

GENETIC RESIDUES OF ANCIENT MIGRATIONS: AN END TO BIOLOGICAL ESSENTIALISM AND THE REIFICATION OF RACE

*William M. Richman**

“He’s a credit to his race—the human race.”

Jimmy Cannon, on Heavy Weight Boxing Champion, Joe Louis¹

Behold, my child, the Nordic man,
And be as like him as you can.
His legs are long; his mind is slow.
His hair is lank and made of tow.

And here we have the Alpine Race.
Oh! What a broad and foolish face!
His skin is of a dirty yellow.
He is a most unpleasant fellow.

The most degraded of them all
Mediterranean we call.
His hair is crisp, and even curls,
And he is saucy with the girls.²

Hilaire Belloc

INTRODUCTION

The two quotes reveal starkly different attitudes toward race and the depth and amount of genetic variation within our species. Belloc’s poem postulates racial essentialism, i.e., that there are large, biologically-based differences among the world’s populations on significant traits—intellect, industry, character. By contrast, according to Cannon’s statement, our species

* Distinguished University Professor, University of Toledo College of Law. Richard Delgado, Richard Epstein, William Reynolds, Dan Steinbock, David Armstrong, John Gray, Phillip Closius, Joseph Slater, Courtney Cahill, David Harris, Patrick Borchers, Richard Komuniecki, Geoffrey Rapp, Carol Richman, Emily Richman and Nathan read earlier drafts of this article and offered very helpful comments and suggestions.

1. Quoted in LEE GREEN, *SPORTSWIT* 270 (1986).

2. Belloc’s poem is an irresistible target for anthropologists and is quoted in WILLIAM W. HOWELLS, *MANKIND IN THE MAKING* 275 (1959). It may be unfair to mulct Belloc with this racist sentiment. The poem was intended as an ironic characterization of the British upper class. Belloc himself may or may not have been a bigot. He opposed philosophical anti-Semitism, but may have entertained anti-Semitic prejudices personally.

is unitary, without important geographically-linked differences on those crucial traits. Adopting the essentialist position yields a ready explanation for the disparate fortunes of different populations in our country and across the globe. Some groups, some nations are richer and more powerful and live longer and healthier lives because they are genetically more fit or better adapted than others on certain key parameters.³ The implication is that remedial efforts, either educational solutions or wealth transfers from First to Third World nations, are pointless or at least destined to produce only minimal results.

According to the opposite “unitary” view, all human populations are very similar biologically, and the limited differences that exist are not well correlated with the popular conception of “race” or continent of ancestry. On that view, the differences in the fortunes of different groups are purely contingent, explainable by historical and socio-cultural factors. But those factors can change. Whether they result from the greed and oppression of the “haves” or the cultural difficulties of the “have-nots” is beside the point at this stage of the argument; what matters according to the unitary theory is that the circumstances that created the disparities are mutable.

The politics of the dichotomy are obvious. The unitary view counsels investing resources in efforts to eliminate the contingent causes of group-wise disparities in health, wealth, and life chances. Racial essentialism is more fatalistic; it might be admirable to invest more in the nation’s and the world’s

3. Some readers might suspect that the position that I have labeled “essentialism” is a caricature, a straw man. But see STEPHEN J. GOULD, *THE MISMEASURE OF MAN* 73-112 (1981) [hereinafter GOULD, *MISMEASURE*] (devoting 40 pages to exposing a cavalcade of scientific racists, among them the foremost scientists of their ages). Robert M. Yerkes, a personal favorite and the author (with others) of the Army I.Q. test, evokes the verse by Belloc with his estimates of the national I.Q.s of different countries. See *id.* at 197-99. The most recent examples are still the subject of current controversy. See RICHARD J. HERRNSTEIN & CHARLES MURRAY, *THE BELL CURVE: INTELLIGENCE AND CLASS STRUCTURE IN AMERICAN LIFE* (1994); ARTHUR R. JENSEN, *THE G FACTOR: THE SCIENCE OF MENTAL ABILITY* (1998); MICHAEL LEVIN, *WHY RACE MATTERS* (1997); J. PHILIPPE RUSHTON, *RACE, EVOLUTION, AND BEHAVIOR* (1995). Rushton’s particular brand depends on the r/K dichotomy and what he calls the Life-History theory. Organisms that are r-selected devote more energy to producing multiple offspring while those that are K-selected produce fewer and put more parental care and energy into each one. RUSHTON at 200-16. The result is a ranking of races with Asians at the top in intelligence and at the bottom in athletic ability and genital size, Africans in the opposite position and Europeans in the middle on all three scales. The supposed reason for these differences is that Asians and Europeans moved out of Africa and had to contend with more challenging environments and thus had to grow smarter. *Id.* at 217-33. You can’t make this stuff up!

For a summary of some older ideas of scientific racism, see STEVE OLSON, *MAPPING HUMAN HISTORY* 178-83 (2002).

have-nots, but the likelihood of success is low. What, if anything, can the scientists, particularly the geneticists, add to help resolve the dichotomy?

This article answers that basic question by resorting to the rapidly expanding field of molecular population genetics. Its goal is to bring to the attention of the legal community, including scholars, judges, legislators, and other law and policymakers, a body of knowledge that is increasingly uncontroversial among anthropologists, population geneticists, and biologists. That knowledge concerns the relationship between the popular and the formerly scientific, notions of “race,” on the one hand, and contemporary understandings of genetics and pre-historic human migrations, on the other. Stripped of its scientific complications, the new research reveals a simple and elegant truth. *There simply are no genetically-based pan-“racial” differences in character, intelligence, or any other set of traits crucial to individual or societal success or position; racial essentialism is intellectually bankrupt.*

The first section of the article briefly introduces the idea of race as one way of accounting for human intra-specific variation. Section II then explains the basic human genetics required to understand the impact of advances in molecular population genetics on our understanding of race. The article then proceeds to show how molecular population geneticists have been able to reconstruct the ancient human migrations (Paleolithic dispersals followed by Neolithic expansions) that are responsible for the current geographical distribution of the world’s populations. With that background, Section III catalogues the ways in which recent advances in genetic research have permitted the reshaping of popular, and formerly scientific, ideas about race and human intra-specific variation. Finally, Section IV treats an aspect of contemporary research that has proved to be the last refuge for the racial essentialist. Most genetic studies show that continental clusters of genes, people, and populations exist; the essentialist position is that these clusters reaffirm the biological reality of the traditional racial divisions. Section IV refutes that claim.

There are two things that this article will not do. First, it will not attempt to persuade lawyers, scholars, and other law and policymakers that there is no legal relevance to the entire subject of race. In the lived experience of millions, race matters. It has had pervasive effects (war, famine, abduction, subjugation) on individuals, groups, and whole societies. The concept of race that matters and that has had such momentous influence, however, is the *socio-cultural construction* of race. Thus, this attack on the *biology* of race should have no effect on legal regimes designed to ameliorate past or present inequities resulting from the *socio-cultural* concept of race. It would be a cruel joke, or at least a sad irony, if good science were used to derail the

programs (e.g., affirmative action) designed to remedy the conditions (e.g., slavery, *de jure* discrimination) that bad science helped to create. The fact that the race concept has no basis in *biology* does not make it less real; in fact (willful) misunderstanding of the biology is one of the causes of much of the misery attendant on the history of the cultural phenomenon of race.

Second, I aim not so much to change outcomes on particular legal issues,⁴ but rather to change the set of background assumptions that lawyers and scholars bring to public policy issues. As the American Legal Realist School of jurisprudence made clear, lawyers are not logic engines, but *people*,⁵ so this story should affect the way that lawyers think about race, human variation, and human uniqueness. It is not new to science that race is not biologically crucial or maybe even relevant. And even some part of the legal community knows it as an abstract proposition. But there is all the difference in the world between a proposition on one hand, and the detailed, textured story of which it is a part, on the other.⁶ It is one thing to believe the proposition that race is merely a cultural construct, but quite another to understand the biology and population genetics that justify that conclusion. The difference, of course, is in psychological impact. The abstract conclusion is not nearly as intellectually convincing or emotionally compelling as the story. This article expands the proposition into the story with the goal of changing the way lawyers, and people generally, think about humanness, human categorization, and perhaps, categorization itself.

I. HUMAN VARIATION AND THE IDEA OF RACE

Race is, of course, one way of treating the variety within the human species. The most rudimentary way to handle a diverse phenomenon is to divide it into discrete groups, and that was the model adopted by early theorists of race. That model had its intellectual origin in the Great Chain of Being, a concept rooted in the union of Christian biblical belief and classical philosophical thought. The idea was that the earth's creatures were arranged

4. There are particular problems, legal and social, where the biology of race matters: the rights to ancestral remains, identification of the dead and living (forensic anthropology), and adoption. In adoption, for example, for good reasons or not, prospective parents may wish to choose to adopt a child based on particular phenotypes or placing agencies might use race to determine suitability of prospective parents. On the conflict over the reality of race among forensic anthropologists, see Diana Smay & George Armelagos, *Galileo Wept: A Critical Assessment of the Use of Race in Forensic Anthropology*, TRANSFORMING ANTHROPOLOGY, 2000, No. 2, at 19, 20.

5. See JEROME FRANK, LAW AND THE MODERN MIND 31 (3d prtg. 1935).

6. See, e.g., *Old Chief v. United States*, 519 U.S. 172 (1997).

in a hierarchy according to the essence of each, and, further, that the hierarchy represented a progression toward greater complexity and worth.⁷

The natural extrapolation of that metaphor to humanity was a division of the species by the great taxonomist Carolus Linnaeus into a set of discrete, hierarchically ranked races. He distinguished races not only by physical characteristics, but also by temperament and means of government.⁸ The result was a four-part taxonomy of white Europeans, red Americans, yellow Asians, and black Africans.⁹ Johann Blumenbach later refined the Linnaean classification and added a fifth race, Malay, that today we would call Oceanians.¹⁰ Finally, in 1962 Carleton Coon established the dominant racial classification that persists today among lay folk and government administrators, distinguishing among Caucasoids, Mongoloids, Australoids, Negroids, and Capoids.¹¹

That system used only a few categories, each race consisting of a discrete group of people with certain diagnostic physical characteristics that qualified them for membership in the group and disqualified those that did not possess the key traits. The diagnostic traits related to skin pigment, hair color and texture, and a few others, such as an epithelial fold of fat around the eye and shovel-shaped incisors. The races had natural (before 1492) bounded ranges¹²

7. See MARK RIDLEY, *THE COOPERATIVE GENE* 45 (2001) [hereinafter RIDLEY, *COOPERATIVE GENE*] (discussing ARTHUR O. LOVEJOY, *THE GREAT CHAIN OF BEING* (1936)). The idea owed much to Platonic essentialism and the Aristotelian notion of “telos” or final cause” (“that for the sake of which”), which followed a model of life on earth as a progression toward a final and most perfect state. See also Donald Braman, *Of Race and Immutability*, 46 UCLA L. REV. 1375 (1999); Steven M. Wise, *The Legal Thinghood of Nonhuman Animals*, 23 B.C. ENVTL. AFF. L. REV. 471 (1996). On the fallacy of essentialism in biological taxonomy, see STEPHEN JAY GOULD, *FULL HOUSE* 38-42 (1996).

8. Braman, *supra* note 7, at 1386 n.25.

9. *Id.*

10. Sarah A. Tishkoff & Kenneth Kidd, *Implications of Biogeography of Human Populations for ‘Race’ and Medicine*, NATURE GENETICS, Nov. 2004, at S21, S21.

11. CARLETON S. COON, *THE ORIGINS OF RACES* 3 (1962) [hereinafter COON, *ORIGINS*]. Earlier, he had proposed a system of six races, with the Pacific race added to the five he settled on in 1962. CARLETON S. COON, *THE HISTORY OF MAN* 190-95 (1955) [hereinafter COON, *HISTORY*]. The 2000 Census used a similar classification. “In October 1997, the Office of Management and Budget (OMB) announced the revised standards for federal data on race and ethnicity. The minimum categories for race are now: American Indian or Alaska Native; Asian; Black or African American; Native Hawaiian or Other Pacific Islander; and White.” U.S. Census Bureau, *Racial and Ethnic Classifications Used in Census 2000 and Beyond*, <http://www.census.gov/population/www/socdemo/race/racefactb.html> (last visited Oct. 28, 2006).

For a discussion of the history of racial categories in the United States, see Eric Lillquist & Charles A. Sullivan, *The Law and Genetics of Racial Profiling in Medicine*, 39 HARV. C.R.-C.L. L. REV. 391, 402-08 (2004).

12. COON, *HISTORY*, *supra* note 11, at 190.

and could be exemplified by pure forms or archetypes (essences)¹³ with variation from the norm treated as exceptions or admixtures.¹⁴ The folk wisdom based on these systematics was that racial categories reflect “dramatic, underlying, essential differences.”¹⁵

It did not require the sequencing of the human genome to show the fallacy of this way of cataloguing human variation. Early work on blood proteins by Richard Lewontin and others showed that human variation was much more complicated and continuous than a simple five or six race matrix could reflect. That steadily accumulating body of work convinced many that human races were meaningless, or at least not biologically based, and that race was merely a social construct. Yet that argument, despite its validity, lacked traction with the laity, whose eyesight and “common sense” told them that race was a biological reality.¹⁶ Further, the argument was dismissed as “political correctness,” and some scientists’ empirical and theoretical work seemed to support that skepticism.¹⁷ Even in the post-genomic era, neo-essentialists continue to argue that humanity divides naturally into discrete groups, basing that conclusion on the finding in most sampling studies that genetic markers cluster by continent. The contention seems to be that the persistent continental clustering ratifies the traditional racial divisions.¹⁸

The great contribution of the genomic era to this recurrent wrangle is that it has moved the problem of race beyond politics and largely beyond debate. Geneticists now can quantify human variation, reduce it to provable assertions about molecular structure, and time its development by retracing the

13. On the identification of races as Platonic types, see Rick A. Kittles & Kenneth M. Weiss, *Race, Ancestry and Genes: Implications for Defining Disease Risk*, ANN. REV. GENOMICS HUM. GENETICS (2003), at 33, 35. On essentialism generally in biology, see ERNST MAYR, WHAT EVOLUTION IS 74-75 (2001) [hereinafter MAYR, EVOLUTION IS]; ERNST MAYR, THIS IS BIOLOGY xii (1997) [hereinafter MAYR, BIOLOGY].

14. This brief exposition is, of course, a gross simplification of a great diversity of thought on the concept of race in America and Europe. For a rich treatment, see Braman, *supra* note 7, at 1384-92.

15. That folk wisdom is still alive and well regardless of what science has to say about its validity, see Pilar Ossorio & Troy Duster, *Race and Genetics*, AM. PSYCHOLOGIST, Jan. 2005, at 115, 116 (2005); George J. Armelagos, *Race, Reason and Rationale*, 4 EVOLUTIONARY ANTHROPOLOGY 103 (1995) (reviewing four books on race, three of which advocate neo-racialist positions) [hereinafter Armelagos, *Race*].

16. Smay & Armalagos, *supra* note 4, at 20; Tishkoff & Kidd, *supra* note 10, at S21.

17. See HERRNSTEIN & MURRAY, *supra* note 3; JENSEN, *supra* note 3; LEVIN, *supra*, note 3; RUSHTON, *supra* note 3; David C. Rowe, *Under the Skin: On the Impartial Treatment of Genetic and Environmental Hypotheses of Racial Differences*, AM. PSYCHOLOGIST, Jan. 2005, at 60, 60-70 (2005).

18. Neil Risch, Esteban Burchard, Elad Ziv & Hua Tang, *Categorization of Humans in Biomedical Research: Genes, Race, and Disease*, GENOME BIOLOGY, July 1, 2002, at 2007.1, 2007.4, <http://genomebiology.com/2002/3/7/comment/2007>.

evolutionary changes and pre-history that produced it. Fortunately, their work also has provided the tools for refuting the cluster-based claims of the neo-essentialists. Exactly how the developing work of the geneticists has transformed the discussion is the subject of the remainder of this article, but first some necessary preliminaries.

II. SOME BASIC HUMAN GENETICS¹⁹

The human genome, replicated in every cell, consists of 46 chromosomes, each of which is a double strand of DNA.²⁰ Each strand is composed of millions of nucleotides, cytosine, guanine, adenine, and thymine, which, for convenience sake, can be termed DNA building blocks or components. The order of the nucleotides (e.g., cctagagact) codes the information for constructing all the cells that compose a human.

Thus, the basic unit of human genetic variation is a difference between two people of one nucleotide at a particular location or “locus.” For example, in the previous sequence, if the third nucleotide were adenine rather than thymine (“a” rather than “t”) the two sequences would differ by a single nucleotide. Humans could be said to be polymorphic at that locus, and the variation would be termed a polymorphism.²¹

The human genome contains roughly 3 billion nucleotides, and no two humans, not even identical twins, have exactly the same set of sequences. The nucleotides can be divided by function; a particular sequence of nucleotides on a chromosome is termed a “gene” if it has a recognizable result in the phenotype (physical structure) of the organism.²² Because human reproduction is sexual, each of us has two copies of each chromosome and of each gene, one from each parent. Each copy of the gene is termed an “allele,” and the two copies, one on each of the chromosomes we inherit from our two parents, can be identical or different. If they are different at a particular locus, their owner is said to be heterozygous for that gene; if they are identical,

19. See generally RIDLEY, COOPERATIVE GENE, *supra* note 7; MARK RIDLEY, EVOLUTION (2d ed. 1996) [hereinafter RIDLEY, EVOLUTION].

20. RIDLEY, EVOLUTION, *supra* note 19, at 23 fig.2.2.

21. *Id.* at 80-81. They come in many forms with a bewildering array of names: single nucleotide polymorphisms, insertions, deletions, short tandem repeats, micro-satellite polymorphisms, etc. Each of these terms represents a particular way in which two genomes can differ; for our purpose it is sufficient to refer to them collectively as polymorphisms or markers (when used to trace migrations). See Sarah A. Tishkoff & Brian C. Verrelli, *Patterns of Human Genetic Diversity*, ANN. REV. GENOMICS HUM. GENETICS, 2003, at 293, 297-301.

