Evolutionary Anthropology 17:88–92 (2008)



# All Roads Lead to. . .Everywhere?

**Is the genetic basis of interesting traits so complex that it loses much of its traditional evolutionary meaning?**

KENNETH M. WEISS

All roads lead to Rome, as the ancient proverb goes (Fig. 1). The Roman Imperial Tourist Bureau probably coined it originally. The Empire contained around 85,000 km of official roads–a lot of sympathy is due to the poor Celts, Gauls, and Goths who, under guard, had to build the roads. These roads were interconnected by nearly 400 links so that, indeed, from any point in the empire, you could get to Rome. If you look at any modern road atlas, you can probably say the same about wherever you are at present. From almost anywhere on your continent, you can get to just about any place you want to go. There are main networks of roads, such as superhighways, that go through hubs, such as major city interchanges, but you can usually avoid the stop-and-go by taking side routes, which may be slower but will still get you there. Whether you're a tourist or in a hurry, you might vary your route based on weather conditions, traffic density, the importance of scenery, construction tie-ups, or the location of museums or restaurants.

With such choices, it doesn't make much sense to ask, "What is the road to Rome?" In a somewhat similar way, rapidly growing knowledge about the nature of genomes and what they do suggests that what's good for the Romans is good for biology as well. Instead of a gene for this and a gene for that, we face the possibility that all genes lead to everywhere, which may have important implications with regard to our understanding of the genetic basis or evolution of traits like the shape of the skull, a skull, or this skull. If all real roads lead to the Circus Maximus, do all our craniofacial genetic roads lead to the foramen magnum?

Department of Anthropology, Pennsylvania State University, 409 Carpenter, University Park, PA 16802. E-mail: kenweiss@psu.edu


DOI 10.1002/evan.20156
Published online in Wiley InterScience (www.interscience.wiley.com).

## STEP BY STEP ALONG A ROCKY ROAD TO UNDERSTANDING

We are in a shifting, if not chaotic time in genetics. The comfortable and tractable idea that genes code for proteins became known as the central dogma of biology and is still widely taught to students. It seemed to account for Mendel's green and yellow peas, blood groups, sickle-cell anemia, tissue transplant matching, and many other traits. Examples like these seemed to exemplify universal biologic causality, which enabled genes to become iconic tokens. We didn't have to know what specific genes were responsible for a shorter face, a wider pelvic outlet, or thinner tooth enamel to assert with confidence that hominids evolved by selection favoring those genes.

But the signs from research are that the central dogma is no longer the straight road to biological understanding. The road clearly has begun to erode, although it still is too early to know how seriously the map of causation will have to be redrawn or whether we need to revise our iconic causal arguments. To obtain a sense of this, we can take a brief tour of the kinds of discoveries that have recently been made.

The evidence that regions of the DNA, of which our chromosomes are comprised, contain nucleotide sequences that code for specific amino acid sequences of proteins is still quite secure. However, the accumulation of knowledge from Mendel's time to ours has taken us to places we never expected. Not so long ago we learned that the classical gene is broken into protein-coding exons and noncoding intervening introns. The whole region is transcribed into RNA, but the intronic parts of the sequence are spliced out and the coding parts joined together to form the mature mRNA that is translated into protein.

Then we learned that in between these protein-coding genes there is DNA that is not involved in coding. Indeed, we thought it did nothing, and was hence invisible to natural selection, so it was cutely referred to as "junk" DNA. Then we learned that a cell can't use a gene unless a DNA sequence near the gene is grabbed by specific regulatory proteins (coded for elsewhere in the genome) to enable the gene's RNA to be transcribed. Then we learned that many such regulatory elements are needed for a given gene to be transcribed.

Then we learned that these regulatory sequence elements can be anywhere before, in, or after the coding part, do not need to be adjacent to each other, and do not even always have to be near the coding part, but can be far away on the same chromosome. We learned that the number and location of regulatory sequence elements varied easily over evolutionary time even for a homologous gene.

