Introduction to Genetic Genealogy

Barbara Rae-Venter, J.D., Ph.D.
GenealogyConsult@gmail.com
www.GenealogyConsult.com
831-718-7994
"Put a scientist to work on your genealogy"

Autosomal DNA testing first became available about six years ago. This test can be used to identify genetic relatives in any of your ancestral lines. The **autosomes are chromosomes 1-22**. Generally the **Most Recent Common Ancestor (MRCA)** with a genetic cousin (a match) identified with this test is related to you within about seven generations or within about the last 210 years.

Depending on your ancestral and ethnic background, you may have anywhere from a few hundred to several thousand autosomal DNA matches. In this presentation I will discuss some of the tools that are available for analyzing your match data from autosomal DNA testing to help you achieve your research goal, whether it is learning more about a particular ancestral line, breaking through a brick wall, finding long lost relatives, determining the truth (or fallacy) of a family tradition, or identifying one or both of your birth parents.

The technique that I will be discussing is called **triangulation** and is based on the fact that **if three or more people match on a particular segment of DNA and they all match each other on that segment of DNA then they have each inherited that shared segment of DNA from a common ancestor**.

The testing companies each make available some tools for analyzing your match data, but these tools generally are limited to comparing results of only a few matches at a time. The tools I am going to discuss allow you to perform large scale comparisons of your match results, and I will also discuss ways to combine test company results for a comprehensive analysis of your match results if you have tested at more than one company..

Each of the three main DNA companies in the US report match results in different formats but all three testing companies provide a means to download your raw DNA data from the testing company's web site and upload it to one or more third party website. Regardless of which DNA testing company you used, it is suggested that you upload your raw DNA data to **www.gedmatch.com**, but especially if you tested at AncestryDNA.

At gedmatch.com you can look for matches with people who have tested at other companies and who have also uploaded their raw data to gedmatch.com, but if you tested at AncestryDNA there is the additional advantage that you will now be able to look at the chromosome profile of your matches. This information is essential for triangulation. Additional utilities include one to many match searches within the gedmatch.com database on both autosomal DNA and X chromosome DNA, one to one comparisons, and several others.

Registration on gedmatch.com is free, although donations are requested, but to use the two Tier 1 utilities **"Segment matching"** and **"Triangulation"** on gedmatch.com, you will need to make

a minimum $10/month donation to gedmatch.com. You will need these two Tier 1 utilities for uploading your gedmatch.com match data to **DNAgedcom.com**.

After uploading your raw DNA data to gedmatch.com, you will be given a kit number. The letter prefix indicates the company at which you tested: A for Ancestry, F for FamilyTreeDNA, and M for 23andMe. You need to upload your raw DNA data file only once; the data does not change with time.

If you have tested at more than one company, you can upload your raw data from each company at which you tested but if you choose to upload from only one company, the preferred data to upload is that from the company you have tested with that uses the most SNPs on their chip. This would be 23andMe if you tested with their V3 chip. If you tested with their V2 or V4 chip, a better choice for the raw data to upload would be that from FamilyTreeDNA. FTDNA and ancestry both use an Illumina OmniExpress chip but the FTDNA chip has more SNPs; 23andMe uses a customized Illumina chip (v4 test).

The number of SNPs on the chips used by each testing company is as follows:

Company/chip	Number of SNPs
23andMe V4 chip	577,382
FamilyTreeDNA	708,092
AncestryDNA	682,549

See: http://www.isogg.org/wiki/Autosomal_DNA_testing_comparison_chart.

A range of different chips have been used for the 23andMe test since the introduction of the service. Below are the dates of introduction and number of SNPs for each of 23andMe's chips:

v1: November 2007
v2: September 2008, 561,846 atDNA SNPs
v3: November 2010, 930,381 atDNA SNPs
v4: November 2013, 577,382 atDNA SNPs

Tests done using the 23andMe V3 chip can be transferred to the Family Tree DNA (FTDNA) FamilyFinder (FF) database for $39 or free if you also have four friends transfer their data. The lower number of autosomal SNPS tested by the 23andMe V4 chip as compared to the V3 chip or FTDNA's own chip is the reason why FTDNA currently will not accept a transfer of raw data generated with 23andMe's V4 chip.

You can also transfer raw data to FTDNA from AncestryDNA. If you have tested at FTDNA, you do not upload another raw data file to FTDNA.

The second site I will discuss is **DNAgedcom.com.** Registration on this site is free, although donations are requested. There are several utilities on this site that I will discuss. The most important utility for doing triangulation of your matches is GWorks which is a tree comparison utility. Prior to using GWorks, however, you need to gather all your match data from FTDNA, AncestryDNA and gedmath.com together.

FTDNA match data can be downloaded directly from FTDNA using an API interface on the DNAgedcom.com website. Check the box below the password box **"Get myFamilyTree List"** which loads myFamilyTree information to GWorks as a series of hyperlinks which are then accessed through GWorks.

