THE CANADA'S HISTORY BEGINNER'S GUIDE TO GENETIC GENEALOGY

Read in sequence or browse as you see fit by clicking on any navigation item below.

Introduction

The "bestest best boy in the land" recently had his DNA tested. It turns out he's a beagle mix with a gazillion undifferentiated hound varieties—basically a nose on legs.

Is DNA testing a gimmick or the real deal? That's the question most often put these days to experienced genealogists by members of the interested public, e.g. the readers of *Canada's History* magazine and of this website. Even dogs are doing it. Should you?

This online guide seeks to show when and how DNA testing might work for you—and why it might not.

We'll cover the basic science and the available tests. We'll review a set of testing strategies for beginners and a recovery guide for those who've already tested and were underwhelmed. Most importantly, we'll see tests in action in a series of case studies. And because there's way more to know than this guide can address, we'll offer plenty of suggestions, many online, for reference and further guidance.

DNA enthusiast, nose on legs and bestest best boy in the land. (Photo: Paul Jones)

A. To test or not

Reasons for testing

People test their DNA for many reasons. This field in its entirety is called "consumer genomics."

Some just want to contribute to scientific knowledge. Others have questions about their genetic predisposition to various diseases or conditions. These interests fall outside the scope of this guide.

Rather, we'll focus on those who wish to use DNA to learn more about their extended families, both ancestors and relatives living today. This field is known as "genetic genealogy."

Some genetic genealogy testers simply wish to explore their regional, ethnic or national origins. Please see here for a discussion of the options for, and limitations inherent in, this "admixture" or "biogeographical ancestry" analysis.

Mostly though, this guide deals with how a tester can use DNA to help identify elusive ancestors, as well as distant and mostly unknown living cousins. The researcher's need might be as pressing as that of an adoptee seeking a biological parent while there's still time, or as frivolous as someone hoping to confirm a family story about a distinct connection to nobility. Read these case studies to see DNA testing in action.

There are some occasions when DNA testing may provide the only practical avenue of genealogical research, e.g. if records were never kept or were later destroyed. For example, any African American or Indigenous North American can expect to find only spotty records of their ancestors in colonial times and virtually none beforehand. And for people subject to comprehensive record-keeping, there's no guarantee documents have survived to the present day, viz. Ireland pre-1922.

Even if you have a well-documented family tree, bear in mind that DNA confirmation of your conjectures may carry even more weight with family and friends. Rightly or wrongly, people may shrug off archival findings while hailing DNA evidence.

Indeed, most experienced genealogists would regard a research project as inadequate if it did not incorporate pertinent and feasible DNA testing. The benchmark of acceptability is the Genealogical Proof Standard, and its provisions require "a reasonably exhaustive" review of all relevant sources. Today, that would include DNA testing for many if not most problems.

Bogus reasons for not testing

First, let's dismiss three bogus reasons for not testing.

Why seek lost cousins?

There are many reasons you might want to seek living cousins with whom you and your immediate family are not in touch. Here are a few of them:

- to satisfy your curiosity about other lines of the family
- to gather information about your common ancestors and relatives
- to find family memorabilia or photos, or to find a more appropriate home for those that you have
- to hear, or gain a different perspective on, family stories
- to organize a family reunion
- to reunite families, e.g. adoptees
- to build a medical pedigree
- to seek compatible organ donors
- to locate heirs

For more information on ***why*** you might want to seek lost cousins, see this column, which ran in *Canada's History* in 2011. And for suggestions as to ***how*** to find them, read this column from 2013.

Fear of the test

No, the tests are not invasive or painful. No one will be drawing your blood. You can spit into a tube or swab your cheek in the privacy of your own bathroom. Instructions are straightforward. More than 20 million people have successfully taken a test. You can too.

"The tests are crap"

And, no, the tests are not "crap." Quality control is not an issue with the major testing companies. Anyone expressing a contrary opinion likely falls into one of these categories:

- someone who tested with a company offering outdated technology or shady business practices, i.e., not those discussed in this Guide
- someone in denial about some aspect of their findings
- a journalist or similar self-appointed guardian of the public good seeking to make a name or a buck by promulgating a provocative or sensationalistic narrative

Price

The Human Genome Project (1990–2003) cost US $2.7 billion. Today you can have your whole genome sequenced for under a thousand dollars. But that's overkill for genetic genealogy. The most popular tests can now be had for under a hundred dollars if you buy wisely. In short, price is no longer an obstacle for any middle-class person for whom genetic genealogy testing is a priority. We will likely see the $99 whole genome sequence within the next decade, maybe sooner.

Substantive reasons for not testing

Privacy concerns

Some people are concerned about possible loss of privacy with DNA testing. Indeed, it's estimated that 60% of North Americans of northern European heritage can already be identified from their DNA, even if they themselves have not been tested. In a sense then, this genie may already be out of the bottle.

Even so, testing companies do their level best to protect your privacy. For example:

- Testing companies typically conform to best practices for data protection—I am not aware of any data breaches.
- They will not disclose information about you to law-enforcement agencies unless presented with a court order. None of the criminal cases recently solved via DNA used the testing companies' databases. (See next page for a discussion of how DNA *did* help solve these cold cases.)
- Testing companies generally allow you to conceal your identity from other testers and to share as much or as little information about yourself and your results as you are comfortable with. (If you're not going to share though, what's the point of participating?)
- With the exception of 23andMe, the testing companies are largely uninterested in medical implications of the results. Indeed, Family Tree DNA does not report findings for genetic markers known to be associated with medical conditions.

So the bottom line is that privacy concerns need not be—and do not appear to be—a deterrent for most people.

But there are exceptions.

If you know or suspect that you have a genetic predisposition to a catastrophic illness, you might want to seek genetic counselling and defer DNA testing for the time being. Wait until you're ready psychologically and have protected yourself from any adverse employment or insurance consequences in the event that your genetic genealogy test inadvertently contains medical information that could be damaging to you.

And if you have a violent criminal in your extended family, your DNA test could potentially provide the needed clue to law enforcement if you choose to make your DNA public. Most people, I think, would welcome this outcome, but some are troubled by the potential for unintended consequences.

Unexpected findings

The main risk in genetic genealogy testing for most people is an unexpected and disturbing finding, usually a hitherto unknown adoption, illegitimacy or some other skeleton in the family closet.

Most families have an illegitimacy somewhere in their past; in some lines, it seems to the rule rather than the exception. Genealogists knew this well before the advent of DNA testing. Even so, a revelation can be startling to someone who has never imagined that grandma could have had a teenage "accident" or that dear old dad had sown more than wild oats in his youth.

Adoptions too can be a fraught subject, rarely or perhaps never discussed in many families. How would you handle clear evidence that a parent, sibling—or even yourself—had been adopted?

Even more extreme, a recent article in the *Washington Post* reported the unexpected discovery through DNA testing of a 1913 mix-up of two babies at a New York City hospital. Generations of both families had sensed something was amiss. The descendants today are working out what it all means.

More shocking still, leading genetic genealogist CeCe Moore remarked in an interview for *MIT Technology Review* that she has encountered a number of cases of "direct DNA evidence of incest" in her analyses of test results.

Adoption. Illegitimacy. Hospital error. Even incest. If you cannot honestly say you could handle any results, no matter how unlikely, you probably shouldn't take a test.

