Genetic Genealogy Journey
Just Fishing or Targeted Testing

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

Fishermen exist all along a spectrum of enthusiasm. On one end, we have the fisherman who is "just fishing," who relaxes by dropping a pole into any stream and waiting for any kind of fish to bite whatever is on the line. At the opposite end, we have the competitive fisherman with "passion for the sport." He knows what kind of fish he wants, how big it should be, the spot in the lake where this type of fish is caught, which weight line is needed, and which bait attracts the fish.

Genetic genealogists cover a wide spectrum of enthusiasm as well. On one end, we have the test-taker who is just fishing, who takes a DNA test because it is new or someone asks him to test. He waits to see what pops up in the match list that seems interesting. He may not do much confirmation of the suggestions made by the testing company or the trees provided by other test-takers. At the opposite end, we have the passionate genetic genealogist who spends all day sorting and analyzing DNA test results in a spreadsheet, comparing family trees, and confirming the trees of DNA matches with documentary research. He knows that the DNA test results of targeted relatives can provide evidence to focus on the most likely family lines where a common ancestor with a new DNA match can be found. Just fishing for DNA cousins may be as much fun as some want. The passionate genetic genealogist spends much more time analyzing test results, and obtains more evidence to confirm the family tree and guide new research based on the DNA clues. Many tools help the genetic genealogist perform differing levels of analysis on the spectrum from just fishing to full-blown passion, even obsession.

Fishing

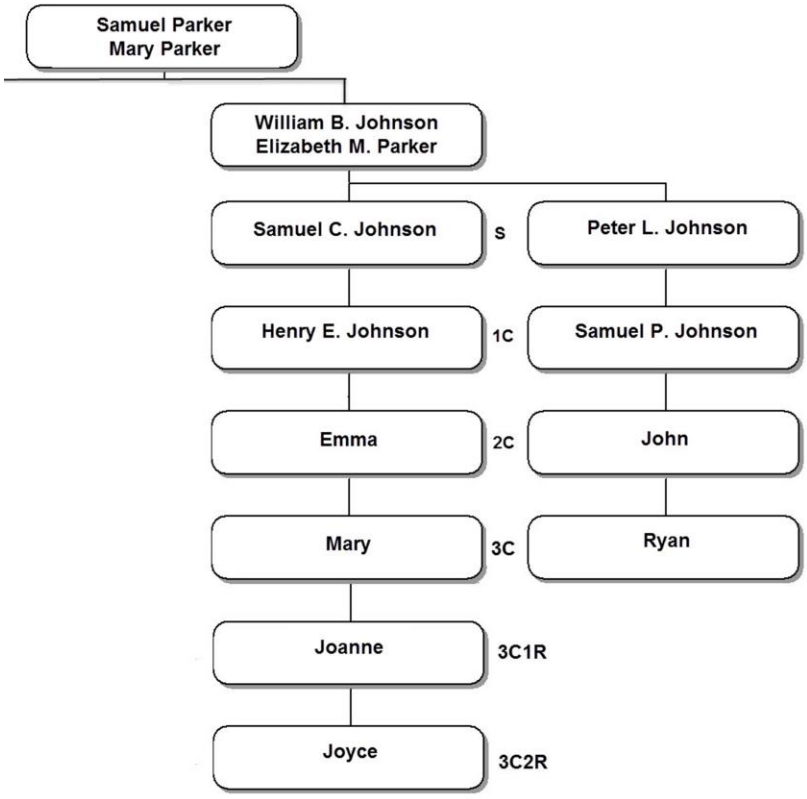
Take the example family shown in figure 1. A new DNA match named Ryan appears in the list of a test-taker named Joyce. Joyce compares her family tree to the tree of Ryan. Joyce and Ryan are descended from two different children of William B. Johnson and Elizabeth M. Parker. If both trees are accurate, Joyce is a third cousin, twice removed to Ryan. The Most Recent Common Ancestor or ancestral couple (MRCA) between Joyce and Ryan is William and Elizabeth.

The genetic genealogist who is "just fishing" may be happy to see a name on a DNA match list and confirm that match has a common ancestor named in his family tree. This is one of the lowest levels of credibility with genetic genealogy. Each step beyond the name comparison can provide more evidence leading to a more credible conclusion. The exact order of the steps given here can vary in different situations, but each step adds more evidence to support or refute the genealogical hypothesis.

