Y-DNA FACT SHEET

By

Bruce A. Crawford

For those not familiar with DNA analysis and particularly Y-DNA, the following explanation may help. DNA is the basic building block of cell information and heredity. it is created as a code in a series of protein chains arginine, cytosine, thymine, and guanine. These are coded as A-C-T-G but can occur in any order. Various chains of these proteins mean a specific thing to the cell. As you may remember from High School biology the sex chromosomes are XX for females and XY for males where one of the chromosomes is contributed by the female and one by the male at conception. Since the female only has X chromosomes she cannot contribute a Y which can only be transferred from male to male offspring down through the generations. In western cultures the surname follows the male line. Thus by looking at the Y chromosome genealogical connections can be deduced.

Y-DNA and Mitochondrial DNA

There are two common types of DNA testing used today in genealogy: (1) Y chromosome DNA and (2) mitochondrial DNA. Concisely, Y chromosome genealogy establishes the genetic pattern in the Y chromosome (males only) to establish an inheritance through paternal descendancy. Mitochondrial DNA establishes a genetic pattern in the mitochondria (genetic material surrounding the chromosomes) to establish a maternal descendancy.

Y-chromosome genealogy is for males only, while mitochondrial analysis can be used on both males and females but establishes links only through the mother. The lineage passing from mother to daughter and son.

Another less common type of DNA genealogy utilizes the X chromosome which is the female analogy to the Y chromosome but is not often used because the mitochondrial genealogy is just as accurate and with broader scope.

Autosomal Analysis

Autosomal analysis refers to analysis of the remaining 22 human chromosomes not related to sex determination. Autosomal testing is used most often in legal situations where paternity is in question or in forensics. It is very accurate in determining relationships between parent and children and all relationships up to second cousins. Beyond second cousin's additional documentation or testing is needed.

Family Tree DNA uses autosomal testing in their Family Finder test. Following is an excerpt from their site:

"The Family Finder test is designed to trace all of your ancestral lines (five generations and beyond) using your autosomal DNA. It will confidently identify relationships for five generations. Family Finder tests thousands of data points on your 22 autosomal chromosomes. Your results are then compared to others in the Family Finder database. The software detects linked blocks of DNA that indicate a common ancestor. The number and size of these segments is used to determine how recently any two people are related.

The strengths of the Family Finder test are that it may be used with equal success by both men and women and it has the ability to find connections on any of your family lines. The challenge is determining which branch of your family tree you share with your cousin. This can be accomplished with traditional genealogical records and by utilizing other types of DNA tests.

Y Chromosome Short Tandem Repeats

The most recent evaluations of Y-chromosomes center around short tandem repeats (STRs) which are DNA segments that are repeated similar to a stutter in speech. What this means is that the coding on the Y chromosome at the STR markers will be passed down from generation to generation exactly the same unless a mutation occurs. A mutation is an error in replicating a portion of the chromosome. It is very rare. In the case of Thorlongus, our hypothetically earliest known ancestor, all of his male offspring from 1100 AD to present will have received the exact same Y chromosome STR marker repeats except when mutations have occurred during the last 915 years. Therefore, it can be seen that it is important to know the rate at which mutations occur if it is to be useful in examining the relationships of surnames and lineages to each other. This relationship has been called Time Till Most Recent Common Ancestor (TMRCA) and is used to estimate how long ago two people with similar Y-DNA may have shared a common ancestor. It is based on how many differences are expressed in their gene markers. Research conducted by a variety of institutions and individuals and used by Family Tree has set the general mutation rate at 1/500 generations per marker where a generation is 20 years (Walsh, 2001)(Zhivotovsky, et al., 2004)(Kerchner, 2008). This equates to a mutation rate of 0.002 or 2×10^{3}.

Since we do not have DNA sampled from ancestors going back to 1100, scientists have estimated the mutation rate by a different method. Weber and Wong (1993) used 20,000 contemporary parent-offspring samples at 28 STR loci on chromosome #19 to check for mutations. They found 47 mutations in the 20,000 events per STR loci per generation (1/833 generations). This assumes that the mutation rate has remained constant through the centuries. This research and other related studies are the basis for determining mutation rates.

Most individuals are interested in determining the TMRCA. As mentioned above and according to Family Tree DNA, based on (Walsh, 2001), a mutation is likely to occur at an STR marker about every 500 generations. The estimate of TMRCA is based on a probability distribution. As one uses more and more markers, the distribution becomes tighter. What this means is that each marker is a probability test in that at that marker there is one chance in 500 that there will be a mutation in a specific generation. However, the more markers tested the greater the likelihood that a mutation will occur somewhere. So if there are 37 markers being tested there are 37 chances in 500 that a mutation will occur. This equates to about one chance in 13.5 generations or 13.5 X 20-25 years equals about 300 years. Since surnames were developed in the 1100s only about 2-3 mutations should be expected within surname lineages since that date. The Family Tree DNA website provides the following general guidance for mutation

distances at 37 loci. The guidance web address from Family Tree DNA is (http://www.familytreedna.com/genetic-distance-markers.aspx?testtype=37) .

