



## Introduction

**Cite this article:** Rubinstein M, de Souza FSJ. 2013 Evolution of transcriptional enhancers and animal diversity. *Phil Trans R Soc B* 368: 20130017.  
<http://dx.doi.org/10.1098/rstb.2013.0017>

One contribution of 12 to a Theme Issue 'Molecular and functional evolution of transcriptional enhancers in animals'.

### Subject Areas:

evolution, genetics, molecular biology

### Keywords:

evolution, gene expression, enhancer, transgenic animals

### Authors for correspondence:

Marcelo Rubinstein  
e-mail: [mrubins@dna.uba.ar](mailto:mrubins@dna.uba.ar)  
Flávio S. J. de Souza  
e-mail: [fsouza.ingebi@gmail.com](mailto:fsouza.ingebi@gmail.com)

# Evolution of transcriptional enhancers and animal diversity

Marcelo Rubinstein<sup>1,2</sup> and Flávio S. J. de Souza<sup>1,2</sup>

<sup>1</sup>Instituto de Investigaciones en Ingeniería Genética y Biología Molecular, Consejo Nacional de Investigaciones Científicas y Técnicas, C1428ADN Buenos Aires, Argentina

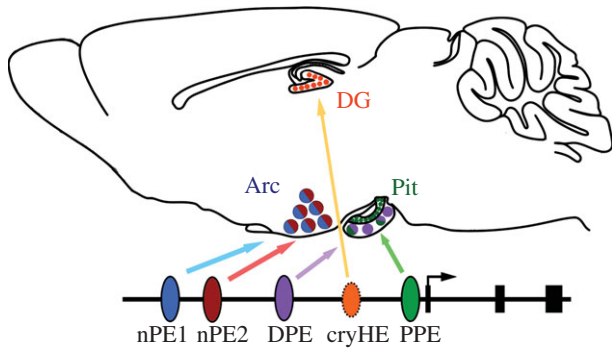
<sup>2</sup>Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, C1428EGA Buenos Aires, Argentina

Deciphering the genetic bases that drive animal diversity is one of the major challenges of modern biology. Although four decades ago it was proposed that animal evolution was mainly driven by changes in *cis*-regulatory DNA elements controlling gene expression rather than in protein-coding sequences, only now are powerful bioinformatics and experimental approaches available to accelerate studies into how the evolution of transcriptional enhancers contributes to novel forms and functions. In the introduction to this Theme Issue, we start by defining the general properties of transcriptional enhancers, such as modularity and the coexistence of tight sequence conservation with transcription factor-binding site shuffling as different mechanisms that maintain the enhancer grammar over evolutionary time. We discuss past and current methods used to identify cell-type-specific enhancers and provide examples of how enhancers originate *de novo*, change and are lost in particular lineages. We then focus in the central part of this Theme Issue on analysing examples of how the molecular evolution of enhancers may change form and function. Throughout this introduction, we present the main findings of the articles, reviews and perspectives contributed to this Theme Issue that together illustrate some of the great advances and current frontiers in the field.

## 1. Introduction

In 1859, the publication of Darwin's *The Origin of Species* provided a powerful natural explanation as to how the *endless forms most beautiful and most wonderful* have been created on this planet. After 150 years of advances in genetics and the molecular principles of heredity, the molecular basis of life diversification can now be understood in its general terms. In the past decade, whole-genome sequencing has allowed direct observation of the genetic changes that separate related species in unprecedented breadth, helping us to understand how genetic variation is generated and opening up the possibility of grasping what genetic changes make, for example, an elephant different from a mouse.

Confirming pioneering observations [1], whole-genome comparisons show that protein-coding sequences and repertoire do not vary much between related organisms, with the exception of proteins related to functions like olfaction, reproduction and immune defence [2,3]. Coding exons are embedded in a sea of intronic and intergenic non-coding sequence, the vast majority of which is devoid of specific functions and constitutes 'junk' DNA. However, the non-coding part of the genome also includes functional regulatory regions, like enhancers. Transcriptional enhancers determine where, when and how much a protein-coding gene is expressed in every animal tissue. Because they encode such critical spatio-temporal and quantitative information, it is expected that their sequences are under strong purifying selection and mutate at a slower rate than flanking neutrally evolving regions. Even though enhancers tend to be evolutionarily conserved, in general they evolve faster than coding regions [4,5], suggesting that changes in regulatory DNA play an important role in evolution. Although this does not mean that physiological and morphological changes cannot be caused by mutations in coding exons [6,7], many characteristics of enhancers and other *cis*-regulatory regions indicate that organismal evolution is mostly driven by changes in gene regulation, as theorized over 40 years



**Figure 1.** Schematic of the modular organization of the expression of mouse *Pomc*. The neuronal enhancers nPE1 (blue) and nPE2 (red) control expression in neurons of the hypothalamic arcuate nucleus (Arc). The proximal pituitary enhancer (PPE; green) drives expression to melanotrophs and corticotrophs of the pituitary (Pit), whereas the distal pituitary enhancer (DPE; purple) is active in pituitary corticotrophs. A cryptic hippocampal enhancer, cryHE (orange), drives reporter gene expression in newborn neurons of the dentate gyrus (DG) of the hippocampus in transgenic mice. *Pomc* exons are in black boxes.

ago [1,8,9]. In fact, as we shall see in §§6 and 7, evidence has accumulated in recent years showing that mutations in regulatory regions are an important source of evolutionary innovation [10–13]. Although the relatively rigid genetic code that determines the amino acid sequence of proteins was deciphered soon after the discovery of DNA structure, the regulatory code of the genome still remains an unsolved mystery and, as such, understanding the complex interplay between regulatory DNA regions and transcription factors (TF), regulatory RNAs, signalling pathways and epigenetic marks that determine gene expression is among the main current challenges of modern molecular genetics. As our knowledge of the regulatory code progresses, the closer we are to understanding how evolution operates at the molecular level to bring about morphological and physiological innovation. This field of research, years ago mainly supported by those interested in the basic mechanisms of gene expression regulation, is experiencing a vigorous growth in medical and human genetics departments interested in understanding human disease, as developmental defects, cancer and other conditions are more and more found to be caused by mutations in enhancers [14–20].

In the introduction to this Theme Issue on *Enhancers Evolution and Animal Diversity*, we shall review some general principles concerning how changes in transcriptional enhancers happen over time and particularly how they can lead to phenotypical innovation. Some principles will be illustrated with our studies on the regulation of the proopiomelanocortin gene (*Pomc*; see figure 1 for a schematic of the structure of mouse *Pomc* and its regulatory regions). *Pomc* encodes a prohormone expressed in the arcuate nucleus of the hypothalamus and in the corticotropes and melanotropes of the pituitary which plays critical roles in the control of energy balance and stress response in vertebrates [21].

## 2. Enhancers: general considerations

In animals, enhancers are one of the main types of transcriptional regulatory regions, others being promoters, promoter-tethering elements, locus control regions, silencers, barrier elements and insulators. Enhancers are *cis*-acting segments of DNA,

usually mapped down to 200–500 bp, that control expression of nearby genes. Enhancer activity is thought to depend on the long-range communication between the enhancer region and the promoter [22], which is achieved by DNA looping mediated by specific proteins like cohesin and the Mediator complex [23], as well as relocation of the active gene from the periphery to the interior of the nucleus [24]. Identified enhancers are often located in the vicinity of the genes they control, but some have been found up to 1 Mb away [14], even within introns of other genes (for a recent excellent collection of papers on distal enhancers, see the Discussion Meeting Issue; [25]).

Since their discovery over 30 years ago [26], enhancers have been found to harbour several transcription factor-binding sites (TFBS) in a particular spatial order, defining what can be called the enhancer grammar. Detailed functional dissection of different enhancers led to the development of two extreme models of enhancer organization, the (i) *enhanceosome* and the (ii) *billboard* models [27]. The enhanceosome model is derived from work on an enhancer that triggers *interferon-β* (*IFN-β*) transcription in response to viral infection. Exhaustive analyses of this enhancer found that it is composed of a tight ensemble of binding sites for several TF that need to be present in a precise order and spacing, as even minor mutations that prevent binding or alter the distance between bound TF cause the enhancer not to function at all. TF in the enhanceosome thus work synergistically as a unit [28]. By contrast, billboard enhancers are composed of TFBS arranged in a looser and more flexible way such that the removal of a binding site diminishes enhancer performance but does not abolish it completely. The enhancer for stripe 2 of *Drosophila even-skipped* (*eve*), for instance, has an arrangement of TFBS for activators (Bicoid, Hunchback) and repressors (Krüppel, Giant) that, when individually mutated, do not abolish its function but rather change stripe width and intensity in transgenic experiments [29]. For many enhancers analysed in less depth, there is evidence that removal of large chunks of conserved enhancer sequence does not abolish enhancer function, also pointing to a flexible, billboard-like organization, as is the case of the neuronal *Pomc* enhancers nPE2 and nPE1 (figure 1; [30,31]). Few enhancers have been studied at the same depth as those for *IFN-β* and *eve* stripe 2, but it is possible that for most enhancers more rigid and more flexible subsegments coexist. The 362-bp core region of the *sparkling* (*spa*) enhancer of *Pax2*, which drives expression to cone cells in the *Drosophila* compound eye, seems to have an intermediate organization, as spacing and order of TFBS are important for enhancer function, as in the enhanceosome, but at the same time one of the critical regions for *spa* function (region 1) can be relocated without affecting expression [32,33]. A third model of TFBS organization was found in enhancers involved in heart development and differentiation [34]. This alternative (iii) *TF collective* mode of enhancer activity operates via the cooperative recruitment of a large number of cardiogenic TFs to activate enhancers with an apparently lax motif grammar, such that motifs for some factors may be absent and some necessary TF are rather recruited via cooperative protein–protein interactions [34].

