



PHYSICAL ANTHROPOLOGY
ORIGINAL READINGS IN METHOD AND PRACTICE

Edited by

Peter N. Peregrine
Lawrence University

Carol R. Ember
Human Relations Area Files

Melvin Ember
Human Relations Area Files



UPPER SADDLE RIVER, NEW JERSEY 07458

2002



GENES, BODIES, AND SPECIES

JONATHAN MARKS

This chapter is about the biological hierarchy through which changes in DNA ultimately become differences between species. The idea of *hierarchy* is an old and fairly nebulous one in evolutionary biology, generally invoked to differentiate between what geneticists study and what anatomists or natural historians study.¹ What geneticists study—genes and chromosomes—is of course so small that you can't see it with the naked eye, while what anatomists study can be sliced and weighed. Likewise, what geneticists study *causes* what anatomists study—that is, genes “code for” bodies.

And yet, such a biological hierarchy is certainly intended to convey something more than simply gross size of the subject matter—big things versus little things! After all, non-genetic information has a major role in the form of the end products of biological development. And genes only “code for” bodies in a very crude way, for bodies also constitute the manner by which the genes are passed into the next generation—so there is an important sense in which bodies make genes as much as genes make bodies.

The object of this chapter is to link together the small and the large through modern evolutionary theory, and to explore the relationships among genes (or cells), phenotypes (or bodies), and populations (or species). The point will be that cells, organisms, and species are all “individuals”—that is to say, singular entities bounded in space and time and composed of interacting parts; as opposed to “classes,” or sets of objects, transcending time and space and composed of members sharing a specific quality or essence.² Each of these kinds of individuals can be studied in a particular fashion, and we will explore the manner in which that is so, and the different kinds of evolutionary knowledge that is accessible from each kind of individual.

GENETICS: CELLS AS INDIVIDUALS AND GENES AS MACHINE CODE

Cells were recognized in the nineteenth century as the most fundamental form of life. Reproduction and development, whose relationship to one another had previously been very unclear, were now seen to be related through the processes of cell division. While the body at all stages was composed of cells, reproduction and development were the results of two different kinds of cell division. Mitosis was the process by which cells divided into two identical cells, and the body thereby grew; and meiosis was the process by which sperm and egg cells were produced to transmit genetic information into the next generation of organisms. Each sperm or egg has half the amount of genetic information of ordinary cells, and the genetic information they carry differs. Thus, ordinary cells have two copies of each gene and are thus *diploid*, while sperm and egg have one copy of each gene and are thus *haploid*. In addition, ordinary cells all carry a set of identical genetic information, while the genetic information carried within each sperm or egg cell is unique and thus distinct from every other sperm or egg.

Cell generations, then, are similar to those of larger organisms. Cells originate, persevere, and die. They exist in space and have a duration in time. They reproduce. They interact with others like themselves and with the physiological environment in which they exist. They are autonomous biological units.

Another powerful recognition came half a century later, as biologists grappled with the precise nature of the genetic information itself. In a book called *What is Life?* published in 1944, physicist Erwin Schrödinger suggested that the genetic information could be considered as a kind of code that the cell's mechanism effectively decodes. This idea of the *genetic code*, evoking powerful images of language and machinery, has subsequently become one of the most powerful metaphors in science.³

The units of genetic information themselves came to be known early in the twentieth century as "genes."⁴ The development of population genetics in the 1930s provided a powerful theoretical model for the processes of evolution, based on a mathematical abstraction, the *gene pool*—a hypothetical summation of all the hereditary elements in a population—a pool in the sense of office pool, not swimming pool.

By focusing on the gene pool, population genetics established a fundamental way of thinking about evolution, a way that completely avoided the recognized biological units: cells and organisms. Population genetics modeled evolution as changes in the gene pool, in which organisms were simply represented by genotypes, transient pairs of genes, each of whose lifespan was far shorter than that of the gene pool itself. The proportions of each genotype were mathematically predictable from the proportions of each gene by a relation that has come to be known as the Hardy-Weinberg Law.⁵