DNA testing and unsolved criminal cases

In 2018 news from the DNA lab momentarily elbowed politics, celebrities and sports from the headlines. Connoisseurs of crime detection were fascinated by the capture of the alleged Golden State Killer, promptly followed by the arrest of a clutch of suspects in comparably ancient cold cases. Researchers using the everyday tools of genetic genealogy had succeeded in unraveling mysteries that had defied solution for decades. To the press and public, it all seemed like something from science fiction.

Yet it shouldn't have been that surprising. If researchers can use DNA to identify the biological parents of aging adoptees, the undocumented fathers of 19th century "baseborn" children, or even the details of a century-old "switched at birth" hospital error, then it should come as no surprise that investigators can figure out the identity of a murderer or rapist clumsy enough to leave his DNA at the scene of a crime just two or three decades ago.

Actually, the process is quite straightforward and is similar to what adoptees do to find biological parents. You do an autosomal test on a sample of interest (in this case, the DNA left at a crime scene) and compare the findings to other testers' results. Using conventional genealogical techniques, you try to figure out how the mystery test fits in with the family trees of its identified matches.

"Already, 60% of Americans of Northern European descent—the primary group using [DNA testing] sites—can be identified through such databases whether or not they've joined one themselves," the *New York Times* reported in October 2018 in an article on a newly published study in the journal *Science*. "Within two or three years, 90% of Americans of European descent will be identifiable from their DNA, researchers found." The figures for Canada are not known but cannot be hugely different.

CeCe Moore, the professional genealogist who has solved most of these cases, predicted to NBC News that we will see dozens of cold cases resolved in upcoming months, and hundreds over the next few years. Jay Cook and Tanya Van Cuylenborg, the victims in one of Moore's earliest cases, were Canadians. It is just a matter of time until a criminal is apprehended by a Canadian police force.

It's important to note that none of the cases solved to date have involved the use of data accessed at, or received from, one of the DNA testing companies. Indeed, the terms of service of AncestryDNA, Family Tree DNA *et al.* preclude access to client information by third parties, including police, without the client's explicit approval. In theory police could obtain a court order for a suspect's data, but that's not what's going on here.

The database used by Barbara Rae-Venter (Golden State Killer case), CeCe Moore and other crime-fighters is GEDmatch, a website where more than a million researchers have voluntarily uploaded their raw autosomal DNA results from their testing companies. (People do this for several reasons. Some are seeking new DNA matches among like-minded people who tested with different companies. Others simply want access to the cutting-edge analytical tools that GEDmatch offers to its users.) Investigators have likewise uploaded appropriately formatted crime-scene DNA and then started looking for matches.

No one expected any perfect matches. But if the investigators found, say, that the perp looked like he was a second cousin to Person A in the database and a third cousin to Person B, then the problem began to resemble the many puzzles, especially adoption cases, that had been solved by researchers in recent years. The solution: build family trees for Persons A & B, work out how they are related to each other, then home in on the spot on their overlapping trees where the perp would have to fall in order to satisfy the relationships estimated from the DNA sample.

It's not an exact science because of statistical uncertainty, and usually there will be a number of prospects (a better word than suspects in this situation)—brothers, cousins and others too close to each other to be ruled out by the rough-and-ready methodology. Next, it's down to everyday detective work, eliminating from consideration those prospects whose ages or places of residence don't fit with the crimes. Finally, it's a matter of getting a sample of the prime suspect's DNA, which won't be on record anywhere or the case would have been solved years earlier. In the case of the Golden State Killer, it was DNA extracted from his car door handle and a discarded tissue. Bear in mind that DNA can be definitively exculpatory if the police have narrowed in on the wrong man—think of all the Death Row convicts pardoned in recent years. No such luck for Joseph James DeAngelo, whose age, career and places of residence tally in every respect with what one might have expected for the Golden State Killer. So does his DNA!

GEDmatch was blindsided by use of its data to solve cold cases and updated its terms of service to explicitly advise users of law-enforcement activities. In addition, GEDmatch has attempted to limit police use of its database to the solving of violent crimes only, e.g. murder and rape. That said, the major deterrent to police use of GEDmatch appears not to be rules of dubious enforceability but the substantial costs inherent in pursuing these cold cases. DNA is only a starting point; typically, a highly skilled genealogist will have to spend many hours building family trees in order to zero in on a suspect with exactly the right profile.

Continued from page 5

Privacy critics are not placated. They argue that hundreds of thousands of GEDmatch users have not given explicit permission to allow police to use their DNA. Their objection is not that violent criminals are being apprehended, which is a welcome development per se. But critics worry about the unintended consequences of making personal DNA widely available, and they are sensitive to the potential for police abuse of a powerful new tool that is unregulated and unmonitored.

"Admixture" or "biogeographical ancestry" analysis

If you're like me, you haven't been able to turn on the TV in recent years without seeing a commercial from a DNA-testing company encouraging you to discover your ethnic, national or regional origins. In one of these ads, the test subject learns that she's 14 per cent Italian, a "fact" hitherto unknown to everyone in her family. Her brother Steve, she decrees, will henceforth be called Stefano.

Cute, but the sad reality is that Stefano might well learn that he should be Esteban or Szcepan if he were to take the same test himself. As genetic genealogy authority Roberta Estes says, "These tests are great at detecting ancestry over 25 per cent—but if you know who your grandparents are, you already have that information… [Sub-continental] results, meaning within Europe, for example, are speculative at best."

Let's think about it. What exactly is an Italian genetic profile? (You can and should ask the same question about any other nationality or ethnicity of interest.) The ancient pre-Roman tribes of the Italian peninsula had various genetic origins. Since Antiquity, there have been many peoples who've left their genetic markers: Greeks, Carthaginians, slaves from across the Roman Empire, the various barbarian invaders, the Lombards (a Germanic people), Normans (a.k.a. Vikings), Aragonese and the French, to name just a few. In more recent times, we've seen Albanians, sub-Saharan Africans and now Syrians joining the mix.

At no point in the past several thousand years so far as we know, was there ever a time when residents of the Italian peninsula were genetically homogeneous.

So when a genetic test identifies someone as having "Italian" origins, what exactly does that mean?

No wonder another authority, Judy G. Russell, dismisses such national estimates as "cocktail party conversation."

Of course, in genealogy there are always exceptions to the general rule. Distinctive, substantial and unexpected findings of ethnic or national origins can help focus the research efforts of those who know nothing about the origins of a parent or grandparent, e.g. someone who is illegitimate or adopted or the child of someone who was illegitimate or adopted. In such instances, admixture analysis can be a useful part of the genealogist's tool kit.

Further, the design and interpretation of ethnic/regional/national analyses are constantly improving as more powerful algorithms and ever larger databases come into play. While national origins may never be anything but a parlour game for the reasons outlined above, truly distinct genetic markers may exist at the level of local communities that have enjoyed relative stability through the ages.

Consider an analogy with surnames. Yes, you can find just about any surname in a big city, but the highest incidence of most surnames generally corresponds to the local areas where the surnames have propagated for centuries. So too it may be with the genetic mutations that occurred in a community and have circulated locally through generations of marriage and reproduction.

Identifying such genetic specificity requires massive samples and powerful algorithms, but at least one of the testing companies, AncestryDNA, may be on the cusp of attaining the critical mass and devising the right tool in its "genetic communities" and "migrations" features.

Among traditional tests for ethnic origins, most experts agree that the company 23andMe may be somewhat more accurate than its peers, although probably not in all regions for all people. Bear in mind though that if you plan to use your DNA results in other ways, the strengths of AncestryDNA, Family Tree DNA or MyHeritage DNA may be more appropriate.

