Genetree

mtDNA MOLECULAR GENEALOGY REPORT

William Walsh

2012-09-25

Your Personal Genetic Report

Genetree

GeneTree (www.genetree.com) is a unique genetic genealogy family networking site designed to help people understand where their personal histories belong within the greater human genetic story. The GeneTree site creates unparalleled opportunities for unlocking genetic heritage and identity, connecting with ancestors and living relatives, and sharing meaningful information and experiences to help preserve family histories. By expanding the concept of family, GeneTree can provide fresh new answers and perspective to the universal questions, "Who am I?" and "Where do I come from?"

Collaborative Power
GeneTree utilizes exceptional tools and resources from:

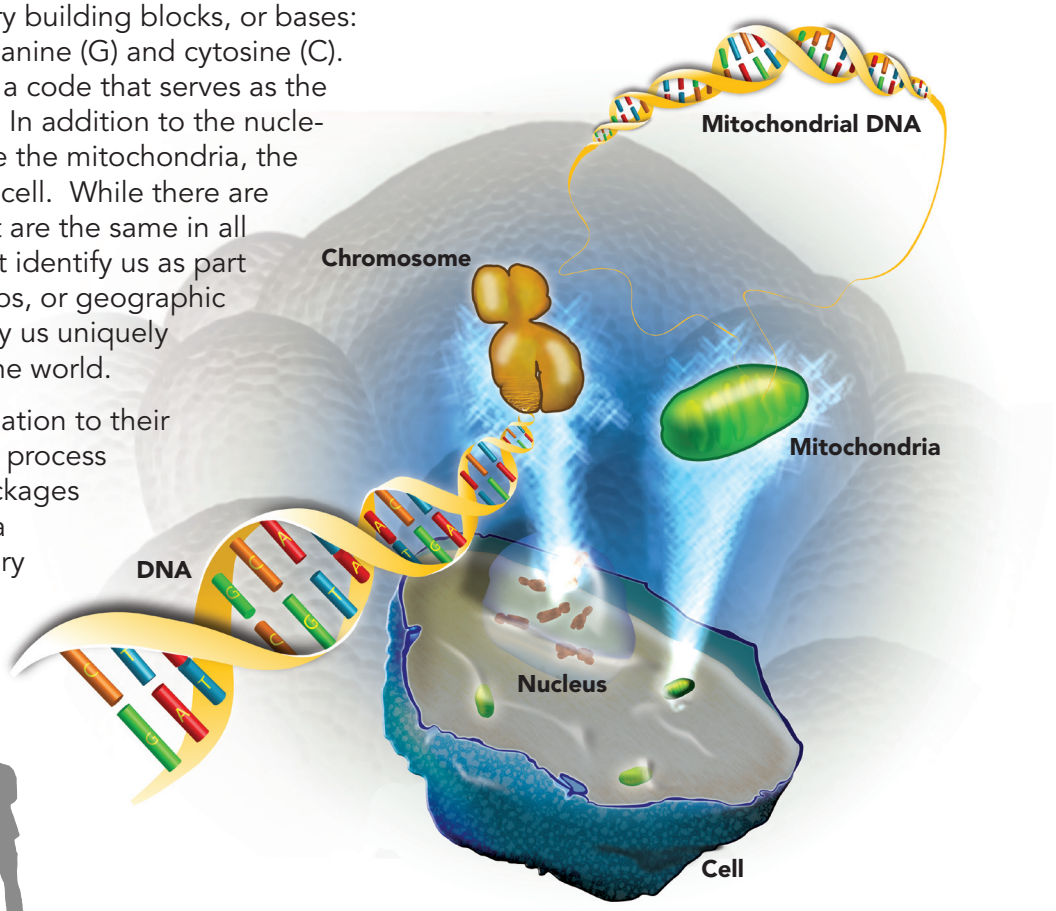
> **The Sorenson Molecular Genealogy Foundation (SMGF)**, a non-profit organization building the world's foremost collection of DNA samples correlated with genealogical information, gathered from individuals in more than 170 nations. www.smgf.org

Get Started
This comprehensive report will provide you with the tools you need to begin making the most of your DNA testing experience. Below is an outline of the report for quick reference.

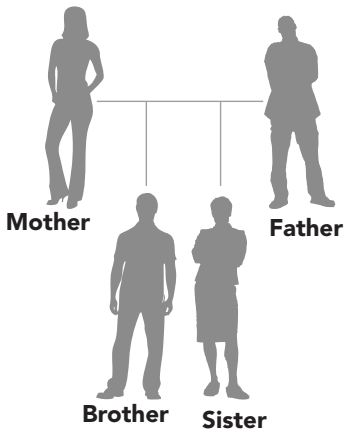
- **Education**
  - → Introduction to DNA
  - → Y chromosome DNA testing
  - → Mitochondrial DNA testing
  - → Education Summary: What is genetic genealogy?
- **Your Genetic Profile**
  - → Test Results
  - → Haplogroup prediction
- **What now? Exploring GeneTree Services**
  - → Database matching
  - → Pedigree Tool
  - → DNA Navigator
  - → Build Your Family Tree

A Brief Introduction to DNA and Heredity

Deoxyribonucleic acid (DNA) is a chemical found mainly in the nucleus of the cell, where it is packaged in units called chromosomes. Each individual has 23 pairs of chromosomes, for a total of 46. Each strand of DNA is made of four primary building blocks, or bases: adenine (A), thymine (T), guanine (G) and cytosine (C). These DNA bases make up a code that serves as the genetic blueprint for all life. In addition to the nucleus, DNA also is found inside the mitochondria, the energy-producing part of a cell. While there are many parts of our DNA that are the same in all humans, there are parts that identify us as part of certain large ethnic groups, or geographic areas, and parts that identify us uniquely apart from anyone else in the world.



Parents pass genetic information to their children via DNA. During a process called meiosis, a parent packages half of his or her DNA into a sperm or egg cell for delivery to their offspring. Therefore, each child receives 23 chromosomes from its father, and 23 different, but corresponding, chromosomes from its mother.

