Y Chromosome and the SNPs STRs

Presented by

Marty Brady, May 16, 2020

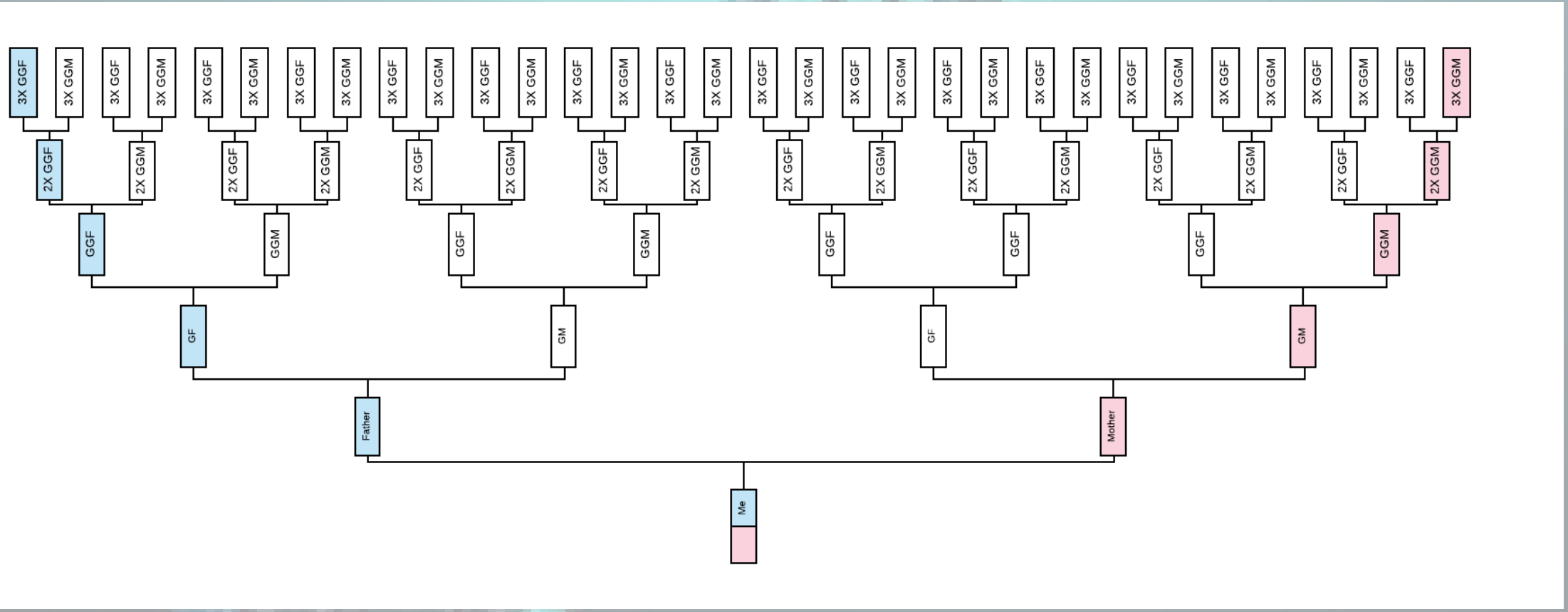
Introduction

- We will talk about:
  - Y chromosome: Its size, its purpose, its genealogical significance
  - Short Tandem Repeats (STRs): what they are, how they come to be, their significance and how they are analyzed
  - Single Nucleotide Polymorphisms (SNPs): what they are, how they come to be, how they are analyzed, and how we use the information

Y Chromosome

- The Y chromosome (Y chr) is about 57 million base pairs (bp) in length, making it the 2nd smallest chromosome. The SRY gene determines if an embryo becomes male.
- Because the studied region (MSY or NRY) does not undergo recombination, it is passed down unchanged from father to son making it suitable for deep ancestry research.
- Y-DNA has been used discover potential distant matches.
- It is used in Surname research. Surname Project Administrators use the information to group members if their projects into subgroups.
- It is also often used to find biological fathers of males. Dr. Blaine Bettinger estimates roughly 30% of males who test their Y-DNA through the Adopted DNA Project at FTDNA are able to identify their likely biological surname through Y-DNA alone.
- Y-DNA has been used to investigate the purported skeletal remains of King Richard III (www.nature.com/ncomms/2014/141202/ncomms6631/full/ncomms6631.html).
- A limitation of Y-DNA is that it can tell if two men are paternally related, but it doesn't tell you how they are related (i.e., father, brother, uncle, etc.)

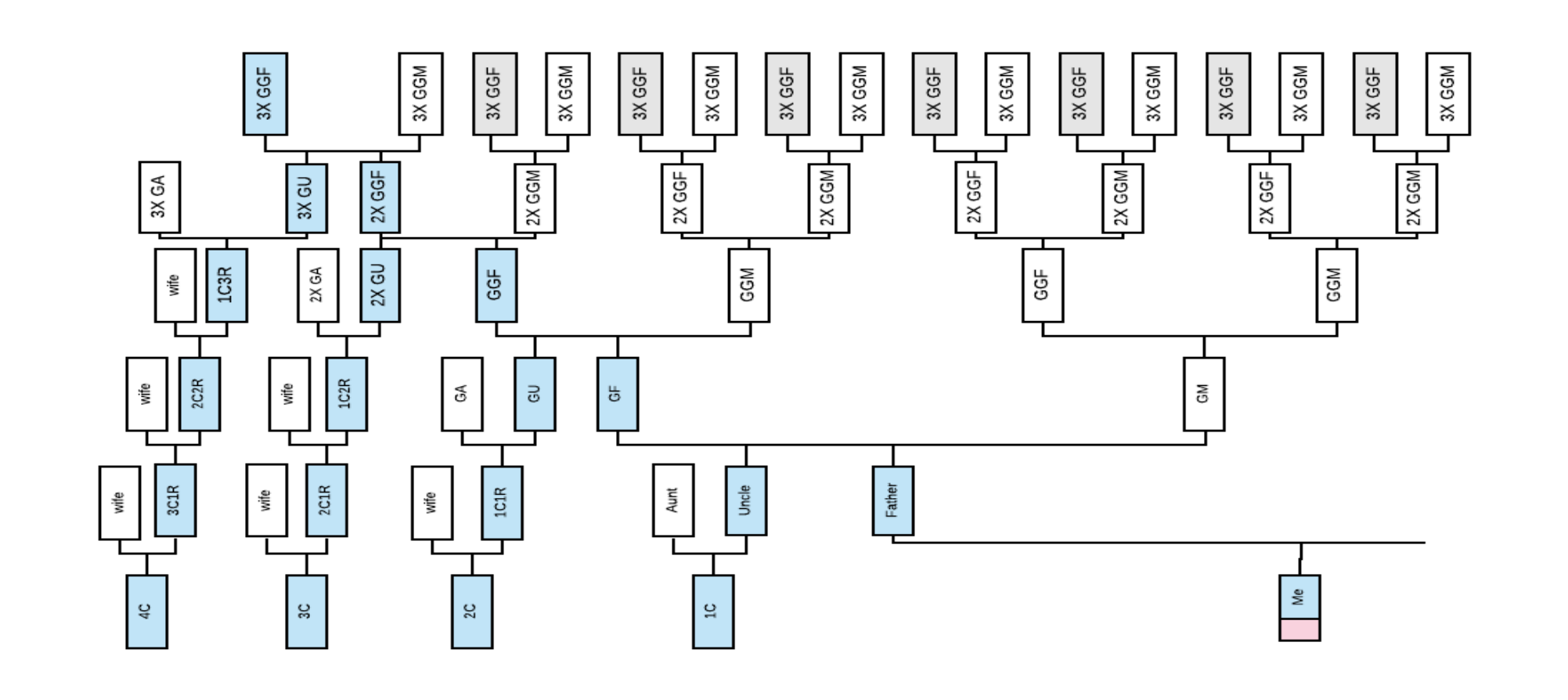
Y DNA Pedigree Chart



Maternal inheritance line (mtDNA) in pink

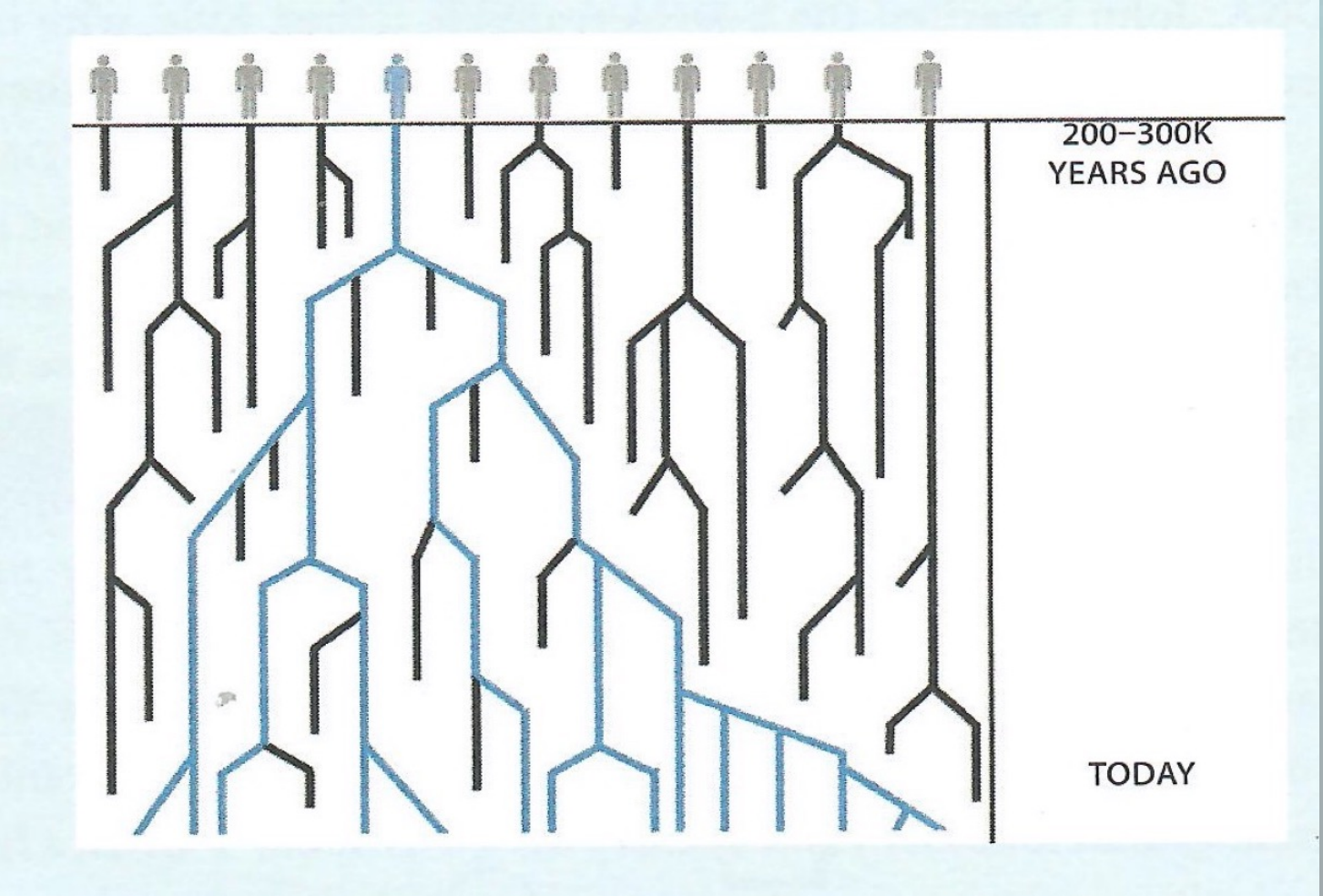
Paternal Inheritance Line (Y DNA) in blue (note how few 3X GGF are included in the Y DNA investigation).

Y DNA Connections



Even though only one of your 3X GGPs has your Y DNA, there are often a lot of male descendants of your 3X GGF (in blue) that would be expected to have your Y DNA. And the 3X GGFs (in gray) also have possible genealogical relevance to you as their male descendants would also be your cousins.

Y-DNA Back to Y-Adam



From The Family Tree Guide to DNA Testing and Genetic Genealogy by Dr. Blaine Bettinger. Depicts the tracing of all Y-DNA in existence today ack to one man referred to as Y-Adam. He was not the only male in existence at that time, however, all other lines of descent died out by not producing an unending string of males to the present time. Dr. Gleeson's video explains that it is estimated that 95% of all ancestral lines have become extinct.

Partial Coverage of Y Chromosome with Big Y-700

Figure 1: *Roughly 23.6 Mbp of chromosome Y (white) are genealogically relevant*

The Big Y-700 test only covers about 40% of the Y chromosome (40% breadth of coverage). The inaccessible region (black) contains a lot of repetitive sequences. The gray regions (PAR1 and PAR2) recombine with the X chromosome and therefore do not make these regions stable enough for paternal heritage information. From FTDNA white paper on FTDNA blog site.

Representation of the “Readable” Y chromosome

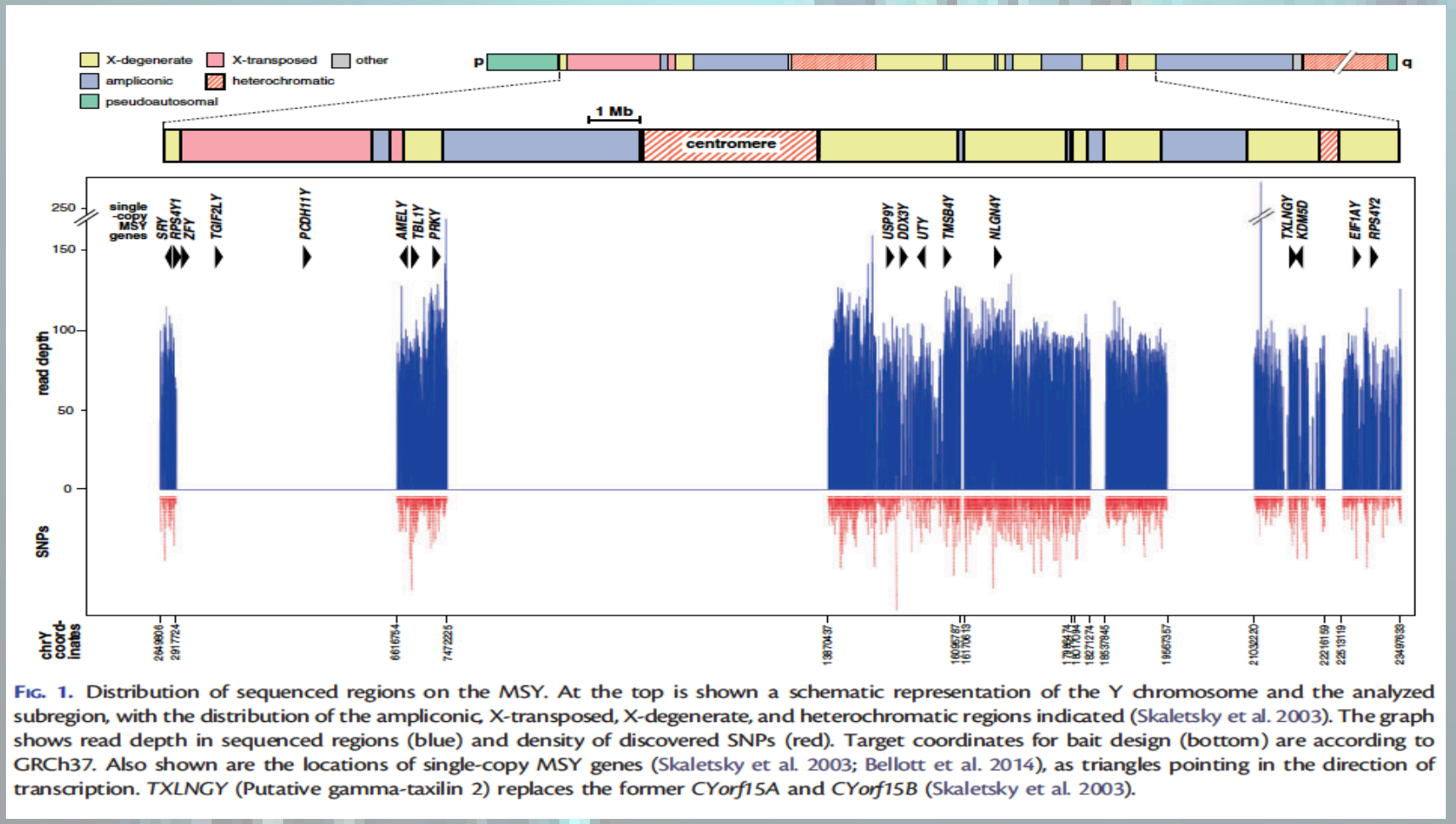


FIG. 1. Distribution of sequenced regions on the MSY. At the top is shown a schematic representation of the Y chromosome and the analyzed subregion, with the distribution of the ampliconic, X-transposed, X-degenerate, and heterochromatic regions indicated (Skaletsky et al. 2003). The graph shows read depth in sequenced regions (blue) and density of discovered SNPs (red). Target coordinates for bait design (bottom) are according to GRCh37. Also shown are the locations of single-copy MSY genes (Skaletsky et al. 2003; Bellott et al. 2014), as triangles pointing in the direction of transcription. *TXLNGY* (Putative gamma-taxilin 2) replaces the former *CYorf15A* and *CYorf15B* (Skaletsky et al. 2003).

From a paper entitled “The Y-chromosome Tree Bursts into Leaf: 13,000 High Confidence SNPs Covering the Majority of Known Clades” (numerous authors); Mol.Biol.Evol. 32(3):661-673.

Definition of an STR

from the FTDNA Learning Center

- An STR is a short tandem repeat. This is a place in your DNA code where a letter sequence is repeated. For example, AGTAAGTAAGTA is three repeats of the sequence AGTA. STRs have a fast mutation rate. Some STRs mutate faster than others. When they change, it is an increase or decrease in the number of repeats. STR values change back (back mutate) more common.
- An STR is not an STR is not an STR. Some are more polymorphic (more alleles throughout the population) than others. STRs also differ in mutation rates

Haplo……Type, Group or Tree

- Haplotype – A 111 marker STR test reveals the number of repeats at 111 individual locations on the Y chromosome. When these are all listed together, it is called a Y-STR haplotype. Haplotypes can be used to predict the haplogroup fairly accurately, but they are conservative predictions (i.e. the predictions don't go very far out on a limb…ha, ha)
- Haplogroup – Related haplotypes belong to the same genetic family which we call a haplogroup. A haplogroup is a major branch on either the maternal or paternal tree of humankind. Haplogroups are further divided into subclades defined as originating with a SNP mutation.
- Haplotree – The tree of humankind (either maternal or paternal) displaying the relationship of all known haplogroups and subclades.