FTDNA data can also be analyzed using one of three spreadsheet formats: **JWorks** for PC or Mac is an Excel. based spreadsheet. JWorks also can be used for preparing a combination spreadsheet of FTDNA and gedmatch.com match data by first combining match data and in common with (ICW) files from the two sources and then running JWorks on the combination files.

A second spreadsheet format provided is **KWorks** a web-based spreadsheet developed by Kitty Cooper.

Both JWorks and KWorks sort the match data by chromosome number and matching segment start position and provide a grid of matches with overlapping segments of DNA showing which of these matches match with each other.

A third spreadsheet option is the **Autosomal Dna Segment Analyzer (ADSA)** spreadsheet which also sorts the match data by chromosome number and matching segment start position. Solid blocks of color on the ADSA spreadsheet indicate which matches with overlapping segments of DNA also match with each other. This spreadsheet can be prepared using either downloaded FTDNA match data or gedmatch.com data uploaded to the DNAgedcom.com site.

AncestryDNA match data is obtained in a two step process. The first step uses a Google Chrome extension, **AncestryDNA Helper**, to scan and download both match data and data on ancestors of matches. This Google Chrome extension was developed by Jeff Snavely and is a free download for both PC and Mac computers.

The AncestryDNA Helper extension adds buttons to your AncestryDNA pages which can be used to perform a variety of tasks, such as additional search options, comparison of two test kits, and more. Best of all the extension enables you to create two files that can be used with DNAgedcom.com's GWorks utility: a CSV file containing all of your AncestryDNA matches and a CSV file containing a list of all of the people in their trees. A button also appears on each individual match's profile page to download a list of people just in that match's tree. Instructions for downloading and using AncestryHelper are provided by Barbara Taylor at **http://www.itstime.com/AncestryDNAHelper.htm**.

Once you have downloaded the two files from the Ancestry DNA site, they can be uploaded to GWorks on the **DNAgedcom.com** where they are stored as hyperlinks to match trees on AncestryDNA.

23andMe match data also can be downloaded directly from 23andMe however this data currently cannot yet be analyzed directly using the tools on **DNAgedcom.com**. There is however an iPad app that allows you to collect data from FTDNA, 23andMe, and **gedmatch.com** all in one place: DNAMatch4iPad by James Collins.

Two Google Chrome extensions are available for use on the 23andMe website: 23++ and 509andYou. These can be downloaded the same way as the AncestryDNA Helper extension.

When doing DNA testing it can be helpful to have relatives from particular family lines of interest also tested as this helps you identify how your matches are related to you and their tree can be used as a part of the triangulation process. Generally you want to choose the oldest relatives available in terms of generations as they are closer to the MRCA. Also, in choosing additional relatives to test, you want relatives that likely will match with you, so you need to keep in mind the amount of DNA that you may share with your various relatives.

As an example, you share about 50% of your DNA with your parents but the amount of autosomal DNA you share with siblings can vary widely, so it is a good idea to test as many siblings as possible. For first cousins the amount of shared autosomal DNA is about 12.5%, second cousins about 3.125%. For the amount of autosomal DNA that you can expect to share with various relatives, see http://www.isogg.org/wiki/Autosomal_DNA_statistics.

X chromosome

The X chromosome can be useful in some situations in identifying how you are related to a match but keep in mind that FTDNA and 23andMe report differently on X chromosome matches.

FTDNA reports X chromosome matches only if you also have an autosomal DNA match with someone who is an X match. FTDNA also only reports an X match if it is at least 1 cM AND has more than 500 SNPS. One of the cousins that I had tested at FTDNA is a female maternal second cousin; our MRCA is our great grandmother and our grandmothers were sisters. So I was surprised that we were not an X match, although she is an X match with both of my brothers. So I ran an X comparison on gedmatch.com and discovered that although we have matching X DNA of over 1 cM we have only 490 SNPs in that matching segment

23andMe reports an X match whether or not you also have an autosomal match with the person. I have several that are quite large X matches (over 20 cM) but no autosomal matching. The MRCA for such matches is likely quite far back in generations.

There are some situations where the X chromosome can be helpful. I have a female paternal first cousin who has tested at FTDNA. Our fathers were brothers and each inherited a random mix of their mother's X chromosomes. My cousin and I share about about 130 cM of X DNA which originated with my paternal grandmother. Since my cousin is a first cousin, she shares all of my paternal ancestry. But if someone has an X match with both me and my cousin I know that the person is matching with me through one of my paternal grandmother's lines as opposed to one of my paternal grandfather's lines. Likewise, if someone is an X match with one of my brothers, I know that the match is related through a maternal line.

If a male has an X chromosome match with you, he must be matching you through one of his maternal lines since he has only one X which came from his mother. If two women share an entire X chromosome then they have a father in common.

Happy hunting!