A relatively recent entrant in the field of admixture analysis is Living DNA, a specialist in the British Isles. As with other admixture tests, some people report very close matches to what they know and have documented; others are puzzled by their findings.

In the future we can probably expect more accuracy and higher specificity from these tests, but interpretation will remain complicated. Some places have enjoyed more stability than others, so it's unlikely there can ever be a "one size fits all" admixture test.

B. The ABCs of DNA testing

In order not to clutter up this guide for beginners, we do not cover minor, rare or irrelevant exceptions to the general rules. Consider everything that follows as accompanied by a generic qualifier to this effect.

The cells in your body contain four types of DNA that can be categorized by their inheritance patterns. These inheritance patterns in turn allow us to confirm or deny conjectured genetic relationships between individuals or to estimate the degree of relationship between two strangers who are identified as having a genetic match.

There are four major testing companies performing three major kinds of genetic genealogy tests on these four types of DNA.

Four major testing companies (listed below in the order in which they began offering consumer DNA testing) collectively constitute the lion's share of the "personal genomics" industry. For a side-by-side comparison of the testing companies' attributes, see the International Society of Genetic Genealogy Wiki (ISOGG Wiki). Click on the titles to visit the home pages.

The four major testing companies (and others)

23andMe

- a company focusing on both the genealogical and medical applications of personal genomics
- traditionally regarded as having the best biogeographical ancestry analysis
- a large but often unresponsive database of testers (more than 5 million per ISOGG Wiki, but many interested only in medical findings, not genealogy)

Family Tree DNA (FTDNA)

- an independent company focused entirely on genetic genealogy and offering state-of-the-art genetic genealogy testing of all kinds
- the only major testing company to offer comprehensive Y- and mitochondrial-DNA testing
- regrettably the smallest (although best informed) autosomal DNA database (ca. 890,000)

AncestryDNA

- a division of Ancestry, a well-established market leader in providing digitized and indexed historical records of relevance to genealogists
- the elephant in the room with close to 15 million tests in its database, many of them relative newcomers to genealogy, genetic or otherwise
- focused on aligning DNA findings with member family trees but criticized by many for its decision not to provide conventional analytic tools such as a chromosome browser

MyHeritageDNA

- a division of MyHeritage, a full-service genealogy company like Ancestry
- known for its "matching" expertise
- a recent but fast-growing entrant in the genetic genealogy market with special strength outside North America (database of 1.75 million and growing)

Two other organizations should be mentioned:

National Geographic Genographic Project

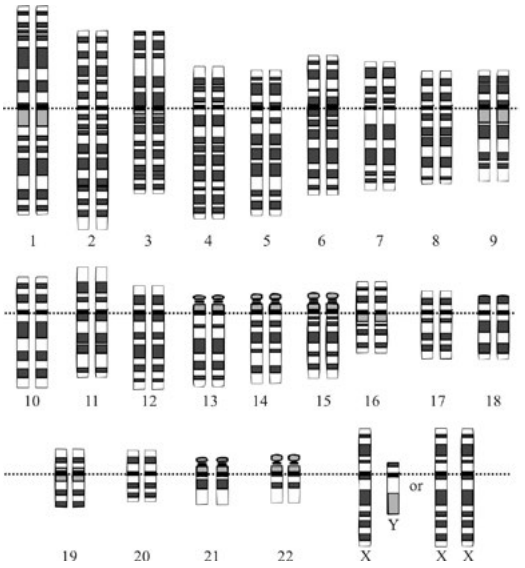
- a pioneer in exploring global pathways of deep-ancestry migration through analysis of mitochondrial and Y DNA
- the scientific rationale largely overtaken in recent years by breakthroughs in the analysis of ancient DNA

Living DNA

- specialists in the genetics of the United Kingdom and Ireland
- currently offering only regional and deep ancestry testing
- matching to other participants in beta testing
- now partnered with established full-service genealogy company Findmypast

It should be noted that there are a number of cutting-edge companies offering next-generation and whole-genome testing to specialist and advanced consumers. If you are a potential customer for such a test, you are presumably sufficiently well informed not to need the advice of this Guide

BUYER BEWARE: Test at your peril with any company touting entry-level DNA testing that is not mentioned above. These sometimes crop up on online couponing sites. There is a good likelihood that you are being offered an inferior or even shoddy product. Be especially careful when encountering extravagant claims (e.g. "find your Cherokee ancestry") or company names that seem suspiciously similar to the legitimate trademarks of the market leader AncestryDNA.



The 23 chromosome pairs in nuclear DNA. (Created by National Human Genome Research Institute; public domain.)

Four types of DNA and the three major genetic genealogy tests

Human cell nuclei contain three types of DNA as illustrated in the following diagram: 22 pairs of "autosomal" chromosomes (numbered 1-22), and either one X and one Y chromosome (males) or two X chromosomes (females). In addition, the mitochondria in our cells contain a small DNA molecule. Each of these types of DNA is discussed below. As genealogists, we are particularly interested in the inheritance patterns for each of these types of DNA.

Mitochondrial DNA (mtDNA)

This DNA is found inside small structures in your cells called mitochondria and is passed along unchanged from mothers to children.

Because mtDNA mutates slowly and is passed whole from generation to generation, it is an excellent way of studying the so-called deep history of your maternal line. In fact, it was the first type of DNA used by scientists to explore our remote genetic past. It is one of the two uniparental types of DNA, i.e., DNA inherited from a single parent, and it is the only type of human DNA that does not reside in cell nuclei.

Note that males inherit their mothers' mtDNA but only daughters can pass it along to future generations. The chart on page 9 illustrates mtDNA inheritance patterns.

Because mtDNA mutates so slowly, it is rarely helpful in differentiating between two people related matrilineally within a genealogical timeframe. Its principal genealogical value lies in confirming or denying a conjectured genetic relationship where the tested parties would otherwise have no reason for sharing identical or near-identical mtDNA apart from the hypothesis under consideration. For an interesting historic case study, read this article from *Nature*.

For a contemporary case study, see this blog post.

Note that it may be possible to use a testing surrogate within your family to render a genealogical problem susceptible to study via mtDNA testing (see the sidebar above).

While several of the major testing companies provide your mtDNA deep ancestry haplogroup (i.e., your ancient maternal line), Family Tree DNA is the only one to provide full mtDNA sequencing to the genealogical community.

Using surrogates in uniparental DNA testing

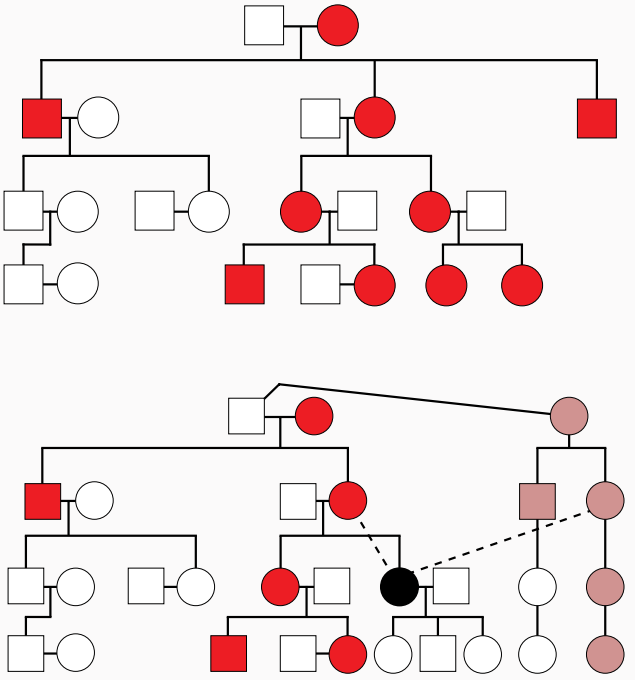
Note that you can expand the scope of meaningful mtDNA or Y-DNA testing by carefully selecting surrogates within your extended family.

