HASKELL JOURNAL

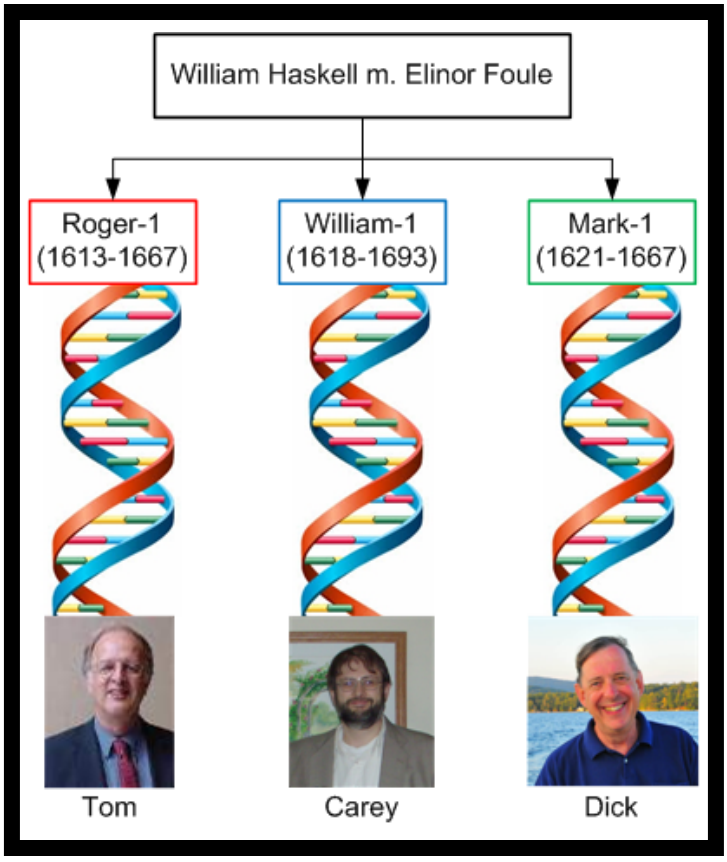
The Journal of the Haskell Family Association

Vol. XXIX — No. 1 **Spring 2013** **Issue 89**

Inside this issue:

The Haskell DNA

by Richard E. Haskell



It is easy to understand how most Haskells in the USA are descended from one of the three brothers, Roger, William, or Mark, and therefore have William Haskell (1577-1630), who married Elinor Foule, as a common ancestor. What is more surprising to many is that recent DNA studies of the Y chromosome (which occurs only in males) point to the fact that all living males today are descended from a single male, nicknamed Y-Adam, who lived in Africa about 60,000 years ago (and maybe up to 100,000 years ago). Similarly, studies of the mitochondrial DNA (which is inherited only from a mother) leads to the conclusion that all females living today are direct descendants through the maternal line of a single female, called mitochondrial Eve, who also lived in Africa, but earlier than Y-Adam, somewhere between 120,000 and 200,000 years ago.

(Continued on page 4)

(Continued from page 1)

Haskell DNA

We will return to Y-Adam and mitochondrial Eve later, but for now, let's go back to the beginning of my DNA experience.

DNA Testing

At the Deer Isle reunion in 2008, Tom Haskell brought with him some Y-DNA 37 Marker kits from Family-TreeDNA to hand out to interested members. Tom was going to use one to test himself, so I thought, "What the heck," and took one of the kits. When I got back to Michigan, I opened the kit, took out the toothbrush-like stick, and swabbed the inside of my cheek. I sealed up the resulting swab and mailed it to FamilyTreeDNA.

A few weeks later I received an email from Family-TreeDNA with a link to my own FamilyTreeDNA web page, which contained my results. From that web page I could print out my certificate, which is shown on page 7. It contains 37 numbers, printed out in three rows. I thought, "What do all these numbers mean?" However, also on my DNA web page, it told me that I matched three other Haskells, including Charles T. Haskell, III—i.e. "Tom". As we'll see, I didn't match Tom exactly on all 37 numbers, but we did match exactly on 35. FamilyTreeDNA said that there was a 98.89% probability that Tom and I shared a common ancestor 10 generations ago. Since I knew that 10 generations ago our common ancestor was, in fact, William Haskell, the father, (Tom is descended from the brother, Roger, and I am descended from the brother, Mark), it was at least gratifying to know that no "paternity surprises" had occurred on either of our family tree branches!

