A Review of the Y-DNA Hubbell/Hubble Surname Project Incorporating New SNP Testing Results

by Richard W. Hubble
2020
Revision 2

Introduction:

This paper presents the results of the 2019-2020 Y-DNA genetic testing performed for the Hubbell/Hubble Family Surname Project. The first section, "Behind the Testing", explains the basic biological processes involved in genetic testing, covering testing concepts and terminology that will be used throughout the paper. Section II presents and discusses the results and implications of the testing data.

What do we hope to discover from a DNA test? There are three major areas of inquiry that can be addressed. The first is whether an individual with the Hubbell name (or any of its various spellings) is descended from Richard Hubball the Immigrant, our common English ancestor who immigrated to the New Haven Colony (CT) around 1640. The second goal is to potentially identify the origin of the Hubball name in England. Is the name derived from Hubbal Grange in Tong Parish; from the Huband family of Ipsely, Warwickshire or is there another as yet undiscovered origin? The third goal is to identify our pre-English roots. Does the family descend from the Vikings (Scandinavian or Danish) or from a Gallic/French origin? Paper records are obscure or non-existent regarding these questions and DNA testing may prove helpful in resolving them.

I) Behind the Testing-Understanding the Biology:

Genetic testing for the general public began around 2000. The original DNA tests were simple screening tests that were not robust enough to conclusively prove that two people shared a recent common ancestor. The original test provided values on only twelve short tandem repeats (STRs). Over the years the tests have improved dramatically. With today's tests (Big Y-700 from Family Tree DNA), 111 STR values are reported along with hundreds of mutations called single nucleotide polymorphisms (SNPs). STRs and SNPs will be discussed, in detail, later in this review. Of course, DNA testing will never replace the paper record and, in-fact, the true power of genetic testing comes with marrying the paper records with the DNA results.

With the new and more powerful tests, the demand for consumer genetic testing is growing explosively. According to a recent MIT Technology Review 26 million people, or more, have taken a genetic ancestry test. The review found, in 2018, the number of tests purchased surpassed sales of all previous years combined. Unfortunately, there are now multiple types of DNA tests offered by multiple vendors resulting in competition between vendors. This has resulted in little or no data sharing between the vendors, leading to a fragmented testing population and increasing the difficulty of assembling comparative databases.

To be able to intelligently choose which DNA test to take and to understand the full scope of the results one will receive will require some basic background information. Presented below are some basic biological concepts to assist the reader.

The Cell and DNA:

Almost every cell in your body contains a structure called the nucleus (Figure 1). Inside the nucleus is your nuclear DNA, consisting of 23 pairs of chromosomes, that is tested in the Y-DNA and the at-DNA test. Another structure in almost every cell is the mitochondria which contains the mitochondrial DNA. This is the DNA that is tested in the m-DNA test.

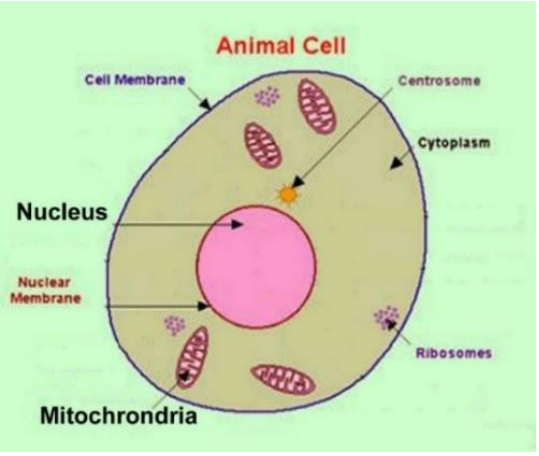


Figure 1: A typical human cell showing the location of the nucleus and the mitochondria along with a few other cellular components. The mitochondria contain m-DNA while the nucleus contains the DNA that is used in the Y-DNA and at-DNA tests.

Each of the 23 pairs of chromosomes inside the nucleus is composed of DNA and supporting proteins. In Figure 2, below, the 23 pairs of chromosomes are illustrated. The 23rd chromosomal pair is the sex chromosome (red circle). For males, the sex chromosome pair is composed of one X and one Y chromosome. For females, the sex chromosome pair is composed of two X chromosomes. The mitochondrial DNA (m-DNA) is composed of a circular DNA strand (not shown).

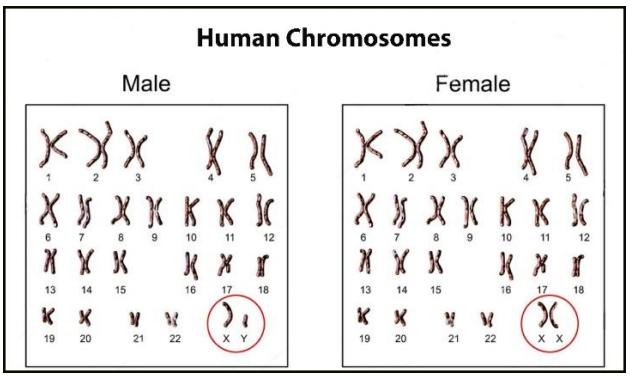


Figure 2: The nucleus contains 23 pairs of chromosomes. The 23rd set of chromosomes (or the sex chromosomes circled in red) in men are composed of one X chromosome and one Y chromosome. The Y chromosome is only present in men. A women's sex chromosome is composed of two X chromosomes. The Big Y700 test screens for mutations on the male Y chromosome only.

Genetic Testing: General Principal:

In principle, genetic testing works by comparing your DNA against another individual's or a large data set. The more closely the DNA matches, the more likely it is that two people are related to each other through a common ancestor. The fewer the matches, the further back in time one must travel to find the most recent common ancestor (MRCA). It is from that common ancestor that the shared DNA segments or regions are inherited.

The Three Tests:

For the genealogist, there are three different types of DNA tests that can be performed. The type of ancestral information one is seeking will determine which test to take.

a) The mitochondrial DNA (m-DNA) test, screens for mutations of the DNA located in the mitochondria and not the nucleus. Because only women pass on the mitochondria, with its DNA, from mother to daughter through the egg, this test screens only for mutations in women. A few Hubbell daughters have had this test performed but the Society has not initiated a formal m-DNA project, so this review will not discuss the m-DNA test in detail.

b) The Y Chromosome DNA (Y-DNA) test, screens only for mutations on the Y chromosome located in the nucleus. Only men possess a Y chromosome, passing it down from father to son. The Y chromosome is especially ideal for genetic testing because the Y chromosome does not undergo recombination during meiosis like the other 22 pairs of chromosomes. The lack of recombination means that mutations occurring in a distant ancestor will be preserved in all his progeny creating a clear link through the generations. This test evaluates SNPs and STRs which will be discussed in more detail below.

c) The autosomal DNA (at-DNA) test, screens for shared DNA segments and SNP mutations on 45 of the 46 chromosomes. This test does not screen for mutations on the sex X chromosome that a female inherits from her father. Nor does it test the Y chromosome that a male inherits from his father (see Figure 2 above). Testing for SNPs on 45 chromosomes would appear to be an advantage over the other tests, however there are limitations. During the production of an egg or sperm, large segments of DNA are exchanged between 22 of the paired chromosomes in a process called recombination. Because of this mixing of the DNA, during meiosis, the autosomal test can only trace ancestors back five to seven generations or approximately 250 years. For more information on recombination during meiosis see Appendix B. SNPs will be discussed later in this review.

The Hubbell DNA Surname Project is a Y-DNA project begun in 2005. Only the details of the Y-DNA test will be discussed in this article at length.

DNA in More Detail:

If we unravel a chromosome, we see that it is composed of a DNA strand (called a double helix) wrapped around supporting proteins called histones (Figure 3).

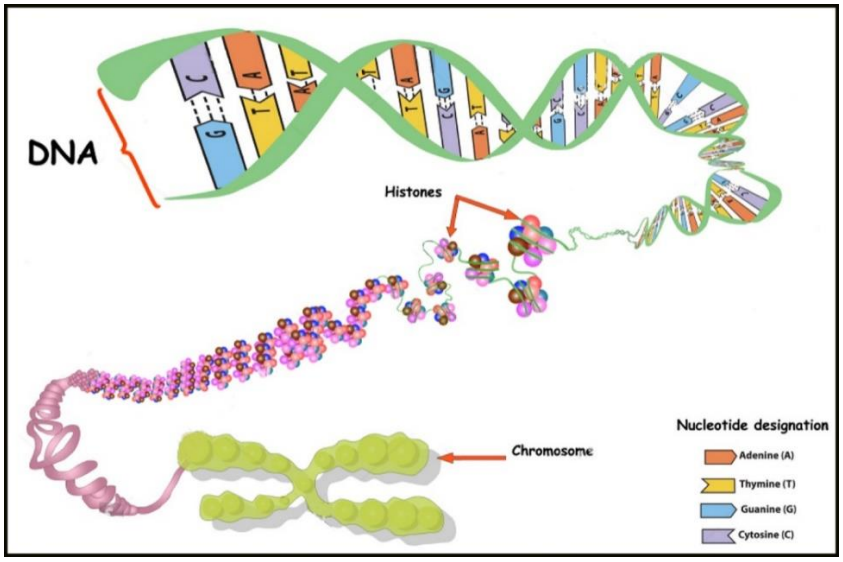


Figure 3: Composition of the chromosome consisting of DNA and supporting proteins called histones.

A closer look at the DNA shows that it is composed of 2 long strands (rails) made of sugar and phosphate, connected together with 4 nucleotides: Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). Thymine will only pair (bind) with Adenine and Guanine will only bind to Cytosine. Each A-T or G-C structure is called a base pair. There are about 3.2 billion of these base pairs in the 23 chromosomes of the human genome.

When illustrating DNA, it is uncoiled to form a ladder shaped structure. Because Thymine (T) will always pair with Adenine (A) and Guanine (G) will always pair with Cytosine (C), there is no need to display both sides of the ladder structure. The information is displayed as a single line representing one side of the ladder with the letters A, T, C and G representing the four nucleotides.

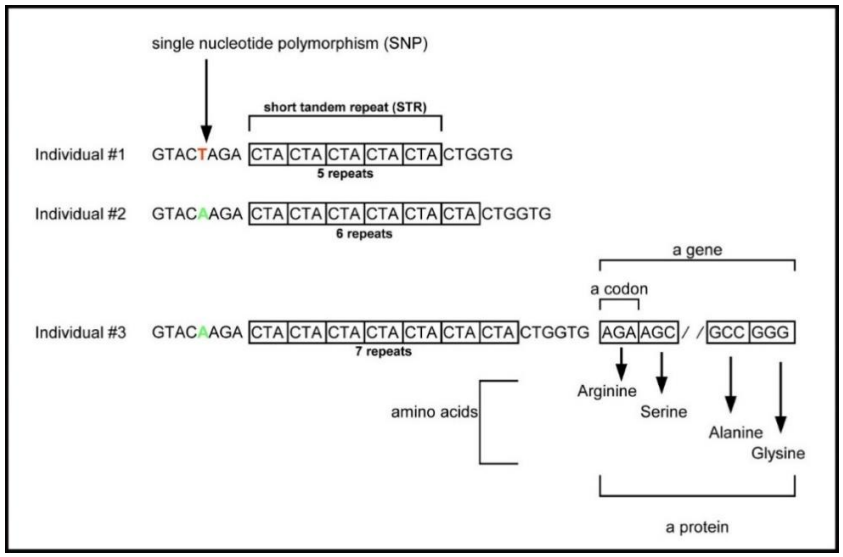


Figure 4: A representation of a segment of DNA. The picture depicts both types of mutations that the Y-DNA test screens for: SNPs (single nucleotide polymorphisms) called "snips" and STRs (short tandem repeats) called "stirs". Also depicted is the structure of a codon and gene that encodes the information to produce proteins that build our body.

The main function of DNA is to code for proteins. This is done through a complicated cellular mechanism that can translate a three-nucleotide sequence into the building blocks of proteins called amino acids. In Figure 4, above, the nucleotide sequence AGA (Adenine-Guanine-Adenine) is the DNA code for the amino acid Arginine. Stringing amino acids together produces a protein; a protein like hemoglobin, a protein that binds oxygen in our blood and carries it to cells throughout your body. A mutation in the DNA coding region of hemoglobin could be fatal to an individual.

However, there are many regions on the human chromosome (and especially the Y-chromosome) that do not code for vital, life sustaining proteins. It is these regions, regions that have been called "junk" DNA in the past, that contain the mutations we are tested for.

The Y-Chromosome:

As pictured in Figure 2 above, humans have 46 chromosomes (or 23 pairs). In men, the last chromosome, (46th chromosome) is known as the Y-chromosome, sometimes called the Y-DNA. The Y chromosome contains approximately 59 million base pairs with mutations spread along its length. The Y-DNA has very few genes that code for proteins, but it has one important gene called the SRY gene. This master switch gene (which activates other genes) converts a human embryo into a male during development.

Two Types of Mutations:

1) Single Nucleotide Polymorphisms (SNPs)

In the early years of DNA testing, it was believed that SNPs were mutations of the DNA that occurred as infrequently as once every 1000 years. Because of this infrequent occurrence, genealogical testing, in the early years, focused solely on STR testing to provide information on relatedness.

SNPs were originally used for identification of very distant historical correlations (>1000 years). However, with new and better testing technologies developed within the last five years, the accepted occurrence of a SNP mutation has been revised down to approximately 83 years. Because SNP mutations are very stable, they have become more useful in identifying differences in the genome than STR testing.

