XV*—HOW TO DECIDE IF RACES EXIST

by Kwame Anthony Appiah

ABSTRACT Through most of the twentieth century, life scientists grew increasingly sceptical of the biological significance of folk classifications of people by race. New work on the human genome has raised the possibility of a resurgence of scientific interest in human races. This paper aims to show that the racial sceptics are right, while also granting that biological information associated with racial categories may be useful.

I

From a very early age, people across cultures classify others on the basis of appearance without any particular encouragement. As Susan Gelman has argued in her fascinating book *The Essential Child*, evidence from developmental psychology shows that by the age of six children treat races as 'possessing inborn features, inherent in the ... person, and passed down from parent to child.'[1] Young children, she argues, also *essentialize* these groups: they believe that the 'outer' characteristics by which they assign people to groups reflect shared 'inner' properties that explain both appearance and behaviour.[2] So there is a large set of ways of classifying people all around the world and throughout history that reflect this cognitive predisposition.

By talk of 'folk races'—and this is just a stipulation—I mean to pick out those folk categories that are based on the idea that membership in the relevant group is determined by intrinsic properties inherited from one's parents, properties that are shared by all normal members of the group. Using this terminology, the hypothesis that there are human folk races is the hypothesis that there are human groups of common ancestry that are (roughly) definable by shared inherited intrinsic properties.

1. Susan Gelman, *The Essential Child*, New York: Oxford University Press, 2003, p. 105.

2. See Gelman 2003, Chapter 11, 'Why Do We Essentialize?'.

*Meeting of the Aristotelian Society, held in Senate House, University of London, on Monday 19 June 2006 at 4.15 pm.

It's a consequence of this stipulation that biological subspecies, at least as many evolutionary biologists have conceived of them, are not likely to be folk races.[3] That's because membership in a subspecies is not an intrinsic property, but a relational one. A subspecies is a kind of biological population. In a sexually reproducing species like ours, a population is a collection of organisms whose members have a significantly higher propensity to reproduce with opposite-sex members of the group than they have to reproduce with organisms outside it. As a result, two organisms that are quite alike in intrinsic biological properties can belong to different populations, and two organisms that are quite dissimilar in properties can belong to the same population. Indeed, you can have two organisms, *A* and *B*, in the same population where *A* is far more different in intrinsic properties from *B* than from *C*, which is not in the population at all. (Imagine a population split in two by the sudden appearance of a new river formed after an earthquake. Consider *A*, *B* and *C*, who were members of the original population before the split. Suppose *A* and *C* are close kin, but *A* and *B* have no recent common ancestor; suppose that *A* and *B* are now on one side of the river and *C* is on the other. Organisms that can't meet can't mate. So *A* and *C* belong to different populations now.)

I advertise this fact—that what I call folk races aren't likely to behave like modern biological classifications—since it *is* the pretty direct result of a stipulation and some well-known biology. For clarity's sake, I'll use the word 'subspecies' for this biological kind. I want to insist that my stipulation isn't arbitrary, though: it is motivated by the fact that folk practices of ethno-racial classification are generally essentialist (in Gelman's sense) because we have the cognitive tendency that Gelman has described so well.

Folk classifications in the modern West are quite typical. We assign people to races in a way that is governed by this rule: if your parents are of the same race, you're of the same race as your parents. Since you get your genetic endowment from your parents, racial identities governed by this rule will sometimes be statistically correlated with genetic characteristics, provided

3. See Philip Kitcher, 'Race, Ethnicity, Biology, Culture', in Leonard Harris (ed.), *Racism*, Amherst, NY: Humanity, 1999, pp. 87–120.

there are genes in the local members of a folk race that are commoner than in the general population. Since people are also often assigned to racial groups in part on the basis of phenotypic characteristics that have a genetic basis, there will often, in fact, be such correlations. But Westerners are inclined to suppose not just that there are biologically-based features of people that are statistically characteristic of their race, but also that those features extend far beyond the superficial characteristics on the basis of which racial categorization is usually based. So we *essentialize* race, in Gelman's sense of that term. And a great deal of what people believe about the biological basis of these deeper differences is false.

