

How to Use DNA in Your Genealogical Research

by William Remus¹

DNA analysis can be used to solve family history mysteries and to fill gaps caused by incomplete and destroyed ecclesiastical and civil records. In this article I first talk about how DNA functions. This leads to the possibilities of analyzing the Y-chromosome (Y-cs) to follow the paternal line or analyze the mitochondrial DNA to follow the maternal line. Case studies are presented to show Y-cs analysis. Details like choosing an analysis DNA lab, getting the samples, and costs are discussed. Also I discuss other interesting results available through DNA analysis.

What genealogical questions can be answered using DNA analysis?

No matter how thorough your work on paper and microfilm records, there are often gaps in your family genealogical records. This is particularly likely in Eastern Europe since wars and destruction have frequently plagued the area. So it is not unusual to find that the needed church books and civil records are missing or destroyed. And if your work is before 1700, the records may have never existed. DNA can help solve such problems.

Also, you may think that you are related to others with the same or a similar surname since there are many similarities between your family and the other family. If you cannot cross the Atlantic with your and their documentation to establish that connection, DNA will do it and find out if there is a common ancestor.

Also there are often family mysteries that you may wish to solve. For example, does one's line result from the usual biological processes or from an adoption? Or are members descended from a famous relative or ancestor like Genghis Khan (eight percent of the men in Central Asia have Genghis as an ancestor). Or perhaps you wonder if the stories about the Native Americans in your family are true. Again DNA can help.

In this article, I will explain how DNA analysis can be used in each of these instances. First, I will provide the basic facts about DNA. Then, I will present two case studies to illustrate how to set up the data gathering to answer these questions. Lastly, I will present the details of the process of getting the analysis done.

How DNA functions

DNA is the blueprint of life. It determines many of the characteristics that people have like the color of their eyes and their height. The DNA of any child comes in equal parts from both parents. Or put another way, the DNA in the nucleus of each cell in a child contains forty-six chromosomes; twenty-three of these come from the child's father (these chromosomes are contained in the father's sperm) and twenty-three come from the mother (these chromosomes are contained in the nucleus of the egg

provided by the mother). So a child has the DNA from both parents.

One of the twenty-three chromosomes provided by the father determines the sex of the child. If one particular type of chromosome is provided (termed the X chromosome), the child will be a girl. If another type of chromosome is provided (termed the Y-cs or YDNA); a boy will result. So boys and girls always differ in one of the forty-six chromosomes. If you are interested in learning more about YDNA, see the book by Bryan Sykes.²

A boy gets his Y-cs from his father, and his father got that same Y-cs from his father (the boy's grandfather). And his grandfather got the same Y-cs from his father (the boy's great grandfather). And so on. So the Y-cs follows the surname line and allows us to establish whether two men with the same surname have a common male ancestor. If they have that common surname ancestor, they will have the same YDNA.

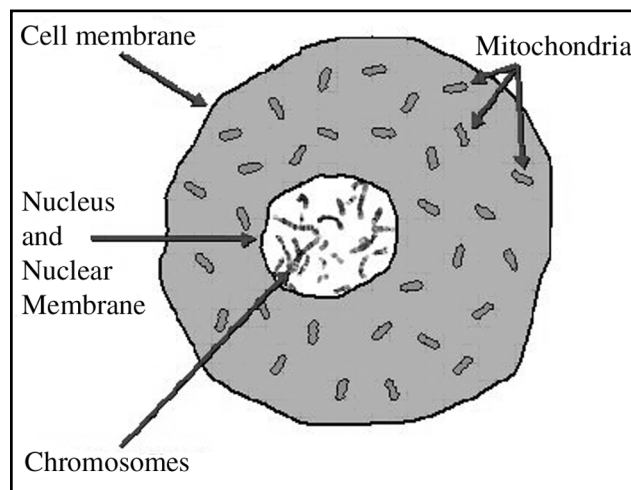


Fig. 1 - A human cell diagram indicating locations of nuclear and mitochondrial DNA

Occasionally, the Y-cs undergoes a small change (termed mutation). Should this happen, then two men with the same surname having a common male ancestor will both have that mutation if the mutation happened prior to the birth of their common ancestor. But the two men with the same surname having a common male ancestor may differ slightly in their YDNA if the mutation occurred after the birth of their common ancestor. This helps us date when the common male ancestor lived.