Improvements in testing technology were the driving force behind the discovery that SNPs occur more frequently than once every thousand years. These new improvements allow more expansive tests that screen regions of the DNA that older tests were unable to read reliably. These new and improved tests are approaching the ideal tool for use in genetic genealogy as they, potentially, allow a resolution of only two or three generations between mutations. This means that a grandfather and grandson, or two paternal cousins could show at least one different SNP on their Y chromosome. As a result, it becomes extremely easy to determine how many generations ago two lineages diverged from one another.

When any point mutation or SNP occurs, it is passed down to all future male generations. Although it is not known for certain, it is generally believed that a point mutation, or SNP, is a permanent mutation and the probability of it mutating back to the original or ancestral nucleotide is extremely low. Some SNPs are proven to have occurred many thousands of years ago and have remained stable over the centuries.

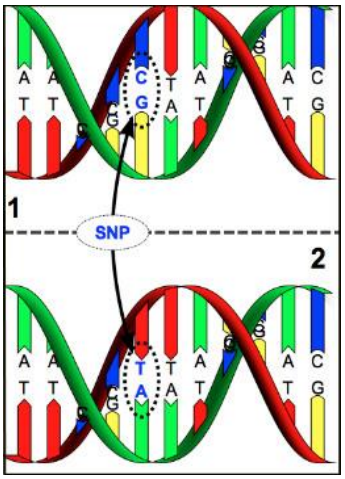


Figure 5: An example of a SNP or point mutation. When sequencing identifies a nucleotide pair that is different from the reference standard it is called a point mutation or single nucleotide mutation. The picture on the left depicts the point mutation A-T (picture #2 below) that is different from the reference standard G-C (picture #1 above)

A new SNP is identified when the nucleotide sequence in the tested DNA sample is different from a reference standard at a single nucleotide. There are thousands of these unique point mutations that have been discovered at this early stage of DNA testing. Mutations that occurred thousands of years ago are found in a large percentage of the population while newer mutations have been identified in only a few individuals. If the mutation is found in only one individual, it is not named but just identified as a "personal or novel variant". The challenge is how to organize all of these SNP mutations. To accomplish this a Y-DNA Haplotree was developed.

The Y-DNA Haplotree:

The construction of the Y-DNA haplotree and its component haplogroups is a mechanism for organizing the thousands of SNP mutations based on the suspected age of the mutation. For more than 20 years, researchers from around the world have been sampling the remains of ancient skeletons and indigenous populations, in an effort to map historical human migration patterns based on the presence of SNPs. From this effort the genetic genealogy and anthropology communities have amassed a large database of single point mutations. By comparing and dating the occurrence of SNPs, researchers have been able to construct a timeline of approximately when and where a mutation occurred. To help organize and display the relationships between these SNPs, a tree structure has been constructed that represent the timeline of SNP occurrence. This effort, began in 2002, has been slowed by a lack of standardization of the different testing technologies and nomenclature.

Each Y-DNA haplogroup has a unique set of SNPs (or markers) that define that haplogroup. Every member of a single haplogroup bears the **same** unique set of SNP Y-DNA mutations which sets them apart from all other haplogroups. Each of these unique markers arose in a single individual, the haplogroup ancestor, a long time ago and has been propagated down through the millennia through his progeny. Thousands of SNP mutations have been discovered.

At this writing (2020), there are approximately 29 major haplogroups in the backbone of the Y-DNA haplotree (see Figure 6). Most of these major haplogroups contain hundreds of additional SNPs that branch out to modern times. Some haplogroups, like the haplogroup A000 that defines the Neanderthal, have become extinct.

All modern human men are descended from Y-Adam an individual that lived in Africa approximately 240,000 BCE (1). As Adam's progeny acquired mutations, they became subdivided into new haplogroups defined by that new mutation. For example, some individuals in the A0-T haplogroup (defined by the mutation L1085) acquired the SNP mutation CTS2809.1, also known as L991.1. These individuals were assigned to a new haplogroup, A0. Other individuals from the A0-T haplogroup acquired a different mutation (P305) and were placed into a different haplogroup, A1.

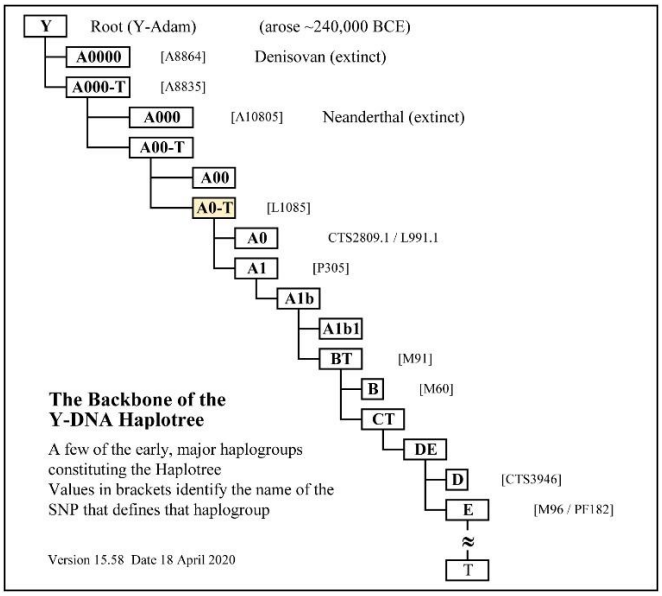


Figure 6: The backbone of the Y-DNA haplotree. The further down the tree, the more recent the mutation. Most haplogroups contain hundreds of SNPs that branch out to modern times. (from the isogg.org website)

(1) Anthropologists use the nomenclature BCE and CE (before the common era and common era) to represent years before the birth of Christ (BCE) and after (CE). This convention has been adopted in this review.

Certain peoples tend to be strongly associated with specific haplogroups due to religious, marital, and social practices. For example, the Y-DNA haplogroup known as Q-M242 (not shown above) is strongly associated with people of Native American descent. The SNP that defines this haplogroup is believed to originate from Central Asia-Siberia approximately 22,000 BCE. Some of the people carrying this mutation likely migrated across the Bering Strait into North America, were it was passed on to their children. Today, this SNP is present in 92.3% of Navajo and 87% of Apache populations. However, while Q-M242 is extremely prevalent in Native Americans, it can also be found in northern Thailand and Indonesia. This suggests that both populations are derived from a common ancestor, living in Central Asia-Siberia, who possessed the Q-M242 SNP.

In 2012, the major and minor branches of the haplotree were defined by about 12,000 SNPs. By 2016, this number was greater than 36,000. Since 2016 the pace of SNP discovery has accelerated adding thousands of new mutations. This scientific onslaught has been termed the "SNP tsunami". With so many SNPs it is difficult to view the haplotree in its entirety. The International Society of Genetic Genealogy maintains an electronic version on their website at: https://isogg.org/tree/index.html

2) Short Tandem Repeats (STRs): (2)

The Y-chromosome contains sequences of DNA called STRs (short tandem repeats) that repeat a sequence of nucleotides multiple times. See Figure 4 for an example of a STR that repeats the nucleotides C, T and A, five, six and seven times. In early studies, it was shown that individuals related to each other shared the identical number of repeats on a number of STRs. The closer the shared ancestor the more STRs showed identical repeats.

When the Hubbell Y-DNA Surname Project began in 2005, the only available test, for the general public, was a test that screened only 12 STRs. STR testing today is performed on 37, 67 or 111 STRs. The genetic testing company Family Tree DNA (FTDNA) has issued guidelines stating that if two men match the same number of repeats on 33 or more of 37 STR markers (90%), they are related to each other within a genealogical time frame (3).

At this early date (2005), SNPs were still believed to occur approximately every 1000 years and were not thought to be relevant for genealogical testing.

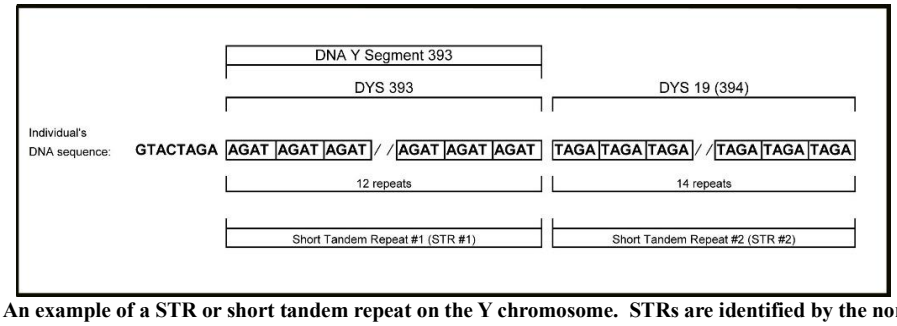


Figure 7: An example of a STR or short tandem repeat on the Y chromosome. STRs are identified by the nomenclature DYS XXX. DYS stands for DNA Y Segment. The number that follows is the location (or address) on the chromosome where the STR is located

The first DNA sequence, pictured above in Figure 7, has the format AGAT-AGAT-AGAT. This DNA sequence, "AGAT", is called a STR marker. This marker (sequence) is repeated 12 times in this example. Researchers have a STR naming convention called the DYS System. DYS stands for **D**NA **Y** **S**egment. The AGAT marker, above, is known as DYS393. The number that follows DYS (in the above example 393) is the location or "address" of the STR on the Y chromosome. In this example the STR marker is repeated 12 times. The number of repeats is represented as "DYS393 = 12".

All modern human males have the same set of DYS markers which are situated in the same order along their Y-chromosome. The difference between individuals is the number of repeats of each of these DYS markers. The reason that all modern males have the same DYS markers is that all human men share Y-Adam as a common distant ancestor (see Figure 6).

_____

2) Originally thought to be "junk" DNA or DNA left over from our evolutionary past, STRs are now recognized as having some biological functions but at this time those functions are not well understood.

3) The genealogical timeframe is the period in which it is possible to find genealogical records relating to individual ancestors which allow the researcher to construct family trees. Definition from the International Society of Genetic Genealogy (isogg.org).

The STR Y-DNA test looks at the number of repeats (values) in a set of DYS markers between any two men. If two individuals share enough identical DYS marker values, they are grouped together in a unique cluster. This set of DYS markers, or clusters, is called a haplotype. Groups of men in the same haplotype can share a common paternal ancestor. How closely they are related depends on the number of identical markers they share.

However, the number of DYS repeats, within a marker, can change (mutate) through the generations. Additionally, some of the markers are more prone to change than others. A father and son should match on all known DYS markers. This is theoretically true since a son inherits a copy of his father's Y-DNA. However, it is possible that even a father and son may differ in a DYS marker value due to a mutation.

To further complicate matters, STRs can mutate back and forth. In the example pictured in Figure 4, it is unclear which of the repeat values came first. Individual #1 has the marker CTA repeated 5 times, but CTA is repeated 6 times in Individual 2 and 7 times in Individual 3. If these three individuals are related to each other, which is the ancestral value? Was it first a 5 that mutated to a 6 in Individual #2 and then to a 7 in Individual #3? Or was it first a 6 that mutated down to a 5 in Individual #1 and then up to a 7 in Individual #3. There is no way to tell based only on this information. Furthermore, STRs can mutate quickly. A STR can mutate up in one generation and back down in another generation making two men look more (or less) closely related than they really are.

Despite STRs instability, they are good predictors of how closely two people are related to each other. The more shared DYS markers, the more likely two people are related. Due to the low cost of performing a STR test, it can be used to initially screen a group of men, who, for example, share the same last name, or similar last names (Hubbell vs Hubble) to ascertain how closely they are related.

Genetic Distance (GD)

STR based relatedness is measured by calculating the genetic distance (GD). The best approach in comparing marker values and relatedness between individuals is to determine the ancestral DNA profile and then compare individuals to this standard. Genetic distance is calculated as the sum of the differences for each marker between the ancestral value and the individual. If an individual deviates from the ancestral value by three repeats, then the genetic distance for that marker is three.

Richard Hubball, the Immigrant, is the progenitor of the Hubbell/Hubble family in North America but a sample of his DNA is not available. Reconstructing his DNA STR profile can be done based on the dominant values of current Hubbell/Hubbles. To determine his profile, all Hubbell/Hubbles who have taken a STR test and are listed in the Millennium Edition of the family genealogy were selected (n=16). The predominant value, for each STR, is designated as the ancestral value. It is more likely that the predominant STR values were passed from Richard the Immigrant to his progeny and the less common values are more recent mutations. In the Table 1 below, the Hubbell Ancestral Value has been determined for the first 12 STR markers (in green).

A large majority of Society members share the same 12 STR values, however, a few show some differences. In the example below, Hubble #1 (who is a descendant of Richard the Immigrant) has two STR values that are different from the ancestral values at DYS 385a and at DYS 392. This member differs by only one value at each STR resulting in a Genetic Distance of 2. With the ancestral value information, all individuals of unknown lineage to Richard the Immigrant, but who carry the Hubbell/Hubble name, can be compared.

	Genetic		STR Markers											
Name	Distance	Haplogroup	DYS393	DYS390	DYS19	DYS391	DYS385a	DYS385b	DYS426	DYS388	DYS439	DYS389i	DYS392	DYS389ii
ancestral value		R-M269	13	25	15	10	11	14	12	12	12	13	13	27
Hubble #1	2	R-M269	13	25	15	10	12	14	12	12	12	13	12	27
Hubble #2	14	I-M253	13	22	14	10	13	14	11	14	11	12	11	28
Hubble #3	17	J-M172	12	23	15	9	13	16	11	16	12	13	11	29

Hubble #2: This Hubble was born and grew up in the London area and now lives in the US
Hubble #3: This Hubble lives in New Zealand and notes his family's origins as England

Table 1: STR marker values for the first 12 STRs showing the ancestral values for the North American Hubbell family. Individuals showing differences from these values are shown in red. The goal of STR value calculations is to assign an individual a haplogroup designation. Members of a haplogroup share a recent common ancestor.

