Comment

What my genome told me - and what it didn't

Gregory A Petsko

Address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA.
Email: petsko@brandeis.edu

Published: 29 June 2009

Genome Biology 2009, **10:**108 (doi:10.1186/gb-2009-10-6-108)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2009/10/6/108

Well, it turns out I'm not descended from Genghis Khan. I'm sure that's as surprising to you as it is to me. I mean, according to what we hear from people who use genomics to track human migrations, a huge percentage of the human race is actually descended from Genghis Khan. But not me.

That's one of the things I learned when I submitted a sample of my DNA for genome-wide single nucleotide polymorphism (SNP) analysis by one of the companies that have sprung up to perform such tests for ordinary individuals for a fee. I was curious to see what sort of information they provide, and wanted to know something about my own genomic makeup, to be honest, so I followed the directions of the company I had selected and spat into a plastic container until I produced the required volume of saliva, mailed it in, and awaited the results. Would I have an allele that doomed me to a rare genetic disorder as I got older? Was I at much higher than normal risk for heart disease, diabetes, or any of the other thousand natural shocks that flesh is heir to? Was I descended from Genghis Khan?

The company I sent my saliva sample to doesn't actually do any DNA sequencing or hybridization itself; it contracts this out to a specialist laboratory. Once the lab received my sample, DNA was extracted from cheek cells in the saliva and amplified by PCR to produce enough material for the genotyping step. Next, the DNA was cut by restriction digestion into smaller, more manageable pieces. These DNA pieces were then applied to a DNA chip, which in this specific case is a small glass slide with millions of microscopic beads on its surface. Attached to each bead are the probes - bits of DNA complementary to those specific sites in the human genome where important SNPs are located. There is a pair of probes for each SNP, corresponding to the 'normal' and 'non-normal' version of each SNP. Hybridization to the particular probe, detected by fluorescence just as in the case of any other DNA chip experiment, serves to identify the allele.

The DNA chip that this particular company uses reads 550,000 SNPs that are spread across the 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 SNPs are specially selected 'tag SNPs' - because many SNPs are linked to one another, the genotype at many SNPs can often be determined by looking at one SNP that 'tags' its group. This tagging procedure maximizes the information from every SNP analyzed, while keeping the cost of analysis low.

In addition, all the DNA analysis companies 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. These SNPs include risk factors for common and rare human diseases, genetic traits such as color blindness, and so on.

Access to the resulting data is through the company's website, which includes the ability to download the entire set of SNP information. Once I was notified that my results were in, I did that, and being a scientist I performed my own bioinformatics on the data, but the website actually does a pretty good job of providing the customer with specific information about alleles for various illnesses, physical traits, and so on.

Here are a few of the things I leaned about myself, physically speaking:

According to my genome, my eye color is likely to be brown (good guess). I should be lactose tolerant (I am). My cytochrome P450 data show that I would be quite sensitive to the anti-clotting drug warfarin if I ever had to take it (which I hope I never do - it's a nasty drug). The SNPs in my androgen receptor gene say that I am considerably decreased in risk for male pattern baldness (I have news for them; I'm getting rather thin on top). I have a SNP in a

dopamine receptor gene that, in one German study, was found to be associated with reduced efficiency in learning to avoid errors (unless I got the facts wrong). According to a single SNP in one gene associated with insulin metabolism, I have increased odds of living to be 100 (that is, if all the mistakes I don't learn to avoid don't get me first). There are a number of SNPs that have, in some studies, been associated with increased athletic performance (faster running, quicker reaction times, and so on). I don't have any of them, which will come as no surprise to any of my gym teachers.

I am at slightly increased risk, relative to the norm, for rheumatoid arthritis and psoriasis (that last one is interesting, because my father suffered from it). I am at slightly decreased risk for Celiac disease, Crohn's disease, type 1 diabetes, and prostate cancer. In all cases, the change is small - less than two-fold difference, and not enough to cause me to consider any lifestyle changes.

But one thing did jump out at me when I looked at my data. I have a guanine at rs1799945, which is located in the gene coding for a protein called HFE. HFE is the protein mutated in hereditary hemochromatosis. Hereditary hemochromatosis, the most common form of iron overload disease, is an autosomal recessive genetic disorder that causes the body to absorb and store too much iron. Excess iron is stored throughout the body in organs and tissues, including the pancreas, liver, and skin. Without treatment, the iron deposits can damage these organs and tissues. There are two primary variants that give rise to this disease.

