http://scotclans-wp.scotclans.dyndns.tv/tracing-your-scottish-ancestry/dna-demystified

DNA Demystified

Dr Bruce Durie

Genetic genealogy is an exciting new development which adds to traditional or "paper" family history research. It is not a replacement, but it can

- supplement information where documentary evidence is missing
- help get around "brick walls"
- identify and connect people with genetic cousins, close and distant
- determine deeper ethnic/genetic ancestry
- possibly indicate geographic origin and track deep ancestral migration
- help to break down brick walls

What DNA testing cannot do is

- pinpoint the "exact" date and identify the specific ancestor shared by two individuals who have a close match
- provide a complete answer to your recent ancestry, even if two participants have a good paper trail
- shed much light, unless compared with the test results of other individuals
- replace "documentary" genealogy

It is necessary to understand the various tests, and how to interpret them. But first…

Some molecular biology

Most cells in the body have a nucleus, which contains 23 pairs of chromosomes – long strands of DNA *(deoxyribo nucleic acid)* wound up with specialised proteins. All cells types in any species have the same DNA, but different parts of it are expressed in, say, skin, muscle, brain, liver etc. This means that almost any cell type from an individual can be used for testing. The cells for inside the cheek scrape off easily, so these tend to be used for testing.

Each of us has inherited one of each pair of chromosomes from our father, one from our mother, and these get *recombined* (they swap pieces with each other) inside us. That's why siblings are not identical unless they grow from the same fertilized egg (identical or monozygotic twins).

One pair of chromosomes has a specialised function – the determination of sex. A male inherits a Y chromosome from the father and its paired X from the mother. A female gets one X from each.

Paternal Y-DNA test

Because a male's Y is inherited only from his father, who got it from his father etc. all the way back, the Y chromosome can track the direct paternal lineage. In theory, this corresponds with the surname, and can also be used to indicate deep (paternal) ancestry. This is a gender-specific test and females will need to identify a male substitute in the particular line – father, brother, male cousin, paternal uncle etc.

Two sorts of Y-DNA tests – 1. STR

The first test analyses the Y chromosome at pre-agreed markers (also called *loci*, meaning 'places') known as short-tandem repeats (STRs). The standard test used today covers 37 markers with the option to increase that to 67 markers and beyond. The more markers tested, the more refined the result. A Y-DNA result on its own is of little use to a genealogist – its value comes when comparing a result with other people's and specifically with men who share the same surname or a variant.

Each of these markers has a tendency to mutate at a slightly different but known rate. This means that the "genetic distance" between two tested males is an indication of how far back in time their

lines branched apart. It is not unusual for men who share their same-surname ancestor 6-10 generations back to have identical Y-DNA results at 37 markers, although the further back in time the shared ancestor lived the more likely that one or other of the tested lines of descendants will have acquired at least one mutational difference at some point since. However, the more people in related lines who test, the more accurate the determination will be of Time to Most Recent Common Ancestor (TMRCA).

It also means that men tested can be assigned to different *haplogroups*, and these can be arranged into a 'family tree'. Testing more markers refines the haplogroup designation from, say, R1b to a branch (*sub-clade*) such as R1b1a2a1a2c1 and distinguish it from another individual who is R1b1a2a1a1b4. The DNA result will not supply the name of the common ancestor and can only indicate when that ancestor lived within a span of a number of generations — but this detail can be determined by traditional documentary research. If almost all tested men within a particular haplogroup have the surname MacSporran, say, but some, whose sub-clade diverged closer to present time, are called MacDribble, that suggests the MacDribbles are a branch of the MacSporrans, or that one particular MacSporran chose to call himself and his family MacDribble from then on. This is exactly what happened with the Kirkpatricks and the Colquhouns.

Name	Markers tested	DYS393	DYS390	DYS19	DYS391	DYS385		DYS426	DYS388	DYS439	DYS389i	DYS392	DYS389ii	DYS458	DYS459		DYS455	DYS454	DYS447	DYS437	DYS448	DYS449	DYS464					HaploGroup
Andrew Durie of Durie	37	13	24	14	11	12	15	12	12	12	13	14	29	16	9	10	11	11	25	15	19	30	15	15	17	18	etc	R1b1a2a1a1b4
Sir David Durie	67	13	24	15	10	11	14	12	12	13	13	13	29	17	9	10	11	11	24	15	19	29	15	15	16	17	etc	R1b1a2
Dr. Bruce Durie	67	15	24	15	10	15	16	11	13	11	13	12	29	16	8	9	11	11	26	15	20	30	11	11	14	15	etc	I2b1a

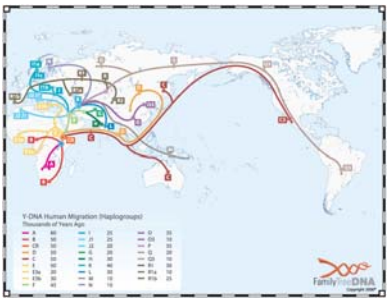
Example of STR results

Name	Markers	Paternal Ancestor	Country	Haplogroup	Short Hand
Andrew Durie of Durie	37	Gilbert de Durie ca 1261 John Durie of Durie ca. 1487	Scotland	R1b1a2a1a1b4	R-L21 R-S145+
Sir David Durie	67	?	Scotland	R1b1a2	R-M269
Dr. Bruce Durie	67	Gilbert de Durie ca 1261 John Durie of Durie ca. 1487	Scotland	I2b1a	I-M284
Robert Durie	67	?	Scotland	R1b1a2	R-M269

Haplogroup confirmed by SNP tests

What are these Haplogroups?

