Below are suggested topics and links to materials on the topic to be discussed at the panel discussion on April 16. **Please post your questions about the topics above on the Google Group website or email them to daytonbreidenbach@gmail.com , so we can focus the panel discussion**.

Preface

Genetic genealogy is a new (15 or so years) information source for genealogists. It doesn't supplant documentary genealogy (traditional sources) but complements it. In fact, genetic genealogy would do little to reveal your family story without the existence and aid of documentary evidence. The two sources complement and reinforce each other. Evidences for this are the tools offered by the testing companies that, if the information is shared by your match, provide quick access to the trees of your matches, lists of their ancestors, their families surnames and places where their ancestors lived. They list shared matches (ICW) matches which may provide clues as to how you relate. More recently they provide further associative tools that graphically relate how various matches are linked to each other and how closely. Examples are Ancestry's Circles and ThruLines, My Heritage's clusters, etc. There are increasingly available, third party tools to assist or extend these capabilities. In the end, both sources work together to reveal previously unknown relationships, confirm (or not) suspected or uncertain relationships and broaden your family story.

This said, it is possible to identify close connections based upon DNA evidence alone. this is particularly evident in the identification of half siblings from non-paternal events (NPE). This is of particular interest to those investigating blind adoptions. Recent news articles have related on the identification of half siblings matching sperm donors. Anyone unwilling to cope with such discoveries should not take a DNA test.

- A REVIEW OF THE BIOLOGY OF CHROMOSOMES AND THEIR INHERITANCE

If you want to better understand the underlying mechanisms of Following are links to short videos on biology of **chromosomes and DNA**: "DNA, How does it work" What is a chromosome and **meiosis**:How meiosis works, meiosis with crossing over, random orientation of chromosomes during meiosis and Phases of meiosis.

From these videos you will understand how the SNPs, described in the next topic **Features of a test**, get passed on from generation to generation. Why siblings are different and inherit different segments different combinations of SNPs and even different chromosomes.

You will understand that after the first meiotic division, where the chromosome pairs are randomly sorted into the two daughter cells, those cells could have 2 combinations out of 8.8 x 10 to the 23 power, possible combinations, ie, more than 8 million. Crossing over further increases the possible number of SNP distributions in one sperm or egg cell. Yet further diversity arises from the random fusion of any of one of the diverse population of sperm cells and any given egg.

You don't need to remember all the details about these processes, **but you do need to understand** that these processes are the basis of most of the rules that govern the correct interpretation of the relationships between you and the matches reported from your test. Another review of these concepts can be found here in the Beginners Guide to Genetic Genealogy. by Kelly Wheaton.

The sites above have other brief presentations if you want to dig deeper. For example the site Stated Clearly has videos like, How genes evolve If you want to understand base pairing and the formation of strands of base pairs What is the RNA world theory will help. There are other interesting videos on the Stated Clearly site, if you are curious.

The important concepts so far are:

1. **The reduction of the number of chromosomes in the gametes 1N=23**
2. **The restoration of the number of chromosomes in the embryo 2N=46**
3. **The shuffling of genetic material (DNA), Recombination of SNPs during this process.**

- FEATURES OF A TEST: WHAT IS TESTED AND HOW.

Features of a test describes the DNA molecules in a test sample, how they are processed, and how the test procedure produces the data you receive from the testing company.

example of raw test data

You have your test results reporting thousands of matches. Just like in conventional genealogy where there are millions of documents and you need to find the ones that answer your question, in genetic genealogy you need means of sorting through your matches to identify those that together with your documentary information help you answer your questions and break through your brick walls. Following are definitions and rules, strategies and tools to help you.

- DEFINITIONS:

1. **SNP**
   1. SNPs, or single-nucleotide polymorphisms, are locations on a chromosome where a mutation has occurred sometime in the past, altering identity of the base from, say an A to a T. Of course the complementary base is altered too. There are thousands of them scattered along a chromosome. SNPs are compared between two people to see if they match. The genetic distance of a consecutive sequence of matching SNPs is measured in cM.

