Expanding the Concept of Family History Through DNA

Ugo A. Perego, Ph.D, looks at the contribution of DNA testing in the field of family history

In 1999, Utah philanthropist, James LeVoy Sorenson, and molecular biologist, Dr. Scott R. Woodward, joined forces and resources to build a comprehensive database of correlated genealogical and genetic data that could be used to assist people in extending and verifying their family history research, as well as promote a greater sense of brotherhood and sisterhood by demonstrating how closely related all humans are.

Sorenson and Dr. Woodward were truly visionaries and pioneers in the field of genetic genealogy, as the work they started more than a decade ago is today corroborated by the fact that millions of people have since included genetic testing as one of the essential components of their genealogical toolbox.

The unique and powerful contribution of DNA testing to the field of family history is unquestionable. Hundreds of genetic genealogy success stories are shared each year through public online forums (Rootsweb, Yahoo, etc.), personal websites, lectures at national genealogical meetings and publications (see, for example, the May/June and November/December 2009 issues of *Family Chronicle*). However, it appears that a large number of individuals who purchased a DNA test are still somewhat confused about what DNA can really tell them about their ancestry. This confusion is evident from the numerous genetic genealogy consultations provided each week by companies like GeneTree.com, where family historians, with their genetic profiles already in hand, are desperately looking for someone who can supply personal and meaningful interpretations to their genetic data.

While some people might have specific family questions they have been trying to resolve through the years, others are simply curious to know how DNA can enlarge what they already know about their familial connections and past.

Wanting to know more about our family history and how we are connected to each other is part of human nature. A simple conversation with a stranger we meet on a business trip will unavoidably lead to a search for common ground. Questions like "Where are you from?", "What is the origin of your surname?", "Do you know so and so?" and "Are you related to…?" are not limited to genealogical circles. We simply want to