Figure 1. On the road to Rome. Left, from Pompeii; right, from Spain. If all of their roads led to the Circus Maximus, do all of ours lead to the foramen magnum? Source: public domain http://en.wikipedia.org/wiki/Roman_road.

We learned that the coding part could be spliced in different ways, with different exons being used in different cellular contexts. Then we learned that multiple splicing occurs even in the same cell in a single context.

Then we learned that DNA sequence elements are responsible for the proteins, called histones, that bind to and wrap up the huge chromosomes to make them fit inside the tiny nucleus of a cell, and then that the way wrapping and unwrapping occurs is based on local DNA sequence signals on the chromosome, so that genes in the right parts are unwrapped in the right cells at the right time, so the genes can be expressed.

Then we learned that there are many "epigenetic" mechanisms that affect gene use. For example, there are sequence-based codes that, in specific contexts, are chemically modified; that is, the sequence stays the same but small modifier molecules become attached to the nucleotides to affect whether regulatory proteins can grab that sequence and, hence, whether a nearby gene will be expressed.

We have also learned that mutations in somatic (that is, body rather than germline) cells can affect how they operate, and are inherited by those cells' descendants in the body, but are not inherited across generations because only mutations in the germline cells are transmitted from parent to offspring. Consequently, perfectly genetic traits need not segregate in families.

Then we learned that small RNA molecules are transcribed that do not code for protein but, in fact, have a sequence complementary to specific mRNA, to which it can bind, making the mRNA vulnerable to degradation by specific proteins in the cell. This prevents or inhibits the translation of the targeted mRNA into protein, providing another way to titrate gene expression levels, by adjusting the amount of mRNA and, hence, the amount of protein.

Then we learned that many or most genes are expressed in multiple tissues at different times during development or adult life. So most genes are pleiotropic rather than trait-specific.

Then we learned that much more of the genome, perhaps even most of it, is transcribed into various small RNA molecules, some of which are known to affect gene expression. However, the function of most of these RNA molecules is unknown.

Then we learned that many genes code for proteins that package other newly formed proteins, or protect them from degradation, or transport them, or modify them in the cell. This constitutes another level of genome-encoded means of regulating gene expression and gene use.

Then we learned that most genes seem not to directly affect a final function, but instead work through signaling cascades in which networks of proteins interact in hierarchically ordered stages. For a function to be fulfilled, all of these genes must be expressed in the relevant spatial and temporal contexts.

Then we learned that there are many roads to Rome: Not all pathways in a network need be used in any given context. Moreover, cells can sometimes respond to an aberrant or absent member of a network by using a different path through the network or even through an entirely different network. In single-celled species, many or even most genes can be experimentally deleted with little effect on the cell. This seems to be less true in complex organisms like mammals, but the jury is still out because, for example, genes can be experimentally knocked out in mice with effects that depend on the strain of mouse that's used, showing the availability of alternate roads to Rome.

We've also learned that many genes are expressed from only one of the maternal or paternal copies of the gene in the genome, rather than being expressed by both copies. And we've learned that a substantial part of the genome, including known functional regions, has different copy numbers even in closely related species, and even among different people, or within the same person, since we're diploid. So the genes in the human genome vary somewhat in number from person to person.

## NATURE'S FUNCTIONAL WAL-MART

The point of this breathless race through recent genetic history is to provide at least a sense of the way that many new findings add to, but never reduce, the degree of potential genetic complexity to biological traits; that is, the complexity of the relationships between phenotypes and their underlying genotypic basis.

In this sense, the central dogma of biology isn't exactly false, but is highly incomplete and oversimplified. The genetic road map has changed to the extent that here is a recent attempt at defining a gene: "a union of genomic sequences encoding a coherent set of potentially overlapping functional products."[1] The term has become, at best, a vague generic description of inherited function that resides in DNA. Today, it is only a brave or perhaps reckless geneticist who still refers to any part of the genome as "junk" DNA.

