Why It's Important to Take a Big Y-700 Test!
By Dr. Daniel C. Hyde
Revised Draft June 7, 2024

I recently looked at the public Family Tree DNA's Hyde DNA Project results at https://www.familytreedna.com/public/Hyde?iframe=yresults. To my dismay, only 41 of the 93 Hyde DNA Project participants have done further Y-DNA testing. Their Y-Haplogroups are in green meaning **"determined."** The other 52 participants' Y-Haplogroups are in red meaning **"estimated"** based on their STR markers tested. I feel strongly that **all** the Hyde DNA Project participants with their Y-Haplogroup listed in red (estimated) should consider upgrading with a Big Y-700 test. This note is to convince people to upgrade to Big Y-700!

Before I go into the reasons why people should upgrade, I want to establish a firm foundation of some of the terms and issues used in genetic genealogy.

What's an STR?

From high school biology you may remember that humans have 23 pairs of Chromosomes made of Nuclear DNA consisting of over 3 billion base pairs or nucleotides. DNA is a long sequence of four nucleotides: Adenine (A), Cytosine (C), Guanine(G), and Thymine(T). Biologists use the letters A-C-G-T to represent portions of this sequence. A man's Y-Chromosome has about 56 million nucleotides or about 2% of his total Nuclear DNA sequence.

Males have a Y-Chromosome and an X-Chromosome while females have two X-Chromosomes. A father passes on his Y-Chromosome to all his sons almost unaltered. A Y-DNA test examines markers on a man's Y-Chromosome so they can be compared to the values of other men's markers. The result for one man on its own reveals little. The key is matches with others.

Genetic scientists selected special locations for the markers called **Short Tandem Repeats** (**STRs**). These **STR**s occur in non-coding regions on a man's Y-Chromosome where a pattern such as "AGAT" is repeated. For example, AGATAGATAGAT has three repeats. On rare occasions, an **STR** on a man's Y-Chromosome has a copy error when he passes his Y-Chromosome to a son. The copy error results in a different number of repeats. **STR** copy errors are useful for genealogy since they mutate relatively fast, e.g., a particular **STR** might mutate in 150 years. Each **STR** is assigned a unique **DYS** number. DYS stands for **D**NA, **Y**-Chromosome, and an unique **S**egment. **DYS 393** is the first marker in a 37, 67, or 111 marker result of a Y-DNA test (called an **STR** signature). Marker **DYS 393** is known to repeat "AGAT" from 9 to 17 times. My distant cousins Jerry Hyde, Herb Hyde, and I all have 13. You can think of the cause of **STR** values changing as the DNA copying mechanism stuttered.

The company **Family Tree DNA** (FTDNA) currently has two versions of their Y-DNA test. One with 37 markers ($119) and a second with 111 makers ($259) that provides more precision.

What's an SNP?

The genetic scientists have determined a second type of **mutation** that may occur as a copy error when a man passes his Y-Chromosome on to a son. However, this copy error is very different from the **STR** copy error we just mentioned. This copy error is a single base pair change, e.g., an "A" to a "G," in the long sequence of about 56 million base pairs on a man's Y-Chromosome. Genetic scientists call this mutation a **SNP**.

SNP stands for **S**ingle **N**ucleotide **P**olymorphism, pronounced "snip" – a single base pair change that occurs within a DNA sequence, e.g., an "A" to a "G." A **SNP** at a particular location occurs very infrequently, over many thousands of years or more, and are used to determine Haplogroups. A **SNP** is where a single nucleotide is different from a standard reference DNA and is called a **Variant**. If many people have the same **Variant**, it is given a name such as **R1** or **R-M269** or **R-M512** and is called a **Named Variant**. If after a Big Y-700 test you have variants which are none of the known **Named Variants**, they are called **Private Variants**.

Note that **SNP**s happen on all 23 pairs of a human's chromosomes. This note discusses only **SNP**s that occur on a man's Y-Chromosome.

What's a Haplogroup?

