Switches within DNA that govern when and where genes are turned on enable genomes to generate the great diversity of animal forms from very similar sets of genes

At first glance, the list of animals could suggest any zoo. There's an elephant, an armadillo, an opossum, a dolphin, a sloth, a hedgehog, big and small bats, a couple of shrews, some fish, a macaque, an orangutan, a chimpanzee and a gorilla—to name a few of the more familiar creatures. But this menagerie is not at all like any zoo that has been constructed before. There are no cages, no concession stands and, in fact, no animals. It is a "virtual" zoo that contains only the DNA sequences of those animals—the hundreds of millions to billions of letters of DNA code that make up the genetic recipe for each species.

The most excited visitors to this new molecular zoo are evolutionary biologists, because within it lies a massive and detailed record of evolution. For many decades, scientists have longed to understand how the great diversity of species has arisen. We have known for half a century that changes in physical traits, from body color to brain size, stem from changes in DNA. Determining precisely what changes to the vast expanse of DNA sequences are responsible for giving animals their unique appearance was out of reach until recently, however.

Biologists are now deciphering the DNA record to locate the instructions that make assorted species of flies, fish or finches look different from one another and that make us humans dif-

Regulating Evolution

By Sean B. Carroll, Benjamin Prud'homme and Nicolas Gompel

ferent from chimpanzees. This quest has led to a profound change in our perspective. For most of the past 40 years or so, researchers have focused most of their attention on genes—the nucleotide sequences in DNA that encode the amino acid chains that form proteins. But to our surprise, it has turned out that differences in appearance are deceiving: very different animals have very similar sets of genes. By following the trail of evolution, devices are being found within DNA—genetic "switches"—that do not encode any proteins but that regulate *when* and *where* genes are used. Changes in these switches are crucial to the evolution of anatomy and provide new insights into how the seemingly endless forms of the animal kingdom have evolved.

Anatomy Genes and the Coding Paradox

For a long time, scientists certainly expected the anatomical differences among animals to be reflected in clear differences among the contents of their genomes. When we compare mammalian genomes such as those of the mouse, rat, dog, human and chimpanzee, however, we see that their respective gene catalogues are remarkably similar. The approximate number of genes in each animal's genome (about 20,000 or so) and the relative positions of many genes have been fairly well maintained over 100 million years of

evolution. That is not to say there are no differences in gene number and location. But at first glance, nothing in these gene inventories shouts out "mouse" or "dog" or "human." When comparing mouse and human genomes, for example, biologists are able to identify a mouse counterpart for at least 99 percent of all our genes.

In other words, we humans do not, as some once assumed, have more genes than our pets, pests, livestock or even a puffer fish. Disappointing, perhaps, but we'll have to get over it.

When biologists look at individual genes in detail, similarity among species is also the rule. The DNA sequences of any two versions of a gene, as well as the proteins they encode, are generally alike to a degree that simply reflects the relative amount of time that has elapsed since the two species diverged from a common ancestor. This preservation of coding sequences over evolutionary time is especially puzzling when one considers the genes involved in body building and body patterning.

Only a small fraction of all genes—fewer than 10 percent—are devoted to the construction and patterning of animal bodies during their development from fertilized egg to adult. The rest are involved in the everyday tasks of cells within various organs and tissues. Anatomical differences among animals—differences in the number, size, shape or color of body parts—must

KEY CONCEPTS

- Because genes encode instructions for building animal bodies, biologists once expected to find significant genetic differences among animals, reflecting their great diversity of forms. Instead very dissimilar animals have turned out to have very similar genes.

- Mutations in DNA "switches" that control body-shaping genes, rather than in the genes themselves, have been a significant source of evolving differences among animals.

- If humans want to understand what distinguishes animals, including ourselves, from one another, we have to look beyond genes.

—*The Editors*

DON FARRALL *Getty Images*

somehow involve the genes for body building. Indeed, the study of the pivotal role played in evolution by genes and processes associated with the development of anatomy has even earned its own nickname: evo-devo. For specialists, like ourselves, in that area of research, the discovery that body-building proteins are even more alike on average than other proteins was especially intriguing because of the paradox it seemed to pose: animals as different as a mouse and an elephant are shaped by a common set of very similar, functionally indistinguishable body-building proteins. The same applies to humans and our closest living relatives—most of our proteins differ from those of the chimpanzee by only one or two of the several hundred amino acids that comprise each protein, and 29 percent of our proteins are exactly identical in sequence. How do we explain this disparity between evolution at the two levels of proteins and anatomy? Somewhere in all of that genomic DNA there must be meaningful differences that have evolved. The trick is to find them, and the trick to doing that has been deciding where to look. It turns out that those places are much harder to locate than are genes themselves.