The Hardy-Weinberg Law has two parts. The first part tells us that, given two genetic variants (alleles) in a gene pool, one with a frequency of p and the other with a frequency of q , the ordinary processes of organismal reproduction will sort these alleles into diploid genotypes in fixed proportions. One homozygote (an organism with two identical alleles) will exist with a frequency of p^2 , the other with a frequency of q^2 , and the heterozygote (an organism with one of each allele) will exist in the population with a predictable frequency of $2pq$. In other words, if 30 percent of the alleles in a gene pool are A1, and 70 percent are A2 (A1 and A2 being two hypothetical variants of the A gene), then we would expect to find 9 percent of the *organisms* to be A1A1 homozygotes, 49 percent to be A2A2 homozygotes, and 42 percent to be A1A2 heterozygotes. The second part of the Hardy-Weinberg Law tells us that these proportions remain constant every generation, as long as no other forces are acting on the population aside from reproduction.

Evolution, then, according to population genetics, is simply a violation of this genetic equilibrium. Studying evolution involves asking a single fundamental question: What causes the proportions of genotypes and genes *not* to remain constant?

Empirically this question can be studied using genetic markers, bits of genetic material whose transmission can be tracked in populations, even though they may not be functionally significant. After all, this is a corpus of statistical theory that eliminates phenotypes, bodies, and organisms from the picture. The function of a gene—what it does, and its effect on the body—is thus irrelevant here.

So what alters the gene pool? Preferential mating with relatives (*inbreeding*) makes a population more homozygous, but doesn't directly affect allele frequencies. Differential reproduction of particular genotypes (*selection*) will directly affect allele frequencies; as will the introgression of genetic material from another population (*gene flow*). Finally, since real populations are finite in size, the laws of chance dictate random deviations from the predicted mathematical constancy (*genetic drift*). Each of these forces has specific effects: Since the environment is what permits different genotypes to reproduce more or less efficiently than their alternatives, selection permits populations to track their environments genetically; thus selection makes populations different from one another adaptively. Gene flow, on the other hand, makes populations more similar to one another, for it reflects the genetic contact between populations. And finally, genetic drift also makes populations different from one another, but in a non-adaptive way.

Thus, "the spread of genes" is a classic way of visualizing evolution, assuming a direct translation from genetical variant to physical/anatomical variant, and modeling the ways in which a population is altered through time upon the emergence of these new genetic variations. But recently, the precise nature of these genetic variations has come to be examined in greater depth, and has revealed a system of far greater complexity than such simple modeling had considered.

THE GENOME AS INFORMATION

The image of genes as blueprints or machine code for the body is immensely powerful. It lies behind the program begun in the 1980s called The Human Genome Project, in which great public and private resources were mustered to generate a comprehensive vision of the blueprints: the DNA sequence of each chromosome. A genome is, formally, the genetic structure of a single reproductive cell: a haploid genetic complement, one of each chromosome.

The chromosomes are visible structures whose purpose seems to be to guide a great mass of DNA through cell division by condensing it into a manageable number of regular structures. Thus, chromosomes are only visible during cell division; at other times the DNA is loose in the cell nucleus, although "anchored" to the nuclear membrane at certain points.

Humans have twenty-three pairs of chromosomes, so the Human Genome Project's goal was to reproduce the detailed linear sequence of each of them. Once we have the gene sequences, and the knowledge that the gene sequence is like machine code for the production of organisms, we will have, as the Human Genome Project's purple prose said, "The Book of Man" or "The Holy Grail of Biology."^{6,7}

This misses a crucial part of genetic physiology, however. As we noted earlier, a normal human body cell is diploid; only gametes are haploid. In fact, the human body is constrained to have two, *and only two*, sets of instructions in its cells. While many domestic plants have multiple sets of chromosomes, a condition known as polyploidy, mammalian cells simply can't work that way. A human with three sets of chromosomes (sixty-nine total) is triploid and cannot survive.

Moreover, having one copy (monosomy) or three copies (trisomy) of any specific chromosome is also bad. This is known as the problem of dosage: You need two doses of DNA, not one and not three. Having one copy of a particular chromosome (making forty-five total) is not compatible with life, unless that chromosome is the X chromosome, which still results in a characteristic pathological condition known as Turner's Syndrome. Having three copies of a particular chromosome (forty-seven total) is slightly better, but not much. One can live normally with an extra Y chromosome (because there is very little on that chromosome): One can also survive, but again with distinctive pathological phenotypes, with an extra X chromosome (Klinefelter's Syndrome) or chromosome 21 (Down's Syndrome). The X chromosome can tolerate variation in number more readily than other chromosomes because it has relatively few genes and has its own special mechanism for regulating dosage—since normal males have one X and normal females have two Xs. Chromosome 21 seems to be able to tolerate trisomy because it is so small and has only about 200 genes, rather fewer than chromosome 22, which is about the same size.