History of STR Testing

- First commercially available Y-DNA tests were in 2000.
- They were low resolution tests of 12 markers. Then 25 markers were introduced, then 37 markers and lastly 67 then 111 markers.
- In order to unify the haplotree, the Y Chromosome Consortium (YCC) was formed in Feb. 2002. At the time, there were 153 branches and 245 variants on the haplotree. As of May 11, there are now 37,044 branches in the Y tree.

Example of an STR (DYS391)

		Result on Report
-TCTATCTATCTATCTATCTATCTA-	$\{TCTA\}_6$	6
-TCTATCTATCTATCTATCTATCTATCTA-	$\{TCTA\}_7$	7
-TCTATCTATCTATCTATCTATCTATCTATCTA-	$\{TCTA\}_8$	8

Alleles are one of two or more alternative forms of a gene that arise by mutation and are found at the same place on a chromosome. The repeat values above (6, 7 and 8) are considered alleles.

DNA Polymerase Slippage Formation of STRs

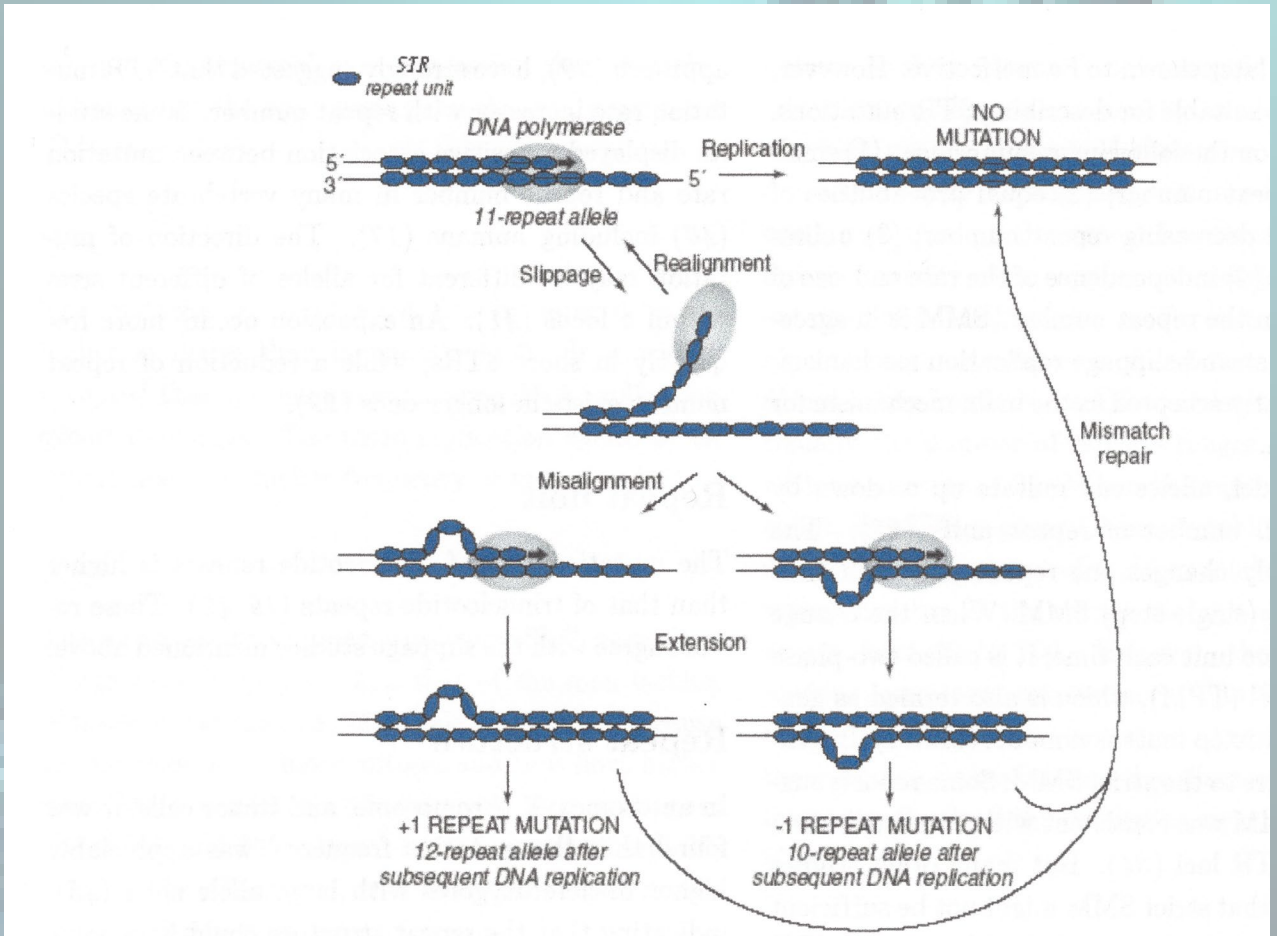
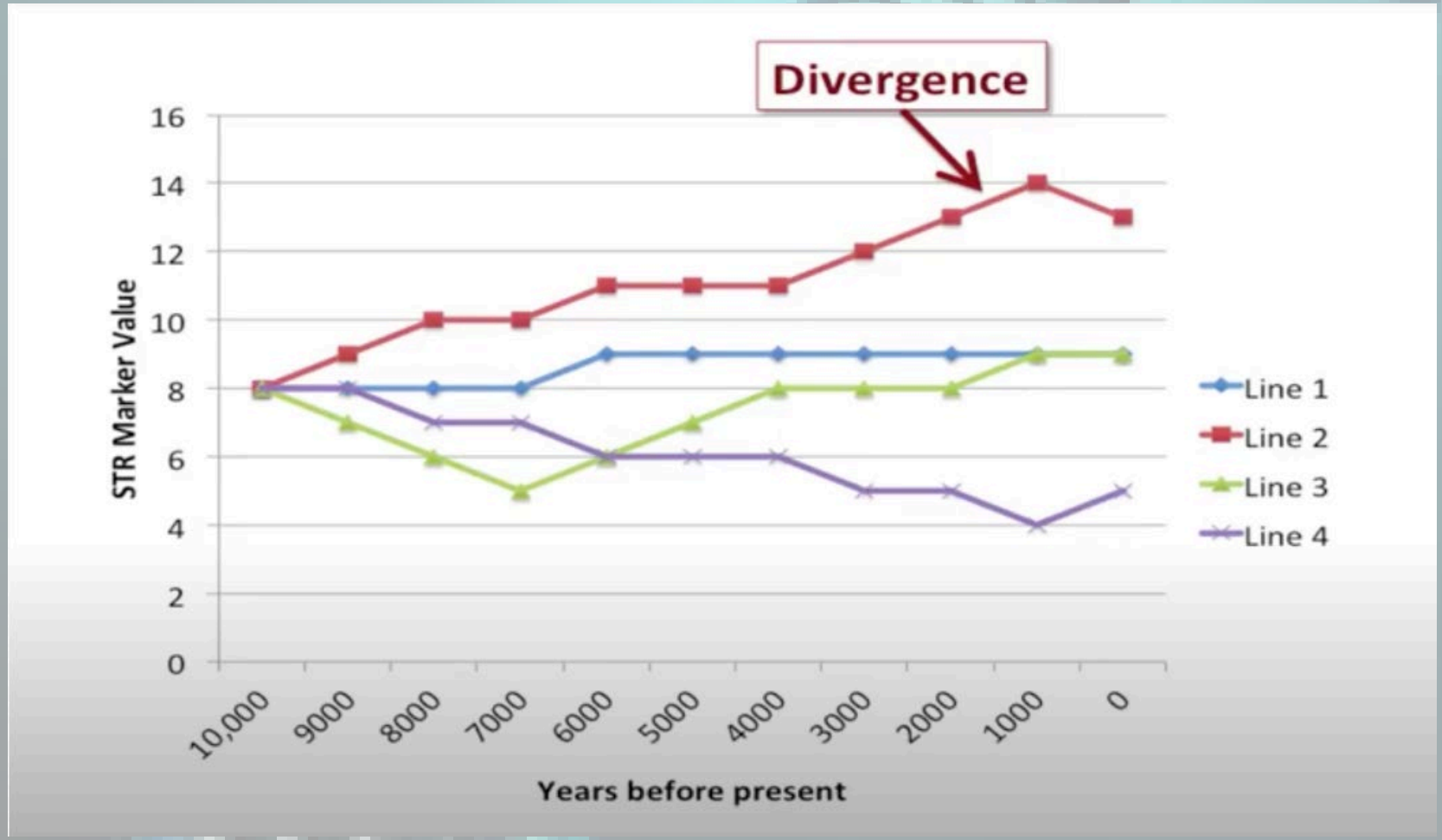


Fig. 1 Schematic illustration of the strand-slippage replication at STR (*24*).

STRs are less reliable for deep ancestry because of their ability to “back mutate” and create “convergence” situations which can be misleading in determining how long ago a common ancestor existed. From a paper entitled “A Brief Review of Short Tandem Repeat Mutation” by Hao Fan & Jia-You Chu, Genomics, Proteomics, Bioinformatics 2007; 5(1): 7-14

Example of Convergence/Divergence From Maurice Gleeson's presentation



Divergence is the expected result of time passage. Convergence reverses that and obscures time passage.

How are STRs Named (FTDNA Learning Center)

What do the DYS, DYZ, DYF, and FTY prefixes on Y-DNA STR names mean?

DYS, DYZ, and DYF Prefixes

The DYS, DYZ, and DYF prefixes are part of the scientific name for a short tandem repeat (STR) found on the Y chromosome. STR markers are named according to guidelines published by the HUGO Gene nomenclature committee (HUGO). For Y-DNA STR tests:

- **D** stands for DNA.
- **Y** stands for Y chromosome.
- **S, Z,** and **F** stand for the complexity of the repeat segment as follows:
  - **S** is a unique segment.
  - **Z** is a number of repetitive segments at one site.
  - **F** is a segment that has multiple copies on the Y chromosome.

FTY Prefix

The FTY prefix stands for Family Tree Y. For now, this prefix acts as a placeholder until HUGO assigns an official prefix to these STRs.

Identification Number

All STRs are given a unique identification number.

For example, DYS393: the **D** indicates that the segment is a DNA segment, the **Y** indicates that the segment is on the Y chromosome, the **S** indicates that it is a unique segment, and the number 393 is the identifier.

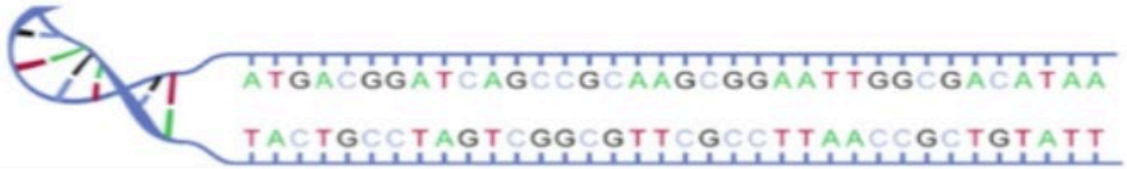
There are at least 4 kinds of STRs

- Simple repeats
- Compound repeats
- Complex repeats
- Complex Hypervariable repeats

- DYS464 and DYS385 are examples of complex repeats used in genealogy, but today we're going to focus on simple repeats for ease of comprehension.
- Also, not all STRs mutate at the same rate. There are known fast mutating STRs and slow mutating STRs. Fast mutating STRs are better for matching in more recent times, while slow mutating STRs are better for deep ancestry.

Some Actual Y Chromosome STRs

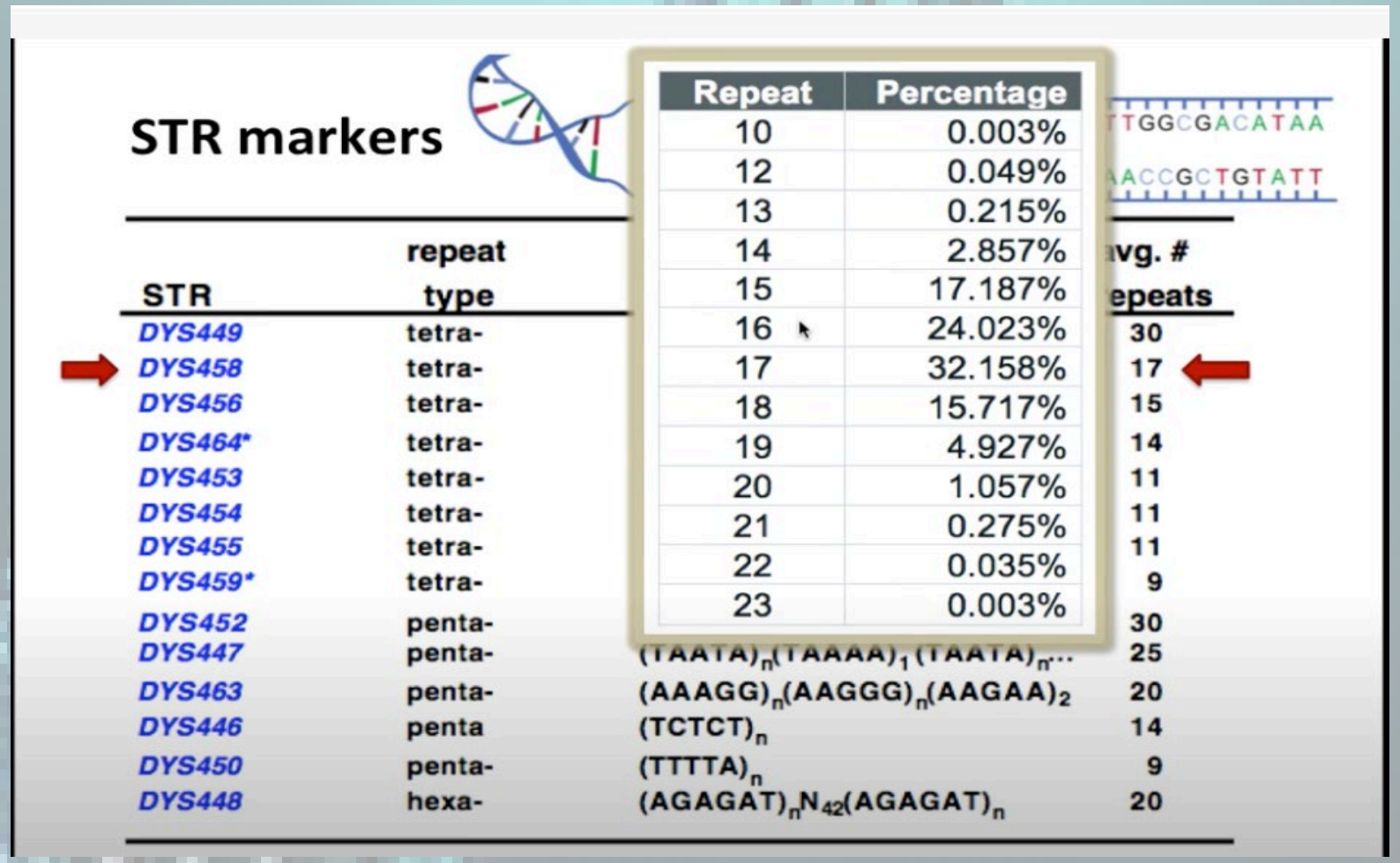
STR markers



STR	repeat type	repeat motif	avg. # of repeats
DYS449	tetra-	$(TTTC)_nN_{50}(TTTC)_n$	30
DYS458	tetra-	$(GAAA)_n$	17
DYS456	tetra-	$(AGAT)_n$	15
DYS464*	tetra-	$(CCTT)_n$	14
DYS453	tetra-	$(AAAT)_n$	11
DYS454	tetra-	$(AAAT)_n$	11
DYS455	tetra-	$(AAAT)_n$	11
DYS459*	tetra-	$(TAAA)_n$	9
DYS452	penta-	$(TATAC)_2(TGTAC)_2(TATAC)_n$...	30
DYS447	penta-	$(TAATA)_n(TAAAA)_1(TAATA)_n$...	25
DYS463	penta-	$(AAAGG)_n(AAGGG)_n(AAGAA)_2$	20
DYS446	penta	$(TCTCT)_n$	14
DYS450	penta-	$(TTTTA)_n$	9
DYS448	hexa-	$(AGAGAT)_nN_{42}(AGAGAT)_n$	20

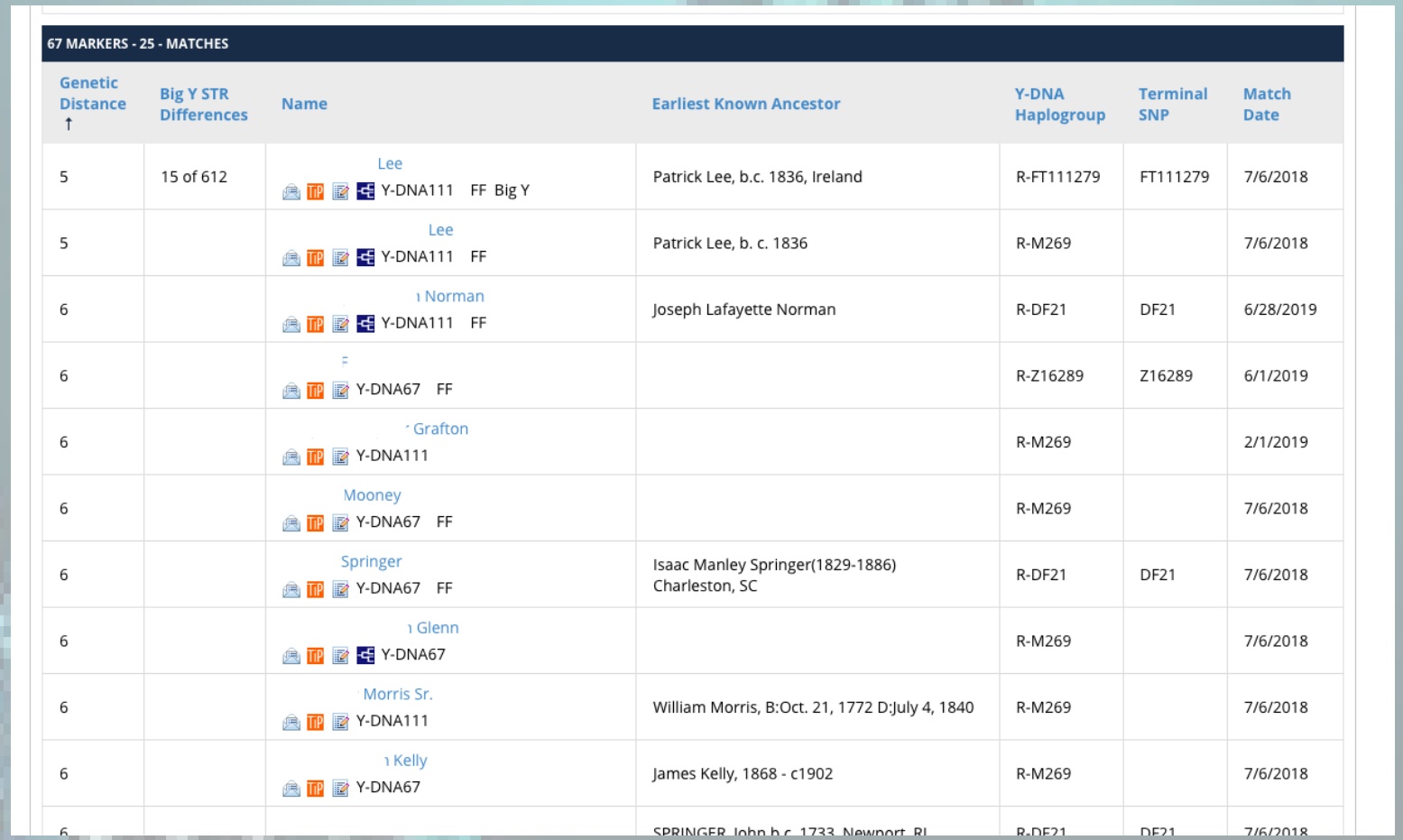
This is a slide from a presentation by Dr. Maurice Gleeson that is available on YouTube entitled “How Y DNA Can Help Your One Name Study”. I highly recommend that you watch it.

