *Drosophila*. *Development* 146, dev178285
88. Williams, T.M. *et al.* (2008) The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* 134, 610–623
89. Sagai, T. *et al.* (2005) Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* 132, 797–803
90. Erickson, P.A. *et al.* (2015) A 190 base pair, TGF-β responsive tooth and fin enhancer is required for stickleback Bmp6 expression. *Dev. Biol.* 401, 310–323
91. Lewis, J.J. *et al.* (2019) Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. *Proc. Natl. Acad. Sci. U. S. A.* 116, 24174–24183

92. Scholes, C. *et al.* (2019) Signal integration by shadow enhancers and enhancer duplications varies across the *Drosophila* embryo. *Cell Rep.* 26, 2407–2418
93. Small, S. *et al.* (1993) Spacing ensures autonomous expression of different stripe enhancers in the *even-skipped* promoter. *Development* 119, 767–772
94. Chan, C.S. *et al.* (1994) A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *EMBO J.* 13, 2553–2564
95. Kassis, J.A. and Brown, J.L. (2013) Polycomb group response elements in *Drosophila* and vertebrates. *Adv. Genet.* 81, 83–118
96. Schuettengruber, B. *et al.* (2011) Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* 12, 799–814
97. Pfeffer, P.L. *et al.* (2002) The activation and maintenance of Pax2 expression at the mid-hindbrain boundary is controlled by separate enhancers. *Development* 129, 307–318
98. Nagy, O. *et al.* (2018) Correlated evolution of two copulatory organs via a single cis-regulatory nucleotide change. *Curr. Biol.* 28, 3450–3457
99. Rowley, A.J. *et al.* (2022) Site pleiotropy of a stickleback Bmp6 enhancer. *Dev. Biol.* 492, 111–118
100. Preger-Ben Noon, E. *et al.* (2018) Comprehensive analysis of a cis-regulatory region reveals pleiotropy in enhancer function. *Cell Rep.* 22, 3021–3031
101. Weaver, L.N. *et al.* (2020) Analysis of Gal4 expression patterns in adult *Drosophila* females. *G3: Genes Genomes Genet.* 10, 4147–4158
102. Montavon, T. *et al.* (2011) A regulatory archipelago controls Hox genes transcription in digits. *Cell* 147, 1132–1145
103. Dickinson, W.J. (1988) On the architecture of regulatory systems: evolutionary insights and implications. *Bioessays* 8, 204–208
104. Frankel, N. *et al.* (2010) Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* 466, 490–493
105. Gao, X. and Lynch, M. (2009) Ubiquitous internal gene duplication and intron creation in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 106, 20818–20823
106. Monteiro, A. and Gupta, M.D. (2016) Identifying coopted networks and causative mutations in the origin of novel complex traits. *Curr. Top. Dev. Biol.* 119, 205–226
107. Fandino, R.A. *et al.* (2024) The *ivory* lncRNA regulates seasonal color patterns in buckeye butterflies. *bioRxiv*, Published online February 19, 2024. https://doi.org/10.1101/2024.02.09.579733
108. Osterwalder, M. *et al.* (2018) Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* 554, 239–243
109. Somma, M.P. *et al.* (1991) The housekeeping promoter from the mouse CpG island HTF9 contains multiple protein-binding elements that are functionally redundant. *Nucleic Acids Res.* 19, 2817–2824
110. Duncan, D.D. *et al.* (1996) Asymmetric redundancy in CD4 silencer function. *Immunity* 4, 301–311
111. Lopez-Rivera, F. *et al.* (2020) A mutation in the *Drosophila melanogaster* eve stripe 2 minimal enhancer is buffered by flanking sequences. *G3 Genes Genomes Genet.* 10, 4473–4482
112. Anania, C. *et al.* (2022) In vivo dissection of a clustered-CTCF domain boundary reveals developmental principles of regulatory insulation. *Nat. Genet.* 54, 1026–1036
113. Crocker, J. *et al.* (2015) Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell* 160, 191–203
114. Xiong, N. *et al.* (2002) Redundant and unique roles of two enhancer elements in the TCRγ locus in gene regulation and γδ T cell development. *Immunity* 16, 453–463
115. Tsai, A. *et al.* (2019) Multi-enhancer transcriptional hubs confer phenotypic robustness. *eLife* 8, e45325
116. Inukai, S. *et al.* (2017) Transcription factor–DNA binding: beyond binding site motifs. *Curr. Opin. Genet. Dev.* 43, 110–119
117. Chakraborty, S. *et al.* (2023) Enhancer-promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness. *Nat. Genet.* 55, 280–290
118. Wagner, A. (1996) Does evolutionary plasticity evolve? *Evolution* 50, 1008
119. McDonald, J.M.C. and Reed, R.D. (2023) Patterns of selection across gene regulatory networks. *Semin. Cell Dev. Biol.* 145, 60–67
120. Félix, M.-A. and Wagner, A. (2008) Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity* 100, 132–140
121. Alonso, C.R. and Wilkins, A.S. (2005) The molecular elements that underlie developmental evolution. *Nat. Rev. Genet.* 6, 709–715
122. Verta, J.-P. and Jacobs, A. (2022) The role of alternative splicing in adaptation and evolution. *Trends Ecol. Evol.* 37, 299–308
123. Gisselbrecht, S.S. *et al.* (2020) Transcriptional silencers in *Drosophila* serve a dual role as transcriptional enhancers in alternate cellular contexts. *Mol. Cell* 77, 324–337
124. Bessis, A. *et al.* (1997) The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. *Proc. Natl. Acad. Sci. U. S. A.* 94, 5906–5911
125. Uhl, J.D. *et al.* (2016) A Hox transcription factor collective binds a highly conserved Distal-less cis-regulatory module to generate robust transcriptional outcomes. *PLoS Genet.* 12, e1005981
126. Kowalczyk, M.S. *et al.* (2012) Intragenic enhancers act as alternative promoters. *Mol. Cell* 45, 447–458
127. Andersson, R. (2015) Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *Bioessays* 37, 314–323
128. Chopra, V.S. *et al.* (2009) Stalled Hox promoters as chromosomal boundaries. *Genes Dev.* 23, 1505–1509
129. Bozhilov, Y.K. *et al.* (2021) A gain-of-function single nucleotide variant creates a new promoter which acts as an orientation-dependent enhancer-blocker. *Nat. Commun.* 12, 3806
130. Raab, J.R. and Kamakaka, R.T. (2010) Insulators and promoters: closer than we think. *Nat. Rev. Genet.* 11, 439–446
131. Kvon, E.Z. (2015) Using transgenic reporter assays to functionally characterize enhancers in animals. *Genomics* 106, 185–192
132. Ma, S. and Zhang, Y. (2020) Profiling chromatin regulatory landscape: insights into the development of ChIP-seq and ATAC-seq. *Mol. Biomed.* 1, 9
133. Lopes, R. *et al.* (2016) Applying CRISPR-Cas9 tools to identify and characterize transcriptional enhancers. *Nat. Rev. Mol. Cell Biol.* 17, 597–604