Introduction to DNA Testing for Genealogy
Wednesday, 12 February 2020
By John Adam Farris
PRIMARY QUESTIONS:

- Why should you spend your hard earned money to buy a DNA test?
- What are the potential benefits?
- What are the potential risks?
- What can you learn?
- Our goal today is to answer these three questions.
GROUNDRULES:

• If you have a question:
• 1. Please ask it while the slide is showing.
• 2. If a later slide will answer your question, The speaker will ask you to wait until that slide is shown to see if you are satisfied.
• 3. Please ask your question in a loud enough voice so that everyone can hear you.

• Thank you for your cooperation: John, Marty & Philip
INTRODUCTION to DNA
MY PRESENTATION OUTLINE

• CAUTION! Don’t test for DNA if your not ready for a surprise!
• What is DNA?
• How is it Inherited?
• Why Buy a DNA Test?
• What Five Companies offer DNA Testing for Family Research?
• Which of the Three Basic DNA Tests Should I Take?
• Can You Trust Your Predicted Ethnicity Results?
WHAT IS DNA?

• Deoxyribonucleic Acid – Molecule, with two long chains with a double-helix structure containing the genetic information of life: Discovered 1953 by Crick & Watson.

• It contains millions of base pairs (A, C, G & T).

• In the nucleus of almost every cell in all living things.

• Not in human RED BLOOD cells, but is in WHITE cells.

• DNA analysis has proved to be very useful in medicine, archaeology, history, forensics & genealogy.
DNA is a New & Exciting & Expanding Science

- 1953 - James Watson & Francis Crick Published their landmark paper on the DNA double helix.
- 1977 - Frederick Sanger develops rapid DNA sequencing.
- 1990 – Human Genome Project begins.
  Alan Savin launched first DNA surname study by a genealogist.
- 1997 – Svante Paabo sequenced the Neanderthal DNA.
- 1999 – First human chromosome decoded.
- 2000 – Bryan Sykes showed he could connect SYKES men using Y-DNA.
  Family Tree DNA (FTDNA) first offered Y-DNA testing services to the public.
- 2003 – Human Genome Project completed - Took 13 Years!
- 2007 – National Genographic & IBM start Genographic Project with FTDNA.
What is DNA & How Do We Use It?

• If you concentrated all of the DNA from a single human body, it would weigh only about 7.5 grams.

• For comparison, the US Quarter coin weighs 5.67 grams.

The DNA from each human cell contains about 3 billion codes. These codes are made up of combinations of only four amino-acid molecules: called A, C, G & T.
What is DNA & How Do We Use It?

All humans are 99.9 % identical in their DNA. However, that 0.1% variation represents about 3,000,000 possible differences between any two people.

For Genealogical purposes, we are concerned with only “Y”, “STR”, “MT”, “X”, “at” & “SNP” type DNA. Currently, “at”, “Y”, “STR”, “MT” & “SNP” are the most commonly used. “MT” stands for “mitochondrial”. These are often referred to as “Y-DNA” & “mtDNA”. “Y” is the male marker and “MT” & “X” are female markers. See later discussion of SNP DNA = snips = autosomal atDNA.
What is DNA & How Do We Use It?

• The DNA markers used for genealogical purposes ("at", "MT", "X", "Y", "STR" & "SNP") are not useful by themselves for legal forensic evidence or paternity evidence because they don’t meet the legal requirements of “Chain of Custody”.

• The DNA testing used in medicine to determine susceptibility to disease isn’t normally useful for genealogy.
What is DNA & How Do We Use It?

DNA is in each cell of every living animal & plant. It is the blueprint that makes each living thing unique. Each species has a distinct pattern of DNA.

The New York Times published a chart in 2007 showing DNA similarities among humans and other mammals. For instance, our human DNA is 98.8% the same as that of a chimpanzee. We both descended from a common ancestor.
What is DNA & How Do We Use It?

When a human egg is fertilized, the atDNA (SNPs) is shared between the egg & the sperm. However, the sharing is never exact. This is why we can see different reflections of the parents/grandparents in each child.

The only exceptions are for the “MT”, “X” & “Y” type DNA. These exceptions remain unique and usually unchanged generation to generation. This is why they are so useful to genealogists. The atDNA is what gets scrambled – more later!
What is DNA & How Do We Use It?

1. Every time a cell divides, the DNA is copied. However, sometimes the copy is not exact. This variation is called a “mutation”. Even though these mutations occur very slowly, they are responsible for the wonderful diversity of all plants & animals, as well as the differences between humans.

2. Without mutations, there would be no evolution.
FAMILY DNA PATTERNS:

• All sons share their father’s “Y” sex chromosome. Thus the “Y-DNA” traces the paternal line which usually (in most cultures) tracks a surname.

• All children share their mother’s “mt-DNA” & her “X” sex chromosome. Thus the “mtDNA” traces the maternal line.

• All of their daughters share their father’s “X” sex chromosome. Thus all daughters always have two “X” sex chromosomes, whereas all sons have an “X” and a “Y”.