In Table 1, two individuals are presented that have a Hubble surname but show large genetic distances (GD) of 14 and 17 from the ancestral values. The participant labelled Hubble#2 has a GD of 14. He matches the ancestral value for the first STR (13) but for the second STR, DYS390, the ancestral value is 25 and Hubble#2 has a value of 22. At this STR he has a difference of 3. At DYS19 the ancestral value is 15 and Hubble#2 has a value of

14 or a GD of 1. Adding up all of the GDs, Hubble #2 has a total genetic distance of 14 for the first 12 STRs. This large genetic distance from the North American ancestral values clearly suggests that both Hubble #2 and Hubble #3 are not related to Richard the Immigrant.

The testing company Family Tree DNA (FTDNA), using a computer algorithm, can assign individuals a position on the haplotree based on their STR values. This placement is not absolute but a calculated probability. To validate this probability, SNP testing must be done. For example, from Table 1 above, using the ancestral Hubbell values of the first 12 STR markers, there is a 99% probability that Richard the Immigrant and the members who shared the same STR values belong to the R-M269 haplogroup. The Hubble#1, with a genetic distance of 2, is also placed in the R-M269 haplogroup, however at a lower probability. For Hubble's #2 and #3, with a GD of 14 and 17 respectively, the computer algorithm assigns them to haplogroups I-M253 and J-M172 respectively. The I and J haplogroups are genetically very distant from the R haplogroup. These 2 Hubbles are not related to Richard the Immigrant. The closest common ancestor between R-M269 and the I-M253 and J-M172 haplogroups occurred over 30,000 years ago.

STR marker evaluation can be a good first step in determining relatedness even with its limitations. But it can only provide a probability of placement within one of the early haplogroups. A SNP test must be done to confirm the placement and assign a more recent, genealogically relevant, haplogroup (more recent haplogroups are referred to as subclades). The above example is a case where a relatively inexpensive STR test can provide useful information, possibly saving the individual from the more expensive SNP test.

II) Results of Testing 2005-2019

In 2005 the Society initiated the Hubbell/Hubble Surname Project using the testing company FTDNA (familytreedna.com.). The testing in those early years was limited to 12 STR markers. Later tests included 37, 67 and 111 STR testing. No SNP testing was performed until 2019.

						Short Tandem Repeat Values (STR Values)												
	ME#	GD	Name	Date Tested	Assigned Haplogroup	DYS393	DYS390	DYS19	DYS391	DYS385a	DYS385b	DYS426	DYS388	DYS439	DYS389i	DYS392	DYS389ii	Comment
			Ancestral Value			13	25	15	10	11	14	12	12	12	13	13	27	
1	15601	0	M.V. Hubbell	2005	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
2	10733	0	Billy John	2005	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
3	10660	0	B. Hubbell, dvm	2005	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
4	17216	1	R.D. Hubbell	2005	R-M269	13	25	15	10	11	14	12	12	11	13	13	27	USA
5	12551	1	Robert Hubble	2005	R-M269	13	25	15	10	11	14	12	12	13	13	13	27	USA
6	19023	1	**B.M. Hubble**	2019	R-FT115348	13	25	15	10	11	14	12	12	13	13	13	27	USA
7	16285	0	**R.W. Hubble**	2017	R-FT115348	13	25	15	10	11	14	12	12	12	13	13	27	USA
8	11859	2	H.R. Hubble	2005	R-M269	13	25	15	10	11	15	12	12	12	13	12	27	USA
9	11859	2	**H.R. Hubble (retest)**	2019	R-FT164982	13	25	15	10	11	15	12	12	12	13	12	27	USA
10	15358	2	**N. Hubble**	2019	R-FT164982	13	25	15	10	11	15	12	12	12	13	12	27	USA
11	13082	2	J.J. Hubbell	2005	R-M269	13	24	15	10	11	15	12	12	12	13	13	27	USA
12	8026	0	T.P. Hubbell	2008	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
13	18282-19236	0	**M.R. Hubbell**	2009	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
14	13905	0	J. Hubbell	2013	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
15	[redacted]	0	**B. Hubbell**	2015	R-M269	13	25	15	10	11	14	12	12	12	13	13	27	USA
16	17520	0	**R.C. Hubbell**	2019	R-FT75886	13	25	15	10	11	14	12	12	12	13	13	27	USA
17	10231	0	**D.C. Hubbell**	2019	R-FT75787	13	25	15	10	11	14	12	12	12	13	13	27	USA
18	na	3	**W. Johnston**	2019	R-FT164981	13	25	15	10	11	14	12	12	12	14	12	28	USA
19	14143	14	R. L. Hubble	2005	I-P37	13	23	15	10	12	15	11	15	12	14	11	30	USA
20	16532	7	D.A.J. Hubbell	2006	R-M269	13	24	14	11	11	14	12	12	12	13	15	29	Canada
21	na	8	K. Hubble **	2006	R-M269	14	24	14	11	11	14	12	12	12	14	13	30	Florida (2010)
22	na	8	S.L. Hubble	2010	R-M269	14	24	14	11	11	14	12	12	12	14	13	30	New Zealand
23	na	8	P.W.B. Hubble	2010	R-M269	14	24	14	11	11	14	12	12	12	14	13	30	Spain
24	na	8	J.P. Hubble	2010	R-M269	14	24	14	11	11	14	12	12	12	14	13	30	England
25	na	14	**F.A Hubble**	2011	I-M253	13	22	14	10	13	14	11	14	11	12	11	28	?
26	na	14	**R.D. Hubble**	2017	I-M253	13	22	14	10	13	14	11	14	11	12	11	28	b: in London
27	na	14	**A. Hubble**	?	I-1	13	22	14	10	13	14	11	14	11	12	11	28	Arkansas
28	na	17	R.J. Hubbell	2006	E-M35	13	25	13	10	16	18	11	12	12	13	11	30	?
29	na	17	R.G. Hubble **	2006	J-M172	12	23	15	9	13	16	11	16	12	13	11	29	New Zealand

** = according to Doug Poulter, K. Hubble and R.G. Hubble are related to Stephen Hubball of England (b:~1790).
ME# = assigned number in the 2017 Millennium Edition: Genealogy and History of the Hubbell Family
GD = Generic Distance
Participants 1-18 in bold type had 111 STRs tested except #13 had only 37 STRs tested. All except #13 and #15 had SNP testing
Participants 25-27 in bold type had 37 STRs tested

Table 2: The first 12 STR marker values for all Hubbell/Hubbles who joined the Hubbell Surname Project. Differences from the ancestral value are in red. ME# are from the 2017 Millennium Edition of the Hubbell Family Genealogy. Participants in bold had more than 12 STRs tested.

12-STR Testing:

Table 2 lists the STR testing results for the first 12 STR markers on all Hubbell participants in the surname project. Highlighted (bold) testers were tested for additional STR markers and SNPs. From the results of STR testing individuals are assigned to a haplogroup with a certain probability.

The first 18 Hubbells listed above have a genetic distance of three or less from the Hubbell ancestral value and all were assigned, with varying probabilities, to the R-M269 haplogroup. FTDNA's computer algorithm calculates the probability of belonging to the R-M269 haplogroup at 99.27% if all 12 values match the ancestral value. If only 10 of the 12 values match (GD=2) the probability of belonging to the R-M269 haplogroup drops to 48.5%.

17 of these 18 participants, including the three individuals with a GD=2, are identified in the 2017 Millennium Edition of the Genealogy and History of the Hubbell Family. Thus, the paper trail firmly establishes these three individuals with a GD=2 as descendants of Richard the Immigrant and consequently belonging to the R-M269 haplogroup. W. Johnston (#18) has a GD of 3. The algorithm's calculation of his belonging to the R-M269 haplogroup is only ~25% yet SNP testing confirmed that he is a genetic descendant of Richard the Immigrant (4).

R.L. Hubble (#19) is listed in the Hubbell Genealogy but exhibits a GD of 14. With a GD of 14, the computer algorithm places this individual in the I-P37 haplogroup, a haplogroup that is genetically far removed from the Hubbell haplogroup of R-M269. R.L. is classified as a NPEs or Non-Paternal Event. There are a few ways an individual can be classified as an NPE, the most common being by an adoption occurring in one of his ancestors.

Individuals #20 through #24 have a genetic distance of 7 or 8 but are assigned to the R-M269 haplogroup. These individuals match less than 50% of the tested markers. It is unlikely that any member of this group is a genetic descendant of Richard the Immigrant. However, results obtained from only 12 STR values is not powerful enough to be conclusive. The minimum level of testing recommended by FTDNA is at the 37 STR level. According to FTDNA, if two men match on 33 or more of 37 STR markers (a GD of 4 or less), they are related to each other within a genealogical timeframe. FTDNA's criteria for a genetic match, regardless of the number of STRs tested, is that two individuals must match at least 90% of their tested STR markers. Whether the descendants of Richard the Immigrant and these five individuals share a much earlier common English Hubball ancestor, further back in time before Richard the Immigrant, can only be determined if they are tested for a higher number of STR markers or through SNP testing. D.A.J. Hubbell (#20), is the only individual in this group that has been assigned a ME# and listed in the Hubbell genealogy. A SNP test will be required to confirm his genetic relatedness to Richard the Immigrant.

The final group of five individuals (#25 through #29) have been classified into genetically distant haplogroups. A common ancestor between the R and the E, I and J haplogroups would have lived 20,000 to 30,000 years ago.

The 29 Hubbells tested above are divided into four distinct haplogroups. Genetically, the oldest haplogroup of the four is E and is represented by R.J. Hubbell (#28) (Refer to Figure 11 below). R.J. displays a GD of 17. The next oldest is the I haplogroup defined by four individuals, #19, #25, #26 and #27, all with a GD of 14. The J haplogroup is defined by individual #29 with a GD of 17. Lastly, there is the R haplogroup, containing 23 individuals, all with a GD of three or less. Of these four haplogroups, it is impossible to draw any conclusions on the E and J haplogroups as they are represented by only one individual each. These two individuals could be NPEs and not represent a separate and unique "family group". The I and R haplogroups, containing four and twenty-three individuals respectively, likely represent two distinct "family groups" using the Hubbell/Hubble name. However, the number of individuals tested is too small to draw any definitive conclusions.

Apart from identifying genetically distinct "family groups" through haplogroup identification, there exists a clear paper record that supports the existence of multiple "family groups". The Society has performed extensive research of the surviving written records in England and this data suggests that there are at least three distinct "family groups" carrying the Hubbell/Hubble name, living in England. Each of these groups can trace their origins to different parts of the European continent. It is likely that additional "family groups" may exist that have yet to be identified through either genetic testing or paper records. SNP testing of Hubbles in England and around the globe would be instrumental in identifying these populations.

(4) W. Johnston was not part of the surname project. He was STR and SNP tested, independently, and based on his SNP results was confirmed as a descendant of Richard the Immigrant. Further investigation revealed that he is descended from ME#4957 James Asa Hubbell. For more information on W. Johnston see the article "A Mystery Solved" in the 2020 Annual.

The usefulness and limitations of STR testing, even at the 12 STR level, became apparent during Doug Poulter's research into the Hubbles of England (5). Doug deduced from the paper record that K. Hubble (#21, living in Florida in 2010) and R.G. Hubble (#29, living in New Zealand) were both descended from Stephen Hubball of Stoke Prior, Worcestershire England. Doug suggested they both take the FTDNA test to confirm that they are related. However, the test showed that they were not genetically related, falling into entirely different haplogroups. Never-the-less, they may both be related to Stephen through an NPE event. As discussed above, it is unlikely that K. Hubble (#21) is related to Richard the Immigrant, but he may be more distantly related. Only SNP testing would confirm this.

37-111 STR Testing:

Twelve participants had more than 12 STR markers tested. Table 6, in Appendix A, lists all STR values for the eight participants who were tested at the 111 level, along with four participants who were tested at the 37 STR level. Table 7 (Appendix A) compares the mutated STR values between related individuals.

The data suggests that the additional STR testing adds little value in defining relatedness among individuals. Regardless of whether one is tested on 12 markers or 111 markers, an individual can only be placed in an early haplogroup. One of the twelve individuals (B. Hubbell #15) tested at the 111 STR marker level was placed, based on STR values only, in the R-M269 haplogroup; the same haplogroup that individuals who were tested with only 12 STR values were placed in. The mutation that defines the haplogroup R-M269 is a very early mutation that occurred about 8,000 BCE and provides little useful genealogical information.

One concern expressed in the literature is the reproducibility of the testing results. H.R. Hubble (#8 & #9) was tested for 12 markers in 2005 and, with a new sample, was retested in 2019. The test results were identical demonstrating the reproducibility of the STR testing results, performed at FTDNA. Additionally, one would expect the results for N. Hubble (#10) to be identical to his father's (H.R.) results. Although it has been reported that a STR mutation can occur between father and son, in this case all 111 markers between father and son were identical. See Tables 6 and 7 in Appendix A.

The one rational for performing a STR test is to ascertain how closely related a group of individuals are to each other. One would expect that closely related individual would closely match STR mutations. In the example above, a father and son showed identical results. In Table 3 below, the STR values between two first cousins are compared. One would expect that their values would be almost identical and yet they are not. There are four mutations in this family in just 2 generations; a rate much higher than suggested in the literature.