Apart from the order and number of TFBS, another important variable in enhancer grammar is TF-binding affinity. Genome-wide analyses indicate that many interactions between TF and DNA are relatively weak, and these are regarded as likely to be non-functional [35] as are also DNA regions bound at low occupancy [36]. However, in this issue, Ramos & Barolo [37] show that a weak interaction

between the TF *Cubitus interruptus* (Ci) and a *decapentaplegic* enhancer in response to Hedgehog signalling is nevertheless very important for the correct interpretation of the Hedgehog gradient in the fly embryo, implying that low-affinity sites can be functional in some circumstances [37]. Enhancer grammar also depends on the interactions between bound TF, and this is reflected by the conservation of distances between TFBS which is necessary to facilitate these interactions. In this issue, Guturu *et al.* [38] take into account DNA sequence and protein structural data to predict regions bound by TF on phylogenetically conserved elements in the genome, generating an original, free database of potential TF complexes for future studies on enhancer organization and function [38].

### 3. Enhancer modularity

During embryonic development and in each tissue of the adult body, different cell types express distinct groups of TF and are exposed to various cell-signalling pathways, giving rise to a unique combinatorial regulatory code that is interpreted by DNA regulatory regions, like enhancers, to ultimately determine whether a particular gene will be on or off in a given cell type. Genes with complex expression patterns have been experimentally shown to have several enhancers, each of which drives expression in particular cell types or time points. This modular organization allows each enhancer to control gene expression in a strict spatial and temporal domain independently of each other and of the basal promoter. For example, the mouse *Sonic hedgehog* homolog gene (*Shh*), which encodes a secreted morphogen, has at least six enhancers controlling its expression in several parts of the neural tube [39], as well as one distal enhancer controlling expression in the limb bud [14,40]. Neuronal enhancers nPE1 and nPE2 drive *Pomc* expression to hypothalamic neurons independently of the pituitary enhancer or promoter, and vice versa (figure 1; [41,42]). The functional modularity of enhancers is best illustrated by targeted mutagenesis experiments. For example, the removal of a distal, limb-specific enhancer of *Shh* causes limb truncations in mutant mice without causing aberrant phenotypes in the notochord or neural tube, which also express *Shh* [40]. It is important to note, however, that enhancers may not be absolutely modular. For instance, some enhancers can synergize with each other, as the case of *Shh* forebrain enhancers in zebrafish [43]. Recently, a thorough characterization of the 200 kb-regulatory landscape of the mouse *Fgf8* gene, which encodes a secreted signalling protein, has shown that the relative position of enhancers in relation to the controlled gene is important to fine-tune the expression pattern of the gene [44]. In any case, enhancer modularity has important evolutionary consequences, as mutations in a particular enhancer will change the expression of a gene in a particular region with no or minor effects in other regions, i.e. with little or no pleiotropic effects associated (see §6). Another important, emerging feature of gene regulation is that many metazoan genes possess more than one enhancer driving partially or completely overlapping expression patterns [45,46]. In flies, many developmental genes carry two enhancers with overlapping functions; the enhancer closer to the gene was called 'primary', whereas the one farther away was named 'shadow' enhancer [47]. The existence of functionally overlapping enhancers is

not restricted to developmental genes. For instance, *Pomc* has two distinct enhancers that drive expression to the same population of hypothalamic neurons [31,42]. Even the expression of *Pomc* in the pituitary, long thought to be driven only by the proximal enhancer and promoter, was recently found to depend also on an enhancer with the same regulatory activity as a proximal enhancer (figure 1; [41]). Functionally overlapping enhancers may confer robustness, buffering gene expression against environmental and genetic disturbances [48,49]. Enhancers with partial redundant activity are also hypothesized to be a potential source of evolutionary novelty, as such systems might be more tolerant to mutations that lead to new expression patterns [47].

### 4. Identification of enhancers and functional studies

Identifying enhancers in genomes is crucial for the understanding of the complexity and mechanisms of gene regulation [50]. Traditionally, regulatory regions have been identified and mapped by cloning candidate sequences upstream of a minimal promoter fused to a reporter gene and testing their transcriptional activity in cell lines or in transgenic organisms. Although more laborious and expensive to make, transgenic animals have the great advantage of providing complete spatio-temporal expression information simultaneously in all tissues and cell types of an overall healthy animal. More and more laboratories map enhancers using 100–200 kb bacterial artificial chromosomes, which allow for the testing of regulatory regions in a context that better resembles the endogenous one, compared with small constructs [51]. Detailed studies using these methods have determined that genes involved in embryonic development are controlled by several enhancers, like the example of the mouse *Shh* gene mentioned in §3 or the chicken *Sox2* gene encoding a TF controlled by at least 11 enhancers arranged in a 50 kb region surrounding the gene [52]. This is not surprising because the expression of developmental genes need to be precisely controlled both spatially and temporally in many different tissues. However, even genes not involved in development may have several enhancers, like *Pomc*, which is controlled by two distal enhancers in the hypothalamus and by one distal and one proximal enhancer in the pituitary (figure 1; [41,42]). These and many other examples indicate that the number of enhancers in every animal genome is substantially larger than the number of protein-coding genes.

Beginning in the early 2000s, the sequencing of the genomes of several species allowed for the identification of enhancers and other non-coding DNA elements by phylogenetic footprinting [53]. This technique is based on the idea that DNA sequences that play a functional role evolve slower than non-functional sequences, which are freer to accumulate neutral mutations [54]. Thus, comparing the genomes of an appropriate set of organisms allows one to identify DNA elements under evolutionary constraint and differentiate them from neutrally evolving sequences. Conserved non-coding sequences (CNE), when experimentally tested, often turn out to display enhancer activity [55–58]. Deeply conserved CNE (i.e. conserved in all vertebrates) are more often found around genes involved in early development, presumably because mutations in enhancers that control the precise regulatory patterns of such genes often cause deleterious developmental phenotypes [59,60].



Recently, the sequencing of 29 mammalian genomes has allowed for the mapping of evolutionary constraint in great detail, showing that 4.2% of mammalian genome sequence is phylogenetically conserved [61]. Within the conserved fraction, around 68% is located in intronic or intragenic regions (i.e. outside exons and promoters), and at least 30% of the conserved fraction overlaps chromatin marks typical of enhancers [61]. Some CNE display extreme levels of conservation (ultraconserved sequences or UCE), reaching 100% sequence identity in mammals [62], and some orthologous CNE have even been found to be present in invertebrates and vertebrates [5,63]. The challenges posed by the study of CNE, in particular the study of their evolutionary dynamics and function, are the topic of a review by Harmston *et al.* in this issue [64], while another review by Maeso *et al.* [65] deals with the challenge of identifying and analysing transphylectic CNE.

In recent years, chromatin immunoprecipitation coupled to microarray hybridization (ChIP-CHIP) or DNA sequencing (ChIP-seq) has allowed for the genome-wide mapping of TFBS and the mapping of chromatin ('epigenetic') marks in the form of specific histone posttranslational modifications [66,67]. Tracing the distribution of such features—as well as regions of open chromatin detected by whole-genome mapping of DNase I hypersensitive sites—can be used to identify potential enhancers at a genome-wide scale [68,69]. Thus, enhancer regions are bound by clusters of TF [70–72] and are often associated with transcriptional cofactors p300/CBP (a histone acetylase) and components of the Mediator complex [23,73]. Active enhancers are also associated with specific histone marks like histone H3 lysine 4 monomethylation (H3K4me1) and histone H3 lysine 27 acetylation (H3K27ac), as well as depletion in H3 lysine 4 trimethylation (H3K4me3), which mark promoters (reviewed in [60,61]). In the human genome, the ENCODE Project identified around 400 000 elements bearing enhancer-like chromatin signatures in the cell lines that were analysed [74], while around 230 000 potential enhancers were found in the mouse genome using similar techniques [75]. The utility of sequence conservation as well as whole-genome ChIP techniques to identify and study the evolution of enhancers is reviewed in this issue by Sakabe & Nobrega [76].

It is important to note, however, that the evidence provided by the genome-wide mapping of genomic features is just an indication of potential enhancer activity and should not be taken as equivalent to regulatory function [77,78]. Indeed, it is expected that the large genome of complex organisms is alive with biochemical activity, including TF binding, with no significant physiological consequences for the organism [77,79–81]. As an indication of this, it has been recently shown that DNA sequences generated at random can display a significant degree of transcriptional regulatory activity in mammalian cells [82], indicating that even expression assays in cells or transgenic models may be deceptive and that more detailed experimental evidence is necessary before assigning enhancer function to a sequence. Around 1 kb upstream of mouse *Pomc*, we have serendipitously found a non-conserved regulatory region which drives reporter expression to the subgranular layer of the dentate gyrus of the hippocampus of transgenic mice (figure 1; [83]) and is being used as a reliable marker of newborn neurons in this brain region [84]. Because *Pomc* is not expressed in the hippocampus, it may well be that this enhancer activity is used by another gene in the vicinity or just represents a

consistent artefact, especially considering that this DNA sequence is not conserved in other mammals.

