

Using the tools at GEDmatch to find Relatives

Presentation by Kitty Cooper, blogging at <http://blog.kittycooper.com>
Slides can be seen at <http://slides.com/kittycooper/gedmatch-10-13#/>
All my presentation slides are at <http://slides.com/kittycooper/>

Introduction

GEDmatch is a free third party site to which you can upload your DNA test results from any of the three main companies and compare them to other results in the GEDmatch database. There are currently 80,000 kits uploaded to GEDmatch. You can also use a number of tools there.

Today's presentation will focus on finding new relatives using these functions:

- **Creating a user and uploading your kits and gedcoms**
- **Analyzing your DNA:** Parents related? Problems in your data?
- **Comparing your DNA results: one to many, one to one, triangulation**
- **Comparing and viewing GEDcoms**

Do you know the definition of cM or SNP or base pair?

Use the ISOGG wiki to look up terms like these

<http://www.isogg.org/>

See this ISOGG wiki article for the relatedness percentages showed in the slide

http://www.isogg.org/wiki/Autosomal_DNA_statistics

and this one from the ISOGG Wiki for the expected number of cMs and segments as shown in table

http://www.isogg.org/wiki/IBD#Ranges_of_total_centimorgans_of_IBD_segments_expected.2C_based_on_family_relationship

Parent/child:	3539-3748 cMs
1st cousins:	548-1034 cMs
1st cousins once removed:	248-638 cMs
2nd cousins:	101-378 cMs
2nd cousins once removed:	43-191 cMs
3rd cousins:	43-150 cMs
3rd cousins once removed:	11.5-99 cMs
4th and more distant cousins:	5-50 cMs

Comparing your DNA results:

- **One to Many**
- **One to One**
- **One to One on the X**
- **An ICW function: People who match one or both kits**

The One to Many:

This is the MOST IMPORTANT of all the tools, check it regularly for new matches. Your newest matches are highlighted in green Looking at the one to many display, these are the columns:

Heading	Explanation
Kit Nbr	Kit number, ancestry kits start with A, 23andme with M, familyTreeDNA with T, generic with Z
Type	The companies change kits occasionally so this tells you which version, V4 is the current version for 23andme
List	Clicking this gives you the one to many for that kit, useful for comparing all your relatives to a new matching kit
Select	You can select a group of kits with this box, then click submit at the top and have many comparison options
Sex	If the user did not put it in you see a U, else F or M
Halogroup	there is a column for the mtDNA haplogroup and the Y haplogroup, self reported.
AUTOSOMAL	
Details	Click the A to get the one-to-one autosomal report
Total cM	For close family, total cM will follow Blaine's charts at ISOGG well, up to 3rd cousins
largest cM	For endogamous groups I recommend sorting by this column (the blue arrow at top)
Gen	Subtract one to get the cousin level. So 4 means a 3rd cousin. A nephew or uncle will be 1.5
X_DNA	
Details	Click the X to get the one-to-one X report.
Total cM	Total matching cMs on the X.
largest cM	Largest segment on the X, as X can reach very far back in time so only look at the large matches initially
Name	the name or pseudonym of the person
Email	email address to use for contact

In the generation listing your parent or child is 1, an aunt or niece is about 1.5, a grandparent is 2, for cousins subtract one from the number so if it shows generations of 4, the suggestion is 3rd cousin. My recommendation is to follow up with anyone who is 4.5 or less. Start with the closer ones. Accuracy in autosomal matching declines after 3rd cousins.

When you first upload to GEDmatch and run a one to many

- Select the top ten or so matches and use submit to get to the multiple kit reports
- Look at the generation report to see who might be related to whom
- Look at the gedcom report to see if any have gedcoms
- Look at any other reports that you prefer, matrices, browsers
- Pick the most interesting two or three to contact (see below for new matches, do that)

What to do with a new likely Match (4.5 gen or less)?

- Use the L to get this kits one to many and see if any known relatives match this person as well, if so check the select and run the various multiple kit reports
- If you collect segment data, use the A and the X to look at shared segments and add that to your master spreadsheet for this kit.
- Send your possible cousin an email with the two kit numbers, a link to your tree and your findings from the above and include any likely shared surnames or expected common lines.

The One to One Comparison

If you have clicked the A from the one to many then the kit numbers will be filled in for you.

Sometimes you may wish to lower the thresholds to see the smaller segments:

- The *SNP count minimum threshold* can be taken down as low as 300
- It is sometimes useful to lower the *minimum segment CM size* to 5 or 6

While the graphics are pretty, it is easier to cut and paste the results from the page with no graphics or to look at a page of only graphics.

How to make a master spreadsheet via GEDMATCH

People who have tested at 23andme or Family Tree DNA can download the data or all their matches from those testing companies or DNAGED.com. However ancestry users need to make their first segment spreadsheet using the Tier 1 utility “Matching Segment Search,” see

<http://blog.kittycooper.com/2016/09/taking-it-to-the-next-level-dna-spreadsheets/>

An ICW function: People who match one or both kits

You can find the people who you match in common with (ICW) another kit by using the function called “People who match one or both kits” on your home page in the analyze section

Advanced uses of the One to Many Grid

The chromosome browsers

The GEDmatch version of a chromosome browser lets you select multiple people from the one to many matrix and display the overlapping segments

- N.B. the color coding indicates the size in cMs since the graph is in terms of base pairs.
- Each person is a different line rather than a different color
- Old branch means that it is match to a previous person, a white line shows where

See the manual for GEDmatch available in my downloads area for details on the other advanced uses of the one to many matrix form

Phasing

- If you have at least one parent tested then you can separate what you got from each parent – this is called phasing
- GEDmatch will make you two new kits that have what you inherited from each parent
- You can then use the phased data for your one to one comparisons or any other function on the web site that takes a kit number
- Your phased kit number will show in your DNA resources box but it will not turn up in other peoples match lists

From the GEDmatch wiki

"The program will look at every data line in the child's raw data file and try to deduce the paternal and maternal value. If the position is homozygotic (that is AA, or CC, or GG or TT) split the data to each phase. If the position is a nocall it will be ignored. If the position is not included in at least one parent's data it will be ignored. If the position shows a data discord it will be ignored. If the maternal or paternal phase can be determined the values will be assigned else the original unphased values will be kept

Analyzing your DNA data

There are lots of fun things you can look at in your DNA results.

- Are your parents related?
- What color are your eyes?
- Any errors in your DNA upload?

Thanks for joining me, send me feedback via my blog

<http://blog.kittycooper.com>