		Genetic Distance	2	2
		Date Tested	2019	2019
STR Name	test #	Ancestral Value	B. Hubbell	R.C. Hubbell
		Haplogroup	R-M269	R-FT75886
DYS557	42	17	17	18
DYS710	59	34	34	33
DYS495	62	16	15	16
DYS715	89	25	24	25

Table 3: Comparison of 111 STR values between first cousins. Within two generations there are four mutations, a rate higher than expected.

Table 7, in Appendix A, further demonstrates the variability of the number of STR mutations across family lines. Specifically, see the comparison between H.R. Hubble, N. Hubble and W. Johnston. They all share a common ancestor only 4 generations in the past (ME# 2542 Steinman Hubbell) but display considerable variability in the number of mutations. Between H.R Hubble and W. Johnston there are eight different mutations in four generations.

Table 4, below, calculates the average STR mutation rate for each of the eight participants tested at the 111 level. The mutation rates range in value from 32 years per mutation to at least 324 years per mutation.

(5) A Sojuourn: The Hubbles of the United Kingdom & Other Commonwealth Countries, Douglas W. Poulter, (2006) available through Lulu.com

Richard Hubball the Immigrant: b: ~1625 d: 1699

	Year of Birth	Years from 1625	# of Mutations since Richard	Ave. time between STR mutations (yrs)
D.C. Hubbell	1930	305	2	151
R.C. Hubbell	1949	324	2	162
B. Hubbell	1957	332	2	166
H.R. Hubble	1933	308	7	44
W. Johnston	1949	324	10	32
N. Hubble	1974	349	7	50
R.W. Hubble	1949	324	0	324
B.M. Hubble	1982	357	3	119
			average:	131

Table 4: STR mutation rates for eight participants tested at the 111 level

One possible explanation for the large discrepancies in the STR mutation rates is the age of the father at the time of conception. The literature suggests that the variable rate of mutations of single point mutations (SNPs) is correlated with the age of the father at the time of conception. Could this also be the case for STR mutations? Table 8 in Appendix A shows the age of the father at the time of the birth of his son. From this data, it cannot be concluded that the age of the father at the time of conception is correlated with a STR mutation. The data does support the definition of a generation as being 30 to 35 years; a lot higher than would be anticipated. The average generation time for the six participants listed in Table 8 is 32 years.

Conclusions:

In conclusion, although STR data has some value in placing individuals on the haplotree, the placement is in the older haplogroups and therefore does not provide the discriminating power necessary to be relevant in a genealogical time frame. However, STR testing can be useful as a first screen to ascertain whether individuals could be genetically related. Additionally, based on the data presented here, it would be impossible to construct an accurate family tree, as some have attempted, based on the different STR marker values due to the randomness of the mutations and the variation in mutation rates observed between related individuals.

To better define the relationships between Hubbell's in the United States, England and elsewhere in the world will require testing for single point mutations or SNPs.

SNP Testing:

In 2019, six Hubbell's agreed to participate in extensive SNP screening using FTDNA's Y-700 DNA test. All six participants can trace their ancestry, through paper documentation, back to Richard Hubball the Immigrant, our common English ancestor (~1625-1699). The results of the SNP testing are shown in Figures 8 and 9.

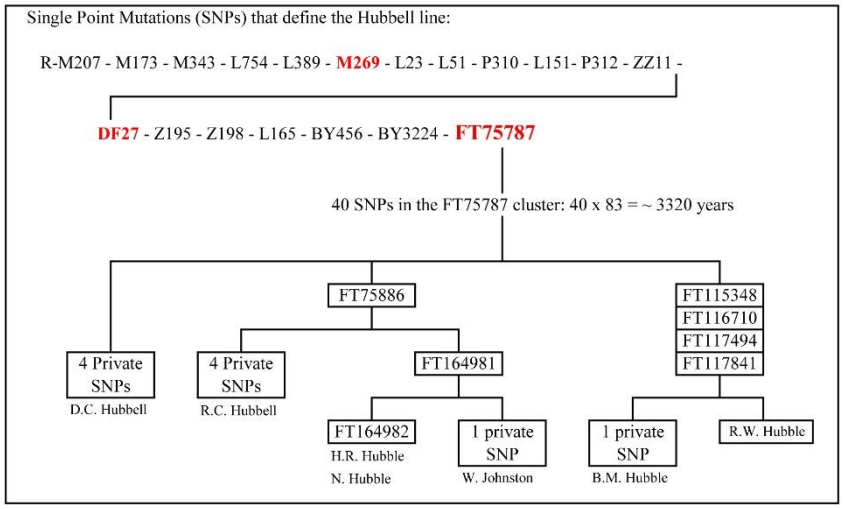


Figure 8: Graphic showing the relationship between SNPs on seven participants. W. Johnston tested positive for all the Hubbell Family markers and is positive for two unique mutations found in the descendants of ME#3 Richard Hubball Jr. W. Johnston has one unique, private SNP. Private SNPs are unique mutations associated with only one individual and have not been named.

Figure 8 displays the progression of mutations that have occurred in the Hubbell family over the last 8-10,000 years. Figure 9 superimposes the most recent SNP mutations on the Hubbell family tree and demonstrates how these mutations can identify the different branches.

In Figure 8, starting with SNP R-M207, the progression of older mutations to more recent mutations proceed to the right. There are older mutations upstream of R-M207. The mutation FT75787 occurred in Richard the Immigrant, or earlier in his family, and is present in all seven participants. All Hubbells descended from Richard the Immigrant will possess this mutation and all up-stream mutations. FT75787 is an arbitrary representative of 40 mutations that occurred after the BY3224 mutation. Of the hundreds of thousands of people tested throughout the world, only seven participants have tested positive for all 40 of these mutations after BY3224. Only further testing will determine the order of occurrence of these 40 mutations.

What is the importance of these 40 mutations? The 40 mutations between BY3224 (including FT75787) and the two mutations found in two of Richard the Immigrants children (FT75886 and FT115348) cover an estimated time span of 3300 years (6). Therefore these 40 mutations occurred approximately between 1666 BCE and 1654 CE. 1654 CE is the birth year of ME# 3 Richard Jr. Identifying individuals possessing some or most of these 40 mutations will identify our early English ancestors predating Richard the Immigrant. The fewer of these 40 shared mutations the further back in time the ancestor. Are we related to the Hubands of Ippsley? If we are then a Huband descendant will possess most of these 40 mutations and confirm our descent from the Huband family.

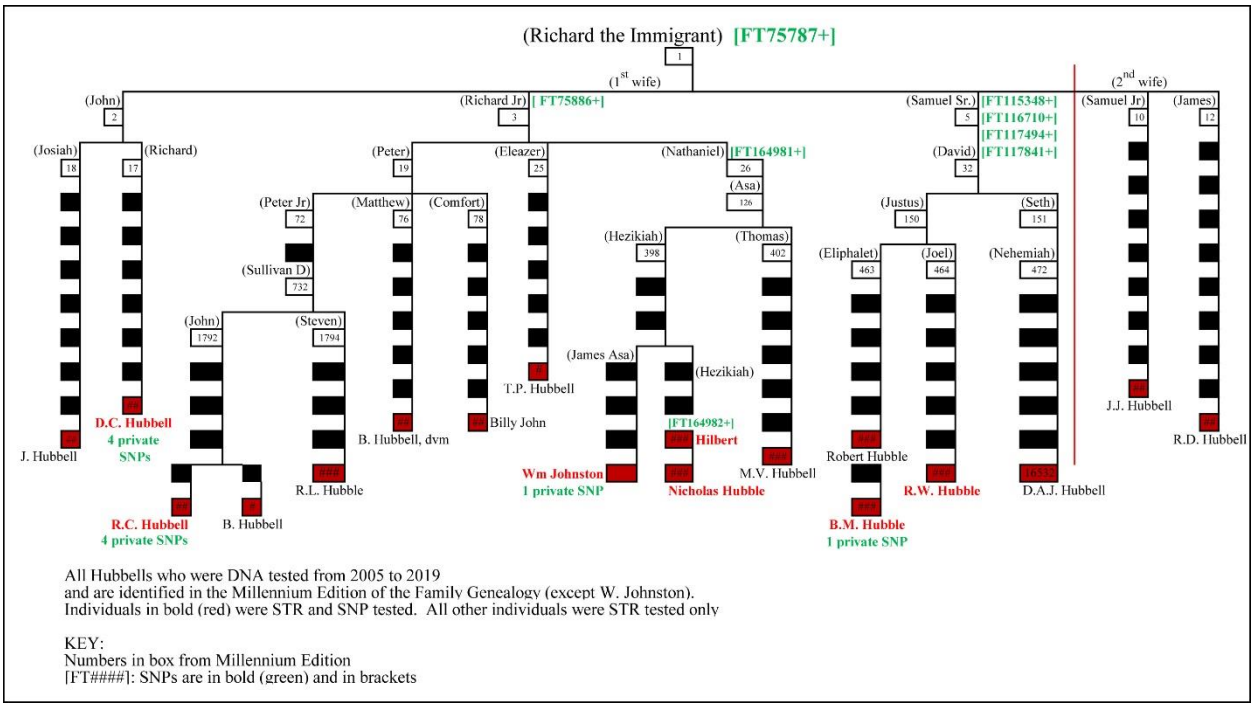


Figure 9: SNP mutations placed within the Hubbell Family Tree. Listed are all the descendants of Richard the Immigrant, who had STR testing performed (from Table 1). Seven individuals, in red, were SNP tested. The name of the SNP mutation is labeled in green. All seven SNP tested participants were positive for the FT75787 mutation. It is not known if this mutation originated with Richard the Immigrant or his father or an earlier ancestor. Private SNPs found in only one person are not named

Rate of SNP Formation:

In the early years of genetic testing the occurrence of a SNP mutation was believed to be rare, occurring once every 1000 years on average. Therefore, STR testing was preferred over SNP testing because of STR marker's higher frequency of occurrence. Today, based on new data, the time between SNP formation is estimated to be, on average, 83 years (6). This shortened time between SNP mutations and SNPs stability make it the preferred test. The time between SNP occurrences could theoretically differentiate individuals every third generation. Most researchers have concluded SNP mutations do not occur uniformly, but in bursts or clusters over time. Additionally, some evidence suggests mutations are more likely to occur in older males reproducing later in life.

(6) Information from recent communications with Tim McLeod suggest that the frequency of SNP mutations may be higher than once every 83 years on average. Tim has suggested that current data may indicate a mutation rate of as low as 40 years on average.

Using the SNP data from Figure 9, the maximum average SNP mutation rate for two Hubbell lines can be calculated (see Table 5). The first and most compelling example involve the four mutations (FT115348, FT116710, FT117494 and FT117841) listed next to Samuel Sr. (#5). These four mutations are not present in the descendants of Samuel Sr.'s (#5) brothers [John (#2) and Richard (#3)] but are present in both of his descendants (B.M. Hubble and R.W. Hubble). Therefore, these four mutations must have occurred in Samuel Sr (#5), David (#32) or Justus (#150). Samuel Sr. was born in 1657 and Justus' oldest son Eliphalet (#463) was born in 1769, a span of 112 years or a mutation rate of 28 years per mutation. This appears to be a burst or cluster of mutations but no correlation with the age of the father at the time of conception could be made. To calculate the mutation rate in this line over a longer time period we make the assumption that one of these four mutations occurred in Samuel Sr. With that assumption, the mutation rate from Samuel Sr. to R.W. Hubble is 73 years per mutation and for B.M. Hubble the mutation rate drops to 65 years, assuming B.M. Hubble's private mutation occurs with him. The mutation rate will be smaller if the assumptions are not met. (Table 5)

In a second example, the four SNP tested descendants of Richard Jr. (#3 in Figure 9) have the FT75886 mutation while the SNP tested descendants of his two brothers, John (#2) and Samuel Sr. (#5), do not, confirming that the FT75886 mutation occurred in Richard Jr. (#3). One of Richard Jr's descendants, R.C. Hubbell, has this mutation and an additional four private mutations that are not present in the descendants of Nathaniel (#26). These four private mutations must have occurred in Peter (#19) or later. We do not know in which descendants of Peter these four mutations occurred.

For this example, it is assumed that one of the 4 private mutations occurred in R.C. Hubbell. If all 4 private mutation occurred before R.C. Hubbell, then the average mutation rate will be smaller. Using these five mutations, from Richard Jr. (#3) to R.C. Hubbell, an average mutation rate of 59 years can be calculated.

	Date of Birth	Years from 1657	# of generations	# of SNPs from 1657	Ave. Years between SNP Formation
R.W. Hubble (#7)	1949	292	9	4	73
Samuel Sr b: 1657					

	Date of Birth	Years from 1657	# of generations	# of SNPs from 1657	between SNP Formation
B.M. Hubble (#6)	1982	325	10	5	65
Samuel Sr b: 1657					

	Date of Birth	Years from 1654	# of generations	# of SNPs from 1654	Ave. Years between SNP Formation
R.C Hubbell (#16)	1949	295	10	5	59
Richard Jr b: 1654					

	Date of Birth	Years from 1702	# of generations	# of SNPs from 1702	Ave. Years between SNP Formation
H.R. Hubble (#9)	1933	231	7	2	115
Nathaniel b: 1702				ave:	78

(#xx) refers to Table 1

Table 5: Rate of mutation of SNPs in two selected lines of the Hubbell Family

In the last example, from Nathaniel (#26) to H.R. Hubble there are only two mutations occurring over a time span of 231 years. As in the other examples, we assume that the FT164981 mutation occurs in Nathaniel and the FT164982 mutation occurs in H.R. Hubble giving an average of one mutation every 115 years.

