DNA Basics

Adapted from a MyHeritage Blog and the International Society of Genetic Genealogy (ISOGG) Wiki by Earl Cory

MyHeritage has started a series to explain DNA, how it works and answer the most common questions. In other words, DNA for non-scientist. Following is a summary of their first postings. For those that are interested in further reading, go to https://blog.myheritage.com/2017/11/new-blog-series-dna-basics-begins-today/

The International Society of Genetic Genealogy (ISOGG) was founded in 2005 by DNA project administrators who shared a common vision: the promotion and education of genetic genealogy. Their mission is to advocate for and educate about the use of genetics as a tool for genealogical research, and to promote a supportive network for genetic genealogists. https://isogg.org/wiki/Wiki_Welcome_Page. They have an extensive BEGINNERS GUIDE TO GENETIC GENEALOGY at https://sites.google.com/site/wheatonsurname/beginners-guide-to-genetic-genealogy

Note: *the best way to absorb these lessons is to read them one at a time and follow the suggested links-- not only will this help with "beginner's head spinning syndrome"-- it will allow you to absorb the material more fully. If you are an advanced beginner or intermediate you might find the terminology and background in Lesson 5 invaluable no matter which lessons are of interest. Also have a look at Lesson 11 which has a list of resources to help you including a new "Cheat Sheet File" category.*

I have merged information from both of these sources to provide an overview of Genetics, DNA, DNA testing and definitions of the common terms.

Genetics 101

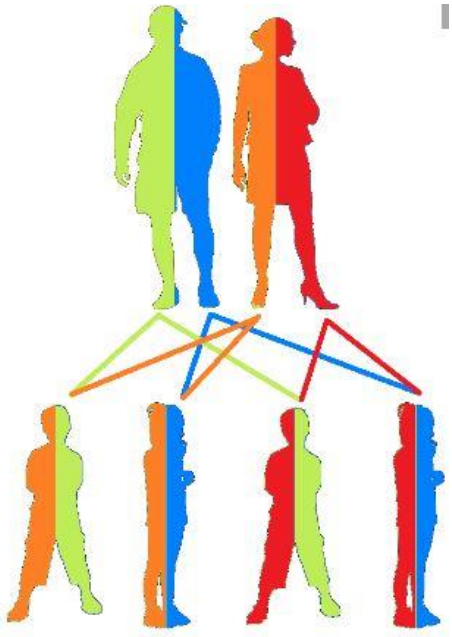
Our bodies are composed of cells. The headquarters of each cell is the nucleus. Inside each nucleus is an identical copy of the unique instruction manual for YOU, written in DNA. There are more than 30 trillion cells in our bodies and, in each nucleus, is a copy of your DNA.

DNA is the molecule that encodes the genetic instructions for building and operating all living things. Humans are 99.9% genetically identical. That is a very important concept to remember. All the differences we see in the way people look, what diseases they may be prone to etc. come from the .1% difference. That doesn't seem like much until you realize that there are about THREE BILLION base pairs in which those differences may be expressed.

No one else's DNA is exactly like yours. So, if you know how to read your DNA you can learn much about who you really are.

Inheriting DNA

We inherit 50% of our DNA from our mother and 50% from our father. They each received 50% of their DNA from their mother, and 50% from their father. So, when we look at your DNA, we're looking at about half of your mom's and half of your dad's, which is about 25% of each of your grandparents, or 12.5% of each of your great-grandparents. In other words, by looking at your DNA, we can see little pieces of all your ancestors who together gave you the exact combination of DNA you have today.



Adapted from The University of Arizona Health Sciences

Sharing DNA

It follows that two people, who have inherited DNA from the same source, will have some common DNA. For example, a daughter who received 50% of her mother's DNA and 50% of her father's DNA, will share some of the same DNA as her brother, who also inherited half of his DNA from each parent. The siblings may have inherited a different 50% of each parent's DNA, and so the siblings won't have identical DNA, but they will have about 50% in common.

That means you have 50% of your DNA in common with your mother, 50% in common with your father, and 50% in common with each of your siblings. Two anonymous DNA samples with 50% overlap might be a brother and a sister, or a child and a parent. To identify the relationship, we first look at the gender of the samples. Males will have a Y chromosome in their DNA sample and females won't. Next, we look at the age of the sample donor, if we have it. That part isn't written in the DNA itself. Knowing those three pieces of information — the amount of DNA in common between the two samples (in the current example, 50%), the donors' genders, and the donors' ages — is enough to give a good estimate of the relationship between the sample donors. Two people with 50% DNA in common, both female, 30 years apart in age, are most likely mother and daughter.

DNA and you

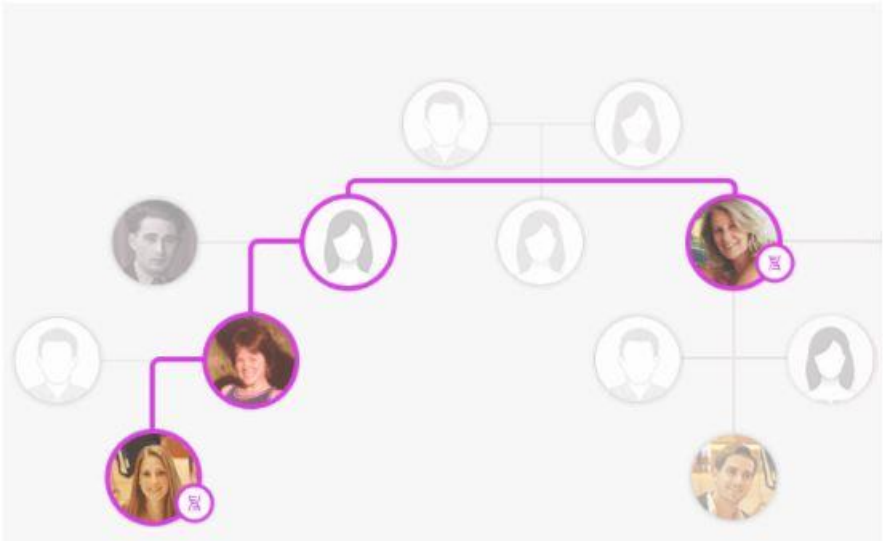
DNA carries the code for YOU — your eye color, hair color, height, even whether you hate the taste of coriander, is written in your DNA. That's why two people with common DNA have common features.

The more DNA in common, the more features you would expect to have in common. If you've always been told you look like your mother, or have your father's nose, or if people who have never met your siblings know you're related the instant they meet them, now you know why!

