

## ESSAY



# Kinship: Race relations

Our notions of family, population and race may need revising in the age of personal genomics, argues **Aravinda Chakravarti**.

Genealogical records are currently the system of choice for people tracing their family history. But in the next decade, we will be able to identify many of our relatives by searching a DNA database of personal genome sequences. There are good reasons for switching to DNA: in general, historical records cover at most the past 500 years; our genomes, in contrast, bear the stamp of tens, if not hundreds of thousands of years of history. Even individuals without genealogical records will be able to correctly create a family tree with connections to known relatives, to those they were unaware of, and to relatives so distant that they stretch the meaning of the word 'family'.

Two developments are making it possible for geneticists to begin homing in on the patterns of relatedness between the world's seemingly diverse people. The first is the discovery that we share very recent common ancestors. Until the late 1980s, our early hominin ancestors, from a few million years ago, had been found in many locations, and modern humans were thought to have arisen from the local evolution of these species in different parts of the world.

In 1987, Allan Wilson and his students presented a new and surprising human history: their comparison of only a fraction of our genome — that found in the mitochondria — in a handful of Africans, Asians and Europeans, showed that all living humans are related via a set of common ancestors who lived in Africa about 200,000 years ago. Other studies have since shown that the world beyond Africa was settled even more recently. From 100,000 years ago, descendants of our African forebears spread out to populate other continents (the New World, perhaps as recently as 25,000 years ago), with the lineages from different settler groups eventually mixing through further migration.

The striking implication of this is that all

living humans are mosaics with ancestry from the many parts of the globe through which our ancestors trekked. In other words, each of us has around 6.7 billion relatives.

The second change that is allowing geneticists to piece together human ancestry is remarkable progress in our ability to study DNA. Driven mainly by the desire to find genetic traits that underpin common diseases, extraordinarily rapid technological advances in DNA and computing analysis are allowing geneticists to compare more than a million markers, or variable DNA sites, in people's genomes, and to make such comparisons between hundreds of thousands of people.

The global picture of relatedness that is emerging from DNA studies, in the context of established facts about our recent common ancestry, stands to shatter many of our beliefs about ourselves. In particular, it calls into question existing ideas about populations and race.

## Extended families

Anthropologists and sociologists have conventionally assessed kinship by asking people about their social relationships with others. This provides a notoriously incomplete, perhaps even erroneous, picture of biological relatedness for living humans, and little more than myth or speculation when used to assess the connections between assumed forebears. As the saying goes, "Maternity is a matter of fact, paternity a matter of opinion." In the United States, findings from organizations that assess paternity, for example to persuade fathers to support their disputed yet biological children, suggest that at least 1 in 20 people

don't know the identity of their genetic father. Also, in the case of isolated people such as the Old Order Amish, all individuals share a small group of common founder ancestors. So a social relationship, such as first cousin or parent, is an imperfect guide to their genetic relatedness.

The only way to assess biological kinship with any certainty is to look for the stories of ancestry marked indelibly in a person's DNA.

In the past decade, technological advances have reduced the costs of examining entire human genomes 1,000-fold or more. These have largely been driven by a desire to identify the genes underlying common chronic diseases or adverse

drug reactions. Already, more than a million marker sites in the human genome, have been examined in some 100,000 people to identify more than 300 novel common disease factors. And, increasingly, researchers are sequencing all 3 billion DNA base pairs of the human genome; one international study, the 1,000 Genomes Project, aims to sequence the genomes of 1,000 people over the next two years.

However, the interactions between myriad genetic and environmental factors seeming to underpin most common diseases are proving to be highly complex. As a result, large-scale comparisons of people's DNA may be bringing geneticists closer to understanding how the world's people are related to one another, than to establishing the causes of common diseases.

DNA studies scour the genomes of 'unrelated' patients for common genetic patterns that are absent in other 'unrelated' people without the disease. For example, my

**"The global picture of relatedness that is emerging from DNA studies stands to shatter many of our beliefs about ourselves."**



G. BECKER

from genetic studies is that we are all multiracial, related to each other only to a greater or lesser extent. More detailed data on genetic variation, along with an improved sampling of humanity, are showing continuity in variation across the globe, not abrupt transitions between population-specific sequence patterns. Differential population growth, about 10,000 years ago, based on the evolution of agriculture, technology and politics seems to have made sparse isolates of our species into the 'major' groups of today. In other words, except for immigrants, kinship between two humans seems to be directly related to the geographical distance between their birthplaces.

