

Genetics and Armenian heritage FAQ

This FAQ is a primer meant to address questions many have regarding genetics as well as its application to the study of the Armenian people.

Note: Because the subject can be highly technical at times, I've also included *italicized text* for simplified answers.

Having encountered so many people interested in genetics and how it relates to Armenians and Armenian origins, there exists a need to properly address the subject, apply it to the genetic analysis of Armenians, and hopefully help make information accessible for most people. As in most subjects related to Armenians, there's a great deal of politicization and the finer points on the science and recent discoveries tend to get lost. This is a neutral attempt at establishing the landscape of genetics. This is not an argument for or against politicized events, nor is it in any way a complete guide to the complexity of genetics as it relates to Armenians. Think of this as a primer for genetics, after which, speculation, assumptions, various theories of Armenian origins can be discussed in an accurate and educated manner. Genetics is a difficult subject, and simplifying it to the most basic level tends to lose nuances and details along the way. This is an attempt at balancing this complex and difficult subject by simplifying some terms, defining others, and hyper linking to other sources for more information should one be interested.

FAQ list:

What is population genetics?

What is Y Chromosome DNA (Y Chr DNA) analysis?

How is ancestral information determined from the Y chromosome?

How diverse are the Y chromosome haplogroups among the Armenian population?

How many haplogroups do Armenians have as a population?

Do Armenians have a dominant Y chromosome haplogroup?

What is the difference between haplogroup frequency and variance?

Does having a high haplogroup frequency mean place of origin for a particular haplogroup?

What is mitochondrial DNA and how is it inherited?

How is ancestral information determined from mitochondrial DNA (mtDNA)?

How diverse is mtDNA in Armenians?

How is it that Armenians account for more mtDNA diversity and yet, the scientific consensus is that humans originated in Africa?

What is autosomal DNA (auDNA)?

How is auDNA analyzed?

What is a reference population?

Why can't you use auDNA for direct ancestry information but can use it for determining relatedness among individuals?

What other things can be determined by aDNA?

What is Whole Genome Sequencing (WGS)?

What is Archaic or Ancient DNA (aDNA)?

How are aDNA and ancestral DNA (mtDNA and Y Chr) used together?

Are there any aDNA results for Armenia?

How is the aDNA analysis related to linguistic/cultural origins?

Is there an Armenian gene?

What do the studies currently say about Armenian origins?

What does this say about the origin of the Armenian language?

How else can uncovering the genetic legacy of Armenians as a population prove useful?

How does the Armenian population compare genetically to neighboring populations such as Turks, Assyrians, Kurds, Georgians, Azeris, and Iranians?

Do we know who the Armenians are with all this recent genetic information?

What does genetics have to say about the antiquated notions of a so called "Armenoid race"?

What is cultural diffusion?

Can you give an example of cultural diffusion?

What is demic diffusion?

Can you give an example of demic diffusion?

What is Neanderthal DNA?

Are there other hybridization events besides Neanderthal?

A glossary of terms:

DNA: The individualized code of each person that is inherited and passed on to subsequent generations.

Chromosome: The storage method for DNA (Humans have 23 pairs in most cells, or 46 Chromosomes in total)

Y-Chromosome DNA: DNA found in the Y chromosome, the one that determines male gender; inherited from the father's side.

Mitochondria: The tiny components of cells that, like a motor, provide energy

Mitochondrial DNA: DNA found in the "motor" of most cells, passed down from the mother to her children

Autosomal DNA: 22 of the 23 pairs of Chromosomes

Sex Chromosomes: the 23rd chromosome pair...the one that determines gender (male or female)

Archaic DNA: Ancient DNA collected from skeletons

Genome: a collection of genes that is personalized to each individual

Gamete: Cells used in Sexual reproduction (sperm and egg cells)...they only have a single set of Chromosomes (23 unpaired).

Paternal: The father's side (father's father's father's...)

Maternal: The mother's side (mother's mother's mother's...)

Nucleus: The “control center” of the cell where the Chromosomes are found.

Nucleated: A cell with a nucleus.

Molecular clock: A method, based on mutation rate, of determining how many generations have passed.

Junk DNA: 90% of DNA that is not used for any particular function (mutations have no consequence).

Haplogroup: A way of grouping people based on shared genetic ancestry (people of the same haplogroup have the same exact mutation and thus same common ancestor).

What is population genetics? Population genetics is the field of study related to how various human populations are related to each other, through the analysis of the respective genomes. How do you determine genetic ancestry? There are multiple methods of informative analysis including Y chromosomal (Y-Chr) DNA, mitochondrial (mtDNA) DNA, autosomal DNA (Au-DNA), and Whole Genome sequencing.

What is Y chromosome DNA analysis? All human (and generally most mammals) males carry a Y chromosome that was passed along the paternal line. Of the 23 pairs of chromosomes that humans carry, the 23rd pair is referred to as the sex determining pair. All nucleated cells in the body, except for gametes (those involved in fertilization), have 23 pairs of chromosomes. That means that there are 46 chromosomes in nearly all cells of the body, other than the gametes. Gametes such as sperm and egg cells only have a single pair of 23 chromosomes and upon fusion (fertilization), the newly formed zygote, a combination of a sperm and egg cell, has once again a complete set of paired chromosomes (other than in the gametic, or sex cells as already mentioned). This form of combination and recombination allows for a potentially unlimited mixture of genes, and makes it very difficult to trace ancestral information. However, the Y-chromosome is a perfect model, because it is inherited intact from father to son. Essentially, recombination is the reason for the near infinite amount of variation in potential offspring, from the same two parents. In this, various genes within the genome can swap places and though it can provide an interesting look at the mechanisms involved in hereditary genetics, it's not very useful in terms of tracing ancestry. The difficulty in detecting when and how the recombination events occurred makes it an expensive and ineffective method of ancestral analysis. It may be useful for health purposes, or uncovering the potential or presence of disease, but not very useful in population genetic studies.

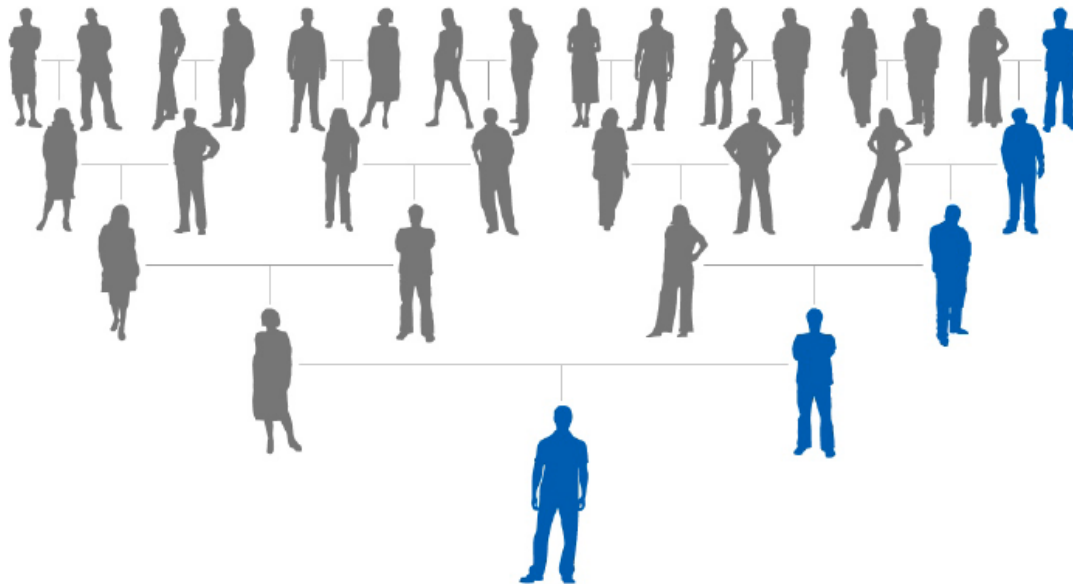
The Y chromosome is a segment of DNA that determines male gender. It's passed down from father to son.