• At conception all of the 22 chromosomes (excluding mt, X & Y) are scrambled (recombined) so that each child get a different combination of at-DNA from their parents. This is why children of the same family will show some strong similarities, but each will look different (except for identical twins) & act different. They can even have different atDNA matches & slightly different ethnic % predictions.
Why Should You Buy a DNA Test?
1 of 2

• I have been made curious by all of the repeated TV ads for DNA testing as well as the TV shows showing the power of DNA in solving Family Research problems & in solving crimes. Over 100 cold cases have now been solved with DNA.

• Because I want to know more about my deep roots.

• Because I want to learn more about my family heritage.

• Because every serious genealogist hits one or more brick walls & hope that DNA results/matches will help break down some of them.

• Because I know that both my maternal & paternal ancestors came from the British Isles & I want more specific locations.
Why Should You Buy a DNA Test?
2 of 2

• Because I am *adopted* & I want to find my biological parents, siblings & other close relatives.

• I want to know my *Neanderthal* heritage.

• I want to see my *health results*.

• Other Reasons?

  • =================================================

• **POTENTIAL RISK:**

• You may uncover old family secrets! BE PREPARED!
There Are 3 BASIC DNA tests Available for Family Research

• mtDNA Tests for Both Women & Men - Traces Maternal Line
• Y-DNA Tests for Men ONLY – Traces Paternal Line (There is also now an advanced test available called Big-Y-700, which several of us have taken) – Only offered by Family Tree DNA
• atDNA Tests for Both Women & Men – Will match kin up to five generations – All five testing companies offer this.

• The selection of the test to take depends on YOUR GOALS
There Are 3 DNA tests Available for Family Research (2 of 5) – **MTDNA Tests**

- Mitochondrial DNA (mtDNA) Tests are only offered by FTDNA.com = Family Tree DNA located in Houston, TX.
- These small cells are the energy source for almost every cell in your body.
- The mother passes it on to all of her children.
- Only her daughters can pass it on to their daughters & sons.
- Traces your maternal line.
- It is very stable & mutations don’t occur often.
- This test also provides your female Haplogroup = Deep Roots!
- I recommend that you get the FULL rather than the partial test.
All Mitochondria Are Maternally Inherited

- Males sharing matrilineal line mitochondrial DNA
- Females sharing matrilineal line mitochondrial DNA
- Males NOT sharing matrilineal mitochondrial DNA
- Females NOT sharing matrilineal line mitochondrial DNA
There Are 3 DNA tests Available for Family Research (3 of 5):

**Y-DNA Tests**

- Y-DNA Tests are Only Offered by FTDNA.com = Family Free DNA.
- Very useful for tracing your PATERNAL LINE.
- Only men have the “Y” Sex Chromosome.
- 37 Markers is the minimum test to do. 111 Markers is even better.
- Provides your male Haplogroup = Your Paternal Deep Roots!
- Mutates more often so that extended families can often be identified.
- Can often save $ by ordering through an existing FTDNA surname project or geographical project.
- I have found over 30 genetic Y-DNA cousins & have met five of them.
## Importance of Testing More Y-DNA Markers

<table>
<thead>
<tr>
<th>Number of STR Markers at FTDNA</th>
<th>Number of My Matches</th>
<th>Aug '18</th>
<th>Feb '19</th>
<th>June '19</th>
<th>Sept '19</th>
<th>Feb ‘20</th>
</tr>
</thead>
<tbody>
<tr>
<td>111 Markers</td>
<td></td>
<td>0</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>67 Markers</td>
<td></td>
<td>21</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>37 Markers</td>
<td></td>
<td>26</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>25 Markers</td>
<td></td>
<td>3,519</td>
<td>3,978</td>
<td>4,076</td>
<td>4,134</td>
<td>4,250</td>
</tr>
<tr>
<td>12 Markers</td>
<td></td>
<td>?</td>
<td>15,179</td>
<td>15,497</td>
<td>15,755</td>
<td>18,767</td>
</tr>
</tbody>
</table>

15 February 2019
Y-DNA Inheritance Patterns, © 2013, Debbie Parker Wayne
There Are 3 DNA tests Available for Family Research (4 of 5): \textit{atDNA Tests}

- Five companies offer this autosomal DNA (atDNA) test – SEE NEXT SLIDE & HANDOUT
- Measures the 600,000 to 700,000 SNPs on all 22 of your non-sex chromosomes.
- Both females & males can take this test.
- \textbf{Good for finding matches with close kin:} children, grandchildren, parents, grandparents, aunts, uncles, nieces, nephews, first cousins, second cousins, & third cousins (90%), etc.
There are Now **FIVE Companies** Offering DNA Testing for Family Research

• I am often asked: Which Company is Best?
• They All Deliver Exactly What They Promise.
• They are all good companies.
• The Selection of the Testing Company Should be Based on **YOUR GOALS**.
Autosomal DNA (atDNA) Testing Companies

