The evolutionary significance of *cis*-regulatory mutations

Gregory A. Wray

Abstract | For decades, evolutionary biologists have argued that changes in *cis*-regulatory sequences constitute an important part of the genetic basis for adaptation. Although originally based on first principles, this claim is now empirically well supported: numerous studies have identified *cis*-regulatory mutations with functionally significant consequences for morphology, physiology and behaviour. The focus has now shifted to considering whether *cis*-regulatory and coding mutations make qualitatively different contributions to phenotypic evolution. Cases in which parallel mutations have produced parallel trait modifications in particular suggest that some phenotypic changes are more likely to result from *cis*-regulatory mutations than from coding mutations.

Cis-regulatory region
A segment of DNA that regulates transcription; such segments typically lie immediately 5′ of the start site of transcription, but are often discontinuous, and individual segments can reside within introns, 5′ and 3′ UTRs, or tens of kilobases on either side of the gene they regulate.

Clade
A group of species that share a unique common ancestor.

Department of Biology and Institute for Genome Sciences & Policy, Duke University, Durham, North Carolina 27708, USA.
e-mail: gwray@duke.edu
doi:10.1038/nrg2063

Ideas about the evolutionary significance of non-coding mutations are nearly as old as the discovery of regulatory sequences themselves. Soon after publishing their ground-breaking paper describing the *lac* operon in 1961 (REF. 1), Jacob and Monod speculated about the unique role that mutations in operators (their term for *cis*-regulatory regions) might have during the course of evolution[2]. They based their arguments on the recognition that the proper function of every gene depends on two distinct components: what its product does and the circumstances under which that product is produced. A perfectly good enzyme can be useless or even counterproductive, they argued, if synthesized under the wrong conditions.

The experimental tools needed to test these ideas lay decades in the future. Nonetheless, two influential papers published during the 1970s argued, on the basis of indirect evidence, that *cis*-regulatory mutations might have an important role in evolution. The first, published by Britten and Davidson in 1971 (REF. 3), was stimulated by the discovery that a substantial proportion of many eukaryotic genomes is composed of repetitive sequences. Britten and Davidson proposed that repetitive sequences regulate transcription; they advanced the first model for the evolution of regulatory sequences and argued that regulatory mutations play a crucial part in phenotypic evolution. The second influential paper, published by King and Wilson in 1975 (REF. 4), was motivated by the realization that homologous proteins in humans and chimpanzees are nearly identical. King and Wilson argued that the modest degree of divergence in protein sequence cannot account for the profound phenotypic differences between the species, and proposed instead that regulatory mutations must be primarily responsible.

With the benefit of 30 years of hindsight, it is clear that some early ideas about regulatory evolution were wrong in their detail. But it is equally evident today that the basic postulate was sound: mutations within *cis*-regulatory regions underlie a variety of interesting and ecologically significant phenotypic differences in morphology, physiology and behaviour (TABLE 1). It is now possible to ask whether coding and regulatory mutations make qualitatively distinct contributions to phenotypic evolution. In this Review, I examine the significance of *cis*-regulatory mutations from the perspective of evolutionary genetics. As such, the focus is on cases in which the genetic basis for a trait is understood in detail and on phenotypic divergence within populations and among closely related species. (Large-scale phenotypic changes, such as modifications in body plan, are notably interesting, but exceedingly difficult to dissect genetically.) I begin by considering how *cis*-regulatory and coding mutations might differ functionally and evolutionarily, then examine several cases in which *cis*-regulatory mutations have altered ecologically relevant traits in three well-studied clades, and end by asking whether *cis*-regulatory mutations have a qualitatively distinct role in phenotypic evolution.

What makes *cis*-regulatory mutations different?

Despite burgeoning interest in the evolution of gene expression during the past several years, most of what we know about the phenotypic impact and fitness

Table 1 | **Cis-regulatory mutations with interesting phenotypic consequences***

Gene	Function of product	Phenotype	Taxon	References
AVPR1A	Vasopressin receptor	Creative dance performance	Humans	82
Avpr1a	Vasopressin receptor	Paternal care	Rodents	83
Cyp6G1	P450 enzyme	Pesticide resistance	Fruitflies	84
DARC	Chemokine receptor	Resistance to infection with malaria	Humans	58,59
e	Pigment synthesis	Colour pattern of abdomen	Fruitflies	25
hsp70	Heat shock protein	Thermal tolerance	Fruitflies	85,86
HTR2A	Serotonin receptor	Obsessive-compulsive behaviour	Humans	87
IL10	Interleukin	Outcome of infection with HIV and infection with leprosy	Humans	88,89
IL10	Interleukin	Susceptibility to schizophrenia	Humans	90
LCT	Digestive enzyme	Lactose persistence	Humans	64,81
LDH	Metabolic enzyme	Cardiac physiology	Killifish	91
ovo/svb	Transcription factor	Bristle pattern on larvae	Fruitflies	34,36
MAOA	Neurotransmitter turnover	Aggressive behaviour	Humans	92,93
MMP3	Matrix metalloprotease	Risk of heart disease	Humans	94,95
PDYN	Neuropeptide	Memory, emotional status	Humans	71,76
pitx1	Transcription factor	Skeletal patterning	Stickleback fish	37,39,40
sc	Transcription factor	Bristle pattern on adult notum	Fruitflies	35,96
SLC6A4	Serotonin transporter	Depression, creativity, anxiety	Humans	82,97
SLC6A4	Serotonin transporter	Dispersal behaviour	Macaques	98
tb	Transcription factor	Branching structure	Maize	99,100
Ubx	Transcription factor	Bristle pattern on adult legs	Fruitflies	101
y	Pigment synthesis	Colour pattern of cuticle Mating behaviour	Fruitflies	25–28 102

*A sampling of cases in which the genetic basis for a trait difference is known to be *cis*-regulatory; well over 100 *cis*-regulatory mutations that segregate in human populations are known to affect diverse aspects of behaviour, physiology and disease susceptibility[103,104], and only a few examples are listed here. *AVPR1A*, arginine vasopressin receptor 1A; *DARC*, Duffy blood group, chemokine receptor; *e*, *ebony*; *hsp70*, *heat shock protein 70*; *HTR2A*, 5-hydroxytryptamine (serotonin) receptor 2A; *IL*, interleukin; *LCT*, lactase; *MAOA*, monoamine oxidase A; *MMP3*, matrix metallopeptidase 3; *PDYN*, prodynorphin; *pitx1*, *paired-like homeodomain transcription factor 1*; *sc*, *scute*; *SLC6A4*, solute carrier family 6 member 4; *tb*, *teosinte branched*; *Ubx*, *Ultrabithorax*; *y*, *yellow*.

consequences of mutations is still based on studies of coding sequences[5–8]. This is because the genetic code makes it easy to identify, accurately and comprehensively, mutations that alter protein sequences from DNA sequence comparisons alone (that is, non-synonymous substitutions, frameshifts, premature stop codons), whereas the same is not true of mutations that alter transcription, splicing, transcript stability and other regulatory processes (FIG. 1). Most regulatory mutations can only be identified through functional or biochemical tests[5,6,8], and are consequently substantially under-represented in evolutionary studies. Of course, not every mutation that alters molecular structure or function also affects organismal phenotype, but the reverse is true: the only mutations that can change traits such as morphology, physiology and behaviour are those that alter the structure and function of macromolecules. The difference in our ability to identify potentially functional mutations from coding and regulatory sequences makes it difficult to estimate their relative contribution to the evolution of organismal phenotypes.

A more interesting and tractable question is whether regulatory mutations make a qualitatively distinct contribution to phenotypic evolution. Several authors have argued this point from examples and first principles[1–8]. Their arguments fall into two basic categories that are not mutually exclusive: that *cis*-regulatory mutations are intrinsically more likely to affect certain kinds of phenotypic traits, and that selection operates more efficiently on *cis*-regulatory mutations.

The first hypothesis proposes that some kinds of phenotypic difference are easier to achieve through *cis*-regulatory mutations than through coding mutations[1,5–7]. The crux of this argument is that transcription is a dynamic process that can be 'fine-tuned' to meet context-dependent functional demands, whereas structure is generally more static. Many aspects of organimsal phenotype require dynamic changes in gene function, including reproduction, development, behaviour, immune responses and resource utilization. Traits associated with dynamic processes such as these might be expected to evolve to some extent more readily through regulatory rather than coding mutations.