22. MAYR, BIOLOGY, *supra* note 13, at 307.

homozygous.²³ If variation of one of a pair of alleles is sufficient to control the organism's phenotype for the trait, that allele is "dominant;" if instead variation of both is required to control phenotype, the allele is recessive.²⁴

Variation, and thus heterozygosity, is caused by mutation, essentially a copying mistake made by the cellular machinery during reproduction.²⁵ It is important to understand that the process of mutation is in no sense "guided." The probability of any mutation's occurrence has nothing to do with whether it will be helpful to the organism; every mutation, helpful or not, is simply a copying mistake.²⁶ While the existence of any single mutation is a pure matter of chance, the fate of the mutation is more complicated.²⁷ If the mutation is sufficiently harmful to the organism's life chances, it will disappear in a single generation because its owner will not live long enough to reach breeding age. Thus, for most of human prehistory, a mutation producing low immune function was very unlikely to persist because its owner was unlikely to survive childhood. If the mutation is not biologically harmful, but simply makes its owner an extremely unattractive mating candidate, it will also disappear in short order. Thus, for most of our tenure on the planet, a mutation producing pronounced facial asymmetry probably would not doom the individual but might make it difficult to find a mate. The first of these winnowing processes is termed natural selection; the second, sexual selection.²⁸

It is, of course, possible, that the mutation will be beneficial in one of two ways by conferring on its owner either a selective survival advantage or a selective sexual advantage. Either sort of advantage increases the probability that the owner will have offspring in future generations.²⁹ But chance still plays an enormous role; throughout most of human history, infant and other prepubescent mortality has been very high, meaning that most humans, whether carriers of a helpful mutation or not, were statistically unlikely to produce offspring. Thus, for example, the carrier of a mutation for higher intelligence has no increased chance of surviving accident, parental neglect, warfare, childhood disease or famine, and should she succumb to one of these, the genius mutation will die with her.

23. Lillquist & Sullivan, *supra* note 11, at 410-11; RIDLEY, *EVOLUTION*, *supra* note 19, at 30-31.

24. MAYR, *EVOLUTION IS*, *supra* note 13, at 93, 290.

25. See RIDLEY, *COOPERATIVE GENE*, *supra* note 7, at 80.

26. See OLSON, *supra* note 3, at 35-37; RIDLEY, *EVOLUTION*, *supra* note 19, at 73-74.

27. LUIGI LUCA CAVALLI-SFORZA & FRANCESCO CAVALLI-SFORZA, *GREAT HUMAN DIASPORAS* 88-92 (1995) [hereinafter *GREAT HUMAN DIASPORAS*]. See RIDLEY, *EVOLUTION*, *supra* note 19, at 73-74.

28. For an explanation of natural selection, see RIDLEY, *EVOLUTION*, *supra* note 19, at 64-66; for sexual selection, see *id.* at 296-307.

29. See, e.g., *id.* at 69.

Furthermore, even if the lucky mutant breeds successfully, pure chance determines whether the mutation will appear in the next generation. The reason, of course, is sexual reproduction. Each parent's chromosomes consist of two strands that separate prior to mating, with the offspring receiving one strand from each parent. The mutation, since it occurs only on one strand of one parent (and there are four total strands) has only a one in four chance of appearing in each offspring of the mating.³⁰

Occasionally, however, despite the long odds, a mutation will survive multiple generations, and then, if it confers a sufficiently important survival or sexual advantage, or if its carriers happen to be very lucky, its frequency will increase in the population. If the frequency reaches one hundred percent, the mutation is said to be "fixed" in the population. Often a combination of these factors will counterbalance each other, and the mutation will become fixed at a particular percentage in the population, e.g., hair and eye color in many human populations.³¹

Pure chance has one more role to play in the genetics of populations rather than individuals. A new population that forms and becomes reproductively isolated from its parent population often is not a representative genetic sample of the parent population. Its founding population may consist of specimens that share a particular trait at higher or lower frequencies than the trait appeared in the parent population. Thus, for instance, only ten percent of the parent population might have had a "widow's peak," but, by pure chance, eighty percent of the few individuals who broke off to form the new population had that feature. The phenomenon, known as genetic drift, or "founder's effect," (really only specialized terms for "chance") would produce

30. RIDLEY, *EVOLUTION*, *supra* note 19, at 127.

31. A classic example is the evolution of the peppered moth in England during the Industrial Revolution. Before industrial pollution, the moth had a peppered coloration (black and white) camouflaging it against the predominant background colors of the branches of the trees in its habitat. As industrial pollutants turned tree branches black, moths bearing melanistic mutations (which formerly had been disadvantageous) now enjoyed a selective advantage and began to increase in frequency. *Id.* at 64. Nearly as well known is the human sickle-cell mutation, which causes a change in the shape of red blood cells, a defect that is usually fatal. Thus, a child that is homozygous for the sickle-cell allele likely will not survive long enough to reproduce. All things being equal then the sickle cell allele should die out in a very few generations. It survives because an individual that is heterozygous for the gene (one normal-cell allele and one sickle-cell allele) does not develop the fatal flaw; moreover, it will enjoy a selective advantage (greater resistance to infection by certain strains of malaria) over those who are homozygous for the normal-cell allele. The frequency of the sickle cell allele in a population becomes a balance of the losses of homozygotes due to sickle-cell anemia versus the gains of heterozygotes from their increased malaria resistance. Thus the frequency of the allele in the population will become fixed at a percentage that is optimal based on that population's exposure to the risk of malaria. *Id.* at 110-12.

in the new population and all of its offspring populations much higher frequencies of “widow’s peak” than existed in the original population.³²

III. TRACING MIGRATIONS VIA GENETICS

A. *The Tools*

Several tools are useful for tracing ancient human migrations and expansions. Archaeology, of course, plays a key role because it shows where ancient peoples have passed by unearthing and studying the artifacts they have left behind. Similarly, linguistics is useful because peoples that share a common origin are likely to speak more closely related languages than peoples who do not.³³ All of this is old hat. The most recent quantum leap in our understanding comes from molecular genetics, the very basics of which are the subject of the prior section. However, to appreciate the contribution of that field to our understanding of human prehistory, it’s necessary to go a bit beyond the basics.

The fundamental premise of genetic investigation of human migrations and dispersals is that the genomes of populations theoretically can show that population’s entire lineage all the way back to the origin of the species.³⁴ It works by tracking mutations.³⁵ If, for instance, an early human emigrant from Africa had a particular mutation and the mutation and its carrier survived the several different types of selection, and the carrier was lucky enough to breed, the mutation may appear in some of the carrier’s offspring. If some of those offspring headed toward Australia and acquired another mutation while another group of offspring headed toward Central Asia and acquired a different mutation, it should be possible theoretically to trace the migration

32. Besides mutation, drift and natural and sexual selection, there is one other obvious process that affects the genetics of a particular population—gene flow from another population. CHRISTOPHER STRINGER & ROBIN MCKIE, *AFRICAN EXODUS* 70 (1997) [hereinafter STRINGER & MCKIE, *AFRICAN EXODUS*].

33. See MERRITT RUHLEN, *THE ORIGIN OF LANGUAGE* (1994). Ruhlen’s reconstruction of ancient migrations via linguistic evidence is largely similar to the reconstruction by the geneticists. See *id.* at 190-94. See also LUIGI LUCA CAVALLI-SFORZA, *GENES, PEOPLES, AND LANGUAGES* 155-65 (2000) [hereinafter *GENES, PEOPLES, AND LANGUAGES*]; STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 141.

34. On the history of using genetic markers to trace migrations, see L. Luca Cavalli-Sforza & Marcus W. Feldman, *The Application of Molecular Genetic Approaches to the Study of Human Evolution*, *NATURE GENETICS*, Mar. 2003, at 263.

35. See OLSON, *supra* note 3, at 37.

patterns of the different populations by the serial mutations in their genomes.³⁶ A crude analogy might be serial stamps on a passport or luggage tags on a piece of baggage to show the traveler's route. Further, if the stamps or tags were dated, it would be possible to know not only the traveler's route, but her departure and arrival dates, and the places and dates of her stops along the way; a complete itinerary, if you will.

Fortunately, the tags that the geneticists use to trace migrations—mutations—can be dated, and thus they supply a way of determining the times at which particular migrations have occurred.³⁷ Mutations occur at a fixed rate (so many per generation) and selectively neutral mutations (those that confer no survival or mating advantage, e.g., the widow's peak) spread through a population in a random process known as genetic drift. Drift also occurs at a fixed rate, and so it is possible to count the mutations, measure their spread, and thus date the migrations.

For itinerary reconstruction, "coding" DNA—that portion of the genome that produces differences in the organism's morphology (form) or physiology (function)—is not as useful as non-coding or "junk" DNA.³⁸ The reason for this counterintuitive result is that coding DNA is, by hypothesis, subject to selection pressure. Different environments and mating regimes will differentially favor some traits, coded for by particular bits of DNA, and not others. Thus, partial immunity from malaria is subject to positive selection pressure in some environments. That is a very good thing for the organism, but an impediment to migration itinerary reconstruction because it makes reading the genetic clock more difficult. Mutations occur at a fixed rate (n per generation) based on probability, and those that are selectively neutral spread randomly at a known rate of genetic drift.³⁹ Thus, it is possible to count and

36. In addition to population, geneticists also have studied the genomes of some of our constant companions—rats, fleas, and *helicobacter pylori* (bacterium responsible for stomach ulcers)—to fill in migrational details. See John Pickrell, *Rat DNA Offers Clues to Pacific Colonization, Study Says*, NATIONAL GEOGRAPHIC NEWS, June 9, 2004, http://news.nationalgeographic.com/news/2004/06/0609_040609_ratdna.html; Noah Kemess Whiteman & Patricia G. Parker, *Using Parasites to Infer Host Population History: A New Rationale for Parasite Conservation*, 8 ANIMAL CONSERVATION 175 (2005). See also Risch et al., *supra* note 18, at 2007.5, for a listing of the various types of genetic markers studied.

37. See GENES, PEOPLES, AND LANGUAGES, *supra* note 33, at 17-19.

38. STRINGER & MCKIE, AFRICAN EXODUS, *supra* note 32, at 133. Coding DNA is DNA that contains instructions for the production of proteins that affect the organism's development and ultimate morphology and physiology. Non-coding DNA is DNA whose function and effect we do not (yet) know; it has no *currently apparent* effect on the organism's development. It had been termed "junk" DNA, but more recent research suggests that it does have significant functions. See Noncoding DNA, http://en.wikipedia.org/wiki/Noncoding_DNA (last visited Oct. 28, 2006).

39. SPENCER WELLS, THE JOURNEY OF MAN: A GENETIC ODYSSEY 21 (2002) [hereinafter WELLS,

compare selectively neutral mutations in a population and determine to a fair degree of precision when (and, by inference, where) each occurred.⁴⁰

For coding DNA, however, selection pressure swamps the effects of drift and thus obscures the timing mechanism. The rate of spread or accumulation of a favored mutation depends not on drift, which can be timed, but on the strength of the selection pressure, which is hard to quantify. Thus, while the widow's peak, with its spread dependent on drift, might take thousands of generations to become prevalent in a population, the malarial immunity might do so in a few hundred generations since it confers such a strong advantage on its carriers. Because it is difficult to determine the strength of selection pressures for mutations occurring in the remote past, it is difficult to time them.⁴¹ Thus, most of the research that has been helpful in reconstructing human migrations and expansions has focused on "junk" or non-coding DNA.

Another feature of the human genome makes it possible to separate and trace individually the male and female ancestors of an individual or population. The human egg is relatively large and contains small specialized structures, or organelles, known as mitochondria.⁴² It is probably a remnant of an ancient parasitic bacterium that took up residence in a larger cell and developed a symbiotic relationship with it. Its current contribution to the cell is the production of energy. Because the mitochondrion started out life as an independent organism,⁴³ it has its own genome, much simpler and different in shape from nuclear DNA. Crucial for purposes of migration reconstruction, mtDNA possesses two additional characteristics. First, it does not recombine during reproduction; thus, mutations are conserved in a traceable pattern rather than scrambled and reshuffled in each generation. Second, it sustains a much higher mutation rate than does nuclear DNA; therefore, it reveals more signs of reconstructing migrations.⁴⁴

JOURNEY].

40. *Id.* at 70-71.

41. *Id.* at 21. The widow's peak example is actually an instance of coding DNA that codes for a selectively neutral feature; actual non-coding DNA has no known function or effect on the organism's development. The same principle of drift, rather than selection, controls the spread of non-coding and selectively neutral coding DNA.

42. See WELLS, JOURNEY, *supra* note 39, at 29; L. LUCA CAVALLI-SFORZA, PAOLO MENOZZI & ALBERTO PIAZZA, *THE HISTORY AND GEOGRAPHY OF HUMAN GENES* 83 (1994).

43. GREAT HUMAN DIASPORAS, *supra* note 27, at 77; WELLS, JOURNEY, *supra* note 39, at 29.

44. WELLS, JOURNEY, *supra* note 39, at 29. The reason it does not recombine is that it is a single circular structure, not a double helix. The reason it is so uniquely traceable is that it passes through the female line only. See the following note and its accompanying text.

The human sperm is much smaller than the egg and lacks mitochondria. That means that all the mitochondrial DNA in each of us comes exclusively from our mothers and that it is possible to trace any individual's maternal lineage (mother to grandmother to great-grandmother, etc.) back to a common maternal ancestor. This feature of human genetics resulted in the recent identification of a mitochondrial "Eve," a female who lived in Africa about 200,000 years ago, and is the common maternal ancestor of all humans alive today.⁴⁵

Fortunately, a complementary feature of the human genome makes it possible to trace male ancestry. But understanding how it works requires more detailed knowledge of sexual reproduction. Humans have twenty-three pairs of chromosomes for a total of forty-six. In females, the pairs are all perfectly matched; one pair, the uniquely female pair, consists of two X-chromosomes. During reproduction, they behave as do all other chromosome pairs. Prior to mating, the individual strands separate. Each of the separated strands then breaks at particular locations and the broken pieces of the two different strands reassemble, i.e., recombine. The strands can recombine, however, only in certain orders; DNA from a particular locus on a strand can switch to the other strand, but it can fit on the other strand only at the appropriate locus. The result is two strands, neither of which is identical to either of the strands of the parent.⁴⁶ Instead, each strand contains some combination of the genes from each of the strands of the parent cell. The cell then divides resulting in an egg, which, unlike other cells, is haploid with twenty-three single chromosomes instead of twenty-three pairs. Thus, it is ready (sometime after dinner and a movie) to be fertilized by a sperm, which also is haploid, i.e., contains twenty-three single strands. At fertilization, the two combine to form a diploid cell, containing two versions of each chromosome, one from each parent.⁴⁷

Males are different. While twenty-two of the chromosome pairs in males are perfectly matched, the twenty-third is not. Instead of consisting of two matched X-chromosomes, it consists of one X and one Y-chromosome. The unmatched pair in the male makes the process of pre-mating recombination different from the process for the female. Prior to mating, the twenty-three chromosome pairs split and twenty-two of them break up and recombine, as just described. The twenty-third pair is not well matched and cannot

45. *Id.* at 30. She was not the only woman alive at the time, just the only one lucky enough to have female offspring that continued to reproduce in an uninterrupted chain to the present.

46. On recombination, see MAYR, *EVOLUTION IS*, *supra* note 13, at 103.

47. GREAT HUMAN DIASPORAS, *supra* note 27, at 84.

recombine, i.e., the broken pieces of the two different strands cannot fit together to form two new strands.⁴⁸ When the cell splits to form two haploid sperms, one gets the intact X-chromosome and the other, the intact Y. If the Y containing sperm fertilizes the egg, the Y will join with the X from the egg, and the offspring will be male; if instead the X containing sperm is the lucky winner, the X-chromosome will join with the X in the egg and the offspring will be female.⁴⁹

The result of this process is that the Y-chromosome passes from grandfather to father to son intact without recombination, and so tracking its serial mutations permits reconstruction of each man's paternal ancestry, just as mitochondrial DNA can be used to trace each person's female line. Surprisingly, the process reveals that the coalescence time (the time of the most recent common ancestor) for the Y lineages of all living males is much later than the corresponding coalescence of mitochondrial DNA lineages. The Y-chromosome "Adam" lived in Africa less than 100,000 years ago, and thus Adam and Eve never actually met, although their DNA surely did, and during Adam's lifetime. Every human alive today is an offspring of that relatively recent African "mating."⁵⁰ Thus, somewhat counter-intuitively, the atypical life cycles of mitochondrial DNA and Y-chromosome DNA make them more useful for migration itinerary reconstruction than is "normal" nuclear DNA.⁵¹

B. *The Results*

1. *Out of Africa, Recently*

The previous section shows the tools that molecular genetics can muster to help us understand the history (really prehistory) of our species: who we are, where we originated, and how we got where we are today. This section summarizes the contributions that genetics has made toward answers to those

48. WELLS, JOURNEY, *supra* note 39, at 42-43.

49. For an explanation of how the Y-chromosome behaves during reproduction see *id.*

50. The reason for the difference in coalescence times is the politics of human mating. Nearly every female can mate and has an upward limit on the number of mating opportunities that is a function of physiology. Not so for males; some do not get the opportunity to mate at all, but others mate with any number of females and produce huge numbers of offspring. An extreme example is Genghis Khan whose Y-chromosome appears in about .5 percent of the world total. Tatiana Zerjal et al., *The Genetic Legacy of the Mongols*, 72 AM. J. HUM. GENETICS 717, 717 (2003).

51. On the use of the Y-chromosome to reconstruct human migrations, see Jaume Bertranpetit, *Genome, Diversity, and Origins: The Y Chromosome as a Storyteller*, 97 PROCEEDINGS NAT'L ACAD. SCI. U.S. 6927 (2000).

questions; the next section considers the implications of the genetic record for our understanding of the biology of race.

The story begins with the emergence of the apes in Africa about twenty-three million years ago.⁵² First, gibbons, and later orangutans, and then gorillas split off and formed separate lineages.⁵³ About six million years ago, the remaining sub-population of apes split into two separate groups, one of which would go on to form the genus “pan” (bonobos and chimpanzees), the other, the genus “homo” (humans).⁵⁴ An early offshoot of this latter group, the Australopithecenes, differentiated about four million years ago and began walking upright and exhibiting considerably larger brains.⁵⁵

A still later group of offshoots, emerging about two million years ago, began using stone to make cutting and hammering tools, and most important for our purposes, began migrating out of Africa, establishing populations in Europe and Asia.⁵⁶ These hominids bear a wide and confusing variety of names (e.g., *Homo habilis*, *Homo erectus*, Java Man, Peking Man), and it is not especially important here to distinguish very carefully among them. It is sufficient to note that they were the first members of our genus, *Homo*, and that their brains were larger than *Australopithecus*’ but not as large as the later-emerging *Homo sapiens*.⁵⁷ The first members of our species, the archaic form of *Homo sapiens*, began appearing in Africa about 500,000 years ago,⁵⁸ and, like *Homo erectus*, spread to other continents and differentiated there to form local climatically adapted subspecies, the most famous of which is Neanderthal Man.⁵⁹ Later, fully modern humans began to emerge in Africa about 200,000 years ago and stayed there exclusively for the next 150,000 years.

52. WELLS, *JOURNEY*, *supra* note 39, at 55. On the genetic proximity of humans and African apes see also STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 20.

53. National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/> (last visited Oct. 29, 2006).

54. JARED DIAMOND, *THE THIRD CHIMPANZEE* 34-39 (1992) [hereinafter DIAMOND, *THIRD CHIMPANZEE*]; OLSON, *supra* note 3, at 19.