## 5. Molecular evolution of enhancers

Although purifying selection keeps transcriptional enhancers evolving at a slow pace in relation to non-functional DNA, enhancers do change by the accumulation of point mutations and small insertions and deletions. In addition, new enhancers can appear by chance when random mutations create clusters of TFBS or lost when large deletions eliminate enhancer sequences. Such processes can be driven by natural selection as well as purely neutral mechanisms [85]. Ultimately, turnover of TFBS in enhancers as well as the birth and elimination of enhancers cause the regulatory element repertoire to change over evolutionary time. An indication of this is that while a third of the coding bases are conserved between mammals and amphioxus (a chordate that belongs to a sister group of vertebrates), less than 1% of CNE bases (a proxy for enhancers) are conserved between these groups [5]. Thus, TFBS reshuffling and enhancer turnover seem to be two pervasive mechanisms setting up hurdles on the way towards the discovery of a universal transcriptional regulatory code, in contrast to the straightforward translational code unravelled in the 1960s that allows predicting protein identity directly from DNA sequence. The variety of rules already found in the regulatory code are perplexing, as illustrated by the titles of two recent papers on the subject: the review 'Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same' [86] and the research paper 'Minor change, major difference: divergent functions of highly conserved *cis*-regulatory elements subsequent to whole genome duplication events' [87]. If the former addresses the concept that the regulatory logic of a functional enhancer can be retained in the absence of detectable sequence conservation, the latter indicates that even small mutations present in highly conserved enhancers may lead to profound functional differences. As understanding the transcriptional code and its impact on the evolution of form and function will be more challenging than anticipated, efforts must be redoubled to study how gene expression is modified by enhancer birth, change or loss.

### (a) The birth of an enhancer

Genetic mechanisms leading to the birth of novel enhancers include: (i) insertion of transposable elements (TEs), (ii) *de novo* mutations and (iii) chromosomal rearrangements that promote enhancer adoption. The high contribution of TE-derived sequences to genome composition provides a substantial amount of raw material with the potential to evolve into novel functional *cis*-regulatory elements. Probably the first convincing case was a sequence derived from a HERV-E retroposon that became co-opted as a parotid-specific enhancer of the human salivary  $\alpha$ -amylase 1C gene (AMY1C; [88]). Interestingly, insertion of this TE in the 5' proximal flanking region of a paralogue copy of the ancestral pancreatic AMY2B occurred in the lineage leading to primates [89] which acquired, therefore, the possibility to taste rewarding sweet sugars produced by the enzymatic processing of starch. More recently, the 29 Mammals Project has found more than 280 000 conserved, TE-derived non-coding elements [61,90]. Although these elements carry potential *cis*-regulatory function, exaptation (co-option) of

TE-derived sequences into transcriptional enhancers has been rigorously demonstrated in transgenic animals in only a handful of cases [77]. For example, we have found that the neuronal *Pomc* enhancers nPE1 and nPE2 originated from two independent exaptation events, an example of convergent evolution of transcriptional enhancers [31,91], whereas nPE2 derived from a CORE-SINE retroposon [30] in the lineage leading to mammals, nPE1 is a more recent placental acquisition derived from a mammalian apparent LTR (MaLR) retroposon [31].

Eichenlaub & Ettwiller [92] have recently discovered that enhancers can originate *de novo* after acquiring minor changes in previously existing non-regulatory sequences [92]. By taking advantage of the massive gene loss following the last whole-genome duplication in teleosts, these authors identified four ancient exons that lost their coding capacity in teleosts and were exapted as transcriptional enhancers of nearby genes, as demonstrated in transgenic medaka embryos [92]. Gene conversion may promote the adoption of an enhancer previously used by another gene, as shown by the group of Mike Levine in the beetle *Tribolium castaneum* in which a cardiac enhancer located in the 3' flanking region of the *ladybird* gene was relocated to the 5' end of the neighbouring gene C15. This inversion, therefore, promotes the cardiac expression of C15 in *Tribolium* [93].

### (b) Enhancers change, function not always

That enhancers can change without modifying their regulatory activity is best illustrated by the *even-skipped* stripe 2 enhancer (S2E) in species of the *Drosophila* genus, in which the shuffling of TFBS within the S2E of different species leads to the same expression pattern in transgenic *Drosophila melanogaster* embryos [94,95]. Even the orthologous *eve* enhancers of sepsid flies, which are distant relatives of *Drosophila* and whose enhancer sequences are very different, can drive equivalent expression in transgenic *D. melanogaster* embryos [96], showing that enhancers with different arrangements of TFBS can correctly interpret a TF code and give rise to the same expression output. Importantly, in complementation experiments, the divergent *Drosophila pseudoobscura* S2E perfectly rescued the embryonic lethal phenotype of mutant *D. melanogaster* carrying homozygous deletions of S2E, proving that two enhancers with highly different sequences can be functionally interchangeable, at least when tested in the laboratory [95]. Thus, stabilizing selection is reminiscent of the politically cynical ideas of the Sicilian Prince of Salinas who in the novel 'The Leopard' by Giuseppe Tomasi di Lampedusa published in 1957 stated that 'everything needs to change, so everything can stay the same'. In vertebrates, this scenario is illustrated by experiments in which mammalian enhancers can drive appropriate expression in zebrafish embryos even in the absence of obvious sequence identity. For example, Fisher *et al.* [97] analysed human enhancers of the *RET* locus in zebrafish and found that 11 out of 13 drove equivalent reporter gene expression in zebrafish, even though no orthologous zebrafish enhancers could be found by sequence comparisons. In this issue, Domené *et al.* [98] show that the mammalian *Pomc* enhancers nPE1 and nPE2 drive expression to POMC neurons of zebrafish embryos, even though their exapted origin from TEs in the early stages of the mammalian radiation [30,31] rules out the possibility that they are orthologous to the teleost enhancers.

Contrarily to the cases described above, there are examples in which conservation and high sequence identity may be functionally deceiving as has been reported for the zebrafish *Shh* enhancers ar-D [5' proximal], ar-A (intron 1) and ar-C (intron 2) that are highly conserved at the sequence and location level with their corresponding mouse orthologues, but which direct expression to different expression territories in each animal, demonstrating that the function of the structurally conserved enhancers has diverged during vertebrate evolution [43]. Similarly, but at the paralogue level, the ar-C enhancers of the fugu *shha* and *shhb* genes carry minor sequence changes that, however, modify the expression pattern of reporter genes in transgenic zebrafish embryos [99].

*Hox* genes also provide a great setting where to look for enhancer evolution between duplicate paralogues. Half a billion years ago, the ancestral *Hox* gene cluster became quadruplicated in the vertebrate lineage, generating novel *Hox* genes paralogous after two consecutive events of tetraploidization. *Hoxa1* and *Hoxb1* are indispensable for proper hindbrain segmentation as has been demonstrated in individual knockout mice. Analysis of these mutant mice showed that no active paralogue was able to compensate for the lack of the loss-of-function copy, suggesting subfunctionalization. Indeed, it is known that *Hoxa1* lost its ability to autoregulate its expression levels, whereas *Hoxb1* lost the responsiveness to retinoic acid. In a reverse evolution experiment, Tvrdik & Capecchi [100] reconstructed the ancestral *Hox1* gene by inserting the autoregulation enhancer of *Hoxb1* into the 5' flanking region of *Hoxa1* in the context of a *Hoxb1* null-allele mutant. The resulting enhancer knockin/gene knockout mice surprisingly showed normal development, demonstrating that subfunctionalized gene paralogues can be reset to the primitive state and replaced with a single copy provided that both paralogue proteins retain equivalent activities.

Of course, enhancer evolution can also lead to divergent expression patterns, possibly driving phenotypical innovation. The introduction of novel TBFS into pre-existing functional enhancers may allow a gene to acquire expression in an additional spatio-temporal domain without affecting its ancestral expression pattern. Rebeiz *et al.* [101] reported that the *neprilysin-1* gene (*Nep1*) of *Drosophila santomea* has gained a novel expression pattern in optic lobe neuroblasts after accumulating four mutations near a pre-existing intronic enhancer responsible for driving *Nep1* expression to ancestral areas such as the ventral ganglion and the retinal field. In this issue, Glassford & Rebeiz [102] expand on their previous work by systematically testing the mutational paths that led to the accumulation of these four mutations. Interestingly, their results indicate that some of the paths were prohibited due to fitness costs associated with epistasis [102].

### (c) Enhancers come and go

Enhancers are not only modified or created *de novo* but may also be lost during the course of evolution as will be further discussed in §6a. Bejerano and co-workers [103] found that hundreds of conserved non-coding genomic regions are independently lost in distinct mammalian genomes, raising the possibility that some of them could be involved in lineage or species-specific morphological variation. The existence of functionally overlapping or near redundant enhancers provides a genetic substrate for this type of mechanism to occur; the surprising lack of detectable differential phenotypes in four different strains of

mutant mice lacking ultraconserved enhancers [104] exemplifies that animals may lose sequences selected for millions of years without notable deleterious effects, at least in the comfortable environment of a laboratory animal facility.

Another interesting evolutionary mechanism that induces loss of function of tissue-specific enhancers is the appearance of a novel transcriptional repressor that silences previously active enhancers, as has been reported in a recent study of the molecular evolution of the vertebrate paralogues *pax2* and *pax8* [105]. In *Xenopus laevis* tadpoles, *pax2* is expressed in several cell-types, whereas *pax8* is expressed only in a subset of *pax2*-expressing tissues. Despite this difference, *pax2* and *pax8* share several paralogous, conserved non-coding elements that, individually, recapitulate a *pax2*-like reporter expression pattern when ligated upstream of a minimal  $\beta$ -actin promoter and tested in transgenic tadpoles. Surprisingly, when the  $\beta$ -actin promoter was replaced by the *Xenopus pax8* proximal promoter, the *pax2* and *pax8* enhancers drove a more restricted *pax8*-like expression pattern, demonstrating that the ancestral enhancers of *pax2* and *pax8* are equally able to direct *pax2*-like, multi-tissue expression, except when placed upstream of a silencer present in the *pax8* proximal promoter that suppresses expression outside the *pax8*-expressing tissues [105].