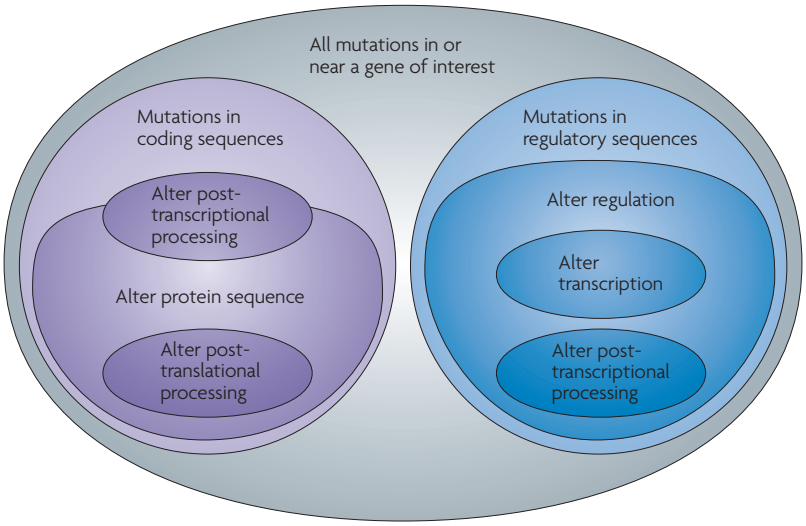

Figure 1 | **Functional classes of mutations.** Although most mutations in and around a gene have no functional consequence (grey area), a subset occur in coding or regulatory regions (purple and blue areas, respectively). Only some of these mutations have functional consequences at a molecular level. Of these, the only mutations that can be reliably and exhaustively identified through sequence comparisons are those that result in an amino-acid substitution. Mutations that affect regulation of all kinds (transcription, mRNA splicing and stability, and post-translational modifications) are generally difficult to identify without functional tests. As a result, the vast majority of evolutionary analyses focus on non-synonymous substitutions, while the evolutionary consequences of other functional classes of mutations remain poorly studied.

Indeed, some authors have argued that the evolution of complex multicellular organisms would have been all but impossible in the absence of *cis*-regulatory systems that allowed context-dependent transcriptional regulation[9,10]. Of course, protein structure is not entirely static: many genes produce distinct isoforms through alternative transcription start sites and splicing and through post-translational modifications. Even so, the number of isoforms is generally small and each represents a discrete state, whereas expression is a continuous variable that can be adjusted in fine increments across a broad dynamic range. C*is*-regulatory mutations might consequently play a disproportionate part in the evolution of quantitative traits and of responses to environmental factors that vary over time, such as stressors, resources and pathogens.

The second hypothesis proposes that natural selection operates differently on mutations in *cis*-regulatory sequences[6–8,11]. This hypothesis is based on two properties of the organization and function of *cis*-regulatory regions. First, allele-specific measures of transcript abundance indicate that each allele in a diploid organism is transcribed largely independently[11–14], suggesting that mutations in *cis*-regulatory regions are often co-dominant. By contrast, many or most coding mutations are recessive[15]. Natural selection operates far more efficiently on co-dominant mutations because they can have fitness consequences as heterozygotes: a new variant is visible to selection immediately rather than requiring drift to raise allele frequencies to the point at which homozygotes begin to appear in the population[11]. To the extent that *cis*-regulatory mutations are more commonly co-dominant than coding mutations, the efficiency of natural selection will differ between these mutational classes. Second, the modular organization of some *cis*-regulatory regions[6,7] means that a mutation in one module might affect only one part of the overall transcription profile[6,7,16]. For instance, the effects of a *cis*-regulatory mutation could be limited to larval anatomy without affecting the adult, or to a single organ or tissue even when the gene is much more widely expressed (examples are discussed below). By contrast, most non-synonymous coding mutations change the resulting protein no matter where it is expressed. (Alternative splicing could in principle limit pleiotropy caused by a non-synonymous substitution in proteins with modular organization, but few clear examples are known.) Reduced pleiotropy allows selection to operate more efficiently by minimizing functional trade-offs[5–7,16]. This might be particularly relevant for genes expressed in a variety of cell types and tissues.

Testing these ideas directly is not straightforward. Nonetheless, we can begin to evaluate whether they are consistent with a rapidly growing body of detailed case studies (TABLE 1). The following sections consider some of the most thorough analyses that have been carried out on the contribution of *cis*-regulatory mutations to organismal phenotypes in three clades. These studies highlight how simple mutations within *cis*-regulatory regions can change transcription in ways that alter ecologically relevant traits.

Pigmentation (and bristles) in fruitflies

The first set of instructive cases concerns the evolution of pigmentation patterns within the cuticle of fruitflies, a set of traits that differ between species and have fitness consequences that are probably based on crypsis, thermoregulation and mate choice[17–20]. The core enzymatic pathway that produces black, brown and tan pigments from tyrosine is well defined[21–25] (FIG. 2a). Work by Carroll and colleagues has shown two genes in this pathway to be particularly interesting from an evolutionary perspective: *yellow* (*y*) and *ebony* (*e*). Yellow protein, which synthesizes black melanin, is produced in a pattern that closely prefigures dark pigmentation[24,25]. This correlation is consistent among species with different pigmentation[26–29], suggesting that the spatial distribution of yellow is a major factor in determining the final pattern of black pigment. By contrast, production of ebony does not generally correlate with pigment patterns[29]; instead, ebony is present at low levels throughout the cuticle and produces tan pigment where black or brown melanins are not synthesized. At higher levels of expression, however, ebony acts antagonistically to yellow, repressing synthesis of black melanin. A number of regulatory inputs affect the activity of the pigmentation pathway[29], including activation by *Abdominal B* (*Abd-B*) and repression by *bric a brac 1/2* (*bab1*/*bab2*) to produce sex-specific patterns[17,20,27].

Co-dominant
A mutation that has an additive phenotypic impact, and is therefore apparent in heterozygotes.

Pleiotropy
The ability of a gene or mutation to alter more than one trait.

Functional trade-off
For many traits, improving one aspect of function might incur a cost in some other aspect of function.

Crypsis
Concealment from predators, usually through shape and colouration of the integument.

Candidate gene
A gene that seems likely, on the basis of its function or a prior association study, to contain a mutation or mutations that underlie a phenotypic trait of interest.

Abdominal pigment patterns. Using a candidate gene approach, Wittkopp and colleagues[25] demonstrated that differences in abdominal pigmentation in the *Drosophila* species *D. melanogaster*, *D. subobscura* and *D. virilis* are due in part to mutations within the *cis*-regulatory sequences of *y*. Transformation of the *y* locus from *D. subobscura* and *D. virilis* into a strain of *D. melanogaster* null for expression of yellow resulted in ectopic yellow production and matching melanin deposition characteristic of these other species. These results suggest that the causal mutations for pigmentation differences are regulatory and lie *cis* to the *y* locus. They also suggest that independent regulatory mutations near the same gene can affect diverse manifestations of the same aspect of phenotype as these species differ in abdominal pigmentation. However, not all mutations affecting species-specific *y* expression are *cis*-regulatory, as shown by reciprocal transformations, which indicate a role for mutations *trans* to the *y* locus. Whether these additional mutations reside in coding or non-coding sequences is not known, but they act by altering *y* transcription.