Distribution of STR Repeat # at Each Location



From Maurice Gleeson's YouTube video Using y DNA to Research Your Surname. Range of values & frequencies at https://yhrd.org/pages/resources/locus_information

Y STR Matches Report

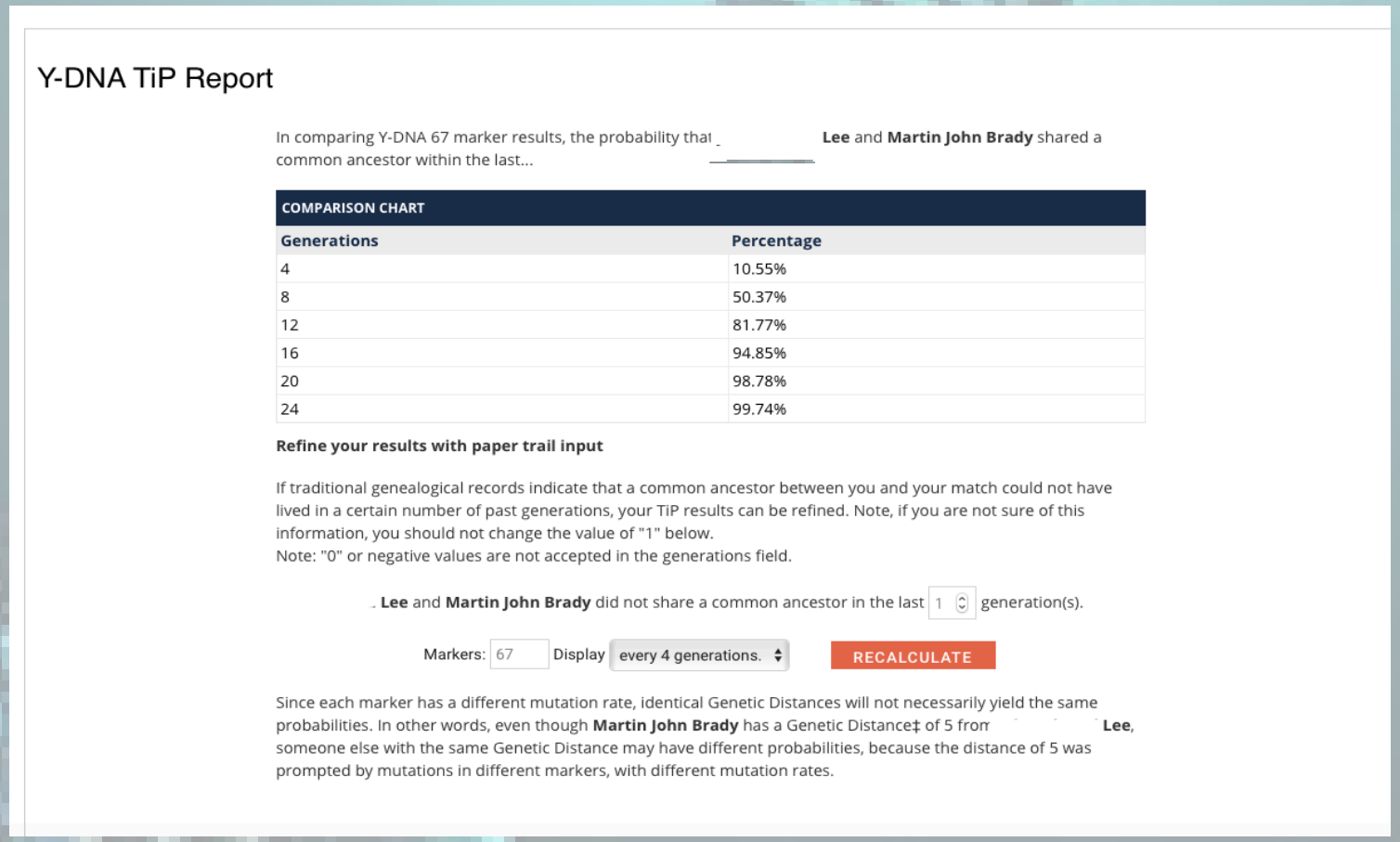


67 MARKERS - 25 - MATCHES

Genetic Distance ↑	Big Y STR Differences	Name	Earliest Known Ancestor	Y-DNA Haplogroup	Terminal SNP	Match Date
5	15 of 612	Lee Y-DNA111 FF Big Y	Patrick Lee, b.c. 1836, Ireland	R-FT111279	FT111279	7/6/2018
5		Lee Y-DNA111 FF	Patrick Lee, b. c. 1836	R-M269		7/6/2018
6		ı Norman Y-DNA111 FF	Joseph Lafayette Norman	R-DF21	DF21	6/28/2019
6		Y-DNA67 FF		R-Z16289	Z16289	6/1/2019
6		Grafton Y-DNA111		R-M269		2/1/2019
6		Mooney Y-DNA67 FF		R-M269		7/6/2018
6		Springer Y-DNA67 FF	Isaac Manley Springer(1829-1886) Charleston, SC	R-DF21	DF21	7/6/2018
6		ı Glenn Y-DNA67		R-M269		7/6/2018
6		Morris Sr. Y-DNA111	William Morris, B:Oct. 21, 1772 D:July 4, 1840	R-M269		7/6/2018
6		ı Kelly Y-DNA67	James Kelly, 1868 - c1902	R-M269		7/6/2018
6			SPRINGER, John b.c. 1733, Newport, RI	R-DF21	DF21	7/6/2018

I have 0 matches at the 111 STR level. I have 25 matches at the 67 STR level, however, the closest has a Genetic Distance of 5. I have redacted the first names, but you can see that none of my matches share my surname.

Time Predicter (TiP) Report



Y-DNA TiP Report

In comparing Y-DNA 67 marker results, the probability that _____ **Lee** and **Martin John Brady** shared a common ancestor within the last...

COMPARISON CHART	
Generations	**Percentage**
4	10.55%
8	50.37%
12	81.77%
16	94.85%
20	98.78%
24	99.74%

Refine your results with paper trail input

If traditional genealogical records indicate that a common ancestor between you and your match could not have lived in a certain number of past generations, your TiP results can be refined. Note, if you are not sure of this information, you should not change the value of "1" below.
Note: "0" or negative values are not accepted in the generations field.

. **Lee** and **Martin John Brady** did not share a common ancestor in the last 1 generation(s).

Markers: 67 Display every 4 generations. RECALCULATE

Since each marker has a different mutation rate, identical Genetic Distances will not necessarily yield the same probabilities. In other words, even though **Martin John Brady** has a Genetic Distance‡ of 5 from **Lee**, someone else with the same Genetic Distance may have different probabilities, because the distance of 5 was prompted by mutations in different markers, with different mutation rates.

Note the Time Predicter calculation takes into account whether or not the differences between you and a match are on fast mutating STRs or slow mutating STRs. The Genetic Distance (GD) determination does not take the mutation rate of the STR into account. GD is simply the total number of repeat differences.

Expected Relationship Match

	Y-DNA12	Y-DNA25	Y-DNA37	Y-DNA67	Y-DNA111	Interpretation
Very Tightly Related	N/A	N/A	0	0	0	Your exact match means your relatedness is extremely close. 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.
Tightly Related	N/A	N/A	1	1-2	1-2	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.
Related	0	0-1	2-3	3-4	3-5	Your degree of matching is within the range of most well-established surname lineages in Western Europe. If you have tested with the Y-DNA12 or Y-DNA25 test, you should consider upgrading to additional STR markers. Doing so will improve your time to common ancestor calculations.
Probably Related	1	2	4	5-6	6-7	Without additional evidence, it is unlikely that you share a common ancestor in recent genealogical times (one to six generations). You may have a connection in more distant genealogical times (less than 15 generations). If you have traditional genealogy records that indicate a relationship, then by testing additional individuals you will either prove or disprove the connection.
Only Possibly Related	2	3	5	7	8-10	It is unlikely that you share a common ancestor in genealogical times (one to 15 generations). Should you have traditional genealogy records that indicate a relationship, then by testing additional individuals you will either prove or disprove the connection. A careful review of your genealogical records is also recommended.
Not Related	3	4	6	>7	>10	You are not related on your Y-chromosome lineage within recent or distant genealogical times (one to 15 generations).

From the FTDNA Learning Center

Connections with STRs

- According to Dr. Gleeson, if you have no differences at 111 markers, you have a 50% chance of being a first cousin or closer.
- DYS393 has a slow mutation rate of 0.00076 or about 1 mutation every 1315 transmission events. DYS439 has a mutation rate of 0.00477 or about 1 mutation in every 210 transmission events. It is likely that the common ancestor of two men who differ by only “fast” markers could be significantly more recent than two men who differ by only “slow” markers. (Blaine Bettinger)

Statistics

- It is important to realize that the TiP feature is a statistical algorithm.
- As such, it deals in probabilities not facts.
- There is no guarantee that a mutation will occur every so many years in any particular lineage.
- There could be several mutations within 4 generations and there could be no mutations within 500 years.
- If you flip a coin, there is a 50/50 chance of getting heads. If the last 100 flips were tails, it doesn't mean that there is an increased likelihood of getting heads on the next flip. The chance of getting heads on the next flip is still 50/50.

Genetic Distance

Panel 1 (1-12)	DYS 393	DYS 390	DYS 19	DYS 391	DYS 385	DYS 426	DYS 388	DYS 439	DYS 389I	DYS 392	DYS 389II	Genetic Distance
Me	13	25	14	10	11-14	12	12	11	13	13	29	
Match 1	13	25	13	10	11-14	12	12	11	12	13	29	2

Panel 2 (13-25)	DYS 458	DYS 459	DYS 455	DYS 454	DYS 447	DYS 437	DYS 448	DYS 449	DYS 464	Genetic Distance
Me	16	9-10	11	11	25	15	19	28	15-15-17-17	
Match 1	16	9-10	13	11	25	15	19	28	15-15-17-17	2

In the first 12 STRs the match and I differ by 2 repeat units (one at DYS19 and one at DYS389I). In the second panel (STRs 13 -25) the match and I differ by 2 repeat units as well (both at DYS455). So, over the first 25 STRs, the match and I differ by 4 repeat units and therefore have a Genetic Distance of 4. If I test the next 12 STRs (26 through 37, a 37-marker test) and have no more differences, this match will show up on my match list because the threshold for displaying matches at the 37-marker level is 4 differences. If we have just 1 more difference, he won't show. The threshold at 111 markers is 10 differences.

Thresholds and matches

- Because of the thresholds used by FTDNA, it is possible for a Y STR match to show up for the 111-marker test when they did not show as a match on the 37-marker test. I have 7 matches at 37 markers, 25 matches at 67 markers and 0 matches at 111 markers.
- Some matches show a GD of 4 at 37 markers and a GD of 6 at 67 markers.
- Surname projects may allow you to see matches that don’t meet the threshold criteria above.

Carroll Surname Project page

Ely Carroll DNA Project - Y-DNA Colorized Chart

For genealogy within the most recent fifteen generations, STR markers help define paternal lineages. Y-DNA STR markers change (mutate) often enough that most men who share the same STR results also share a recent paternal lineage. This page displays Y-Chromosome DNA (Y-DNA) STR results for the project. It uses the colorized format. The columns display each project member's kit number, paternal ancestry information according to project settings, the paternal tree branch (haplogroup), and actual STR marker results. The color coding of STR names is explained here. In the haplogroups column, haplogroups in green are confirmed by SNP testing. Haplogroups in red are predicted. Above each subgroup, we display the minimum, maximum and mode values for each STR marker in the subgroup. STR marker values that differ from the mode values are color-coded. You can read about the coding system here. You may learn more about Y-DNA STRs on the Understanding Y-DNA STRs learning page.

Markers: Y-DNA12 **Page Size:** 500 Show All Columns

Kit Number	Name	Paternal Ancestor Name	Haplogroup	DYS393	DYS390	DYS19	DYS391	DYS385	DYS426	DYS388	DYS439	DYS389i	DYS392	DYS389ii	DYS458	DYS459	DYS455	DYS454	DYS447	DYS437	DYS448	DYS449	DYS464	DYS460	Y-GATA-H4	YCAII	DYS456	DYS607	DYS576	DYS570	CDY	DYS442	DYS438
100 R-Z16291>Z16284>Z18012>BY3301>FGC63585>FGC63582 Egan																																	
MIN				13	25	14	11	11-14	12	12	12	13	13	30	16	9-10	11	11	25	16	19	27	15-15-17-17	11	11	19-23	15	15	19	17	37-37	11	12
MAX				13	25	14	11	11-14	12	12	12	13	13	30	16	9-10	11	11	25	16	19	27	15-15-17-18	11	11	19-23	15	15	19	17	37-38	11	12
MODE				13	25	14	11	11-14	12	12	12	13	13	30	16	9-10	11	11	25	16	19	27	15-15-17-18	11	11	19-23	15	15	19	17	37-38	11	12
3601	Egan	Michael Egan, Clonaslee, Laois	R-FGC63582	13	25	14	11	11-14	12	12	12	13	13	30	16	9-10	11	11	25	16	19	27	15-15-17-17	11	11	19-23	15	15	19	17	37-37	11	12
555223	Egan	Thomas Andrew Egan, 1875-1931, Farnans, Laois, Ire	R-FGC63582	13	25	14	11	11-14	12	12	12	13	13	30	16	9-10	11	11	25	16	19	27	15-15-17-18	11	11	19-23	15	15	19	17	37-38	11	12
105 R-Z16291>Z16284>Z18012>BY3301>FGC63585>FT67708 Flanagan																																	
MIN				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-16-17-17	11	11	19-23	15	15	19	17	37-38	12	12
MAX				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-16-17-17	11	11	19-23	15	15	19	17	37-38	12	12
MODE				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-16-17-17	11	11	19-23	15	15	19	17	37-38	12	12
917007	Best	Thomas Flanagan, b. 1821, d. 1887	R-FT67708	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-16-17-17	11	11	19-23	15	15	19	17	37-38	12	12
110 R-Z16291>Z16284>Z18012>BY3301>FGC63585 Flanagan																																	
MIN				13	25	14	11	11-14	12	12	12	13	14	29	16	9-10	11	11	25	15	19	28	15-15-16-17	11	11	19-23	15	15	19	17	36-38	12	12
MAX				13	25	14	11	11-14	12	12	12	13	14	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	19	17	36-38	12	12
MODE				13	25	14	11	11-14	12	12	12	13	14	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	19	17	36-38	12	12
IN32480	Flanagan	Paul Flanagan, b. 1953	R-M269	13	25	14	11	11-14	12	12	12	13	14	29	16	9-10	11	11	25	15	19	28	15-15-16-17	11	11	19-23	15	15	19	17	36-38	12	12
MK64948	OShaughnessy	Thomas Shaughnessy, born 15/02/1877, Rathangan, Co	R-M269	13	25	14	11	11-14	12	12	12	13	14	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	19	17	36-38	12	12
3722	Irving	John Flanagan, b. 1815	R-FGC63585	13	25	14	11	11-14	12	12	12	13	14	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	19	17	36-38	12	12
120 R-Z16291>Z16284>Z18012>BY3301>BY30489>BY84846 Carroll 2																																	
MIN				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
MAX				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
MODE				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
913044	Rose		R-M269	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
283359	Carroll	Douglas Carroll, b. 1765 and d. 1829	R-BY84846	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
6781	Rose	George Rose, b. 1740 and d. 1853	R-BY84846	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
130 R-Z16291>Z16284>Z18012>BY3301>BY30489 Carroll 2																																	
MIN				13	24	14	11	11-14	12	12	12	13	13	29	16	9-10	10	11	25	14	19	28	15-15-15-17	11	11	19-21	15	15	18	16	36-37	12	12
MAX				13	25	14	11	11-14	12	12	12	13	13	29	17	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	16	15	21	17	37-39	12	12
MODE				13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
229066	CARROLL	John Carroll, IRE B 1620 D 1714 Virginia	R-M269	13	24	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	36-38	12	12
810430	Carroll	William Carroll, b. 1810 and d. 1880	R-M269	13	24	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
112777	Carroll	Patrick Carroll	R-M269	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	10	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	19	17	36-37	12	12
14073	Carroll	Michael Joseph Patrick Carroll, b. 1855	R-P312	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	14	19	28	15-15-17-17	11	11	19-23	15	15	19	16	37-39	12	12
107055	Carroll	Edward Carroll, 1824 - 1900	R-Z16281	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-15-17	11	11	19-21	15	15	18	17	36-38	12	12
120553	Carroll	James Madison CARROLL, b. 1829 and d. 1882	R-M269	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
93928	Carroll		R-M269	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
B216219	Harrington	Nancy Wilson, b. 1946	R-M269	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	20	16	37-38	12	12
291857	Cairel	William Carroll,b1815	R-DF21	13	25	14	11	11-14	12	12	12	13	13	29	16	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	15	15	21	16	36-38	12	12
190806	Carroll	Daniel Carroll b. 1807 Limerick	R-BY30489	13	25	14	11	11-14	12	12	12	13	13	29	17	9-10	11	11	25	15	19	28	15-15-17-17	11	11	19-23	16	15	19	16	37-38	12	12
140 R-Z16291>Z16284>Z18012>FGC61823>FGC61825 Mitchell																																	
MIN				13	25	14	10	11-14	12	12	12	13	13	29	17	9-10	11	11	25	15	19	28	15-15-17-17	11	10	19-22	15	15	18	17	37-38	12	12
MAX				13	25	14	10	11-14	12	12	12	13	13	29	17	9-10	11	11	25	15	19	28	15-15-17-17	11	10	19-22	15	15	18	17	37-38	12	12
MODE				13	25	14	10	11-14	12	12	12	13	13	29	17	9-10	11	11	25	15	19	28	15-15-17-17	11	10	19-22	15	15	18	17	37-38	12	12