The point is that normal development depends not just on the DNA sequence, but specifically on the interactions of *two* DNA sequences. Using the

metaphor of DNA as a blueprint or code, it may be easy to miss the importance of the dosage of DNA required; the complex interactions between two *sets* of instructions that seem to add up to more than the sum of their parts.

They say that when the only tool you have is a hammer, everything tends to look like a nail. Unfortunately we don't really have tools for understanding the interaction between two genomes that results in a normal, healthy organism. But we do have the ability to study the structure of an individual genome in its finest detail, and so we do. From the standpoint of evolution, this leads us to the process of mutation. Mutation is a change in the genetic material; alleles differ from one another ultimately because of mutation, and consequently mutation is the source of all genetic differences between organisms and between species. Since the 1980s, our increasing knowledge of the genome has greatly affected our view of mutation.

The genome is composed of DNA, a long series of elementary subunits known as adenine, guanine, cytosine, and thymine, abbreviated A, G, C, and T. These are known as nucleotides or bases. When we speak of a "DNA sequence" we mean simply a long, specific arrangement of those four bases: for example, AAGCTATATCCAGCA.

The human genome is essentially 3.2 billion of those letters, divided among twenty-three chromosomes. Genes are simply particular regions of the genome that have some kind of function. Here, however, is where a major conceptual revolution in the past quarter-century has occurred. Where it used to be thought that the genes were arrayed like beads on a string—the genome being the string and the genes being the beads—it is now clear that (1) the beads are simply "special cases" of string; and (2) there are very few beads.⁸ In other words, the genes are simply functional bits of genome; and genes are rare within the genome. The human genome sequence, published in early 2001, shows that less than 25 percent of the genome consists of genes (30,000 genes, each on average 27,000 bases long), and even within each gene less than 10 percent is actually functional. The Human Genome Project reports that only between 1–2 percent of the DNA actually comprises *coding sequence*.⁹

Moreover, there are complex patterns of redundancy within the genome. These patterns of redundancy are themselves products of the mutational machinery, and contribute to the perpetuation of that redundancy. For example, the most basic kind of mutation is the substitution of one base for another, such as AAGCT to AAGCC. However, a different mutational process called *strand slippage* can insert or delete additional bases, such as altering AAGAGGCT to AAGAGAGAGAGGCT. That of, course, creates a fundamental pattern of redundancy of bases. It may also increase in intensity over the generations, as an auto-catalytic process. This has now been found to be the cause of a significant class of genetic diseases, which includes Huntington's chorea and fragile-X syndrome.

Another kind of redundancy is created by a widespread "rubber-stamping" process in the genome, whereby one DNA sequence is simply copied

next to itself. If this DNA sequence includes a functional gene, it will produce two functional genes next to each other, where formerly there was only one. Three things can happen to this new copy over many generations: (1) It can continue to do the same thing, so the body now has twice as much gene product, which, if beneficial, would then be preserved; (2) if its presence makes no difference to the body, it can degrade by mutation to a state of nonfunction, now called a *pseudogene*; (3) rather than becoming nonfunctional, mutations can alter the properties of the second gene product so that it does something different, which again might be beneficial.

And indeed we find that genes in the genome are found in clusters, some copies of which do identical things, some of which do slightly different things, and some of which do nothing at all. The genes coding for hemoglobin, the best-known genetic system, are located in two clusters, one at the tip of chromosome 11 and the other at the tip of chromosome 16. The one at the tip of chromosome 16 codes for the alpha component of hemoglobin, a protein that is 141 amino acids long, and has two identical genes, side-by-side, churning out the raw material for alpha-globin. There is also, however, a gene just a little ways away, which also makes a 141-amino-acid-long protein, but only does so early in embryonic life, when the needs for gas transport in the tissues are considerably different. And there are also DNA stretches that bear strong similarities to the functional genes but do not themselves do anything: pseudogenes, the result of archaic duplications of DNA that didn't help and didn't hurt, a genetic experiment that neither succeeded nor failed, but whose record remains.