Genetic Switches

In humans, the protein-coding stretches of DNA make up only about 1.5 percent of our genome, so genes are really like little islands of information in a vast sea of DNA sequence. Much of the remaining noncoding DNA does nothing that we know of, but some of those sequences participate in the very important task of regulating gene expression. And these regulatory sequences are key to evolution.

The expression of a gene entails the transcription of the DNA sequence into a messenger RNA (mRNA) version and the translation of that mRNA into a protein sequence. The expression of most genes is regulated at the transcriptional level—cells do not waste energy making mRNAs and proteins they do not need. Many genes are therefore expressed only in an organ-, tissue- or cell type–specific manner. Certain noncoding DNA sequences play a critical part in directing when and where that happens. They are components of "genetic switches" that turn genes on or off at the right time and place in the body. Sequence-specific DNA-binding proteins called transcription factors, which are the other components of the switch, recognize those DNA sequences, often called enhancers. The binding of the transcription factors to the enhancer within

a cell nucleus determines whether the switch and the gene are on or off in that cell.

Every gene has at least one enhancer. Unlike the genes themselves, whose coding regions are readily identified because of the genetic code's fairly simple grammar, enhancers cannot be recognized solely on the basis of their DNA sequences and must be identified experimentally. Enhancers are usually hundreds of base pairs in length and may be located on either side of a gene or even within a noncoding stretch inside a gene. They can also be thousands of nucleotides away from the gene itself.

Most important to our discussion here is the fact that some genes have many separate enhancers. This is particularly true for genes that encode proteins that shape anatomy. Each enhancer independently regulates the expression of the gene in different parts of the body and at different times in the animal's life cycle, such that the complete expression of a gene is a patchwork of multiple, independently controlled sites of expression. These enhancers enable the same gene to be used many times in different contexts and thus greatly expand the functional versatility of individual genes.
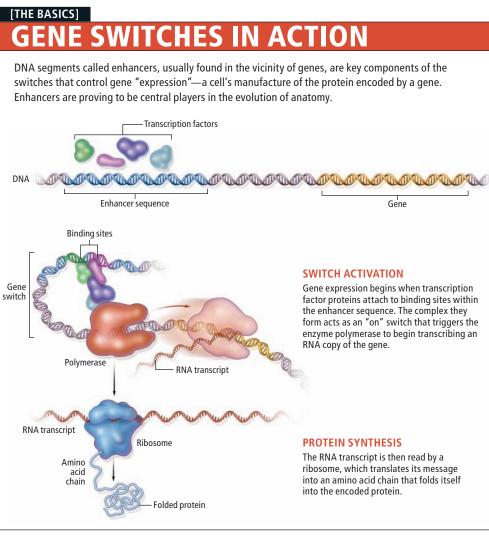
A gene involved in coloring the body parts of the fruit fly illustrates the modular logic of this gene regulation system. The somewhat confusingly named *Yellow* gene encodes a protein that promotes the formation of black pigmentation (mutant flies without this protein are yellow). The *Yellow* gene has separate enhancers that activate it during the development of a variety of fly body parts, including the wings and abdomen.

Because the *Yellow* gene plays a role during the development of so many tissues, mutations in the gene itself could be disastrous if they alter or disable the function of the protein; they would affect the function of the Yellow pigmentation protein throughout the organism. In contrast, changes in just one of the gene's enhancers will affect only the function of that enhancer and only the *Yellow* gene expression governed by that enhancer, leaving the expression and function of the protein in other tissues unchanged.

The evolutionary implications of the modular regulation of body-patterning genes are profound. In theory, mutations in enhancers would allow individual body traits to be selectively modified without changing genes or proteins themselves. And in the past few years, direct evidence has emerged that this is frequently how the evolution of various body parts and patterns has occurred.

Study of the pivotal role played in evolution by genes and processes associated with the development of anatomy has earned its own nickname: evo-devo.

GENE SWITCHES IN ACTION

DNA segments called enhancers, usually found in the vicinity of genes, are key components of the switches that control gene "expression"—a cell's manufacture of the protein encoded by a gene. Enhancers are proving to be central players in the evolution of anatomy.

