Robin Bhattacharya
MSc in Race, Ethnicity and Post-Colonial Studies
London School of Economics and Political Science

Human Population Categories in Genomic Studies and Racialisation

August 2010

Contents:

Introduction

The difficulty in writing about 'race' and 'genes' is that both terms have ambiguous meanings and are used differently, depending on their discursive field. And in these cases the discursive fields are numerous ('politics', sociology, biology as well as the media and colloquial conversation) - separated by educational level, discipline, interests and political history, but also full of surprising interconnections.

Additionally, to compound things, both terms have a common history, but this was seemingly radically split by the mid-20th century, in the wake of the Nazi genocides.

In order to comprehensively understand the current interrelation between race and genetics, I will de-contextualise racial, ethnic and national categories and analyse how race and the genetic imaginary are co-constituents to each other.

To make this methodological critique of the use of race in genetics more tangible, I will examine the process at work in a case study on my own genome.

My contrapuntal reading of the results from my genomic testing and their scientific interpretation, reveals a familiar pattern of *exception*, *intervention* and *instrumentalisation* in the production of knowledge that is intrinsic to modernity.

The Thing with Race in Genomics

In popular understanding, the relationship between race and genetics is reduced to a twofold interlink, expressed in the question whether 'race' is 'biological' or not. This misses the point that both 'race' and the 'biological' are socially constructed categories, developed in sync during the 18th century.

Both came to prominence in Europe amid the concerns brought about by the biopolitics of its age: The dwindling birth-rate as discovered by the newly introduced census and the encounter with populations perceived to be 'other' from the imperialist European. As Michel Foucault notes, there was no "bio-" before bio-politics, no "life itself" (Rose, 2007) before the intervention in what is seen as the factness of life itself. As Foucault focuses on the institutionalisation of biopolitics in the European metropolis, the significance of 'race' as a constituent 'other' to the population of the homeland is secondary to his analysis in 'History of Sexuality'. But in his lectures at the College de France (2008), where he specifically engaged with the issue of what is 'biopolitics', Foucault recognised the centrality of 'race' in the establishment of biopolitics, including the science of biology, at the beginning of European modernity.

From that perspective, it would not seem surprising that the Nazis, who put biology at the centre of its ideology, turned all politics into bio-politics and created the most 'modern' nation-state of its time, where every body would be incorporated into the national project – or, at the point of their incorporability, their life be exterminated, but not without first extracting all possible use-value from these bodies, be that slave-labour, gold-teeth and hair, or scientific knowledge from using them in

biological experiments.

But Nazi Germany, while having the unenviable/sad position of leading in the implementation of eugenic policy by the 1930s, its biological science was not extraordinary in the sense of being fundamentally different from the work of other biologists of the time interested in 'genetics', the scientific study of the newly discovered smallest unit in biology called 'gene'.
In the US, the experiments of Charles Davenport of the Coldspring Harbor laboratory resulted not only in the forced sterilization of 60'000 persons, but his recommendations on 'miscegenation' led to the formulation of US immigration laws (Immigration Act of 1924) that stayed in place until 1965 when national quotas for immigration were abolished.

After World War II biology sought to vehemently distance itself from the horrors produced by scientists in Nazi Germany, which led to the publication of the UNESCO statements on race. These partly and selectively condemned the use of race in biology, specifically in regards to racial hierarchy. A political project, it was spurned by sociologists and anthropologists and aimed at conclusively doing away with racial discrimination by addressing its scientific basis. But as more biologists were involved in the 1st revision, it became clear that racial categories were too deeply embedded in the biological science even Post-WWII and so the statements fell short of denouncing the use of race per se. In its place, the division of discourse was institutionalised in the statement, by declaring that 'races' might be used in biology, that they *do* signify differences among humans – but that these differences should only be of relevance to biological scientists and remain inconsequential to

politics and society as a whole (Reardon, 2005). Such a discursive division in regards to a concept, is symptomatic of the understanding of science as separate from politics and society at large - and persists to this day when scientists using racial categories in their research claim to do so in a non-political way.

A decade ago the first results from the sequencing of 'the' human genome, as it was referred to at the time, were published. The Human Genome Project (HGP) was and still is ($ 2.7 billion), along with the Large Hadron Collider (€ 3 billion), one of the largest publicly funded research initiatives ever undertaken.

Its initiators hoped that once we were able to read the sequence of the nucleotide bases, we could understand it and discover the 'genes' – those elusive little units first proposed by Gregor Mendel that were supposed to determine who we are, what we look like, our physical and mental abilities and our susceptibility to diseases. Albeit, as it was based on a reductionist functionalist view of human nature, this hope was not shared by all scientists (Lewontin, 2000).

When the 'complete' human genome was finally presented in 2003, it raised as many questions as it answered. The issue of whose genome this actually was became apparent: In fact it was not of one person, but a composite from several donor's DNA.

In the process of sequencing it was claimed that 99.9% of the genome was shared by all humans, confirming the common African ancestry of all humanity previously hypothesised from fossil findings such as Lucy. But ever since the availability of

DNA sequencing technologies, population geneticists had focused on the 0.01% differences between human genomes to uncover the routes of early human migration and find correlations to linguistic families (Cavalli-Sforza et al., 1994). What became a complimentary project to the HGP, the Human Genome Diversity Project (HGDP) aimed at mapping the differences between human 'populations', a geneticist's term for a group of humans that are genetically similar and therefore related. To do this, the teams of Cavalli-Sforza set out to collect samples of human groups that were seen as having evolved isolated from others, and from the distribution of differences found between their genomes, the pattern how one group evolved/migrated out of the other was to be implied. While they managed to do this to a certain extent, in order to deepen the hierarchy of the 'haplogroups' differentiated, more genetic samples had to be collected. But the collection of samples from 'genetic isolates', often indigenous peoples in former European colonial territories, was met with more opposition than the researchers had anticipated, both from supposed sample donors, as well as the anthropologists who were supposed to collect them. This although they put a lot of weight in the development of their ELSI (Ethical, Legal, Social Issues) guidelines, which have become the standard in genomic studies with public participation. But these failed to address the issue of agency posed by the assignment of the roles of researcher and object-of-research, along with the ensuing methodological and authorship problems, that was at the core of the criticisms levelled at the HGDP (Reardon, 2005).

The project was eventually moved into and replaced by the international HapMap, which takes the form of a public online library: The sampling and sequencing of

more genetic data has devolved into various trans-local and national genomic diversity projects, which use the library and add to its collection, while private commercial companies such as ancestry testing services also download its data to complement their own, proprietary genetic databases.

Although the results from the research leading up to the HGDP were interesting – and received a lot of media attention - they were ultimately less consequential than they could have been, because they failed to grapple their own methodological issues. If the issues underlying the opposition, which led to the projects failure, were taken into account, the reality of human genetic variation could have been set against the reification of race. Instead the biggest change effected by the HGP and HGDP was the push in the development in genotyping technologies caused by their subsidisation on such a massive scale and their subsequent reduction in price. This in turn has led to the propensity in genetic variation studies as it is now – including this study that could not have been made ten years ago and is contingent on the availability of the commercial genomic ancestry testing companies that have sprung up in the meantime.

Over the last two decades, hundreds of scientists have conducted studies on the genetic variation of humans, mapping out the migration patterns of our departure from Africa. (Bolnick, 2008; P.71). These studies found that all humans are genetically a subgroup of Africans and, following the migration out of Africa, all further variation is geographically evenly – or in geneticists' terms 'clinally' – distributed, when not impeded by natural or cultural factors ('bottlenecks' and 'founder effects'), as far as can be statistically inferred.

The reality of human genomic variation is endlessly chaotic and only by statistical clustering of data previously attributed to socially constructed labels can distinct patterns be inferred.

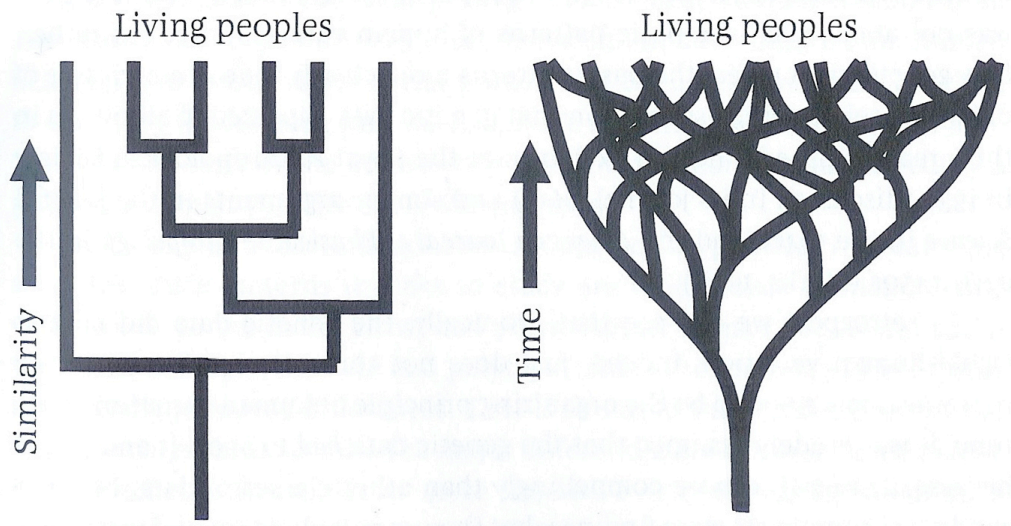

FIGURE 1-1. *Left*, a bifurcating tree, the product of applying clustering methods to genetic data derived from discretely defined populations, to estimate the patterns of similarity among gene pools. *Right*, a reticulating tree, depicting the complex biological histories of human populations.

Image 1: Our current understanding of and the reality of human genetic variation. From: Marks (2008). In Koenig et al. (2008) p. 29

However, a few studies that have received a lot of attention (Rosenberg et al, 2002; Bamshad et al, 2003) have tried to map the distribution of human genetic diversity unto 'major geographic regions' – Africa, Eurasia, East Asia, Oceania and America - that largely correspond to the place of origin of races according to their contemporary (US) folk understanding as in Black, Caucasian, Asian, Aboriginal and Native American. In their genome-wide association studies, bioinformatics software

is used to find clusters of correlations across all the variations in base pairs, called Single Nucleotide Polymorphisms (SNPs), and not just the 'genes' (the functional units they form together) in the genomic data sets available. While it was once stated that the differences between genetic groupings were only a fraction of the differences found between groups (Lewontin 1984, 2000), this technique hones in ever further on the minute differences found between groups of peoples. According to their interpretation of the results (Rosenberg, et al 2002; Bamshad et al, 2003), which are statistical approximations, the genomic data can be mapped unto these geographic regions. But as the methodological review of these studies by Deborah Bolnick (2008) shows, the same data, with different software settings, maps just as well or even better unto other, higher numbers of genetic groupings. Thus there is no evidence that genetic population groups correspond to race, ethnicity or culture – and Bolnick in consequence suggests that

> "these particular results have been emphasized simply because they fit the general notion in our society that continental groupings are biologically significant. This notion is a legacy of traditional racial thought and seems to persist even when not clearly supported by biological data." (Bolnick, 2008, in Koenig et al, 2008, p. 77).