2. **SEGMENT**
   A "segment" refers to a section or block of contiguous SNPs. A "matching segment" is the same sequential order of SNP at the same sequence of addresses on a chromosome for tests of two or more people. **But remember this rule**: Just because the test says the sequence is the same doesn't alone prove that it is. You should now realize that because of how the test is done, the test can't tell whether a matching base A at a particular address,came from your mother"s chromosome or your father's chromosome. The same is true for the next base and so on, so the match could be a totally false match of bases some from one side and some from the other side of your parentage. This fact cannot be stressed too much. Such a segment is termed IBS (identical by state or chance). Restated from the earlier discussion in the document **Features of a test**, the longer the match the less likely that it is false. Conclusion: Start with your longest matches and matches with persons you know are genealogically related. Strategies will arise from this to identify the true or false relationships with other matches.

3. **CENTIMORGAN** (abbreviated cM) is a measure of genetic linkage. Think of it as a measure of DNA information within a chromosome or a segment.

   The link above provides a more detailed definition and there is another discussion by here. If it is confusing to you don't worry because it is not

essential to interpreting your results. Note that all the major testing companies now report shared cM for your matches; based upon the amount of shared cM you can (roughly) estimate the closeness of the connection. **What is important here** is CM measure the strength of a match, how closely you are likely related

Each chromosome contains different amounts of information. Chromosome 1 contains 281.5 cM of information. Chromosome 2 has 263.7 cM. Chromosome 21 has only 70.2 cM.[1]

4. **In Common With(ICW) or Shared Matches**
   All Matches between 3 or more people are ICW (shared matches) whether they are on the same allele (sequence of SNPs) on one chromosome (either from mom or from dad), the corresponding allele on the other chromosome of the pair, or the three or more individuals who match on different alleles on the same chromosome or even who match each other on different chromosomes. These should be all be considered ICW until their status as triangulated can be verified.

   ICW matches can be important matches if they provide clues to their relationship to you. For example, if they share a common surname, a common time frame, or a common location or are part of a triangulation group that you can explore. Testing sites and various third party cells provide filters or sorting routines to identify ICW matches with such clues.
   Possible or probable Triangulations, as identified on FTDna, My Heritage,etc. are ICW,(not Triangulated), until they can demonstrated to be the same allele on the same chromosome of the pair of chromosomes. ie. phased.

   If each of three or more individuals match every other member of the presumed Triangulation Group (TG) on a significantly overlapping sequence (12 cM or more), it is almost certain they all share the same MRCA or ancestral line.
   If they also all match one parent of a match with the known MRCA, then they are proven to triangulate.

5. Triangulation is when three or more individual are verified to match on the same same allele on the same chromosome, either mom or dads, and are all descendants of a MRCA couple or their ancestral lines on one parental side or the other.
   If a MRCA or ancestral line is known for 2 or more of the group and one the matches is 12cM or more, the probability that they all share the same MRCA pair is very very high, but any member of the triangulation group who hasn't been phased or whose relationship is shown by documentation, could still possibly be from another ancestral line. This is especially true, if there is any chance of endogamy between the lines.

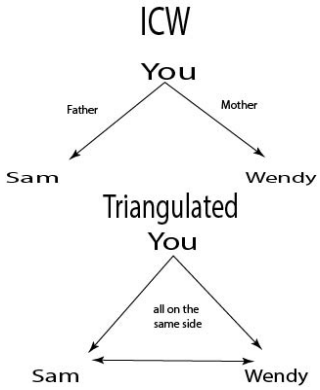
[1] This definition attributed to Jared Smith

We are always playing a probability game, unless the relationships are very close and positively known. A very clear discussion of this topic can be found here . When you are comfortable with the concepts of ICW and Triangulation. There is a more detailed discussion in part 2 of Debbie Kennetts blog.

To restate this very IMPORTANT concept: Triangulation is when three or more people who all match each other the same segment, or a significant portion of the same segment, on the same chromosome, meaning they all inherited that segment from the same ancestor couple MRCA or their ancestral lines. . Each individual must match both of the other members of the troika on the same segment and this must be confirmed.

Blaine Bettinger has a paper triangulation-intervention that details this concept and there is another good discussion in: Who are you made of by Amanda.

In summary, it is **important** to understand that a chromosome segment where three or more relatives are shown to overlap on a chromosome browser DOES NOT AUTOMATICALLY MEAN THAT THEY TRIANGULATE. It means that the individuals share the same sequence of SNP at the overlapped loci. These should be considered Shared or ICW matches until Triangulation is confirmed. If person A matches B and A matches C it must be shown that B and C also match on the same sequence on the same chromosome of the pair of chromosomes. Documentary evidence, if proven, can establish that they are all descendents of the same MRCA couple or their ancestral lines.

