

An Introduction to DNA for Family Historians

Handout

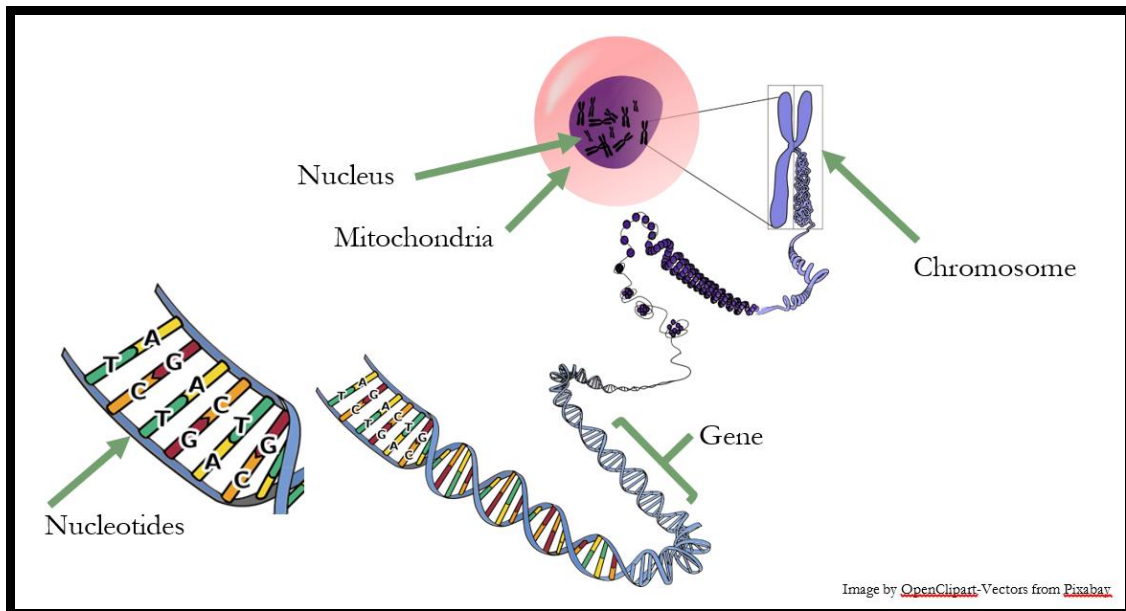
This handout supports the 'An Introduction to DNA for Family Historians' talk presented online for West Middlesex FHS on 17 August 2023.

What is DNA?

There are explanations of what is DNA at:

- BBC Bitesize: <https://www.bbc.co.uk/bitesize/topics/zpffr82/articles/zvwbcj6>
- Scitable: <https://www.nature.com/scitable/topicpage/introduction-what-is-dna-6579978/> and <https://www.nature.com/scitable/topicpage/DNA-Is-a-Structure-that-Encodes-Information-6493050/>
- BBC Science Focus Magazine: <https://www.sciencefocus.com/the-human-body/what-is-dna/>

DNA contains the blueprint for who you are. There are 23 pairs of chromosomes (46 total; half of each pair from each of your parents). 22 pairs are 'autosomal' DNA and one pair is the sex chromosomes. Men (e.g. born male) have an X and a Y sex chromosome and women (e.g. born female) have two X sex chromosomes. Each of the 23 chromosomes is made up of genes (sections of DNA that identify a specific characteristic) which are made up of nucleotides held together in a double helix formation. The chromosomes are stored within the nucleus of cells and outside the nucleus is the mitochondria. The mitochondria contains a small section of DNA that is specifically inherited from your mother. DNA testing looks at patterns in nucleotides in locations that are normally different between humans. By identifying locations that share the same DNA and examining how many of these there are, the amount of shared DNA can be assessed.



Why use DNA tests for Family History?

DNA is inherited from your biological parents. Therefore, your cousins will inherit some of the same DNA as you have inherited. This is true to second cousins and can apply to much more distant cousins (up to about 8th cousins) although not in all cases and in much smaller amounts. However, this means that you can identify genetic cousins with a DNA test who you should share a common ancestor with. By using the information from DNA tests and traditional genealogy research, it is possible in many cases to work out who the common ancestors are.