<table>
<thead>
<tr>
<th>COMPANY</th>
<th>23andMe</th>
<th>Ancestry</th>
<th>Family Tree DNA</th>
<th>My Heritage</th>
<th>Living DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phone</td>
<td>800-239-5230</td>
<td>800-615-6560</td>
<td>713-868-1438</td>
<td>800-987-9000</td>
<td>????</td>
</tr>
<tr>
<td>WEB Site</td>
<td>23andme.com</td>
<td>ancestry.com</td>
<td>ftdna.com</td>
<td>myheritage.com</td>
<td>livingdna.com</td>
</tr>
<tr>
<td>Date Started</td>
<td>2006</td>
<td>2012</td>
<td>2010</td>
<td>2016</td>
<td>2016</td>
</tr>
<tr>
<td>CURRENT SALES</td>
<td>$79 &amp; $129</td>
<td>$59</td>
<td>$49</td>
<td>$99 W/Health</td>
<td>$79 &amp; $149</td>
</tr>
<tr>
<td>Location</td>
<td>Los Angeles</td>
<td>Provo, UT</td>
<td>Houston, TX</td>
<td>Israel</td>
<td>England</td>
</tr>
<tr>
<td>Matches Provided</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Finally Now - a Few</td>
</tr>
<tr>
<td>Matches connect trees/circles</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>Matches presented as cM's</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Ethnicity % Info</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ancestry General Location</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ancestry Detailed Location</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>M&amp;F Haplogroups Given</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Analysis Tools Available</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Not Yet</td>
</tr>
<tr>
<td>Medical Information</td>
<td>Yes +$</td>
<td>Yes +$ New</td>
<td>Yes +$ New</td>
<td>Yes +$</td>
<td>No *</td>
</tr>
<tr>
<td>Neanderthol Info</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Direct contact with matches</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Est. Database Size-30 Dec 2019*</td>
<td>~10,000,000</td>
<td>~16,000,000</td>
<td>~1,150,000</td>
<td>3,770,000</td>
<td>No Data</td>
</tr>
</tbody>
</table>

NOTES:

*Total Tested = >30,920,000

1. The atDNA results from each of these companies can be transferred for FREE to [GEDmatch.com](https://gedmatch.com), which has many analysis tools.
2. For more details go to [www.isogg.org](https://www.isogg.org) & select WIKI & select: Autosomal DNA testing comparison tool.
3. For matches in cMs, enter for FREE the value into [dnapainter.com](https://dnapainter.com) by J. Seabright & all the possible relationships appear.

Prices are per their WEB Sites, but check for sales around holidays

All of these atDNA tests can be taken by both Women & Men.

John AFarris@comcast.net

By: John Adam FARRIS
<table>
<thead>
<tr>
<th>Country</th>
<th>Testing Company</th>
<th>23andMe</th>
<th>Ancestry DNA</th>
<th>FTDNA</th>
<th>Living DNA</th>
<th>MyHeritage</th>
<th>Nat’l Geographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>European</td>
<td></td>
<td>99.8-99.9-99.8</td>
<td>99-100-99</td>
<td>99</td>
<td>100-100</td>
<td>100-100</td>
<td>98</td>
</tr>
<tr>
<td>British Isles</td>
<td></td>
<td>69.7-67.9-61.6</td>
<td>59-89-95</td>
<td>27</td>
<td>99.3-95.9</td>
<td>73-73</td>
<td>8</td>
</tr>
<tr>
<td>England</td>
<td></td>
<td>?-?-64</td>
<td>57.5-52.0</td>
<td>?-38.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td></td>
<td>?-?-31</td>
<td>13.5-16.3</td>
<td>?-34.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scotland</td>
<td></td>
<td>Ire &amp; Scot</td>
<td>11.1-9.9</td>
<td>Ire &amp; Scot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulster</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West, Central &amp; North</td>
<td></td>
<td>18.5-28.4-20.6</td>
<td>0-0-2</td>
<td>61</td>
<td>4-4</td>
<td>27-27</td>
<td>57</td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.4-0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scandinavia</td>
<td></td>
<td>6.3-1.7-1.4</td>
<td>16-0-2</td>
<td>11</td>
<td>0-0</td>
<td>0-0</td>
<td>33</td>
</tr>
<tr>
<td>Italy &amp; Greece</td>
<td></td>
<td>1.4-1-0</td>
<td>8-0-0</td>
<td>0</td>
<td>0-0</td>
<td>0-0</td>
<td>0</td>
</tr>
<tr>
<td>Iberian Peninsula</td>
<td></td>
<td>0.9-0.4-3</td>
<td>1-0-0</td>
<td>0</td>
<td>0-0</td>
<td>0-0</td>
<td>0</td>
</tr>
<tr>
<td>Misc.</td>
<td></td>
<td>3-0.5-0</td>
<td>3-0-0</td>
<td>&lt;1.0</td>
<td>0-0</td>
<td>0-0</td>
<td>2</td>
</tr>
</tbody>
</table>
NEXT SPEAKER: Marty Brady, Our AGS President
SNPs, Chips and NGS Clips

DNA Analysis for Genealogical Methods
Introduction

• We will talk about:
  • SNPs: what they are, how they are made, how they are analyzed, and how we use the information.
  • Chips: what are DNA microarrays.
  • NGS: what is Next Generation Sequencing, how is it used in genealogy (the Big Y-700 test, possibly artifact testing).
Single Nucleotide Polymorphism (SNP)

- A **SNP** is a substitution of a single nucleotide that occurs at a specific position in the genome. Where each variation is present at a level of 1% in the population.

