



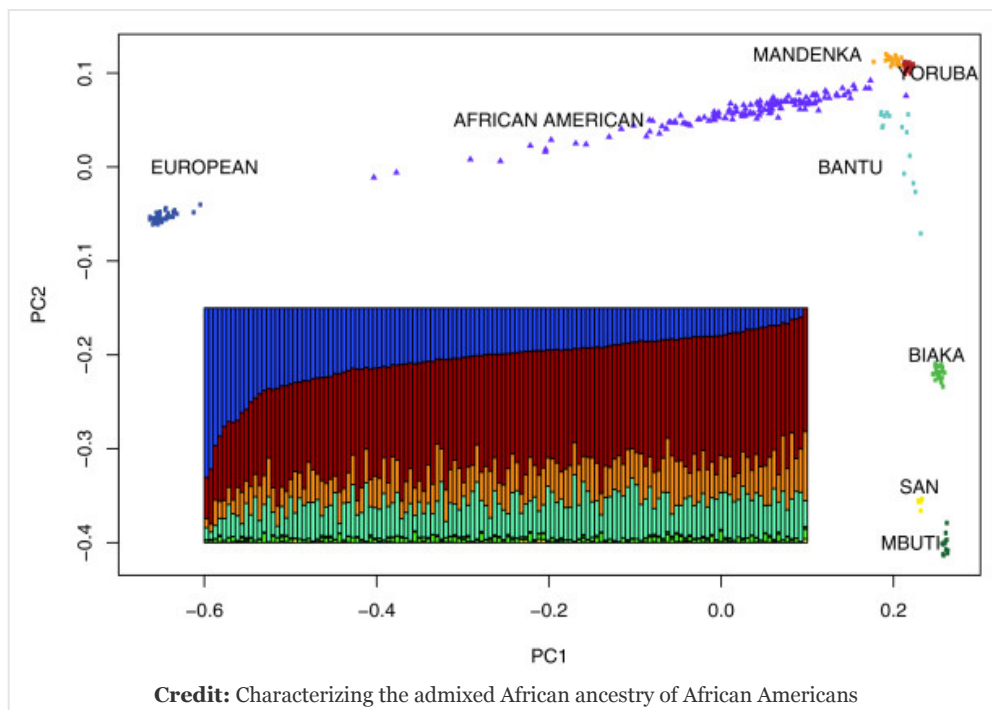
Gene Expression

« Angelina Jolie, Myriad Genetics, & patents on genes

Open thread, 5/19/2013 »

Why race as a biological construct matters

By Razib Khan | May 16, 2013 6:29 am



My own inclination has been to not get bogged down in the latest race and IQ controversy because I don't have that much time, and the core readership here is probably not going to get any new information from me, since this is not an area of hot novel research. But that doesn't mean the rest of the world isn't talking, and I think perhaps it might be useful for people if I stepped a bit into this discussion between [Andrew Sullivan](#) and [Ta-Nehisi Coates](#) specifically. My primary concern is that here we have two literary intellectuals arguing about a complex topic which spans the humanities and **the sciences**. Ta-Nehisi, as one who studies history, feels confident that he can dismiss the utility of racial population structure categorization because as he says, “no **coherent, fixed** definition of race actually exists.” I am actually more of a history guy than a math guy, not because I love history more than math, but because I am not very good at math. And I've even read books such as [The Rise and Fall of the Caucasian Race](#) and [The History of White People](#) (as well as [biographies](#) of older racial theorists, such as [Madison Grant](#)). So I am not entirely ignorant of Ta-Nehisi's bailiwick, but, I think it would be prudent for the hoarders of old texts to become a touch more familiar with the crisp formalities of the natural sciences.

In his posts on this topic Ta-Nehisi repeatedly points to the real diversity in physical type and ancestry among African Americans, despite acknowledging implicitly the shared preponderant history. **But today with genomic methods we have a rather better idea of the distribution of ancestry among African Americans.** The above plot is from [Characterizing the admixed African ancestry of African Americans](#), a 2009 paper with 94 Africans of diverse geographic origins, 136 African Americans, and 38 European Americans. They looked at 450,000 genetic variants (SNPs) per person (there are somewhat more than 10 million SNPs in the human genome). Obviously individuals and populations

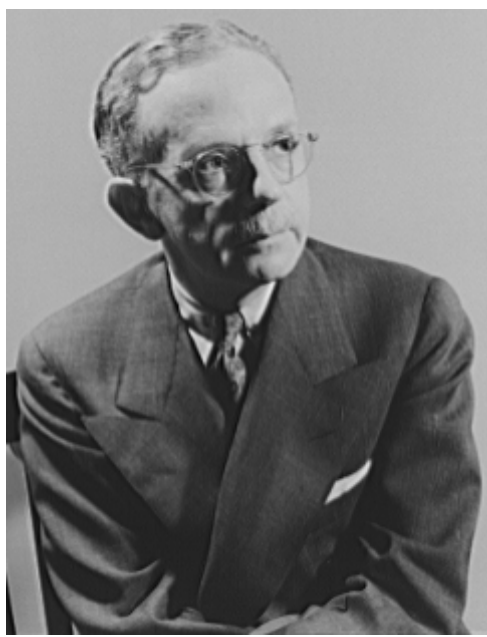
exhibit genetic relationships to each other contingent upon the patterns of the variation of base pairs (A, C, G, and T) across the genomes of individuals, but there's no reasonable way to comprehend this "by eye" when you're talking about hundreds of thousands of markers. The authors used two simple methods to infer clustering within the data set.