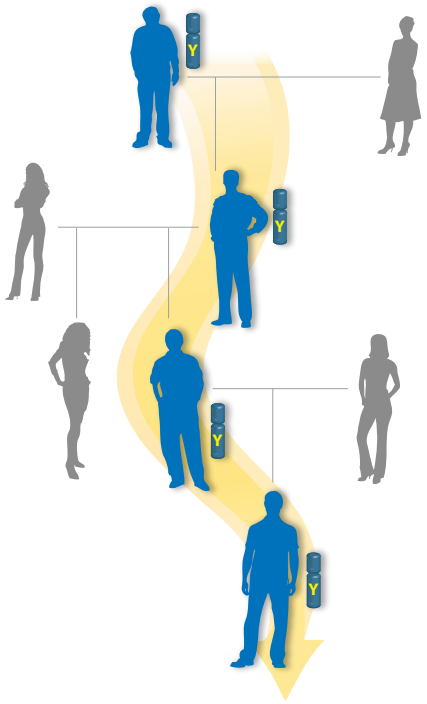


Because each parent's DNA was likewise inherited from their parents, the offspring of two individuals actually carries the DNA of not only their parents, but of their grandparents, and great-grandparents, and so on into the past.

A **family tree**, also called a pedigree, is a diagram of relationships in a family. It serves as an informational framework for understanding genetic relationships and can allow one to see a great deal of information about a family at a glance.

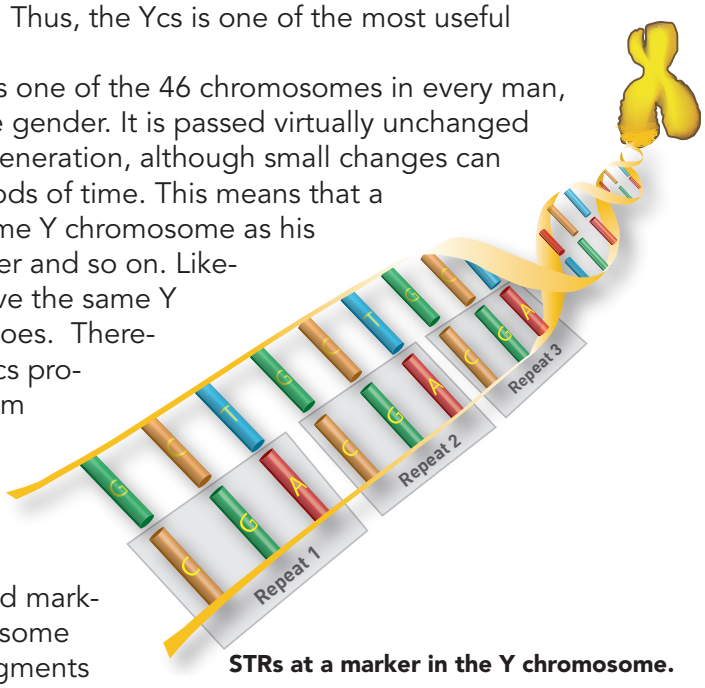
Detecting Paternal Relationships Using The Y Chromosome

Information contained in the Y chromosome (Ycs) can be used to identify male ancestral lines. Because the Ycs of paternally-related individuals are nearly identical, scientists can use information from this chromosome to detect relationships in a male line. Thus, the Ycs is one of the most useful chromosomes for genealogical studies.



The Y chromosome is passed from generation to generation in males.

The Y chromosome is one of the 46 chromosomes in every man, and determines male gender. It is passed virtually unchanged from generation to generation, although small changes can occur over long periods of time. This means that a man will have the same Y chromosome as his father and grandfather and so on. Likewise, his sons will have the same Y chromosome as he does. Therefore, while a man's Ycs profile cannot identify him uniquely, it does connect him to his direct paternal line.



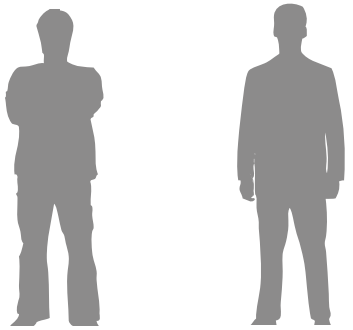
STRs at a marker in the Y chromosome.

Many locations, called markers, on the Y chromosome contain repetitive segments of DNA called **Short Tandem Repeats** (STRs). These are small segments of DNA, usually 2-5 bases long that repeat the same pattern numerous times. All males have these STR regions and some locations have the same number of STRs in most people while others show more variability. By selecting locations on the Ycs that vary, scientists can use them to distinguish paternal ancestral lines.

These variations in the number of repeated segments at each marker are called **alleles**. For example, at Marker B, one individual might have 11 STRs, called allele 11, while another might have 14 STRs, or allele 14. These differences are particularly found in people who are not related.

A set of allele values at particular markers along a chromosome is called a **haplotype**. To find genetic ancestors, GeneTree compares haplotypes, looking for matches. If you are male, the STRs on your Y chromosome were analyzed to determine your haplotype.

Our partner SMGF is actively engaged in Y-DNA research and has developed the world's foremost repository of Y-DNA data and correlated genealogies. GeneTree is a portal to accessing this database, though you can search it independently for no cost at www.smgf.org.



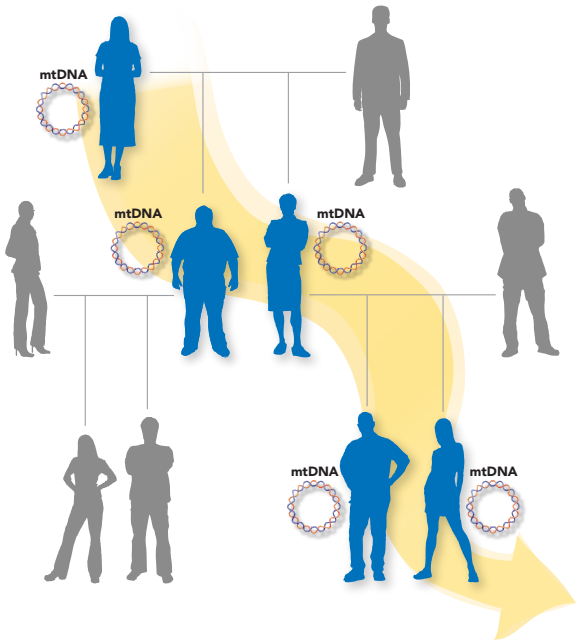
Individual One	Individual Two
Marker A: Allele 15	Marker A: Allele 10
Marker B: Allele 11	Marker B: Allele 14
Marker C: Allele 9	Marker C: Allele 9

Detecting Maternal Relationships Using Mitochondrial DNA

Mitochondrial DNA (mtDNA) is widely used in determining maternal ancestral lines. This small (16,000 bases) circular segment of DNA is contained in the mitochondrion, which are found in the mother's egg but not in the father's sperm.