Carroll Project Y STR Results for me

MODE				13	25	14	11	11-14	12	12	12	13	13
256801	Redmond	James Redmond b 1833 d 1908	R-BY20009	13	25	14	11	11-14	12	12	12	13	13
230 R-Z16291>Z16284>FT14437>Z16289 Tracey 1 520=19													
MIN				13	25	14	10	11-14	12	12	12	13	13
MAX				13	25	14	11	11-14	12	12	13	13	13
MODE				13	25	14	11	11-14	12	12	12	13	13
622693	Treacy	Martin Treacy b 1800	R-M269	13	25	14	10	11-14	12	12	12	13	13
318497	Tracey	Tracey	R-M269	13	25	14	11	11-14	12	12	12	13	13
578385	Treacy	Michael Treacy b. 1815 Kiltullagh, Galway	R-M269	13	25	14	11	11-14	12	12	12	13	13
271750	Treacy	Michael Treacy, born 1815- 1876 Kiltullagh, Kilker	R-DF21	13	25	14	11	11-14	12	12	12	13	13
335793	Tracey	Tracey	R-M269	13	25	14	11	11-14	12	12	12	13	13
289639	Tracey	Lawrence Treacy 1813-1882	R-DF21	13	25	14	11	11-14	12	12	12	13	13
45013	Tracey	Lawrence Tracey b.c.1825 Glenamaddy Galway Ireland	R-Z16289	13	25	14	11	11-14	12	12	12	13	13
164932	Tracy		R-M269	13	25	14	11	11-14	12	12	13	13	13
129222	Tracy	Dennis Tracy, b.c. 1795, Irelanddied between 1835-	R-L21	13	25	14	11	11-14	12	12	13	13	13
240 R-Z16291>Z16284>BY4011>BY4005>BY19556 Meagher 594=11													
MIN				13	25	14	11	11-13	12	12	12	13	13
MAX				13	25	14	12	12-14	12	12	13	14	13
MODE				13	25	14	11	11-14	12	12	12	13	13
N73615	Maher	Thomas Maher	R-M269	13	25	14	11	11-13	12	12	12	13	13
8099	Mahar	Thomas Mahar (1834-1904) Co. Queens, Ireland	R-M269	13	25	14	11	11-14	12	12	12	13	13
430970	Maher	Edward Maher, 1816, Conahy, Kilkenny	R-M269	13	25	14	11	11-14	12	12	12	13	13
509147	Maher	Bartholomew Maher	R-M269	13	25	14	11	11-14	12	12	12	13	13
N144276	McCormack	Richard CORMACK b1782 Killenaule Tipperary	R-BY19556	13	25	14	11	11-14	12	12	12	13	13
22400	Maher	William Maher, c.1810 - 1872	R-M269	13	25	14	11	11-14	12	12	12	14	13
396950	Maher	Philip Maher, b. 1808 and d. 1875	R-BY19556	13	25	14	11	11-14	12	12	13	13	13
B34187	Warner		R-M269	13	25	14	11	11-14	12	12	13	13	13
26665	Mahar	Laurence Maher b. 1819 d. 1876	R-BY4005	13	25	14	11	12-14	12	12	12	13	13
633251	Harvey	Greg Bateman	R-M269	13	25	14	12	11-14	12	12	12	13	13
245 R-Z16291>Z16284>BY4011>BY4005>FGC65865>BY194635>BY193379 Meagher 594=11, 19=15													
MIN				13	25	15	11	11-14	12	12	12	12	13
MAX				13	25	15	11	11-14	12	12	12	12	13
MODE				13	25	15	11	11-14	12	12	12	12	13
B75307	Maher	William Maher, b. abt 1788, Killough, Templemore,	R-FT183081	13	25	15	11	11-14	12	12	12	12	13
250 R-Z16291>Z16284>BY4011>BY4005>FGC65865>BY194635 Meagher 594=11, 19=15													
MIN				13	25	15	11	11-14	12	12	12	12	13
MAX				13	25	15	11	11-14	12	12	12	12	13
MODE				13	25	15	11	11-14	12	12	12	12	13
B387131	Meagher	Stephen Maher (Meagher), b. 1807 and d. 1879	R-BY194635	13	25	15	11	11-14	12	12	12	12	13
260 R-Z16291>Z16284>BY4011>BY4005>FGC65865 Meagher 594=11, 19=15													
MIN				13	25	14	10	11-14	12	12	12	12	13
MAX				13	25	15	12	11-14	12	12	12	14	13
MODE				13	25	15	11	11-14	12	12	12	13	13
385132	Maher	John Maher b.c. 1775	R-BY4005	13	25	14	11	11-14	12	12	12	14	13
24434	O'Meagher	Patrick Maher, c1810-unk, Tullaherin Kilkenny	R-FGC65865	13	25	15	10	11-14	12	12	12	13	13
897152	Maher	James Maher, b. 1817, d. 1872	R-M269	13	25	15	11	11-14	12	12	12	12	13
897247	Threlfall		R-M269	13	25	15	11	11-14	12	12	12	13	13
375935	Maher	Eamon Maher, b. 1800 and d. 1878	R-M269	13	25	15	12	11-14	12	12	12	13	13
270 R-Z16291>Z16284>BY4011 Tracey 2 594=11													
MIN				13	25	14	10	11-13	12	12	12	13	13
MAX				13	25	14	11	11-14	12	12	12	13	13
MODE				13	25	14	11	11-14	12	12	12	13	13
185725	Tracy	Eamon Treacy 1790- 1844	R-DF21	13	25	14	10	11-14	12	12	12	13	13
439049	Harrison/Tracy	John Treacy, b. 1836, Bulgaden, Co. Limerick, Irel	R-BY4011	13	25	14	11	11-13	12	12	12	13	13
273 R-Z16291>Z16284>FT14437 Brady													
MIN				13	25	14	10	11-14	12	12	11	13	13
MAX				13	25	14	10	11-14	12	12	11	13	13
MODE				13	25	14	10	11-14	12	12	11	13	13
529257	Brady	Thomas Brady, 1806-1873	R-FT14437	13	25	14	10	11-14	12	12	11	13	13
276 R-Z16291>Z16284 Carey 1 520=19													
MIN				13	25	14	11	11-14	12	12	13	13	13
MAX				13	25	14	11	11-14	12	12	13	13	13
MODE				13	25	14	11	11-14	12	12	13	13	13
298623	Carey	John Carey b. 1817 Ireland d. 1891	R-Z16284	13	25	14	11	11-14	12	12	13	13	13
280 R-Z16291>BY3829>BY3303>BY116174 Bohan 460=10, 511=11													

Carroll Project Y STR Results for FT14437

160 R-Z16291>Z16284>Z18012 Bowe									
MIN				13	24	14	11	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
276544	Bowes	Phillip Bowes, 1734-1813, Isle of Wight, England	R-DF21	13	24	14	11	11-14	12
177452	Bowe	Edmund Bowe 1726-1794, Lough, Kilkenny, Ireland	R-M269	13	25	14	11	11-14	12
203814	Bowe	Michael Bowe, Ballycuddihy, Johnstown, Kilkenny, I	R-M269	13	25	14	11	11-14	12
146114	Bowe	William Bowe, Kilkenny (most likely) Ireland	R-Z18012	13	25	14	11	11-14	12
174346	Bowes	Denis Boe, bef. 1800, prob. Muckalee Parish, Kilke	R-DF21	13	25	14	11	11-14	12
170 R-Z16291>Z16284>FT14437>Z16289>BY20009>BY20010>BY95483 Carroll 1 520=19									
MIN				13	24	14	11	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
359921	Carroll	John Carrell, b. 1700-1710, prob. Virginia	R-BY95483	13	24	14	11	11-14	12
MK45929	Carroll	Sterling Carroll	R-BY95483	13	25	14	11	11-14	12
175 R-Z16291>Z16284>FT14437>Z16289>BY20009>BY20010>FT111279 Carroll 1 520=19									
MIN				13	25	14	10	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
54646	Lee	Patrick Lee, b.c. 1836, Ireland	R-FT111279	13	25	14	10	11-14	12
71400	Carroll	Denis Carroll 1809, NL Canada	R-FT111279	13	25	14	11	11-14	12
180 R-Z16291>Z16284>FT14437>Z16289>BY20009>BY20010 Carroll 1 520=19									
MIN				13	24	14	10	11-14	11
MAX				13	25	14	12	11-14	12
MODE				13	25	14	11	11-14	12
355417	Carroll	William Carrell, b UNk and d1811, Johnston , NC	R-M269	13	24	14	11	11-14	12
394786	Carroll	William Carroll, b. 1702 and d. 1754	R-M269	13	24	14	11	11-14	12
621368	Lee	Patrick Lee, b. c. 1836	R-M269	13	25	14	10	11-14	12
23133	Linville	Dudley Linville b1844 VA d 1873	R-BY20010	13	25	14	11	11-14	11
185954	Carroll	John Carroll, 1836-1910	R-DF21	13	25	14	11	11-14	12
B85668	Caudle	Moses Caudle, b. 1765 d.1853	R-M269	13	25	14	11	11-14	12
471460	Carroll	Elijah Carroll b. 1795 and d. 1875	R-M269	13	25	14	11	11-14	12
B5459	Brown	George W Brown 1858 -1912 unk father probable NPE	R-Z16281	13	25	14	11	11-14	12
300355	Carroll	Poss. Elijah, James Franklin Carroll 1857-1889	R-Z16289	13	25	14	11	11-14	12
112378	Carroll		R-P312	13	25	14	11	11-14	12
794034	Grafton		R-M269	13	25	14	11	11-14	12
112059	Carroll	John Carroll 1790 Tennessee	R-L21	13	25	14	11	11-14	12
599898	Carroll	Cosmas Carroll (1808 - 1878)	R-M269	13	25	14	11	11-14	12
738989	Carroll	Daniel O'Carroll b. 1642 and d. 1688	R-BY20010	13	25	14	12	11-14	12
190 R-Z16291>Z16284>FT14437>Z16289>BY20009>BY20011>BY50769 Crow 520=19									
MIN				13	25	14	11	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
823375	McGee		R-BY50769	13	25	14	11	11-14	12
178413	Davidson	BV Davidson, but non-paternity event in this line	R-Z16289	13	25	14	11	11-14	12
48750	Crow	CROW, George Washington b. 1822, KY d. 1892 MO	R-BY50769	13	25	14	11	11-14	12
200 R-Z16291>Z16284>FT14437>Z16289>BY20009>BY20011 Dooley 520=19									
MIN				13	25	14	11	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
115408	Dooley	H02/02 -Thomas Dooley, b.1804, Ballymack, Kilkenny	R-BY20011	13	25	14	11	11-14	12
N54552	Dillon	Daniel Dillon, Crusheen, Ire.; to U.S. 1900	R-M269	13	25	14	11	11-14	12
366379	Dooley	Martin Michael Dooley, b. 1824 and d. 1902	R-DF21	13	25	14	11	11-14	12
168720	Dooley	H - Jeremiah Dooley, b.c. 1790 Ireland	R-BY20011	13	25	14	11	11-14	12
210 R-Z16291>Z16284>FT14437>Z16289>BY20009 Murphy 520=19									
MIN				13	25	14	11	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
223130	Murphy	Neil McMurfie Ballynamony, Lurgan Co Armagh 1631	R-DF21	13	25	14	11	11-14	12
194206	Murphy	Edward Murphy lived Lurgan 1690 died 1705 Ireland	R-BY20009	13	25	14	11	11-14	12
220 R-Z16291>Z16284>FT14437>Z16289>BY20009 Redmond 520=19									
MIN				13	25	14	11	11-14	12
MAX				13	25	14	11	11-14	12
MODE				13	25	14	11	11-14	12
256801	Redmond	James Redmond b 1833 d 1908	R-BY20009	13	25	14	11	11-14	12
230 R-Z16291>Z16284>FT14437>Z16289 Tracey 1 520=19									

Leo Little's Spreadsheet

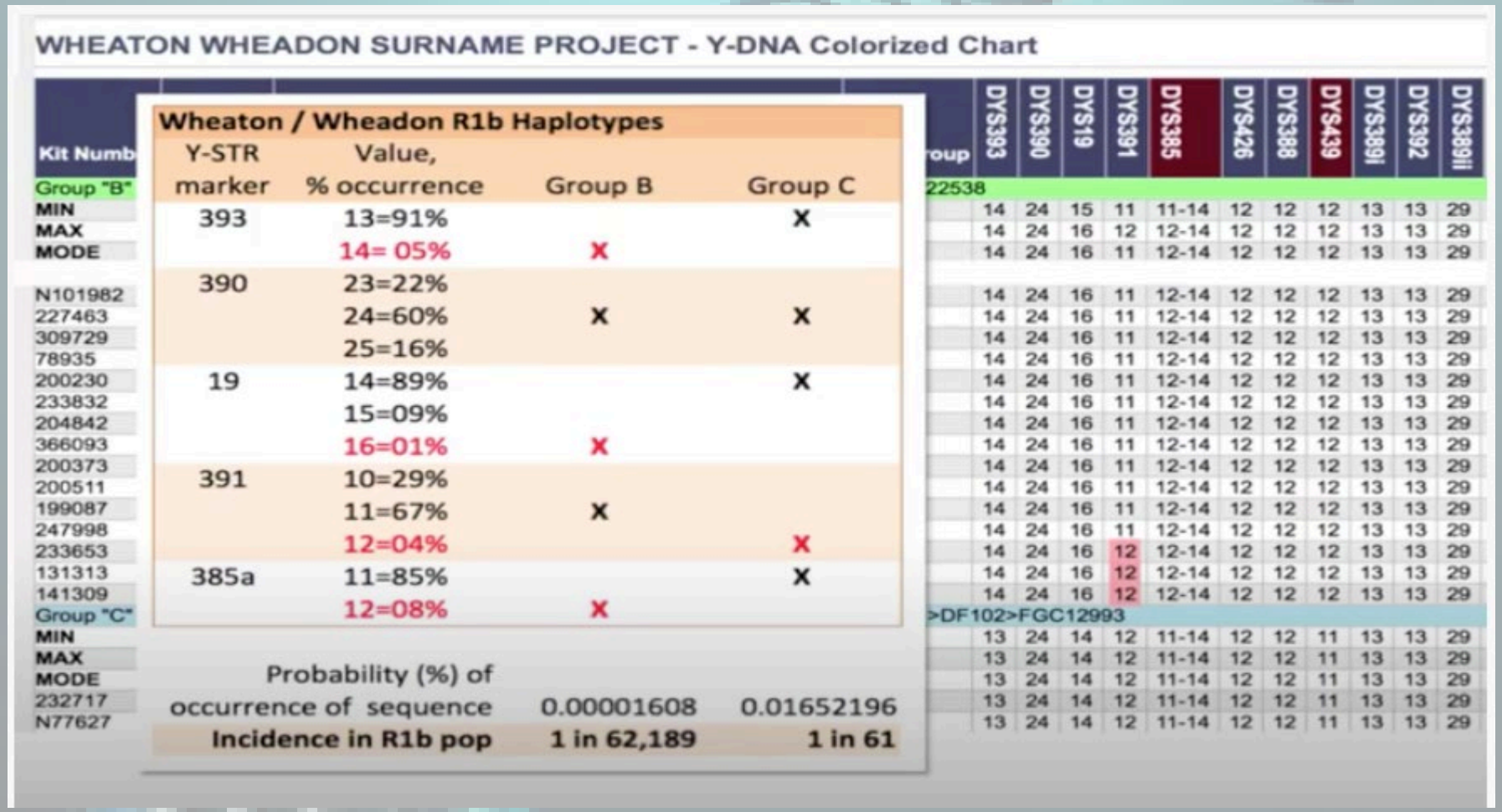
http://freepages.genealogy.rootsweb.ancestry.com/~geneticgenealogy/yfreq.htm

1	E3a	393	330							.02	.55	.27	.15	.01																											
1	E3b	393	1185							.04	.80	.14	.02																														
1	G	393	454							.01	.25	.64	.10		.00																													
1	I	393	5700							.03	.68	.17	.11	.00																														
1	J2	393	915						.01	.89	.10	.00																																
1	R1a	393	1574							.01	.95	.05																																
1	R1b	393	22129							.04	.91	.05	.00																																
2	hg	locus	count	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	
2	E3a	390	330															.01	.92	.05	.02																						
2	E3b	390	1185																	.02	.15	.67	.16	.01																		
2	G	390	454															.00	.05	.75	.16	.03																				
2	I	390	5700															.00	.01	.39	.45	.11	.03	.01																		
2	J2	390	915																.00	.14	.45	.36	.05	.00																		
2	R1a	390	1574																		.01	.15	.76	.08	.00																		
2	R1b	390	22129																.00	.01	.22	.60	.16	.01																		
3	hg	locus	count	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	
3	E3a	19	330									.02	.58	.23	.17																												
3	E3b	19	1185						.00	.00	.85	.12	.02	.01																														
3	G	19	454									.03	.89	.07	.01																												
3	I	19	5700								.01	.51	.33	.11	.04	.00																												
3	J2	19	915								.01	.44	.46	.08	.01																												
3	R1a	19	1574									.01	.51	.38	.10																												

Using STR Frequency Data to Examine My Y Results

- After finding out about the frequency distributions, I checked my 67-marker STR results.
- For DYS492 I had a result of 11 repeats. Only 1% of the R1b haplogroup have this number of repeats for DYS492.
- Nobody in the Brady Surname Project had this value.
- Almost everybody in the Carroll Surname Project has this value.
- Many of the other low-frequency STRs in my results are also in the Carroll Project and not in the Brady Project.
- This requires further investigation.

