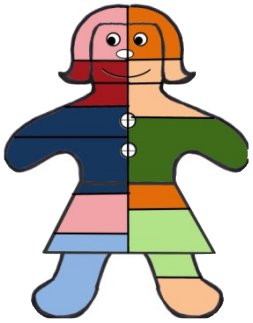


Genetic Genealogy Tools to Achieve DNA Discoveries

Debbie Parker Wayne, CG[SM], CGL[SM]

One look at the results we see once our DNA test is completed makes it clear we need to learn more about analysis, we need to be organized to make sense of the enormous amount of data, and we need good tools to help us achieve our goals. Articles and books we read, webinars and videos we study, classes and lectures we attend all help us learn more. Some tools help us be more organized; we also need to structure our file system folders in a way to make it easier to find specific data. Other tools help us in analyzing the DNA data. Many more tools are available than can be discussed in one article, but this will introduce you to some useful tools.

In this article the term “focus person” identifies the person whose DNA account or profile is being used for the analysis. The term “test-taker” identifies the DNA matches of the focus person.

MATCH LISTS

Each DNA testing company and many third-party tool compares one test-taker’s DNA (the focus person) to the DNA of other test-takers in the database or folder. When the focus person and a test-taker share enough DNA to meet a threshold that is imposed by the company or tool, the name of each will be on the match list of the other. These “DNA matches” are

---

Original publication:
Wayne, Debbie Parker. “Genetic Genealogy Tools to Achieve DNA Discoveries.” *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

related in some way even if the family tree is not complete enough to name the common ancestor.[1]

The total amount of DNA shared by the focus person and the test-taker is used to predict how the two may be related. The companies display the total amount of shared DNA in percentages or centimorgans (cM). Centimorgans are a logical measurement of the "length" of the shared DNA (this is a simplified definition for those new to genetic genealogy).[2] Centimorgans are not a physical measurement where one cM is always the same length; some 15 cM blocks are physically longer than other 15 cM blocks in a different place on the chromosomes, but they are logically the same–they represent the same value for the analysis we are doing. The percentage is the total of the DNA tested that matches between the two test-takers.[3]

This match threshold is important to understand. It means you can have two people who are known cousins that, due to random recombination, do not appear on the match list as cousins. To date, all known second cousins share enough DNA to be on each other's match list. Beyond the second cousin level, it becomes possible that both did not inherit enough DNA from the common ancestor to meet the company matching threshold.

---

[1]All websites accessed 13 August 2017.

There may be a small number of persons on a match list who are not really cousins (false positives) and there will be some who share an ancestor so far back in time it cannot be proven with documentary research. The testing companies strive to provide as many real matches as possible and no false positives. The false positives may be the result of misread DNA data causing the comparison algorithm to see a DNA match where none exists. These false positives can sometimes be eliminated by testing parents, children, or siblings, and comparing their results to the suspected false match. False positives are more likely to appear in distant relationships than in close relationships.

[2] For more complete definitions see Blaine T. Bettinger and Debbie Parker Wayne, *Genetic Genealogy in Practice* (Arlington, VA: National Genealogical Society, 2016), 140 and "centiMorgan," *ISOGG Wiki* (https://isogg.org/wiki/CentiMorgan).

[3] Test used today do not test the whole genome of a person–all of a person's DNA. Only a small amount of DNA is tested (500,000 to 700,000 DNA locations are tested out of the 3 billion or so possible locations).

Original publication:
Wayne, Debbie Parker. "Genetic Genealogy Tools to Achieve DNA Discoveries." *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

This does not mean the two test-takers are not cousins or that the family tree is incorrect. This is a limitation of the level of testing available today and the technology used in the labs. Researchers must understand these limitations in order to accurately interpret DNA test findings for application as evidence in a genealogical problem. Some think we will have better match predictions once we are testing more DNA (as will be done once we have reasonably-priced whole genome sequencing). When and if lab procedures are able to differentiate which of a test-taker's DNA came from each parent (phased DNA data that has been assigned to a specific parent) that will be a major improvement allowing the matching algorithms to be much more accurate.