Although the dataset and timeframes are small, this data demonstrates that the average rate of SNP mutation of 78 years in the Hubbell family is close to the 83 years reported in the literature. The possibility that the SNP mutation rate is influenced by the father's age, at the time of conception was reviewed but no conclusions could be reached. See Table 8 in the Appendix A.

How Does this Genetic Information Help Us Unravel Our Story?

Knowing the sequence of our Y-chromosome mutations, as stand-alone information, adds little to our knowledge and understanding of our family's history and genealogy. Our genetic data must be compared with thousands of others before any meaningful interpretations of the data can be proposed. To this end there are a few groups and organizations worldwide constructing SNP databases. Unfortunately, the effort is fragmented. All of the genetic genealogy organizations (FTDNA, 23andMe, Ancestry, etc.) are profit based companies with little incentive to share data with competitors. Of these companies, probably the most extensive Y-DNA SNP database has been compiled by FTDNA.

FTDNA, in concert with citizen scientists, are incorporating this new genetic information with the vast knowledgebase built up over the decades by historians, archeologists and anthropologists. This effort is relatively new so the story is incomplete but given time, with more widespread testing, a more complete picture of the story of mankind's ancient DNA history will emerge.

The narrative that follows is the most current (2018) understanding of the "history" of SNP formation. It is only a framework, attempting to combine classical archaeology and anthropology with modern genetic analysis. It is important to note that the proposed dates of occurrence and location of these mutations is controversial. Conclusions based on the analysis of the existing data fluctuates wildly from author to author. As more information is acquired in the years to come, a clearer picture will emerge (7). To assist the reader, a timeline of historical events has been created in Figure 10 below.

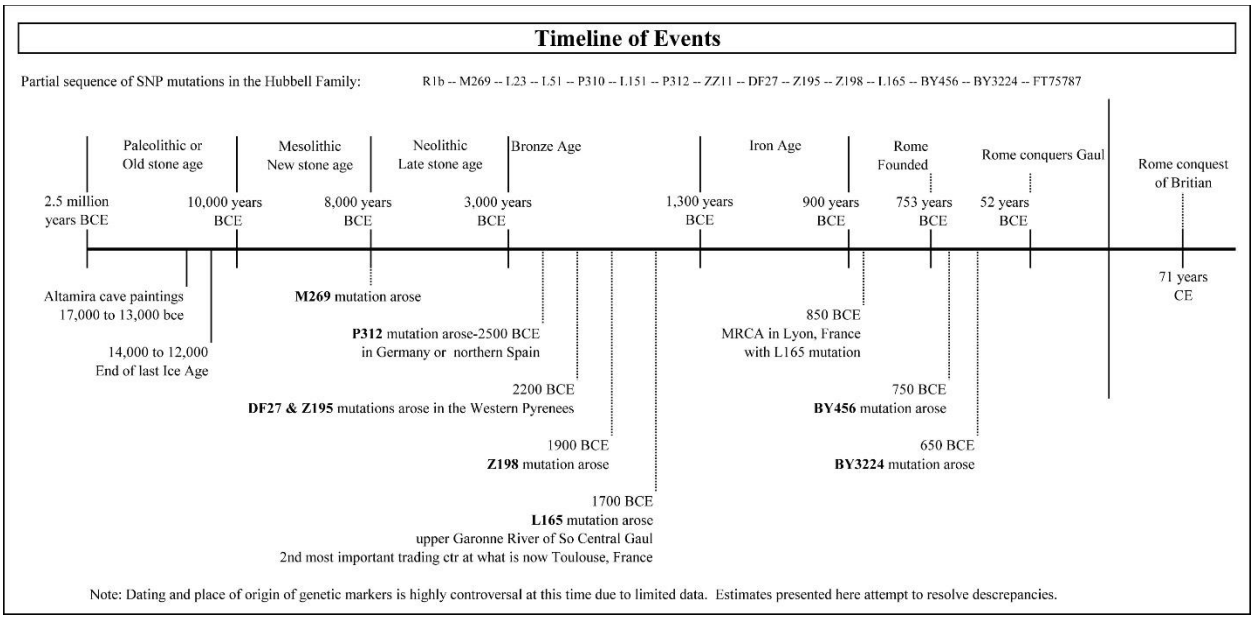


Figure 10: Timeline of Historical Events

All modern males can trace their ancestry back to Africa, starting with Y-chromosome Adam about 240,000 BCE. This does not mean that "Adam" was the only male in Africa at that time. It means that his genetic code is the only male code to survive into the present. All other male lines have since died out. Recent evidence (2019) suggests that humans had migrated from Africa into Eurasia as early as 160,000 BCE but it was not until about 60,000 BCE that the major migration from Africa into Eurasia began.

From 240,000 to about 30,000 BCE "Adam's" Y-DNA accumulated many new SNP mutations that have become so widely dispersed in the human population that they are now the major haplogroups that categorize the human race. However, it was not until about 30,000 BCE that the mutation **M207** occurred forming the "R" haplogroup. Hubbells, descending from Richard the Immigrant, belong to the R haplogroup. (Refer to Figure 11 below)

This mutation was first identified in the remains of a 24,000 year old boy from south-central Siberia. It is believed this individual belonged to a tribe of mammoth hunters that roamed across Siberia and parts of Eastern Europe during the last Ice Age.

From our descendants carrying the **M207** mutation, the **M173** mutation occurred forming the R1 haplogroup. Mutations between **M207** and **M173** are considered "minor" mutations. These minor mutations have not been characterized as to time or location of occurrence and are often omitted from timelines to conserve space. From the R1 haplogroup, there arose two mutations that define two groups of tribes speaking a Proto-Indo-European language. These two groups of men are the R1a haplogroup, defined by the **M420** mutation, and the R1b

(7) The information presented below was compiled from numerous web sources but primarily from: https://www.eupedia.com/europe/Haplogroup_R1b_Y-DNA.shtml. This is an excellent site that presents information on a number of the major haplogroups.

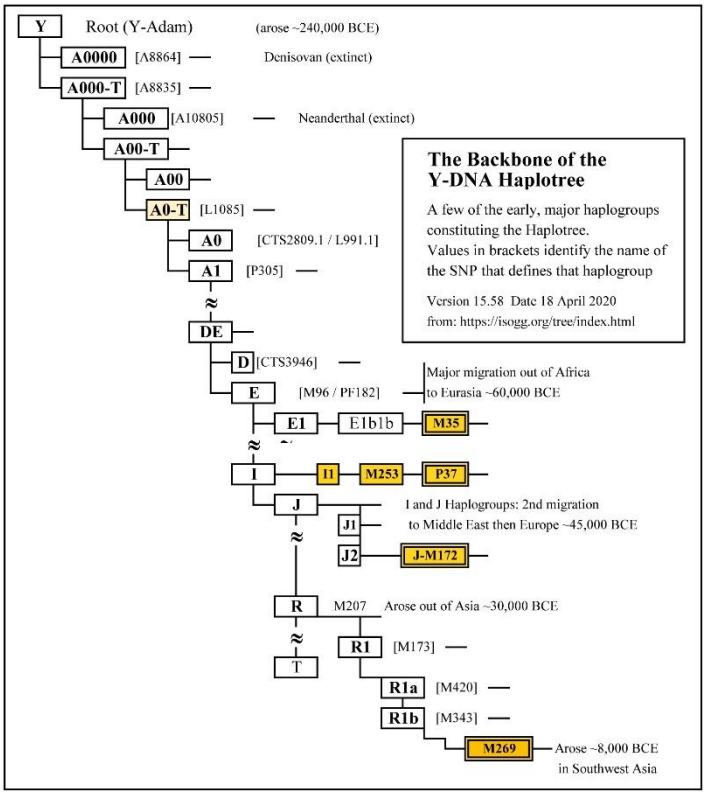
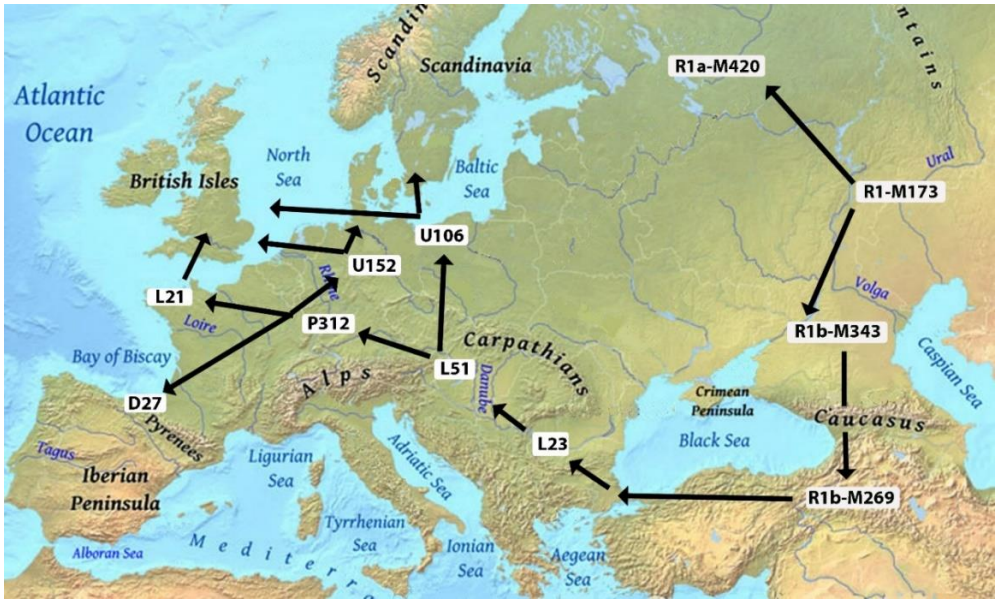


Figure 11: Progression of Haplogroup formation starting with Y-Chromosome Adam. Each haplogroup is represented by a SNP mutation. For clarity, some major, and all minor haplogroups have been omitted. The "Hubbell" haplogroup R is defined by the SNP mutation M207 (sometimes written R-M207)

haplogroup, defined by the **M343** mutation. The American Hubbell family descend from the R1b haplogroup. As R1b descended people, we are all **M207, M173, M343** positive. Those are essentially the markers or signposts for R1b. They have been very stable mutations over many centuries.



Map A: Likely migration route of R1b-DF27 people from Western Asia, across Europe to the Pyrenees Mountains of Northern Spain and Southern France

Tribes belonging mainly to the R1a haplogroup reportedly occupied the northern part of the Asian steppe (forest-steppe and tundra) in Central Asia, while in the southern part (open steppe) was predominately occupied by the haplogroup R1b (See Map A above). These southern nomadic herders, descendants of the mammoth hunters, carried the **R1b-M343** mutation. When the mammoths disappeared, along with the ice sheet, these R1b+ people

learned to domesticate animals (cows, goats). Classical anthropological studies have uncovered the earliest evidence of animal domestication in southeastern Turkey and northern Iraq. This area is now considered to be the "original homeland" of R1b peoples.

The next major mutation in R1b+ individuals was **R1b-M269**. It is thought this mutation arose in the early Neolithic age, around 8,000 BCE around the Caspian Sea but authors differ widely on the dating of this mutation.

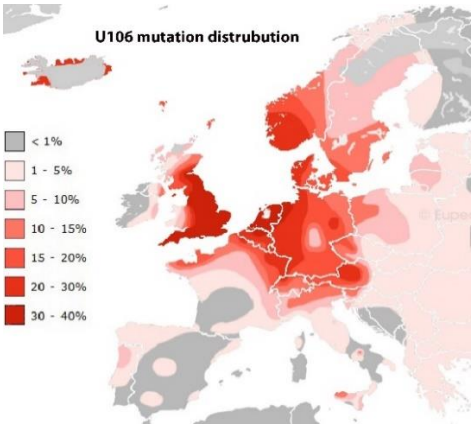
For the next 4000 years, these **R1b-M269+** people slowly migrated westward through the Middle East (Iraq, Turkey). Around 4200 BCE, **R1b-M269+** people migrated into Southeastern Europe, mostly in the Balkans (Romania, Bulgaria, Serbia). The Balkans were well settled by this time and possessed some of the world's largest towns. Easy plunder for the mounted, nomadic warriors from the East. Around 4000 BCE, Northwest Europe including France, England and Germany remained sparsely settled supporting small tribal societies of hunter-gatherers practicing only limited agriculture. There was little incentive for the R1b+ nomads to leave the comfort of the wealthy and populous civilizations in the Balkans for the harsh living conditions that lay to the West. These late Neolithic Age / early Bronze Age people coveted tin, copper, and gold, of which the Balkans had plenty, riches that had not yet been discovered in Western Europe.

During this slow migration into Eastern Europe, two major mutations of the **M269+** population occurred: **L23** and **L51**. The **L23** mutation is thought to have appeared around 4,500 BCE. While the **L51** mutation is thought to have occurred northwest of the Balkans, in East Central Europe (Hungary and Austria) or present-day Germany around 2500 BCE.

