

### Y-DNA MOLECULAR GENEALOGY REPORT

# Antonio Cantu

2010-06-11

MAPPING THE WORLD'S DNA AND EXTENDING THE CONCEPT OF FAMILY

# Your Personal Genetic Report



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

- → Y chromosome DNA testing
- Mitochondrial DNA testing
- ---> Education Summary: What is genetic genealogy?

#### ■ Your Genetic Profile

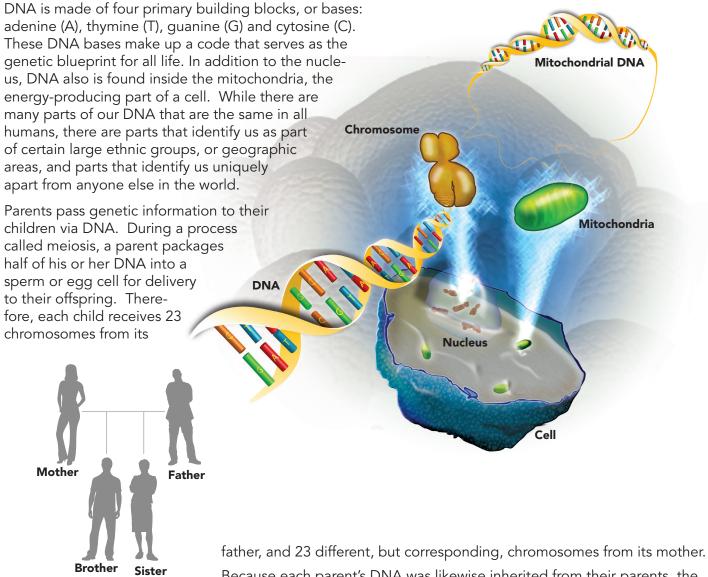
- Test Results
- Haplogroup prediction

#### 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



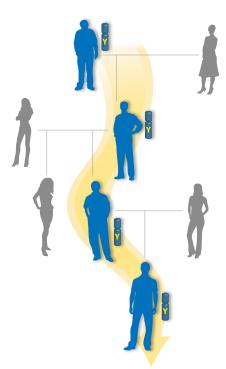
Because each parent's DNA was likewise inherited from their parents, the offspring of two individuals actually carries the DNA of not only their par-

ents, 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.

#### Many locations, called markers, on the Y chromosome contain repetitive segments of DNA called **Short Tandem**

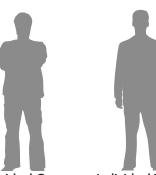
STRs at a marker in the Y chromosome.

**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 Marker A: Allele 15 Marker B: Allele 11 Marker C: Allele 9

Individual Two Marker A: Allele 10 Marker B: Allele 14 Marker C: Allele 9

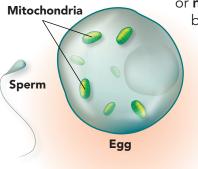
## Detecting Maternal Relationships Using Mitochondrial DNA

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.

Mitochondrial DNA test results are reported as the differences,



or mutations, in the sequence of bases when compared to a standard reference sequence called

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

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.

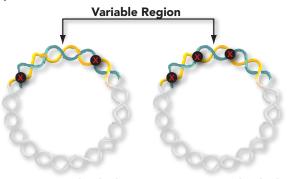
Mutations occur less frequently in mitochondrial DNA than in Y chromosome DNA. For example, in 12,000

**Mitochondrial DNA originates** solely from the mother's egg.

pairs of individuals we would expect to see one mutation in about every 100 generations. Usually between 1 and 15 muta-

tion differences in mtDNA will be found between any two randomly chosen individuals.

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.



Mutations in Individual A's Mutations in Individual B's mitochondrial DNA

mitochondrial DNA

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.