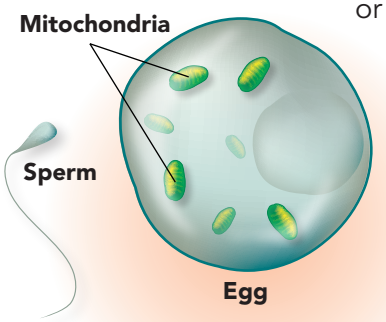
The mitochondrial inheritance system functions in a manner similar to the Y chromosome system, in that mtDNA is passed on relatively unchanged from one generation to the next. Like Y chromosome information, it also is relatively easy to analyze and use.

Unlike the Y chromosome, which is passed on from a father to his sons, mitochondrial DNA is transmitted from a mother to both her daughters and sons. However, only her daughters (not her sons) will pass mtDNA on to their children. This enables genetic genealogists to use mtDNA to reconstruct maternal ancestral relationships.



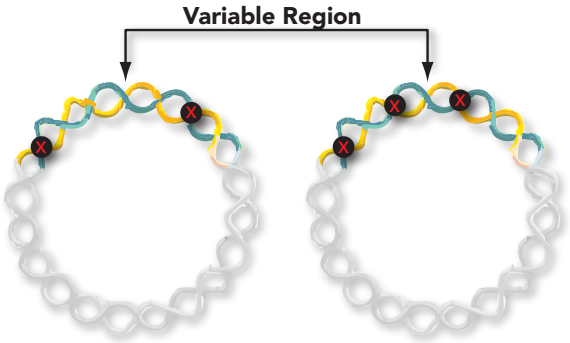
Females pass mitochondrial DNA on to their male and female children.

Mitochondrial DNA test results are reported as the differences, or **mutations**, in the sequence of bases when compared to a standard reference sequence called the Cambridge Reference Sequence (CRS). Mutations occur when a single base is substituted for another in the DNA sequence, such as an A for a T. Bases also can be inserted into or deleted from the DNA sequence during a mutation event.



Mitochondrial DNA originates solely from the mother's egg.

Mutations occur less frequently in mitochondrial DNA than in Y chromosome DNA. For example, in 12,000 pairs of individuals we would expect to see one mutation in about every 100 generations. Usually between 1 and 15 mutation differences in mtDNA will be found between any two randomly chosen individuals.



Mutations in Individual A's mitochondrial DNA

Mutations in Individual B's mitochondrial DNA

Most commercial mtDNA test cover bout 1,500 bases in regions known to be highly variable in the mitochondrial DNA called Hypervariable Region I (HVRI), Hypervariable Region II (HVRII) and Hypervariable Region III (HVRIII). Taken together, the differences you have from the CRS in this region are called your mitochondrial profile or haplotype. When you search the GeneTree database, you will be looking for other individuals who share the same or a similar mtDNA profile.

In Summary: What is Genetic Genealogy?

When first introduced to the idea of genetic genealogy, many people wonder how we are able to match their genetic profile with people who lived many generations ago. Some wonder if we have gone around the world sampling DNA from mummies and crypts!

Thankfully, that is not necessary because each of us contains in our DNA a complete record of who we are and where we came from. The field of genetic genealogy taps into that record to pull out useful information to help with reconstructing our genealogy.

Currently there are two main genealogical lines that can be accessed using DNA testing:

Your Direct Maternal Line

You have mitochondrial DNA (mtDNA) you inherited from your mother. It is the same as the mtDNA of your mother's mother. It is the same as the mtDNA as your mother's mother's mother...and so on.

Your Direct Paternal Line

If you are a male, you have a Y chromosome that was inherited from your father. It is the same Y chromosome that was inherited from your father's father. It is the same as the Y chromosome that was inherited from your father's father's father...and so on.

Armed with one or both of these DNA profiles for your own lineage(s), you can then rely on a simple generational mathematical calculation (two parents, four grandparents, etc.) that reveals how every living person today should have more than one billion *possible* ancestors 30 generations (750 years) ago. This is clearly impossible (aside from the simple logistics of just fitting everyone on the planet) since experts estimate that the world population 750 years ago was only about 400 million.

Thus, it is readily apparent that at some point in the recent past our own ancestors converge dramatically. That is, the same ancestors appear on many different branches of the same family tree — often more recently than we might realize. This fact becomes key in identifying common ancestors through DNA sampling.

> *"The genealogies of all living humans overlap in remarkable ways in the recent past."*
>
> *Dr. Douglas L. T. Rohde*
> *Massachusetts Institute of Technology*

What this really means is that if you have the same DNA signature as someone else, there is a high probability that the DNA signature was inherited from a common ancestor. This genetic information, combined with traditional genealogical methods, will then lead you in the right direction toward extending your family tree.

