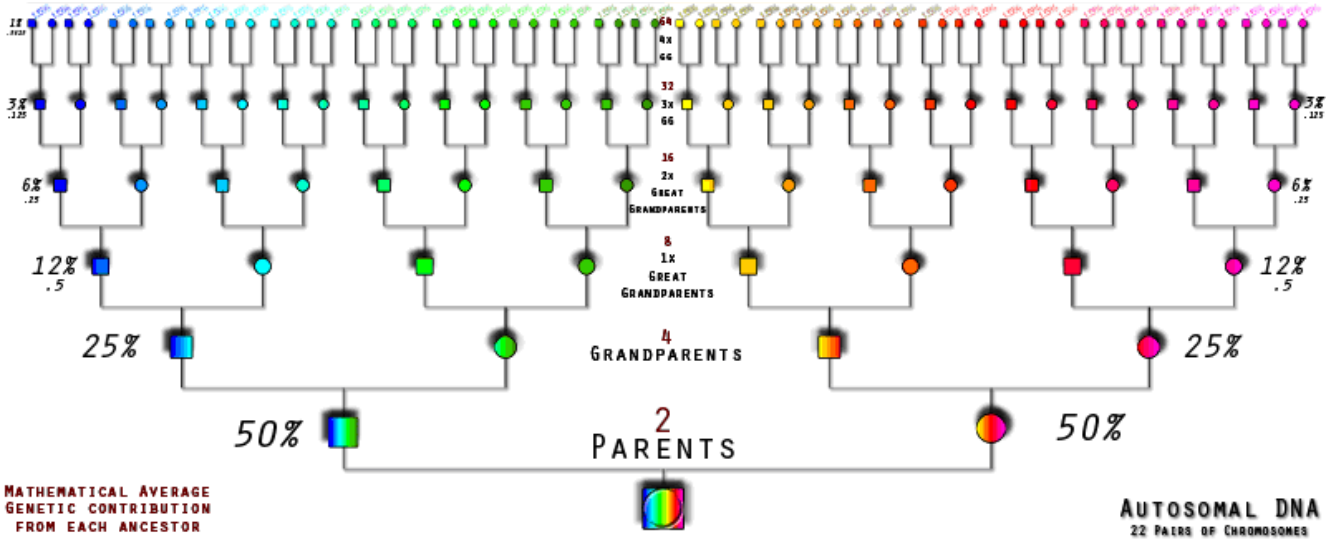
DNA in Practice

Class 1: How to Plan a DNA Research Project

Pam Holland, pam.holland@nehgs.org

Biology basics

We all have twenty-three pairs of chromosomes: twenty-two pairs of non-sex chromosomes, and one pair of sex chromosomes, XX (female) or XY (male). You get half of your DNA from each parent, 25% from your grandparents, about 12.5% from your great-grandparents, about 6.25% from your 2nd great-grandparents and so on.



Y-DNA

Only found in males since the Y chromosome is not present in females. It is used to trace your paternal ancestry and in one name studies.

Mitochondrial or mtDNA

This is not part of the chromosomes and is passed down relatively unchanged from a mother to all her children. Only her daughters (not her sons) can continue to pass the mitochondrial down to future generations. It is used to trace your maternal ancestry and confirm if two people share a maternal ancestor.

Haplogroups

If you take a Y or mitochondrial DNA test you will be assigned a paternal or maternal haplogroup. It is an ancient ancestral grouping which shows the particular migration path your ancient ancestors travelled when they left Africa.

Autosomal or atDNA

The autosomal test is the most popular DNA test taken. It examines the 22 non-sex chromosome pairs and the X chromosome on chromosome 23. It is the random recombination of DNA inherited from all your ancestors.

- The more DNA you share, the closer the relationship. Each company estimates the closeness of the relationship between yourself and a match.
- It is used in genealogy to identify relationships in the past five or six generations or to at least the 2nd great grandparent or the third cousin level.
- If a parent has tested, you can determine which chromosomes are from your mother or father.
- Test your older relatives whenever possible. This will allow you to identify matches further back in your family tree. For example, if you can test a parent or aunt or uncle you are going back five generations in their family tree which is six generations in your family tree.

X-DNA

Found in both males and females. Because of the way it is passed down X-DNA follows only certain maternal lines in your family tree. See the resource list for fan charts with these lines.