Genetic variant 1 (C282Y/rs1800562) is in the *HFE* gene. The *HFE* gene makes a membrane protein that is similar structurally to MHC class I-type proteins and associates with β2-microglobulin. It is thought that HFE helps cells in the intestines, liver, and immune system control iron absorption by regulating the interaction of the transferrin receptor with transferrin. The C282Y substitution disrupts the interaction between HFE and its β2-microglobulin light chain and prevents cell-surface expression. Pamela Bjorkman's 2.6 Å resolution crystal structure of HFE confirms that, as predicted from its sequence, Cys282 (residue 260 in the mature form of the protein) is involved in a disulfide bridge analogous to those found in class I MHC α3 domains. Loss of the disulfide destabilizes the native fold of the protein. The second most common variant is also in the HFE gene. It is a change of histidine 63 to aspartic acid. In the crystal structure of HFE, His63 (41 in the sequence of the mature form) is involved in a crucial salt bridge, which would be destroyed by mutation to a negatively charged residue, thereby also destabilizing the protein. Thus, like so many other hereditary disorders, hemochromatosis is a protein conformational disease.

In the US, variant 1 is the more frequent. The 'normal' Cys282 allele has guanine in both strands and is found in about 876 out of 1,000 people of European ancestry. The most common form of hereditary hemochromatosis is typically associated with people homozygous for an adenine in both positions; this occurs in about 4 out of 1,000 people of European ancestry (0.4%). However, penetrance is incomplete: only about a third to a half of the homozygotes will show elevated iron levels and perhaps fewer than 10% of the males (and 1 to 2% of the females) will develop the full clinical symptoms of the disease, which include joint pain, fatigue, abdominal pain, liver dysfunction, and heart problems. As Ernest Beutler has pointed out, the hemochromatosis mutation is relatively common; the hemochromatosis disease is rare. Mutation of the HFE gene is a necessary, but not sufficient, condition. The challenge, as in the case of so many diseases in the age of genomics, is to understand what other genetic, epigenetic, and environmental factors determine why a few homozygotes for the C282Y (or H63D) mutations develop severe iron-storage disease, while the majority go through life pretty much unscathed by this genotype.

Heterozygotes for C282Y have an adenine in only one strand and represent about 120 in 1,000 people of European ancestry; they almost never develop clinical symptoms. Heterozygotes for H63D are less common, but also are unlikely to develop clinical symptoms. Like one person in ten in the United States, I am a hemochromatosis carrier. I am heterozygous for H63D.

Now that I know that, what does it do for me? Not much, I guess, but I will remember it, and should I ever develop any of the symptoms of iron overload, I will probably tell my physician to check my iron levels. Maybe that's worth knowing.

But if you go to the website of the company that did my analysis, you will find that the sort of information I've talked about here is actually not that prominently displayed. What is displayed are all manner of data connecting to ancestry. I spoke to the CEO of the company, and she confirmed that, much to their surprise, people who use their service are much more interested in tracing their roots, genetically speaking, than they are in things related to their health or physical condition. The site offers several tools for connecting yourself with others who share your ancestry, genetically speaking. In other words, for the present, the primary use of personal genome-wide SNP analysis is social networking.

My maternal haplogroup is T2b2. Haplogroup T originated about 33,000 years ago in the Middle East, as modern humans first expanded out of eastern Africa. Its present-day geographic distribution is strongly influenced by multiple migrations out of the Middle East into Europe, India, and eastern Africa after about 15,000 years ago. T2 is currently widespread in northern Africa and Europe. My mother's family most recently came from Italy, so I guess this makes sense. You can find famous people with your haplotype on the site: for example, if your maternal haplotype is H4a, you have the same type as Warren Buffet, one of the richest men

in the world. You'll be delighted - and perhaps not surprised - to know that the only famous person the site lists as sharing my maternal haplotype is the notorious old-west outlaw Jesse James.

My paternal haplotype is I2. Haplogroup I2 is most abundant in eastern Europe and on the Mediterranean island of Sardinia, where it is found in 40% of the male population. Like its brother haplogroup, I1, I2 expanded northward at the end of the Ice Age about 12,000 to 14,000 years ago. But unlike I1, which expanded from the Iberian peninsula into northwestern Europe, I2 radiated outward from the Balkans and southwestern Russia into the eastern half of the continent. That again makes sense, as my father's family were Cossacks. If my paternal haplotype were the extremely common C3, I would be descended from Genghis Khan. I'm not. If it were type T, I would share paternal lineage with the great American president and founding father Thomas Jefferson. I don't. In fact, the company website doesn't list a single famous person with paternal haplotype I2 (unless you count me, of course).

So now, thanks to my own personal genome SNP analysis, I know that I'm not likely to be exceptionally athletic and that I'm not a blue-eyed balding blonde, neither of which comes as any surprise whatsoever. But I have also learned that I'm not descended from Genghis Khan. So I've got that going for me. Which is something, I suppose.