DNA testing provides you with three types of results:

- DNA matches – these are genetic relatives who share some of your DNA and have also taken a DNA test
- Ethnicity estimates – these provide an estimate of where your origins are from in the past ~1000 years
- Ancient origins (haplogroup) – this provides a prediction of your deep origins on either your male or female direct line (not available on all tests)

An Introduction to DNA for Family Historians – Handout

The DNA matches allow you to identify common ancestors which in turn:

- Helps verify known family relationships
- Expands your family tree to living people
- Helps break down brickwalls
- Provides new avenues for research

Remember the reason people take a DNA test is not solely to help research their family history. They may be interested in discovering their heritage or trying to find biological parent(s). This needs to be taken into account when interacting with your DNA matches. Often now the taking of a DNA test is the starting point for doing family history research.

Types of DNA

There are three different types of DNA test:

- Autosomal
- Y chromosome (Y-DNA)
- Mitochondrial (mtDNA)

Autosomal DNA

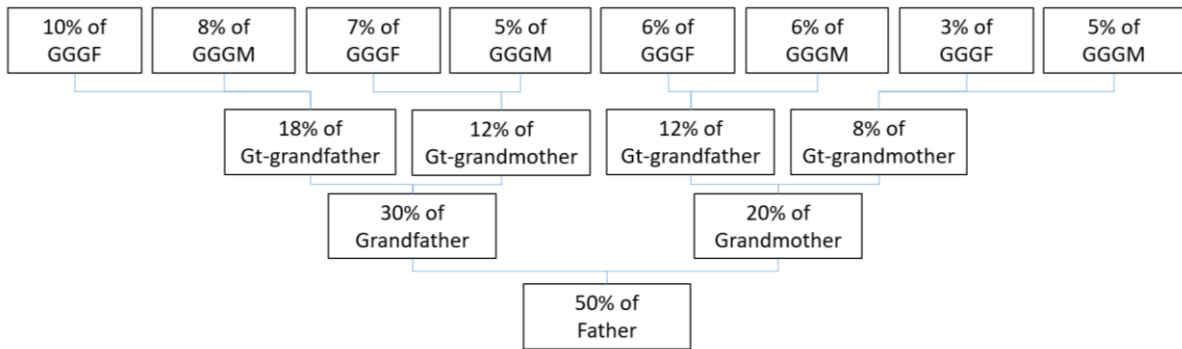
This is the DNA that is tested at all the main commercial DNA companies. If you are not sure what test you have taken, it will be an autosomal test. This tests 22 of your chromosome pairs (and sometimes the X chromosome(s) on the sex pair). As the DNA comes from both your parents and it came from both their parents and so on, it can help identify common ancestors (back to at least great-grandparents) on all your ancestral lines (provided a descendant has tested). 50% of the tested DNA comes from your father and 50% from your mother. The DNA however gets diluted as it comes down the generations. This is due to a phenomenon called 'Recombination'. This means that the 50% you inherit from one of your parents is a random selection of their parents across both chromosome pairs and as such you do not inherit exactly 25% of your grandparent's DNA.

This diagram shows a chromosome pair. The four colours represent your four grandparents' DNA and how they have arrived from your parents and ended up on your chromosome pair.



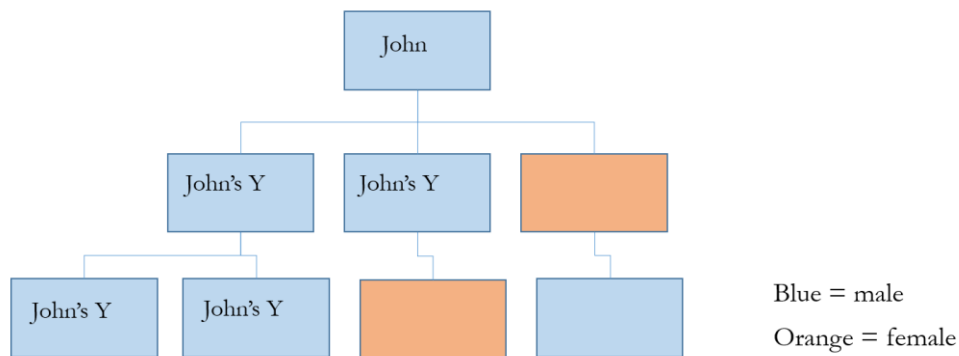
The mixing results in you getting more DNA from some generations further back than from others. Therefore, you may match with sixth cousins on some lines but, on another line, there is a chance you will not match some of your third cousins. An example of how much you have inherited from each ancestor back to great-great grandparents is shown below. Note that the amount inherited is different in every case and these splits are at the extreme end of the range (the differences are normally a little closer but these were chosen to help illustrate the point).