We have moved steadily, for example, from speaking of a gene *for* Mendel's pea color to genetic *effects on* pea color. There are many ways to start, stop, time, or titer the concentration of a coded protein, depending on the circumstances of a cell. Currently, it is not clear how finely tuned this typically is, but we know it varies greatly from system to system, cell to cell, and context to context. The rewards for exaggeration are great these days, but unless recent findings are being heavily over-interpreted, the genome is a Wal-Mart of functions. We have to ask whether this functional cornucopia we are finding in genomes should put us to sleep from boredom or should awaken us to different ways of thinking about the traits we care about. The answer is not yet clear.

A traditional interest of evolutionary anthropology is the evolution of primate skulls. Many changes have occurred, including adaptation for upright posture, omnivorous dentition, and three-dimensional vision. The foramen magnum has moved under the skull, the face shortened, the calvarium increased to accommodate an enlarging brain, and the stuff in that brain become more capable of abstract thought. Our visual sense has improved, while olfaction has diminished. Changes in locomotion have been associated with changes in head orientation, the balance system in the inner ear, and so on. Until recently, we didn't know about or even suspect the genetic facts listed here, so it did not bother us to imagine the genes for craniofacial structure and speak of them as black-box icons by confidently asserting that mutations in face-shape genes that caused shortened hominid faces were preferred by selection. That was never totally satisfying scientifically, because the evolutionary road from there to here might have meandered in various ways within and between species that depend on the nature or complexity of those genetic mechanisms.

We now have many powerful statistical and molecular means of identifying the phenogenetic architecture of our favorite traits. Can we finally be freed from our speculative assumptions about the genetic basis of traditional traits of interest and identify that basis directly? That is not an easy question, and may turn out to be somewhat naïve; it may be that we should change the question itself rather than answering it.

## WHAT DO WE WANT TO KNOW TO "UNDERSTAND" THE PHENOTYPE?

Modern methods document genetic effects in several ways that can be illustrated with growing knowledge of craniofacial shape and evolution.[2–5] The findings for other traits differ in all their details, but the general picture is consistent. One can easily use keywords to search the literature, and especially data bases on human disease (www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM) and mouse mutations (www.informatics.jax.org/), to identify genes that, when mutated, have been shown to produce serious disease or dysgenesis. Hundreds of such genes are known to affect craniofacial development. But this is only the first cut that we can make with available knowledge, and barely includes the plethora of new genetic functions described earlier.

In addition, hundreds of studies have been done to document gene expression in craniofacial development. Embryonic head, jaw, and dental tissues have been obtained, their RNA extracted from the cells and sequenced to identify the genes, as defined by traditional mRNA, that are used in each tissue or developmental stage. For example, a growing data base at Genepaint.org shows the tissue-specific expression pattern of more than 15,000 different genes in a mouse embryo at a stage equivalent to about a month's gestation in humans. Genes of interest can then be manipulated in experiments to identify details of their stage and tissue-specific expression, or the effects of engineered mutations. From such work, we have learned of many different signaling and developmental networks, each containing many genes.[2–5] Much is known about the way genes interact, such as through signaling proteins, the cell-surface receptors that detect them, and the hierarchy of other proteins that interact when the signal is detected to cause tissue to differentiate, such as inducing new bone forming centers, secreting calcification proteins, forming sutures, and so on. These processes are often characterized as "systems" or "networks," one aspect of which is that in networks of interaction there are major hub genes, or limiting factors, which must be in place for proper development, as well as many genes with lesser effects.

Another basic approach to identifying the genetic basis of a trait is gene mapping. Gene mapping studies scan the genome in human families, or experimental crosses in animal models like mice, to find regions of chromosomes in which sequence variation is associated with variation in the trait, such as face length or width or the size or angle of the orbits. Mapping studies typically identify a few genes that have statistically significant effects on variation. But these genes often do not include the same genes that are reported in dysgenesis data bases or ones that are known from experiments to be major genes in developmental networks. If the "major" role played by these genes is being correctly understood, perhaps those key genes are so important that they rarely can vary and remain compatible with normal development. Thus even while those critical genes led the way for developmental biology to identify and characterize developmental genetic networks, they may be under such tight selective constraint that, although they may be functionally more vital, they may be evolutionarily *less* important to the less fundamental changes that constitute most adaptations than are other contributing genes.