How is ancestral information determined from the Y chromosome? Since the Y chromosome is passed down paternally, from father to son largely intact and unchanged, all one has to do is look for the occasional mutation in a sequence of DNA on the Y chromosome that has no genetic function (90% of our genes are called Junk DNA for this reason), and determine the rate of

[mutation](#) per generation. This is called a [molecular clock](#) and can be used to date the mutations on any particular sequence of DNA. If we determine that a certain region of DNA mutates every 5 generations, and if we detect 5 mutations, we know that 25 generations have passed. Thus, as we compare the Y chromosome of various human males with each other, we find that approximately 150,000 years ago (130,000- 170,000 years are the most recent estimate, though only 5 years ago, it was calculated to be about 90,000 years old), all human males alive today had a common male ancestor who gave rise to all the Y chromosomes found in the world. This is otherwise known as [Y-chromosome Adam](#). We group Y chromosomes due to mutations into [haplogroups](#), and even [subclades](#) (subgroups) within each haplogroup.

Since the Y chromosome is passed down from father to son, it can be traced back down the male line to the distant past. All male humans in the world can be placed into about 20 different Y chromosome groups, called haplogroups. The haplogroups also have their own subgroup.

Human Chromosomes (Y-DNA)



Because of its unique pattern of inheritance, Y chromosome DNA is only informative about a small fraction of ancestors - just those in the direct paternal line.

How diverse are the Y chromosome haplogroups among the Armenian population? Armenians, as a [population](#), have one of the highest Y chromosome haplogroup diversities in the world. This generally means that the paternal origin of Armenians is from multiple sources of both autochthonous and migratory groups. As Armenians formed into an ethno-linguistic group, they coalesced from various peoples who had migrated into the Armenian highlands since humans left Africa (modern humans began their wave like migrations out of Africa approximately 70,000 – 100,000 years ago). In essence, the Armenian Highlands, and surrounding areas, such as the Cappadocian Highlands (aka Anatolian plateau) and Iranian Plateau, were a very early melting

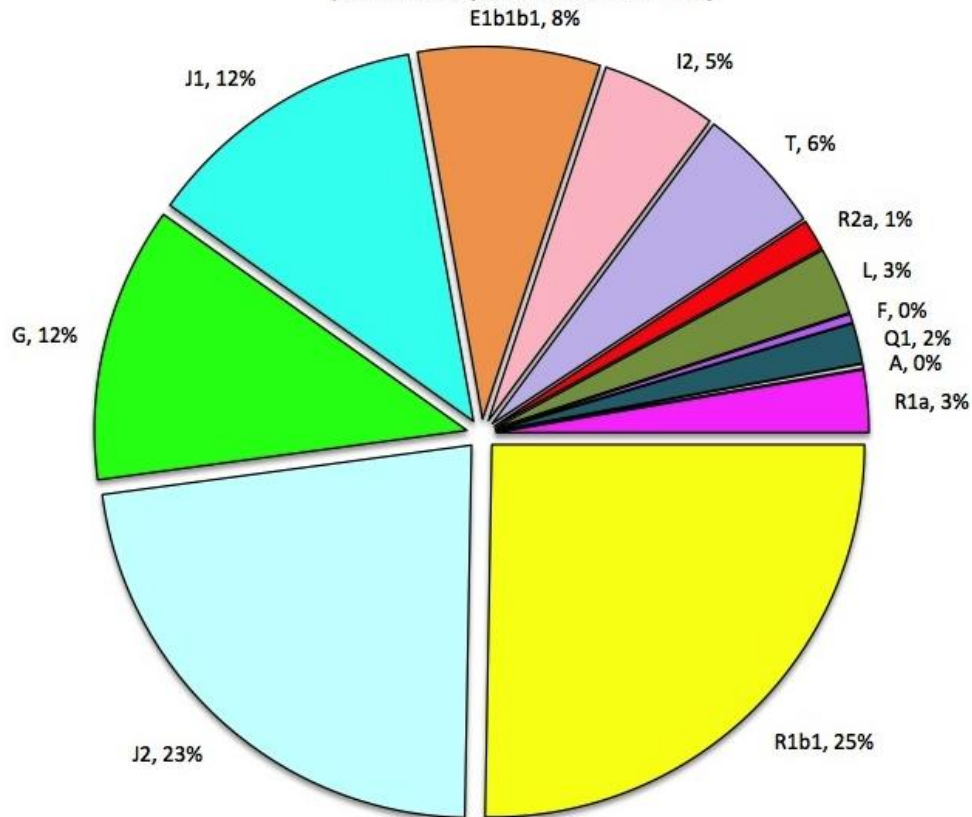
pot and migratory hub outside of Africa from which some of the most important Eurasian haplogroups such as [G](#), [J1](#), [J2](#) and [I](#) to name but a few originated. In fact, a high diversity of Y-chromosomal DNA is present among nearly all West Asian and Central Asian populations.

Armenians have a very high diversity of Y chromosomes, accounting for nearly 70% of the haplogroups. In fact, Armenians also have a high diversity of haplogroup subgroups (subclades).

Armenian DNA Project - Y-DNA Haplogroup Distribution

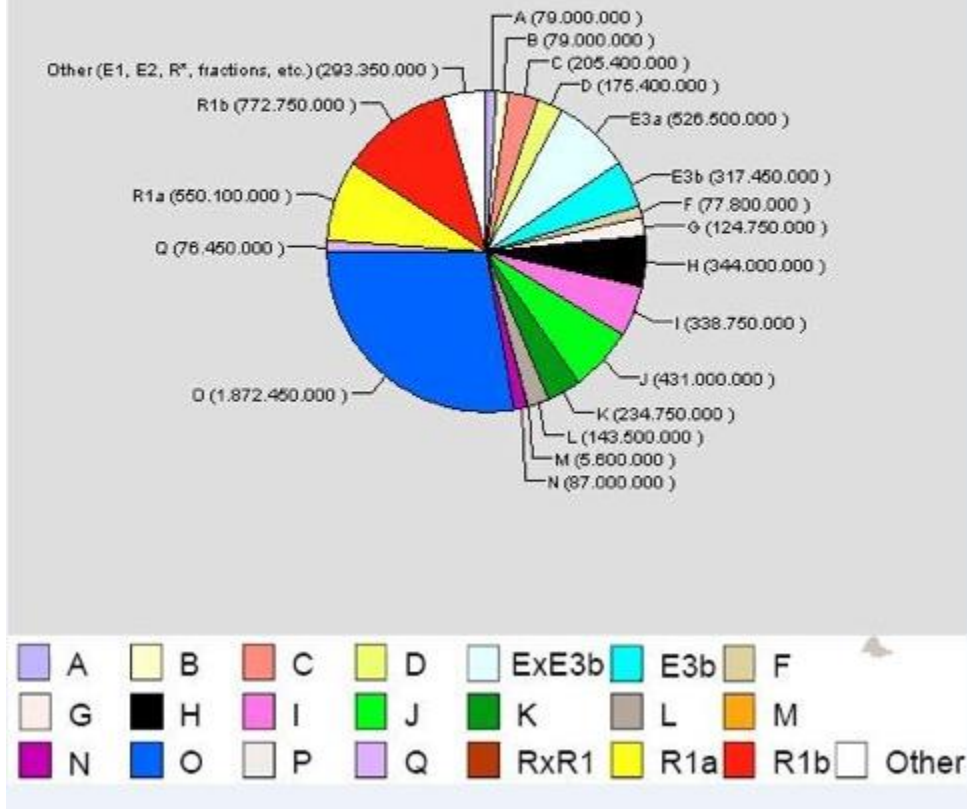
1 september 2013 / n = 547

(less 29 known paternal cousins: n = 518)