Over the years, geneticists have arranged the various haplogroups into a tree, starting from A (our original African ancestors), and branching into different populations, as human migrated, separated and no longer interbred. We can estimate the times when these divergences happed, which ties in with other evidence – archaeological, anthropological, historical, etc., as shown. This means that the ancestral migration of a particular group can be traced – so it is possible to say whose male line in Scotland is originally Pict, British, Gael, Viking, Anglo-Norman, French and so on. For instance, someone within some branch of the haplogroup R1b is probably of (male) Irish-Gaelic origin, while I2b1a indicates (male) ancestry of pre-Roman, indigenous origin in Scotland ('Pictish').



Human migrations and haplogroups (courtesy of FamilyTreeDNA)

Two sorts of Y-DNA tests – 2. SNP

A different kind of analysis, SNP (*single nucleotide polymorphism*) has allowed further refining of large haplogroups and sub-clades into even more fine-grained patterns. Sub-clades are typically known by their SNP identifier i.e. M284. The letter before a SNP number designates the laboratory where it was first discovered or described. For example, S145 (also known as L21) is the classic Scottish indigenous marker, and corresponds to the STR designation R1b. "Downstream" (more modern) mutations include M222, often called 'Naill of the Nine hostages', for a preseumed ancestor, which spread from North of Ireland to Scotland. By contrast, L165/S68 is Norse Viking, and typically found in the West and Northern Highlands.

Your DNA test results will typically give you both an STR and an SNP designation. The SMPs are winning out, as they are more straightforward – R1b1a2a1a1b4, while hierarchically meaningful (you can see how it descends from R1b1a2a1a1, and how that's different from R1b1a2a1a2) is considered to be a bit of a mouthful.

There is a new test from FamilyTreeDNA (see below) called 'The Big Y' which does a huge range of analyses on a single microchip, and is further helping to group people into smaller 'tribes' and closer together in time.

How meaningful is all this?

At a macro level. It can indicate ancestral links. For example, L165/S68 (Norse Viking) is shared by MacNeils of Barra, the Buies of Jura, some MacDonalds and the core MacLeod lineage. Other MacNeils are M222 (Naill of the Nine hostages). This has implications for the supposed single origin of all MacNeil/O'Neil surname groups.

Also, by comparing results with others tested, it is possible to identify links and TMRCA between individuals. This may links people or entire groups together in ways they hadn't realised. For instance, my results show a match with someone of a different surname, but from the same area as my family. The TMRCA is about 5 generations ago, which suggests that one of us has a GGGGrandparent with a guilty secret! But it could have been an adoption, a change of name to that of a step-parent, or some other innocent explanation.

BUT PLEASE REMEMBER… your Y chromosome is a tiny part of your genetic inheritance – about 1/1000th – and your overall genetic admixture might be very different. Imagine if a Spaniard was shipwrecked in the Orkneys 500 years ago, stayed and got married, and all the subsequent generations intermarried with other "native" Orcadians down to yourself – how much "Spanish" is actually left in you? Not much. But your Y-test will scream 'SPANISH!' at you.

It's useful, but it's not everything. Therefore, if you want to know anything genealogically useful, DO NOT test with a company that only provides a Y-based prediction of ethnic ancestry and has no database to match against. For example, I'm intrigued that my furthest-back forebears were probably 'Pict', and from the same part of Scotland where my family has lived for the intervening 3,000 years, but it doesn't help with my genealogical researches. However, being able to match with others in the database has linked me to other individuals – and other surname groups – in ways that I may not have thought to investigate.

Y-DNA: What is a match?

69720-MacDonald	12	25	15	10	11-13	12	12	12	12	13	28	17	9-9	11	11	25	15	20	30	14-15-16-17	11	11	19-23	16	15	18	16	37-39	11	12
46281-McDonald	12	25	15	10	12-13	12	12	12	12	13	28	17	9-9	11	11	25	15	20	30	14-15-16-17	11	11	19-23	16	15	17	16	37-39	11	12

These two haplotypes show a two marker mismatch on 37 markers. This is written as 35/37 match. It sounds close – 35/37 is about 85% after all. But it doesn't mean that. Because we know the rate at which these two mutations would have occurred, the probability that individual 69720 shared a common ancestor with 46281 in the last 8 generations is only about 57%, rising to 86% within 12 generations and 96% within 16 generations. But that may well tie up with the known history – "My ancestors left Scotland 300 years ago" implies a 12-generation gap.

What about females?

