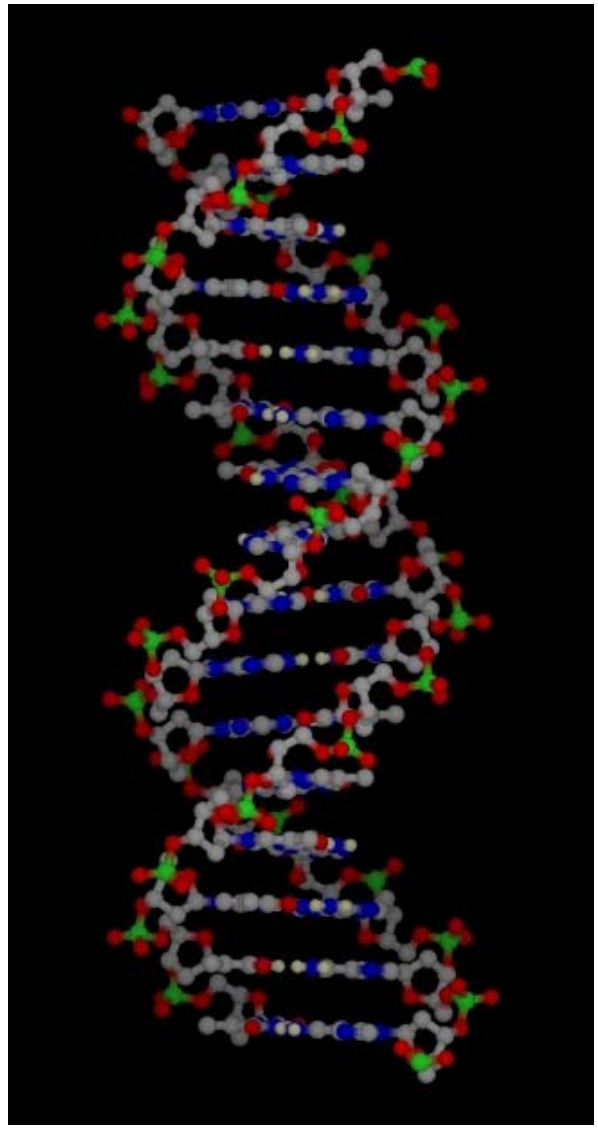
What's in Your DNA?