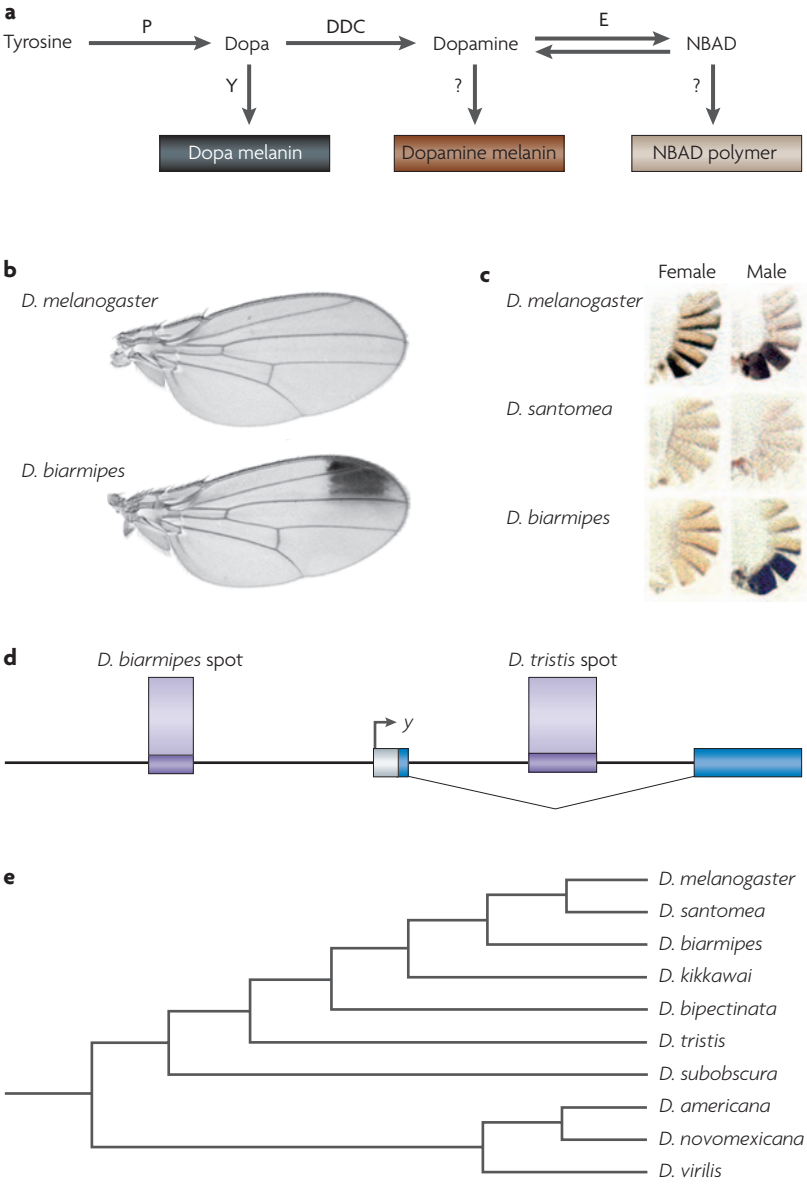

Figure 2 | **Evolution of cuticular pigmentation in *Drosophila* species. a** | Pigment synthesis pathway. Dark pigmentation in fruitflies is predominantly due to black and brown melanins, which are derived from dopa and dopamine, respectively, whereas lighter, tan pigmentation is a polymer of *N*-β-alanyl dopamine (NBAD). **b**,**c** | Species-specific differences in pigmentation in wings and abdomens within the *Drosophila* genus[26,27]. Both sites of pigmentation are sexually dimorphic in some species. **d** | Genomic organization of the *yellow* (*y*) locus. Blue shading indicates exons, grey shading indicates the 5′ UTR. The location of *cis*-regulatory mutations that resulted in the evolution of the anterodorsal wing spot (as in *Drosophila tristis*) are indicated by the purple regions. **e** | Phylogenetic relationships among the *Drosophila* species discussed in the text[28,80]. DDC, Dopa decarboxylase; E, ebony; P, Pale. Images in panel **b** reproduced with permission from *Nature* REF. 26 © (2005) Macmillan Publishers Ltd. Images in panel **c** reproduced with permission from REF. 27 © (2006) Cell Press.

The divergence between *D. virilis* and *D. melanogaster* occurred ~65 million years ago[30], making it difficult to reconstruct the sequence of events that led to divergence in pigment patterns. Analysing closely related species can help to resolve how these differences became established. Abdominal pigmentation differs between *Drosophila americana* and *Drosophila novomexicana*, which diverged just a few million years ago[31] (FIG. 2e). In this case, a combination of QTL mapping and a candidate gene approach ruled out the possibility that mutations in *y* are responsible. Instead, four other loci account for most of the differences, one of which is *e*[32]. Further examples that are particularly interesting involve evolutionary changes in sexual dimorphism of abdominal pigmentation, which are due in part to changes in *bab1*/*bab2* and *Abd-B* transcription[17,19]. Although most pigmentation differences among species correlate well with altered *bab1*/*bab2* transcription, a few do not, indicating that mutations at other loci also affect this trait. In none of these comparisons of closely related species is it known whether the causal mutations are regulatory or coding, but they clearly act by altering transcription of *y*. They also demonstrate that mutations at multiple points in the pigmentation pathway and in its regulatory inputs can produce parallel phenotypic changes.

The precise genetic basis for pigmentation differences is clearer in the Oriental *D. melanogaster* species group. Male-specific dark abdominal pigmentation is widespread but not universal within this group: phylogenetic relationships suggest that this trait is ancestral for the Oriental species group but was subsequently lost at least three times (FIG. 2c). Jeong and colleagues transformed the *cis*-regulatory region of *y* from various species linked to a GFP reporter into *D. melanogaster*, which has sexually dimorphic abdominal pigmentation[27]. The *cis*-regulatory region from *Drosophila biarmipes*, which also has male-specific pigmentation, drives sexually dimorphic GFP expression, indicating that the set of *trans*-acting factors driving sex-specific transcription has been conserved between the two species. More surprising were the results from three species reconstructed as having lost sexually dimorphic pigmentation. The *cis*-regulatory regions from *Drosophila bipectinata* (with sexually monomorphic pigmentation) and *Drosophila santomea* (with no abdominal pigmentation) both drove sexually dimorphic GFP expression, whereas that of *Drosophila kikkawai* (also lacking abdominal pigmentation)

produced no GFP expression. Through additional experiments, Jeong *et al.* demonstrated that the loss of male-specific transcription in *D. kikkawai* is due in part to mutations within a specific ABD-B binding site (as well as other mutations) within the *cis*-regulatory region of *y*. By contrast, sexually monomorphic transcription in *D. bipectinata* is probably due to a change in the expression of *bab1*/*bab2*, with little or no contribution from mutations in the *cis*-regulatory region of *y*. These results demonstrate that the loss of sexually dimorphic pigmentation can happen through changes *cis* or *trans* to *y* that affect its transcription, and that parallel phenotypic changes can arise in different ways.

Wing pigment patterns. *Cis*-regulatory mutations have also been important in the evolution of wing pigmentation. Although the wings of *D. melanogaster* are notably plain, the genus encompasses a considerable diversity of pigmentation patterns (FIG. 2b). A phylogeny of the *melanogaster* and *obscura* species groups[28] indicates that anterodistal wing spots have evolved independently at least once within each group, and were secondarily lost at least four times within the *melanogaster* group. Gompel and colleagues[26] examined the genetic basis for the origin of a male-specific spot on the distal part of the wing of *D. biarmipes*. As with abdominal pigmentation, they found that expression of yellow precisely prefigures the wing spot. Using transformations, the authors demonstrated that the genetic basis for this change in *y* transcription resides within a specific portion of its *cis*-regulatory region (FIG. 2d). They identified specific mutations along the lineage leading to *D. biarmipes* that created *de novo* binding sites for transcription factors. However, ectopic *y* transcription is not sufficient to produce a pigment spot in transformed *D. melanogaster*, indicating that mutations *trans* to *y* were also important in the evolutionary origin of the spot; localized downregulation of ebony, allowing greater accumulation of black melanin, is part of this *trans* component. Taken together, the evidence suggests that the evolution of the anterodistal spot required mutations at several loci, at least some of which are *cis*-regulatory in nature.

Prud'homme and colleagues subsequently examined the independent gain of an anterodistal wing spot in *Drosophila tristis*, a member of the *subobscura* group[28]. Using transformations, they found that *cis*-regulatory changes affecting *y* transcription were again involved in producing the pigment spot. In this case, however, the causal mutations reside within a different module of the *cis*-regulatory region from that driving spot-specific transcription in *D. biarmipes* (FIG. 2d). So, what initially appears to be yet another example of a parallel genetic basis for a convergent phenotype on closer examination proves to be more subtle and striking: functionally similar changes in transcription evolved through mutations in different modules of the *cis*-regulatory apparatus of the same locus. Therefore, even if the genetic basis for a parallel phenotypic change maps to the same locus in two different clades, this does not mean that the underlying molecular basis is necessarily the same.

Trans
Located far away from the gene of interest; in practical terms, anywhere in the genome except nearby.

Macrochaete
The largest bristles on flies; their function is mechanosensory.