Say, for example, you're hot on the trail of a pair of great-grandparents who have to date eluded every attempt to identify them. You finally have a candidate couple, the known ancestors of a putative second cousin you've identified through previous autosomal testing. Even better, this person descends matrilineally from the couple of interest. Problem is, you don't descend matrilineally from them. Even if you have correctly identified your great-grandparents, mtDNA testing will not be of use in confirming a valid match between you and your putative second cousin.

But it's not game over. You're a genealogist and you know your family inside out. The chances are good that you have a first or second cousin who's known to be descended matrilineally from a daughter of the mystery ancestors. Instead of testing yourself, test this person.

Note that precisely the same argument could be made in this case if we were to replace matrilineal with patrilineal descent and mtDNA with Y-DNA testing.

Such uniparental DNA matches are particularly telling in establishing a relationship because, unlike autosomal DNA, they can only arise in narrowly defined circumstances, i.e., strict matrilineal or patrilineal descent.



Three-generation descendancy chart for mtDNA. This chart shows how mtDNA testing can be used to determine whether a descendant was born to a first or second wife. Squares represent males, circles represent females. (Courtesy of Debbie Parker Wayne)

Y DNA

The Y is a sex chromosome found in the nuclei of the cells of males only. It is passed along unchanged from fathers to sons. Women do not have Y DNA but can study their paternal lines through Y-DNA surrogates such as their brothers, fathers, paternal uncles, etc. Because Y DNA mutates slowly and is passed whole from generation to generation, it is an excellent way of studying the deep history of your paternal line. As with mtDNA, it is a uniparental kind of DNA. This link illustrates Y-DNA inheritance patterns.

Note that Y DNA cannot be used directly to research many of your ancestors, i.e., those who are not related patrilineally to you or your family. Even so, and as with mtDNA testing discussed above, it may be possible to convert a genealogical problem into one susceptible to Y-DNA testing through the use of a surrogate tester within your extended family who carries the Y DNA relevant to the problem at hand.

For a case study involving Y DNA, read this article.

While several of the major testing companies provide your Y-DNA haplogroup (i.e., your ancient paternal line), Family Tree DNA is the only one to provide high-resolution Y-DNA testing for genetic genealogy purposes.

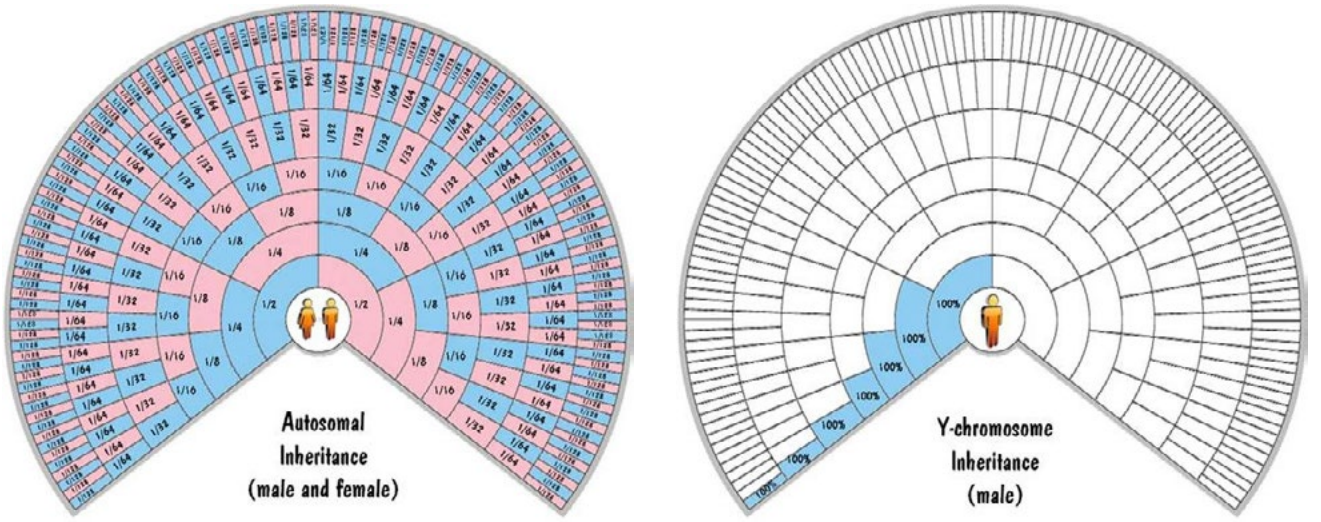
Autosomal DNA (atDNA)

Autosomes are the 22 pairs of chromosomes in cell nuclei unrelated to the gender of the carrier (i.e., neither X nor Y). You inherit one chromosome in each autosomal pair from mom and one from dad; each is a mash-up ("recombination") of the corresponding chromosomes your parents in turn inherited from your grandparents. Here's a schematic chart showing how a brother and sister inherit their autosomal DNA:

While you get 50 per cent of your autosomal DNA from each parent, recombination of grandparental DNA ensures the following:

- siblings each get half of their DNA from mom and half from dad, but they don't get exactly the same DNA as each other (unless they're identical twins)
- the proportion of DNA you get from each grandparent is subject to chance, the average being 25 per cent but with considerable variation

As you can appreciate, algorithms to analyze autosomal DNA depend on statistical analysis and probabilities. Testing companies typically look at 600–700K markers (i.e., known points of mutation) on your autosomal DNA; they compare you with every other person in their database at each marker and apply statistical algorithms to estimate your degree of relatedness to those whom you match most closely. Your test results will include a list of your matches, each with an estimated relationship and ancillary information that varies from one testing company to another.



Left, autosomal DNA inheritance. Right, Y-DNA inheritance. (Courtesy of Stephen P. Morse, based on an idea by Blaine Bettinger)

If you are going to make any progress with atDNA, you will have to master a few basic concepts. Here is a sampling:

