Y-DNA

The value of Y-DNA genealogical tests lies in their ability to provide information from that one direct line of paternal inheritance beyond the five to six generations covered by autosomal DNA with no recombination to obscure the connection.

Y-chromosomal DNA passes only from father to son, and only part of the Y chromosome recombines (mixes) with the X chromosome which contains DNA from other ancestors. Testing avoids that portion of the chromosome, making Y-DNA uniquely useful in tracing paternal ancestry beyond the five to six generations that autosomal DNA covers. The Big Y expands and refines the reach of Y-chromosomal testing and helps to isolate the family's specific branch on the Y-genetic tree, regardless of the tester's surname.

Useful terms

- **Sequenced DNA test:** Explores mutations on a segment of DNA as opposed to looking for specific mutations at specific locations.
- **Base pair:** Match-up of nucleotides adenine (A), cytosine (C), guanine (G), tyrosine (T) that comprise DNA along the double-helix strand, such as A-T, C-G.
- **Mutation:** change from one nucleotide base to another.
- **Phylogeny:** Lines of descent or relationship showing evolution of a group or species.
- **STRs - Short Tandem Repeats:**
    - An STR is the number of times a sequence of nucleotides repeats at a specific location on a chromosome. In genetic genealogy, usually refers to Y-DNA.
    - In Y-DNA traditionally used for "close" matching to testers sharing a paternal ancestor.
    - Number of STR differences between two Y-DNA testers known as "genetic distance," a measure of how closely testers may be related.

- **SNPs - Single Nucleotide Polymorphisms:**
  - A SNP is a change (mutation) from one base (i.e., G-C, A-T) at a single location on the chromosome.
  - Phylogenetic position on the Y-DNA haplotree defines haplogroups and subclades.
  - Corresponds to migratory paths in Y and mitochondrial DNA.
  - Analyzed in all four types of DNA (mtDNA, Y-DNAY-DNA, X-DNA and autosomal DNA).
- **Synonymous SNPs:** In Y-DNA, same mutation with two or more different names.
- **Equivalent SNPs:** In Y-DNA, different SNP name, different mutation, but on the same Y-DNA haplotree branch. Not enough information to place on a different branch yet.
- **Reoccurring SNPs:** Same Y-DNA SNP mutation, same SNP name, different places on the haplotree.
- **Ancestral:** The original state at a location on the chromosome, no mutation has occurred. The tester is negative for the SNP.
- **Derived:** A change from one base to another at a location on the chromosome. A mutation has occurred, and the tester is positive for the SNP.

STRs - Short Tandem Repeats

  - "Stutter" sequences in Y-chromosomal, autosomal, and X-chromosomal DNA.
  - Y-STRs useful for "close" matching when tested in panels.
  - **Levels:**
    - 37- and 111-marker STRs.
    - 12-marker available in group projects or through FTDNA Customer Service, 67-marker available through FTDNA Customer Service.
    - 25-marker discontinued in January 2020.
    - Big Y-500 Discontinued but results still display, and additional STRs were added in 2020 to meet 700 levels. Includes minimum 389 additional STRs out of potential 727.
    - Big Y-700 - Minimum 589 additional out of potential 727.
- **Genetic Distance**
  - The number of STR mismatches is called Genetic Distance (GD).

- Different GD threshold at each level for matches to be displayed.
    - 12-marker matches must be exact, though a GD of 1 displays if in same project.
    - To show at 25 – GD 2, 37 – GD 4, 67 – GD7, 111 – GD 10.
- Typically, the higher the genetic distance, the more distant the relationship.
- Typically, comparisons beyond GD thresholds are not as genealogically relevant with some exceptions such as nulls and deletions.
- STRs in Big Y-500 and 700 not used currently but ar used for automated matching, must be manually compared between testers.

SNPs

- Individual SNPs often available for purchase.
- SNP Packs - Curated groups of SNPs from a specific haplogroup or subclade but are being phased out at FTDNA so are not being updated.
- Big Y and other NGS sequenced tests.

- **Haplogroups**
    - Haplogroups are broad categories of migration patterns identified by letter designations, combinations of letters and numbers, or named for the SNP that defines that branch of the tree of humankind. For example, R1b is identified as R-M269.
    - Each one has many subgroups that shared common ancestors thousands of years ago but diverged from each other in more subtle ways that we can tell apart.
    - Y-STRs used to predict haplogroup.
        - FTDNA uses Y-12 matches to predict basic haplogroup.
        - Administrators or citizen scientists use STR patterns to predict haplogroup.
    - Y-SNPs used to confirm haplogroup and determine subclade.
    - FTDNA SNP Assurance Program (SAP) guarantees that if haplogroup cannot be determined with 100% confidence, a backbone test will be run to ascertain the correct haplogroup. Not to be confused with SNP packs labeled as "backbone."
    - Haplotrees (both Y and mtDNA) are formed from phylogenetically arranged SNPs from most distant to youngest along each branch.