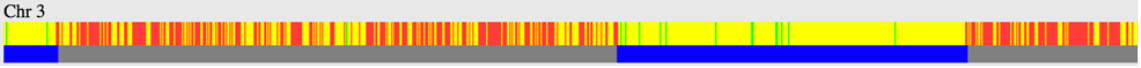


6. MRCA Most Recent Common Ancestor

Most recent common Ancestor couple for two people with a true match

7. Chromosome Browser.
   A tool that provides detail information about the properties of a matching segment. Most testing companies and third party sites provide browsers in some form. At a minimum a browser shows the chromosome number, the segment length in CM, the address of its starting and ending SNP, and the number of SNP's.

Chr	Start Location	End Location	Centimorgans (cM)	SNPs
3	36,495	5,168,135	15.8	2,114
3	104,270,146	168,695,458	57.2	13,878



   There is often an option to display the chromosome graphically, showing the segments.

The blue bars indicate two segments that match on Chromosome 3 between two people. The table indicates Start and End Locations and the cM and number of matching SNPs in each segment. You'll notice that the start location for the first segment is 36,495 instead of 0 even though it appears at the beginning of the chromosome - this is because not all markers in a chromosome, especially those near the ends, are tested. The blue identifies matching segments. Read Jared's article for what the other colors portray.

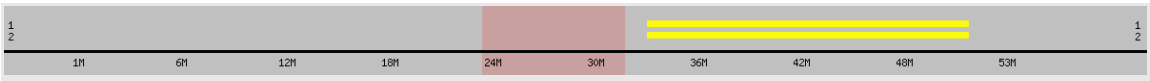
Gedmatch has recently become Gedmatch-Genesis and the format of many of the features Smith references has changed somewhat, but using his guide and instructions on the site should allow you to learn to use the sites tools. Genesis is so new that the Genetic genealogy community and its blogs have not quite caught up yet. We are sure better discussions will come soon.

Below is the Genesis display of the browser.

Chr 19

Match ID	Name	Matching segments on Chromosome 19	Overlap with previous match
1	sam	33767568 - 52157686 (32.993 cM)	root
2	jim	33767568 - 52157686 (32.993 cM)	33767568 - 52157686

Chr 19



Note the red green and yellow as in the older browser are gone. there are now new visualization options to show these properties.

8. Pileups **Unimportant matches:** You may find a especially large group of matches, 10cM or less that have nearly the same starting and ending points. These are likely to be pileups.There may be cousins in this group, but more likely this group is made up of a lot of IBC matches or IBP (Identical by Population) matches because of endogamy where a sequence has been fixed in a population from an isolated population with a lot of intermarriage. It can also arise if the SNP's in the sequence are close (linked) to a highly important (evolutionarily conserved) gene. In any case these groups if identified should go to the end of your to do list, except in specialized research cases.

9. **Important Matches for Id'ing MRCA and ancestral lines.**

    i. **Test your oldest and closest relatives ASAP (**if you test on Ancestry transfer your test to another testing company and Gedmatch. If you tested elsewhere transfer to Gedmatch (as technology advances you may be able to test older ancestors from saliva on old letters, hair clay pipes, etc.

        1. **Grandparents**

2. The earlier the tested generation the further back in your genetic ancestry you can go

1. Parents

3. If your parents are tested then every segment of your dna can be assigned to either your paternal or maternal side and you can identify segments that were passed forward from your grandparents.

1. Aunts and Uncles

4. Your aunts and uncles will have segments from your grandparents that your parents may not have and can stand as surrogates for an untested parent.

1. First cousins

5. If none of the above are tested they can help assign Paternal and Maternal sides to matches.

1. Second cousins

6. Triangulated matches with 2nd cousins with you and your great grandparents are either MRCA or someone in their ancestral line is.

7. Nth cousins

8. Can be assigned to relationships by triangulation with other matches who's side or MRCA is established. Statistically the most important matches in this group are the largest matches because they are more likely closer relationships to you and therefore more likely to have the relationship discovered through trees and other documentary evidence. This will be elaborated on when we get to discussions of testing company and third party tools as well as data management.