Bristle patterns. Before leaving *Drosophila*, another aspect of cuticle phenotype deserves particular mention, namely species-specific differences in the distribution of bristles[33]. Mutations within the *cis*-regulatory regions of several genes provide additional examples of evolutionary significance: in *Ultrabithorax* (*Ubx*) altering bristles on legs[34], in the *achaete–scute* (*ac–sc*) complex changing macrochaetes on the notum[33,35], and in *ovo*/*shaven baby* (*ovo*/*svb*) affecting the denticle bands of larvae[34,36]. The last case is especially interesting, because independent *ovo*/*svb* *cis*-regulatory mutations underlie parallel changes in larval bristle patterns[36], hinting at a situation that is similar to *y* and its contribution to parallel changes in pigmentation. Unlike pigmentation, however, causal nucleotides have not yet been identified for species-specific differences in bristle patterns.

Skeletal reduction in stickleback fish

Bones have long provided an important source of information about the grand sweep of vertebrate evolution; more recently they have begun to provide insights into microevolutionary processes as well. The second set of examples illustrating the evolutionary significance of *cis*-regulatory mutations come from a clade of small bony fish called sticklebacks.

Threespine sticklebacks. Kingsley and colleagues have established the threespine stickleback, *Gasterosteus aculeatus*, as a genetic model for vertebrate evolution[37], complementing an extensive body of behavioural and ecological studies[38]. *Gasterosteus aculeatus* is widely distributed in high temperate regions of the northern hemisphere (FIG. 3b). Although most individuals live in the ocean, populations have repeatedly invaded freshwater habitats. During the last glacial retreat (~10,000–20,000 years ago), hundreds of these populations were isolated in newly formed lakes, where they have independently adapted to a diversity of local environmental conditions from a common ancestral marine form[38–40].

Some of these independently isolated populations show a similar pattern of reduction or loss in skeletal armour, involving the dorsal spines and pelvic girdle (FIG. 3a), which is ecologically associated with reduced calcium and fewer large-gape predators. Shapiro and colleagues[40] crossed a marine, fully armoured fish with a freshwater fish from Paxton Lake in British Colombia that had reduced pelvic armour, and used F2 hybrids to map the genetic basis for pelvic reduction. One locus of major effect, in the vicinity of *paired-like homeodomain transcription factor 1* (*pitx1*), explains about half the variation in pelvic structures, and four modifier loci explain most of the remainder. *Pitx1* encodes a transcription factor expressed in the hindlimbs but not the forelimbs of mouse embryos[41], and whose null phenotype is a loss of hindlimbs[42]. The sequence of *pitx1* in Paxton Lake fish lacks non-synonymous substitutions relative to the marine population, implying that the causal mutation is regulatory. This conclusion is supported by mRNA localizations that show *pitx1* transcription in the pelvic region of marine larvae but not of those larvae from Paxton Lake[40]. Interestingly, other domains of *pitx1*

transcription, including the thymus and olfactory pits, are indistinguishable between the two populations. The specific mutation (or mutations) responsible for pelvic reduction in the Paxton Lake population has not been identified, but it is clearly *cis*-regulatory.

The existence of multiple, independently derived stickleback populations with pelvic reduction has provided an opportunity to examine the genetic basis for the parallel evolution of anatomical traits. Populations of *G. aculeatus* in Lake Vifilsstadavatn in Iceland and in Boot Lake, Whale Lake and Bear Paw Lake in Alaska must have evolved reduced pelvic structures independently because they were never part of the same lake system (FIG. 3b,c). Offspring from crosses between these populations also have highly reduced pelvic structures[39,40]. This failure to complement strongly suggests that mutations in the same genes are responsible for the derived phenotypes in each lake. Mapping studies with the three Alaskan populations point to a single locus of major effect in the vicinity of *pitx1*, reinforcing this conclusion[39]. Gene expression was not examined in any of these populations, so it remains possible that the causal mutations disrupt *pitx1* function by altering splicing or protein function. Nonetheless, it seems clear that closely linked (possibly even identical) mutations of major effect are primarily responsible for the parallel anatomical changes in these independently isolated lake populations.

Ninespine sticklebacks. Even more striking is a similar result from ninespine sticklebacks (*Pungitius pungitius*), where pelvic reduction has evolved independently in a biogeographical context that parallels that of threespine sticklebacks. *Pungitius pungitius* also has circumpolar marine populations that have given rise to multiple, recent freshwater populations isolated in lakes and ponds throughout the boreal northern hemisphere (for example, Fox Holes Lake). Threespine and ninespine sticklebacks both belong to the family Gasterosteidae[43], but last shared a common ancestor ~10 million years ago[44]. When individuals showing pelvic reduction from each of the two species were crossed, the offspring also showed pelvic reduction[45]. Because *cis*-regulatory mutations affecting *pitx1* transcription are causal in *G. aculeatus*, this result implies a similar specific causal basis in *P. pungitius* as well. Indeed, the amino-acid sequence encoded by *pitx1* is identical in the marine and pelvic-reduced populations of *P. pungitius*, but transcription is absent or very low in the pelvic-reduced populations, confirming that the causal mutations are *cis*-regulatory. Mutations in *pitx1* might not account for all cases of pelvic reduction in teleosts, however. In the pufferfish *Takifugu rubripes*, the basis appears to reside in the expression of *homeobox D9a* (*hoxd9a*), which acts upstream of *pitx1* during limb development[46].

Not surprisingly, the genetic basis for pelvic reduction in sticklebacks has been worked out in less detail than that of some of the pigmentation changes in fruitflies. On the other hand, the ecological context for pelvic reduction is clearer in sticklebacks and has been replicated in many 'natural experiments': parallel anatomical changes in ecologically similar lakes were probably driven by similar selective forces and have evolved over approximately the same interval of time. Another important contribution of the stickleback system has been demonstrating how *cis*-regulatory variants segregating

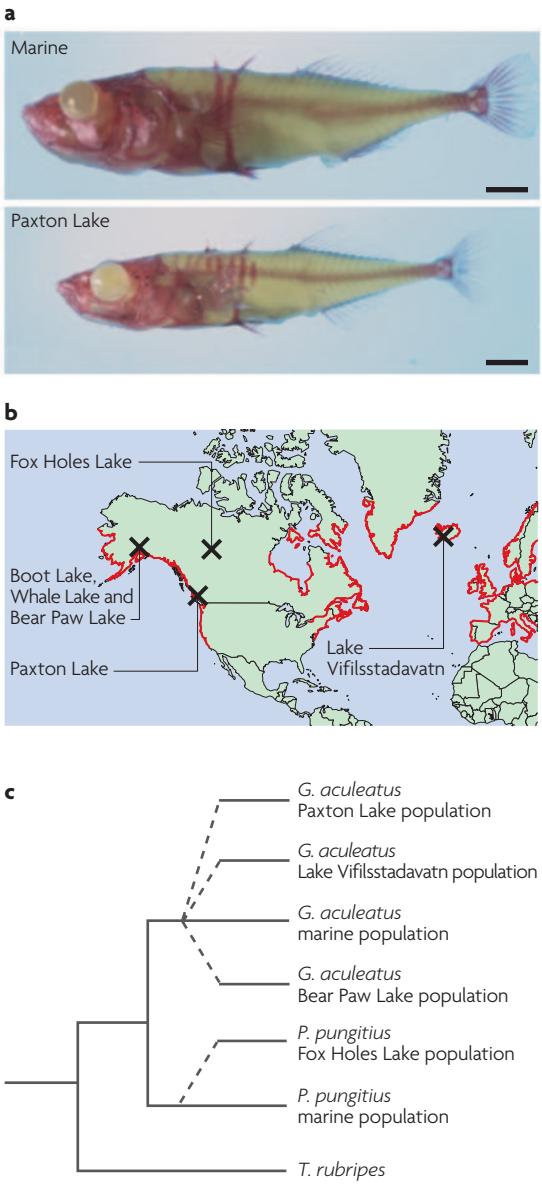

Figure 3 | **Evolution of armour reduction in stickleback fish. a** | Differences in pelvic armour among *Gasterosteus aculeatus* from the marine population and from Paxton Lake (British Columbia, Canada). **b** | Biogeographical distribution of *G. aculeatus* marine populations (red) and the location of the lakes discussed in the text where pelvic reduction has occurred independently in *G. aculeatus* and *Pungitius pungitius*. **c** | Phylogenetic relationships among *G. aculeatus* marine and lake populations, *P. pungitius* marine and lake populations, and *Takifugu rubripes*[43,44]. Note that each of the lake populations of *G. aculeatus* shown was derived independently as a peripheral isolate of the marine population[38–40]. Images in panel **a** reproduced with permission from *Nature* REF. 37 © (2001) Macmillan Publishers Ltd.

