

DNA AND FAMILY HISTORY

**A lecture presented to Canterbury Festival on 21 October 2009
by Dr Richard C F Baker, FHG**

It is always difficult to know quite how to pitch a presentation of this nature, largely because it draws together quite diverse and disparate ideas, and therefore attracts people from a wide range of backgrounds.

First I'll start with a quick tutorial on genetics – the science of inherited characteristics. Geneticists, please bear with me as this is going to be a gross oversimplification of a complex but fascinating subject; hopefully it will be enough for non-scientists to understand the basic processes.

It is no accident that the two words genetics and genealogy derive from a common etymological root. Genetics is the study of the mechanism of heredity, whereas genealogy is the study of the effects of heredity.

The mind-sets required for these two pursuits – Genetics and Genealogy – are curiously similar; clearly to be a geneticist, you need to be a scientist, someone who is disciplined in research; but successful genealogists have to be disciplined researchers, too, and it should not be too much of a surprise to learn that many of the leading names in Genealogy are, in fact, highly qualified scientists. As Maureen Lipman in her role as Beattie said a number of years ago in the TV advertisement for BT, “You’ve got an – ology, you’re a scientist!”

GENETICS

Many years ago, the vexed subject of sex education was often broached in terms of the birds and the bees – and I am sure that many of us found this a rather odd metaphor. But curiously enough, during the twentieth century, it became clear that virtually all living things on this planet did, in fact, use the same blueprint as a mechanism for passing on characteristics from parent to offspring.

Discovery of this blueprint was based on the work of a Czechoslovakian monk named Gregor Mendel who conducted experiments on variation in pea plants in the 19th century, and that of the American Thomas Hunt Morgan on the inheritance of distinctive characters by the fruit fly *Drosophila melanogaster* early in the twentieth century.

And that blueprint is called deoxyribonucleic acid – known more familiarly simply as DNA. The structure of DNA was famously elicited by James Watson and Francis Crick working at the Cavendish Laboratory in Cambridge in 1953; this work earned them the Nobel Prize for Chemistry in 1962.

I am sure that everyone will have seen representations of DNA in the form of a double helix – a twisted ladder, if you will – comprising two parallel strands of ribose sugar molecules (the sides of the ladder), linked by pairs of nucleotide bases (the rungs of the ladder).

Now there are just four types of nucleotide base – A (Adenosine), C (Cytosine), G (Guanine) and T (Thymine) –and what is really clever is that Adenosine and Thymine fit perfectly together, as do Cytosine and Guanine, so that A always pairs with T and C always pairs with G which make all the rungs of the ladder exactly equal in length. And knowing that A always pairs with T and C always pairs with G, a very elegant means of reproduction becomes obvious; if the ladder is split down the middle, for example when cells divide, and new strands are built up on each half, the new strands will exactly mirror the old ones.

Thus we have a brilliant code for passing on information - and the information that is passed on from one DNA strand to the next is a set of instructions necessary to create a cell, so that new cells generated by cell division should be identical to the original cell – it is, effectively, a blueprint for life.

In fact, it is such a superlative code that only about 10% of the DNA of any advanced species is used to define the next generation, the remaining 90% being derisively called “junk” as, until recently, we had no idea what its function might be.

The replication of DNA is not quite perfect, however, because slight errors sometimes arise, for a variety of reasons. These errors are mutations, but because of the implications associated with that word (conferred by popular novelists), geneticists prefer the term polymorphisms. And once a polymorphism has occurred, it normally remains unchanged. Those that arise in the active 10% of our DNA ultimately account for all of the differences that we see between each human being – including genetic traits such as blood group, hair colour or the propensity to inherit genetic diseases.

But, such mutations also occur in the “junk” regions, where they have no detectable effect on the individual - but they do create a permanent unique DNA signature for every individual (identical twins apart) as well as recording our ancestral history. There are two basic types – single nucleotide polymorphisms (or SNPs, usually called “Snips”) in which a single nucleotide base is affected, and then there are Short Tandem Repeats (STRs, which don’t seem to have a pet name. Let’s see if I can start a trend by calling them “Stars”!), in which groups of two or three base pairs are added or omitted. These are most significant for us, because it is these changes that are analysed during genealogical DNA tests.