Only sometimes do the same genes appear in different mapping studies of the same trait. Moreover, these genes typically account for only a fraction of the overall genetic contribution to the trait as measured by the trait-correlation among relatives (known as heritability). What this appears to mean is that the bulk of genetic effects on craniofacial variation is exerted by many genes making contributions to the trait that are important in the aggregate but too small to detect individually. Presumably, these include the kinds of newly identified genetic elements listed earlier. All of this would be consistent with the idea that evolution molds traits only very gradually, with the large array of newly found mechanisms available for use in the process.

Another finding is that traits like the skull manifest "morphological integration." What that means is that the trait is divided into subregions that developmentally are more or less independently organized, each having its own underlying genetic basis with the characteristics just described. These regions are loosely correlated with each other. In the skull, such regions include the basicranium, brain case, and face; the mandible is developmentally and genetically regionalized in a similar manner.[6–15]

A final fact that is clear from experimental, observational, DNA sequence, and functional data is that developmental networks are highly conserved phylogenetically. The networks' main genes that we see today are ancient. All the genes vary. Some do come and go, but there is substantial conservation in the basic developmental genetic structure. A frog, a chick, and you all used very homologous networks and control systems to progress from egg to adult.

The upshot is that traits like the skull, as well as other classical physical and behavioral traits of interest to anthropology, are controlled by variation in a great many functional regions of the genome. As a result, many different genotypes, which are combinations of DNA sequence variants at different functional regions in different individuals, produce the same trait, such as a long face in a monkey or ape (or in different individuals in a species), or are involved, for example, in the production of tooth enamel. This means more than just that there is redundancy in the system to protect against otherwise harmful mutations. It means that while natural selection may have constrained the trait overall, selection tolerates enough variation that with many contributing genes, each potentially varying, different genotypes can produce a long face or thick enamel.

The many newly discovered genetic functions have revealed that, at least in principle, a much larger number of factors can be involved in a complex trait than just the classical protein-coding genes. These additional mechanisms mean that an increased potential DNA target for mutations is available: More encoded mechanisms in the genome mean that more DNA sequence is involved, and each of the nucleotides is always vulnerable to mutation. More mutations, means more variation, which means more complexity. An important question for genetic research to address is the extent to which these mechanisms do, in fact, play a central role in the kinds of traits, like the skeleton and dentition, that are the traditional interests of organismal biology and anthropology.

If the effect of selection is distributed across numerous contributing genetic mechanisms, it may have only a small impact on the individual variants at any one of them, so that their frequencies are largely free to change over time just by the chance aspects of survival and reproduction (that is, genetic drift). Some individual variants are even being lost while new ones are introduced by mutation. Over time or in different populations, selection may have had different genetic variation to work with, the result being that even when we see convergent phenotypic evolution, with similar traits in different individuals, groups, or species, we find divergent evolution when we look at the underlying genotypes. The more complex the genetic architecture and the weaker the selection, the more likely this is to be true. But such varying genetic basis is, for example, what we see in the case of the response to the very strong selective force of endemic malaria in Asia versus India or Africa.

## WATCHING LIFE GO BY

Maybe because all roads lead to Rome, it is said that if you sit at a café in the Via Veneto, you will eventually see everyone you know. But just because you see your friends doesn't mean you can know how they got there. They may have taken a meandering route, gotten lost a few times, or perhaps they don't even know how they arrived!

