Code of Many Colors

Can researchers see race in the genome?

Christen Brownlee


Historian Frank W. Sweet of the University of Florida in Gainesville recounts the classic rags-to-riches tale of Louetta Chassereau, an early 20th-century socialite. As a baby, Chassereau was adopted from an orphanage by a well-to-do white couple. She later married a wealthy man, and her children attended the best white-only schools.

However, a dilemma developed when Chassereau's husband died, leaving everything in his will to his beloved wife. Enraged, her husband's relatives contested the will. The reason? Although people in her community had always thought of her as white, "Louetta had started life as a Black baby," says Sweet in a recent essay. Because Chassereau was born of black parents, according to an antimiscegenation law of the time, Chassereau legally could marry only a black man. The white family claimed that she had no right to the fortune.

Although the courts ruled in Chassereau's favor in 1940, saying that her life's path had made her "irrevocably white," her in-laws remained unconvinced.

In the past 65 years, defining race hasn't become less ambiguous. While it's abundantly clear that race exists from a sociological standpoint—racism wouldn't take place without it—does that categorization also exist biologically?

Current genetic research hasn't yet come up with a black-and-white answer. Nevertheless, understanding the biology underlying perceptions of race could have dramatic implications.

Racy subject

It's difficult to get most scientists even to say the word race when referring to people. That's because in traditional scientific language, races are synonymous with subspecies—organisms in the same species that can interbreed but nevertheless are distinctive genetically.

Many species split into subspecies after being separated geographically for an extended amount of time. During generations of genetic mixing within but not between the isolated groups, some of each group's genes develop slightly different versions, or alleles. Scientists often use a rule called Wright's F statistic to judge whether separate groups are actually subspecies. If 25 percent or more of one group's alleles are different from another's, then by F-statistic standards, the two groups are considered subspecies. A difference of 100 percent would separate them into distinct species.

Subspecies, or races, exist for many animals—for example, the alleles in some populations of grey wolves score up to 70 percent on the F-statistic scale. However, the groups of people considered to be of different races have allelic differences of at most 15 percent, too little to constitute subspecies.

To the nonscientist, however, race clearly is a meaningful term, says Vivian Ota Wang of the Ethical, Legal, and Social Implications Research Program at the National Human Genome Research Institute (NHGRI) in Bethesda, Md. The concept seems to depend on a collection of physical features, "like a checklist," she says, "so that people can categorize each other into groups." Items on the list might
include skin tone, hair texture, and the shapes of eyes, noses, or lips.

Most people don’t carry a conscious perception of the checklist. Wang says that race has a lot in common with Supreme Court Justice Potter Stewart’s famous definition of pornography: We know it when we see it.

About 100,000 years ago, defining race wasn’t an issue—all early humans lived in Africa and had similar characteristics. That relatively small population of recently evolved humans carried the majority of alleles present in people today.

But over the next 50,000 years or so, as humans separated into groups, slight differences among populations crept into the genome. First, as waves of emigrants left Africa and spread throughout the world, our ancestors took slightly different groups of alleles with them. Just as each handful of jellybeans scooped out of a jar might have a different mix of flavors, every group of migrant humans carried a slightly different array of alleles.

Later, when roaming humans settled into permanent residences on different continents, new genetic mutations gradually built up within groups as they adapted to their distinct environments. Because people mated most frequently with others from the same region, each population developed its own set of mutational differences, some influencing survival and some being just genetic quirks.

Share, share alike

According to Lisa Brooks, a geneticist at the NHGRI, the genetic differences within each population take several forms. Some people have certain segments of DNA wedged within stretches that run without interruption in other people. Conversely, genetic pieces present in many people are missing in others. Also, stretches of DNA can be flipped so that they read backwards, or they might contain small repeated segments, called microsatellites, that vary in number from person to person.

The genetic variation that most interests Brooks is called a single nucleotide polymorphism, or SNP (pronounced “snip”). It's a one-letter change in the string of DNA components that go by the letters A, C, G, and T. For example, where one person might have a section that reads TAAACA, another person's section might read TAAAGA.

Most SNPs occur in places in the genome that aren't used for making proteins—the so-called junk DNA. But the few SNPs that land squarely in a gene or in a regulatory region near a gene can alter characteristics influenced by that gene—for example, physical appearance or propensity for disease.