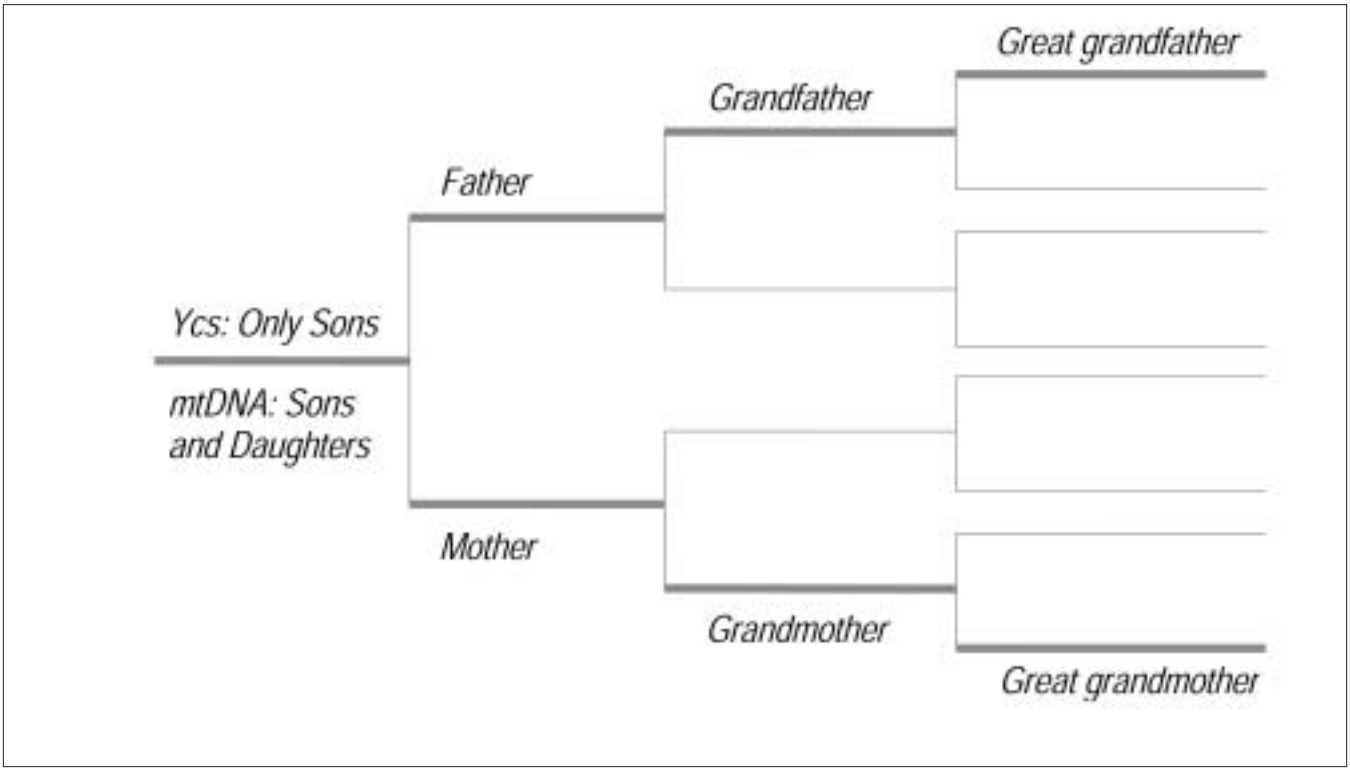

Figure 1: Pedigree representing Y chromosone (YCS) and mitochondrial DNA (mtDNA) paternal and maternal inheritance patterns respectively.