DNA and your family

While you have a large portion of DNA in common with close relatives, you have a smaller portion in common with more distant relatives. Two distant cousins who have never met but have a common ancestor will have a small piece of DNA inherited from that ancestor in common. That's the basis for using DNA tests to discover relatives you otherwise wouldn't know about.

How small a portion of common DNA depends on how many generations back the common ancestor was. Combined with databases of family trees where we have the types of clues discussed above (such as gender and age).



DNA and your family's past

Generations ago, people didn't move around as much as we do today. They didn't fly cross-country or across oceans. Most people married and died in the same region where they were born. Because people married other people from the same region and had children who then married in the same region, a correlation can be seen between DNA and geographic location. This was especially true for geographically isolated places like the island of Ireland.

Sometimes this insular DNA inheritance wasn't necessarily geographic — it may have been cultural. For example, Jews or Mennonite Christians married and had children within their own groups for generation after generation, creating a correlation between DNA and cultural identity.

When you have a DNA test, they compare your DNA to models of DNA from different ethnicities to see which ones you match. For example, MyHeritage Ethnicity Estimates include 42 ethnicities.

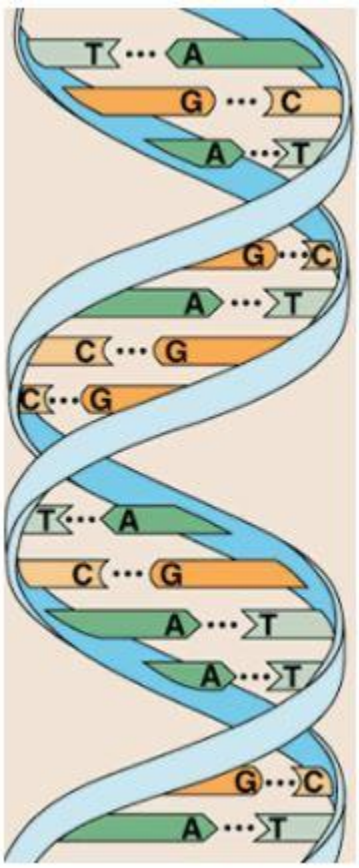
The **human genome** is the complete set of human genetic information. It is located within the 23 pairs of chromosomes. Each half of the pair represents our mother or our father. These 23 chromosomes reside within the nucleus of our cells. There is also a small DNA molecule found within individual parts called mitochondria.

The Structure of DNA

Let's take a closer look at how this looks under a microscope.

Breaking down DNA: From double helix to nucleotides

Our bodies are made up of more than 30 trillion cells. The headquarters of the cell is the nucleus. Not every single cell actually has a nucleus — red blood cells don't, for example. But almost all of our cells do have a nucleus and that's where DNA is stored.



DNA takes the shape of a double helix — think of two long (untied!) shoelaces, pressed together along their entire length, and then wrapped around your finger. Each of the two shoelaces is made up of a series of little blocks called nucleotides. There are four nucleotide types, abbreviated as A, T, G and C. The same four nucleotides appear over and over again in different orders to make up the entirety of your individual DNA sequence.

In our analogy of the "instruction manual for you", think of DNA as the text written in the manual that the body reads in order to make you, you. Nucleotides are the alphabet used to write that text. Nucleotides are the letters, every set of three consecutive letters is called a codon — that's like the words. Groups of codons make sentences and all together, the whole entire text — all of your DNA together — makes up the whole set of instructions.

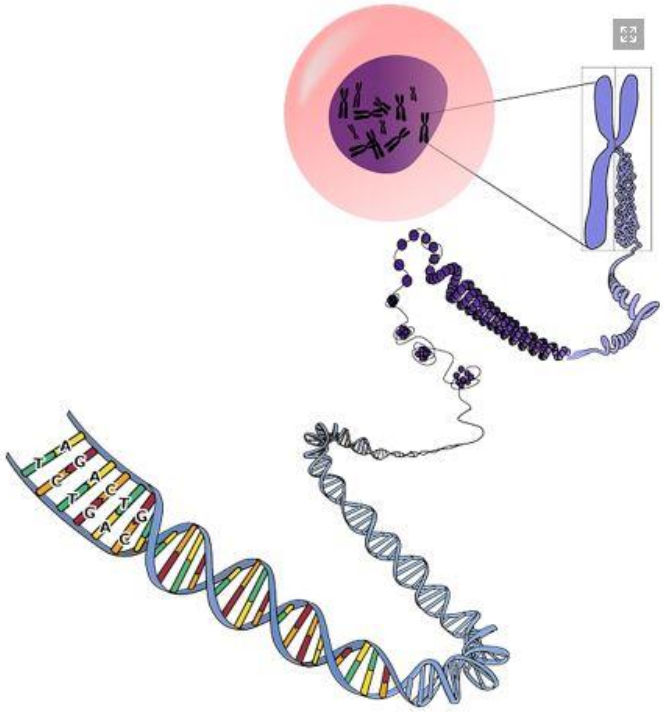
When you hold the two shoelaces next to each other, an A on one will always match up with a T on the other; a G on one will always match up with a C on the other. So really, by looking at one shoestring you already know what is written on the other — every place one has an A, the other has a T; every place one has a T, the other has an A. Similarly, every place one has a G, the other has a C; and, every place one has a C the other has a G.

About 99.9% of this very long sequence is identical in every person on earth. The 0.1% that varies from person to person carries the part of the instruction manual that makes us each unique — from the different colors of our skin to our height, and beyond.

Organizing DNA: From histones to chromosomes

The human genome contains approximately 3 billion pairs of nucleotides — that's two very long shoelaces! Because this is the instruction manual to you, it is extremely important that they are well-organized and don't get tangled up. There are a number of levels of organization involved, but here we'll get to know just a few.

Because the strands of DNA are very long, segments are wrapped around proteins called histones, similar to a spool of thread. These spools are then packaged into chromosomes. Think of it as an instruction manual whose text is so long that it is published in many volumes — each volume is a chromosome.



Autosomal DNA (atDNA)

Autosomal DNA is a term used in genetic genealogy to describe DNA which is inherited from the autosomal chromosomes. An autosome is any of the numbered chromosomes, as opposed to the sex chromosomes. Humans have 22 pairs of autosomes and one pair of sex chromosomes (the X chromosome and the Y chromosome). Autosomes are numbered roughly in relation to their sizes. That is, Chromosome 1 has approximately 2,800 genes, while chromosome 22 has approximately 750 genes. There is no established abbreviation for autosomal DNA: atDNA (more common) and auDNA are used.