There is more DNA in a cell than that which resides in the nucleus of the cell. There is also DNA in the mitochondria inside the cell but outside the nucleus (see fig. 1). In conception, both mother and father provide chromosomes for the DNA in the nucleus but the

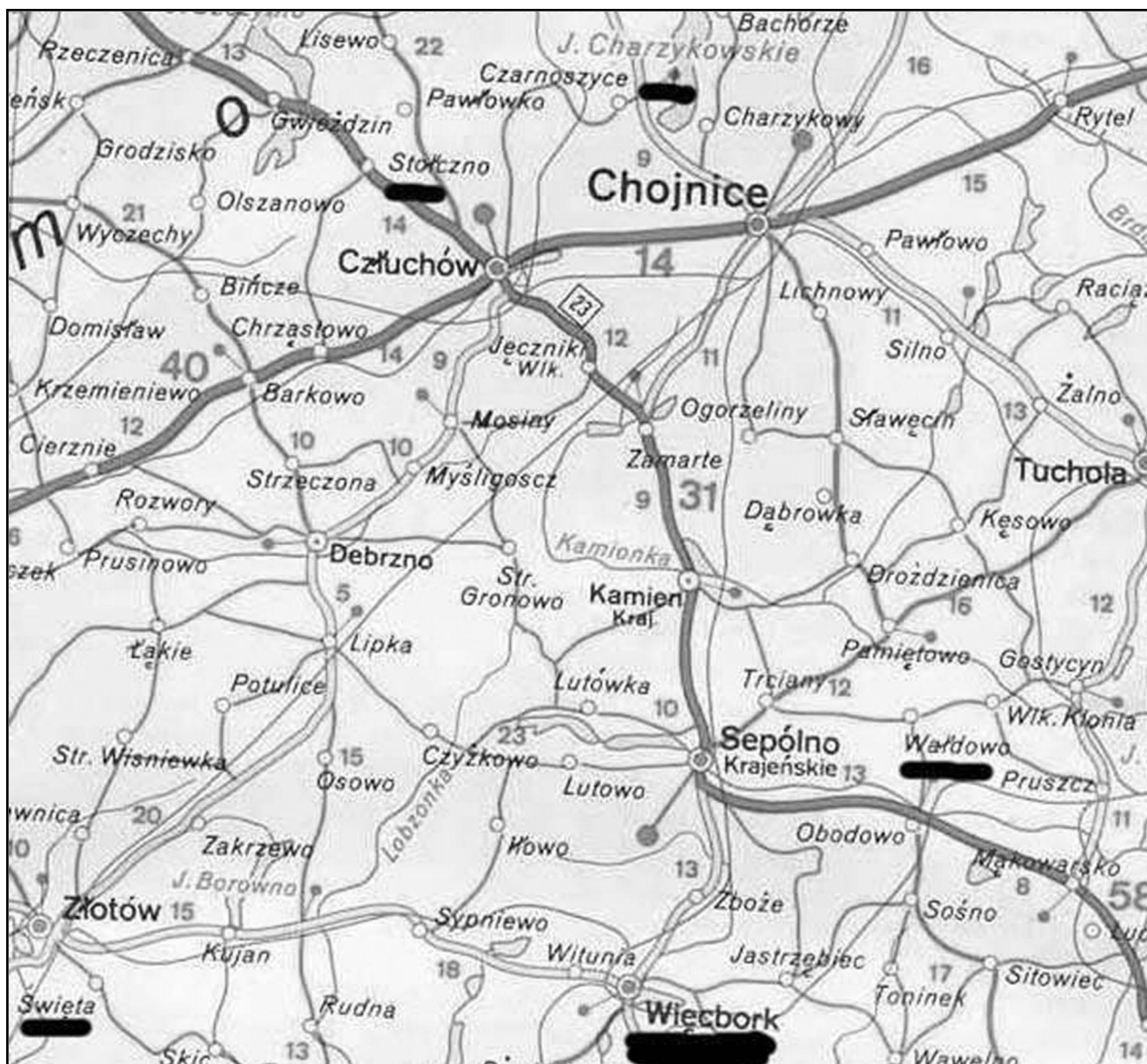


Fig. 2 - Map showing probable environs of Franz Remus' father

mitochondrial DNA (mDNA) is in the egg provided by the mother. So the mDNA for a boy or girl comes from his or her mother. And the mother's mDNA comes from her mother (the children's maternal grandmother). And so on. Thus the mDNA for either a boy or girl follows the maternal line and allows us to establish if two people have a common female ancestor along their maternal lines.