- So, technically if a variation occurs at a level less than 1%, it is only a variation and not a SNP. But many people use the terms interchangeably.
How do we get SNPs

• 222 billion to 242 billion cells produced daily (so there are 222,000,000,000 X 3,200,000,000 = 710,400,000,000,000,000,000 opportunities for mistakes). (Million, billion, trillion, quadrillion, quintrillion, sextillion.)

• Many mistakes are made each day, but most don’t get inherited because they don’t make it into the germ cell line (eggs and sperm).

• Average male will produce roughly 525 billion sperm cells in a lifetime, (so this doubles the opportunities for mistakes above, but over a lifetime).

• There are 4 to 5 million SNPs in the human genome.

• So the question isn’t why are there so many SNPs, the question is why are there so few.
DNA Replication

- DNA Polymerase is the **enzyme** responsible for synthesizing new strands of DNA (i.e., making copies, replication).

- **Enzymes** are protein molecules in cells which work as biological catalysts. Enzymes speed up chemical reactions in the body, but do not get used up in the process, therefore can be used over and over again. Almost all biochemical reactions in living things need enzymes.

- DNA Polymerase works in one direction adding new nucleotides (dNTP) to the 3 position of deoxyribose backbone of the DNA strand.

- A **nucleotide** is the building block unit of nucleic acids such as DNA. When we talk about sequencing, we are talking about the sequential order of the different nucleotides as we progress down the DNA strand.

- **Complementarity** is maintained. Complementarity refers to the fact that each nucleotide has a base (A, C, T or G), and each base on a DNA strand can pair up with the appropriate (complementary) base (A with T and G with C) from the opposing nucleotide on the second DNA strand.
DNA Polymerase Activity and Mistake Correction

1. DNA template
   - 5' G C A T C G A 3'
   - 3' C G A T C G 5'
   - DNA polymerase
   - Nucleotide triphosphate

2. DNA extension
   - 5' G C A T C G A 3'
   - 3' C G A T C G 5'
   - DNA polymerase

3. Misincorporation
   - 5' G C A T C G A 3'
   - 3' C G A T C G 5'
   - DNA polymerase

4. Proofreading via 3' to 5' exonuclease activity
   - 5' G C A T C G A 3'
   - 3' C G A T C G 5'
   - DNA polymerase

5. DNA extension with correct base pair
   - 5' G C A T C G A 3'
   - 3' C G A T C G 5'
   - DNA polymerase
DNA Microarrays

• So, DNA Polymerase has proofreading capability which helps reduce the number of mutations or misincorporations (i.e., SNPs) that occur during DNA replication.

• However, some mutations make it through to the final product (newly copied strand of DNA) and so we have SNPs or variants and we test for those SNPs using DNA microarrays.

• An array is an ordered series or arrangement. An excel spreadsheet is an array that orders data into columns and rows so that each cell has a specific address.

• DNA microarrays are microscope slides that are printed with thousands of tiny spots in defined positions, with each spot containing a known, small nucleotide sequence (i.e., an oligonucleotide).

• An oligonucleotide is several nucleotides connected together to form a small segment of DNA.

• Annealing or hybridization is when 2 complementary strands of DNA come together and pair up into a double strand of DNA.
DNA Microarray (chip) Visual (with oligonucleotide)
Example of Completed DNA Microarray

Each dot on the plate represents one SNP.
Review of the DNA Microarray (chip) Process

• Determine where a target SNP exists (the SNP we are trying to detect).
• Determine the nucleotide sequence preceding the SNP (~10 or 20 dNTP).
• Attach complementary fragments of DNA (oligonucleotides) to a plate bead.
• Drop the donor sample onto the plate and allow fragments to anneal or hybridize to the attached oligonucleotide.
• DNA Polymerase adds one nucleotide (ddNTP) to the attached fragment. Each ddNTP has a fluorescent dye attached to it (i.e., in my drawing, T has red dye attached and we might use the following scheme for the other ddNTPs: A - blue dye, C - green dye and G - yellow dye).
• Record which ddNTP was incorporated based on what dye color is detected.
23andMe “Chip” versions

- **Chip versions**
  - v1: November 2007.
  - v2: September 2008, ~555K SNPs.
  - v3: November 2010, >900K SNPs.
  - v4: November 2013, ~570K SNPs.
  - v5 August 2017, ~640K SNPs (change made to Illumina Global Screening Array BeadChip)
Ancestry “chip” versions

- AncestryDNA v1 chip had about 701,400 SNPs.
- AncestryDNA v2 chip (2016) has 669,000 SNPs.
- About 300,000 “low performing” SNPs were changed.
  - Low performing means SNPs that were not as good at predicting ethnicity or medical conditions.
  - v2 Optimized for medical and ethnicity.
Significance of SNP Selection

• Slight variation in ethnicity projections (database variance also affects ethnicity estimates).
• Slight variance in shared Centimorgan (cM) determinations.
• A cM is a measure of distance between two points on a chromosome. One cM is the distance in which there is a 1% chance that a recombination event will occur in a single generation. In humans, 1 cM is equivalent, on average, to 1 million base pairs.
• Recombination is a DNA exchange event usually between two copies of the same chromosome (chr) at similar positions on the chromosome (i.e., the end of chr2 inherited from one parent might swap with the end of chr2 inherited from the other parent).
• Possible problems with transferring DNA raw data to another company through the use of imputation. Imputation is the assignment of a value to something by inference from associated data.
• Possible overlap problems for GEDmatch site.
Imputation example with words

• Your brain can perform amazing statistical calculations, and fill in the blanks!