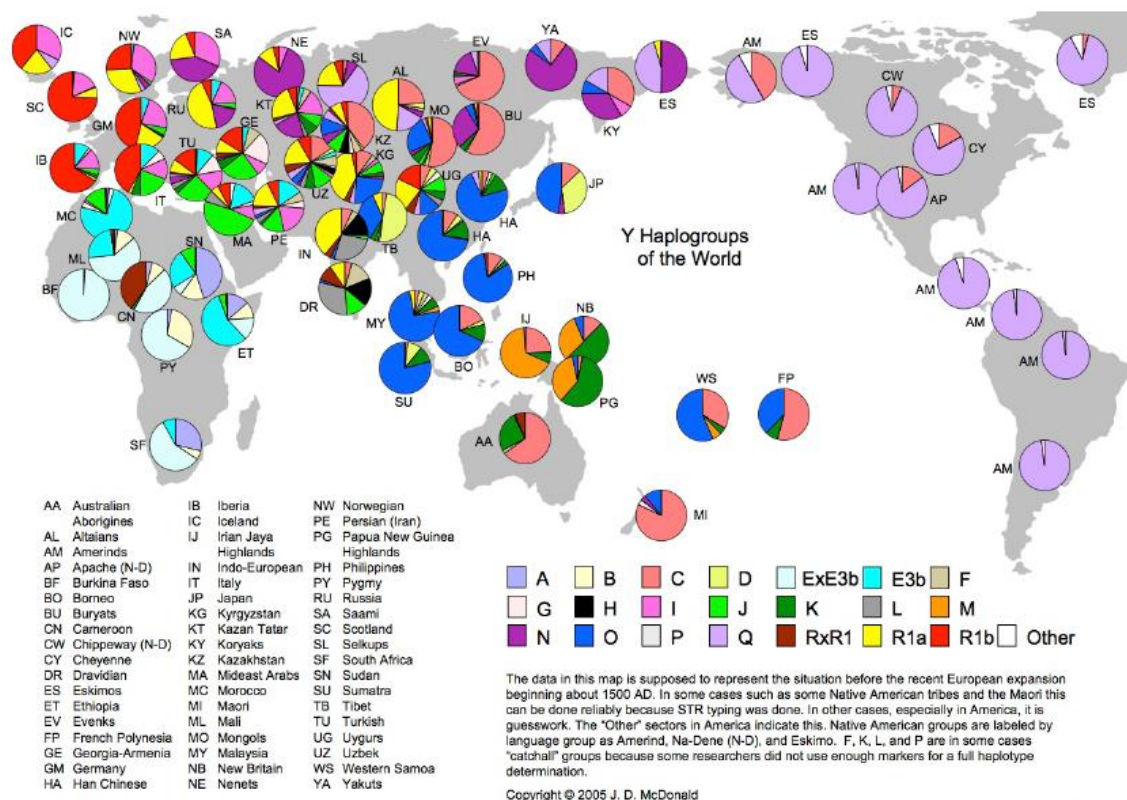


■ R1b1 ■ J2 ■ G ■ J1 ■ E1b1b1 ■ I2 ■ T ■ R2a ■ L ■ F ■ Q1 ■ A ■ R1a

Y-dna of the world's 6.7 billion people



How many haplogroups do Armenians have as a population? Out of the 20 or so major haplogroups and haplogroup branches present in the world, Armenians belong to two thirds of them (namely [E](#), [F](#), [G](#), [I](#), [J1](#), [J2](#), [K](#), [L](#), [Q](#), [T](#), [R1a](#), [R1b](#), [R2](#)). There are no Armenians in haplogroups [B](#), [C](#), [D](#), [M](#), [O](#) and only singletons in haplogroups [A](#), [H](#) and [N](#).



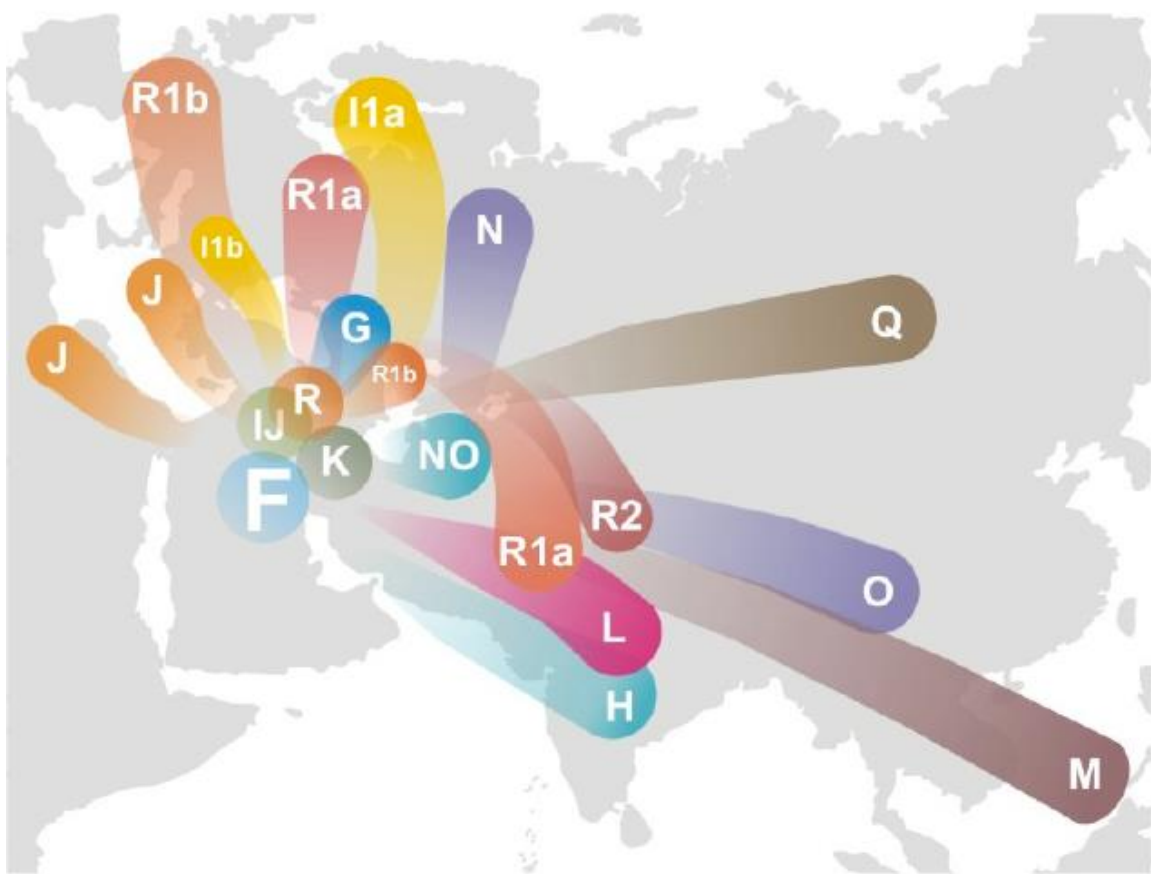
26

Do Armenians have a dominant Y-chromosome haplogroup? Two haplogroups could be viewed as codominant (R1b1 and J2), as they generally form around 25% of the male genome. An additional 3 haplogroups are found to be moderately widespread, present in around 10% of the population (G, J1, and E). Two more that are low to moderately widespread (I2 and T), present in around 6% of Armenian males. The rest are found in low frequencies, around 3% or less with isolated cases of a few other haplogroups.

Armenians have two haplogroups that occur in high frequencies, three more in moderate frequencies, two in lower frequencies, and the rest in low to very limited frequencies.

The distribution of Eurasian Y-DNA from ancestor F

The Geographical Relationship of Macro-Haplogroup F to its Descendant Haplogroups



What is the difference between haplogroup frequency and variance? Haplogroup frequency generally implies the amount of people found to have a particular haplogroup. Variance relates to the diversity present of that particular haplogroup. As an example, of the codominant Y haplogroups mentioned, two called J2 and G, and the diversity of subclades found in Armenians belonging to J2 and G are quite high.