Other portions of the genome appear to be immensely long stretches of simple sequences—just a few letters repeated millions of times, the result of a process (or processes) of massive tandem duplication.

Still other kinds of mutation involve the movement of specific DNA bits from one place to another, known as transposition, or the creation of several copies of a particular DNA bit, and the integration of these DNA copies in many places throughout the genome, known as retrotransposition. The most famous of these DNA segments is known as *Alu*, and consists of a fairly specific stretch of DNA about 300 bases long and intercalated seemingly at random through the genome, over a million times.

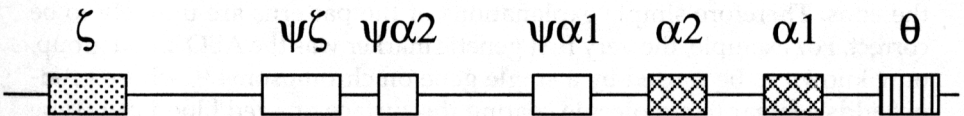


FIGURE 2-1

Schematic diagram of the alpha-globin gene cluster at the tip of human chromosome 16. From left to right, zeta is a functional gene in embryonic life; three pseudogenes are inactive; the two alphas contain identical coding sequences and are functional throughout the entire lifespan; and theta is of unknown function. The entire region covers about 30,000 bases.

Since the genome is far more complex than had been thought by an earlier generation, our ideas about the genome have had to be revised, particularly in relation to evolution. We have tended to conceptualize mutation as a simple process of nucleotide substitution in functional genes in our formal models, but it is now clear that the genome is a dynamic landscape, and can tolerate considerable change without apparent adverse effect. With mutation being the basic source of biological novelty, the implication is that the sources of novelty for evolution are considerably more diverse than considered by earlier generations of geneticists. How do we use this new information to understand the biological history of species?

REPRESENTATION

From the standpoint of evolution, these different parts of the genome can convey different kinds of information. It is axiomatic, for example, that since most of the genome is nonfunctional, it is not affected by natural selection. Thus, most mutations have no effect on the phenotype and are not weeded out by natural selection, while those that alter the structure of genes generally do so adversely and are consequently eliminated. Thus inter-genic DNA evolves most rapidly, and genic DNA most slowly.

However, in the time since, say, human and chimp have been separate species, very few changes have occurred in the genome. We are over 98 percent genetically identical to chimpanzees and (at least to a first approximation) 100 percent genetically identical to each other; consequently it is hard to use DNA to study the relationships among human beings. However, a particularly rapidly evolving part of the genome is found outside the cell nucleus and is known as mitochondrial DNA or mtDNA. Here, humans and chimpanzees are 9 percent different, and two random humans are 0.2 percent different from each other. Consequently this can be used as a genetic marker for studying the relationships among groups of people.

Genetic markers can be tricky, however. Most of the genome doesn't vary from person to person, and what does vary is often the product of complex mutational processes in the cell and complex population dynamics over the eons. Therefore simple explanations of the patterns are unlikely to be correct. For example, the very first genetic marker was the ABO blood group, now known to be caused by a single gene on chromosome 9, whose product adds a sugar to a molecule coating the surface of a red blood cell. Type A adds one sugar, type B adds a different sugar, and type O adds none. Simple enough.

The distribution of the three alleles, however, is far from simple. All populations of the world have all three of them but in different proportions, with O being universally the most common. Clustering human populations on the basis of the frequencies of these alleles puts very different peoples together

who share the same frequencies by accident, such the Poles and Chinese. Moreover, in spite of the fact that most people are type O, chimpanzees are overwhelmingly type A, and gorillas are overwhelmingly type B. It is simply not clear just what is going on, or how can we make easy evolutionary sense of it.

It is also important to bear in mind that these are genetic markers, and not genes *for* physical attributes, such as skin color or facial contour—which even today we have no access to.

By the 1970s it was evident that one could pool several different genetic markers and use high-powered statistics to determine which populations had the most similar clusters of allele frequencies, and might therefore be considered closest relatives. It turned out, however, that the results were always very unstable, and highly sensitive to just which genetic markers were used, just what computer algorithm was used, the demographic history of the population (whether expanding, contracting, or interbreeding), and even the specific people chosen to represent the population, and the specific populations chosen to represent the region.¹⁰

In one infamous anecdote, a geneticist drew blood from pygmies in central Africa but rejected anyone he thought was too tall; thus the test tubes he brought back from the field didn't reflect the real population, but rather his preconceptions of what the population should have been. Other studies generalized extravagantly about "Africans" from two or three local tribes. Thus, representation is a key issue: An accurate relationship between the samples the geneticist has and what or whom those samples are supposed to reflect cannot be taken for granted.