Match lists may contain thousands of matches. A researcher will be most efficient by focusing first on the matches where it is more likely a common ancestor can be identified. Start by scanning the match lists looking for test-takers

(1) who are known cousins (maybe known from documentary research done before DNA testing was available),

(2) where the test-taker's surname matches an ancestral surname,

(3) where the test-taker has listed an ancestral surname of interest (in a surname list or as part of a family tree),

(4) who have ancestors of any surname in the same location at the same time as ancestors of the researcher, or

(5) who share the most DNA with the researcher.

Identifying the shared common ancestor with one person on a match list can help identify the relationship to others who also share the same common ancestor and who inherited the same segment of DNA. Once we

---

Original publication:
Wayne, Debbie Parker. "Genetic Genealogy Tools to Achieve DNA Discoveries." *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

“map” a segment of DNA to an ancestor, matches with new test-takers on the same segment can be focused on that ancestral line when searching for the common ancestor.

MATCH LIST NOTES

Most companies allow a researcher to save notes online about the relationship to another test-taker. Before spending hours making notes on a company server, take some time to determine if there is a way this information can be saved on your own device. The major testing companies may not go out of business any time soon, but they periodically make changes to their user interfaces that do not always improves things for the user. Many genetic genealogists prefer to save notes on their own device in a spreadsheet, word processor document, or a tool such as *Genome Mate Pro* (described below).

IN COMMON WITH (SHARED MATCHES)

The companies and third-party tools generally provide a way for two test-takers to determine who else in the match list shares DNA with both the focus person and another test-taker. This is generally named “In Common With” (ICW) or a phrase such as “Shared Matches.”

Shared matches are a powerful tool, but there are some limitations. Most of the tools only compare the match lists to determine which names of other test-takers are duplicated for two focus test-takers. That is, a researcher logs in to the DNA account of a focus person, chooses a second person from the match list, then selects options to display the list of other test-takers who appear on the match list for the focus person and the second person selected.

Original publication:
Wayne, Debbie Parker. “Genetic Genealogy Tools to Achieve DNA Discoveries.” *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

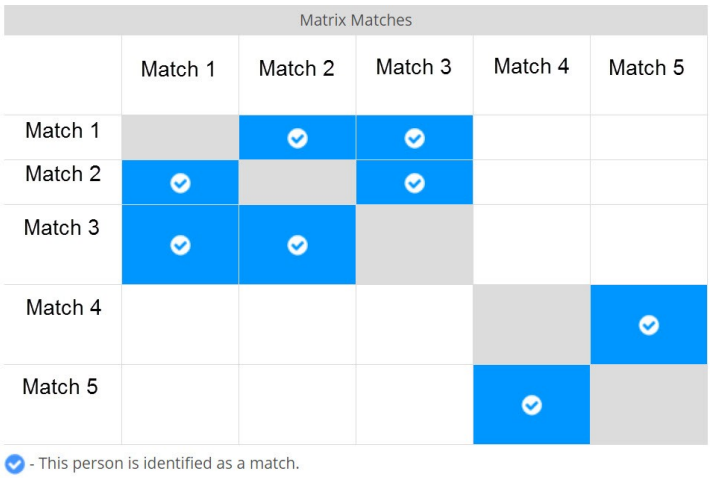
Some tools are not determining if everyone on the shared match list shares the same DNA segment, only that some DNA is shared. This means it is possible for the focus person and the second-person to be related to those on the shared match list through different ancestral lines. The focus person might even be related to the second-person on a maternal line and to someone else on the shared match list on a paternal line. Some of these tools output a list and some a graphic format.