55. *Id.* at 33-34.

56. GREAT HUMAN DIASPORAS, *supra* note 27, at 42-45; WILLIAM HOWELLS, *GETTING HERE* 102 (1997).

57. DIAMOND, *THIRD CHIMPANZEE*, *supra* note 54, at 37.

58. *Id.*

59. See STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 47. For more on Neanderthals, see HOWELLS, *supra* note 56, at 141-55.

2. *Paleolithic Dispersals*

About 150,000 years ago, the population began to differentiate and spread throughout Africa in several waves.⁶⁰ One of the earliest splits resulted in three major mitochondrial lineages. L1, the most ancient of the three,⁶¹ survives at high frequencies today among several southern and eastern African populations that speak click languages: Bushmen (also called San) Khoi, Hadza, Sandawe. Also displaying high frequencies of the lineage are the Biaka (western) pygmies of Cameroon and the Central African Republic.⁶² Additional mitochondrial lineages, L2 and L3, separated from L1 about 80,000 years ago; L2 is common in many West African Bantu-speaking populations and in Mbuti (eastern) Pygmies, who now live near border between Uganda and the Democratic Republic of Congo. L3, common in East Africa but infrequent in West and South Africa, occurs at high frequency among Ethiopians and is probably the ancestor of all Eurasian populations.⁶³

60. See Peter Forster, *Ice Ages and the Mitochondrial DNA Chronology of Human Dispersals: A Review*, 359 *PHILOSOPHICAL TRANSACTIONS ROYAL SOC'Y LONDON* 255, 257 (2004); Tishkoff & Verrelli, *supra* note 21, at 307. The term "Paleolithic" means literally "old stone age," and refers to a culture which was the only human lifestyle for about 150,000 years. There was no metallurgy, and most tools were fashioned from stone initially and later from animal bone or sinew; food was obtained by foraging. "Neolithic" refers to a more developed culture that also lacked metallurgy, but incorporated agriculture and herding. Tools were of natural materials but became more sophisticated and specialized.

61. Lev A. Zhivotovsky et al., *Features of Evolution and Expansion of Modern Human Populations from Genomewide Microsatellite Markers*, 72 *AM. J. HUM. GENETICS* 1171, 1182 (2003). Before this population split, the species probably underwent a significant expansion event, starting with a population of about 6,600. See David E. Reich & David B. Goldstein, *Genetic Evidence for a Paleolithic Human Population Expansion in Africa*, 95 *PROC. NAT'L ACAD. SCI. U.S.* 8119, 8123 (1998).

62. Pygmies no longer have their own language but speak the languages of the surrounding farmers. *GREAT HUMAN DIASPORAS*, *supra* note 27, at 9.

63. Tishkoff & Verrelli, *supra* note 21, at 309. For more on Africa, see generally Reich & Goldstein, *supra* note 61; Antonio Salas et al., *The Making of the African mtDNA Landscape*, 71 *AM. SOC. HUM. GENETICS* 1082 (2002). This view of human dispersal, now widely accepted is known as Recent African Origin (RAO); the opposite view, multi-regionalism posits an early migration of *Erectus* out of Africa and parallel evolution of modern humans on several different continents. The difference between the two theories is highly relevant to the subject of this paper as the RAO view postulates racial differences that are only 100,000 years old and thus relatively superficial. On the two theories, see *GENES, PEOPLES, AND LANGUAGES*, *supra* note 33. The MR view depends upon a million-year-old separation, which would allow for fundamental racial differences. See STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 59-61. Some evidence for the MR view exists in the shared anatomic features of archaic and modern humans in particular geographic regions, e.g., shovel-shaped incisors among archaic and modern Asians. The principal difficulties with MR are (1) the genetic evidence suggesting RAO, (2) the improbability of nearly exact convergent evolution of archaics into moderns in several different, widely separated locales, (3) the improbability of significant genetic exchange among small, widely separated groups of archaics (the only alternative to coincidental convergent evolution), and (4) the high degree of genetic difference among African populations compared to non-African ones. See Recent Single-Origin Hypothesis, <http://>

In addition to these mitochondrial data, evidence from the Y-chromosome also indicates that the current click language speakers are descendants of the most ancient human lineage and that current Eurasian populations derive from African populations.⁶⁴

The concentration of Eurasian ancestor markers in Ethiopia suggests that the original migration out of Africa occurred about 60,000 years ago via the Horn of Africa and into the Arabian Peninsula rather than through the Middle East, as had been thought previously.⁶⁵ Climatological and genetic evidence confirm the hypothesis.⁶⁶ Immediately before the African exodus, the continent was in the midst of a drought that caused the rich savannas of East Africa to become more desert-like, leaving savannas only in narrow strips near the coast.⁶⁷ Coastal environments present rich foraging opportunities for hunter-gatherer populations, and middens (ancient garbage piles) unearthed in East African coastal regions suggest that populations in that region exploited those resources.⁶⁸ The world-wide lowering of sea levels made a migration from Africa to the Arabian Peninsula feasible and the genetic evidence suggests that just such a migration did occur.⁶⁹

The first non-Africans were descendants of the L3 mitochondrial (mother's side) lineage and a Y-chromosome (father's side) lineage with the rather prosaic name of M168.⁷⁰ Descendants of that union split soon after arriving in Arabia with a small portion moving east and hugging the coast, a route that permitted exploiting the same resource base that had supported them

en.wikipedia.org/wiki/Recent_single-origin_hypothesis (last visited Dec. 1, 2006). The RAO model is clearly dominant. Milford Wolpoff is the most well-known proponent of multiregionalism. See MILFORD WOLPOFF & RACHEL CASPARI, *RACE AND HUMAN EVOLUTION: A FATAL ATTRACTION* (1997).

64. Tishkoff & Verrelli, *supra* note 21, at 309.

65. There is, however, some evidence for migration along both routes, from the Horn and through the Levant. See J.R. Luis et al., *The Levant Versus the Horn of Africa: Evidence for Bidirectional Corridors of Human Migrations*, 74 AM. J. HUM. GENETICS 532 (2004).

66. Modern humans had moved out of Africa and into the Middle East earlier, but did not stay long. Cyclic warming and cooling of the area made it sometimes more hospitable to modern humans, and other times more habitable by cold-adapted Neanderthals. WELLS, JOURNEY, *supra* note 39, at 98-99.

67. See *id.* at 68.

68. *Id.*

69. About ten miles distance separates the Horn of Africa from Yemen and there are close genetic links between the populations of those areas, suggesting two-way migration over many years. The narrow inlet is called the Bab-el-Mandeb or "Gate of Tears." See Toomas Kivisild et al., *Ethiopian Mitochondrial DNA Heritage: Tracking Gene Flow Across and Around the Gate of Tears*, 75 AM. J. HUM. GENETICS 752, 752 (2004). This is known as the Out of Africa model, as contrasted with Multiregionalism. See J. Lynn B. Jorde & Stephen P. Wooding, *Genetic Variation, Classification and 'Race'*, NATURE GENETICS, Nov. 2004, at 28.

70. WELLS, JOURNEY, *supra* note 39, at 72.

in coastal East Africa. That group, identified by the mitochondrial marker M⁷¹ and the Y-chromosome marker M130, proceeded with astonishing speed to colonize the coastal portions of Arabia, Yemen, Iraq, Persia, India and eventually Malaysia, New Guinea and Australia, resulting in human population of Australia as early as 50,000 years ago.⁷² Their genetic markers, while prevalent in some places along the coastal route, are virtually absent in Europe and other parts of Asia.⁷³

There are two difficulties with this coastal migration scenario, but both resolve relatively easily. First, the key markers appear only at moderate frequencies in current populations along many parts of the route, reaching majority or even unanimous representation only among Aboriginal Australians.⁷⁴ Second, along many parts of the route there is very little archaeological evidence of this early migration.⁷⁵ Changing sea levels account for this paucity of archaeological remains; sea levels at the time were 300 feet lower than today, so any remnants of the settlements of those early coastal migrants would lie under hundreds of feet of ocean.⁷⁶

The first difficulty, the low frequencies of the tell-tale coastal migration markers among many populations currently living along the route, results from subsequent migrations from the north that have swamped the genetic contribution of the early coastal migrants.⁷⁷ A south/north gradient in frequency of those markers in many places along the route confirms that hypothesis as does the presence of several population isolates: the so-called Negritos of the Philippines, the Semang of Malaysia and the Andaman Islanders, all of whom are closer genetically to Africans and Australians than to their closer neighbors.⁷⁸

71. See OLSON, *supra* note 3, at 129.

72. Max Ingman & Ulf Gyllensten, *Mitochondrial Genome Variation and Evolutionary History of Australian and New Guinean Aborigines*, 13 GENOME RESEARCH 1600 (2003).

73. WELLS, JOURNEY, *supra* note 39, at 72-74.

74. *Id.* at 78.

75. *Id.*

76. *Id.* at 74.

77. *Id.* at 113; Peter A. Underhill, *A Synopsis of Extant Y Chromosome Diversity in East Asia and Oceania*, in THE PEOPLING OF EAST ASIA: PUTTING TOGETHER ARCHAEOLOGY, LINGUISTICS AND GENETICS 313 (2004) [hereinafter Underhill, *Synopsis*]; P.A. Underhill, *Inferring Human History: Clues from Y-Chromosome Haplotypes*, 68 COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY 487, 489 fig.1 (2003) [hereinafter Underhill, *Inferring*].

78. See Kumarasamy Thangaraj et al., *Genetic Affinities of the Andaman Islanders, a Vanishing Human Population*, 13 CURRENT BIOLOGY 86, 86 (2003). For more on Australia, see Underhill, *Synopsis*, *supra* note 77, and Ingman & Gyllensten, *supra* note 72.

So much then for the smaller portion of the initial “out of Africa” migration. The larger portion contained individuals with high frequencies of the mitochondrial marker N and the Y-chromosome marker M89; their descendants went on to populate most of the rest of the world.⁷⁹ Very likely, the initial move was north on the Arabian peninsula, exploiting coastal resources there as well as inland hunting opportunities made possible by good grazing for their principal prey species.⁸⁰ Alternatively, the migration may have followed the Nile through the Sahara, which at the time was a relatively moist savanna, and spread through the Sinai peninsula.⁸¹ By either or both routes, the migrants had reached the Middle East by about 50,000 years ago.⁸²

At that point, the migrants were committed to Eurasia; back migration into Africa became extremely difficult because drying weather had turned the Sahara and Arabian Deserts into a nearly impenetrable barrier for the migrants.⁸³ The migration then split with a small portion (identified by a Y-chromosome mutation designated M172) penetrating Europe only minimally via the Balkans.⁸⁴ This route was not ideal because Europe was forested and the hunter-gatherers’ principal food resource consisted of herds of grazers that required savanna to flourish in large numbers.⁸⁵ That savanna environment existed plentifully, however, throughout southwestern Asia, and the largest portion of the human migration—this one identified by a different Y-chromosome marker (M9)—followed the herds east into the heart of south central Asia.⁸⁶

Following that belt of savanna, the descendants of M9 pushed east until they encountered the Hindu Kush Mountains,⁸⁷ rugged, glaciated and largely impenetrable to them and, just as important, to the savanna-dwelling grazers that were their principal food resource.⁸⁸ Faced with that barrier, one group headed south into India and Pakistan, where their male descendants are recognizable by the M20 (Indian Clan) mutation that they carry on their Y-chromosomes.⁸⁹

79. WELLS, JOURNEY, *supra* note 39, at 112-13.

80. *Id.* at 108.

81. *Id.*

82. *See id.*

83. *Id.*

84. *See id.* at 182-83 fig.10.

85. *Id.* at 110-11.

86. *See id.* at 111.

87. *Id.* at 111.

88. *Id.* at 112.

89. *Id.* at 112-13. For more on the genetic landscape of India, see T. Kivisild et al., *The Genetic*

Moving south in India, they encountered and mixed with the earlier coastal migration out of Africa that had populated the southern portion of the subcontinent thousands of years earlier. The genetic evidence here shows an asymmetrical mating pattern that occurs with distressing regularity in cases of successive migrations or conquests. The mitochondrial (female side) marker M, characteristic of the coastal migrants, appears commonly in the current South Indian population, but the corresponding Y-chromosome coastal marker (M130) does not. Instead, the prevalent Y-chromosome marker is M20, the tell-tale of the later land migration through Asia.⁹⁰ The inference, of course, is that the later invaders mated with the females of the coastal clan but either killed, drove off or merely reduced the mating opportunities of the coastal males.⁹¹

Another branch of the eastward-moving Eurasian (M9) group headed north into Central Asia rather than south into India.⁹² It faced two additional restrictions on its migration: the way east into Western China was cut off by the Tien Shan Mountains,⁹³ and the way due west back through Iran was cut off by desertification brought on by a period of cooling weather and advancing glaciation.⁹⁴ One portion of the M9-bearing group detoured around the mountains and then headed south into China. A mutation in their Y-chromosome, known as M175 arose at that point and distinctively marks that lineage.⁹⁵ At that point, members of the M175 group encountered the earlier coastal migrants, who had settled East Asia from south to north. The two populations inter-mated to form the ancestral population of current East

Heritage of the Earliest Settlers Persists Both in Indian Tribal and Caste Populations, 72 AM. J. HUM. GENETICS 313 (2003); Malliya gounder Palanichamy et al., *Phylogeny of Mitochondrial DNA Macrohaplogroup N in India, Based on Complete Sequencing: Implications for the Peopling of South Asia*, 75 AM. J. HUM. GENETICS 966 (2004).

90. WELLS, JOURNEY, *supra* note 39, at 113.

91. *Id.* There are many examples of this asymmetrical mating pattern, including replacements of hunter/gatherer populations by Neolithic farmers, replacement of Native Americans by European colonizers and replacement of Australian Aborigines by Europeans. It may be function as a genocide marker. Not every asymmetry signals a genocide or even a conquest, however. When the asymmetry is weaker, a possibility is a dual-gender migration rather than a male-only conquest. See Bo Wen et al., *Analyses of Genetic Structure of Tibeto-Burman Populations Reveals Sex-Biased Admixture in Southern Tibeto-Burmans*, 74 AM. J. HUM. GENETICS 856 (2004).

92. WELLS, JOURNEY, *supra* note 39, at 182 fig.10.

93. *Id.* at 112.

94. *Id.* at 118.

95. See WELLS, JOURNEY, *supra* note 39, at 118-19. For more on the genetic landscape of East Asia, see Bo Wen et al., *supra* note 91; Yong-Gang Yao et al., *Phylogeographic Differentiation of Mitochondrial DNA in Han Chinese*, 70 AM. J. HUM. GENETICS 635 (2002); Yuan-Chun Ding et al., *Population Structure and History in East Asia*, 97 PROC. NAT'L ACAD. SCI. U.S. 14003 (2000).

Asians, which today still displays a north/south gradient of genetic markers.⁹⁶ Much later, M122 arose on some of their M175 Y-chromosomes and may be a marker for Neolithic expansion.⁹⁷

Not all of the Eurasian M9 lineage rounded the mountains into China. Another portion, which had acquired an additional Y-chromosome mutation, M45, also found itself in Southern Siberia. There it split again, one part of the M45-bearing group moving northwest around the deserts into Europe, another heading northeast around the mountains into Siberia.⁹⁸ Once isolated from each other, each part of the split population began accumulating different mutations and thus acquired its own unique Y-chromosome marker, M173 in the case of the future Europeans and M242 in the case of the soon-to-be central and Eastern Siberians.⁹⁹

Migrating west of the Urals, the European clan encountered another older human population that had migrated out of Africa as much as 300,000 years earlier.¹⁰⁰ The Neanderthals had been living in Europe for thousands of years and had developed physiological adaptations to the cold climate, stocky bodies and short limbs among others.¹⁰¹ Moderns and Neanderthals co-occupied Europe for only a brief period, with the Neanderthals all but disappearing by 30,000 years ago.¹⁰² It is unclear whether moderns had any role in the disappearance of the Neanderthals, either by direct conflict or simply by out-competing them or forcing them into marginal territories.¹⁰³ It is clear that there was no significant genetic mixture between the two populations.¹⁰⁴ Modern Europeans do not have Neanderthals among their ancestors.¹⁰⁵

96. See WELLS, JOURNEY, *supra* note 39, at 120-21.

97. *Id.* at 156-57; for more on the Neolithic expansions, see *infra* Section II.B.3.

98. See R. Spencer Wells et al., *The Eurasian Heartland: A Continental Perspective on Y-Chromosome Diversity*, 98 PROC. NAT'L ACAD. SCI. U.S. 10244, 10247-48 (2001) (fixing M45 as the source marker for the European and Native American populations) [hereinafter *Heartland*].

99. Mark Seielstad et al., *A Novel Y-Chromosome Variant Puts an Upper Limit on the Timing of First Entry into the Americas*, 73 AM. J. HUM. GENETICS 700 (2003).

100. See Peter Beerli & Scott V. Edwards, *When Did Neanderthals and Modern Humans Diverge?*, 2002 EVOLUTIONARY ANTHROPOLOGY SUPPLEMENT 60.

101. The best estimate is that the ancestors of modern humans and Neanderthals diverged genetically about 700,000 years ago. *Id.*

102. Forster, *supra* note 60, at 260.

103. See STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 109-11.

104. See *id.* at 53.

105. Forster, *supra* note 60, at 260. Moreover, neither the Neanderthals nor any other population of archaic Homo sapiens has descendants among current human populations. Although they survived successfully for hundreds of thousands of years, they form an evolutionary dead end. For more on the making of the genetic landscape of Europe, see Lluís Quintana-Murci et al., *Where West Meets East: The Complex mtDNA Landscape of the Southwest and Central Asian Corridor*, 74 AM. J. HUM. GENETICS 827

There is fairly wide agreement that the Siberian clan, those bearing the M242 marker, would eventually colonize the Americas.¹⁰⁶ That marker is found widely in the new world as is its descendant M3, which is found there almost exclusively,¹⁰⁷ suggesting that it arose after the American migration left Asia. While the basics are not controversial, there is a good deal of controversy over three closely related issues: when were the Americas colonized, how many waves of migration participated, and what is the likely Asian source population?