By 2500 BCE, agrarian towns had started to develop in Northwest Europe. Gold and copper had begun to be mined. The prospects for migration were now far more appealing. At about this time, from Austria, the R1b+ invaders, mounted on horses and carrying copper weapons, migrated into what is today Germany where they easily vanquished the indigenous peoples. Ultimately, they reached the Atlantic Ocean, north of the Pyrenees Mountains, around 2200 BCE. In contrast, around this time, contemporary Egyptians completed building the Great Pyramid of Giza (2600 BCE). It was during this time that mutations in **M269+ L23+ L51+** individuals produced two major lines defined by the mutations **U106** and **P312**, thought to have occurred about 2500 BCE. Recent DNA tests on ancient human remains (2012, 2015) suggest that all three mutations **L51, U106** and **P312**, were present in Germany during this period (8).

U106 Mutation:

Around 2000 BCE, in what is now Central Germany, **U106+** individuals migrated north. Today, **U106+** individuals mainly cluster in the Low countries of present-day Belgium and Northwest Germany (Map B). It appears that the **U106** mutation, in Britain, identifies "Anglo-Saxon" derived peoples. Norway is about two thirds **U106+** and the homeland of the Anglo-Saxon's (Friesland-modern day Belgium and Netherlands) is about 75% **U106+**. Ancient Norse populations appear to belong mostly to Y-DNA haplogroups I, R1a and R1b (**U106+**). However, there are great disparities between the regions of Scandinavia. Over 40% of Swedes belong to haplogroup I1a and another 10% to haplogroup I1c.



Map B: Distribution of the U106 mutation in modern Europe (from eupedia.com)

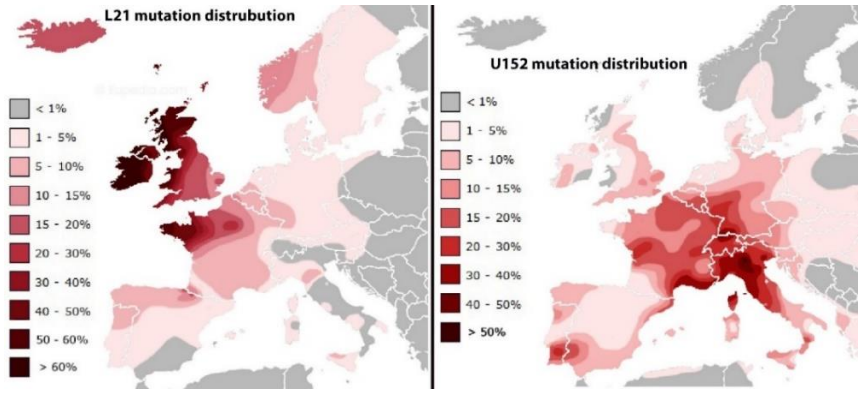
(8) Other authors place the occurrence of the P312 mutation in northern Spain.

P312 Mutation:

The other major branch, **P312** (also known as S116), discovered in 2008, is the branch the Hubbell family descends from. Researchers differ on where this mutation first appeared, either in Germany or Northern Spain (9). From this population of people there arose three major branch mutations: **L21, U152** and **DF27**

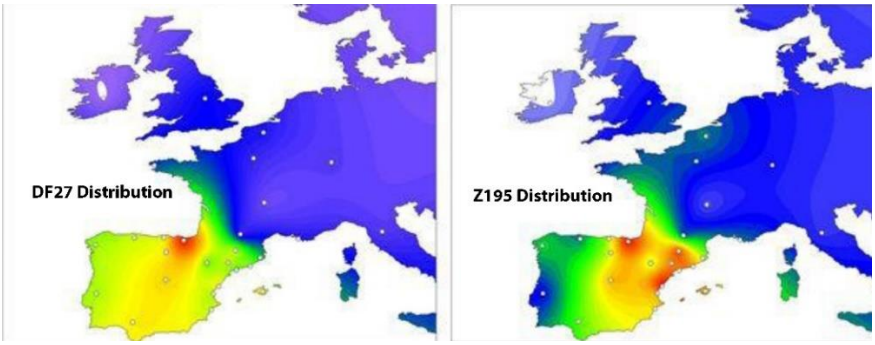
The **L21** mutation is the most common marker in the British Isles with significant concentrations in France but it is also found in low frequency across Western and Central Europe (Map C). Individuals possessing the **L21** mutation arrived in France and the Low countries around 2200 BCE; into Britain by 2100 BCE and into Ireland by 2000 BCE. By 1700 BCE, **L21** positive peoples had migrated into Scandinavia and mingled with the **U106+** population. In France, **L21** is mainly present in Brittany and Normandy. Additionally, it is believed that the presence of this mutation in England is due to Frisians and Anglo-Saxon migrations that took place in the 3rd through 10th centuries CE.

The **U152** mutation is seen further east in Northern Europe and Northern Italy. In England, this mutation is predominantly found in Eastern England in what is known as the "Danelaw" (10). This **U152** mutation appears to identify descendants of the Danish Viking in those who possess it. Some Norse (Southern Sweden and Norway) will be **U152+** (Map C).



Map C: Distributions of the mutations L21 and U152 in modern Europe (from eupedia.com)

The third mutation found in **P312+** peoples is **DF27**, a major Hubbell Family marker. The **DF27** mutation is thought to have occurred around 2000 BCE, first appearing in or near the town of Bilbao, Spain in the Pyrenees Mountains. This mutation is accompanied by the **Z195** mutation that appears to have originated at about the same time as the **DF27** mutation.



Map D: Heat map of the concentration of DF27+ and Z195+ people in modern Spain. The study screened over 1000 individuals. The mutations DF27 and Z195 occurred about the same time around 2,000 BCE in Northern Spain. The mutations are found in the highest concentrations in the red areas and the least concentrated in the blue areas. (from Neus Solé-Morata et al)

(9) Laura Valverde, et al; New Clues to the Evolutionary History of the Main European Paternal Lineage M269: Dissection of the Y-SNP S116 (P312) in Atlantic Europe and Iberia; European Journal of Human Genetics (2016) 24, p. 437-444

(10) The Danelaw is that portion of Eastern Britain conquered and settled by the Vikings from 800 to 1000 CE.

Today, the highest frequencies of the **DF27** mutation are found in Native Basques and Western Iberian populations such as Asturias and Galicia in the Pyrenees Mountains of Spain and Portugal (Map D). The mutation is found in small percentages in Southern France and Great Britain but seems to be rare or absent elsewhere. **DF27** was found in 30-50% of the modern Iberian population (mean of 42%), with the notable exception of native Basques, where it reaches 74%. In France, the percentages drop to 6-20% (mean of 11%). Elsewhere, it is 15% in Britain, but less than 1% in Ireland. Tuscany's population is 8% positive. After evaluating genetic drift and STR data, Neus Solé-Morata concluded that a local Iberian origin of the mutation is the most plausible hypotheses (11). For the **Z195** mutation, the percentages in the Basque Country and in East Iberia (Catalonia, Valencia) range from 20 to 41% of the population.

Cultural anthropologists believe these early Cantabrians lived in small, settled bands comprised of several family groups numbering up to 30 individuals. In contrast to city-states and empires in the East, archaeological evidence suggested there was little, if any, true warfare in Ancient Gaul (France) or on the Northern Iberian Peninsula (Spain), during this time. It was not until a widespread famine occurred around 800-750 BCE that forced small, agrarian bands and clans to consolidate into proto-tribes. During this period, a "severe cold and wet period" (possibly the result of a volcanic eruption), causes crop failures and famine.

From Spain to England: A narrative on the migration of our DF27+ ancestors across Europe, 2000 BCE to 1600 CE.

Around 1900 BCE, after the **Z195** mutation, another mutation, **Z198** occurred, in the same region. From the **Z198** mutation, about 3500 years elapse until the **FT75787** mutation that is found in Richard the Immigrant living in the Midlands of England at the beginning of the 17th century CE (12). In this timespan of about 3500 years there occurred an additional 44 mutations.

These 44 mutations in our ancestors, from **Z198** to **FT75787**, provide little information as nothing about their time or place of occurrence has been published. So how did the peoples that would become the Hubball clan migrate from Spain to England? Were they mariners; fishing and/or trading along the coast of Spain and France? Perhaps, even trading with and finally settling with the early inhabitants of Southern England? Or did they migrate north, slowly, overland, across Gaul (France) finally ending up in England?

Map E: Proposed migration routes from northern Spain to England for R-Z198-L165+ people as proposed by Ayton and McLeod in their manuscript "R1b-L165: A Reconstructed History of the Haplogroup and Its Principal Sublineages (c.2500 BCE to c.1500 CE)"

To address this lack of information, Bill Ayton and Tim McLeod published a speculative, but well-reasoned analysis based on existing anthropological, historical and genetic data (Map E). A few highlights from their narrative will be reproduced here but the reader is encouraged to read their article (13). It is important to remember

---

(11) Neus Solé-Morata et el; Analysis of the R1b-DF haplogroup shows that a large fraction of Iberian Y-Chromosome lineages originated recently in situ; Scientific Reports, 2017; 7: 7341 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5544771/)
(12) Using 83 years per SNP mutation gives 3652 years. Using 40 years, as suggested by Tim McLeod, results in a span of only 1760 years.
(13) William Ayton and Tim McLeod, R1b-L165: A Reconstructed History of the Haplogroup and Its Principal Sublineages (c.2500 BCE to c.1500 CE), 2018. Article can be accessed at: http://mcleod-cabin.net/ayton/

that the Ayton/McLeod explanation is not based on hard facts but merely an attempt to integrate what little we know about pre-historic Gaul and where the descendants of the **Z198+** people of Spain reside today.

The underlying rational for the Ayton-McLeod narrative is an attempt to explain how certain branches of the **Z198-L165+** people migrated to their present locations. For example, how did the descendants of the **BY129+** branch end up as part of the Viking invaders in England? Or how did the **BY67716+** Bidoli clan end up settling in Italy? (refer to Figure 12 below).

The author's first assumption is that our ancestors were successful traders slowly migrating along the ancient trade routes across pre-historic Gaul. They first theorize that our ancestors, carrying the **DF27**, **Z195** and **Z198** mutations, slowly migrated East along the northern side of the Pyrenees Mountains following the East-West trans-Gaul trade route through this region. This trade route was one of the two most important inter-regional trade routes in Bronze-Age Southern Gaul.

They speculate that these traders likely lived, for a while, in the vicinity of Tolosa or modern-day Toulouse, France. Tolosa was the second largest trade center in Bronze-Age Gaul and was the hub of several important middle Bronze-Age trade routes. It was here, around 1700 BCE that the next family mutation, **L165**, occurred.

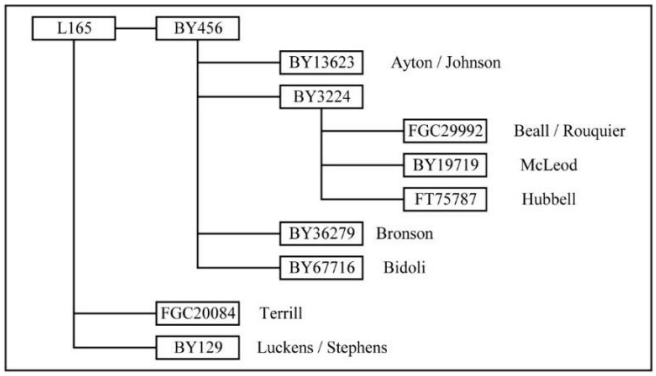


Figure 12: Condensed version of SNP formation in the R-L165 subclade. For a more detailed flowchart see Figure 13 in the Appendix A

Over the next 1000 years, these **L165+** peoples slowly migrated eastward along this major East-West trade route eventually making their way up the Rhône River Valley, settling in Lyon, France around 850 BCE. The North-South Rhône River route was the single most important inter-regional trade route in pre-Celtic and post-Celtic Gaul, and it led to the town of Lugdunum (Lyon) the most important trade center west of the Alps in pre- and post-Roman Italy.

By 750 BCE our ancestors, carrying the **L165** mutation, were divided into three genetic subclades or branches due to additional mutations. The mutation **BY456** is thought to be the oldest of the mutations of **L165** and defines our family line (Hubbell). Mutation **FGC20084** (Terrill branch) and the **BY129** mutation occurred shortly afterward. Still congregating together, these three **L165+** branches continued as traders migrating north from Lyon along the major North-South flowing Saône River into the heart of ancient Gaul. (See Figure 12 and Map E)

Approximately two generations later (~650 BCE), The **BY456+** branch mutated into 4 new subclades:

* BY456* (unmutated): the proto-Lewis/Valentine lineages
* BY13623+: the proto Ayton/ Johnson lineages
* BY3224+: the proto-Beall/Elliott/McLeod/Hubbell lineages
* BY36279+: the proto-Bronson English lineage
* BY67716+: the proto-Bidoli lineage

By 650 BCE, it is theorized that members of most of the **L165** branches resided in Confederation-ruled, petty kingdoms on the Upper Saône and Upper Seine Rivers in Northeastern Gaul. These kingdoms maintained fortified "cities" along the major North-South trade route. Called Senones, these tribes were part of the La Tène culture.

Then, in the mid-500s BCE, the Celtic Cubi, emerge from the forests of Germania, and quickly wrested control of Central Gaul away from the Senones and other tribes, and found their large, fortified capital Avaricum on the riverside site of today's Bourges, France. In time the Celts ruled a great confederacy of Celtic tribes that

includes much of Northern and Eastern Gaul and Southern Britain. Evidently, the Senones (including its surviving lineages of **L165**+ people) became part of this confederacy.

From Central Gaul, the **L165-BY129+** people, around 500 BCE, migrated into the Belgic territory (modern Belgium) where, it is theorized, they formed alliances with Viking chieftains and eventually joined Viking expeditions that spread the **BY129** mutation into the Scottish Hebrides, Sweden, Normandy and England.