Because the central beliefs of many people about folk races are mistaken in these ways, we cannot explain how people are assigned to races by discovering some folk theory and supposing it to be roughly true. So—since folk races are, like it or not, an important feature of our social landscape—we need an account of the racial categories actually in place that is consistent with the pervasiveness of erroneous beliefs.

II

Here is such an account.[4] It begins by supposing that folk race is an important kind of social identity. That's because I think that folk races are of interest to us largely because they *are* forms of social identity. They continue to be interesting in that way whether or not they are interesting for biological purposes.

My explication of social identities is nominalist: it explains how the identities work by talking about the labels for them. The main motivation for the nominalism is that it allows us to leave open the question of whether the empirical presuppositions of a labelling practice are correct. Since many social identities are like folk races in being shot through with false belief, this is a decided advantage. So, take a representative label, X, for some identity.

4. See K. A. Appiah and Amy Gutmann, *Color Conscious: The Political Morality of Race*, Princeton, NJ: Princeton University Press, 1998, and K. A. Appiah, *The Ethics of Identity*, Princeton, NJ: Princeton University Press, 2005.

IDENTITY: There will be criteria of ascription for the term '*X*'; some people will identify as *X*s; some people will treat others as *X*s; and there will be norms of identification.

Each of these notions—*ascription*, *identification*, *treatment* and *norms of identification*—requires brief commentary.

III

A person's criteria of ascription for '*X*' are properties on the basis of which she sorts people into those to whom she does and those to whom she doesn't apply the label '*X*'. The criteria of ascription need not be the same for every user of the term; indeed, there will rarely be a socially agreed set of properties individually necessary and jointly sufficient for being an *X*.[5]

Here is what characterizes competence with the term '*X*'. There will be certain kinds of people—we can call them 'prototypical *X*s'—such that your criteria of ascription must pick them out as *X*s. There will be other kinds—'antitypes', let us call them—that your criteria of ascription must exclude. A prototype is not an actual person: it is a specification of conditions sufficient for being an *X*; just so, an antitype specifies conditions sufficient for not-being one. But something may be neither a prototype nor an antitype of an *X*. A Cuban-American, most of whose ancestors came to Cuba before the eighteenth century, and who arrived in Florida in 1950 is a prototype of a Latino. A normal European or African who does not speak Spanish or Portuguese and does not come from the Iberian peninsula is an antitype. List all the prototypes and antitypes and you may find that they do not divide logical space into two classes.

Because prototypes and antitypes don't always divide logical space in two, criteria of ascription need not divide actual people into *X*s and not-*X*s, either. Rather, they must divide

5. For those who want to go this way, I suggest the best chance you have is to suppose that someone is competent if their conception picks out most of the *X*s in their social environment; where what it is to be an *X* is explicated in terms of the best scientific account of what it is most users are talking about. One reason I don't favour this approach is that I think that for some social identities the best scientific account is that they're not referring to anything; but then that would make no users competent, if they thought there were any *X*s at all.

all actual people roughly into three classes, which we can call (modelling our classification on Max Black's account of metaphor) the positive, negative and neutral classes. That is, they must make some people, in the positive class, *X*s; some people, in the negative class, not-*X*s; and they may leave some people, in the neutral class, as neither determinately *X*s nor determinately not-*X*s.[6] Let me underline that, whereas prototypes are abstract, these classes are classes of actual people. I am not trying to get at the way the predicate works across possibilia.

This is what competence consists in; but people do not need to know that their criteria of ascription have these features to be competent. And, in general, they won't know what the relationship is between their criteria of ascription and the total human population. So they may well think, for example, that they can divide the world precisely into *X*s and not-*X*s, even though there do in fact exist people (people they have not met) who would be in the neutral class for them, if they did know about them. I shall say that someone who has criteria of ascription for an identity-term '*X*' that meet the conditions for competence has a *conception* of an *X*.