An Introduction to DNA for Family Historians – Handout



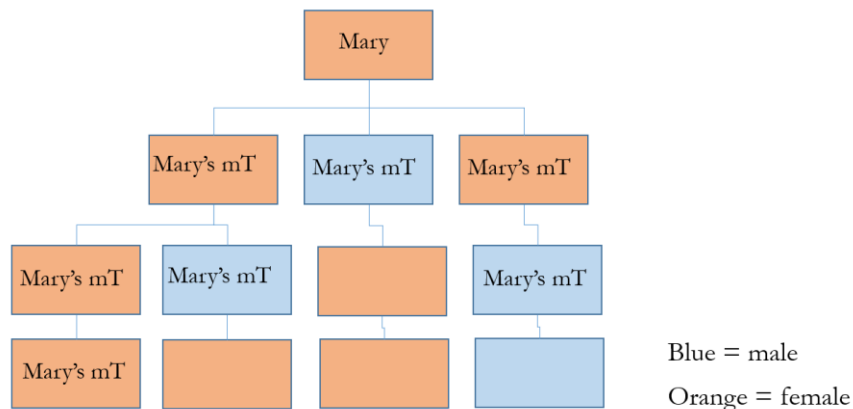
Y-DNA

Y-DNA tests are looking at the Y chromosome that is one half of the sex chromosome pair in men. Therefore, only men can take it. Y-DNA has a single inheritance path from father to son and is therefore useful for tracing the male (surname) line. Y-DNA only mutates (changes) slowly so most close relations have identical or very similar Y-DNA. It is therefore also useful for tracing your ancestry back multiple generations as the matches provided within your test results will be descended from a common direct male line ancestor. The closeness of your match determines how many generations back this common ancestor is likely to be. Y-DNA testing also provides you with a haplogroup. This is a prediction of the ancient origins of your male line. This diagram illustrates how Y-DNA is passed on:



mtDNA

Mitochondrial (mtDNA) tests are looking at the DNA contained within the mitochondria that surrounds the nucleus of a cell. It is passed on from a mother to all children. Therefore, men and women can take a mtDNA test. mtDNA mutates very slowly so most close relations on your direct maternal line have identical or very similar mtDNA. However, as the surname changes at most generations, it has limited uses for genealogical research. mtDNA testing also provides you with a haplogroup. This is a prediction of the ancient origins of your female line. This diagram illustrates how mtDNA is passed on:



An Introduction to DNA for Family Historians – Handout

DNA Testing Results

Autosomal tests provide two types of information:

- Ethnicity Estimate
- DNA Match List

Y-DNA and mtDNA tests provide two types of information:

- DNA Match List
- Haplogroup (ancient origin)

Ethnicity Estimates

These are just estimates and they change as more information becomes available to the testing companies. All the companies calculate it slightly differently. In addition, the calculations can only use the DNA that you have inherited. As three siblings inherit different DNA from their parents (e.g. not the same 50% unless they are identical twins), their ethnicity estimates will be different (and sometimes very different) even though they descend from all the same ancestors. The calculations are also different at each company because:

- The Reference Population used is different. This is the data from a sample of people who share a certain number of grandparents or great-grandparents from the same location (e.g. they have not moved areas for a long time). The companies compare this information with your DNA to predict where your DNA is likely to come from.
- The definition of Regions is different. What is meant by 'Northern Europe' at one company may mean a different selection of countries at another. It is always worth looking at the map provided to see exactly where the 'region' is covering.
- The algorithms (mathematical calculations) that they use to determine your ethnicity are different as many are proprietary.