Little is speculated about the **L165-FGC20084**+ (Terrill) branch, only citing its earliest identifiable ancestor; Roger Terrill b: 1586 in Middlesex, England.

Around 500-400 BCE the Senones/Celts, during the first or second Gallic expansion, sent an army into Italy. It is speculated that the **BY456-BY67716** proto-Bidoli lineage was part of this army and they eventually settled in the Tuscany region of Italy where their descendants live to this day.

Using educated guesswork and the history of the Senones culture, Ayton and McLeod theorize that the **BY456-BY13623** (proto-Ayton/Johnson) and the **BY456-BY3224** (proto Beall/Elliott/Hubbell lineages) lived near the Melun island fortress located on the Seine River 45 miles downriver from Sens during the 400s BCE (14).

According to Ayton and McLeod, based on archaeological evidence, the **BY456-BY13623** Ayton/Johnson lineages left Gaul, around 400 BCE, and establish a "colony" in eastern Yorkshire, England.

They postulate that the **BY456-BY3224** proto Beall/Elliott/Hubbell clan was part of this migration, settling in the Firth of Forth area around what today is Edinburgh, Scotland. Descendants with the clan name of Elliott are found in this area as early as the 1400s CE. The earliest known Beall **(FGC29985**) and Bell **(FGC2000**0) ancestors were born in the 1600s CE in Scotland.

However, this scenario does not fit with what is known of the Hubball ancestors. There is no evidence that the Hubball clan lived in Scotland. Substantial documentation places the Hubballs in the English Midlands since the early 1500s CE and circumstantial evidence suggests that they have been in the Midlands of England at least since the time of William the Conqueror (1068 CE). The most probable scenario is that the **BY3224+** Beall/Elliott/Hubbell people slowly migrated from France into Southern England with some branches staying in Central England while others migrated further north into Scotland.

Additionally, this land migration scenario does not take into account the differences between the **BY456-BY3224-FGC29992+** Beall and Rouquier families. The Beall's appear in Scotland while the Rouquier family is from the town of Rouquier in southern France (see Map E). This does suggest that the common ancestor of the Beall's and Rouquier's lived somewhere in France, perhaps in the Senones/Celt controlled areas of Central Gaul.

The remaining branch of the **BY456-BY3224** subclade, **BY19719**-McLeod, is thought to have followed ancient marine trade routes between Aquitaine (15) and England and Scotland establishing themselves in the Orkney Islands of Scotland, prior to 787 CE.

The above scenario, of a trans-Gaul migration along the major waterways and trade routes of ancient Gaul, as proposed by Ayton and McLeod, is a working hypothesis based on educated guesses. No hard evidence for this migration exists. Migration by sea, as a possible alternative to the land route was not discussed in their paper.

Maritime trade between The English (Alba) Islands and the Mediterranean has flourished for thousands of years. The earliest written record of a journey from present day Marseille (France) to Iceland, through the Irish Sea and the Hebrides Islands of Scotland, was undertaken by Pytheas in 330 BCE (16). However, there is ample evidence to suggest that routine trade has occurred since at least 1000 BCE. This trade route, leaving the Mediterranean through the Straits of Gibraltar, would have followed the coastline of Spain and France taking it near the town of Bilbo, Spain in the Bay of Biscay. Whether our ancestors were traders or fisherman, this trade route provided the easiest and most direct route to Britain.

---

(14) The city of Sens is said to have been one of the oppida (fortified "cities") of the Senones, one of the oldest Celtic tribes living in Gaul. It is mentioned as Agedincum by Julius Caesar several times in his *Commentarii de Bello Gallico*. (Wikipedia)

(15) The Aquitaine region is in the southwest corner of France and is part of the Pyrenees Mountains. This assumption suggests that one of the subclades of BY3224 (BY19719) did not migrate into eastern central Gaul with the rest of the L165 population but instead remained behind on the Spain/France border.

(16) Mowat, Farley, The Farfarers-Before the Norse; 1998 published by Key Porter Books.

Whether by land or by sea, the real test for any migration theory must take into consideration significant Y-DNA testing of additional British and French populations. Unfortunately, at this time in France, unlike the U.S. and U.K., home-styled genealogical DNA testing has been banned since 1994. The only recourse for a legal DNA test, in France, is with a medical approval or a court order.

Acknowledgement: The author would like to thank Andrew H. Hubble for critiquing and editing this document. His numerous comments and suggestions substantially improved the narrative.

Additional Reading:

1) Dorit, Robert et al, Absence of Polymorphism at the ZFY Locus on the Human Y Chromosome, Science, Vol 268, May 26, 1995, pp 1183-1185

2) Ambrose, Stanley H., Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans, Journal of Human Evolution, Amsterdam, the Netherlands: Elsevier 1998

3) Discovery of Western European R1b1a2 Y Chromosome Variants in 1000 Genomes Project Data: An Online Community Approach, Rocca, Richard A et al, 2012, Plos One Biology. (pioneering study of DF27)

4) Britain's Genes, The fine-scale genetic structure of the British population, Stephen Leslie et al, Nature Magazine, March 19, 2015, p309

5) The Study of human Y Chromosome variation through ancient DNA by Toomas Kivisild (2017), Human Genetics, volume 136, p 529-546: https://link.springer.com/article/10.1007/s00439-017-1773-z?shared-article-renderer#Fig7

6) Genetic History of France, Aude Saint Pierre et al, 2019, bioRxiv website at: https://www.biorxiv.org/content/10.1101/712497v2.full

Appendix A

		Genetic Distance	0	3	2	2	2	7	7	10	2	Arkansas		English
		Date Tested	2019	2019	2019	2019	2019	2019	2019	2019	2009	42	42	48
													2011	
STR Name	test #	Ancestral Value	R.W. Hubble	B.M. Hubble	B.Hubbell	R.C. Hubbell	D.C. Hubbell	H.R. Hubble	N. Hubble	W. Johnston	M.R. Hubbell	A. Hubble	F.A. Hubble Jr	R.D. Hubble
		Haplogroup	R-FT75787	R-FT115348	R-M269	R-FT75886	R-FT75787	R-FT164982	R-FT164982	R-FT75787	R-M269	I-1	I-M253	I-M253
DYS393	1	13	13	13	13	13	13	13	13	13	13	13	13	13
DYS390	2	25	25	25	25	25	25	25	25	25	25	22	22	22
DYS19	3	15	15	15	15	15	15	15	15	15	15	14	14	14
DYS391	4	10	10	10	10	10	10	10	10	10	10	10	10	10
DYS385	5	11-14	11-14	11-14	11-14	11–14	11-14	11-15	11-15	11-14	11-14	13-14	13-14	13-14
DYS426	6	12	12	12	12	12	12	12	12	12	12	11	11	11
DYS388	7	12	12	12	12	12	12	12	12	12	12	14	14	14
DYS439	8	12	12	13	12	12	12	12	12	12	12	11	11	11
DYS389i	9	13	13	13	13	13	13	13	13	14	13	12	12	12
DYS392	10	13	13	13	13	13	13	12	12	12	13	11	11	11
DYS389ii	11	27	27	27	27	27	27	27	27	28	27	28	28	28
DYS458	12	18	18	18	18	18	18	17	17	18	18	18	18	15
DYS459	13	9-10	9-10	9-10	9-10	9-10	9-10	9-10	9-10	9-10	9-10	8-9	8–9	8-9
DYS455	14	11	11	11	11	11	11	11	11	11	11	8	8	8
DYS454	15	11	11	11	11	11	11	11	11	11	11	11	11	11
DYS447	16	25	25	25	25	25	25	25	25	25	25	24	24	22
DYS437	17	15	15	15	15	15	15	15	15	15	15	16	16	16
DYS448	18	19	19	19	19	19	19	19	19	19	19	20	20	20
DYS449	19	29	29	29	29	29	29	29	29	29	29	28	28	30
DYS464	20	14-15-17-17	14-15-17-17	14-15-17-17	14-15-17-17	14-15-17-17	14-15-17-17	14-15-17-17	14-15-17-17	15-15-17-17	14-15-17-17	12-14-15-16	12-14-15-16	14-14-15-16
DYS460	21	11	11	11	11	11	11	11	11	11	11	10	10	10
Y-GATA-H4	22	11	11	11	11	11	11	11	11	11	11	10	10	10
YCAII	23	19-22	19-22	19-22	19-22	19-22	19-22	19-22	19-22	19-22	19-22	19-21	19-21	19-21
DYS456	24	17	17	17	17	17	17	17	17	17	17	15	15	14
DYS607	25	15	15	15	15	15	15	15	15	15	15	14	14	14
DYS576	26	18	18	18	18	18	18	18	18	18	19	16	16	17
DYS570	27	17	17	17	17	17	17	17	17	17	17	19	19	18
CDY	28	36-37	36-37	36-38	36-37	36-37	36-37	36-37	36-37	37-37	37-37	36-36	36-36	33-38
DYS442	29	12	12	12	12	12	12	12	12	12	12	12	12	13
DYS438	30	12	12	12	12	12	12	12	12	12	12	10	10	10
DYS531	31	11	11	11	11	11	11	11	11	11				
DYS578	32	9	9	9	9	9	9	9	9	9			GD=0	GD=0 GD=18
DYF395S1	33	15-16	15-16	15-16	15-16	15-16	15-16	15-16	15-16	15-16				
DYS590	34	8	8	8	8	8	8	8	8	8				
DYS537	35	10	10	10	10	10	10	10	10	10				
DYS641	36	10	10	10	10	10	10	10	10	10				
DYS472	37	8	8	8	8	8	8	8	8	8				
DYF406S1	38	10	10	10	10	10	10	10	10	10				
DYS511	39	10	10	10	10	10	10	10	10	10				
DYS425	40	12	12	12	12	12	12	12	12	12				
DYS413	41	23-23	23-23	23-23	23-23	23-23	23-23	23-23	23-23	23-23				
DYS557	42	17	17	17	17	18	17	17	17	17				
DYS594	43	10	10	10	10	10	10	10	10	10				
DYS436	44	12	12	12	12	12	12	12	12	12				
DYS490	45	12	12	12	12	12	12	12	12	12				
DYS534	46	15	15	15	15	15	15	15	15	15				
DYS450	47	8	8	8	8	8	8	8	8	8				
DYS444	48	12	12	12	12	12	12	12	12	13				
DYS481	49	22	22	22	22	22	22	22	22	22				
DYS520	50	20	20	20	20	20	20	20	20	20				
DYS446	51	14	14	14	14	14	15	14	14	14				
DYS617	52	12	12	12	12	12	12	12	12	12				
DYS568	53	9	9	9	9	9	9	9	9	9				
DYS487	54	13	13	13	13	13	13	13	13	13				
DYS572	55	11	11	10	11	11	11	11	11	11				
DYS640	56	11	11	11	11	11	11	11	11	11				
DYS492	57	12	12	12	12	12	12	12	12	12				
DYS565	58	11	11	11	11	11	11	11	11	11				
DYS710	59	34	34	34	34	33	34	34	34	33				
DYS485	60	15	15	15	15	15	15	15	15	15				
DYS632	61	9	9	9	9	9	9	9	9	9				
DYS495	62	16	16	16	15	16	16	16	16	16				
DYS540	63	12	12	12	12	12	12	12	12	12				
DYS714	64	27	27	27	27	27	27	27	27	27				
DYS716	65	27	27	27	27	27	27	27	27	27				
DYS717	66	18	18	18	18	18	18	18	18	18				
DYS505	67	12	12	12	12	12	12	12	12	12				
DYS556	68	11	11	11	11	11	11	11	11	11				
DYS549*	69	12 or 13	13	13	13	13	12	12	12	12				
DYS589	70	12	12	12	12	12	12	12	12	12				
DYS522	71	10	10	10	10	10	10	10	10	10				
DYS494	72	9	9	9	9	9	9	9	9	9				
DYS533	73	13	13	13	13	13	13	13	13	13				
DYS636	74	12	12	12	12	12	12	12	12	12				
DYS575	75	10	10	10	10	10	10	10	10	10				
DYS638	76	11	11	11	11	11	11	11	11	11				
DYS462	77	11	11	11	11	11	11	11	11	11				
DYS452	78	30	30	30	30	30	30	30	30	30				
DYS445	79	12	12	12	12	12	12	12	12	12				
Y-GATA-A10	80	13	13	13	13	13	13	13	13	13				
DYS463	81	24	24	24	24	24	24	24	24	24				
DYS441	82	13	13	13	13	13	13	13	13	13				
Y-GGAAT-1B07	83	10	10	10	10	10	10	10	10	10				
DYS525	84	10	10	10	10	10	10	10	10	10				
DYS712	85	20	20	20	20	20	20	20	20	20				
DYS593	86	15	15	15	15	15	15	15	15	15				
DYS650	87	22	22	22	22	22	22	20	20	21				
DYS532	88	13	13	13	13	13	13	14	14	14				
DYS715	89	25	25	25	24	25	25	25	25	25				
DYS504	90	17	17	17	17	17	17	17	17	17				
DYS513	91	12	12	12	12	12	12	12	12	12				
DYS561	92	15	15	15	15	15	15	15	15	15				
DYS552	93	24	24	24	24	24	24	24	24	24				
DYS726	94	12	12	12	12	12	12	12	12	12				
DYS635	95	23	23	23	23	23	23	23	23	23				
DYS587	96	18	18	18	18	18	18	18	18	18				
DYS643	97	10	10	10	10	10	10	10	10	10				
DYS497	98	14	14	14	14	14	14	14	14	14				
DYS510	99	17	17	17	17	17	17	17	17	17				
DYS434	100	9	9	9	9	9	9	9	9	9				
DYS461	101	12	12	12	12	12	12	12	12	12				
DYS435	102	11	11	11	11	11	11	11	11	11				