First, you see a PCA plot. This method is one where the independent dimensions of variation within the data set of the markers are pulled out. They are rank ordered in terms of how much variation they can explain (dimension 1 by convention explains the most, dimension 2 explains the second most, and so forth). Each dimension can be thought of as having a value proportional to its explanatory power. Each individual then has a value position on the dimension, dependent on how that individual relates to the others. **When you take multiple dimensions and transpose the data geometrically you quickly see population structure fall out of the data set.** Notice above that the first dimension of variation (PC1) separates the Europeans from all the African populations. The second dimension of variation (PC2) separates the hunter-gatherer populations of Africa from the agriculturalists. While the Mandenka are from Senegal, the Yoruba are from Nigeria, at opposite ends of what is traditionally termed West Africa. This was the presumed source of most of the African slaves who arrived in the United States. Once these slaves came to the United States some of them had children with white Americans. It turns out that the average African ancestral contribution to African Americans is ~80%, with the balance being mostly European (there is some Native American, but not much). In fact this is very close to the estimates which were produced by genealogists. The concordance of these methods is reassuring, since the underlying phenomena is the same.

Notice that on the PCA plot no African American falls in the Mandenka-Yoruba cluster. That is because almost no African American whose ancestors are not recent immigrants from Africa lack white ancestry. This is entirely reasonable when you consider that the vast majority of their ancestors were resident in the colonies before the Revolutionary War. Admixture events would have percolated throughout the genealogical tree in subsequent generations. The African Americans are distributed almost perfectly along a line between the West African populations and the European Americans. Observe that the density seems to decrease as you approach the European American cluster.

Now we can move to the second visualization technique. While the PCA does not posit any hypothesis of population structure (it just "fell out" of the genetic variation due to the shared history of some individuals via their common ancestors), the second method is "model based," in that the authors posited seven ancestral populations to match the seven populations which African Americans may be derived from. In a way this is rigging the game; if you force the method to squeeze out particular numbers of populations it may act strangely. But in this case we have prior expectations, so this number of populations is not unreasonable. **Above each bar plot represents an African American individual, with each fraction of shading an ancestral element.** The results from the PCA are reproduced nearly perfectly by this differing method. The average ancestral quantum of African heritage in this sample is ~80%. And, you see more cleanly the variation in European ancestry among African Americans. Less than 10% of African Americans are like Barack Obama, at least 50% (or more) of European ancestry. The African ancestry excludes the hunter-gatherer populations which is reasonable since the slaves were from the Congo in the east (where some were Bantu) as far as Senegal in the west.

Ta-Nehisi has used an image of **Walter White**, the first African American head of the NAACP, to illustrate the pliability of the black identity. It certainly shows that there are no **fixed definitions** of race which are particularly useful. But that is a misconception of biological science, which is rife with exceptions and boundary conditions, and characterized by an **instrumental** perspective. The data above suggests that self-identified African Americans are characterized by **some** African ancestry, but over 90% are more than 50% African in ancestry. Walter White, who had five black great great great grandparents and 27 white ones, was almost certainly less than 20% African in ancestry. **There are such people**



The first black head of the NAACP

even today, but they are not typical, and do not disprove the reality that African Americans are predominantly of African ancestry.

From a **scientific** perspective in **biology** there are not ultimate and fundamental taxonomic facts. There are simply useful ideas and concepts to illustrate and explore the objective phenomena of the natural world. The **Species Concepts** debate shows us this reality well, as even species can be tendentious. But the debate often shakes out along disciplinary lines. Many more ecological scientists seem to be taken by the ecological species concept, while evolutionary geneticists are more keen on the biological species concept. That is because they are choosing the framework most useful for their ends. There is nothing “Post Modern” in this in that it denies reality. Rather, we are disputing the **ideas** which we use to capture the essence of real phenomena in compact semantic relations suitable for symbolic representation (whether with math or language).

Prior to the modern systematic era of biology humans did attempt to classify themselves. Generally they looked at a few informative features. For example the Chinese referred to both South and Southeast Asians as “black,” not because they thought they were African, but because they had brown or dark brown skins. Similarly, Arab ethnographers differentiated between ruddy peoples to the north, black ones to the south, and black ones to the east (Indians). And so on. **This is almost certainly an elaboration of our innate cognitive ‘folk biology.’** By this, I mean that we as humans tend to classify organisms. Why this is adaptive is trivially obvious. When humans meet new organisms which resemble those which they have familiarity with prior, they simply reformulate the novel creatures as variants of the familiar ones. For example the Tasmanian Tiger was no tiger. It was not even a placental mammal. But through convergent evolution it resembled placental carnivores. Analogously, when Europeans first met the straight haired brown skinned native peoples of the New World they termed them “Indians,” a straight haired brown skinned population of the Old World. When they met the very dark and kinky haired peoples of the western Pacific they assumed they were some relation to Africans, and these became “Melanesians” (which means “black islanders”).