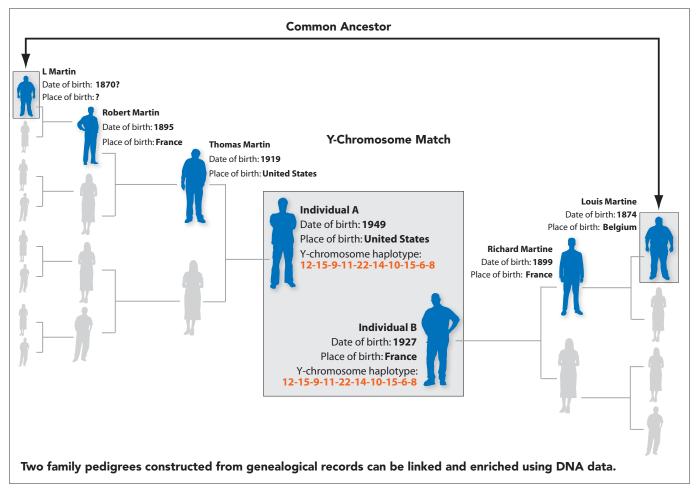
# Molecular Genealogy Uses DNA Combined with Written Records

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.



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.

### Y Chromosome Genetic Profile

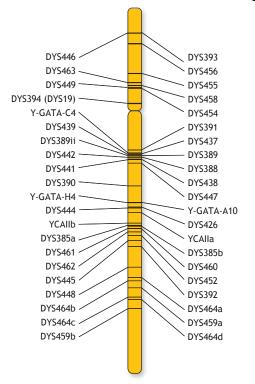
#### **Test Results:**

Genetic markers have been analyzed on your Y chromosome to establish a Y chromosome genetic profile (haplotype). This haplotype is the set of alleles (repeat values) observed at each of the markers, and is shown in the result table below.

| Marker     | Allele | Marker     | Allele | Marker    | Allele |
|------------|--------|------------|--------|-----------|--------|
| DYS385a    | 11     | DYS385b    | 14     | DYS388    | 12     |
| DYS389I    | 13     | DYS389II   | 31     | DYS390    | 24     |
| DYS391     | 10     | DYS392     | 13     | DYS393    | 13     |
| DYS394/19a | 14     | DYS394/19b | -      | DYS426    | 12     |
| DYS437     | 15     | DYS438     | 12     | DYS439    | 12     |
| DYS441     | 14     | DYS442     | 17     | DYS444    | 12     |
| DYS445     | 12     | DYS446     | 13     | DYS447    | 25     |
| DYS448     | 19     | DYS449     | 29     | DYS452    | 30     |
| DYS454     | 11     | DYS455     | 11     | DYS456    | 16     |
| DYS458     | 16     | DYS459a    | 9      | DYS459b   | 9      |
| DYS460     | 11     | DYS461     | 12     | DYS462    | 11     |
| DYS463     | 24     | DYS464a    | 15     | DYS464b   | 15     |
| DYS464c    | 17     | DYS464d    | 17     | DYS464e   | -      |
| DYS464f    | -      | GATA H4.1  | 12     | GGAAT1B07 | 10     |
| YCAlla     | 17     | YCAIIb     | 23     | YGATAA10  | 15     |
| YGATAC4    | 23     |            |        |           |        |

Notes on specific markers:

- DYS19 is the same as DYS394.
- DYS385, YCAII, DYS459, and DYS464 are multi-allelic and are reported with a comma separating the allele values.
- DYS389I and DYS 389II are separate, but nested, markers.
- Y-GATA-C4 is the same as DYS635, which may be used in test reports from other sources.
- Y-GATA-H4.1 is reported by counting the AGAT repeat motif, according to the ISFG 2006 recommendations.
- More information on the markers used by SMGF can be found at www.smgf.org.



### Haplogroup Predictions

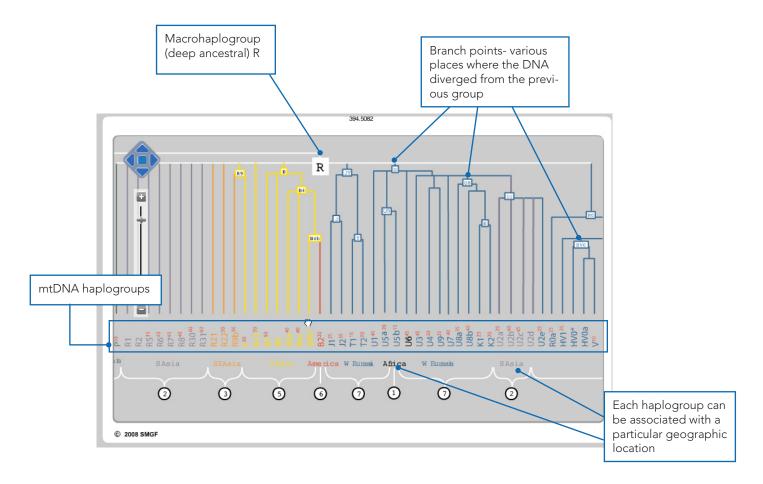
One way to think about haplotypes and haplogroups is that a haplogroup is a branch on the human tree and a haplotype represents a leaf of the tree. All the haplotypes that belong to a haplogroup are leaves on the same branch. When a mutation occurs, the tree splits into branches. Various branches often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



