THE LOCUS OF EVOLUTION: EVO DEVO AND THE GENETICS OF ADAPTATION

Hopi E. Hoekstra[1,2] and Jerry A. Coyne[3,4]

[1]*Department of Organismic and Evolutionary Biology and the Museum of Comparative Zoology, Harvard University,*

26 Oxford Street, Cambridge, Massachusetts 02138

  [2]*E-mail: hoekstra@oeb.harvard.edu*

[3]*Department of Ecology and Evolution, The University of Chicago, 1101 E. 57 Street, Chicago, Illinois 60637*

  [4]*E-mail: j-coyne@uchicago.edu*

An important tenet of evolutionary developmental biology ("evo devo") is that adaptive mutations affecting morphology are more likely to occur in the *cis*-regulatory regions than in the protein-coding regions of genes. This argument rests on two claims: (1) the modular nature of *cis*-regulatory elements largely frees them from deleterious pleiotropic effects, and (2) a growing body of empirical evidence appears to support the predominant role of gene regulatory change in adaptation, especially morphological adaptation. Here we discuss and critique these assertions. We first show that there is no theoretical or empirical basis for the evo devo contention that adaptations involving morphology evolve by genetic mechanisms different from those involving physiology and other traits. In addition, some forms of protein evolution can avoid the negative consequences of pleiotropy, most notably via gene duplication. In light of evo devo claims, we then examine the substantial data on the genetic basis of adaptation from both genome-wide surveys and single-locus studies. Genomic studies lend little support to the *cis*-regulatory theory: many of these have detected adaptation in protein-coding regions, including transcription factors, whereas few have examined regulatory regions. Turning to single-locus studies, we note that the most widely cited examples of adaptive *cis*-regulatory mutations focus on trait loss rather than gain, and none have yet pinpointed an evolved regulatory site. In contrast, there are many studies that have both identified structural mutations and functionally verified their contribution to adaptation and speciation. Neither the theoretical arguments nor the data from nature, then, support the claim for a predominance of *cis*-regulatory mutations in evolution. Although this claim may be true, it is at best premature. Adaptation and speciation probably proceed through a combination of *cis*-regulatory and structural mutations, with a substantial contribution of the latter.

KEY WORDS: *Cis*-regulation, evolution of development, gene regulation, phenotypic evolution. structural gene.

As new areas of research have been folded into the Modern Synthesis, each has claimed to offer unique and revolutionary insights into the evolutionary process. Punctuated equilibrium, for example, proposed novel and non-Darwinian explanations for a seemingly discontinuous fossil record. These included the fixation of nonadaptive macromutations by genetic drift in small populations, and the operation of "species selection," producing macroevolutionary trends via the differential splitting and extinction of entire taxa (Eldredge and Gould 1972; Gould and Eldredge 1977, 1993; Gould 1980).

Some advocates of "evo devo" (the new field that fuses developmental and evolutionary biology) also claim to have revolutionized the study of macro- and microevolution. Like advocates of punctuated equilibrium, adherents to evo devo extrapolate from

pattern to process. Their novel evolutionary theories include the notion that the new body plans (i.e., phyla) arise by mutations different from those distinguishing populations or species (Davidson and Erwin 2006); the idea that evolution involves the transformation of developmental "modules" that are relatively independent of each other genetically (Breuker et al. 2006); the view that evolution itself establishes traits that promote *future* evolution ("evolvability;" Kirshner and Gerhardt 1998); and the idea that most important evolution involves alterations in the regulation of genes rather than in their structure.

The emphasis on gene regulation is evo devo's most famous and widely accepted contribution to evolutionary theory. It began with the work of Jacob and Monod on bacterial operons (1961), and was formalized by Jacob (1977) in a now-famous paper suggesting that evolution acts as a "tinkerer," assembling new adaptations by puttering about with gene regulation. Around the same time, King and Wilson (1975), noting the similarity in protein and DNA sequence between humans and chimps, suggested that minor changes in gene regulation could yield major phenotypic change between taxa. Wilson and colleagues expanded this view in a series of papers (e.g., Wilson et al. 1974a,b; Wilson 1975). The emphasis on gene regulation was also a major theme of influential work by Britten and Davidson (1969, 1971), who suggested that morphological evolution resulted more from changes at "integrator" and "receptor" genes than from "producer" genes (categories that presumably correspond, respectively, to transcription factors, promoters, and structural genes).

As evo devo matured, the focus on gene regulation narrowed to a single one of its forms: that involving *cis*-regulatory elements (short, noncoding DNA sequences that control expression of a nearby gene). For various reasons, which we discuss below, *cis*-regulatory elements are now seen as not only the most likely target for the evolution of gene regulation, but also as the site of most important evolutionary change, at least for morphology.

Perhaps the first detailed argument for the importance of *cis*-regulation was made by Stern (2000). But the most vigorous advocate of this view has been Carroll, who, in a series of papers, scholarly books, and popular books (Carroll 2000; 2005a,b; Carroll et al. 2001, 2006), has repeatedly emphasized that the evolution of animal form and other macroevolutionary features resulted largely from changes at *cis*-regulatory sites:

> In the final chapter of this book [titled "From DNA to Diversity: The Primacy of Regulatory Evolution"], we consider why regulatory evolution is the creative force underlying morphological diversity across the evolutionary spectrum, from variation within species to body plans. The link involves evolution at the DNA level and phenotypic diversity involves the *cis*-regulatory elements acting as units of evolutionary change (Carroll 2001, p. 173).

It has required several decades to obtain evidence that regulatory sequences are so often the basis for the evolution of form that, when considering the evolution of anatomy (including neural circuitry), regulatory sequence evolution should be the primary hypothesis considered (Carroll 2005a, p. 1165).

> This regulatory DNA [noncoding promoter regions] contains the instructions for building anatomy, and evolutionary changes within this regulatory DNA lead to the diversity of form (Carroll 2005b, p. 12).

These conclusions are delivered without caveats. The popular book *Endless Forms Most Beautiful* (Carroll 2005b), for example, begins with a quote from the Beatles' song "Revolution 1." In case the reader misses its significance, Carroll quickly explains (p. *x*):

> Over the past two decades, a new revolution has unfolded in biology. Advances in developmental biology (dubbed "evo devo") have revealed a great deal about the invisible genes and some simple rules that shape animal form and evolution. Much of what we have learned has been so stunning and unexpected that it has profoundly reshaped our picture of how evolution works.

Although Carroll's views have been by far the most influential in this area, other workers have also taken up the cudgels, showing the same unwavering confidence about the genetic basis of evolutionary change:

> The conclusion we draw from these inferences is that the evolution of plant form will be most readily accomplished by changes in the *cis*-regulatory regions of transcriptional regulators (Doebley and Lukens, 1998, p. 1081).

> For anyone interested in mechanism, there is in fact no other way to conceive of the basis of evolutionary change in bilaterian form than by change in the underlying developmental gene regulatory networks. This of course means change in the *cis*-regulatory DNA linkages that determine the functional architecture of all such networks" (Davidson 2001, p. 201).

> From what is already known, it is evident that the evolution of regulatory gene systems, rather than of structural alleles, has been chiefly responsible for the sorts of major morphological innovations revealed by the fossil record. . . Indeed, for the origin of bodyplans, involving the patterning of novel architectures, evolution of *cis*-regulatory elements appears to have been preeminent (Valentine 2004, pp. 77, 104).

(See also Wray et al. 2003).

But are these claims supportable? Considerable data now exist documenting the types of DNA changes underlying adaptive differences among species and higher taxa. Here we review these data. We will conclude that evo devo's enthusiasm for *cis*-regulatory changes is unfounded and premature. There is no evidence at present that *cis*-regulatory changes play a major role—much less a pre-eminent one—in adaptive evolution. We hasten to add, however, that future work may indeed show *cis*-regulatory change to be an important feature of evolution, and, as Carroll and

others suggest, one that should be studied carefully. At present, however, we can conclude only that changes in *both* the structure and regulation of genes have been important in adaptation, that their relative importance will not be known for a considerable time, and that the role of structural mutations in morphological evolution—and other adaptive change—is unlikely to be trivial.

The argument for the ubiquity of *cis*-regulatory evolution rests on two pillars. The first is a theoretical claim: the nature of gene regulation makes promoter elements perforce the most likely site of evolutionary change. Moreover, the involvement of promoters is said to have been far more important in the evolution of anatomical traits than of other sorts of traits. The second argument is empirical: *cis*-regulatory evolution has actually been the most important cause of adaptation. We will examine these arguments separately, but first we address two related questions: Do we expect a difference between the genetic basis of anatomical versus physiological evolution? And what is a regulatory change?

Form versus Function

It is a curious aspect of evo devo theory that *cis*-regulatory evolution is said to be enormously important for the evolution of body plans and anatomy, but not necessary for other types of adaptations. Thus, the "theory" of gene regulation largely ignores adaptations affecting behavior, biochemistry, metabolism, and physiology.

It is not clear why this is so. Although advocates of evo devo certainly make a sharp distinction between the evolution of anatomy on one hand and the evolution of all other traits on the other, which they lump together as "physiological" (e.g., Britten and Davidson 2001; Carroll 2005b), they have offered no biological justification for this distinction. Certainly it cannot be because nonanatomical changes are unimportant in evolution. It must be the case that many major evolutionary innovations and transitions involved changes that were not reflected in body form. Think, for example, of the transition from water to land, which involved innovations in respiration, behavior, and reproduction. The evolution of new phyla certainly involved more than just the changes in body plan documented in the fossil record, as we can see from examining adaptations of living phyla.

We suspect that there are two reasons for omitting nonanatomical traits from evo devo theory. First, many practitioners are interested in macroevolutionary changes that *can* be studied in the fossil record, and these of course are limited to changes in form. This appears to have promoted the view that changes in form are the most important of all adaptations. As Carroll (2005b) notes:

> The evolution of form is the main drama of life's story, both as found in the fossil record and in the diversity of living species (p. 294).

We do not address other forms of innovation, though they are fascinating in their own right, such as the evolution of physiological adaptations through protein evolution (for example, antifreeze proteins, lens crystallins, keratins, lactose synthesis, immune systems), because they do not concern morphological evolution per se (p. 160).

But the omission of "physiological" traits from the theory fails to acknowledge the tremendous amount of already-existing data showing that the adaptive evolution of such traits usually involves changes in structural regions. This, in fact, is acknowledged by evo-devotees. For example:

> There is ample evidence from studies of the evolution of proteins directly involved in animal vision, respiration, digestive metabolism, and host defense, that the evolution of coding sequences plays a key role in some (but not all) important physiological differences between species. In contrast, the relative contribution of coding or regulatory sequence evolution to the evolution of anatomy stands as the more open question, and will be my primary focus (Carroll 2005a, p. 245; references given in text omitted).

But why *should* there be a difference between the types of changes involved in the evolution of form versus function? Is there really an important evolutionary difference between making a bone long and making it strong? After all, physiological and biochemical changes are tissue- and organ-specific in exactly the same way as are anatomical changes, and both types of change occur within developmental networks. Indeed, the same impediments to protein evolution that are said to lead to *cis*-regulatory-based change of anatomy—the deleterious pleiotropic effects of protein-coding mutations—would seem to be at least as strong for physiological and biochemical innovations as for anatomical innovations.

We can find only one explicit biological rationale for distinguishing between the evolution of anatomy versus physiology:

> One absolutely crucial difference, then, between proteins involved in physiology and those involved in body-building, concerns the consequences of mutations that alter these proteins. A mutation that alters an opsin protein may affect the spectrum of light detected in either rods or cones in the eye. However, a mutation in a tool-kit protein [a transcription factor] may abolish the eye altogether, as well as affect other body parts. For this reason, mutations that alter tool-kit proteins are often catastrophic and have no chance of being passed on. The important consequence is that the evolution of form occurs more often by changing how tool-kit proteins are used, rather than by changing the tool-kit proteins themselves (Carroll 2006, p. 204).