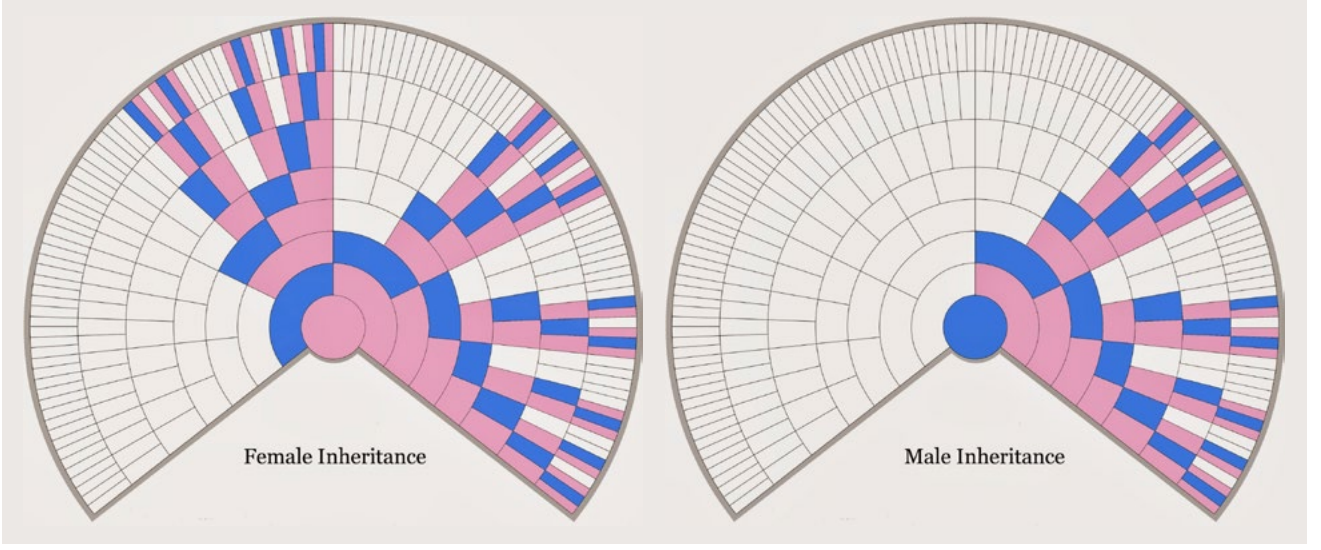
- the distinction between Inheritance by State (IBS) and Inheritance by Descent (IBD)
- the impact of endogamy on testing results
- what you should expect to find when relatives test, as studied in the Shared cM Project

All the major testing companies offer atDNA testing. It is fair to say that using the results from Ancestry is a qualitatively different experience from using those from other companies. For a more detailed discussion of this topic, read the sidebar on page 11.

X DNA

The X is a sex chromosome with very complicated inheritance patterns (see below). Females have two X chromosomes, one inherited whole from their fathers, the other a mash-up of their mothers' two X chromosomes. Males have one X chromosome, a mash-up of their mothers' two X chromosomes. Exploring ancestry through X DNA is a complex topic beyond the scope of a beginner's guide.

All the major testing companies test the X chromosome as a by-product of autosomal testing. Each has a different way of analyzing and reporting the results, thereby adding a layer of complexity to an already difficult topic.



Fan charts showing male and female inheritance of X DNA. (Courtesy of Blaine Bettinger, *The Genetic Genealogist* blog.)

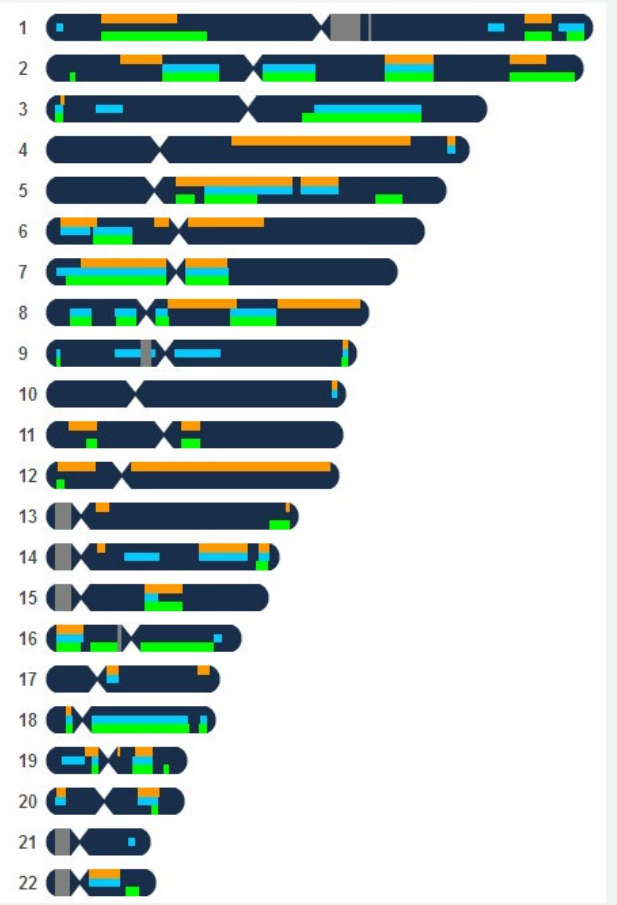
Setting a genealogical goal for DNA testing

Is there a genealogical hypothesis you wish to test? The most powerful use of DNA testing is to see if a conjectured genealogical relationship between two living people can be confirmed. Usually one of these living people is yourself or a surrogate within your family; the other is typically the descendant of someone you conjecture may be an ancestor. Confirming a match would go a long way to confirming the conjecture. No match would require a new session at the drawing board. See case studies 1 and 2.

For a less complex but still structured challenge, how about seeing if you can confirm that you are indeed the genetic descendant of each of your four grandparents, or more ambitiously each of your eight great-grandparents? (You were warned that DNA testing doesn't respect social decorum!) Taking each of your forebears one at a time, your challenge is to identify among your matches a tester, or preferably a cluster of them, who can be documented to descend from someone ancestral to your forebear. If you strike out on one of your (great-)grandparents, then that becomes a research project in its own right.

If you have no other hypothesis to test, how about trying to resolve a genealogical brick wall? Do you have an unknown ancestor in the last four or five generations? Many of us do, even genealogists who've been working on their ancestry for decades. How about setting the identification of that ancestor as the goal for your DNA testing program? See case study 4.

If you're one of the lucky few with no recent genealogical brick walls, let's set a goal of finding a surprise, something you had no idea of. Stay alert to the appearance on your match list of someone whose presence is unexpected. See case study 5.



An example of a chromosome browser: a comparison of the DNA of a tester (dark blue) and shared segments with each of three first cousins (tan, light blue, green), who are themselves siblings. Image has been truncated for display here. [David Pike, ISOGG Wiki / License)

Summarizing the tests

As noted above, there are three kinds of genetic genealogy tests: mtDNA, Y DNA, atDNA. There is no separate X-DNA test, as information about the X chromosome is typically included in autosomal results.

Some autosomal tests also give basic information about mtDNA and Y DNA, usually enough to establish basic haplogroups, sometimes a little more.

All the major testing companies offer autosomal testing. Only FTDNA does comprehensive mtDNA and Y-DNA testing.

More experienced testers will be aware that this Guide has simplified Y-DNA testing. There are actually two different approaches, the so-called STR ("short tandem repeat") and SNP ("single nucleotide polymorphism") tests. If you hear someone talking about matching at 37 markers or having a genetic distance of two at 67 markers, it's the more common STR test that's being discussed. The future of Y-DNA testing lies in the more precise, but costlier SNP tests. For a discussion of the two different kinds of tests, read here.

Please note that the major companies do not conduct separate ethnicity tests, more properly known as admixture or biographical ancestry analysis. Findings of ethnicity/national origin/etc. result from the application of proprietary algorithms to information gleaned from autosomal testing. For a more extensive discussion of ethnicity testing, read the sidebar about "admixture" on page 6.

C. How to proceed

Testing strategies for beginners

1. **Test earlier generations first.** Unless you yourself are in poor health, test surviving members of earlier generations before you test yourself. Once their DNA is gone, it's gone.

2. **Test now, analyze later if necessary.** Testing and analyzing test results are two distinct activities. If you don't have the time today for the analysis, test anyway. You can do the analysis in three months or three years or whenever. And if something should happen to you in the interim, at least your DNA is on record.