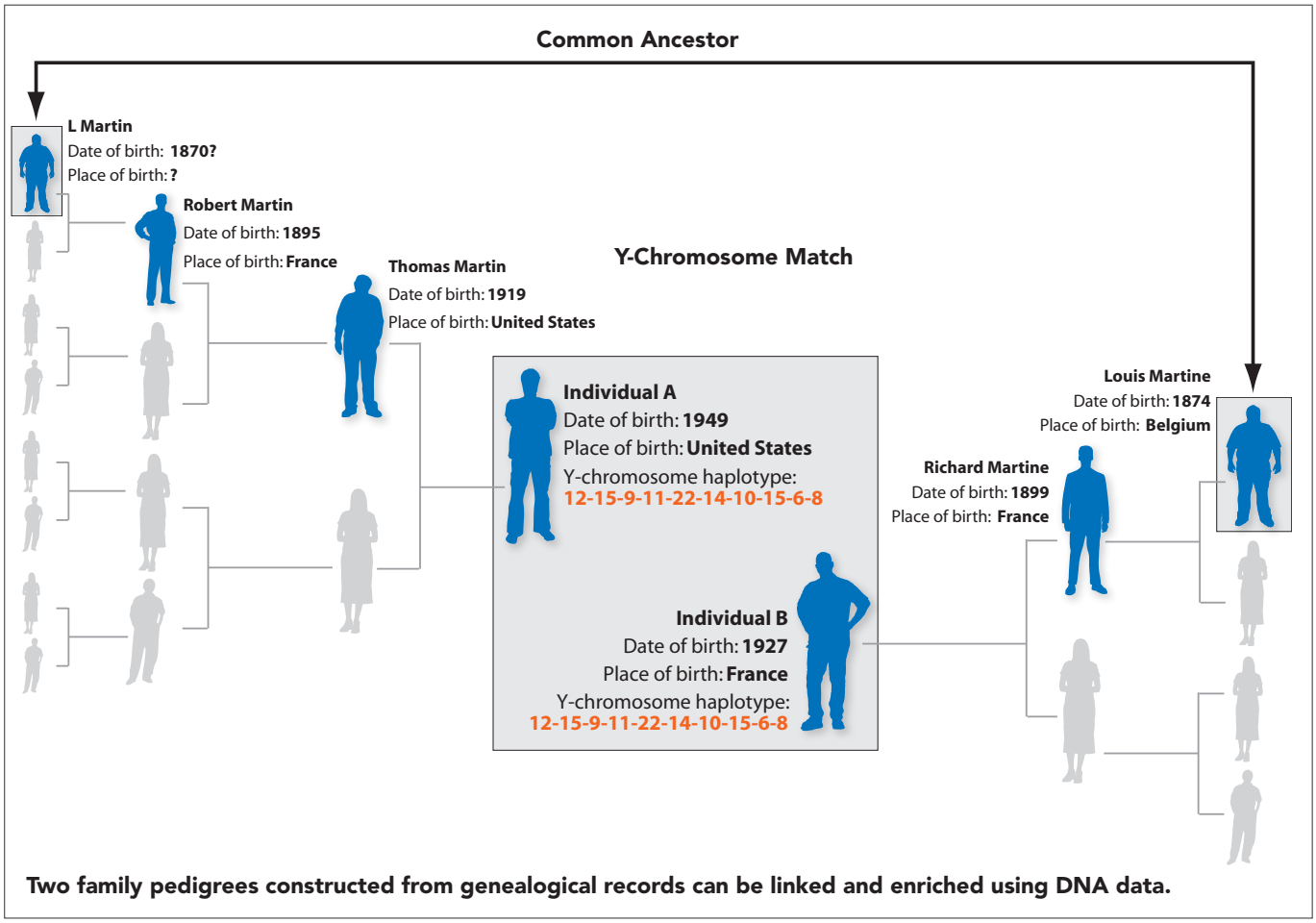
Molecular Genealogy Uses DNA Combined with Written Records

Most people investigating their family roots utilizing conventional genealogical research methods hit a significant barrier fairly quickly. This is due to the inherent limitations of records-based research, even when utilizing computerized information.

Molecular genealogy (also known as genetic genealogy) is a new method of research designed to enhance traditional genealogical efforts. This method utilizes DNA in association with written records.

The diagram below illustrates this method for two hypothetical families. While this example uses the paternal line and the Y chromosome DNA, this same principle applies to the maternal line and mitochondrial DNA.

In this example, Individual A used traditional genealogical methods to trace his ancestors back three generations before encountering incomplete information. After obtaining his DNA data, he searched the GeneTree Y-Chromosome Database to look for matches with his haplotype. He found an exact match with Individual B, who had used traditional genealogical methods to trace his lineage back two generations. From this information, Individual A was able to deduce that his "L Martin" ancestor might be the same as Individual B's "Louis Martine" ancestor, allowing him to fill in the missing information. In addition, when Individual B searches the database he will now find a new connection on his family tree.



Two family pedigrees constructed from genealogical records can be linked and enriched using DNA data.

Rapid advances in DNA research over the past two decades have provided scientists with the ability to positively identify biological relationships across generations. Combining this new science with conventional genealogical records overcomes the limitations of records-based methods and provides powerful new ways to help people identify their ancestors and verify genetic family relationships.

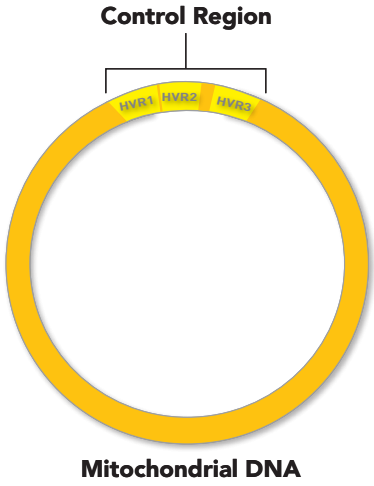
Mitochondrial DNA Genetic Profile

Your DNA was analyzed and differences from the CRS (Cambridge Reference Sequence) were found at the locations listed in the charts below. Each location number is followed by the base (A, T, G or C) present in your mtDNA.

location	16126	16294	16519	73	152	263	315.1
your sequence	C	T	C	G	C	G	C
reference	T	C	T	A	T	A	--

Mitochondrial DNA is a continuous circle of 16,569 bases, each appearing at a distinct location. Locations are designated by number from 00001 to 16569.

Locations on the mtDNA circle from 16001 to 00579 are most useful for recent genealogy research. This range of locations is called the "control region" or the "D-loop". The control region can be divided into three regions called HVR1 (locations 16024 to 16365), HVR2 (locations 00073 to 00340), and HVR3 (locations 00438 to 00574).



Each location consists of a single DNA base value of A, T, G or C. For example, location 16073 might have a base value of "C" and would be listed in your mtDNA results as 16073C.

Most mtDNA tests cover a minimum of 400 locations; GeneTree tests over a thousand locations. Displaying base values at this many locations is impractical, so SMGF and other companies only report base values that differ from the "CRS".

The Cambridge Reference Sequence (CRS) is a set of mtDNA locations and base values that are used universally as a reference. Your test report will show how your mtDNA differs from the CRS; locations that do not differ from the CRS are not listed. For example, an mtDNA report might list your values as 16184T and 16399G. This means that at location 16184 you have a T, which is different than the DNA base value for that location in the CRS. Likewise, your location 16399 has a G, which differs from the CRS.

The Cambridge Reference Sequence (CRS) for human mitochondrial DNA was first published in 1981 prior to the human genome project.