• Your brain can perform amazing statistical calculations, and fill in the blanks!
**Imputation example with DNA Loci**

<table>
<thead>
<tr>
<th>Chr</th>
<th>Locus</th>
<th>Identifier</th>
<th>Mom allele</th>
<th>Dad allele</th>
<th>Locus</th>
<th>Identifier</th>
<th>Mom allele</th>
<th>Dad allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>731546</td>
<td>rs5071251</td>
<td>C</td>
<td>C</td>
<td>731546</td>
<td>rs5071251</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>7889960</td>
<td>rs15251</td>
<td>T</td>
<td>C</td>
<td>7889960</td>
<td>rs15251</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>7</td>
<td>918110955</td>
<td>rs651315251</td>
<td>A</td>
<td>T</td>
<td>918110955</td>
<td>rs651315251</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>7</td>
<td>114554361</td>
<td>rs1049036655</td>
<td>A</td>
<td>A</td>
<td>114554361</td>
<td>rs1049036655</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>7</td>
<td>114554370</td>
<td>rs929989090</td>
<td>G</td>
<td>C</td>
<td>114554370</td>
<td>rs929989090</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>7</td>
<td>114554389</td>
<td>rs147984339</td>
<td>T</td>
<td>C</td>
<td>114554389</td>
<td>rs147984339</td>
<td>C</td>
<td>T</td>
</tr>
</tbody>
</table>

Company A chip

Company B chip

SNPs in red were not tested by the Company listed below them, but were inferred using the tested SNPs shown in black.

Locus is a location on the chromosome, as in nucleotide # 731546 on chromosome 7. (Loci is plural for locus).

Alleles are one of two or more alternative forms of a gene that arise by mutation and are found at the same place on a chromosome. We inherit one from mom and one from dad.
### Possible Centimorgan (cM) Differences

<table>
<thead>
<tr>
<th>Chr</th>
<th>Locus</th>
<th>Identifier</th>
<th>Mom allele</th>
<th>Dad allele</th>
<th>Locus</th>
<th>Identifier</th>
<th>Mom allele</th>
<th>Dad allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>114554361</td>
<td>rs1049036655</td>
<td>A</td>
<td>A</td>
<td>114554361</td>
<td>rs1049036655</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>114554370</td>
<td>rs929989090</td>
<td>G</td>
<td>C</td>
<td>114554370</td>
<td>rs929989090</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>7</td>
<td>114554389</td>
<td>rs147984339</td>
<td>T</td>
<td>C</td>
<td>114554389</td>
<td>rs147984339</td>
<td>T</td>
<td>C</td>
</tr>
</tbody>
</table>

DNA matches with Company A  DNA matches with Company B

The locus in red above was not tested by Company A, but was tested by Company B. Since the tested individuals differ at this locus the DNA matches may appear to share larger contiguous segments (more cM) of DNA with Company A than with Company B.
Next Generation Sequencing (NGS)

• AKA Massively Parallel Sequencing or Second Generation Sequencing
NGS Process (watch the youtube video listed in the next slide)

• The first step in NGS is fragmenting the DNA into small fragments.
• Fragments of the appropriate length are isolated. (My Big Y-700 fragments were about 150 base pairs (bp) in length.)
• Identifier Oligonucleotides are attached to each end of the fragment.
• The fragments are applied to a flow cell that has complementary oligonucleotides attached to its surface.
• DNA Polymerase is added (along with the requisite dNTPs).
• The fragments are amplified in clusters by bridge amplification.
• Then the process of Sequencing by Synthesis is begun. Each dNTP added is recorded and tracked by the instrument.
• Finally, the forward and reverse reads are assembled into a sequence (the segment can be sequenced in either direction – forward or reverse).
Next Generation Sequencing (NGS)

Illumina next generation sequencing method.

m.youtube.com/watch?v=fCd6B5HRaZ8
NGS

- Typically 200 to 500 million segments are sequenced per sample at one time. (Can be up to one billion segment reads.)
- \((200,000,000 \times 150\text{bp} = 30,000,000,000\text{ bits of information per run})\)
- So, you need large computer systems to crunch all the data.
- DNA sequencing data production speed doubles about every 6 months.
- Computer processing speed doubles about every 2 years.
Where is NGS used

• NGS is ideal for artifacts and samples with degraded DNA because the first step in NGS is fragmenting the DNA and degraded DNA is already fragmented.

• FTDNA may be the only company using NGS for genealogical purposes (Big Y-700).

• NGS detects new/unknown SNPs (as well as known SNPs) and is not restricted like microarrays.