Inside every normal cell, we find a nucleus, populated by chromosomes which are made up of DNA and protein. Also in the cell are other, smaller organelles called mitochondria, which are the power houses of the cell, providing energy, and they also contain a little bit

of DNA – just the same as nuclear DNA, but with the ends of the strand joined to form a circle.

In the nuclei of normal human cells, there are 46 chromosomes, arranged as 22 pairs and two odd bits known as the X and the Y chromosomes, which are of different shape and cannot form a pair. Each of the 22 pairs carries about 1500 genes (the active bits of the DNA), but the X and Y carry only 2 active genes, so they are particularly useful for DNA analysis as they are mostly “junk” which, as I said, is where we look for the Snips and the Stars.

And, it is the X and Y chromosomes that determine the gender of the offspring. Normally a woman has two X chromosomes in the nucleus of each cell in her body (one inherited from each of her parents); a man will have one X chromosome (inherited from his mother) and one Y chromosome (inherited from his father). When reproductive cells are being formed, they each carry only 23 chromosomes, half of the normal cell’s complement. All of a woman’s eggs have a single X chromosome, but only 50% of a man’s sperm cells have a single X chromosome, the other 50% having a single Y chromosome). When an X sperm fertilizes an egg, the embryo gets the X from the sperm and another X from the egg and, having two Xs is, therefore, female. When a Y sperm fertilizes an egg, the embryo receives one X and one Y chromosome and is, therefore male.

The important thing here is that the Y chromosome is passed from father to son, generation after generation, completely unchanged – or I should say almost completely unchanged; if there were no changes at all, all men would have the identical Y chromosomes, and the DNA test would be unable to distinguish between them!

Similarly, the DNA in the mitochondria - mt-DNA as it is known - is passed from a female to her offspring, which means that it is unchanged in the maternal line; it is however, too short a strand to allow the same sort of analysis that is possible with Y-DNA (mt-DNA contains about 16 thousand base pairs, whereas Y-DNA has about 5 billion), and so it is of no real use to genealogists, but it does allow us to make generalised inferences regarding the racial origins of an individual. The book *The Seven Daughters of Eve* by Bryan Sykes is the key text exploration of mt-DNA analysis.

So for these and other reasons, it is the Y-chromosomal DNA that is routinely tested by the genealogical DNA companies, and it is the Y-DNA test upon which this presentation is based.

FAMILY HISTORY

Genealogy encompasses the acquisition of information necessary to compile what we call a PEDIGREE or FAMILY TREE. It is a chart representing relationships between individual people, and it will often include basic details such as dates and places of birth, marriage and death.

Genealogy, it is said, is the second oldest profession. The earliest civilisations in the Euphrates and Nile valleys recorded successions of dynasties reaching back into antiquity, usually to establish rights to title and property. Indeed, right up to the 19th century, scant attention was paid to describing technique; the mere compilation of results was the sole objective. However, in 1828, Stacey Grimaldi published his *Origines Genealogicae*, which was the first practical guide to research. But until the middle of the twentieth century, family history was largely the province of antiquarians tracing the genealogies of the great and the good.

In 1957, Cecil Humphery-Smith started a school of family history, and moved it to Canterbury as the forerunner of this Institute of Heraldic and Genealogical Studies which was founded in 1961. The publication of the book *Roots* by Arthur Hayley started to arouse real interest in the subject, and then in 1979 there was a BBC TV series on the family history of newsreader Gordon Honeycombe – which inspired a lot of people to try tracing their ancestors; and many family history societies started to be formed. And so information about what sources to use (GRO, Census, PRs, Wills, Service Records etc) was disseminated to the public for the first time.

There followed a general increase in the number of historical television series, and the family history movement reached a sort of critical mass, generated by popular television shows, in particular the BBCs WDYTYA which makes it all look so easy, and fuelled by growth of information available in the home through the personal computer, so that the number of people getting into family history almost literally exploded. And Family History is still growing, with the advent of computer sites such as Genes Reunited allowing people across the globe to compare and discuss their findings and to pool their information. It is certainly true to say that there has never been such an upsurge of interest in family history as there is now.