Frequency of a haplogroup means the amount of people of a particular haplogroup encountered; variance is the diversity of the haplogroup (many different subgroups).

Does having a high haplogroup frequency mean source of origin for a particular haplogroup? To find the source of haplogroup origin, or origin of any particular DNA segment, gene, or even species, we must look for high diversity. The highest diversity is the source of likely origin. For example, we know that potatoes originated in the Andes mountain range, because the highest diversity of wild potatoes, as well as domesticated potatoes, happens to be in the Andes Mountains between Peru and Bolivia. Even if potatoes have a higher frequency of consumption in say Ireland, we know Ireland cannot be the origin of potatoes, because the diversity is so low, and there are no wild types of potatoes. There are perhaps 3 varieties of potatoes that are grown

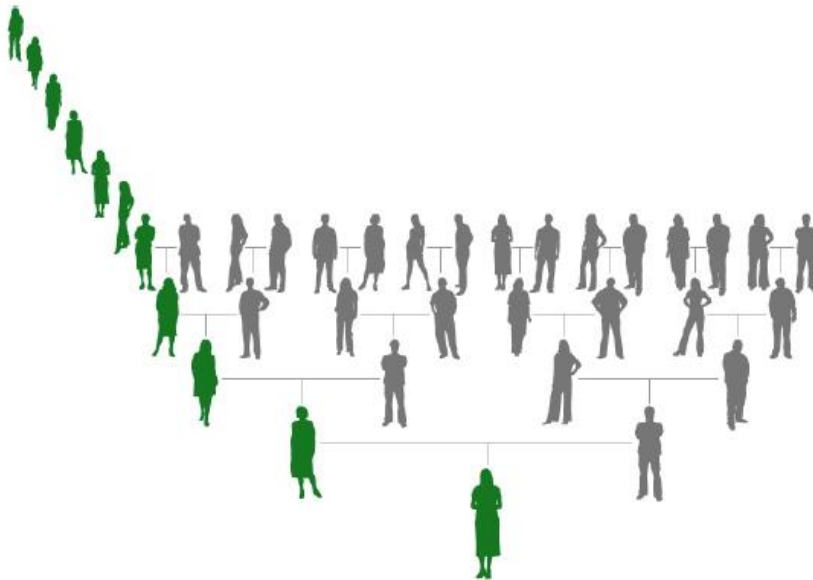
in Ireland, and thus, must have been introduced there. Similarly, even if certain human populations have a very high frequency of a particular haplogroup, we must look at the diversity to determine origin. Haplogroup J2, from the previous example, is found in highest frequencies among some populations in the Caucasus (such as Chechens and Dagestani), but the variance, or diversity of J2 is much lower than it is in populations derived from the Highlands (such as Armenians, Kurds, Assyrians, Turks, Northern Syrians, etc.) This points to a “[founder effect](#)”, in the sense that just as a few varieties of potatoes found in Ireland indicates which species were successfully introduced there, we can use a similar method for human populations. The same can be said about the high concentrations of G2 in Georgians and J1 among the tribes of Dagestan in the Caucasus.

No, high haplogroup frequency is not always associated with haplogroup origin. A few related males, such as brothers, can move to a village, replace all the males, and over time, if they're successful reproducers, the whole area can have the same male relative. For example, if these brothers had a hypothetical haplogroup Z of their Y chromosome, then over time, their offspring would all have haplogroup Z. This would mean that the region would have a high frequency of haplogroup Z.

What is mitochondrial DNA (mtDNA) and how is it inherited? Unlike DNA found in chromosomes, which are present in the nucleus (control center) of most cells, mitochondrial DNA is found in another component (organelle) of a cell called the [mitochondria](#). These mitochondria can be viewed as the engine, or powerhouse of the cell, as they provide the energy necessary to run things. The mitochondria also have their own DNA, shaped like a ring (as opposed to the X shape of the chromosomes). Mitochondrial DNA (mtDNA), unlike Y chromosome DNA, is passed on down the [maternal](#) line, from mother to offspring. Both male and female offspring inherit their mother's mitochondrial DNA, but only females can further pass it along to their respective offspring. The reason for this is that the egg cell produced by the female, has nearly all the cellular functions of other cells, and only requires the 23 single chromosomes of the sperm to be fertilized. In it is a nearly complete house, including its own engine (the mitochondria) with only the father's chromosomes necessary to start human life. Thus, we all inherit the same mitochondrial chromosomes present in our mother's cells, which are in turn inherited along the maternal line back into the distant past.

Mitochondrial DNA is a type of DNA that comes from another part of the cell (mitochondria), unrelated to DNA from the chromosomes (where most of human DNA is found). It's passed down from a mother to all her children (both males and females), but only her female children can pass it down to her children. Basically, it's passed down from the mother's line.

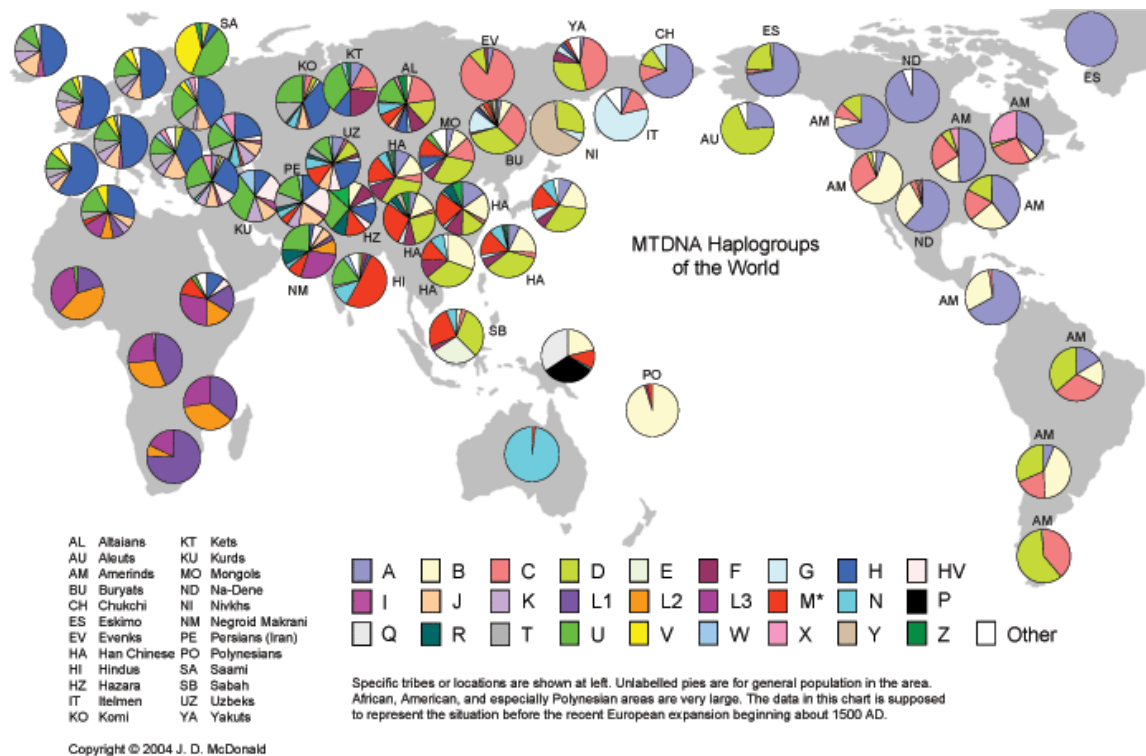
Mitochondrial DNA (mtDNA)



However, since it passes from generation to generation with very little change, mitochondrial DNA is a rich source of information for deep ancestry, going back thousands of years.