Rarity of Alleles Used for Assigning Test Takers to a Group



WHEATON WHEADON SURNAME PROJECT - Y-DNA Colorized Chart

Wheaton / Wheadon R1b Haplotypes

Y-STR marker	Value, % occurrence	Group B	Group C
393	13=91%		X
	14= 05%	X	
390	23=22%		
	24=60%	X	X
	25=16%		
19	14=89%		X
	15=09%		
	16=01%	X	
391	10=29%		
	11=67%	X	
	12=04%		X
385a	11=85%		X
	12=08%	X	
Probability (%) of occurrence of sequence		0.00001608	0.01652196
Incidence in R1b pop		**1 in 62,189**	**1 in 61**

Kit Numb	DYS393	DYS390	DYS19	DYS391	DYS385	DYS426	DYS388	DYS439	DYS389i	DYS392	DYS389ii
Group "B"	roup 22538										
MIN		14	24	15	11	11-14	12	12	12	13	13	29
MAX		14	24	16	12	12-14	12	12	12	13	13	29
MODE		14	24	16	11	12-14	12	12	12	13	13	29
N101982		14	24	16	11	12-14	12	12	12	13	13	29
227463		14	24	16	11	12-14	12	12	12	13	13	29
309729		14	24	16	11	12-14	12	12	12	13	13	29
78935		14	24	16	11	12-14	12	12	12	13	13	29
200230		14	24	16	11	12-14	12	12	12	13	13	29
233832		14	24	16	11	12-14	12	12	12	13	13	29
204842		14	24	16	11	12-14	12	12	12	13	13	29
366093		14	24	16	11	12-14	12	12	12	13	13	29
200373		14	24	16	11	12-14	12	12	12	13	13	29
200511		14	24	16	11	12-14	12	12	12	13	13	29
199087		14	24	16	11	12-14	12	12	12	13	13	29
247998		14	24	16	11	12-14	12	12	12	13	13	29
233653		14	24	16	12	12-14	12	12	12	13	13	29
131313		14	24	16	12	12-14	12	12	12	13	13	29
141309		14	24	16	12	12-14	12	12	12	13	13	29
Group "C"	>DF102>FGC12993										
MIN		13	24	14	12	11-14	12	12	11	13	13	29
MAX		13	24	14	12	11-14	12	12	11	13	13	29
MODE		13	24	14	12	11-14	12	12	11	13	13	29
232717		13	24	14	12	11-14	12	12	11	13	13	29
N77627		13	24	14	12	11-14	12	12	11	13	13	29

https://sites.google.com/site/wheatonsurname/beginners-guide-to-genetic-genealogy/lesson-14-more-with-the-y

The Y-DNA-12 marker test is sufficient to allocate some members to Group B (per Dr. Maurice Gleeson)

Big Y-700 STRs

- There were at least 500 STRs in the Big Y-500 test and there are at least 700 STRs in the Big Y-700 test (111 + 589), however, the additional 589 are currently extraneous information as the matching system for those STRs is not yet fully developed.

Differences between STRs & SNPs

from the FTDNA Learning Center

- An STR is a short tandem repeat. This is a place in your DNA code where a letter sequence is repeated. For example, AGTAAGTAAGTA is three repeats of the sequence AGTA. STRs have a fast mutation rate. Some STRs mutate faster than others. When they change, it is an increase or decrease in the number of repeats. STR values change back (back mutate) more common.
- A SNP is a single nucleotide polymorphisms. That means that it is a single small change in your DNA code. These changes are rare. Once they happen, they seldom change back (back mutate).

Big Y or Targeted SNPs

- You can test your (or a male relatives Y-DNA) with the Big Y-700 or targeted SNPs. Targeted SNPs are cheaper, but if you don't know what SNPs to target, you may waste time and money.
- The Big Y-700 uses Next Generation Sequencing (NGS) technology to discover new SNPs and detect known SNPs. While the FTDNA Learning Center states the below statement, the new Big Y-700 is discovering new SNPs at a rapid rate and if all males tested, we would know where every male sat on the haplotree.
  - "The Big Y test is intended for expert users with an interest in advancing science. It may also be of great interest to genealogy researchers of a specific lineage. However, it is not a test for matching you to one or more men with the same surname in the way that our other Y-STR tests do, such as Y-37, Y-67 or Y-111."

Terminal SNP

- Terminal SNP - A terminal SNP determines the terminal (final) subbranch on the Y-DNA Tree to which someone belongs. It is the SNP that is the farthest out on a limb of the Y-DNA Tree.
- For a variant to be placed on the Y-DNA Tree, at least two people have to be derived for that particular variant.
- Private Variants are “Singletons” and as such are not on the Y-DNA Tree.
- As additional men test, your private variants will almost certainly be paired with another and will then be placed on the Y-DNA Tree which will change your “Terminal” SNP.

SNP Discovery

- At my last look (May 12, 2020), there were almost 300,000 variants on the haplotree and growing quickly. This tidal wave of SNP discovery enabled by the use of new technology known as Next Generation Sequencing (NGS) is referred to as the SNP Tsunami.

Frequency of SNPs on the Y Chromosome

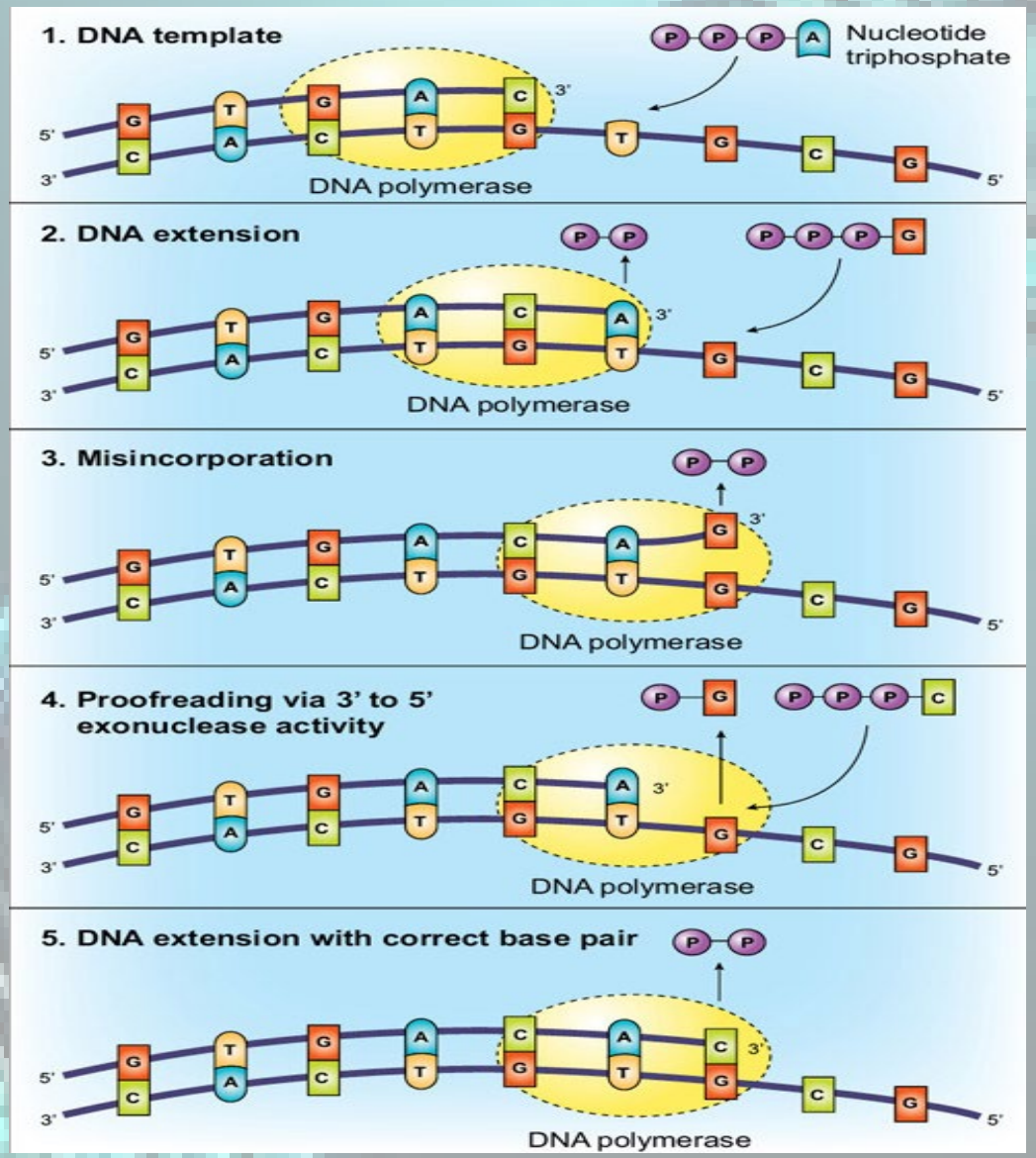
- 1.25 billion years for each variant at each nucleotide position.
- Since Big Y-500 tested 10 million nucleotides, the expected frequency of variants in this test is therefore one every 125 years.
- 1,250,000,000/10,000,000 = 125 years/SNP
- The Big Y-700 tests ~15 million nucleotides, so
- 1,250,000,000/15,000,000 = 84 years/SNP
- Full Genome now offers a Long-Read test of 20,000,000 Y nucleotides.

As we test more of the Y chromosome, we can refine our estimation of time between SNPs. 84 years is starting to get within genealogical timeframe. Statistics from “The Future of Y-DNA Testing for Genealogy” YouTube video by Dr. Iain McDonald

DNA Replication

- DNA Polymerase is the enzyme responsible for synthesizing new strands of DNA (i.e., making copies, replication).
- Enzymes are protein molecules in cells which work as biological catalysts. Enzymes speed up chemical reactions in the body, but do not get used up in the process, therefore can be used over and over again. Almost all biochemical reactions in living things need enzymes.
- DNA Polymerase works in one direction adding new nucleotides (dNTP) to the 3 position of deoxyribose backbone of the DNA strand.
- A nucleotide is the building block unit of nucleic acids such as DNA. When we talk about sequencing, we are talking about the sequential order of the different nucleotides as we progress down the DNA strand.
- Complementarity is maintained. Complementarity refers to the fact that each nucleotide has a base (A, C, T or G), and each base on a DNA strand can pair up with the appropriate (complementary) base (A with T and G with C) from the opposing nucleotide on the second DNA strand.

DNA Polymerase Activity and Mistake Correction



Naming System for SNPs

From “Origins of the Irish, Scottish, Welsh and English R1b-M222 population”

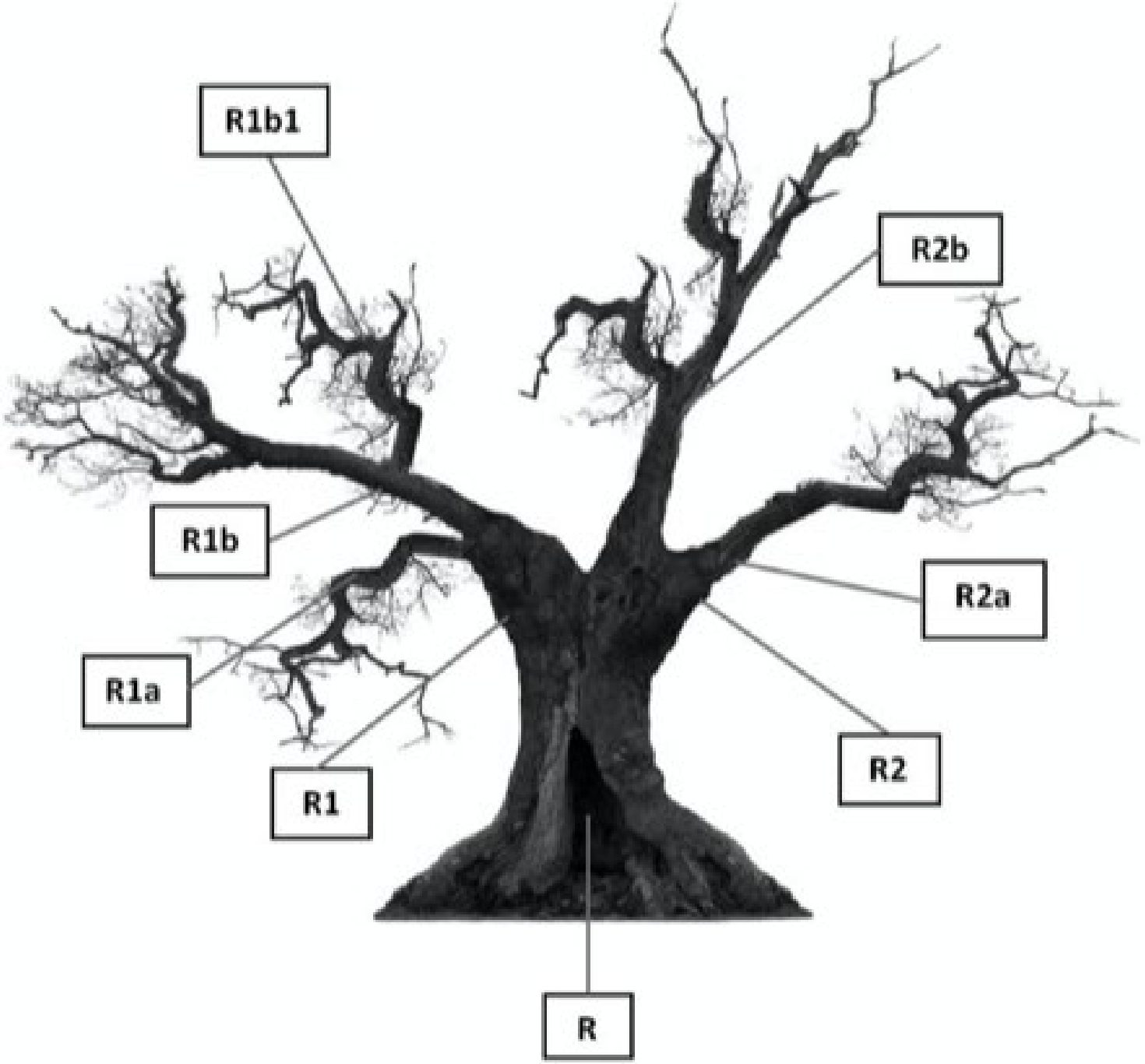
Old Naming System Lineage System	New Naming System Mutation System	Subclade appearance
R1b	M343	~16,000 ybp
R1b1a2	M269	~7,000 ybp
R1b1a2a	L23	~6,200 ybp
R1b1a2a1a	L51	~5,300 ybp
R1b1a2a1a1	L11	~4800 ybp
R1b1a2a1a1a	U106	
R1b1a2a1a1b	P312	
R1b1a2a1a1b3	U152	4125+/-450 ybp
R1b1a2a1a1b4	L21	3750+/-400 ybp

List of SNP Prefixes from ISOGG website

ALK = Ahmad Al Khuraiji
AM or AMM = Laboratory of Forensic Genetics and Molecular Archaeology, UZ Leuven, Leuven, Belgium
B = Estonian Genome Centre
BY = Big Y testing (next generation sequencing) discovered with the BigY-500, Family Tree DNA, Houston, Texas
BZ = Q-M242 Project, Family Tree DNA, Houston, TX. SNPs named in honor of Barry Zwick.
CTS = Chris Tyler-Smith, Ph.D., The Wellcome Trust Sanger Institute, Hinxton, England
DC = Dál Cais, an Irish group believed to be descended from Cas, b. CE 347, related to SNP R-L226; Dennis Wright
DF = anonymous researcher using publicly available full-genome-sequence data, including 1000 Genomes Project data; named in honor of the DNA-Foru
E = Bulat Muratov
F = Li Jin, Ph.D., Fudan University, Shanghai, China
F* = Chuan-Chao Wang, Hui Li, Fudan University, Shanghai, China (Beginning letter F; second letter Haplogroup, i.e. FI is Fudan Haplogroup I)
FGC = Full Genomes Corp. of Virginia and Maryland
FT = Big Y testing (next generation sequencing)discovered with the Big Y-700, Family Tree DNA, Houston, Texas
G = Verónica Gomes, IPATIMUP Instituto de Patologia e Imunologia Molecular da Universidade do Porto (Institute of Molecular Pathology and Immunol
GG=Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia
IMS-JST = Institute of Medical Science-Japan Science and Technology Agency
JD = David Stedman using Big Y and other NGS sources.
JFS = John Sloan
JN = Jakob Nortsedt-Moberg
K = Youngmin JeongAhn, Ph.D; Education: Seoul National University and the University of Arizona
KHS = Functional Genomics Research Center, Korea Research Institute of Bioscience and Biotechnology
KL = Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai, China
KMS = Segdul Kodzhakov; Albert Katchiev; Anatole Klyosov; Astrid Krahn; Thomas Krahn; Bulat Muratov; Chris Morley; Ramil Suyunov; Vadim Sozi
Science; Prof. Elsa Khusnutdinova, Sc.D. of Biological Sciences, Laboratory of Molecular Human Genetics, Institute of Biochemistry and Genetics, Ufa F
L = Thomas Krahn, MSc (Dipl.-Ing.) formerly of Family Tree DNA's Genomics Research Center; snps named in honor of the late Leo Little
M = Peter Underhill, Ph.D. of Stanford University
MC = Christopher McCown, University of Florida; Thomas Krahn, MSc (Dipl.-Ing.), YSEQ.net, Berlin, Germany