Ethnicity Estimates provide a good feel for where you come from. If you have a specific ethnicity such as African or Asian within very recent generations, this should show up. However if you had a German great-grandmother and you do not show any German ethnicity, this does not mean you do not have a German great-grandmother, just that not enough of the DNA that indicates German ethnicity at the company you tested at has been inherited by you. Use ethnicity estimates to the nearest 25% rather than considering small amounts under 5%.

The big benefit of ethnicity estimates in research comes if you do have at least one 25% that is distinctive from the rest of your ancestry. When reviewing matches, people who also have that ethnicity are more likely to come from that ancestral line. However be careful as this could be a coincidence if they have a different ancestor that also has that ethnicity.

Ancient Origins (Haplogroups)

A predicted haplogroup gives an indication of where either your direct male line or direct female line originated. A location, approximate timescale and indication of likely migration routes are normally presented. These are provided if you take a Y-DNA or mtDNA test at FamilyTreeDNA or an autosomal test at 23andMe or LivingDNA (not uploads). Note that only men have a Y-DNA haplogroup.

DNA Match List

This is a list of people, who have taken the same test type and at the same company, who share some DNA with you. These 'matches' will share a common ancestor with you. For each of these people you get the following information:

- Amount of shared DNA (see below)
- Predicted relationship based on the amount of shared DNA (note however that the exact relationship could be different from the prediction especially if only a small amount of DNA is shared)
- Parental side – at AncestryDNA, each match is assigned as Parent 1 or Parent 2 (or Both Sides/Unassigned). Note Parent 1 can be either paternal or maternal; Parent 2 will be the opposite of Parent 1. More information at <https://support.ancestry.co.uk/s/article/DNA-Matches-Split-Up-By-Parent>.
- Family Tree or Surname information (depends on company and if the 'match' has added information)
- Shared Matches – these are other 'matches' who also share DNA with your match (e.g. all three of you share common DNA so are likely to all be related to the same common ancestor) (see below)
- Predicted common ancestor (only some companies provide a prediction and they are only a good guess based on your trees and your matches' trees (therefore you will need to verify them yourself))

An Introduction to DNA for Family Historians – Handout

For Y-DNA and mtDNA, the type of information provided is similar however the information is presented in a slightly different way. They do not provide a predicted common ancestor.

When reviewing your matches, you are looking for clues to allow you to try to find a common ancestor. If they have a well developed family tree (available at some testing companies), this is the easiest way to find a common ancestor. By reviewing the amount of shared DNA, you can work out possible relationships and hence which generation the common ancestor is likely to be at. Review the match's information to see if there are any shared surnames or places. Ultimately you may need to research out their tree to find a common ancestor. As you progress, you should keep a record of what you have found. In addition to your own records, there is the ability to record notes at most companies. All companies allow you to communicate with your matches.

Groups/Labels

AncestryDNA and MyHeritage provide a facility for assigning a 'Group' or 'Label' to your matches. These are often used for grouping your matches based on shared matches (see below) but can be used to tag your DNA matches in any way you need. The Groups/Labels allow you to filter your DNA matches.

- Ancestry Groups (Coloured Dots) – <https://support.ancestry.com/s/article/Grouping-and-Filtering-AncestryDNA-Matches>
- MyHeritageDNA Labels – <https://blog.myheritage.com/2021/12/labels-for-dna-matches-now-improved/>

Shared DNA

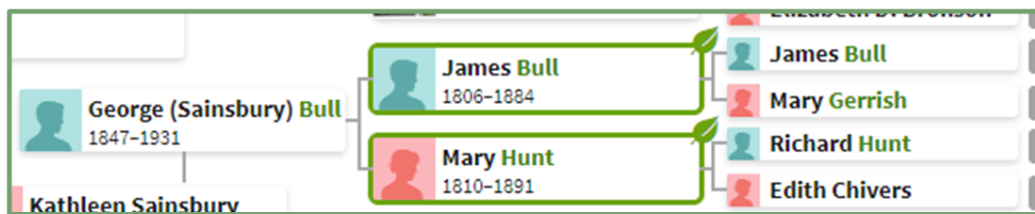
The amount of autosomal DNA you share with another person is measured in centiMorgans (cM). The amount you share will depend on how much each of you has inherited from your common ancestor. As a consequence, the amount of DNA you will share depends on how many generations it has passed through on each line and how diluted the DNA has become on each line. Therefore, for most shared amounts of DNA, there are a number of different possible relationships.