Transcription factors

DNA

Enhancer sequence

Gene

Binding sites

Gene switch

Polymerase

RNA transcript

SWITCH ACTIVATION

Gene expression begins when transcription factor proteins attach to binding sites within the enhancer sequence. The complex they form acts as an "on" switch that triggers the enzyme polymerase to begin transcribing an RNA copy of the gene.

RNA transcript

Ribosome

Amino acid chain

Folded protein

PROTEIN SYNTHESIS

The RNA transcript is then read by a ribosome, which translates its message into an amino acid chain that folds itself into the encoded protein.

DETECTING A SWITCH

To understand when and where an enhancer regulates a gene's expression, scientists engineer a DNA fragment containing the enhancer sequence and a reporter gene that will produce a visible signal when it is active. After this DNA construct is injected into a single-celled embryo, it will integrate into the animal's genome and be present in every cell of the developing body. The reporter gene's activation reveals the enhancer's role in body-building processes during development.

Engineered DNA fragment

Enhancer sequence

Reporter gene

Single-celled embryo

Developing embryo

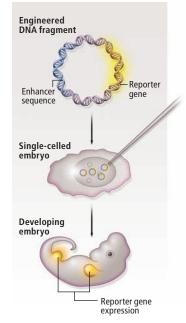Reporter gene expression

Evolving Switches

One of the most important strategies in biology is to identify the simplest experimental models of the phenomenon one wishes to understand. With respect to the evolution of body pattern, coloration fits the bill. It is one of the most obvious features of animals and plays a major role in how animals interact with their environment and with one another. Body-color patterns in fruit flies have diversified rapidly among closely related species, and analyses of how fruit flies got their spots and stripes illustrate how and why the evolution of genetic switches shapes the evolution of anatomy.

In some species, the males have intense black spots on the edges of their wings, whereas other species lack these spots. In some of these same species, males have a very dark abdomen (which is how the most famous fruit fly, *Drosophila melanogaster,* got its name: *melanogaster* means "black belly"), whereas males of other species lack this black band. In wing-spotted species,

the male displays his spots to the female as he courts her with a dance. We have found that in spotted species, the Yellow protein is produced at very high levels in the cells that will make the spot and at low levels in the rest of the wing cells. In unspotted species, Yellow is made only at low levels throughout the wing, generating just a light dusting of black pigment.

To figure out how Yellow is produced in a wing spot in some species and not others, we searched the DNA sequences around the *Yellow* gene for the enhancers that control its expression in various body parts. In unspotted species, there is an enhancer that drives *Yellow* expression in a low uniform pattern in the wing. This wing-enhancer activity generates the fly wings' light-gray color. When the corresponding piece of DNA was analyzed from a spotted species, we found that it drives both this low-level pattern and the intense spot pattern of gene expression. What has happened in the course of evolution of spotted species is that new binding sites for tran-

TAMI TOLPA

MODULAR SPOTS AND STRIPES

When multiple enhancers control the expression of a gene in different parts of the body, a change to one enhancer can alter the gene's activity in a specific place without affecting it elsewhere. A fruit fly gene called *Yellow,* for example, produces black pigment in a fly's developing body and wings, but various species have evolved distinct pigmentation patterns through changes to their enhancer sequences.

ANCESTRAL PATTERN

In a species representing an ancestral version of fruit flies, the enhancer that controls *Yellow* activity in the wings drives low gene expression, yielding a light-gray coloring, but in the abdomen a different enhancer drives high gene expression, producing a dark-black band.

FEATURE GAIN

Some species have evolved black wing spots by gaining a new transcription factor binding site in the wing enhancer sequence, which drives high expression of the *Yellow* gene in certain cells during wing development.

FEATURE LOSS

Other species have lost the abdominal band by losing a binding site in the corresponding enhancer sequence.

WING ENHANCER   ABDOMEN ENHANCER

Transcription factor

Black abdominal band

Yellow gene

Binding sites

Wing spot

New binding site

Lost binding site

scription factors made in the wing evolved in the *Yellow* wing-enhancer DNA sequence. These changes created an expression pattern—wing spots—without altering where the Yellow protein is made or how it functions elsewhere in the body [*see box above*].