But this argument is flawed on two grounds. First, taken at face value, it explains only *why transcription factors may evolve more slowly than other types of proteins*. It does not explain why physiology should evolve by changes in protein structure and anatomy by changes in *cis*-regulatory elements. After all, the

expression of both "physiology" and "anatomy" genes involves transcription factors and promoters, and so should be equally constrained. And there is no evidence that these two classes of genes are regulated in different ways. The study of comparative gene regulation is in its infancy, and although there are hints that different classes of genes may have different types of promoters (e.g., McNutt et al. 2005; Yang et al. 2006), these partitions neither include form versus function, nor say anything about the evolutionary *potential* of different classes of genes. The second problem with this argument is there is no necessary relation between the potential effects of mutations at a locus and the rate of adaptive evolution at that locus. We do not expect, a priori, that loci which can mutate to more lethal alleles (e.g., transcription factors) will evolve more slowly than loci whose extreme effects are more benign (e.g., genes producing structural proteins).

The artificiality of separating form and physiology becomes most evident when considering the evolution of pigmentation, which, although clearly involving physiological and metabolic processes, is nevertheless seen as an aspect of form:

> Changing the size, shape, number, or color patterns of physical traits is fundamentally different from changing the chemistry of physiological processes (Carroll 2005a, p. 1159).

The reason, then, why the evolution of anatomy is a "more open" question than that of physiology is not because there is some fundamental biological difference between the two classes of traits. It is *only because we have less evidence about the nature of change affecting form*, and therefore are less constrained by facts in speculating about its genetic basis. Because there is no clear theoretical reason for expecting different types of evolutionary changes for form than for physiology, we will, when dealing with the data, lump together both types of adaptations.

What is a Regulatory Change?

Historically, the literature on evo devo has conflated two concepts: regulatory genes and regulatory mutations. We will show that while trying to define a regulatory *gene* leads one into a tangled semantic thicket, one can define regulatory *mutations* (i.e., *cis*-regulatory changes) in a consistent way that allows us to address and evaluate the claims of evo devo.

On some level it can be argued that most genes regulate something, whether that something be a protein, a pathway or a biochemical product. True, the primary function of some genes is clearly regulatory. The main role of transcription factors, for example, is to bind to DNA elsewhere in the genome and thereby regulate the spatiotemporal expression of genes. Likewise, some genes have a distinct structural function. They may, for example, contribute to the physical structure of chromosomes and cells. One example is keratin, an insoluble fibrous protein found in hair, feathers, and scales.

There are, however, many cases in which it is hard to draw a simple dichotomy between "structural" and "regulatory" function of genes. Histones, for example, form nucleosomes, which act as spools around which DNA is coiled, maintaining its helical structure and forming chromatin. Although histones were once thought to have a purely structural function, their posttranscriptional modification also allows them to act in more diverse biological processes, including gene regulation (Strahl and Allis 2000). Similarly, the protein beta-catenin has dual regulatory and structural functions (Perez-Moreno and Fuchs 2006). As a structural protein, it is an essential component of cellular adhesion in the cytoskeleton. As a regulatory protein, it acts as a transcriptional coactivator in the *Wnt* signaling cascade. Because of the domain structure of the beta-catenin protein, these two functions can be separated; that is, mutations can alter beta catenin's regulatory function while maintaining its structural role, and vice versa (Bremback et al. 2006). Finally, other proteins have structural and regulatory functions that are inseparable. *SATB1* organizes chromosomes into distinct loop domains, and thus acts as a traditional structural gene. But this structural aspect has a regulatory end: *SATB1* orchestrates gene expression by remodeling chromatin at specific genomic locations, allowing enzymes access to target DNA for regulating DNA transcription (Yasui et al. 2002).

We have not singled out histones, beta-catenin and *SATB1* because they are among only a few genes having both structural and regulatory properties. We could give many similar examples. And when we understand development more fully, it seems likely that many "structural" proteins will act together with transcription factors to regulate gene expression.

A related issue is whether mutations *within* a gene should be classified as regulatory or structural. This question, too, is not straightforward. For example, amino-acid ("structural") substitutions in transcription-factor proteins may be more common than previously appreciated, and these can alter gene regulation. Like *cis*-regulatory elements, many transcription factors are modular in structure (having several functional elements that can act independently of one another), and there is increasing evidence that their coding changes can alter expression of a subset of downstream target genes without completely disrupting downstream pathways (Hsia and McGinnis 2003). In fact, Levine and Tjian (2003) suggest that the diversification of activation sites of transcription factors—whose DNA binding domains nevertheless remain conserved—also contribute to organismal diversity.

One example involves homeobox (Hox) genes, the most famous class of transcription factors, which help specify segmentation patterns along the anterior–posterior body axis of animals. Although the DNA sequences of Hox genes are largely conserved among major animal groups, some coding changes in the Hox gene *ultrabithorax* (*Ubx*) affect *its* ability to regulate downstream transcription levels and ultimately its ability to repress limb formation.

(In vitro studies implicate the loss of serine phosphorylation sites.) Thus, coding mutations in a transcription factor might be involved in a "macroevolutionary" change in animal body plan (Ronshaugen et al. 2002).

Likewise, in different groups of insects, evolution has exchanged binding motifs in the coding region of the Hox gene *ftz*. This swap has changed *ftz*'s downstream binding targets and hence its regulatory role. These swapped motifs may be associated with the diversity of body segments (Lohr et al. 2005). Should we consider such mutations regulatory—because they alter the expression levels of downstream genes—or structural—because they alter the structure of the transcription factor? And should we classify as structural or regulatory those amino acid changes in a protein that affect *its own* regulation (e.g., mutations in G-protein coupled receptors that downregulate the receptor [Benya et al. 2000; Rathz et al. 2002])? What about coding-region mutations that affect mRNA folding or stability (e.g., Wisdom and Lee 1991; Schiavi et al. 1994; Shen et al. 1999), protein level (Carlini and Stephan 2003) or tissue-specific expression pattern (e.g., Nakayama and Setoguchi 1992)? Or silent mutations in the coding regions that affect translation rate and protein function (Kimchi-Sarfaty et al. 2007)?

To escape this semantic tangle, we take two approaches. First, we refrain from classifying *genes* as either structural or regulatory, although some bits of DNA, like promoters, are clearly regulatory. Second, we classify *mutations* based on their physical location. Mutations must lie either inside or outside the coding region of a gene (either DNA that is transcribed into mRNA and translated into a protein, or "functional" RNA molecules such as ribosomal RNA, ribozymes, or antiviral RNA). If a mutation affecting a phenotype lies within the coding region, we consider it a structural mutation. Conversely, mutations that lie outside the coding region (including mutations in introns) are considered regulatory. Although regulatory elements are often poorly delineated, we can infer that if a noncoding mutation causes a change in phenotype, it usually occurs in a functionally important *cis*-regulatory element (e.g., enhancer, promoter core element, or other transcriptionally relevant element).

When considering a "causal locus" affecting a phenotypic difference, our distinction between regulatory and structural mutations covers all possible changes, and we no longer need to distinguish between *cis*- and *trans*-regulation. For example, if a *cis*-regulatory change alters the expression of gene A, which then has downstream effects on unlinked gene B, and the effects of gene B alter the phenotype, then the *causal* change is *cis*-acting for A, *trans*-acting for B, but is still a regulatory mutation in our classification. Finally, we will not distinguish here between the various types of regulatory elements (for a description see Alonso and Wilkins 2005), as this is irrelevant to our discussion.

Our distinction between structural and regulatory mutations comports with much current usage in evo devo. Of course, while it is easy to construct such a dichotomy, it is much harder to *identify* the mutation or mutations associated with a gene that causes an important phenotypic change.

Theory

Carroll (2005a,b; 2006) outlines what we call the "theoretical imperative" for *cis*-regulatory evolution. This derives from what we know about the nature of gene regulation (see Levine and Tjian 2003; Wray et al. 2003), and so a brief review is in order.

Eukaryotic genes are under the control of noncoding DNA sequences (e.g., *cis*-regulatory elements), including promoters usually located "upstream" (in the 5′ direction) from the start codon of a gene. Core promoter sequences are the sites where transcription is initiated. A gene can be controlled by several independent promoters (indicating different transcriptional start sites), which may or may not be close to each other. The default state of a gene is "off" (no expression or low basal expression), and mRNA transcription begins when RNA polymerase binds to the gene's promoter region. The binding of RNA polymerase is mediated by transcription factors, regulatory proteins that may themselves require other transcription factors or organic molecules for accurate binding.

By and large, transcription factors are evolutionarily conserved in both structure and function; the classic example is Hox genes, which are conserved in their amino acid homeodomains, genomic organization, and expression patterns among animals (Hill et al. 1989; Doboule and Dolle 1989). In addition, promoter regions often work together with other *cis*-regulatory elements (e.g., enhancers, silencers, insulators, etc.) to control the expression of the gene in a specific tissue or at a specific time. For example, enhancers (sequences a few hundred base pairs long) usually bind sequence-specific transcription factors to mediate expression within a specific tissue or cell type. Silencers, on the other hand, bind transcription factors that block or reduce transcription levels by impeding RNA polymerase binding. Both enhancers and silencers can be up to 100 kilobases away from their core promoter, making them difficult to identify. Taken together, these *cis*-regulatory elements are modular: that is, different *cis*-regulatory elements can independently affect the expression of a transcript at different times and places. Consequently, diversity in gene regulation can be achieved by different combinations of *cis*-regulatory elements working independently of one another to direct composite patterns of gene expression.

The fact that each gene is controlled by a set of modular *cis*-regulatory elements leads to the most important consequence for evo devo theory. Whereas a change in a protein sequence may have deleterious pleiotropic effects (proteins interact with other

proteins through the ramifying network of development, and a sequence change could affect every such interaction), a change in a *cis*-regulatory element may affect only the specific temporal or spatial expression of its single attendant gene. *Cis*-regulatory changes are therefore thought to be relatively free of negative pleiotropic effects on fitness. The problem with protein-sequence change seems even worse if the protein is a transcription factor, because every gene regulated by such a factor might show altered expression.

All other things being equal, then, a change in a *cis*-regulatory region is supposed to have a higher probability of being adaptive than is a random change in a structural gene or transcription factor. Moreover, if a mutation in a *cis*-regulatory element brings a gene under the control of a new transcription factor, a radical co-option of function can take place. Such co-option is said to underlie evolutionary innovations such as body segmentation and diversification of those segments (Carroll et al. 2001).

The final factor said to promote regulatory evolution is "the combinatorial action of the transcription factor repertoire in cells" (Carroll et al. 2001, p. 190). As Carroll et al. explain (p. 190), "The transcription factor repertoire is sufficiently diverse and the stringency of DNA binding [to the promoter region] sufficiently relaxed such that sites for most transcription factors can evolve at significant frequency in animal genomes." This idea—effectively that promoters have a higher rate of adaptive nucleotide substitution—could produce the differences in timing or tissue expression said to be involved in most evolutionary innovations.

Taken together, these facts about gene regulation underlie the theoretical imperative for *cis*-regulation:

> It [the nature of *cis*-regulatory regions] constitutes pervasive evidence that the diversification of regulatory DNA, while preserving coding function, is the most available and most frequently exploited mode of genetic diversification in animal evolution" (Carroll 2005b, p. 231).

However, there are several other ways to obviate the negative consequences of pleiotropy besides changing *cis*-regulatory elements. The most obvious is gene duplication followed by divergence of the duplicated copies (termed "paralogs"). This process allows a protein to retain an ancestral function while its paralog or paralogs evolve to new functions. (Gene duplication, of course, can also create new *cis*-regulatory regions that may likewise diverge adaptively.) In addition, a gene can mutate to new forms by creating alternative splicing sites or recruiting new coding domains while still allowing production of the ancestral protein; these two processes are relatively rare. The evolutionary fixation of duplications, however, appears to be fairly common. Nevertheless, Carroll (2005a) argues that duplications are established too rarely to play an important role in micro- and macroevolution, citing Lynch and Conery's (2000) calculation of one duplication fixed (or nearly fixed) per gene per 100 million years.