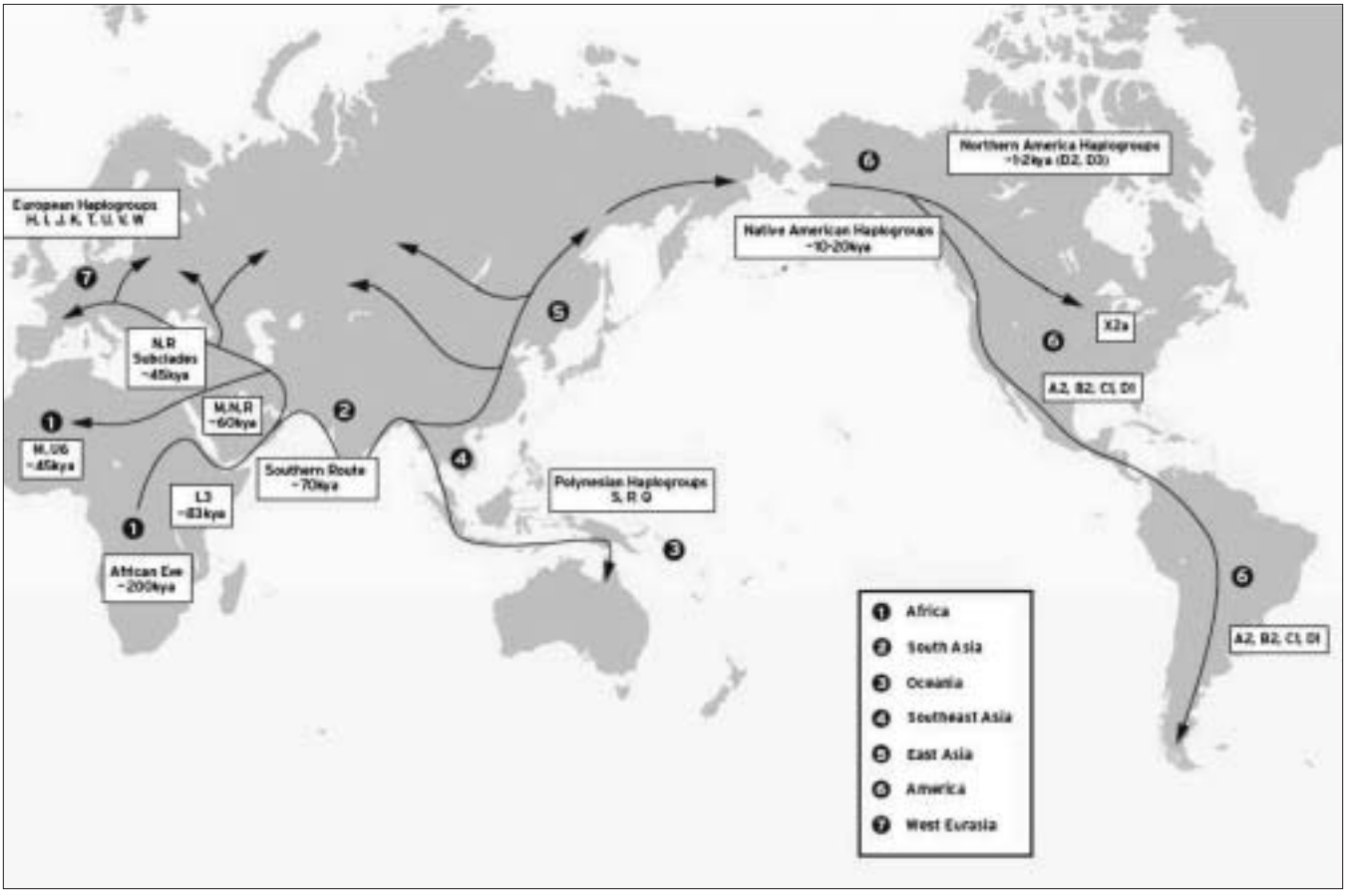

Figure 2: MtDNA World Migrations Map.

know how we fit in the world family and we often feel limited by the amount of information we have available to make these links. It is no surprise then that genealogy is the most popular hobby in North America, that family history websites are among the most visited on the Internet and that online social networking services, such as Facebook or Linkedin, are flourishing. However, it is possible that the concept and potential of using DNA to increase our web of connections, and expand on the traditional concept of family as we know it, is largely unknown to most people.

DNA testing for genealogical and ancestral purposes can be summarized in three main categories (Figure 1): A) Y chromosome (Ycs) testing to trace unbroken paternal lineages, B) mitochondrial DNA (mtDNA) testing to study strict maternal lineages and C) autosomal DNA testing to learn something about our past through the genetic contribution of all our ancestors.

Each method has the capability to explore different parts of our genetic family tree and provide specific information that could shed more light about our ancestry. There are plenty of free tutorials online to learn how each system works and how it could be used in our genealogical research (for example, *www.genetree.com /tutorials* for Ycs and mtDNA or *www.23andme.com/ancestry* for autosomal DNA).

In this article, I would like to focus on information obtained from mtDNA testing and use my own family to demonstrate the concept of expanding connections, relationships and ancestry. A similar approach can be taken by everyone and with any of the genetic methods available.

Two peculiar characteristics of mtDNA are particularly helpful in its application to study the history of humankind. As already stated, mtDNA is inherited exclusively along the maternal line, from a mother to all of her children. Both males and females carry mtDNA in their cells, but only women will pass their mtDNA profile (called haplotype) to the next generation.

The second feature of mtDNA is that it does not mix with any of the genetic material found in the cell's nucleus. This means that the mtDNA haplotype of all people represents the genetic legacy of hundreds of generations back in time following the unbroken maternal line. Additionally, when our first ancestors migrated from place to place during the early expansion processes that resulted in the colonization of the whole planet, their mtDNAs accumulated some small but significant changes (called mutations) that could be linked to specific geographic locations around the world (Figure 2).

These ancient mutations have survived in our DNA after thousands of years and can be used to assign people living today to large family clusters called mtDNA haplogroups. Therefore, through a combined interpretation of both mtDNA profiles and haplogroup affiliation, it is possible to reconstruct the maternal legacy of all modern people and plot such

information on a comprehensive worldwide mtDNA tree (called a phylogeny) where literally every mtDNA profile has its own exact place.

When the first haplogroups were identified about two decades ago, scientists adopted a simple nomenclature system following the letter of the alphabe. As new sub-lineages are identified, letters are alternated with numbers in a straightforward, but universal, fashion. A complete and up to date mtDNA tree is available at, *www.phylotree.org/tree/main.htm*.

Figure 3 includes a portion of the larger worldwide mtDNA tree that focuses only on two specific lineages called haplogroup U and its sub-lineage K. The data plotted on this tree come from partial and complete mtDNA profiles gathered over the past few months from my own family and from the publicly accessible database GenBank, *www.ncbi.nlm.nih.gov/nuccore*. Below the tree is my three-generation family tree, which includes me, my parents and my four grandparents (listed with our first names followed by a number). DNA samples were collected from all the people on this pedigree chart with the exception of my maternal grandfather, Francesco, who passed away when I was only a couple of months old. Each individual mtDNA lineage is labeled on the tree with numbers corresponding to the individuals on my pedigree chart (starting with myself as #1) or with their GenBank IDs. My mother and my maternal grandmother share the same mtDNA signature as I do, as mtDNA is inherited exclusively along the maternal line, and therefore we are represented by the same branch on the schematic U/K mtDNA tree with numbers 1, 3, and 7.