This does not imply that scientists who commit these methodological shortcomings do so on conscious ideologically racist grounds.

Instead it has to be seen in terms of the raciological order that permeates modernity, in both scientific enterprise and political discourse.

While the two studies have been influential, human population genetics make up only a fraction of genetic research that is concerned with racial categories. Most reification of race in genomics comes from pharmacogenomics. The high attention paid to the studies of Rosenberg et al. and Bamshad et al. must be understood amid the rapid, steady growth in medical publications on biotech research aimed at finding genetic causes of common complex diseases, so as to create a marketable cure for them, and which uses racial categories as variables (Outram and Ellison, 2010, p. 92).

My Race in my Genome

Methods

There are numerous genetic ancestry testing (also called personal genomics or direct-to-consumer genetic genealogy) services available to the general public, usually on the internet.[1] In a 2007 review of such websites, Greely counts sixteen

[1] Although National Geographic's 'Genographic Project Public

commercial and one non-profit genetic genealogy companies (Greely, 2008). I selected four companies to which I sent my DNA for testing my 'ancestry'. These are: 23andMe, Family Tree DNA, National Geographic's 'Genographic Project' and Oxford Ancestors. Three of them are among the one's reviewed by Greely, while the first one only came into existence recently, though immediately to prominence, as they also offer genetic testing for health risk prediction.

Together they represent the most famous in terms of controversy (23andMe); one of the oldest and biggest by their own claim (Family Tree DNA); the largest in terms of sample collection scope (National Geographic); and the best-known British company, which offers a slightly different interpretation of the data (Oxford Ancestors).

All of them either provide mtDNA and/or Y-DNA sequencing, which is the current standard in genomic ancestry sold to the public and there are particular issues involved in the limits of these tests.

Only '23andMe' additionally offers genome-wide SNP genotyping that uses data gathered from all chromosomes, not just the Y-chromosome or mt-DNA. Thus it corresponds to the more complex GWAS studies mentioned in the previous chapter and for that reason I will elaborate a bit more on the methods employed and the results provided by '23andMe'.

Finally, using these 4 companies, I am able to analyse all of the currently commercially available, state-of-the-art DNA testing procedures used in genetic

---

Participation Kit' can also be bought at National Geographic's 'flagship store' on Regent Street, London W1

ancestry research for the public and can compare the results they yield.

Patrilineal and Matrilineal Ancestry Testing

In the case of mtDNA or Y-DNA tests, the male customer first has to make a sex-based choice: To either test just the mtDNA, whose results indicate the test subject's matrilineal ancestry, meaning information about one's mother, one's mother's mother, one's mother's mother's mother and so on.
Or, instead or in addition to that, males only can test the Y-DNA that yields information on one's father, father's father and father's father's father etc.
In both cases, it becomes clear, the ancestry that can be researched with such lineage tests is limited to a fraction of a person's actual ancestry, as every person, male or female, has or has had 4 great-grand-mothers and 4 great-grandfathers – of whom only 1 is reflected in mtDNA and 1 in the Y-DNA respectively. A few generations back, this discrepancy becomes ever more wide and 10 generations ago all of us have 1024 ancestral lines – 1022 of them being invisible in the tests on offer. (Greely, 2007; p. 225)

For obvious commercial reasons, none of the companies I used highlight and only few explicitly explain the uncertainty inherent to such genetic testing. A fact that has led to critique and calls for regulations of the commercial DNA testing industry

and resulted in recommendations for genetic testing services, such as the 2008 Ancestry Testing Statement by the American Society of Human Genetics (ASHG) and in the UK the recently published 'framework of principles' by the Human Genetics Commission (HGC) (2010). But as they both consist solely of voluntary codes of practice and there is only commercial interest in describing the positive characteristics, but not to "describe … the limitations of the tests offered", as is stipulated (HGC, 2010, Principle 2.2), unsurprisingly the genetic testing companies have not implemented them (yet).

Additionally, while most testing companies are based in the United States – in my study 3 of the 4 companies used[2] – the statement of the ASHG (2008) refrains from putting the burden of information on the test providers alone. Even though it makes the rather generalising statement that: "greater efforts are needed on the part of both industry and academia to make the limitations of ancestry estimation clearer to consumers, the scientific community, and the public at large." In a typically American individualist liberal perspective, it also relieves the US genetic ancestry testing industry from public pressures by demanding from potential customers to conduct independent research in this highly technical and fast evolving field: "In turn, the public has the responsibility to avail themselves of information regarding ancestry testing and strive to better understand the implications and limitations of these assessments." (ASHG, 2008; Recommendation 1).

In the light of this, it should not come as a surprise that a recent investigation by the U.S. Government Accountability Office (2010) uncovered gross negligence in the

[2] all except Oxford Ancestry

handling of information in customer relations by the commercial genomic testing services. Currently a Senate subcommittee is deciding whether or not to personal genomic's companies will require licensing from the FDA – a move strongly feared by the financial stakeholders in such enterprises (Wall Street Journal, 2010).

Online Disclaimers

What should be considered here is that all of these testing services are marketed and sold over the internet. Navigating online, downloading and installing applications, the average online user is habituated to encounter 'disclaimers', be they age-related warnings or more commonly, copyright notices and is used to scroll down and click on 'accept'. In that moment the burden of responsibility is placed on the user and the assumption is made that she or he has actually read, understood and accepted the 'terms and conditions' displayed on the screen. But the reality of online usage shows that up to 88% of internet users usually do not read such texts but prefer to click 'accept' and move on (GottaBeMobile, 2010).
In the case of the companies I examined, their disclaimers are numerous, lengthy and written in small print and legal language. To the effect that, even if the disclaimers actually do contain information on the tests' technical limitations and we assume that customers have an interest in the issues revolving around their genome that is higher than in other online services, they seem to be playing into the aforementioned pattern of online users' behaviour and it is unlikely that they are

read by many more than other 'terms and conditions'.

National Geographic

One of the biggest test-providers is the National Geographic's 'Genographic Project', which offers the cheapest genotyping service available at $99. Such a discounted price includes only a limited Haplogroup analysis of either one's mtDNA or Y-DNA to a depth of 12 markers. The actual genotyping is conducted by Family Tree DNA, one of the oldest companies specialised in genomic ancestry research, which also offers more detailed testing (at an increased price).

The idea behind the cheap offer by National Geographic is that as many people as possible can participate, so as to have a wider basis of data for the subsequent comparison of the genomic data. The Genographic Project not only sells these 'participation kits' to the paying public, but also gathers genomic data from "indigenous and traditional people" in field studies. As such it represents a – privately funded – successor to the Human Genome Diversity Project. The HGDP set out to collect genomic samples from cultures around the world that represent humanity's genomic diversity, but was abandoned after widespread opposition to the way the 'vampire-project' reduced so-called indigenous cultures to the very bodies of its people by extracting genomic data and using it without regard and benefit to the actual living conditions of the people studied (Reardon, 2005). To counter such criticisms, part of the proceeds from the sale of the 'participation kits'

go towards the Genographic Legacy Fund, which *"supports indigenous conservation and revitalization projects"* (such as for example recording and perpetuating the endangered Minyak language of Tibet[3]).

The project was initiated by author and TV-presenter Dr. Spencer Wells with the support of National Geographic and was sponsored by IBM, who has sold their consumer computer department and is now focusing on large-scale computing such as is used in genomics. The project is relatively exposed to any criticisms regarding the reification of race in their research, especially as it has received a lot of media attention (*NBC Today*, *Oprah*, *The Daily Show*, etc).

To preempt a racist (mis-)interpretation of the results they provide, these are illustrated on the project website with example persons that correspond to the resulting haplogroups (See image 2). Crucially, these example persons are not meant to stand in for their respective 'race', but should represent the 'genetic diversity' of their ancestral origins, as they each come from another part of the

---

[3] "The Minyak people are descended from an ancient tribe that included the Xixia Kingdom in present-day Gansu Province in China. Fifty thousand people spread across four counties once spoke their language, but the inexorable forces of modernization have led Tibetan and Chinese to supplant the local language. Now, the number of Minyak speakers has dwindled to less than twelve thousand - and these are found almost exclusively in remote rural areas in the Sichuan Province of China. In December 2007, the Genographic Legacy Fund awarded a grant to the Kham Aid Foundation (KAF) to launch a project with the goal of preserving and perpetuating the Minyak language in Tibet. In partnership with the Minyak Cultural and Environmental Service Group, KAF aimed to collect Minyak words, phrases, and stories, develop a writing system, and reconcile differences in dialects to standardize the language. And they are seeing success. This community led project has increased ethnic pride among the Minyak people and improved the chances that future generations of Minyak people will value and pass on both their language and other unique cultural traditions."
From: http://blogs.nationalgeographic.com/blogs/genographic/
Accessed: 13/08/2010

planet.

The fact that they *do* each have a different race inscribed in their bodies makes it in practice impossible to separate the unseen genetic difference from the apparent phenotypic difference. After all, this phenotypic difference was precisely selected to reinforce the factor of *difference* that separates the genomic Haplogroups. By demonstrating how all Haplogroups can be found on a single street in multicultural New York City, the genetic diversity is framed as analogous to the racial and ethnic diversity of that city.

While, theoretically, the different Haplogroups might all have been represented by people with ostensibly African ancestry, i.e. seemingly black people, such a portrayal of genetic diversity would have been counter-intuitive to racial thinking and make for a less catchy TV-programming. As the procedure and results of the Genographic Project are aimed at producing a marketable product for National Geographic's various media outlets (TV, magazine, web), such considerations must be assumed to be of relevance to the form of the project's public presentation.

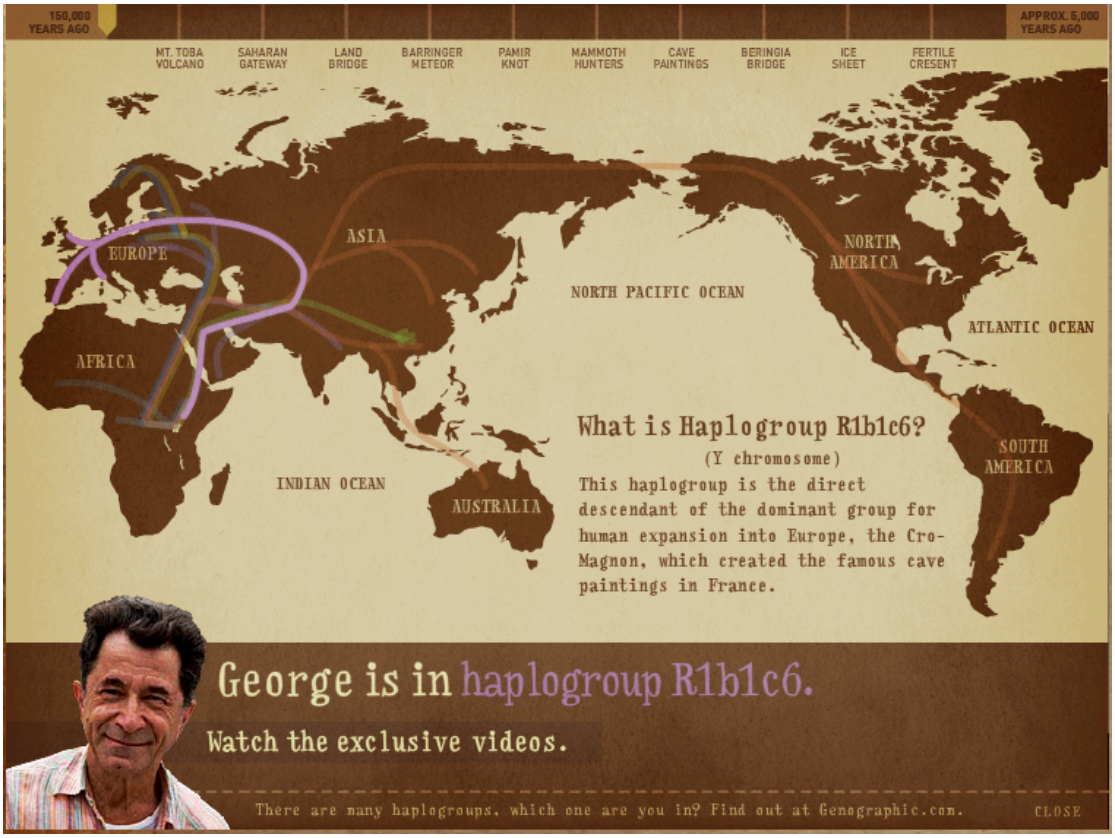

Image 2: National Geographic website - Example representative of Haplogroup R1b1c6.