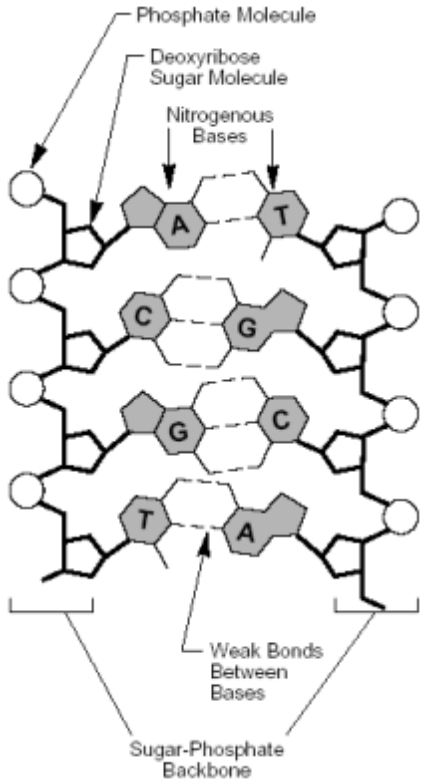
Sergio Sanchez

2009 Charlottesville Family History Conference



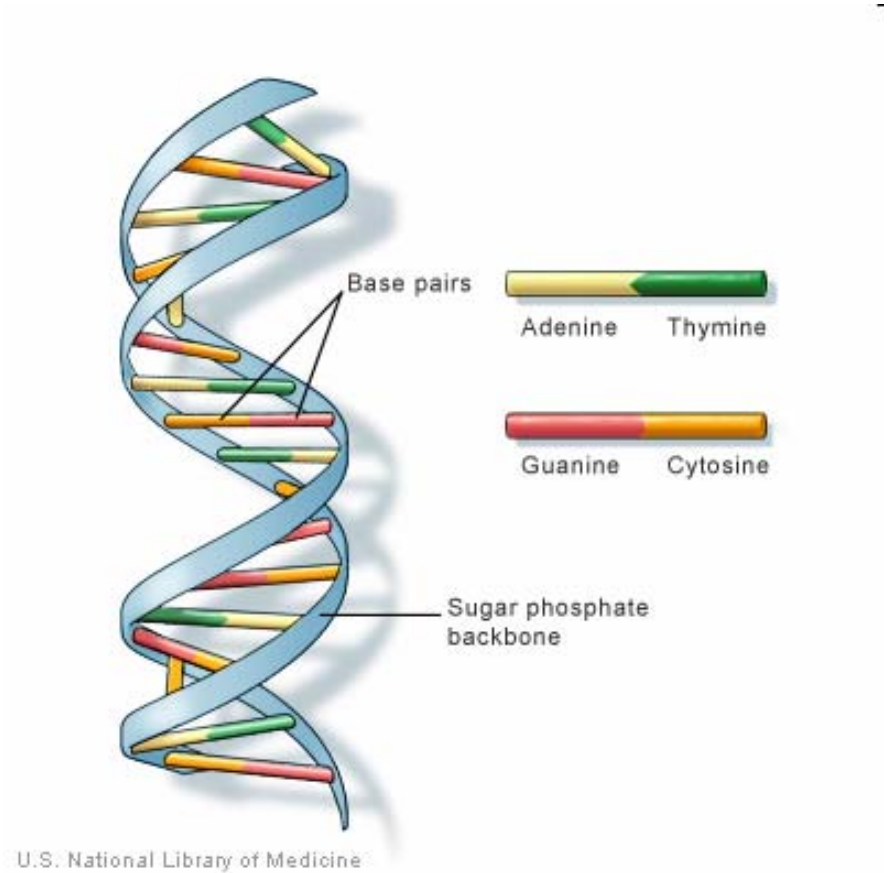
What is DNA?

- DNA → **D**eoxyribo**N**ucleic **A**cid
- DNA contains all of the genetic instructions for the development and functioning of living organisms.
- DNA is the long-term storage of information in a cell and is often described as the set of blueprints for each cell.
- DNA is found in all “living organisms.”



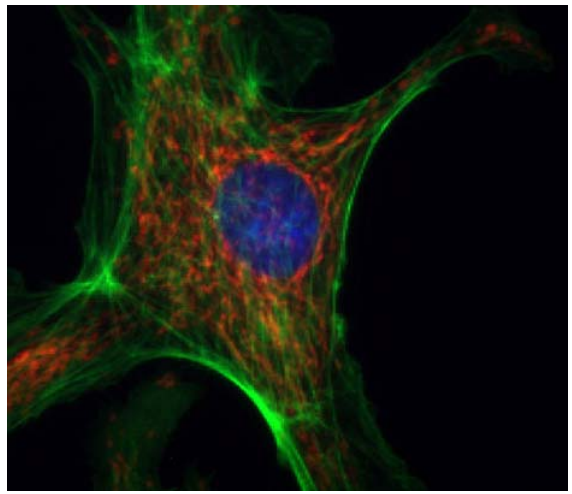
What is DNA?

- DNA is made up of 4 nucleotide bases.
  - Adenine (A)
  - Cytosine (C)
  - Guanine (G)
  - Thymine (T)
- Nucleotides are paired together.
  - A – T
  - C – G



What is DNA?

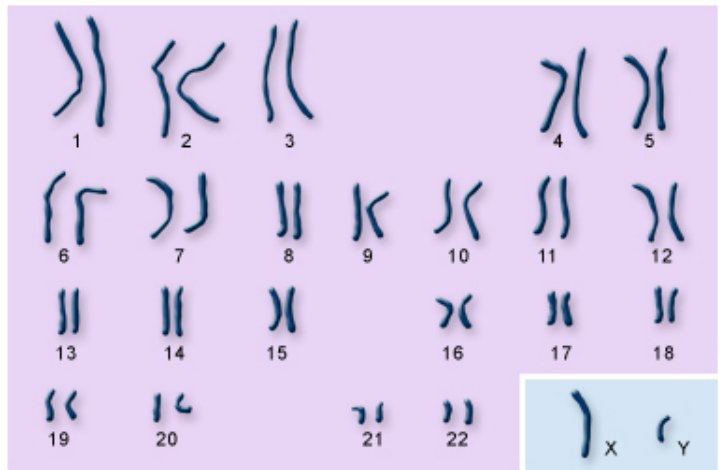
- DNA is found in the nucleus of most human cells (6ft long!)
- Each human cell contains a full set of 46 chromosomes.
  - 44 Autosomal Chromosomes
  - 2 Sex Chromosomes
- 23 chromosomes are inherited from each parent.
  - 22 Autosomal
  - 1 Sex chromosome