A similar story applies to the evolution of the black band in the abdomen, but with a twist. Whereas we are naturally inclined to think that the presence of a feature in one species and its absence in another related species is the result of a gain by the first, that is often not the case. A flip side to evolution, the loss of features, is very common, though much less appreciated. The loss of body characters perhaps best illustrates why the evolution of enhancers is the more likely path for the evolution of anatomy.

One enhancer of the *Yellow* gene governs its expression in the abdomen. In males of species with the black band, this enhancer drives the expression of the *Yellow* gene at high levels in cells at the rear of the abdomen. But some species, such as *Drosophila kikkawai*, lost this band of pigmentation in the course of evolution. In *D. kikkawai*, the enhancer can no longer drive high levels of *Yellow* expression in the rear of the abdomen because a few mutations have disrupted some of its transcription factor binding sites.
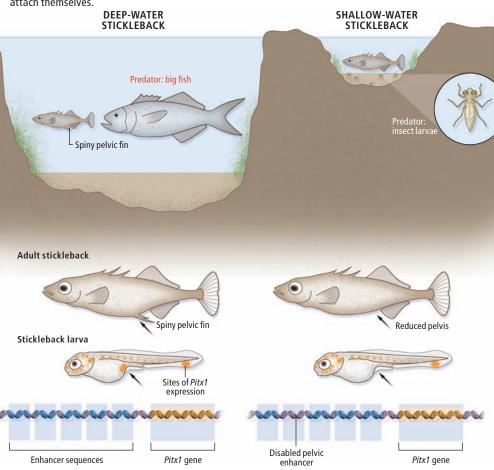
It is important to emphasize that the *Yellow* gene remains active elsewhere in the body and that its biochemical function is intact. Although one path to losing the black band could have been through mutations that inactivate the *Yellow* gene and its protein, this path is not permitted by natural selection, because the loss of Yellow function elsewhere in the body would have additional, detrimental consequences.

Losses of features may or may not be beneficial for survival or greater reproductive success, but some losses are adaptive because they facilitate some change in lifestyle. Hind limbs, for example, have been lost many times in verte-

brates—by snakes, lizards, whales and manatees—and those losses are associated with adaptation to different habitats and means of locomotion. The evolutionary forerunners of the hind limbs of four-legged vertebrates are the pelvic fins of fish. Dramatic differences in pelvic fin anatomy have also evolved in closely related fish populations. The three-spined stickleback fish occurs in two forms in many lakes in North America—an open-water form that has a full spiny pelvis, and a shallow-water, bottom-dwelling form with a dramatically reduced pelvis and

shrunken spines. In open water, the long spines help to protect the fish from being swallowed by larger predators. But on the lake bottom, those spines are a liability because dragonfly larvae that feed on the young fish can grasp them.

The differences in pelvic morphology among these fish have evolved repeatedly in just 10,000 years since the last Ice Age. Long-spined oceanic sticklebacks colonized many separate lakes, and the reduced form evolved independently several times. Because the fish are so closely related and interbreed in the laboratory, geneticists can

Mutations in regulatory sequences are not the exclusive mode of evolution— they are just the more likely path when a gene has multiple roles and only one of them is modified.

[CASE STUDY]

A BENEFICIAL LOSS

The three-spined stickleback offers another vivid example of adaptation through the evolution of a gene-regulating enhancer sequence. These fish take one of two forms, depending on where they live and there-fore which predator threatens them most: deep-water sticklebacks have a prominent spiny pelvic fin on their underside, which makes them more difficult for larger fish to swallow; shallow-water sticklebacks have lost the pelvic fin, making it harder for bottom-dwelling insect larvae that feed on the young fish to attach themselves.

DEEP-WATER STICKLEBACK

SHALLOW-WATER STICKLEBACK

Predator: big fish

Spiny pelvic fin

Predator: insect larvae

Adult stickleback

Spiny pelvic fin

Reduced pelvis

Stickleback larva

Sites of *Pitx1* expression

Enhancer sequences *Pitx1* gene

Disabled pelvic enhancer *Pitx1* gene

The gene that regulates pelvic-fin development, called *Pitx1,* also contributes to the development of many other major structures in the fish, each of which is regulated by a separate enhancer. In the shallow-water sticklebacks (*right*), a mutation has disabled only the enhancer responsible for pelvic-fin development, leaving *Pitx1* and its other functions intact.