3. **Set a genealogical goal.** Most first-time DNA testers have no idea why they're doing it. They may have bought the kit on a lark or received it as a gift. A family historian may be succumbing to peer pressure or a growing sense that you're not really on the ball if you don't "do" your DNA. As the old saying goes, if you don't know where you're going, you'll never get there. A goal will give structure to your testing. Doing a DNA test with no goal in mind is a sure-fire shortcut to a "so what?" outcome. See the previous page for a discussion of how to set a genealogical goal that may be achievable with DNA testing.

4. **Select a test.** Take an autosomal test unless you wish:

    - to compare two or more people who are conjectured to be related on a strictly patrilineal bloodline, in which case do Y-DNA STR testing, with a minimum of 37 and preferably 67 or more markers, then follow up with autosomal testing as soon as funds allow

    - to compare two or more people who are conjectured to be related on a strictly matrilineal bloodline (apart from the testers themselves, who can be male or female), in which case do full-sequence mtDNA testing, then follow up with autosomal testing as soon as funds allow

    - to prospect for other living descendants of unknown patrilineal ancestors, in which case augment with autosomal testing as funds allow if the unknown ancestry carried into the past 200 years

5. **Choose a test supplier.** If you're doing Y-DNA or mtDNA testing, you will have to purchase your tests from Family Tree DNA (FTDNA). If you're doing autosomal testing, test with as many companies as you can afford. If you can afford to test with only one company, you're best off with Ancestry because of its huge database. Exceptions:

    - If you're Amish, a Newfoundlander, Jewish or a member of any other group that has been reproductively isolated for generations by culture or geography, do your testing with a company that provides a chromosome browser and access to segment-matching information, i.e., not Ancestry. In such populations, there is an unusually large amount of shared DNA irrespective of the familial closeness of two people. This is called endogamy. You will need to use advanced analytic tools if you hope to cut through the confusion.

    - If the test subject has difficulty producing saliva, test with a company that uses cheek swabs, i.e., Family Tree DNA or MyHeritage. Conversely someone who is feeble and may have difficulty in extracting a usable sample with a cheek swab might be better off with a saliva test, such as those available from AncestryDNA or 23andMe.

6. **Take advantage of holiday sales.** All the testing companies offer particularly good deals at certain predictable times of the year. Wait for a sale from the company you've selected, especially around Mother's/Father's Day, July 1st/4th, Black Friday/Cyber Monday, Christmas/New Year's.

7. **Test other family members.** In the long run, you will need more information to identify how you relate to some of your matches. An excellent strategy is to test other family members; given the luck of the genetic draw, some of them will have important matches that you don't. In addition, you can simplify the task of identifying your connection to your mystery DNA matches by seeing which of them you share in common with your tested first cousins, or even better, second cousins. First cousins have value because matches you have in common with them can usually be identified with confidence as belonging to either your maternal or your paternal side, thereby eliminating half the challenge of figuring out where they fit in. Second cousins are even better because any matches in common with them can usually be attributed to a connection ancestral to just one of your four grandparents.

8. **Fish in more than one pond.** If you are comfortable with the privacy implications, you should download your raw test data from your testing company, then upload to GEDmatch, where you will discover new matches who tested with other companies. Some testing companies, but not market-leader Ancestry, will allow uploads of test data from other companies for direct comparison with their databases. Policies vary from company to company and change from time to time; uploading and/or gaining access to matches may be subject to a modest fee.

9. **Use more than one type of DNA.** Think laterally about your problems, and don't fall into the rut of trying to solve every problem the same way. In case studies 2 and 4, Eliza-

beth used Y and mtDNA, as well as atDNA and reams of documentary research. In case study 3, I used atDNA to supplement my understanding of findings that Y DNA had supplied.

10. **Use Ancestry results differently from the findings from other companies.** It's fair to say that working with test results on Ancestry is qualitatively different from working with the other companies. For a more detailed discussion, read page 17.

11. **Keep a spreadsheet of your matches.** Keep a spreadsheet summarizing all key information about the matches you've already investigated and those you plan to pursue. Be sure to record not only test findings but also details of your communications with, and any conclusions about each match, including notes on shared ancestry, shared DNA segments, etc. You will not be able to remember all these details.

12. **Go back to basics, if necessary.** If you tested without an objective in mind or if these steps haven't got you launched for some other reason, it's time to go back to basics by rebooting.

Recovery guide for those who tested and were underwhelmed

So! You finally ordered a DNA test from a reputable organization. You swabbed your cheeks or deposited saliva in a tube, depending on the needs of your testing company. And you sent the collection kit to the lab, then waited, and waited.

At long last you got an email to say that your results were ready for online inspection. The first thing you saw was your ethnicity analysis. Unless you're an adoptee or have some other reason why you don't know the origins of one or more of your grandparents, you should have set this aside for the next time you're invited to a cocktail party.

Then you probably clicked on the link to your DNA matches. This is where you'd been told the genetic rubber hits the road. You had no specific expectations, no lost cousin in particular you were looking for. You weren't hoping to test a genealogical conjecture, just following the leads where they took you and maybe forming some hypotheses in the process.

And you cast your eyes down a list of people whose identities were masked by online handles such as "BananaLover" or "Anonymous Male." Each had an estimated relationship to you such as "2nd to 4th cousins," and maybe statistics about the DNA you share expressed in segments, percentages and/or centiMorgans.

Congratulations! You've "done" your DNA. But what the heck does it mean? Where do you go from here?

You weren't the first person to throw up your hands at this point and decide that DNA is not for you. If you had no clear objective in testing, there's no obvious next step. As the adage says, if you don't know where you're going, any road will do—and you won't get there anyway. The way out of this morass is to choose a road—or even construct one if necessary.

Here are some suggested next steps. They won't necessarily bring about a breakthrough, but you should at least be entertained enough to keep going until you start to get the hang of it and see some tangible results.

1. **Take the great-grandparental descendancy challenge.** Confirm that you are indeed descended from each of your eight putative great-grandparents. Go back to page 11 (second paragraph) for a brief discussion.

2. **Do you have anyone in your match list who is estimated to be closer than a fourth cousin?** If not, then you truly are unlucky in genetic genealogy; try testing other relatives who may have been luckier in the genetic draw. If you do have matches closer than 4th cousins…

3. **Quickly scan your close matches for a surprise.** Do you see someone familiar who is not to your knowledge a relative? Do not make the mistake of saying, "Oh, there's no way I'm related to Hazel. What a load of nonsense this is." You must take the match seriously. A different kind of surprise would be the discovery of a close relative you've never heard of, e.g. an uncle or aunt, first cousin or even half-sibling. Any of these mysterious matches would be a great place to start your inquiries. Proceed to Step 4 below. Probably though, you won't be this lucky, so…

4. **Go back to the first person in your match list.** Can you work out who this is from the clues at hand? Do they provide a family tree where you recognize a surname that also arises in your ancestry? Similarly, is there a geographic location in their ancestral past that also appears in yours? If you work out how you and this person are related, proceed to the next match on the list and repeat the process. Sooner or later, you will be stumped, it's time to…

5. **Send your match a note via the internal messaging system provided by the testing company**: "Hi, I see we're close genetic matches on [Testing Company] and I'm trying to figure out how we're related. My grandparents were Max Flax from Halifax, Mary Carey from Tipperary, Hamish Beamish from Squamish, and Mavis Davis from Nevis. Do any of these names or places ring a bell with you?"