How is ancestral information determined from mtDNA? Just as neutral DNA on the Y chromosome was investigated for mutations and mutation rate, certain neutral segments of mtDNA are also looked at. From this analysis, we can determine that all mtDNA observed in the world originated from one particular woman, called [mitochondrial Eve](#), approximately 150,000 – 200,000 years ago. Just as in Y Chr Adam, mtDNA Eve lived in Africa (we know this from the highest levels of genetic diversity in humans being found in Africa).

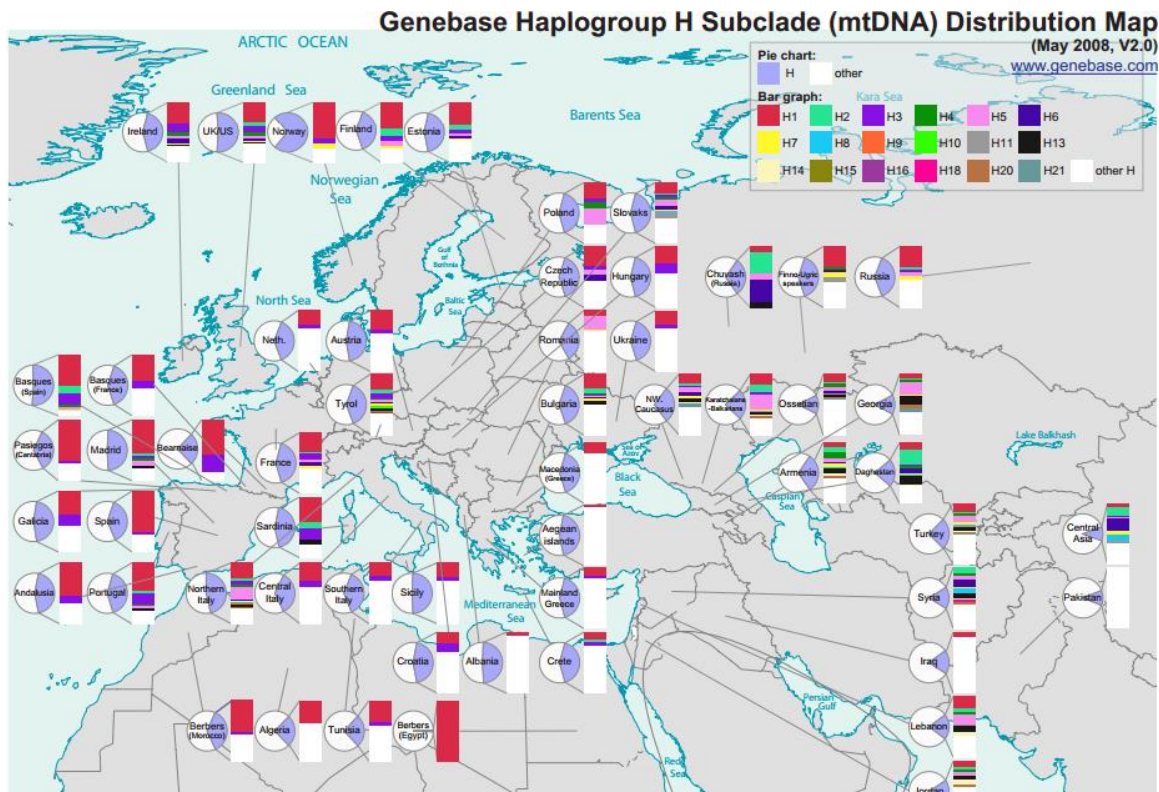
Just as with the Y chromosome, all mitochondrial DNA in the world can be grouped into haplogroups and subgroups of a haplogroup



How diverse is mtDNA in Armenians? Of the 24 or so mtDNA haplogroups present in the world, Armenians account for half of them (namely H, HV, I, J, K, N, R, T, U, V, W, X). There are no Armenians in haplogroups A, B, D, G, L, M, P, Q, Y, Z and only singletons in haplogroups C and F. Of these, one accounts for nearly 30% of haplogroups, three between 10-15%, four more between 5-10%, and the rest account for less than 3%. The dominant Eurasian haplogroup (relative to the others), called H, also has a very high variance in Armenians. This means that Armenians have many different subclades of H (H1, H2, H4, H5, H6, H7, H8, H13, H14, and H20); while populations like the Basque, who are truly dominant with this haplogroup, have but two dominant subgroups (H1 and H3).

Armenians account for nearly half of all mitochondrial DNA haplogroups that occur in world.

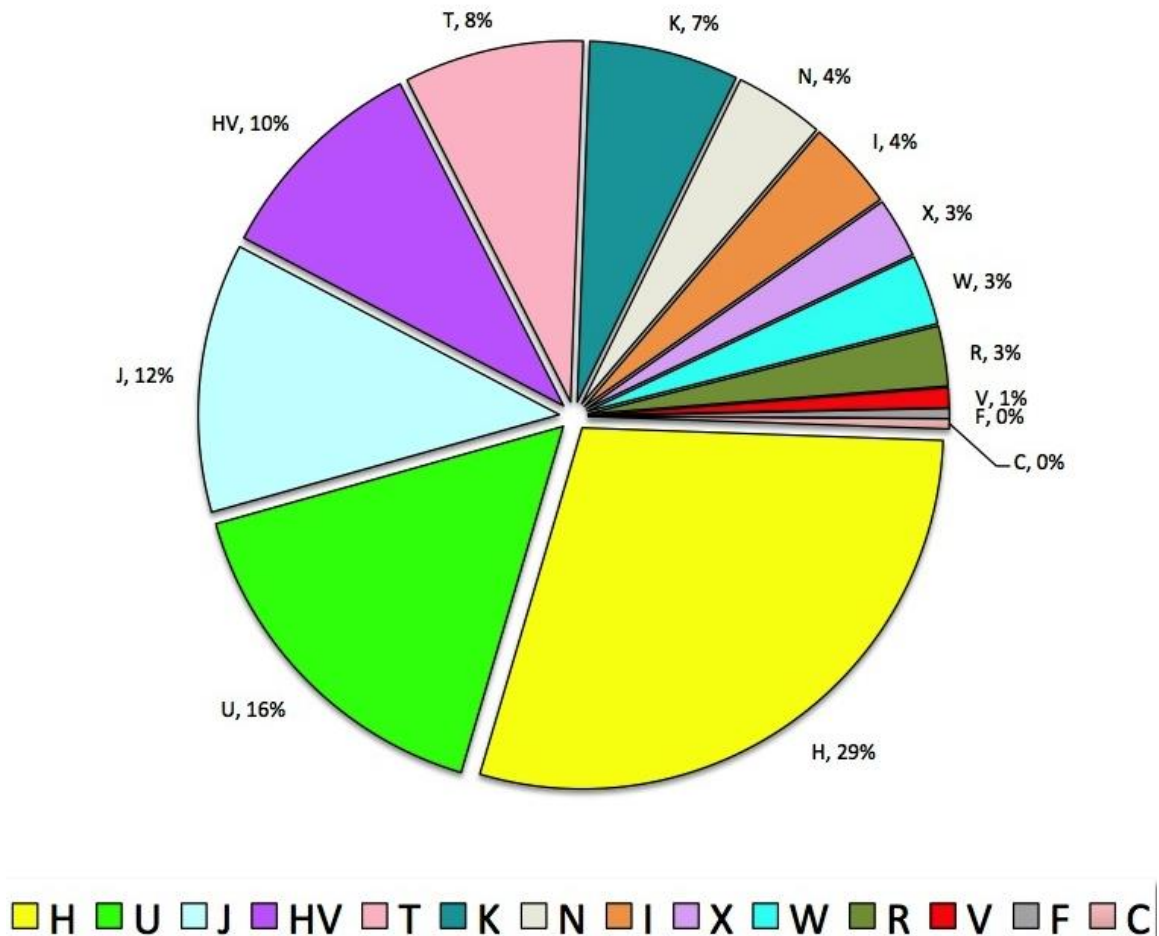
With Haplogroup H as an example, the following distribution map illustrates just how much subclade variability is present in West and Central Asia, compared to other areas of Eurasia. Note how evenly distributed the H subgroups are within Armenians compared to most other populations. This is likely a result of West Asia, particularly the Armenian Highlands, as a migratory hub outside of Africa. In essence, the bands of early humans who settled in West Asia, then eventually migrated elsewhere, with their genetic roots and traces evident across the world. Armenians are not only a good representation of the genetic variability present thousands, if not tens of thousands of years before, they also present a population of that era that has largely preserved the genetic legacy of the first migrants to settle outside of Africa.