The theoretical argument for the importance of *cis*-regulation thus rests on eliminating evolutionary alternatives: changes in structural genes affecting anatomy must either be deleterious themselves or accompanied by deleterious pleiotropic effects, and recruitment of coding domains, alternative splicing, and gene duplication are rare. We are then left with *cis*-regulatory regions as the most likely site of adaptive change.

This logic, however, is not convincing in light of what we know about the population genetics of new mutations. The rate of fixation of *cis*-regulatory versus structural mutations depends on three factors: (1) their relative mutation rates, (2) their relative chances of being adaptive (positively selected; Fisher 1930), and (3) the relative sizes of the selection coefficients (Kimura 1983; Orr 2003). Even if a *cis*-regulatory mutation is less likely to have deleterious pleiotropic effects, this does not necessarily mean it is more likely to be fixed, because such mutations may be less likely to occur or their net selection coefficients may make them less likely to be fixed. For example, *cis*-regulatory sites at a given gene may be less numerous than protein-coding sites, and their mutation rate correspondingly lower. Moreover, it is easy to imagine that expressing a protein at a new time or place could have effects just as deleterious as—or more deleterious than—changes in protein sequence. Is it so clear that activating a gene in a new part of the body, or making twice as much of an enzyme, is more likely to be adaptive than, say, a single substitution of valine for leucine in an enzyme?

What about gene duplications? Are they, as Carroll maintains, too infrequent to explain much adaptation? This seems unlikely. It is curious that the paper cited by Carroll supporting the *infrequency* of adaptive change by gene duplication—that of Lynch and Conery (2000)—actually claims that duplications are not only frequent, but make important contributions to speciation and species-level differences. One duplication per gene per 100 million years is a low *per-gene* mutation rate, but not necessarily a low *per-genome* mutation rate; and it is the latter that is important for adaptation and speciation. As Lynch and Conery (2000 p. 1154) note, "With rates of establishment of 0.002 to 0.02 duplicates per gene per million years and a moderate genome size of 15,000 genes, we can expect on the order of 60–600 duplicate genes to arise in a pair of sister taxa per million years . . . " Moreover, gene copy number polymorphisms within species are well documented in the one species that has been extensively studied—humans (Sebat et al. 2004)—and are likely to be found in other groups.

One need only peruse Ohno's (1970) book to see the pervasiveness and potential importance of gene duplication, one of the few ways that new genes can actually arise in evolution. After all, nearly every gene can be considered a duplicate or chimera of earlier genes, and the origin of new genes must therefore have been important in adaptation. It is almost superfluous to list the gene families and adaptations deriving from duplications: they include

globins, the immune system, olfaction, opsins and, indeed, transcription factors themselves.

The multiply-and-diversify model of evolution does not depend solely on the duplication of single genes: the evolution of tetrapods probably involved at least two bouts of *whole-genome duplication* (Dehal and Boore 2005). Moreover, it is estimated that between 47% and 70% of angiosperms are polyploids (Ramsey and Schemske 1998), and thus harbor duplicated genes. Otto and Whitten (2000) calculated that ploidy changes represent between 2% and 4% of speciation events in flowering plants, although polyploidy is far rarer in animals. In view of these facts, it seems unwise to deny a priori that structural genes could play a major role in the evolution of plants and animals.

Moreover, there are other ways besides gene duplications that novel and useful structural genes can arise. These include gene fusion and fission (e.g., mammalian fatty acid synthase), recruitment of old genes to new functions (e.g., the antifreeze proteins permitting fish to live in frigid waters), exon shuffling (e.g., involved in the evolution of blood clotting), and the addition of transposons to coding sequences. In a review on the origin of new genes, Long et al. (2003) describes these and many other processes. Given the diverse ways that useful new genes can arise, one should be cautious about making sweeping evolutionary statements about likelihood. Before concluding, for example, that the difference between a man and a mouse rests largely on the nature of their promoters, one should realize that 21% of human protein-coding genes have *no known homologs* (gene copies related by descent) in mice (available from http://eugenes.org:7072/all/homologies/hgsummary-2005.html; Don Gilbert, Genome Informatics Laboratory, Indiana University).

Given the contrast between evo devo theory and the evidence that there has indeed been dramatic change in structural genes during evolution, it is no surprise that some have taken a position completely antithetical to the *cis*-regulatory imperative, viz. Li's statement (1997, p. 269) that "there is now ample evidence that gene duplication is the most important mechanism for generating new genes and new biochemical processes that have facilitated the evolution of complex organisms from primitive ones."

In the end, such back-and-forth assertions seem Talmudically irresolvable, at least on the basis of a priori considerations. The real way to settle the issue of the importance of *cis*-regulatory evolution is to look at the data. How often have new adaptations involved evolutionary changes of promoters versus changes in coding sequences? We now turn to the empirical evidence on the molecular basis of adaptation.

The Facts
GENOMIC STUDIES OF *cis* AND *trans* MUTATIONS
It is appropriate to begin our discussion with recent genomic data, because much of the inspiration for "regulatory gene" theories arose from early genomic data on the similarity of protein sequences between humans and chimps.

The recent production of complete genome sequences from many species has allowed far more refined analysis of adaptation using genome-wide patterns. Several studies are relevant to the question of *cis*-regulatory evolution. Andolfatto (2005), for example, showed that patterns of nucleotide variation in untranslated regions (UTRs) of the *Drosophila* genome are consistent with the view that changes in these regions affect fitness. He estimates that changes in UTRs probably contribute at least equally, if not more, to adaptation than do changes in coding regions. However, a major drawback of this approach, inherent in most genomic studies (discussed below), is that genome-wide surveys are conducted in the absence of phenotypic information, limiting our ability to identify the specific DNA mutations affect phenotype and are truly adaptive.

In the spirit of King and Wilson (1975), most of these genomic studies have focused on identifying genetic regions showing rapid evolution in the human lineage; the implicit goal is to discover mutations contributing to "human-ness." Early and highly publicized estimates that the DNA of chimps and humans is 99% identical led King and Wilson (1975, p. 115) to conclude that "a relatively small number of genetic changes in systems controlling the expression of genes may account for the major organismal differences between humans and chimpanzees." However, a 99% identity of DNA sequence still translates into a considerable difference in protein sequence, a conclusion confirmed by the data of Glazko et al. (2005) showing that 80% of the proteins of humans and chimps differ by at least one amino acid. Regulatory change, then, may not be necessary to explain the phenotypic differences between these species.

Several other studies have identified rapidly evolving proteins (and hence, structural mutations) that may have been involved in adaptive evolution in primates. For example, hundreds of genes show evidence of positive selection in the hominid lineage (Clark et al. 2003). Two more recent studies also showed evidence for rapid evolution of amino acid sequences (ca. 5–9% of genes under analysis), including genes involved in sensory perception and immune defenses (Dorus et al. 2004; Nielsen et al. 2005). In fact, one study (Bustamante et al. 2005) identified transcription factors as a particularly rapidly evolving class of proteins, contradicting the evo devo assertion that antagonistic pleiotropy precludes changes in the amino acid sequence of transcription factors. Indeed, the results of Bustamante and colleagues suggest that even if differences in gene expression played a prominent role in the divergence of humans from chimps, the ultimate cause may often involve structural mutations.

Only a few studies, however, have simultaneously compared regulatory with structural evolution. A recent one identified DNA elements in both coding and noncoding regions that showed rapid

divergence along the human lineage, elements termed "human accelerated regions" (HARs; Pollard et al. 2006a,b). The authors conclude that the majority of HARs are: (1) in noncoding regions, (2) contiguous to coding regions, and (3) if within coding regions, often in transcription factors. Together these results raise the possibility that *cis*-regulatory changes contribute disproportionately to human-specific traits. Unfortunately, because these genomic studies are conducted without reference to the phenotype, it is impossible without further work to determine which mutations in HARs contributed to adaptive evolution.

Recent technological advances allow us to gauge the relative contributions of *cis* versus *trans* mutations to interspecific changes in gene expression at many loci. In an elegant study, Wittkopp et al. (2004) examined the contributions of *cis*- and *trans*-acting factors to species-level divergence of gene expression in F_1 hybrids of *Drosophila melanogaster* and *D. simulans*. (In this study, the distinction between *cis*- and *trans*- mutations is not identical to our own distinction between regulatory and structural mutations). Wittkopp and colleagues clearly show that *cis*-acting factors are an important part of interspecific divergence in gene expression in *Drosophila*. Again, however, we do not know the phenotypic effects of any of the 29 genes analyzed. It is thus unclear whether any of the species differences in gene expression have an adaptive (or even phenotypic) effect, and, if so, which proportion of such adaptive changes were are caused by structural versus regulatory mutations.

Despite these problems, the studies discussed here raise the possibility that future genomic studies could address the relative contribution of *cis*-regulatory and structural mutations to biological diversity. At present, however, the genomic data are ambiguous. We turn now to the data from individual loci—data that constitute main bulwark of *cis*-regulatory theory.

SINGLE-LOCUS DATA

We begin by discussing the criteria for deciding whether a change in phenotype is caused by changes in *cis*-regulatory elements, protein structure, or both. We then describe the experiments necessary to demonstrate the relative contributions of *cis*-regulatory versus structural changes to adaptive variation at single loci.

A common method for determining the role of gene regulation in evolutionary change—a method that is the foundation of the evo devo approach—involves simultaneously comparing differences in a phenotype among species (often distantly related ones) with the pattern of expression of a single gene thought to influence that phenotype. Although this approach has successfully identified important pathways involved in phenotypic change (e.g., the calmodium pathway involved in beak-size evolution of Darwin's finches [Abzhanov et al. 2004, 2006] and *Notch/Distal-less* in the formation of butterfly wingspots [Beldade et al. 2002; Reed and Serfas 2004]), it does not give us the

complete story because the source of phenotypic differences are not pinpointed:

> In many cases a gene required for the development of a trait in one species shows a difference in expression in other species that correlates with a difference in that trait ... A causal relationship is plausible but not proven in these cases, because comparisons of gene expression cannot by themselves demonstrate that a change in transcriptional regulation is the genetic basis for a phenotypic difference (Wray et al. 2003, p. 1378).

The common methods of observing spatiotemporal patterns of gene expression (e.g., in situ hybridization, quantitative PCR) and experimentally manipulating protein levels (e.g., ectopic or misexpression studies) can do no more than show an association of gene expression with phenotype and perhaps implicate the developmental pathway in which the causal mutation lies. Although the causal mutation may be located in the *cis*-regulatory region of the protein of interest, it is equally likely, if not more likely, to lie somewhere upstream of the gene of interest, somewhere in the panoply of *trans*-regulatory factors or cofactors that affect regulation of the gene. Although these correlational studies often proclaim that change in *gene regulation* contributes to phenotypic diversity, this result is neither novel nor surprising. In fact, it would be surprising if a mutation did *not* affect gene regulation, for most mutations (either structural or *cis*-regulatory) have effects on the regulation of gene products downstream in their respective pathways.

Likewise, structural mutations can also explain divergence in gene expression among species. Differences in the amount of mRNA or protein may reflect structural rather than regulatory changes if they result from differential stability of the gene product. For example, in *D. melanogaster* the replacement of certain synonymous codons in the coding sequence of *alcohol dehydrogenase* causes a significant decrease of enzyme production (Carlini and Stephan 2003). Others have noted that experiments documenting changes in gene expression do not necessarily implicate regulatory changes:

> Many comparative studies that use in situ hybridization interpret different probe patterns as an indication of transcriptional changes in enhancer [in the *c*is-regulatory region], openly ignoring the possibility of post-transcriptional events that alter mRNA stability or changes in splicing profiles that affect the sequences detected by (often) a single probe (Alonso and Wilkins 2005, p. 713).