X Chromosome

The X chromosome (X-DNA) is one of the two sex-determining chromosomes in humans (the other is the Y chromosome). The X chromosome was named for its unique properties by early researchers, which resulted in the naming of its counterpart Y chromosome, for the next letter in the alphabet, after it was discovered later.

The sex chromosomes X X are one of the 23 homologous pairs of chromosomes in a female. The X chromosome spans more than 153 million base pairs and represents about 5% of the total DNA in women's cells, 2.5% in men's.

Each person normally has one pair of sex chromosomes in each cell. Females have two X chromosomes, whereas males have one X and one Y chromosome. Both males and females retain one of their mother's X chromosomes, and females retain their second X chromosome from their father. Since the father retains his X chromosome from his mother, a human female has one X chromosome from her paternal grandmother (father's side), and one X chromosome from her mother.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. The X chromosome contains about 2000 genes compared to the Y chromosome containing 78

genes, out of the estimated 20,000 to 25,000 total genes in the human genome. Genetic disorders that are due to mutations in genes on the X chromosome are described as X linked.

Y Chromosome

The human **Y chromosome** is a male-specific sex chromosome. Nearly all humans who possess a Y chromosome will be morphologically male. Although the Y chromosome is situated in the cell nucleus, it only recombines with the X-chromosome at the ends of the Y chromosome; the vast majority of the Y chromosome (95%) does not recombine. When mutations (errors in the copying process) arise in the Y chromosome in the form of single-nucleotide polymorphisms) or short tandem repeats, they are passed down directly from father to son in a direct male line of descent. This line is known as the patriline.

Passing on DNA, one chromosome at a time

Humans have 23 pairs of chromosomes. Or, continuing our instruction manual analogy, we have two editions of the 23-volume instruction manual. For each volume in the series, or each chromosome, we get one of our mother's two editions and one of our father's two editions, which they in turn got from their parents. DNA tests can help you figure out which editions you got from your parents.

By the way, this is also why your Ethnicity Estimate may be different from your siblings'! For example, if your mother is 50% Japanese and 50% Irish, and your father is 50% Italian and 50% Scandinavian, you might end up mostly Japanese-Italian while your sister might be more Irish-Scandinavian. Of course, in reality, the differences are not usually that dramatic. But it is possible, even likely, that siblings will inherit at least slightly different percentages of different ethnicities in their DNA.

More likely, if your mom is 40% Irish and 60% Scandinavian, you might have gotten 30% Irish and 20% Scandinavian (adding up to 50% of your own ethnicities being from your mother), while your brother got 50% Scandinavian and 0% Irish (still giving him 50% of his own ethnicities being from your mother). This is how two biological siblings get different Ethnicity Estimates. Half of your total ethnicities did come from each parent, but you didn't necessarily get half of each of their ethnicities.

It also works this way with most inheritable traits, not just ethnicity. If one parent is very tall and the other is very short, you and your siblings probably inherited different combinations of your parents' heights, which is why you will all be different heights from one another, even though your height genes all came from the same original pool. You simply got different combinations from the available genetic options.

DNA Expression

Let's review starting with the smallest unit and working to the largest.

- We have **four nucleotides:** A, T, C and G the building blocks of our DNA
- The **nucleotides** form **base pairs** (AT) or (CG)
- **SNP**s are base pairs where a variation (mutation) has occurred
- Segments of base pairs that have a specific function are called **genes**
- Genes are parts of a very long string of DNA called a **chromosome**
- We all have **23 pairs of chromosomes**
- Which reside in the **nucleus** of our cells
- All of our **cells** together form our body
- The variations and the combination of those variations are what makes each of us unique.

To summarize briefly:

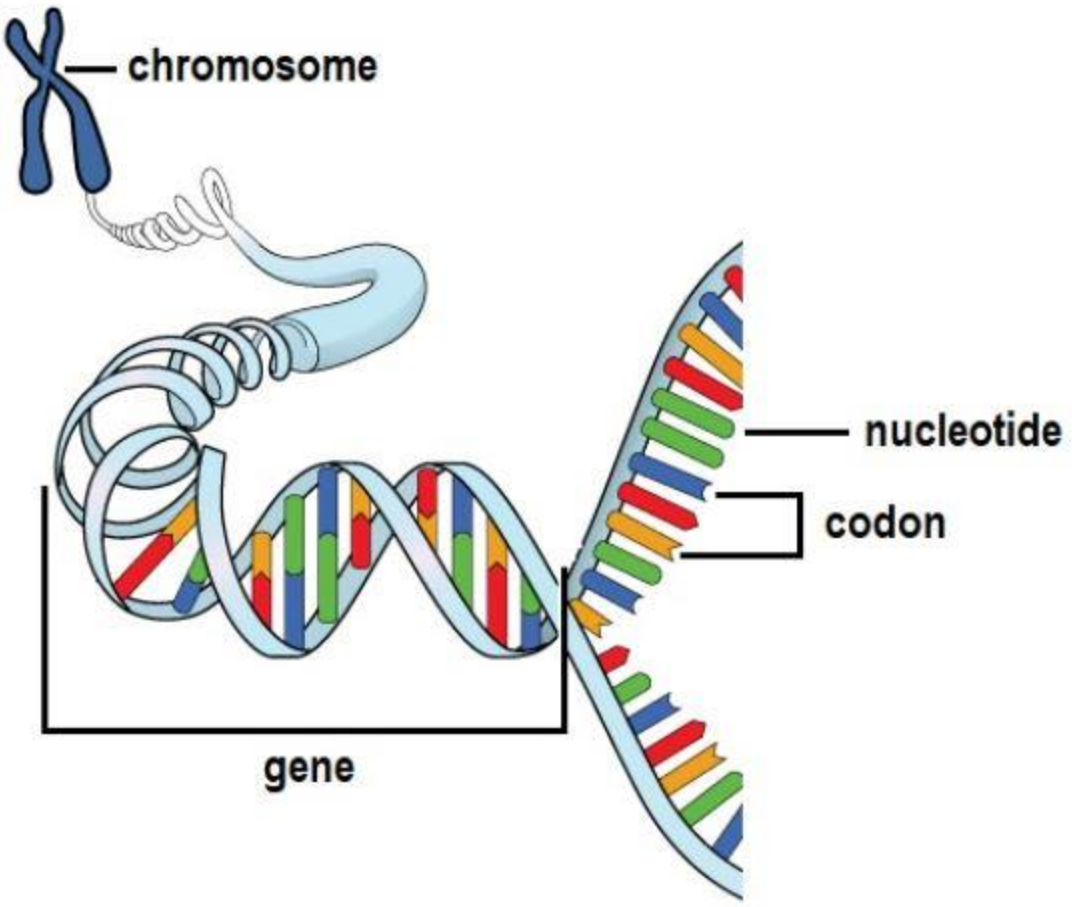
DNA is the instruction manual to each of us.