The time issue implicates the “Clovis First” debate. Nearly all authorities agree that humans had entered the New World by 11,000 years ago (11 KYA), as evidenced by distinctively carved stone projectile points first found near Clovis, New Mexico, and later throughout North and South America. Some, however, argue for a much earlier entry (as early as 30 KYA), citing archaeological finds at Meadowcroft, Pennsylvania, and Monte Verde, Chile. Dating to before 13 KYA, they would require hypothesizing an entry as least as early as 15 KYA.¹⁰⁸ At present, there is little genetic evidence for an entry before 15 KYA.¹⁰⁹

Evidence from linguistics, geology and genetics has figured importantly in the “number of waves” issue. Joseph Greenberg, who produced the first coherent taxonomy of Native American languages hypothesized three separates waves of migration: one about 15 KYA composed of speakers of a widely dispersed family of languages known as Amerind; another, about 9 KYA composed of Na-Dene speakers (Navaho, Apache, and Athabaskan); and a final wave of speakers of Eskimo-Aleut languages that entered later yet.¹¹⁰

(2004); S. Plaza et al., *Joining the Pillars of Hercules: mtDNA Sequences Show Multidirectional Gene Flow in the Western Mediterranean*, 67 ANNALS HUM. GENETICS 312 (2003); Martin Richards et al., *In Search of Geographical Patterns in European Mitochondrial DNA*, 71 AM. J. HUM. GENETICS 1168 (2002); Siiri Rootsi et al., *Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe*, 75 AM. J. HUM. GENETICS 128 (2004); Kristiina Tambets et al., *The Western and Eastern Roots of the Saami—the Story of Genetic “Outliers” Told by Mitochondrial DNA and Y Chromosomes*, 74 AM. J. HUM. GENETICS. 661 (2004); *Heartland*, *supra* note 98.

106. The isolation of the marker is described in Seielstad, *supra* note 99.

107. There are some reports of M3 in extreme Northeast Siberia, suggesting a back migration from Beringia. See Maria-Catira Bortolini et al., *Y-Chromosome Evidence for Different Demographic Histories in the Americas*, 73 AM. J. HUM. GENETICS. 524 (2003).

108. WELLS, JOURNEY, *supra* note 39, at 136. See generally J.M. ADOVASIO WITH JAKE PAGE, *THE FIRST AMERICANS* (2002); THOMAS D. DILLEHAY, *THE SETTLEMENT OF THE AMERICAS* (2000).

109. Mark Sielstad et al., Letter to the Editor, *A Novel Y-Chromosome Variant Puts an Upper Limit on the Timing of First Entry into the Americas*, 73 AM. J. HUM. GENETICS 700 (2003) (fixing the date of entry by finding marker M242 between M45 and M3 on the Y-chromosome).

110. RUHLEN, *supra* note 33 at 169-70.

Subsequent genetic studies have suggested fewer waves, perhaps only one.¹¹¹ Further, all of this is complicated by the geological history of arctic Asia and North America, some periods permitting easy cross-Bering migrations, others not, and some permitting glacier-free corridors from Alaska into the Great Plains and others not.

One way to harmonize much of the conflicting data is to understand that the migration did not utilize a “land bridge” across the Bering Straits.¹¹² Rather, at the glacial maximum, sea levels were low enough that Beringia itself was a homeland, at least for several thousand years, rather than just a migration corridor. It was a mini continent hundreds of miles from west to east and nearly a *thousand* miles from north to south. This “Out of Beringia” model hypothesizes a migration into Beringia from Siberia, followed by a lengthy sojourn there during which populations diverged biologically, culturally and linguistically. Eventually, as sea levels rose there were a series of migrations, timed by the availability of interglacial access routes, by these differentiated populations into North America and back into Asia.¹¹³

The final unresolved issue is the Asian source population for the new Americans. The current residents of far Northeast Siberia (Chuckchi excepted) are not closely related to Native American populations. The inference is that current Northeast Siberians settled there after the migration to the Americas had already occurred. Rather, most studies identify the current inhabitants of the Altai Mountains in South Central Siberia as the likely predecessor population of most Native Americans.¹¹⁴ Some evidence points to an additional founder population in far Eastern Siberia, which may

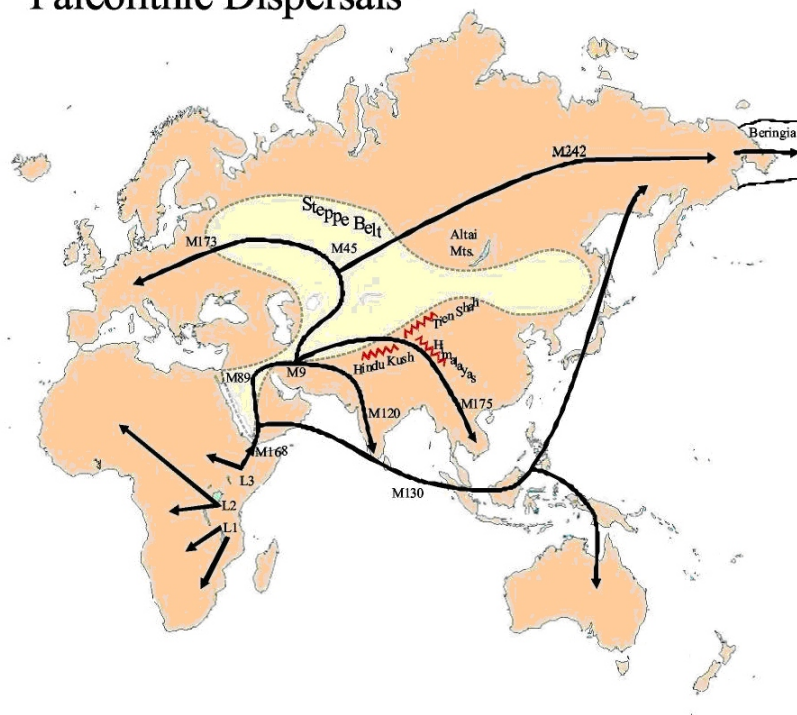
111. Jeffrey T. Leil et al., *The Dual Origin and Siberian Affinities of Native American Y Chromosomes*, 70 AM. J. HUM. GENETICS 192 (2002) (finding two waves of migration, one from South Central Asia and one from coastal East Asia); Ripan S. Mahli et al., *The Structure of Diversity within New World Mitochondrial DNA Haplogroups: Implications for the Prehistory of North America*, 70 AM. J. HUM. GENETICS 905 (2002) (finding the population came in a single wave from a single source in the Altai Mountains of South Central Siberia); Connie J. Mulligan et al., *Population Genetics, History, and Health Patterns in Native Americans*, 5 ANN. REV. GENOMIC & HUM. GENETICS 295 (2004) (finding American population results from a single migration); Stephen L. Zegura et al., *High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y Chromosomes into the Americas*, 21 MOLECULAR BIOLOGY & EVOLUTION 164 (2004) (finding one migration event about 17,000 years ago with a remote source in the Altai Mountains).

112. See Sandro Bonatto & Francisco M. Salzano, *A Single and Early Migration for the Peopling of the Americas Supported by Mitochondrial DNA Sequence Data*, 94 PROC. NAT'L ACAD. SCI. U.S. 1866 (1997) (originating from the Beringia model).

113. H.J. Bandelt et al., *Identification of Native American Founder mtDNAs Through the Analysis of Complete mtDNA Sequences: Some Caveats*, 67 ANNALS HUM. GENETICS 512 (2003) (finding support for a staggered series of migrations from Siberia to Beringia to the Americas).

114. See, e.g., Mulligan et al., *supra* note 111.

Paleolithic Dispersals*



Age of Markers

L1 - @ 150 kya

L2 - @ 80 kya

L3 - @ 80 kya

M168 - @ 60 kya

M130 - @ 50 kya

M89 - @ 45 kya

M9 - @ 40 kya

M45 - @ 35 kya

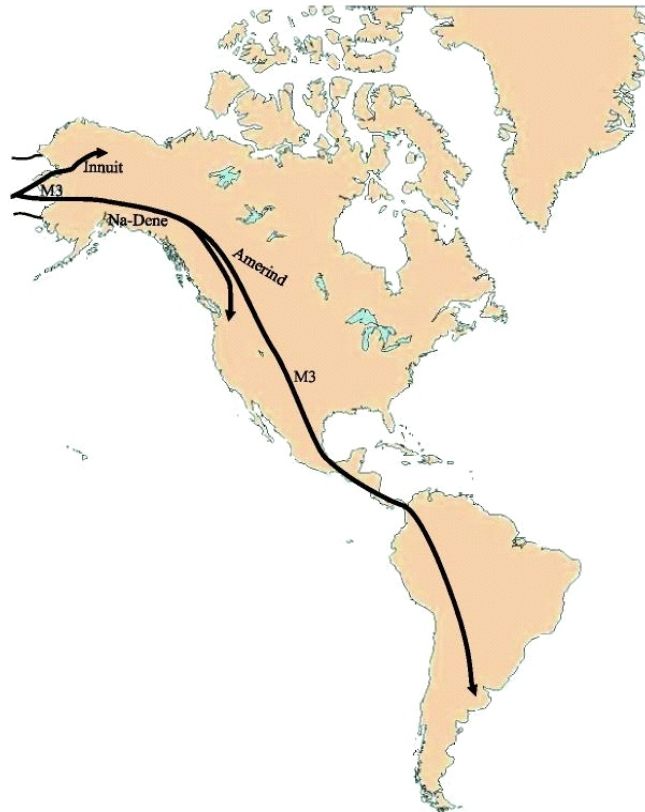
M173 - @ 30 kya

M20 - @ 30 kya

M242 - @ 20 kya

M3 - @ 15 kya

* After Wells, *Journey*, *supra* note 37.



represent a separate migration wave of the ancestors of today's Na-Dene and Eskimo-Aleut speaking populations.¹¹⁵

Whatever the eventual resolution of the number-of-waves, source-population and date-of-entry issues, there is wide agreement that humans had occupied all of North and South America only a few thousand years after the original entry.¹¹⁶

3. Neolithic Expansions

Thus, by about 14,000 years ago, humans had occupied all the continental land masses on the globe, leaving only the oceanic islands, which would wait another 10,000 years to be settled. 14,000 years ago also marks the time of maximum human genetic diversity. When the original human population expanded, first within and then outside of Africa, the sub-populations separated; each then evolved to meet climatic changes and experienced genetic drift, resulting in genetic differences among populations. As we will see later the divergences are not that great, but, such as they are, their greatest extent existed then. Since then, the process has operated in reverse. The last 14,000 years have seen a steady diminution in human genetic diversity caused by population consolidations, technological developments, wars, conquests, nation-founding, colonizations, epidemics, localized extinctions, and slave trading.¹¹⁷

The first major development was the invention of food production, agriculture and herding. After 150,000 years as hunter-gatherers, humans began producing food about 11,000 years ago.¹¹⁸ There is clear archaeological evidence that this development occurred independently in at least five locations at varying times: Southwest Asia (11 KYA), China (9.5 KYA), Mesoamerica (5.5 KYA), the Andes and Amazonia (5.5 KYA), and the Eastern United States (4.5 KYA). In four other locations, the Sahel (7 KYA), Tropical West Africa (5 KYA), Ethiopia, and New Guinea (9 KYA),

115. WELLS, JOURNEY, *supra* note 39, at 142.

116. For more on the population of the Americas, see Bonatto & Salzano, *supra* note 112; Keith Hunley & Jeffrey C. Long, *Gene Flow across Linguistic Boundaries in Native North American Populations*, 102 PROC. NAT'L. ACAD. SCI. U.S. 1312 (2005); Elena B. Starikovskaya et al., *Mitochondrial DNA Diversity in Indigenous Populations of the Southern Extent of Siberia, and the Origins of Native American Haplogroups*, 2005 ANNALS HUM. GENETICS 67, 85-86 (comparing Amerind, Na Dene, Eskimo, and Chukchi populations, all of which were inhabitants of Beringia).

117. See sources cited in notes 120-59, *infra*.

118. See generally George J. Armelagos & Kristen N. Harper, *Genomics at the Origins of Agriculture*, 14 EVOLUTIONARY ANTHROPOLOGY 68 (2005).

agriculture may have been invented independently or may have been imported; the evidence is inconclusive.¹¹⁹ Finally, in three additional areas, Western Europe (8-5.5 KYA), the Indus Valley (9 KYA) and Egypt (8 KYA), agriculture was introduced by importation of “founder” crops from elsewhere and then followed by subsequent domestication¹²⁰ of indigenous wild plants and animals.¹²¹

Agriculture had a huge effect on human demographics. Most notably, it caused a population boom.¹²² Hunter-gatherers occupy the land at very low densities. The animals worth hunting and the plants worth gathering form only a tiny fraction of the total fauna and flora. These low densities mean that hunter-gatherers must range far and wide to find adequate nutrition. Selecting and nurturing those plants and animals most useful for human consumption changes the equation dramatically. An acre of land devoted to food

119. JARED DIAMOND, GUNS, GERMS AND STEEL 98-99 (1997) [hereinafter DIAMOND, GUNS].

120. Domestication means altering the genetic composition of a plant or animal in order to make it more useful to humans. *Id.* at 14.

121. *Id.* at 100. The reason why some areas developed food production early, some late and some never at all (Australia), has little if anything to do with the differences in various human populations and very much to do with the native suite of domesticable plants and animals available to them during the Upper Paleolithic period. Put simply, some plants and animals are simply much better candidates for domestication than others, and those upper Paleolithic hunter-gatherers fortunate enough to find themselves in the vicinity of the wild ancestors of future domesticates benefited from the luck of the draw.

What made a wild plant a good candidate for domestication by upper Paleolithic hunter gatherers? First it would have to be edible in the wild state and give high yields compared to other gathered plants. Ease of cultivation, by mere sowing or planting, was required because humans had no more sophisticated horticultural techniques. A short growing time minimized the delay of the return on the energy investment, and ease of storage provided a steady year-round source of nutrition.

In addition to these obvious physical traits, two more subtle biological properties are crucial. A good potential domesticate should be self pollinating and thus able to breed true and retain its advantage in future generations. Plants lacking this feature might lose their selected genetic advantage in the process of random pollination by less suitable wild specimens. Moreover in an ideal domestication candidate, these six desirable crop traits should be obtainable with relatively little genetic modification of the wild precursor. *See id.* at 124.

Fortunately upper Paleolithic hunter-gatherers did not need to know enough about horticulture to choose consciously the wild plants that had these characteristics. Hunter-gatherers did not gather indiscriminately (if they had they would not have lived long enough to pass on their genes); rather they consciously and unconsciously sought out those plants that were easy to find, nutritious, and required little time and effort to transform into edibility. These they consumed selectively, i.e., largest, sweetest, most resistant to climate extremes, and excreted their seeds nearby, thus increasing the frequency of the choice specimens in the next generation in the wild. Eventually purposeful selective propagation followed, and by evolution and later selective breeding, these wild precursors became the staple crops of the earliest Neolithic farmers. *See id.* at 123-30.

122. GREAT HUMAN DIASPORAS, *supra* note 27, at 134. The population of the world is estimated at 6 million at the dawn of the age of agriculture, and by the year 1700 it had reached 500 million. OLSON, *supra* note 3, at 101.

production (farming and herding) can support 10 to 100 times the population as the same acre devoted to foraging.¹²³

Domestication of animals¹²⁴ increased the advantage by providing protein (flesh and milk) and fertilizer. The use of domesticates for draft and transportation made farming more efficient and made trade and transport much more accessible. A better food supply meant not only greater population density, but more permanent sedentary settlements than hunter-gatherers could support. In a feedback loop, sedentary settlements allowed even greater population densities because women in settled agricultural societies can bear children at about twice the rate as in hunter-gatherer societies, where birth frequency is limited by the ability to carry infants on periodic migrations.¹²⁵

A direct result of food production was surplus, and that development along with fixed settlements permitted food storage, which in turn gave rise to the first specialists.¹²⁶ In hunter-gatherer society, which is largely egalitarian, everyone must forage;¹²⁷ no one can devote substantial time to specialties such as metallurgy, governing or soldiering. Food production economies, by contrast, permit specialization and thus conquest, especially of

123. DIAMOND, GUNS, *supra* note 119, at 88; DIAMOND, THIRD CHIMPANZEE, *supra* note 54, at 189.

124. Chance again played a huge role in determining which societies would domesticate animals and how early in their history that development would occur. Once again, some animals are much better candidates for domestication than others, and the societies located in the natural ranges of these candidates developed herding early.

Again it is important to define domestication. It consists of much more than mere taming. In addition to making the animal tolerate the presence and activities of humans, domestication requires the ability of the animal to reproduce in captivity and the ability of its human captors to breed it selectively. Thus in many cases (the cheetah and the elephant) animals have played a role in human history without ever being truly domesticated; rather each new specimen is caught wild. See DIAMOND, GUNS, *supra* note 119, at 159, 169-70.

In order to be fully useful to human agriculturalists a candidate must be large—large enough to provide enough food to make the project worthwhile and large enough to be useful for draft duties. It must be efficient to feed; otherwise the result of keeping the animal would be a net loss rather than a gain; as a result few carnivores have been bred for food. It must mature quickly because few incipient herders are willing to invest the massive amount of food required to raise an animal to an exploitable age of more than a few years. So much for elephants, which are caught wild and tamed by Indians and Africans, rather than bred in captivity. The animal must be willing to breed in captivity, and that excludes some animals that even modern zoo keepers must inseminate artificially. Another required characteristic is a relatively mild and tractable disposition; bears and hippopotamuses fail spectacularly, but so, less obviously, do zebras and elk. A hyperactive startle reflex also precludes useful domestication of species like gazelles and most types of deer. Finally herding animals with defined dominance hierarchies and overlapping ranges are much better candidates than other herding species. *Id.* at 173.

125. GREAT HUMAN DIASPORAS, *supra* note 27, at 133-34.

126. See OLSON, *supra* note 3, at 103.

127. GREAT HUMAN DIASPORAS, *supra* note 27, at 9.

neighboring hunter-gatherer peoples, who usually are vastly outnumbered and “outgunned.”¹²⁸

Moreover, agricultural and herding societies, not only surpass their non-agricultural neighbors in numbers, organization and metallurgy, they also possess a secret biological weapon. The combination of denser settlements and close association with domesticated animals produces a new suite of human pathogens. Over time, these co-evolve with their animal vectors and human victims so that they become less deadly, and the human populations develop increasing immunities. When immunized agriculturalists encountered previously unexposed hunter-gatherer populations, the latter died in droves by new plagues for which their immune systems were unprepared.¹²⁹

The result of the competitive advantage of food producers over foragers was a vast expansion of the territories of the agriculturalists. Several variables controlled the speed of the expansion. Among the most important was the geographic axis of the land mass because crops move much more easily along east/west axes than along north/south ones.¹³⁰ Toleration of temperature extremes is the most obvious reason for this, but amount of rainfall and its seasonal distribution are also important. The axis issue explains in part why agriculture spread much sooner and more rapidly in Eurasia than in Africa or the Americas.