General Guidance for Mutation Differences for 37 loci tested.

Tightly Related: If you have one mutation distance (36/37) and you share the same surname (or a variant) with another male and you mismatch by only one 'point' at only one marker the two are tightly matched. Very few people achieve this close level of a match. Your mismatch is within the range of most well established surname lineages in Western Europe or about 1100 AD.

Related: You share the same surname (or a variant) with another male and you mismatch by only two or three points a 35/37 match or 34/37 match. Your mismatch is likely within the range of most well established surname lineages in Western Europe or about 1100 AD.

Probably Related: You share the same surname (or a variant) with another male and you mismatch by four 'points' --a 33/37 match. Because of the volatility within some of the markers this is about the same as being 11/12 and its most likely that you matched 23/25 or 24/25 on previous Y-DNA tests. The only way to confirm is to test additional family lines and find where the mutations took place.

Not Related: Persons with marker distances of more than 4 are probably not related and those with more than 6 have no common ancestor for thousands of years.

General Guidance for Mutation Differences for 67 loci tested.

Very Tightly Related A 67/67 match between two men who share a common surname (or variant) means they share a common male ancestor within the genealogical time frame. Their relatedness is extremely close. All confidence levels are well within the time frame that surnames were adopted in Western Europe with the common ancestor predicted, 50% of the time, in three generations or less and with a 90% probability within five generations. Very few people achieve this close level of a match.

Tightly Related A 65/67 or 66/67 match between two men who share the same surname (or a variant) indicates a close relationship. It is most likely that they matched 36/37 or 37/37 on a previous Y-DNA test. Very few people achieve this close level of a match. All confidence levels are well within the time frame that surnames were adopted in Western Europe about 1100 AD.

Related. A 63/67 or 64/67 match between two men who share the same surname (or a variant) means that they are likely to share a common ancestor within the genealogical time frame (about 1100 AD). The common ancestor is probably not extremely recent but is likely within the range of most well-established surname lineages in Western Europe. It is most likely that they matched 24/25, 36/37, or 37/37 on previous Y-DNA tests, and mismatches are within DYS458, DYS459, DYS449, DYS464, DYS576, DYS570, and CDY.

Related. A 61/67 or 62/67 match between two men who share the same surname (or a variant) means that they may share a common ancestor within the genealogical time frame. The common ancestor is

probably not recent, but may still be within the range of most well-established surname lineages in Western Europe about 1100 AD. It is most likely that they matched 24/25, 36/37 or 37/37 on previous Y-DNA tests. Mismatches are within DYS458, DYS459, DYS449, DYS464, DYS576, DYS570, and CDY.

Probably Related. A 60/67 match between two men who share the same surname (or a variant) means that they may share a common ancestor within the genealogical time frame about 1100 AD. Because of the volatility within some of the markers, this is about the same as being 11/12, and it is most likely that they matched 23/25 or 24/25 or 33-34/37 on previous Y-DNA tests. If they test additional individuals, they will most likely find that their DNA falls in between the persons who are seven apart demonstrating relatedness within this family cluster or haplotype. If several or many generations have passed, it is likely that these two lines are related through distant family lines. The only way to confirm the relationship is to test additional family lines and to find where the mutations took place. By testing additional family members, you can find the person in between them. This 'in betweener' is essential, and without him, the possibility of a match exists but cannot be confirmed.

Only Possibly Related. A 58/67 or 59/67 match between two men who share the same surname (or a variant) means it is possible but unlikely that they share a common ancestor within the genealogical time frame (1100 AD to present). If you test additional individuals, you may find the person whose DNA results falls in between the persons that are eight or nine apart demonstrating relatedness within this family cluster or haplotype. It is most likely that they did not match 24-25/25 or 35-37/37 in previous Y-DNA tests. If several or many generations have passed, it is possible that they are related through other family members. The only way to confirm or deny the relationship is to test additional family lines and find where the mutation took place. By testing additional family members, you may find the person in between. This 'in betweener' is essential. Without him, only the distant possibility of a match exists.

Not Related. A 56/67 or 57/67 match between two people means they are not related within the genealogical time frame (1100 AD to present). The odds greatly favor that the two men have not shared a common male ancestor within thousands of years.

Not Related. Greater than 11 mutations different The two men are totally unrelated within the genealogical time frame on their direct paternal line. Their shared ancestry is deeply anthropological and dates to the common African heritage of the human race.

General Guidance for Mutation Differences for 111 loci tested.

Very Tightly Related 111/111 match indicates a very close or immediate relationship. Most exact matches are 3rd cousins or closer, and over half are related within two generations (1st cousins).

Tightly Related 110/111 match indicates a close relationship. Most one-off matches are 5th or more recent cousins, and over half are 2nd cousins or closer.

Tightly Related 109/111 match indicates a close relationship. Most matches are 7th cousins or closer, and over half are 4th or more recent cousins.

Related 108/111 match indicates a genealogical relationship (1100 AD to present). Most matches at this level are related as 9th cousins or closer, and over half will be 5th or more recent cousins. This is well within the range of traditional genealogy.

