OLAQ

OLA Quarterly

Volume 7
Number 4
Trends in Geneaology for Librarians (Winter 2001)

July 2014

The Molecular Genaology Research Project

Ugo A. Perego

Natalie M. Myres

Scott Woodward

Recommended Citation
Perego, U. A., Myres, N. M., & Woodward, S. (2014). The Molecular Genaology Research Project. *OLA Quarterly, 7*(4), 18-22. http://dx.doi.org/10.7710/1093-7374.1599

OLA Quarterly is an official publication of the Oregon Library Association | ISSN 1093-7374

The Molecular Genealogy Research Project

by Ugo A. Perego, MS,
Natalie M. Myres, MS
and
Scott R. Woodward, PhD

Introduction

A couple of months ago, we received an email with the following request, "In 1986, my husband and I adopted a 5 month old boy who was abandoned at birth, which means we have no background information on the biological parents. He is biracial and they could not tell us his nationality other than he is possibly Asian. Now that our son is 15 years old, he would like to know what nationality he actually is as he gets questions all the time from friends. Is there anything in your research program that could help us to determine his nationality? I think it would be almost impossible to find the parents, but he wants to know his nationality."

This request is not unique. Many people find themselves at a dead end with their genealogical research. Written records are simply not available, or not reliable. No matter how willing they are to search for the needed information, there is almost no hope to find the missing link with their past. Sometimes, these precious ties with the past are destroyed by fire or flood. Other times, migrations or adoptions form a huge barrier between the branches (the present) and the roots (the past) of our family trees. What can be done when these situations arise for those who are desperately searching for their own origins and personal identity?

Genetic research with the purpose of tracing genealogies by using DNA is currently underway at Brigham Young University. This study is known as the Molecular Genealogy Research Project (MGRP) and its main goal is to develop a database containing correlated genealogical and genetic data from all over the world. This instrument will then be used to assist those having problems finding information about their family history.

What is DNA?

DNA (Deoxyribonucleic Acid) is the genetic material containing all of the genetic information necessary for living organisms and is the repository of hereditary traits. The human genome (DNA) is the complete genetic blueprint of a person, and it consists of over three billion chemical pairs. It is found mainly in the nucleus of cells, in structures known as chromosomes: 23 received from the father and 23 received from the mother. In addition to nuclear DNA, there is also genetic material found inside mitochondria (mitochondria are energy-producing organelles found in the cytoplasm). These genetic components contain all of the necessary information for the foundation and the sustaining of human life. The color of our eyes, our height, our predisposition to certain diseases are just a few examples of what is contained in the DNA of our cells.

What is Molecular Genealogy?

DNA is transmitted from one generation to the next. Some parts are passed almost unchanged, while others experience a high rate of recombination. This mode of transmission from parents to children creates an unbreakable link between generations and it can be of great help in reconstructing our family histories. Molecular Genealogy is therefore a new way to do genealogy, where DNA is used in association with traditional written records. Since we have inherited our genetic material from our ancestors, our relatives share with us a portion of this information. The closer the relative, the greater the amount of genetic information that is shared. This means that inside any family unit, the members of that unit share a greater quantity of genetic material than those outside of it. In other words, even though the entire DNA sequence of an individual is unique to that individual, similar genetic information can be found among those that descend from common ancestors. The objective of Molecular Genealogy is to establish family links among individuals, families, tribes and populations by using the information encoded in DNA.

Why Do Molecular Genealogy?

For some people genealogy is a hobby, while for others it is part of finding out who they are. In general, there is great interest in understanding the origin and history of different populations. Written records, pictures, archeological findings, and other material can provide important information used to elucidate personal and world history. History books are the result of these types of research. Unfortunately, because of the lack of such material, it is not always possible to reconstruct the history of an individual or of a people. This can become a serious obstacle for those who are trying to reconstruct their own origins. Because each of us has received our genetic material from those that came before, we are literally a living record of our family history. The use of DNA in genealogy will bring additional information to those who have reached a wall in their search for ancestors.

How Do You Do Molecular Genealogy?

DNA is extracted from samples donated by volunteers using simple techniques. Although there are many ways to obtain genetic material, the MGRP uses blood samples. Blood samples are used because they yield DNA of superior quality and higher quantity than other sampling methods. Each sample is analyzed at approximately 250 loci (sections of DNA) across the entire genome. This genetic information is correlated to the genealogical data provided by the donors. The only genealogical information gathered is places and dates of birth. Names are substituted with codes to protect participants' confidentiality. People who belong to the same family groups or that share common places of origin will also share common genetic identifiers. This process will allow the MGRP to reconstruct a genetic map of the world that will be used to help people with blocked genealogies in tracing their family origins.