In either the case of YDNA or mDNA, the DNA is extracted from the cell and sequenced. That is, the DNA is characterized by counting genetic markers on the DNA strands. The most common YDNA test finds the frequency that the twelve most important markers occur; other versions of the YDNA test find the frequency of the twenty-five or thirty-seven most important markers. People with a common ancestor should have almost a perfect match in their marker frequency.

What DNA can tell you

Case study one: who is Franz Remus' father?

Often times it is difficult if not impossible to have paper documentation linking people of the same surname. Consider the case of Franz Remus who was postmaster and sometime *Schultz* (village head) of the town of Vandsburg, *Kries Flatow*, West Prussia; he was born in 1755 according to an article of compiled genealogy published in the *Deutsche Geschlechterbuch*.³ He married the daughter of the *Schultz* of Runowo, a village located about ten miles southwest of Vandsburg. The article provides that family tree for many generations of Remus family members, but there are other Remus families not documented in the article (including mine). So it would be very interesting to find Franz's ancestors including an ancestor that we might have in common.

The first step in this process is gathering all the ordinary genealogical data to try to link Franz the Postmaster's family to the other Remus families of West Prussia. Since Franz was born in 1755, his father would be in the West Prussia 1772 Land Census. Thus, my first step is to consult the online West Prussia 1772 land census at <www.odessa3.org/collections/land/wprussia> to find candidates who might be the father of Franz. The possible candidates for Franz's father were within fifty miles of Vandsburg and were selected by consulting the church books for each village and seeing if they had a son named Franz who was born around 1755. Here are the candidates:

- Peter Remus the Shepherd in Waldowo
- Franz Remus the relatively wealthy farmer and sometime *Schultz* in Stretzin
- *Schultz* Martin Remus, village head of Schwente

All three of the above did and other nearby candidates from the census did not.

The map in fig. 2 shows clockwise from the top Stretzin (home of Franz the farmer around 1760); Rozollen where that Franz the farmer retired around 1790; Waldowo the home of Peter the Shepherd; Wiecbork (Vandsburg) the home of Franz the Postmaster; and Swiete (Schwente) the home of Martin of Schwente. The distances are in kilometers.

Now consider what we know without DNA analysis.

Candidate one: Peter the shepherd lived nearby (about twenty kilometers north) but shepherds were usually poor and their children had little chance for an education. His son Franz did marry a villager and had two children. Then he and his family disappear from the records. But the disappearance is before the marriage of Franz the Postmaster, so maybe Franz the son of Peter might be Franz the postmaster and take a new wife and have a new family. So in theory it is possible but would the daughter of a village head marry a shepherd (recall Franz the postmaster was well married)? Would a likely illiterate shepherd be a postmaster and *Schultz*?

Candidate two: There was also a Franz Remus who was a relatively wealthy farmer and sometime *Schultz* in Stretzin (about seventy kilometers northwest). He had a son Franz who was about the right age. There are no records of his son Franz marrying or having children. So this Franz might have moved to Vandsburg, married the daughter of the *Schultz* of Runowo and become the postmaster of Vandsburg. Certainly the marriage would have been appropriate and his father Franz was in a financial position to provide money to assist the Prussians in selecting his son Franz as postmaster. And son Franz might have had some opportunity for some education. However, when we look at the birth records for children of Franz the postmaster, we find no godparents from the family of Franz the farmer. Franz the farmer's other children were still living and well married and one would expect them to appear as godparents at least occasionally. Also we do not find Franz the Postmaster as witness to any

of the children of those who would have been his nieces and nephews if he were the son of Franz the farmer.