First published: Debbie Parker Wayne, "Just Fishing or Targeted Testing," National Genealogical Society *NGS Magazine* 42 (April-June 2016): 40-43.

Confirm the Family Tree

Beyond the common ancestor named in the trees, the next step could be confirming the documentary research to ensure both trees are accurate. Joyce should confirm the accuracy of Ryan's tree and her own tree. If Ryan does not have a well-documented tree, it may be necessary for Joyce to research his tree to prove the kinship links. Even if Ryan's tree were well-documented, Joyce would be wise to analyze the cited documents and ensure the evidence of kinship has been properly interpreted. Ryan may have less experience as a researcher or have made an error during data entry resulted in a missing generation or an incorrect parental link in the GEDCOM file uploaded to an online site. Or, perhaps Ryan's tree is accurate and allows Joyce to correct an error in her own tree. Incorrect trees can prevent accurate interpretation of the DNA test results.



Amount of Shared DNA

More data, correctly interpreted, helps ensure our conclusions are accurate. Relationships beyond second cousin can sometimes be more difficult to confirm with DNA. Random recombination and inheritance patterns mean that cousins that are more distant sometimes share little or no DNA at a level that meets a company's threshold. When DNA is shared between two distant cousins, we need more evidence to confirm this is a true match and not a false positive. Each testing company uses different algorithms and thresholds designed to reduce false matches.

First published: Debbie Parker Wayne, "Just Fishing or Targeted Testing," National Genealogical Society *NGS Magazine* 42 (April-June 2016): 40-43.

Genetic genealogy is a relatively new field of study. As the field matures and there is more real family data for comparison, the algorithms and thresholds may be adjusted to reduce false positives (someone listed as a match who is not a biological relative) as well as false negatives (someone not on a match list who is known to be a cousin).

The next step could be determining if the amount of shared DNA is consistent with the actual relationship between the test-taker and the DNA match. The *ISOGG Wiki* documents statistically predicted amounts of shared DNA for specific relationships; the Shared cM Project documents percentages seen in real families.[1]

Joyce should compare the total amount of DNA indicated in her match list as shared with Ryan. Assume that Joyce and Ryan share 40.59 total centimorgans (cM) of DNA. The actual relationship in our example is third cousins, twice removed. Third cousins, twice removed share, on average, 13.28 cM with a range of 16.4 to 122 cM seen in the Shared cM Project.[2] When counting only the segments of 5cM or larger, Joyce and Ryan share 20.31cM. Based on the amount of data available today for known family relationships, many genetic genealogists believe that the total amount of shared DNA reported should be reduced by eliminating segments smaller than 5cM from the total count.

The total shared DNA between Joyce and Ryan is more than statistically expected, but within the range reported in the Shared cM Project. If either person's family tree is incorrect or has blank places where an ancestor is unknown, there might be more than one common ancestor between these two. In this case, we may know that cousins Samuel and Mary Parker were Elizabeth's parents, resulting in more shared DNA than expected between descendants. Joyce's written conclusion should explore and discuss the resolution of this potential conflict.

Multiple Cousins Descended from the Same MRCA

The next step could be to gather more data by finding additional cousins who share DNA with Joyce and/or Ryan and who share the same common ancestor. A search of the match list for both might indicate another DNA match naming William B. Johnson and Elizabeth M. Parker or an ancestor of William or Elizabeth. Joyce or Ryan needs to be lucky enough to have a cousin who has tested, who shares DNA, and who names William, Elizabeth, or one of their ancestors in an online tree.

Even if none of the DNA matches uploaded a family tree, the testing company may provide tools to locate test-takers who share DNA with Joyce and Ryan. Family Tree DNA offers a chromosome browser, "In Common With (ICW)," and "Matrix" tools. AncestryDNA offers "Shared Matches," "Shared Ancestor Hints," "DNA Circles," and "New Ancestor Discoveries (NAD)" tools. 23andMe is in the process of changing their offerings, but should continue to offer

