**Scientific American: How Accurate Are Online DNA Tests?**
**Geneticist and author Adam Rutherford examines the evidence**
**By Adam Rutherford on October 15, 2018**

The age of consumer genomics has arrived. Nowadays you can send a vial of your spit in the mail and pay to see how your unique genetic code relates to all manner of human activity—from sports to certain diets to skin cream to a preference for fine wines, even to dating.The most widespread and popular companies in this market analyze ancestry, and the biggest of these are 23andMe and AncestryDNA, both with more than five million users in their databases. These numbers dwarf the numbers of human genomes in scientific databases. Genetic genealogy is big business, and has gone mainstream. But how accurate are these tests—truly?

First, a bit of genetics 101. DNA is the code in your cells. It is the richest but also most complex treasure trove of information that we've ever attempted to understand. Three billion individual letters of DNA, roughly, organized into 23 pairs of chromosomes—although one of those pairs is not a pair half the time (men are XY, women are XX). The DNA is arranged in around 20,000 genes (even though debate remains about what the definition of a gene actually is). And rather than genes, almost all of your DNA—97 percent—is a smorgasbord of control regions, scaffolding and huge chunks of repeated sections. Some of it is just garbage, left over from billions of years of evolution.

Modern genetics has unveiled a picture of immense complexity, one that we don't fully understand—although we are certainly a long way from Mendel and his pea experiments, which first identified the units of inheritance we know as genes. Throughout the course of the 20th century we gained a firm grasp of the basics of biological inheritance: how genes are passed from one generation to another and how they encode the proteins that all life is built of, or by. In the 1980s we identified genes that had mutated, making faulty proteins, which could cause terrible diseases such as cystic fibrosis or muscular dystrophy, for example.

By 2003, the Human Genome Project had delivered the human DNA sequence in its entirety. One of the most important by-products of that endeavor was the advent of technology that allowed us to read DNA at unprecedented speed and for ever-decreasing costs. We can now pump out the genomes of hundreds of thousands of people for peanuts, and with that data comes greater and greater perspicacity into the profound questions of inheritance, evolution and disease. There's effectively infinite variation in human genomes, and scrutinizing our DNA helps us to understand what makes us human as a species and as individuals.

With the plummeting costs of gene sequencing came commercial interests. All of a sudden any company could set up shop, and in exchange for some cash and a vial of saliva, could extract your DNA from the cells in your mouth and sequence your genome. Alongside the behemoths 23andMe and AncestryDNA, dozens of companies have done just that.

There are two potential issues arising from the question of their results' accuracy. The first is somewhat trivial: Has the sequencing been done well? In critiquing this business, it seems fair to assume the data generated is accurate. But there have been some bizarre cases of failure, such as the company that failed to identify the sample DNA as coming not from a human, but from a dog. One recent analysis found 40 percent of variants associated with specific diseases from "direct to consumer" (DTC) genetic tests were shown to be false positives when the raw data was reanalyzed.

Assuming the tests are done accurately, some discrepancies can still arise from differences in the companies' DNA databases. Almost every DTC genetic test does not sequence your entire genome, but instead looks at positions in your DNA that are known to be of interest. When I was tested by 23andMe, they proclaimed I do not carry a version of a gene that is associated strongly with red hair. Another ancestry company said I did. This merely reflects the fact one company was looking at different variants of the gene that code for ginger hair.

If we assume the data generated is accurate, then the second question that arises is on the interpretation. And this is where it gets murky. Many of the positions of interest in your DNA are determined by experiments known as Genome Wide Association Studies, or GWAS (pronounced gee-woz). Take a bunch of people, as many as possible, that have a shared characteristic. This could be a disease, like cystic fibrosis (CF) or a normal trait, say, red hair. When you sequence all their genes, you look out for individual places in their DNA that are more similar within the test group than in another population. For CF, you would see a big spike in chromosome 7 because the majority of cases of CF are caused by a mutation in one gene. For redheads, you'd see 16 or 17 spikes very close to one another, because there are multiple variants in the same gene that all bestow ginger locks. But for complex traits like taste or ones relating to diet or exercise, dozens of variants will emerge, and all of them only offer a probability of a predisposition toward a certain behavior as a result of your DNA, as measured in a population. This even applies to something as seemingly straightforward as eye color: A gene variant that is associated with blue eyes is still only a probability that you will have blue eyes, and it is perfectly possible to have two blue-eyed genes and not have blue eyes.