Candidate three: There was also a Franz Remus who was the son of a relatively wealthy farmer and *Schultz* Martin Remus of Schwente (about thirty kilometers west). Martin had a son Franz who was about the right age. Martin's son Franz married the daughter of the *Schultz* of Schmirdowo and had one child. He worked for the local Polish nobility as a scribe and also inherited and then sold the right to be *Schultz* of Schwente after his father and older brother's death. So this Franz might have moved to Vandsburg, married the daughter of the *Schultz* of Runowo and become postmaster of Vandsburg assuming his first wife died. Certainly the marriage would have been appropriate and this Franz had the money to motivate the Prussians to select him as postmaster. And this Franz was fully literate. When we look at the birth records for children of Franz the postmaster, we find no godparents from the family of Martin. Also we do not find Franz the postmaster as witness to any of the children of those who would have been his nieces and nephews if he were the son of Martin of Schwente. It is fair to say, however, that most of his siblings died young.

My reading of the evidence above ranks the candidates for father of Franz the postmaster in this order of likeliness:

- Most likely: *Schultz* Martin Remus the village head of Schwente
- A possibility: Franz Remus the relatively wealthy farmer and sometime *Schultz* in Stretzin
- Unlikely: Peter Remus the Shepherd in Waldowo

Now let's see what the YDNA analysis tells us about this question. First we need to get YDNA samples from living male descendants of each line. Fortunately, I correspond with artist Marc Remus of Köln and he was willing to provide a YDNA sample from the line of Franz of Vandsburg. And I am from the line of Peter the Shepherd. Also mathematician Horst Remus is from the line of Martin Remus of Schwente (his ancestor would have been another son of Martin). I cannot find a living male descendent of Franz of Stretzin.

The samples revealed that Marc Remus of Köln, the descendent of Franz the Postmaster, and Horst Remus, the descendent of Martin of Schwente, had the same number of each of the twenty-five YDNA markers. I, a descendent of Peter the Shepherd, differ greatly from them.

Interpreting the results

It is almost certain (more on this in a moment) that Franz the postmaster shares a common ancestor with Franz the son of Martin of Schwente. The genealogical evidence allows us to conclude that Franz the Postmaster and Franz the son of Martin maybe one and the same person. To determine how likely this is, we need to consult the following chart from Family Tree DNA (fig. 3).

The probability is estimated by choosing the curve for the number of markers used in the YDNA test (twenty-five in our case so use the middle line) and then selecting the

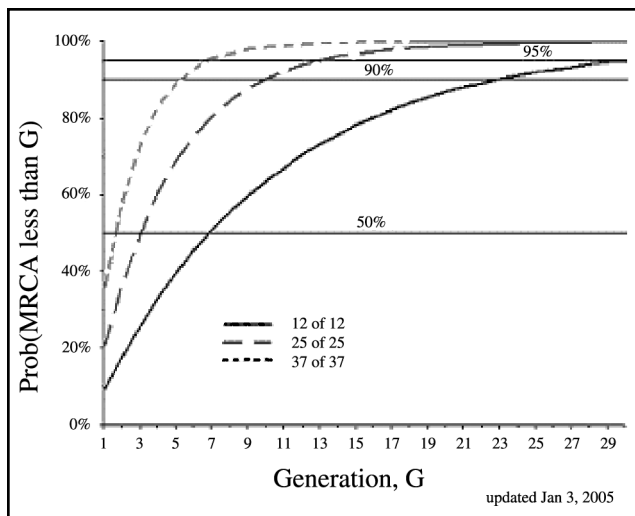


Fig. 3 - Family Tree DNA probability chart

number of generations ago the event occurred (nine in our case). The probability is then read off the vertical axis (in our case about ninety-five percent). That is, we are ninety-five percent certain that they have a common ancestor who is from our paper genealogy Martin Remus of Schwente.

Could a grandson rather than a son of Martin be Franz the postmaster instead? Yes, that is theoretically possible since the grandson would have the same YDNA. But we find no such grandson in the genealogical records who was born at about the right time and have the name Franz.

Could another nephew or grand-nephew of Martin be Franz the Postmaster instead? Yes, that is theoretically possible since the nephew or grand-nephew would have the same YDNA. But we find no such nephew or grand-nephew in the genealogical records who were born at about the right time and have the name Franz *unless* Franz of Stretzin is the brother of Martin and Franz of Stretzin's son Franz is Franz the postmaster. Then the common ancestor (the father of Martin of Schwente and Franz of Stretzin) is back eleven generations.