Based on a number of mtDNA profiles available in the public domain, a deeper and more precise classification for haplogroup K is now possible, indicating the existence of two major sub-branches (K1 and K2) which in turn split in numerous additional sub-lineages. One of them is called K1a4a1 and is the one where I belong, together with my mother and maternal grandmother. My paternal grandfather (Alessandro, #4) sits alone on the K1b lineage, while my paternal grandmother (Alma, #5) is part of the U1b branch together with my father (Alberto, #2).

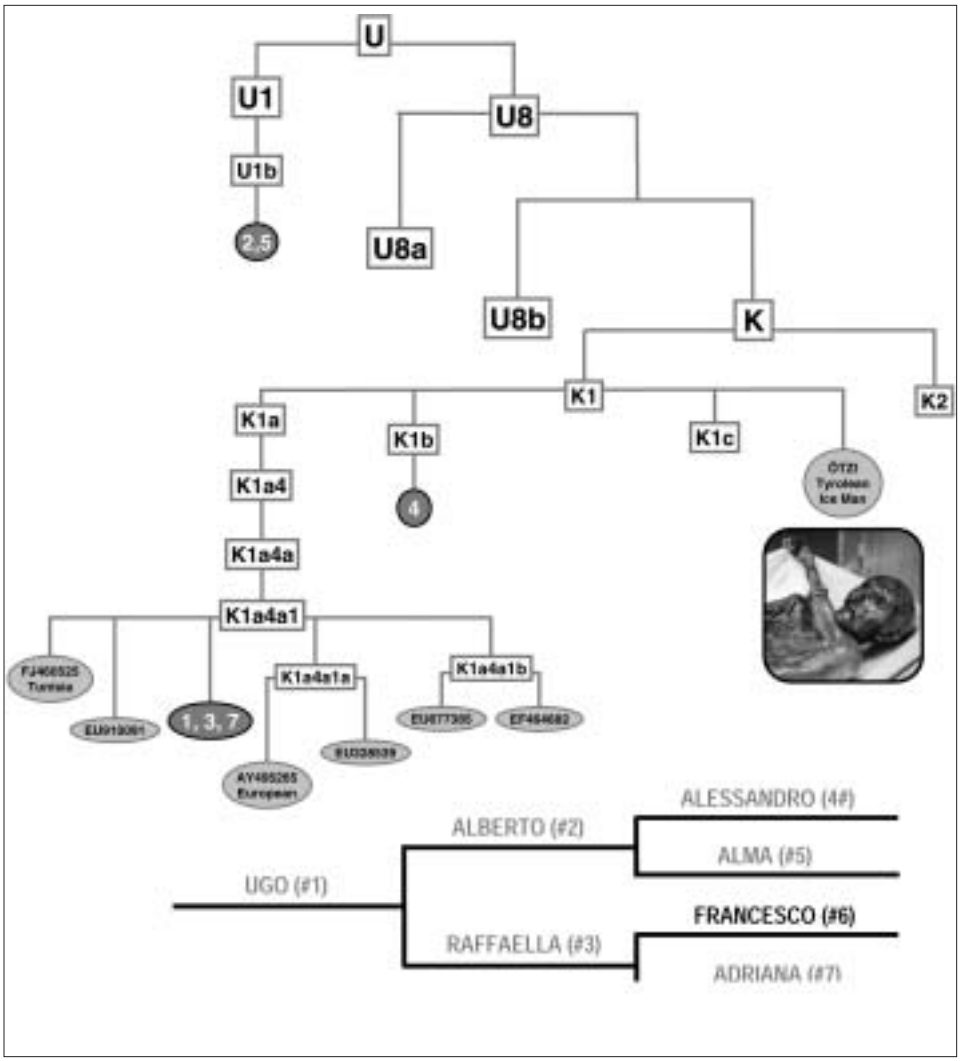

Figure 3: Schematic mtDNA tree for lineages U and K.

A few valuable observations can be drawn from Figure 3 that should expand on the concept of genetic versus genealogical relationships:

1) Although there are only two generations separating me from my two paternal grandparents, from a strict mtDNA point of view, we are not even closely related. In fact, I share a more recent common ancestry with six unknown individuals whose mtDNA appears in the GenBank database. These individuals and I (including my mother and my maternal grandmother) are part of the same "mtDNA family" called K1a4a1, as we share all the mutations characteristic of this specific branch of the mtDNA tree. Four of these genetic contributions come from individuals who purchased an mtDNA profile from a private laboratory and made a direct submission to GenBank (the other two came from research papers where their ethnic/geographic group was also available).