Armenian DNA Project - mtDNA Haplogroup Distribution

1 september 2013 / n = 230

(less 6 known maternal cousins: n = 224)



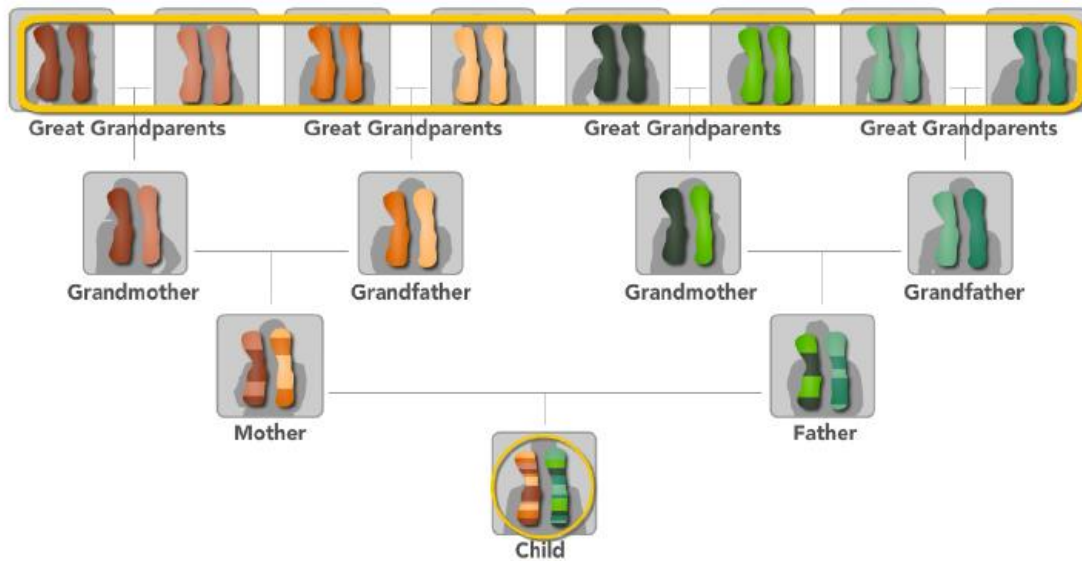
How is it that Armenians account for more mtDNA diversity while the scientific consensus is that humans originated in Africa? When we look at [genetic diversity](#) in full, we don't just look at one particular type of DNA, but the whole genome (more on this later). When we compare DNA as a whole, Africans have the highest diversity, followed by Near Easterners, Central Asians, Europeans, Austronesians, East Asians, and ultimately Native Americans, who have the lowest diversity.

Usually, high diversity of a species means that it had its origins in that region, however, mitochondrial DNA or the Y chromosome is only a small part of the whole human genome. Mitochondrial DNA and the Y chromosome account for only a very small fraction of DNA (in a male, together they account for less than .03%, while in females, without a Y chromosome, it's even less). When all the DNA of humans is looked at, the greatest diversity is located in Africa, which is the region of origin for modern humans.

What is autosomal DNA (auDNA)? Having mentioned that all nucleated cells (some cells, like red blood cells, lose their nucleus upon maturity and thus don't have any chromosomes) have 23 pairs of Chromosomes, the autosomal Chromosomes account for the first 22 of the 23 pairs (the 23rd pair being the aforementioned sex chromosomes). Most of our functional genes, as well as most of our junk DNA are found among the autosomal chromosomes.

Autosomal DNA is the DNA located in 22 of the 23 pairs of chromosomes in humans. The 23rd pair, called sex chromosomes are not considered autosomes.

Human Chromosomes



Their autosomes were recombined before they were passed to your parents, and so on. So your autosomes are random mixtures of your great-grandparents' autosomes.

How is auDNA analyzed? Generally, a few segments of various genes present on autosomal chromosomes are looked at across human populations to determine similarity. It's a type of population comparison to determine relatedness. For best results, a reference population is used as a comparison.

Only a few selected parts of the autosomes (the first 22 pairs of chromosomes) are looked at for variations between human beings. These parts include genes for certain physical traits like hair and eye color, as well as genes for health and disease.

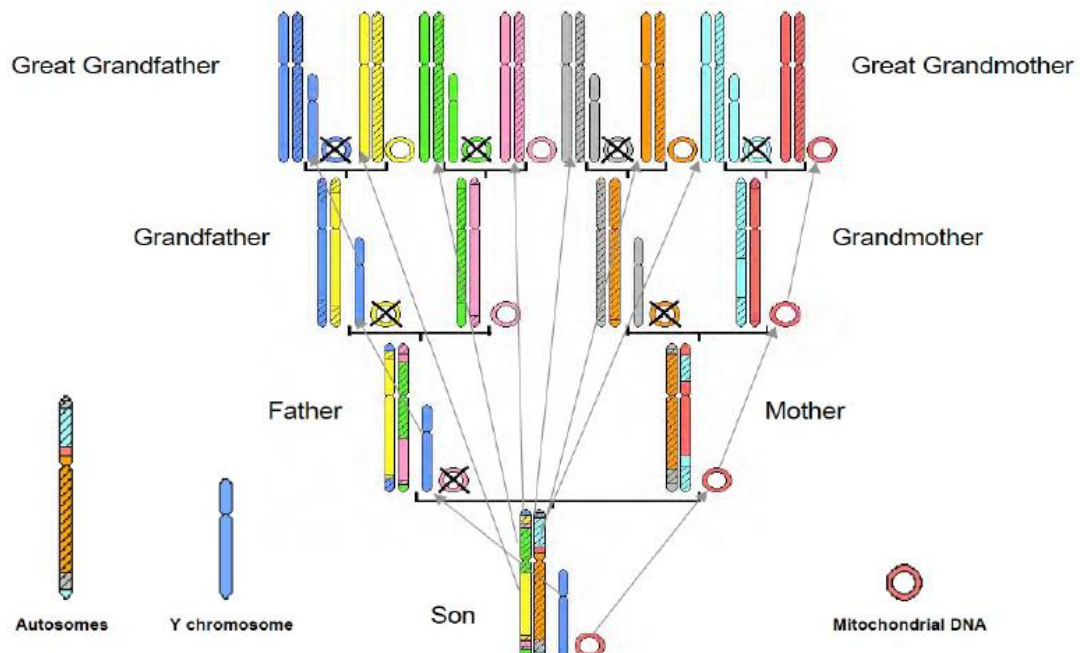
What is a [reference population](#)? A reference population allows for the comparison of individuals to various population groups to determine possible origin. A good reference population is built up from data collected by individuals of a particular conserved group. For example, an optimal reference example for an Armenian would be a person who has all four of their grandparents originating in a particular historic region of Armenia, say the city of Van. Repeat this collection for a few hundred individuals who have a specific regional affiliation and you can use this as a reference population to compare others. The more specific the region, the more detailed the reference analysis. For a starting point, collecting from individuals with 4 Armenian grandparents is instrumental, but to further obtain detailed analysis, ethnic group isn't enough. Instead, one must gather specific regional data (optimally 4 grandparents from the same village) and do this for a number of people from the same location, reducing statistical error and noise.