From: http://channel.nationalgeographic.com/channel/human-family-tree-3706-interactive

Accessed: 01/08/2010

About one week after placing my online order, I was sent the 'Genographic Project Public Participation Kit'. This was a small box with the logo of National Geographic's Genographic Project containing a DVD with a documentary by Dr. Spencer Wells, an information booklet and two buccal swabs.

With these swabs I had to scrape the inside of my cheeks for at least one minute to collect cells. By pushing on an inner stick at the end of the swab, the heads of the swabs with the cheek-cells are then detached and each dropped into a small plastic

container tube with my 'participation number' on it.

These I placed in the return-envelope addressed to the Family Tree DNA lab in Houston, Texas, and sent them via mail.

Family Tree DNA

Family Tree DNA is one of the oldest (since 1999) and best established genome testing companies that cater to the public. According to their website, they are the largest genetic genealogy company in the world.
In fact, Family Tree DNA is also tresponsible for the actual genotyping for the National Geographic's Genographic Project.

Currently their most detailed test on offer is the 'Comprehensive Genome Test' for $ 843. Instead of just either mtDNA or Y-DNA, this test includes both patrilineal and matrilineal ancestry up to a depth of 67-markers each.

In the description of its products on the company website, Family Tree DNA is very straightforward about what it is they do: "Results identify the ethnic and geographic origin of the maternal and paternal lines."
Whereas certain population geneticists go to great lengths to differentiate between the two terms *ethnicity* and *geographic* or *ancestral origin* (Shriver and Kittles, 2008), Family Tree DNA's product description demonstrates how, in the everyday practice

of commercial genetic testing for the public, the two concepts are interchanged - whether this is scientifically correct or not.

As the laboratory of Family Tree DNA is the same that does the testing for National Geographic, the buccal swabs, plastic container tubes and return envelope were the exact same - but did not contain a DVD or leaflet, just the tools to collect and return my DNA sample.

Oxford Ancestors

The British genotyping company Oxford Ancestors offers the same testing procedures as the above companies, but presents the results in a different format. The company is owned by Prof. Edward Sykes, geneticist and author of the 'The Seven Daughters of Eve', a best-selling novel inspired by genomic ancestry studies. In the book he describes the fictional lives of theoretical women who represent the haplogroups of the mtDNA, or "Clans" in Sykes vocabulary, and he gives them names associated with later cultures from the presumed geographic area where they lived.

The test-results from Oxford Ancestors thus connect the genome-sample to the 'Clan' corresponding to one's haplogroup. The customer receives a 'certificate of ancestry' that states the name of one's Clan mother (and/or father in the additional Paternal Clan test).

The collection of DNA took the same procedure as with National Geographic and

Family Tree DNA, and I mailed the samples back to Oxford Ancestor's lab in Oxford, UK.

23andMe

The fourth testing company I used is 23andMe, a genomic testing service that has encountered stark criticism because in addition to ancestry, it offers to test one's genetic health risks.

The company was only launched in 2008, having been promoted by celebrities with a "spit-party" during New York Fashion Week[4]. It was immediately named "invention of the year" by *Time* (2008).

It offers probably the most 'advanced' genomic ancestry testing in that it not only maps markers on the mtDNA and Y-DNA, but selected SNPs[5] across the whole genome, over 550'000 in total. These can then be compared to the publicly available genome samples, and from the overlap the ancestry is inferred.

---

[4] Gisli Palsson (as gislipalsson) in comment to blog entry by Ian Hacking. From: http://onthehuman.org/2009/03/current-controversies-ian-hacking/

[5] Single Nucleotide Polymorphisms, pronounced 'Snips'. Refers to tiny variations in base pairs at specific locations and has replaced the *gene* in importance

This method takes on the issue that in normal, commercial mtDNA and YDNA testing only a fraction of a person's ancestry is actually reflected.

The procedure is the same as in genome wide association studies, which is currently the academic standard in genomic population studies.

In terms of collecting one's DNA, 23andMe is the only company in my knowledge that does not use cheek-cells, but saliva instead. For that purpose the customer is supplied with a *spittoon*, a tool that has such an antiquated name and interesting history, that it has been the object of ridicule (Gisli, 2010).

I had to repeatedly spit into the spittoon until my saliva reached a certain marked level. Once filled, I detached the receptacle, resealed the container tube and placed it in the return-envelope. (See image 2.)

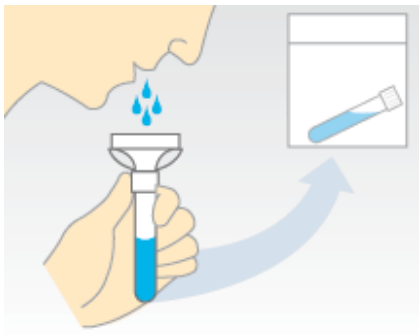

Image 3: Use of the Spittoon. From 'how it works' on website of 23andMe

From: https://www.23andme.com/howitworks/

Accessed: 05/08/2010

Results

Sometime between 2-6 weeks after sending my DNA samples to the respective companies I received the results from their analyses.

Table 1 (see page sums up all the different tests I conducted using the four genetic ancestry testing companies.
The first column lists the type of DNA under analysis, whether mt-DNA for matrilineal ancestry testing, Y-DNA for patrilineal or the genome wide test of 23andMe. The next two columns contain the names of the genetic ancestry testing services and the prices I paid for all the combined tests from each company.
Under 'test' the type of genetic test conducted is listed, usually the haplogroup determination, but sometimes using another name, or in the case of 23andMe, an entirely different test. This is followed by the respective results in column 5.
The last column, entitled 'recent common ancestry', lists national states and geographic regions where persons have lived or currently live who have tested their DNA and whose results 'match' (up to a certain percentage) mine.

National Geographic

The result of the testing by National Geographic was the first one I received back and the most underwhelming, which might be partly due to its relatively low price. As I had to choose to either analyse my patrilineal or matrilineal ancestry only, I

selected the Y-DNA, so as to get information on my father's ancestry.

What I received was a code giving me access to a sub-section of the National Geographic's Genographic Project website. There I could see the haplogroup I was assigned to, albeit only up to a certain degree; the result was: R1a1.

This was accompanied by a map wherein the migration routes of my earliest known ancestors with whom I share markers on my Y-chromosome are traced, and a 7-page text that lists the corresponding timeframes of these past migrations along with descriptions of the assumed lifestyles of these early humans.

The first page of the text also names the current distribution of people who have the same haplogroup as me:

> "Today a large concentration -around 40 percent- of the men living in the Czech Republic across the steppes to Siberia, and south throughout Central Asia are members of haplogroup R1a1. In India, around 35 percent of the men in Hindi-speaking populations belong to this group. The M17 marker is found in only five to ten percent of Middle Eastern men. The marker is also found in relatively high frequency -around 35 percent- among men living on the eastern side of present-day Iran."

Such a result now could be interpreted as absolutely 'correct' at first sight, as my father is indeed from India.

But going into a bit more depth, I would have to ask what is the 'Hindi-speaking population'? My father did learn Hindi at school and can speak it, but it is definitely not his mother-tongue and neither was it of his father. My Indian family's language

is Bengali, also official language of Bangladesh. In fact, my family's homeland was in present-day Bangladesh and it was only after the partition of British India into India and East Pakistan that my father's father had to leave our farmlands in what is now Bangladesh and settled in Calcutta alone.

This part of my family's history, which is marked by war and exile, forms a strong political undercurrent to my ancestral narrative and is probably of the greatest importance for my ethnic and kinship affiliations and identity shaping. But this is in no way represented in my genetic ancestry test results.

Family Tree DNA

An email I received from Family Tree DNA announcing that my results of the 'Comprehensive Genome Test' are available, contained the login-details for the members-only section of their website.

This is divided into subsections for the mt-DNA test results, Y-DNA test results and an additional section with tools for contacting other people of the same haplogroup and a link to a PDF eBook on how to interpret the genetic ancestry test results.

The Y-DNA test result is, unsurprisingly, again R1a1. This is displayed in a colourful, interactive representation of the haplogroup tree (Image 4).

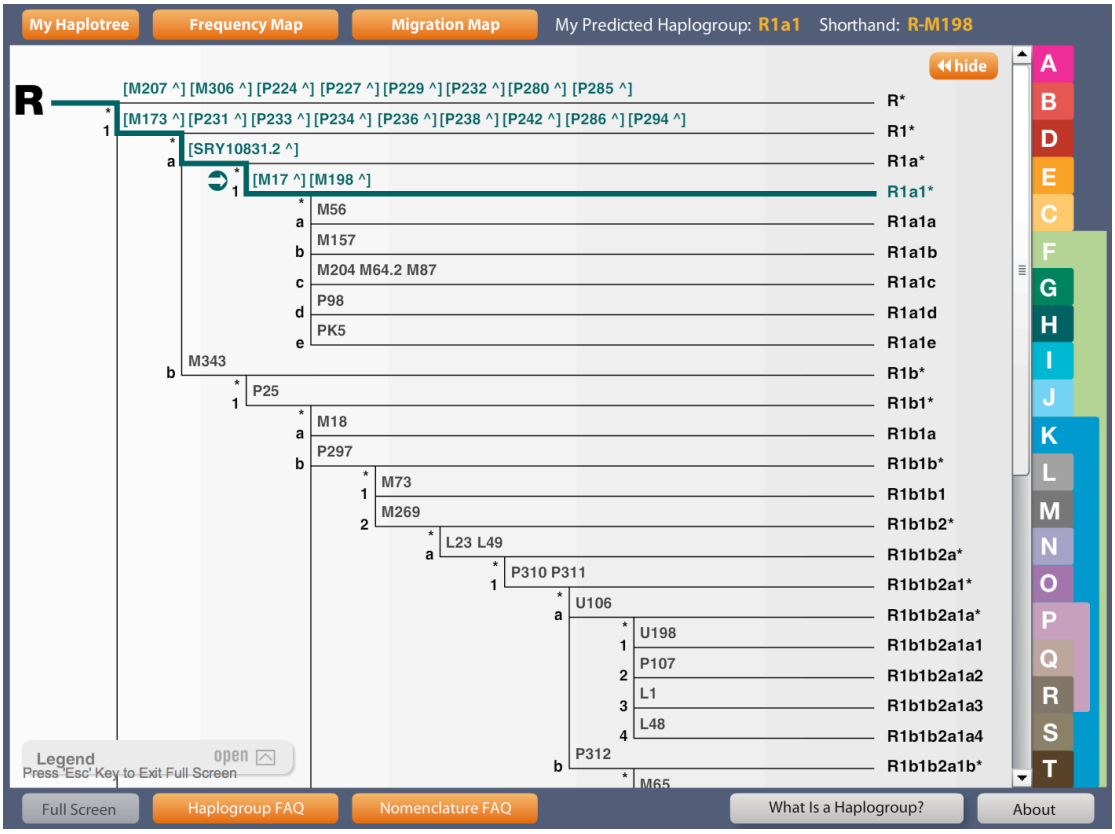

Image 4: Family Tree DNA - Visualisation of the haplogroup tree branch R, with the sub-branch of my result highlighted

From: http://www.familytreedna.com/my-ftdna/y-dna-haplotree.aspx

(restricted access)

Accessed: 22/08/2010

As opposed to the National Geographic test where I was provided with just the short text outlining the general knowledge available on the distribution of the haplogroup R1a1, the full Family Tree DNA service accesses public databases such as the HapMap, along with its proprietary database of customer's genetic information to render a list of countries where persons with matching Y-DNA markers live (Table

2[6]).