## 6. Enhancer evolution and animal diversity

What about real data indicating that mutations in enhancer sequences underlie evolution of important traits? The number of detailed studies of phenotypic diversification driven by enhancers is still very small, and the difficulty in setting up such experiments lies in the necessity of having (i) an easy-to-spot phenotype that displays intra- or interspecific variation; (ii) knowledge of gene(s) that control(s) such phenotype; (iii) detailed data on the DNA sequences (promoter and enhancers) that regulate the genes in the relevant organ/structure; (iv) sequence information from related species or populations in which the phenotype shows a relevant variation and (v) an amenable experimental system with which to perform functional studies with the different versions of the enhancer, usually by transgenesis.

Owing to this complexity, most studies on the importance of *cis*-regulatory variation in species evolution and adaptation do not go very deep. Phenotypic variation at many interesting traits is known to be controlled by *cis*-regulatory divergence, but usually the DNA sequences behind the variability are unknown. For instance, it is well established that colour pattern in wings of *Heliconius* butterflies is linked to divergent regulation of the *optix* gene, but the *cis*-regulatory sequences are still to be found [106,107]. In other examples, nucleotide polymorphisms linked to a trait are known but their regulatory role has not been analysed in great detail. One example of the latter case is *lactase* expression in human populations. Several pastoralist human populations around the world have independently evolved the capacity to express *lactase* during adulthood, which allows digestion of milk and dairy products. Many polymorphisms linked to the persistence of *lactase* expression in adults are known, all of them located on an intron of an adjacent gene [108]. The intronic region is conserved only in primates and displays enhancer activity in cell culture assays, but a detailed characterization of *lactase* regulatory elements and its variants in a

more physiological model, like transgenic mice, is still lacking [108,109]. In other cases, suggestive differences in enhancer sequence and activity between species are known, but the derived expression pattern cannot be linked to a trait.

A detailed understanding of the relation between enhancer evolution and trait variation is only possible by series of studies of a particular locus in several species or populations. Table 1 shows a list of studies that provide compelling evidence as to how enhancer sequence evolution has contributed to specific traits in animals. The examples illustrate various mechanisms that lead to divergence of enhancer function: deletions, modification in the TFBS repertoire and acquisition of new enhancer regions.

### (a) Enhancer deletion and phenotypic variation

A dramatic mechanism of enhancer evolution is simply the deletion of an enhancer, which leads to the loss of gene expression in a particular region of the body. A good example concerns a pelvic enhancer of the *Pitx1* gene in the three-spined stickleback, a teleost fish that lives in the sea as well as in North American and Northern European freshwater lakes. Marine sticklebacks possess bony spines in the pelvic region, presumably as protection from predators; freshwater populations, on the other hand, usually lack large pelvic armour. Development of such spines depends on the activity of a TF, Pitx1, which is expressed in the pelvic region under the control of a specific enhancer [111,112]. Several freshwater stickleback populations lack spines due to deletions of the pelvic enhancer; interestingly, deletions have happened independently several times in isolated populations of freshwater sticklebacks [111,112].

Other likely cases of enhancer deletions during evolution have been uncovered by McLean *et al.* [114], who identified 509 instances of non-coding DNA regions which are conserved in mammals but have been deleted since the split between humans and chimpanzees. One of the sequences absent in humans is an enhancer of the *androgen receptor (AR)* gene that drives expression to vibrissae (sensory hairs in the face) and the genital tubercle during development [114]. Testosterone is necessary for sensory vibrissae growth and for the development of penile spines, structures that are present in rodents and non-human primates but absent in our species. Thus, a reasonable hypothesis is that the loss of an *AR* enhancer in our close ancestors led to morphological change in our lineage [114]. Thus, the work on teleost *Pitx1* and mammalian *AR* shows that deletion of enhancers that control the expression of a TF in a particular region can cause evolution to happen by large steps, eliminating whole morphological structures provided that the lack of such structures (like pelvic and penile spines) are not selected against. Another enhancer deletion with potential functional significance uncovered by McLean *et al.* [114] affects the expression of the tumour-suppressor gene *GADD45G*, which might be related to human-specific neuronal proliferation in the forebrain.

### (b) Evolution by mutations in TFBS

Another, more subtle way of enhancer evolution is via point mutations and small indels that either create or eliminate TFBS within an enhancer, thereby changing its activity. Several enhancers that control pigmentation in flies of the genus *Drosophila* are known to have evolved in this way. The *yellow* gene, which encodes an enzyme involved in pigment synthesis, is expressed in the *Drosophila* wing under the control



**Table 1.** Enhancers involved in evolutionary innovation. n.d., No data.

trait	species/populations	gene	TF	enhancer variation	references
vertebrates	mouse versus chicken	<i>Hoxc8</i>	n.d.	point mutations	[110]
	sticklebacks	<i>Pitx1</i>	n.d.	deletion	[111, 112]
	bats versus mice	<i>Prx1</i>	n.d.	point mutations	[113]
	humans versus apes and mice	<i>androgen receptor</i>	n.d.	deletion	[114]
invertebrates	teleosts versus tetrapods	<i>Hoxd13</i>	n.d.	acquisition of a new limb enhancer	[115]
	snakes versus mammals	<i>Myf5</i>	Pax3, Hoxa10, Hoxb6	point mutations	[116]
	<i>D. melanogaster</i> , <i>biarmipes</i> , <i>tristis</i> , <i>gunungcola</i> , <i>elegans</i> , <i>mimetic</i>	<i>yellow</i>	Engrailed, Distalless (Dll)	point mutations	[117–119]
	various <i>Drosophila</i> species	<i>yellow</i>	Abdominal B (ABD-B)	point mutations	[120]
invertebrates	<i>D. melanogaster</i> , <i>sechellia</i> , <i>ezoana</i> , <i>viridis</i> , <i>littoralis</i>	<i>shavenba by/owo</i>	n.d.	point mutations	[121–123]
	<i>D. melanogaster</i> and <i>D. santomea</i>	<i>tan</i>	n.d.	point mutations, deletions	[124]
	Ugandan <i>D. melanogaster</i> populations	<i>ebony</i>	n.d.	point mutations	[125]
	abdominal pigmentation				

of the *spot* enhancer [117]. In *D. melanogaster*, *yellow* expression is low and steady throughout the wing, leading to uniform pigmentation. In *Drosophila biarmipes*, in contrast, *yellow* is highly expressed in a corner of the wing, creating a dark spot [117]. Dark spots on the wings may have a function during wooing, when males execute an elaborate dance to females. Analyses of enhancer sequence differences and transgenic assays indicate that the *spot* enhancer in *D. biarmipes* has sites for a transcriptional activator (recently identified as Distalless) [118] as well as for the repressor Engrailed, the latter responsible for setting the posterior boundary of the *yellow* spot [117]. In a subsequent work, Prud'homme *et al.* [119] checked for the presence of wing spots throughout *Drosophila* phylogeny and found that independent gains and losses of the wing spot have happened a couple of times during the evolution of the genus. Loss of the wing spot in *Drosophila gunungcola* and *Drosophila mimetica* happened by mutations in the same *spot* enhancer, in a case of convergent phenotypic evolution driven by change in the same regulatory element [119]. Notably, gain of wing spot in *Drosophila tristis*, which primitively lacked a spot, happened not by mutations in the *spot* enhancer, as in *D. biarmipes*, but in another enhancer located in a *yellow* intron that directs expression to wing veins [119]. Thus, a remarkably similar phenotype—the dark spot on the upper right corner of the wing—was independently generated by the co-option of two different *yellow* enhancers during fly evolution [119].

*Yellow* is also involved in the male-specific pigmentation of the abdomen, where its expression is controlled by the *body* enhancer [120]. In *D. melanogaster* and *D. pseudoobscura*, *body* enhancer activity depends on the binding of TF Abdominal-B (Abd-B), but the TFBS responsible for binding have been lost in the *body* enhancer of *Drosophila kikkawai*, which lacks abdominal pigmentation [120]. In another species lacking abdominal pigmentation, *D. santomea*, it is the expression of *tan* that is altered [124]. *tan* is an enzyme with functions in pigmentation and vision. Its expression depends on an enhancer that is conserved in the *Drosophila* genus, but in *D. santomea* at least three different kinds of mutations, including point mutations and deletions, render the enhancer non-functional in the abdomen, which stays unpigmented [124]. Also variation in an enhancer of *ebony*, which encodes an enzyme related to pigmentation, was found to be related to abdominal pigmentation in African *D. melanogaster* populations [125].

Another variable phenotype in flies which has been studied in great depth is trichome development in *Drosophila* larvae. The precise pattern of trichome distribution depends on several enhancers that control the expression of the TF *shavenbaby* (*svb*), as illustrated by Stern & Frankel [126] in this issue, who summarize 13 years of continuous work on the evolution of *shavenbaby* expression in several *Drosophila* species. Quarternary trichome loss in *Drosophila sechellia* larvae is due to several (at least five) point mutations in one particular enhancer, *E6*, which reduce *svb* expression in the area of the larvae that gives rise to quarternary trichomes [121]. Interestingly, each mutation in the *E6* has a small effect *per se*, so that significant reduction in *svb* expression is only achieved when all sites are mutated. Thus, loss of enhancer activity can be brought about not only by enhancer deletion but also by the accumulation of point mutations with small effect, possibly to avoid pleiotropic effects caused by complete enhancer loss [121].