The Y Chromosome

The surname is not the only thing that is passed from father to son. One of the 46 chromosomes in every male is known as the Y chromosome (Ycs). This chromosome is passed almost unchanged from father to son. Unless there is a case of adoption or illegitimacy, the Ycs follows a strictly paternal line. This portion of DNA is relatively small compared to the entire genome, but it is widely used in establishing relationships among individuals sharing the same, or similar last names. Fig.1 shows the inheritance pattern of the Ycs from one

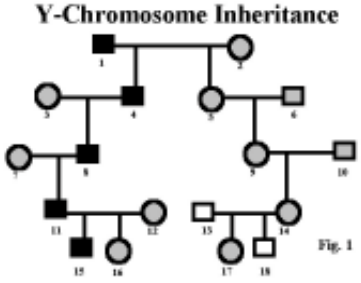

Fig. 1

generation to the next, following the paternal line. Males are represented by squares, while circles represent females. The Ycs of individual 1 is represented in black. Males 4, 8, 11 and 15 have the same Ycs inherited from ancestor 1. Individuals 13 and 18 have different Ycs because they belong to a different male line. This method was used to compare individuals sharing a common paternal ancestor with President Thomas Jefferson with descendants of Sally Hemmings (one of President Jefferson's slaves). The Ycs analysis revealed that Sally's descendants share a common Ycs with President Jefferson's descendants, and therefore belong to the same family.

The Mitochondrial DNA

As explained above, the nucleus is not the only part of the cell that contains genetic

material. Inside the mitochondria there is a small portion of DNA (mtDNA) that is transmitted from a mother to all of her children, males and females. However, only daughters will pass the mtDNA to their children. Because of this property, the mtDNA is also widely used in genealogical research. Fig.2 shows the inheritance pattern of the mtDNA through a number of generations, this time following

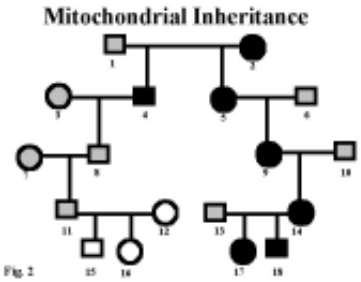

Fig. 2

a strictly maternal line. The color black is used to show all the people sharing the same mtDNA. They all inherited the mtDNA from the common ancestor 2. Individuals 15 and 16 have mtDNA different from individuals 17 and 18 because they belong to a different maternal line.

The Remaining Chromosomes

The Ycs and the mtDNA are relatively easy to use because they experience little recombination and they follow strict inheritance patterns. However, their use places serious limitations on our search for ancestors. Within five generations, only two of sixteen great-great-grandparents can be identified with these two methods (Fig.3). The majority of one's DNA (and one's family history) resides in the rest of the genome—the remaining chromosomes. To better use genetics in genealogical research it will be fundamentally important to use the remaining chromo-

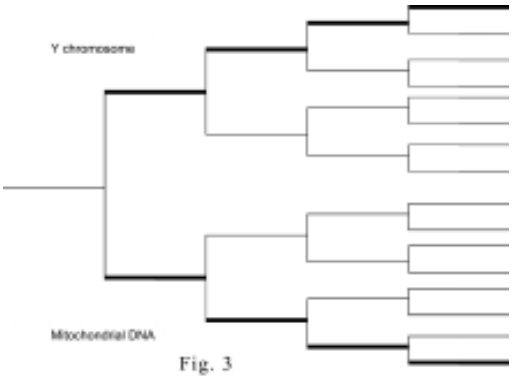

Fig. 3

somes. Unfortunately, these chromosomes are not transmitted in as clear a manner as the Ycs and the mtDNA. In order to trace the movement of these chromosomes through time and space, it is necessary to develop a very large database that will represent world populations.

The Application

Every person that is born has (or had) two parents, four grandparents, eight great-grandparents, and so on. At each generation the number of ancestors doubles. This means that at the thirtieth generation (ca. 750 years) every person alive today should have had over a billion ancestors (Fig.4). However, the estimated world population for the year A.D. 1250 is 400 million. This discrepancy is explained by the fact that at some point in the past we all begin to share ancestors (Fig.5). Therefore the total number of actual ancestors is much smaller than the number of possible ancestors (Fig.6). This means that we are all related to different degrees. These family ties are continually lost at each generation. For example, few people know their third or fourth degree cousins, even though they share ancestors relatively close in time.