But I still wondered what these 37 numbers really meant. To find out, I had to learn a little about DNA.

A Little DNA Background

DNA stands for DeoxyriboNucleic Acid (dee-ox-ee-rye-boh-new-clee-ic acid) and has the double-helix structure shown on the cover page. This structure was discovered by James Watson and Francis Crick in 1953. The DNA molecule is very long and skinny — if you could unwrap it, it would be about 6 feet long. But it is wrapped up and twisted into a small size that fits inside the nucleus of every cell in your body. This is called nuclear DNA. There is another small amount of DNA, called mitochondrial DNA, or mtDNA, which is found in the mitochondria outside of the nucleus. We will return to mtDNA later.

If you unwrap the helix shapes shown on the cover page, the structure would look like a ladder. The sides of the ladder are made up of alternating sugar (Deoxyribose)-phosphate pairs and form the backbone of the DNA molecule. The rungs of the ladder are made up of *base pairs* (or *nucleotides*). There are four possible bases, called *adenine* (A), *guanine* (G), *thymine* (T), and *cytosine* (C). However, the base A can only be connected to the base T, and the base G can only be connected to the base C. On the DNA structure shown on the cover page, you will notice that the rungs of the ladder are either a yellow (A) − green (T) pair, or a red (G) − blue (C) pair. Because of this pairing requirement, if you know the sequence of bases along one of the backbone strands, you will automatically know the sequence along the other, since A will always be opposite T, and G will always be opposite C, and vice versa. This is important for DNA *replication*, where the DNA molecule is "unzipped" between the two bases on the ladder rungs, and new "complementary bases" come in along with another backbone side to form two new copies of the original DNA molecule.

A sequence of bases along one of the backbones is therefore enough to define the entire DNA molecule. The two backbone sides have directions associated with them and are anti-parallel. The DNA sequence is read in a particular direction. Thus, a DNA sequence is just a sequence of the four letters, A, T, G, and C, such as, TCGATCCTAGCCTA. The human DNA consists of about 3 billion base pairs, over 99% of which are the same in all people.

The human DNA is packaged into 46 chromosomes, made up of 23 chromosome pairs. The lengths of these chromosome pairs vary from about 16 mm for chromosome number 21 with over 48 million base pairs to about 85 mm with over 248 million base pairs for chromosome number 1. The first 22 chromosome pairs (1—22) are called *autosomes*, and you inherit one copy from your father and one copy from your mother. Chromosome number 23 is the sex chromosome. If you are a male, you have one X chromosome and one Y chromosome (XY). If you are a female, you have two X chromosomes (XX). You inherit one of these from your mother and one from your father. However, since your mother had two X chromosomes, you must get an X chromosome from her. However, your father has an X and a Y, so you have a 50-50 chance of getting an X or a Y from your father. If you get the X, you are a female and if you get the Y, you are a male.

The DNA within a chromosome is further divided into sequences called *genes*, that contain the genetic code to produce various proteins. The entire set of 46 chromosomes is called the *genome*. The process of copying the genetic code in a gene (the string of A, T, C, and Gs) to an RNA molecule, called *transcription*, and then form-

(Continued on page 5)

(Continued from page 4)

Haskell DNA

ing a specific protein from 3-letter-coded amino acids is a fascinating one, which is worth watching on a number of YouTube videos. The following link is one good example showing a computer-generated animation in real time:

http://www.youtube.com/watch?v=yqESR7E4b_8

The U.S. Human Genome Project began in 1990 and was completed in 2003. Some of the goals of the project were to determine the sequences that make up the 3 billion base pairs of the human DNA and to identify the approximately 20,000—25,000 genes in the human DNA. It was found that less than 2% of the genome provides the codes used for generating proteins, and that repeated sequences that are not codes for producing proteins, called "junk DNA," make up at least 50% of the human genome.

The Haskell DNA Results

We have seen that if you are a male, you got your Y chromosome from your father. And your father got his Y chromosome from his father, and so forth back through the direct male line of ancestry. If everyone's Y chromosome were the same, it would be useless for genealogy purposes. However, while the DNA replication process is very accurate, every now and then a mistake will be made, called a *mutation*, that will be passed on to the next generation. A Y-DNA *marker* is the location along the Y chromosome where these mutations occur. The DYS numbers on my certificate on page 7 are numbers that refer to a particular marker, or location along the Y chromosome. The list of 37 numbers under each DYS number is called my *haplotype*. Each of these numbers is referred to as an *allele*.