• I had my NGS Whole Genome Sequencing (WGS) done by Dante labs recently, but it is mainly for health purposes. We need WGS databases before it is of use to genealogists (and the price needs to come down a bit, but it is on the horizon).
One Private Variant in My Big Y-700 Results (position 10981829)

Sequenced segments are aligned with the reference genome at top. Forward reads are in blue. Reverse reads are in green. The long line of pink in the center of the screen is the private variant (T instead of a C as in the reference genome). It has been confirmed in well over 40 reads/segments (depth of coverage over 40X). There are some isolated pink variants that only appear in one read/segment and so those may be attributed to analytical error. This is why depth of coverage is important in reducing misreads.
Partial Coverage of Y Chromosome with Big Y-700

The Big Y-700 test only covers about 40% of the Y chromosome (40% breadth of coverage). The inaccessible region (black) contains a lot of repetitive sequences. The gray regions (PAR1 and PAR2) recombine with the X chromosome and therefore do not make these regions stable enough for paternal heritage information.
Summary

• We talked about what a SNP is.
• We talked about how SNPs are created (mutations/misincorporations).
• We talked about how we test for SNPs (using DNA microarrays, aka Chips).
• We looked at SNP profiles from different companies.
• We talked about Next Generation Sequencing (NGS).
• How NGS is used to detect new SNPs (Big Y-700).
• How NGS might be ideal for artifacts.
Adenosine

3 Parts of a Nucleotide

Phosphate group

Nitrogenous base (adenine)

Pentose Sugar
DNA Developments
Philip Spivey
NEW SEASON

Long Lost Family

WATCH NOW
Joseph James DeAngelo
Accused Golden State Killer
Sacramento County Sheriff’s Office mugshot
Phasing & Triangulation
Phasing

Separating Paternal and Maternal Matches
Hello, Philip

This test is shown to match as ppspivey74, linked to Philip Charles Spivey.

A healthy addition to the Ancestry family.
Get health insights you can take action on.

DNA Story

Ethnicity Estimate

- 55% England, Wales & Northwestern Europe
- 25% Ireland & Scotland
- 2 Other regions

Discover the places, history, and cultures that shaped who you are today—using just your DNA.

DNA Matches

- 0 Starred matches
- 1000+ 4th cousins or closer

View All DNA Matches

ThruLines™

ThruLines uses Ancestry trees to suggest how you may be related to your DNA matches through common ancestors.

Explore ThruLines
<table>
<thead>
<tr>
<th>Name</th>
<th>Match Date</th>
<th>Relationship Range</th>
<th>Shared cM</th>
<th>Longest Block</th>
<th>Ancestral Surnames</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>07/02/2015</td>
<td>Half Sister, Grandmother/Granddaughter, Aunt, Niece</td>
<td>1,878</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>02/11/2017</td>
<td>2nd Cousin - 3rd Cousin</td>
<td>181</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>01/04/2020</td>
<td>2nd Cousin - 4th Cousin</td>
<td>108</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>08/28/2019</td>
<td>2nd Cousin - 4th Cousin</td>
<td>93</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>08/26/2014</td>
<td>2nd Cousin - 4th Cousin</td>
<td>90</td>
<td>29</td>
<td>Crowell / King / Mashburn / Smith / King / Sr / Mashburn / I / Mashburn / II /</td>
</tr>
<tr>
<td></td>
<td>09/21/2019</td>
<td>2nd Cousin - 4th Cousin</td>
<td>89</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>
## Ancestry-Matches

### Philip Spivey's DNA Matches

#### Close Family

<table>
<thead>
<tr>
<th>Name</th>
<th>Relation</th>
<th>Shared DNA</th>
<th>Unlinked Tree</th>
<th>Add/edit groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nona Bertschy</td>
<td>Close Family–1st Cousin</td>
<td>1,870 cM across 55 segments</td>
<td>No Trees</td>
<td>Add/edit groups</td>
</tr>
<tr>
<td>M.N.</td>
<td>Close Family–1st Cousin</td>
<td>1,553 cM across 56 segments</td>
<td></td>
<td>Add/edit groups</td>
</tr>
</tbody>
</table>

#### 2nd Cousin

<table>
<thead>
<tr>
<th>Name</th>
<th>Relation</th>
<th>Shared DNA</th>
<th>2,348 People</th>
<th>Add/edit groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>melissa hughes</td>
<td>1st–2nd Cousin</td>
<td>541 cM across 23 segments</td>
<td>Common ancestor</td>
<td>Add/edit groups</td>
</tr>
</tbody>
</table>
# Ancestry-Matches

## Philip Spivey's DNA Matches

### Close Family

<table>
<thead>
<tr>
<th>User</th>
<th>Relationship</th>
<th>Shared DNA:</th>
<th>Unlinked Tree</th>
<th>Add/edit groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nona Bertschy</td>
<td>Close Family–1st Cousin</td>
<td>1,870 cM across 55 segments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M.N.</td>
<td>Close Family–1st Cousin</td>
<td>1,553 cM across 56 segments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2nd Cousin