While most of my Y-DNA matches are found in Greece, these matches are still only 0.5% of Greek people who tested their genetic ancestry – and as such amount to only 3 persons. This data is therefore statistically insignificant, as is indicated in the table's legend.

My matrilineal ancestry was determined by the mtDNA test as placing me in the haplogroup T2b – what exactly this means, I am left to guess.
The list of countries with some matches contains is headed by England, with 3 persons matching some of my genetic markers… But also two in Greece (again), Italy, Poland and Scotland and each one in Estonia, Germany and in a not further specified nation of the 'United Kingdom'.

That my mother stems from one of the oldest families of burghers of Berne, with citizenship accorded in 1312 is obviously not reflected due the limitations of linear ancestry. But also her mother being from a Swiss noble family, and her mother, my great-grand-mother, who was one of the first women in Switzerland to drive a car and told me story of how electricity came to Zürich, forms my 'matrilineal' genealogy – which remains without resonance in the data that Family Tree DNA supplies me from my DNA.

---

[6] As only the first 11 countries listed have people with exactly matching Y-DNA markers, they are the only one's I copied to table 1.

Oxford Ancestors

The results from Oxford Ancestors came to me by standard mail, in the form of two paper folders, each containing a colourful 'certificate', along with a letter and explanatory leaflets. The certificate shows a tree-like structure, a coloured circle at the end of each branch, and one with golden star on it. My 'Paternal Clan' is said to be 'Sigurd', about which the accompanying text says:

> "The clan of Sigurd, at 25'000 years old, is younger than the related clan of Oisin. Sigurd's descendants are concentrated in Northern Europe, particularly Scandinavia where the clan reaches its highest frequency (30%) in Norway and also Iceland (20%). The presence of the clan is a particularly valuable signal of Norse Viking colonisation, not only in Iceland, but also in Britain where the clan reaches its highest concentration outside Scandinavia in the former Norse strongholds of the Orkney and Shetland off the North Coast of Scotland. The albeit rare occurrences of Sigurd's descendants in Normandy and Southern Britain are likely to be an echo of Viking annexation of Normandy from the King of France in the tenth century AD and the Norman Conquest of Britain that followed a hundred years later."

It continues by explaining the name of this hypothetical person:

> "The clan Sigurd is named after one of the most prominent heroic figures of Norse mythology. Known as the Dragon-Slayer, Sigurd won his

title by killing the dragon Fafnirand taking the treasure that it guarded. His sword was forged by his friend and ally, the master swordsmith Regin."...

This narrative equates my genetic data with Viking ancestry, something of a surprise considering that my father is supposed to come from an uninterrupted line of Bengali Brahmins.

The second folder contained 'The Seven Daughters of Eve' certificate, pointing to 'Tara' as my mt-DNA clan mother. It is accompanied by a four-page extract from Bryan Sykes book (Sykes, 2003) that bears the title 'Tara's Story' and tells the life of a girl in the Neolithic: A hunter-gatherer, she meets another 'band' of humans after finding a dolphin-carcass on the shore and joins that group, changing

"from one band whose main territory was the wooded hills of Tuscany to another, which hunted further up the coast. Four years later, she was pregnant and the first of her two daughters was born. As soon as the baby appeared, it was obvious that she had inherited her father's flame-red hair. By the time she was a year old, it was also clear that she had inherited Tara's independent streak. ... Tara was a good mother and a welcome new member of the band."

Also in the package is a leaflet 'Interpreting your Matriline Certificates', wherein a short paragraph is dedicated to each 'clan mother':

"Tara (Gaelic for rocky hill)

Tara lived in Tuscany about 17'000 years ago. At the time, Europe was in the grip of the last Ice Age and the only parts of the continent where life was

possible were the lands bordering the Mediterranean. Then, the Tuscan hills were a very different place. No vines grew; no Bougainvillaea decorated the hillside farms. Instead, they were thickly forested with pine and birch. The streams held small trout and crayfish, which helped Tara to raise her family and held the pangs of hunger at bay when the men folk failed to kill a deer or wild boar.

> As the Ice Age loosened its grip, Tara's descendants moved round the coast into France and joined the great band of hunters following the big game across the tundra that then covered Northern Europe. Eventually, Tara's descendants walked across the dry land that was to become the English Channel and moved right across to Ireland, from which Celtic kingdom the clan takes its name."

What starts out in Italy turns into a Trans-European migration narrative that ends up 'celtic'. Such freedom from 'scientific' claim is charming and might be a safety buffer from over-interpretation. As Sykes deliberately mythologises a hypothetical Person instead of trying to historicise her, this could be the basis for a non-national, non-racial, human counter-narrative.

Nevertheless the unspoken whiteness that permeates the protagonist can be just as misleading and the geographic territories mentioned play a specific enough role to ground Tara as 'European'.

23andMe

3 weeks after sending my 'Spit-Kit™' to Mountain View, California, I receive an email telling me *"Data is available for: Robin Bhattacharya – Start exploring your health and ancestry today."* Upon logging into the company's website using the 'Claim Code' I was supplied, I am greeted by the "welcome to you", about which Ian Hacking says: "It is "You", that's Me, all the way. The implication is clear, you are about to learn about the real you" (Hacking, 2009).

The website is structured into 4 main sections: "My Health", "My Ancestry", "Sharing and Community" and "23andWe".

While "My Health" is made up of subsections for "Disease Risk", "Carrier Status", "Drug Response", "Traits" and "Health Labs", "My ancestry" is divided into "Maternal Line", "Paternal Line", "Relative Finder", "Ancestry Painting", "Global Similarity" and "Ancestry Labs" - all of which are different ways to display and interpret the genetic data gathered from my DNA sample.

"Sharing and Community" offers tools to compare my test-results with other people and make contact with them – functions that have been likened to a Facebook of biosociality (Hacking, 2009).

"23andWe" asks the consumer to participate in surveys on health-related subjects – which are then aggregated with the genetic information on file and result in 'studies' that reflect the risk of the health issue in question for persons with the same mutation on a given gene.

Maternal Line

The mtDNA's 'Maternal Line' result is given in more detail than the previous tests as: T2b4. But that does not result in more information about my ancestry. Again I can read a short text about the prehistoric migrations of said haplogroup, telling me how these early pastoralists came from the Near East into Europe, India and eastern Africa. The current distribution of T2 is indicated in a side-box as among 'Northern Europeans' and 'Spanish', and a sub-section below tells me that 'Jesse James' shared this haplogroup with my mother and me.

In addition a map is displayed that shows me the pre-colonial geographic distribution of my (rough) haplogroup (image 5).

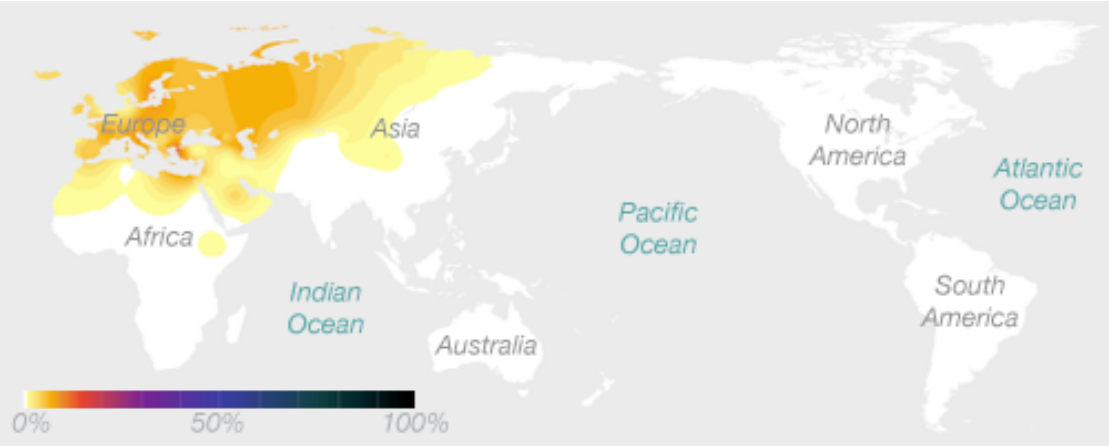

Image 5: 23andMe - "Locations of haplogroup T2 circa 500 years ago, before the era of intercontinental travel."

From: https://www.23andme.com/you/haplogroup/maternal/ (restricted access)

Accessed 15/07/2010

Paternal Line

The results for my 'Paternal Line' are displayed in the exact same way. The haplogroup is given as: R1a1a*. For the first time determined up to the 'deep clade', meaning also the last digit, in my case an asterisk*, that places me on a specific sub-branch of the haplogroup tree. What difference this makes for my 'identity', I cannot determine. The description tells me that

> "[t]he haplogroup is most common in a swath from Ukraine and the Balkans north and west into Scandinavia, along the path of the men who followed the receding glaciers into Europe. It is also common near its presumed point of origin in south-central Asia."

The box on the side summarises this as "Populations: Ukranians, Indians, Poles" and indeed the corresponding regions are coloured dark on the illustrative map (image 6).