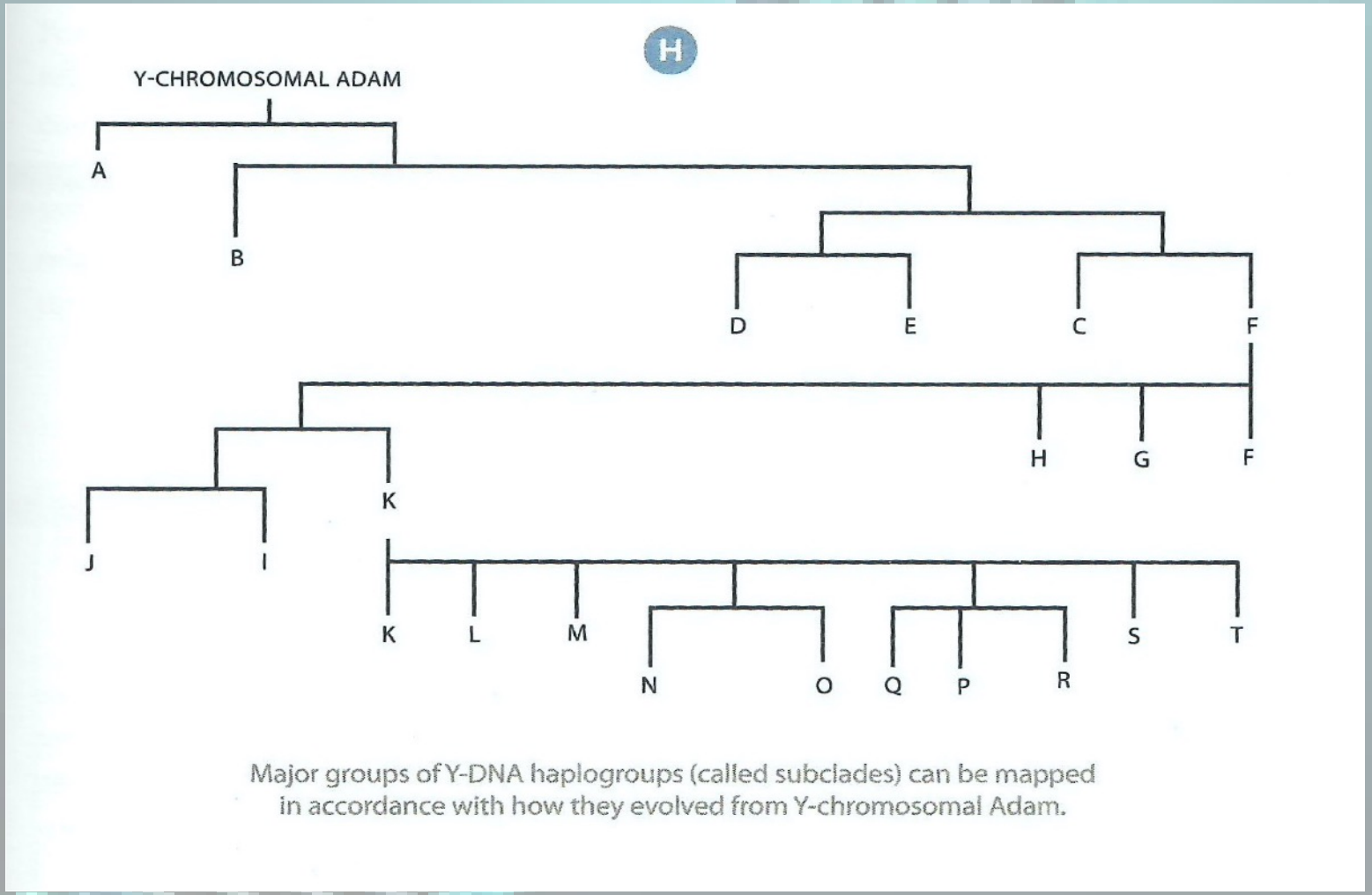
Typical representation of a tree.

from “Genetic Genealogy: The Basics and Beyond” by Emily D. Aulicino



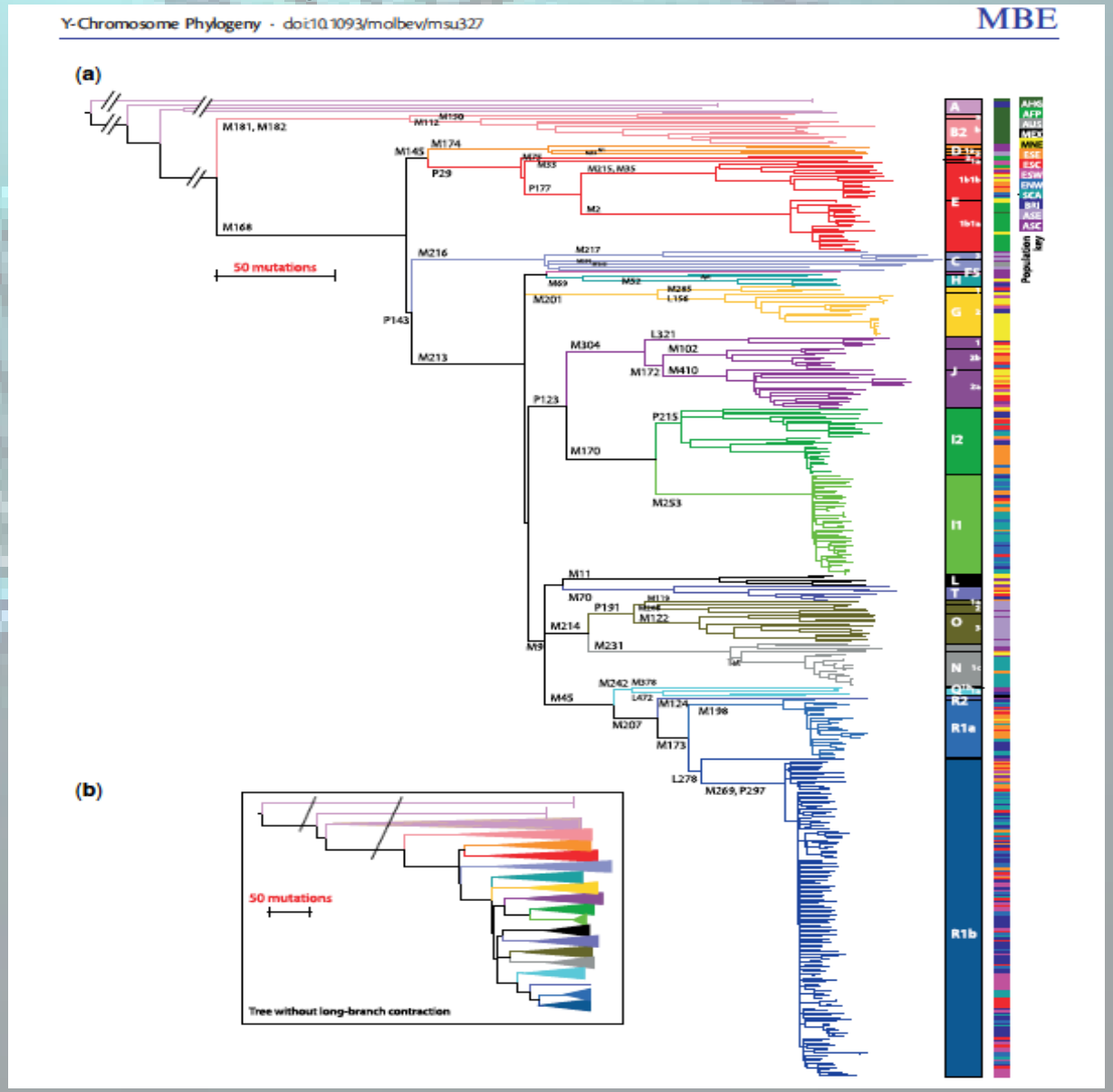
Y Haplotree

From “The Family Tree Guide to DNA Testing and Genetic Genealogy” by Blaine Bettinger



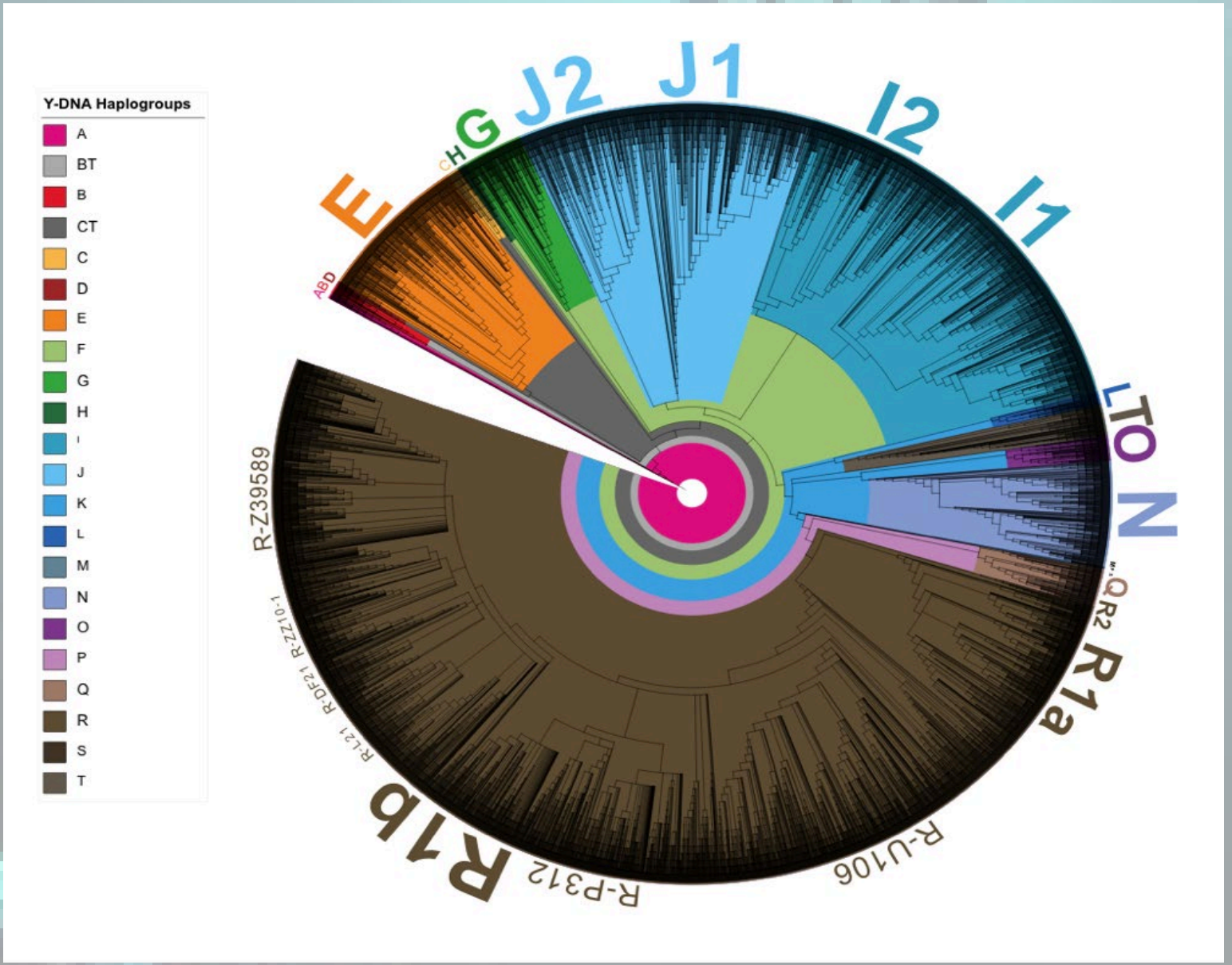
Major groups of Y-DNA haplogroups (called subclades) can be mapped in accordance with how they evolved from Y-chromosomal Adam.

Rectangular Phylogenetic (Haplotree) Tree

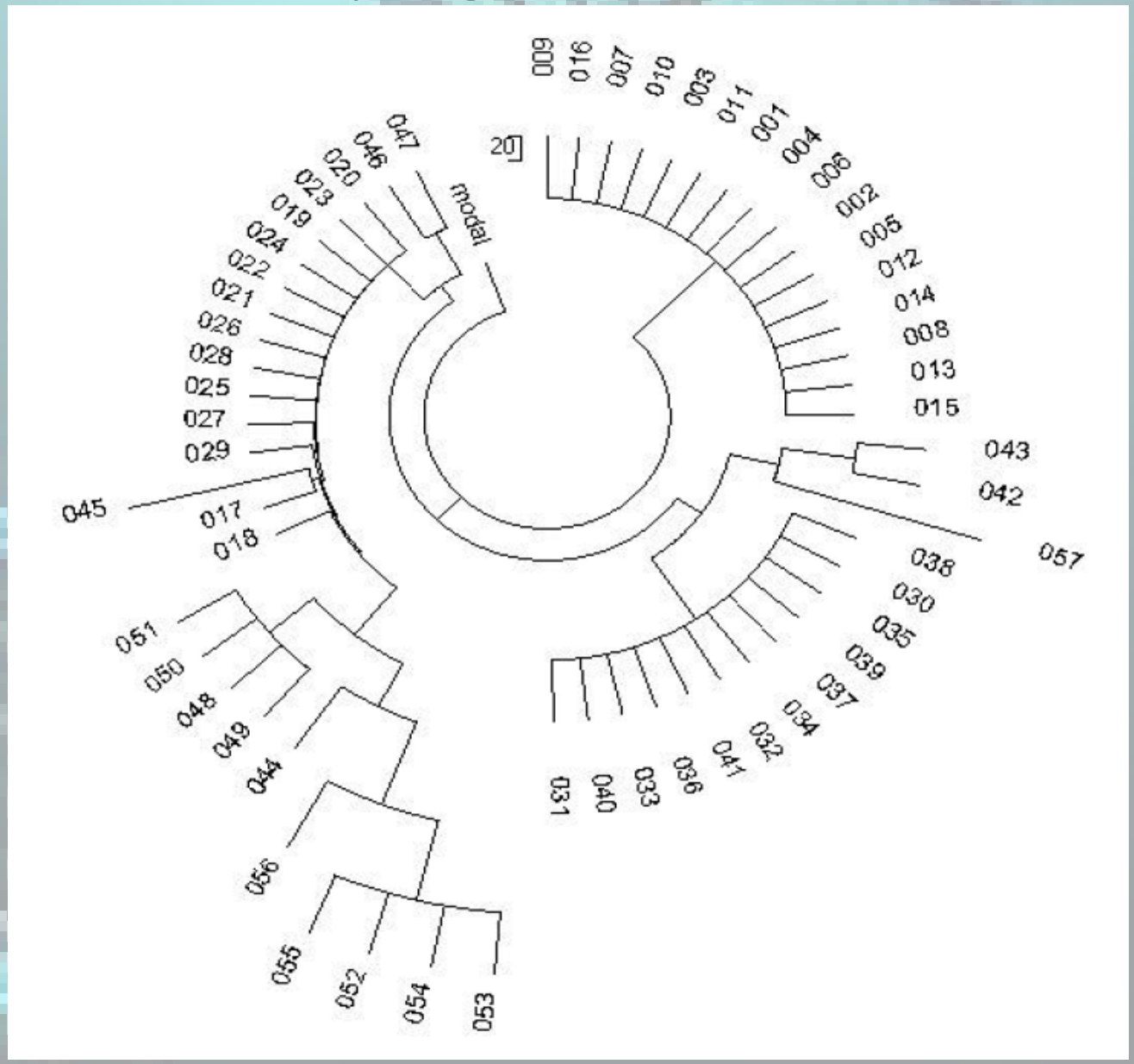


From “The Y-chromosome Tree Bursts into Leaf: 13,000 High Confidence SNPs Covering the Majority of Known Clades (numerous authors); Mol.Biol.Evol. 32(3):661-673.

Circular Phylogenetic Tree (Y-Haplotree) FTDNA blog page

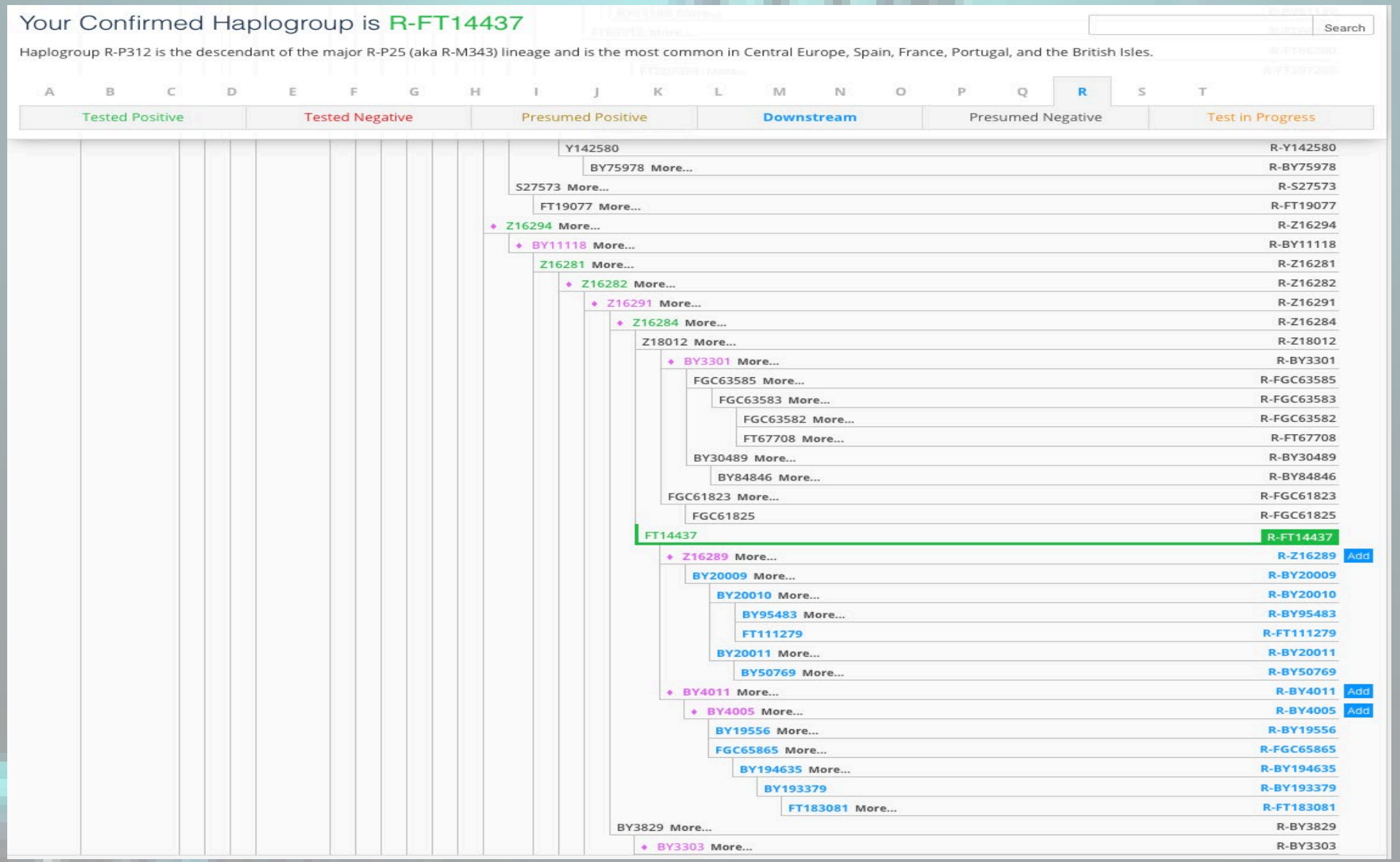


Less Complex Circular Phylogenetic Tree



From Phylogenetic Trees Made Easy" by Barry G. Hall

My Portion of the Y Haplotree from FTDNA



SNP Tsunami Part 2 (Big Y-700)

FT SNP index

From ISOGG Wiki

The **FT SNP index** lists Big Y-700 SNPs that have been named by Family Tree DNA as they have been considered for inclusion to the Y chromosome haplotree. It supersedes the BY SNP index.