Related 107/111 match indicates a genealogical relationship (1100 AD to present). Most matches at this level are related as 10th or more recent cousins, and over half will be 6th or more recent cousins. This is well within the range of traditional genealogy.

Related 106/111 match indicates a genealogical relationship (1100 AD to present). Most matches at this level are related as 12th cousins or more recently, and over half will be 7th cousins or closer. This is well within the range of traditional genealogy.

Related 105/111 match indicates a more distant genealogical relationship. Over half of matches will be 9th cousins or closer, and most matches at this level are related as or more recently than 14th cousins. If there is a tradition of a recent genealogical relationship, the best way to confirm it is to test additional family lines. By testing additional family lines, you can find the person in between who is a closer match to each of the others tested. This 'in betweener' is essential for you to find as their match proves the connection between the more distant matches.

Probably Related 104/111 match indicates a more distant genealogical relationship. Over half of matches at this level are related as 10th cousins or closer. Most matches at this level are related as 16th cousins or more recently. If there is a tradition of a recent genealogical relationship, the best way to confirm it is to test additional family lines. By testing additional family lines, you can find the person in between who is a closer match to each of the others tested. This 'in betweener' is essential for you to find as their match proves the connection between the more distant matches.

Only Possibly Related 103/111 match indicates a distant cousinship with only a chance of a genealogical relationship. Over half of matches at this level are related as 12th cousins or more recently. Most matches at this level are related as 18th cousins or more recently. The connections here can be highly informative for relationships with historic groups and events. If there is a tradition of a recent genealogical relationship, the best way to confirm it is to test additional family lines. By testing additional family lines, you can find the person in between who is a closer match to each of the others tested. This 'in betweener' is essential for you to find as their match proves the connection between the more distant matches.

Only Possibly Related 102/111 match indicates a distant cousinship with a chance of a genealogical relationship (1100 AD to present). Over half of matches at this level are related as 13th cousins or closer. Most matches at this level are related as or more recently than 20th cousins. The connections here can be highly informative for relationships with historic groups and events. If there is a tradition of a recent genealogical relationship, the best way to confirm it is to test additional family lines. By testing additional family lines, you can find the person in between who is a closer match to each of the others tested. This 'in betweener' is essential for you to find as their match proves the connection between the more distant matches.

Only Possibly Related 101/111 match indicates a distant cousinship with some chance of a distant genealogical relationship. Over half of matches at this level are related as 15th cousins or closer. Most matches at this level are related as 22nd cousins or more recently. The connections here can be highly

informative for relationships with historic groups and events. If there is a tradition of a recent genealogical relationship, the best way to confirm it is to test additional family lines. By testing additional family lines, you can find the person in between who is a closer match to each of the others tested. This 'in betweener' is essential for you to find as their match proves the connection between the more distant matches.

Not Related Ten or more mutations. The two men are totally unrelated within the genealogical time-frame on their direct paternal line. Their shared ancestry is historical or anthropological.----

Table 1 below is a simple example of six Y chromosome STR markers. Marker 393 is a specific segment of the Y chromosome with 13 repeats of the same code while marker 390 has 24 repeats. As can be seen, markers 391 and 385b are different between the two individuals. There have been three mutations, one at marker 391 and two at marker 385b. If we know the mutation rate, we can then calculate the TMRCA. Since each marker is a probability equation, a computer program is used to calculate the time. The more markers tested, the more accurate the information.

MARKER	393	390	19	391	385a	385b
Individual A	13	24	14	11	11	14
Individual B	13	24	14	10	11	16

How To Join The Crawford Surname Project

To join the Crawford Surname Project go to https://www.familytreedna.com/group-join.aspx?Group=Crawford Once at the site you can choose 12, 25, 37, 67, or 111 markers. We highly recommend that you test at least 37 markers to optimize your chances of being placed in a lineage.

All persons in the project previously tested for 12 or 25 loci should seriously consider upgrading to 37, 67, or 111 as your current information is marginally useful.

If you have already had your DNA tested by Ancestry.com, Genealogy.com or other laboratories, you can still join the Crawford surname project by transferring your information to Family Tree. This can be done at the same web page listed above at the bottom of the page under "Third Party" for about $19.00.

Y Chromosome Single Nucleotide Polymorphisms (SNPs)

A SNP is a genetic marker at a particular position within a DNA sequence, consisting of a single nucleotide variation in DNA code. It is the most common type of stable genetic variation. SNPs can result from a base transition (A for G, T for C), a transversion (G or A for T or C) or single base deletion. SNPs are the smallest possible change in DNA. Most changes do not change an organism's appearance, though some do. On an individual level a mutation may cause an

inherited disease. Geneticists have now continued to map the occurrence of SNPs on the Y chromosome and have the ability to detect unique SNPs in Crawford individuals. The occurrence of SNPs have enabled geneticists to divide the entire male population into haplogroups lettered A through T and determine their origin and migration patterns across the world. They have now been able to evaluate sub-units known as clades and sub-clades. This breakdown of the world population coupled with STRs has enabled geneticists to begin to link near term genealogy with ancient major haplogroups and clades.

Last updated: August 7, 2016