Genetics is a probabilistic science, and there are no genes "for" anything in particular. I have severe reservations about the utility of genetic tests that indicate one individual's propensity for certain conditions outside of a clinical setting; if you don't have a PhD in genetics, these results can be misleading or even troubling. Even if, as I do, you carry a version of a gene which increases the probability of developing Alzheimer's disease, most people with this variant do not develop the disorder, which is also profoundly influenced by many lifestyle choices and some blind luck. There is little a geneticist can tell you with this information that will outweigh standard lifestyle advice: Don't smoke, eat a balanced diet, exercise regularly and wear sunscreen.

When it comes to ancestry, DNA is very good at determining close family relations such as siblings or parents, and dozens of stories are emerging that reunite or identify lost close family members (or indeed criminals). For deeper family roots, these tests do not really tell you where

your ancestors came from. They say where DNA like yours can be found on Earth today. By inference, we are to assume that significant proportions of our deep family came from those places. But to say that you are 20 percent Irish, 4 percent Native American or 12 percent Scandinavian is fun, trivial and has very little scientific meaning. We all have thousands of ancestors, and our family trees become matted webs as we go back in time, which means that before long, our ancestors become everyone's ancestors. Humankind is fascinatingly closely related, and DNA will tell you little about your culture, history and identity.

*The newsmagazine of the American Historical Association*
PERSPECTIVES ON HISTORY

**WHAT ARE YOU?**
**Historians Confront Race, Genealogy, and Genetics**
**Sadie Bergen | Feb 16, 2018**

The advertisement opens on a young woman named Lezlie. A collage of family pictures fills the screen, and she explains that all her life people have asked her, "What are you, what are you, what are you . . . Asian, or Moroccan, or something else?" Lezlie is selling AncestryDNA, a service from the genealogy website ancestry.com that promises to reveal one's extended family tree and genetic background, broken down into a handy pie chart of ethnic percentages. Lezlie, it turns out, is 46 percent African, 25 percent British, 2 percent Asian, and 17 percent "other." Even with that nebulous 17 percent, Lezlie proclaims a newfound sense of identity.

The idea that our biology contains definitive answers about who we are is seductive. Genetic testing services like AncestryDNA bring a veneer of scientific objectivity to racial and ethnic identity, cementing what many people already intuitively believe. Science, for its part, has done little to counter the notion that race is just another iteration of biological human difference; in our current "genomic era," scientists and geneticists routinely use race as a tool to make sense of medical risks and genetic distinctions within human populations. For historians, humanists, and social scientists, this acceptance of biological essentialism, both in popular culture and in the scientific realm, presents a unique challenge: How to convincingly articulate to the public that race is a social construct rather than a biological fact?

In conversations during the 2018 AHA annual meeting in Washington, DC, historians and social scientists pushed back against the narrative that genes are carriers of essential truths about identity and heritage. In the social sciences and humanities, a consensus that racial categories are socially determined was reached in the mid-20th century. But this followed a long era of

eugenics, when humans were racially categorized based on perceived physical differences and supposed genetic intellectual capacities. Since then, social scientists and humanists have spelled out the implications of using biology to map personal histories, affirm social hierarchies, and construct identities. Historians, especially, can attest to the dark consequences of viewing race as biological.

Yet the idea has lingered in cultural and scientific spheres and was reinforced in the late 20th century by medical geneticists' efforts to identify genomic risks for disease. At one annual meeting session, "Science and Difference in History: Biology, Genetics, and the Politics of Race," sociologist Joan Fujimura (Univ. of Wisconsin–Madison) explained that cultural assumptions about race were embedded in late 20th-century technological advancements in genetics research like the Human Genome Project, even as social scientists increasingly emphasized the social construction of race. The Human Genome Project, between 1990 and 2000, mapped the complete DNA sequence of the human genome and identified its component genes.

When searching for genetic markers hypothesized to be associated with complex diseases like asthma, diabetes, and cancer, "geneticists built shortcuts," said Fujimura, in order to draw connections between genes, diseases, and the human populations they coalesce within. One "shortcut" was to draw upon prevailing, socially determined racial categories to organize, analyze, and interpret research findings. The "populations" geneticists studied, however, were not neutrally constructed, but, like all knowledge, a "product of situated priorities, actions, and decisions," said Fujimura. In other words, the decisions scientists made when drawing connections between race and genetics emerged out of a particular historical context.