When people purchase their DNA results and then go through the trouble of uploading them

somewhere online, it is a clear demonstration that they want to be found and that they are looking for others to whom they might be related. Just as much as I care about my two paternal grandparents and want to know everything I can about them, do I have the same desire to connect and learn about these six strangers that are part of my mtDNA family group? The answer is a definitely "yes" and I am fairly confident that the others on this tree feel the same way I do.

2) Another interesting point from Figure 3 is the distant connection I share with Ötzi, the Tyrolean Ice Man, who was found on the border between Italy and Austria few years ago. Carbon dating indicated that he lived approximately 5,000 years ago. Because his body was frozen all this time, soft tissue was available to obtain good DNA and produce one of the very few complete mtDNA sequences obtained from ancient remains. It turned out that Ötzi also belongs to haplogroup K. However, there are three known K1 lineages found in the modern population (K1a, K1b, and K1c) and Ötzi does not belong to any of them (he has been temporarily placed in his own group called K1ö).

No other mtDNA sample tested to date shares all the same characteristics as those from Ötzi's mtDNA, most likely because his lineage became extinct over time, although it is also possible that others in this same lineage will be revealed as more DNA sequences are produced. However, I find it exciting to have a "VIP" on my mtDNA tree and I am definitely interested in learning everything I can about this "ancient cousin" and to boast a connection with him. Likewise, many famous people have now had their genetic profiles reconstructed through genetic genealogy by collecting and analyzing DNA samples from their known descendants. A list of famous people DNA and other helpful links are found at, *www.isogg.org/famousdna.htm*.

3) The last observation has to do with my maternal grandfather (Francesco, #6). He is nowhere to be found on the mtDNA tree. As stated earlier, he passed away when I was an infant. I have no recollections of him and no one ever collected a DNA sample from him to know what mtDNA haplogroup he belonged to. Because mtDNA is maternally inherited, none of his children carry his same mtDNA signature and, as far as I know, no other maternal relatives are living today from whom a DNA sample could be obtained to infer his genetic profile.

The nearly seven billion people living today are the product of no more than 400 million people who lived approximately seven centuries ago.

Although he is my biological and genealogical grandparent, I know nothing about the genetic connection we share. In fact, there could be mtDNA profiles of distant and unknown maternal relatives floating around in public databases right under my nose, but I would not be able to recognize them because I don't have the reference genetic signature from my own grandparent.

These simple observations, based on a schematic tree built using exclusively genetic data obtained from mtDNA profiling, are powerful examples of the expanded view on the classic concept of family relationships and genealogical connections that are familiar to us. This new approach is not meant to replace traditional genealogical research and pedigrees based on family history data, but stands as an encouragement to enlarge our understanding of the concept of family by seeking previously unknown relationships revealed through DNA testing and welcoming them in our lives.

After all, we are truly all members of the same human family, sharing just a few hundred thousand common ancestors within the last couple of millennia. Although each person has over a million ancestors within just 20 generations and a billion ancestors in 30 generations, the reality is that these ancestors are not all unique.

In fact, mathematical reconstruction of the number of actual common ancestors versus potential ancestors provides a startling reality: the nearly seven billion people living today are the product of no more than 400 million people who lived approximately seven centuries ago. We are all more closely related than what we ever thought to be and, thanks to genealogical research and the new field of genetic genealogy, we are narrowing this gap by continually revealing such connections.

After a decade of genetic genealogy, we can now enjoy a variety of DNA testing options developed for the family historian and several large freely accessible databases to search for matches and shared ancestry. While many individuals are hoping that DNA will knock down some of the brickwalls in their genealogical research, one added value that should be kept in mind when staring at our genetic results is the possibility to find new connections beyond the traditional concept of family. The genetic legacy of the common ancestors we share is found today in our cells, waiting to be unlocked. That information can then be used to find previously unknown "genetic cousins" as well as revealing important clues about our distant past and the ancient origin of our progenitors.

FC

Ugo A. Perego, Ph.D, is a senior researcher with the Sorenson Molecular Genealogy Foundation and a scientific consultant for GeneTree.com. You can contact him at, *ugo@smgf.org*.