Tools that do *not* indicate whether the shared DNA is on the same segment include Family Tree DNA's[4] In Common With and Matrix; AncestryDNA's[5] Shared Matches and DNA Circles; The Autosomal DNA Segment Analyzer (ADSA, hosted on DNAgedcom)[6] which creates a graphic display showing ICW matches displayed side-by-side with the chromosome segments each test-taker shares with the focus person. ADSA allows more test-takers to be listed in one display than the company tools allow.

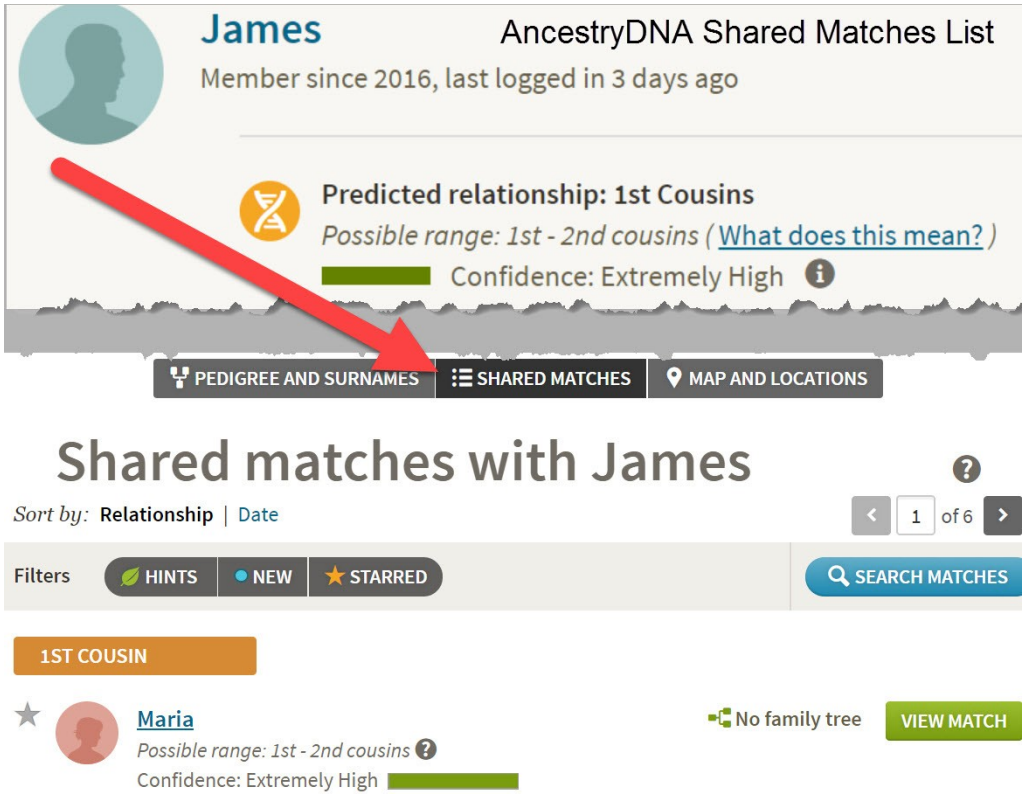
DNAgedcom[7] also hosts JWorks, an Excel-based tool, and KWorks, a browser-based tool, to analyze match lists and ICW files. These tools eliminate small segment matches and group matches into ICW segments.

[4] *Family Tree DNA* (http://www.familytreedna.com/).
[5] *AncestryDNA* (http://www.ancestrydna.com/).
[6] *DNAgedcom* (http://dnagedcom.com/).
[7] *DNAgedcom* (http://dnagedcom.com/).