A reference population is a group of people from a particular region (population) that others are compared to for results. To simplify this, think of a reference population as a solid color, like red, yellow, or blue, and all other colors compared to these three. So a purple would be compared to a red and blue, an orange to a red and yellow, etc. Similarly, a reference population allows the comparison of a person's genetics to various references and determines closeness.

Why can't you use auDNA for direct ancestry information but can use it for determining relatedness among individuals? Since most DNA is shuffled across generations (other than Y Chr and mtDNA), it's impossible to determine direct ancestral DNA from any one Autosomal Chromosome. Statistically, by translocations, gene swapping, and other genetic mechanisms, Chromosomal DNA is too messy for direct ancestral information. Yet, since two siblings will share roughly 50% of their DNA, the two siblings will have half the same autosomal DNA segments. Furthermore, a grandparent or 1st cousin will have ¼ of the same autosomal DNA segments. This way, relatedness among individuals can be measured. Further, by comparing the unrelated autosomal DNA of one individual to another, their genetic similarity can be compared. For example, comparing the auDNA of a German with that of a Swiss will show that they are more closely related than that of a German and a Mongolian.

Since all chromosomes, besides the Y chromosome mix with their pair and swap genes, it's impossible to trace one chromosome to any ancestor. The reason the mitochondrial DNA and Y chromosome make for such accurate ancestral information is due to the fact that other than the occasional change due to mutations, they don't have any other DNA to swap genes with.

Y-DNA & mtDNA & autosomal DNA



The Armenian DNA Project analyzes (a) the male & paternally inherited Y-chromosome, (b) the female & maternally inherited mitochondrial DNA, and (c) the autosomal DNA of 100% Armenians.

What other things can be determined by auDNA? Certain traits among [expressed genes](#) (those 10% of genes not considered “junk DNA”) can be compared in individuals. For example, the trait for nearsightedness, present in a few autosomal genetic locations, can be compared and useful in determining whether the individual has the condition or not. In essence, Au-DNA can provide a wealth of information on health and physical traits (even complex ones such as hair and eye color), with a predictive value based on the amount of genetic locations of such traits (polygenic traits).

Physical traits such as hair color, eye color, nearsightedness, the ability to taste certain compounds, as well as health related functions like the increased or decreased likelihood of certain diseases can all be traced to auDNA.

What is Whole Genome Sequencing (WGS)? Affordable Whole Genome Sequencing is the ultimate goal of Genetic research. Instead of looking at a few isolated genes, or particular genetic vehicles (such as Y-Chr, or Mt-DNA), WGS looks at the whole genetic sequence found within the 23 pairs of human chromosomes. This is a very costly and time consuming endeavor, but as the technology improves and costs are reduced, ultimately, WGS may be accessible for everyone. With this, better health information can be determined; [Drug Metabolizing Enzymes](#) (DME) can be fully mapped, which will allow for personalized pharmaceutical treatment for a given individual's detailed genetic profile and response rate. Further, it will allow for [Genome Wide Association Studies](#) (GWAS) among populations, such as Armenians, to determine higher causes of mortality and morbidity and improve the efficacy of treatment. To a rudimentary extent, autosomal DNA analysis can also be used for these GWAS and DME studies, but ultimately, more detailed WGS analysis will provide even more accurate personalized medical outcomes.

Whole Genome Sequencing attempts to analyze all the DNA in full, not just a few regions of the autosomes. It looks at all 23 pairs of chromosomes, as well as mitochondrial DNA.

What is Archaic or [Ancient DNA](#) (aDNA)? aDNA is DNA that's extracted from skeletons found in archaeological or paleontological sites. DNA is notorious for degrading quickly and the conditions required to preserve it in an amount useful for genetic analysis are very particular. Cool and dry conditions best preserve aDNA. Another complication for collecting aDNA is contamination, as anyone or anything (with its own DNA) that comes into contact with the sample can accidentally introduce his or her own DNA and ruin the analysis. Human aDNA is very useful because it can show the genetic population change in a particular region with time. The technology has only recently become available and is still working its way to be cost effective enough for mass analysis.

Archaic DNA is DNA that's extracted from ancient skeletons and used to compare change over time for modern human population of a particular region.

How are aDNA and ancestral DNA (mtDNA and Y Chr) used together? aDNA and ancestral DNA (Y-Chr, and mt-DNA) can be used in conjunction to reinforce each other and contribute to the development of population origins. Modern population DNA, through aforementioned genetic variability, can only infer past migratory events and haplogroup origins, but ancestral DNA can confirm or falsify such inferences.

Archaic DNA can be used to directly validate assumptions (prove or disprove) and implications that come from comparing mitochondrial and Y chromosome DNA in populations.

Are there any aDNA results for Armenia? Currently, as of this writing, there are no aDNA results available from Armenia, Historical Armenia, or regions previously inhabited by Armenians. Unfortunately, it's a highly politicized issue and many governments fear the results may cause political and social upheaval. However, there are a few samples currently undergoing final analysis and haplogroup typing, with hundreds of others awaiting initial analysis. Unfortunately, the limitations are financial. Further, the core of Historic Armenia, the Armenian Highlands itself, is largely ignored, as the samples currently undergoing and awaiting analysis are from the modern Republic of Armenia (a small sliver of Historical Armenia). There is a project called [BEAN](#) (bridging the European and Anatolian Neolithic) that is attempting to collect, analyze, and catalogue aDNA samples from the Balkans to the Highlands, but none of the present results are anywhere near the Highlands. Most of the samples from present day Turkey are coastal samples that will not help elucidate terrestrial migrations during pre and post Neolithic times, or make a dent in determining the origins of the Armenian people.

How is the aDNA analysis related to linguistic/cultural origins? Population origins are very complex, and genetic studies alone can never fully explain population origins. It takes a combination of archaeology, linguistics, history, and genetics to even come close to starting to answer origin questions. Genetics is but one tool in analyzing various populations and ultimately determining origins.

Is there an Armenian gene? There is no such thing as an Armenian Gene, as Armenians carry the genetic legacy of all ancestral populations that [coalesced](#) to form the Armenian ethnic entity. Many of the neighboring populations of Armenians also share similar ancestral origins; regardless of linguistic or cultural divides.

What do the studies currently say about Armenian origins? Unfortunately, not enough information is available to speak definitively, but due to current data, some things can be predicted with reasonable accuracy. Genetically, Armenians as a population are largely indigenous to the areas of inhabitation (the Highlands), and the amount of recent distant [genetic admixture](#) is highly limited. As a whole, due to the conserved nature of the population, Armenians

are a good representation of the various tribes and peoples that populated the Near East thousands of years ago. Further, Y Chr, and mtDNA studies seem to indicate that the Highlands were an important migratory hub for early humans, particularly after the Neolithic.

Current studies seem to indicate that although Armenians are genetically diverse; their recent genetic history shows very little traces of mixture from external sources, particularly distant ones. This implies that something among the Armenian population, whether it was the Armenian Church/development of religious institutions, increased social order, or some other mechanisms (such as the escape into mountains to preserve heritage and culture), prevented most gene flow within the last 3000 years.

What does this say about the origin of the Armenian language? Unfortunately, genetics cannot be used (yet?) to reconstruct language or linguistic origin for Armenians. Since Armenians are genetically so variable, with no highly dominant haplogroups, it's hard to elucidate which of the groups was responsible for introducing or natively producing the proto-Armenian language. Historically, it's known that Armenian was spoken by the Achemenid Persian dynasty (500BCE). Anything before this is conjecture and as of yet, has no genetic basis in determination.

How else can uncovering the genetic legacy of Armenians as a population prove useful? Since much of the Eurasian population is a result of migratory paths out of the [Near East](#), and Armenians represent a conserved legacy of the genetic admixture found in the historical migrations and inhabitation of the Near East, particularly the [Fertile Crescent](#), uncovering detailed analysis of Armenians can help research into various other populations throughout Eurasia and even retroactively back into Africa. An early conserved melting pot is a rare commodity in population genetics and the better and more detailed the reference data collected from Armenians, the better analysis of historical migration within, around, and distant to the Highlands.