6. **Repeat until you're tired of doing this.** Work your way methodically through all your close matches. Even if you can only assign a match to your maternal or paternal line, that will help down the road. With any luck, you now have a few lines of inquiry under active investigation.

D. Case studies

Note: *Some names have been slightly disguised at the request of those who've generously allowed their stories to be told.*

CASE STUDY 1

Confirming a hypothesis with autosomal DNA

Consider my friend Joy and her mysterious grandfather Moses Levine. He died in Alberta in the 1940s but his origins were always unclear. Family lore and credible records both affirmed that, as a young man in the 1880s, he had made his home in the Dakota Territory.

US censuses of the time identified him as having French Canadian heritage, but no record of him as a child or youth could be found in Canadian censuses or in Quebec parish records. The first breakthrough came when a young Moses Lavine of French Canadian parentage was located in the 1870 and 1880 censuses, a resident of upstate New York just a couple of miles from the Canadian border.

This "New York Moses" disappeared from the documentary record after 1880, just in time for the appearance of "Dakota Moses." Were they the same person?

Autosomal DNA to the rescue! In Joy we had a direct descendant of Dakota Moses. If we could find a living person known to be genetically related to New York Moses, we could then use autosomal testing to see if there was any biological connection between the two lines.

After several false starts, we eventually identified Gloria, a living great-granddaughter of the youngest sister of New York Moses. Fortunately, Gloria set aside her misgivings and agreed to take an autosomal DNA test. After a suspense-filled wait for the findings, the testing company revealed that Gloria and Joy share DNA in an amount expected of second or third cousins.

Bullseye! According to our conjectured family tree, Gloria and Joy are second cousins once removed, genetically halfway between second and third cousins.

CASE STUDY 2

Refuting a hypothesis with autosomal DNA

My friend Elizabeth, an adoptee, had good reason to believe that her biological father was Leslie Elmwood, an Englishman who had died in the 1990s.

Information about his life corresponded in every way to the snippets of truths, half-truths and fading reminiscences that Elizabeth had gleaned from Vi, her taciturn biological mother. A decorated nursing officer, Vi had served in the renowned Burma Campaign during the later years of the Second World War. That's where Elizabeth had been conceived in 1945.

Shortly after autosomal DNA testing first became affordably available to the general public, Elizabeth quickly established that she bore no genetic relationship to Leslie's surviving brother.

In an earlier era, Elizabeth would have continued indefinitely in the mistaken belief that Leslie Elmwood was her biological father. It's no wonder relevant genetic evidence is now expected as part of any genealogical research project worthy of the name, viz. the Genealogical Proof Standard.

On average a putative uncle (such as Leslie's brother) and niece (Elizabeth) should expect to share 25 per cent of their autosomal DNA. The absence of any genetic relationship at all, while shocking, was decisive. Don't cry for Elizabeth, though. The telling of her story has barely begun.

CASE STUDY 3

Confirming a hypothesis with Y DNA

My grandfather Jones was born out of wedlock in Dorset, England in 1891. The father was not identified in any record, and my grandfather was given his mother's surname. Family lore was mute on the subject of the father's name but did provide some meagre clues, such as his supposed occupation.

After decades of research interspersed with long periods of inactivity, I developed a working hypothesis that a man named Charles Davis was my grandfather's father. The case was persuasive but circumstantial, hardly the final word on the subject.

For at least a decade, it had been clear that DNA testing might offer hope — if we could find a living descendant of Charles Davis. The problem was that we had no idea what had happened to Charles Davis after his appearance in the 1891 census.

There the matter rested until Wiltshire parish records were digitized and published on Ancestry.co.uk. The bad news was that Charles Davis was killed in an accident in Wiltshire just months before the 1901 census.

The good news is that he had married in 1895 and had fathered four children. Three of these children themselves had children, two of them quite large broods. After much further research, I eventually located a living patrilineal grandson of Charles Davis who was willing to be tested.

Alan Davis, my putative half first cousin once removed, was connected to Charles Davis through an all-male line, just as I had conjectured that I am. If the conjecture was correct, we should be very nearly perfect Y-DNA matches. And so it transpired! On FTDNA's basic Y-DNA test, Alan Davis and I matched perfectly at 67 out of 67 markers.

For a more famous use of Y DNA, read this article from *The New York Times*.

CASE STUDY 4

Developing (and then confirming) a hypothesis with autosomal DNA

So how did Elizabeth (case study 2) identify her real biological father? It's a story that deserves a book, not a couple of hundred words, but here's the gist.

Noticing that she had a large number of autosomal matches from the Southern US, unlike her maternal half-siblings, Elizabeth hypothesized that her biological father had a genetic connection to this area. She compiled a candidate database of roughly 300 American servicemen, mostly pilots and ambulance drivers, whose stationing in the spring of 1945 might have brought them into contact with her mother.

Elizabeth researched and constructed family trees for several dozen of the men who seemed most promising, especially those from the Southern US. At the same time, she compiled a compendium of surnames found in the family trees of her autosomal DNA matches, in particular those seemingly from her paternal side.

At this point then, Elizabeth had two surname lists, one from her DNA matches, one from servicemen who were candidates to be her biological father. Elizabeth then looked for commonalities between the two sets of names.

After untold hours of work, Elizabeth identified a US pilot who had ancestral surnames on both sides of his family that corresponded to names in the compendium prepared from Elizabeth's genetic matches. This is exactly the finding we would expect if the man was Elizabeth's biological father.

Still, suspicion was not proof. Regrettably the man had died in the 1990s, and Elizabeth was unable to test his DNA directly. But he had two daughters who were alive and well in Florida.

To settle the matter beyond any reasonable doubt, Elizabeth persuaded the daughters to undergo autosomal testing. Both were shown beyond any statistical doubt to be half-siblings to Elizabeth. Game, set and match!

Here's another adoption story that involved a lot of sleuthing, testing and conventional genealogical research.

CASE STUDY 5

Developing a completely unexpected hypothesis from autosomal DNA

My friend Linda recently discovered a pair of brothers who were genetic matches not only to herself but also to her mother and several other relatives.

Indeed, these brothers were estimated to be Linda's mother's first cousins or first cousins once removed or some relationship of similar degree of closeness. Yet no one knew who they were or how they were connected to the family.

After a painstaking review of all the available information, including time-lines, DNA connections to other tested relatives, and family information from the brothers, it seems likely that there had been an illegitimacy involving Linda's maternal grandfather or his brother.

At this late date, a more precise explanation of the findings is unlikely to emerge. Still, without DNA testing, Linda's connection to the brothers and the need for an explanation of any kind would have been beyond all imagining.

E. Assorted observations on interpreting DNA tests

- DNA does not lie, although it may prevaricate. A finding of close genetic relatedness is not something that can be dismissed as nonsense, as some are disposed to do if the finding conflicts with their world-view.

- DNA does not observe social decorum. If you don't want to learn about skeletons in the family closet, don't get tested.

- Contrary to popular belief, DNA results do not provide absolute proof or disproof of any hypotheses. In Elizabeth's initial search for her biological father (case study 2), there was another possible, albeit improbable, explanation for the finding that she had no genetic connection to her putative father Leslie's brother. What if either Leslie or his brother had been adopted? Fortunately, we have…

- … Occam's Razor. Given several explanations for a set of facts, the simplest one is most likely to be correct. How does one define simple in this context? The simplest explanation is the one that requires the fewest assumptions.