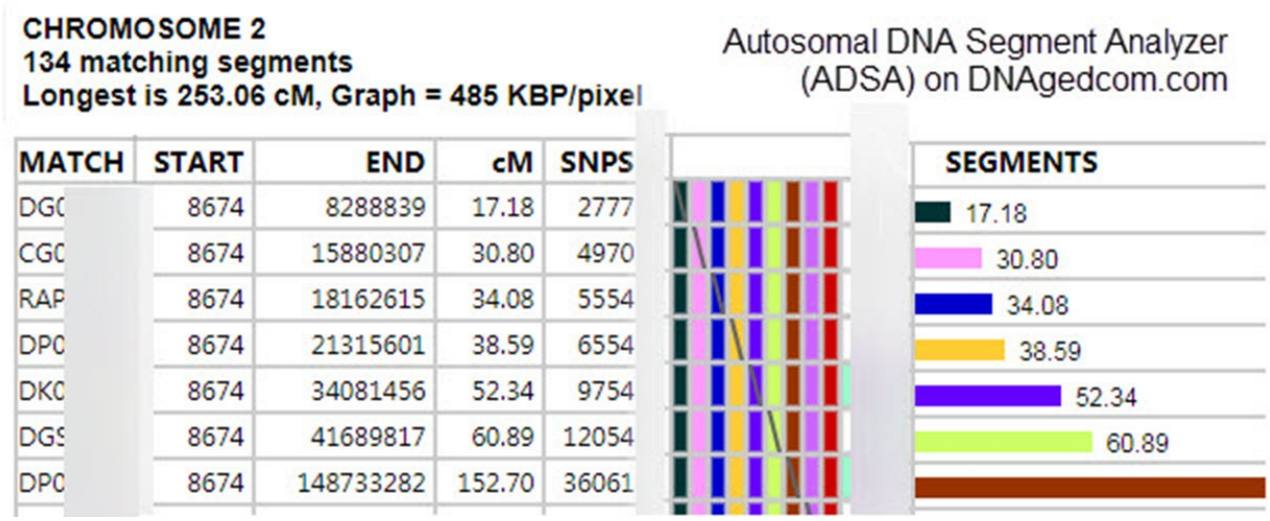
Original publication:
Wayne, Debbie Parker. "Genetic Genealogy Tools to Achieve DNA Discoveries." *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

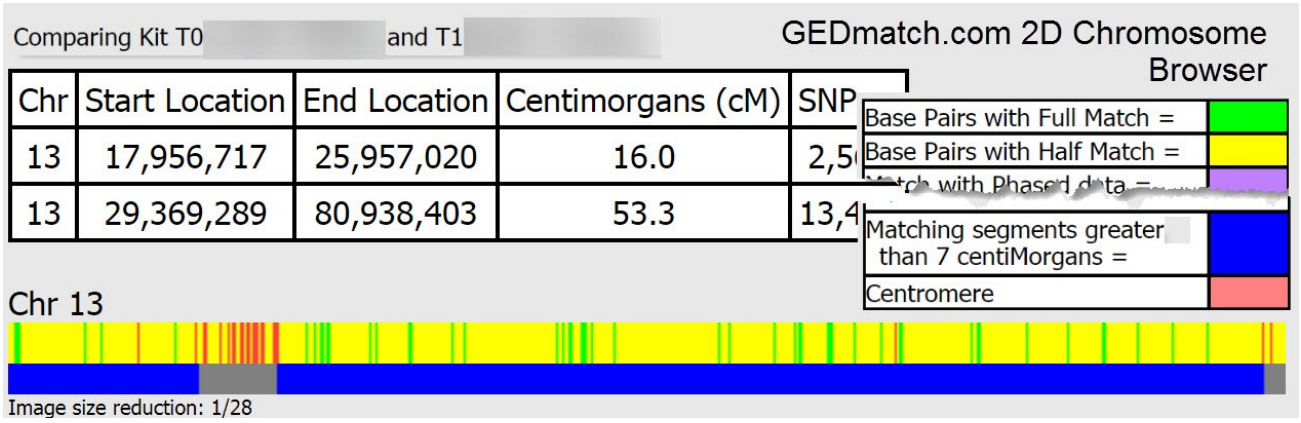


Family Tree DNA Matrix
(Showing Matches In Common with Shaded Boxes with Check Marks)