Following are some links to articles and videos that review and expand on what has been presented so far.

Jared Smiths "an Absolute Beginners Guide", although focused on "Genetic Genealogy using Gedmatch" and is cited again later under the topic Gedmatch-genesis, provides a complementary description of chromosome basics and chromosome inheritance. You should read this article at this point, if you haven't already, at least to the end of his section on definitions. We will use the rest of the article to learn how to use the third party tool Gedmatch. Much of the information is applicable to using match information in any site having a Chromosome Browser. Also see

Genetic Genealogy to Breakthrough Brickwalls presentation by Jonathan Long for more details

Diane Southard has a free online video from the 2019 Rootstech conference that is a good "intermediate" level discussion of how to use genetic genealogy (autosomal tests specifically) to address brickwalls in your family tree:

If you are still with us to this point, the next sections describe some of the testing companies and third party tools that will allow you to exploit genetic genealogy as a

means to open new vistas for your family history research. The advances in technology that simplify genealogical/genetic genealogy research are truly amazing.

Your genealogical ancestors and your genetic genealogy ancestors and all of their descendants can be defined as a multidimensional network. As a consequence they can be analyzed using theories of networks and artificial Intelligence to identify relational associations among individuals in the multitudes of people in the network. New tools based on these concepts are popping up by the day at both the testing company sites and third party sites. Many of these have just been announced at the 2019 Rootstock convention in Salt Lake City. Ancestry has just announced “Thrulines” and My Heritage announced “Theories of Relativity” as well as “Cluster analysis.” Third party sites such as DNADNA, Nodex, Genetic Affairs and others provide similar tools using network and graphic analysis algorithms to analyze DNA or tree data or both to reveal clusters of ICW matches.

These tools then suggest how your DNA matches may be related to a particular ancestor or ancestor line and even suggest possible ancestors you didn’t even know you had. This can lead you past the barrier imposed by the vast array of matches, that seem almost impossible to relate to your family history. This is revolutionary. That being said the relationships being suggested by these tools are hypothesis and hypothesis only. They are hypotheses that rely heavily on trees in the respective databases and we all have encountered trees that are terribly wrong. Roberta Estes has carefully cautioned this in a blog “Ancestry’s ThruLines Dissected: How to Use and Not Get Bit by the ‘Gators Similar cautions apply to My Heritage’s Theories of Relativity.

All these cautions aside, these hypothesis allow you to group your matches with your own tree data to confirm relationships and formulate new relationships and ancestral lines. This is a game changer. If you can find members of a cluster who have tested with other testing companies or have uploaded to Gedmatch, My Heritage, etc. where there is a chromosome browser, you can Triangulate matches, know which ancestral line they pertain to and map ancestral DNA to specific chromosomes and segments with tools like DNApainter or GenomeMate Pro. Exciting!!!!

Now to specifics on testing companies and third party sites.

- **Testing companies for genetic genealogy**
  ISOGG has a good summary of the major testing companies here:
  https://isogg.org/wiki/Autosomal_DNA_testing_comparison_chart

  also reviewed here *Last Updated 4-2017*

Ancestry

Note that Ancestry has the largest pool of matches, especially in the U.S. (although they have branched out to other parts of the world in recent years), and those matches are usually interested in genealogy. But AncestryDNA does lack a chromosome browser tool which is very useful for validating the nature of a match. It does have other very useful tools, such as shared matches (fourth cousins or closer), DNA "circles" of people who share DNA and a common ancestor, and "new ancestor discoveries" based upon matches with people who share a common ancestor but don't connect to your own family tree. Crista Cowan Ancestry at Rootstech

Note also that AncestryDNA does include some Y DNA SNPs that can be mined for clues about Y haplogroups. Link to further discussion of Ancestry site

- **FamilyTreeDNA**

  FTDNA has been doing genetic genealogy the longest and is particularly suited to genetic genealogy, since it offers three kinds of genetic tests (autosomal, Y, and mitochondrial) as well as a chromosome browser tool. For detailed Y and mitochondrial testing, FamilyTreeDNA is often the best choice. It offers a number of tools to use in sorting through your matches to try to assign them to your P or M side and to MRCA discussed by Roberta Estes

- **MyHeritage**

  A relative newcomer to the field of genetic testing, they already have more people who have tested than any companies except Ancestry and 23&Me, and a larger percentage of them live outside the US than is true for other companies. They have a strong focus on genealogy, encourage uploading of family trees, and provide excellent tools to compare your tree with those of your matches. Roberta Estes has a recent blog on a new feature on My Heritage that appear to be a very promising tool, called "Theory of Relativity."