<table>
<thead>
<tr>
<th>User</th>
<th>Relationship</th>
<th>Shared DNA:</th>
<th>2,348 People</th>
<th>Add/edit groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>melissa hughes</td>
<td>1st–2nd Cousin</td>
<td>541 cM across 23 segments</td>
<td>Common ancestor</td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Match Date</td>
<td>Relationship Range</td>
<td>Shared cM</td>
<td>Longest Block</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>-------------------------------------</td>
<td>-----------</td>
<td>---------------</td>
</tr>
<tr>
<td>Nona Bertschy</td>
<td>07/02/2015</td>
<td>Half Sister, Grandmother/Granddaughter, Aunt, Niece</td>
<td>1,878</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>02/11/2017</td>
<td>2nd Cousin - 3rd Cousin</td>
<td>181</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>01/04/2020</td>
<td>2nd Cousin - 4th Cousin</td>
<td>108</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>08/28/2019</td>
<td>2nd Cousin - 4th Cousin</td>
<td>93</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>08/26/2014</td>
<td>2nd Cousin - 4th Cousin</td>
<td>90</td>
<td>29</td>
</tr>
</tbody>
</table>
DNA Relatives

Get started with your predicted relationships, then connect and message to learn more.

Showing 1312 of 1312 Relatives

<table>
<thead>
<tr>
<th>Filters</th>
<th>Sort by</th>
<th>First Cousin</th>
<th>First Cousin</th>
<th>First to Second Cousin</th>
<th>Second Cousin</th>
<th>Second Cousin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>11.2% DNA shared, 31 segments</td>
<td>7.52% DNA shared, 25 segments</td>
<td>7.00% DNA shared, 17 segments</td>
<td>4.22% DNA shared, 11 segments</td>
<td>3.86% DNA shared, 14 segments</td>
</tr>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notifications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profile features and activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ancestor birthplaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mom's side / Dad's side</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Your genetic relationship

Niece

You and Christine share DNA that was passed down from your parents (her grandparents).

View DNA details ➔
You:

**H4a1**
See full report

**H4a**
See full report

You:

**R-L51**
See full report

Although women inherit roughly 50% of their DNA from their fathers, they do not inherit Y chromosomes and, as a result, do not have paternal haplogroups.

---

**Neanderthal Ancestry**

You:

**271**
Variants

You have more Neanderthal variants than 44% of 23andMe customers.

See full report

Christine has more Neanderthal variants than 40% of 23andMe customers.

**268**
Variants

---

You and Christine have Relatives in Common

Finding common relatives can help you piece together your family story.

[View Relatives in Common]

---

Help us improve this experience! Answer a few quick questions

---

You aren't sharing Health reports.

Invite Christine to also view one another's detailed Health Predisposition, Carrier Status, Traits, and Wellness reports. Once your request is accepted, you will be able to compare Health reports. If Christine upgrades to Health, your reports will be shared with Christine.

Late-Onset Alzheimer's Disease, Parkinson's Disease, BRCA1/BRCA2, and MUTYH-Associated Polyposis are never included in sharing. Learn how sharing works.

[Share Health reports]
Triangulation
Triangulation

• Segment Triangulation:
  3 people share the same DNA segment

• Tree Triangulation:
  3 people share the same CA
# Philip Spivey's DNA Matches

## Close Family

<table>
<thead>
<tr>
<th>Name</th>
<th>Relationship</th>
<th>DNA Information</th>
<th>Tree Information</th>
<th>Add/-edit Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nona Bertschy</td>
<td>Close Family–1st Cousin</td>
<td>Shared DNA: 1,870 cM across 55 segments</td>
<td>Unlinked Tree</td>
<td></td>
</tr>
<tr>
<td>M.N.</td>
<td>Close Family–1st Cousin</td>
<td>Shared DNA: 1,553 cM across 56 segments</td>
<td>No Trees</td>
<td></td>
</tr>
</tbody>
</table>

## 2nd Cousin

<table>
<thead>
<tr>
<th>Name</th>
<th>Relationship</th>
<th>DNA Information</th>
<th>Add/-edit Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>melissa hughes</td>
<td>1st–2nd Cousin</td>
<td>Shared DNA: 541 cM across 23 segments</td>
<td></td>
</tr>
</tbody>
</table>
# The Shared cM Project – Version 3.0

**August 2017**

Blaine T. Bettinger  
[www.TheGeneticGenealogist.com](http://www.TheGeneticGenealogist.com)  
CC 4.0 Attribution License

## How to read this chart:

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Average</th>
<th>Range (low-high)</th>
<th>(99% Percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aunt/Uncle</td>
<td>1750</td>
<td>1349 - 2175</td>
<td></td>
</tr>
<tr>
<td>Great-Great-Grandparent</td>
<td>881</td>
<td>464 - 1486</td>
<td></td>
</tr>
<tr>
<td>Great-Great-Grandparent</td>
<td>427</td>
<td>191 - 885</td>
<td></td>
</tr>
<tr>
<td>GGG-Aunt/Uncle</td>
<td>914</td>
<td>251 - 2108</td>
<td></td>
</tr>
<tr>
<td>Great-Aunt/Uncle</td>
<td>187</td>
<td>12 - 383</td>
<td></td>
</tr>
</tbody>
</table>