This is, no doubt, too abstract; so let me just exemplify. Take the term 'Asian' as used by Johnny from Cornwall, who has met very few people from anywhere in Asia and very few British Asians either. Johnny says 'Asians are a race' and ascribes the term 'Asian' to everyone who looks a certain way, in fact the sort of way most movie stars in Bollywood movies would look to him. (I'll call this 'looking Asian to' Johnny.) He also thinks that the label is properly applied to anyone whose ancestors for many generations have come from India, because he supposes that everybody in those countries would look Asian to him. Now Johnny will get all the prototypes and antitypes right. Give him a Bangladeshi? 'Asian.' Give him most Finns or Congolese? 'Not Asian.' So he's competent. But presented with a Kirghiz or a Kazakh (people, let us suppose, of whose existence he is currently unaware) he might not know what to say. So his conception has a neutral class, even though he doesn't know this. He may also have false beliefs—such as that almost everyone in Asia

6. I say 'roughly' to acknowledge a complication that I will ignore from now on: these classes will usually each be fuzzy.

looks roughly the way Indian people look—even though most people in Asia do not: a couple of billion people in China and South-East Asia, for example.

IV

By itself a way of classifying people that works in this way by ascription would not produce a social identity. What makes it a social identity of the relevant kind is not just that people suppose themselves or others to be *X*s but that being-an-*X* figures in a certain typical way in their thoughts, feelings and acts. When a person thinks of herself as an *X* in the relevant way, she *identifies as an X*. What this means is that she sometimes *feels like an X* or *acts as an X*.

An agent *acts as an X* when the thought 'because I am an *X*' figures in her reasons for acting or abstaining. Perhaps you never act as a British person (hereafter 'Brit'). But feelings can constitute identification too. You discover that hundreds of thousands of Brits responded to the Asian tsunami by sending money. You feel proud to be British. To *feel like an X* is for your being an *X* to figure in the intentional content of your feeling. The intentional content doesn't have to be *that you're an X*, though: you may feel proud of Mary, a fellow Brit, say. Here your being British figures in the intentional content of the feeling, because part of the intentional structure of the feeling is that Mary is *British like me*, even though you're not proud *that you're British.*

Similarly, our treatment of and feelings about other people reflect identity. You treat *A* as an *X* when 'because *A* is an *X*' figures in your reason for doing something to *A*. Supererogatory kindness is a common form of treatment-as directed towards fellow in-group members. Morally opprobrious unkindness is, alas, a horribly frequent form of treatment-as directed towards out-group members. It takes ascription, identification and treatment for a label to be functioning as the label for a social identity of the sort that I am explicating.

One reason identities are useful is that they allow us to predict how people will behave. This is not just because the existence of criteria of ascription entails that members of the group have or tend to have certain properties. It is also because social identities

are associated with *norms* for *X*s. That is the final element of my explication of the notion.

There are things that, *qua X*s, people ought and ought not to do. The 'ought' here is the general practical ought, not some special moral one. Here are some examples. Negatively: men ought not to wear dresses; gay men ought not to fall in love with women; blacks ought not to embarrass the race. Positively: men ought to open doors for women; gay people ought to come out; blacks ought to support affirmative action. To say that these norms exist is evidently not to endorse them. I don't myself endorse any of the norms I just listed. The existence of a norm that *X*s ought to *A* amounts only to its being widely thought—and widely known to be thought—that many people believe that *X*s ought to *A*.[7]

V

I should underline how many and various are the predicates of persons that fit this general rubric. I started with racial and ethnic terms; and I mentioned a nationality, British. But I could also have mentioned professional identities, vocations, affiliations, formal and informal (like Man U fan or Conservative), and other more airy labels ... dandy, say, or cosmopolitan.