FILTER MATCHES

Show Matches For: [The Entire Database ▾] Markers [111 ▾] Distance [All ▾] Matches Per Page: [25 ▾] Display Only Matches With Big Y: ☐
1   2   3   4
Last Name Starts With: [_____] (Optional) New Since: [_____] [Run Report]

111 MARKERS - 24 - MATCHES

Genetic Distance ↑	Big Y STR Differences	Name	Earliest Known Ancestor	Y-DNA Haplogroup	Terminal SNP	Match Date
0	4 of 650	: Rhinehart 🔲🔲🔲🔲 Y-DNA111 FF Big Y	John Adam Rinehart, b. 1739 and d. 1818	R-BY39278	BY39278	12/19/2017
1	3 of 597	Nixon 🔲🔲🔲🔲 Y-DNA111 FF Big Y	William Nixon b. 1679 Bew Castle, England	R-BY177482	BY177482	12/4/2016
1	4 of 640	Johnson 🔲🔲🔲🔲 Y-DNA111 FF Big Y	Elijah S. Johnson, b. 1850 (MO), d. 1929 (TX)	R-BY140649	BY140649	6/5/2016
1		Cloud 🔲🔲🔲🔲 Y-DNA111		R-M269		8/1/2018

STR Match Page – numbers below correspond to numbers on image above

1. Compare your matches against the entire database or an individual group.
2. Select match level. Do not forget the lower levels, since some people may not have tested to the same level you have.
3. Select a specific Genetic Distance.
4. Check the box to view only testers who have taken Big Y.
5. First column is Genetic Distance in that level of STRs, up to 111.
6. Second column is differences in only Big Y STRs, from 111 and up, showing how many mismatches and for how many markers the match had called results.
7. Columns are sortable by header. Default sort is by Genetic Distance but click the header to sort by date, haplogroup, name, etc.

When reviewing matches, remember that the Most Recent Common Ancestor (MRCA) for Y-DNA matches:

a. May be from the time before surnames were commonly used.
b. May be from cultures that used patronymic naming systems so the names may change with each generation.
c. May be misattributed paternity.
d. May not have been enough testers from that surname who have tested yet.

What is the Big Y?

An exploratory test that sequences between 16-22 million base pairs of the non-recombining portion of the Y-chromosome.

- Includes SNPs and STRs.
- Provides data to build the Y-DNA haplotree.
- Maps the migration path of the tester's direct paternal line.
- Confirms and refines haplogroup.
- Reveals SNPs unique to your family branch.
- Helps to estimate TMRCA to matches.
- Contributes to research and the advancement of science

Big Y Components

- **Named Variants** - SNPs that are already known and have been given names.
- **Private Variants** - SNPs that are newer mutations and have not been named yet.
  - Personal, family, or clan mutations.
  - When found in at least 2 or more individuals in high confidence, SNP is named and placed on the haplotree.
  - Personal, family, and clan mutations can be used similarly to STRs.
- **Y-Chromosome Browser Tool** - Displays each SNP or Position, and reads for both the Named and Unnamed Variants results.
- **Big Y Matches**
  - Provided only for SNPs at this time. STR matching stops at 111 markers.
  - Available matches, if any, displayed at the tester's terminal SNP and the four SNP/branches above it.
  - Must share 30 or fewer SNP differences (Named and Unnamed Variants combined).
  - Matches' names and contact information displayed and searchable.
  - Non-Matching Variants displayed. Currently includes variants in noisy regions. Number of Shared Variants with each match.