Sets of adjacent SNP alleles in the same chromosome region are called haplotypes. Through a collaboration called the International HapMap Project, researchers around the world are recording and analyzing the SNPs present in four populations: Utah residents with European ancestry, a Nigerian population called the Yoruba, Han Chinese in Beijing, and Japanese inhabitants of Tokyo.

Although the study wasn't designed to probe the genetics of race per se, one of the major findings from HapMap so far, says Brooks, is the enormous similarity between the four groups’ SNP patterns. In an interview, Brooks illustrated her point by drawing on a piece of paper four largely overlapping circles, with only slim slivers of each ring peeking out at the edge.
Brooks points to the large area where the circles overlap and explains that about 85 percent of variation is shared by all populations. By default, the genetic variation in the slivers includes the alleles that lead to differences among populations, including those on the typical racial checklist. "Superficial traits, like skin color or hair texture, aren't typical in their patterns of variation of [most of] the genome," says Brooks.

This concept can be hard to grasp for people who believe that racial groups are fundamentally different genetically, says Georgia Dunston of Howard University's National Human Genome Center in Washington, D.C.

Dunston studies how the human immune system distinguishes between a person's tissues and foreign material, such as a splinter, a bacterium, or a transplanted organ. The genes responsible for this recognition are called histocompatibility genes. Having similar histocompatibility genes is a major factor in successful organ transplants.

In tissue matching, a bastion of genetic differences between people, Dunston finds that race is not the determining factor. "We have this thinking in America that there are some deep differences in biology between whites and blacks, that tissue in whites is more similar to [tissue in] whites than tissue in blacks," she says. "But when we look at the genetics, because of the tremendous variation in all groups, and especially in the group called 'black,' it's not uncommon at all to find two blacks who could be very different from each other."

In some cases, a black organ donor, Dunston adds, can better match a white recipient than a black one.

**What's the difference?**

Nevertheless, no geneticist can overlook the slim fringes on Brooks' overlapping-circles diagram. According to Brooks, the longer people in a population have mated in close association, and so share the same ancestors and kin, the more similar will be their genes—and all the traits they encode—in that 15 percent of the genome.

Race and family origin aren't entirely synonymous in modern times, when people can relocate around the globe. However, many researchers have found that the distribution of certain genetic variations can lump people into ancient ancestral groups uncannily similar to what nonscientists call races.

For example, Noah Rosenberg of the University of Southern California in Los Angeles and his colleagues published a study in 2002 that analyzed the number and type of microsatellite variations in the DNA of 1,056 people from 52 populations around the world.

Rosenberg's team masked any information about the study volunteers' ancestral backgrounds and then plugged the microsatellite information into a computer program that clusters people by genetic similarities. Six main clusters emerged.

After restoring individuals' ancestry data to the files, the researchers found that five of the six microsatellite clusters corresponded with geographic regions: Africa, Eurasia (Europe, the Middle East, central and south Asia), east Asia, Oceania (islands of the central and South Pacific), and the Americas (specifically native Americans). The sixth and smallest cluster linked to an isolated group of mountain-dwelling Pakistanis known as the Kalash.

The scientists weren't surprised that people's genetic mutations usually lump them into continental groups. For much of history, people have been land bound and so have mated mostly with people from the same continent.

However, Rosenberg says that he was surprised that he and his colleagues found it impossible to predict with certainty which combination of gene variants any specific person in each cluster had. The computer runs couldn't determine, for example, exact shades of skin color or types of hair texture for individuals.

"In a lot of classical anthropological views of race, race is thought to be a quality predictive of a large variety of traits about a person. We found that for any given person, it's not possible to predict accurately which [variant] they have at any particular site in the genome based on their group membership," Rosenberg says.

Neil Risch of the Stanford University School of Medicine and his colleagues recently used a similar method to come to a very different conclusion. Using microsatellite information from another study that had looked for a genetic link with hypertension in several U.S. populations, Risch's team ran data from 3,636 people through a computer...
program similar to Rosenberg's. However, instead of searching for clusters based on geography, Risch and colleagues compared clusters from the genetic data with self-described race/ethnicity categories.

The genetic data sorted into four categories—white, African American, east Asian, and Hispanic—which neatly matched what each person had checked on a form at the beginning of the study. Only five people had results inconsistent with their self-described race/ethnicity, giving an error rate of 0.14 percent, the team reports in the February *American Journal of Human Genetics*.