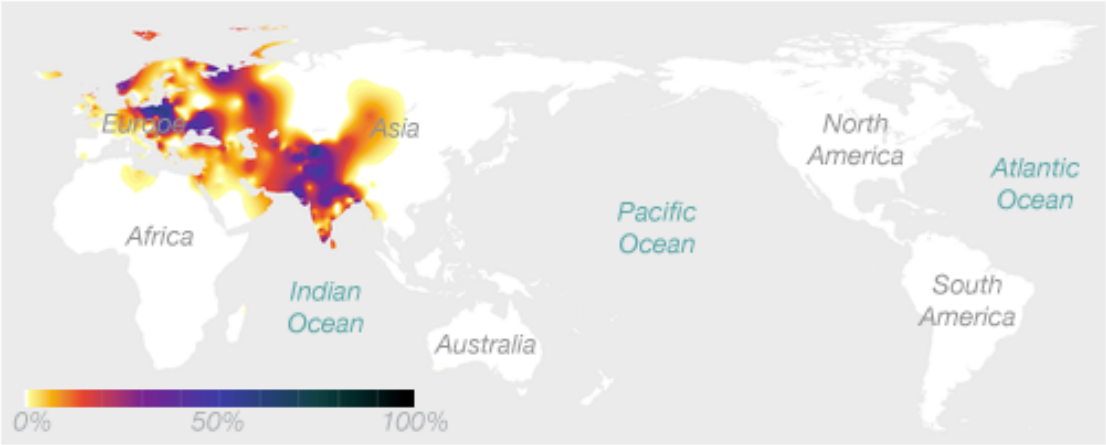

Image 6: 23andMe - "Locations of haplogroup R1a1a circa 500 years ago, before the era of intercontinental travel."

From: https://www.23andme.com/you/haplogroup/paternal/ (restricted access).

Accessed 15/07/2010

'Ancestry Painting'

The feature that differentiates 23andMe from others genetic testing services is their genome-wide analysis and comparison of my DNA.

The "ancestry painting" gives a view of all my chromosomes – not just mt-DNA or Y-DNA – and compares it to 3 reference populations, these being 'European', 'Asian' and 'African'.

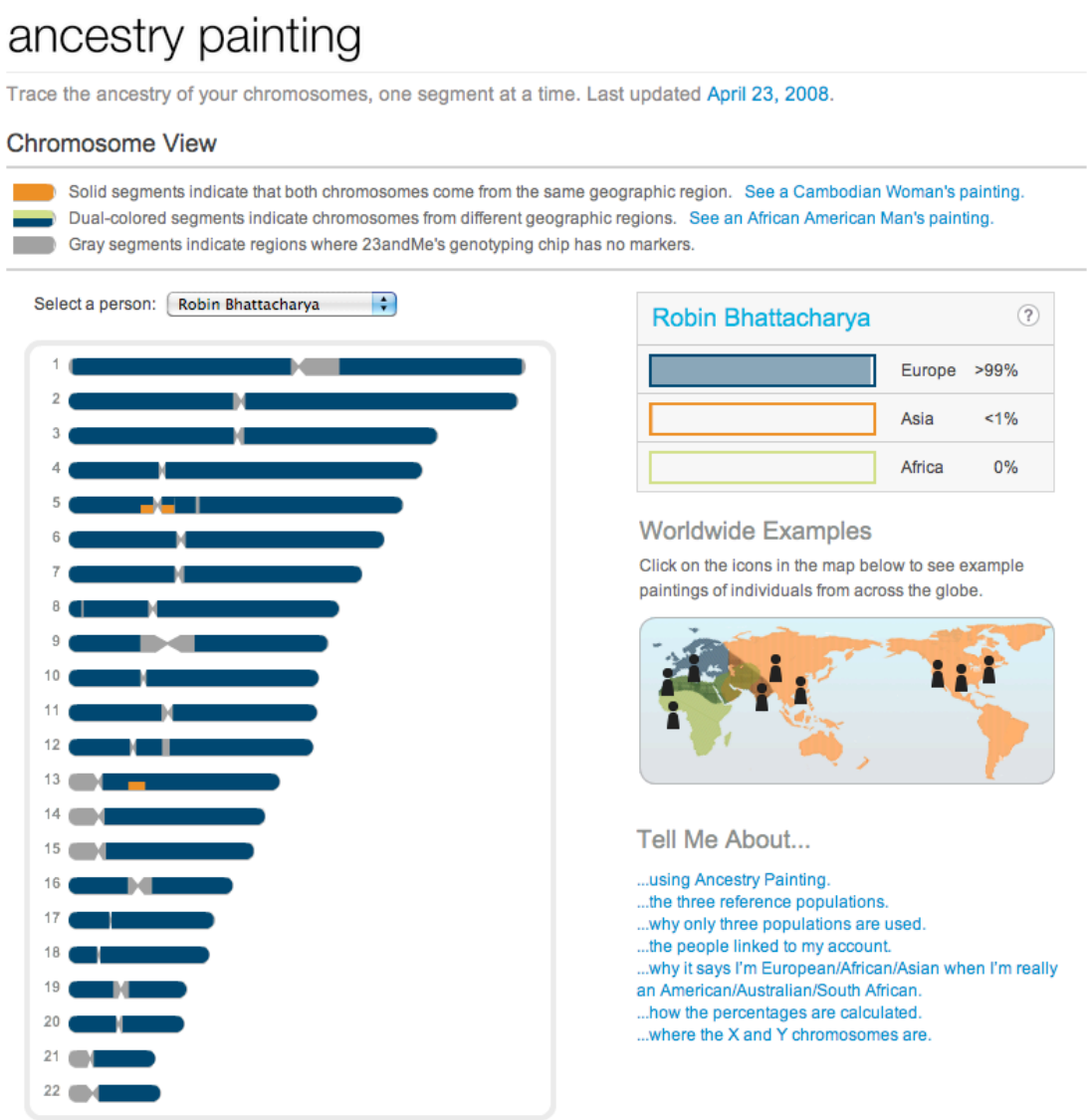

Image 7: 23andMe - 'ancestry painting'

From: https://www.23andme.com/you/ancestry/paint/ (restricted access)

Accessed 12/08/2010

My result, as can be seen from the amount of dark blue in the 'Chromosome View', presents itself as "> 99% European" and "<1% Asian", placing me firmly in the European continent. By clicking on the links at the bottom right of the page, more information can be found about these reference populations and why the were selected: Apparently 23andMe, like Family Tree DNA, also uses the International

HapMap project's database[7] for comparing genetic data. And as the HapMap is the successor to previous public DNA collections, which it incorporated, such as the data from CEPH (Centre pour l'Etude des Polymorphismes Humains) and HGDP (Human Genome Diversity Project), and this data is sorted along said reference populations – the final picture given by 23andMe reflects these, highly controversial, because arbitrary, groupings as given "corners" of human genetic diversity[8].

---

7 23andMe - What reference populations are used and why?
23andMe takes advantage of publicly available data for four populations studied extensively via the International HapMap project (hapmap.org). That project obtained the genotypes for 60 individuals of western European descent from Utah, 60 western African individuals from Nigeria, and 90 eastern Asian individuals, 45 from each of Japan and China. Because the two eastern Asian populations are geographically near one another and relatively similar at the genetic level, 23andMe combines these to form a single eastern Asian reference population. For more information on why these regions were used, please see (Why are these three populations used?)
From: https://www.23andme.com/you/faqwin/whyrefpops/ (restricted access)
Accessed 08/08/2010

8 23andMe - Why are only three populations used?
There are a few answers to this question, one practical, and two more theoretical. The practical reason is that the ideal dataset for ancestry painting doesn't exist yet. (We're working on it!) The International HapMap Project (hapmap.org) is an excellent resource, but as more data meeting our quality standards becomes available we will incorporate it.
One theoretical reason is that the three geographic regions in HapMap were selected to be good representatives of all human diversity. Genetic population studies suggest that clusters of human diversity often correspond roughly to geographic obstacles to human migration, such as the Sahara desert, the Himalayas, and the Bering Strait. So by using samples from the hearts of these clusters, as HapMap does, we are using the "corners" of human diversity so that any human genotype will lie somewhere between them.
A second theoretical reason is that the problem of inferring ancestry becomes more difficult when less genetic information is used. In order to infer the ancestry of some specific stretch along one of your chromosomes, we naturally limit ourselves to the SNPs lying within that stretch. It is often difficult to distinguish people from genetically similar populations, such as French and German, even using the full genome; it stands to reason that making such distinctions within a short stretch of one chromosome is even harder. Therefore, it is often the case that a given stretch of chromosome is no more "French" than it is "German", so it is fairest to simply call such stretches "European".
From: https://www.23andme.com/you/faqwin/whythreepops/ (restricted access)
Accessed 08/08/2010

Global Similarity

Another feature unique to 23andMe is entitled 'Global Similarity'. It uses the same data as the 'ancestry painting' to calculate a geographic location, based on the similarity of my genome to that of other populations. The overview map (image 8) places me on the edge of the 'Central/South Asian' population cluster, almost bordering the 'European'.

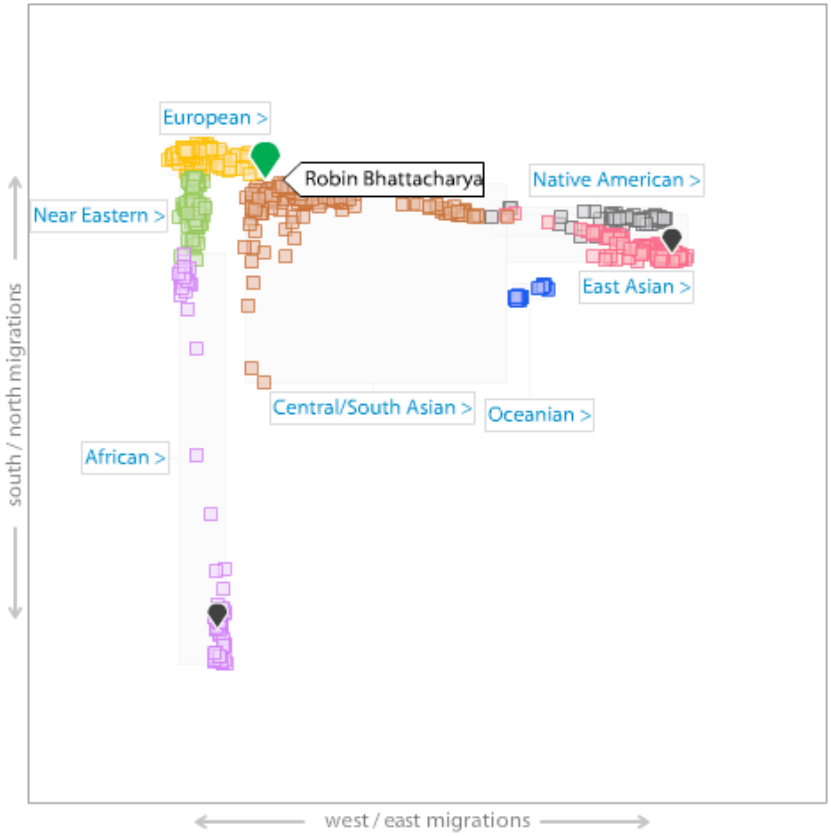

Image 8: 23andMe - 'Global Similarity advanced view' overview

From: https://www.23andme.com/you/globalsim/advanced (restricted access)

Accessed: 12/08/2010

Upon clicking on the population name I can zoom in, but within 'European' there is

no match. Looking at the 'Central/South Asian' cluster though, one finds correlations of my DNA to that of the "Pathan", population (image 9). The information box to the side informs me that the "Pathan" or "Pashtun" are "Sunni Muslims who live in northern Pakistan, Afghanistan, and parts of India".[9](This would explain why, as a dark-skinned, bearded man living in Europe post-9/11, I have often been likened to a Taliban. But this might have just as much to do with the social construction of our selves in opposition to an other, which after 9/11, for the 'West' has become an indistinct Muslim 'other'.)