Another significant question is the quality of the founder crops. When they were good—nutritious and easy to cultivate—the transition from foraging to food production was more rapid. Again, Eurasia, with its high concentration of large-seeded grasses (wheat and barley) and pulses (peas and soybeans) had a large advantage.¹³¹ By contrast, the goosefoot, squash and sunflower native to the Eastern United States were not sufficiently nutritious to warrant abandoning the more profitable foraging the area offered.¹³²

Geographic boundaries had a powerful influence. It is very difficult for cultivation to spread across deserts, high mountains, rain forests and large bodies of water, explaining why Mediterranean crops did not spread across the Sahara into the Sahel and why Asian crops did not spread from Indonesia to Australia.¹³³ A related factor is isolation and fragmentation of useful habitat. Arable land in New Guinea, for instance, is limited to the interior highlands

128. See DIAMOND, GUNS, *supra* note 119, at 112.

129. See *id.* at 211.

130. *Id.*

131. *Id.*

132. *Id.* at 150-51.

133. See *id.* at 186.

(4,000 to 8,000 feet) far from other agricultural societies, and separated from them by environments that are unfriendly to farming.¹³⁴ Similarly, the Andean civilizations with potatoes and llamas were separated from the Mesoamerican farmers, who had corn, by the tropical rainforests of the Central American Isthmus where neither set of domesticates could thrive.¹³⁵ Also related to geography is the size of the interface between two groups of people. In Eurasia, contact was continuous across a vast east-west corridor, which stretched a thousand miles wide, north to south.¹³⁶ By contrast, the connection between New Guinea and Australia consisted of a chain of small islands across the Straits of Torres.¹³⁷

Another important issue is the method of the expansion: One possibility—the demographic model—is that the spread of farming involved a replacement of hunter-gatherer *populations* by agriculturalists.¹³⁸ The opposite hypothesis—the borrowed technology model—is that the spread of farming involved mostly the spread of *technology*, with indigenous hunter-gatherers adopting the food producing techniques of their Neolithic neighbors.¹³⁹ The difference between the two models resolves into a single key question: was the advance of farming a result of the spread of people or ideas?

The demographic model does not require the assumption of vast prehistoric genocidal battles, nor does it exclude it. The most likely scenario is that farmers expanded as population density increased; inevitably the advancing farmers encountered the territory of indigenous hunter-gatherers and took it over, sometimes violently, sometimes not, by weight of numbers and introduced pathogens.

Undoubtedly, this scenario occurred repeatedly during the Neolithic expansions.¹⁴⁰ It almost certainly accounts for the expansion of the Bantu—from an initial homeland in West Africa (Cameroon and Nigeria)—

134. *Id.*

135. *See id.* at 187.

136. *Id.*

137. *See id.* at 315.

138. *See Heartland, supra* note 98, at 5 (stating that marker M172 is the signal of the spread of Neolithic farmers through Europe).

139. *See WELLS, JOURNEY, supra* note 39, at 152.

140. Its most dramatic examples would occur thousands of years later when European agriculturalists displaced Native American, African and Australian hunter-gatherers. In some places where European crops flourished (North America, the Cape area of Africa, and parts of Australia and Tasmania) the replacement was almost total. In other areas where European crops and animals fared less well (Australia and parts of South America), the replacement was less thorough.

eastward and southward nearly to the Cape.¹⁴¹ About 5,000 years ago, their agricultural package, millet, sorghum, African yams and cattle, allowed them to displace or engulf three separate sets of hunter-gatherers. Thus, the Bantu expansion caused a drastic contraction of the ancestral range of the Bushmen and Khoi-khoi (Hottentots) in central, eastern, and southern Africa until it remained only in the far south of Africa, where the Bantu crops did not thrive.¹⁴² The other two populations of hunter-gatherers—the two (genetically separate) groups of Pygmies (Biaka and Mbuti)¹⁴³—became surrounded and concentrated in the rain forest habitats, again unsuited to Bantu food production.¹⁴⁴

Similarly in Asia, agriculture expanded from two centers (millet in the North and rice in the South), and the spread was primarily one of people rather than culture. Chinese rice¹⁴⁵ propagators first displaced the indigenous population of Southeast Asia, who were the descendants of the original coastal migration out of Africa.¹⁴⁶ Then a subgroup, which had settled in Taiwan or Malaysia, began to expand through Indonesia and Malaysia. There, they adapted their temperate, rice-based agricultural package to one more suited to tropical lands¹⁴⁷ (taro, pigs, chicken and dogs), developed an impressive suite of boat-building and navigational skills, and ultimately moved west to colonize Madagascar and east to settle all of the habitable islands of Polynesia.¹⁴⁸

141. For genetic evidence of the expansion, see Stephanie Plaza et al., *Insights into the Western Bantu Dispersal: mtDNA Lineage Analysis in Angola*, 115 HUM. GENETICS 439 (2004); Salas et al., *supra* note 63. See also Onella Semino et al., *Ethiopians and Khoisan Share the Deepest Clades of the Human Y-Chromosome Phylogeny*, 70 AM. J. HUM. GENETICS 265 (2002).

142. DIAMOND, GUNS, *supra* note 119 at 396-97.

143. Unlike the click language speakers, the two groups of pygmies lost their native languages and now speak the languages of their Bantu neighbors. See *supra* note 62.

144. DIAMOND, GUNS, *supra* note 119, at 394.

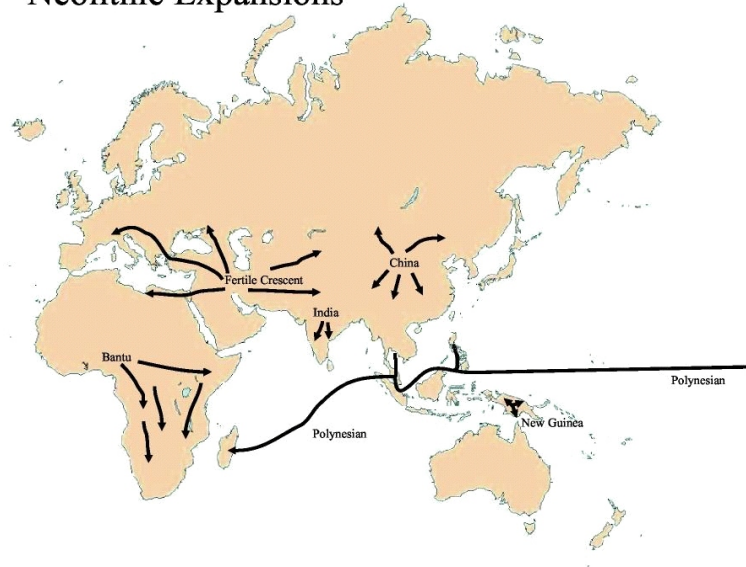
145. WELLS, JOURNEY, *supra* note 39, at 156.

146. They bear the Y-chromosome marker M130. See *supra* notes 71-78 and accompanying text. The remnants of that early migration exist now in Asia as population and linguistic isolates such as the Andaman Islanders and the Negritos of the Philippines. WELLS, JOURNEY, *supra* note 39, at 74-75. Many admixed populations display that marker as do native Australians and New Guineans.

147. WELLS, JOURNEY, *supra* note 39, at 179.

148. Bing Su et al., *Polynesian Origins: Insights from the Y Chromosome*, 97 PROC. NAT'L. ACAD. SCI. 8225 (2000).

Neolithic Expansions*

**Beginning Dates**

Fertile Crescent 11 kya

India 9 kya

China 10 kya

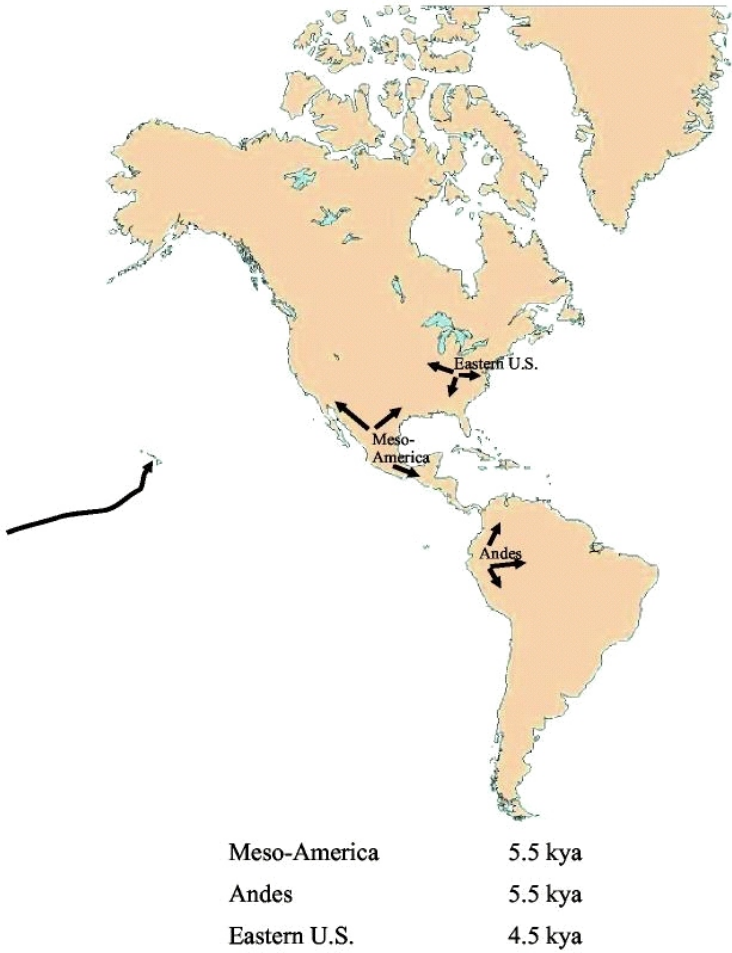
New Guinea 6 kya

Polynesia 6-2 kya

Polynesians to Madagascar 1.5 kya

Bantu 5-2.5 kya

* After Diamond, *Guns*, *supra* note 91.



When agriculture spreads by population replacement, it often leaves a characteristic genetic signature in the resulting population. The signature comes in the form of a mismatch between the genetic composition of the resultant population's mitochondrial DNA and its Y-chromosome DNA. Usually the mixed population will show a much higher representation of indigenous hunter-gatherer genes in its mitochondrial DNA and a higher concentration of agricultural/invaser genes in its Y-chromosome DNA.¹⁴⁹ The explanation is that the replacement of hunter-gatherer males was more complete than that of females. The most plausible reason is that the invaders killed or drove off most of the native men and mated with at least some portion of the women.¹⁵⁰ The pattern occurs also in later population replacements (South American Indians) and also in conquests and enslavements (African-American survivors of the Middle Passage).

The spread of food production occurred differently in Europe. An early hypothesis had been that the expansion in Europe was demographic as it was in Africa and Asia,¹⁵¹ but subsequent genetic research has shown that demographic expansion of agriculture occurred only in the areas very close to the Mediterranean. The subsequent expansion of food production into central and Northern Europe occurred by cultural diffusion, with indigenous hunter-gatherers gradually adopting agriculture and herding. The result is that eighty percent of the current population of Europe traces back to the original invasion by Paleolithic hunter-gatherers, who came forty thousand years ago via Asia, rather than to an invasion of Middle Eastern agriculturalists that occurred thirty thousand years later.¹⁵²

The important point here about the Neolithic expansion of food producers is that it radically reduced the genetic diversity of the world's human population. Those hunter-gatherers that acquired agriculture early on succeeded in spreading their genes over vast areas and in the process replaced countless other less lucky lineages. Concomitantly, language replacement has occurred, thus significantly reducing the total of languages world-wide and leaving some language isolates surrounded by a sea of speakers of a more wide-spread tongue.¹⁵³ To use an appropriate agricultural metaphor, the effect

149. For evidence of this pattern in the Bantu expansion into Southern and Eastern Africa, see Elizabeth T. Wood et al., *Contrasting Patterns of Y Chromosome and mtDNA Variation in Africa: Evidence for Sex-Biased Demographic Processes*, 13 EUR. J. HUM. GENETICS 867, 875 (2005).

150. WELLS, JOURNEY, *supra* note 39, at 113; DIAMOND, GUNS, *supra* note 119, at 291.

151. GREAT HUMAN DIASPORAS, *supra* note 27, at 138-39; Joaquim Fort, Toni Pujol & Luigi Luca Cavalli-Sforza, *Paleolithic Populations and Waves of Advance*, 14 CAMBRIDGE ARCH. J. 53 (2004).

152. WELLS, JOURNEY, *supra* note 39, at 154.

153. WELLS, JOURNEY, *supra* note 39, at 169.

on human populations has been similar to the effect on plant species' distributions. When wild savanna or forest is cleared and replaced by crop land, the effect is a monoculture; the absolute amount of vegetation per acre does not decrease and, in fact, often increases; but the variety of species crashes from a full distribution of native flora to one or, at most, a few of the favored agricultural crops. This point will be highly relevant later on in the discussion of the total amount of genetic variation among human populations.¹⁵⁴

4. *Historical Expansions*

The steady diminution in human genetic diversity only accelerated during historical times. There is space here only for the grossest of developments. Early empire building in Mesopotamia and Egypt had a huge effect. Later conquests spread Hellenic genes throughout the Eastern Mediterranean all the way to India, and Imperial Roman genes as far as Northern Europe and Britain. Religious movements—principally Christianity and Islam—spread genes as well as doctrine. Great waves of conquest from East and Central Asia, the Mongols in particular, also added to the homogenization.¹⁵⁵ Perhaps the most significant development was European colonialism, which wiped out numerous indigenous lineages in America and Australia, and also mixed European genes with the genes of Amerindians, Africans (via the slave trade)¹⁵⁶ and Asians. It also led to the transport of African slaves to the New World, creating African/European and African/Amerindian mixed populations.

On a less momentous scale, nationalism produced consolidations in one European country after another (first France, then Spain and later, Italy and Germany). The result was that a Breton no longer had to marry a “foreigner”

154. See *infra* Part IV.A. An interesting satellite inquiry is whether the lifestyle of the remaining hunter-gatherers changed after the development of agriculture by their neighbors. Are today's remaining foragers remnants of an earlier way of life that was once ubiquitous, or are they occupying niches that did not even exist before their neighbors adopted food production? See Frank W. Marlowe, *Hunter-Gatherers and Human Evolution*, 14 *EVOLUTIONARY ANTHROPOLOGY* 54 (2005).

155. Tatiana Zerjal et al., *The Genetic Legacy of the Mongols*, 72 *AM. J. HUM. GENETICS* 717, 720 (2003) (about 8% of the region's male population and .5% of the world's male population carry the Y-chromosome of Genghis Khan and his male-line descendants).

156. See Antonio Salas et al., *The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade*, 74 *AM. J. HUM. GENETICS* 454, 463-64 (2004). Most of the transported populations (about 60%) came from western Africa with smaller portions coming from west-central Africa (about 30%) and southeastern Africa (less than 10%). The pervasive effect of the Bantu dispersals makes the original genetic homelands of those populations more difficult to pinpoint.

to wed a Marseillaise because both were “French.” Symptomatic of this development has been the steady loss of regional languages and the rise of a relatively few “national” languages. Later the consolidation progressed to super-nations—large amalgams like the USSR—that linked peoples as disparate as Balts, Tadjiks and Chuckchis. More recent advances in transportation and communication, and international cooperation (the EU), as well as globalization, have accelerated the process.¹⁵⁷

IV. GENETICS, MIGRATIONS AND THE CONCEPT OF RACE

A. The Contributions of Molecular Genetics

What, then, does this complex story of past migrations have to say about the concept of race? The short answer is “quite a bit,” but before considering its contribution to the subject, it is useful to reprise briefly the model of race accepted by science before the advent of modern genetic and genomic research. The traditional model divided humanity into a set of discrete, hierarchically-ranked races, distinguished by different physical characteristics, as well as different temperaments and means of government.¹⁵⁸ The eventual result was a six-part taxonomy of Europeans, Americans, Asians, Africans, Australians and Oceanians.¹⁵⁹

In that system, each race consisted of a discrete group of people with certain diagnostic physical characteristics that qualified them uniquely for membership in the group and disqualified those that did not possess the key traits. The races had natural (before 1492) bounded ranges¹⁶⁰ and could be exemplified by pure forms or archetypes (essences),¹⁶¹ with variation from the norm treated as exceptions or admixtures.¹⁶² The folk wisdom based on these

157. WELLS, JOURNEY, *supra* note 39 at 184-87.

158. Braman, *supra* note 7, at 1386 n.25.

159. Tishkoff & Kidd, *supra* note 10.

160. COON, HISTORY, *supra* note 11, at 190.

161. On Platonic types, see Kittles & Weiss, *supra* note 13.

162. This brief exposition is, of course, a gross simplification of a great diversity of thought on the history of the concept of race in America and Europe. For more complete treatment, see George J. Armelagos & Alan H. Goodman, *Race, Racism and Anthropology*, in BUILDING A NEW BIOCULTURAL SYNTHESIS: POLITICAL-ECONOMIC PERSPECTIVES ON HUMAN BIOLOGY 359 (Alan H. Goodman & Thomas L. Leatheman eds., 1998) [hereinafter Armelagos & Goodman]; Braman, *supra* note 7, at 1384-93; Lillquist & Sullivan, *supra* note 11, at 408-26.

systematics was that racial categories reflected “dramatic, underlying, essential differences.”¹⁶³

Early serological studies, conducted by Richard Lewontin and others began to erode the race model, demonstrating that human variation was much more complicated and continuous than a division into a few large groups could reflect. Many concluded as a result that human races were biologically meaningless, merely a social construct. But the conclusion was not universal with some rejecting it as political correctness and others finding it contrary to common sense.¹⁶⁴ Furthermore, some scientists’ empirical¹⁶⁵ and theoretical¹⁶⁶ work seemed to support that skepticism.

The signal accomplishment of the genomic era has been to end that debate by providing sophisticated methods to *quantify* human variation, show its molecular causes, and time its development by retracing the migrational and evolutionary events that produced it. The previous section showed how molecular genetics helps to explain the spread of humanity out of Africa and across the globe; this section describes in greater detail the contributions of the developing body of knowledge to the understanding of race.

The first contribution of genomics is to situate humanity within the animal kingdom. Genetically, humans are apes, and closely related to common chimpanzees and bonobos (sometimes called “pygmy chimpanzees”).¹⁶⁷ In fact we share 98.4 percent of our DNA with the two species of chimps.¹⁶⁸ That is about the same genetic distance as exists between wolves and jackals or between horses and zebras.¹⁶⁹ Moreover, both species of chimps are more closely related to us than either is to gorillas or orangutans.¹⁷⁰ The genetic distances allow reconstruction of the history of the divergence of the primates, showing that monkeys diverged from apes about 25 MYA (million years ago), orangutans from the common ancestor of the other large apes (gorillas, chimps and humans) about 14 MYA, gorillas from

163. Ossorio & Duster, *supra* note 15, at 116. The folk wisdom is alive and well and sometimes rationalized by pseudo science. See sources cited *supra* note 3. See also Armelagos, *Race*, *supra* note 15.