The Shared cM Project is the easiest way to find out what the possible relationships are (as well as their likelihood) given a specific amount of shared DNA. It is located at: <https://dnainter.com/tools/sharedcmv4>. If you type in the amount of DNA you share with someone (given on your match list), it indicates the percentage chance of various relationships. The lower the amount of shared DNA you have, the more options there are for how you are related. If you share less than 10cM, there is a higher likelihood that you will share the DNA by chance (e.g. not 'Inherited by Descent') and that the match is 'false'. The Shared cM Project also indicates the furthest back generation you are expected to find your common ancestor.

Common Ancestors and ThruLines at Ancestry

For Common Ancestors and ThruLines to work you need to link your DNA to yourself in your tree. To do this, go to the DNA home page and click on 'Settings' in the top right corner. Under 'Link Tree', select the family tree you want to use to connect your DNA (if you have more than one tree). Start to type in your name (if the DNA kit is yours) and then select yourself from the dropdown.

Once you have linked your DNA, Ancestry will start to generate 'Common Ancestors' (provided your tree is searchable – it can be public or private). These are suggestions for your potential common ancestors with the match and should be treated as hints not fact. If a common ancestor(s) is identified, these are listed in a box on the left-hand side of the page. In addition, if a common ancestor has been identified, this will show as a green leaf in the corner of the person's box on their family tree. Any surnames that appear in your tree as well are coloured green. An example extract:



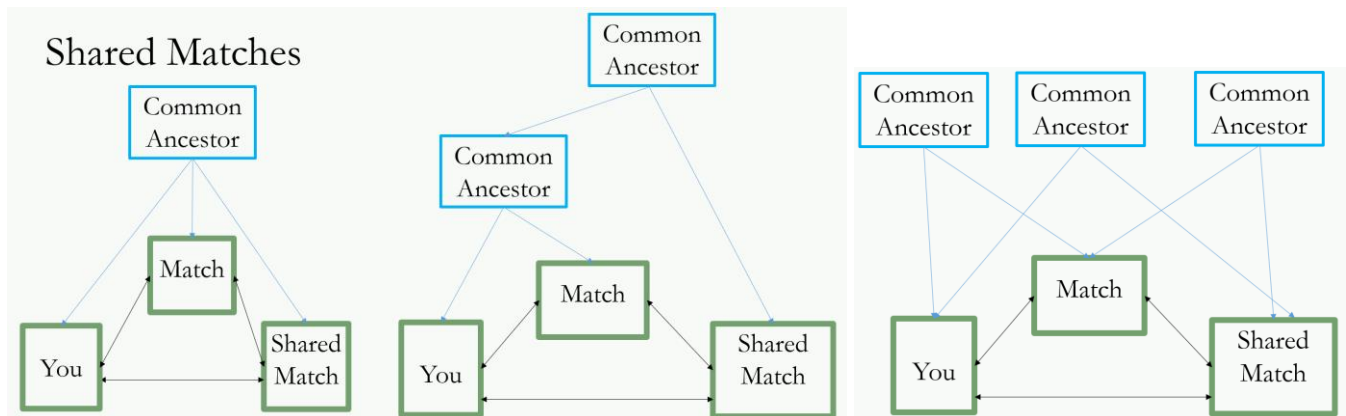
If you click on one of the common ancestor suggestions, the possible connection is listed. These also link into a system called ThruLines which finds all paths to that common ancestor within all your DNA matches. Note these paths are only suggestions created from your tree and other trees on Ancestry and are only as good as those trees. There is a chance that

An Introduction to DNA for Family Historians – Handout

the actual DNA relationship is inherited from a different path. Be careful of suggestions with less than 10cM of shared DNA. Although these do not prove that you inherited the shared DNA from the identified common ancestor, they do provide useful genealogical paper connections. On ThruLines, confirm that most matches are shared matches with each other to give additional confidence that they all share a common ancestor (although note the path may still be wrong for all of them). Beware if you have speculated in your tree – the Common Ancestors/ThruLines feature will give you confirmation bias (e.g. appear to confirm the speculation but it is only doing that because you put in the speculation).