### Y-DNA Haplogroup R



#### Your predicted Y-DNA Haplogroup is: R-M207 Subgroup: R1b1b2a\*-S128

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.

Haplogroup R is defined by a DNA marker known as M207. Everyone who carries this marker today descends from a common paternal ancestor who lived about 30,000 years ago in west Asia. To date, over thirty subclades of haplogroup R have been identified, of which, R1a1-M17 and R1b1b2-M269 (historically called R1b3) are the most well described. Both of these subgroups are indicators of European ancestry with haplogroup R1a1-M17 most representative of Eastern Europeans and R1b1b2-M269 most characteristic of Western Europeans.

#### HISTORY AND GEOGRAPHY OF HAPLOGROUP R

As members of the human family, all people living today can trace their earliest paternal ancestors to populations that lived approximately 100,000 years ago in eastern Africa. These early humans became spread throughout the African continent, and beginning ~50,000 years ago, a series of complex migrations moved them out of Africa into regions of Asia and beyond to eventually populate every major area of the world.

Approximately 26-30,000 years ago haplogroup R emerged in west Asia descending from a widespread Asian haplogroup known as K-M9. Representatives of haplogroup R expanded to the west and south migrating throughout central and southwest Asia into India, the Middle East, and Europe. During this time, subgroups R1-M173 became frequent throughout central and western Eurasia while R2-M124 distributed mainly within regions of Central and South Asia. Today, members of haplogroup R\*-M207 who carry the R-M207 marker, but neither R1-M173 nor R2-M124 markers, can still be found in Central and South Asia.

Haplogroup R2-M124 expanded in the Asian subcontinent following the last Ice Age which manifested its maximum severity ~18,000 years ago and today is found in India, Pakistan, Central Asia, and Turkey with decreasing frequencies respectively. Little is currently known about the subgroups of R2-M124.

Haplogroup R1-M173, represented primarily by subgroups R1a1-M17 and R1b1b2-M269, became well established in Europe during the Upper Paleolithic Period (Late Stone Age) contributing to the spread of Aurignacian or Gravettian culture, characterized by advanced artistic and technological achievements. During this period, populations were ranging across Europe and subsisted mostly by hunting and gathering. An impending Ice Age, reaching its maximum (the LGM=Last Glacial Maximum) beginning ~18-20,000 years ago and lasting several thousands of years, subsequently covered most of Europe with massive glaciers forcing populations to retreat to geographically isolated refuge areas where they were able to persist. Individuals of R1a1-M17 withdrew primarily to refuges located in present day Ukraine and Asia Minor while representatives of R1b1b2-M269 became most concentrated in the Iberian refuge.

As the Ice Age began to abate ~13,000 years ago and living conditions slowly improved across Europe, localized populations migrated from refuge areas to repopulate Europe and areas of Asia. Specifically, representatives of haplogroup R1a1-M17 expanded into eastern Europe and into parts of Asia whereas those of R1b1b2-M269 moved primarily throughout western Europe, creating opposite geographic distributions which are still evident today. Currently, R1a1-M17 is rare across Western Europe but is found throughout eastern Europe in countries such as Poland, Hungary, and Ukraine; in central Asia with highest prevalence in Altaic-speaking populations of eastern Kyrgyzstan and Indo-European-speaking populations of Tajikistan; and in areas of northwest India. In contrast, R1b1b2-M269 is found at very high frequencies (50-80%) throughout western Europe particularly in the Ireland, Wales, Scotland, England, Portugal, France, Germany, and northern Italy, with diminishing frequencies towards the east although representation stretches as far as central Asia and India.