- If you have an X-DNA match you are related on the lines shown on the male or female chart.
- However, the converse is not true. If you do not have an X-DNA match it could mean a distant ancestor's X-DNA is not showing up and you might still be related on your X-DNA lines.
- All the companies test X-DNA as part of autosomal testing but not all provide access to it.

DNA at the testing companies

The testing companies use genetic markers to compare your DNA to other testers and population groups. For autosomal tests, they will report back to you:

- Your ethnicity estimates.
    - These are speculative and only as good as the reference populations they are compared to.
    - The identified regions may be from ancestors in your family that are 500 or more years ago.
- Your matches to other individuals in their database.
    - They will predict your possible relationship to a match.
    - New matches will appear as more people take tests so check back from time to time.

What DNA Can Do:

- Find relatives and cousins who have tested.
- Test a hypothesis.
- Confirm a matrilineal or patrilineal origin.
- Provide evidence of recent non-paternity events.
- Deliver unexpected results.

What DNA Can't Do:

- Solve a brick wall ancestor without any genealogical research,
- Determine a Native American tribe.
- Be positive proof of a suspected distant ancestor.

Goal	DNA Test(s) to take
Learn ethnic origins	Autosomal
Connect with genetic cousins	Autosomal
Connect with biological family	Autosomal
Identify father of a recent male ancestor (within 5 generations)	Autosomal and Y-DNA
Identify father of an historic male ancestor (5 – 15 generations back)	Y-DNA
Identify mother of a recent female ancestor (within 5 generations)	Autosomal and mtDNA
Identify mother of an historic female ancestor (5 – 15 generations back)	mtDNA
Learn health predispositions	Autosomal (23andMe)

Choosing a Testing company

- FamilyTreeDNA: https://www.familytreedna.com/ – Only company currently offering Y-DNA and mtDNA testing; their autosomal test is called Family Finder
- AncestryDNA: http://dna.ancestry.com/ - only offers Autosomal
- MyHeritage: https://www.myheritage.com/ - only offers Autosomal
- 23andMe: https://www.23andme.com/ - only offers Autosomal
- LivingDNA: https://livingdna.com/ - only offers Autosomal

	Ancestry DNA	23andMe	MyHeritage	FamilyTree DNA
Autosomal participants	+21 million	+12 million	+5.5 million	+1.2 million
Autosomal test cost	$99	$99	$89	$79
Collection method	Saliva	Saliva	Cheek Swab	Cheek Swab
Medical information option	No*	+ $100	No	No
Shared matches	Yes	Yes	Yes	Yes
Search by surname	Yes	Yes	Yes	Yes
Keep private notes	Yes	Yes	Yes	Yes
DNA download	Yes	Yes	Yes	Yes
DNA upload	No	No	Yes	Yes
Keeps sample	Yes	Yes	Yes	Yes
Linked tree (*link matches)	Yes*	No	Yes	Yes*
Ethnic communities	Yes	No	Yes	No
Ethnicity by chromosome	Yes	Yes	No	Yes
X-DNA viewable	No	Yes	No	Yes
Chromosome Browser	No	Yes	Yes	Yes
Name Projects	No	No	No	Yes
Y & mtDNA haplogroups	No	Yes	No	Yes*
Sells DNA for research with opt-in	Yes	Yes	Yes	Yes
Law Enforcement Cooperation	No	No	No	Yes

Also see https://isogg.org/wiki/Autosomal_DNA_testing_comparison_chart.

Unique Tools

	Ancestry DNA	23andMe	MyHeritage	FamilyTree DNA
Common Ancestors	Yes			
ThruLines	Yes			
Theory of Family Relativity			Yes	
Smart Matches			Yes	
AutoClusters			Yes	
Family Tree with predicted relationships		Yes		
DNA groups or labels (self-created)	Colored Dots		Colored Dots	
Search common location in match's tree	Birth location	Country	Country	

Best Practices

- Provide a well-researched and documented tree at the testing company. Make sure to include all your direct line ancestors and any identified collateral relatives.
- Try to go back 8 generations to your 6th great-grandparents where possible.
- Make sure your test is linked to your tree at Ancestry and MyHeritage.
- Update your Profile with surname and location information.
- Test close relatives and different branches of your family tree – the older the better.
- Test at multiple companies to find the most matches.