---

[9] 23andMe – Pathan: The 40 million Pathan (or Pashtun, as they are known in Afghanistan) are Sunni Muslims who live in northern Pakistan, Afghanistan, and parts of India. They speak Pushto, one of many Iranian languages that are spoken between Turkey to the west and Pakistan to the east.
References to the Pathan in classical sources date back to the 10th century AD. Recent genetics research suggests Pathans are Indo-Europeans and related to other Iranian groups. Their ancestors probably began expending from somewhere near the city of Kandahar over a thousand years ago. Though they constitute a clearly defined ethnic group with a distinct language and culture, many Pathans have family trees that include non-Pathans absorbed generations ago by dominant Pathan groups.
Pathans/Pashtuns are directly linked with the recent history of modern Pakistan and Afghanistan. Pathans from both countries mounted active opposition against Soviet invasion in Afghanistan in 1979 and, more recently, comprised the primary ethnic group of the Taliban. The current president of Afghanistan, Hamid Karzai, is Pashtun.
23andMe's reference database includes 24 representatives of the Pathan population from Peshwar in northwestern Pakistan.
From: https://www.23andme.com/you/globalsim/advanced/ (restricted access)
Accessed: 12/08/2010

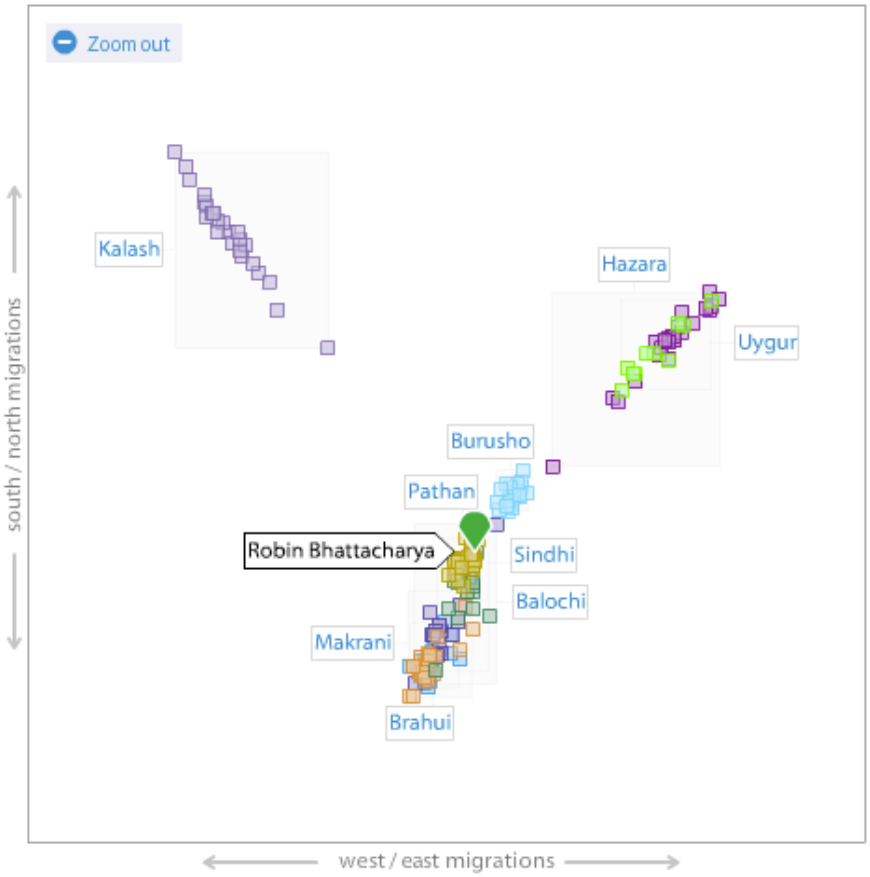

Image 9: 23andMe - 'Global Similarity advanced view' detailed view

From: https://www.23andme.com/you/globalsim/advanced/ (restricted access)

Accessed: 12/08/2010

But to be fair, this specific result provided by a genetic ancestry testing company, is the first one entirely mitigated by my self-ascribed, racial phenotype: At the bottom of the 'global similarity' page, in the list under "Tell Me About..." is a link to a info-text called "...interpreting my plot if I'm of mixed ancestry".[10] Therein I am told that

---

10 23andMe - How do I interpret my plot if I'm of mixed ancestry?

The Global Similarity Advanced feature positions you on a map of the global gene pool according to a 'weighted average' of your ancestry. So if your parents come from widely separated populations you will end up between them, and most likely among people who do not share your ancestry!

For example, imagine someone with one English parent, and one Chinese parent. In the World view of Global Similarity Advanced, the English parent will appear amid

if my parents or even grandparents "come from widely separated populations" I will not be accurately mapped unto the 'global similarity' graph.

As I can assume this to be the case in light of my 'mixed-race', post-colonial family history, once again I must take the results I received as being statistically insignificant.

Indeed the website even contains a link to a PDF White-Paper containing all the technical details about the 'Global Similarity's Genetic Similarity Map', or GSGSM. This includes details about the reference populations from CEPH-HGDP - and lists Rosenberg's studies as its main methodological reference (23andMe White Paper 23-04, 2008).

European reference individuals (yellow squares) and the Chinese parent amid East Asian reference individuals (pink squares). Their offspring will appear about halfway between them. So even though their offspring's ancestry traces to England and China, he or she will appear amid Central and South Asian reference individuals (brown squares).

It's not just individuals with one generation of mixing who end up being placed on the map according to the average of their ancestry. For instance, someone with three Irish grandparents and one German grandparent will appear in the Northern European view between the Irish and German clusters, although nearer the Irish cluster than the German one.
From: https://www.23andme.com/you/faqwin/recentadmixture/ (restricted access) Accessed 08/08/2010

Discussion

Limitations

Not (yet) Sequenced my Entire Genome

This study is self-funded, with the cost of DNA testing amounting to £ 1'267.58, and due to these financial restrictions, I was unable to have my entire genome sequenced.

With approximately 10 Mio. SNPs genotyped, this is offered for example by Knome inc. for $99'000.00 [approx. £ 63'727.00] (Hacking, 2009).

But this does not seem to be of too big a consequence for my ancestry testing. As 'relevant' sections of the genome have been identified and thus the comparison is made based on these 'tag SNPs', as they are known[11], the result should be more or

---

[11] 23andMe - How does 23andMe genotype my DNA?
The process by which 23andMe's contracted laboratory genotypes your DNA uses the latest in DNA technology.
Once the lab receives your sample, DNA is extracted from cheek cells in your saliva. Your DNA is then copied many times so that there is enough DNA to use for the genotyping step. Next, the DNA is cut into smaller, more manageable pieces. These DNA pieces are then applied to a DNA "chip." The DNA chip is a small glass slide with millions of microscopic beads on its surface. Attached to each bead are "probes"—bits of DNA complementary to sites in your genome where SNPs are located. There is a pair of probes for each SNP, corresponding to the two versions of each SNP. Because two complementary pieces of DNA stick together, your DNA sticks to whichever probes match your versions of a SNP.
To tell which versions you have, your DNA is extended in a way similar to the process of DNA replication inside your cells. But on the chip the process adds not just DNA but a fluorescent marker as well. By determining which beads are glowing we can tell which versions of a SNP you have.
The DNA chip that we use genotypes hundreds of thousands of SNPs at one time. It actually reads 550,000 SNPs that are spread across your entire genome. Although this is still only a fraction of the 10 million SNPs that are estimated to be in the human genome, these 550,000

less equivalent to a genome-wide association study with a higher number of SNPs. (23andMe has tested my DNA for 550'000 SNPs, while the HapMap samples have approx. 1.4 Mio genotyped SNPs.)

No more mtDNA / Y-DNA tests

As the mtDNA/Y-DNA tests have consistently shown the same results in all 4 tests by the different companies (R1a1, T2b), I saw it as unnecessary to conduct further similar tests.

No Afro- or Jewish- specific tests conducted

Certain smaller genotyping companies offer tests aimed at specific groups, such as *roots4real* or *African Ancestry* for finding African or Jewish ancestry. As in my personal biography there is no indication for me being African other than being human and thus having migrated out of Africa like all homo sapiens at some point in the past, I refrained from testing my genome for such to that end. Similarly, as I have no reason to think of myself as Jewish, I did not send my DNA to any company that offers tests to that end.

SNPs are specially selected "tag SNPs." Because many SNPs are linked to one another, we can often learn about the genotype at many SNPs at a time just by looking at one SNP that "tags" its group. This maximizes the information we can get from every SNP we analyze, while keeping the cost low.
In addition, we have hand-picked tens of thousands of additional SNPs of particular interest from the scientific literature and added their corresponding probes to the DNA chip. As a result, we can provide you personal genetic information available only through 23andMe.

From: https://www.23andme.com/you/faqwin/chip/ (restricted access)
Accessed 05/08/2010

But not Askenazi- and not Native American

That said, amongst the tests included in the package by 23andMe are an examination for Askenazi-Jewish specific markers, i.e. those genome-markers relevant for the genetic basis of Tay-Sachs disease. According to that test, I am not Askenazi-Jewish and do not have an increased risk for Tay-Sachs.

Also part of the 23andMe offer is a test for 'Native American DNA'. According to that, my genome does not seem to have patterns related to those found among Native American indians.

This test result in itself is problematic, as TallBear (2008) writes, because it assumes a biological basis for being admitted to an Indian tribe, supplanting the very different structure of kinship of Indians with a Eurocentric logic of blood-ties, reframed in the language of genomics.

Key Findings

- mtDNA, Y-DNA results with little meaning

  haplotype R1? T2?

The testing of my mt-DNA and my Y-DNA, which gives me information about certain lines of ancestry that relate back to prehistoric times, may be interesting up to a certain degree, but ultimately of little meaning. The distribution of R1a1a and T2b4 is far too wide to allow for identification without mythologising.

- other tests either: oversimplify (European) or

  are inconclusive, because of my own 'admixture'

Technically more advanced tests that use genome wide association to infer a genetic 'ancestry painting' or my 'global similarity' to other people in the world, either oversimplify the genetic data (and as a result my identity), reducing it to something as scientifically vague, but historically marked, as 'European'. Or it is inconclusive because of my own 'admixture' – which says more about the statistical process than about my genome.

Of course many of the companies offering genetic ancestry testing note that these services should be understood as *complimentary* to other genealogical methods like

archival research, a fact pointed out by its advocates (Shriver and Kittles, 2008; Nelson, 2008a). So what the results of my genomic ancestry testing show most clearly is the *interpretative flexibility* (Wajcman, 2004) that I am given in their uncertainty and unspecificity. Using these I could read a 'factness' in my genome that testifies to my identity (Nelson, 2008b), as 'European', 'Indian', or even 'mixed-race', depending on my choice.

But were my family history and personal identity narrative different, I could just as easily read into the same data that my mother is from the Middle East and that my father is Polish. Nothing in the genome points to Berne, Calcutta or somewhere in rural Bangladesh, the places were my actual family history took place and my 'culture' is located.