In tetrapods, limb anatomy and length is partially controlled by the TF *Prx1*. Forelimbs in bats are much longer

than those of mice, and Cretekos *et al.* [113] showed that replacing a mouse *Prx1* limb enhancer with the orthologous enhancer from a bat caused the limb of mutant mice to be a little longer, indicating that mutations in the enhancer alter its activity and contribute to the evolution of forelimb anatomy [113]. Another recent example of enhancer mutations driving innovation is the work by Guerreiro *et al.* [116] on the axial anatomy of snakes. Vertebrae in tetrapods can usually be subdivided into several types (cervical, thoracic, lumbar, sacral and caudal). Snakes, on the other hand, are characterized by a homogeneous vertebral anatomy, with a great number of thoracic-like, ribbed vertebrae. In vertebrates, rib formation at posterior levels of the column is repressed by TF Hoxa10, but in snakes this repression does not take place [127]. Studying an enhancer of *Myf5*, a gene necessary for proper axial development, Guerreiro *et al.* [116] identified sites that can bind to TF Pax3 and Hoxb6, which act as activators, and Hoxa10, which is a repressor [128]. In snakes, there is a single nucleotide substitution in the enhancer that prevents Hoxa10 binding without interfering with activator binding, which likely explains why ribbed vertebrae are formed in snakes despite of Hoxa10 expression not being altered in this group [116,129].

### (c) Acquisition of new enhancers and expression territories

During evolution, it is certain that new regulatory activities originate not only by tinkering with pre-existing enhancers but also by the appearance of new enhancers. Patterns of CNE distribution suggest that new enhancers played a role in several stages of vertebrate evolution [130], while a burst of co-option of transposons as new *cis*-regulatory elements is associated with the reproductive innovations of placental mammals [131,132]. Recently, mapping of enhancer chromatin marks (H3K27ac) at equivalent stages of limb development in mouse, macaque and human uncovered many putative enhancers which seem to be active only in the developing human limb and which might contribute to the specific limb phenotype of our species [133].

An important event in the history of vertebrates was the morphological transition from a fin to a limb in the ancestor of tetrapods. Expression of genes of the *Hoxd* group are more intense in the developing limbs of tetrapods compared with developing fins of fish, and it has been hypothesized that this feature might be related to the differences in the morphology of these structures. Recently, Freitas *et al.* [115] found that overexpression of *Hoxd13* in the zebrafish developing fin causes increased proliferation of the chondroskeleton, which acquires anatomical and molecular characteristics similar to a tetrapod fin. In addition, they show that a tetrapod-specific *Hoxd* enhancer, *CsC*, drives robust expression to the fins of transgenic zebrafish, showing that the *trans*-acting factors ready to overexpress *Hoxd* were in place before the appearance of the tetrapod limb [115,134].

Many of the studies described above that pinpoint enhancer mutations leading to morphological innovation began by mapping a large-effect genetic locus responsible for variation in closely related species or populations [135–138]. Mapping may also detect genomic regions with signs of a selective sweep, an indication that natural selection is fixing a variant in the population. In this issue, Glaser-Schmitt *et al.* [139] describe a selective sweep around an enhancer of the *CG9509*

gene (encoding an enzyme of unknown function) that is responsible for a consistently higher expression of this gene in European versus sub-Saharan African populations of *D. melanogaster*. Increased expression of the enzyme outside sub-Saharan Africa likely indicates that this phenotype has fitness value, perhaps because the enzyme may have detoxifying functions [139].

## 7. Enhancer evolution in humans

In their seminal paper, King & Wilson [1] observed that chimpanzee and human proteins were exceptionally similar in sequence and deduced that the differences between these species should mainly be due to mutations in regulatory DNA. Testing this idea is not an easy task due to the lack of appropriate model organisms to manipulate [140], but the work of McLean *et al.* [114] discussed above indicates that, at least for some traits, highly suggestive experimental evidence can be obtained for the identification of human-specific, non-coding mutations that helped mould our species. In this regard, sequences displaying signs of accelerated evolution in humans often have enhancer activity and might be responsible for human-specific phenotypic adaptations [141]. One example is element *HACNS1*, a limb enhancer showing evidence of accelerated evolution in the human lineage [142]. In transgenic mice, human *HACNS1* shows an increased transcriptional activity compared with the chimpanzee and macaque orthologous enhancers [142], probably due to mutations that prevent the binding of repressors of the gene [135]. Although it has been hypothesized that this might be linked to the evolution of limb morphology, the gene controlled by *HACNS1* is unknown and no experimental evidence for a function in limb development is available [142]. In this issue, Capra *et al.* [143] combine data on non-coding human accelerated regions with genome-wide surveys of chromatin marks and binding of transcription factors and cofactors to identify potential enhancers. Experimental testing of a set of human and chimpanzee enhancer orthologues indicates that most can work as enhancers in transgenic mice and, consistent with the evidence for positive selection in humans, many of them have different regulatory activity between the species [143]. These human accelerated regulatory sequences, together with other putative human-specific enhancers found in genome-wide surveys [142,144], are a rich dataset for the discovery of the peculiar gene regulation features that make us human. Taking advantage of some of these databases, the group of Lucía Franchini identified the genes carrying the highest number of human-specific accelerated sequences [141]. The developmental brain gene *NPAS3* is at the top of this ranking with up to 14 elements that are highly conserved in mammals, including primates, but carry human-specific nucleotide substitutions. In this issue, Kamm *et al.* [145] study the accelerated element 2xHAR142 present in intron 5 of human *NPAS3* and show that the mouse and chimp 2xHAR142 orthologues behave as transcriptional enhancers in transgenic mice driving *lacZ* expression to similar regions of the central nervous system where mouse *Npas3* is normally expressed. Interestingly, the human 2xHAR142 orthologue extends the area of *lacZ* expression to the developing anterior telencephalon, providing an example of human-specific heterotopy promoted by an accelerated transcriptional enhancer that could have



contributed to the characteristic enlargement of this brain area in humans [145].

## 8. Concluding remarks

Deciphering the genetic mechanisms that operated during the past 650 Myr of animal evolution to create the seemingly infinite variety of animals that populate our planet has been one of the holy grails in biology. However, it is only recently that scientists have developed a number of experimental tools that allow study of these problems that were much more difficult to tackle in the pre-genomic era (before *ca* 2005). The idea of putting together this Theme Issue comes at a critical time in which the massive availability of genome sequences from several species, coupled to single- and multiple-gene functional studies in transgenic organisms and sophisticated maps of chromatin features have produced an unprecedented impact in the identification of potential cell-type-specific enhancers with functional roles in different lineages or species. We are just beginning to understand the different possible mechanisms by which enhancers evolve and contribute to novel forms and functions. The limited number of examples documented in the literature, most of which are listed above, are certainly but the tip of the iceberg, as hundreds of thousands of candidate enhancer sequences have been identified by sequence conservation and enhancer-associated chromatin marks in many different animal genomes. However, to make sense of whole-genome surveys and bioinformatic predictions, it will be important to further develop the methods to evaluate regulatory function *in vivo*, because the data derived from genome-wide studies are descriptive in nature. In this regard, novel techniques have been recently developed to test the regulatory activity of

DNA elements at a large scale (see for instance ref. [82]), something that should help functional studies to keep pace with genome-wide surveys. In addition to reporter gene assays, it will also be desirable to accelerate the study of enhancers by loss-of-function assays, which is perhaps the best way of evaluating the function of a DNA sequence. In recent years, promising novel genome-editing techniques based on the use of prokaryotic CRISP/Cas9 nuclease [146,147] and the transcription activator-like effector nucleases [148–150] have emerged and may be rapidly adapted to make loss-of-function studies of enhancers in several vertebrate and invertebrate species, including those organisms in which regulatory regions knockouts were, until now, not amenable, as have been recently achieved in zebrafish [148,151,152] and medaka [150]. The use of these novel techniques will certainly accelerate discoveries in the field of enhancer evolution and animal diversity.

The articles, reviews and perspectives that follow this introduction shed more light in the still murky and mysterious world of gene expression evolution in the animal kingdom. Each of these papers, in one way or another, consolidates the idea that there will probably be no fixed law, like gravity, to explain at the molecular level *how endless forms most beautiful and most wonderful have been, and are being evolved*. It rather seems that a wide variety of peculiar molecular mechanisms perform, together, the complex task of putting the genome in action, in each cell type of each animal species, at every moment in life and under every possible physiological and environmental circumstance.

**Funding statement.** This work was supported by an NIH grant no. DK068400 (M.R.), Agencia Nacional de Promoción Científica y Tecnológica, Argentina (M.R. and F.J.S.) and Universidad de Buenos Aires (M.R.).