Shared Matches

Each DNA company identifies for each match (e.g. person you share DNA with) all the other matches that also share DNA with that first match. This gives a list of people that share DNA with both yourself and your match (whose Match Page you are on). At Ancestry, to appear on the list, they must share greater than 20cM with both you and your match. Having a shared match indicates that it is very likely that you all three share a common ancestor (first scenario below) (noting that two of you may share a more recent common ancestor (second scenario)). There is a chance (although a lot less likely) that you each share a different common ancestor (third scenario). These diagrams show the three scenarios.



Finding Shared Matches

- AncestryDNA – see ‘Shared Matches’ tab on Match page. To appear on the list, the Shared Match must share greater than 20cM with you and share greater than 20cM with your match.
- MyHeritageDNA – scroll down to ‘Shared DNA Matches’ on Match page – click on ‘Show more DNA matches’ at bottom of list to show more
- FTDNA – click on the people symbol on right hand side of match list and choose ‘In Common With’
- 23andMe – scroll down to ‘Find Relatives in Common’ and click button to show shared matches

MyHeritageDNA and 23andMe also give the amount of DNA the shared match shares with your match.

Analysing DNA Results

The big benefit of using autosomal DNA is by identifying your common ancestor with as many matches as possible. Matches can be grouped (or clustered) into family lines by looking at matches’ Shared Matches as all the Shared Matches should be related in some way on that family line. To help find how your matches’ fit together, you will probably need to build their family trees out. These are often done quickly without the normal rigour applied to research and are called Quick and Dirty Trees. If you are trying to solve a brickwall or unknown parentage case, once you have identified the family groups, build the trees back down to try to identify where the family groups interact (normally via a marriage).

Y-DNA

To understand Y-DNA results, an understanding of what aspect of the Y-DNA they are comparing is required.

Y-DNA tests at FamilyTreeDNA are available at a number of different levels (the higher the number, the more precise the results but the more expensive the test is). The numbers in the test names represent a number of ‘markers’ (e.g. Y-DNA37 test compares 37 markers; Y-DNA111 test compares 111 markers). ‘Markers’ are specifically chosen different sections of DNA that contain patterns of DNA called ‘Single Tandem Repeats’ (STRs). At each ‘marker’, there is a repeated pattern of nucleotides and the number of times it is repeated can be counted. This number is recorded for each

An Introduction to DNA for Family Historians – Handout

tester at each marker and then compared against each other. Each difference in number of repeats represents a mutation and these happen relatively slowly.

If two testers have very similar number of repeats at markers, then they are a match as there have been very few mutations of the Y-DNA. The measurement of Shared DNA for a Y-DNA test (and mtDNA test) is called the 'Genetic Distance'. The genetic distance is calculated as the sum of the total number of differences at each marker that has been tested by both parties. So on the extract below showing the number of repeats at each marker for you and your match, the fourth column has a difference of 1 (14 and 15) and the seventh column has a difference of 2 (18 and 16). Assuming that all the other 25 markers (not shown here) have an identical number of repeats, the genetic distance is 3. This is the sum of the two differences; 1 + 2.