- DNA tests are not perfect. There can be false positives when shared DNA is assumed to derive from a common ancestor but actually depends on chance or a limited genetic pool (endogamy). More likely are false negatives, although this is something of a misnomer. Current autosomal tests miss about 10 per cent of third cousins, 50 per cent of fourth cousins and substantially higher proportions of more remote relations. This does not arise from any imperfection in the test but for the simple reason that our genetic family tree is a subset of all our ancestors. Not all our ancestors have bequeathed us their DNA.

- For this reason, current autosomal DNA tests have limited use when scouting for specific ancestors more than about 200 years ago. There's no hard and fast rule here. Some of your ancestors from 150 years ago may not have bequeathed you any of their DNA, and you may have an extra dollop from someone 100 years their senior. But by and large, autosomal DNA runs out of steam after a certain number of recombinations. Certainly, you have DNA from very many ancestors, but these inheritances are often in small enough chunks that they cannot be reliably assigned to any one person.

- Notwithstanding the previous statement, autosomal DNA may have use in older problems where the documentary record is sparse or non-existent, viz. Irish ancestry in the 18th century. Persistent or unusual segments of atDNA, even if tiny, may help narrow your research to one or two surnames or locales. It would help of course if you could reframe your problem so as to bring Y or mtDNA into play.

- Don't fall into the trap of compartmentalizing your DNA testing from your other genealogical research. The two should go hand in hand. When you encounter a problem in your conventional research, you should always question whether DNA testing might help resolve it. And when you're puzzling over your DNA results, ask yourself what aspects of the case could be resolved by recourse to conventional documentary research.

- Bear in mind that relationships estimated by DNA testing companies are only approximate. There's a natural range of variation and several very different kinds of relationship may produce similar DNA findings. For example, your DNA match with a half-sibling would be pretty much indistinguishable from that with a niece or nephew or aunt or uncle. That's quite a range. And the farther out you go, the wider the range of possible relationships. For helpful guidance on this subject, visit the Shared cM Project.

The Shared cM Project – Version 3.0
August 2017

Cluster	Relationships
Cluster #1	Siblings
Cluster #2	Half Sibling, Aunt/Uncle/Niece/Nephew, and Grandparent/Grandchild
Cluster #3	1C, Half Aunt/Uncle/Niece/Nephew, Great-Grandparent/Great-Grandchild and Great-Aunt/Uncle/Niece/Nephew
Cluster #4	1C1R, Half 1C, Half Great-Aunt/Uncle/Niece/Nephew, and Great-Great Aunt/Uncle/Niece/Nephew
Cluster #5	1C2R, Half 1C1R, 2C, and Half Great-Great-Aunt/Uncle/Niece/Nephew
Cluster #6	1C3R, Half 1C2R, Half 2C, and 2C1R
Cluster #7	Half 1C3R, Half 2C1R, 2C2R, and 3C
Cluster #8	Half 2C2R, 2C3R, Half 3C, and 3C1R
Cluster #9	Half 3C1R, 3C2R, and 4C
Cluster #10	Half 3C2R, 3C3R, Half 4C, and 4C1R

Table of relationship clusters (truncated) in Shared cM Project. Courtesy of Blaine Bettinger, "The Genetic Genealogist." CC4.0.

- Even full-time professional geneticists have difficulty keeping up with all the developments in the field, so you shouldn't feel inadequate if you find this all overwhelming. My top pieces of advice are: spend a lot of time poking around on the ISOGG Wiki; join a DNA working group, either online or in your community if one exists; follow one or more genetic genealogy bloggers.

- Make provision in your will for who will serve as the custodian of your genetic genealogy information. If you do not do so, your test results could pass eternity in electronic limbo of no use to anyone.

Analyzing DNA matches at Ancestry and elsewhere

The conventional first step with all testing companies is to search among your matches for people who share ancestral surnames and locations. Very often with Ancestry, this technique alone will be sufficient to make breakthroughs. It does, after all, boast 60 per cent of the total number of autosomal tests conducted to date.

To resolve mysteries at companies other than Ancestry, the focus is on matching segments of DNA via tools such as chromosome browsers (offered by all companies except Ancestry), as well as third-party utilities like those provided by GEDmatch, DNA Painter and others.

The goal is to achieve some or all of the following objectives, which can lead to identification of specific relationships and shared ancestors:

- identify the ancestral origin of each segment of your DNA (preferably at least as far back as your four grandparents)

- reconstruct the genomes of your parents and grandparents, to the degree possible, not only from your tests, but also those of other close family members

- categorize your matches by the DNA segments they share with you and therefore the quadrant of your ancestry (i.e., grandparent) where your trees overlap

- build sets of matches who share overlapping segments with each other and who are thus all interrelated

Combine these insights about who is related to whom with conventional genealogical research, and you are well on your way to discovering the common ancestors you share with your matches.

Ancestry's approach is dramatically different. The company believes that if testers provide their most complete family trees linked to their DNA test results, Ancestry algorithms can yield spectacular results when applied across their huge database.

Consider "DNA circles," i.e., webs of testers identified by Ancestry as connected by overlapping DNA and family trees. "Genetic communities" are groups of testers "who likely descend from a population of common ancestors," whether through continuous residency in a localized area or a shared past migration. Either of these tools could provide in a moment the breakthrough you need to resolve a long-standing genealogical conundrum.

Metaphorically, Ancestry doesn't want you to have to learn to fly the plane—and there's a lot to be said for that if you're a busy person without the time to become an expert. But if you need to get to an out-of-the-way location, you won't be able to fly there on Ancestry. (Let's bail out of this metaphor right now before it crashes.)

In short both approaches have relative strengths. The ideal is to be sufficiently knowledgeable that you can switch effortlessly between them as circumstances demand.

F. More resources

Selected recent publications

NB: Check WorldCat.org for holdings in libraries; confer with your local librarian about interlibrary loans. To purchase online, check with the author's website, Global Genealogy, Chapters-Indigo, or Amazon.

Bettinger, Blaine T. *The Family Tree Guide to DNA Testing and Genetic Genealogy*. Cincinnati, OH: Family Tree Books, an imprint of F+W Media Inc., 2016. ISBN 10: 1440345325 ISBN 13: 9781440345326

Bettinger, Blaine T. and Debbie Parker Wayne. *Genetic Genealogy in Practice*. Arlington, VA: National Genealogical Society, 2016. ISBN 10: 1935815229 ISBN 13: 9781935815228

Gleeson, Maurice, MB. *DNA & Your Genealogy*. Ajax, ON: Moorshead Magazines Ltd., 2018. Visit Your Genealogy History Store and search for "Tracing Your Ancestors Series." Available in both print and PDF.

Basic information about genetic genealogy

International Society of Genetic Genealogists (ISOGG) Wiki
Cyndi's List, "Genetics and Family Health"

Blogs by notable genetic genealogists (a selective list)

Blaine Bettinger, "The Genetic Genealogist"
Kitty Cooper, "Kitty Cooper's Blog"
Roberta Estes, "DNAeXplained"
Richard Hill, "DNA Testing Update"
Debbie Kennett, "Cruwys news"
Leah Larkin, "The DNA Geek"
CeCe Moore, "Your Genetic Genealogist"
Judy Russell, "The Legal Genealogist"

Tools and utilities

Genetics glossary at ISOGG Wiki
Autosomal DNA transfers
Cousin statistics
The Shared cM Project
GEDmatch
DNA Painter
Canine DNA testing (Wisdom Panel 4.0)