

Genetic Genealogy Journey

DNA Projects

by Debbie Parker Wayne, CGSM, CGLSM

Genealogy can be a solitary pursuit. Genealogists sometimes collaborate to work on common lines, but lone researchers can perform documentary research without collaborating with others. Genetic genealogy, on the other hand, requires active collaboration to make the best progress towards research goals.

Participation in a DNA project is one method of collaboration. Several types of projects may help us find information to apply to our genealogical questions. Projects may be specific to a family surname, shared ancestral locations, shared DNA haplogroups, or descendants of a specific person or couple. Large society-wide projects may also provide information that can be applied to genealogy as well as to deep ancestral origins. We still have so much to learn about DNA and how we may be able to use it in the future. We contribute to that knowledge base by participating in projects.

What is a DNA project?

A DNA project is simply a group of test-takers with some DNA in common who share DNA information and genealogy information. The shared information advances the knowledge of links between the group members. That knowledge may be directly applicable to our own family research or it may contribute to a more global understanding of our story as humans.

Finding a DNA project

This article lists some of the advantages in joining projects; many more advantages exist than can be covered in this space. Some examples of project types are also described. An Internet search will turn up projects not named here that may be significant to test-takers. See figure 1.

Of the companies currently used most often for genetic genealogy testing, only Family Tree DNA currently supports DNA projects. A Project Search box allows users to find appropriate DNA projects. Changing the search type from Equals to Contains often returns more relevant project names.¹ See figure 2.

Why join a DNA project?

While we might think of our individual match lists as a mini-project, we can get distinct advantages by joining a formal project. Project administrators are often the most knowledgeable researchers on the subject of the project whether that subject is a family or surname, a Y-DNA haplogroup, an mtDNA haplogroup, or a regional area. Project administrators often have access to tools for comparing everyone in the project that are not accessible to us as individual test-

Websites were accessed on 15 November 2016.

¹ "Surname & Geographical Projects," *Family Tree DNA* (<https://www.familytreedna.com/projects.aspx>).

takers. These two factors make a project administrator one of our best research avenues when using genetics for genealogy. In addition to access to project administrators, project members often have access to more information on the shared DNA with their matches within a project. As an example of this additional information refer to the Y-DNA STR² discussion below, in particular figure 4 where the STR marker values are shown as compared to figure 3 where only the genetic distance is shown.



Figure 1. Google search terms to locate a Parker surname project

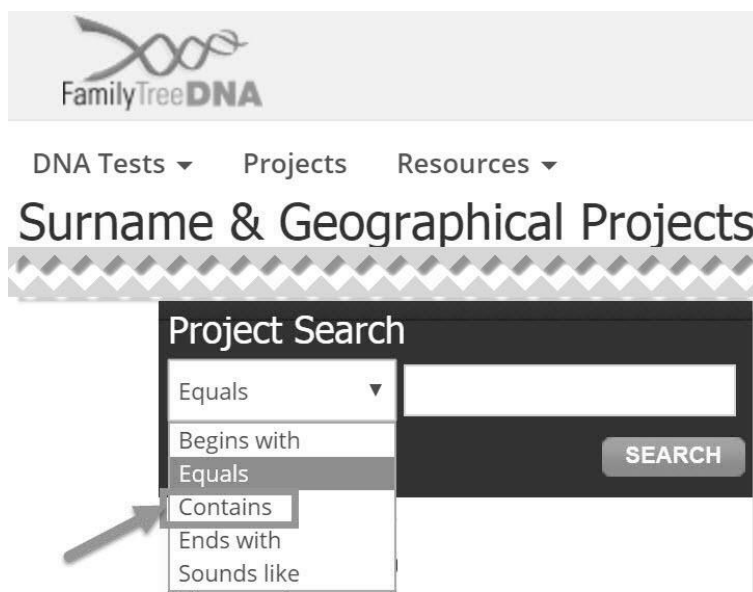


Figure 2. Family Tree DNA project search box

² "Short tandem repeat," *ISOGG Wiki* (http://isogg.org/wiki/Short_tandem_repeat).

Y-DNA surname projects

Surname projects were the first types of DNA projects formed. Test-takers must be male as only men have a Y chromosome. Females can sponsor a male family member to take the test to learn about Y-DNA lines.

Surname projects started with Y-DNA STR marker test results and some now incorporate Y-DNA SNP³ test results. Project administrators typically sort test-takers into groups of those who share marker values and likely share a common patrilineal line, even when the documentary evidence has not yet named the common ancestor. The project members with close DNA matches can then collaborate on documentary research to locate the common ancestor. Clues provided by project members whose family trees are more complete can guide the collaborators to productive research avenues.