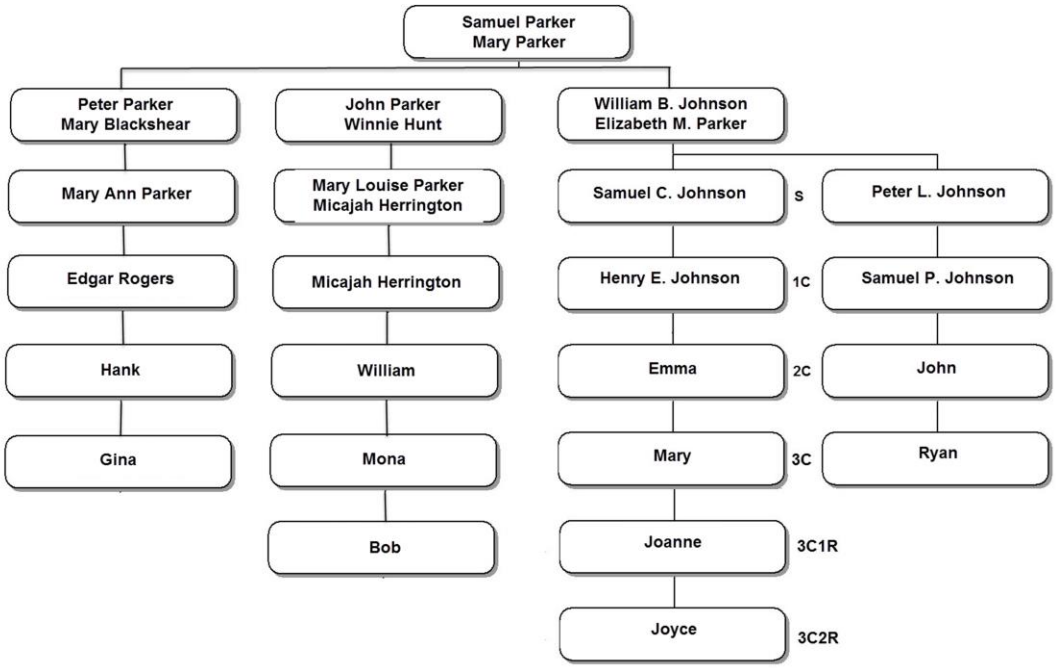
[1] "Autosomal DNA statistics," *ISOGG Wiki* (http://isogg.org/wiki/Autosomal_DNA_statistics). The statistical predictions are based purely on mathematical averages. The Shared cM Project records self-reported shared amounts of DNA seen in real family projects. The accuracy of the Shared cM Project depends on users making no typos during data entry and knowing the correct relationships of two test-takers.

[2] "Autosomal DNA statistics," *ISOGG Wiki* (http://isogg.org/wiki/Autosomal_DNA_statistics).

First published: Debbie Parker Wayne, "Just Fishing or Targeted Testing," National Genealogical Society *NGS Magazine* 42 (April-June 2016): 40-43.

a chromosome browser. GEDmatch, DNAgedcom, Genome Mate Pro, and other third-party groups offer tools that help find additional cousins related on the same ancestral lines.[3]

To determine which people in a database share DNA with both Joyce and Ryan, Joyce could select Ryan and run Family Tree DNA's ICW tool or AncestryDNA's Shared Matches tool. A new list names those in the company database who share enough DNA with both Joyce and Ryan to meet a threshold. In our example, Bob and Gina are on the ICW list and both have a family tree indicating descent from siblings of Elizabeth Parker as shown in figure 2. Bob is a fifth cousin, once removed to Joyce. Gina is a fourth cousin, twice removed from Joyce.



Family Tree DNA's Matrix tool allows one additional comparison. Up to ten people can be selected from those who we believe match both Joyce and Ryan. A grid will indicate which of the selected persons shares DNA with each of the others. While logged into Joyce's account, Ryan, Bob, and Gina are selected in the matrix tool. A display similar to table 1 is generated. The pale gray blocks indicate the same person is named in the row and column. Here we use Xs to indicate whether two of the test-takers are on each other's match list. On the Family Tree DNA website this is done with colored blocks.

Table 1. Matrix display of Joyce			
	Ryan	Bob	Gina

[3] "Autosomal DNA tools," *ISOGG Wiki* (http://isogg.org/wiki/Autosomal_DNA_tools).