- **YSEQ**

  YSEQ (established by FTDNA specialists who left the company) is a cost effective tool for advanced testing (such as full sequence mitochondrial and advanced Y SNP testing). But you would need to work with a project coordinator or experienced genetic genealogist to get much from the results.

- **23andme**

  23andme offers 3-way tests (Y, mitochondrial, and autosomal) at relatively low cost, and it offers useful tools such as matching, matches-in-common, and a chromosome browser. It has a relatively large pool of users, but many of them have little interest or background in genealogy, relative to the other companies.

- **LivingDNA**
  This company specializes in the British Isles but is expanding to other countries. It offers relatively refined Y and mitochondrial haplogroup assignments. LivingDNA is adding matching functions in "Beta" testing now; until that expands, its utility for genetic genealogy will be very limited. But they are accepting free transfers of data from other companies, so it could be wise to transfer or do their test.

- tribecode ?
- **Notes on "3 in 1 tests"**
  23andme and LivingDNA are useful in that they are "three-in-one" tests that include autosomal, Y DNA, and mitochondrial DNA for a relatively low cost.

- **Third party tools for analyzing and interpreting your match data**

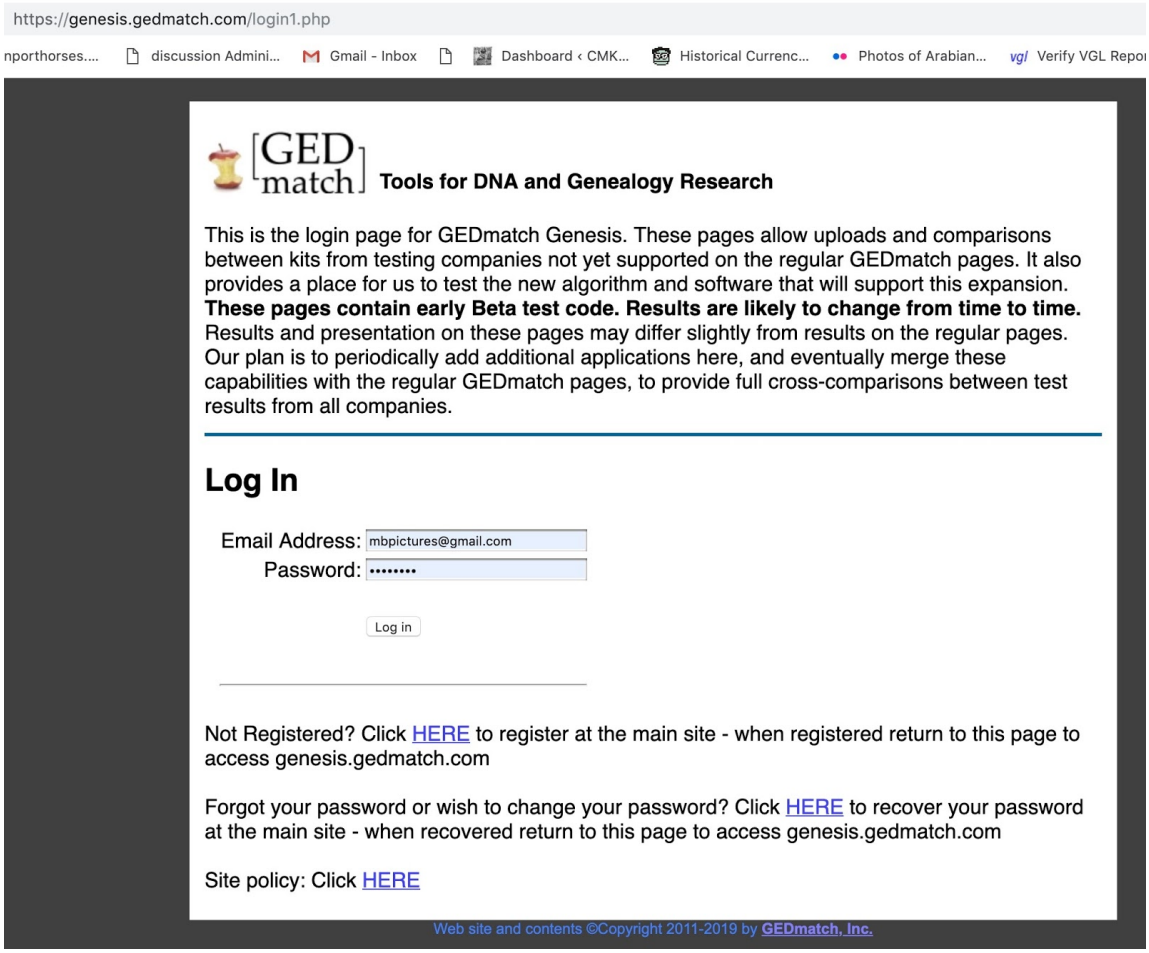
- **Gedmatch>Genesis**: Is a very useful and important third party site and is especially useful, if you can get your matches on Ancestry to upload their test to Gedmatch.