During the 1970s a group of researchers at Cambridge University sequenced the mitochondrial genome of one individual of European descent. The sequence consisted of about 16,568 base pairs and contained some 37 genes. A revised version of the sequence was published in 1999 and serves as a practical way to report mutations in mitochondrial DNA, primarily for genealogical purposes.

Haplogroup Predictions

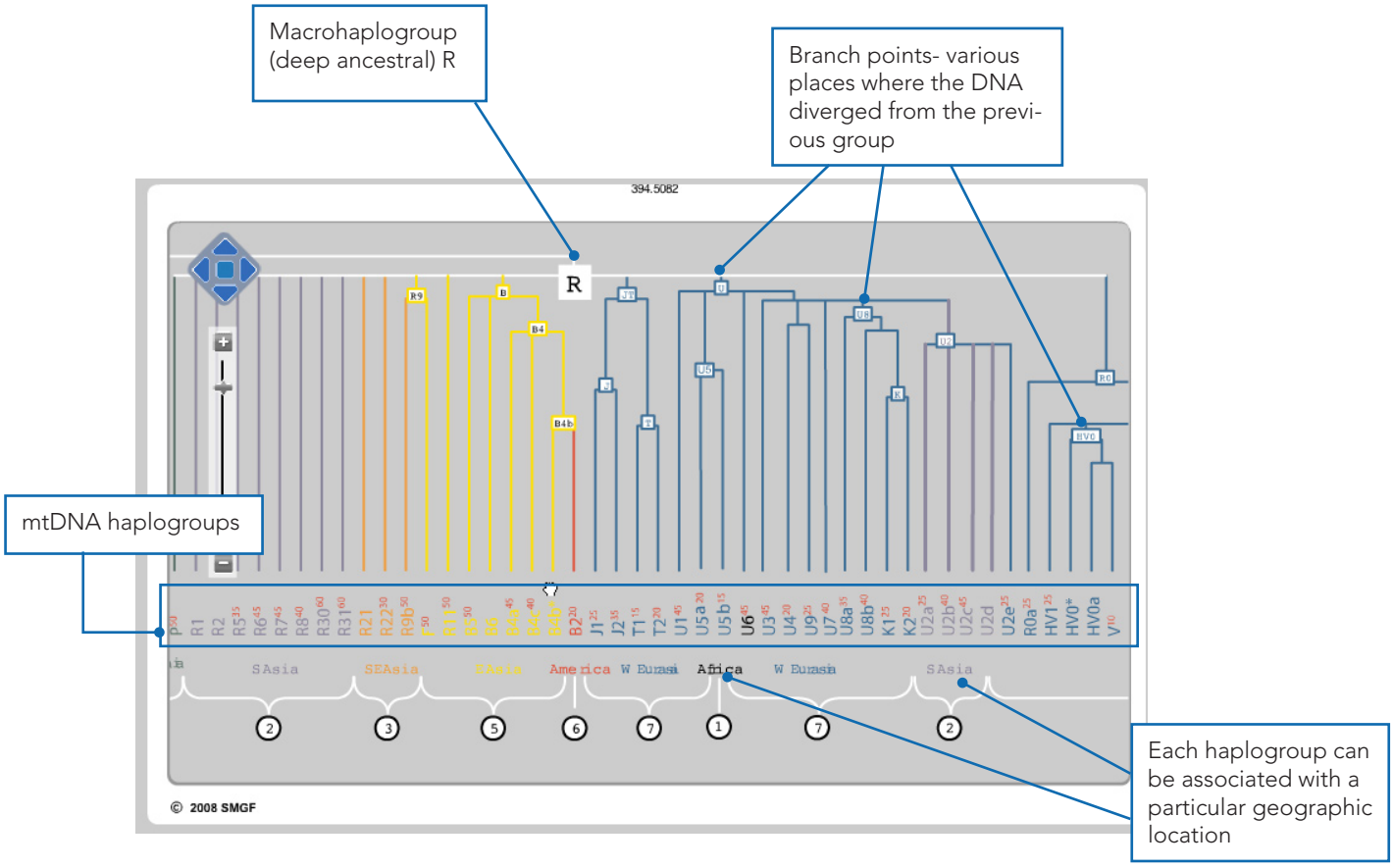
One way to think about haplotypes and haplogroups is that a haplogroup is a branch on the human tree and a haplotype represents a leaf of the tree. All the haplotypes that belong to a haplogroup are leaves on the same branch. When a mutation occurs, the tree splits into branches. Various branches often correlate with geographic world regions.

Therefore shared *haplotypes* indicate recent common ancestry, whereas shared *haplogroups* indicate deeper ancestral groupings.

To definitively determine a person's haplogroup, either on the Y chromosome or the mitochondrial DNA, another kind of DNA test must be performed. Most individuals do not have this direct test performed, but rely on a computer program to provide a prediction of the haplogroup based on the known haplotype data.