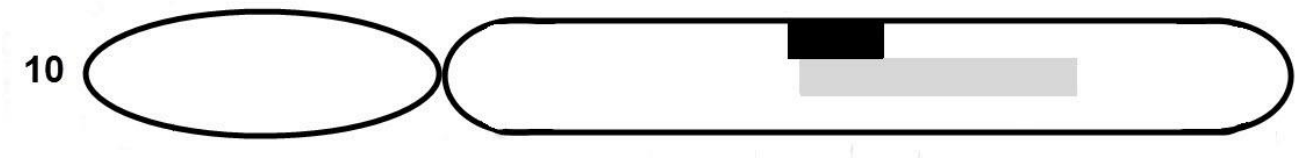
First published: Debbie Parker Wayne, "Just Fishing or Targeted Testing," National Genealogical Society *NGS Magazine* 42 (April-June 2016): 40-43.

Ryan		X	
Bob	X		X
Gina		X	

Joyce is not shown in the display, but the genetic genealogist should always remember to include the account owner. From the matrix we see that Joyce shares DNA with Ryan, Bob, and Gina (because we are logged into her account). Ryan shares DNA with Joyce and Bob, but not Gina. Bob shares DNA with Joyce, Ryan, and Gina. Gina shares DNA with Bob and Joyce, but not Ryan. This could mean Gina is related to the others through a different common ancestor. Assuming Gina's family tree in figure 2 is accurate, it is likely that Gina and Ryan inherited different DNA segments from the common ancestors or random recombination may have broken the DNA into small segments that do not meet the company threshold. From the matrix display, we cannot determine if any or all of the cousins share the same or an overlapping DNA segment. All we know is that one person appears on the match list of the other indicating they share some DNA meeting the company threshold when the corresponding block contains an X (or is colored on the company website).[4]

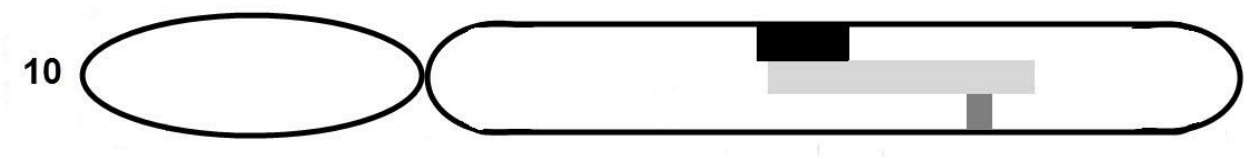
Shared Segments

The next step could be to view the shared segments in the chromosome browser. Figure 3 shows a shared segment on chromosome 10.[5] The threshold is set to a minimum segment size of 5 centimorgans (cM). The black segment shared by Joyce and Ryan is 11.78cM long, starting at location 79.1 million (rounded), ending at location 89.6 million. The light gray segment shared by Joyce and Bob is 30.96cM long, starting at location 80.3 million, ending at location 112.9 million. However, if the display threshold is set to 1cM we see the display shown in figure 4. Segments smaller than 5–7cM are not generally used in analysis of relationships, but here it shows how a segment might have been lost due to random recombination. Shown in dark gray, Gina and Joyce share a tiny portion of the same segment that Joyce shares with Ryan and Bob: a 1.44cM segment starting at location 105.7 million, ending at location 107.6 million. Random recombination could have eliminated most of the segment in Gina that seems to have been inherited by the other cousins from the common ancestors.



[4] This is one situation where it can be advantageous to convince a cousin to upload their DNA data to GEDmatch.com where we can do a comparison using lower thresholds than the testing company uses.
[5] There are additional shared segments between the cousins in this example. Only one segment is shown here.

First published: Debbie Parker Wayne, "Just Fishing or Targeted Testing," National Genealogical Society *NGS Magazine* 42 (April-June 2016): 40-43.



What Next

The next steps could be triangulation to determine whether Bob, Gina, and Ryan share the same segment with each other that each shares with Joyce. Targeted testing of relatives may also help with the analysis. These will be continued in a future column.

Conclusion

Random recombination and inheritance may mean some DNA is not shared by all cousins even when test-takers share common ancestors. More data leads to more credible conclusions.

Debbie Parker Wayne, CG, CGL, is experienced using DNA analysis, as well as more traditional techniques, for genealogical research in Texas, the South and 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. See http://debbiewayne.com/ for more information.

First published: Debbie Parker Wayne, "Just Fishing or Targeted Testing," National Genealogical Society *NGS Magazine* 42 (April-June 2016): 40-43.