- FT1-FT9999
- FT10000-FT19999
- FT20000-FT29999
- FT30000-FT39999
- FT40000-FT49999
- FT50000-FT59999
- FT60000-FT69999
- FT70000-FT79999
- FT80000-FT89999
- FT90000-FT99999
- FT100000-FT109999
- FT110000-FT119999
- FT120000-FT129999
- FT130000-FT139999
- FT140000-FT149999
- FT150000-FT159999
- FT160000-FT169999
- FT170000-FT179999
- FT180000-FT189999
- FT190000-FT199999
- FT200000-FT209999
- FT210000-FT219999
- FT220000-FT229999
- FT230000-FT239999
- FT240000-FT249999
- FT250000-FT259999

- Download Spreadsheet

SNP Tsunami (Part 2 continued)

FT SNPs 250K

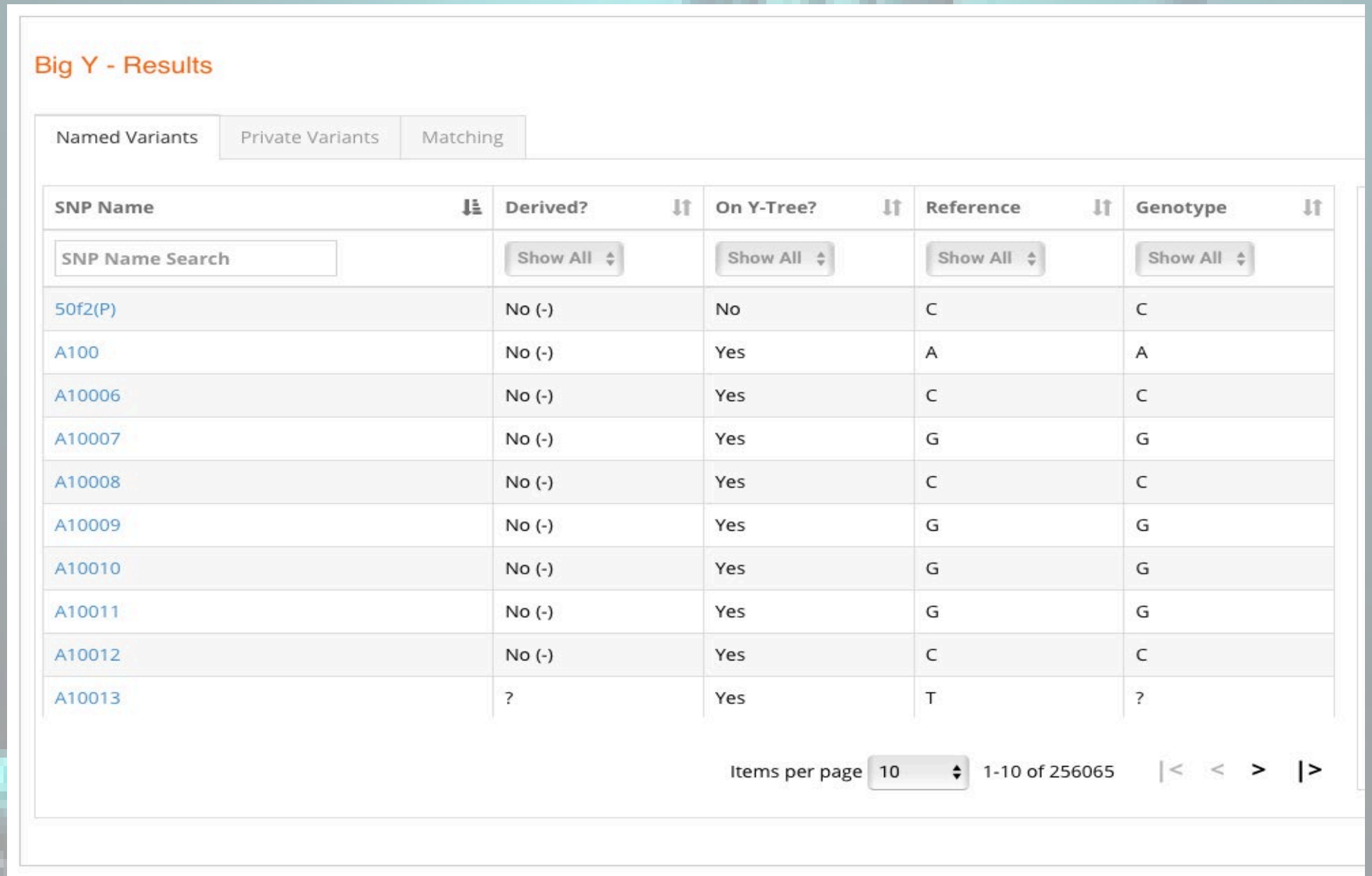
From ISOGG Wiki

The **FT SNP index** lists Big Y-700 SNPs that have been named by Family Tree DNA as they have been considered for inclusion to the Y chromosome haplotree.

Previous List (FT240000-FT249999)

SNP	Position (hg38)	Mutation
FT250000	19555075	G to A
FT250001	19557143	A to G
FT250002	19565666	T to C
FT250003	19573986	T to C
FT250004	19574180	G to T
FT250005	19578239	T to C
FT250006	19579067	T to C
FT250007	19579360	G to A
FT250008	19596519	A to T
FT250009	19598930	G to T
FT250010	19612559	G to A
FT250011	19612815	T to C
FT250012	19633658	C to T
FT250013	19634645	G to A
FT250014	19634961	G to A
FT250015	19636561	A to G
FT250016	19645649	C to A
FT250017	19651822	C to G
FT250018	19654009	C to T
FT250019	19660669	T to C

Named Variants – Showing All



Big Y - Results

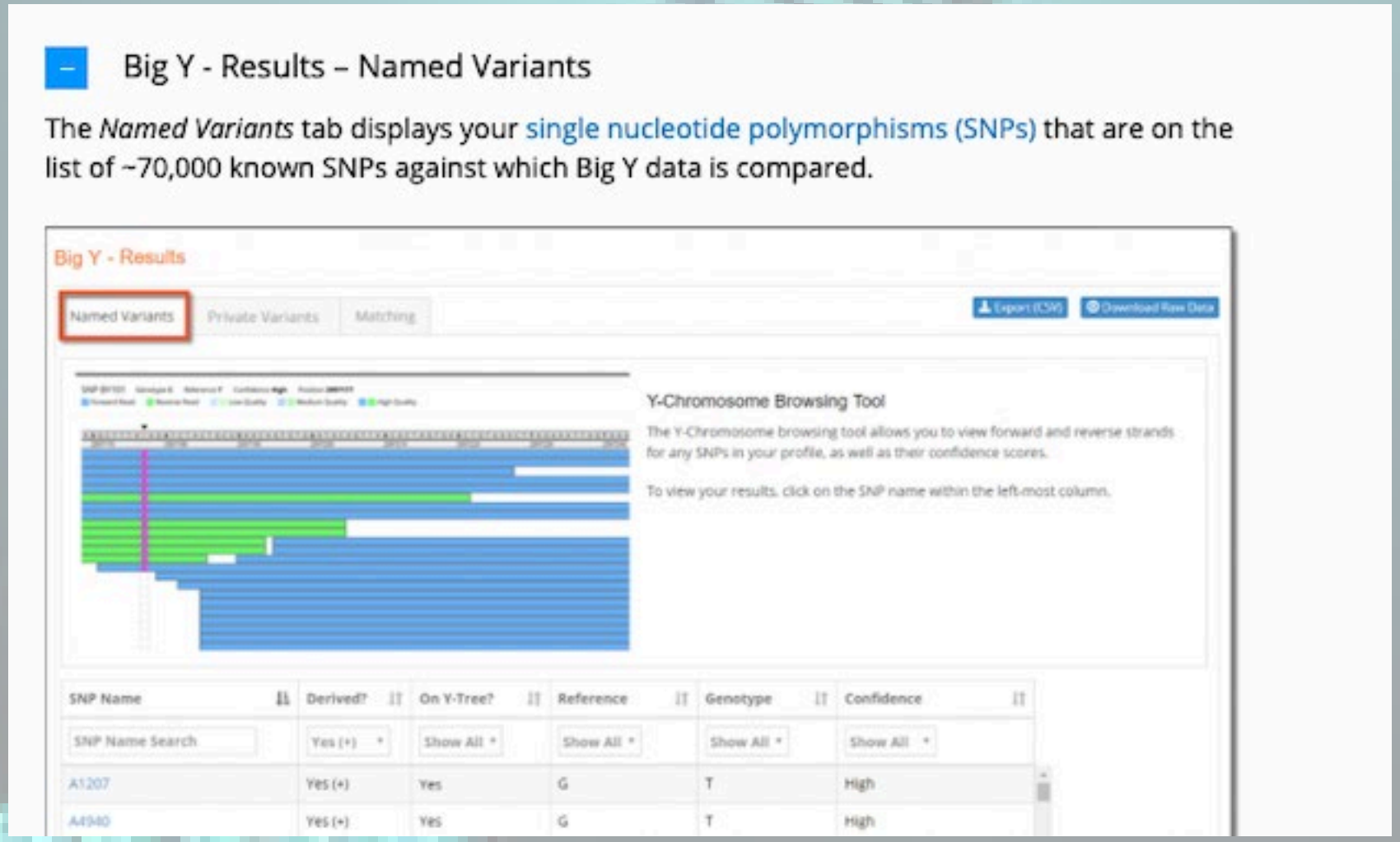
Named Variants | Private Variants | Matching

SNP Name	Derived?	On Y-Tree?	Reference	Genotype
SNP Name Search	Show All	Show All	Show All	Show All
50f2(P)	No (-)	No	C	C
A100	No (-)	Yes	A	A
A10006	No (-)	Yes	C	C
A10007	No (-)	Yes	G	G
A10008	No (-)	Yes	C	C
A10009	No (-)	Yes	G	G
A10010	No (-)	Yes	G	G
A10011	No (-)	Yes	G	G
A10012	No (-)	Yes	C	C
A10013	?	Yes	T	?

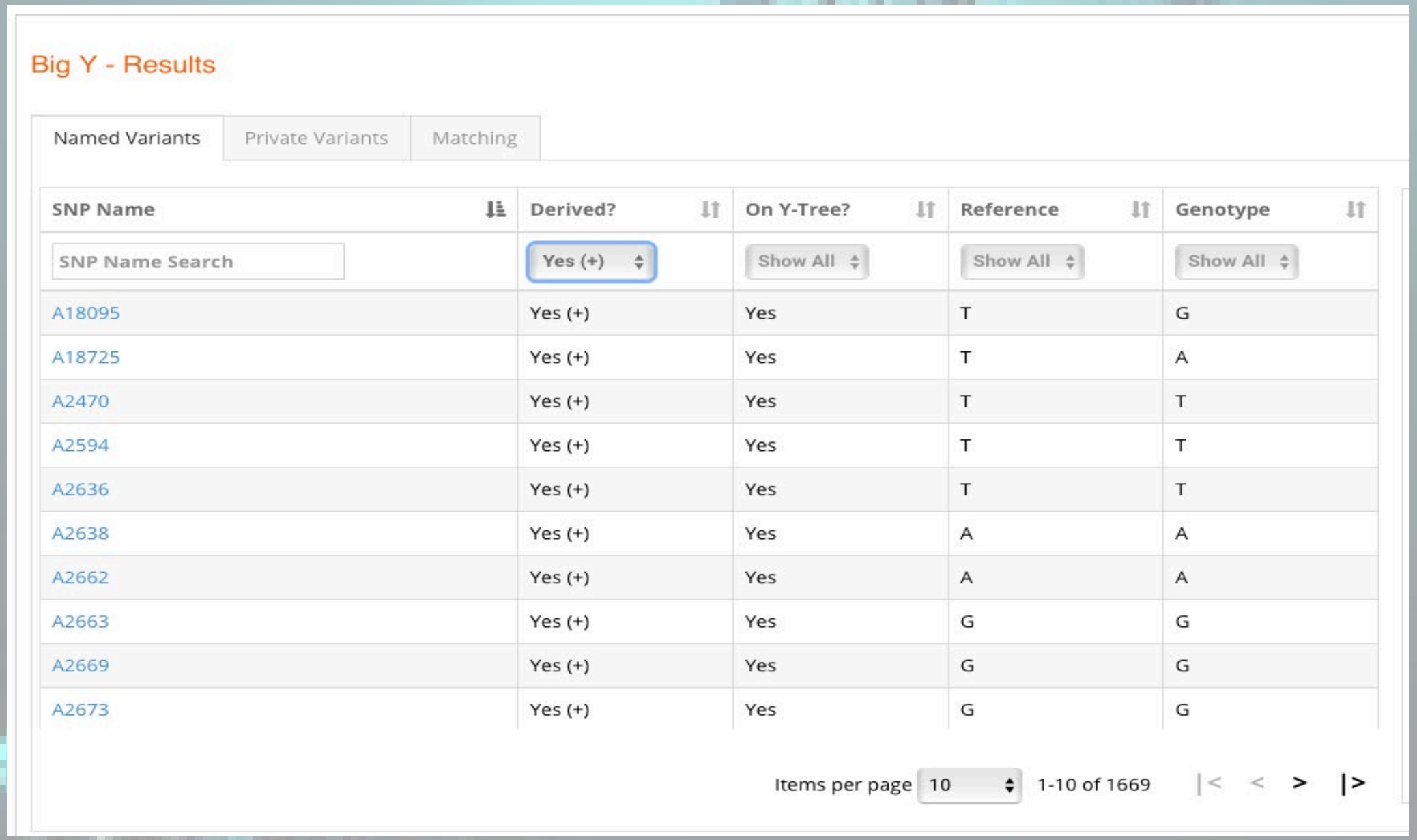
Items per page 10 1-10 of 256065 |< < > >|

FTDNA Learning Center said only ~70K SNPs on Haplotree as of 3/9/2020

Corrected to ~260,000 as of 3/11/2020



Named Variants – Showing Only Those I Have in my Y chr



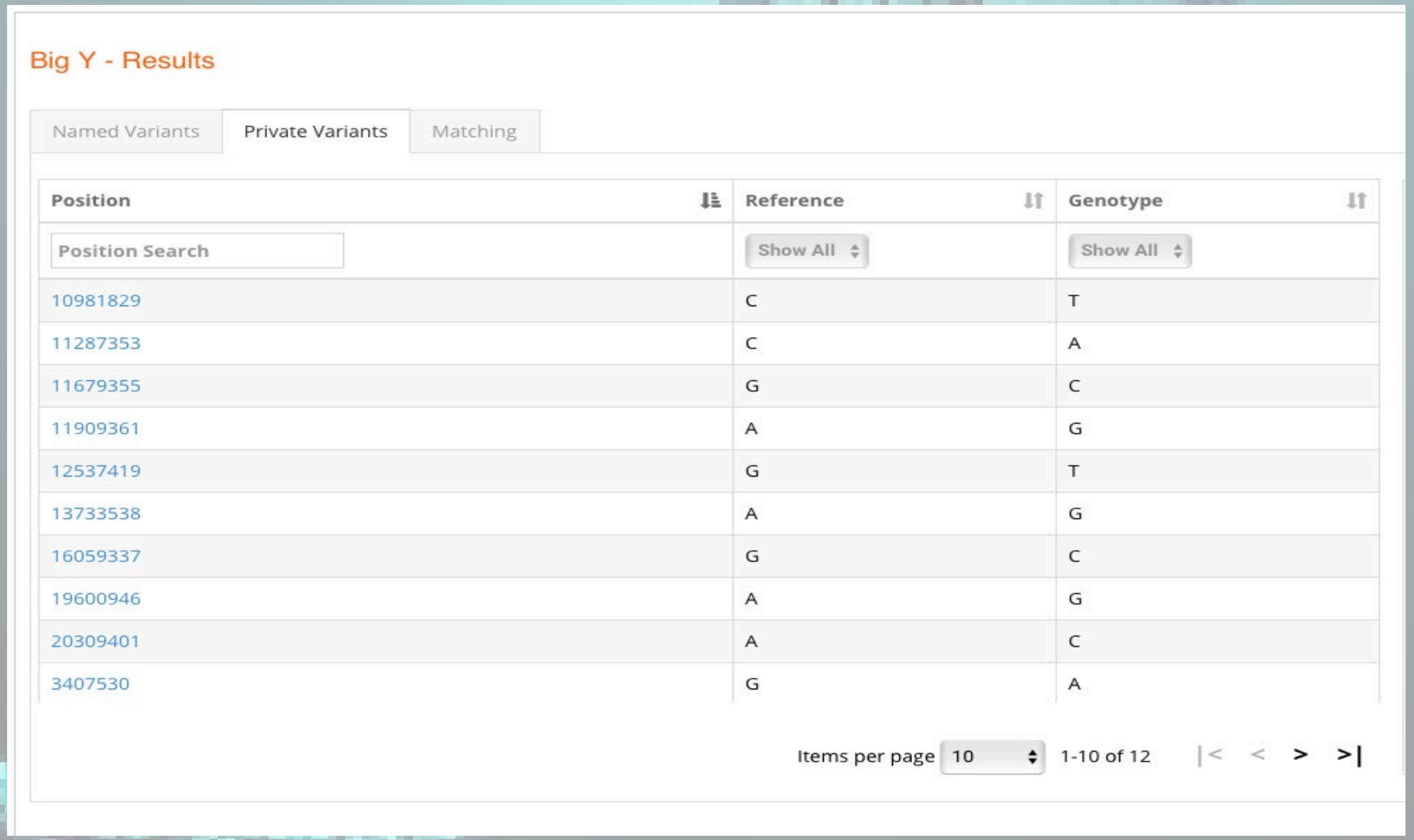
Big Y - Results

Named Variants | Private Variants | Matching

SNP Name	Derived?	On Y-Tree?	Reference	Genotype
SNP Name Search	Yes (+)	Show All	Show All	Show All
A18095	Yes (+)	Yes	T	G
A18725	Yes (+)	Yes	T	A
A2470	Yes (+)	Yes	T	T
A2594	Yes (+)	Yes	T	T
A2636	Yes (+)	Yes	T	T
A2638	Yes (+)	Yes	A	A
A2662	Yes (+)	Yes	A	A
A2663	Yes (+)	Yes	G	G
A2669	Yes (+)	Yes	G	G
A2673	Yes (+)	Yes	G	G

Items per page 10 1-10 of 1669 |< < > |>

My Private Variants



Close up view of Big Y-700 result for position 10981829

You can see the reference genome sequence across the top of the frame. The paired-end segment “reads” (forward and reverse) have been aligned to this sequence. SNP is listed in pink. Coverage depth is about 62X (all reads not shown). Coverage is important as aberrant singleton pink variants are apparent, but since they are not repeated, they are disregarded as spurious.