We all – male and female alike – inherit our mothers' mitochondria. These are the energy-generators in cells, and contain their own smaller amount of DNA, separate from the chromosomes in the cell nucleus. This means that mtDNA, as it's called, tracks from a male or female to his or her mother, the maternal grandmother, her mother etc.

This is not surname-linked in the same way, and is not as definitive as Y-DNA testing (the mutations are less frequent) but is a useful adjunct to the male-specific test. If possible, go for a Full Sequence test as the results are better, and there are more chances of finding a match.

Autosomal DNA testing.

The 22 chromosome pairs, called the *autosomes*, that are not XY (and not mitochondria) are the summary of your genetic history, all the way back to the very origin of life. Scientists use analyses of these to determine, for example, the degree of relatedness of bees and bananas (not much!), various species and varieties of animals and plants, and so on. In humans, this is a useful way of summarising all your ancestral lines – your 'ancestral barcode', if you like. It will produce two things:

1. matches to "close" cousins (as far back as 5 generations or about 150-250 years) in all lines
2. your deep ethnic admixture – 70% Northern European, 20% Chinese, 10% Martian, or whatever)

Which test?

It depends what you want to find out. Most people start with a 37 or 67 marker STR test, which gives a haplogroup prediction and also allows matching with others tested. That may be all the information required ('Good grief! I'm not French after all!') You can then choose to upgrade, using the same submitted sample, to specific SNP tests, autosomal testing, mtDNA etc.

Often a combination of Y-STRs, carefully-chosen SNPs and a FamilyFinder (autosomal) will be more than enough.

- Y-DNA (SNP) and mtDNA = 'deep' ancestry B.C to c.1400 (before adoption of surnames)
- Y-DNA (STR) and mtDNA (full sequence) = 'Genealogical ' or 'Documentary' time, c.1400 to present, which often ties in with other evidence and oral tradition
- Autosomal DNA (Family Finder & Relative Finder) = Close ancestry - 4th or 5th cousins, 150-250 years ago to present day – but also gives an indication of deeper ethnic ancestry in all lines

Which testing company?

There are a number of companies, but my personal recommendation would be FamilyTreeDNA (www.ftdna.com). They have the largest matching database, the best understanding of the needs of the genealogical community, links to the National Geographic 'Genographic' project, a long track record, a great reputation for security and privacy and excellent educational materials on their website.

You are also encouraged to join the Scottish DNA Group Project and any of the Haplogroup or Surname Projects at FTDNA.

Finally…Will DNA testing tell me if I'm descended from Robert the Bruce?

Don't worry – you probably are. Because we ALL are. At its simplest, the argument goes like this:

- Robert Bruce had at least 12 children
- Assume each one of them had an average of 2 children, in every generation
- 700 years = 28 generations, which means 12 x 2^28 descendants today
- That's about 3.2 billion descendants – take out India and China, and it's almost everyone else on the planet!

The same goes for descendants of Edward II, Charlemagne and all the rest. Most Europeans are descended form all of these to some degree, so it's an unremarkable claim. But can you prove it with documents?

Y-DNA Glossary

- **Allele value** = the count of the number of base pair tandem repeats
- **DNA signature** = haplotype that belongs to an individual DNA study participant
- **DYS marker** = a number given to each marker and simply indicates the locus number in sequence of discovery
- **Haplotype** = a string of numerical values derived from different markers
- **Haplogroup** = the group related haplotypes belong to
- **Marker** = position measured on Y-DNA structure
- **Modal** = most commonly found value among a range of values
- **TMRCA** = Time to Most Recent Common Ancestor
- **SNP** = mutation which has occurred only once at a specific position in a particular chromosome in a single individual
- **STR** = short tandem repeat markers

Links

1. FTDNA – www.ftdna.com
2. Scottish DNA Project: http://scottishdna.net
3. ISOGG – International Society of Genetic Genealogy http://www.isogg.org/
4. IGOGG wiki http://www.isogg.org/wiki/Wiki_Welcome_Page
5. Dean McGee's Y-Utility - http://www.mymcgee.com/tools/yutility111.html
6. Jim Cullen's World Haplogroup Predictor - http://members.bex.net/jtcullen515/haplotest.htm
7. Ysearch - http://www.ysearch.org/
8. Mitosearch - http://www.mitosearch.org/
9. Genebase - http://www.genebase.com/
10. JOGG – Journal of Genetic Genealogy http://www.jogg.info/
11. 2011 Human evolution, migration and history audio files https://royalsociety.org/2011-Human-evolution-DM-audio-/
12. Learn Genetics - http://learn.genetics.utah.edu/content/extras/molgen/index.html

This information was provided by:

Dr. Bruce Durie BSc (Hons) PhD OMLJ FSAScot FCollT FIGRS FHEA
Genealogist, Author, Broadcaster, Lecturer
Shenachie to the Chief of Durie
Shennachie to COSCA
Honorary Fellow, University of Strathclyde
Member, *Académie internationale de généalogie*
E: bruce@durie.scot
W: www.brucedurie.co.uk

- Test kits can be purchased at the discounted project prices and paid for in $USA or £Sterling, contact …. [SCOTCLANS SHOULD BE SELLING THESE AS AGENTS]