Can we rule out the possibility that Franz of Stretzin is the brother of Martin and Franz's son Franz is Franz the postmaster? No, we can't rule it out using DNA. It is true that the genealogical paper records are not supportive. We cannot prove Martin and Franz of Stretzin are brothers and additionally we cannot provide that Franz's son Franz is Franz the Postmaster. To do that we would need a YDNA sample from a male descendant of the Franz of Stretzin line (I have yet to identify such a person).

It is clear that the descendants of my line from Peter the Shepherd do not share the DNA of Martin and Franz. This could mean that Peter's ancestors adopted the Remus name independently and so do not have a common ancestor with Martin and Franz. It could also mean that someone in the line between me and Peter was adopted or illegitimate. But the line between me and Peter looks pretty clean so this is not likely. But, nevertheless, could Franz the Postmaster be the

son of Peter the Shepherd? Yes, it is possible that Franz the Postmaster is a biological descendant of Peter (excluding my possible adopted or illegitimate line) might have common YDNA and thus be related. This would mean Peter the Shepherd, Martin of Schwente, and Franz the Postmaster would all be related. Again we have to look to paper genealogical evidence which finds little support for this. But it is possible.

Notice in the above analysis the YDNA is much better in disproving rather than proving relationships. That is, there are always possibilities that are hard to rule out with YDNA.

One of the big issues in DNA research is getting the right DNA sample. Consider my family tree below:

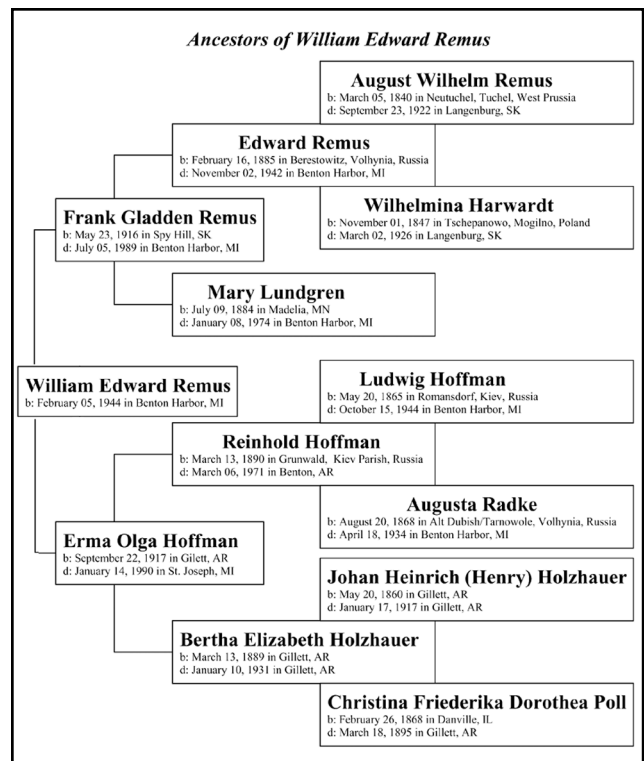


Fig. 4 - Family tree of author

To follow a surname like Remus, you need a sample of a male along that surname line. Being a male with the surname Remus, I am the right person to give a Remus sample. However, I do not have YDNA from my mother's Hoffman line. So to get the right YDNA sample for the Hoffmann surname line, I would need to ask my maternal grandfather, uncles, or nephews with a Hoffman surname for a sample. If I was interested in my paternal grandmother's Lundgren line, I would need a great grandfather, great uncle, uncle, or nephew with the Lundgren surname to provide a sample.

To follow my maternal line, I already have the right DNA in the form of my mDNA. This mDNA is common with my mother Erma Hoffman, my maternal grandmother Bertha Holzhauser, and all great grandmothers along that maternal line. As with the YDNA analysis, things get a little

more complicated when off the main maternal or paternal line. For example, one of my maternal great grandmothers is Bertha Radke. To follow her mDNA maternal line, I would need a sample from a male or female descendant of the Radke family since they would have the right mDNA.