While studies of gene expression at either individual candidate loci or many loci simultaneously (e.g., microarrays) can test developmental pathways involved in phenotypic variation, determining whether variation in gene expression among species involves structural versus regulatory changes usually requires genetic analyses. Genetic crosses (e.g., quantitative trait locus

mapping) and genetic complementation approaches (e.g., deletion mapping) can initially be used to localize genomic regions, often containing hundreds of genes, of which one or more contain mutations that contribute to the phenotypic difference of interest.

For candidate loci, the challenge is then to determine whether causal mutations occur in the structural or in the *cis*-regulatory regions. This requires identifying mutations that are functionally important as well as excluding mutations that are not. For amino acid changes, functional assays (e.g., cell-culture based or enzyme assays) can provide evidence for the role of particular mutations on protein behavior.

For regulatory changes, *cis*-regulatory elements can be tested for their ability to drive the expression of reporter constructs (e.g., green fluorescent protein; GFP) to determine the association between particular *cis*-regulatory regions and the spatial pattern of expression. These experiments, however, are still one step removed from the phenotype. For both structural and regulatory changes, the ultimate test of mutational effect is the use of transgenics together with a thorough examination of the phenotype of interest. In such tests, a construct containing the mutation(s) of interest is expressed in the appropriate genetic background to determine if it affects the phenotype of interest—and not just the protein activity or gene expression pattern. Of course, such experiments are not always technically feasible.

Excluding structural or regulatory mutations that do not play a role in the phenotypic difference can be even more difficult than identifying the causal mutations themselves. If there are no nucleotide differences in the entire coding region between individuals that differ phenotypically, then one can rule out structural mutations at that gene. Unfortunately, interspecific comparisons, (which predominate in evo devo) usually show some nucleotide differences, making it difficult to pinpoint the causal substitution(s) in a sea of irrelevant substitutions.

Determining the relative contribution of structural and *cis*-regulatory changes at any genetic locus is challenging, and most studies have not examined both types of change. Nonetheless, Table 1 lists and describes mutations implicated in adaptive change between closely related taxa. Each mutation is accompanied by a description of the adaptive nature of the change, its effect on protein function, the evidence supporting the causal link between mutation and phenotype, and the information still needed to fully characterize the mutation's effects.

We chose these examples because each demonstrates fairly rigorously both the adaptive nature of the genetic change as well as whether that change is *cis*-regulatory, structural, or both. We have omitted examples of structural gene changes between distantly related groups that are undoubtedly adaptive (e.g., α vs. β vs. fetal hemoglobin). Including such cases would strengthen the evidence for structural versus *cis*-regulatory evolution, for while the adaptive significance of amino acid substitutions in many of these cases is fairly clear, we know little about changes in *cis*-regulation.

Although there have been many arguments (some verging on the philosophical) about how to define and recognize an "adaptation," in Table 1 we have used a fairly loose criterion: if a trait is generally recognized to increase fitness or is maintained by selection, we regard it as an adaptation. So, for example, Table 1 includes features like cryptic coloration, antifreeze proteins in ectotherms, polymorphisms apparently maintained by balancing selection, and clinally varying traits. Some traits have shown the expected fitness effects in laboratory or field tests, while others have not been rigorously tested. Indeed, for one trait—a species difference in larval bristle pattern in *Drosophila*—we have no idea of its adaptive significance (see below); we include this trait because it is an oft-cited example of *cis*-regulatory change in evolution.

We do not claim that this table shows the relative importance of structural versus *cis*-regulatory change in evolution. There is almost certainly an ascertainment bias in favor of structural changes, because these are far easier to detect than changes in promoter regions. Differences in protein structure, for example, can be identified by simply comparing nucleotide or cDNA sequences. In contrast, most regulatory elements are small, not strictly conserved, and often far removed from the gene, making them difficult to identify and to pinpoint their functionally relevant sites. Also, although we know something about mutational effects in protein-coding regions based on the type of DNA or amino acid change (e.g., nonsynonymous vs. synonymous, conservative vs. radical, hydrophobic vs. hydrophilic) and its location (e.g., conserved motifs, active sites), the functional effects of mutations in regulatory elements remain largely unknown. Of course, identifying *cis*-regulatory changes would be facilitated by a better understanding of gene regulation. Finally, we hasten to add that the examples given in Table 1 are not exhaustive: we have inevitably missed relevant studies. Nevertheless, the list gives an idea of what we know at present about the molecular genetics of adaptation, and how much empirical evidence supports a claim for the importance of *cis*-regulatory variation.

Table 1 clearly shows that we have far more evidence for structural than for *cis*-regulatory changes. While the most well-supported examples of *cis*-regulatory based adaptation (the first three entries of Table 1) have not yet identified precise causal mutations, there are, in contrast, many examples of individual structural mutations contributing to adaptation. Moreover, most of the *cis*-regulatory examples involve loss of an ancestral trait (usually via loss-of-function alleles), whereas the structural mutations involve both gains and losses of traits. We discuss the two types of mutations, cis-regulatory and structural, below.

Empirical evidence for cis-regulatory adaptation

The claim that adaptive change is predominantly driven by *cis*-regulatory mutations rests on a handful of elegant but still incomplete studies. The three most relevant analyses, which focus respectively on skeletal armor in threespine sticklebacks (Shapiro et al. 2004), pigmentation on *Drosophila* wings (Gompel et al. 2005; Prud'homme et al. 2006), and dorsal bristle (trichome) density on *Drosophila* larvae (Sucena and Stern 2000), have been repeatedly cited as exemplars of *cis*-regulatory evolution. We will show, however, that in each case additional data are needed to identify the molecular basis of phenotypic change. It is also important to note that these three studies focus primarily on the genetic dissection of trait *loss*, so from the outset these data may be biased by a specific type of phenotypic change, and thus quite possibly by a specific type of mutational change. (It may, for example, be much easier for a *cis*-regulatory change to eliminate a trait than to create a new one.)

Perhaps the most comprehensive study of "*cis*-regulatory" adaptation comes from comparing pelvic spine morphology in marine versus benthic sticklebacks (*Gasterosteus aculeastus*; Shapiro et al. 2004). This work involved genetic analysis of phenotypic differences between an ancestral marine form, clad with armor plating and rigid spines that protect against predators, and a derived benthic form having reduced pelvic spines and very little armor. The use of genome-wide molecular markers allowed Shapiro and colleagues to map the difference in pelvic morphology to several chromosomal regions, one of which contains the candidate gene *Pitx1*. Because there is no difference in amino acid sequence between the *Pitx1* proteins of the two phenotypes, we can rule out the possibility that amino acid change in the *Pitx1* transcript caused the loss of pelvic spines. However, the absence of amino acid variation does not by default *prove* that the causal mutation(s) is located in an upstream *cis*-regulatory element, as there are alternative hypotheses (e.g., mutations in closely linked loci).

To examine divergence in *Pitx1* gene expression, Shapiro et al. (2004, 2006) used in situ hybridizations to compare *Pitx1* transcript levels between marine and benthic fish. This experiment yielded two important results. First, in some structures, like the mouth and jaw, the spatial expression pattern of *Pitx1* is conserved between the marine and benthic phenotypes. Second, in the pelvis, the *Pitx1* transcript is undetectable in the less-armored benthic form, and thus its absence is correlated with the absence of pelvic spines. These results suggest that there has been tissue-specific divergence in the regulation of *Pitx1*. Based on these two patterns of *Pitx1* expression, it is possible that benthic fish have undergone an inactivating mutation in a *cis*-regulatory element specific responsible for pelvic expression. However, additional data, including identifying the precise mutation(s), are necessary to prove that a *cis*-regulatory mutation(s) contributes to the adaptive pelvic re-

duction. The crucial experiments (undoubtedly underway) include the following:

(1) fine-scale mapping to exclude the contribution of neighboring genes (e.g., transcription factors, miRNAs) to protein expression and ultimately to morphological variation.

(2) identifying and verifying through functional analysis (e.g., transgenic experiments) the causal mutations in the cis-regulatory region.

Moreover, to support the ancillary hypothesis that modularity in the *cis*-regulatory region promotes evolutionary change, it will be necessary to identify multiple *cis*-regulatory elements and demonstrate that each element, by binding distinct transcription factors, independently controls tissue-specific expression. (A second locus, *ectodysplasin* (*Eda*), has been implicated in the loss of lateral plates in freshwater populations of sticklebacks [Colossimo et al. 2005]. However, nothing is yet known about the relative role of structural versus regulatory mutations at this locus.)

Two other examples of *cis*-regulatory evolution come from *Drosophila,* one on larval trichome loss and the other on pigmentation. These studies both compared divergent *Drosophila* species, one or more of which experienced the loss of a trait during their evolutionary history. The first pair of studies examined the role of species-specific differences in the expression of the *yellow* protein in the formation of male wing spots that may play a role in courtship behavior (Gompel et al. 2005; Prud'homme et al. 2006). Most notably, this work used transgenic methods to test individual sub-regions of the 5′ *yellow* promoter and to determine which regions drove reporter expression in the developing wing. Together with sequence data, these experiments show that the gain and loss of binding sites in the *cis*-regulatory region affect the expression of *yellow* protein among species. Although in this case the promoter region clearly contains regulatory modules controlling the spatial expression of *yellow*, the direct link between genotype and phenotype is not complete. This is because changes in the *cis*-regulatory elements of *yellow* alone are not sufficient to produce the phenotype of interest—the pigmented wing spot (Gompel et al. 2005). Additional loci must therefore be involved. Although it is not surprising that different *cis*-regulatory elements in the *yellow* promoter affect *yellow* expression, a critical piece of evidence is still missing: the demonstration that species-specific *cis*-regulatory elements produce the species-specific difference in the wing spot.

A third study, that of Sucena and Stern (2000), used genetic mapping (genetic crosses with visible markers, deletion mapping, and single-gene complementation) to pinpoint the gene *ovo*/*shavenbaby* (*svb*) as the cause of differences in trichome pattern between species. While larvae from most species in the *Drosophila melanogaster* subgroup have robust denticles and a

lawn of fine hairs on their abdominal segments, *Drosophila sechellia* (and four other species) maintain the rows of robust denticles but have lost the fine hairs, thus acquiring a naked cuticle. The adaptive significance of the interspecific difference in trichome pattern, if any, is unknown. There are several lines of evidence that the causal mutation(s) is regulatory. First, expression of *svb* mRNA is correlated with phenotypic variation in trichome pattern. That is, in *D. melanogaster* the *svb* transcript is abundant in cells that form robust denticles and less abundant (but still present) in cells that produce fine hairs. In *D. sechellia*, the *svb* transcript is similarly abundant in cells that produce denticles but absent in cells producing fine hairs. Second, transgenic assays show that a 60 kb region upstream of the *svb* coding region contains several sites influencing the species difference in trichome pattern (D. Stern, pers. comm.). Third, the *sechellia* naked-cuticle phenotype is not consistent with a null mutant at the *svb* locus itself, because such mutants completely lack trichomes (both the robust denticles and the fine hairs). Finally, recombination studies show that the coding region of *svb* is not responsible for the phenotypic difference (D. Stern, pers. comm.). Nevertheless, the precise locations of the DNA changes that produce the interspecific difference in trichome pattern remain elusive.

The three sets of studies described above are the strongest (and most widely cited) cases used to show the evolutionary importance of *cis*-regulatory mutation. None of them has yet identified an individual mutation in a *cis*-regulatory element, or functionally verified via transgenics that that mutation contributes to the phenotypic difference of interest. By raising these issues, we do not mean to criticize these studies, for the conclusive experiments are almost certainly in progress and may well show that *cis*-regulatory evolution is involved. Indeed, this seems likely for the cases of *Pitx1* in sticklebacks and *ovo/shavenbaby* in *Drosophila*. We claim only that, at present, these studies cannot serve as formal demonstrations of *cis*-regulatory change in evolution. Moreover, even if all of these cases do prove to involve *cis*-regulatory change, we are still left with only a handful of such examples compared to the much larger amount of data implicating structural changes. Finally, we must recall that these three studies focus primarily on the *loss* of traits (pelvic spines, wing spots, and trichomes). Supporting the evo devo claim that *cis*-regulatory changes are responsible for morphological innovations requires showing that promoters are important in the evolution of *new* traits, not just the losses of old ones.

