Eastman's Online Genealogy Newsletter

"DNA and Family History, Ten Years On" by Chris Pomery

Posted on 10 March 2009

Time certainly flies. I find it hard to believe that genetic testing aimed at genealogists has been around for ten years. A decade ago DNA tests were rarely seen on TV and broadly unused outside of academia. Roll forward to today, and no one researching their family history can remain unaware of DNA testing for very long. Even Ancestry has finally started offering cut-price DNA tests, chasing the coat tails of the market leader, Houston-based Family Tree DNA, which boasts nearly 150,000 Y-chromosome results and more than 5,000 surname projects.

Yet despite all this activity by family historians, there's a great deal that we've been taking on trust during the past decade.

Firstly, there have been very few write ups of surname-linked DNA projects, even in academia. Most activity has been led by family historians outside of it, but only a few of us have ever tried to write up our results in a way that allows outsiders to compare one surname project with another. And then there are lots of assumptions and variables within the process of analysing a set of DNA results that still have to be quantified or proved.

That situation is in the process of changing, but even as it improves, it reminds us again that DNA testing is just a tool used by family historians, not an end in itself.

The original link between surnames and Y-chromosome DNA testing was laid out in a paper published in 2000 by an Oxford professor, Bryan Sykes, and a colleague, Catherine Irven. They tested 48 men who shared the Sykes name and lived in the area of northern England where the name was understood to originate. Their surprising conclusion was that the Sykes surname had a single common ancestor. That is to say that there was a single man from whom all men with the surname Sykes are linked within a single family tree stretching back no more than 700 years.

I say surprising because Sykes is not an uncommon surname in England. In the 1881 census of England and Wales, it was ranked in the top 300 surnames, and I suspect that there are more than thirty thousand name-bearers in the UK today. My guess is that if you asked most British people in the street if they know someone called Sykes, then they'd say yes. I was so surprised by the paper's conclusion that I emailed the professor to tell him so, which was my first involvement in the field of DNA testing.

This original Sykes paper has been cited ever since, though its methodology looks crude by comparison to what we can do today with much more powerful tests, more results to compare with, and more complex classifications and definitions of results at our fingertips.

From a commercial lab, I can readily buy an off-the-shelf Y-chromosome test that probes 67 markers on my Y-chromosome: the Sykes study analysed just four. The haplogroup results we get back from the lab – the ones which hint whether our male ancestor was an early post-Ice Age coloniser in Britain or a later arrival -- now run to as many as eleven sub-divisions defined by a global committee. Bryan Sykes made his up and called them haplogroup I and II. (A year later, when his book The Seven Daughters of Eve became a best-seller, he'd learned how to sex up his classifications a bit more!)

Since then academic geneticists have generally avoided looking at or using surnames, focusing broadly on different types of migration studies (and hosting TV series!). A notable exception is a privately-financed project by Trinity College Dublin, which produced some fine results in 2006. After looking at a range of surnames, they found that men were 30 times more likely to share the same DNA if they have the same surname. They also concluded that surnames such as O'Sullivan and Ryan have a single ancestor at their head while others, such as Murphy and Kelly, do not.

One group in Leicester University, however, is now taking a closer look at the relationship between DNA results and surname formation, and their results are both interesting and challenging. Last year they published the results of a study using surnames found in the areas of west Lancashire and the Wirral, traditionally the scenes of Viking settlement on the north-west coast of England, to see if surnames of Scandinavian origin contained a strong hint of Scandinavian DNA. Which they did. Right now they are repeating the exercise in Yorkshire and the counties along the north-east coast of England to try and differentiate genetically between Norwegian and Danish Vikings.

A few weeks ago this group released a paper with the exotic title of "Founders, drift and infidelity: the relationship between Y chromosome diversity and patrilineal surnames." This is the first major paper to evaluate the linkage between English surnames and DNA testing since the original Sykes paper ten years ago, and some of its conclusions are different.

The team analysed 40 surnames, ranging from rarely-found ones to very common ones. As the Irish team had found, they identified a "remarkably high degree of coancestry that generally increases as surnames become rarer." In other words, if you hold a rare surname, you are much more likely to share your DNA with your other name bearers than you are if you hold a common surname.

So far so obvious, but another of their conclusions – that "modern patterns" of DNA signatures within a surname "provide little reliable information about the original founders of surnames some 700 years ago" – is more challenging for family historians running their own DNA projects. The team concluded that while many surname

DNA projects will report a modal result – that is, they will identify one DNA result or signature that appears to dominate among the modern descendants bearing the same surname – the fact that such a result is commonly found doesn't mean that this DNA signature belonged to the original founder. In other words, the leap of faith made in the Sykes paper back in 2000, that a strong modal result indicates both a single ancestral origin to the surname and the genetic identify of that founder, was a leap too far.

One reason for their conclusion is that the type of surnames among the forty they studied that reported strong modal results aren't always those that one would first suspect of having a single ancestor at their head. One is Herrick, which is derived from a first name. Another reason is that their calculations suggest that the time when the single common ancestor at the head of several surnames seems to have lived is much more recent than was expected.

There's no doubt that the thousands of surname DNA projects under way are usefully using their results to reveal which men within a surname belong within the same tree. But this new paper is saying that the pattern of results found within any surname can't guarantee to find the genetic identity of any original founder. The reason for this is one that will be familiar to anyone who has reconstructed a large, old family tree: while we know that most lines do eventually daughter out, or fail to produce any offspring at all, we simply don't know how often lines die out. What the paper suggests is that this process, known as "variation in reproductive success," is the most influential force in determining the pattern of DNA signatures found in any surname.

So how do we, as family historians, get around this problem? Well, we do what we were doing already: we carry on reconstructing the family trees through traditional documentary means. And that's the key point for all of us ten years on: a DNA result is not your history; your family tree is your history. Collecting DNA results in a big spreadsheet is a great starting point, but the real work lies in documenting those trees together.

Within my own Pomeroy project I'm much less concerned today about the single vs. multiple ancestor question than I used to be. In our early days Bryan Sykes confidently said that our surname had multiple ancestors. However, ten years later the trees we're recreating hint strongly that, in the end, they will all eventually link up. But we'll only know for sure when we finish documenting it.

References

Bryan Sykes & Catherine Irven
Surnames and the Y Chromosome
www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1288207

Brian McEvoy & Dan Bradley
Y-chromosomes and the extent of patrilineal ancestry in Irish surnames
www.ncbi.nlm.nih.gov/pubmed/16408222

Turi King & Mark Jobling
Founders, drift & infidelity: the relationship between Y chromosome diversity and patrilineal surnames
http://mbe.oxfordjournals.org/cgi/content/abstract/msp022

Biographical Description

Chris Pomery has published two books on DNA testing with the imprint of The National Archives in London and lectures regularly on DNA testing and surname reconstruction. He has run the Pomeroy DNA project since its inception in 2000. The website for the Pomeroy surname project is at www.pomeroyfamilyhistory.com