Finding other Participants

You may need to confirm or reject a hypothesis that your DNA will not answer. In that case you will need to approach potential strangers and ask them to test.

- Use your genealogical research skills to find living descendants of the target ancestor.
- Carefully explain what you want and provide lots of information.
- Offer to pay for their test.
- Be sensitive when asking and don't badger them if they decline to test.

Organizing your Results

- Keep notes for each match at the testing company.
- Create colored dot groups (Ancestry) and labels (MyHeritage).

Interpreting your Results

- Sort into ancestral lines.
    - Can first use the Leeds Method for 4 grandparent lines.
    - Work back to later generation lines.
- Identify how you are related to a match.
    - Look at their trees, shared surnames, common locations, etc.
    - Look for Common Ancestor (Ancestry) or Theory of Family Relativity (MyHeritage) hints.
    - Build out a match's tree to find a common ancestor
    - Evaluate shared matches and triangulate
- Use any company tools available like AutoClusters (My Heritage) and Family Tree (23andMe).

Genetic Genealogy Resources:

- International Society of Genetic Genealogists or ISOGG wiki: http://www.isogg.org/wiki/
- X-DNA: http://www.thegeneticgenealogist.com/2008/12/21/unlocking-the-genealogical-secrets-of-the-x-chromosome/ & http://thegeneticgenealogist.com/2009/01/12/more-x-chromosome-charts/
- Percent of DNA shared: http://www.isogg.org/wiki/File:Cousin_tree_(with_genetic_kinship).png
- Shared cMs for various relationships: http://www.isogg.org/wiki/File:Shared-cM-Project-Image-2.png also https://thegeneticgenealogist.com/2015/05/29/the-shared-cm-project/
- Paul Woodbury, Genetic Genealogy Success: Profiles and Family Trees: https://www.legacytree.com/blog/foundations-for-genetic-genealogy-success-profiles-family-trees
- Dana, Leeds, The Leeds Method: https://www.danaleeds.com/the-leeds-method/
- Triangulation
    - https://www.yourdnaguide.com/ydgblog/dna-triangulation
    - https://dna-explained.com/2019/11/06/triangulation-in-action-at-family-tree-dna/
- Blogs
    - Roberta Estes, DNAeXplained - Genetic Genealogy http://dna-explained.com/
    - Judy Russell, The Legal Genealogist http://www.legalgenealogist.com/category/dna/
    - Leah Larkin, The DNA Geek: https://thednageek.com/blog/
    - Diahan Southard, Your DNA Guide: https://www.yourdnaguide.com/ydgblog
    - Kitty cooper, Musings on genealogy and genetics http://blog.kittycooper.com/
    - Debbie Parker Wayne, Deb's Delvings http://debsdelvings.blogspot.com/
- Books and Articles
    - Blaine T. Bettinger, *The Family Tree Guide to DNA Testing and Genetic Genealogy, second edition* (Cincinnati: Family Tree Books, 2019)
    - Blaine T. Bettinger & Debbie Parker Wayne, *Genetic Genealogy in Practice* (Arlington, Va.: NGS, 2016)
    - Bryan Sykes, *DNA USA: A Genetic Portrait of America* (New York: Liveright Pub. Corps., 2012).
    - Bryan Sykes, *The Seven Daughters of Eve* (New York: Norton, 2002)
    - P. Gill, "Romanovs find closure in DNA," *Nature Genetics* 6 (1994):130-35.
    - C. C. Child, "Using mtDNA to identify remains: Twenty years in review," *American Ancestors* 14.3 (2013):55-56 (and sources cited therein).

Glossary

Autosomal DNA (atDNA) - Autosomal DNA is DNA from one of our chromosomes located in the cell nucleus. It generally excludes the sex chromosomes. Humans have 22 pairs of autosomal chromosomes and a pair of sex chromosomes.

centiMorgan - A centiMorgan (cM) is a measurement of how likely a segment of DNA is to recombine from one generation to the next. A single centiMorgan is considered equivalent to a 1% (1/100) chance that a segment of DNA will crossover or recombine within one generation.

Chromosome - A chromosome is a structure found in the nucleus of a cell that contains genetic material. Humans have 23 pairs of chromosomes: 22 pairs of autosomes and one pair of sex chromosomes.