164. Tishkoff & Kidd, *supra* note 10, at S21; Smay & Armelagos, *supra* note 4, at 26.

165. See Armand Marie Leroi, Op-Ed., *A Family Tree in Every Gene*, N.Y. TIMES, Mar. 14, 2005, at A21; Hua Tang et al., *Genetic Structure, Self-Identified Race/Ethnicity, and Confounding in Case-Control Association Studies*, 76 AM. J. HUM. GENETICS 268, 273-74 (2005) [hereinafter Hua Tang]; Rowe, *supra* note 17, at 60; Risch et al., *supra* note 18, at 2007.1.

166. See sources cited *supra* note 3.

167. DIAMOND, GUNS, *supra* note 119, at 36.

168. DIAMOND, THIRD CHIMPANZEE, *supra* note 54, at 23.

169. STRINGER & MCKIE, AFRICAN EXODUS, *supra* note 32, at 20.

170. They are more closely related to each other than to us, but we are only twice as different genetically from them as they are from each other. DIAMOND, THIRD CHIMPANZEE, *supra* note 54, at 24.

chimps and humans about nine MYA, and that we parted company with the chimps only about six MYA.¹⁷¹

The next important contribution is to situate our species along a continuum of species based on the total amount of intra-specific genetic variation in each. Compared to other animal species, humans display very little intra-specific genetic variation. Because genetic distance is a function of time of separation, the youth of our species helps to explain this result. The best estimate is that our species emerged in Africa about 200 KYA (thousand years ago).¹⁷² Comparing that figure with the ages of the other primates as well as those of other species generally makes it clear that our species is *very* young and thus had little time to diversify genetically.¹⁷³

Two additional factors have helped to produce the low genetic variability found among humans. The founding population of the species was very small—probably only a few thousand adults.¹⁷⁴ Moreover, since humanity's initial migrational dispersal, the Neolithic agricultural expansions¹⁷⁵ and subsequent historical events¹⁷⁶ have drastically reduced the total number of extant human lineages. Fewer lineages mean less total diversity. These factors have combined to make humanity very non-diverse compared to other animal species. We are, for instance, four times less diverse than chimpanzees.¹⁷⁷ Indeed, at the molecular level, humans are 99.9% identical.¹⁷⁸

171. *Id.*

172. This view is known as the Recent African Origin Model. An alternative is the increasingly disfavored Multiregional Theory, which holds that homo-erectus and archaic humans migrated out of Africa and colonized other continents where each separate group independently evolved into different races of fully modern homo sapiens. See Cavalli-Sforza & Feldman, *supra* note 34.

173. STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 117.

174. *Id.* at 20; Henry C. Harpending & Elise Eller, *Human Diversity and Its History*, in BIODIVERSITY (M. Kato & N. Takahata eds.) (forthcoming 2006) (manuscript at 2, available at http://www.continuitas.com/harpending_humandiversity.pdf); Henry C. Harpending et al., *Genetic Traces of Ancient Demography*, 95 PROC. NAT'L. ACAD. SCI. 1961, 1961 (1998).

175. See *supra* Part III.B.2.

176. See *supra* Part III.B.3.

177. Ossorio & Duster, *supra* note 15, at 117; See also OLSON, *supra* note 3, at 29 (chimpanzees have more than twice as much variety in their mitochondrial DNA as do all six billion people); see generally Anne Fischer et al., *Evidence for a Complex Demographic History of Chimpanzees*, 21 MOLECULAR BIOLOGY & EVOLUTION 799 (2004) (characterizing patterns of genomic variation in central chimpanzees).

178. See Jorde & Wooding, *supra* note 69, at S28. For more on the homogeneity of human genomes, see E. Ya. Tetushkin, *Genetics and the Origin of Human 'Races'*, 37 RUSSIAN J. GENETICS 853 (2001). Shortly before this article went to press, a research team led by J. Craig Ventner published the entire six-billion-letter sequence of one human, Ventner. See Samuel Levy, Granger Sutton, Pauline C. Ng, Lars Feuk, Aaron L. Halpern et al., *The Diploid Genome Sequence of an Individual Human*, Vol. 5, No. 10 PLoS BIOLOGY 1371 (2007). One of the key findings that emerged from that project is that there is much more genetic variation among humans than had been thought previously. This finding runs counter to the

In other words, the most genetically diverse humans differ from each other in only one in one thousand nucleotides or 0.1%.

The numbers are dry and hard to compare for the scientifically challenged, so an image may help. Consider two different groups: (1) a group of humans composed of a Yanomama Indian (Amazon Basin), an Inupiak Inuit, a Finn, a Basque, a Nigerian, a Biaka Pygmy, a Nepalese Sherpa, a native Hawaiian, an aboriginal Australian and a Han Chinese, and (2) a group of chimpanzees from a single relatively large troop in Central Africa. Even though the chimps look very similar to us and the humans look wildly different, the genetic variation found within the group of chimps will vastly exceed that in found in the group of humans. Despite our extensive geographical range, our myriad languages and cultures, our different skin hues, our considerable differences in size and body shape and features,¹⁷⁹ a comparison of human variation to variation within other species shows unequivocally that humans are all alike.

A third and highly significant contribution of molecular genetics to our understanding of race is its ability to apportion the total amount of human variation between intra-group differences on the one hand and inter-group differences on the other. In other words, it can tell us what portion of the total 0.1 percent of variation is attributable to “race” and what part is not. A consistent finding, first determined by analysis of proteins¹⁸⁰ and recently confirmed at the nucleotide level is that more than 85% of the total variation among humans is found *within any* population or group (racial, ethnic, religious, etc.) of unrelated individuals.¹⁸¹ The genomes of the students in a typical first-year civil procedure class, for instance, probably will represent more than 85% of the total genetic variation—itsself very small—in the entire human species.

The same studies show that the remaining 15% of variation splits between variation by continent of ancestry (a proxy for “race”) and variation by region

results of nearly all previous studies on human genetic variation, from the early work of Lewontin on proteins, *see infra* note 180, to contemporary genomic studies, *see, e.g.*, Jorde & Wooding, *supra* note 69, but the finding has yet to be replicated. Similarly the implications of this finding for intra-ethnic versus inter-ethnic genetic variation have yet to be determined.

179. Physical anthropologists have obsessed over measuring and cataloguing every sort of difference among humans. *See* GOULD, *MISMEASURE*, *supra* note 3; Olson, *supra* note 3, at 179.

180. *See* Richard C. Lewontin, *Confusions About Human Races*, Apr. 20, 2005, <http://raceandgenomics.ssrc.org/Lewontin/>.

181. Ossorio & Duster, *supra* note 15, at 117; Cavalli-Sforza & Feldman, *supra* note 34, at 268-69. *See also* Noah A. Rosenberg et al., *Genetic Structure of Human Populations*, 298 Sci. 2381 (2002) (placing the amount of intra-group variation as high as 95%).

within a continent.¹⁸² Thus, for example, a random sample of Belgian school children likely will contain 85% of the total genetic diversity within the entire human species; adding a similar sample of Romanian children (same continent, different region) will add an additional 7%, and adding a similar sample of Senegalese children (different continent) would supply the remainder.¹⁸³ Of course, it is important to remember that the apportionment percentages are percentages of a very small quantity of total human diversity; multiplying the total amount of human genetic diversity (0.1%) by the fraction attributable to race, (5%) yields .007% as the total *amount* of human genetic diversity attributable to race.

At this point, a clarification is important; the percentage apportionment comparison does not mean that two people from the same continent are necessarily more similar than two people from different continents. The comparison simply sets a limit on the *maximum* genetic difference between a pair of people from the same continental “race” and the *minimum* genetic difference of a pair from different races. Because continent of origin accounts for about 7% of the total possible human genetic difference, the pair from the same continental race typically will display no more than 93% of the total available human genetic diversity. And because continent and region together account for 15% of total human variation, the pair from different continents typically will display at least that 15%. Thus the apportionment percentages are perfectly consistent with an outcome in which the genetic difference between a pair of Europeans, say Tony Blair and Vladimir Putin, equals 90% of the total possible human genetic diversity, while the difference between a Briton and a Ghanan, say Blair and Kofi Anan, equals only 20%. In other words, the percentage of genetic difference attributable to race is perfectly consistent with Blair being vastly more similar genetically to Anan than to Putin.¹⁸⁴ This is really just a complicated way to reiterate the finding that the

182. While there is some agreement on the finding of 15% variation among human groups there is less agreement on the proportion of that variation that can be attributed to differences between continents versus differences between populations within each continent. Compare Ossorio & Duster, *supra* note 15, at 117 (attributing only 5% to inter-continental difference and 10% to intra-continental variation) with Robert J. Sternberg, Elena L. Grigorienco & Kenneth K. Kidd, *Intelligence, Race, and Genetics*, 60 AM. PSYCHOLOGIST 46, 54 (2005) (finding 9% of the variation within continental races and 6-7% between them), and Lewontin, *supra* note 180 (finding the 15% variation about evenly split between inter-continental and intra-continental differences). In the remainder of this article, I use this last estimate; although it is the most conservative, it still provides strong support for the thesis.

183. See Ossorio & Duster, *supra* note 15, at 117; Lewontin, *supra* note 180, at 1; Cavalli-Sforza & Feldman, *supra* note 34, at 269.

184. All humans are 99.9% similar. If the Blair/Putin pair displays 90% of the total human diversity (.1%), they will be 99.910% [$99.9\% + (10\% \text{ of } .1\%)$] similar. If the Blair/Anan pair displays 20% of the

most significant variation in our species is found among individuals *within* any population, not between two different populations.

Another way to evaluate the comparative significance of the apportionment percentages is to consider the concept of sub-species or race in zoological taxonomy. Sub-species and race, used interchangeably by zoological taxonomists, both refer to a “clade”—that is a group of living things with some shared morphology (shape) and phylogeny (evolutionary origin). Other examples of clades are the familiar phylum, class, order, family, genus and species. As the terms suggest, sub-species and race refer to a cladistic division below the species level. An alternative approach to intra-specific variation is not to subdivide a species into discrete units, but instead to treat variations across the species’ geographic range as a gradient or “cline.” Thus brown bears could be grouped into subspecies, for example grizzly and Kodiak, or, instead, the variation could be described as a south/north cline of increasing size. The tendency to prefer clines over clades or vice versa divides taxonomists into groups of their own, “lumpers” and “splitters.”¹⁸⁵

The decision to use a cline as opposed to a clade, though conceptually arbitrary, is practically objective; taxonomists use a convention to determine whether it is justified, or for that matter, worthwhile to use the concept of subspecies or race in accounting for intra-specific variation. If the amount of genetic variation that can be ascribed to subspecies or race is greater than 25% of the total genetic variation in the species, the use of the additional clade is justified; otherwise not.¹⁸⁶ As is apparent, the percentage of human genetic

available human diversity, they will be 99.980% [$99.9\% + (80\% \text{ of } .1\%)$] similar. Thus the Blair/Putin difference ($100\% - .01\% = .99\%$) will be 4.5 times as great as the Blair/Anan difference ($100\% - .02\% = .98\%$); i.e., Blair will be 4.5 times more similar to Anan than to Putin.

Genetic diversity occurs not only in the most visible “racial” traits such as skin color and hair form, but also with a host of “invisible” differences having to do with the production of hormones such as those controlling growth, adrenaline production, height, resistance to tooth decay, dyslexia, resistance to various kinds of cancer, etc. The point is that the genetic variation that gives rise to the visible “racial” markers is only a small part of the total variation. Thus, while Blair is closer to Putin on the visible “racial” traits, he might be closer to Anan on a host of less visible but more significant traits, the distribution of which is not a function of geography. This useful clarification was suggested by David F. Armstrong.

185. On lumpers and splitters, see generally GOULD, *MISMEASURE*, *supra* note 3. The extreme among splitters was a group of taxonomists known as polygenists (literally, “many beginnings”) that attributed separate creation of the several races. The opposite group monogenists believed in a single creation, but arranged races along a hierarchy. Both groups embraced biological essentialism and thus racism. *Id.* at 39.

186. OLSON, *supra* note 3, at 63. The measure is of most interest to conservation biologists who must determine what, if any, unique genetic stock exists in a particular threatened habitat. S.O.Y. Keita et al., *Conceptualizing Human Variation*, 36 *NATURE GENETICS SUPPLEMENT* S17 (2005).

variation apportionable to continental “race” is less than a third of the amount typically used to justify a cladistic treatment (5% compared to 25%). By comparison, the figure for sub-groups of African elephants is 40% and 75% for sub-groups of North American grey wolves.¹⁸⁷ Thus not only do humans exhibit less intra-specific genetic variation than other animals, but also the variation that does exist is less cladistic, less separable into discrete lumps.

A fourth contribution of molecular genetics has been a clarification of the kind of variation that exists among humans. Lawyers are accustomed to thinking essentially—in terms of elements or necessary and sufficient conditions. If conduct is a touching, is harmful or offensive and is not consented to, it is a battery. There is no battery without all the elements and no conjunction of all the elements without a battery. Continental “races” do not work that way. There is no genetic variation—not one—that is *diagnostic* for membership in a particular continental race. In other words, there is no genetic or morphological feature possessed by all members of a particular race and absent in all members of other races. All differences among “races” are differences in the *frequency* with which particular genetic variations appear.¹⁸⁸ Standard examples are the classic Mendelian diseases (i.e., those with a simple genetic cause); sickle cell disease, for instance, appears at much greater frequencies among Africans,¹⁸⁹ and cystic fibrosis, among Europeans. Sinodonty (a particular pattern of tooth shape) occurs at high frequencies in Asia yet only rarely in Europe.¹⁹⁰

But clearly, there are morphological differences among humans,¹⁹¹ and just as clearly they align according to particular geographic patterns. The modal (most common) skin tone among Kenyans is a good deal darker than that of Swedes; the modal hair shape of Mediterraneans is curlier than for Vietnamese; the median height of Mayans is shorter than that of Massai. Nevertheless those physical differences do not warrant a cladisitic division of humans into continental races for several reasons.

First, the premise of any racial typology—that a few visible markers can reveal deep fundamental differences in human morphology, physiology or

187. OLSON, *supra* note 3, at 63.

188. Lillquist & Sullivan, *supra* note 11, at 409-10.

189. *Id.* at 410.

190. See WELLS, JOURNEY, *supra* note 39 at 191-92.

191. OLSON, *supra* note 3, at 49. Some writers have overstated this point and concluded that all human groupings are social constructs. That exaggeration, so obviously belied by observation (pygmies really are short), has produced skepticism about the truth and political motivation for the very different, and factually correct, assertion that continental race is not a necessary or even a useful way to sort human genetic variation.

behavior—is false. The characteristics used to differentiate “races” are largely external appearance—skin tone, hair color and texture, body size and shape. There are also differences in the *frequencies* of certain rare disease-causing alleles. No genetic evidence has emerged, however, either before or after the genomic revolution, that supports the essentialist view that geographic variations (skin color, etc.) can predict *anything* about the intellect or character of particular individuals or populations.¹⁹² Humans vary on those characteristics, to be sure, and some component of that variation is genetic, but it is not geographical. Those fundamental features of the species were formed in the one hundred thousand years of evolution that occurred before humans expanded beyond Africa. Nothing in the subsequent migrational history suggests that crucial differences to the basic blueprint occurred along the way.¹⁹³

Most geographically-sensitive human variation is caused by environmental adaptation. Dark skin, for example, is an advantage in the tropics; small size is helpful in the tropical rainforest where it minimizes volume and thus heat creation while maximizing surface area and thus heat dispersion; and resistance to particular pathogens is adaptive where those pathogens are endemic. But it is readily apparent that these superficial characteristics are not phylogenetic, but rather examples of adaptive evolution.¹⁹⁴ Thus peoples exposed to the same environments often display similar morphology even though they are not especially close genetically (phylogenetically).¹⁹⁵ The parallel process by which evolution solves a

192. See Smay & Armelagos, *supra* note 4, at 23.

193. Tishkoff & Kidd, *supra* note 10, at S24, S25 tbl.1. One commentator suggests that the greater climatic hardships endured by Asian migrants forced them to develop better intellects in order to survive. Africans, who remained in the relatively benign tropics, did not need the extra brain power. RUSHTON, *supra* note 3, at 263. Of course there is no empirical data to support this hypothesis, nor any way of quantifying the difficulties that different environments pose for human settlement. Nor is there any mention of the difficulty that greater population density and resource competition posed for Africans. Finally, if Asians endured climatic hardships, Native Americans presumably endured worse during hundreds of generations in Northern Siberia and Berrengia, but Rushton does not place them at the top of the intellectual pyramid. This is not science. Armelagos, *Race*, *supra* note 15, at 106-07.

194. Other factors also help explain the different appearances of different peoples. For example, sexual selection may explain differences that center on features that have cultural norms of beauty (breast, lips and buttocks shape). DIAMOND, *THIRD CHIMPANZEE*, *supra* note 54, at 120. Chance also plays a role. If a disproportionate percentage of a small founding population possesses a trait, a similarly disproportionate number of the much larger descendant population will possess it also. This may explain the concentration of blonds in far northern Europe and its absence in similar regions in Asia.

195. Pygmies, Negritos and Amazonian Indians all have developed small size but are phylogenetically distant.

particular environmental problem is referred to as convergence,¹⁹⁶ a famous example being the fusiform shape of sharks and dolphins, one a fish and the other a mammal. More subtle examples have fooled taxonomists for decades; for instance, old world vultures descended from hawks, and were for years grouped with new world vultures, who, it turns out, are descended from storks.¹⁹⁷

Next, a cladistic division requires geographical discontinuities, lines where the trait changes. Variation in some species does show such a pattern where inter-group migration and mating is cut off by barriers, often geographical.¹⁹⁸ But, humans have spread widely and for the most part contiguously; further, we are very mobile and frequently have mated across population groups.¹⁹⁹ Our world-wide distribution thus contains very few discontinuities. Instead of grouping into discrete clades, particular traits vary by cline. For instance, in many places there is a low latitude/high latitude cline for skin pigment. In others there are genetic admixture clines; in India, for example, there is a north/south cline of ancestry, caused by an initial invasion of coastal horn-of-Africa migrants followed by later migrations out of Africa through the Middle East and into India from the north.²⁰⁰

The world's population, rather than representing a series of clades, more closely resembles an array of overlapping and crisscrossing clines. Further, the clines are not concordant; a north/south cline for skin pigment will not coincide with a forest/savanna cline for height, or a temperature cline for body mass, or a high/low altitude cline for lung capacity, or an east/west phylogenetic cline for an epitheal fold. Moreover these and other clines are specific to certain locales; the skin pigment cline, so obvious in Africa and Europe is much more subtle in Asia and the Americas and seems to be absent in Australia. The attempt to pigeonhole the clinal pattern of human variation into a discrete set of races forces purely arbitrary decisions about where in the

196. See MAYR, *EVOLUTION IS*, *supra* note 13, at 156.

197. SCOTT WIEDENSAUL, *RAPTORS: THE BIRDS OF PREY* 6-7 (1995).

198. Several species of large mammals have developed sub-speciation because their former ranges have been split by human development. Examples are North American cougars, <http://www.cotf.edu/ete/modules/everglades/FEpanther.html>; gorillas, <http://nationalzoo.si.edu/Animals/Primates/Facts/FactSheets/Gorillas/default.cfm>; and worldwide populations of grey wolves, <http://www.kerwoodwildlife.com/GRAYWOLFSSUBSPECIES.htm>. For a general discussion of isolation mechanisms that can produce speciation, see OLSON, *supra* note 3, at 21-23.