Nucleotides are the letters;

Sets of three consecutive nucleotides called *codons* are the words;

Groups of codons called *genes* are the sentences; and,

Chromosomes are complete volumes of the 23-volume manual — one edition of which is inherited from each parent.



RNA

Ribonucleic acid (RNA) is similar to DNA in many ways. RNA's main function is to "read" DNA. Copies of sections of DNA are made in RNA. These copies are used to assemble proteins.

A single instruction in DNA can be transcribed into RNA and then, different sections can be erased, or cut out; the remaining sections are then stitched together. This mechanism is called *alternative splicing.*

In order to function, our bodies need a lot of different kinds of proteins. If every single protein had its own instruction, we would have far more DNA than the 3 billion nucleotide pairs we already have! There's a limit to how much material can be contained and maintained. At the same time, if each piece of DNA could only produce a single protein, we wouldn't have nearly all the proteins we need to carry out essential functions. The fact that the middle man between DNA and proteins, RNA, can be edited, is an ingenious solution.

Proteins

Once the proteins are assembled, they go through some additional modifications, like being folded correctly. Then they are transported to wherever inside or outside of the cell they are meant to carry out the original instruction.

Variability across ethnicities

Interestingly, in some cases, protein levels, differ between ethnicities.

To give you an Ethnicity Estimate, your DNA is read a data file is produced with the information. They don't read every part of your DNA, which amounts to about 3 billion points. This is an expensive method called whole genome sequencing, currently reserved for specific clinical and research applications. Instead they focus on reading approximately 700,000 locations in your DNA that are known to vary between individuals, called **s**ingle-**n**ucleotide **p**olymorphisms (SNPs, pronounced "snips"). This method is called genotyping and it produces a data file that lists each SNP that they read, its position in your DNA, and the two alleles found there (i.e. the A, T, G, or C you inherited from each parent). By analyzing your genotypes for certain SNPs, they can estimate what percentage of your DNA is from each ethnicity.

The allele frequency of certain SNPs, e.g., how likely you are to have an A or a T in given positions, is associated with different ethnicities. Some SNPs are associated with how much of a certain gene tends to be expressed. To give just one example, there is a specific protein that is generally found in different amounts in British Caucasian people versus in Jamaican people due to different, ethnicity-associated genotypes.

The Dogma

The process of gene expression — from DNA to RNA to proteins — is called "The Dogma" by scientists.

Mitochondria DNA

Mitochondria are small organelles that lie in the cytoplasm of eukaryotic cells, such as those of humans. Their primary purpose is to provide energy to the cell. Mitochondria are thought to be the vestigial remains of symbiotic bacteria that were once free living. One indication that mitochondria were once free living is that they contain a relatively small circular segment of DNA, called **mitochondrial DNA** (mtDNA). The overwhelming majority of a human's DNA is contained in chromosomes in the nucleus of the cell, but mtDNA is an exception. Individuals inherit their cytoplasm and the organelles it contains exclusively from their mothers, as these are derived from the ovum (egg cell) only, not from the sperm. Males inherit mtDNA from their mothers but do not pass it on to their children. Consequently, when a mutation arises in the mtDNA molecule, the mutation is passed on in a direct female line of descent. These rare mutations are derived from copying mistakes – when the DNA is copied it is possible that a single mistake occurs in the DNA sequence.

Haplotype

The term for the set of numbers that consists of your Y-chromosome or mitochondrial DNA results. A haplotype is also known as a genetic signature or a DNA signature. It is a set of markers (polymorphisms) on a single chromosome that tend to be inherited together. In genetic genealogy the term is normally applied to the letters or numbers obtained from the results of a genealogical DNA test. Haplotypes can consist of varying numbers of markers depending on the test taken and therefore exist at different resolutions.

Haplogroup

A group of similar haplotypes that share a common ancestor with a SNP mutation. A haplogroup is a genetic population group of people who share a common ancestor on the patriline or the matriline. Haplogroups are assigned letters of the alphabet, and refinements consist of additional number and letter combinations. Example: R1b1a. Y-chromosome and mitochondrial DNA haplogroups have different haplogroup designations. Haplogroups pertain to your deep ancestral origins dating back thousands of years.

Y-chromosome DNA haplogroups

Y-chromosome DNA (Y-DNA) haplogroups are determined by single-nucleotide polymorphism (SNP) tests. SNPs are locations on the DNA where one nucleotide has "mutated" or "switched" to a different nucleotide.

Because a haplogroup consists of similar haplotypes, it is possible to predict a haplogroup from the haplotype. A SNP test is required to confirm the haplogroup prediction. Not all the testing companies offer SNP testing, and consequently their customers' haplogroup predictions are sometimes inaccurate.

Mitochdondrial DNA haplogroups

When you have taken a mitochondrial DNA (mtDNA) test most companies will give you a prediction of your mtDNA haplogroup. The haplogroup can however only be confirmed by testing specific branch-defining single-nucleotide polymorphisms (SNPs), some of which are only found in the coding region. If no SNP tests are done there is a possibility that the prediction will be incorrect. mtDNA tests from Family Tree DNA (HVR1 and HVR2) and the Genographic Project version 1 (HVR1) include a panel of 22 SNPs to confirm the base haplogroup, but rarely can determine the specific subclade. Similarly, 23andMe tests a set of SNPs that predict haplogroup assignments according to a now-outdated version of the mtDNA tree. Therefore, for those who wish to discover or confirm their exact haplogroup (subclade) assignment and to know exactly where they fit on the latest version of the mtDNA tree (applicable to all future versions of the tree), they will need the mtFull-Sequence test (available from FTDNA) either for themselves or a relative in their direct maternal line.

DNA Testing

A DNA test is a tool that genealogists use for answering the questions. There are three basic types kits being used. One involves spitting into a tube and the others are done by swabbing the inside of the cheek with a small brush. In the one the brush is left to dry in the other it is deposited in a tube. All are easy to use. The latter type has some advantages for those with little saliva production or those failing to get results with the other kind of test. They are also cheaper to mail overseas.

DNA tests for genealogical purposes do not involve needles, blood or urine. They are not the same tests used by law enforcement (CODIS) for crime identification purposes.