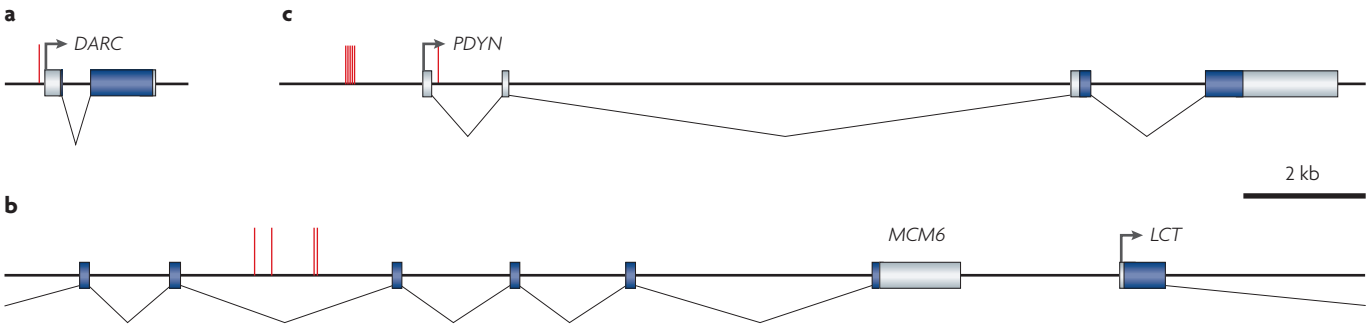

Figure 4 | ***Cis*-regulatory mutations with phenotypic consequences in humans. a** | The *DARC* (Duffy blood group, chemokine receptor) locus, showing the location of a mutation at –46 (red bar) that abolishes transcription in red blood cells[57,58]. **b** | The *LCT* (lactase) locus, showing the location of four mutations (red bars) that result in lactase persistence embedded within an intron of the *MCM6* (minichromosome maintenance deficient 6 homologue) locus[64,81]. **c** | The *PDYN* (prodynorphin) locus, showing the location of five mutations within the 68-bp repeat region and one mutation within the DREAM binding site (red bars). Genes are shown to the same scale; start sites of transcription are indicated by bent arrows, UTRs are indicated by grey boxes and coding regions are indicated by blue boxes.

in a wild population can contribute to phenotypic differences over microevolutionary timescales — a process that is poorly understood for *y* and fly pigmentation, but one that has also been studied in the next set of examples.

Physiology in humans

Humans and the other great apes provide the third set of instructive examples. Although these are difficult organisms to work with for a variety of reasons, they provide some of the most precisely delineated examples of mutations in *cis*-regulatory regions that have contributed to within- and between-species trait differences. This is partly because of the enormous resources directed towards biomedical research, which provide an unparalleled knowledge base regarding the genetic basis for complex trait variation. But, just as importantly, genetic divergence between and within species is small in comparison with many other well-studied organisms, facilitating the identification of causal mutations once a locus has been implicated. There are other reasons for studying the genetic basis for phenotypic evolution within the great apes, of course, chief among them a desire to understand our own origins[47,48].

Immune responses. One of the clearest examples of a significant *cis*-regulatory mutation in human evolution concerns the Duffy blood group, chemokine receptor (*DARC*, formerly known as *FY*) locus, which encodes a receptor that binds interleukin 8 (IL8), and other signalling molecules of the immune system[49–51]. *DARC* is transcribed in several tissues and cell types including erythrocytes, where it provides a point of entry for the malarial parasite *Plasmodium vivax*[52]. Certain *DARC* haplotypes segregating in modern human populations provide almost complete resistance to infection with *P. vivax*[53,54]. Resistance is due to the absence of Duffy protein expression in erythrocytes, but not in several other cells where it is normally expressed[55,56]. Individuals lacking *DARC* expression in erythrocytes show no adverse health consequences. The causal mutation is a *cis*-regulatory SNP (FIG. 4a), which disrupts binding of the transcription factor GATA binding protein 1 (GATA1) (REFS 57,58). This case provides a striking illustration of limited pleiotropy, involving a dramatic change in transcription within one expression domain and little or no impact on other expression domains of the same gene. It is also notable in that a single SNP results in a phenotype expected to accrue a substantial fitness gain. Consistent with this expectation, genetic variation in and around the *DARC* locus bears evidence of recent, strong positive selection within geographical regions where malaria is endemic[59,60]. *Cis*-regulatory mutations in genes encoding several other immune system components also show evidence of positive or balancing selection as well as phenotypic associations with pathogen interactions, including tumour-necrosis factor-α (*TNFA*), *IL4*, and *IL10* (TABLE 1).

Dietary changes. *Cis*-regulatory mutations have also contributed to traits that distinguish humans from other species. Among the great apes, the human diet is an outlier, involving a shift from largely herbivorous to omnivorous habits. An important recent human dietary adaptation is the ability to digest lactose as an adult, a condition known as lactase persistence[61]. This trait involves more than the mere ability to extract energy from lactose, the primary carbohydrate in milk; by eliminating indigestion due to lactose fermentation in the gut, it allows the substantially more nutritious lipid and protein components to be utilized without complications. Lactose persistence is thought to have evolved in association with pastoralism sometime during the past 2,000–20,000 years[62]. The enzyme that can catalyse this reaction is lactase-phlorizin hydrolase, making lactase (*LCT*), the gene that encodes it, an obvious candidate locus for lactose persistence. Ennatah and colleagues[62] identified the genetic basis for lactose persistence in northern Europeans as a SNP residing in an intron of minichromosome maintenance deficient 6 homologue (*MCM6*), the next gene 5′ of *LCT* (FIG. 4b). Experimental tests demonstrate that this SNP elevates *LCT* transcription[63]. However, this mutation is absent in some other

Pastoralism
The practice of tending domesticated animals for the milk they produce.

pastoral societies. Tishkoff and colleagues[64] identified three additional SNPs within the same *MCM6* intron that are genetically associated with lactose persistence in East African populations (FIG. 4b). By cloning several haplotypes of the *MCM6* intron independently fused to the core promoter into an expression vector and measuring expression in a relevant cell line, the authors demonstrated that each of these three mutations also increases *LCT* transcription. As with *y* in *Drosophila*, this case provides clear evidence that several independent mutations within the *cis*-regulatory region of a key enzyme have contributed to changes in an ecologically relevant trait.

Behaviour and cognition. Looking deeper into evolutionary time, some of the most interesting traits in human evolution concern behaviour and cognition, including language with syntax, extensive tool use, abstract reasoning and creativity expressed in myriad ways[65,66]. Several studies have compared transcript abundance from the brains of deceased chimpanzees and humans using microarrays. The first such analysis by Enard and colleagues indicated a greater degree of evolutionary change in transcription in brain than in liver, specifically on the human branch of a three-species comparison (human compared with chimpanzee, and rooted with macaque)[67]. Subsequent microrray studies have found that ~10% of genes examined differ in brain expression between humans and chimpanzees[68–70]. This is probably an underestimate, because only a few brain regions have been examined, and because microarrays can only reliably detect moderately large differences in transcript abundance. Also relevant to evolutionary analyses, microarrays do not indicate where the genetic basis for an expression difference resides: *cis* to the affected gene, *trans* to the gene in a transcription factor that regulates its expression, or even further upstream genetically.

In the case of prodynorphin (*PDYN*), the genetic basis for an expression difference between humans and chimpanzees has been identified as *cis*-regulatory[71]. *PDYN* encodes a precursor protein that is cleaved to release dynorphin, a neuropeptide with roles in memory, emotional status and perception of pain[72,73]. In humans, decreased *PDYN* expression is functionally associated with schizophrenia and bipolar disorder[74,75] and is genetically associated with schizophrenia and temporal lobe epilepsy[76,77]. Rockman and colleagues use expression assays in cultured neurons to demonstrate that human and chimpanzee haplotypes containing the *cis*-regulatory region each drive different levels of both constitutive and induced transcription[71]. Functional sites within the *cis*-regulatory region of *PDYN* sustained several mutations during human origins: five substitutions within a 68-bp region that regulates constitutive transcription and one mutation within the binding site for the transcription factor DREAM (also known as KCNIP3) that regulates induced expression (FIG. 4c). By testing chimeric human and chimpanzee *cis*-regulatory regions, Rockman and colleagues demonstrated that these six human branch-specific mutations account for most of the differences in constitutive and induced expression. Mutations within the *cis*-regulatory region of *PDYN* show signatures of positive selection during human evolution, as well as ongoing balancing selection among populations[71]. Several other genes in humans harbour *cis*-regulatory mutations that affect cognitive or behavioural traits, including arginine vasopressin receptor 1A (*AVPR1A*), 5-hydroxytryptamine (serotonin) receptor 2A (*HTR2A*), monoamine oxidase A (*MAOA*) and solute carrier family 6 member 4 (*SLC6A4*) (TABLE 1), although interspecies functional analyses have not been carried out.