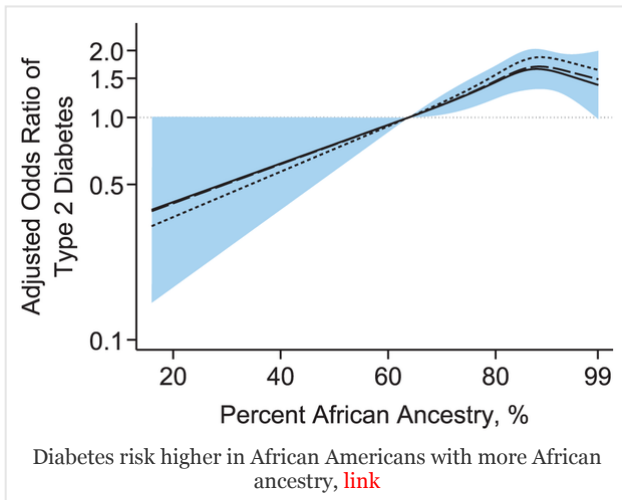
A second component of human nature which Coates alludes to is our tendency **to cohere into groups with narratives of internal identity set apart from the Other.** In the pre-modern world these inter-group cleavages would be marked by accent, dress, and tattoos. In the early modern world they would be correlated with religion or nationality. The dynamic at issue here is that **extremely genetically close populations which would be indistinguishable naked had to generate salient cultural markers.** In the case of the ancient Hebrews one could argue that circumcision was exactly the sort of marker which would persist even when naked!* This does not mean that there were **no** detectable genetic differences between adjacent small scale societies; there are after all **detectable genetic differences** across European villages today. But for particular technical evolutionary reasons (far more within group variance than between group variance in regards to genetics) it is likely that for inter-group competition cultural forces reigned supreme over biology, and were determinative of identity.

Both of these parameters are from our deep history as a species characterized by life as hunter-gatherers in bands. The next force is more recent, and historically contingent. As I suggested above non-European and pre-modern peoples had a vague conception of race on the continental scale. The Classical Greeks even distinguished the various brown peoples, the Egyptians and the Indians of the north, and black peoples, the Ethiopians and the Indians of the south. The fact that the initial explorers who arrived in the New World labeled the indigenous people Indians, and not Chinese or Africans, shows an awareness of

global diversity (in contrast, the British referred to the Australian Aboriginals as blacks). When the British first arrived in India as supplicants to the Great Mughal they differentiated between the diverse races of the subcontinent. The black and brown natives, and a portion of the elites who were white (West Asian Persians and Turks).

This changed over the centuries, **and after 1800 the age of European supremacy and the rise of systematic science produced the sort of racial nationalism which serves as the backdrop to our understanding of race more generally.** Whereas the pre-modern folk biological taxonomies were coarse, but generally accurate up to a point, the age of white supremacy produced a somewhat schizophrenic science of precision and exaggeration. By this, I mean that the attempt to be formally scientific resulted in a plethora of categorizations and grades of hierarchy. But, the reality of white supremacy generated a taxonomy of dominion, where all the races of color were aggregated into an amorphous whole. Perhaps these two countervailing tendencies explains the juxtaposition of quasi-fixed racial characters with a bizarrely elastic definition of the Other, the non-white. Few moderns agree with Lothrop Stoddard's **The Rising Tide of Color Against White World-Supremacy**, but many implicitly accept the framework of whites and a coalition of "people of color."

So there you have it. An underlying biological reality which is a reflection of deep history. It may not be real or factual, but it is consistent and coherent. Then there are innate faculties which lead us toward categorization of humans into various kinds, for deeply adaptive purposes. Finally, there are historically contingent events which warp our perception of categories so as to fit into power relations in a straightforward sense. But wait, there's more!



The biological aspect above focused on ancestry and history. But this is not academic detail. **The history of a population affects its genome, and its genome effects the nature of its traits and diseases.**

Because of differences across populations statistical geneticists with medical aims routinely restrict their data set to individuals of one population. And, within groups like African Americans which are admixed **there is variation in disease risk by genomic fraction.** Though an individual with 60 percent African ancestry may feel and say they are no more or no less African American than someone who is 80 percent African in ancestry, there are differences in disease susceptibilities.

There is no Platonic sense where there are perfect categories with ideal uses. Rather, we muddle on, making usage of heuristics and frameworks which are serviceable for the moment. We lose our way when we ignore the multi-textured nature of the issues.

* Though many of the neighboring peoples practiced circumcision, so this is more of an apocryphal illustration than a real instance of functional traits on a cultural level in societies.

CATEGORIZED UNDER: **GENETICS, GENOMICS**

MORE ABOUT: **RACE**

+