You	10	12	19-22	14	15	19	18	36-38	12	11
Match	10	12	19-22	15	15	19	16	36-38	12	11

Sample of 12 out of 37 markers

The smaller the genetic distance, the closer your Y-DNA match is. A genetic distance of zero means that there have been no mutations of the Y-DNA between you and your common ancestor and between your match and your common ancestor. Note however that even a father and son could have a genetic distance of 1 as a mutation has to take place at some point however this does not happen at every generation and does not happen often. The more markers that are tested, the more locations a mutation could be identified by the test so a higher genetic distance at the bigger tests (67 and 111 markers) will be closer than a smaller genetic distance at a Y-DNA37 test. The normal rule of thumb is a genetic distance of 2 for a Y-DNA37 test (or genetic distance of 5 for a Y-DNA111 test) probably indicates that the common ancestor is within about 10 generations. The more markers you test, the more expensive the test is. You can take a Y-DNA37 test and then upgrade to more markers at a later date if your matches at Y-DNA37 indicate that there is benefit in upgrading to more markers.

When can a DNA test help your research?

- Verify known family history
- Build family tree based on DNA results
- Breakdown (some) brickwalls
- Identify unknown parents/ancestors

All the above require a significant amount of traditional genealogy research to be carried out in conjunction to the DNA analysis.

Considerations on whether DNA analysis will help include:

- Research question is relevant to DNA (identifying relationships, births, locations)
- Time period is relevant to DNA (5-6 generations for solving brickwalls/unknown parentage unless on the direct male line)
- Appropriate test taker(s) if DNA test not yet taken (or more data required)

Choosing a Test

There are five commercial DNA testing companies for genealogy. They are:

- AncestryDNA – www.ancestry.co.uk/dna
- MyHeritageDNA – www.myheritage.com/dna
- 23andMe – www.23andme.com
- FamilyTreeDNA – www.familytreedna.com
- LivingDNA – www.livingdna.com/uk

There are a number of factors to consider when deciding which test to take:

An Introduction to DNA for Family Historians – Handout

- **Type of test** – autosomal, Y-DNA or mtDNA. Autosomal allows you to find DNA matches with common ancestors on all ancestral lines. Autosomal tests are available at all companies. Y-DNA covers direct male line. mtDNA covers direct female line. Y-DNA and mtDNA tests are only available at FamilyTreeDNA.
- **Size/Coverage of Database** – the sizes of the databases are given in the table below. The larger the database, the more matches you will have. Although most of the databases are USA dominated, luckily a lot of our ancestors had siblings who moved there and hence there are a lot of genetic cousins in the USA. MyHeritageDNA has good coverage in Europe particularly Scandinavia and the Netherlands and within the Jewish population. LivingDNA is UK based and is likely to have a larger percentage of UK testers. However, the database is believed to be much smaller so the total number of UK testers will still be significantly lower than the main commercial databases.
- **Method of Testing** – depending on the company, the test will either be a cheek swab or a spit test. Be aware that a few people struggle to give a spit sample.
- **Cost** – the basic cost for an autosomal test is similar across the companies. All regularly hold sales (linked to national holidays, Mother's/Father's Day, DNA Day, Black Friday, etc. or at events such as family history shows) so it is worth awaiting a sale to purchase.
- **Uploads** - For MyHeritageDNA, FamilyTreeDNA and LivingDNA, you can take your autosomal results output from any of the other companies and upload them to add your autosomal DNA into those databases for free. Payment is required to enable more detailed information and tools to be used at these companies following upload however this is not essential to get some useful information. This does mean that if you want to be in the Ancestry or 23andMe database, the only way is to take a test at that company.
- **Upgrades** – this only applies to FamilyTreeDNA. If you take a test (not via an upload) for either FamilyFinder (autosomal), Y-DNA (various levels e.g. Y-37, Y-111, Big-700) or mtDNA, then you can upgrade it at a later stage to any of the other tests without doing another physical cheek swab test.
- **Additional Benefits** – these include family trees, predictions of common ancestors, medical/health add-ons, etc. The table below lists these additional benefits under Tools and Extras.
- **Other Family Members** – If cousins/other family have tested at a specific company, it is worth testing at the same company as any shared matches will be on the line they are on. This will help with identifying the line/common ancestor for those DNA matches.
- **Potential concerns** – There are a number of issues to be aware of when taking a DNA test.
 - **Unexpected results** – there is the potential that the results may uncover different parentage than expected or additional relationships (e.g. an unknown half sibling) either for yourself or a relation (e.g. for your parent or grandparents).
 - **Limited results** - there is the potential that you will only get limited results that help your own family history research (and potentially no results for Y-DNA or mtDNA) although this does not happen often. Once you have taken the test, as more people take tests, your number of matches will increase.
 - **Other uses of databases** – most of the companies provide the ability to participate in anonymised research however it is completely your own choice. Most of the companies offer medical/health profiling as an additional service (it is completely up to you but is unlikely to aid your genealogical research). There have been occurrences of DNA databases being used to aid police investigations in the USA however the only companies covered here that has had an involvement is MyHeritageDNA and FamilyTreeDNA. They have both since improved their Terms and Conditions to make it clear what is and is not allowed so you are fully consenting to how your information is used. The key website that has been used for police investigations is GEDmatch and this website has not been discussed in this presentation.