I am pointing to this range not just because, like a well-bred philosopher, I am interested in generality, but also because this range invites an obvious question. *Why* do we have such a diverse range of social identities and relations? One answer, an aetiological one, will talk about our evolution as a social species and the fact that we are designed evolutionarily for the social game of coalition-building in search of food, mates and protection. This is, I think, a good explanation for our having the sort of psychology of in-group and out-group solidarities and antagonisms that social and developmental psychologists, like Susan Gelman, have been exploring for the last half-century.

But the psychologies that evolution has given us mean that there is a way the world looks from the inside, from the point of view of a creature with that psychology. And from that point

7. I put it this way because I think it sometimes turns out that hardly anybody really believes in the norm; still, it exists if people mostly think most people endorse it.

of view I think there is another, equally persuasive answer. Each of us has to make a life and to try to make it go well, and we need identities to make our human lives.

We make our lives, that is, *as* men and *as* women; *as* Americans and *as* Brits; *as* philosophers and novelists. Morality—by which I mean what we owe to one another—is part of the scaffolding on which we make that construction. So are various projects that we voluntarily undertake: Voltaire's garden—the one, perhaps, to whose cultivation he consigned his *picaro* Candide—shaped the last part of his life. But identities are another central resource for making our lives. Identities are diverse and extensive, I think, because people need an enormous diversity of tools for making their lives. Each person needs many options. And, because people are various, the range of options that would be sufficient for each of us won't be sufficient for us all.

VI

There are positive, negative and neutral classes for each competent speaker: that is, there is a way she would assign everybody on the planet roughly to one of these three classes, if that person showed up in her environment and answered truthfully questions about herself. The prototypes and antitypes define the socially permissible limits of individual positive and negative classes. So we might ask whether there is an interesting property—intrinsic or relational, simple or logically compound—shared by (most) prototypes that is not shared by (most) antitypes. Can we tell a story about racial identity, for example, that shows it to correspond roughly, in this way, to a biological property of genuine interest? If so, folk races are, in a sense, biologically real.

It's in answering this question that new work on the human genome strikes some people as helpful. Genomics teaches us not only what genes are, but also how they tend to be associated with each other. This offers the prospect of associating certain social groups statistically with genomic features. And where those statistical correlations are distinctive enough of the group and the genomic feature is of importance—for example, for medical reasons—there can be an obvious sense in which biological

claims about the group can turn out to be statistical truths. This has been part of folk wisdom for quite a while for a few cases: sickle-cell disease, glucose-6-phosphatase dehydrogenase deficiency, and Tay-Sachs disease, for example, are both rare in human beings generally and much more frequent in some groups of common ancestry than in others.

Sometimes the groups in question are quite small: there are alleles that have been found in certain families and nowhere else. Sometimes the groups are large: Yoruba people, of whom there are more than thirty million in south-west Nigeria, have a 6% frequency of the gene for haemoglobin C (which produces a relatively mild blood disease even in heterozygotes, who carry two copies of it); and 25% of the population of Nigeria as a whole carries the gene for haemoglobin S, which produces the classic and serious form of sickle-cell disease in heterozygotes.[8] A normal haemoglobin molecule is made up of four subunits, two α and two β chains; each chain is produced by a distinct gene, and there are many variants of both the α and the β chains. Since the α and β chains are required in equal numbers to form normal haemoglobin, there is also a range of genetic diseases associated with non-standard haemoglobins—the thalassaemias—in which one or other chain is produced in too small a quantity. 39% of Nigerians have some form of α-thalassaemia, the diseases produced when you have an under-production of α chains. These disorders—sickle-cell and thalassaemia—can be inherited both separately and together, producing a dazzling array of blood diseases, and so there is a very wide range of clinical contexts in which it is relevant to know if someone has Nigerian ancestry.

Of course, it's the differences in frequency between populations that make these correlations significant. As a standard discussion of blood diseases points out:

> α-Thalassemia is perhaps the most common single-gene disorder in the world. The frequency of α-thalassemia alleles is 5–10% in persons from the Mediterranean basin, 20–30% in portions of West Africa, and as high as 68% in the southwest Pacific.