Taking stock

We now have numerous clear examples of *cis*-regulatory mutations that have contributed to functionally significant and ecologically relevant traits. Besides the cases discussed above, these include mutations that affect a wide range of morphological, physiological and behavioural traits representing a broad taxonomic diversity (TABLE 1). Many of these cases were described only recently, and the list of cases is likely to grow rapidly as methods for identifying functional *cis*-regulatory mutations become more powerful and, particularly, as investigators become more attuned to looking for them. Evaluating the relative quantitative contribution of *cis*-regulatory and coding mutations to phenotypic evolution is largely beside the point: both have been important and mutations of either kind can contribute to a wide range of traits.

The more interesting question at this point is whether *cis*-regulatory mutations are qualitatively distinct in evolutionary terms. With the benefit of several well-defined cases, it is useful to revisit the arguments for such a role that were considered near the beginning of this article. One of these arguments is that some kinds of phenotype might be easier to achieve through *cis*-regulatory mutations than through coding mutations. The observation that parallel phenotypic changes in pigmentation and bristles in fruitflies, pelvic reduction in fish, and lactose persistence in humans have all been caused by parallel mutations in *cis*-regulatory regions adds weight to this argument. For these phenotypes, at least, it would seem that *cis*-regulatory mutations are either mutationally more likely or functionally more effective than coding mutations. The prediction that various mutations within the same *cis*-regulatory apparatus can achieve similar phenotypic consequences is particularly well illustrated by the different genetic bases for parallel abdominal and wing pigment patterns and by the multiple SNPs that can independently increase *LCT* transcription. Of course, this will not always be the case: parallel increases in overall melanization in vertebrates are often due to coding mutations in melanocortin 1 receptor (MC1R)[78,79]. (This contrast could reflect the different developmental bases for pigment patterns in the two groups, which are established by regulating enzyme synthesis in insects and through cell migration in vertebrates.) For some traits in some clades, however, the genetic basis seems to be primarily *cis*-regulatory.

The second argument is that selection could generally operate more efficiently on *cis*-regulatory mutations than on coding mutations for reasons of reduced pleiotropy

and incomplete dominance. Limited pleiotropy is evident in several of the examples considered. With *DARC*, an advantageous point mutation apparently affects transcription in just one crucial cell type out of several where it is expressed. Both *y* and *e* are crucial for pigment synthesis throughout the fly cuticle, but *cis*-regulatory mutations often alter restricted aspects of the overall pigment pattern; the same is true of *ovo/svb* and *sc* mutations affecting bristle patterns. Similarly, the *pitx1* haplotype that is primarily responsible for pelvic reduction in Paxton Lake sticklebacks has little effect on transcription in other domains of expression. Co-dominance is also clear in some cases. Individuals who are heterozygous for the *DARC* mutation that abolishes erythrocyte transcription show partial protection from malarial infection, while the degree of lactase persistence in humans and pelvic reduction in sticklebacks are a function of both *cis*-regulatory haplotypes. Therefore, reduced pleiotropy and incomplete dominance have been observed in cases in which *cis*-regulatory mutations have an ecologically significant phenotypic impact.

A conspicuous theme that runs through many of the cases discussed here is the degree to which parallel phenotypic changes can have parallel genetic bases. Not only are mutations in the same gene often responsible, but they are often repeatedly *cis*-regulatory rather than coding in nature. Independent mutations in the *cis*-regulatory regions of *y* and *LCT* have been responsible for parallel changes in morphology and physiology, respectively; the same is probably true of *pitx1*, *bab1*/*bab2* and *ovo/svb* although the precise mutations have not been identified in these cases. Interestingly, however, there are exceptions in certain instances: some parallel changes in fly pigmentation map to genes other than *y*, *e* and *bab1*/*bab2*, and at least one other case of pelvic reduction in teleosts involves changes genetically upstream of *pitx1* during hindlimb development[46]. These exceptions are a useful reminder that phenotypic evolution is a probabilistic process, even if parallel traits are sometimes caused by mutations in the same gene.

Future directions

The number of well-documented cases in which *cis*-regulatory mutations have contributed to interesting organismal traits has grown rapidly during the past few years. Future studies can now address gaps in our understanding of their evolutionary significance. One pressing issue is generality, as most existing examples come from a handful of systems and phenotypes (TABLE 1). Are these cases unusual, or will it be possible to expand the phylogenetic and phenotypic breadth of well-documented cases? A second important issue is the genetic and molecular basis for changes in gene expression in more cases. Although causal mutations have been identified in some cases, more often a trait difference is known to be *cis*-regulatory but the genetic and mechanistic bases remain unknown. This information can help to clarify evolutionary processes (was one mutation or several required?) and molecular processes (how does a change in transcription produce trait differences?). A third pressing issue is the ecological and population genetic contexts that lead to the causal mutation or mutations becoming established in a population. This information, which is lacking in most of the cases listed in TABLE 1, provides insights into why particular trait differences evolved. Fortunately, all three of these issues are tractable using existing technology. Connecting the dots between segregating variation, molecular consequences, organismal traits, evolutionary mechanisms and ecological contexts will significantly enrich our understanding of the evolutionary significance of *cis*-regulatory mutations, and of the genetic basis for biological diversity in general.

1. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
2. Monod, J. & Jacob, F. General conclusions — teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 389–401 (1961).
3. Britten, R. J. & Davidson, E. H. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q. Rev. Biol.* **46**, 111–138 (1971).
4. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
This classic paper proposed that phenotypic differences between humans and chimpanzees are largely due to *cis*-regulatory mutations.
5. Carroll, S. B., Grenier, J. & Weatherbee, S. *From DNA to Diversity* (Blackwell Publishing, Malden, 2004).
6. Stern, D. L. Evolutionary developmental biology and the problem of variation. *Evolution* **54**, 1079–1091 (2000).
7. Wilkins, A. S. *The Evolution of Developmental Pathways* (Sinauer Associates, Sunderland, 2002).
8. Wray, G. A. *et al.* The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**, 1377–1419 (2003).
9. Davidson, E. H. *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Academic, Burlington, 2006).
10. Gerhart, J. & Kirschner, M. *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability* (Blackwell Science, Malden, 1997).
11. Ruvkun, G., Wightman, B., Burglin, T. & Arasu, P. Dominant gain-of-function mutations that lead to misregulation of the *C. elegans* heterochronic gene *lin-14*, and the evolutionary implications of dominant mutations in pattern-formation genes. *Dev. Suppl.* **1**, 47–54 (1991).
12. Pastinen, T. *et al.* A survey of genetic and epigenetic variation affecting human gene expression. *Physiol. Genomics* **16**, 184–193 (2003).
13. Ronald, J., Brem, R. B., Whittle, J. & Kruglyak, L. Local regulatory variation in *Saccharomyces cerevisiae*. *PLoS Genet.* **1**, e25 (2005).
14. Wittkopp, P. J., Haerum, B. K. & Clark, A. G. Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430**, 85–88 (2004).
15. Li, W. H. *Molecular Evolution* (Sinauer Associates, Sunderland, 1997).
References 13–15 measured the relative contributions of mutations in *cis* and *trans* to differences in transcription.
16. Force, A. *et al.* Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999).
17. Gompel, N. & Carroll, S. B. Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. *Nature* **424**, 931–935 (2003).
18. Hollocher, H., Hatcher, J. L. & Dyreson, E. G. Genetic and developmental analysis of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. *Evolution* **54**, 2057–2071 (2000).
19. Kopp, A., Duncan, I., Godt, D. & Carroll, S. B. Genetic control and evolution of sexually dimorphic characters in *Drosophila*. *Nature* **408**, 553–559 (2000).
20. Kopp, A. & True, J. R. Evolution of male sexual characters in the Oriental *Drosophila melanogaster* species group. *Evol. Dev.* **4**, 278–291 (2002).
21. Hovemann, B. Tissue specific expression of the *ebony* gene. *J. Neurogenet.* **7**, 128–128 (1991).
22. Hovemann, B. T. *et al.* The *Drosophila ebony* gene is closely related to microbial peptide synthetases and shows specific cuticle and nervous system expression. *Gene* **221**, 1–9 (1998).
23. True, J. R. *et al. Drosophila tan* encodes a novel hydrolase required in pigmentation and vision. *PLoS Genet.* **1**, e63 (2005).
24. Walter, M. F. *et al.* Temporal and spatial expression of the *yellow* gene in correlation with cuticle formation and dopa decarboxylase activity in *Drosophila* development. *Dev. Biol.* **147**, 32–45 (1991).
25. Wittkopp, P. J., True, J. R. & Carroll, S. B. Reciprocal functions of the *Drosophila* yellow and ebony proteins in the development and evolution of pigment patterns. *Development* **129**, 1849–1858 (2002).
26. Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. & Carroll, S. B. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**, 481–487 (2005).
27. Jeong, S., Rokas, A. & Carroll, S. B. Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* **125**, 1387–1399 (2006).
This study illustrates the power of using a combination of detailed molecular genetic analyses with an in-depth sampling of related species.