Combining the use of DNA testing and Genealogy is a relatively young method of family research. A common question is “What will DNA testing tell me?” It, depends on the type of test taken. For example, if you are a male or test a male relative's Y-DNA, if there are any matches in the database of the testing company, it will show the people you match. Depending upon the company you test with, other information may be made available to you, like your ethnic origin. Whether you have a Cohanim gene or have Native American ancestry. The first 12-markers of a Y-DNA test are considered "deep ancestry" markers and may show matches with people of different surnames. These matches generally indicate that you shared a common ancestor hundreds of years before surnames even came into use.

Free DNA Tests

Free DNA tests are sometimes available to encourage participation in surname projects. Offers are usually restricted to Y-DNA tests with sponsorship being provided by the relevant surname project. Some projects will underwrite the entire cost of a DNA test. Other projects will contribute towards the cost of a test or pay for tests for a limited number of markers. In order to qualify for the offer, it is usually necessary to supply a list of your paternal line ancestors for at least three or more generations. A list of Projects offering a free DNA test can be found at: https://isogg.org/wiki/Free_DNA_tests. The Cory project at Family Tree DNA is not currently on this list.

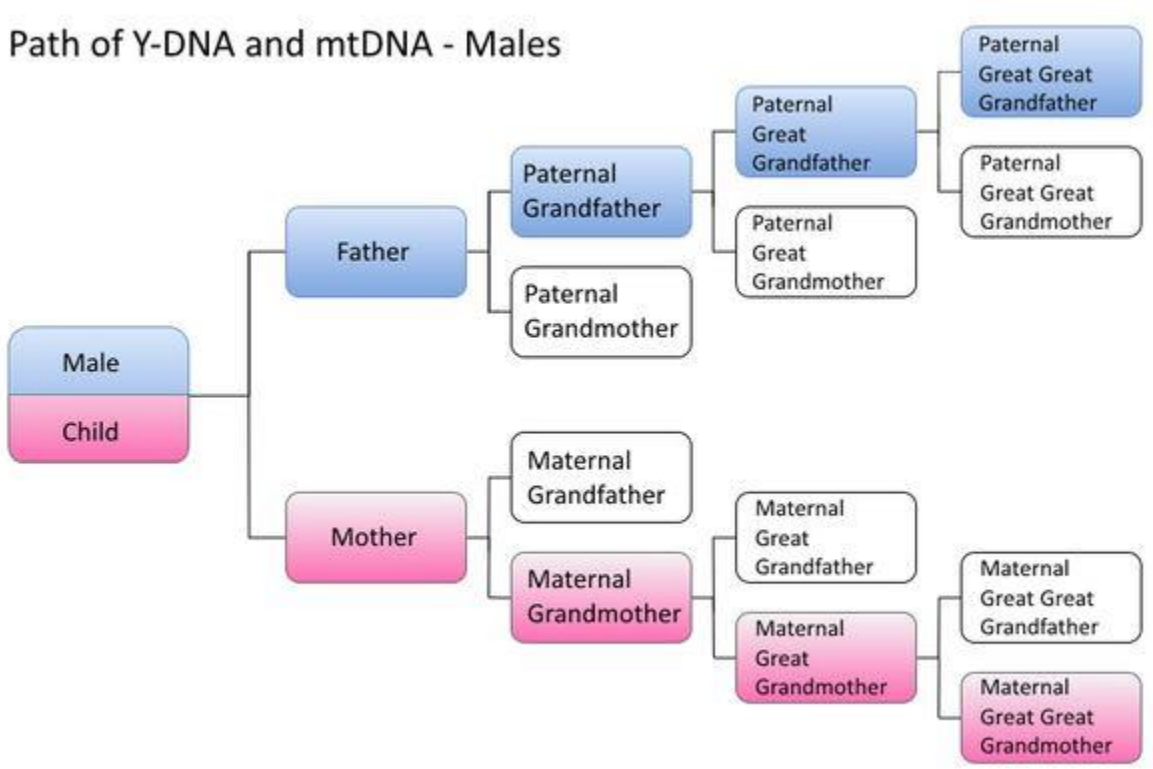
Mitochondrial DNA tests

A mitochondrial DNA test (mtDNA test) traces a person's matrilineal or mother-line ancestry using the DNA in his or her mitochondria. mtDNA is passed down by the mother unchanged, to all her children, both male and female. A mitochondrial DNA test, can therefore be taken by both men and women. If a

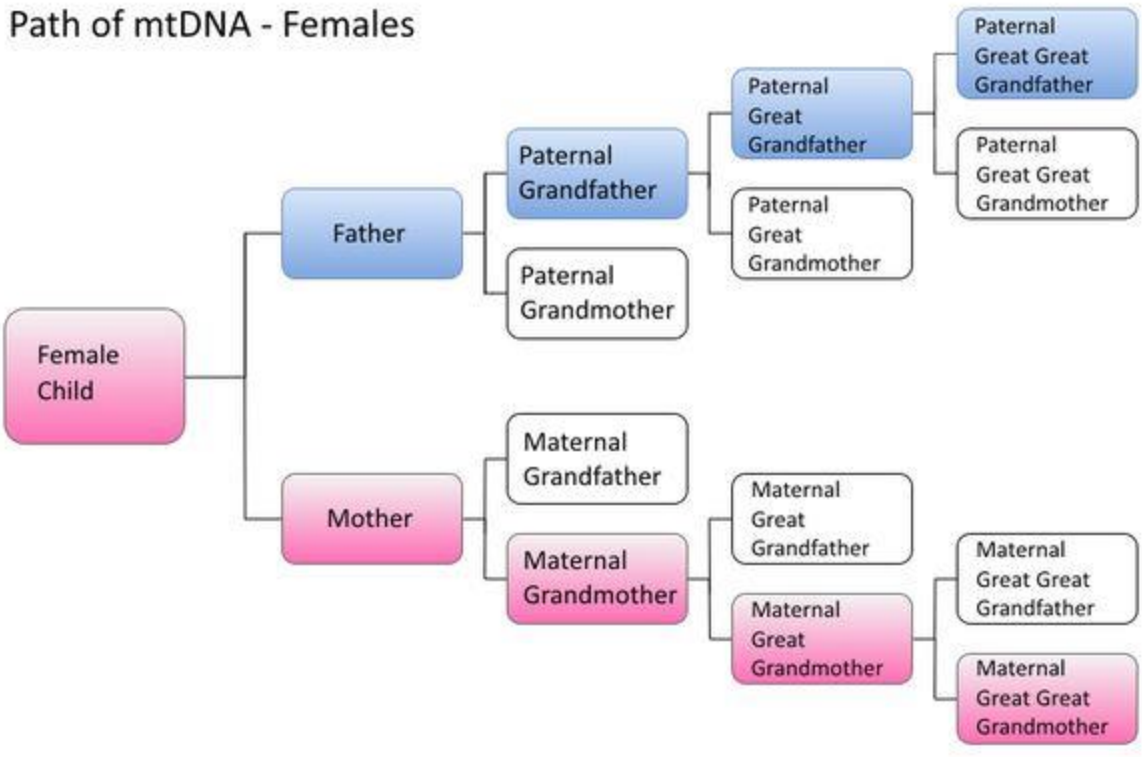
perfect match is found to another person's mtDNA test results, one may find a common ancestor in the other relative's (matrilineal) "information table".