8. See O. O. Akinyanju, 'A Profile of Sickle Cell Disease in Nigeria', *Annals of the New York Academy of Sciences*, 565.1, 1989, pp. 126–36; and Kenneth R. Bridges, *Information Center for Sickle Cell and Thalassemic Disorders*, http://sickle.bwh.harvard.edu/index.html.

> The frequency of heterozygote carrier status among the Chinese population has been reported to vary from 5–15%. The frequency of α-thalassemia is less than 0.01% in Great Britain, Iceland, and Japan.[9]

What is most obviously distinctive of reproductively isolated biological populations is the frequency with which variant alleles occur in that population. As we saw earlier, an *individual* in one biological population could, in principle, have almost the same genotype as an individual in another. That is, in essence, why attempts to define biological populations by biological properties shared by their members won't work. A population is a collection of organisms defined, as I said earlier, by the fact that they have a significantly higher probability of reproducing with opposite-sex members of the group than they have of reproducing with organisms outside it. This is a relational property—though it is one that is sometimes explained by an intrinsic property: some sub-populations of *Drosophila* have male genitalia that don't work with the genitalia of females in others. Sometimes the explanation is not an intrinsic property of the organism: populations may just be separated by a mountain range. And sometimes, in humans at least, the explanation could be cultural. If two human populations had ever lived side by side for a long time with no exchange of genes, indeed that would be the most likely explanation. History does not, so far as I know, afford examples of total reproductive isolation of this kind.

So, for example, a majority of members of the folk race of African-Americans have relatively dark skin for genetic reasons. Biological remains that contain some of the genes that characteristically account for this darker skin colour can therefore reasonably be identified for forensic purposes as (socially) African-American. Here there is a genuine biological trait that can be used to identify a genuine social trait, even though the social trait is not identical with any intrinsic biological property. So the utility of genomic properties in identifying a social group doesn't entail that the social group is a subspecies.

9. Alexandra C. Cherva, Afshin Ameri and Ashok Raj, 'Hemoglobin H Disease', *eMedicine*, http://www.emedicine.com/ped/topic955.htm. Last updated: April 2 2002.

This is all consistent with recognizing that many African-Americans do not bear the genes that produce darker skin; that there are other genomic characteristics statistically distinctive of African populations that a person of African ancestry may share without having the skin-colour genes; and that you can be an African-American while having many fewer of the genomic characteristics statistically distinctive of an African population than many people who are identified as white.

Perhaps all this is obvious. But I find in discussion that people seem not to grasp these points intuitively, so perhaps they are worth making. And if they are worth making, perhaps it is also worth filling in some of the conceptual background.

VII

As we all now know, genes consist of sequences of bases, and each sequence of three such bases (a *triplet* or DNA *codon*) has a functional significance in determining what protein is produced. Mutations in genes occur when one base is replaced with another. Because the relationship between codons and amino acids is many–one, some such substitutions make little functional difference, since the same polypeptide sequences result and the same proteins are formed. Other substitutions change the polypeptide sequence, by substituting one amino acid for another, but make little difference to the biological functioning of the resulting protein: enzymes, for example, characteristically have certain active regions that are important to their functioning, while other sequences are structural supports for the active regions.

Where a mutation has a functional significance, it is most likely to have a negative effect on the organisms that carry it: we are complex wholes with interdependent parts adapted to one another over a relatively long period in a relatively stable environment, and in general a change in the functioning of one element of this complex stable whole will reduce, not increase, our overall fitness. But where a mutation has little or no functional significance it can survive. There will be no selection pressure against it. And so there will be single nucleotide polymorphisms—DNA sequences that differ in just one base from each other—that produce different forms of a gene

that are nevertheless functionally equivalent. ('Single nucleotide polymorphism' is a long expression for a short change. Usually it's abbreviated to SNP, pronounced 'snip'.) A SNP refers both to a site on a chromosome which is occupied in different people by different bases and to the various bases that can be there.[10] Most loci on most genes are the same in everybody: many of the base sequences it takes to be a functioning organism are identical, because changes in most base sequences don't produce a functioning individual. But it's usually estimated that 0.1% of the DNA consists of sites where SNPs can occur in living people. By October 2005, about 3.6 million SNPs had been 'validated'.[11]