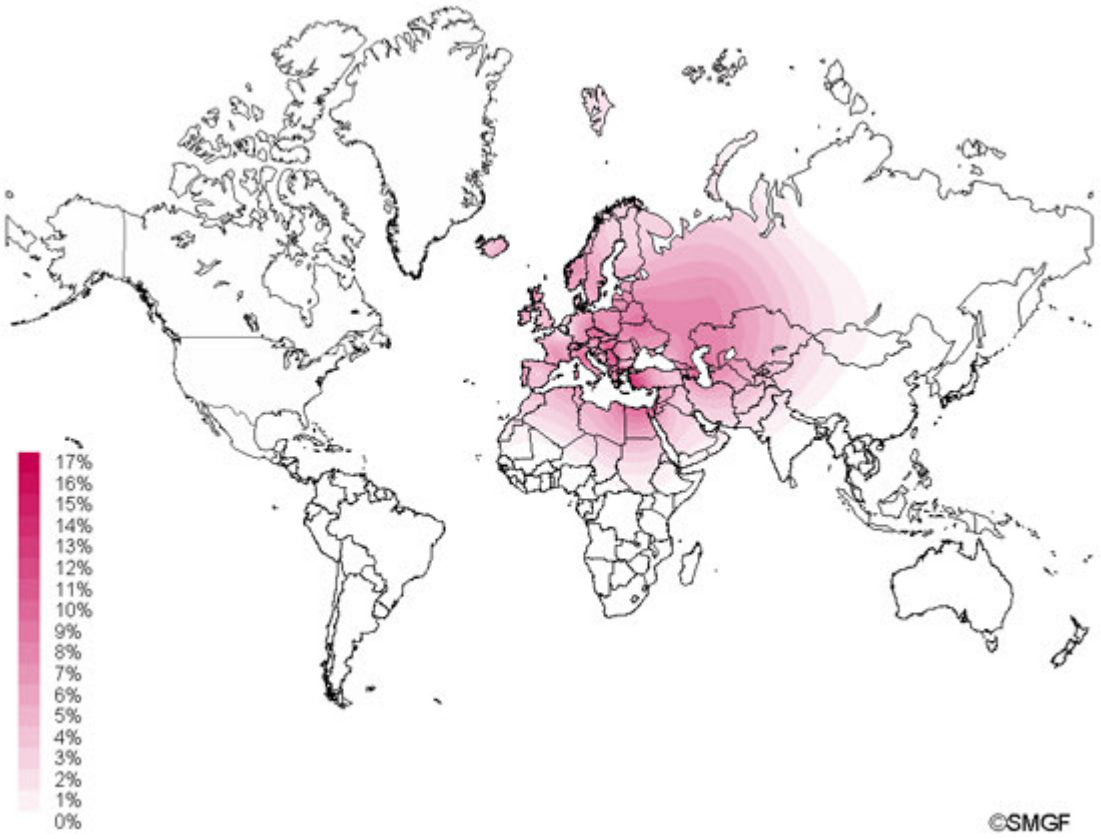
All of the current DNA tests from GeneTree come with a mitochondrial or Y haplogroup prediction. Much scientific effort has been made directly on the part of SMGF to make these predictions as accurate as possible. However, in order to be certain of your haplogroup, and to obtain the most detailed sub-haplogroup information possible, you would need to have a separate DNA test performed.

Below is a portion of the mtDNA phylogenetic tree. A similar tree has been identified for the Y chromosome DNA haplogroups. More information about each of the Y chromosome and mtDNA haplogroups predicted by GeneTree can be found under the DNA Education tab once you login at www.genetree.com.



Haplogroup T

Your predicted mtDNA Haplogroup is: T2



This map shows the geographic distribution of this Haplogroup as measured in various geographically and ethnically defined populations from around the world. The value displayed at any geographical location indicates the percentage of individuals at that location who belong to this Haplogroup.

Mitochondrial haplogroup T is common in European populations, and is found throughout Western Eurasia, including North Africa, the Arabian Peninsula, and occasionally in the Near East. About 10% of Europeans carry haplogroup T. This haplogroup is a sister to haplogroup J as both carry a similar variation pattern and originated about 30,000 years ago probably in the Near East.

Phylogeography of T

Haplogroup T is divided into two main groups: T1 and T2. The T1 subclade shows a lower occurrence in overall European populations, but is more common in the southern Caucasus regions, peaking around 5% among Armenians. This subclade appears to be a more recent development (about 10,000 years old), and is considered a good marker for the expansion of farmers from the Near East into Europe. As populations carrying the T1 subclade came into

Europe, they brought both their farming and domestication techniques as well as their gene pool.

The T2 subclade is most frequent in Europe with the highest frequencies in the Mediterranean and Near East areas. This subclade appears to be older than the T1 subclade, and has higher levels of occurrence among Palestinians.

Famous Members of T

Famous historical figures within haplogroup T include Russian Tsar Nicholas II and his brother, the Grand Duke George Alexandrovitch Romanov. When genetic testing was completed on the bones of Nicholas Romanov to authenticate his remains, it was determined he was part of the T2 subclade.

Additionally, the notorious American outlaw, Jesse James, has been classified into the T2 subclade. James' presumed remains were exhumed from Mt. Olivet Cemetary, and the haplotype exactly matched those of two known, living relatives of James.

References

1. Richards M, Macaulay V, et al. Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet. (Nov. 2000)
2. Palanichamy MG, Sun C, et al. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. Am J Hum Genet. (Dec. 2004)

What now?

In order to maximize the value of your DNA results, you should log on to www.genetree.com and explore the many tools, both genetic and genealogical, that you can use to increase your knowledge of genetic genealogy and extend your family tree.

A few of these tools are explored on the following pages, including:

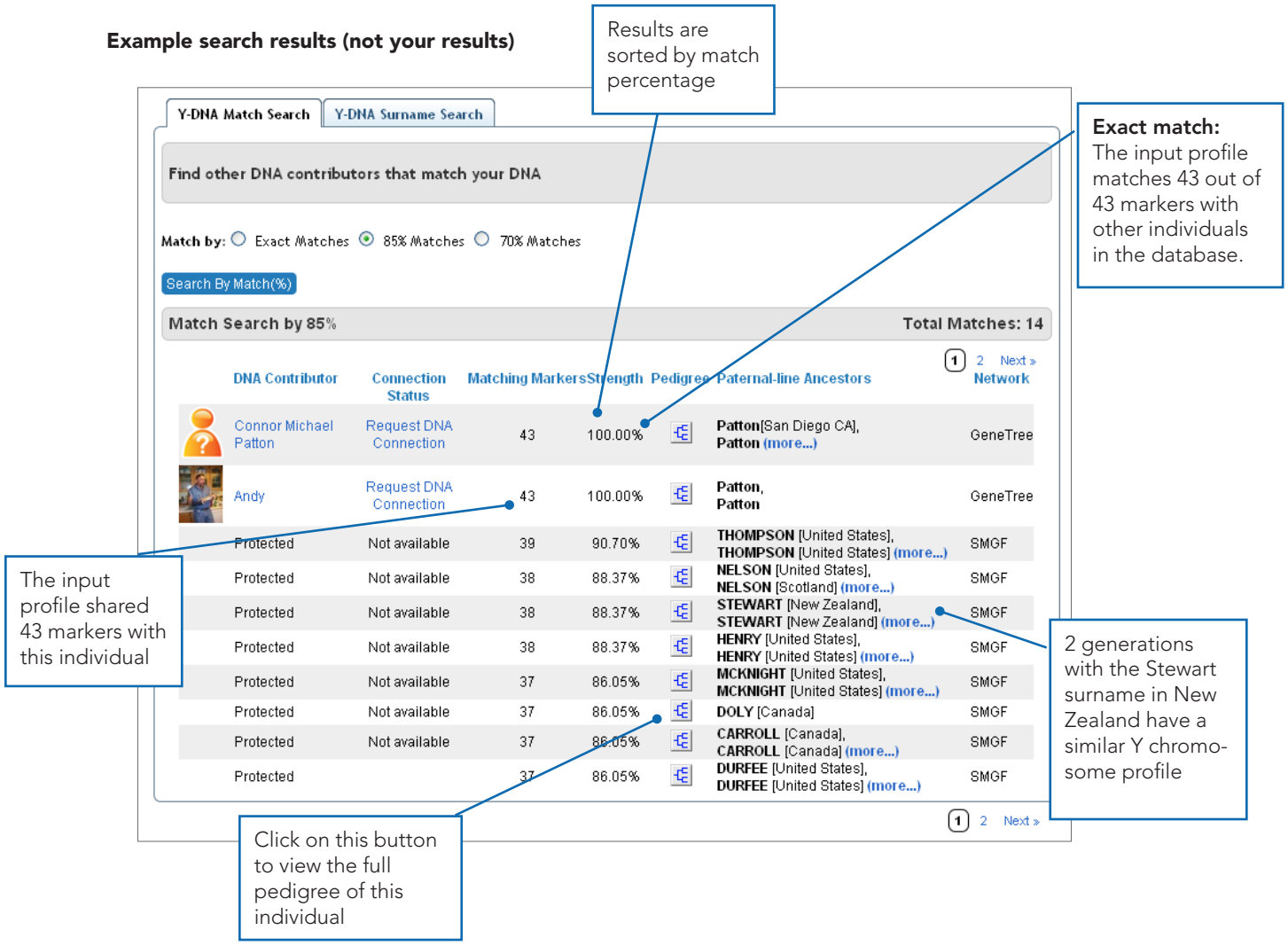
- *Database matching*
- *Pedigree Tool*
- *DNA Navigator*
- *Build Your Family Tree*