Empirical evidence for structural adaptation

In contrast to the dearth of evidence for *cis*-regulatory changes are the many cases in which an adaptation has involved changes in a structural region (see Table 1). Except for insecticide resistance, all of these are "natural" adaptations that do not involve human intervention or selection. (Although we did not use examples from animal or plant breeding, we included genes involved in insecticide resistance because such adaptations still take place in a semi-natural environment with all of its constraints, and because evolutionary responses to insecticides highlight the diverse ways that the genome handles the adaptive challenge of toxicity).

Inspecting these data yields several conclusions. The first is obvious: there are many examples of simple changes in amino acid sequence contributing to adaptive evolution. Some cases involve changes in morphological traits (e.g., *Mc1r* in pigmentation), while most involve physiological traits (e.g., lysozymes in digestion). It is important to add that not all of these structural mutations have yet been tested using functional assays.

As we emphasized above, the larger number of documented structural changes may partly result from a bias in our understanding of underlying molecular pathways. While most of us have seen the detailed physiological pathways illustrated in textbooks (e.g., the Krebs cycle), there is little similar information about the genetic network for morphology. Therefore, candidate loci for physiological traits are more readily identified and their coding regions more readily sequenced. In contrast, traditional evo devo studies are motivated by understanding differences in morphology (body plan) and use comparisons of gene expression pattern as their primary tool.

Several examples of amino acid substitutions are clearly involved in species adopting new ways of life, that is, occupying new "adaptive niches." These include changes in hemoglobin structure that allow birds to migrate over high mountain ranges, in "antifreeze" proteins of fish that permit them to inhabit frigid waters, and in pancreatic RNAase in monkeys associated with increased herbivory. Finally, virtually every change in the color of animals and plants analyzed so far appears to involve changes in the coding regions of genes (see Hoekstra 2006), even though pigmentation is often considered to be a "form" trait, and thus hypothesized to evolve by changes in *cis*-regulation.

"Speciation" genes

Evo devo research is often explicitly motivated by a desire to explain the generation of biological diversity. While adaptation within lineages (anagenesis) represents part of the story, speciation and the generation of new lineages (cladogenesis) is the other part, without which morphological diversity would not be preserved. Here we briefly discuss what is known about the contribution of *cis*-regulatory and structural mutations to reproductive isolation.

The study of "speciation genes" (the name we use for any gene causing reproductive isolation between related taxa, even though some of these must have evolved after rather than during speciation [Coyne and Orr 2004]) is in its infancy, and hence only a handful of genes contributing to reproductive isolation have been identified. However, several patterns are already emerging (reviewed in Orr et al. 2004; Orr 2005; Noor and Feder 2006). One

is that all known speciation genes whose divergence in DNA sequence causes hybrid sterility or inviability (e.g., *OdsH*, *Hmr*, *Lhr*, *Nup96*) show evidence of rapid evolution and the signature of positive selection on mutations in coding regions (see Orr et al. 2004; Brideau et al. 2006). In addition, it is clear that *cis*-regulatory changes alone cannot be the cause of postzygotic isolation: for example, complementation tests show that *Nup96*'s effect on hybrid inviability is probably due to divergence in the protein itself (Presgraves et al. 2003). Indeed, *none* of these cases have found a contribution of *cis*-regulatory changes to reproductive isolation. And, unlike studies of genes involved in adaptation, the methods for identifying speciation genes do not suffer from the same ascertainment bias: there is no a priori expectation that genes causing inviability of hybrids, for example, should be "physiological" rather than "anatomical."

Conclusions

While the study of *cis*-regulatory evolution is an important endeavor, justifiably championed by Carroll and others, our survey of the theory and empirical data shows that the widespread enthusiasm for the importance of *cis*-regulatory change in evolution is at best premature. Analyzing the verbal theory, one finds no compelling reason to draw a distinction between the genetic basis of anatomical versus physiological evolution. Nor is there good reason to accept the a priori argument that—for either anatomy or physiology—changes in *cis*-regulatory genes are more likely to be fixed in evolution than are changes in the coding region of genes.

The data, though they may suffer from ascertainment bias, also show no strong evidence for important *cis*-regulatory change in evolution. In contrast to the many known adaptive changes in protein structure (some of which may have opened new ways of life for animals), there are only a handful of examples that are probable cases of adaptive *cis*-regulatory evolution. And, in contrast to the evidence for structural change, all three of the most widely cited cases have not yet produced definitive evidence that *cis*-regulation is involved. Moreover, these three cases focus on losses of traits rather than the origin of new traits, and in only one of the three (loss of pelvic structures in stickleback fish) is there a clear adaptive explanation for the trait loss. Obviously, we still cannot make sound generalizations about the molecular basis of adaptation. What we *can* say is that adaptations of both form and physiology are likely to involve a mixture of structural and *cis*-regulatory changes, and that structural changes are unlikely to be negligible.

At present, then, we should neither draw conclusions stronger than this nor represent to the general public that we fully understand the genetic basis of adaptation. Those who feel otherwise would do well to remember Carl Sagan's (1987, p. 45) testy remark when pressed to give an opinion about the probability of extraterrestrial intelligence: "Really, it's okay to reserve judgment until the evidence is in."

NOTE ADDED IN PROOF

Since this paper was accepted, four additional relevant studies have been published. Contrary to our view, Wray (2007) concludes that there is ample empirical evidence to support the claim that *cis*-regulatory mutations are more important than structural mutations in phenotypic evolution. However, empirical studies continue to support the importance of structural mutations in adaptive evolution. Tang et al. (2007) describe a genome-wide survey of polymorphism in humans, estimating that 10–13% of amino acid substitutions between humans and chimpanzee may be adaptive. Demuth et al. (2006) show that in humans and chimpanzees at least 6% (1,418 of 22,000 genes) of the genes in one species has no known homologue in the other, suggesting that gene duplication and gene loss occur frequently and contribute to the genetic (and perhaps phenotypic) differences between even closely related species. Both of these genomic studies, then, point to a potentially important role of structural mutations in human evolution. Finally, one other study provides yet another example of structural mutations in phenotypic evolution: loss-of-function mutations in the structural region of *anothcyanin2* (*An2*) have evolved five times independently (through five different mutations causing premature stop codons or frame-shifts), leading to an adaptive shift in pollinator syndrome in Petunia.

Demuth J. P., T. D. Bie J. E. Stajich, N. Cristianini and M. W. Hahn. 2006. The evolution of mammalian gene families. PLoS ONE 1(1): e85

Hoballah, M. E., T. Gubitz, J. Stuurman, L. Broger, M. Barone, T. Mandel, A. Dell'Olivo, M. Arnold and C. Kuhlemeir. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. Plant Cell. In press.

Tang, J. G. H., J. M. Akey and C.-I. Wu 2007. Adaptive evolution in humans revealed by the negative correlation between the polymorphism and fixation phases of evolution. Proc. Natl. Acad. Sci. U.S.A. 104:3907–3912.

Wray, G. A. 2007. The evolutionary significance of *cis*-regulatory mutations. Nature Rev. Gen. 8:206–216.

ACKNOWLEDGMENTS

LITERATURE CITED

Abzhanov, A., M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin. 2004. *Bmp4* and morphological variation of beaks in Darwin's finches. Science 305:1462–1465.

Abzhanov, A., W. P. Kuo, C. Hartmann, B. R. Grant, P. R. Grant, and C. J. Tabin. 2006. The *calmodulin* pathway and evolution of elongated beak morphology in Darwin's finches. Nature 442:563–567.

Alonso, C. R., and A. S. Wilkins. 2005. The molecular elements that underlie developmental evolution. Nat. Rev. Genet. 6:709–715.

Aminetzach, Y. T., J. M. Macpherson, and D. A. Petrov. 2005. Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. Science 309:764–767.

Andolfatto, P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. Nature 437:1149–1152.

Beldade, P., P. M. Brakefield, and A. D. Long. 2002. Contribution of *Distal-less* to quantitative variation in butterfly eyespots. Nature 415:315–318.

Benya, R. V., T. Kusui, T. Katsuno, T. Tsuda, S. A. Mantey, J. F. Battey, and R. T. Jensen. 2000. Glycosylation of the gastrin-releasing peptide receptor and its effect on expression, G-protein coupling, and receptor modulatory processes. Mol. Pharmacol. 58:1490–1501.

Brembeck, F. H., M. Rosario, and W. Birchmeier. 2006. Balancing cell adhesion and *Wnt* signaling, the key role of *beta-catenin*. Curr. Opin. Genet. Dev. 16:51–59.

Breuker, C. J., V. Debat, and C. P. Klingenberg. 2006. Functional evo-devo. Trends Ecol. Evol. 21:488–492.

Brideau, N., H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash. 2006. Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. Science 314:1292–1295.

Britten, R. J., and E. H. Davidson. 1969. Gene regulation for higher cells: a theory. Science 165:349–357.

———. 1971. Repetitive and non-repetitive DNA sequences and a speculation on origins of evolutionary novelty. Q. Rev. Biol. 46:111–138.

Budar, F., and G. Pelletier. 2001. Male sterility in plants: occurrence, determinism, significance and use. C. R. Life Sciences 324:543–550.

Budar, F., P. Touzet, and R. De Paepe. 2003. The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. Genetica 117:3–16.

Bustamante, C. D., A. Fledel-Alon, S. Williamson, R. Nielsen, M. T. Hubisz, S. Glanowski, D. M. Tanenbaum, T. J. White, J. J. Sninsky, R. D. Hernandez, D. Civello, M. D. Adams, M. Cargill, and A. G. Clark. 2005. Natural selection on protein-coding genes in the human genome. Nature 437:1153–1157.

Carlini, D. B., and W. Stephan. 2003. In vivo introduction of unpreferred synonymous codons into the *Drosophila* Adh gene results in reduced levels of ADH protein. Genetics 163:239–243.

Carroll, S. B. 2000. Endless forms: the evolution of gene regulation and morphological diversity. Cell 101:577–580.

Carroll, S. B., J. K. Grenier, and S. D. Weatherbee. 2001. From DNA to diversity: molecular genetics and the evolution of animal design. Blackwell Publishing, Malden, MA.

———. 2005a. Evolution at two levels: on genes and form. PLoS Biology 3:1159–1166.

———. 2005b. Endless forms most beautiful: the new science of evo devo. W. W. Norton & Co., New York.

———. 2006. The making of the fittest: DNA and the ultimate forensic record of evolution. W. W. Norton & Co., New York.

Chang, S. M., Y. Q. Lu, and M. D. Rausher. 2005. Neutral evolution of the non-binding region of the anthocyanin regulatory gene *Ipmyb1* in *Ipomoea*. Genetics 170:1967–1978.

Cheng, C. H. C. 1998. Evolution of the diverse antifreeze proteins. Curr. Opin. Genet. Dev. 8:715–720.

Chung, H., M. R. Bogwitz, C. McCart, A. Andrianopoulos, R. H. ffrench-Constant, P. Batterham, and P. J. Daborn. 2007. *Cis*-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. Genetics: *In press*.

Clark, A. G., S. Glanowski, R. Nielsen, P. D. Thomas, A. Kejariwal, M. A. Todd, D. M. Tanenbaum, D. Civello, F. Lu, B. Murphy, S. Ferriera, G. Wang, X. G. Zheng, T. J. White, J. J. Sninsky, M. D. Adams, and M. Cargill. 2003. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. Science 302:1960–1963.

Colossimo, P. F., K. E. Hoemann, S. Balabhadra, G. Villarreal, M. Dickinson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley.