  The way it works is, you download your raw data from the testing company and upload it to Gedmatch. Here is how . Once your test data is processed at Gedmatch you will be able to see everyone who matches you, who also uploaded their test data from any testing company (including Ancestry) to Gedmatch. Gedmatch has received a lot of publicity lately because it has been used increasingly by law enforcement to solve cold cases.

  Jared Smith provides an excellent description of how to use the site after you have uploaded your data. Although the details differ the principles for using the chromosome browser and other tools are similar at various testing company and some other third party sites. Kitty Cooper describes some of the features in Genesis compared to Gedmatch and there is a tutorial for beginners

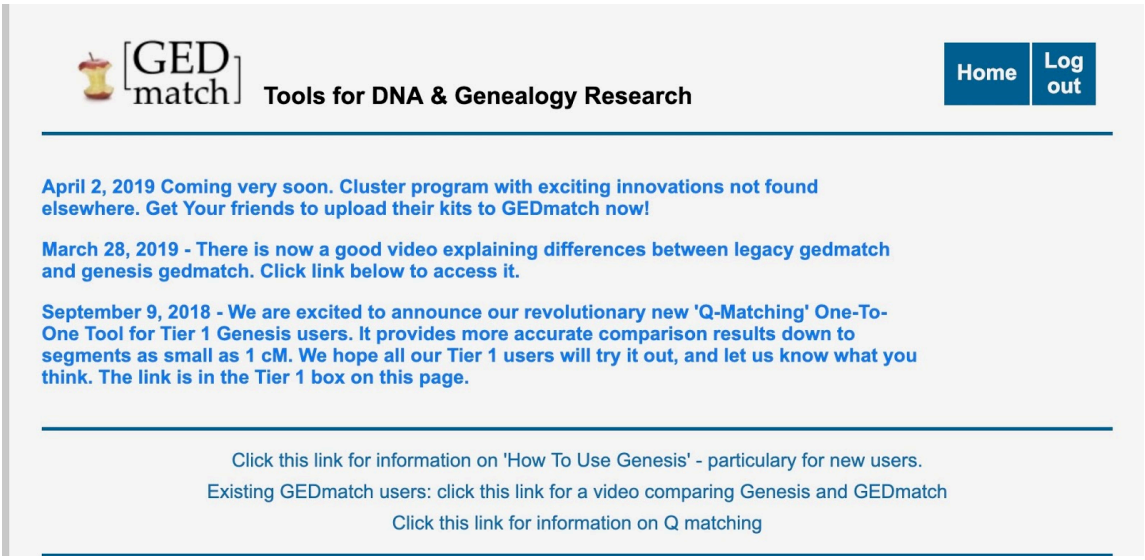
  Here's the login page for Genesis; note that the URL is visible at the top of the screen. If you've already been working with GEDmatch you are ready to go at Genesis and can reach this page after you sign in on your GEDmatch ID.