The particular category of DNA mutations used in the Y-DNA 37 test is called *Short Tandem Repeats*, or STRs. My allele number is the number of repeats of a particular DNA sequence at that marker location. As an example, marker DYS#391 repeats the DNA sequence TCTA from 6 to 14 times. From my certificate on page 7, the fourth entry on the first row shows my allele value for this marker to be 10. Therefore, my sequence would look something like the following:

GTCTGTCTG/**TCTA/TCTA/TCTA/TCTA/TCTA/ TCTA/TCTA/TCTA/TCTA/TCTA/**TCTGCCTATC

where the sequence TCTA repeats 10 times.

Studies of over 36,000 samples showed that an allele value of 10 is the most common for this marker, occurring 55.8% of the time. The studies also showed that the mutation rate for this marker is 0.00265, or 0.265%. This means, on the average, for this particular marker you would expect one change in 377 generations.

Other markers have different mutation rates. For example, marker DYS# 464 has an average mutation rate of 0.00566, which corresponds to one change in 176 generations, and marker CDY (also know as DYS# 724) has an average mutation rate of 0.03531, which corresponds to one change in 28 generations. Remember, these are all averages, and actual rates for a particular ancestry line could vary considerably. When comparing two individuals to see if they have a high probability of a common ancestor within so many generations, it is necessary to compare multiple markers. Early Y-DNA testing often used 12 or 25 markers. Now the most common is 37 markers, which is the test I took with the results shown on page 7.

As I mentioned earlier, FamilyTreeDNA told me that I matched three other Haskells with a high probability of a common ancestor within 10 generations. The differences between my haplotype and the other three occurred in the DSY# 439, DSY# 464, and CDY markers. The DSY# 464 and CDY markers are so-called "multi-copy" markers and are known to have a higher mutation rate than other markers. The CDY marker has two copies (*a* and *b*), and the DSY# 464 marker has four copies (*a*, *b*, *c*, and *d*). The allele values of the DSY# 439 and the two multi-copy markers for the four Haskell tests are shown in Table 1. On all other markers, we matched 100%.

Table 1 Comparison of Haskell Alleles

Marker	A	B	C	D
439	12	12	12	11
CDYa	36	38	37	36
CDYb	37	38	37	37
464a	15	15	15	15
464b	15	15	15	15
464c	17	17	16	17
464d	18	18	18	18

In Table 1, I am Haskell-A, descended from the brother, Mark and Tom is Haskell-B, descended from the brother, Roger. The other two Haskells, labeled C and D are both descended from the brother William. Haskell-C is descended through the William-Joseph line and Haskell-D is descended through the William-

(Continued on page 6)

(Continued from page 5)

Haskell DNA

William line. The allele values that differ from mine in column 1 are circled in Table 1. Note that the only difference between A and D is marker 439, which differs by 12 − 11 = 1. The *genetic distance* between A and D is therefore 1. There are two marker differences between A and C, CDYa and 464c. Each has an absolute value difference of 1 and, therefore, the total genetic distance between A and C is 2. The difference between A and B occurs only with the multi-copy marker, CDY. The difference in CDYa is 2 and the difference in CDYb is 1. FamilyTreeDNA calls this a total genetic distance of 3. However, some researchers believe that a mutation will occur in all copies of a multi-copy mutation as one event, and would classify the genetic distance between A and B as 1. Another property of the multi-copy markers is that the allele values are always reported in sorted order from smallest to largest. This means that apparent differences may not be real differences. As long as any of the copies have the same allele value, that difference will be zero.

From Table 1 there are hints, but no proof, of what the allele values would have been for William Haskell, the father of the three brothers. For marker 439 it looks as if D is the odd man out and that 12 is the original value for William. For the CDY marker, it is not so clear. The original value may have been 37 for both copies and B increased both by 1, while A and D decreased CDYa by 1. But there could be other explanations, and it will take more Haskell DNA samples to determine when and where in the family tree particular mutations occurred. Again, from the limited four samples we have, a good candidate for the four 464 samples would be 15-15-17-18. Only column C in Table 1 differs from this sequence.