199. For example, there is considerable *supra*-Saharan (especially Moroccan) admixture in some West African populations, including some that eventually ended up in the Americas via the Middle Passage. See generally Alexandra Rosa et al., *MtDNA Profile of West Africa Guineans: Towards a Better Understanding of the Senegambia Region*, 68 *ANNALS HUM. GENETICS* 340 (2004).

200. WELLS, *JOURNEY*, *supra* note 39 at 113, 166.

continua to draw the lines, and that results in different classifiers' "finding" different numbers of races. Some have chosen as few as 3 and others opt for as many as 40.²⁰¹ But in the end, it is a fool's errand; you cannot slice soup. No matter where you make the cuts and no matter how many cuts you make, it is still soup.

Not only does the cladistic system of continental races oversimplify the distribution of human genetic diversity, it also obscures one of the most striking features of that distribution. World-wide sampling consistently shows that the total amount of human genetic diversity is not evenly distributed geographically. Rather, human populations are vastly more diverse in some places than in others. Like other genetic traits, genetic diversity varies by cline; in this case, a cline radiating away from Africa.²⁰² The most genetically diverse populations live in Africa, particularly Southern and Eastern Africa. Outside Africa, diversity decreases proportionally, with the least diverse populations inhabiting the Americas and the oceanic islands of the Pacific.

The migrational history described in Part II.B above, provides an explanation for this world-wide distribution of diversity. The group that made the initial out-of-Africa migration, crossing from the horn of Africa into the Arabian Peninsula, exhibited only a fraction of the total genetic diversity found at the time within African populations. Since all other non-African populations are descendants of that original group of founding migrants, they display only that narrow slice of the total of African diversity,²⁰³ as noted above,²⁰⁴ this diminution in total diversity is called the founder's effect. Moreover, subsequent migrations and founder's events have compounded the diversity-shrinking effect. The initial migrants out of Asia, first into Australia, then into Europe, the Americas and Polynesia similarly contained only a narrow slice of the total human diversity in their Asian parent populations, and so have contributed to the cline of decreasing diversity

201. Lillquist & Sullivan, *supra* note 11, at 420.

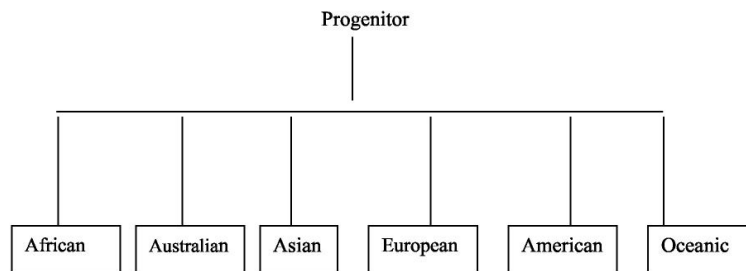
202. Tishkoff & Kidd, *supra* note 10, at S22; Jeffrey Long & Rick A. Kittles, *Human Genetic Diversity and the Nonexistence of Biological Races*, 75 HUM. BIOLOGY 449, 467 (2003); Tishkoff & Verrelli, *supra* note 21, at 305.

203. They do, however, display a component of human diversity not contained in Africa. It is composed of the mutations that occurred during the migrations that peopled the rest of the world. The migrational history of those non-African populations is short (about 60,000 years) compared to the age of African populations, which continued to accumulate mutations during the same period. Thus, while there is a uniquely non-African component of total human diversity, it is much smaller than the diversity component that is uniquely African.

204. See *supra* Part III.B.

radiating away from Africa. Grouping humans into discrete continental races tends to obscure rather than highlight this cline.

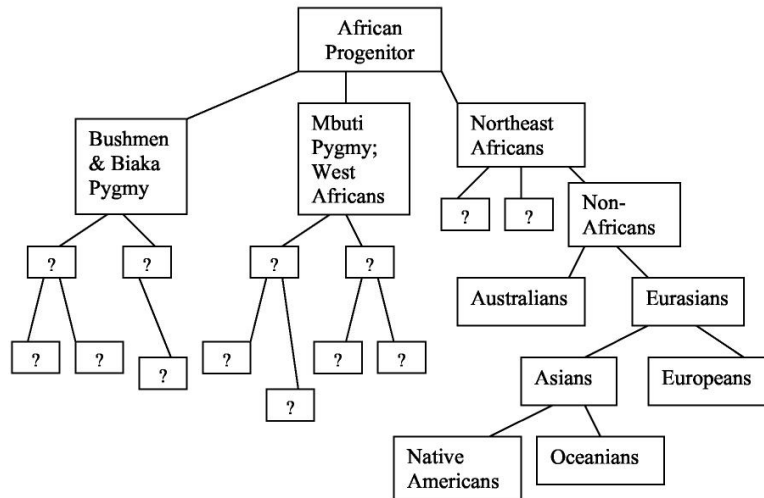
A fifth contribution of molecular genetics to our understanding of human variation is to reveal the phylogeny, or evolutionary history, of that variation. As a consequence it throws into sharp relief the lopsided disproportionality and non-parallelism of the traditional six-race model. That model suggests six parallel genetically equidistant races all at the same phylogenetic level. It could be represented by the following dendrogram.



But this drawing substantially misrepresents the relationships between human populations and the degree of genetic diversity within each. First, the genetic diversity within Africa is greater than that in the entire rest of the world, and the splits between African populations are older and the genetic distances among them correspondingly greater than for other populations. A phylogenetic classification would separate the population of Africa into three major lineages: (1) Bushmen and other click language speakers, and Biaka Pygmies, (2) Mbuti Pygmies and other West Africans, and (3) Northeast Africans. The entire remainder of the world's populations would be grouped simply as descendants of group (3). In other words, there is greater genetic difference between the two small groups of African pygmies than among the combined populations of Europe, Asia, Australia, America and Oceania.

Moreover, the classical division also misrepresents the relationships among the world's non-African populations. The split between Australians and Eurasians occurred prior to the division of the Eurasian population into Europeans and Asians and accordingly involves greater genetic distances. Similarly, the Europe/Asia split predates the subdivision of the ancestral Asian population into modern Asians, Americans and Polynesians and again involves greater genetic distances than exist among those Asian daughter

populations. Thus, the following is a more accurate dendrogram of the genetic relationships of the world's populations.²⁰⁵



The diagram shows that the traditional six-race grouping amplifies small phylogenetic differences and minimizes large ones.²⁰⁶ The categories are not arbitrary and non-parallel; rather, some are subdivisions of others. Moreover,

205. The question marks in the diagram are designed to show that human genetic differentiation did not stop at the identified groups; it continued on the left-hand branches of the diagram much as it did on the right. The vagueness of the descriptions on the left side results from much less study of the substructure and phylogeny of those populations.

While this diagram is more accurate than the six-race model, it is still flawed. First it is very poor in resolution. Compare the much greater resolution in Underhill, *Inferring*, *supra* note 77, at 487, 489 fig.1. Moreover, the diagram purports to represent human *phylogeny* only, not the entire range of human genetic or morphological variation. A diagram more representative of total human genetic variation would have much fuzzier, more overlapping groups and many more crisscrossing lines, revealing the constant history of population intermixture. A Venn diagram could capture the relationship of human genetic variation to continent of ancestry. Imagine six nearly congruent circles (representing the six inhabited continents), each about six inches in diameter. Now suppose that the center points of each of the six fit within a one millimeter square. The huge area common to all six circles would represent the portion of the human genome that *does not vary by continent* of ancestry. The slivers near the outside boundary of the figure would be shaped by their inclusion within one circle but not the others. These slivers would then portray proportionally the amount of human genetic variation that *is accounted for by continent of origin*. Such a diagram would show accurately the contribution of continent of origin to total human genetic diversity, but it would not represent phylogeny, which the dendrogram does.

206. Traditionally, taxonomists used phenotype to classify organisms. Today, however, scientists have the additional tools of genomics and generally accept the primacy of phylogeny (evolutionary history) over phenotype. See generally Keita, *supra* note 186.

there is no reason for making gross distinctions in one clade and fine distinctions in another. Basically, the traditional six-race grouping is ungrammatical—analogueous to dividing mammals into six groups: placental mammals, cats, dogs, Jack Russell terriers, Labrador retrievers and yellow Labrador retrievers.

A final contribution of the genomic era is the insight it grants into the groupings we treat as “race” within the United States, showing that those groupings, while socio-culturally meaningful, are biologically incoherent.²⁰⁷ A corollary benefit is not the solution of an intractable problem but rather the its exposure as a pseudo-problem caused by our muddled use of racial terminology. In the United States, we acknowledge five main “racial groups”: European-Americans, African-Americans, Native Americans, Hispanics or Latinos, and Asian-Americans. As socially constructed entities, these groupings are not especially problematic;²⁰⁸ individuals are capable of self-identifying, and self-identification largely is compatible with identification by others. There are probably no precise definitions for the groups in terms of necessary and sufficient conditions for membership, but society seems to tolerate the imprecision relatively well.

The serious conceptual problem arises from attributing an unwarranted biological significance to these socially-constructed entities. In point of fact, they are, some more and some less, biologically incoherent.²⁰⁹ The biological incoherence of the Hispanic group was obvious long before the insights of molecular genetics. The imprecise socio-cultural criteria for group membership—some combination of one or more Hispanic ancestors, a Hispanic surname, Hispanic cultural identification, current or historical discrimination because of group membership, and use of the Spanish language—clearly are not biological. If the group were confined to the admixed population of European and Native Mezo-Americans, it might have some genetic significance, but the socio-cultural classification includes populations with Afro-Caribbean lineages (e.g., Dominicans, Puerto Ricans); Spanish Europeans, and other Europeans whose parents or grandparents spent a generation or two in South or Central America before coming to the U.S.²¹⁰

207. *Id.* at S18.

208. Multi-ethnic classifications do add an additional layer of complication. It is, of course, only the groupings that are not problematic; no one would claim that there are no social problems generated by and incident to race in the United States.

209. By “incoherent” I do not mean “not understandable” or “confused,” but rather the older literal meaning of the term in which a group is incoherent when its members do not cohere; in this case they do not cohere genetically.

210. Hua Tang, *supra* note 165, at 273; Keita, *supra* note 186, at S19.

Similarly, the socio-cultural term “Asian American” applies to such highly divergent groups as Indonesians, Mongolians, Chinese, Southeast Asians, Koreans, Japanese and Philipinos, all with or without some European admixture. The migrational histories of these groups, however, are very different.²¹¹ Thus, it would be surprising if they shared many substantial and significant phenotypic variations.

Surprisingly, internal diversity among African Americans²¹² is even greater. The bulk of the African American population arrived by the Middle Passage and thus has West and Central African ancestry. The parent West and Central African populations are some of the oldest and most internally diverse on earth.²¹³ The current American population, moreover, varies substantially and regionally in degree of European admixture, with an average of about 30%.²¹⁴ In addition to this diversity, the rule of hypo-descent, known colloquially as the “one-drop” rule, likely adds many more people who self-identify as African American but whose African genetic heritage is minimal.²¹⁵

Native Americans are the most homogeneous of the peoples of the United States because the settling of the Americas by Paleo-Siberians is the most recent of the Paleolithic human migrations, occurring about 15 KYA, and because the founding population was very small. However, the current population of Native Americans is very small, and intermarriage with African-Americans and European-Americans has been very frequent, so the resultant census group contains quite a bit of internal diversity imported from those other groups.²¹⁶

And what of the “majority” European-American population? While Europeans are more homogeneous than most other Old World populations,²¹⁷ they still represent an admixture, albeit one that is much older than the more

211. See *supra* Part III.B.2.

212. Ossorio & Duster, *supra* note 15, at 18; Keita, *supra* note 186, at S19; *cf.* Hua Tang, *supra* note 165, at 268 (continuous range of European admixture averaging 10-20%).

213. See *generally* Rosa, *supra* note 199 (detailing the many lineages that make up the population of one of many West African regions).

214. That is a nationwide average; but admixture differs substantially by region, averaging 10% in the South and 50% in the North. There is also some Native American and Hispanic admixture. GENES, PEOPLES, AND LANGUAGES, *supra* note 33, at 74-75.

215. Historically it was advantageous for some such people to violate the “one-drop” rule and “pass” as white.

216. The criteria for membership in the socio-cultural category is fuzzy and is determined by each tribe. Some tribes require only one third ancestry. See Race (United States Census), http://en.wikipedia.org/wiki/Native_American_%28U.S._Census%29.

217. GENES, PEOPLES AND LANGUAGES, *supra* note 33, at 23.

recent historical mixture of African-Americans. Recall from the migrational history that Europe is the most recently settled of the Old World continents; its population must have come from somewhere else, with the only available candidates being Asia and Africa. While there have been multiple Neolithic and historical migrations of Middle Easterners and Central Asians into Europe, the principal mixture occurred even earlier (about 40 KYA). As a result of that admixture, Europeans appear to be about one third African and two thirds Asian.²¹⁸

The key point is to understand the consequence of the genetic incoherence of the “race” divisions within the United States. Start with the well-accepted finding that humans are all remarkably similar genetically. Add a system of classification that is largely biologically incoherent, and it is easy to conclude that it is time to stop searching for genetic explanations for inter-group differences on major elements of physiology or behavior. Of course, there will be differences, many due to environmental factors that differ for the various groups. And there will be genetic variations, most dealing with different frequencies of rare (disease-causing) alleles. But the search for major, biological group-wise differences is a fool’s errand when the groups lack biological coherence at the required order of magnitude. Thus a significant contribution of molecular genetics is not to resolve but to dissolve the question whether there are genetic, group-wide differences in IQ. It does not make sense to ask a genetic question of a genetically incoherent typology.

B. The Cluster Ruckus

1. Clusters

Biological essentialism has the tenacity of original sin; to quote a James Bond villain, it “appear[s] with the tedious inevitability of an unloved season.”²¹⁹ But how could a concept of race that is rooted in the pre-scientific Great Chain of Being²²⁰ persist in the post-genomic era? Its toe-hold on survival consists of a persistent finding of continental clustering in almost all sampling studies. Allele frequency studies assess the frequencies with which certain genetic variations appear in different populations. Their results show

218. *Id.* at 76.

219. The villain is Hugo Drax; the movie is *Moonraker* (1979); Drax is addressing James Bond, and referring to Bond’s repeated efforts to thwart Drax’s plans.

220. See *supra* note 7 and accompanying text.

consistently that populations “cluster” by continent.²²¹ In other words, populations from Asia display one set of frequencies of certain genetic markers while populations from Europe display another. Furthermore, people, as well as populations, cluster; sorting a random sample of people by computer²²² into groups based on genetic similarity usually results in 4-6 clusters that align more or less with the continents. Finally, the same studies show strong correspondence between a subject’s self-identified race or ethnicity and the continental cluster into which he or she fits. In other words, people, using their common sense assessment of their own race or ethnicity, usually group themselves in the same continental cluster that the computer program does.²²³

The cluster findings are widely replicated and not subject to serious dispute. The remaining question then is: what do they mean? The essentialist response is that they reinvigorate the traditional cladisitic view that humanity divides naturally into a few, distinct, homogeneous, continentally-bounded groups, and that this grouping correctly captures some underlying biological reality. Thus this neo-essentialist position represents a fundamental challenge to the current consensus that race is merely a social construct; it relies on the robust finding of continental clustering to support the conclusion that the “social construct” model, however well-intentioned, is simply incorrect.²²⁴

2. Sampling Problems

There are two main flaws in this reasoning, one is methodological, the other conceptual. The methodological flaw concerns two features of the design of the clustering studies: (1) populations sampled; (2) the number of clusters selected for the computer query.²²⁵ The populations sampled in a

221. Lewontin, *supra* note 180, ¶ 6; Tishkoff & Kidd, *supra* note 10, at S26 (37 populations studied using over 80 independent genome loci); Hua Tang, *supra* note 165 (10,527 participants grouped using 326 genetic markers).

222. The program, known as “Structure,” is succinctly described by Tishkoff & Kidd, *supra* note 10, at S25.

223. See Hua Tang, *supra* note 165; Joanna L. Mountain & Neil Risch, *Assessing Genetic Contributions to Phenotypic Differences Among ‘Racial’ and ‘Ethnic’ Groups*, 36 NATURE GENETICS SUPPLEMENT S48, S48 (2004); Risch et al., *supra* note 18, at 2007.

224. Leroi, *supra* note 165 (“[O]ne thing was clear; the consensus about social constructs was unraveling.”); Risch et al., *supra* note 18; Rowe, *supra* note 17.

225. There is yet another variable that may affect the extent of continental clustering observed in these studies: the genetic markers selected for study. When studies are limited to coding regions of the DNA (those that have a recognizable effect on the organism’s structure and function), continental clustering is not very pronounced; when instead the selected markers are non-coding—especially those that show the

study will have a substantial effect on the degree of “clustering” in the results.²²⁶ Sampling in the United States often centers on three main groups, African Americans, European Americans and Asian Americans.²²⁷ These are groups whose genetic homes are discontinuous, in fact very widely separated, so it would be odd if the results did not show discrete bundles.²²⁸ The skewed sampling is not a result of a grand conspiracy to produce discrete clusters, but rather stems from an availability heuristic. It is simply easier and less expensive to acquire samples from groups that are already conveniently at hand or whose genetic material can be obtained from existing cell and tissue repositories.²²⁹ Just as we look for our keys under the lamppost, we obtain genetic samples from accessible areas and groups.

But different sampling strategies will produce considerably different clustering patterns. For instance, studies that include more samples from geographically intermediate areas (Ethiopia²³⁰ and other parts of Northeast Africa, the Middle East, India, particularly its Southern states, Central Asia, Sicily, Sardinia and Portugal) would produce genetic plots that were less tidy and isolated. Instead of resembling separate clusters, the results would display continuous frequency variations arranged along geographic gradients.²³¹ To see the point clearly, imagine an experiment plotting populations composed of shades of colors of the spectrum. If the population samples all came from

highest inter-population variability—clustering is more evident. See Tishkoff & Kidd, *supra* note 10, at S24; JOSEPH L. GRAVES, JR., THE RACE MYTH: WHY WE PRETEND RACE EXISTS IN AMERICA 112-17 (2004) [hereinafter GRAVES, MYTH]; Lewontin, *supra* note 180, ¶ 7; Kittles & Weiss, *supra* note 13, at 48-53.