The match list of a test-taker with Y-DNA STR results names the person with matching DNA and indicates the genetic distance to the match, but does not list the specific marker values for comparison. See figure 3. In-depth analysis requires the names and values of the markers for each test-taker under analysis. The marker names and values are usually available to project members as well as project administrators. See figure 4. Without access through the project, each test-taker on the match list must be individually contacted and must choose to share test results before in-depth analysis can be done.

| Genetic Distance | Name | Most Distant Ancestor | Y-DNA Haplogroup |
|------------------|---|---|------------------|
| 3 | H (Identifying name information has been deleted to preserve privacy.) Parker | Lewis Parker, b. about 1772 and d. about 1859 | R-M198 |
| 3 | M ight | | R-M512 |
| 3 | J Parker | Henry Parker, Sr., b 1782 SC, d 1859 AR | R-M512 |
| 3 | D Parker | James Parker | R-M198 |
| 4 | C D ights | | R-M512 |

Figure 3. Sample Y-DNA match list for test-taker showing genetic distance

³ “Single-nucleotide polymorphism,” *ISOGG Wiki* (http://isogg.org/wiki/Single-nucleotide_polymorphism).

| Kit # | Paternal Ancestor Name | Haplo-group | D | D | D | D | D |
|--------|---|-------------|----|----|----|----|-------------|
| | | | Y | Y | Y | Y | Y |
| | | | S | S | S | S | S |
| | | | 3 | 3 | 3 | 4 | 4 |
| | | | 9 | 9 | 1 | 4 | 6 |
| | | | 3 | 0 | 9 | 9 | 4 |
| 286924 | John Parker d. 1788 Duplin Co., NC | R-M512 | 13 | 25 | 15 | | |
| 26171 | | R-M512 | 13 | 25 | 15 | 30 | 12-15-16 16 |
| 116364 | Fredrick Nance Parker, b. 1866, Springfield, IL | R-M198 | 13 | 25 | 15 | 31 | 12-15-15-16 |
| 40466 | | R-M512 | 13 | 25 | 15 | 31 | 12-15-15-16 |
| 214794 | Henry Parker, Sr., b 1782 SC, d 1859 AR | R-M512 | 13 | 25 | 15 | 31 | 12-15-15-16 |
| 29635 | Henry Parker Sr., 1780 S.C.-abt. 1859 Ar. | R-M512 | 13 | 25 | 15 | 31 | 12-15-15-16 |

Figure 4. Sample Y-DNA surname project list showing STR marker values

Y-DNA SNP projects and Y-DNA and mtDNA haplogroup (lineage) projects

Some projects are focused on specific haplogroups⁴ and some on only those haplogroup members with a specific SNP value. Haplogroups place a test-taker on a branch of the Y-DNA tree or mtDNA tree. Haplogroups are defined by SNPs. Haplogroup projects contribute to knowledge of our ancient ancestry. Exact and close matches may also indicate relationships within a genealogical timeframe (generally defined as the timeframe for which we can find documentary confirmation of relationships). Haplogroup projects are sometimes named “lineage projects” in the project list at Family Tree DNA.

Years ago, Y-DNA SNPs were thought to be useful only for deep or ancient ancestral origins. Advances in the science now show Y-DNA SNPs can be useful in a genealogical timeframe. Separate SNP projects exist and SNPs are also incorporated into some surname projects to refine the matches defined by STR markers. Many researchers today believe both STR and SNP markers should be used to obtain the most reliable information using Y-DNA.

For mtDNA, the full mitochondrial sequence test is often required before joining a haplogroup project. Because mtDNA mutates less often than Y-DNA STRs, more complete testing provides more information that may be useful in a genealogical timeframe. The full mitochondrial sequence provides access to every base pair within the mitochondrial DNA molecule. While many project members have taken a full mitochondrial sequence test, public result pages only display the results for Hyper-Variable Region 1 (HVR1) and HVR2. The coding region on the mitochondrial molecule contains genes that could reveal medically-relevant information so it is not publicly displayed. The “U5 mtDNA” project, for example, focuses on test-takers whose mtDNA is in haplogroup U5.⁵

Geographic and state-level DNA projects

⁴ “Haplogroup,” *ISOGG Wiki* (<http://isogg.org/wiki/Haplogroup>).

⁵ Anne M. Berge and Gail Tonnesen, project administrators, “U5 mtDNA,” *Family Tree DNA* (<https://www.familytreedna.com/groups/u-5b/about>).

Many projects exist for specific geographic areas. For example, there are separate Y-DNA and mtDNA projects for families “whose ancestors settled in or passed through the counties surrounding the Cumberland Gap” encompassing counties in Tennessee, Virginia, and Kentucky.⁶ Many other regions have similar projects.

Several US states have formed projects using DNA to learn more about their pioneer settlers. Only a few are described below. An Internet search for the state of interest and the words “dna project” may reveal a beneficial project.