TAMI TOLPA

[THE AUTHORS]

Sean B. Carroll, Benjamin Prud'homme and **Nicolas Gompel** have worked together for several years to decipher how the evolution of regulatory DNA sequences shapes animal morphology. Carroll is a Howard Hughes Medical Institute investigator and professor of molecular biology and genetics at the University of Wisconsin–Madison, as well as the author of two popular books about evolution. Prud'homme and Gompel, both former postdoctoral fellows in Carroll's laboratory, now investigate the evolution of animal forms and behavior in their own laboratory in France, at the Developmental Biology Institute of Marseille Luminy.

map the genes involved in the reduction of the stickleback pelvis. David M. Kingsley of Stanford University, along with Dolph Schluter of the University of British Columbia and colleagues, has shown that changes in the expression of a gene involved in the building of the pelvic skeleton are associated with the pelvic reduction. Like most other body-building genes, the *Pitx1* gene has multiple jobs in the development of the fish. But its expression is selectively lost in the area of the fish that will give rise to the pelvic-fin bud and spines. Once again, evolutionary changes in an enhancer are responsible. There are no coding changes in the Pitx1 protein between different forms of the stickleback.

Yellow, Pitx1 and most other body-building and body-patterning genes are said to be pleiotropic, in that they influence the formation or appearance of multiple traits. Mutations in the coding sequence of a pleiotropic gene have multiple effects on all the traits controlled by this gene, and that drastic amount of change is unlikely to be tolerated by natural selection. The key lesson from the evolution of spots, stripes and skeletons is that mutations in regulatory sequences circumvent the pleiotropic effects of mutations in coding sequences and allow for the selective modification of individual body parts. Mutations in regulatory sequences are not the

exclusive mode of evolution—they are just the more likely path when a gene has multiple roles and one of those roles is selectively modified.
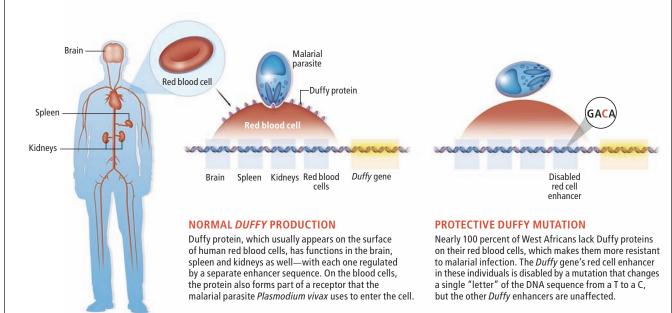
Common Genes, Endless Variety
The evolution of enhancers is not at all limited to genes affecting body form nor just to fruit flies and weird fish. Quite a few examples of evolutionary changes in regulatory sequences that alter gene expression have been demonstrated for human traits as well.

One of the more striking cases in recent human evolution represents an adaptation, through selective loss of gene expression, to an environment where malaria is endemic. In addition to the familiar A, B and O blood types, other so-called minor blood types have been well studied. The status of a protein called Duffy, present on the surface of red blood cells, defines one of these types. The Duffy protein constitutes part of the receptor that is used by a malaria-causing parasite, *Plasmodium vivax,* to infect red cells, but in West Africa the protein is absent from the blood cells of almost 100 percent of the population, making individuals resistant to infection. The *Duffy* gene is also expressed in several other body tissues, including cells of the spleen, the kidneys and the brain. In the African population, *Duffy* expression in

[CASE STUDY]

HUMAN DIVERSITY

The human genome, like those of flies and fish, also displays evidence of evolution through changes to enhancer DNA. One example is the adaptive loss of a protein called Duffy on red blood cells in a West African population living in areas where malaria is endemic.

NORMAL *DUFFY* PRODUCTION
Duffy protein, which usually appears on the surface of human red blood cells, has functions in the brain, spleen and kidneys as well—with each one regulated by a separate enhancer sequence. On the blood cells, the protein also forms part of a receptor that the malarial parasite *Plasmodium vivax* uses to enter the cell.

PROTECTIVE DUFFY MUTATION
Nearly 100 percent of West Africans lack Duffy proteins on their red blood cells, which makes them more resistant to malarial infection. The *Duffy* gene's red cell enhancer in these individuals is disabled by a mutation that changes a single "letter" of the DNA sequence from a T to a C, but the other *Duffy* enhancers are unaffected.