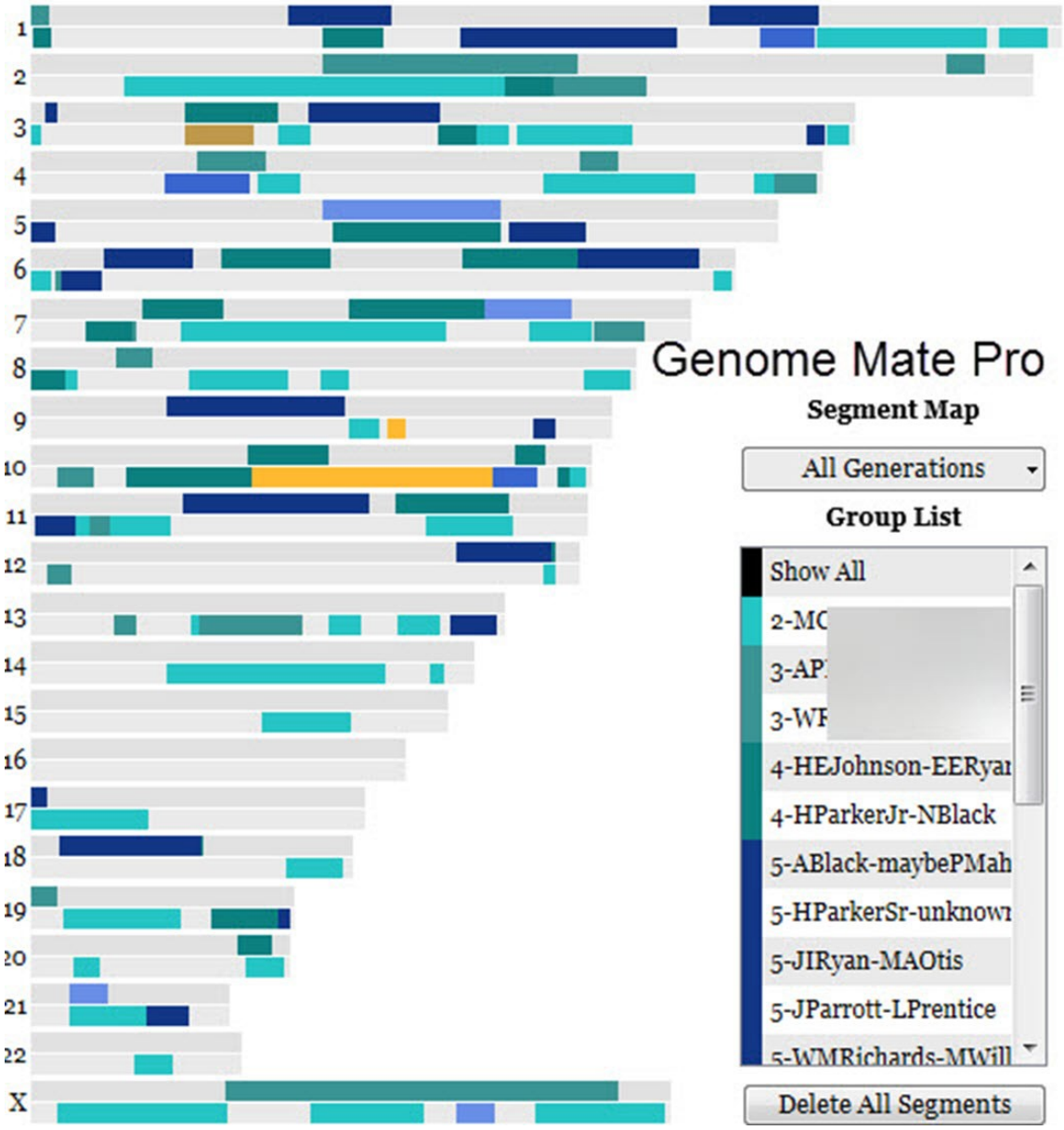


Original publication:
Wayne, Debbie Parker. “Genetic Genealogy Tools to Achieve DNA Discoveries.” *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).