12

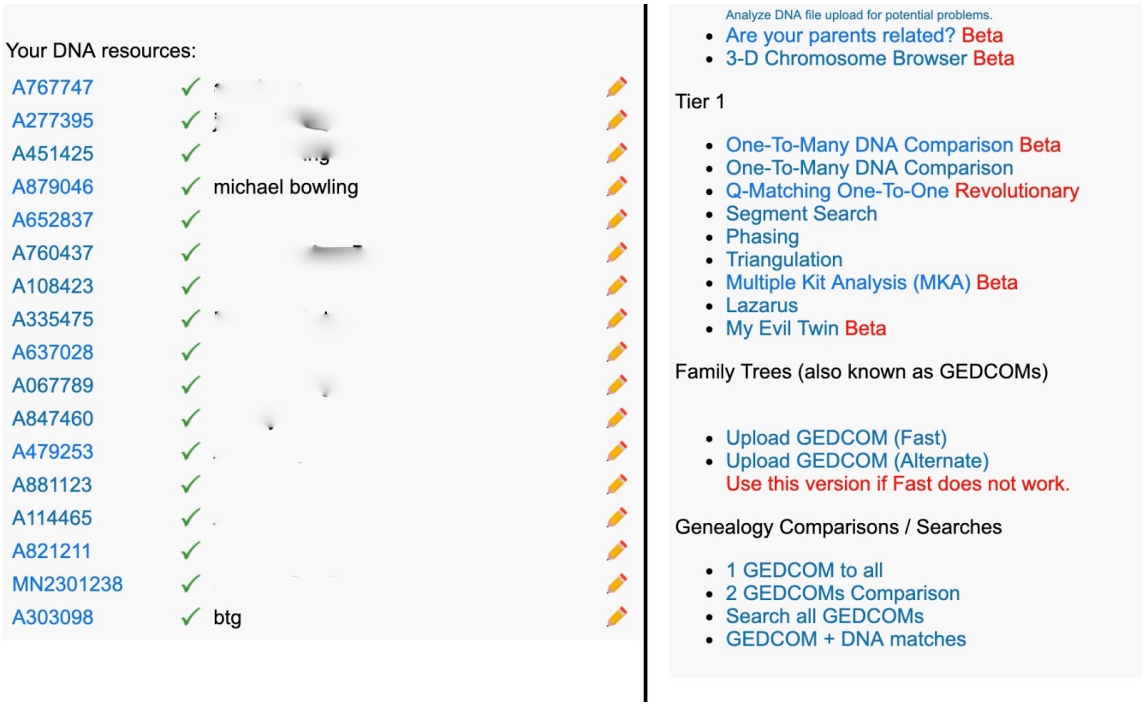


As Bill has made clear, this whole area is changing even as we watch; I will try to familiarize myself with Q-matching before the presentation, but I can't promise anything.

13



The options for working with Genesis are listed on the page after you're logged in. We will discuss some of the basic possibilities, but there's a lot more and the functions are always expanding.



User Profile(88975):
Name: Michael Bowling
Email: mbpictures@gmail.com

Tier 1 Member

Tier 1 Sustainer

View/Change/Delete your profile (password, email, groups)

The number of online users is 178

LEGEND:	
(Status indicators shown to the right of each kit below)	
✓	Kit has completed all processing and has good status
	Click on pencil if you wish to EDIT or DELETE kit profile
	Kit has not yet completed matching with other kits
?	Unknown Status
Click on blue kit number to go directly to one-to-many results	

Information:

- Welcome to Genesis BETA
- User Lookup - Find information on your matches.
- Information area about Genesis
- Take me back to the main GEDmatch site

Upload your DNA files:

- Generic Uploads (23andme, FTDNA, AncestryDNA, most others)
- Upload if generic upload fails

DNA Applications:

- One-To-Many Beta - give it a try
- One-To-Many DNA Comparison Result
- One-to-One Autosomal DNA Comparison
- One-to-One X-DNA Comparison NEW
- Admixture (heritage)
- Admixture / Oracle with Population Search NEW
- People who match both, or 1 of 2 kits NEW
- DNA File Diagnostic Utility
  Analyze DNA file upload for potential problems.

Your Family Trees (Also known as GEDComs) below:

1330424 2015-04-09 George Michael BOWLING

Click on the GEDCOM number above to go to the individual detail page for the point person. To change the point person, or to create a link between a DNA kit and a person in your GEDcom: Go to their individual detail page in the online tree and at the bottom of that page there is a box for linking a DNA kit to that GEDCOM, or for changing the point person.

Click HERE to manage GEDCOM resources.