Standard Y-DNA testing, e.g., Y-37, Y-67, and Y-111, may help to distinguish different male lines back for a few hundred years but what about ancient ancestral lines back thousands of years? Genetic scientists have devised different DNA techniques to explore ancient ancestral lines. These techniques are aimed at placing individuals in **Haplogroups**. There are male Haplogroups (Y-Haplogroups) and female Haplogroups. A Y-Haplogroup is a group of males with a common (possibly ancient) direct-male ancestor who had a particular **SNP** mutation. By direct-male lines we mean along the male line, e.g., father to son, to his son, to his son, etc. This is because a father passes his Y-Chromosome to his sons and *never* to his daughters. Females don't have a Y-Chromosome. Since this note only deals with Y-Haplogroups, we will drop the "Y-." Think of the structure of a Haplogroup as a large inverted tree with the ancestor at the root, all his sons below him, and all their sons, below them, etc.

Now with the foundational definitions out of the way, I want to discuss several important issues surrounding Y-DNA testing.

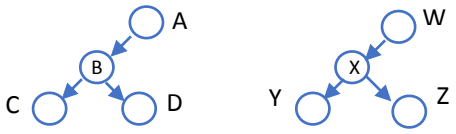
Why we can't sort out surname lines with Y-test markers (STRs) alone.

I believed for a long time that one could sort out surname lines by using only the **STR**s (**S**hort **T**andem **R**epeats) markers of Y-67 and Y-111 DNA test results. However, when I tried to use this approach on my own Hyde ancestors, I encountered strange inconsistencies. I was very confused! However, I'm not alone! Many have discovered that using only the **STR**s to sort out relationships is *problematic*! Why? The issues of **reverse** and **parallel mutations** of the **STR**s cause significant problems and alter incorrectly the genetic distance in a match! Let's look at both of them in detail.

Reverse Mutations of STRs

First, a particular **STR** may be a value, say 13, it mutates to 14 and then it mutates back to 13. This may falsely indicate that two individuals who were closely matched had no change in that **STR** when really there were two mutations. This is called a **reverse mutation**. Unfortunately, **STRs** revert back to a previous value fairly often! On the other hand, **SNP**s used to determine Haplogroups are more stable and **don't seem to ever revert back to a previous value!** Note a *particular* location of the 56 million locations on a Y-Chromosome has a **SNP** event very rarely, once over many thousands of years. Whereas a *particular* **STR** marker may mutate much quicker, say 150 years.

Parallel Mutations of STRs



Second, the *same* **STR** marker may change to the *same* value in one person's ancestral line and independently in another person's ancestral line. For example, let's assume DYS393 changed from 12 to 13, when one man (A) passed on his Y-Chromosome to his son (B) and when a different man (W) passed on his Y-Chromosome to his son (X) there was a mutation that resulted in the same value of DYS393 being 13. This is called a **parallel mutation**. This may cause a researcher to group all the descendants of individuals (C, D, Y and Z) with the change into one ancestral group, e.g., assume there is a common ancestor, where in reality the changes were parallel independent events and should not be grouped. Unfortunately, this happens with **STRs** fairly often as well.

The Ugly Truth!

Unfortunately, a researcher can't identify if either of the two **STR** issues—reverse and parallel mutations—have occurred **without further independent information**. Paper trails may provide the necessary independent information. Or **SNP**s, say from Big Y-700 tests, may be used to sort out the relationships. The expert genetic scientists use the **SNP**s to place individuals into their different Y-Haplogroups. A genealogy researcher can use the Haplogroups to identify possible occurrences of the above two **STR** issues. Since the **SNP**s are more stable, if an inconsistency in sorting out different surname lines is discovered, the **SNP**s can be used to confirm one or both of the two **STR** issues had occurred.

Testing at the Y-111 level may help to sort out surname lines but it is NOT enough to identify and resolve the two **STR** issues of reverse and parallel mutations. Doing a Big Y-700 test provides the same Y-111 results (as well as at least 590 additional **STR**s) AND the **SNPs** needed to place a man in his terminal Haplogroup. Then the **SNP**s can be used to identify and correct issues due to reverse and parallel mutations of the **STR**s.