Original publication:
Wayne, Debbie Parker. “Genetic Genealogy Tools to Achieve DNA Discoveries.” *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).



Original publication:
Wayne, Debbie Parker. “Genetic Genealogy Tools to Achieve DNA Discoveries.” *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

Double Match Triangulator Tool

Names 1 2 3

	CHR	LEAD	L	START	END	CM-AB	T	TRAIL	CM-A	CM-B	STATUS	22	23	24	25	26	27	28	29	30	3
DJF BJ(DP	10	88087	a	21840983	29033807	8.25	a	29268325	52.71	8.25	Full Triangulati	X	X	X	X	X	X	X	X		
DJF BJ(Da	10	21840983		21840983	30313037	10.33		30313037	10.33	10.33	Full Triangulati	X	X	X	X	X	X	X	X	X	X
DJF BJ(Ro	10	20918456	b	22679713	45461843	20.34		45461843	20.34	22.52	Full Triangulati	X	X	X	X	X	X	X	X	X	X
DJF BJ(Ba	10	13748888	b	22679713	53469337	26.04		53469337	26.04	39.54	Full Triangulati	X	X	X	X	X	X	X	X	X	X
DJF BJ(Va	10	21118428	b	23517068	29883839	8.23	b	30313037	8.23	11.56	Full Triangulati	b	X	X	X	X	X	X	X	b	
DJF BJ(Da	10	21398148	b	23517068	45461843	19.13		45461843	19.13	21.37	Full Triangulati	b	X	X	X	X	X	X	X	X	X
DJF BJ(Pa	10	21840983	b	23517068	45461843	19.13		45461843	19.13	20.78	Full Triangulati	b	X	X	X	X	X	X	X	X	X
DJF BJ(Sh	10	21840983	b	23517068	53469337	24.84	b	54083560	24.84	27.29	Full Triangulati	b	X	X	X	X	X	X	X	X	X
DJF BJ(PC	10	18942808	b	23517068	53742907	25.13		53742907	25.13	31.14	Full Triangulati	b	X	X	X	X	X	X	X	X	X
DJF BJ04 f	10	0	.	23517068	59875823	31.25	.	249999999	31.25	31.25	Base a-b	.	X	X	X	X	X	X	X	X	X
DJF BJ(RA	10	8656133	a	23517068	59875823	31.25	a	118417769	119.05	31.25	Full Triangulati	a	X	X	X	X	X	X	X	X	X
DJF BJ(ZK	10	88087	a	23517068	61164043	32.31	a	88598789	108.21	32.31	Full Triangulati	a	X	X	X	X	X	X	X	X	X
DJF BJ(KO	10	88087	a	23517068	61164043	32.31	a	135327873	176.25	32.31	Full Triangulati	a	X	X	X	X	X	X	X	X	X
DJF BJ(Mi	10	24656853		24656853	30616944	7.76		30616944	7.76	7.76	Full Triangulati			X	X	X	X	X	X	X	
DJF BJ(Ca	10	29270919		29270919	53469337	17.52	b	53742907	17.52	17.81	Full Triangulati								X	X	X
DJF BJ(AC	10	88087	a	29410436	59875823	23.68	a	135327873	176.25	23.68	Full Triangulati	a	a	a	a	a	a	a	X	X	X

Original publication:
Wayne, Debbie Parker. “Genetic Genealogy Tools to Achieve DNA Discoveries.” *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

CHROMOSOME BROWSERS AND SHARED DNA SEGMENTS

Chromosome browsers offer a graphic display of shared DNA segments between a small number of test-takers. AncestryDNA does not provide a chromosome browser or the necessary data for segment analysis. Other testing companies and tools display the DNA segments shared by the focus person and those test-takers in the match list.