- **Block Tree -** Layout based on the Big Tree by Alex Williamson, volunteer administrator of the R-P312 Project. Provides alternate view of data and matches, incorporating information from the Y-DNA Public Haplotree.
  - Tester's branch shown in white.
  - Child branches shown below parent branches.
  - Sibling branches represented beside each other, below their parent branch.
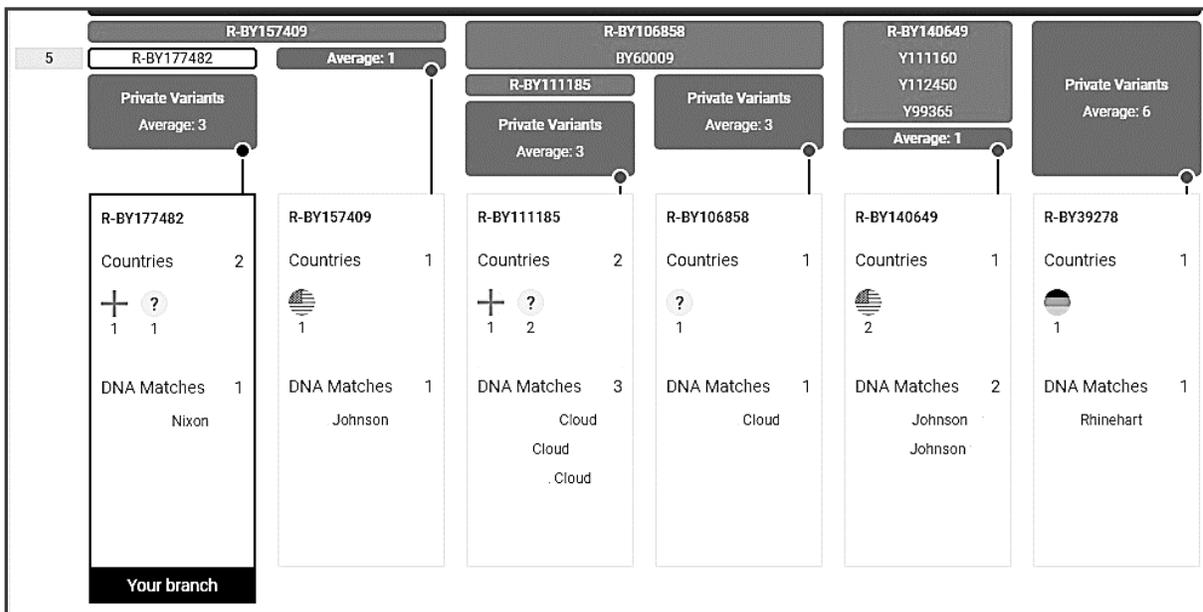
Group Projects

Surname projects, haplogroup projects, and even geographical projects can provide Y-DNA testers context and additional clues about their genetic paternal lines. Volunteer administrators run group projects, usually to further their own research, and they analyze data to determine relationships among project members as well as origins of various Y-DNA lines. Projects allow testers to pool resources and share information to which they might not otherwise have access.

- **Types of Y-DNA Group Projects**

  - Surname - Studies the origins and variants of surnames.
  - Haplogroup - Studies an entire haplogroup or specific subclades.

○ Geographical - Studies the genetics of a region varying from a county or parish to a country or continent (also dual, with mtDNA).

● **Y-DNA Projects and Your Research**
   ○ Group Project Administrators may provide:
      ■ Data comparison for grouping in project.
      ■ Data analysis within matches and within project.
      ■ Data integration with paper trail research.
   ○ Administrators and project members often help with:
      ■ Recruiting test candidates to further research.
      ■ Financing tests - Donations can be made to the Group General Fund.

<u>Resources</u>

Davis, C., Sager, M., Runfeldt, G., Greenspan, E., Bormans, A., Greenspan, B., & Bormans, C., Big Y-700 [White paper] 2019, https://blog.familytreedna.com/big-y-700-white-paper/

Biology Dictionary: https://biologydictionary.net/dna-sequencing/

Estes, Roberta. DNAeXplained http://dna-explained.com

FamilyTreeDNA Learning Center: https://learn.familytreedna.com/

FamilyTreeDNA Public Y-DNA Haplotree: https://www.familytreedna.com/public/y-dna-haplotree

McDonald, Dr. Iain. Recent human genetic anthropology:
http://www.jb.man.ac.uk/~mcdonald/genetics.html

Vance, David – DNA Concepts for Genealogy videos
 -Y-DNA part 1 https://www.youtube.com/watch?v=RqSN1A44lYU&t=16s
 - Y-DNA part 2 https://www.youtube.com/watch?v=mhBYXD7XufI&t=355s
 - Y-DNA part 3 https://www.youtube.com/watch?v=03hRXVg9i1k&t=4s

Vance, David The Genealogist's Guide to Y-DNA Testing for Genetic Genealogy: Self-published, 2020

Y-DNA Warehouse: https://ydna-warehouse.org/ and https://ydna-warehouse.org/statistics.html