Just as construction companies continue to build new roads, research continues to identify new kinds of genetic contributions to biological traits. Sometimes it seems that in attempting to find the genes "for" a trait of the kinds we're usually interested in—behavioral, morphological, or otherwise—we'll either find no genes at all, because most contributing genes have such a small individual effect that we can't detect it, or we'll eventually find that because genes interact with each other in so many ways from conception to adult that all genes, like roads, connect to each other. Hundreds of genes, defined in the modern way to include the panoply of functional units encoded in DNA, are expressed in the development of any complex organ. Complex traits are produced by webs of genetic pathways; that is, of interactions among proteins coded by many different genes. These networks provide many alternate potential routes to a given end, whose use may evolve to vary among or even within species. Thus, as in Figure 2, when we look at a lineup of mangabey or baboon or ape heads or fossil skulls, even from the same population, we can probably say what genetic networks were involved. But we may not be able to say which specific genes or their sequence variants were responsible.

But it may not matter. For what purposes would one want to know such details for a given specimen? Except for the pure and considerable joy of genetics research, which seeks to understand developmental mechanisms and their evolution, it is, most of the time, the trait itself, not its particular genetic architecture, that is important. An adaptive difference between two species or specimens may usually not be attributable to a

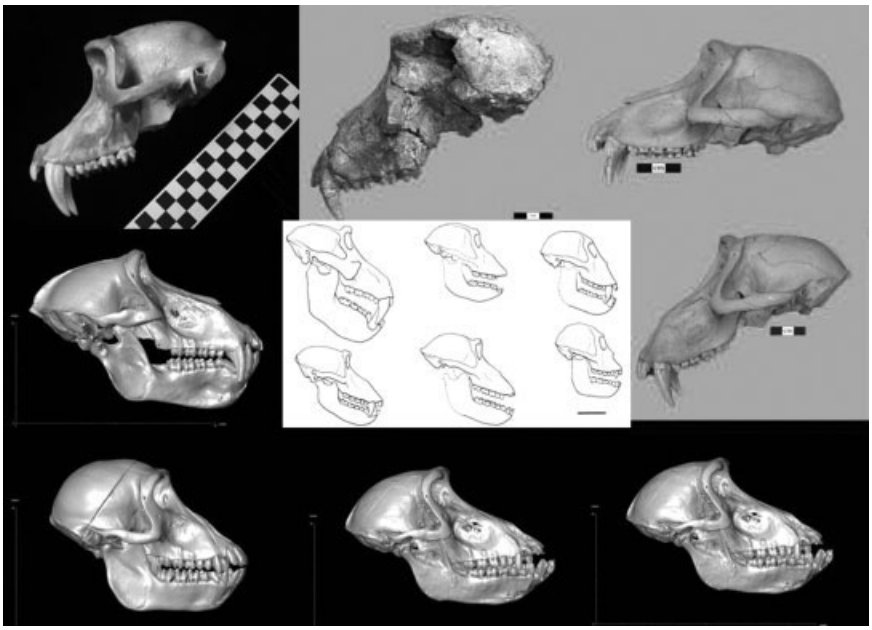

Figure 2. Papionins on the march. Clockwise from top left: A modern and fossil gelada, two Hamadryas males from the same contemporary population, CT scans of male and female African baboons from a single genealogy. Center, a sketch of the range of papionin skulls. For sources, see Acknowledgments.

specific genetic change. If we gene-mapped a trait in a population of ancestral baboons from eons ago, then did it again among modern baboons, we might find that the genetic scaffold would be the same, but its parts replaced, modified, or even removed.

When enough DNA sequence is compared in a study, it seems able to provide reliable taxonomic relationships between species, or even among populations within widely dispersed species, including humans or other primates. This is because the accumulation of genetic variation over time is only indirectly related to any specific traits such as the skull that we may see in the fossil record. However, categorical (e.g, cladistic) approaches to systematics may have to beware of phenotypic convergence in genetically divergent lineages; that is, physical similarities that are homoplasies rather than homologies at the genetic level.

Maybe, for most purposes, our old black-box explanations are still as valid or sufficient as they have always been. If you were in Gaul and wanted to go the capital, there are many roads you could take. But when you finally meet up with your friends, what's important is that they're all there.