DYS549*	69	12 or 13	arbitrarily chose 13 as ancestral DYS value: 4 individuals have a value of 12 and 4 have a value of 13

Table 6: STR marker values for 12 of the participants of the Hubbell Surname Project. DYS549 (test #69) had a value of 12 in four individuals and a value of 13 in four individuals. An ancestral value was not assigned

STR Name	test #	Ancestral Value	H.R. Hubble	N. Hubble	W. Johnston
Genetic Distance			7	7	10
Date Tested			2019	2019	2019
Haplogroup			R-FT164982	R-FT164982	R-FT75787
DYS385	5	11-14	11-15	11-15	11-14
DYS389i	9	13	13	13	14
DYS392	10	13	12	12	12
DYS389ii	11	27	27	27	28
DYS458	12	18	17	17	18
DYS464	20	14-15-17-17	14-15-17-17	14-15-17-17	15-15-17-17
CDY	28	36-37	36-37	36-37	37-37
DYS444	48	12	12	12	13
DYS710	59	34	34	34	33
DYS549	69	13	12	12	12
DYS650	87	22	20	20	21
DYS532	88	13	14	14	14

STR Name	test #	Ancestral Value	B. Hubbell	R.C. Hubbell
Genetic Distance			2	2
Date Tested			2019	2019
Haplogroup			R-M269	R-FT75886
DYS557	42	17	17	18
DYS710	59	34	34	33
DYS495	62	16	15	16
DYS715	89	25	24	25

STR Name	test #	Ancestral Value	R.W. Hubble	B.M. Hubble
Genetic Distance			0	3
Date Tested			2019	2019
Haplogroup			R-FT75787	R-FT115348
DYS439	8	12	12	13
CDY	28	36-37	36-37	36-38
DYS572	55	11	11	10

STR Name	test #	Ancestral Value	M.R. Hubbell
Genetic Distance			2
Date Tested			2009
Haplogroup			R-M269
DYS576	26	18	19
CDY	28	36-37	37-37

STR Name	test #	Ancestral Value	D.C. Hubbell
Genetic Distance			2
Date Tested			2019
Haplogroup			R-FT75787
DYS446	51	14	15
DYS549	69	12 or 13	12

Summary

1) H.R. & N. (father & son) have 102 identical STR values
2) H.R., N. and W. Johnston all descend from #2542 Steinman. W. Johnston shares 3 identical STR mutations with H.R. and N. (#10, 69, 88); has 7 unique mutations (#9,11,20,28,48,59,87) while H.R. & N. have 3 mutations that W. Johnston does not (#5,12,87)
3) B. & R.C. are 1st cousins: they have 4 unique STRs between them
4) R.W. & B.M. have 3 different STRs between them and they share a common ancestor 7 generations in the past (Justus Hubble)
5) D.C. shows only 2 STR mutations from ancestral value
6) M.R. was only tested for 37 STR values

Table 7: A comparison of the mutated STR markers between related individuals

Age of Father When Son Was Born

ME#		age of father when son was born
	D.C. b: 1930	
6553	Russell b: 1897	33
3543	Sylvester b: 1867	30
1591	Augustus b: 1827	40
650	Robert b: ~1786	41
209	Richard b: 1751	35
55	John b: 1709	42
17	Richard b: 1683	26
2	John b: ~1652	29
1	Richard b: 1625	27
	ave:	34

ME#		age of father when son was born
	R.C. b: 1949	
14081	Charles Arnold b: 1913	36
10479	Clyde b: 1883	30
7018	William Henry b: 1856	27
3876	Stephen L b: 1825	31
1792	John G b: 1791	34
732	Sullivan b: 1769	22
239	John b: 1746	23
72	Peter b: 1715	31
19	Peter b: 1686	29
3	Richard b: 1654	32
1	Richard b: 1625	29
	ave:	29

ME#		age of father when son was born
	B. b: 1957	
14085	Basil b: 1922	35
10478a	Clyde b: 1883	39
7018	William Henry b: 1856	27
3876	Stephen L b: 1825	31
1792	John G b: 1791	34
732	Sullivan b: 1769	22
239	John b: 1746	23
72	Peter b: 1715	31
19	Peter b: 1686	29
3	Richard b: 1654	32
1	Richard b: 1625	29
	ave:	30

ME#		age of father when son was born
	N. b: 1974	
	Hilbert b:1933	41
8240	Burt Samuel b: 1907	26
4959	Hezekiah R b: 1843	64
2542	Steinman b: 1819	24
1088	Asa b: ~1780	39
398	Hezekiah B b: 1760	20
126	Asa b: 1736	24
26	Nathaniel b: 1702	34
3	Richard b: 1654	48
1	Richard b: 1625	29
	ave:	35

ME#		age of father when son was born
	R.W. b: 1949	
12667	William b: 1910	39
9022	Claude b: 1884	26
5524	William F b: 1852	32
2903	Edward b:1826	26
1239	William H b: 1795	31
464	Joel b:1771	24
150	Justus b: 1744	27
32	David b: 1698	46
5	Samuel b: 1657	41
1	Richard b: 1625	32
	ave:	32

ME#		age of father when son was born
	B.M. b: 1982	
16151	Michael R b: 1957	25
12551	Robert D b: 1933	24
8892	Gilbert b: 1899	32
5460	Tiberius b: 1877	22
2869	Henry T b: 1850	27
1228	William S b: 1793	57
463	Eliphalet b:1769	24
150	Justus b: 1744	29
32	David b: 1698	46
5	Samuel b: 1657	41
1	Richard b: 1625	32
	ave:	33

Average of all six values: 32

Table 8: Age of father at the time of the birth of his son.

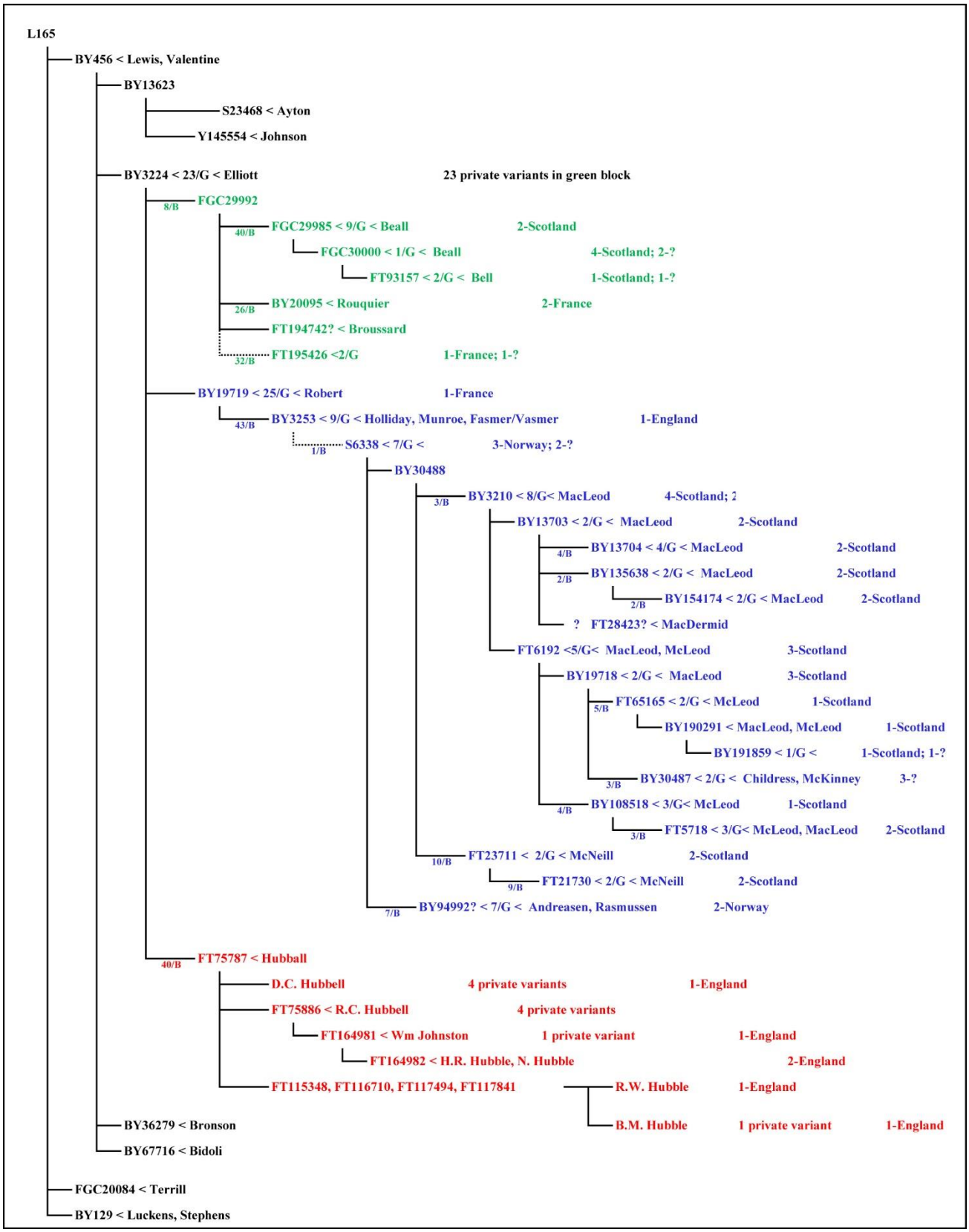


Figure 13: Haplotree of the R-L165 haplogroup displaying all currently known mutations of L165 (2020) and all known subclades with associated Family Names. The Hubbell line is from L-165 to BY456 to BY3224 to FT75787 (in red). The number preceding a mutation represents the number of mutations that occur concurrently. For example, between BY3224 and FT75787 there are 40 additional mutations. Original diagram courtesy of Tim McLeod.

Appendix B: Meiosis and Recombination

Although the cells from the inside of your mouth are the cells that are tested, they originate from your parent's reproductive cells through a process called Meiosis.

Meiosis begins like mitosis where the cell makes a copy of each chromosome (DNA duplication). In mitosis this cell, with double the amount of DNA, would divide forming two cells, each with the normal amount of DNA. However, during meiosis the duplicated chromosomes line up and exchange large segments of DNA in a process called **recombination (DNA cross-over)**.

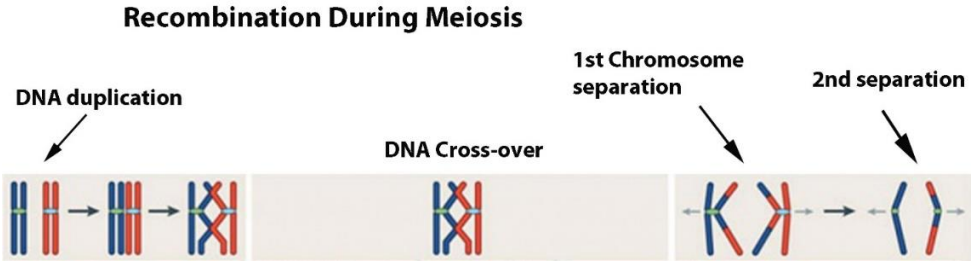


Figure 14: Recombination During Meiosis

The newly recombined chromosome pairs are then divided into two daughter cells (1st chromosome separation). Then the paired chromosomes are pulled apart (2nd separation) into a total of four reproductive cells (either eggs or sperm). Each of these cells has one copy each of the 23 chromosomes, all with a unique combination of DNA segments. During fertilization, the egg and sperm each contribute one copy of each chromosome so in the fertilized egg the chromosomes are paired again. The Hubbell website has a video of meiosis that can be viewed at: www.hubbell.org/DNA/DNA: Mitosis vs Meiosis.

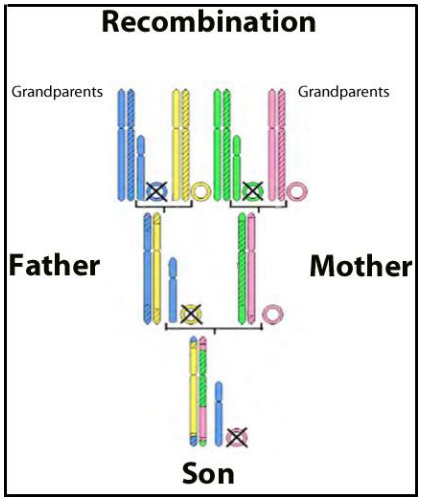


Figure 15: Illustrating the impact of recombination.

Figure 15 illustrates the impact of recombination on inherited DNA. The Mother's DNA contains DNA segments from both her mother (pink solid and pick hashed) and her father (green solid and green hashed). She also inherited her mother's mitochondria containing the m-DNA intact (pink circle).

The Son has inherited some of his maternal Grandmother's DNA (pink) and some of his maternal Grandfather's DNA (green). He has also inherited some DNA from his paternal Grandmother (yellow) and paternal Grandfather (blue). He also inherited his father's intact Y-chromosome (short blue).

In fact, the Son would have fragments of DNA from both his father's ancestors and his mother's ancestors going back about 5 to 7 generations. Further back than that, the older segments of DNA have been diluted out. It is for this reason that an at-DNA test can only detect shared DNA similarities for a few hundred years into the past.