## References

- King MC, Wilson AC. 1975 Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116. (doi:10.1126/science.1090005)
- Waterston RH *et al.* 2002 Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562. (doi:10.1038/nature01262)
- Chimpanzee Sequencing and Analysis Consortium. 2005 Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87. (doi:10.1038/nature04072)
- Prabhakar S *et al.* 2006 Close sequence comparisons are sufficient to identify human *cis*-regulatory elements. *Genome Res.* **16**, 855–863. (doi:10.1101/gr.4717506)
- Clarke SL, VanderMeer JE, Wenger AM, Schaar BT, Ahituv N, Bejerano G. 2012 Human developmental enhancers conserved between deuterostomes and protostomes. *PLoS Genet.* **8**, e1002852. (doi:10.1371/journal.pgen.1002852)
- Konopka G *et al.* 2009 Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature* **462**, 213–217. (doi:10.1038/nature08549)
- Hoekstra HE, Coyne JA. 2007 The locus of evolution, *evo devo* and the genetics of adaptation. *Evolution* **61**, 995–1016. (doi:10.1111/j.1558-5646.2007.00105.x)
- Britten RJ, Davidson EH. 1969 Gene regulation for higher cells, a theory. *Science* **165**, 349–357. (doi:10.1126/science.165.3891.349)
- Ohno S. 1972 Simplicity of mammalian regulatory systems. *Dev. Biol.* **27**, 131–136. (doi:10.1016/0012-1606(72)90117-0)
- Carroll SB. 2005 Evolution at two levels, on genes and form. *PLoS Biol.* **3**, e245. (doi:10.1371/journal.pbio.0030245)
- Wray GA. 2007 The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* **8**, 206–216. (doi:10.1038/nrg2063)
- Carroll SB. 2008 *Evo-devo* and an expanding evolutionary synthesis, a genetic theory of morphological evolution. *Cell* **134**, 25–36. (doi:10.1016/j.cell.2008.06.030)
- Wittkopp PJ, Kalay G. 2011 *Cis*-regulatory elements, molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69.
- Lettice LA *et al.* 2003 A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735. (doi:10.1093/hmg/ddg180)
- Jeong Y *et al.* 2008 Regulation of a remote Shh forebrain enhancer by the Six3 homeoprotein. *Nat. Genet.* **40**, 1348–1353. (doi:10.1038/ng.230)
- Ghasvand NM, Rudolph DD, Mashayekhi M, Brzezinski JAT, Goldman D, Glaser T. 2011 Deletion of a remote enhancer near ATOH7 disrupts retinal neurogenesis, causing NCRNA disease. *Nat. Neurosci.* **14**, 578–586. (doi:10.1038/nn.2798)
- Spielmann M *et al.* 2012 Homeotic arm-to-leg transformation associated with genomic rearrangements at the PITX1 locus. *Am. J. Hum. Genet.* **91**, 629–635. (doi:10.1016/j.ajhg.2012.08.014)
- Sur IK *et al.* 2012 Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. *Science* **338**, 1360–1363. (doi:10.1126/science.1228606)
- Ragvin A *et al.* 2010 Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc. Natl Acad. Sci. USA* **107**, 775–780. (doi:10.1073/pnas.0911591107)
- Maurano MT *et al.* 2012 Systematic localization of common disease-associated variation in regulatory

- DNA. *Science* **337**, 1190–1195. (doi:10.1126/science.1222794)
21. Hadley ME, Haskell-Luevano C. 1999 The proopiomelanocortin system. *Ann. NY Acad. Sci.* **885**, 1–21. (doi:10.1111/j.1749-6632.1999.tb08662.x)
  22. Krivega I, Dean A. 2012 Enhancer and promoter interactions—long distance calls. *Curr. Opin. Genet. Dev.* **22**, 79–85. (doi:10.1016/j.gde.2011.11.001)
  23. Kagey MH *et al.* 2010 Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435. (doi:10.1038/nature09380)
  24. Kind J, van Steensel B. 2010 Genome—nuclear lamina interactions and gene regulation. *Curr. Opin. Cell Biol.* **22**, 320–355. (doi:10.1016/j.ceb.2010.04.002)
  25. van Heyningen V, Bickmore W. 2013 Regulation from a distance: long-range control of gene expression in development and disease. *Phil. Trans. R. Soc. B* **368**, 20120372. (doi:10.1098/rstb.2012.0372)
  26. Banerji J, Rusconi S, Schaffner W. 1981 Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308. (doi:10.1016/0092-8674(81)90413-X)
  27. Arnosti DN, Kulkarni MM. 2005 Transcriptional enhancers, intelligent enhanceosomes or flexible billboards? *J. Cell Biochem.* **94**, 890–898. (doi:10.1002/jcb.20352)
  28. Merika M, Thanos D. 2001 Enhanceosomes. *Curr. Opin. Genet. Dev.* **11**, 205–208. (doi:10.1016/S0959-437X(00)00180-5)
  29. Arnosti DN, Barolo S, Levine M, Small S. 1996 The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development* **122**, 205–214.
  30. Santangelo AM, de Souza FS, Franchini LF, Bumashny VF, Low MJ, Rubinstein M. 2007 Ancient exaptation of a CORE-SINE retroposon into a highly conserved mammalian neuronal enhancer of the proopiomelanocortin gene. *PLoS Genet.* **3**, 1813–1826. (doi:10.1371/journal.pgen.0030166)
  31. Franchini LF, Lopez-Leal R, Nasif S, Beati P, Gelman DM, Low MJ, de Souza FJS, Rubinstein M. 2011 Convergent evolution of two mammalian neuronal enhancers by sequential exaptation of unrelated retroposons. *Proc. Natl Acad. Sci. USA* **108**, 15 270–15 275. (doi:10.1073/pnas.1104997108)
  32. Swanson CI, Evans NC, Barolo S. 2010 Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer. *Dev. Cell* **18**, 359–370. (doi:10.1016/j.devcel.2009.12.026)
  33. Evans NC, Swanson CI, Barolo S. 2012 Sparkling insights into enhancer structure, function, and evolution. *Curr. Top. Dev. Biol.* **98**, 97–120. (doi:10.1016/B978-0-12-386499-4.00004-5)
  34. Junion G, Spivakov M, Girardot C, Braun M, Gustafson EH, Birney E, Furlong EE. 2012 A transcription factor collective defines cardiac cell fate and reflects lineage history. *Cell* **148**, 473–486. (doi:10.1016/j.cell.2012.01.030)
  35. Li XY *et al.* 2008 Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6**, e27. (doi:10.1371/journal.pbio.0060027)
  36. Fisher WW *et al.* 2012 DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*. *Proc. Natl Acad. Sci. USA* **109**, 21 330–21 335. (doi:10.1073/pnas.1209589110)
  37. Ramos AI, Barolo S. 2013 Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution. *Phil. Trans. R. Soc. B* **368**, 20130018. (doi:10.1098/rstb.2013.0018)
  38. Guturu H, Doxey AC, Wenger AM, Bejerano G. 2013 Structure-aided prediction of mammalian transcription factor complexes in conserved non-coding elements. *Phil. Trans. R. Soc. B* **368**, 20130029. (doi:10.1098/rstb.2013.0029)
  39. Jeong Y, El-Jaick K, Roessler E, Muenke M, Epstein DJ. 2006 A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* **133**, 761–772. (doi:10.1242/dev.02239)
  40. Sagai T, Hosoya M, Mizushima Y, Tamura M, Shiroishi T. 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. (doi:10.1242/dev.01613)
  41. Langlais D, Couture C, Sylvain-Drolet G, Drouin J. 2011 A pituitary-specific enhancer of the POMC gene with preferential activity in corticotrope cells. *Mol. Endocrinol.* **25**, 348–359. (doi:10.1210/me.2010-0422)
  42. de Souza FS, Santangelo AM, Bumashny V, Avale ME, Smart JL, Low MJ, Rubinstein M. 2005 Identification of neuronal enhancers of the proopiomelanocortin gene by transgenic mouse analysis and phylogenetic footprinting. *Mol. Cell Biol.* **25**, 3076–3086. (doi:10.1128/MCB.25.8.3076-3086.2005)
  43. Ertzer R, Muller F, Hadzhiev Y, Rathnam S, Fischer N, Rastegar S, Strahle U. 2007 Cooperation of sonic hedgehog enhancers in midline expression. *Dev. Biol.* **301**, 578–589. (doi:10.1016/j.ydbio.2006.11.004)
  44. Marinic M, Aktas T, Ruf S, Spitz F. 2013 An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530–542. (doi:10.1016/j.devcel.2013.01.025)
  45. Barolo S. 2012 Shadow enhancers, frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *Bioessays* **34**, 135–141. (doi:10.1002/bies.201100121)
  46. Frankel N. 2012 Multiple layers of complexity in cis-regulatory regions of developmental genes. *Dev. Dyn.* **241**, 1857–1866. (doi:10.1002/dvdy.23871)
  47. Hong JW, Hendrix DA, Levine MS. 2008 Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314. (doi:10.1126/science.1160631)
  48. Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. 2010 Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493. (doi:10.1038/nature09158)
  49. Perry MW, Boettiger AN, Bothma JP, Levine M. 2010 Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20**, 1562–1567. (doi:10.1016/j.cub.2010.07.043)
  50. Haeussler M, Joly JS. 2011 When needles look like hay: how to find tissue-specific enhancers in model organism genomes. *Dev. Biol.* **350**, 239–254. (doi:10.1016/j.ydbio.2010.11.026)
  51. Heintz N. 2000 Analysis of mammalian central nervous system gene expression and function using bacterial artificial chromosome-mediated transgenesis. *Hum. Mol. Genet.* **9**, 937–943. (doi:10.1093/hmg/9.6.937)
  52. Uchikawa M, Ishida Y, Takemoto T, Kamachi Y, Kondoh H. 2003 Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev. Cell* **4**, 509–519. (doi:10.1016/S1534-5807(03)00088-1)
  53. Wasserman WW, Sandelin A. 2004 Applied bioinformatics for the identification of regulatory elements. *Nat. Rev. Genet.* **5**, 276–287. (doi:10.1038/nrg1315)
  54. Tagle DA, Koop BF, Goodman M, Slightom JL, Hess DL, Jones RT. 1988 Embryonic epsilon and gamma globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *J. Mol. Biol.* **203**, 439–455. (doi:10.1016/0022-2836(88)90011-3)
  55. Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. 2003 Scanning human gene deserts for long-range enhancers. *Science* **302**, 413. (doi:10.1126/science.1088328)
  56. Woolfe A *et al.* 2005 Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol.* **3**, e7. (doi:10.1371/journal.pbio.0030007)
  57. Pennacchio LA *et al.* 2006 In vivo enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502. (doi:10.1038/nature05295)
  58. Shin JT *et al.* 2005 Human—zebrafish non-coding conserved elements act in vivo to regulate transcription. *Nucleic Acids Res.* **33**, 5437–5445. (doi:10.1093/nar/gki853)
  59. Plessy C, Dickmeis T, Chalmel F, Strahle U. 2005 Enhancer sequence conservation between vertebrates is favoured in developmental regulator genes. *Trends Genet.* **21**, 207–210. (doi:10.1016/j.tig.2005.02.006)
  60. Woolfe A, Elgar G. 2008 Organization of conserved elements near key developmental regulators in vertebrate genomes. *Adv. Genet.* **61**, 307–338. (doi:10.1016/S0065-2660(07)00012-0)
  61. Lindblad-Toh K *et al.* 2011 A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482. (doi:10.1038/nature10530)
  62. Bejerano G *et al.* 2004 Ultraconserved elements in the human genome. *Science* **304**, 1321–1325. (doi:10.1126/science.1098119)
  63. Royo JL *et al.* 2011 Transphyletic conservation of developmental regulatory state in animal evolution. *Proc. Natl Acad. Sci. USA* **108**, 14186–14191. (doi:10.1073/pnas.1109037108)
  64. Harmston N, Barešić A, Lenhard B. 2013 The mystery of extreme non-coding conservation. *Phil. Trans. R. Soc. B* **368**, 20130021. (doi:10.1098/rstb.2013.0021)
  65. Maeso I, Irimia M, Tena JJ, Casares F, Gómez-Skarmeta JL. 2013 Deep conservation of cis-regulatory elements in metazoans. *Phil. Trans. R. Soc. B* **368**, 20130020. (doi:10.1098/rstb.2013.0020)