Haplogroups

In addition to sending me the certificate on page 7, FamilyTreeDNA told me that my *haplogroup* was R1b1a2, also known as R-M269. A haplogroup is a group of similar haplotypes that have a common distant ancestor. The haplogroups are defined in terms of a different kind of mutation, called a *Single Nucleotide Polmorphism*, or SNP. In this case, only a single base in the Y-DNA (A, T, C, or G) changes by chance. For example, what should be a T might get passed on to the next generation as a C. Such an occurrence is very rare and it normally happens only once at a particular loci, or location, on the Y chromosome. This means that all future descendants of the person who first gets this mutation will have this mutation. If most of the population has a particular SNP, it must be a very ancient mutation.

The SNP mutations are designated by a letter and number, and the R1b1a2 haplogroup is characterized by having the M269 SNP mutation. It is estimated that this mutation occurred between 4,000 and 10,000 years ago, and is a sub-group of the R1b haplogroup, which is the most common haplogroup in European populations, is associated with the marker M343, and dates from about 25,000 years ago. The R1b haplogroup is itself a sub-group of the Haplogroup R, which has the marker M207 and is believed to originate in central or south Asia 20,000—34,000 years ago. Four or five additional markers from even earlier times takes us back to our Y-Adam in Africa over 60,000 years ago. A map showing the migration of the different haplogroups out of Africa and into the rest of the world is shown on page 7.

Other DNA Tests

We have stressed the Y-DNA test, because that is the one I took and follows the male line of ancestry. There is another test, called the mtDNA test, which follows the maternal line of ancestry. The mitochondria DNA is passed to the next generation only through the mother's egg. This test is primarily used to find ethnic ancestry and produces migration maps of haplogroups similar to the figure on page 7.

There are also DNA tests that you can take that measure SNP mutations in the autosomal chromosomes and are therefore not limited to the paternal or maternal line, but can include mutations that come from any one of your ancestors. The National Geographic Society is sponsoring the Genographic Project, and their Geno 2.0 DNA kit is designed to find your deep ancestry. It will tell you what percent of you ancestry is European, African, and East-Asia/Native American. FamilyTreeDNA has a Family Finder autosomal DNA test, which will provide the same information, and in addition, allow you to find relationships within the past 4 - 5 generations. (This is their most expensive test!)

The Haskell HFA DNA Project

The Haskell Family Association has started its own DNA project within FamilyTreeDNA. Such a project provides discounts for the Y-DNA 37 test. There is another DNA project within FamilyTreeDNA, which includes Haskell, Peterson, and Soule surnames, and you may wish to join that project as well. The main goals of the Haskell HFA DNA projects are

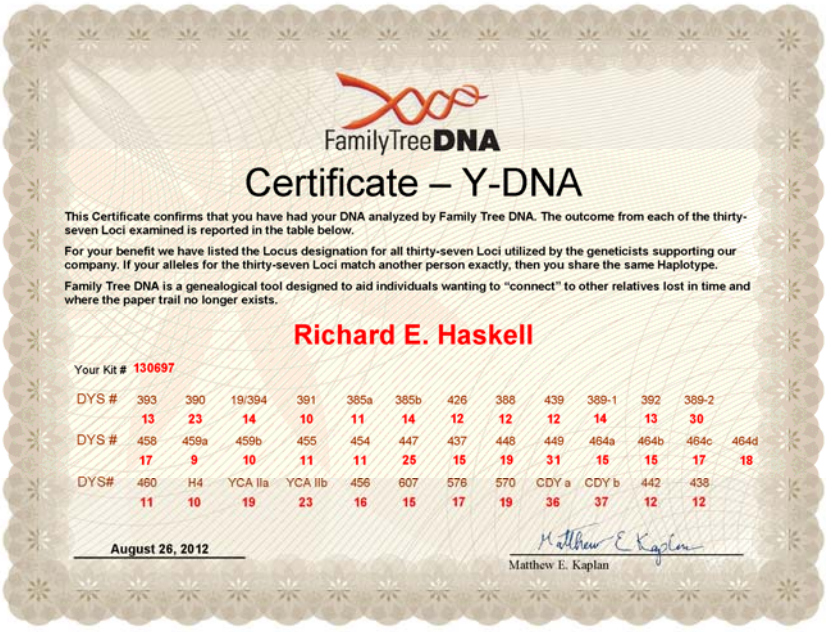
- Use DNA testing to confirm the lines of descent from William Haskell and Elinor Foule in our extensive database.