An even clearer, and unbiased, picture of humanity's genetic diversity and relationships would emerge if geneticists focused on individuals instead of populations. This may involve sampling humans randomly across a grid, and then assessing their individual and group features (such as birth place, parental birth places, language and group affiliations). Genome-wide studies carried out in this way could result in individual identity and kinship coming to define populations rather than the other way around. We could test once and for all whether genetic race is a credible concept.

This would be tremendously exciting. It is bound to stir up our deeply held notions of who we are, where we came from, our history and thus our politics. More often than not, the views of society have shaped science rather than the other way around. In this instance, it may be time for science to reshape the views of society. By dismantling our notions of race and population, we may better appreciate our common, shared and recent history, and perhaps

more importantly, our shared future. Overhauling such concepts in the light of genetic research is particularly important if we are to accommodate the changing face of human groups around the world thanks to increased immigration to distant lands. Such migration has happened before, as our genomes show, only slowly, over 150,000 years. ■

**Aravinda Chakravarti** is at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Broadway Research Building, Room 579, 733 North Broadway, Baltimore, Maryland 21205, USA. e-mail: aravinda@jhmi.edu

See <http://tinyurl.com/8ub4f4> for further reading. For more on Being Human, see [www.nature.com/nature/focus/beinghuman](http://www.nature.com/nature/focus/beinghuman).

colleagues and I have examined the genomes of 16,000 people to identify a region on chromosome 1 that harbours a gene affecting the risk of sudden cardiac death (in combination with many other genetic and environmental factors). But our finding that markers in this region are over-represented among patients compared with those who are disease-free indicates that the disease arises, in part, from shared ancestry. In other words, the disease runs in families, even though the family links may be thousands of years old.

Even putting aside the disease factor, we can uncover both the proximal and remote ancestral relationships of any two of these 16,000 people by comparing the degree and pattern of similarity across their genomes. Indeed, comparisons of millions of markers, and certainly of entire genomes, will identify far more specific relationships between strangers than has been uncovered by the ancestry tests now in vogue.

Increasingly, customers pay companies to convert their DNA into ancestry information. But most if not all such pictures of relatedness are based on markers on the mitochondrial genome (inherited only from mothers) or on the Y chromosome (inherited only from fathers). These represent a minute fraction of our genetic inheritance (less than 1%), so give a highly incomplete picture of relatedness. Companies also tend to compare people's DNA to sequences held in their own private databases, which are currently too small to uncover more than the continental origin of a person's ancestors.

It is not inconceivable that studies investigating the genetic basis of diseases will reveal people's previously unknown cousins, siblings, half-siblings or even parents. Human geneticists are bound by consent forms not to reveal cases of mistaken paternity if they discover them. But if databases of DNA sequence information become publicly available, just as genealogical records are now, people will be able to compare their own genome sequence with those of millions of others. Children born to mothers artificially inseminated by an anonymous donor could

potentially discover their numerous half-siblings. Even in the case of remote relationships, people may interact, perhaps through online social networking, with newly found, distant relatives regardless of their culture, politics and race. Such a scenario is increasingly plausible given people's willingness to share personal information online, for example in social-networking websites.

### Only skin deep

Perhaps the most striking consequence of more and more people having their entire genome examined for genetic variation is the blurring of our concept of discrete human populations. Current thinking, championed by anthropologists and buttressed by old genetic data, is that human populations are intact groups that have had their own language and culture for eons. In fact, the population is thought to define an individual's genetic identity, and kinship between individuals is considered only within the context of this or that group. It's within this context that people trace their ancestry using genealogies or ancestry tests, and that the discovery that President Barack Obama is related to Dick Cheney makes news.

Currently, the population view dominates in genetics because researchers sample clusters of individuals from distantly related groups. The clearly observable, or measurable, physical and genetic differences between people are especially marked when people from the peripheries of the spectrum of human variation are compared — so, for example, when Africans are compared with Europeans or Asians.

Race has long been a socio-political construct. But by focusing on the effects of natural selection on genes whose effects are visible, and sampling people from the extremes of human diversity, geneticists have unwittingly (and sometimes wittingly) added credence to society's views on separateness by genetically characterizing racial categories.

However, the current picture emerging

**"We are all multiracial, related to each other only to a greater or lesser extent."**