The *genotype* of a person is a specification of every pair of alleles that she carries for every locus on the genome. Consider two people, each of whom carries the same two alleles at the same two sites: say, *Aa* and *Bb*. But suppose in John *A* and *B* are on one chromosome and *a* and *b* are on another, while in James *A* and *b* are on the same chromosome and *a* and *B* are on another. Suppose that these sites are close together on the same chromosome: as a result the alleles that they carry are extremely unlikely to be separated in cell division.

Now consider the results of sex with a partner whose genotype is *AABB*. With John, she will have offspring *AABB* or *AaBb*. With James, the options are *AABb* or *AaBB*. While John's and James's genotypes are the same, the genotypes of their offspring with the same partner will be different. We will be able to tell, in particular, if we come across one of these offspring, which of the two males was their father simply by looking at two loci, *even though, for those loci, the two potential fathers have the same genotype*. What determines your propensity to produce offspring of a certain genotype, simply put, isn't just your genotype, it's the way in which that genotype is placed on your chromosomes.

That's why the notion of a haplotype—or haploid genotype—is useful in tracing ancestry. It's the specification not of your genotype, but of the sequence of genes on just one of each pair of your chromosomes. Each individual can be thought of

10. This is like the word 'gene', which is used to refer both to a locus on the chromosome and to the various alleles that can occur at that locus.

11. The International HapMap Consortium, 'A Haplotype Map of the Human Genome', *Nature*, 437, 27 October 2005, p. 1316.

genetically, then, as having two haplotypes. Of course, because there are twenty-three chromosome pairs, you could specify the haplotype in 2^{23} ways: but once you had picked one such way—by selecting one from each pair of chromosomes—you would also have fixed which other haplotype you needed to specify.

One non-arbitrary way to pick a way of specifying the haplotype would be to specify the sequence of alleles on the chromosomes derived from the mother's egg and then specify the sequence on the chromosomes derived from the father's sperm. In the process of meiosis—the type of cell division that produces sex cells—material can be swapped between the two versions of a chromosome carried in a normal somatic cell, in the process called 'crossing over'. But if crossing over did not occur, you could think of a person as the combination of a maternal and a paternal haplotype, since without crossing over each person would get exactly one chromosome of each homologous pair of chromosomes from each parent. (Bear in mind, though, that there are 2^{23}—or 8,388,608—possible haplotypes derivable from each parent without crossing over; that's one reason why children of the same parents would be different from one another even if there were no recombination of genes in meiosis.)[12]

The word 'haplotype' is also used to refer to classes of haplotypes in the sense I have just defined: namely, a class of haplotypes that are identical in some sequence of alleles close to each other on a single chromosome, often, more particularly, a set of genes for proteins that carry out related activities. More precisely, a haplotype in this second sense is fixed by the sequence of alleles on a relatively short continuous stretch of a chromosome (modulo a few SNPs that have little functional significance). From now on I'll use 'haplotypes' in this second sense. So to say two people have the same haplotype is to say, roughly, that they share an interesting collection of genes on a single chromosome.

Since the genes in short regions of a chromosome seldom get separated in cell division, your haplotype in this sense is almost always derived from a single parent. As a result, when a SNP

12. This is the reason haplotypes are called haplotypes: the spermatozoa and the oocytes are haploid—they have only one member of each type of chromosome—unlike most somatic cells, which are diploid, having two of each.

arises by mutation in an ancestral chromosome, it provides a marker for descendants of that ancestor, so long as that SNP does not undergo further mutation and the sequence of genes that includes it does not get broken by crossing over. And this is the basis on which African-Americans are now seeking to identify ancestral ties to particular places in Africa.