Partial list of my SNP Chronology from GeneticHomeland.com

27	M269	PF6517 R1b1a1a2 rs9786153	Designates WAMH major European branch of R1b. Arose about 11,000 bce. hg38 Ref does not match ancestral allele value. Britain's DNA labeled this branch: Anatolian. Sometimes labeled as R1b1a1a2 haplogroup in literature using older nomenclature.
28	L23	PF6534 S141 R1b1a1a2a rs9785971	Arose about 6,000 bce. hg38 Ref does not match ancestral allele value. Sometimes labeled as R1b1a1a2a haplogroup in literature using older nomenclature.
29	L51	M412 PF6536 S167 rs9786140	In Europe, almost entirely west of the Danube river. hg38 Ref does not match ancestral allele value.
30	P310	PF6546 S129 rs9786283	Under R1b M269. Believed coincident with P311, CTS 7650, L52, YSC0000082. Changed Ref value per YSeq 11/22/2017
31	L151	PF6542 rs2082033	Under R1b M269 > L23 > L51 according to FTDNA. hg38 Ref does not match ancestral allele value.
32	P312	PF6547 S116 MF52579 rs34276300	Major block under R1b. Arose about 5,000 bce. hg38 Ref does not match ancestral allele value. Britain's DNA labeled this branch: Beaker Folk and formerly Bell Beaker.
33	Z290	S461 rs146019383	Largest branch under under M269
34	L21	M529 S145 rs11799226	Largest European group under R1b P312. Highly correlated with geography of ancient Celts. Britain's DNA labeled this branch: Pretani.
35	DF13	CTS241 S521 rs373989227	Directly under R1b L21.
36	DF21	S192 rs138322855	Under R1b L21 > DF13
37	S5488	Y11277 rs928913967	Branch under L21 DF13 DF21
38	Z16294	rs962837871	Under R1b L21 > DF21 > S5488
39	BY11118		Under R1b L21 > DF21 > S5488 > Z16294
40	Z16281		
41	Z16282	rs908110799	Z16282 marker found in descendant of Charles Carroll, signer of U.S. Declaration of Independence and believed to be descended from medieval Irish kings of Éile O'Carroll.
42	Z16291		
43	Z16284	rs995580140	downstream of DF21
44	FT14437	R-FT14437	

Examples of Haplotree Changes with new SNPs

From FTDNA White Paper on blog site.

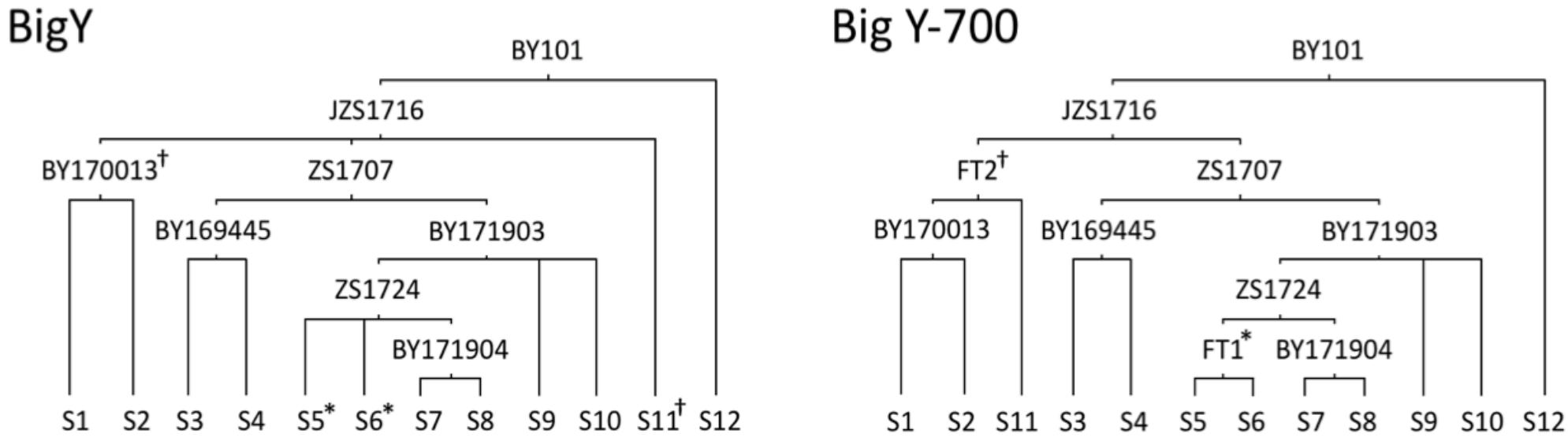


Figure 4: *Novel branch SNPs FT1 and FT2 identified in clade of 11 samples (TMRCA 1-1.5 kya). During placement of the novel branch SNPs, samples S5 and S6 (*) in panel Big Y reorganized under FT1 as shown in panel Big Y-700. Similarly, sample S11 and branch SNP BY170013 (†) reorganized under FT2.*

Other Resources for Y-DNA Analysis

- You can have your Big Y-700 raw data sent to other companies for analysis, such as YFull.com (Russian company), Full Genomes Corp, Y-DNA Data Warehouse, Alex Williamson's "Big Tree" (mostly for R1b at present)
- However, FTDNA currently has the largest database.

Summary

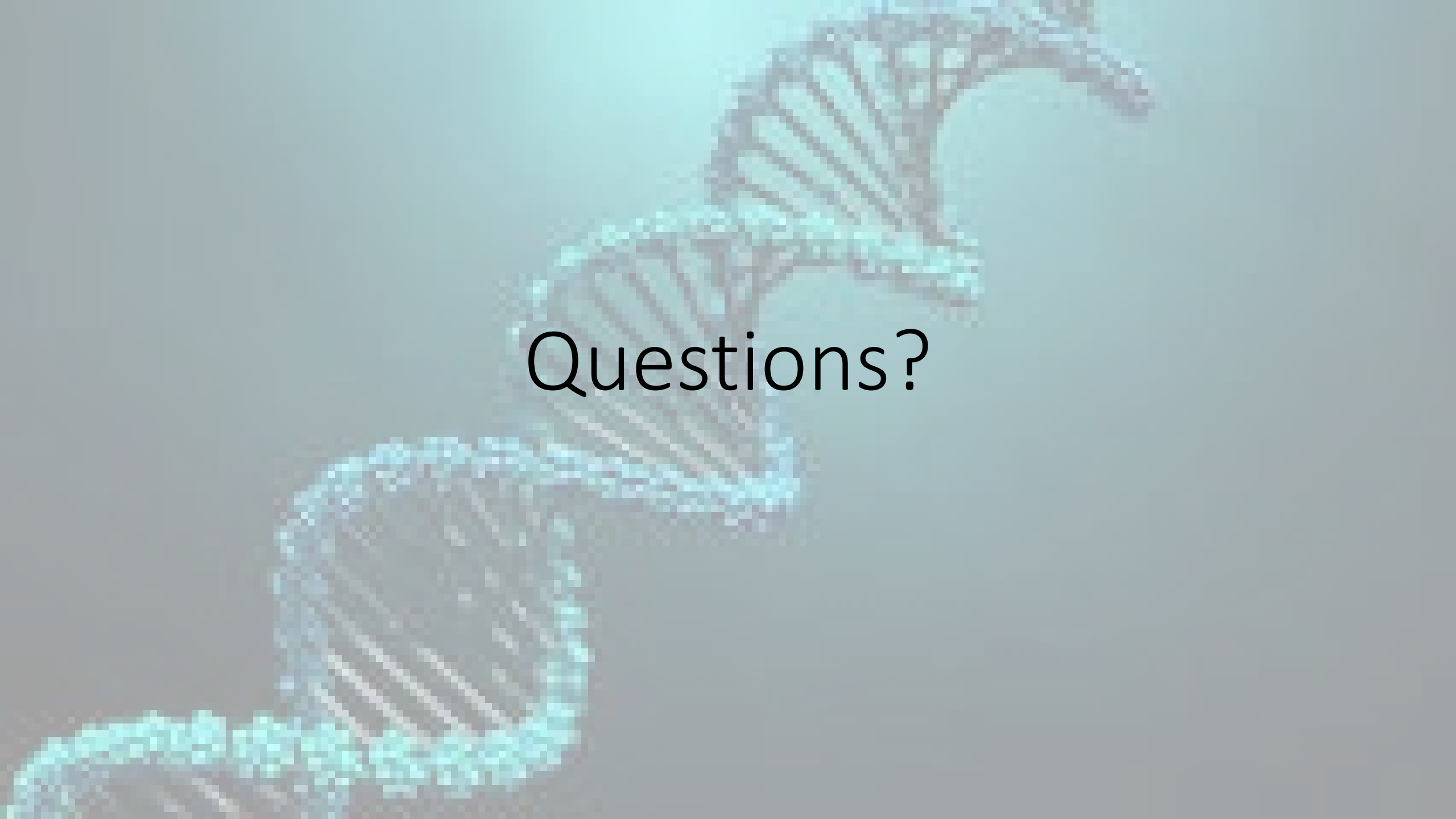
- We discussed features of the Y chromosome/SRY gene.
- We discussed how the Y-DNA is used genealogically.
- We discussed how STRs come to be.
- We discussed features of STRs and how to use the information.
- We discussed how SNPs arise (mutations/misincorporations).
- We discussed features of SNPs.
- We discussed the haplotree.

Useful Resources

- YouTube videos
  - Dr. Maurice Gleeson
    - How Y DNA can Help Your One Name Study
    - Research Your Clan Using DNA & Documentary Records
    - Using Y DNA to Research Your Surname
  - Dr. Michael Sager
    - The Tree of Mankind
  - Dr. Iain McDonald
    - Big Y-700: the cutting edge of Y-DNA testing
- Books
  - Tracing Your Ancestors Using DNA (Y DNA chapter by Alasdair F. Macdonald & John Cleary)
  - The Family Tree Guide to DNA Testing and Genetic Genealogy by Blaine Bettinger
  - Advanced Genetic Genealogy edited by Debbie Parker Wayne

Useful Resources 2

- A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups by the Y Chromosome Consortium, Genome Research, www.genome.org


Questions?

Using STRs to estimate ages

Each STR has its own mutation rate

CDY is probably the fastest, at one mutation per 15 generations (~525 years)
Many mutate slower than once per 10,000 generations (~350,000 years)
By totalling the mutation rates, we can estimate how long before any marker will mutate.

Test	Average time between mutations
Y-12	1400 years/mutation
Y-25	560 years/mutation
Y-37	270 years/mutation
Y-67	200 years/mutation
Y-111	120 years/mutation
Y-500	~86 years/mutation
Y-700	~39 years/mutation

(Still working out exact rates for new STRs)

Example:

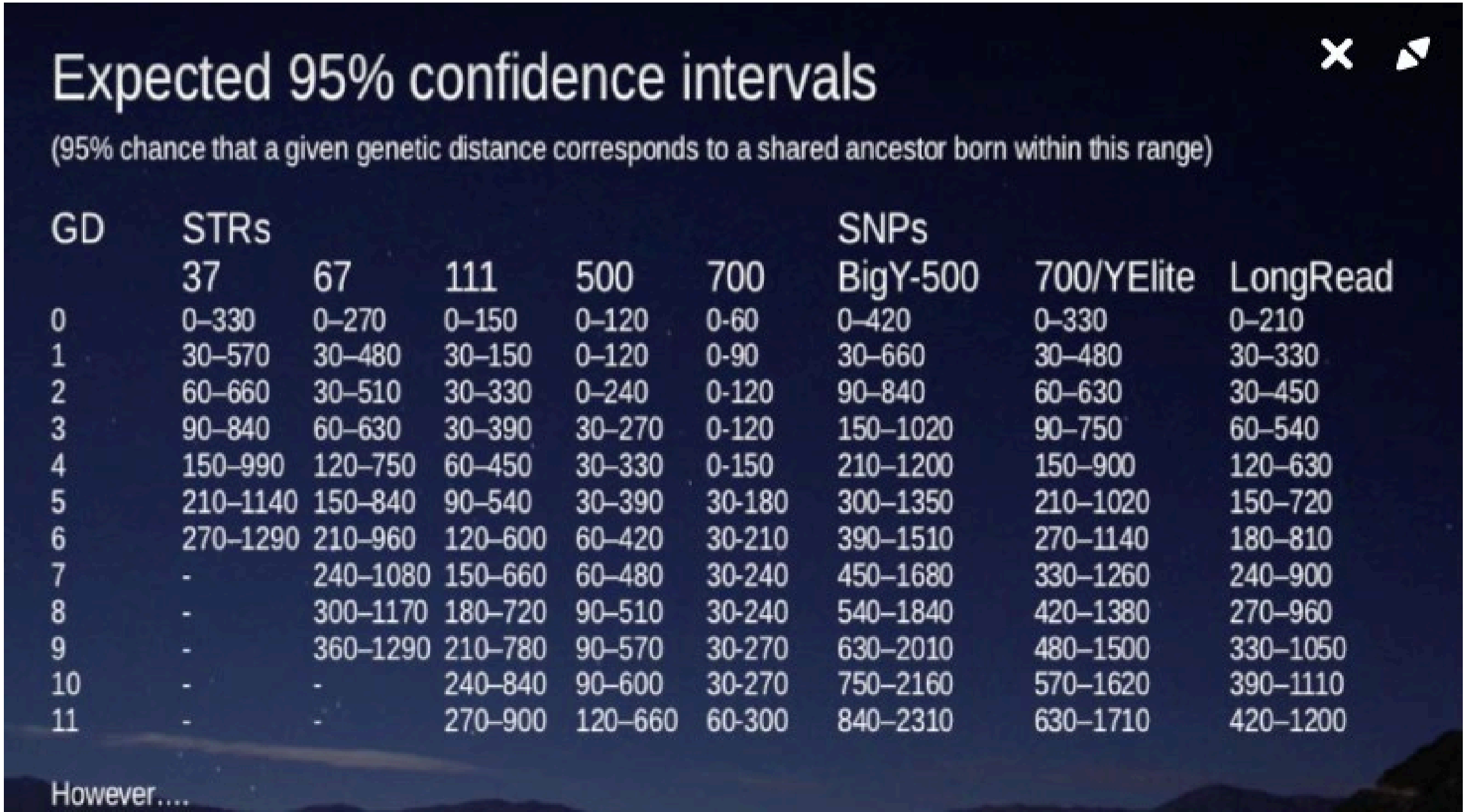
6	5 of 438	[redacted] Donald Y-DNA111 Big Y

Two tests have a genetic distance of 6 in 111 markers.
→ 3 STR mutations on average
Time since their most-recent common ancestor (TMRCA) =
3 mutations * 120 years/mutation = 360 years.

Also, genetic distance of 11 in 552 markers.
→ 5.5 STR mutations on average
TMRCA = 5.5 mutations * 86 years/mutation = 473 years.

From Dr. Iain McDonald's video "Exploring New Y-DNA Horizons with Big Y-700

From Dr. Iain McDonald's video "Exploring New Y-DNA Horizons with Big Y-700



GD	STRs					SNPs		
	37	67	111	500	700	BigY-500	700/YElite	LongRead
0	0–330	0–270	0–150	0–120	0-60	0–420	0–330	0–210
1	30–570	30–480	30–150	0–120	0-90	30–660	30–480	30–330
2	60–660	30–510	30–330	0–240	0-120	90–840	60–630	30–450
3	90–840	60–630	30–390	30–270	0-120	150–1020	90–750	60–540
4	150–990	120–750	60–450	30–330	0-150	210–1200	150–900	120–630
5	210–1140	150–840	90–540	30–390	30-180	300–1350	210–1020	150–720
6	270–1290	210–960	120–600	60–420	30-210	390–1510	270–1140	180–810
7	-	240–1080	150–660	60–480	30-240	450–1680	330–1260	240–900
8	-	300–1170	180–720	90–510	30-240	540–1840	420–1380	270–960
9	-	360–1290	210–780	90–570	30-270	630–2010	480–1500	330–1050
10	-	-	240–840	90–600	30-270	750–2160	570–1620	390–1110
11	-	-	270–900	120–660	60-300	840–2310	630–1710	420–1200