The genetic and genealogical work being conducted at BYU through the MGRP will result in the construction of a world-

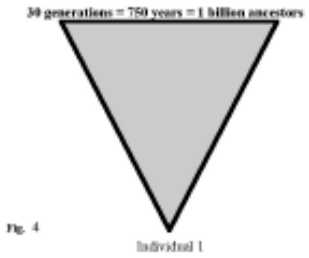

Fig. 4

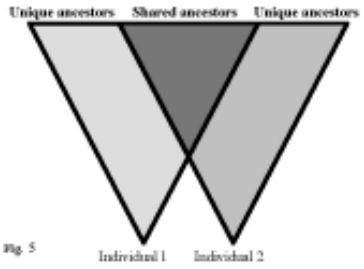

Fig. 5

Fig. 6

wide family tree that will link people together based on the genetic information inherited from common ancestors. This knowledge will be of great use to those people living in countries that do not have written records. For example, researchers from BYU collected over 6,500 DNA samples among different population groups in Peru, with the objective of increasing our understanding concerning the complex history of this country. These samples were collected in large cities along the coast and in small villages in the mountains. Among these samples were some collected from the ancient mummies in San Josè de Moro.

In Peru, most people know little about their past because of the absence of written records, and because of migrations within the country, particularly toward the larger cities. Within just a few generations, these movements have caused a gap between the present population and the past generations. Through the analysis of modern and ancient DNA, it was possible to link together the mummies of San Josè de Moro with the modern population of Cajamarca (Fig.7). Similarly, a number of genetic characteristics of the people living on the shore of Lake Titicaca are also found in other villages. These similar genetic traits shared by people of diverse locations could have resulted from Inca conquests that displaced local peoples to other parts of the country, in order to maintain control over them (Fig.7). This is just one example of how it is possible to shed light on the history of a people with limited written records through genetic analysis.[i]

The Database

In the past eighteen months, over 20,000 people have volunteered to participate in the MGRP by donating blood samples and four-generation pedigree charts. These samples have come mainly from people that

Fig. 7

live in the United States. Other sample collections have taken place in Oceania, Europe, South America and the Middle East.

The first objective of the MGRP is to create a database of correlated genetic and genealogical information representative of worldwide populations. To accomplish this goal, the MGRP will collect at least 100,000 samples from all over the world. The database will continue to grow until every population of the earth is properly represented, both genealogically and genetically. The collection and the analysis of these first 100,000 samples will take at least four years. Participants in the first stage of the project will not receive any information back because of confidentiality issues. As the database progresses and genetic markers of specific populations are identified, the MGRP will use this information to assist individuals with genealogical questions that cannot be answered using traditional written methods.

Special Cases

As explained in this article, the main goal of the MGRP is to complete the worldwide database. This is stage one of the project. Many answers regarding the migration of populations and the reconstruction of family trees will come with the availability of the database. In the meantime, the MGRP can assist in solving a certain number of family situations defined as "Special Cases." At the moment, many of these cases are limited to the use of the Ycs and mtDNA analysis. As the database grows, it will be possible to work on more complex cases, such as the one cited in the introduction. Participating in the construction of the database does not imply necessarily qualification as a "special case." Those individuals who think they might have a genealogical situation that can be solved with the use of genetics should contact the MGRP.

Conclusion

Blocked genealogies, adoptions, and records that are missing or unreliable are all situations in which the MGRP might provide some help. The success of this project is based on the participation of individuals with known genealogies from all over the world in the construction of the database. The MGRP is willing to work with groups that can help in accelerating the process of developing the database. A packet of instructions will be sent to all those interested in organizing such groups. Participants need to be eighteen or older and provide a biological pedigree chart in the form of paper or GEDCOM file.

Qualified medical personnel will perform the blood draws and the MGRP will provide all the necessary equipment for the collection of the samples. Those interested in learning more about the MGRP, or who would like to organize a group to be included in the database can use one of the following contacts:

Write: Molecular Genealogy Research Project 775 WIDB Brigham Young University, Provo, UT 84602
Email: molecular-genealogy@byu.edu
Fax: (801) 378-1576

For more information concerning the progress of the research, visit the Web site at: http://molecular-genealogy.byu.edu/.

---

[i]Joel E. Myres, PhD (1969–2001) was responsible for the collection and analysis of the Peruvian samples. At the time of his premature death, Dr. Myres was in charge of the special cases and co-principal investigator for the MGRP.

Ugo A. Perego has many responsibilites with the Molecular Genealogy Research Group. He is the Director of Public Relations and his main job is to find the 100,000 participants for the project worldwide. He organizes the presentations and blood draws, and writes articles and answers to inquires pertaining to the project. He also maintains the Web site, offers presentations and helps with the special cases. He has a background in accounting and health sciences and a masters in Health Science from BYU.