VIII

Many contemporary African-Americans have come to take an interest in Yoruba religion, especially in the forms mediated by Haitian *vodou* and the Afro-Brazilian traditions of Bahia. To discover that you have SNPs associated with a haplotype distinctive of contemporary Yorubaland would be, for many African-Americans, therefore, an exciting discovery. But Yoruba identity provides a good paradigm of the difficulties faced by those seeking an African identity through the human genome project.

The HapMap Project has a site in Ibadan in Nigeria, a city that is predominantly Yoruba, and the ninety or so individuals in thirty families whose genes were sampled there identified themselves as having four Yoruba grandparents. The theory is simple enough. Find SNPs (or sets of them) in haplotypes that are common in Ibadan today, and that have not been found elsewhere. While there will be contemporary Yoruba people who don't have this polymorphism, it is extremely unlikely that anyone that does carry it does not share ancestry with those that do. For someone not descended from the ancestor to have both the haplotype and the SNP, they would both have to have both the same sequence of alleles and have an ancestor who had the same SNP produced by a mutation at exactly the same locus. With 3.6 million SNPs already validated, that is extremely unlikely.

The empirical conditions under which this sort of thing can be reliably done are quite constraining, however. You must first be sure that you have identified SNPs that are in fact distinctive of a certain population. To do that, you have, of course, not only to have detailed knowledge of the genome in Yorubaland, but also knowledge of the genome in other (especially nearby) places. That is the knowledge that the HapMap aims to provide.

Notice that if a SNP originated with a mutation, say a thousand or even five hundred years ago, it may in fact be quite widely dispersed. So, for example, some significant number of the contemporary descendants of that common ancestor might have been living hundreds of miles west of their distant cousins for several centuries. Suppose that the reason you share the Yoruba haplotype is that you are descended from someone who was born in what is now the country of Benin in the early eighteenth century. Then, while your ancestor had cousins in what is now Yorubaland, he never identified as Yoruba. For despite the antiquity of many Yoruba traditions, Yoruba identity itself was developed largely in the last hundred years.

Of course, the city of Ife, now regarded as the origin and heartland of the Yoruba people, was founded at least a millennium ago. But the city-state that was there in the eleventh century was superseded in the fourteenth century by the kingdoms of Oyo and Benin (each of which traced the ancestry of its royal lineage to Ife). As Benin declined, Oyo became the dominant state in the region; by the eighteenth century the kings of Oyo were being paid tribute by the kings of Dahomey, a practice that continued well into the nineteenth century. As a result of warfare and trade in the region—including the trade in slaves—some men travelled widely and took wives from, or had children in, political communities other than their own. Dahomey, a major slave-trading state, sold people from Oyo or Benin into the slave trade. But it was only in the twentieth century that people in south-western Nigeria who spoke related dialects of the Yoruba language, began to think of themselves as a single Yoruba nation. Suppose that your haplotype with some of its distinctive SNPs is very likely derived from someone who has many descendants in Ibadan today. Even if your ancestor had been taken from near Ibadan in the eighteenth century, he would not have thought of himself as Yoruba.

Simply put, the interpretation of haplotype data requires that you know some non-biological history. A couple of thousand years ago, iron-smelting people moved south from somewhere north of the Bight of Biafra, started migrating south and east into equatorial Africa. We call this the Bantu migration because in many of the languages spoken by their descendants from Congo south to the Cape, the word for people is 'Bantu'.

Haplotypes distinctive of that ancestral population could be spread across half the continent. The Ndebele of southern Zimbabwe are largely descendants of migrants from Zululand who escaped from Shaka in the early nineteenth century. Haplotypes distinctive of Zululand might be found in a person whose ancestor was taken into slavery from Zimbabwe and exported through Angola to Brazil.