## Other Relationships

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6C</td>
<td>21</td>
<td>0 - 86</td>
</tr>
<tr>
<td>6C1R</td>
<td>16</td>
<td>0 - 72</td>
</tr>
<tr>
<td>6C2R</td>
<td>17</td>
<td>0 - 75</td>
</tr>
<tr>
<td>7C</td>
<td>13</td>
<td>0 - 57</td>
</tr>
<tr>
<td>7C1R</td>
<td>13</td>
<td>0 - 53</td>
</tr>
<tr>
<td>8C</td>
<td>12</td>
<td>0 - 50</td>
</tr>
</tbody>
</table>

Minimum was automatically set to 0 cM for relationships more distant than Half 2C, and averages were determined only for submissions in which DNA was shared.
### Detailed segment data

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Chrom.</th>
<th>Start Position</th>
<th>End Position</th>
<th>Genetic Distance (cM)</th>
<th>Number of SNPs</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Philip Spivey / melissa Hughes</td>
<td>1</td>
<td>119452758</td>
<td>193603469</td>
<td>57.23</td>
<td>9047</td>
<td>Half</td>
</tr>
<tr>
<td>Philip Spivey / melissa Hughes</td>
<td>1</td>
<td>215925167</td>
<td>223603809</td>
<td>8.68</td>
<td>1483</td>
<td>Half</td>
</tr>
<tr>
<td>Philip Spivey / melissa Hughes</td>
<td>2</td>
<td>199134456</td>
<td>237592752</td>
<td>50.53</td>
<td>8009</td>
<td>Half</td>
</tr>
<tr>
<td>Philip Spivey / melissa Hughes</td>
<td>3</td>
<td>129944129</td>
<td>171041055</td>
<td>36.34</td>
<td>7448</td>
<td>Half</td>
</tr>
<tr>
<td>Philip Spivey / melissa Hughes</td>
<td>5</td>
<td>80361698</td>
<td>132488440</td>
<td>42.18</td>
<td>9370</td>
<td>Half</td>
</tr>
</tbody>
</table>
### GEDmatch Triangulation of Selected Kits -- V0.3

Triangulation with Kit XL194899C1 – Philip Spivey.

All kits shown in columns Kit1 and Kit2 are taken from the selected kit matches to XL194899C1.

3-Way (Triangulated) segment matches shown in green. This is an indication of common ancestry.

Segments shown are larger than 7.0 cM and between 200 and 400 SNPs.

**Triangulated results sorted by Chromosome, Start Position**

<table>
<thead>
<tr>
<th>Chr</th>
<th>Kit1</th>
<th>Kit2</th>
<th>B37 Start</th>
<th>B37 End</th>
<th>cM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>M726777</td>
<td>T575566</td>
<td>205,057,607</td>
<td>234,204,113</td>
<td>41.4</td>
</tr>
<tr>
<td>6</td>
<td>M726777</td>
<td>T575566</td>
<td>61,963,172</td>
<td>91,844,387</td>
<td>18.7</td>
</tr>
<tr>
<td>10</td>
<td>M726777</td>
<td>T575566</td>
<td>127,791,184</td>
<td>135,434,303</td>
<td>19.6</td>
</tr>
<tr>
<td>20</td>
<td>M726777</td>
<td>T575566</td>
<td>63,799</td>
<td>5,223,014</td>
<td>15.3</td>
</tr>
</tbody>
</table>

**Triangulated results sorted by Kit Number, Chromosome, Start Position:**

Select 2 or more kits, and click this button for additional display and processing options.

<table>
<thead>
<tr>
<th>Chr</th>
<th>Kit1</th>
<th>Kit2</th>
<th>B37 Start</th>
<th>B37 End</th>
<th>cM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>M726777</td>
<td>T575566</td>
<td>205,057,607</td>
<td>234,204,113</td>
<td>41.4</td>
</tr>
<tr>
<td>6</td>
<td>M726777</td>
<td>T575566</td>
<td>61,963,172</td>
<td>91,844,387</td>
<td>18.7</td>
</tr>
<tr>
<td>10</td>
<td>M726777</td>
<td>T575566</td>
<td>127,791,184</td>
<td>135,434,303</td>
<td>19.6</td>
</tr>
<tr>
<td>20</td>
<td>M726777</td>
<td>T575566</td>
<td>63,799</td>
<td>5,223,014</td>
<td>15.3</td>
</tr>
</tbody>
</table>
Triangulation
DNApainter
So what does this triangulation tell us?
Triangulation

- William and Hulda Byrd are our common ancestors
Triangulation

- William and Hulda Byrd are our common ancestors
- The DNA I share with Melissa and Matt all came down to us from William and Hulda.
Triangulation

- William and Hulda Byrd are our common ancestors
- The DNA I share with Melissa and Matt all came down to us from William and Hulda.
- Bonnie Cora Byrd is likely my grandmother because she is related to all three of us.
Summary

• Phasing is used to separate paternal matches from maternal matches

• Triangulation is used to confirm relationships