Stating it again! Even Y-111 is not enough to sort out surname lines! We must have further independent information such as Big Y-700 test results! Therefore, whenever we have the opportunity, we should encourage our Y-37, Y-67, and Y-111 matches to take a Big Y-700 test.

What's a haplotree?

Your Y-Haplogroups are determined by your **SNP**s on your Y-Chromosome. A Haplogroup *name* is associated with each **SNP** event. These **SNP** events happen at very different times, some very ancient, e.g., thousands of years ago, and some only a few years ago. Whenever a man passes his Y-Chromosome onto one of his son with a **SNP** mutation, this creates a new Haplogroup, i.e., a new branch. The son is in all of his father's Haplogroups plus the new one. All the man's other sons without the new **SNP** will be in the same Haplogroups as their father. All the Haplogroup *names* form a tree structure called a haplotree. Note there are no men in the haplotree but the names of Haplogroups, e.g., R1, R-M512, R-Z159, and R-A6093.

How to construct a haplotree? Start at the root (geneticists call the root the Haplogroup "Genetic Adam") and add branches for the **SNP** events in the correct order in time. The hard part is determining the correct order in time when each **SNP** copy error occurred. Many genetic scientists around the world are working on constructing the Y-DNA Haplotree for all of humanity. We will leave the task of constructing this tree-structured *history* of **SNP** copy errors to the genetic scientists. This tree-structured history of **SNP**s is called the **Y-DNA Haplotree**, and also called the **Big Y Tree**, or the **Full Y Tree**. This Big Y Tree is growing almost daily as more men perform Big Y-700 tests.

It has been estimated that the man associated with "Genetic Adam" lived about 150,000 years ago. This man was not the earliest man to live. He is the most recent common ancestor of all the men on the planet who have been Y-DNA tested.

On December 6, 2021, Family Tree DNA announced "Y-DNA Haplotree Reaches 50,000 Branches, a Milestone for Genealogy." Ref: https://blog.familytreedna.com/y-dna-haplotree-reaches-50000-branches/. The article is worth reading!

Why are FTDNA's estimated Y-Haplogroups almost useless?

When a person tests at Family Tree DNA (FTDNA) with a Y-37, Y-67, or Y-111 DNA test, FTDNA estimates the person's Y-Haplogroup by using the person's **STR**s. These estimated Y-Haplogroup values, such as **R-M269** and **R-M512,** are rather useless. It's like saying a guy's address is "USA." It's correct but useless in mailing him a letter or finding his home to visit him. One needs a lot more specificity like a street address, a city, and a state. The same with Y-Haplogroups. To sort out surname lines, one needs a lot more specificity in their Y-Haplogroup. This specificity is provided by their Big Y-700 test results.

When most of you took your Y-test, you probably found the Y Haplogroup part of the results had little or no utility. But today, a *refined* Y-Haplogroup plays a significant role in genetic genealogy. This is the result of two major developments in genetic genealogy.

The first major development was the realization of the disturbing but a profound fact that one can NOT sort out surname lines with Y-test markers (STRs) **alone**! This was discussed in detail above. Even if every man took a Y-test with 111 markers, one can NOT sort out surname lines. To sort out surname lines, one needs further information beyond STR markers such as good paper trails or refined Y-Haplogroups as provided by the Big Y-700 test. This shook a lot of genealogy researchers up but, unfortunately, many have not learned this fact! Some genealogists abandoned Y-testing and moved their focus to Autosomal testing. But luckily, Big Y-700 provides a good solution! Now we need to educate people about this game changer!

The second major development was the significant advancements in DNA testing technology in the last few years. With Big Y-700 testing, the refinement of Y-Haplogroups is no longer confined to ancient time thousands of years ago but into genealogical time (last 600 years). Genealogists have been able to identify specific individuals by name where a SNP event occurred between father and son. That means the son was the first to possess that specific Y-Haplogroup! Remember Y-Haplogroups are based on **SNPs** on a man's Y-Chromosome. Genetic genealogists have determined that on *average* a **SNP** occurs about once every 83 years in a male line. That means one of the 56 million base pairs on the Y-Chromosome changes on average about every 83 years. But we don't know which one! This has the potential to provide the ability to sort out surname lines with great precision! These are exciting times!