2005. Widespread parallel evolution in sticklebacks by repeated fixation of ecotdysplasin alleles. Science 307:1928–1933.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil, S. Jeffers, N. Tijet, T. Perry, D. Heckel, P. Batterham, R. Feyereisen, T. G. Wilson, and R. H. ffrench-Constant. 2002. A single P450 allele associated with insecticide resistance in *Drosophila*. Science 297:2253–2256.

Davidson, E. H. 2001. Genomic regulatory systems. Academic Press, San Diego, CA.

Davidson, E. H., and D. H. Erwin. 2006. Gene regulatory networks and the evolution of animal body plans. Science 311:796–800.

Davies, P. L., and C. L. Hew. 1990. Biochemistry of fish antifreeze proteins. FASEB J. 4:2460–2468.

Dehal, P., and J. L. Boore. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3:1700–1708.

Demuth J. P., T. D. Bie, J. E. Stajich, N. Cristianini and M. W. Hahn. 2006. The evolution of mammalian gene families. PLoS ONE 1(1): e85.

Devonshire, A. L., L. M. Field, S. P. Foster, G. D. Moores, M. S. Williamson, and R. L. Blackman. 1998. The evolution of insecticide resistance in the peach-potato aphid, Myzus persicae. Phil. Trans. R. Soc. Lond. B 353:1677–1684.

Dorus, S., E. J. Vallender, P. D. Evans, J. R. Anderson, S. L. Gilbert, M. Mahowald, G. J. Wyckoff, C. M. Malcom, and B. T. Lahn. 2004. Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. Cell 119:1027–1040.

Doebley, J., and L. Lukens. 1998. Transcriptional regulators and the evolution of plant form. Plant Cell 10:1075–1082.

Duboule, D., and P. Dolle. 1989. The structural and functional organization of the murine Hox gene family resembles that of *Drosophila* homeotic genes. EMBO J. 8:1497–1505.

Duman, J. G. 2001. Antifreeze and ice nucleator proteins in terrestrial arthropods. Annu. Rev. Physiol. 63:327–357.

Durbin, M. L., K. E. Lundy, P. L. Morrell, C. L. Torres-Martinez, and M. T. Clegg. 2003. Genes that determine flower color: the role of regulatory changes in the evolution of phenotypic adaptations. Mol. Phylogenet. Evol. 29:507–518.

Eizirik, E., N. Yuhki, W. E. Johnson, M. Menotti-Raymond, S. S. Hannah, and S. J. O'Brien. 2003. Molecular genetics and evolution of melanism in the cat family. Curr. Biol. 13:448–453.

El-Assal, S. E. D., C. Alonso-Blanco, A. J. M. Peeters, V. Raz, and M. Koornneef. 2001. A QTL for flowering time in *Arabidopsis* reveals a novel allele of CRY2. Nat. Genet. 29:435–440.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pp. 82–115 *in* T. J. M. Schopf, ed. Models in paleobiology. Freeman, Cooper, San Francisco, CA.

Escriva, H., S. Bertrand, P. Germain, M. Robinson-Rechavi, M. Umbhauer, J. Cartry, M. Duffraisse, L. Holland, H. Gronemeyer, and V. Laudet. 2006. Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet. 2:955–965.

Fernald, R. D. 2000. Evolution of eyes. Curr. Opin. Neurobiol. 10:444–450.

ffrench-Constant, R. H., T. A. Rocheleau, J. C. Steichen, and A. E. Chalmers. 1993. A point mutation in a *Drosophila* GABA receptor confers insecticide resistance. Nature 363:448–451.

———. 1994. The molecular and population genetics of cyclodiene insecticide resistance. Insect Biochem. Mol. Biol. 24:335–345.

ffrench-Constant, R. H., P. J. Daborn, and G. Le Goff. 2004. The genetics and genomics of insecticide resistance. Trends Genet. 20:163–170.

Field, L. M., R. L. Blackman, C. Tyler-Smith, and A. L. Devonshire. 1999. Relationship between amount of esterase and gene copy number in insecticide-resistant *Myzus persicae* (Sulzer). Biochem. J. 339:737–742.

Field, L. M., and A. L. Devonshire. 1998. Evidence that the E4 and FE4 esterase genes responsible for insecticide resistance in the aphid *Myzus persicae* (Sulzer) are part of a gene family. Biochem. J. 330:169–173.

Field, L. M., A. L. Devonshire, and B. G. Forde. 1988. Molecular evidence that insecticide resistance in peach potato aphids (Myzus persicae Sulzer) results from amplification of an esterase gene. Biochem. J. 251:309–312.

Fisher, R. A. 1930. The genetical theory of natural selection: a complete variorum edition. Oxford Univ. Press, Oxford, U.K.

Fletcher, G. L., C. L. Hew, and P. L. Davies. 2001. Antifreeze proteins of teleost fishes. Annu. Rev. Physiol. 63:359–390.

Glazko, G., V. Veeramachaneni, M. Nei, and W. Makayowski. 2005. Eighty percent of proteins are different between humans and chimpanzees. Gene 346:215–219.

Gompel, N., B. Prud'homme, P. J. Wittkopp, V. A. Kassner, and S. B. Carroll. 2005. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. Nature 433:481–487.

Gould, S. J., and N. Eldredge. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology 3:115–151.

———. 1980. Is a new and general theory of evolution emerging? Paleobiology 6:119–130.

Gould, S. J., and N. Eldredge. 1993. Punctuated equilibrium comes of age. Nature 366:223–227.

Gratten, J., D. Beraldi, B. V. Lowder, A. F. McRae, P. M. Visscher, J. M. Pemberton, and J. Slate. 2006. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for a coat-colour polymorphism in a free-living population of Soay sheep. Proc. R. Soc. Lond. B 274:619–626.

Hill, R. E., P. F. Jones, A. R. Rees, C. M. Sime, M. J. Justice, N. G. Copeland, N. A. Jenkins, E. Graham, and D. R. Davidson. 1989. A new family of mouse homeo box-containing genes—molecular structure, chromosomal location, and developmental expression of Hox-7.1. Genes Dev. 3:26–37.

Hoballah, M. E., T. Gubitz, J. Stuurman, L. Broger, M. Barone, T. Mandel, A. Dell'Olivo, M. Arnold and C. Kuhlemeir. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. Plant Cell. In press.

Hoekstra, H. E. 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity 97:222–234.

Hoekstra, H. E., R. J. Hirschmann, R. A. Bundey, P. A. Insel, and J. P. Crossland. 2006. A single amino acid contributes to adaptive color pattern in beach mice. Science 313:101–104.

Hsia, C. C., and W. McGinnis. 2003. Evolution of transcription factor function. Curr. Opin. Gen. Dev. 13:199–206.

Jacob, F., and J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of protein. J. Mol. Biol. 8:318–356.

———. 1977. Evolution and tinkering. Science 196:1161–1166.

Jessen, T. H., R. E. Weber, G. Fermi, J. Tame, and G. Braunitzer. 1991. Adaptation of bird hemoglobins to high altitudes: demonstration of molecular mechanism by protein engineering. Proc. Nat. Acad. Sci. USA 88:6519–6522.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino, and C. Dean. 2000. Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290:344–347.

Kimchi-Sarfaty, C., J. M. Oh, I.-W. Kim, Z. E. Sauna, A. M. Calcagno, S. V. Ambudkar, and M. M. Gottesman. 2007. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science 315:525–528.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge, U.K.

King, M. C., and A. C. Wilson. 1975. Evolution at two levels in humans and chimpanzees. Science 188:107–116.

Kirschner, M., and J. Gerhart. 1998. Evolvability. Proc. Natl. Acad. Sci. USA 95:8420–8427.

Kleinschmidt, T., J. Marz, K. D. Jurgens, and G. Braunitzer. 1986. Interaction of allosteric effectors with alpha-globin chains and high altitude respiration of mammals—the primary structure of 2 *Tylopoda* hemoglobins with high oxygen affinity—Vicuna (*Lama vicugna*) and Alpaca (*Lama pacos*). Biol. Chem. Hoppe-Seyler 367:153–160.

Levine, M., and R. Tjian. 2003. Transcription regulation and animal diversity. Nature 424:147–151.

Li, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, MA.

Lohr, U., and L. Pick. 2005. Cofactor-interaction motifs and the cooption of a homeotic hox protein into the segmentation pathway of *Drosophila* melanogaster. Curr. Biol. 15:643–649.

Long, M., E. Betran, K. Thornton, and W. Wang. 2003. The origin of new genes: glimpses from the young and old. Nat. Rev. Genet. 4:865–875.

Lynch, M., and J. S. Conery. 2000. The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155.

McNutt, M. C., R. Tongbai, W. W. Cui, I. Collins, W. J. Freebern, I. Montano, C. M. Haggerty, G. V. R. Chandramouli, and K. Gardner. 2005. Human promoter genomic composition demonstrates non-random groupings that reflect general cellular function. BMC Bioinformatics 6:259.

Miyazaki, M., K. Ohyama, D. Y. Dunlap, and F. Matsumura. 1996. Cloning and sequencing of the para-type sodium channel gene from susceptible and kdr-resistant German cockroaches (*Blattella germanica*) and house fly (*Musca domestica*). Mol. Gen. Genet. 252:61–68.

Mundy, N. I., N. S. Badcock, T. Hart, K. Scribner, K. Janssen, and N. J. Nadeau. 2004. Conserved genetic basis of a quantitative plumage trait involved in mate choice. Science 303:1870–1873.

Nachman, M. W., H. E. Hoekstra, and S. L. D'Agostino. 2003. The genetic basis of adaptive melanism in pocket mice. Proc. Natl. Acad. Sci. USA 100:5268–5273.

Nakayama, T., and Y. Setoguchi. 1992. Structural change of a particular *H2b* histone gene possibly results in differences in its transcriptional regulation in different chicken tissues. Nuc. Acids Res. 20:1309–1314.

Newcomb, R. D., P. M. Campbell, D. L. Ollis, E. Cheah, R. J. Russell, and J. G. Oakeshott. 1997. A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. Proc. Nat. Acad. Sci. U.S.A. 94:7464–7468.

Nielsen, R., C. Bustamante, A. G. Clark, S. Glanowski, T. B. Sackton, M. J. Hubisz, A. Fledel-Alon, D. M. Tanenbaum, D. Civello, T. J. White, J. J. Sninsky, M. D. Adams, and M. Cargill. 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol. 3:976–985.

Noor, M. A. F., and J. L. Feder. 2006. Speciation genetics: evolving approaches. Nat. Rev. Genet. 7:851–861.

Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, New York, NY.

Orr, H. A. 2003. The distribution of fitness effects among beneficial mutations. Genetics 163:1519–1526.

———. 2005. The genetic basis of reproductive isolation: insights from *Drosophila*. Proc. Nat. Acad. Sci. USA 102:6522–6526.

Orr, H. A., J. P. Masly, and D. C. Presgraves. 2004. Speciation genes. Curr. Opin. Genet. Dev. 14:675–679.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. Ann. Rev. Genet. 34:401–437.

Perez-Moreno, M., and E. Fuchs. 2006. Catenins: keeping cells from getting their signals crossed. Dev. Cell 11:601–612.

Perutz, M. F. 1983. Species adaptation in a protein molecule. Mol. Biol. Evol. 1:1–28.

Piatigorsky, J. 2003. Gene sharing, lens crystallins and speculations on an eye/ear evolutionary relationship. Integrative Compar. Biol. 43: 492–499.

Piccinini, M., T. Kleinschmidt, K. D. Jurgens, and G. Braunitzer. 1990. Primary structure and oxygen-binding properties of the hemoglobin from guanaco (*Lama guanacoe*, Tylopoda). Biol. Chem. Hoppe-Seyler 371:641–648.

Pollard, K. S., S. R. Salama, B. King, A. D. Kern, T. Dreszer, S. Katzman, A. Siepel, J. S. Pedersen, G. Bejerano, R. Baertsch, K. R. Rosenbloom, J. Kent, and D. Haussler. 2006a. Forces shaping the fastest evolving regions in the human genome. PLoS Genet. 2:1599–1611.

