

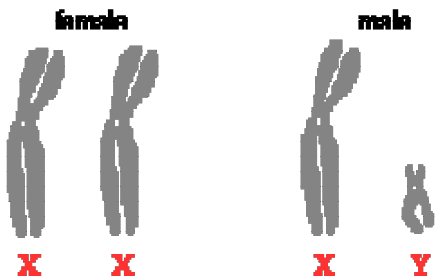
CHROMOSOME DNA TESTING

This is the second of three articles about DNA and its use in genealogical research. The first article dealt with the amazing world of DNA and what it does. The science was quite challenging but this second article, about Y-chromosome DNA testing will be easier to follow. As before, you can contact me at: ronald.east@btinternet.com if there are any points you would like to discuss.

First, a recap:

A unique set of 46 chromosomes were created at the moment when you were conceived; 23 were contained in your mother's egg and 23 were contained in the sperm. Amazingly, this set of 46 chromosomes was then perfectly copied into almost every one of the cells in your body, all 100 thousand billion of them!

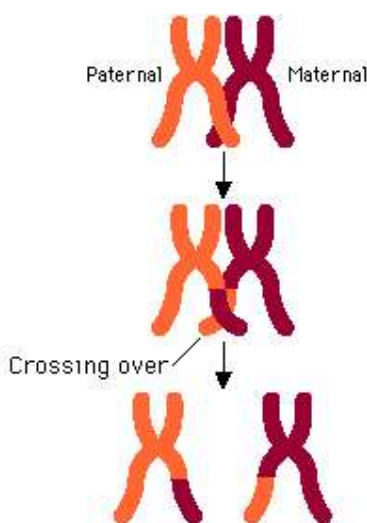
The chromosomes are paired: chromosome number one from the sperm is paired with



chromosome number one from the egg, two with two and so on but the 23rd pair is not always matched. In a male, one looks like a letter X and the other looks like a letter Y. They form an XY pair and it is the Y-chromosome that codes for 'maleness'. The Y-chromosome came from his father and so a male has an XY pair in each cell. In a female the 23rd pair is matched. They both look like a letter X and so every female has an XX pair in each cell.

To move on:

Matched pairs of chromosomes can swap bits of information before they become eggs or sperm



(this is called '**crossing over**'). In this way we can inherit features like shades of eye colour rather than the actual eye colour of mother or father. In a woman every pair of chromosomes can swap bits of information because every pair is matched but in a man the 23rd pair cannot swap information because the X-chromosome does not match the Y-chromosome. The Y-chromosome therefore passes from generation to generation unchanged (except by **mutation** which we will look at below).

The Y-chromosome is therefore passed unchanged from father to son, together with the paternal surname, but is never passed on to a daughter.

Amazingly, all of our chromosomes, including the Y-chromosome, contain just a few genes (about 4% of our DNA) and the remaining DNA has no known function – it is called **junk**

DNA. So, along with the genes for maleness that a boy inherits from his father on the Y-chromosome he also inherits all the junk DNA.

Within this junk DNA are certain locations where a short segment of DNA will *stutter*, repeating itself a number of times. This is known as a **STR** (or Short Tandem Repeat) and its location on the Y-chromosome is called its **locus** (plural **loci**). Geneticists refer to it by the letters DYS (**DNA Y**-chromosome Segment) and a number that identifies the locus. This is also called a **marker**.

For example, DYS 393 contains the bases AGAT repeated between 9 and 17 times in the Y-chromosomes of different men. It occurs in the junk DNA at 3191134 base pairs along the Y-chromosome.

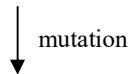
Therefore DYS 393 (9) means AGAT is repeated 9 times at locus 393:

DYS 393 (9) AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT

The number of times that the stuttering occurs at each marker is called an **allele value** and it stays more or less constant when passed from father to son over many generations. But very occasionally a change will occur in the number of stutters. This is called a mutation:

Father

DYS 393 (9): AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT



Son

DYS 393 (10): AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT AGAT

The son will then pass this mutation at DYS 393 to his son and so on, probably for many generations until another mutation may occur at DYS 393.

So the allele value at DYS 393 could differ between two men who were, in fact, related because of this mutation.

The East Family History Society therefore asks that 37 different loci (not just DYS 393) are examined when comparing the Y-DNA of two men.

If all 37 loci match for both men (i.e. each locus has the same allele value for the two men) then they are most certainly related in the paternal line.

If 36 out of 37 loci match (i.e. one locus differs in the allele value for the two men) then they are most probably related in the paternal line. There could have been a mutation in an earlier generation in the paternal line.

It is, of course, possible to be even more certain of a paternal match by examining more loci but a 37 marker test gives a reliable result when set against the extra cost involved.

The set of allele values for the loci investigated for a man is called his *haplotype*. Here is my haplotype:



FamilyTreeDNA

Certificate – Y-DNA

This Certificate confirms that you have had your DNA analyzed by Family Tree DNA. The outcome from each of the thirty-seven Loci examined is reported in the table below.