An Example of Success!

To show you what can be accomplished using Big Y-700 results, here is an example. Several years ago, Joe Flood spearheaded an effort to recruit as many members of the R-A6093 Y-Haplogroup as we could find and get them to take Big Y-700 DNA tests. I'm one of the R-A6093 members and a co-administrator of the FTDNA Group **R-A6093 North Mercia** with Joe. The Project's Website: https://www.familytreedna.com/groups/r-a6093-north-mercia/about

The 26 members are all positive for the **R-A6093/FGC17294** Y-Haplogroup, which dates to about 100 BC, and most of us have a common ancestor in the 1300s, possibly in the period of recovery after the Black Death. Members are part of a Y-cluster where their Y-67 matches many of the others. The most common surnames so far are BLOOD, HYDE/HITE, one branch of REDDICK, WILLETT, CHEATHAM, probably COKER, and a number of singletons.

The North Mercia refers to an area in England surrounding Nottingham where many of our members's ancestors seemed to have lived. BLOOD is from Derby/Nottinghamshire, CHEATHAM and ASHMORE from Derbyshire, HYDE is an outer suburb of Manchester, etc.

We have encouraged 15 members to take a Big Y-700 DNA test and they have agreed to share their data with the project. This effort provided enough data to determine the following table.

Main splits in the R-A6093 Y-Haplogroup

Number of SNPs (N)	Split	Approx date*
22	Formation, Polish split (FT178522/A6093)	130AD
14	Jackson line (FGC17309)	800AD
11	Ashmore-Coker line (FGC17310)	1000AD
9	Ashmore and Coker split	1200AD
8	Hyde/Blood split (FT85184, FT200000, FT198830)	1300AD
6	Blood Nottingham/Derby (FT124365, FT85084)	1450AD
7	Highfield/Chittam-Elliott split (FT199987)	1370AD
6	Derby/Nottingham Hydes (FT198853)	1450AD

Note * Approximate time from present calculated as N*83+50 where N is the number of variants or SNPs.

The members of the **R-A6093 North Mercia Project** continue our investigation and are hopeful of further results.

The Y-DNA testing game has changed significantly!

All participants in the Hyde DNA Project and other surname projects must realize the Y-DNA testing game has changed significantly. The above mentioned two major developments —first, the serious issues of reverse and parallel mutations in STRs, and second, the recent advances in Y-DNA testing— mean you should upgrade with a Big Y-700 test to be where the action is!

Even though a close relative, e.g., a brother, father, or first cousin has taken the Big Y-700 DNA test, you should still take one. With only a few individuals in a terminal Haplogroup, FTDNA can't sort out the Private Variants. If you tested and other appropriate male relatives, such as second cousins, tested, that would provide information that might cause a branch and create a new Haplogroup that splits off your line. Remember a mutation could have happened between yourself and your father Or between your father and your grandfather, etc. These mutations happen at the birth of a son on average once every 83 years. That's about two or three generations. This is really useful for genealogy in sorting out surname lines but only IF many related men take the Big Y-700 test.

And another reason you and other appropriate male relatives should take the Big Y-700 test is the wealth of information in the results you receive. These include over 700 STRs AND a list of your Private Variants. Genealogy researchers are using the Private Variants and the 700 STRs to sort the different lines within a common terminal Haplogroup (one determined by Big Y-700).

I hope this convinces you to take the Big Y-700 test. FTDNA is currently running a sale. To upgrade from Y-67 to Big Y-700, the sale price is $259; and to upgrade from Y-111 to Big Y-700 the sale price is $229. FTDNA's 2024 Father's Day Sale ends June 17, 2024, at 11:59 pm PST.

Dr. Daniel C. Hyde, Co-Administrator of the FTDNA Hyde DNA Project