Case study two (in progress)

Are Gottlieb Krassin and Martin Krassin lines related? I am related to a Krassin family by marriage. The family traveled with my mother's to Arkansas from Volhynia in 1892 and before that from the Poznan area to Volhynia in 1833. When I work backward through the church books, the earliest Krassin birth I can find is Martin Krassin born 7 October 1782 in Schokken (modern Skoki) near Poznan, Poland to Johan Krassin and Christina Pflugard.

There is however another Krassin family in America. This family, the Gottlieb Krassins, came from Radwanki (about ten kilometers) from Schokken and settled in Waseca County Minnesota. Try hard as I could, I cannot find church books that show a common Krassin ancestor for the two lines. The 1772 Netze area census finds only one Krassin within 100 miles and he is located in Rattai near Schokken and Radwanki. So it makes sense that the two families would be related since they have an uncommon surname and lived very near each other in Poland. Since we are interested in two families with the same surname, YDNA analysis should be able to establish if the two families are related. Getting such a sample has been a problem as is discussed in the next section. For a completed DNA study of this kind, see either Bryan Sykes⁴ or Megan Smolenyak-Smolenyak.⁵

Once you have willing donors, getting the sample is easy. The DNA processing company provides the tools; it is like brushing your teeth except you brush the inside of your cheek for sixty seconds two times (no sperm samples).

The lengthy plastic brush gathers the sample. The head of the brush is broken off inside the small storage bottle. This is done twice and mailed for processing.

Another impediment to doing DNA analysis is the cost of each sample. For the Krassin study, only a twelve sample YDNA test is needed so the total cost is two times \$120. Usually the people interested in the answers pay some or all of the costs. In both cases, I was the interested party and covered the costs.

Choosing the test and DNA laboratory to do the work is another issue. Usually the choice of a mDNA or YDNA test is determined by the genealogical question you are asking. If YDNA, the number of markers you want to use depends on both the genealogical question and the confidence you want in the results. In study two, the Krassin study, we are seeking evidence that the families had a common ancestor and, if so, the YDNA should match. So a twelve marker test will do. In study one, the Remus study, it was possible that these family lines actually merged at some early date. So a twenty-five marker test would be more sensitive to those varying family lines and when the common ancestor occurred and well as giving more confidence in the results.

Choosing the right processor of the DNA is not as complicated as it seems. There are just a few processors and they seem to have differing special interests (for example, African American ancestry). All the processing will be

DYS 393	DYS 390	DYS 19/394	DYS 391	DYS 385a	DYS 385b	DYS 426	DYS 388	DYS 439	DYS 389-1
13	24	15	10	11	14	12	12	11	13
DYS 392	DYS 389-2								
13	29								

Fig. 5 - A sample of YDNA results

The steps in doing the DNA analysis

In this section I will detail both the steps and issues in doing DNA analysis. The first big issue is getting the DNA sample. Getting the right samples for a project is theoretically simple but actually the most difficult part. For example, in study two above, just one sample was needed from a male with a Krassin surname from each Krassin line. This is the major problem if, like me, you do not have strong relationships with the males involved. There is something very personal and maybe even mystical about DNA samples. Further, a man might worry that the sample could lead to a problem if he had a "sowing the wild oats" phase (a needless worry but a haunting worry). Perhaps the easiest solution is to use samples from male children of the right surname that are too young for wild oats and haunting memories.

competently done so it amounts to price which is roughly the same and the kinds of support they provide in terms of interpreting and comparing the results (more on this later).

The DNA processor will provide you with a list of the DNA markers and the frequency that these DNA markers occur (see fig. 5 above).

In the above table you can see the marker names and below them the frequency that those markers occurred. If two men have the same YDNA, the table will be identical for both. The more the tables differ the more unlikely that they have a common ancestor. In study two, the Krassin study, the two lines are likely to either march perfectly or to be different. Because of this, the results are not likely to be ambiguous.

You may also get updates from your DNA lab as others do DNA testing. For example, I found several people with

DNA similar to me (recall I am a descendent of Peter the Shepherd who lived in Northern Poland). Although I thought my paternal line was German (with some pretty Polish girls for spice), the matches were almost all Slavic (Poles and Czechs). Well, that works since the Remus family left Saxony in the 15th century and the people of Saxony are a mix of German and Wends (a Slavic group). But, I have a twenty-five marker perfect fit as a Kasubian (a Slavic group living in Northern Poland). So maybe Peter the Shepherd's ancestors took Remus as a family name while some people in Saxony were also taking that family name.