Since the GeneTree Database, and the Sorenson Molecular Genealogy Foundation Database, are growing rapidly, you may wish to check them regularly to see if more matches appear.

GeneTree Matching Service

Below is an example of the Y chromosome search results screen. A similar screen displays the results for the mitochondrial search results.

Example search results (not your results)



Matches are ranked on two criteria: the number of common markers compared between the haplotypes, and the number of matches (or match percentage) between these pairs of common markers. In general, the higher these two values are, the closer the two individuals are likely to be related

Common genealogical information, in addition to high values for both the number of common markers compared and the match percentage, will generally indicate a common ancestor in the time that genealogical records have been kept. High values for the matching criteria but with different genealogical information may be coincidence, but likely still inidcates a shared genealogical connection and should be thouroghly investigated.

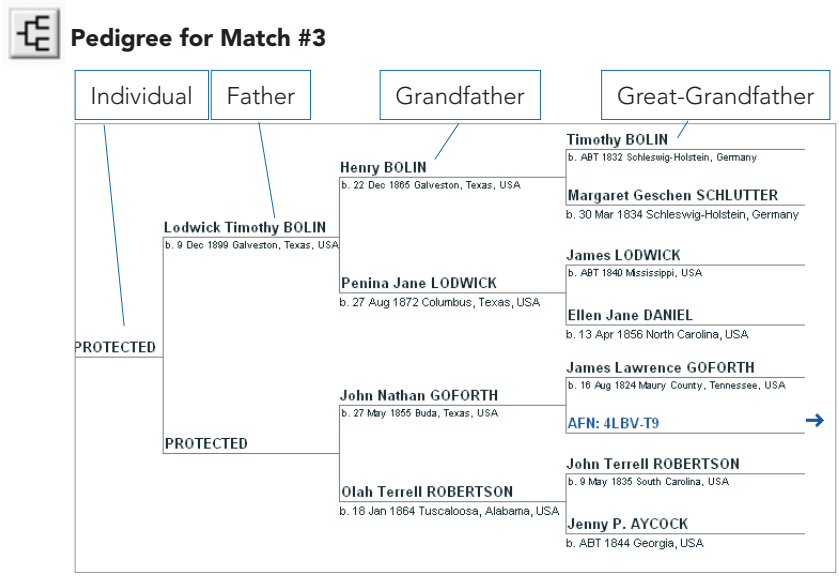
Lower values for the match criteria generally indicate a connection before the time of genealogical records. All matches should be interpreted in the context of individual family history. Molecular genealogy is based on probabilities, and like forecasting the weather, is not an exact science.

The Pedigree Tool

Clicking on the pedigree button next to a match on the search result screen reveals the chart below. To protect the privacy of our participants, we display the word LIVING or PROTECTED for individuals who are still alive. The arrows on the right side of the family tree are links to pedigrees for earlier generations.

Because the Y chromosome is inherited unchanged (with few exceptions) from father to son, and the mitochondrial DNA is inherited unchanged (with a few exceptions) from mother to child, then the individuals in the direct maternal or direct paternal line of this pedigree (depending on which DNA test you had performed), will have the same genetic profile values as the individual who was tested.

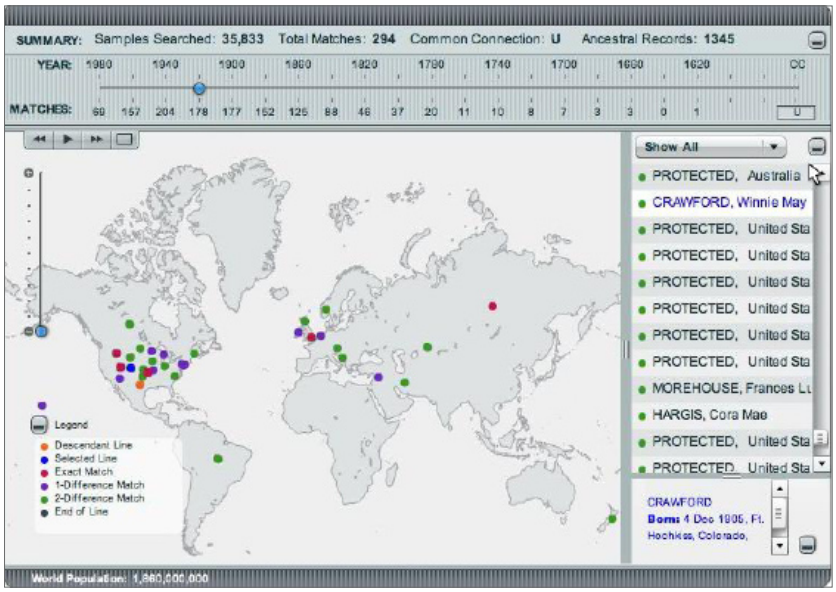
If this were your profile, you might use the information shown on this pedigree to enrich your own pedigree.