Scanning for Switches

One of the main limits on the pace of discovery of human enhancers has been the difficulty of identifying where they reside in the human genome's vast noncoding regions. Biologists are now using the preservative power of natural selection to sniff out stretches of noncoding DNA that have been unusually well conserved over long stretches of evolutionary time in the hope of detecting enhancers.

In this article we have been emphasizing changes in enhancers that account for differences among organisms. But it should be easy to appreciate that some enhancers carry out functions that have not changed. While the steady pace of mutation erodes the overall similarity of DNA sequences among species as they diverge, natural selection will maintain the sequences of enhancers that maintain their function, sometimes to an extraordinary degree.

It is common knowledge that lawyers and sharks have a lot of similarities. But who would have guessed that extends to the level of DNA? Yet that is essentially what researchers at the Institute of Molecular and Cell Biology in Singapore and the J. Craig Venter Institute in Rockville, Md., have demonstrated. The team found that despite more than 500 million years of evolution separating sharks and people, we share nearly 5,000 elements in noncoding regions near genes that appear to be enhancers. Remarkably, most of these highly preserved elements are located in the vicinity of body-building genes, reflecting the shared overall body architecture of vertebrates.

Every vertebrate has anatomical features—organs, tissues, cell types, and so forth—that have been preserved throughout their diversification. Over shorter evolutionary distances, the number of shared elements and degree of similarity increases.

The genome-comparison approach is thus rapidly expanding the catalogue of known human, mammalian and vertebrate enhancers and could lead to the identification of sequences involved in the divergence of body forms.

—*S.B.C., B.P. and N.G.*

those other tissues is preserved. Not surprisingly, these Duffy-negative individuals carry a mutation in an enhancer of the *Duffy* gene that knocks out the binding site for a transcription factor that activates *Duffy* expression in red blood cell precursors but that has no effect on Duffy production elsewhere in the body.

Gregory A. Wray of Duke University and his collaborators have identified other aspects of human biology that have evolved through mutations in enhancers in different human genes. One of the most intriguing associations revealed thus far involves divergence in the great ape and human regulatory sequences controlling the *Prodynorphin* gene, which encodes a set of small opioid proteins produced in the brain and involved in perception, behavior and memory. The human gene is more highly expressed in response to stimuli than is the chimpanzee version, and strong evidence suggests that the human regulatory sequence evolved under natural selection—that is, it was retained because it was advantageous.

As these examples illustrate, mutations in regulatory DNA have undoubtedly played a role in human evolution and regulatory variation may be an important source of physical and health differences among individuals as well. Because scientists cannot tinker with the DNA of living humans the way we can with flies and fish, it is somewhat harder to study certain examples of regulatory DNA changes responsible for our divergence from other species, although some new methods for analyzing genomes are producing encouraging leads [*see box above*].

These are still early days for research into the evolution of gene-regulating DNA sequences. And hundreds of thousands of genetic switches in the virtual zoo of genomes have yet to be discovered or investigated. Biologists are already learning new principles, however, that have predictive value for future studies: evolutionary changes to anatomy, particularly those involving pleiotropic genes, are more likely to happen via changes to gene enhancers than to the genes themselves.

This phenomenon also reveals how very diverse groups of animals can share most, if not all, the genes involved in body building and body patterning—contrary to scientists' early expectations; it is mostly a matter of how and when those genes are used that shapes the different forms of the animal kingdom. If we really want to understand what makes the human form different from that of other apes or what makes an elephant distinct from a mouse, for that matter, much of that information lies not in our respective genes and proteins but in an entirely different realm of our genomes that remains to be explored. ∎

➥ **MORE TO EXPLORE**

Evolution at Two Levels: On Genes and Form. Sean B. Carroll in *PLoS Biology,* Vol. 3, Issue 7, pages 1159–1166; July 2005.

Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom. Sean B. Carroll. W. W. Norton, 2005.

The Making of the Fittest: DNA and the Ultimate Forensic Record of Evolution. Sean B. Carroll. W. W. Norton, 2006.

The Evolutionary Significance of cis-Regulatory Mutations. Gregory A. Wray in *Nature Reviews Genetics,* Vol. 8, pages 206–216; March 2007.

Emerging Principles of Regulatory Evolution. Benjamin Prud'homme, Nicolas Gompel and Sean B. Carroll in *Proceedings of the National Academy of Sciences USA,* Vol. 104, Supplement 1, pages 8605–8612; May 15, 2007.

For links to teaching resources, see **www.seanbcarroll.com**