Before You Test: DNA Basics You Need to Know

Drew Smith, MLS
drew@ahaseminars.com

What is DNA?

DNA is a complex chemical molecule found in essentially all living organisms that determines the biological features of the organism and that is imperfectly copied from an individual to its offspring.

What are genomes and chromosomes?

For any individual, independent structures containing DNA are known as chromosomes (these structures may be linear or circular), and one complete set of these chromosomes is known as the individual's genome.

How is human DNA useful to genealogists?

DNA from an individual can be scientifically tested to see how it compares with the DNA of other individuals who have already been tested. The more similar the DNA, the more closely related the individuals are. Test results can be used to provide evidence that two individuals are closely related or not closely related. When combined with other genealogical research, DNA evidence can be used to prove or disprove claims of relatedness.

Where in the human body is DNA found?

DNA is found in every human cell with the exception of red blood cells. (This is also true of other mammals.)

Types of human DNA:

- Nuclear DNA, found in the cell's nucleus and linear in structure.
- Mitochondrial DNA (mtDNA), found in mitochondria outside the cell's nucleus and circular in structure

Nuclear DNA:

- The typical human being has 23 pairs of nuclear chromosomes (for a total of 46 individual chromosomes).
- The 23 pairs consist of 22 pairs of autosomal chromosomes and 1 pair of allosomal (sex-determining) chromosomes. In each pair, one chromosome came from the father and one from the mother.

Sex-determining chromosomes:

- Two types: X and Y
- Men have 1 of each type
- Women have 2 X chromosomes
- Important to remember: Y-DNA is found only in men.

Y-DNA and surnames:

- In cultures where males inherit their surname from their father, the Y-DNA and the surname tend to go together.
- There are existing groups of Y-DNA test subjects organized by surnames. These are referred to as surname projects.
- Men with the same surname will not necessarily be related to all other men with that surname, because their ancestors may have independently adopted the same surname. As a result, typical surname projects may have subgroups of individuals who are related to each other but not to the individuals in the other subgroups.
- Documented male descendants of a common male ancestor may share a surname but not share the same Y-DNA. This can result from a number of causes, such as adoptions (both formal or informal) or non-paternal events.
- Two men with different surnames may share the same Y-DNA. This can result from a number of causes, such as adoptions, non-paternal events, or name changes.

Process for Y-DNA (surname) testing:

1. Complete sufficient traditional research on surname line.
2. Search for a surname project (Family Tree DNA, Cyndi's List, Google, etc.). If none, consider starting a project.
3. Have appropriate individual tested (living male relative from earliest generation in surname line). Move this up in priority if there is only one such individual.
4. Compare results with others who have been tested.

What is a mutation?

Because the copying of DNA from generation to generation is not a perfect process, some changes in the DNA will occur. These changes are known as mutations. Mutations that affect the DNA responsible for vital biological systems can result in deadly or harmful genetic diseases. But because the human genome includes at least some DNA that does not appear to affect any biological function, this non-coding (or "junk") DNA contains harmless mutations.

What is a marker?

Scientists have identified specific locations on the non-coding part of the Y-chromosome where the sequence of DNA at each location is prone to mutation. These DNA sequences are known as markers. For a typical Y-DNA test, as few as 12 or as many as 111 different markers may be tested, depending upon the test that was ordered.

What is an STR?

At each marker location on the Y-chromosome, the DNA repeats itself some number of times (usually a number between 6 and 38, depending upon the specific marker). This is known as a short tandem repeat (STR). Individual markers are uniquely named, usually with the letters DYS (DNA Y-chromosome Segment) followed immediately by a unique number.

Mitochondrial DNA testing:

- mtDNA is passed from a mother to her children (so both men and women have it and can be tested for it).
- Not normally as useful as Y-DNA testing because there are no surname groups to serve as a comparison group.
- Primarily used in unique situations where a common maternal lineage needs to be supported or refuted.
- Tests can be for the non-coding Hyper Variable Regions (HVR1 and HVR2) and for the Coding Region (but the Coding Region may identify medical issues).

SNPs and haplogroups:

- Y-DNA and mtDNA are subject to SNPs (pronounced "snips"), which are single-nucleotide polymorphisms (mutations in a single letter of the genetic code).
- As new subgroups are created with each mutation, these subgroups are known as haplogroups. Y-DNA haplogroups are identified by a letter (A through T) followed by an alternating sequence of numbers and letters. For example, the most common Western European Y-DNA haplogroup is R1b1a2. (Current practice is to identify the Y-DNA haplogroup by the name of the mutation that caused it.) mtDNA haplogroups also use a letter and numbering system, but there is no relationship between the letters used by the two different sets of haplogroups.
- Individuals from around the world have been tested in order to create haplogroup maps for both Y-DNA and mtDNA. The map is used to estimate approximately when and where each mutation took place, providing a geographical origin for that haplogroup.

Autosomal DNA testing:

- The 22 pairs of nuclear chromosomes that are not the sex-determining (X and Y) chromosomes are the autosomes.
- Each person inherits a unique mix of autosomal DNA from their parents. 50% is inherited from each parent. But the 50% provided by each parent is a random mix of the grandparents' DNA. This means (only on average) 25% from each grandparent, 12.5% from each great-grandparent, and so forth.
- Siblings will typically have the most DNA in common. First cousins will have less in common, second cousins will have even less, and so forth.
- Autosomal DNA can be tested to look for relatedness on all ancestral lines but is only reliable back to about 5th cousins.

Ethnicity testing:

- Y-DNA or mtDNA haplogroup information provides information about only one ancestral line.
- Autosomal testing can provide information about the "admixture" of different ethnic groups from all of your ancestral lines.

For more information:

Bettinger, Blaine T. *The Family Tree Guide to DNA Testing and Genetic Genealogy*. Family Tree Books, 2016.

© Andrew M. Smith. All rights reserved in all media. Publication, copying, electronic transmission, or reproduction of any part of this handout material is not permitted without advance written permission.