(Continued on page 7)

(Continued from page 6)

Haskell DNA

- Identify the location within our family tree of the various STR mutations which have occurred since the time of William Haskell (1577-1630). Such a detailed classification of DNA results may help to place some of our "Haskell orphans" in our Haskell family tree.
- Compare Haskell North American DNA tests with tests from Haskells in the United Kingdom and see if the Haskell line can be pushed back to the 12th or 13th century. In the latest issue of the IHFS Newsletter, Peter P. Haskell has an interesting article about Housecarls and suggests that the apple tree in the Haskell crest may be associated with the region of England known as Appletree where Henry de Ferrers lived in the 11th century, and he and his descendants had "housecarls." It would be interesting to know if there are any direct male descendants from Henry de Ferrers. If those descendants and others whose ancestors have lived in the Appletree region of England took the Y-DNA 37 test, and if the results showed a distant common ancestor to Haskells, then the theory of the Housecarl origin of the Haskell name would gain credence. That would be very interesting!



FamilyTreeDNA

Certificate – Y-DNA

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

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

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

Richard E. Haskell

Your Kit # 130697

DYS #	393	390	19/394	391	385a	385b	426	388	439	389-1	392	389-2	
	13	23	14	10	11	14	12	12	12	14	13	30	
DYS #	458	459a	459b	455	454	447	437	448	449	464a	464b	464c	464d
	17	9	10	11	11	25	15	19	31	15	15	17	18
DYS#	460	H4	YCA IIa	YCA IIb	456	607	576	570	CDY a	CDY b	442	438	
	11	10	19	23	16	15	17	19	36	37	12	12	

August 26, 2012

Matthew E. Kaplan

My Y-DNA Certificate from FamilyTreeDNA

More Reading:

There are three books that I found particularly interesting related to DNA testing for genealogy.

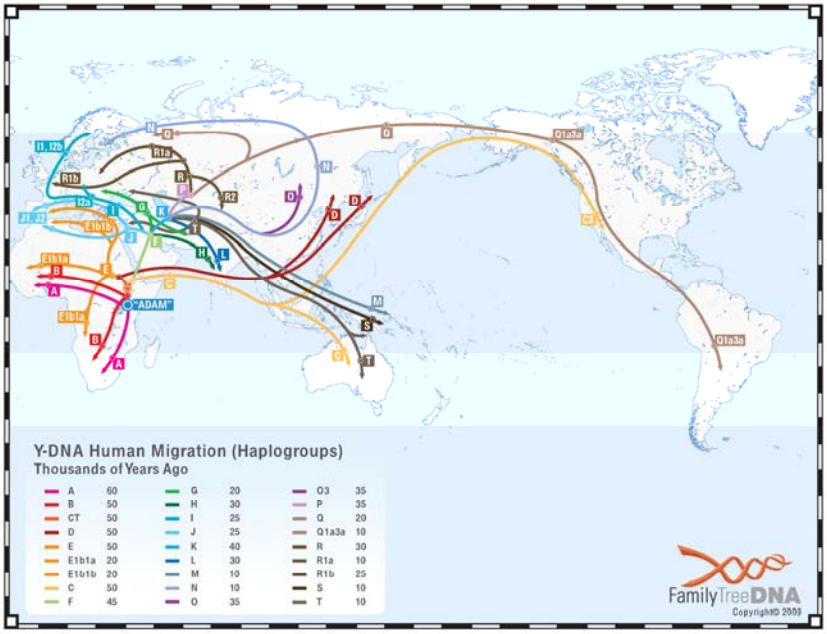
1. *Trace Your Roots with DNA—Using Genetic Tests to Explore Your Family Tree*, by Megan Smolenyak Smolenyak and Ann Turner, Rodale, Inc., 2004.

2. *The Journey of Man—A Genetic Odyssey*, by Spencer Wells, Random House, NY, 2003.

3. *DNA USA—A Genetic Portrait of America*, by Bryan Sykes, Liveright Publishing Corp., NY, 2012.

The following website provides more information on deep ancestry DNA, and includes a page and three short videos on the genetic science behind the Genographic Project:

https://genographic.nationalgeographic.com/science-behind/



Reproduced with permission from FamilyTreeDNA

To order a Y-DNA 37 kit and join the Haskell HFA project, go to www.familytreeDNA.com and click on the Projects tab. In the Y-DNA SURNAME PROJECTS section, click on H and scroll down until you find **Haskell HFA**. Click on this link and order the Y-DNA 37 kit. This link will give you the current discounted project price.