How does the Armenian population compare to neighboring populations like Turks, Assyrians, Kurds, Georgians, Azeris, and Iranians? Regardless of the various roots of the languages spoken in the region, the closest genetic population to Armenians happens to be regional neighbors. As an example, though the Turkic language and peoples migrated from Central Asia into the Near East, the modern populations of Turkey or Azerbaijan have more similarity to the natives of the region than they do with Central Asia. Though a sizable percentage of the genome of Turks and Azeris (on average, 3-10% of the [genetic markers](#) of an individual in Turkey or Azerbaijan are Central Asian) contain genetic traces (markers) present in Central Asia, the majority of their genome is indigenous. This information points to [language replacement](#) as a model of cultural diffusion (see "[Are Turks acculturated Armenians](#)" for more information on this). In essence, a small group of individuals migrate to a region, assimilate it by dominating culturally, politically, and socially, eventually replacing the original culture with the newly introduced one. One example is the Spanish domination of the Americas. Though the Spaniards left limited

genetic traces in the native population of the Americas, large swaths of Central and South America now retain the culture, language, and politics of Spain.

Do we know who the Armenians are with all this recent genetic information? There is no short answer to this question. As a population, Armenians are an indigenous people, broadly native to the Highlands and surrounding areas and [relatively unmixed with the wave of invaders](#) who passed through or settled in the area within the last few thousand years. The ancestral groups that ultimately formed the Armenian national identity also contributed to the national identity of neighboring populations. In fact, because Armenians historically originate from such a large geographic area, it's important to investigate the population in a very detailed manner, involving the genetic characteristics of not just historical Armenian regions, but even perhaps at the micro-region level. Once a genetic map of the Armenian Highland is created, and a solid reference population for Armenians is established, it will then be possible to have detailed comparisons of individuals to various reference populations. For example, a person of Near Eastern heritage will be able to compare their genome with reference populations of Armenians, Assyrians, Bedouins, Palestinians, and Druze, and ultimately determine which groups they are more closely related with. An even more intriguing possibility, should detailed regional data be collected in sufficient amounts, an Armenian with unknown historical roots may be able to recover their regionally specific geographic legacy.

What does genetics have to say about the antiquated notions of a so called "Armenoid race"?

Most attempts at racial classification are based on outdated and obsolete models. Since the advent of genetics and DNA analysis, closer and more accurate scrutiny of human beings can be made. Genetics has shown that the concept of race is not applicable to human beings. It's better to refer to different groups of human beings as a population, since all humans are essentially of the same species. Because physical anthropologists and scientists of earlier eras did not have the tools to investigate genetics, or even the concept of genetics, they frequently classified human beings based on physical traits and features alone. Though this was a crude method, it was the best they had available for their time. Due to certain cranial and facial features (morphology), different groups of humans were classified into different "races". Thus, the so-called Armenoid racial type was created. This specifically described a combination of features that was present among Near Easterners, particularly Armenians, Syrians, Lebanese, and Anatolian Turks. These features are a result of shared ancestry and certain environmental constraints that these groups experienced. Most of these cultures practiced a form of infant binding, called swaddling, that would often result in exaggerated flatness of the occipitus (back of the head). Thus, a combination of a physical trait, plus cultural ones, resulted in people with very rounded heads. In essence then, using racial classification to describe modern human beings would be like using a flamethrower to light a birthday candle.

What is [cultural diffusion](#)? Cultural diffusion is the spread of technology and culture via adaptation of a particular culture by the indigenous (local) inhabitants of a region.

Can you give an example of cultural diffusion? Though as already mentioned, the potato had its origins in South America; it has been readily incorporated into the cuisine of people around the world. We don't see the genetic legacy of Peruvian/Bolivian people in the places that potato consumption has spread. There is no trace of Peruvian ancestry in high potato consuming regions of Ireland, thus indicating cultural diffusion.

What is [demic diffusion](#)? Demic diffusion is the spread of people into a particular region, replacing and displacing the original inhabitants of the area.

Can you give an example of demic diffusion? The widespread explosion of farming across the Fertile Crescent and the Mediterranean Basin resulted in the introduction of genes directly by successful farmers. By compare aDNA found in Bronze Aged skeletons in southern and Eastern Europe with later dates, we find that there was a shift in population approximately 5000, 3000, and 2000 years ago. Though history can explain the potential shift 2000 years ago, as the Roman Empire's expansion lead to migration into newly acquired regions, the shifts before that, during the Iron Age and late Bronze Age are related to other factors, perhaps metal work, farming innovations, more efficient agricultural practices, and others that resulted in demic diffusion.

What is [Neanderthal DNA](#)? As aDNA techniques have improved, it's possible to compare the genome of Neanderthals (a cousin branch to modern humans that died out approximately 40,000 years ago) with that of modern Humans and determine whether there was cross-mixing or hybridization (ingression). Approximately 600,000 years ago, in Africa, a proto human population (called [Homo Heidelbergensis](#)) split, with one group migrating into Eurasia, and the other staying behind in Africa. Over time, the group that migrated to Eurasia evolved into Neanderthals, while the group that stayed in Africa evolved into modern humans. When some bands of modern humans started leaving Africa in waves (70,000 - 100,000 years ago), they likely encountered Neanderthal populations in the Near East. All modern human beings outside sub Saharan Africa retain the traces of having interbred with Neanderthals...in essence, they have Neanderthal genes. Approximately 1-4% of the genes of humans outside of sub Saharan Africa are from Neanderthals, with the highest concentration being in the Near East (as that was the likely source of interbreeding).

Are there other hybridization events besides Neanderthal? There was another species of humans, an offshoot of Neanderthals called [Denisovans](#) that have left their traces in the people of [Austronesia](#). This was recently found in the aDNA extracted from a single tooth and small finger bone found in a Siberian cave. In addition, there is recent evidence for another archaic human species present in southern Spain, as well as traces within our own genome of archaic populations in Africa. These are more difficult to determine, because we don't have aDNA available in large areas of Africa. Much of Africa is too warm and moist for the long-term preservation of DNA.

References and further reading:

<https://www.familytreedna.com/public/ArmeniaDNAProject>

<https://www.facebook.com/groups/armenianDNAproject/>

<http://dienekes.blogspot.com/>

<http://blogs.discovermagazine.com/gnXP/tag/population-genetics/#.VANbzzKSzzI>

<http://digitalcommons.wayne.edu/humbiol/vol84/iss4/2/> (Paternal Lineage Analysis Supports An Armenian Rather Than A Central Asian Genetic Origin Of The Hamshenis.)

<http://link.springer.com/article/10.1007%2Fs00439-001-0627-9?LI=true> (Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group.)

<http://admixturemap.paintmychromosomes.com/>

http://www.genebase.com/doc/mtdnaHaplogroup_H_Subclade_Distribution_Map.pdf

<http://www.genebase.com/learning/article/38>

Commercial projects that specialize in DNA analysis.

<https://www.familytreedna.com/>

<https://www.23andme.com/>

<https://genographic.nationalgeographic.com/>

--Armen Martirosian

For questions, concerns, and/or comments, you can contact me via email:

martiansolstice@hotmail.com

Acknowledgements:

Peter Ara Hrechdakian for his critique and suggestions. Admin for the Armenian DNA Project.

Levon Yepiskoposyan and his team at the Molecular Biology Institute in Yerevan for their work in population genetics and their acceptance of a curious diasporan.

Mihran Nazaretyan and Sevak Avagyan at the Armenian Bone Marrow Donor Registry for their insights into the future of health and medicine in Armenia.

The ABS academic trust who were helpful, supportive, and instrumental as an early sounding board. (Hovann Simonian, also an admin for the Armenian DNA project, as well as LA, PCV, and LM).

AM for her cultural anthropological and ethnographic perspective.