66. Visel A *et al.* 2009 ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858. (doi:10.1038/nature07730)
67. Heintzman ND *et al.* 2009 Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112. (doi:10.1038/nature07829)
68. Zentner GE, Scacheri PC. 2012 The chromatin fingerprint of gene enhancer elements. *J. Biol. Chem.* **287**, 30 888–30 896. (doi:10.1074/jbc.R111.296491)
69. Harmston N, Lenhard B. 2013 Chromatin and epigenetic features of long-range gene regulation. *Nucleic Acids Res.* **41**, 7185–7199. (doi:10.1093/nar/gkt499)
70. He A, Kong SW, Ma Q, Pu WT. 2011 Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc. Natl Acad. Sci. USA* **108**, 5632–5637. (doi:10.1073/pnas.1016959108)
71. Gotea V, Visel A, Westlund JM, Nobrega MA, Pennacchio LA, Ovcharenko I. 2010 Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Res.* **20**, 565–577. (doi:10.1101/gr.104471.109)
72. Goke J *et al.* 2011 Combinatorial binding in human and mouse embryonic stem cells identifies conserved enhancers active in early embryonic development. *PLoS Comput. Biol.* **7**, e1002304. (doi:10.1371/journal.pcbi.1002304)
73. Whyte WA *et al.* 2013 Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319. (doi:10.1016/j.cell.2013.03.035)
74. Bernstein BE *et al.* 2012 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74. (doi:10.1038/nature11247)
75. Shen Y *et al.* 2012 A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **488**, 116–120. (doi:10.1038/nature11243)
76. Sakabe NJ, Nobrega MA. 2013 Beyond the ENCODE project: using genomics and epigenomics strategies to study enhancer evolution. *Phil. Trans. R. Soc. B* **368**, 20130022. (doi:10.1098/rstb.2013.0022)
77. de Souza FS, Franchini LF, Rubinstein M. 2013 Exaptation of transposable elements into novel *cis*-regulatory elements: is the evidence always strong? *Mol. Biol. Evol.* **30**, 1239–1251. (doi:10.1093/molbev/mst045)
78. Graur D, Zheng Y, Price N, Azevedo RB, Zufall RA, Elhaik E. 2013 On the immortality of television sets, ‘function’ in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol. Evol.* **5**, 578–590. (doi:10.1093/gbe/evt028)
79. Struhl K. 2007 Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* **14**, 103–105. (doi:10.1038/nsmb0207-103)
80. Biggin MD. 2011 Animal transcription networks as highly connected, quantitative continua. *Dev. Cell* **21**, 611–626. (doi:10.1016/j.devcel.2011.09.008)
81. Doolittle WF. 2013 Is junk DNA bunk? A critique of ENCODE. *Proc. Natl Acad. Sci. USA* **110**, 5294–300. (doi:10.1073/pnas.1221376110)
82. White MA, Myers CA, Corbo JC, Cohen BA. 2013 Massively parallel in vivo enhancer assay reveals that highly local features determine the *cis*-regulatory function of ChIP-seq peaks. *Proc. Natl Acad. Sci. USA* **110**, 11 952–11 957. (doi:10.1073/pnas.1307449110)
83. Overstreet LS *et al.* 2004 A transgenic marker for newly born granule cells in dentate gyrus. *J. Neurosci.* **24**, 3251–3259. (doi:10.1523/JNEUROSCI.5173-03.2004)
84. Perederiy JV, Luikart BW, Washburn EK, Schnell E, Westbrook GL. 2013 Neural injury alters proliferation and integration of adult-generated neurons in the dentate gyrus. *J. Neurosci.* **33**, 4754–4767. (doi:10.1523/JNEUROSCI.4785-12.2013)
85. Lynch M. 2007 The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc. Natl Acad. Sci. USA* **104**, 8597–8604. (doi:10.1073/pnas.0702207104)
86. Weirauch MT, Hughes TR. 2010 Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. *Trends Genet.* **26**, 66–74. (doi:10.1016/j.tig.2009.12.002)
87. Goode DK, Callaway HA, Cerda GA, Lewis KE, Elgar G. 2011 Minor change, major difference: divergent functions of highly conserved *cis*-regulatory elements subsequent to whole genome duplication events. *Development* **138**, 879–884. (doi:10.1242/dev.055996)
88. Ting CN, Rosenberg MP, Snow CM, Samuelson LC, Meisler MH. 1992 Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev.* **6**, 1457–1465. (doi:10.1101/gad.6.8.1457)
89. Keller SA, Rosenberg MP, Johnson TM, Howard G, Meisler MH. 1990 Regulation of amylase gene expression in diabetic mice is mediated by a *cis*-acting upstream element close to the pancreas-specific enhancer. *Genes Dev.* **4**, 1316–1321. (doi:10.1101/gad.4.8.1316)
90. Lowe CB, Haussler D. 2012 29 mammalian genomes reveal novel exaptations of mobile elements for likely regulatory functions in the human genome. *PLoS ONE* **7**, e43128. (doi:10.1371/journal.pone.0043128)
91. Franchini LF, de Souza FS, Low MJ, Rubinstein M. 2012 Positive selection of co-opted mobile genetic elements in a mammalian gene: if you can’t beat them, join them. *Mob. Genet. Elem.* **2**, 106–109. (doi:10.4161/mge.20267)
92. Eichenlaub MP, Ettwiller L. 2011 *De novo* genesis of enhancers in vertebrates. *PLoS Biol.* **9**, e1001188. (doi:10.1371/journal.pbio.1001188)
93. Cande JD, Chopra VS, Levine M. 2009 Evolving enhancer–promoter interactions within the tinman complex of the flour beetle, *Tribolium castaneum*. *Development* **136**, 3153–3160. (doi:10.1242/dev.038034)
94. Ludwig MZ, Bergman C, Patel NH, Kreitman M. 2000 Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**, 564–567. (doi:10.1038/35000615)
95. Ludwig MZ, Palsson A, Alekseeva E, Bergman CM, Nathan J, Kreitman M. 2005 Functional evolution of a *cis*-regulatory module. *PLoS Biol.* **3**, e93. (doi:10.1371/journal.pbio.0030093)
96. Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB. 2008 Sepsid even-skipped enhancers are functionally conserved in *Drosophila* despite lack of sequence conservation. *PLoS Genet.* **4**, e1000106. (doi:10.1371/journal.pgen.1000106)
97. Fisher S, Grice EA, Vinton RM, Bessling SL, McCallion AS. 2006 Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science* **312**, 276–279. (doi:10.1126/science.1124070)
98. Domené S, Bumashny VF, de Souza FSJ, Franchini LF, Nasif S, Low MJ, Rubinstein M. 2013 Enhancer turnover and conserved regulatory function in vertebrate evolution. *Phil. Trans. R. Soc. B* **368**, 20130027. (doi:10.1098/rstb.2013.0027)
99. Hadzhiev Y, Lang M, Ertzer R, Meyer A, Strahle U, Muller F. 2007 Functional diversification of sonic hedgehog paralog enhancers identified by phylogenomic reconstruction. *Genome Biol.* **8**, R106. (doi:10.1186/gb-2007-8-6-r106)
100. Tvrdik P, Capecchi MR. 2006 Reversal of Hox1 gene subfunctionalization in the mouse. *Dev. Cell* **11**, 239–250. (doi:10.1016/j.devcel.2006.06.016)
101. Rebeiz M, Jikomes N, Kassner VA, Carroll SB. 2011 Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences. *Proc. Natl Acad. Sci. USA* **108**, 10 036–10 043. (doi:10.1073/pnas.1105937108)
102. Glassford WJ, Rebeiz M. 2013 Assessing constraints on the path of regulatory sequence evolution. *Phil. Trans. R. Soc. B* **368**, 20130026. (doi:10.1098/rstb.2013.0026)
103. Hiller M, Schaar BT, Bejerano G. 2012 Hundreds of conserved non-coding genomic regions are independently lost in mammals. *Nucleic Acids Res.* **40**, 11 463–11 476. (doi:10.1093/nar/gks905)
104. Ahituv N *et al.* 2007 Deletion of ultraconserved elements yields viable mice. *PLoS Biol.* **5**, e234. (doi:10.1371/journal.pbio.0050234)
105. Ochi H, Tamai T, Nagano H, Kawaguchi A, Sudou N, Ogino H. 2012 Evolution of a tissue-specific silencer underlies divergence in the expression of pax2 and pax8 paralogues. *Nat. Commun.* **3**, 848. (doi:10.1038/ncomms1851)
106. Reed RD *et al.* 2011 Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333**, 1137–1141. (doi:10.1126/science.1208227)
107. Supple MA *et al.* 2013 Genomic architecture of adaptive color pattern divergence and convergence in *Heliconius* butterflies. *Genome Res.* **23**, 1248–1257. (doi:10.1101/gr.150615.112)
108. Ingram CJ, Mulcare CA, Itan Y, Thomas MG, Swallow DM. 2009 Lactose digestion and the evolutionary genetics of lactase persistence. *Hum. Genet.* **124**, 579–591. (doi:10.1007/s00439-008-0593-6)
109. Tishkoff SA *et al.* 2007 Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* **39**, 31–40. (doi:10.1038/ng1946)
110. Belting HG, Shashikant CS, Ruddle FH. 1998 Modification of expression and *cis*-regulation of