At another session, titled "Racial Sciences, Old and New," historian Daniel Smail (Harvard Univ.) called the Human Genome Project a "productive failure." Its results, he noted, were not what people expected. Rather than unveiling the singularity of the human species and its vast internal variations, scientists found that the human genome was neither as large as nor as distinct from other living organisms as they had expected. Humans share more than 99 percent of their genes in common—the only race revealed by the genes was the human race.

Contests over the scientific reification of racial categories, however, began long before modern genetics research. Michael Yudell (Drexel Univ.) elucidated part of this history at the session "Science and Difference in History" with the example of W.E.B. Du Bois, who questioned the idea that health disparities stemmed from racial differences at the turn of the 20th century. Du Bois instead suggested that the racialized "conditions of life"—poverty, a lack of education, unsanitary living and working conditions—led to racial disparities in health. Thus, while race is not biologically determinative, it can masquerade as such when employed reflexively in scientific research and medical practice. Yudell explained that scholars who criticize the use of racial categories as proxies for genetic diversity are reinvigorating Du Bois's critiques, revealing the historical inertia of the biological concept of race despite a century of scientific and technological advancements.

Cultural assumptions about race were embedded in late 20th-century technological advancements in genetics research like the Human Genome Project.
Direct-to-consumer genetic testing services have profited from this inertia. Companies like AncestryDNA and 23andMe can only communicate their findings via preexisting cultural scripts that assume race, ethnic heritage, genealogy, and geography to overlap seamlessly. In his presentation during “Racial Sciences, Old and New,” historian Patrick Geary (Institute for Advanced Study) argued that genetic testing services are “too eager to equate ethnic with geographic origins.” Once the layers are peeled back, it becomes clear that the neat pie charts of ethnic and geographic compositions that testing services like AncestryDNA generate are mired in historical contingency.

Migration and border fluctuations have been constants throughout human history, and the regional labels we have divided up the continents with—eastern Europe, South Asia, and so on—are similarly context-dependent. If DNA proves, for example, that one’s ancestors were eastern European, then a host of follow-up questions would need to be answered to determine exactly what that means: how far back does the test reach? How do these tests define the point of origin—spatially and temporally—for a family line? Depending on the answers to these and other questions, the label “eastern European” could point to any number of ethnic groups and associated heritage, history, and custom—all important components of what we call identity.

Further, as Warwick Anderson (Univ. of Sydney) explained in his presentation during “Science and Difference in History,” understandings of race and ethnicity, including its biological dimensions, change when the locus of knowledge production shifts to the global South. Anderson studies the “intellectual currents” of racial science: the ways race has been conceived and human biology interpreted in the global South. Taking an alternative geographic and social perspective, argues Anderson, unsettles the historical narrative of racial science and genetics that traditionally relies on Western knowledge production. Anderson has found that while approaches to understanding human differences were still racialized in the global South, they focused less on strict categorization of the races. In his article “Racial Conceptions in the Global South” (Isis, 2014), Anderson writes that, instead, settler societies shared an interest in “the intermingling and plasticity of races and inquiries into the formation of new races.” For instance, in 20th-century Brazil, European immigration and interracial reproduction were considered positive ways of “‘whitening’ the nation.”

Historian of science and medicine Keith Wailoo (Princeton Univ.) has pointed out in Genetics and the Unsettled Past: The Collision of DNA, Race, and History (2014) that genetic testing services are “as much about making meaning in the present as . . . about the past.” People get DNA from both their parents, yet genetic analysis often draws upon only one of these lines. Thus from the outset, people curious about their heritage must make a choice about the histories they want to claim. Results are also constrained by the databases they draw upon. Like any information system built by humans, genetic data and the biases behind its collection are circumscribed by historical context. Limited genetic testing in non-Western or

underdeveloped parts of the world means that less data from those regions is available for comparison. As Wailoo writes, the databases themselves “shape what past will be found.”

For all the ways genetic tests are more complex than their taglines promise, they can also provide an unexpected avenue for reckoning with the past. Sociologist Alondra Nelson’s (Columbia Univ.) presentation during “Science and Difference in History: Biology, Genetics, and the Politics of Race” focused on marginalized communities who have been “most subjugated by scientific racism” and yet have found “pockets of agency” by using genetic testing as a way to stake political claims. Such “reconciliation” projects include the #GU272, which attempts to identify through genetic testing and genealogical research the descendants of the 272 enslaved persons sold in 1838 to help finance Georgetown University. As a result of this effort, the university has promised an “edge” in admissions to descendants.