All family historians start out by becoming genealogists - they use the records of birth, marriage and death – or hatch, match and despatch as genealogists like to say – to compile pedigrees with the sole objective apparently of boring everyone who will listen by telling them how far back they've got. Much like those early genealogists, in fact! Then we develop into Family Historians by going further to flesh out the bones of their skeletal pedigrees by finding out more about each of our ancestors – what their life was like, their aspirations and motivations – and thus bring them to life as real people rather than as mere names on a chart.

But be warned, Family History is an insidious disease – not only is it incurable, but you will find that it pervades your life 24 hours a day, 7 days a week – you may well find that you awaken in the middle of the night with some flash of inspiration as to where to seek your forebears – you will find yourselves buying or making new bookshelves to accommodate a rapidly expanding genealogical library. Another symptom of this disease called Family History, of course, is that you will start planning future holidays around your ancestral researches!

Why do people “do” family history? Quite a variety of reasons – often the acquisition of knowledge for its own sake is enough, running down stories of lost fortunes, or perhaps a long-term aim of employment. Some seek famous, or better still, infamous ancestors, ancestors who made history.

But if you think about it, Nelson for example may have taken all the credit for the famous naval victory at the Battle of Trafalgar, but he wouldn't have been able to do it without the foresters and sawyers, joiners and shipwrights, iron and brass founders, chandlers and merchants, mariners and powder-monkeys, sail and rope-makers and so forth, amongst whom we might well find our ancestors; they may not have been famous, their names may not have been recorded in this context, but they were real people and they each had an influence on the result of that battle against the Franco-Spanish fleet off Cape Trafalgar two centuries ago. Every one of our ancestors helped to make history!

And don't think that, just because you have no belted earls or coats of arms in your family, as far as you know, that there never were any. Family fortunes do fluctuate; the 7th son of the 7th son of a Duke could quite possibly be an agricultural labourer; but family fortunes, like investments, can go up as well as down - Queen Anne's mother, for example, was of humble stock.

WHAT GENETICS DOES FOR GENEALOGY

DNA testing is now a part of everyday life, taking an ever more important role in forensic criminal investigations to identify the perpetrator of some crime, or on the other hand to identify the victim. I am sure that you are all aware of the number of popular TV shows based on forensic investigation – *Silent Witness* and *CSI*, for example. These shows might exaggerate the efficacy of some of the procedures, but the DNA testing is a very real weapon in the forensic detective's arsenal – only this week in Australia, an armed robber has been brought to justice, eight years after the event, because his DNA matched that in the blood taken from a leech found at the scene of the crime in 2001. DNA analysis has also been used to identify the remains of the children of the Russian Tsar Nicholas, and prove conclusively that all had been killed during the revolution in 1917, including Anastasia. In another celebrated case, Dr Crippen was executed for murdering his wife, but recent DNA testing on the victim's remains showed that it was not his wife's remains that were found buried in the cellar; so was Crippen innocent? Well, his wife was never seen again, and that other body in the cellar would have been difficult to explain....

It is this ability of DNA testing to provide a unique profile of any individual that makes it a tool of great value to Genealogists; genealogy, after all, is about identifying individuals and determining their relationships to one another.

DNA testing, put simply, involves the subject providing a sample of bodily tissue, which is sent off to the testing laboratory, where the DNA is extracted and analysed in a large machine.

Many people cite one of three reasons for not taking a DNA test:

- 1) It is dangerous: the test does not involve any form of blood-letting; in fact, all you usually have to do to provide a sample, is to rub a swab similar to a cotton bud or Q-tip against the inside of your cheek; a number of dead cells, each with a nucleus containing chromosomes is transferred. Providing the sample is significantly less dangerous than cutting your finger nails! And the most complicated thing about it is in filling on the forms!
- 2) Once the testing company has got my DNA I have no control over what happens to it: well, no – depositing DNA with a testing company has been likened to depositing money in a bank – they look after it, but it is still your own; you can even ask them to destroy it once testing is complete. Testing companies are bound by a very strict code of ethics not to disclose your personal information to third parties; those forms that I mentioned limit what the company can do – and they would surely be closed down if they transgressed.
- 3) Insurance companies might get hold of my DNA profile and use it to raise my premiums: in fact they would have absolutely no use for an individual's genealogical DNA profile – as I said, testing is done on the so-called “junk” DNA, not on the active DNA which is where all mutations causing genetic diseases occur.