DNA segment - A DNA segment is any continuous run or length of DNA. It is described by the place where it starts and the place where it stops.

GEDCOM file (Genealogical Data Communication) - A Genealogical Data Communication (GEDCOM) file is a special file format that was developed to provide a standard for encoding genealogical data. It is not used by most family tree software packages but most can import and export to GEDCOM format. Because of this, it is today used by many genealogists to exchange pedigree data files.

Genetic Distance - There are two meanings for Genetic Distance:

1. The number of differences, or mutations, between two sets of results. A genetic distance of zero means there are no differences in the results being compared against one another, i.e., an exact match. This is the meaning when comparing Y-chromosome DNA or mitochondrial DNA.
2. For autosomal DNA comparisons, genetic distance may refer to the size of a DNA segment. The genetic distance is then the length of the segment in centiMorgans.

Genetic Genealogy - Genetic genealogy is the use of your DNA to solve genealogy puzzles.

Genome - A genome is the entire complement of an organism's genetic material. This may refer to the DNA of a gamete, organelles (mitochondria and chloroplasts), organism, or species. The human nuclear genome is composed of 46 chromosomes (23 pairs). They contain a total of 3 billion base pairs. The human mitochondrial genome is composed of a single circular DNA sequence that contains 16,569 base pairs.

Haplogroup - A haplogroup is a major branch on either the maternal or paternal tree of humankind. Haplogroups are associated with early human migrations. Today these can be associated with a geographic region or regions.

Haplotype - The set of values for a set of DNA values. For example, the results of the Y-DNA12 test for one person is their haplotype. Two individuals that match exactly on all markers have the same haplotype.

HVR1 (Hyper-variable Control Region 1) - One of two segments of mitochondrial DNA which changes relatively quickly.

HVR2 (Hyper-variable Control Region 2) - A second segment of mitochondrial DNA which changes relatively quickly.

Identical by Descent (IBD) - IBD stands for identical by descent. This means the DNA matches because it comes from a common ancestor. IBD can refer to a single mutation or to a segment of DNA.

Identical by State (IBS) IBS stands for identical by state, meaning the DNA matches by coincidence. When two individuals share numerous individual results without being related, those results are IBS.

In Common With (ICW) -- Identifies DNA matches who share DNA with yourself and another match.

Longest Block - The Longest Block in the autosomal Family Finder test refers to the longest continuous segment of autosomal DNA that is shared between two individuals.

Marker - A physical location on the chromosome. Often used colloquially in genetic genealogy to refer to a short tandem repeat (STR). For example, "The Y-DNA67 test is a panel of 67 markers."

Mitochondrial DNA - The genetic material found in mitochondria. It is passed down from females to both sons and daughters, but sons do not pass down their mother's mtDNA to their children.

MRCA - The most recent common ancestor which you and a DNA match share.

Phasing - A tool used to separate the autosomal DNA received from the father and autosomal DNA received from the mother.

Recombination - The mixing of DNA on each chromosome that you receive from your mother and father. Different chromosomes and different parts of each chromosome are more or less likely to recombine in a single generation.

RSRS (Reconstructed Sapiens Reference Sequence) - A recent effort by scientists to represent one ancestral genome for the entire human population.

Sex Chromosome - X or Y chromosome. Normally males have one X and one Y; females have two Xs.

Short Tandem Repeat (STR) - A short DNA motif (pattern) repeated in tandem. ATGC repeated eleven times would give the marker a value or allele of 11**.**

SNP (Single Nucleotide Polymorphism) - A single nucleotide of DNA of which there are four types; (A) adenine, (C) cytosine, (G) guanine, and (T) thymine

Triangulation - Segment Triangulation: Using DNA segments you share with a known relative to identify common ancestors of other matches who share DNA on the same segments. Tree Triangulation: Comparing family trees of DNA matches to locate a common ancestor.

X Chromosome - One of the two sex chromosomes, X and Y. The X chromosome is passed down from mother to child. X is the sex chromosome that is present in both sexes, singly in males and doubly in females.

Y Chromosome - One of the two sex chromosomes. The Y chromosome passes down from father to son. Females do not receive it. Valuable for surname-based genealogy studies.