Pollard, K. S., S. R. Salama, N. Lambert, M. A. Lambot, S. Coppens, J. S. Pedersen, S. Katzman, B. King, C. Onodera, A. Siepel, A. D. Kern, C. Dehay, H. Igel, M. Ares, P. Vanderhaeghen, and D. Haussler. 2006b. An RNA gene expressed during cortical development evolved rapidly in humans. Nature 443:167–172.

Prager, E. M., and A. C. Wilson. 1988. Ancient origin of lactalbumin from lysozyme: analysis of DNA and amino acid sequences. J. Mol. Evol. 27:326–335.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature 423:715–719.

Protas, M. E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, L. I. Zon, R. Borowsky, and C. J. Tabin. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nature Genet. 38:107–111.

Prud'homme, B., N. Gompel, A. Rokas, V. A. Kassner, T. M. Williams, S. D. Yeh, J. R. True, and S. B. Carroll. 2006. Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. Nature 440:1050–1053.

Quattrocchio, F., J. Wing, K. van der Woude, E. Souer, N. de Vetten, J. Mol, and R. Koes. 1999. Molecular analysis of the *anthocyanin*2 gene of petunia and its role in the evolution of flower color. Plant Cell 11:1433–1444.

Ramsey, J. M., and D. W. Schemske. 1998. The dynamics of polyploid formation and establishment in flowering plants. Ann. Rev. Ecol. Syst. 29:467–501.

Rathz, D. A., K. M. Brown, L. A. Kramer, and S. B. Liggett. 2002. Amino acid 49 polymorphisms of the human beta(1)-adrenergic receptor affect agonist-promoted trafficking. J. Cardiovasc. Pharmacol. 39:155–160.

Reed, R. D., and M. S. Serfas. 2004. Butterfly wing pattern evolution is associated with changes in a *Notch/Distal-less* temporal pattern formation process. Curr. Biol. 14:1159–1166.

Ronshaugen, M., N. McGinnis, and W. McGinnis. 2002. Hox protein mutation and macroevolution of the insect body plan. Nature 415:914–917.

Rosenblum, E. B., H. E. Hoekstra, and M. W. Nachman. 2004. Adaptive reptile color variation and the evolution of the MC1R gene. Evolution 58:1794–1808.

Ross, J. 1995. Messenger RNA stability in mammalian cells. Microbiol. Rev. 59:423–450.

Sagan, C. 1987. The burden of skepticism. Skeptical Inquirer 12:38–46.

Schiavi, S. C., C. L. Wellington, A. B. Shyu, C. Y. A. Chen, M. E. Greenberg, and J. G. Belasco. 1994. Multiple elements in the *c-fos* protein coding region facilitate messenger RNA deadenylation and decay by a mechanism coupled to translation. J. Biol. Chem. 269:3441–3448.

Schnable, P. S., and R. P. Wise. 1998. The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant. Sci. 3:175–180.

Sebat, J., B. Lakshmi, J. Troge, J. Alexander, J. Young, et al. 2004. Large-scale copy number polymorphism in the human genome. Science 305:525—528.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jonsson, D. Schluter, and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428:717–723.

———. 2006. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 439:1014.

Shen, L. X., J. P. Basilion, and V. P. Stanton. 1999. Single-nucleotide polymorphisms can cause different structural folds of mRNA. Proc Nat. Acad. Sci. U.S.A. 96:7871–7876.

Soderlund, D. M., and D. C. Knipple. 2003. The molecular biology of knockdown resistance to pyrethroid insecticides. Insect Biochem. Mol. Biol. 33:563–577.

Stern, D. L. 1998. A role of *Ultrabithorax* in morphological differences between *Drosophila* species. Nature 396:463–466.

———. 2000. Perspective: Evolutionary developmental biology and the problem of variation. Evolution 54:1079–1091.

Stone, J. R., and G. A. Wray. 2001. Rapid evolution of *cis*-regulatory sequences via local point mutations. Mol. Biol. Evol. 18:1764–1770.

Strahl, B. D., and C. D. Allis. 2000. The language of covalent histone modifications. Nature 403:41–45.

Sucena, E., and D. L. Stern. 2000. Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by *cis*-regulatory evolution of ovo/shaven-baby. Proc. Nat. Acad. Sci. U.S.A. 97:4530–4534.

Tang, J. G. H., J. M. Akey and C.-I. Wu. 2007. Adaptive evolution in humans revealed by the negative correlation between the polymorphism and fixation phases of evolution. Proc. Natl. Acad. Sci. U.S.A. 104:3907–3912.

Theron, E., K. Hawkins, E. Bermingham, R. E. Ricklefs, and N. I. Mundy. 2001. 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.

Tomarev, S. I., and J. Piatigorsky. 1996. Lens crystallins of invertebrates: diversity and recruitment from detoxification enzymes and novel proteins. Eur. J. Biochem. 235:449–465.

Valentine, J. W. 2004. On the origin of phyla. Univ. of Chicago Press, Chicago, IL.

Weber, R. E., H. Ostojic, A. Fago, S. Dewilde, M. L. Van Hauwaert, L. Moens, and C. Monge. 2002. Novel mechanism for high-altitude adaptation in hemoglobin of the Andean frog, *Telmatobius peruvianus*. Am. J. Physiol. Regul. Integ. Comp. Physiol. 283:R1052–R1060.

Weill, M., P. Fort, A. Berthomieu, M. P. Dubois, N. Pasteur, and M. Raymond. 2002. A novel acetylcholinesterase gene in mosquitoes codes for the insecticide target and is non-homologous to the *ace* gene in *Drosophila*. Proc. Roy. Soc. Lond. B 269:2007–2016.

Weill, M., G. Lutfalla, K. Mogensen, F. Chandre, A. Berthomieu, C. Berticat, N. Pasteur, A. Philips, P. Fort, and M. Raymond. 2003. Insecticide resistance in mosquito vectors. Nature 423:136–137.

Werner, J. D., J. O. Borevitz, N. H. Uhlenhaut, J. R. Ecker, J. Chory, and D. Weigel. 2005. FRIGIDA-independent variation in flowering time of natural Arabidopsis thaliana accessions. Genetics 170:1197–1207.

Werner, J. D., J. O. Borevitz, N. Warthmann, G. T. Trainer, J. R. Ecker, J. Chory, and D. Weigel. 2005. Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural floweringtime variation. Proc. Nat. Acad. Sci. U.S.A. 102:2460–2465.

Wilkins, A. S. 2002. The evolution of developmental pathways. Sinauer Associates, Sunderland, MA.

Williamson, M. S., D. Martinez Torres, C. A. Hick, and A. L. Devonshire. 1996. Identification of mutations in the housefly para-type sodium channel

gene associated with knockdown resistance (kdr) to pyrethroid insecticides. Mol. Gen. Genetics 252:51–60.

Wilson, A. C., L. R. Maxson, and V. M. Sarich. 1974a. Two types of molecular evolution: evidence from studies of interspecific hybridization. Proc. Natl. Acad. Sci. USA 71:2843–2847.

———. 1974b. Importance of gene rearrangement in evolution: evidence from studies on rates of chromosomal, protein, and anatomical evolution. Proc. Nat. Acad. Sci. USA 71:3028–3030.

Wilson, A. C. 1975. Evolutionary importance of gene regulation. Stadler Symp. 7:117–133.

Wisdom, R., and W. Lee. 1991. The protein coding region of c-myc messenger RNA contains a sequence that specifies rapid messenger RNA turnover and induction by protein synthesis inhibitors. Genes Dev. 5:232–243.

Wistow, G. J., J. W. M. Mulders, and W. W. Dejong. 1987. The enzyme lactate-dehydrogenase as a structural protein in avian and crocodilian lenses. Nature 326:622–624.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in *cis* and *trans* gene regulation. Nature 430:85–88.

Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman, and L. A. Romano. 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20:1377–1419.

Wray, G. A. 2007. The evolutionary significance of *cis*-regulatory mutations. Nature Rev. Gen. 8:206–216.

Yang, C., E. Bolotina, T. Jiang, F. M. Sladekc, and E. Martinez. 2006. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene: *In press*.

Yasui, D., M. Miyano, S. T. Cai, P. Varga-Weisz, and T. Kohwi-Shigematsu. 2002. SATB1 targets chromatin remodeling to regulate genes over long distances. Nature 419:641–645.

Yokoyama, S., H. Zhang, F. B. Radlwimmer, and N. S. Blow. 1999. Adaptive evolution of color vision of the Comoran coelacanth (*Latimeria chalumnae*). Proc. Nat. Acad. Sci. USA 96:6279–6284.

Yokoyama, S., F. B. Radlwimmer, and N. S. Blow. 2000. Ultraviolet pigments in birds evolved from violet pigments by a single amino acid change. Proc. Nat. Acad. Sci. USA 97:7366–7371.

Zhang, J. Z. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nature Genet. 38:819–823.

Zufall, R. A., and M. D. Rausher. 2003. The genetic basis of a flower color polymorphism in the common morning glory (*Ipomoea purpurea*). J. Heredity 94:442–448.

———. 2004. Genetic changes associated with floral adaptation restrict future evolutionary potential. Nature 428:847–850.

Associate Editor: M. Rausher

Appendix. Mutations in *cis*-regulatory or structural regions that contribute to adaptation.

Gene	Structural/ regulatory	Gain/ Loss[†]	Main result	Evidence	Missing evidence	Reference
MORPHOLOGY						
Pitx1 (transcription factor)	R	L	Loss of pelvic spines associated with novel habitat (predators) in threespine sticklebacks (*Gasterosteus aculeatus*).	QTL mapping; in situ hybridization; candidate gene information; amino acid sequence comparison.	Neighboring genes not ruled out; functional mutations not identified; transgenic assays needed.	Shapiro et al. 2004, 2006
Ovo/shavenbaby	R	L	Loss of trichomes in several species of *Drosophila* larvae. *D simulans/D. sechellia*.	Genetic mapping; deletion mapping; gene complementation.	Functional mutations not identified; adaptive significance unknown.	Sucena and Stern 2000; Stern (pers. comm.)
Yellow (pigment enzyme)	R	G/L	Male wing spots used in courtship displays in some *Drosophila* species.	Transgenic assays with different *cis*-regulatory elements associated with changes in *yellow* protein expression.	a.a. changes not ruled out; functional mutations not identified; association with pigmentation changes not shown.	Gompel et al. 2005; Prud'homme et al. 2006
Eda (transcription factor)	R/S?	L	Loss of lateral plates associated with novel habitat (predators) in threespine sticklebacks (*Gasterosteus aculeatus*).	QTL and LD mapping; candidate gene information; amino acid sequencing; transgenic results in recovery of some plating.	Functional mutation not identified (only haplotype); contribution of mutations in coding region not ruled out.	Colossimo et al. 2005
Ubx (Hox transcription factor)	S	G	Gain of limbs via loss of serine phosphorylation sites in C terminus results in loss of limb repression in *Artemia* crustaceans versus *Drosophila*.	In vitro assay using site directed mutagenesis shows loss of binding and pathway disruption.	In vivo transgenic assay.	Ronshaugen et al. 2003
Oca2 (ocular albinism enzyme)	S	L	Loss of pigment in three independent cave populations (*Astyanax fasciatus*).	Deletions in different exons in different populations; loss of expression using cell-based functional assay.	Mutation in *Oca2* not identified in 1 of the 3 populations; transgenics assays needed.	Protas et al. 2006
Tyrp1 (tyrosinase-related protein 1)	S	L	Single a.a. mutation associated with light-colored island sheep (*Ovis aires*) maintained as a balanced polymorphism.	One a.a. change is perfectly correlated with color in pedigree of 500 sheep; no change in expression level between color morphs.	Functional assay showing the effect of a.a. substitution on protein function and phenotype; selective agent unclear.	Gratten et al. 2006

Continued.