The actual process of DNA analysis is actually simple enough to be automated, but we don't need to know what is involved there – in much the same way that you don't need to know how a computer is built simply to send an email.

What you do need to know, however, is how to interpret your test results.

This is what they might look like for a basic 10-marker test; you will get 2 rows of numbers – the top line is the identification code of the marker – markers are the particular sites on the chromosome that have been selected by the testing company - and the number beneath each is the number of SNIPs or STARS found in that particular bit of Y-chromosomal DNA.

This row of numbers is your DNA signature.

DNA Test Result

Marker	19	388	390	391	392	393	399i	399ii	425	426
Richard B	13	11	24	10	14	13	13	29	12	12

It might seem pretty meaningless, and even more so when you discover that there may be hundreds of thousands of men on this planet with an identical signature!

A more discriminating test covering up to 67 markers will cut down the number of men sharing the same signature; but for most family historians, a medium resolution test of 25 markers is regarded as adequate – that would provide a signature shared by perhaps 50,000 men in Britain.

Now, even if your name is Smith, very few of those 50,000 will share the same surname as you – and in genealogy, they are the very people in whom you will be interested. Statistically, it is very probable that someone sharing your DNA signature AND your surname is closely related to you.

Which means that the value of genealogical DNA testing is going to be realised when you compare your result with others of the same surname.

Comparative DNA Test Results

Marker	19	388	390	391	392	393	399i	399ii	425	426
Richard B	13	11	24	10	14	13	13	29	12	12
John B	13	11	24	10	14	13	13	29	12	12
Tony B	13	12	24	10	14	13	13	29	12	12
James B	13	11	23	11	13	14	11	27	12	12
David P	13	11	24	10	14	13	13	29	12	12
Charles P	13	11	23	11	13	14	11	27	12	12

Firstly Richard B and John B have identical results – the numerical values on each marker are the same. Tony B's result is almost the same except that on marker 388 he has the value 12 instead of 11. James B's result, however, differs from Richard and John on seven of the ten markers. Turning to David P, his result is identical with Richard and John; Charles P's result is different from David's, but the same as James B's.

It's fairly clear that Richard and John are closely related as their results are identical AND they share the same surname. Having a very similar pattern of results and the same surname, Tony is evidently related, too, but perhaps not quite as closely. David P has the same DNA signature as Richard and John, but he has a different surname, which indicates that, if they did have a common ancestor, it was before the period when surnames were being formed. They are far more likely to have different ancestors, however, because, as I said earlier, there could be many thousands of men with identical DNA signatures but different surnames.

It is clear that in this test, James B is not genetically related to the other three Bs, so his B-line would seem to have had a different paternal ancestor who adopted the B surname independently. But if James has been documented as a member of John and Richard's family then we need to explain why his DNA is different. It is possible that one of James's male ancestors was illegitimate, or that one of his male ancestors' wives was unfaithful, but it would be unwise to leap to conclusions, as there are numerous other explanations – a relative might have adopted the B surname in a legal move to claim an inheritance, for instance.

We cannot infer that James B and Charles P share a paternal ancestor simply because they have the same DNA signature; because they have different surnames, they are probably unrelated.

This shows us that the genealogical Y-DNA test will tell us whether or not two individuals with the same surname are likely to be closely related - or not. The more markers that are used, the more discriminating the test.

Where genealogical DNA testing is of particular benefit to family historians is in the area of One-Name Studies. A One Name Study, as its title implies, concentrates on individuals bearing a single surname, drawing together all discoverable documentary references to the name, with the object of linking all or as many as possible in a single pedigree or a few pedigrees.

In this context, DNA testing can be used

- To prove or disprove a connection with a known historical person.

- To test individuals with the same surname across the world to match emigrant families with their relatives in Europe.

- To establish whether two family trees might be linked through a common ancestor.

There are several hundred ongoing One-Name DNA studies, many with over 200 individuals tested and their DNA signatures recorded.

The first of these was the Sykes DNA Study, led by Oxford University's Prof Bryan Sykes, the results of which were announced in 2000. He concluded that all of the thousands of men bearing the surname Sykes essentially stem from a single common

ancestor of that name. Professor George Redmonds even came up with the probable identity of the original Mr Sykes in Yorkshire in the 13th century.