Now all this might be entertaining and interesting to some, but the crucial methodological question that arises in the process is: Why is there the tendency to categorise DNA samples along racial, ethnic, national, cultural lines? Why such a pressure to align this messy data that indicates something about a pre-historic past, along well-worn labels?

In a sense it is a circular argument: To infer early human migration from genetic data, it is necessary to collect data that is assigned to a place, a culture, any stand-in for the assumed genetic population, therein lies the *use value* of what Cavalli-Sforza (1994) named 'genetic isolates'. Until other technologies are developed to geo-tag genetic samples – say, by a future convergence of the currently separate Facebook(-like) functions of geo-tagging and genotyping – it is the only way to *make sense* of the genetic data, can shed light on paleo-anthropological questions and help us in

understanding how the great diversity of human cultures evolved during thousands of years of migrations, conviviality and sex.
But in that moment it is crucial to keep in mind how genetic ancestry inference is a bio-technologically assisted process that is based on socially constructed variables.

The reverse inference does not work: Of course I may share genetic similarity with this or that reference population, but that does not mean I share anything other than certain SNPs with that reference population, because the correlation comes not from a shared present - say me actually being Pashtun, Italian, European, Polish or Indian - but our theoretical shared ancestor who was a hunter-gatherer millennia ago.

It is symptomatic that a mathematically produced visualisation like the 'GSGSM' of 23andMe is not possible with my genome due to it being of 'mixed ancestry'/race: The data used presupposes the same genetic purity it relies on for its statistical inference.
A reading of such results that relates them back to actual existing populations or regions necessarily falls into reifying socially constructed categories like race (European), nationality (Indian) or ethnicity (Pashtun).

Thus it amounts to the same "Biologistical Construction of Race" that Fullwiley (2008) has observed in laboratories conducting genetic admixture studies with the aim of finding genetic causes to racial health inequalities.

The necessary distinction made between genetic populations and reference

populations, does not differ from how the more educated Nazis distinguished between the ideal type of 'race' and real existing 'nations' and ethnicities:

> "Whoever tries [in this way] to visualize the nature of a human race must immediately admit that it is hardly possible to find a race anywhere in the world as a self-enclosed human group. ... All Western nations are mixtures of races which include, in certain percentages, pure and mixed, all the races of Europe, or in which, at least, several European races are represented. ... popular usage will continue for a long time to speak of a "Jewish race," even though educated persons will long have recognised that the Jews, like other peoples, represent a mixture of races."
>
> (Hans F. K. Günther, *Kleine Rassenkunde des deutschen Volkes,1929.* In Mosse, 1966; p. 63)

The reason for this reproduction of race and ethnicity is not to be found in a racist conspiracy of medical practitioners (Fullwiley, 2008), just as the methodological critiques of genotyping (Duster, 2003; Fullwiley, 2008; Nelson 2008a, 2008b; and TallBear, 2008) cannot simply be dismissed as left-wing bias or 'political correctness' (Burchard in Fulwiley, 2008; Marks, 2008: P.28).

Agamben is acutely aware of the complex relationship between biopower and the knowledge produced in in the life sciences, when he points out that:

> "Nazism, contrary to a common prejudice, did not limit itself to using and twisting scientific concepts for its own ends. The relationship between National Socialist ideology and the social and biological sciences of the time – in

particular, genetics – is more intimate and complex and, at the same time, more disturbing"

(Agamben,1998; P.85)

The goal of Dr. Spencer Wells to discover how all of the big, human family is related together, is certainly a laudable one and of great historical interest. But while tales like those of Sigurd, Tara and other haplogroup ancestors might suffice to sell books and DVDs, they do not retain enough specificity to create the same kind of identification as hundreds of years of tradition.

The same is true about the bio-social groups (Rabinow, 1992; Hacking, 2006) that emerge in the user-pages of 23andMe, Family Tree DNA and personal blogs on genetic genealogy[12]: While they certainly have the same level of identification as any other FaceBook group, they do not seem to go anywhere beyond that, which led Hacking to call them "*almost a parody of the idea of biosocial identity"* (Hacking, 2009).

A shared history of Othering however, of living as a dark-skinned person in a white society, has had a far more profound effect on my self-identification than my membership of haplogroup R1a1a* or T2b4 could ever have.

---

[12] Haplogroup R1a1 FaceBook group:
http://www.facebook.com/group.php?gid=14443500485

Haplogroup T (subclade T2 mtDNA) - the Genographic Project FaceBook Group:
http://www.facebook.com/group.php?v=wall&gid=16968521635

The Genetic Genealogist:
http://www.thegeneticgenealogist.com/

Genomes Unzipped – Public Personal Genomics:
http://www.genomesunzipped.org/

Just as my position of privilege in that set of relations that defines whiteness, contributes to my personal identity, i.e. I am also white/European when in India, or 'mixed-race' in a multi-cultural context that asks for such a kind of identification (like UK census data).

So to make sense of this uncertain, unspecific, impure mass of data, for this paleo-anthropological information to be *useful*, we have to relate it (back) to actual existing social relations and that is precisely where racialisation and the reification of ethnic and national categories in genomics takes place. Hence all of the genomic ancestry testing companies in this study rely in some form on racial, ethnic and national categories and also reify them in their genomic test results and interpretations.

Conclusion

This case study shows how genomic data and its interpretation as provided by commercial direct-to-consumer genetic testing companies, be that mt-DNA, Y-DNA or genome-wide association results, suffer from the same shortcomings as other population genetics: In order to study the differences, the *exceptions* that define a genetic population, it is necessary to *intervene* on the data by relying on socially constructed reference populations, which are subsequently reified, biologised, geneticised in the process.

While this may be done with good intentions, out of paleo-anthropological or, as is the case more often, pharmacological reasons, the research is never conducted without self-interest: The knowledge is instrumentalised as it is being produced, following the dictate of *utilitarianism*. This can be to present certain indigenous peoples as descendants from an uninterrupted line of ancestry to create interesting material for a TV-documentary[13]. Or to sell a cure for a condition that affects a specific biologised population, like BiDill (Duster, 2003; Outram and Ellison, 2010). Or to sell a service that promises to reveal something about their identity at a time when – despite the discourse around new biosocial groups - race, ethnicity and nationality continue to be the salient modes of identification.

Thus, we can see a pattern in the studies I described and the tests I conducted:

---

[13] Journey of Man (2005) [DVD] Directed by Dr. Spencer Wells. USA: National Geographic

Exception - intervention – utilitarianism.

First, there is a focus on the *exception*: Even though humans are for the vastly biggest part genetically identical, it is the *differences* that are of most interest. This is the case for obvious reasons in medical studies, although even these show that only a fraction of diseases and other health issues are genetically determined and rather unfold in complex exchanges with epigenetic factors.

But also human diversity projects, like the Genographic Project and the ever more elaborate personal genomic ancestry testing services, hone in on increasingly minute details of differentiation.

Secondly, in the very moment of doing so, of collecting the sample and labelling it, they *intervene* and act upon the object of research, classifying it and thereby reifying its socially constructed 'biologistic' objecthood.

This need not take such an extreme form, as when a sample donor from Kazakhstan replies to Dr. Spencer Well's announcement that his Y-DNA shows his family has been here for 2000 generations with the words: "Yes, my family is very pure, thank you."

It can also be as superficial as a Facebook group of people looking for 'Viking' or 'European' traces in their DNA. While for some this may bolster their sense of identity in a 'positive' way, such as when people from the black Trans-atlantic diaspora find confirmation of their 'race' (Shriver and Kittles, 2008), any detailed information about one's 'culture' is purely speculative 'factness' (Nelson, 2008b).

But in any case, the research investment in the field is large and growing and the reason for this, and this is the third point, is the perceived the *utility* of genotyping,

as with all classificatory information.

The founder of 23andMe is the wife of the CEO of Google who has as its mission "to organize the worlds information"[14] so as to make a profit from it. And so it would make sense to include all personal genomic information of all people, organised in a FaceBook-like fashion and with all the data aggregated so as to be able to make ever new studies, based around genomic markers as variables.
It just happens that in order to motivate people to contribute to such data collections, and even pay for it themselves, there must be a reward in the form of relatable, useful information, and so this is what the genomic testing services provide in the form of racialised and ethnicised categories, even if it is speculative up to the point of endangering lives. (Marks, 2008; p. 35)]

This structuring dynamic of biologisation and genetisation of human traits and the subsequent intervention and instrumentalisation are the major shifts inaugurating the genetic era (Franklin and Lury, 2000). It constitutes part of the secularised ideology with Christian and Aristotelian underpinnings that Haraway recognises as at the heart of what she calls the 'material-semiotic apparatus' of modern science (Haraway, 1997).

Even though biotech innovations like the ones described in this study certainly are technologies of the self, the biopower mobilised therein goes beyond the institutions analysed by Foucault (2008). Racialised genomic discourse centres directly on the question of bare life and how political territory is derived from the

[14] From: http://www.google.com/intl/en/corporate/facts.html Accessed: 24/08/2010

logic of blood-ties, framed in the genetic imaginary. The way in which the exception informs the rule by which value is attributed to lives, is the "The fundamental biopolitical structure of modernity – the decision on the value (or nonvalue) of life as such" (Agamben, 1998; p. 80). The systematic reification of racial categories in genomics must thus be understood as neither an aberration from 'good' scientific practice, nor apolitical pragmatism, but as intrinsic to how scientific knowledge in modernity is produced for its use value.

A different interpretation of the same data could also be used to undermine raciological orders and ethnic essentialism, something that has been tried to some extent in certain cases[15]. This could have been the real merit of genomic studies like the HGDP, were they configured differently: To use the insights of our shared ancestry to shape a new transient, post-racial and post-ethnic genomic imaginary. But to base political reconciliation on a preceding biologisation of differences is probably not the best way to erode existing social divisions. Maybe, and this is what I hope, new technologies such as the possible convergences I alluded to in this text, will allow for more interpretative flexibility and provide us with tools to do just that.