For your benefit we have listed the Locus designation for all thirty-seven Loci utilized by the geneticists supporting our company. If your alleles for the thirty-seven Loci match another person exactly, then you share the same Haplotype.

Family Tree DNA is a genealogical tool designed to aid individuals wanting to "connect" to other relatives lost in time and where the paper trail no longer exists.

Ronald D. East

Your sample # **29206**

DYS #	393	390	19/394	391	385a	385b	426	388	439	389-1	392	389-2	
Allele	14	26	15	10	11	14	12	12	10	14	11	31	
DYS #	458	459a	459b	455	454	447	437	448	449	464a	464b	464c	464d
Allele	14	9	10	11	11	24	14	20	32	12	14	15	15
DYS#	460	H4	YCA IIa	YCA IIb	456	607	576	570	CDY a	CDY b	442	438	
Allele	12	12	19	22	15	16	18	20	34	38	12	11	

March 17, 2005

Matthew E. Kaplan
Matthew E. Kaplan

You can see that 37 loci have been tested and for each one, the number of repeats (the allele value) is given.

Look at DYS 449. It can be found at 8278048 base pairs along the Y-chromosome in the junk DNA. The repeat structure for DYS 449 is TTTC which occurs between 26 and 36 times. My allele value for DYS 449 is 32 repeats of TTTC.

Therefore my son, Ben, and I would have a perfect match at this and all the other 37 loci unless a mutation has occurred (or unless the family got the milk delivered free that week!) but mutations happen only very rarely.

Assuming that a typical STR mutation rate is about once in every 370 generations (it varies between about 250 and 500 generations) then if 37 loci are tested, there is about a 1 in 10 chance that one of them might mutate every generation. Taking a generation to be 25 years then there is one chance that every 250 years that a mutation is observed. Not very probable, you will agree, which is why Y-STRs are so useful for genealogical research.

Note, however, that these mutations are also very useful, otherwise every male would have the same Y-chromosome.

Because it is known roughly how often these mutations occur, it is possible to estimate when the **MRCA** (Most Recent Common Ancestor) lived. So, if two men have a recent MRCA say, the same grandfather, then their haplotypes are probably identical (or almost identical since in only two generations it is unlikely that any mutations have occurred).

Note that the word ‘probable’ keeps occurring. Statistically it can be shown that if 37 out of 37 markers match then there is a 95% chance that the MRCA is no more than 7 generations ago. If 35 out of 37 match, there is a 95% chance that the MRCA is no more than 14 generations ago.

However, if the MRCA for two men is 40 generations back (about 1000 years), then their haplotypes will probably show several mismatches because of mutations.

The message is there is no certainty in genealogical research!

A Post Script

Sometimes it is possible to make a tentative suggestion about where, geographically, a male family line may have originated (this should really be done by looking at **SNPs** or Single Nucleotide Polymorphisms – but that is another story)

For example, there is a geographical population known as **haplogroup** R1a1 that spread from an area around the Black Sea at the end of the last Ice Age, about 15 000 years ago. This population has a most common allele value of 12 at DYS 426 and a value of 11 at DYS 392. These allele values happen to match those in my haplotype which suggests that my male ancestors may have come from this area.

Haplogroup R1a1 is found in Hungary (60% of the male population), Poland (56%) and Russia where one out of two men has this haplogroup. High frequencies are also found in Norway, Sweden and Iceland (23%) and it is believed to have been spread across Europe by later migrations of Vikings, which accounts for its existence in the British Isles.

Glossary

allele the number of times that a sequence of bases repeats itself in a marker

haplogroup the geographical population from which a family is descended

haplotype the collection of allele values for a set of markers or loci (usually 37)

locus this identifies where on the Y-DNA the repeat sequence of bases is found but it is also used interchangeably with the term **marker**

<i>marker</i>	a sequence of bases that repeats itself in junk DNA and is used to establish matches in genealogical research
<i>MRCA</i>	Most Recent Common Ancestor is the most recent individual from whom a group is directly descended
<i>mutation</i>	a change in the number of repeats in a STR (or a change in the sequence of bases in a SNP)
<i>SNP</i>	a Single Nucleotide Polymorphism – a mutation in which one of the bases changes in the repeat sequence so that, for example, the repeat sequence TTTC becomes TTAC; this happens much less often than a STR mutation
<i>STR</i>	a Single Tandem Repeat – a sequence of bases, such as TTTC, that repeats itself several times in the junk DNA

TIP from (George) Clarie EAST

I am a member of the East FHS - George C East - and if you remember my life story was published in recent issues of 'Points East' I am also a member of a few other societies and recently I was informed about a web site which shows members of HM Forces in Australia in both World Wars.

I have just glanced at it again there are 158 EASTS recorded and a few were born in the UK. I do not know if you are aware of this but I think it is worth a mention in the next magazine.

The web site is www.ww2roll.gov.au.