The problem of representation is crucial when we begin to appreciate that populations are composed of diverse organisms that may, and usually do, have complex histories. Looking, for example, at the relationships among human populations, one is always laboring under the shadow of gene flow. Two human populations may be genetically similar because they diverged recently, or because they have interbred. Human populations are also fluid and symbolically-defined: Marriage, adoption, raiding, alliance, and trade all serve to make the boundaries between them far less distinct genetically than culturally. Thus, if one wished to study the genetic relationships among the Germans, Swiss, and Italians, there may be no real scientific answer to such a scientific-sounding question. The answer you get will depend crucially on who is selected to represent those populations, given that Switzerland lies geographically in between Germany and Italy and has a long and complex social history in relation to them.

The important point is that all comparisons are not equal, nor are all evolutionary conclusions based on genetic data self-evident. Simply the choice of specimen, from which the DNA is ultimately isolated, may have a major effect on the conclusions of a study, if the hierarchical relationships between person, population, and cell are ignored. On the other hand, this problem might seem to be mitigated if you study the relations among different species.

Above the species level, gene pools can only diverge; there can be no inter-species breeding (that is generally what we mean by species, after all). Thus, the relations among humans, chimpanzees, and gorillas might at face value seem an easier question.

SPECIATION

Population genetics is a classically *transformationist* field, tracking a single gene pool as it changes through time. The diversity of life, however, is a product of divergence, as well as descent. How does one evolutionary lineage become two, so that divergence, presumably by a combination of genetic drift and natural selection, can occur?

The multiplication of lineages is the proliferation or reproduction of species, a series of processes known collectively as speciation. This can be considered analogous to the mitosis of cells or the reproduction of (asexual) organisms—the generation of two individuals or biological units where formerly there was but one.

Speciation requires the division of the gene pool, or more precisely it requires the segregation of organisms into populations that are not in genetic contact with one another. The maintenance of genetic contact (gene flow) acts in opposition to speciation. Once the population is divided into two or more segments, they diverge through time genetically both at random (via genetic drift), and in conjunction with their new environments (natural selection). The key question, however, is how the organisms of those populations ultimately come to regard one another as *different*—that is to say, not as potential mates. Dogs, after all, can be exceedingly promiscuous with other dogs, even with dogs that look quite different, but one never sees them mate with cats.

What is it that permits an organism to recognize another as a potential mate, and thereby worthy of reproductive attention and effort, so that it doesn't squander its time, resources, and energy? How does a rhesus monkey know it's a rhesus monkey and not a pig? And more importantly, how does it recognize another rhesus monkey as a potential mate?

The processes of species formation involve not just genetic divergence, but the development of new specific mate recognition systems.¹¹ These may be visual, olfactory, or behavioral signals, but they comprise a crucial part of the evolutionary process. Often, one population becomes physiologically incapable of reproducing with another population before they recognize each other as being different species—they may be fertile at different times of the year, or recognize different mating signals, or simply have physically incompatible genitalia. Or it may be that the structure of the chromosomes of one population has changed, so that they may mate and hybridize, but the hybrid is infertile—precisely the situation that exists between horse and donkey.

Three sets of changes, all ultimately genetic, but not easily related to one another, occur together during speciation. The first are the clocklike genetic changes all through the genome, so that regardless of the function (if any) of a particular DNA segment, it can be used as a record of the biological history of the cell, organism, and species it came from. The second are the physical changes to the body, itself a responsive and reactive system. And the third are the physiological, reproductive signals that unite a reproductive community, and distinguish it from others. Species that have been separated for a long period of time are different in all three ways; but recent divergences, or even incomplete divergences, yield important insights into the manner by which molecular, anatomical, and reproductive divergences or incompatibilities are generated in populations separately, but in parallel.