---

[15] BBC News Online: "Scientists have analysed the genomes of five southern Africans, including Archbishop Desmond Tutu, to study their genetic diversity and health. ... Desmond Tutu: 'It would be disastrous if scientists were to ignore the diversity of the human race because this is the greatest asset of humanity'" From:
http://news.bbc.co.uk/2/hi/science/nature/8519954.stm
Accessed: 24/08/2010

GENOMIC DATA SOURCE	COMPANY	PRICE	TEST	RESULT	RECENT COMMON ANCESTRY
mtDNA	23andMe	$ 499*	Haplogroup	T2b4	Northern Europeans, Spanish
	Family Tree DNA	$ 843*	Haplogroup	T2b	England, Estonia, Germany, Greece, Italy, Poland, Scotland, United Kingdom
	Oxford Ancestors	£ 340*	World Clan / Seven Daughters of Eve	Tara	Tuscany- France- England- Ireland**
Y-DNA	23andMe	$ 499*	Haplogroup	R1a1**a**	Ukranians, Indians, Poles
	Family Tree DNA	$ 843*	Haplogroup	R1a1	Greece, Hungary, India, Ireland, Latvia, Lithuania, Norway, Poland, Portugal, Russian Federation, Taiwan, Ukraine
	National Geographic	$ 99	Haplogroup	R1a1	Czech Republic, Siberia, Central Asia, Middle East, India, Iran
	Oxford Ancestors	£ 340*	Paternal Clan	Sigurd (R1a)	Norway, Iceland, Orkney and Shetland Islands
Genome	23andMe	$ 499*	Genotyping	European	(US),

– Results from genetic ancestry testing companies

Legend for Table 1:

* Combined tests price

** Temporally differentiated (last = latest)

*** Post-colonial states, usually excluded

**** Depending on settings: Segment size > 10 cM (centi-Morgans) / < 10 cM

Table 2 – Y-DNA Recent Ancestral Origins according to Family Tree DNA

Exact Matches					
Country	Your Matches	Comment	Match Total	Country Total	Percentage
Greece	3	-	3	635	0.5%
Hungary	1	-	1	1,018	0.1%
India	2	-	2	1,236	0.2%
Ireland	1	-	1	12,366	< 0.1%
Lithuania	1	-	1	882	0.1%
Norway	1	-	1	1,195	0.1%
Poland	2	-	2	3,213	0.1%
Portugal	3	-	3	694	0.4%
Russian Federation	1	-	1	2,715	< 0.1%
Taiwan, Province of China	1	Han	1	157	0.6%
Ukraine	1	-	1	1,356	0.1%
One Step Mutations					
Country	Your Matches	Comment	Match Total	Country Total	Percentage
Austria	1	-	1	539	0.2%
Belgium	1	-	1	455	0.2%
Czech Republic	1	Czechoslovakia	1	606	0.2%
Denmark	1	-	1	743	0.1%
England	11	-	11	21,195	0.1%
Finland	1	-	1	1,510	0.1%
Georgia	1	-	1	36	N/A

Germany	15	-	15	10,757	0.1%
Greece	3	-	3	635	0.5%
Hungary	3	-	3	1,018	0.3%
Iceland	1	-	1	143	0.7%
India	7	-	7	1,236	0.6%
Indonesia	1	-	1	654	0.2%
Ireland	3	-	3	12,366	< 0.1%
Italy	1	-	2	3,062	0.1%
	1	Sicily			
Lithuania	4	-	4	882	0.5%
Macedonia	1	-	1	46	N/A
Mongolia	1	-	1	579	0.2%
Myanmar	1	-	1	3	N/A
Netherlands	1	-	1	1,492	0.1%
Norway	3	-	3	1,195	0.3%
Poland	7	-	8	3,213	0.2%
	1	Prussia			
Romania	1	-	1	491	0.2%
Russian Federation	8	-	8	2,715	0.3%
Saudi Arabia	1	-	2	399	0.5%
	1	Arab			
Scotland	10	-	10	9,925	0.1%
Serbia	1	-	1	67	N/A
Slovakia	2	-	2	458	0.4%
Slovenia	1	-	1	131	0.8%
Sweden	7	-	7	1,460	0.5%

Switzerland	1	-	1	1,576	0.1%
Syrian Arab Republic	1	Arab	1	186	0.5%
Turkey	1	-	1	420	0.2%
Ukraine	2	-	2	1,356	0.1%
United Kingdom	5	-	5	9,425	0.1%
Uzbekistan	2	-	2	156	1.3%

The chart displays:

-Each country from which you have matches

-The number of people you match for each country and comment combination

-Any additional information your matches provided about their origins

-The total number of people you match from that country

-The total number of people who have reported this as their country of origin

-The percent of the people we have tested from this country who match you.

Please note if the number of people reporting a particular country is too small, no percentage will be calculated and N/A will appear instead.

Percentages above 2% are considered **significant** indicators of your family's origins.

Percentages above 4% may be interpreted as **highly significant** indicators of your family's origins.

To update or view how your ancestral origin is recorded in our database, click on the link above titled Update Contact Information.

How to read this chart:

The following is an example of how to correctly interpret your matches below:

You match 3 person out of 635 people from Greece, this is 0.5% of the population tested from Greece.

From: https://www.familytreedna.com/my-ftdna/y-dna-recent-ancestral-origins.aspx (restricted access)

Accessed 12/08/2010

Table 3: mt-DNA Recent Ancestral Origins according to Family Tree DNA

LOW RESOLUTION (HVR1) MATCHES					
Country	Your Matches	Comment	Match Total	Country Total	Percentage
England	3	-	3	6,737	< 0.1%
Estonia	1	-	1	60	N/A
Germany	1	Baden-Württemberg	1	7,563	< 0.1%
Greece	2	-	2	391	0.5%
Italy	2	-	2	2,513	0.1%
Poland	2	-	2	2,811	0.1%
Scotland	2	-	2	2,803	0.1%
United Kingdom	1	-	1	5,270	< 0.1%

HIGH RESOLUTION (HVR1+HVR2)					
Country	Your Matches	Comment	Match Total	Country Total	Percentage
No Matches.					

FULL GENOMIC SEQUENCE MATCH					
Country	Your	Comment	Match	Country	Percentage

Matches	Total	Total

No Matches.

The chart displays:

-Each country from which you have matches

-The number of people you match for each country and comment combination

-Any additional information your matches provided about their origins

-The total number of people you match from that country

-The total number of people who have reported this as their country of origin

-The percent of the people we have tested from this country who match you

How to read this chart:

The following is an example of how to correctly interpret your matches below:

You match 3 person out of 6,737 people from England, this is < 0.1% of the population tested from England.

From: https://www.familytreedna.com/my-ftdna/mtdna-ancestral-origins.aspx

(restricted access)

Accessed 12/08/2010

References

23andMe White Paper 23-04 (2008). Global Similarity's Genetic Similarity Map. <https://www.23andme.com/res/pdf/23-04_Genetic_Similarity_Map.pdf> Accessed: 17/08/2010

Agamben, G. (1998). Homo sacer : sovereign power and bare life. Stanford, Calif., Stanford University Press.

American Society of Human Genetics (ASHG) (2008). Ancestry testing statement. <http://www.ashg.org/pdf/ASHGAncestryTestingStatement_FINAL.pdf> Accessend: 24/08/2010

Bamshad, M. J. et al. (2003). Human population genetic structure and inference of group membership. American Journal of Human Genetics, 72, p. 578-589.

Bolnick, D. (2008). In Koenig, B. A., S. S.-J. Lee, et al. (2008). Revisiting race in a genomic age. Piscataway, N.J., Rutgers University Press ; London : Eurospan [distributor]. Ch. 4.

Cavalli-Sforza, L. L., P. Menozzi, et al. (1994). The history and geography of human genes. Princeton, N.J. ; Chichester, Princeton University Press.

Duster, T. (2003). Backdoor to eugenics. New York ; London, Routledge.

Foucault, M. and M. Senellart (2008). The birth of biopolitics : lectures at the Collège de France, 1978-79. Basingstoke, Palgrave Macmillan.

Franklin, S., C. Lury, et al. (2000). Global nature, global culture. London, SAGE.

Fullwiley, D. (2008). The biologistical construction of race: 'admixture' technology and the new genetic medicine. Social Studies of Science, 38(5), p. 695-735.

Gisli, P. (2010). Spitting Image. <http://anthronow.com/articles/spitting-image> Accessed: 20/06/2010

GottaBeMobile, (2010). Who Owns Your Soul…Do You Read Terms and Conditions?. <http://www.gottabemobile.com/2010/04/19/who-owns-your-soul-do-you-terms-and-conditions/> Accessed: 20/08/2010

Greely, H. T. (2008). Genetic Genealogy: Genetics Meets the Marketplace. In Koenig, B. A., S. S.-J. Lee, et al. (2008). Revisiting race in a genomic age. Piscataway, N.J., Rutgers University Press ; London : Eurospan [distributor]. Ch. 11

Hacking, I. (2009). Current Controversies: Ian Hacking. <http://onthehuman.org/2009/03/current-controversies-ian-hacking/> Accessed: 20/06/2010.

Haraway, D. J. (1997). Modest_Witness@Second_Millennium.FemaleMan_Meets_OncoMouse : feminism and technoscience. New York ; London, Routledge.

Human Genetics Commission (HGC) (2010). A common framework of principles for direct-to-consumer genetic testing services. <http://www.hgc.gov.uk/Client/document.asp?DocId=280&CAtegoryId=10> Accessed: 05/08/2010

Lewontin, R. C., S. Rose, et al. (1984). Not in our genes : biology, ideology, and

human nature. New York, Pantheon Books.

Lewontin, R. C. (2000). It ain't necessarily so : the dream of the human genome and other illusions. London, Granta.

Marks, J. (2008). Race: past, present, and future. In Koenig, B. A., S. S.-J. Lee, et al. (2008). Revisiting race in a genomic age. Piscataway, N.J., Rutgers University Press ; London : Eurospan [distributor]. Ch. 1

Mosse, G. L. (1966). Nazi culture : intellectual,cultural and social life in the Third Reich, W.H.Allen.

Nelson, A. (2008a). Bio science: genetic genealogy testing and the pursuit of african ancestry. Social Studies of Science, 38(5), p.759-783.

Nelson, A. (2008b). The factness of diaspora: social sources of genetic genealogy. In Koenig, B. A., S. S.-J. Lee, et al. (2008). Revisiting race in a genomic age. Piscataway, N.J., Rutgers University Press ; London : Eurospan [distributor]. Ch. 13.

Outram, S. M., Ellisson, G. T. H. (2010) Arguments against the use of racialised categories as genetic variables in biomedical research: what are they, and why are they being ignored?. In Whitmarsh, I., Jones, D. S. (2010) What's the use of race?. Cambridge, Massachusetts ; London, MIT Press. Ch. 5.

Rabinow, P. (1992). Artificiality and enlightenment: from sociobiology to biosociality. In Crary, J. ed., Zone 6: Incorporations. Cambridge, Massachusetts ; MIT Press.

Reardon, J. (2005). Race to the finish : identity and governance in an age of

genomics. Princeton, N.J. ; Oxford, Princeton University Press.

Rose, N. (2007). Politics of life itself : biomedicine, power and subjectivity in the twenty-first century. Princeton, N.J. ; Oxford, Princeton University Press.

Rosenberg, N. A. et al. (2002). Genetic Structure of Human Populations. Science, 298, p. 2381-2385.

Shriver, M. D., Kittles, R. A. (2008). In Koenig, B. A., S. S.-J. Lee, et al. (2008). Revisiting race in a genomic age. Piscataway, N.J., Rutgers University Press ; London : Eurospan [distributor]. Ch. 10.

TallBear, K. (2008). Native-american-dna.com: in search of native american race and tribe. In Koenig, B. A., S. S.-J. Lee, et al. (2008). Revisiting race in a genomic age. Piscataway, N.J., Rutgers University Press ; London : Eurospan [distributor]. Ch. 12

U.S. Government Accountability Office (2010). Direct-to-consumer genetic tests: misleading test results are further complicated by deceptive marketing and other questionable practices. <http://www.gao.gov/products/GAO-10-847T> Accessed 20/08/2010

Wajcman, J. (2004). TechnoFeminism. Cambridge ; Malden, MA, Polity.

Wall Street Journal (2010). FDA officials not sure agency will approve direct-to-consumer genetic tests <http://online.wsj.com/article/BT-CO-20100716-711161.html> Accessed: 20/08/2010