- Hox8 in the evolution of diverged axial morphology. *Proc. Natl Acad. Sci. USA* **95**, 2355–2360. (doi:10.1073/pnas.95.5.2355)
111. Chan YF *et al.* 2010 Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302–305. (doi:10.1126/science.1182213)
112. Shapiro MD *et al.* 2004 Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**, 717–723. (doi:10.1038/nature02415)
113. Cretekos CJ, Wang Y, Green ED, Martin JF, Rasweiler JT, Behringer RR. 2008 Regulatory divergence modifies limb length between mammals. *Genes Dev.* **22**, 141–151. (doi:10.1101/gad.1620408)
114. McLean CY *et al.* 2011 Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* **471**, 216–219. (doi:10.1038/nature09774)
115. Freitas R, Gomez-Marin C, Wilson JM, Casares F, Gomez-Skarmeta JL. 2012 Hoxd13 contribution to the evolution of vertebrate appendages. *Dev. Cell* **23**, 1219–1229. (doi:10.1016/j.devcel.2012.10.015)
116. Guerreiro I *et al.* 2013 Role of a polymorphism in a Hox/Pax-responsive enhancer in the evolution of the vertebrate spine. *Proc. Natl Acad. Sci. USA* **110**, 10 682–10 686. (doi:10.1073/pnas.1300592110)
117. Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. 2005 Chance caught on the wing, *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**, 481–487. (doi:10.1038/nature03235)
118. Arnould L *et al.* 2013 Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* **339**, 1423–1426. (doi:10.1126/science.1233749)
119. Prud'homme B *et al.* 2006 Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050–1053. (doi:10.1038/nature04597)
120. Jeong S, Rokas A, Carroll SB. 2006 Regulation of body pigmentation by the abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* **125**, 1387–1399. (doi:10.1016/j.cell.2006.04.043)
121. Frankel N, Erezylmaz DF, McGregor AP, Wang S, Payre F, Stern DL. 2011 Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* **474**, 598–603. (doi:10.1038/nature10200)
122. McGregor AP *et al.* 2007 Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* **448**, 587–590. (doi:10.1038/nature05988)
123. Frankel N, Wang S, Stern DL. 2012 Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc. Natl Acad. Sci. USA* **109**, 20 975–20 979. (doi:10.1073/pnas.1207715109)
124. Jeong S, Rebeiz M, Andolfatto P, Werner T, True J, Carroll SB. 2008 The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* **132**, 783–793. (doi:10.1016/j.cell.2008.01.014)
125. Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB. 2009 Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* **326**, 1663–1667. (doi:10.1126/science.1178357)
126. Stern DL, Frankel N. 2013 The structure and evolution of *cis*-regulatory regions: the *shavenbaby* story. *Phil. Trans. R. Soc. B* **368**, 20130028. (doi:10.1098/rstb.2013.0028)
127. Woltering JM *et al.* 2009 Axial patterning in snakes and caecilians, evidence for an alternative interpretation of the Hox code. *Dev. Biol.* **332**, 82–89. (doi:10.1016/j.ydbio.2009.04.031)
128. Vinagre T, Moncaut N, Carapuco M, Novoa A, Bom J, Mallo M. 2010 Evidence for a myotomal Hox/Myf cascade governing nonautonomous control of rib specification within global vertebral domains. *Dev. Cell* **18**, 655–661. (doi:10.1016/j.devcel.2010.02.011)
129. Mansfield JH. 2013 *Cis*-regulatory change associated with snake body plan evolution. *Proc. Natl Acad. Sci. USA* **110**, 10 473–10 474. (doi:10.1073/pnas.1307781110)
130. Lowe CB *et al.* 2011 Three periods of regulatory innovation during vertebrate evolution. *Science* **333**, 1019–1024. (doi:10.1126/science.1202702)
131. Lynch VJ, Leclerc RD, May G, Wagner GP. 2011 Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* **43**, 1154–1159. (doi:10.1038/ng.917)
132. Emera D, Wagner GP. 2012 Transposable element recruitments in the mammalian placenta, impacts and mechanisms. *Brief Funct. Genomics* **11**, 267–276. (doi:10.1093/bfpg/els013)
133. Cotney J *et al.* 2013 The evolution of lineage-specific regulatory activities in the human embryonic limb. *Cell* **154**, 185–196. (doi:10.1016/j.cell.2013.05.056)
134. Schneider I, Shubin NH. 2012 Making limbs from fins. *Dev. Cell* **23**, 1121–1122. (doi:10.1016/j.devcel.2012.11.011)
135. Sumiyama K, Saitou N. 2011 Loss-of-function mutation in a repressor module of human-specifically activated enhancer HACNS1. *Mol. Biol. Evol.* **28**, 3005–3007. (doi:10.1093/molbev/msr231)
136. Sucena E, Stern DL. 2000 Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by *cis*-regulatory evolution of *ovo/shaven-baby*. *Proc. Natl Acad. Sci. USA* **97**, 4530–4534. (doi:10.1073/pnas.97.9.4530)
137. Carbone MA, Llopart A, deAngelis M, Coyne JA, Mackay TF. 2005 Quantitative trait loci affecting the difference in pigmentation between *Drosophila yakuba* and *D. santomea*. *Genetics* **171**, 211–225. (doi:10.1534/genetics.105.044412)
138. Pool JE, Aquadro CF. 2007 The genetic basis of adaptive pigmentation variation in *Drosophila melanogaster*. *Mol. Ecol.* **16**, 2844–2851. (doi:10.1111/j.1365-294X.2007.03324.x)
139. Glaser-Schmitt A, Catalán A, Parsch J. 2013 Adaptive divergence of a transcriptional enhancer between populations of *Drosophila melanogaster*. *Phil. Trans. R. Soc. B* **368**, 20130024. (doi:10.1098/rstb.2013.0024)
140. Preuss TM. 2012 Human brain evolution, from gene discovery to phenotype discovery. *Proc. Natl Acad. Sci. USA* **109**(Suppl. 1), 10 709–10 716. (doi:10.1073/pnas.1201894109)
141. Kamm GB, Pisciotto F, Kliger R, Franchini LF. 2013 The developmental brain gene NPAS3 contains the largest number of accelerated regulatory sequences in the human genome. *Mol. Biol. Evol.* **30**, 1088–1102. (doi:10.1093/molbev/mst023)
142. Prabhakar S *et al.* 2008 Human-specific gain of function in a developmental enhancer. *Science* **321**, 1346–1350. (doi:10.1126/science.1159974)
143. Capra JA, Erwin GD, McKinsey G, Rubenstein JLR, Pollard KS. 2013 Many human accelerated regions are developmental enhancers. *Phil. Trans. R. Soc. B* **368**, 20130025. (doi:10.1098/rstb.2013.0025)
144. May D *et al.* 2011 Large-scale discovery of enhancers from human heart tissue. *Nat. Genet.* **44**, 89–93. (doi:10.1038/ng.1006)
145. Kamm GB, López-Leal R, Lorenzo JR, Franchini LF. 2013 A fast-evolving human NPAS3 enhancer gained reporter expression in the developing forebrain of transgenic mice. *Phil. Trans. R. Soc. B* **368**, 20130019. (doi:10.1098/rstb.2013.0019)
146. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. 2013 One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918. (doi:10.1016/j.cell.2013.04.025)
147. Gratz SJ *et al.* 2013 Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029–1035. (doi:10.1534/genetics.113.152710)
148. Cade L *et al.* 2012 Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.* **40**, 8001–8010. (doi:10.1093/nar/gks518)
149. Sung YH, Baek IJ, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim JS, Lee HW. 2013 Knockout mice created by TALEN-mediated gene targeting. *Nat. Biotechnol.* **31**, 23–24. (doi:10.1038/nbt.2477)
150. Ansay S, Sakuma T, Yamamoto T, Ariga H, Uemura N, Takahashi R, Kinoshita M. 2013 Efficient targeted mutagenesis in medaka using custom-designed transcription activator-like effector nucleases. *Genetics* **193**, 739–749. (doi:10.1534/genetics.112.147645)
151. Jao LE, Wente SR, Chen W. 2013 Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl Acad. Sci. USA* **110**, 13 904–13 909. (doi:10.1073/pnas.1308335110)
152. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK. 2013 Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229. (doi:10.1038/nbt.2501)