Because pre-existing ethnic solidarities were strongly discouraged among slaves in the New World, they were deliberately introduced into groups of multiple origins and discouraged from holding on to their mother tongues. As a result, by the nineteenth century many slaves in the western Atlantic would have had ancestors from a variety of African societies. Finding that one has ancestry in one place is interesting, I suppose. But, given those facts, it seems odd to insist that this is where one is really from. More than this, the population that we call African-American is likely to have eighteenth-century ancestors from many parts of Europe and from Native American Indian populations as well. The converse is also true. It has been estimated that there are as many US citizens who identify as white descended from American slaves as there are who identify as African-American. This is a consequence of two things: the fact that you may claim African-American ancestry if just one of your parents is African-American, and the fact that many people who could have claimed that ancestry chose, beginning in the nineteenth century, to identify as white, because their skins were light enough for them to be able to 'pass'. As a result, while not many white Americans are going to go hunting for Yoruba haplotypes in their genomes, perhaps thirty or forty million of them in fact have haplotypes derived from ancestors born in Africa in the last four hundred years.

If you grasp these points you are likely to notice that racial identities in social life tend to be configured in a way that takes account of these sorts of complexities, even while people announce commitments to folk biological theories that are inconsistent with them. In practice, for example, race-like social identities in local contexts are important to patterns of solidarity: in these contexts, people whose (partially genetically determined) physical appearance doesn't fit the physical stereotype of the group are counted in or out in part

on the basis of whether they identify with the interests of the group, in part by their utility to the group. As claims to be able to settle issues of ancestry by genomic analysis become more common, it will be interesting to see whether the appeal of the determinateness and objectivity of scientific claims will come to override more flexible and interest-relative folk understandings; or whether, on the other hand, people will become increasingly clear about the gap between folk races and the interests of biology.[13]

IX

We live in a scientistic civilization. That is one reason, I suspect, that people want the categories they care about to be 'scientific'. There are, as I have suggested, ways in which folk race might be connected with biological facts. But current biology, even after the genome project, is very unlikely to endorse race-like categories that are essentialized (in the psychologist's sense); or to find much interest in human subspecies, given the rather low barriers to gene flow between human groups over the evolutionary timescale. If you want to say there are races, understand race as a social identity, I suggest. But know that as biological and historical knowledge about them is diffused, the criteria of ascription associated with them are likely to change. Know also that as long as they are essentialized they won't correspond to classifications that are likely to be central to theoretical biology, though the statistical distribution of their haplotypes may, from time to time, be of medical interest.

University Center for Human Values
Princeton University
Princeton, NJ 08544
USA
kappiah@princeton.edu

13. In thinking about their ancestral roots, the descendants of my English grandparents will have to bear in mind that most of Granny and Grandpa's haplotypes had descendant tokens in at least England, Ghana, Kenya, Namibia, Nigeria, Thailand and the United States, in the bodies of people with haplotypes recently derived from England, Ghana, India, Kenya, Nigeria and Norway, less than fifty years after they died.

REFERENCES

Akinyanju, O. O. 1989: 'A Profile of Sickle Cell Disease in Nigeria'. *Annals of the New York Academy of Sciences*, 565.1, pp. 126–36.
Appiah, K. A. 2005: *The Ethics of Identity*. Princeton, NJ: Princeton University Press.
Appiah, K. A. and Amy Gutmann 1998: *Color Conscious: The Political Morality of Race*. Princeton, NJ: Princeton University Press.
Bridges, Kenneth R.: *Information Center for Sickle Cell and Thalassemic Disorders*. http://sickle.bwh.harvard.edu/index.html.
Cherva, Alexandra C., Afshin Ameri and Ashok Raj 2002: 'Hemoglobin H Disease'. *eMedicine*. http://www.emedicine.com/ped/topic955.htm
Gelman, Susan 2003: *The Essential Child*. New York: Oxford University Press.
International HapMap Consortium 2005: 'A Haplotype Map of the Human Genome'. *Nature*, 437, 27 October 2005.
Kitcher, Philip 1999: 'Race, Ethnicity, Biology, Culture'. In Leonard Harris (ed.), *Racism*, Amherst, NY: Humanity, 1999.