Join Tier1

- Basic GEDmatch® programs remain free and we plan to keep them this way. Purchase of a Tier 1 membership helps cover the costs associated with running this site, and will provide you with the benefi of using the additional Tier 1 tools for a period of time equal to one month for every $10.
- You may use the 'Join Tier1' button below, for a one-time payment of any amount, or the 'Monthy Auto Renewal' button to establish a recurring $10 per month amount.
- Information on canceling a recurring payments can be found HERE.
- If you do not wish to use PayPal, you can send your check or money order to GEDmatch, c/o Curtis Rogers at 710 First Avenue South, Lake Worth, FL 33460. Please write your GEDmatch login email on your check.
- Thank you for your support of GEDmatch!

Time Period

1 Month $10.00 USD

Join Tier 1 Or Monthly Auto Renewal

- **Wikitree**
  Wikitree is a public platform that is building a single “world family tree” to which individuals contribute. Notably, people can report if they have done genetic genealogy tests. Most of those tests are autosomal results, but wikitree can be very helpful in identifying patrilineal or matrilineal descendants in lines of interest to you, as well as other genealogists who may be able to support your searches. They also enforce rules on the verification of entries into the tree. They just celebrated (4/19) reaching 20 million profiles on their site.

Cluster analysis tools: using shared or ICW matches

- There are a number of new tools all based on the idea of using various associative algorithms to group matches and discover patterns of inheritance from ICW also known as shared matches. This is especially valuable for Ancestry matches where you do not have a chromosome browser and cannot triangulate matches, but can be helpful for analyzing shared matches from other testing companies as well. Blaine Bettinger first provded an analysis of the general principles of cluster analysis in 2017 when Ancestry introduced shared matches information on it’s website. Since then tools to exploit this concept have proliferated.
  - **Genetic Affairs autocluster**
    The Intrepid Sleuth provides a step by step process for submitting your data to Genetic Affairs Autocluster for analysis. Your autoclustered results are emailed to you. My Heritage will also email autoclustered data from your shared matches in their database.

- DNA Painter Blaine Bettinger onDNAPainter
-

- NodeXL
  NodeXL Basic – a free template created by the ‘Social Media Research Foundation’ – which gives Excel the ability to create network charts. It allows the user to utilize the tools of the branch of mathematics that analyzes complex networks relationships graphically. Twigs of Yore

- DNADNA
  another graphic analysis of relationships among ICW groups of matches

- **Managing your information:**

1. **Spreadsheets**
2. Many genetic genealogists manage their data using spreadsheets, Excel or Google Sheets, and a cloud based word processor such as Google Docs. The value of the cloud based systems is having web addresses for each document which can thereby be linked to. Roberta Estes (DNAeXplained) gives a detailed description spreadsheet sorting and of her spread sheet system. Every Genetic genealogist using spreadsheets etc probably has a unique system.
3. **Genome Mate Pro**
   GenomeMate Pro is a downloadable program on your computer that provides a central place to organize all of your genetic genealogy information from all tests and provides many tools to manage, visualize and analyze it. It has a bit of a learning curve, but if you are careful to read and follow the instructions in **Getting Started with GenomeMatePro** provided by the DNAGeek Blog site. An advantage of GMP is that it provides most of the tools discussed above and more that can be applied to the data from all of your tests as well as third party data from Gedmatch Genesis

   (IMPORTANT But likely to change. At this Time 3/1/19 The format One to Many data of Gedcom changed with the switch to Genesis and the proper import template to adjust for this change has not been added to GenomeMatePro yet. The work around has been posted on the GenomeMatePro Facebook group site.)

   Currently your test (one to Many match) imports: you must first open Genesis and use the 2nd item one to many under the heading DNA Applications. It will take you to the screen shown below. THIS IS THE ONLY ONE TO MANY THAT WILL WORK WITH THE STANDARD TEMPLATES for importing data in GMP.

   ...See More at this link

- xDNA

  Roberta Estes has a case study which is a good example of the strategies for using xDNA to help solve a question; which ancestors are the most likely to have introduced native american heritage into her genome.

- Y DNA

- Strategies for testing hypothesized ancestral connections efficiently:
    1) Form a testable hypothesis
    2) Choose the right test (and company) to evaluate the hypothesis (consider Y-DNA and mitochondrial DNA, and X DNA, in addition to autosomal DNA)
    3) Find useful individuals to do the test (or who have done one that you can use)