28. Prud'homme, B. *et al.* Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050–1053 (2006).
References 26 and 28 illustrate a remarkable case of similar genetic bases for a parallel change in wing pigmentation in flies.
29. Wittkopp, P. J., Vaccaro, K. & Carroll, S. B. Evolution of *yellow* gene regulation and pigmentation in *Drosophila*. *Curr. Biol.* **12**, 1547–1556 (2002).
30. Beverley, S. M. & Wilson, A. C. Molecular evolution in *Drosophila* and the higher Diptera II. A time scale for fly evolution. *J. Mol. Evol.* **21**, 1–13 (1984).
31. Vieira, J., Vieira, C. P., Hartl, D. L. & Lozovskaya, E. R. A framework physical map of *Drosophila virilis* based on P1 clones: applications in genome evolution. *Chromosoma* **106**, 99–107 (1997).
32. Wittkopp, P. J., Williams, B. L., Selegue, J. E. & Carroll, S. B. *Drosophila* pigmentation evolution: divergent genotypes underlying convergent phenotypes. *Proc. Natl Acad. Sci. USA* **100**, 1808–1813 (2003).
This study demonstrated that parallel morphological changes do not always have a similar underlying genetic basis.
33. Simpson, P., Woehl, R. & Usui, K. The development and evolution of bristle patterns in Diptera. *Development* **126**, 1349–1364 (1999).
34. Sucena, E. & Stern, D. L. 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 (2000).
35. Skaer, N. & Simpson, P. Genetic analysis of bristle loss in hybrids between *Drosophila melanogaster* and *D. simulans* provides evidence for divergence of *cis*-regulatory sequences in the *achaete–scute* gene complex. *Dev. Biol.* **221**, 148–167 (2000).
36. Sucena, E., Delon, I., Jones, I., Payre, F. & Stern, D. L. Regulatory evolution of *shavenbaby/ovo* underlies multiple cases of morphological parallelism. *Nature* **424**, 935–938 (2003).
37. Peichel, C. L. *et al.* The genetic architecture of divergence between threespine stickleback species. *Nature* **414**, 901–905 (2001).
38. Bell, M. A. & Foster, S. A. (eds) *The Evolutionary Biology Of The Threespine Stickleback* (Oxford Univ. Press, New York, 1994).
39. Cresko, W. A. *et al.* Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc. Natl Acad. Sci. USA* **101**, 6050–6055 (2004).
40. Shapiro, M. D. *et al.* Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**, 717–723 (2004).
This study used a combination of genetic mapping and expression analyses to demonstrate that *cis*-regulatory mutation(s) constitute the primary genetic basis for loss of pelvic armour in sticklebacks.
41. Shang, J., Luo, Y. & Clayton, D. A. *Backfoot* is a novel homeobox gene expressed in the mesenchyme of developing hind limb. *Dev. Dyn.* **209**, 242–253 (1997).
42. Marcil, A., Dumontier, E., Chamberland, M., Camper, S. A. & Drouin, J. *Pitx1* and *Pitx2* are required for development of hindlimb buds. *Development* **130**, 45–55 (2003).
43. McLennan, D. A. & Mattern, M. Y. The phylogeny of the Gasterosteidae: combining behavioral and morphological data sets. *Cladistics* **17**, 11–27 (2001).
44. Bell, M. A., Baumgartner, J. V. & Olson, E. C. Patterns of temporal change in single morphological characters of a miocene stickleback fish. *Paleobiology* **11**, 258–271 (1985).
45. Shapiro, M. D., Bell, M. A. & Kingsley, D. M. Parallel origins of pelvic reduction in vertebrates. *Proc. Natl Acad. Sci. USA* **103**, 3753–3758 (2006).
46. Tanaka, M. *et al.* Developmental genetic basis for the evolution of pelvic fin loss in the pufferfish *Takifugu rubripes*. *Dev. Biol.* **281**, 227–239 (2005).
47. Carroll, S. B. Genetics and the making of *Homo sapiens*. *Nature* **422**, 849–857 (2003).
48. Vallender, E. J. & Lahn, B. T. Positive selection on the human genome. *Hum. Mol. Genet.* **13**, R245–R254 (2004).
49. Chaudhuri, A. *et al.* Expression of the Duffy antigen in K562 cells. Evidence that it is the human erythrocyte chemokine receptor. *J. Biol. Chem.* **269**, 7835–7838 (1994).
50. Horuk, R. *et al.* A receptor for the malarial parasite *Plasmodium vivax*: the erythrocyte chemokine receptor. *Science* **261**, 1182–1184 (1993).
51. Tournamille, C. *et al.* Sequence, evolution and ligand binding properties of mammalian Duffy antigen/receptor for chemokines. *Immunogenetics* **55**, 682–694 (2004).
52. Pogo, A. O. & Chaudhuri, A. The Duffy protein: a malarial and chemokine receptor. *Semin. Hematol.* **37**, 122–129 (2000).
53. Hadley, T. J. & Peiper, S. C. From malaria to chemokine receptor: the emerging physiologic role of the Duffy blood group antigen. *Blood* **89**, 3077–3091 (1997).
54. Miller, L. H., Mason, S. J., Clyde, D. F. & McGinniss, M. H. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N. Engl. J. Med.* **295**, 302–304 (1976).
55. Chaudhuri, A., Polyakova, J., Zbrzezna, V. & Pogo, A. O. The coding sequence of *Duffy* blood group gene in humans and simians: restriction fragment length polymorphism, antibody and malarial parasite specificities, and expression in nonerythroid tissues in Duffy-negative individuals. *Blood* **85**, 615–621 (1995).
56. Peiper, S. C. *et al.* The Duffy antigen/receptor for chemokines (DARC) is expressed in endothelial cells of Duffy negative individuals who lack the erythrocyte receptor. *J. Exp. Med.* **181**, 1311–1317 (1995).
57. Iwamoto, S., Li, J., Sugimoto, N., Okuda, H. & Kajii, E. Characterization of the Duffy gene promoter: evidence for tissue-specific abolishment of expression in Fy(a-b-) of black individuals. *Biochem. Biophys. Res. Commun.* **222**, 852–859 (1996).
58. Tournamille, C., Colin, Y., Cartron, J. P. & Le Van Kim, C. Disruption of a GATA motif in the *Duffy* gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nature Genet.* **10**, 224–228 (1995).
59. Hamblin, M. T. & Di Rienzo, A. Detection of the signature of natural selection in humans: evidence from the *Duffy* blood group locus. *Am. J. Hum. Genet.* **66**, 1669–1679 (2000).
References 58 and 59 provided evidence regarding the molecular and evolutionary mechanisms, respectively, for adaptive evolution of a regulatory SNP that provides protection from malarial infection.
60. Hamblin, M. T., Thompson, E. E. & Di Rienzo, A. Complex signatures of natural selection at the *Duffy* blood group locus. *Am. J. Hum. Genet.* **70**, 369–383 (2002).
61. Swallow, D. M. Genetics of lactase persistence and lactose intolerance. *Annu. Rev. Genet.* **37**, 197–219 (2003).
62. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
63. Olds, L. C. & Sibley, E. Lactase persistence DNA variant enhances lactase promoter activity *in vitro*: functional role as a *cis* regulatory element. *Hum. Mol. Genet.* **12**, 2333–2340 (2003).
64. Tishkoff, S. A. *et al.* Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genet.* **39**, 31–40 (2007).
This study demonstrated a nearly parallel genetic basis for an important physiological adaptation during recent human evolution.
65. Fisher, S. E. & Marcus, G. F. The eloquent ape: genes, brains and the evolution of language. *Nature Rev. Genet.* **7**, 9–20 (2006).
66. Sikela, J. M. The jewels of our genome: the search for the genomic changes underlying the evolutionarily unique capacities of the human brain. *PLoS Genet.* **2**, e80 (2006).
67. Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* **296**, 340–343 (2002).
The authors demonstrated an elevated rate of change in the expression of genes within the brain on the branch leading to humans relative to that leading to chimpanzees.
68. Caceres, M. *et al.* Elevated gene expression levels distinguish human from non-human primate brains. *Proc. Natl Acad. Sci. USA* **100**, 13030–13035 (2003).
69. Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**, 1850–1854 (2005).
70. Khaitovich, P. *et al.* Regional patterns of gene expression in human and chimpanzee brains. *Genome Res.* **14**, 1462–1473 (2004).
71. Rockman, M. V. *et al.* Ancient and recent positive selection transformed opioid *cis*-regulation in humans. *PLoS Biol.* **3**, e387 (2005).
This study found evidence of positive selection on specific *cis*-regulatory mutations at a locus encoding a neuropeptide with roles in human cognition.
72. Rodgers, R. J. & Cooper, S. J. (eds) *Endorphins, Opiates, and Behavioural Processes* (Wiley, New York, 1988).
73. Wagner, J. J., Terman, G. W. & Chavkin, C. Endogenous dynorphins inhibit excitatory neurotransmission and block LTP induction in the hippocampus. *Nature* **363**, 451–454 (1993).
74. Hurd, Y. L. Subjects with major depression or bipolar disorder show reduction of prodynorphin mRNA expression in discrete nuclei of the amygdaloid complex. *Mol. Psychiatry* **7**, 75–81 (2002).
75. Peckys, D. & Hurd, Y. L. Prodynorphin and κ opioid receptor mRNA expression in the cingulate and prefrontal cortices of subjects diagnosed with schizophrenia or affective disorders. *Brain Res. Bull.* **55**, 619–624 (2001).
76. Stogmann, E. *et al.* A functional polymorphisms in the prodynorphin gene promoter is associated with temporal lobe epilepsy. *Ann. Neurol.* **51**, 260–263 (2002).
77. Ventriglia, M. *et al.* Allelic variation in the human prodynorphin gene promoter and schizophrenia. *Neuropsychobiology* **46**, 17–21 (2002).
78. Nachman, M. W., Hoekstra, H. E. & D'Agostino, S. L. The genetic basis of adaptive melanism in pocket mice. *Proc. Natl Acad. Sci. USA* **100**, 5268–5273 (2003).
79. Theron, E., Hawkins, K., Bermingham, E., Ricklefs, R. E. & Mundy, N. I. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. *Curr. Biol.* **11**, 550–557 (2001).
80. Kopp, A. Basal relationships in the *Drosophila melanogaster* species group. *Mol. Phylogenet. Evol.* **39**, 787–798 (2006).
81. Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nature Genet.* **30**, 233–237 (2002).
82. Bachner-Melman, R. *et al. AVPR1a* and *SLC6A4* gene polymorphisms are associated with creative dance performance. *PLoS Genet.* **1**, e42 (2005).
83. Hammock, E. A. & Young, L. J. Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* **308**, 1630–1634 (2005).
84. Daborn, P. J. *et al.* A single p450 allele associated with insecticide resistance in *Drosophila*. *Science* **297**, 2253–2256 (2002).
85. Lerman, D. N. & Feder, M. E. Naturally occurring transposable elements disrupt *hsp70* promoter function in *Drosophila melanogaster*. *Mol. Biol. Evol.* **22**, 776–783 (2005).
86. Lerman, D. N., Michalak, P., Helin, A. B., Bettencourt, B. R. & Feder, M. E. Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. *Mol. Biol. Evol.* **20**, 135–144 (2003).
87. Enoch, M. A. *et al. 5-HT2A* promoter polymorphism –1438G/A, anorexia nervosa, and obsessive-compulsive disorder. *Lancet* **351**, 1785–1786 (1998).
88. Moraes, M. O. *et al.* Interleukin-10 promoter single-nucleotide polymorphisms as markers for disease susceptibility and disease severity in leprosy. *Genes Immun.* **5**, 592–595 (2004).
89. Shin, H. D. *et al.* Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of *IL10*. *Proc. Natl Acad. Sci. USA* **97**, 14467–14472 (2000).
90. He, G. *et al.* Interleukin-10 –1082 promoter polymorphism is associated with schizophrenia in a Han Chinese sib-pair study. *Neurosci. Lett.* **394**, 1–4 (2006).
91. Crawford, D. L., Segal, J. A. & Barnett, J. L. Evolutionary analysis of TATA-less proximal promoter function. *Mol. Biol. Evol.* **16**, 194–207 (1999).
92. Caspi, A. *et al.* Role of genotype in the cycle of violence in maltreated children. *Science* **297**, 851–854 (2002).
93. Kim-Cohen, J. *et al. MAOA*, maltreatment, and gene–environment interaction predicting children's mental health: new evidence and a meta-analysis. *Mol. Psychiatry* **11**, 903–913 (2006).