An Introduction to DNA for Family Historians – Handout

All companies have their own Terms and Conditions which should be read and agreed to before buying the test or uploading to that company. If you are asking someone else to test (or you are being asked to test by someone else), it is advisable that you use an Informed Consent form. An example is available at:

<https://thegeneticgenealogist.com/2018/02/15/informed-consent-agreement-and-beneficiary-agreement/>

The table below summarises the information regarding these considerations for each company.

Company	Size	Collection Method	Price	Tools	Extras	Uploads accepted?	Concerns
AncestryDNA	~ 21 million	Spit	£79 £49*	ThruLines; Family Trees	Traits	No	None
MyHeritageDNA	5.7 million +	Swab	£79 £34*	Theory of Family Relativity; Family Trees; Chromosome Browser		Yes	Past breaches of T&C
23andMe	12 million +	Spit	£79 £39*	DNA FamilyTree; Chromosome Browser	Traits; Health	No	None
FamilyTreeDNA	1.7 million +	Swab	\$79 (£62^) \$49* (£38^)	Chromosome Browser; Family Trees	Y-DNA & mt-DNA upgrades	Yes	Past breaches of T&C; Limited Law enforcement use
LivingDNA	~ 0.3 million	Swab	£99 £69*	Chromosome Browser	Wellbeing Kit	Yes	Limited database size

Database sizes courtesy
of www.thednageek.com

*Current or recent sale price
^UK price based on exchange
rate 16 August 2023

Further Reading

- Books
 - *The Family Tree Guide to DNA Testing and Genetic Genealogy*; Blaine T Bettinger; Family Tree Books, USA; 2nd Edition, August 2019.
 - *Tracing Your Ancestors Using DNA: A Guide for Family Historians*; Edited by Graham S Holton, contributed by John Cleary, Michelle Leonard, Iain McDonald & Alasdair F MacDonald; Pen & Sword Family History, UK; 1st Edition, June 2019.
 - *Your DNA Guide - The Book*; Diahna Southard; Your DNA Guide, USA; 1st Edition, February 2020.
- ISOOG wiki - https://isogg.org/wiki/Wiki_Welcome_Page
- Websites/'How to' pages
 - <https://www.familysearch.org/blog/en/genealogy-dna-test/>
 - <https://www.whodoyouthinkyouaremagazine.com/tutorials/dna/>
 - <https://www.family-tree.co.uk/dna-testing>
- Blogs
 - Memories in Time – Fiona Brooker - <https://memoriesintime.co.nz/blogs/news>
 - The DNA Geek – Leah Larkin - <http://thednageek.com/blog/>
 - Your DNA Guide – Diahna Southard - <https://www.yourdnaguide.com/ydgblog>
 - Crewys News – Debbie Kennett - <https://cruwys.blogspot.co.uk/>
 - The Genetic Genealogist – Blaine T Bettinger - <https://thegeneticgenealogist.com/>
- Facebook groups
 - DNA Help for Genealogy (UK) - <https://www.facebook.com/groups/AncestryUKDNA>
 - Genetic Genealogy Tips and Techniques - <https://www.facebook.com/groups/geneticgenealogytipsandtechniques>
 - The DNA Roundtable - <https://www.facebook.com/groups/1587685288089570>

Contact Me

Mia Bennett

miagenealogy@gmail.com

www.miagenealogy.com/about