226. See GRAVES, MYTH, *supra* note 225, at 112-17.

227. Joseph L. Graves, Jr., *What We Know and What We Don't Know: Human Genetic Variation and the Social Construction of Race*, ¶ 12, Apr. 25, 2005, <http://raceandgenomics.ssrc.org/Graves/Variation> [hereinafter Graves, *Variation*].

228. See Jorde & Wooding, *supra* note 69, at S29; Tishkoff & Kidd, *supra* note 10, at S25; Kittles & Weiss, *supra* note 13, at 38.

229. Troy Duster, *Race and Reification in Science*, 307 SCI. 1050, 1050-51 (2005).

230. Note, for instance, the close genetic relationship of the populations of Northeast Ethiopia and the portion of Yemen just across the Red Sea at Bab-el-Mandeb. Kivisild, *supra* note 69.

231. See Kittles & Weiss, *supra* note 13, at 38; Tishkoff & Kidd, *supra* note 10, at S25; Jorde & Wooding, *supra* note 69, at S29. At least some evidence for this prediction comes from studies that show that Middle Easterners and Central Asians had partial membership in multiple clusters. Tishkoff & Kidd, *supra* note 10, at S25. Strangely, few studies sample Indians, prompting one pair of commentators to ask, “Who are those pesky billion? One race? A mix of the already-sampled races? A multiplicity of races as has often been suggested?,” Kittles & Weiss, *supra* note 13, at 38. Also artificially exaggerating the clustering effect is the inclusion of only one or two African populations, in spite of the well-recognized finding that intra-African genetic diversity exceeds that of all the rest of the world’s populations combined. Charles N. Rotimi, *Are Medical and Nonmedical Uses of Large-Scale Genomic Markers Conflating Genetics and ‘Race’?*, 36 NATURE GENETICS SUPPLEMENT S43, S44 (2004).

among the reds, the greens and the violets, the plotted results naturally would fall into three discontinuous groups; if instead samples are added from among the oranges and yellows, and the blues and indigos, the result will be the familiar roaygbiv spectrum—a continuous variation of shades arranged along a gradient of light wavelength.

The choice of the computer query also affects the congruence between the statistical clusters and the traditional continental “races.” The program used to generate the clusters separates the samples into some number of groups (the particular number represented by the symbol K); it is up to the investigator to specify the value of K , and the choice is largely arbitrary. For some values of K , the correspondence between the clusters generated and the continental races is good, for others it is not. Thus in one study, specifying K as 5 yielded a group of clusters that corresponded fairly well to the standard continental races; however, at $K = 2$, the clusters were not so intuitive, one being anchored in Africa and the other in the (pre-Columbian) Americas; and at $K = 6$, the clusters separated out the five continental races, but the sixth group was composed of individuals from the isolated Kalash population of Northwest Pakistan.²³² The point is simply that relatively arbitrary choices have substantial effects on the robustness of the cluster-“race” relationship; when those choices reflect a preexisting conception of a species divided into a specific set of sub-groups, it is not all that surprising that the resulting clusters correspond with the preconceived groups.²³³

3. *Clusters and Races*

The argument from the clustering results to the reification of essentialist races has a conceptual flaw that is far more serious than the methodological quibbles with the clustering studies. The problem is that the argument equates clusters with traditional essentialist races, and they are very different concepts. Part IV.A, above, explained the common conception of race as a few, large,

232. Rosenberg et al., *supra* note 181, at 2382.

233. Jonathon Marks summarizes the methodological difficulties associated with using cluster studies to establish the reality of races:

[A] cluster analysis . . . is sensitive to the population samples chosen, the individual people representing them, the demographic history of the populations, the assumptions of the particular algorithm, and the patterns of contact among the populations. In other words, the species still doesn't come prepackaged . . . ; you still have to decide, given the fact of difference, how much and what kind is meaningful and how much and what kind is not.

Jonathan Marks, *The Realities of Races*, ¶ 14, Apr. 20, 2005, <http://raceandgenomics.ssrc.org/Marks/>.

discrete, geographically bounded groups, each defined by common diagnostic genetic traits that are absent in non-members.

In order to see why the continental clusters are not races, it is useful to consider exactly what the clusters are. There are two types: clusters of populations and clusters of people. A continental cluster of populations is a group of populations that have similar frequencies for several genetic markers, frequencies that distinguish them from populations that fit in the other clusters. For example, suppose the following groups have the varying percentages of members who display the marker V.

- Albanians – 29%,
- Germans – 26%
- French – 21%
- British – 20%
- Bushman – 41%
- Ethiopians – 44%
- Ghanaians – 49%
- Kenyans – 48%
- Thais – 70%
- Vietnamese – 74%
- Burmese – 71%
- Han Chinese – 79%

The frequency values seem to separate into three different groups with different frequency ranges: 20%-30%, 40%-50%, and 70%-80%. Thus, these values might well contribute to a finding that three clusters exist, one based on each continent.

The continental clustering studies rely, of course, on many different markers with similar frequency distribution patterns. Nevertheless, the studies do not justify the conclusion that “races” exist. In order to see why this is true, it is crucial to understand that the studies do not predict that the clustering pattern would persist for all markers, a majority of markers, or even a decent size minority of markers. They show only that a continental grouping of populations accounts for more of the inter-population variation in marker frequency than does any other small set of large groups. Note further that only a very small minority of markers varies in frequency by population; thus, the continental clustering reveals very little about the genetic substructure of the species.

The studies also reveal clusters of people as well as clusters of populations. A person is assigned to a cluster if her population-differentiating markers fit the frequency pattern of one continental cluster better than that of

any other.²³⁴ Thus, suppose that Mathilda has markers L, M, N, O, and P and that the populations of the continents display the following frequencies of those markers:

	Australia	Americas	Africa
L	40%	10%	2%
M	80%	40%	40%
N	15%	1%	1%
O	20%	10%	15%
P	50%	2%	30%

Mathilda is likely an Australian because she has markers L, M, N, O, and P, which are more prevalent on that continent than on the others. It is possible that she is an American or an African, but the *probability* is that she is Australian. If the same result pattern obtained for 50 or 500 markers instead of only five, the probability of non-Australian ancestry would be very low. Note that the fact that Mathilda's membership in the Australian cluster does not mean that she will align with Australians on every marker; indeed, she might have marker Q, which occurs much less frequently in Australia than elsewhere.²³⁵ Thus, while it is possible to assign Mathilda to the Australian cluster by examining enough of her markers, it is not possible to predict any single marker of hers from her membership in that cluster.²³⁶ All that her membership in the cluster means is that Australian frequencies fit her pattern of population-differentiating markers better than any other cluster.²³⁷

The examples show the most crucial way in which clusters differ from races; in the traditional essentialist model, geographically distributed traits vary concordantly; so in the first example, the fact that the populations divide continentally on marker V would mean that they also divide that way on every other marker, or at least every marker that shows geographic variation.²³⁸ In the second example, the fact that Mathilda fits in the Australian cluster would permit many confident predictions about the specific markers she will display at other geographically variable sites. In fact, however, neither proposition is

234. See Keita, *supra* note 186, at S19.

235. The tendency to assume that the markers cluster concordantly relies on the fallacy that because differences are genetic that they also are pan-racial. Armelagos & Goodman, *supra* note 162, at 369.

236. Christen Brownlee, *Code of Many Colors: Can Researchers See Race in the Genome?*, 167 SCI. NEWS 232 (2005), available at <http://www.sciencenews.org/articles/20050416/bob.asp> ¶ 30 (2005) (paraphrasing Rosenberg et al., *supra* note 181).

237. Indeed, studies often do not assign an individual to a cluster absolutely; rather, they affix a probability of, say 90%, to membership in the assigned group and some smaller percentage probability of membership in one or more of the other groups. Jorde & Wooding, *supra* note 69, at S32.

238. See text following note 202, *supra*.

true. As indicated in Part III.A, geographical variation is often non-concordant, and an individual's membership in a cluster permits only imprecise and tentative predictions about which geographically variable markers she will display and little if anything about the rest of her genome.

The empirical clusters also differ from socially-constructed races in the predictable ways that statistical entities differ from ideal ones. Thus there can be no archetype of a cluster, as there are for races, nor can there be a "pure specimen" of a cluster. Each member of a cluster represents a unique node in an array of many different crisscrossing genetic gradients. We could choose the geographical center of some gradient, but it probably would not be the center of others. Thus, it would be completely arbitrary to privilege one intersection over all the others as "pure" or "unmixed." The notion of purity or archetype makes no sense in the empirical universe in which clusters exist. Finally, compared to the classical races, the continental clusters have indistinct, fuzzy boundaries. There are no major genetic discontinuities;²³⁹ in some places, gradients are much steeper than others, but all variation is continual rather than distinct and interrupted.

If the clusters are not races, what are they? They are a statistical summary of the gross spatial distribution of human genetic variation. Hence they serve as a reminder that some portion of human intra-specific variation is geographical, but we knew that already. A persistent result in protein studies, later confirmed by research at the nucleotide level, is that geography accounts for about 15% of the total variation among humans, and that continental divisions account for about half of that fifteen percent.²⁴⁰

Although the clusters reflect some of the geographically-linked variation in our species, they are not the most informative conceptual apparatus for summarizing that variation. The most dominant geographic pattern of human genetic variation is known as "isolation by distance."²⁴¹ Humans do not mate randomly; people who live close to each other are much more likely to mate than those who live far apart. Similarly, people who share a common language inter-mate more frequently than those who do not. Also at work is the phenomenon of assortive mating, (the tendency of people to find mates similar to themselves), which explains why married couples often resemble

239. See Jorde & Wooding, *supra* note 69, at S30 (no clear boundaries); Long & Kittles, *supra* note 202, at 468; Rotimi, *supra* note 231, at S44.

240. See *supra* Part III.A.

241. For an explanation of the phenomenon, see Tishkoff & Kidd, *supra* note 10, at S23; Graves, *Variation*, *supra* note 227, ¶ 17.

each other. Finally, certain geographic, ethnic, and religious populations practice endogamy (a requirement to mate within the group).

The result of hundreds and thousands of generations mating in this way is that humans are more similar genetically to their neighbors than to distant populations, producing a positive correlation between genetic and geographical distance. Moreover, isolation by distance does not work in a single direction; unless restricted by geographical barriers, it works in all directions. Finally, most geographically variable traits vary by degree rather than discretely; thus the enumerable shades of skin and hair hue, fading one into another. The result of these processes, when combined with geographically-mediated natural and sexual selection, is that geographically sensitive genetic variation is arranged in a complicated pattern of overlapping and crisscrossing non-concordant clines.²⁴²

A small group of continental clusters is not a very efficient or informative way to represent such a pattern of variation; a system of gradients does a far better job. Consider the analogy of elevation above and below sea level. The world's locales could be divided into a series of discrete groups: deep depressions, sea level plains, elevations between sea level and 3,000 feet, elevations between 3,000 and 10,000 feet, and elevations above 10,000 feet, but a topographical relief map (a pictorial display of gradients) would reveal much more information about the relationship between geographic location and elevation.

So continental clusters are not equivalent to traditional essentialist races, nor are they particularly informative constructs for plotting human genetic diversity. Why then are they so interesting? Perhaps we have been asking the wrong question about clusters; what we really want to know about them is not what they are, but rather what their significance is in the seemingly endless debate on the biological reality of race? The clusters fascinate and have provoked so much discussion because they are the closest analogues to traditional races for which there is *any* biological support; “race-lite,” if you will. Race is a social construct, but the clusters are real *biologically*; their “reality” consists in their ability to capture, accurately but at low resolution, about half of the 15% of human genetic variation that can be attributed to geography.

If the socio-cultural concept of race had not played so momentous a role in human history, the clusters would still “exist,” but would anyone care? After all, they are simply one way—and not an especially elegant or

242. See *supra* Part III.A.

informative way—to describe the relationship between geography and human genetic variation. Much better, more descriptive, and more quantitative ways exist. The clusters derive their cache from two features: their indisputable statistical reality and their superficial resemblance to race. For those who, because of politics or merely long intellectual habit, cannot break completely from biological essentialism, they represent a last remnant that is not embarrassed completely by current science.

IMPLICATIONS: LAW, POLICY AND WORLD VIEW

What, then, are the implications of this brief discussion of molecular genetics? What legal or policy results does it change? It has a few practical consequences for some relatively minor legal/scientific issues²⁴³ and one major implication for the central problem of distributive justice that currently faces the world's wealthy industrialized societies. Simply stated, that question is: to what extent is it morally compelled or justified, and practically wise, to undertake programs of wealth re-distribution within and between the world's populations. Do programs like affirmative action, designed to deal with intra-societal disparities, and foreign aid, designed to deal with world-wide disparities, make sense?

The essentialist position is that such programs, well-intentioned or not, are destined to fail. Fundamental genetic differences among human populations explain the different fortunes of nations and the population groups within them. Remedial efforts, even if justified, are hopeless. Moreover, the position is not a straw man; scientific racism is alive and well in the work of J. Phillippe Rushton, Arthur Jensen, Hernstein and Murray and Michael Levin. It has the tenacity of original sin. Nowadays it surfaces more in sly hints than in straight-forward argument. On the international scene: "Africa is hopeless!" Within our country: "They just aren't equipped to benefit from these programs (affirmative action)."

The genomic revolution provides a conclusive refutation. Its most basic message is that the essentialist explanation is nonsense; disparities in wealth, opportunity, education medical care—general life chances, in other words—among various populations in our country and across the globe have

243. It has spawned a debate in the medical and pharmacological communities, over whether the geographically-mediated differences in disease frequency and drug susceptibility within our species are sufficient to justify medical treatment and pharmacological research tailored to particular groups. It also affects the way forensic anthropologists go about the task of identifying unknown bodies and the way genetic identity evidence can be used in criminal cases.

nothing to do with the genetic differences among those populations. *There simply are no genetically-based pan-“racial” differences in character, intelligence or any other set of traits crucial to individual or societal success or position.* Without a specious genetic argument to rely on explicitly or hint at slyly, policymakers will have to confront the purely contingent causes of the disparities.

Even if we can agree that the reasons for the different fortunes of different groups are socio-cultural rather than biological, that consensus still does not end the moral debate about racial justice. The crucial socio-cultural factors might be, on the one hand, the continuing oppression (as well as its residual consequences) of some groups by others or, on the other hand, the cultural failings of the disadvantaged groups. Genetics, even the refined techniques of contemporary molecular genetics, cannot resolve that dispute. Systematic research by the social scientists might shed some light on the question, but in the final analysis, fault is somewhat beside the point; the remedy is what matters. If there are no specious genetic explanations to fall back on, the remedy must lie in some set of policy initiatives designed either to eliminate the oppression and its consequences, on the one hand, or the cultural failure, on the other. Wise policymakers might well choose to work the problem in both ways. Perhaps eliminating the distraction of essentialist explanations for racial disparities will improve the prospects for success.

In addition to its implications on the issue of distributive justice, awareness of the current state of molecular population genetics has a more general re-conceptualizing effect. In that sense, it doesn't change anything; it changes *everything*. Who are we? Where do we come from? How are we different from the animals, from each other? These are fundamental questions of the human condition. The answers to those questions thus affect the vantage point from which we perceive all socio-ethical questions. Other writers have explained some of the contributions of molecular genetics to our understanding of specific legal issues (immutability,²⁴⁴ identification, medical screening,²⁴⁵ ancestral remains) but they have been too modest. By tying the implications of the field to particular legal problems, they have undersold its massive re-conceptualizing effect. I will leave to them and others the task of exploring the range of specific legal problems implicated by the new research. I wish to concentrate, instead, on the fundamental shift it requires in our common concept of what it means to be human, of what our species is like.

244. Braman, *supra* note 7.

245. Lillquist & Sullivan, *supra* note 11.

Our notions of history, prehistory, science, progress, war, genocide are all implicated. Lawyers deal with race, and human variation generally, in a wide variety of contexts, and, if I am right about the re-conceptualizing effect of the advances in molecular population genetics, they will change lawyers' view of what race is and thus how it should function in many of those contexts. More importantly, it should affect the way policymakers view the vast disparities in wealth, health, and life chances that exist within our nation and among populations around the world.

Yet isn't this fundamental transformation of our view of human variation old news? Forty years ago, long before the genomic era, we knew the basic story of human variation, dispersals, and race. Indeed, many of us grew up with the article of faith that "race is merely a social construct." What, then, does a brief exposition of molecular genetics add? My view is that it removes the politics from the biology-of-race debate; it's not about ethnocentrism versus multiculturalism or racialism versus inclusion. Instead it's about science, demonstrable, provable, quantitative propositions about how and how much humans vary and how that variation came to be. That knowledge cannot fix the socio-cultural problem of race; it can, however, end the debate about the biology; furthermore, it can serve as a kind of intellectual inoculation against the revival of essentialist thought in the legal or popular literature. Inevitably, some time in the near future, a piece of essentialist science—complete with pseudo-biological jargon—that identifies a new trait Z among human populations will emerge. The claim will be that existence or degree of Z in some populations compared to others can reveal deep fundamental differences among human groups and that those differences are responsible for the different fates of societies and individuals. But even the minimal dose of molecular genetics supplied here can immunize against that sort of sophistry because it shows conclusively that there simply are no deep fundamental differences to reveal.²⁴⁶

246. What then has driven a socio-cultural construction of race that is so far removed from any biologically justified model? One of many reasons (the one important to the theme of this article) is historical; modernity in science coincided with the era of European exploration and colonialization. Journeys of thousands of miles by ship gave the explorers and their accompanying naturalists a very different view of human variation than previous forays conducted over land. Land travelers see the country side between their homes and their destinations in detail and at close resolution; human variation appears continuous and clinal. Populations vary, of course, but the pattern of that variation appears as a gradient, not as a set of discontinuities separated by clear boundaries. By contrast long ocean voyages do not reveal clines of human variation; they drop the traveler into a completely different geographical environment populated by humans who look very different indeed. That artifact of modern travel and exploration goes a long way toward explaining the tendency of Europeans to see humanity as set of discrete and

fundamentally distinct clades. *See* STRINGER & MCKIE, *AFRICAN EXODUS*, *supra* note 32, at 191-92.

At the same time political events encouraged racial essentialism. The enlightenment, the rise of capitalism, egalitarianism and meritocracy led Jefferson, among others, to proclaim all men equal and equally possessed of natural rights. That noble sentiment did not fit in the same moral universe created by European colonization in Asia, Africa and the Americas and the development of a new and especially brutal form of chattel slavery in America. Racial essentialism was an ideal doctrine to cure the psychic dissonance. If there were fundamental differences in talent and character among groups, then even titular egalitarians could justify different treatment. Thus race, while biologically tenuous, became culturally constructed and entrenched.