Path of mtDNA transmission

Males inherit mtDNA (shown in pink) from their mother but do not pass it on to their children. Males inherit Y-DNA (shown in blue) from their father. They pass on Y-DNA to their sons but not their daughters.



Females inherit mtDNA (shown in pink) from their mother. They pass on mtDNA to both their male and female children. Females do not inherit Y-DNA (shown in blue) from their father.



See also the animation on mitochondrial DNA on the Learn Genetics website which provides a good explanation of the inheritance of mtDNA.

mtDNA by current conventions is divided into three regions. They are the coding region (00577-16023) and two hyper-variable regions (HVR1 [16024-16569], and HVR2 [00001-00576]). All test results are compared to the mtDNA of a European in haplogroup H2a2a. The two most common mtDNA tests are a sequence of HVR1 and a sequence of both HVR1 and HVR2. Some mtDNA tests may only analyze a partial range in these regions. With falling prices, the preferred option is now to test both HVR1 and HVR2 as a minimum.

Many people are now choosing to have the full mitochondrial genome sequenced. This is at times called the full mitochondrial sequence (FMS) test An FMS test will report results for all 16,569 bases in the mtGenome and will provide the most detailed subclade assignment. If two people have an exact FMS match they will generally share a common ancestor within the last 22 generations (about 550 years.) Family Tree DNA and GeneBase are the only two companies that currently offer the FMS test. Some companies and organizations, such as the Genographic Project, 23andMe and Living DNA, use microarray chips which provide a scan of selected SNPs from across the entire mtDNA genome. The SNPs are chosen for the phylogenetic relevance and can provide a fairly detailed haplogroup assignment. However, the results from these tests are not compatible with the genealogical matching databases.

A MtDNA testing comparison chart can be found at: https://isogg.org/wiki/MtDNA_testing_comparison_chart Family Tree DNA also processes tests for a number of affiliate companies: African DNA, iGENEA, DNA Ancestry & Family Origin and DNA Worldwide. Participants who join Family Tree DNA's database from these companies will have kit numbers that begin with 'A' from African DNA, 'E' from iGENEA, 'M' from DNA Ancestry & Family Origin, and 'U' from DNA Worldwide. National Geographic Genographic kit transfers begin with 'N'.

Family Tree DNA acquired DNA Heritage in April 2011. DNA Heritage customers were given the option of transferring their results to the FTDNA database. DNA Heritage kit numbers begin with the letter 'H'.

Y-DNA test

A Y chromosome DNA test (Y-DNA test) is a genealogical DNA test which is used to explore a man's patrilineal or direct father's-line ancestry. The Y chromosome, like the patrilineal surname, passes down virtually unchanged from father to son. Every now and then occasional mistakes in the copying process occur, and these mutations can be used to estimate the time frame in which the two individuals share a most recent common ancestor or MRCA. If their test results are a perfect or nearly perfect match, they are related within a genealogical time frame. Each person can then look at the other's father-line information, typically the names of each patrilineal ancestor and his spouse, together with the dates and places of their marriage and of both spouses' births and deaths. The two matched persons may find a common ancestor or MRCA, as well as whatever information the other already has about their joint patriline or father's line prior to the MRCA. Y-DNA tests are typically coordinated in a surname DNA project. And each receives the other's contact information if the other chose to allow this.

Women who wish to determine their direct paternal DNA ancestry can ask their father, brother, paternal uncle, paternal grandfather, or a cousin who shares the same surname lineage (the same Y-DNA) to take a test for them.

Y-DNA testing involves looking at Y-STR segments of DNA on the Y chromosome. The STR segments which are examined are referred to as genetic markers and occur in what is considered non-coding DNA or "junk" DNA.

STR markers

A chromosome contains sequences of repeating nucleotides known as short tandem repeats (STRs). The number of repetitions varies from one person to another and a particular number of repetitions is known as an allele of the marker. An STR on the Y chromosome is designated by a **DYS** number (**D**NA **Y**-chromosome **S**egment number).

SNP markers

SNP is a change to a single nucleotide in a DNA sequence. The relative mutation rate for an SNP is extremely low. This makes them ideal for marking the history of the human genetic tree. SNPs are named with a letter code and a number. The letter indicates the lab or research team that discovered the SNP. The number indicates the order in which it was discovered. For example, M173 is the 173rd SNP documented by the Human Population Genetics Laboratory at Stanford University, which uses the letter M.

Understanding test results

Y-DNA tests generally examine 10-111 STR markers on the Y chromosome, but hundreds of markers are available. STR test results provide the personal haplotype. SNP results indicate the haplogroup.

Autosomal DNA testing

Autosomal DNA testing for genetic genealogy purposes is provided by the following five companies: 23andMe, AncestryDNA, MyHeritage, Family Tree DNA and the Genographic Project. The genetic testing project for the Cory Family Society is coordinated by Thomas Corey at Family Tree DNA. A comparison chart of the various companies and what they offer can be found at: https://isogg.org/wiki/Autosomal_DNA_testing_comparison_chart

Company	23andMe	Family Tree DNA's Family Finder Test	Ancestry.com's Ancestry DNA Test	National Geographic Genographic Project Geno 2.O Next Generation Test	MyHeritage
GEDCOM file upload allowed	No	Yes	Link is created to Ancestry.com pedigree charts	No	Yes
Upload of raw data file allowed from other companies	No	Yes, but 23andMe Version 3 and Version 4 files (c. November 2010 to August 2017) and Ancestry.com files only	No	No	Yes, from 23andMe, FTDNA's Family Finder, Ancestry.com, and WeGene

The output from an atDNA test is often called the RAW DNA file. It is simply a list of SNPs and the reported value of those SNPs. Here is an example showing the first few reported SNPs on Chromosome 1:

rsid#	Chrm.	Position	Genotype
rs4477212	1	72017	AA
rs3094315	1	742429	AA
rs3131972	1	742584	GG

In this example the rs numbers (rsid) are the name of the SNP, the position where it is located on the chromosome and the **genotype** ("AA") is the reported value for that SNP. Occasionally you will have a "No Call" where they were unable to declare what values were there. Ancestry and FTDNA have eliminated many of the medically relevant SNPs so that no information is reported. This involves privacy for those who might have significant inherited medical conditions.