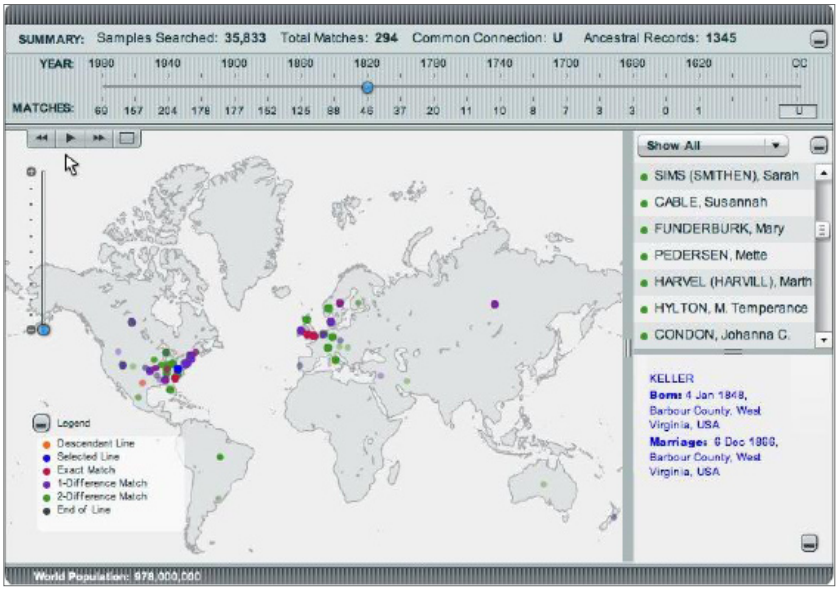
DNAvigator™: A Service of GeneTree

Another unique tool found in the DNA section of GeneTree is DNAvigator™. DNAvigator™ is a unique tool that allows you to evaluate your DNA matches based not only on the genetics, but the genealogy as well. The DNAvigator™ tool places the maternal ancestors of the individuals you match on the map, and allows you to view, in twenty year increments, where each one lived. See the example below.

Start by clicking on the year 1920 above the map and you will see the distribution at right:



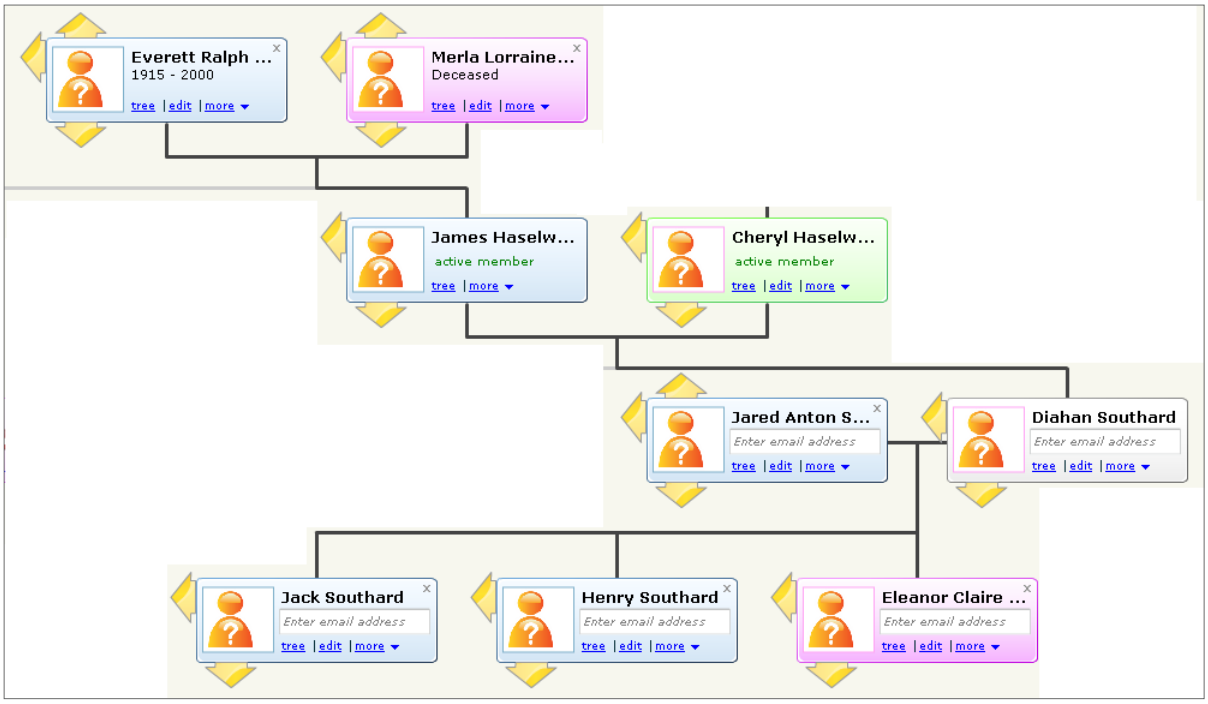
Now click on the year 1820 and watch as the dots representing the ancestors of your DNA matches fly across continents and oceans to reach their residences of that year.



You can also click on the play button in the upper left corner of the map to see the ancestral progression. This tool enables you to view your DNA matches and see if they have ancestors who lived in similar regions to your own. A shared locale and a shared DNA signature may mean that you share a common ancestor around that time period, and should be investigated further.

The Family Tree Tool

Using the tools on GeneTree you can also keep a collaborative record of your own family tree. You can add individuals to your family tree, and invite them to be GeneTree members. Then they can log in and link their pedigree information to yours.



Credits

Results interpretation: SORENSON MOLECULAR GENEALOGY FOUNDATION

Introduction: Genetic Science Learning Center at the University of Utah
Learn.Genetics.utah.edu