Researchers must understand the chromosome browser only illustrates the DNA shared between the focus person and that person's DNA matches. There is no guarantee that the matches will match each other at all or in the same location where each matches the focus person. Each person inherits one copy of each autosomal chromosome from his or her mother and one from the father. The chromosome browsers do not differentiate between those who match the chromosome 1 inherited from a focus person's father and the chromosome 1 inherited from the focus person's mother. Additional analysis to compare the matches to each other may help separate the maternal and paternal matches.

TRIANGULATED DNA SEGMENTS

As of July 2017, 23andMe[8] and GEDmatch[9] provide a way for a researcher to easily determine whether three test-takers all share the same segment of DNA. When three or more test-takers all share the same segment of DNA it more likely that that DNA came from the same common ancestor.

[8] DNA Relatives tool, *23andMe* (http://www.23andme.com/).
[9] One-to-One, X One-to-One, Matching Segment Search, and Triangulation tools, *GEDmatch* (http://gedmatch.com/).

Original publication:
Wayne, Debbie Parker. "Genetic Genealogy Tools to Achieve DNA Discoveries." *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

If a researcher has access to the DNA data files of multiple test-takers, the Double Match Triangulator (DMT)[10] tool will be useful. DMT uses the chromosome browser file (the file listing the chromosome number with start and end location of each matching DNA segment) of two focus persons to determine which DNA segments are shared by the two and *other* test-takers in the database. DMT identifies both triangulated and un-triangulated segments. DMT is a powerful tool, but requires careful reading of the user information to correctly interpret the output of the tool.

GENOME MATE PRO

Genome Mate Pro (GMP)[11] is a DNA data management application that runs on your computing device (Windows, Linux, or Mac). GMP allows data from multiple testing companies and websites, and from multiple test-takers, to be combined in one database for analysis. Many of the functions described above are available in GMP.

The GMP database is contained on a personal computing device, not stored on a website. The disadvantage of this approach is that it takes some time to download the data from the testing company or tool website and import the data to a GMP database. The process must be repeated periodically to access new matches found after the initial import.

Some find it difficult to get started with GMP. If the instructions are carefully followed, the data import process is smoother. It does take several hours to get all of the data needed into GMP, but once the data is in the GMP database the power of the application is impressive. Updates to imported data go more quickly than the initial import.

[10] Louis Kessler, "Double Match Triangulator," *Behold Genealogy* (http://www.beholdgenealogy.com/dmt.php).
[11] Rebecca Walker, *Genome Mate Pro* (https://www.getgmp.com/); *Genome Mate Pro* Facebook Group (https://www.facebook.com/groups/816785941743656/). Review by Leah LaPerle Larkin, "Product Review: Genome Mate Pro," *Journal of Genetic Genealogy* 8(1) 2016: xviii-xxiii (https://jogg.info/pages/vol8/editorial/larkin/Larkin-GenomeMatePro.pdf).

Original publication:
Wayne, Debbie Parker. "Genetic Genealogy Tools to Achieve DNA Discoveries." *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).

Debbie Parker Wayne, CG, CGL, is experienced using DNA analysis as well as more traditional techniques for genealogical research in Texas, the South and the West. She coordinates the genetic genealogy courses at several genealogy institutes and is the Texas State Genealogical Society's DNA Project director. She wrote the NGS *Continuing Genealogical Studies: Autosomal DNA* course and co-authored the first genetic genealogy workbook, *Genetic Genealogy in Practice*, published by NGS. See http://debbiewayne.com/ for more information.

Original publication:
Wayne, Debbie Parker. "Genetic Genealogy Tools to Achieve DNA Discoveries." *Stirpes* 56 (Sep 2017): 6–9; *DNA Tools, Methods, Experiences* theme issue; online (http://debbiewayne.com/pubs/pub_stirpes_201712_tools.pdf).