New Mexico has a project exploring Y-DNA and mtDNA test results and genealogies to “better understand New Mexican ancestry.”⁷

The Iowa DNA Project is “for those who have direct ancestry in Iowa, as well as those researching collateral lines that lived in Iowa.” The project focuses on autosomal DNA while also reporting on Y-DNA and mtDNA origins. Project statistics indicate there are twenty-six database connections, on average, between project members and 989 Iowa surnames are included.⁸ This state project was one of the first to use the graphic tools offered by DNA Genealogy Experiment (dnagen.net)⁹ to map the interconnections between project members.

The Early Texans DNA Project formed by the Texas State Genealogical Society will use autosomal DNA, Y-DNA, and mtDNA to explore the settlers into the state through the end of the nineteenth century. The project is in its infancy. The long-term goal is to map chromosome segments to particular ancestral couples that can provide clues for new researchers on potential ancestral lines. The short-term goal is to explore the origins of the Texas settlers by comparing haplogroups and ethnic origins to scientific databases.¹⁰ Most Texas projects focus on the settlers in Texas prior to statehood (19 February 1846); the Early Texans DNA Project accepts descendants whose ancestors were in Texas by 31 December 1900.

Family autosomal DNA projects

As autosomal DNA testing becomes more popular, many families have formed projects open only to those descended from a named ancestral couple. Such a project allows the project administrator to confirm triangulated DNA segments without contacting each test-taker individually. An astute researcher may also be able to see likely relationships between descendants who did not inherit enough DNA from an ancestral couple to meet the threshold to be on each other’s match list.

⁶ Janet Crain, et al., project administrators, “CumberlandGap-mtDNA,” *Family Tree DNA* (<https://www.familytreedna.com/groups/cumberlandgap-mtdna/about>) and “CumberlandGap-YDNA,” *Family Tree DNA* (<https://www.familytreedna.com/groups/cumberland-gap-ydna/about>).

⁷ Miguel A. Tórréz, “The New Mexico Genealogical Society’s DNA Project,” *New Mexico Genealogical Society* (<http://www.nmgs.org/dna>).

⁸ Lori Alexander, Project Administrator, “Iowa DNA Project Home Page,” *Iowa DNA Project* blog, 24 August 2016 (<https://iowadnaproject.wordpress.com/>) and “Iowa DNA,” *Family Tree DNA* (<https://www.familytreedna.com/groups/iowa/about>).

⁹ Göran Runfeldt and Staffan Betné, “Family Finder Graphs,” *DNA Genealogy Experiment* (<https://dnagen.net>).

¹⁰ Debbie Parker Wayne, DNA Project Chair, “Early Texans DNA Project,” *Texas State Genealogical Society* (<http://www.txsgs.org/programs/dna-project/early-texans>). This project is managed by the author.

Research DNA projects

Many academic and scientific research projects enable researchers to contribute to understanding DNA beyond their personal genealogy. These projects include

- The Genographic Project (<https://genographic.nationalgeographic.com/>), studying human origins and deep ancestry (sponsored by the National Geographic Society);
- DNA Land (<https://dna.land/>), providing tools for test-takers to learn more about their own DNA while sharing the data with scientists to enable new discoveries (a collaborative effort of the New York Genome Center and Columbia University);
- Personal Genome Project (<http://www.personalgenomes.org/>), inviting participants to share their full genome sequence in an open data network accessible by scientists and citizen scientists (sponsored by Harvard Medical School).

Limitations of projects

DNA projects do not solve every problem related to DNA analysis. Some test-takers do not understand the usefulness of projects. Sometimes a close DNA match is not willing to join a specific project. Each researcher must still do analysis of the DNA data, but more in-depth data is usually accessible.

Due to limited screen space, DNA projects often list only the most distant known ancestor in a particular line, especially the Y-DNA line or the mtDNA line where only one ancestor in each generation contributed to a test-taker's DNA. It would be much more helpful to researchers if the projects could display the entire lineage. For Y-DNA, display of the full lineage would be helpful for test-takers who are descended from a common ancestor, but one of the test-takers has not researched the line as far back as the other. When the line is more recent the common ancestor's name may not be recognized. For mtDNA, display of the full lineage would be helpful because in most western cultures the female's surname changes with marriage.

Conclusion

Many DNA projects may be available to help test-takers learn more about applying DNA test results to recent genealogical research and to their deep ancestry thousands of years ago. Explore the available projects to find those most useful for the research goals.

Additional information on joining projects can be found in blog posts¹¹ and in the Family Tree DNA Learning Center.¹²

¹¹ Roberta Estes, "Surname Projects," *DNA eXplained—Genetic Genealogy* blog, 7 August 2014 (<https://dna-explained.com/2014/08/07/surname-projects>).

¹² Frequently Asked Questions, Group Projects, *Family Tree DNA* (<https://www.familytreedna.com/learn/dna-group-projects>). This now has the title "The Family Tree DNA Learning Center BETA"

© 2017, Debbie Parker Wayne, All rights reserved.

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.