Once the Ice Age ended haplogroup R would continue to play significant roles in subsequent phases of European history, including the adoption and spread of farming and agriculture, Greek and Roman influences, and the Middle Ages. These and other historic events are currently being studied within the context of Y-DNA haplogroups. This ongoing research will surely enrich our current understanding of world history as well as our own personal and family histories.

#### **IBERIAN R1b HAPLOGROUPS**

Haplogroups R1b1b2a2c-M153 and R1b1b2a2d-M167 are considered indigenous to the Iberian Peninsula which encompasses Portugal, Spain, Andorra, Gibraltar, and a portion of France. Evidence suggests that R1b1b2a2c-M153 originated ~18,000 thousand years ago among the Iberian Basques, an ethnic group whose non-IndoEuropean language, Euskara is

the oldest surviving language in Europe. Haplogroup R1b1b2a2d-M167 also originated in Iberia but evolved among its non-Basque population.

#### HAPLOTYPES OF HAPLOGROUP R

A Y-DNA haplotype consists of a series of STR (Short Tandem Repeat) markers located along the Y-chromosome. Each STR marker has a very high mutation rate and therefore changes rather quickly through time. Because of their high variability STR haplotypes can identify recent relationships within a haplogroup. One well known example is the 6-marker Atlantic Modal Haplotype§ (AMH), which is shared at high frequencies by members of haplogroup R1b1b2-M269 living in the European Atlantic facade, specifically Celtic-speaking populations of Ireland and Wales, Scandinavian countries of Netherlands and Norway, and the Basque population. Additionally, a 17-marker haplotype known as the Irish Modal Haplotype§§ (IMH) accounts for 17% of haplogroup R1b1b2-M269 members in northwest Ireland.

The Y-STR marker, DYS458 is commonly tested by genetic-genealogy companies and is included in many public databases. An unusual "0.2" value (ex. DYS458 = 17.2) is relatively common and typically associated with the Middle Eastern haplogroup J1-M267. However, DYS458.2 values have recently been found to also exist within haplogroups R1b1b2-M269 and R1b1b2a1-M405. Haplogroups J and R do not share this unusual marker because of common ancestry but due to chance independent mutation events. This result underscores the weakness of using a single STR marker or too few STR markers as indicators of common ancestry, even when those values are relatively rare.

#### FAMOUS MEMBERS OF HAPLOGROUP R

One of Scotland's greatest warriors, Somerled of Argyll, who is credited with driving the Vikings from Scotland, belonged to haplogroup R1a1-M17. Members of Clan Donald, which Somerled founded, were tested to determine Somerled's Y-chromosome membership. In 2005, Oxford University further found that Somerled possibly had 500,000 living descendants. If that is correct, he would be the second most common ancestor after Genghis Khan.

Possible descendants of Niall of the Nine Hostages, an Irish king in the Dark Ages, are members of R1b1b2a2e-M222. These descendants are associated with the U' NŽill clan.

American presidents John Adams and Franklin Pierce belonged to Y-DNA haplogroup R1b1b2-M269.

Joseph Smith Jr., founder of the Mormon Church with current worldwide membership exceeding 13 million, belonged to haplogroup R1b1b2a2e-M222.

Anderson Cooper, an Emmy Award winning American journalist, author, and anchor of the CNN news show Anderson Cooper 360° belongs to haplogroup R1a1-M17.

#### HAPLOGROUP R AND THE GENETREE DATABASE

The GeneTree haplogroup predictor reports the following R-haplogroups: R1a1-M17, R1b-M343, R1b-M343, R1b1b2\*-M269, R1b1b2-M269, R1b1b2a\*-S127, R1b1b2a1\*-M405, R1b1b2a1-M405, R1b1b2a1a-M467, R1b1b2a2e-M222, and R1b1b2a2g-U152, with more on the way!

Approximately 45% of the GeneTree and SMGF databases are members of haplogroup R.