Since these kinds of changes occur together, although at different rates, the relationships of closely related species are no more easy to disentangle with molecular data than they are with classical anatomical data. Returning to the question of chimpanzees, gorillas, and humans, we find that some molecular data appear to link chimps and people, some link chimps and gorillas, and most yield thorough ambiguity. We are consequently obliged to regard the relationships among chimps, gorillas, and humans as a three-way-split or trichotomy.¹²

A question arises when we study the relationships among groups of organisms, whether they are parts of one species, or of several: Is one representative sufficient? Can a single specimen stand as a synecdoche (a metaphor in which the part substitutes for the whole) of a biological group?

Actually we can show quite easily the inadequacy of such an assumption at the level of the species. Imagine a single ancestral species, broadly distributed geographically. Perhaps they are protochimpanzees in equatorial Africa. One group fissions off to the west, and over a few hundred thousand years they become a species of protogorillas. Another group fissions off to the east and over a similar span of time they become protohominids. To ask which pair of species a few million years later are closest relatives is a nonsensical question: The protochimpanzees would be the ancestors of all three living forms (gorillas, chimps, and humans), the early gene pools of the species would have overlapped extensively, and the relations of their modern descendants would be ambiguous.¹³ Indeed, the answer you obtained might depend very specifically on which specimens you chose to represent the species: The chimpanzees from one region might turn out to be slightly more closely related to protogorillas and those from another might be slightly more closely related to protohumans. After all, according to this simple model, they were ancestral to both.

And once again, we find that at places where genetic diversity has been sampled in the apes, it is generally considerably greater than anticipated (especially considering that the apes are all endangered, and living in small ranges in Africa), and also considerably larger than in humans, who have nevertheless expanded to fill up the entire planet.¹⁴

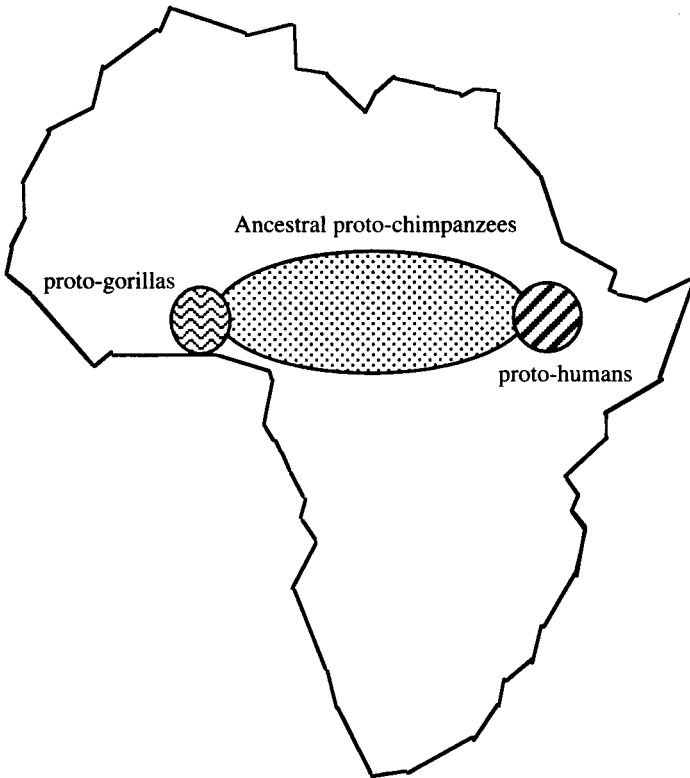


FIGURE 2-2

If an ancestral population of protochimpanzees gave rise to both protogorillas in the west and protohumans in the east at around the same time, the phylogenetic relationships among them would be vague from modern species in precisely the way we find them to be. And it wouldn't matter whether you looked at DNA or bodies—the problem would be modeling the biological history adequately.

So, just like the problems of understanding the relationships among human populations if the hierarchical relations of cell, body, and population are ignored, there exists a similar set of problems that accrue if one tries with comparable naivete to understand relationships among species.

SEEING EVOLUTION: THE EYE OF THE BEHOLDER

The processes of evolution to a molecular geneticist involve principally mutation and genetic drift. Random changes to the DNA occur at a calculable rate, and since most of the genome has no function, the changes are generally either unexpressed or neutral (functionally equivalent to their alternatives), and will thus be unaffected by selection. Even in functional regions there is

considerable “slop”—insulin derived from the pancreas of a pig, while structurally different, nevertheless still functions well for a human.