94. Beyzade, S. *et al.* Influences of matrix metalloproteinase-3 gene variation on extent of coronary atherosclerosis and risk of myocardial infarction. *J. Am. Coll. Cardiol.* **41**, 2130–2137 (2003).
95. Ye, S. *et al.* Progression of coronary atherosclerosis is associated with a common genetic variant of the human stromelysin-1 promoter which results in reduced gene expression. *J. Biol. Chem.* **271**, 13055–13060 (1996).
96. Marcellini, S. & Simpson, P. Two or four bristles: functional evolution of an enhancer of scute in drosophilidae. *PLoS Biol.* **4**, e386 (2006).
97. Hariri, A. R. *et al.* Serotonin transporter genetic variation and the response of the human amygdala. *Science* **297**, 400–403 (2002).
98. Trefilov, A., Berard, J., Krawczak, M. & Schmidtke, J. Natal dispersal in rhesus macaques is related to serotonin transporter gene promoter variation. *Behav. Genet.* **30**, 295–301 (2000).
99. Clark, R. M., Wagler, T. N., Quijada, P. & Doebley, J. A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. *Nature Genet.* **38**, 594–597 (2006).
100. Wang, R. L., Stec, A., Hey, J., Lukens, L. & Doebley, J. The limits of selection during maize domestication. *Nature* **398**, 236–239 (1999).
101. Stern, D. L. A role of Ultrabithorax in morphological differences between *Drosophila* species. *Nature* **396**, 463–466 (1998).
102. Drapeau, M. D., Cyran, S. A., Viering, M. M., Geyer, P. K. & Long, A. D. A *cis*-regulatory sequence within the *yellow* locus of *Drosophila melanogaster* required for normal male mating success. *Genetics* **172**, 1009–1030 (2006).
103. Knight, J. C. Regulatory polymorphisms underlying complex disease traits. *J. Mol. Med.* **83**, 97–109 (2005).
104. Rockman, M. V. & Wray, G. A. Abundant raw material for *cis*-regulatory evolution in humans. *Mol. Biol. Evol.* **19**, 1991–2004 (2002).

Acknowledgements
Thanks to C. Babbitt, D. Garfield, J. Tung, and three anonymous reviewers for helpful comments. G.A.W.'s research is supported by the US National Science Foundation and the Institute for Genome Sciences & Policy at Duke University.

Competing interests statement
The author declares no competing financial interests.

DATABASES

The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
DARC | GATA1 | IL8 | *LCT* | *MCM6* | *PDYN*
FlyBase: http://flybase.bio.indiana.edu
Abd-B | *ac* | *e* | *ovo* | *sc* | *Ubx* | *y*

FURTHER INFORMATION

The Wray Laboratory homepage:
http://www.biology.duke.edu/wraylab
Access to this links box is available online.