

Interpreting African Genetic Diversity

Rick Kittles¹ and S. O. Y. Keita²

Explanations of human biological variation in extant African populations have historically been shaped by a racial paradigm. In this paradigm deep genetic differences are assumed to exist between so-called "racial" groups or types, which are viewed as being composed of nearly uniform individuals and are taken as "natural" fundamental taxonomic units of *Homo sapiens*. These types were originally based on specific aspects of external morphology. Carleton Coon (1962), a physical anthropologist who contributed to this paradigm, proposed a theory of human evolution which postulated the independent emergence of five primordial races or subspecies from distinct, separate, *Homo erectus* ancestors. Two of these purported primordial races originated in Africa. They were called Congoid ("Pygmies") and Capoid (Khoisan or "Bushmen"). In Coon's conception other continental Africans, which exhibit a broad range of biological diversity, were considered to be primarily or only the result of various levels of admixture between these two groups with each other or with immigrants deemed to be of European or Asian origin and generally called "Caucasian." Although the causal aspects of Coon's theory have been rejected, racial thinking still persists today using Coon's categories and is quite evident in the methods and interpretation of genetic studies of African populations (Keita and Kittles, 1997). Here we briefly discuss the implications of modern genetic studies of African and world populations for understanding African bio-history. We propose an evolutionary biogeographical model as an alternative to "racial" explanatory models of African biological diversity. A rational definition of African in a biological sense should be derived from biogeography and not be based on uninformed (and biased) traditions of practice.

Racial models, as traditionally presented, are static, and obligate one to postulate gene flow as the primary explanation for variation. The racial paradigm is manifested in the early writings on Africa by Seligman (1930) and Coon (1962, 1965) and is still operationalized in varying degrees by anthropologists and geneticists, who otherwise reject Coon's thinking, to explain the vast biological

¹Department of Microbiology, College of Medicine, Howard University, Washington, DC 20059.

²Field Museum of Chicago, Chicago, Illinois 60605.

diversity observed in the continent. While it is a formidable task to evaluate the extent of genetic diversity in Africa, it is quite evident that in many of these studies there is the lack of an evolutionary perspective. This is especially the case when northern Africa and the Horn are considered. The variation in these regions has generally been explained as being due primarily to admixture between "Africans" and "non-Africans." In fact, supra-Saharan African populations have frequently been conceptualized as being *derived* from "European" ancestors and, hence, non-African (see Seligman, 1930). The genetic profiles of supra-Saharan populations are indeed in a relative sense "intermediate" to those of various sub-Saharan groups (stereotypically defined) and "Eurasians" (see Guglielmino *et al.*, 1987; Chibani *et al.*, 1985; Tartaglia *et al.*, 1996; Merghoub *et al.*, 1997). But is this genetic "intermediateness" due primarily to supra-Saharan populations being foundationally an admixed group (the result of gene flow between two distinct "races")? Or is their biology reflective of factors acting on indigenous modern *Homo sapien* populations in Africa?

Invoking admixture to explain human variation is related to another polemic generally seen in many genetic studies of African populations. This is the persistence of the socially constructed normative view of the African as only the "Forest Negro" type, or the so-called "Pygmy" from Central Africa, in line with Coon's (1962, 1965) thinking. These populations are used as the representative African in many studies (see Bowcock *et al.*, 1991, 1994; Cavalli-Sforza *et al.*, 1994). Northern African populations are rarely used in research under the designation "African." Also, studies which utilize pooled samples of individuals from distinct African populations as a representative "race" are problematic. If modern humans evolved in Africa 150,000 to 200,000 years ago, and dispersed outside of Africa only about 100,000 years ago, why should it be assumed that those populations which remained in Africa became stagnant and homogeneous? While African populations share a common ancestry, they also have evolved separately for a long time. Evidence for isolation by distance operating in Africa is seen by genetic distinctions among eastern, western, northern, central, and southern Africa (Excoffier *et al.*, 1987). These genetic differences broadly correlate with geographic distances. Various populations in Africa have interacted via migrations during past history. One striking and most apparent signature of migration is the dramatic eastern-to-western Africa cline of mitochondrial DNA (mtDNA) haplogroup L3a frequencies (Watson *et al.*, 1997). Haplogroup L3a is closely related to a mtDNA haplotype common in European populations [the Cambridge Reference Sequence (Anderson *et al.*, 1981)]. A subgroup of related mtDNA haplotypes appears to be East African specific and may represent a common ancestral sequence for most of Europe and Eurasia. In other words, the mtDNA diversity observed in non-African populations is a subset of African mtDNA variation. We note that while this group of mtDNA haplotypes is common in Eastern Africa, it represents only a subset of the total mtDNA diversity observed throughout the African continent. A similar pattern is observed for nuclear (Tishkoff *et al.*, 1996, 1998) and Y chromosome

(Passarino *et al.*, 1998) variation in Eastern Africa. There are several implications for these observations. First, it provides evidence for an African (specifically, Eastern African) origin for Eurasians. Second it suggests that before major migrations occurred out of the continent, populations were diverging. These observations deconstruct racial thinking, especially the concept of “racial divergence.”