"This shows that people's self-identified race/ethnicity is a nearly perfect indicator of their genetic background," says Risch.

**Racism realism**

Risch's results have stirred up controversy among many geneticists. For instance, Mark Shriver of Pennsylvania State University in State College says that Risch's method "can overcluster people," making associations between individuals and their race that don't exist with other types of analyses. Shriver and others haven't found similar clusters when they applied a different computer program to similar data.

Shriver also contends that the study's separation of people into four racial groups shrinks the natural range of genetic variation, making people within each group seem more alike than they really are.

Rather than there being clear racial lines, says Shriver, "there's really a continuum of variation across the globe." If researchers sampled only people in Africa and Sweden, the genetic differences between the two groups would be striking. However, a sampling of people from Africa, Sweden, and everywhere in between would reveal only small differences between each population and its neighbors. "You won't see a place where you'll say, 'There's the racial divide,'" says Shriver.

Nevertheless, Shriver works with a company that uses what variation there is among populations to trace people's ancestry. The company, DNAPrint Genomics in Sarasota, Fla., starts with DNA from a customer's inner cheek. After comparing the sample's genetic markers with those in a data set collected from people around the world, the company estimates what percentage of the person's ancestry is African, East Asian, European, or Native American.

The results can be surprising. When Shriver, who considers himself to be white, analyzed his own DNA, he found that it contained the Duffy null allele, found only in descendants of sub-Saharan Africans. "The test estimated that I have 11 percent west African ancestry," says Shriver.

In spring 2003, Shriver and his colleagues applied the test to an urgent task—they were instrumental in catching a Louisiana serial killer. After analyzing DNA from semen at the crime scenes, Shriver and his colleagues estimated that the killer was 85 percent African and 15 percent Native American. Officers eventually arrested Derrick Todd Lee, a black man whose DNA matched that left at the scenes. As testament to the uncertainty of eyewitnesses, Lee was convicted although several people had reported seeing a white man at the scene of several of the murders.

Although an ancestry test had put police on the right track in this case, Shriver expresses concern about the test's potential for misuse in other social realms. One danger, he notes, is attempting to correlate ancestry with qualities such as intelligence, athletic performance, or musical ability. "It's hard to know what the right use is. We have to be vigilant," he says.

Researchers may never nail down a precise connection between race and genetics, but there's little chance that the concept of race will ever go away, says Charles Rotimi of Howard University's genome center. Over the centuries, some people have used race to set discriminatory categories and then give themselves privileges and take privileges from others. People who have been advantaged by racism aren't likely to give it up, Rotimi says.

"Look at the Hutus and the Tutsis," he adds, referring to two Rwandan tribes that have been fighting each other for decades. "You don't need genetics to be racist."

*Part II: "The Race to Prescribe," is available at* http://www.sciencenews.org/articles/20050416/bob8.asp.
Your article gives a simplistic and generally inaccurate account of the relationship between Fst [also called Wright's F statistic] and race/subspecies/species. Fst reflects the relative amount of total genetic variation between populations. While there is bound to be a correlation between Fst and species status, Fst is not normally used to define species. An Fst can be lower between populations of North American grey wolves and coyotes, two different species, than between many populations of the same species, including humans, depending on which genetic loci are considered.

John Goodrum
Port Angeles, Wash.

Researcher Sarah Tishkoff of the University of Maryland in College Park, a source for the article, agrees that Fst is not a good indicator of subspecies. She also says the article misrepresented Fst, or Wright's F statistic, in that an Fst of 15 percent does not mean that two populations differ in 15 percent of their genome. She adds, however, that the story is accurate in saying that the genetic variability between races is much less than that between subspecies in other animals.—C. Brownlee

If you have a comment on this article that you would like considered for publication in Science News, send it to editors@sciencenews.org. Please include your name and location.

To subscribe to Science News (print), go to https://www.kable.com/pub/scnw/subServices.asp.

To sign up for the free weekly e-LETTER from Science News, go to http://www.sciencenews.org/pages/subscribe_form.asp.

References:


Further Readings:


Additional information about DNAPrint genomics can be found at http://www.dnaprint.com/.

Sources:

Lisa D. Brooks
National Human Genome Research Institute
National Institutes of Health
5635 Fishers Lane, Suite 4076
Mailstop Code 9305
Bethesda, MD 20892-9305