The action of selection here is principally to weed out genetic variations that function poorly. Thus, given the structure of a protein and its physiological necessities, randomly altering its structure (by mutation) is far more likely to compromise its function than to improve it. Imagine opening the hood of your car and randomly hitting your engine with a hammer (the equivalent of mutation). While it is theoretically conceivable that you might improve the engine’s performance, it is exceedingly unlikely. Most likely you’ll screw it up; and if you’re lucky you just won’t do any serious damage.

A geneticist therefore finds two DNA sequences from different species to differ in very specific ways. Because of the constant pressure of genetic mutation, differences will always be there, but you expect to find the most differences in functionless regions, and proportionately fewer in regions with important genes; and within each gene, you expect to find more differences in specific regions that compromise the structure and function of the protein product the least.

In other words, one expects to find difference when comparing DNA regions across species; when “too much” similarity is found, that is taken as evidence for strong constraints on the particular DNA region, indicative of functional importance. That is precisely how the *homeobox*—a stretch of 180 nucleotides in a small class of genes involved in early embryonic development—was discovered: It was there, almost intact, in the genomes of flies and mice. Since flies and mice have little else in common, this suggested an exceptional physiological importance for the region.

Morphologists see evolution, in the classical sense, quite differently. They anticipate stability of structure, because they work with adapted bodies, not with sloppy genes. Therefore they expect related organisms to be *similar* to one another, and attempt to explain the *differences* by recourse to selection. This is precisely the opposite of what the geneticists do, expecting divergence of DNA sequence and explaining the situation in which sequences are more similar than anticipated!

A geneticist and a morphologist look at the same animals: orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and humans (*Homo sapiens*). The geneticist finds that most of the genome differs by about 2 percent between humans and chimps, and by about 5 percent between either of them and the orangutan. But certain regions are identical: The geneticist thus frames questions such as, “Why haven’t mutation and drift caused this region to diverge more through time? Why is this region so important? What does it do that constrains it from tolerating any change?” The morphologist, by contrast, sees two creatures with hairy bodies, long arms, and small brains, and asks, “Why has the other species lost body hair, changed limb proportions, and expanded its brain?” The morphologist does *not* ask, “Why does the orangutan and chimpanzee *retain* body hair, long arms, and small brains?” The answer

is simple: Those features work. The morphologist expects preservation of functional body systems due to adaptation, and in contrast the geneticist expects decay of DNA sequences due to the constant pressure of mutation.

The geneticist and the morphologist thus see different patterns in their data, see change occurring at different rates, in different modes, and ask different evolutionary questions to explain their findings. While genes, bodies, and species all evolve together (obviously!), it is frequently difficult to unify them into a coherent evolutionary narrative. This is not altogether surprising: The big question in genetics is the complex relationship between genotypes (i.e., genetic constitutions or DNA sequences) and phenotypes (the outward appearances of organisms).

In the simplest case, early geneticists following Mendel found that one phenotype was dominant over another, but a pea heterozygous for a gene (Aa) looked indistinguishable from a pea homozygous for one of the genes (AA). Thus, from what the pea looks like, you cannot tell what alleles it has.

In more complex cases, a phenotype results from the interaction of several genes operating in physiological systems. Not only that, but the developing body is reactive and sensitive to the conditions of growth. Thus, someone may be short-statured because of a particular combination of alleles, or because of nutritional stress during childhood. The body is thus developmentally plastic: A given genetic background may result in different physical forms under different circumstances. The body is sensitive to the conditions of life; it is adaptable.

More than that, the body is also pulled in the opposite direction. Not only does the body change in harmony with the circumstances of growth, but in other ways the body is also very *insensitive* to genetic variation or environmental shock. The developmental geneticist C. H. Waddington called this property *canalization*—the tendency of the genetic system to be buffered, so that the same phenotype may result from different genetic backgrounds or in different environments.¹⁵

In other words, the same genotype can produce different phenotypes, and the same phenotype can be the product of different genotypes. With such a level of disconnect between genes and bodies, it's no wonder that molecular and morphological specialists see their subject matter in such different ways.

CONCLUSIONS

The ultimate source of all evolutionary difference is mutation, the change in DNA, located in cells. Mutation is a more complex process than earlier generations conceived it to be. The genome is a complex landscape, only being mapped now, and the study of *comparative genomics* is in its infancy.