The term *racial divergence* fails to describe the process responsible for producing the variation that exists as a continuum in the human species. So-called racial divergence dates reflect the times of differentiation of genes within populations used in a given analysis and should be interpreted as such. “Racial divergence” time estimates have been used to infer the age of the common ancestor between sampled groups, but this is definitely not the time of origin for the so-called “racial groups,” which traditionally have been defined by morphology, nor is it the time of “origin” for the sampled populations. Time, geography, and other data help elucidate the larger meaning of genetic studies. A date of 156,000 years ago has been suggested by Goldstein *et al.* (1995) for the separation of African (stereotypically defined) and non-African populations. Given that there is no fossil evidence for modern humans anywhere at 150,000 years ago except in Africa, this does not represent an African/non-African “split.” This date is actually an estimate of the initial genetic divergence that occurred *within* the continent of Africa among our modern human ancestors. After the expansion of modern humans out of Africa, subsets of the resultant genetic variation were distributed throughout the other continents. Most genetic variants observed outside of Africa are also found within Africa at various frequencies (and are of indigenous origin); this clearly indicates that “Africans” are not monomorphic. Northern African genetics, when this information is considered carefully with palaeontological data, would not seem to be explicable as simply “hybrids” or lost Eurasians. A different perspective, having more explanatory power and consistent with the available data, is that northern and Horn of Africa populations constitute gradients of differentiation largely reflective of African biohistorical processes.

The data of Tishkoff *et al.* (1996, 1998) are best interpreted as suggesting that a model of *in situ* evolutionary differentiation explains the bulk of variation in Africa (which includes supra-Saharan, Saharan, and sub-Saharan regions). Markers from mitochondrial DNA and the Y chromosome, which define maternal and paternal lineages, respectively, also reveal significant differentiation of ancestral African populations (Watson *et al.*, 1996, 1997; Passarino *et al.*, 1998; Malaspina *et al.*, 1998). Another likely example of internal differentiation within Africa is to be found in the “unusual” genetic makeup of the Khoisan, a southern African population with a high frequency of purported “Asian” alleles (Cavalli-Sforza *et al.*, 1994). Some suggest admixture with Asians to explain the Khoisan genetic profile. However, the Khoisan speakers may be descendants of a generalized ancestral human population from which modern Asian and African populations were ultimately derived. The Khoisan have been shown to be less diverse than other African populations for a variety of genetic loci (Excoffier *et al.*, 1987). This may be due to two

reasons: historical isolation and a smaller effective population size. The Khoisan have historically remained relatively isolated from other African populations due to cultural and linguistic differences. Also, the smaller effective population size is likely the result of the hunter-gatherer culture of the Khoisan. These cultural and subsistence patterns could promote significant differentiation of the population.

The overall pattern of genetic variation that exists within and outside of Africa suggests an African origin of modern humans and a recent common origin of non-African populations. The pattern of human genetic diversity supportive of this position is observed in studies which examine genetic signatures left behind from population expansions and population bottlenecks (Rogers and Harpending, 1992; Harpending *et al.*, 1993, 1998; Shriver *et al.*, 1997; Reich and Goldstein, 1998; Kimmel *et al.*, 1998; Jorde *et al.*, 1997). These studies consistently reveal an initial population expansion within Africa prior to out-migration into other continents. What is actually a population expansion has been mistakenly termed "racial divergence," which implies morphological differentiation into the recognizable entities now labeled "races." In reality, it represents the early genetic divergence of ancestral *Homo sapiens*. From the genetic data we find evidence of *in situ* differentiation (or genetic divergence) of mid-Pleistocene populations in Africa and subsequent migrations out of Africa into Europe and Asia, with continued drift due to isolation by distance and founder effects, which abruptly end when expansion in population size and frequent migrations occurred. Also, numerous environmental zones exerted their own selective pressures on the populations. The populations remaining in Africa are the primary ancestors of present Africans. This model is more consistent with archaeological and molecular evidence and can be tested using various data sets. This biogeographical perspective better explains the main source of variation within Africa and obviates the need for racial thinking and its associated baggage.

REFERENCES CITED

- Anderson, S., Bankier, A., Barrell, B., de Bruijn, M., Coulson, A., Drouin, J., Eperon, I., Nierlich, D., Roe, B., and Young, I. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**: 457-465.
- Bowcock, A., Kidd, J., Mountain, J., Herbert, J., and Cavalli-Sforza, L. L. (1991). Drift, admixture, and selection in human evolution: A study with DNA polymorphisms. *Proceedings of the National Academy of Sciences USA* **88**: 839-843.
- Bowcock, A., *et al.* (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**: 455-457.
- Cavalli-Sforza, L. L., Piazza, A., Menozzi, P., and Mountain J. (1988). Reconstruction of human evolution: Bringing together genetic archaeological, and linguistic data. *Proceedings of the National Academy of Sciences USA* **85**: 6002-6006.
- Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. (1994). *The History and Geography of Human Genes*, Princeton University Press, Princeton, NJ.
- Chibani, J., Lefranc, G., and Constans, J. (1985). Serum protein polymorphism among Tunisian Berbers: Haptoglobin, transferrin and group-specific component subtypes, C3 and BF types. *Annals of Human Biology* **5**: 449-462.