Notes:

\* A haplogroup designation followed by a "\*" signifies the presence of a haplogroup defining marker but the absence of any other known subclade markers. For example, R1\*-M173 is a member of haplogroup R1 but not a member of its subgroup R1a-SRY1532.2. § AMH: DYS388,DYS393,DYD392,DYS19,DYS390,DYS391: 14,24,11,13,13,14 §§ IMH: DYS19, DYS388, DYS390, DYS391, DYS392, DYS393, DYS434, DYS435, DYS436, DYS437, DYS438, DYS439, DYS389I,DYS389B, DYS460, DYS461, DYS462: 14,12,25,11,14,13,9,11,12,15,12,12,13,16,11,10,11

### 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.

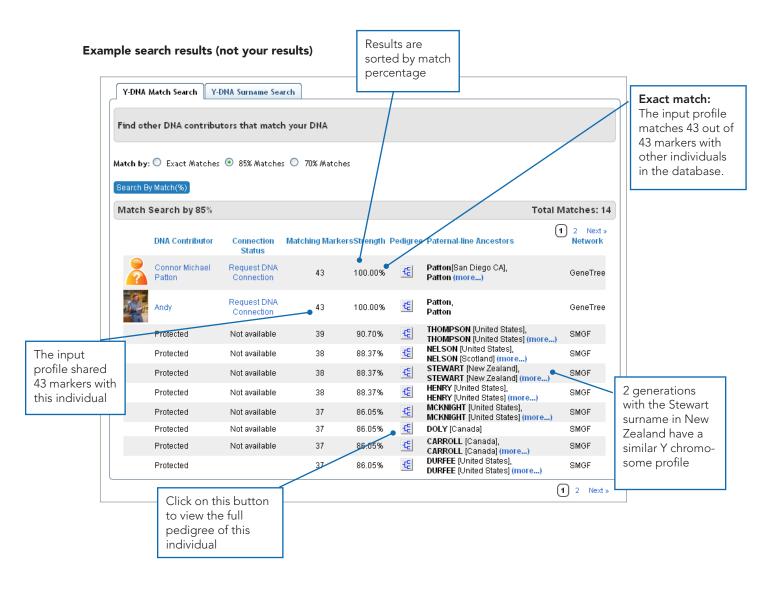
A few of these tools are explored on the following pages, including:

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.



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.

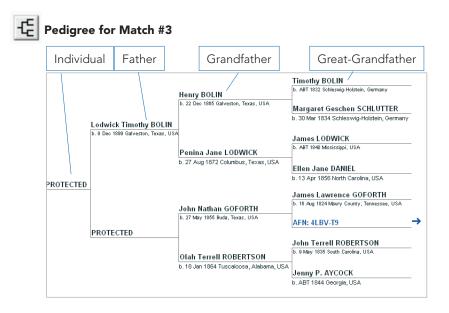
Lower values for the match criteria generally indicate a connection before the time of genealogical records. All matches should be interpreted in the context of individual family history. Molecular genealogy is based on probabilities, and like forecasting the weather, is not an exact science.

### 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.

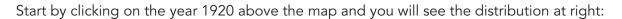
Because the Y chromosome is inherited unchanged (with few exceptions) from father to son, and the mitochondrial DNA is inherited unchanged (with a few exceptions) from mother to child, then the individuals in the direct maternal or direct paternal line of this pedigree (depending on which DNA test you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

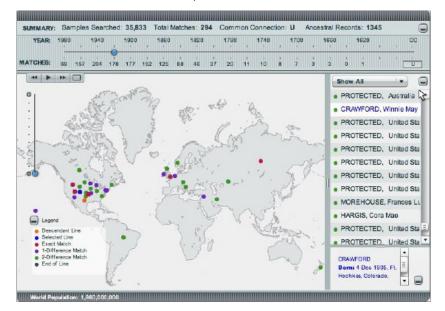
If this were your profile, you might use the information shown on this pedigree to enrich your own pedigree.



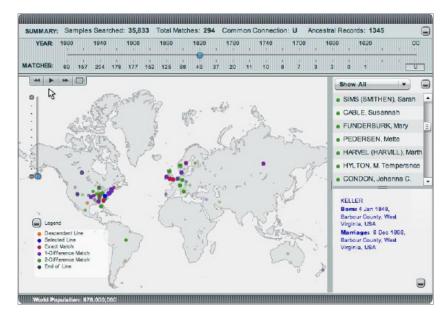
### DNAvigator<sup>™</sup>: A Service of GeneTree

Another unique tool found in the DNA section of GeneTree is DNAvigator<sup>TM</sup>. DNAvigator<sup>TM</sup> 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<sup>TM</sup> 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.





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.