The fundamental changes that occur initially in a cell ultimately accrue to the body the cell is a part of. Mutations that aid in the survival and reproduction of those bodies—or at least don't hurt them—are perpetuated, and thus become disproportionately represented in the gene pool. New species arise when the gene pool is partitioned, and organisms cease to identify each other as potential mates.

Cells, organisms, and species are the units of contemporary biology—molecular or cellular biology in the first case, physiology and anatomy in the second, and systematic biology in the third. Their connections are real, but sometimes difficult to understand; and it has often been easier to ignore the other levels than to grapple with the complexity they impart to the study of evolution. The marvel is that somehow, the same history of life is inscribed into species in each of their manifold components. It is etched into each body and each cell; the trick is to decode it.

NOTES

1. M. C. King and A. C. Wilson, "Evolution at Two Levels in Humans and Chimpanzees," *Science* 188 (1975): 107–116.
2. D. Hull, "Units of Evolution: A Metaphysical Essay," in U. L. Jensen and R. Harré, eds., *The Philosophy of Evolution*. (Brighton: Harvester Press, 1981), pp. 23–44.
3. S. Sarkar, "Decoding "Coding"—Information and DNA," *BioSystems* 46 (1996): 857–864.
4. This was in honor of Darwin, who paradoxically had nothing to do with it. Darwin had proposed a theory of heredity called "pangenesis," which never gained wide acceptance. Darwin himself, of course, became a biological icon for his theory of evolution by natural selection. Decades later, while searching for a name for the elementary particles of heredity first characterized by Mendel, the evolutionary geneticist Hugo De Vries proposed the Darwinian "pangenes." The second syllable stuck. E. F. Keller, *The Century of the Gene* (Cambridge, MA: Harvard University Press, 2000).
5. William B. Provine, *The Origins of Theoretical Population Genetics*, 2d ed. (Chicago: University of Chicago Press, 2001).
6. W. Bodmer and Robin McKie, *The Book of Man* (London: Little, Brown, 1994).
7. Daniel J. Kevles and L. Hood, eds., *The Code of Codes* (Cambridge, MA: Harvard University Press, 1992).
8. J. Marks, "Beads and String: The Genome in Evolutionary Theory, in E. J. Devor, ed., *Molecular Applications in Biological Anthropology* (New York: Cambridge University Press, 1992), pp. 234–255.
9. J. C. Venter, et al., "The Sequence of the Human Genome," *Science* 291 (2001): 1304–1351.
10. J. Marks, *Human Biodiversity: Genes, Race, and History* (Hawthorne, NY: Aldine de Gruyter, 1995).
11. L. R. Godfrey and J. Marks, "The Nature and Origins of Primate Species," *Yearbook of Physical Anthropology* 34 (1991): 39–68.

12. J. Marks, "Blood Will Tell (Won't It?): A Century of Molecular Discourse in Anthropological Systematics," *American Journal of Physical Anthropology* 94 (1994): 59–80.
13. Amos S. Deinard, "The Evolutionary Genetics of the Chimpanzees" (Ph.D. Thesis, Department of Anthropology, Yale University, 1997).
14. G. Ruano, Jeffrey A. Rogers, Anne C. Ferguson-Smith, and Kenneth K. Kidd, "DNA Sequence Polymorphism within Hominoid Species Exceeds the Number of Phylogenetically Informative Characters for a HOX2 Locus," *Molecular Biology and Evolution* 9, no. 4 (1992): 575–586.
15. C. H. Waddington, "Evolutionary Adaptation," in S. Tax, ed., *Evolution after Darwin* (Chicago: University of Chicago Press, 1960), vol. 1: 381–402.

SUGGESTED READINGS

- Lewontin, R. *The Triple Helix: Gene, Organism, and Environment* (Cambridge, MA: Harvard University Press, 2000).
- Marks, J. *Human Biodiversity: Genes, Race, and History* (Hawthorne, NY: Aldine de Gruyter, 1995).
- Marks, J. *What it Means to Be 98% Chimpanzee* (Berkeley: University of California Press, 2002).
- Monod, J. *Chance and Necessity* (New York: Knopf, 1971).
- Simpson, G. G. *The Meaning of Evolution* (New Haven: Yale University Press, 1951).