- Clark, J. D. (1989). The Origin and Spread of Modern Humans: A broad perspective on the African evidence. In Mellars, P., and Stringer, C. (eds.), *The Human Revolution: Behavioral and Biological Perspectives on the Origins of Modern Humans*, Edinburgh University Press, Edinburgh, pp. 565–588.
- Coon, C. (1962). *The Origin of Races*, Alfred Knopf, New York.
- Coon, C. (1965). *The Living Races of Man*, Alfred Knopf, New York.
- Excoffier, L., Pellegrini, B., Sanchez-Mazas, A., Simon, C., and Langaney, A. (1987). Genetics and history of sub-saharan Africa. *Yearbook of Physical Anthropology* 30: 151–194.
- Goldstein, D., Ruiz-Linares, A., Cavalli-Sforza, L. L., and Feldman, M. (1995). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences USA* 92: 6723–6727.
- Guglielmino, CR., Menozzi, P., Piazza, A., and Cavalli-Sforza, L. L. (1987). Measures of genetic admixture for North African populations. *Atti Associazione Genetica Italiana* 33: 177–178.
- Harpending, H., Sherry, S., Rogers, A., and Stoneking, M. (1993). The genetic structure of ancient human populations. *Current Anthropology* 34: 483–496.
- Harpending, H., Batzer, M., Gurven, M., Jorde L., Rogers A., and Sherry S. (1998). Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences USA* 95: 1961–1967.
- Jorde, L., Rogers, A., Bamshad, M., Watkins, A., Krakowiak, P., Sung, S., Kere, J., and Harpending, H. (1997). Microsatellite diversity and the demographic history of modern humans. *Proceedings of the National Academy of Sciences USA* 94: 3100–3103.
- Keita, S. O. Y., and Kittles, R. A. (1997). The persistence of racial thinking and the myth of racial divergence. *American Anthropologist* 99(3): 534–544.
- Kimmel, M., Chakraborty, R., King, J., Bamshad, M., Watkins, W., and Jorde, L. (1998). Signatures of population expansion in microsatellite repeat data. *Genetics* 148: 1921–1930.
- Malaspina, P., et al. (1998). Network analyses of Y chromosome types in Europe, Northern Africa, and Western Asia reveal specific patterns of geographic distribution. *American Journal of Human Genetics* 63: 847–860.
- Merghoub, T., Sanchez-Mazas, A., Tamouza, R., Lu, C., Bouzid, K., Ardjoun, F., Labie, D., Lapoumeroulie, C., and Elion, J. (1997). Haemoglobin D-Ouled Rabah among the Mozabites: A relevant variant to trace the origin of Berber-speaking populations. *European Journal of Human Genetics* 5: 390–396.
- Nei, M., and Ohta, T. (1991). Evolutionary relationships of human populations at the molecular level. In Osawa, S., and Honjo, T. (eds.), *Evolution of Life*, Springer, Tokyo, pp. 415–428.
- Nei, M., and Roychoudhury, K. (1993). Evolutionary relationships of human populations on a global scale. *Molecular Biology and Evolution* 10(5): 927–943.
- Passarino, G., Semino, O., Quintana-Murci, L., Excoffier, L., Hammer, M., and Santachiara-Benerecetti, S. (1998). Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *American Journal of Human Genetics* 62: 420–434.
- Reich, D., and Goldstein, D. (1998). Genetic evidence for a Paleolithic human population expansion in Africa. USA. *Proceedings of the National Academy of Sciences* 95: 8119–8123.
- Rogers, A., and Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9(3): 552–569.
- Seligman, C. (1930). *The Races of Africa*, Oxford University Press, London.
- Shriver, M., Jin, L., Ferrell, R., and Deka, R. (1997). Microsatellite data support an early population expansion in Africa. *Genome Research* 7: 586–591.
- Tartaglia, M., Scano, G., and De Stefano, G. (1996). An anthropogenetic study on the Oromo and Amhara. *American Journal of Human Biology* 8: 505–516.
- Tishkoff, S., et al. (1996). Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* 271: 1380–1387.
- Tishkoff, S., et al. (1998). A global haplotype analysis of the myotonic dystrophy locus: implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations. *American Journal of Human Genetics* 62: 1389–1402.
- Watson, E., Bauer, K., Aman, R., Weiss, G., von Haeseler, A., and Paabo, S. (1996). mtDNA sequence diversity in Africa. *American Journal of Human Genetics* 59: 437–444.
- Watson, E., et al. (1997). Mitochondrial footprints of human expansions in Africa. *American Journal of Human Genetics* 61: 693–704.