As the work of #GU272 reveals, it can be productive to leverage the complex ways genetics and race overlap. But interpretations of genetic test results must be grounded in historical detail and contingency. As the annual meeting presenters demonstrated, race and genetics are bound up together, and historians are ideally positioned to begin untangling them. The popularity of services like AncestryDNA demonstrates that people like Lezlie are searching for answers to “who they are.” But the answers they currently get only reinforce the spurious connection between race and genetics, something the world of social scientists and humanists has long abandoned. Historians must provide better, alternative answers, pushing back against the biological concept of race while honoring the roles that heritage, genealogy, and geographical origins play in determining “who we are.”

Sadie Bergen is a PhD student in the Center for the History and Ethics of Public Health at Columbia University.

NewScientist

**The father of all men is 340,000 years old**
**Magazine issue 2908 , published 16 March 2013**
**By Colin Barras**

Albert Perry carried a secret in his DNA: a Y chromosome so distinctive that it reveals new information about the origin of our species. It shows that the last common male ancestor down the paternal line of our species is over twice as old as we thought.

One possible explanation is that hundreds of thousands of years ago, modern and archaic humans in central Africa interbred, adding to known examples of interbreeding – with

Neanderthals in the Middle East, and with the enigmatic Denisovans somewhere in southeast Asia.

Perry, recently deceased, was an African-American who lived in South Carolina. A few years ago, one of his female relatives submitted a sample of his DNA to a company called Family Tree DNA for genealogical analysis.

Geneticists can use such samples to work out how we are related to one another. Hundreds of thousands of people have now had their DNA tested. The data from these tests had shown that all men gained their Y chromosome from a common male ancestor. This genetic “Adam” lived between 60,000 and 140,000 years ago.

All men except Perry, that is. When Family Tree DNA’s technicians tried to place Perry on the Y-chromosome family tree, they just couldn’t. His Y chromosome was like no other so far analysed.

Deeper roots

Michael Hammer, a geneticist at the University of Arizona in Tucson, heard about Perry’s unusual Y chromosome and did some further testing. His team’s research revealed something extraordinary: Perry did not descend from the genetic Adam. In fact, his Y chromosome was so distinct that his male lineage probably separated from all others about 338,000 years ago.

“The Y-chromosome tree is much older than we thought,” says Chris Tyler-Smith at the Wellcome Trust Sanger Institute in Hinxton, UK, who was not involved in the study. He says further work will be needed to confirm exactly how much older.

“It’s a cool discovery,” says Jon Wilkins of the Ronin Institute in Montclair, New Jersey. “We geneticists have been looking at Y chromosomes about as long as we’ve been looking at anything. Changing where the root of the Y-chromosome tree is at this point is extremely surprising.”

Digging deeper, Hammer’s team examined an African database of nearly 6000 Y chromosomes and found similarities between Perry’s and those in samples taken from 11 men, all living in one village in Cameroon. This may indicate where in Africa Perry’s ancestors hailed from.

Older than humanity

The first anatomically modern human fossils date back only 195,000 years, so Perry’s Y chromosome lineage split from the rest of humanity long before our species appeared.

What are the implications? One possibility is that Perry’s Y chromosome may have been inherited from an archaic human population that has since gone extinct. If that’s the case, then

some time within the last 195,000 years, anatomically modern humans interbred with an ancient African human.

There is some supporting evidence for this scenario. In 2011, researchers examined human fossils from a Nigerian site called Iwo Eleru. The fossils showed a strange mix of ancient and modern features, which also suggested interbreeding between modern and archaic humans. “The Cameroon village with an unusual genetic signature is right on the border with Nigeria, and Iwo Eleru is not too far away,” says Hammer.

Chris Stringer at the Natural History Museum, London, was involved in the Iwo Eleru analysis, and says the new Y chromosome result highlights the need for more genetic data from modern-day sub-Saharan Africans. “The oldest known fossil humans in both West Africa at Iwo Eleru and Central Africa at Ishango [in Democratic Republic of the Congo] show unexpectedly archaic features, so it certainly looks like we have a more complex scenario for the evolution of modern humans in Africa.”

Journal reference: American Journal of Human Genetics, doi.org/kp4