## NOTES

I welcome comments on this column: kenweiss@psu.edu. I have a feedback and supplemental material page at http://www.anthro.psu.edu/weiss_lab/index.shtml. I thank Anne Buchanan and John Fleagle for critically reading this manuscript. This column is written with financial assistance from funds provided to Penn State Evan Pugh professors. SFBR CT scans are from a project supported by award BCS 0523637 and BCS 0725227 from the NSF. See www.getahead.psu.edu/hominid_index.html for participating authors.

## ACKNOWLEDGMENTS

Sources for Fig. 2: Geladas provided courtesy of Nina Jablonski, are specimen UCB.LHES.108, courtesy Laboratory of Human Evolutionary Studies, University of California, Berkeley, and 18925ab_bw, courtesy and copyright Department of Paleontology, National Museums of Kenya; Hamadryas specimens courtesy Nina Jablonski, courtesy and copyright, Department of Osteology, National Museums of Kenya; living baboons, from the Southwest Foundation for Biomedical Research. Inset, courtesy and copyright Eric Delson.

## REFERENCES

**1** Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbel JO, Emanuelsson O, Zhang ZD, Weissman S, Snyder M. 2007. What is a gene, post-ENCODE? History and updated definition. Genome Res 17:669–681.

**2** Depew MJ, Tucker AS, Sharpe PT. 2002. Craniofacial development. In: Rossant J, Tam PPL, editors. Mouse development: patterning, morphogensis, and organogenesis. San Diego, CA: Academic Press.

**3** Francis-West PH, Robson L, Evans DJR. 2003. Craniofacial development: the tissue and molecular interactions that control development of the head. Berlin: Springer.

**4** Hall BK, Miyake T. 2000. All for one and one for all: condensations and the initiation of skeletal development. Bioessays 22:138–147.

**5** Helms JA, Cordero D, Tapadia MD. 2005. New insights into craniofacial morphogenesis. Development 132:851–861.

**6** Ackermann RR, Cheverud JM. 2000. Phenotypic covariance structure in tamarins (genus *Saguinus*): a comparison of variation patterns using matrix correlation and common principal component analysis. Am J Phys Anthropol 111: 489–501.

**7** Cheverud JM, Hartman SE, Richtsmeier JT, Atchley WR. 1991. A quantitative genetic analysis of localized morphology in mandibles of inbred mice using finite element scaling analysis. J Craniofac Genet Dev Biol 11:122–137.

**8** Ehrich TH, Vaughn TT, Koreishi SF, Linsey RB, Pletscher LS, Cheverud JM. 2003. Pleiotropic effects on mandibular morphology I. Developmental morphological integration and differential dominance. J Exp Zool Part B Mol Dev Evol 296:58–79.

**9** Hallgrimsson B, Willmore K, Dorval C, Cooper DM. 2004. Craniofacial variability and modularity in macaques and mice. J Exp Zool B Mol Dev Evol 302:207–225.

**10** Klingenberg CP, Leamy LJ, Cheverud J. 2004. Integration and modularity of quantitative trait locus effects on geometric shape in the mouse mandible. Genetics 166:1909–1921.

**11** Leigh SR. 2006. Cranial ontogeny of Papio baboons (*Papio hamadryas*). Am J Phys Anthropol 130:71–84.

**12** Lieberman DE, Krovitz GE, McBratney-Owen B. 2004. Testing hypotheses about tinkering in the fossil record: the case of the human skull. J Exp Zool B Mol Dev Evol 302:284–301.

**13** Cheverud J. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:1737–1747.

**14** Cheverud J. 2004. Modular pleiotropic effects of quantitative trait loci on morphological traits. In: Schlosser G, Wagner G, editors. Modularity in development and evolution. Chicago: University of Chicago. pp 132–153.

**15** Halgrimsson B, Lieberman D, Young N, Parsons T, Wat S. 2007. Evolution of covariance in the mammalian skull. In: Bock G, Goode J, editors. Tinkering: the microevolution of development. Novartis Foundation Symposium 284. Chichester, UK: Wiley.