Another use for DNA testing is currently in progress at Fromelles, where in 2008 a mass grave of some 250 unidentified British and Australian soldiers buried by the Germans following the Battle of Fromelles in 1916 was found. The objective is to discover the identity of each man so that they can be buried with full honours in marked graves; the project is expected to be concluded by February next year.

WHAT GENEALOGY DOES FOR GENETICS

The process of replication of DNA, taken together with the processes of reproduction in living things, means that offspring inherit characteristics from their parents.

If both of your parents have red hair and blue eyes, you too will have those characteristics; if only one of your parents has them, then it is quite possible that you will too.

But there are many inherited characteristics that are not overtly obvious, in particular, in the present context, the propensity to inherit diseases. These so called genetic diseases are caused by transcriptional errors during replication of DNA – mutations or polymorphs and, as I suggested earlier, these changes are permanent, so that they do tend to run in families.

Genetic diseases have, perhaps, been brought into prominence by the twentieth century successes of sanitation and hygiene in virtually eradicating infectious diseases such as tuberculosis, cholera, typhoid and dysentery. Now the big killers are coronary heart disease, stroke and cancer, all of which have a genetic component; this emphasises the importance of genealogists and family historians in maintaining a record of the causes of death of their forebears.

Some genetic diseases are famous, especially haemophilia, of which Queen Victoria was a carrier, and porphyria which had been in the Royal Family for several generations longer.

The application of genealogical research techniques to studies in inherited diseases is not new; our own involvement commenced in 1984, arising from earlier studies on the familial inheritance of Huntington's chorea and retinoblastoma. Our researchers have since then traced families with a wide range of cancerous conditions and illnesses affecting the Central Nervous System.

Ethical considerations mean that we may only conduct research into those families that medical research teams ask us to investigate; normally these will demonstrate a clear

pattern of inheritance of the propensity to a specific disease. Now, of course, most specialised medical teams have their own resident genealogist; this is in itself a cause of concern, for hardly any of them have any genealogical qualification whatsoever. Naturally, because this Institute is the prime educational body in genealogy and our qualifications are recognised throughout the profession, I believe that all genealogists involved in this kind of work should be properly qualified. Well, yes, I would say that; but think of the consequences of, for example, telling someone that they have a genetic predisposition to a fatal condition when they have not – or worse, telling them that they do not have such a predisposition when they do!

Whilst technology has made analysis of DNA almost routine, how can we determine who ought to be screened? Genealogy can provide the answer, because identifying familial linkages between individuals is what genealogists do. Our voluntary research commitment to the medical teams that asked for our help enabled them to resume studies in areas in which they had become “stuck”. Genealogists have the expertise to trace a family back through a number of generations, and in many cases it is possible to track the defective gene by reference to cause of death.

The descendants can then be screened by DNA analysis, and identifying the genetic defect at a pre-clinical stage there can be enhanced prognosis for treatment, or at least counselling at an appropriately early stage.

Sometimes we have the great fortune to trace back two separate families exhibiting the same condition. It happened, for example with two families suffering from a thankfully rare form of Alzheimer’s disease, which had a very early onset at age 35-40 and death age 45-50, the cause being described in the 19th century as “paralysis of the insane”. The two families descended from a common ancestor born in 1838. Today’s sufferers are 6th or 7th cousins and share about 1% of their DNA – thus molecular biologists have only a tiny proportion of the DNA on which to focus their attention.

Occasionally, our research has an international dimension – we were asked by the Royal Hobart Hospital in Tasmania to trace the English relatives of the progenitor of some 1400 Tasmanians who suffer from one particular form of Multiple Endocrine Neoplasia, a poor devil who was transported to Van Diemen’s Land in 1842. Curiously, and perhaps thankfully, too, none of his father’s descendants in Leicestershire exhibit the propensity for this malady, so it appears that the defect started with our transportee.

To conclude, I have brought in concepts that may be outside the experience of many of you, so I hope that I was able to demonstrate a curious symbiosis between two ostensibly unrelated disciplines, in which each can make valuable contribution to the other - and that I was able to make it all understandable.