Planning your DNA study

I think the best course of action for any kind of study is to carefully plan your study so it clearly answers genealogical questions that are important to you. In case study one, the Remus study, I wanted to link together several large Remus family groups in West Prussia and hoped to find their common ancestor if there was one. In study two, the Krassin study, I had tried to link the two Krassin families. So DNA would be appropriate for either case.

Another important point to consider is that the DNA analysis may provide answers that you really don't want to know. For example, would you want to know that your biological father wasn't the man who raised you (and mom kept this secret)? Or perhaps a family believes that they have an American Indian ancestor and that is important to them. What if it turns out that this is not true?

Many people do not have clear objectives and do a DNA test based on curiosity. The problem here is that you often don't find out much other than that others have similar DNA. And thus the DNA study is interesting but your family research does not progress. Once, however, you have answered the research question you want, like study one, the Remus study, or two, the Krassin study, it is certainly interesting to find out more about other people based on DNA. There are basically two kinds of these studies that I call exploratory studies.

Exploratory study type one:

Who has similar DNA and what implications result?

In this situation, you choose a popular DNA processor and have your DNA analyzed. Then you use their Web site to attract others with your surname to participate. In my case, I used Family Tree DNA and have this Web site: <www.familytreedna.com/surname_join.asp?code=X53983&special=true>. Others have contacted me particularly if they had similar surnames (like a possible Polish version of the family name Remus) and DNA testing will tell them if their family is related to mine. I also have a Web page on my personal genealogical Web site should people visiting my site wish to be tested. It is <www.cba.hawaii.edu/remus/genes/dna>. However interesting this has been, it has not moved my personal family research forward.

Exploratory study type two: does my DNA fall into a broad category that will tell me more about my family history in ancient times, and what implications result?

These kinds of studies are really interesting. For example, it has been argued that in spite of the great variety of current mDNA in the cells of the people of the world, there were originally only seven original mDNA and hence seven women from which all people descend (Eves if you wish to think Biblically). This research finds which Eve you descend from and tells where that Eve lived and how she differs from the other Eves. See Sykes book for more details.⁶

Another variation on this is based on the YDNA. The idea here is that any man's YDNA comes from an initial Adam. But over the millennia, mutations have arisen. So while it is unlikely a boy and his father differ in YDNA, a boy and his ancestors may differ in YDNA due to mutations. If you research the history of mutations, then you can estimate the approximate time and place that those mutations occurred. So a man's paternal line's history is coded into his YDNA mutations. Decode the YDNA and you have the path your YDNA took from Adam to today. This is the gist of the National Geographic Genomonic Program (<www3.nationalgeographic.com/genographic>). Incidentally, the DNA processor for the National Geographic project is Family Tree DNA, so samples provided to either can also be used in either or both programs.

Another interesting DNA variation reflects the genealogical interests of African-Americans. African Ancestry (<www.africanancestry.com>) sampled the mDNA and YDNA of many tribal groups in Africa. They then match people with the African tribal groups based on YDNA or mDNA or both.

So as you can see, these DNA projects are very interesting indeed. However, they don't add any branches to your family tree. Only a planned study with specific objectives will do so.

Conclusions

DNA analysis is a very good tool for solving various genealogical mysteries. And if properly used, it can help to link family groups together. It also provides interesting studies of ancient family history.

Endnotes

1. Bill Remus is an Emeritus Professor of Information Technology Management at University of Hawaii. His area of expertise is how people make economic predictions. He has written over 100 articles on scientific subjects. For more about him, see <www.cba.hawaii.edu/remus>.
2. Bryan Sykes, *Adam's Curse* (New York: Norton, 2004).
3. *Deutsche Geschlechterbuch*, 1929, vol. 62.
4. Sykes, *op. cit.*
5. Megan Smolenyak-Smolenyak, *Trace Your Roots with DNA* (Emmaus: Rodale Press, 2004)
6. Sykes, *op. cit.*, or visit <www.oxfordancestors.com>.