WARNING: Please do not try to print your RAW DNA file. It is over 20,000 pages long!

To give you an idea of what these mutations might mean let's look at the SNPS for lactose intolerance as reported at 23andme.

rs4988235= AA= likely tolerant of lactose (such as found in cow's milk)

r4988235= AG = likely lactose tolerant

rs4988235= GG = lactose intolerant

What this means is that individuals with "AA" or "AG" at SNP rs4988235 are likely to be tolerant of lactose. And those with either "GG" are likely to be lactose intolerant.

Ancestry does not allow uploading raw data from other companies, while Family Tree DNA and MyHeritage do. The most cost-effective way to do Autosomal DNA is to use Ancestry, then upload to Family Tree DNA and MyHeritage. If you are interested in other tests, then it is best to start with Family Tree DNA, since Family Tree DNA is the only company that offers Additional DNA testing options. It has multiple options for Y chromosome STR and SNP testing, complete or partial mitochondrial DNA testing, and complete genome testing

All but National Geographic allow Upload of raw data file allowed by GEDmatch into the GEDmatch database.

If you tested at:		AncestryDNA v1, v2	23andMe v1, v2	23andMe v3, v4	23andMe v5	Family Tree DNA	MyHeritage	WeGene	Living DNA
You can transfer to	AncestryDNA	—	NO	NO	NO	NO	NO	NO	NO
	23andMe	NO	—	NO	NO	NO	NO	NO	NO
	Family Tree DNA	*Yes but*	YES	YES	*Not yet*	—	YES	NO	NO
	MyHeritage	YES	YES	YES	*Not yet*	YES	—	NO	NO
	WeGene	YES	YES	YES	YES	NO	NO	—	NO
	Living DNA	YES	YES	YES	YES	YES	YES	NO	—
	DNA.Land	YES	YES	YES	YES	YES	YES	NO	*No?*
	GEDmatch	YES	YES	YES	NO	YES	YES	YES	NO
	GEDmatch Genesis	YES	YES	YES	YES	YES	YES	YES	YES

The V_1, V_2, etc. refer to the ISOGG Wiki Chip Version used in testing. Each version used a different number of SNPs and locations.

There are about 15 million known SNPs (aka Ancestry Informative Markers) for genetic ancestry but only 1-to-5 million are utilized by advanced genetic studies, and much less (~700,000) by vendors like DTC DNA testing companies (i.e. 23andMe).

According to ISOGG Wiki Chip Versions, here are the different microarray chip versions utilized by 23andMe for genotyping since the debut of its DTC personal genome service:

- v1: November 2007
- v2: September 2008, ~555K SNPs (Illumina)
- v3: November 2010, >900K SNPs (Illumina OmniExpress)
- v4: November 2013, ~570K SNPs (Illumina OmniExpress)
- v5: August 2017, ~640K SNPs (Illumina Global Screening Array)

Tests done using the version 3 and version 4 chips can be transferred free of charge to the Family Tree DNA Family Finder database. For details see the FTDNA page on the autosomal DNA transfer program.

Rebekah Canada has produced heat maps showing the distribution of SNPs on the 23andMe v2, v3 and v4 chips:

- Exploring microarray chips (http://haplogroup.org/exploring-microarray-chips/)
- 23andMe v2 chip
- 23andMe v3 chip
- 23andMe v4 chip

For a list of the Y-SNPs on the 23andMe chips (v2, v3, v4 and v5) see the spreadsheet provided by Tim Janzen which can be downloaded from this link..

SNP Density

The number of SNPs per million base pairs DNA tell us the density of SNP coverage. Density matters, because low density areas are a potential risk for false positives. This is especially true as we move back past 1st and 2nd cousin matches, and the DNA shared with them.

Thus, the SNP density is something we should consider when we look at a DNA segment shared with one or more matches.

Chip Comparison

The same thing is true for SNPs in common between any two chips. For those who use GEDMatch and other matching sites, this is important. You are matched on SNPs in common with both chips.

Recombination Maps

The autosomal chromosomes and the X-Chromosome recombine each generation. Each person gets half of each parent's autosomal DNA. Men get a mixed single copy their mother's X-Chromosomes. Women get an exact copy of their father's X-Chromosome and a mixed single copy of their mother's X-Chromosomes.

Recombination is random, but some places across the autosomal and X-Chromosomes that are more often recombinant points than others. This is why DNA segment length is measured in centiMorgans based on the recombination rates and not direct base pair counts.

Recombination rates can also be population specific. The 1000 Genomes Project has centiMorgan mapping data for twenty of the populations they have sequenced.

Allele Frequency Mapping

The SNP markers on the microarray chips that we use almost always have exactly two possible values. The more common one is the major allele and the less common one is the minor allele. It is more likely that you will randomly match. It is important for evaluating the quality of a shared DNA segment, because there is a greater chance of randomly matching someone –a false positive– in places where you have mostly or entirely the major allele.

Which value is major, and which one is minor can vary by population.

SNP No-Call rates

What is a No-Call? It is a place where the microarray chip fails to return a result for a SNP. Sometimes this is completely random. Sometimes a single SNP marker will have high No-Call rate. Mapping the frequencies of No-Calls may expose interesting patterns.

Comparison of 23 and Me and Family Finder

A comparison of 23 and Me and Family Finder can be found at: https://isogg.org/wiki/Family_Finder_versus_DNA_Relatives. Some of the most important points follow:

Databases

In October 2009, prior to the launch of the 23andMe Relative Finder service, the company announced that it had '30,000 "active" genomes in its database. By July 2010 the database had grown to 50,000 customers. As of June 2011, 23andMe had over 100,000 customers in their database. The company does not provide a breakdown of customers by country. Anecdotal evidence suggests that the vast majority of their customers, and perhaps in excess of 90%, are in the United States.