Appendix. Continued

Gene	Structural/ regulatory	Gain/ Loss[†]	Main result	Evidence	Missing evidence	Reference
Mc1r (melanocortin receptor)	S	L (partial)	Cryptic pigment pattern in beach mice (*Peromyscus polionotus*).	One derived a.a. in coding region associated with patterning in large cross; a.a. change shown to reduce receptor signaling and ligand binding in cell culture assays.	Neighboring genes not ruled out (but no changes in mRNA expression levels observed between color morphs).	Hoekstra et al. 2006
Mc1r (melanocortin receptor)	S	G	Melanism in lava-dwelling pocket mice (*Chaetodipus intermedius*).	Four linked a.a. changes in coding region perfectly associated with color.	Individual effects of a.a. mutations unknown; functional assay needed; regulatory regions not ruled out.	Nachman et al. 2003
Mc1r (melanocortin receptor)	S	G (additive)	Sexually selected plumage pattern in snow geese (*Anser caerulescens*) involved in assortative mating.	Two a.a. in coding region correlated with quantitative plumage variation.	Functional assay needed; regulatory regions not ruled out.	Mundy et al. 2004
Mc1r (melanocortin receptor)	S	G (additive)	Plumage pattern in and artic skua (*Stercorarius parasiticus*).	One a.a. in coding region correlated with quantitative plumage variation.	Functional assay needed; regulatory regions not ruled out.	Mundy et al. 2004
Mc1r (melanocortin receptor)	S	G	Melanism in bananaquits (*Coerebra flaveola*); clinal variation.	One a.a. in coding region perfectly correlated with melanism.	Functional assay needed; regulatory regions not ruled out; selective agent unknown.	Theron et al. 2001
Mc1r (melanocortin receptor)	S	L (additive)	Cryptic pigmentation in little striped whiptail (*Aspidocelis inornata*).	One a.a. change statistically correlated with blanched color.	Functional assay needed; regulatory regions and linked genes not ruled out.	Rosenblum et al. 2004
Mc1r (melanocortin receptor)	S	L (additive)	Cryptic pigmentation in eastern fence lizards (*Sceloporus undulatus*).	One a.a. change statistically correlated with blanched color.	Functional assay needed; regulatory regions and linked genes not ruled out.	Rosenblum et al. 2004
Mc1r (melanocortin receptor)	S	L (additive)	Cryptic pigmentation in lesser earless lizard (*Holbrookia maculata*).	One a.a. change statistically correlated with blanched color.	Need to control for population structure; functional assay needed; regulatory regions and linked genes not ruled out.	Rosenblum et al. 2004

Continued.

Appendix. Continued

Gene	Structural/ regulatory	Gain/ Loss[†]	Main result	Evidence	Missing evidence	Reference
Mc1r (melanocortin receptor)	S	G	Association with melanism in jaguars (*Panthera onca*).	Deletion in first transmembrane region perfectly associated with melanism; segregation in pedigree.	Functional assay needed; regulatory regions and linked genes not ruled out; selective. agent unclear	Eizirik et al. 2003
Mc1r (melanocortin receptor)	S	G (additive)	Association with melanism in jagurundis (*Herpailurus yaguarondi*).	Single a.a. change perfectly associated with melanism; segregation in pedigree.	Functional assay needed; regulatory regions and linked genes not ruled out; selective agent unclear.	Eizirik et al. 2003
Ipmyb1 (*myb* transcription factor)	S	L	Blue to white flower color change in morning glory (*Ipomoea purpura*), balanced polymorphism.	Transposition causing a frame-shift mutation.	Functional assay needed; regulatory regions and linked genes not ruled out.	Chang et al. 2005
An2, homologous to *Impmyb1* (transcription factor)	S	L (additive)	Purple to white flower color change in petunias (*Petunea axilaris*) associated with a change in pollination syndrome.	Transposition mediated deletion/frameshift and in an independently derived allele a nonsense mutation are both large contributors to phenotypic divergence.	Structural changes in enzymes may also contribute; functional assay needed; regulatory regions and linked genes not ruled out.	Quattrocchio et al. 1999
F3′H and *DFR* (anthocyanin enzymes)	S	L	Blue to red flower color change in morning glories (*Ipomoea quamoclit*) associated with change in pollination syndrome.	Knockout allele of *F3′H* enzyme and amino acid change in *DFR* enzyme (independent effects of each gene mutation are unknown).	Functional assays needed; regulatory regions and linked genes not ruled out (note that reduced *F3′H* expression can also lead to color change but casual mutation unknown).	Zufall and Rausher 2003, 2004
PLANT LIFE HISTORY						
Flm (flower timing transcription factor)	S	L	Early flowering time in isolates from nature (*Arabidopsis thaliana*).	QTL; microarrays; locus deletion; isogenic lines; quantitative transgenic complementation.	Looked in accessions; functional assay needed; regulatory regions and linked genes not ruled out.	Werner et al. 2005
Flc (flower timing transcription factor)	S	L	Clinal variation in leafy phenotype (*Arabidopsis thaliana*).	Nonsense a.a. mutation associated with early flower timing and alternative splicing disrupts function (null allele).	Looked in accessions; functional assay needed; regulatory regions and linked genes not ruled out.	Werner et al. 2005

Continued.

Appendix. Continued

Gene	Structural/ regulatory	Gain/ Loss[†]	Main result	Evidence	Missing evidence	Reference
Cry2	S	R	Clinal variation in flowering time (*Arabidopsis thaliana*).	QTL and positional cloning; protein polymorphism; single a.a. change induces light-induced downregulation of Cry2.	Looked in accessions; functional assay needed; regulatory regions and linked genes not ruled out.	El-Assal et al. 2001
Frigida (flower timing transcription factor)	S	L	Clinal variation in flowering time (*Arabidopsis thaliana*).	Deletion polymorphism disrupting open reading frame associated with early flowering times.	Looked in accessions; functional assay needed; regulatory regions and linked genes not ruled out.	Johanson et al. 2000
ALTITUDINAL PHYSIOLOGY						
Hb (hemoglobin tetramer)	S	G	Comparison high-altitude bar-headed goose (*Anser indicus*) and Andean goose (*Chloephaga melanoptera*) to low-altitude greylag goose (*Anser anser*).	Single a.a. change; increased O_2 affinity binding assays.	Functional assay needed; regulatory regions and linked genes not ruled out.	Perutz 1983; Jessen et al 1991
Hb (hemoglobin tetramer)	S	G	Comparison of high-altitude camelids [llama (*Lama glama*), alpaca (*L. pacos*), guanaco (*L. guanacoe*), vicuna (*L. vicugna*)] relative to low-land camelids (genus *Camelus*).	a.a. change; O_2 affinity binding assay.	Functional assay needed; regulatory regions and linked genes not ruled out.	Perutz 1983; Kleinschmidt et al. 1986; Piccinini et al. 1990
Hb (hemoglobin tetramer)	S	G	Analysis of high-altitude Andean frog (*Telmatobius peruvianus*.)	a.a. changes in alpha chains to reduce chloride binding; O_2 affinity binding assay.	Regulatory regions and linked genes not ruled out.	Weber et al. 2002
INSECTICIDE RESISTANCE						
Rdl (GABA receptor)	S	L (binding)	Resistance to insecticide dieldrin in *D. melanogaster* (& five other species of insects).	Single a.a. change in coding region (ala -> ser or gly). In *D. melanogaster*, RNA injected into frog oocytes renders them less sensitive to dieldrin. In five other species, exact same substitution associated with resistance.	Regulatory regions not ruled out.	ffrench-Constant 1994; ffrench-Constant et al. 1993, 2004.

Continued.

Appendix. Continued

Gene	Structural/regulatory	Gain/Loss†	Main result	Evidence	Missing evidence	Reference
kdr (knockdown resistance) sodium channel gene	S	L	Resistance to DDT in house flies (Musca domestica & six other species of insects).	1–2 a.a. subs. depending on allele. Correlation in species between substitution and resistance.	Functional assay needed; regulatory regions and linked genes not ruled out.	Williamson et al. 1996; Miyazaki et al. 1996; Soderlund and Knipple 2003
Ace-1 (acetyl cholinesterase enzyme)	S	L	Resistance to insecticide in mosquitoes (Anopheles gambia, Culex pipiens).	Single a. a. substitution, same in both species. Correlation between substitution and resistance.	Functional assay needed; regulatory regions and linked genes not ruled out. Small sample of strains.	Weill et al. 2003
Cy6g1 (cytochrome P450 enzyme)	R	G (overtranscription)	Resistance to DDT in Drosophila melanogaster	Insertion of Accord transposon at 5′ end outside of gene, causes overtranscription of gene; Perfect correlation between transposons and resistance; genetic manipulation by overtranscription; transgenic constructs containing the TE in flies.	None	Daborn et al. 2002; Chung et al. 2007
LcαE7 (esterase)	S	G (amplification)	Resistance to organophosphate insecticides in sheep blowfly (Lucilia cuprina).	Single a.a. replacement; recombinant mutant enzyme has increased organophosphatehydrolysis.	Show change in biochemistry associated with resistance but do not higher resistance of genetically engineered flies.	Newcomb et al. 1997
E4 (esterase)	S	G (amplification)	Mutant protein sequesters insecticide in peach potato aphid (Myzus persicae).	Either truncated protein (FE4) or amplified protein (E4). Genes are regulated in different ways, too. Correlation between resistance and no. of genes, as well as known mechanism of sequestration of insecticide by enzyme.	No experiments to show whether regulation difference is adaptive, although duplicated genes have different regulation.	Field et al. 1988, 1998, 1999; Devonshire et al. 1998

Continued.

Appendix. Continued

Gene	Structural/ regulatory	Gain/ Loss†	Main result	Evidence	Missing evidence	Reference
CKKovI (choline kinase)	S	G (novel protein)	Resistance to pesticide AZM in *D. melanogaster*.	Allele makes a truncated peptide by insertion of TE in coding region followed by 7 a.a. changes in remaining peptide. Correlation of genotype with pesticide resistance.	Functional assay needed; regulatory regions and linked genes not ruled out.	Aminetzach et al. 2005
VISUAL PIGMENTS						
SWS1 opsin	S	G (novel protein)	UV sensitivity in bird vison.	Change of single a.a. makes violet pigment into UV pigment; determined by absorption spectrum of purified pigment.	Regulatory regions not ruled out.	Yokoyama et al. 2000
RH1 and RH2 opsins	S	G (novel protein)	Blue sensitivity of coelocanth vision (*Latimeria chalumnae*).	Change of two a.a.'s in each of two pigments changes sensitivity in expected direction; determined by absorption spectrum of purified pigment.	Regulatory regions not ruled out.	Yokoyama et al. 1999
OTHER						
AFPs (antifreeze proteins) and AFGPs (antifreeze glycoproteins)	S	G (novel protein—duplication & a.a. subs).	Resistance of cytoplasm to freezing in various fish, insects, and plants.	Changes in duplicated genes (often involving repeated a.a.'s), confers resistance to ice crystals as shown in functional studies of purified proteins.	Regulatory regions not ruled out: no information about concordant changes in regulation of antifreeze genes.	Cheng 1998; Duman 2001; Fletcher et al. 2001
pancreatic RNAse RNAASE1B	S	G (duplication & a.a. subs)	Ability to recycle nitrogen in the small intestine by increased RNAase activity in the guereza monkey (*Colobus guereza*) and douc langur (*Pygathrix nemaeus*).	pHs examined of recombinant purified proteins: new proteins operated better at lower pHs (as in small intestine).	Regulatory regions not ruled out.	Zhang 2006
lens crystallin	S	G [co-opted (and often altered) proteins & enzymes]	Recruitment and change of enzymes and proteins in vertebrates and invertebrates (many of them products of duplicate genes) into eye helps focus light.	Proteins function as an intraocular matrix for focusing light.	Regulatory regions not ruled out.	Wistow et al. 1987; Tomarev and Piatigorsky 1996; Fernald 2004

† Change in protein expression or function associated with mutations causing gain of function (e.g., increased or altered expression or activity) or loss of function (e.g., decreased or absence of expression, reduced activity).