Family Tree DNA's Family Finder test was launched in February 2010. Although Family Tree DNA publicize figures for the size of their Y-DNA and mtDNA databases they have not publicly disclosed the size of their Family Finder database. However, a subscriber to the Rootsweb Autosomal DNA mailing list reported that Bennett Greenspan, the CEO of FTDNA, had announced at a Jewish genealogy meeting in September 2011 that the Family Finder database had passed the 20,000 milestone. Family Tree DNA has a much more international database than 23andMe, largely thanks to their involvement with the Genographic Project. The company does not provide a breakdown of customers by country. It is estimated that over 30% of their customers live outside the US, with the United Kingdom being the second largest population in the database after the US.

Match menu

Because DNA Relatives incorporates both Ancestral and Health tests, their match menu does not give the names of your matches. You are given the Predicted Relationship, Relationship Range, whether the match is male or female, percentage of DNA shared and number of shared segments. From that menu you have a Make Contact button to send a message to your match to see if they wish to correspond further. All contact goes through an internal mailing system within 23andMe.

Family Finder is an ancestral test only, so when you access the match menu you are able to see the names of your matches, as well as the Suggested Relationship, Relationship Range, Shared cM, Longest Block and Ancestral Surnames. Under your matches name you have a notes option where you can input notes regarding this match, as well as an email option where you can contact your match. This contact goes through your computer email program apart from FTDNA.

Match responsiveness

The responsiveness of matches at 23andMe is not very high. Many who order the tests have ordered for the health side and are not interested in genealogy. Though 23andMe has a larger database at this time, many have not researched their family and do not respond when you contact them. Even if they respond, they usually have little information regarding their family trees. The response rate is higher for Family Finder as most of the initial customers were Group Administrators running genetic genealogy projects. The Family Finder test is for ancestral research only and attracts a different audience than DNA Relatives. Because of this, it is expected that Family Finder will have a better response rate in general and the contacts will be more fruitful.

Customer service

Family Tree DNA is known for its excellent customer service and response. Questions are answered in a timely manner either by email or phone call. Currently, 23andMe is not very responsive to its customers. It is not unusual to wait several days for a response to a question either by email or by phone call.

Ancestry

AncestryDNA is the genetic genealogy database service of myfamily.com (the owner of Ancestry.com). AncestryDNA offers an autosomal DNA test. The test was first launched in the US in 2012. It became available in the United Kingdom, Ireland, Australia, New Zealand and Canada in 2015. It was launched in a further 29 countries in February 2016.

Note: Ancestry no longer sells Y-DNA and mtDNA tests.

The Ancestry autosomal test has a unique feature whereby possible common ancestors are identified by Shared Ancestor Hints as can be seen from the screenshot below. Results can also be filtered for those who have a "hint". A number of ISOGG members in the US have reported receiving a large number of meaningful

matches thanks to this feature. See also the blog post from Ancestry.com AncestryDNA discoveries made easier with the help of the shaky leaf.

Transfers

AncestryDNA results can be uploaded free-of-charge to the third-party website GedMatch where many additional tools are available for advanced analysis. Ancestry does not accept transfers.

AncestryDNA results can also be transferred to Family Tree DNA's Family Finder database where you can search for additional matches, join projects and use additional tools to compare results. The transfer is free, but a small fee is required to unlock additional features. Details can be found at www.familytreedna.com/AutosomalTransfer.

The most cost-effective way to do Autosomal DNA is to use Ancestry, then upload to Family Tree DNA and MyHeritage. If you are interested in other tests, then it is best to start with Family Tree DNA, since Family Tree DNA is the only company that offers Additional DNA testing options. It has multiple options for Y chromosome STR and SNP testing, complete or partial mitochondrial DNA testing, and complete genome testing

Extracting Y-DNA data

The chip used by AncestryDNA includes a small selection of Y-chromosome SNPs. Ancestry does not provide a Y-DNA haplogroup report, but it is possible to get a haplogroup assignment from the raw data. See the blog post by Steven Frank Updated method to get YDNA haplogroup from AncestryDNA results for details of how to do this.

Comparing Ancestry and FTDNA results

The way in which testing companies report the values for the various markers is not yet consistent though eventually all companies will conform to the NIST (US National Institute of Standards and Technology) standards. Ancestry's markers now conform to the NIST standards but Family Tree DNA, with their much larger database, have not yet completed the upgrade.

AncestryDNA results can be uploaded free of charge to Y-search, a public Y-DNA database sponsored by Family Tree DNA. Results can also be compared with those on other public DNA databases such as the Sorenson Molecular Genealogy Foundation's database and the Y Chromosome Haplotype Reference Database.

In August 2011 Family Tree DNA announced that it would accept third-party transfers from people who had taken a Y-DNA test with companies that used the Sorenson Genomics Laboratories. Ancestry customers are therefore now eligible to transfer their results to the larger FTDNA database and join the various surname, haplogroup and geographical projects. It is however necessary to pay a small fee for the transfer to preserve the integrity of the database. Y-DNA results from third-party Sorenson transfers are prefixed in the FTDNA database with the letter B.

Contacting AncestryDNA

Ancestry.com has toll-free/freephone telephone contact numbers in six countries: United States, UK, Australia, New Zealand, Canada and Sweden. They provide a telephone service for DNA enquiries in five countries: United States, UK, Australia, New Zealand, Canada. The contact numbers are the same as the Ancestry.com numbers but outside North America the operating hours are much shorter. ISOGG members in the UK have reported that if they ring the UK Ancestry telephone number with a DNA query they are transferred free of charge to the US site if the UK representative is not able to answer the question. The call must be made within the appropriate hours.

Autosomal Transfers

If you or a family member have previously tested your autosomal DNA at 23andMe©, AncestryDNA™, or MyHeritage™, you can transfer your results to Family Tree DNA by uploading your raw data file. After transferring your file, your autosomal data is uploaded to our database, one of the world's largest genetic genealogy databases.

When you transfer, for free, you will receive a list of your autosomal matches from our database and have access to our Family Finder – Matrix. The Matrix feature allows you to select and compare the autosomal DNA relationship between up to ten of your matches at one time.

Note: Family Finder is the name of their autosomal DNA test.

After transferring, you can unlock all Family Finder features, which include the Chromosome Browser, myOrigins, and ancientOrigins for only $19.

Please note that you can only transfer the following file versions from 23andMe© and AncestryDNA™ :

- 23andMe© V3
- 23andMe© V4
- AncestryDNA™ V1
- AncestryDNA™ V2

Unfortunately, at this time, you cannot transfer 23andMe© V1 or 23andMe© V2 results.

The process for transferring can be found at: https://www.familytreedna.com/learn/imports/transfer-autosomal-ancestry/family-tree-dna-family-finder-transfer-program/