Blue: DNA in nucleus

Red: Mitochondria

Green: Cellular skeleton

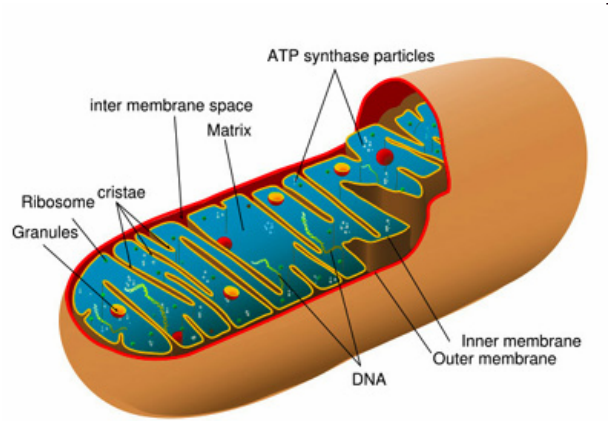


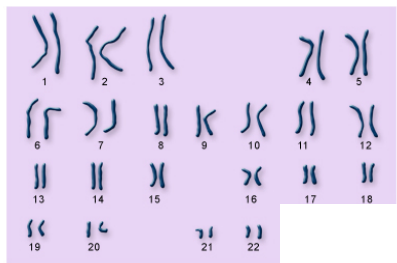
U.S. National Library of Medicine

DNA and Genealogy

- For genealogic analysis, DNA is classified into three subsets
  - Y-Chromosome DNA (Y-DNA)
  - Mitochondrial DNA (mtDNA)
  - Autosomal DNA

sex chromosomes

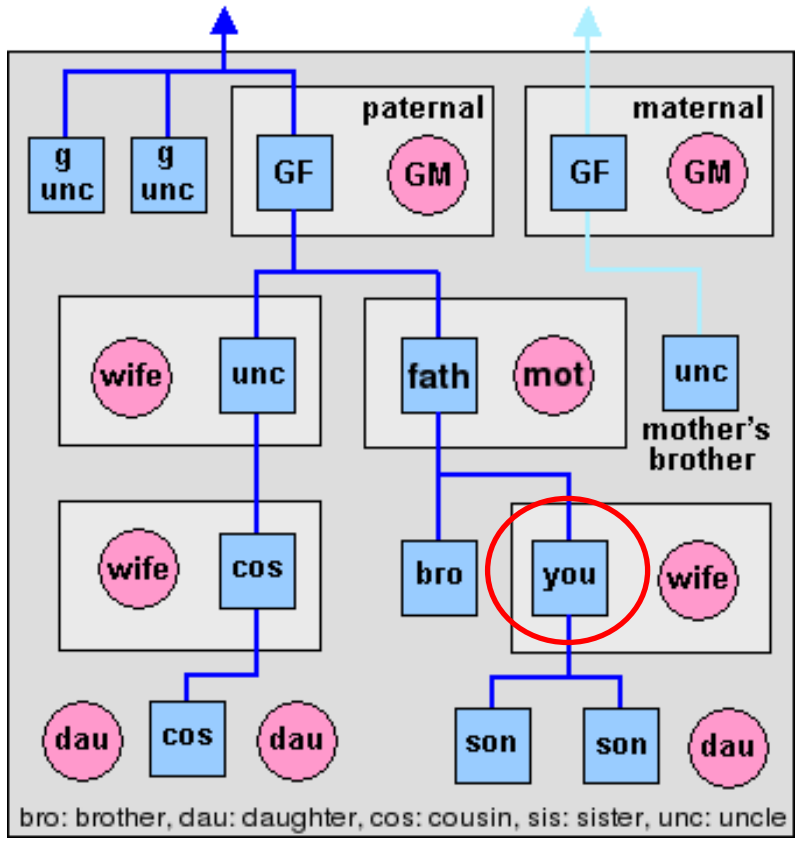




autosomes
U.S. National Library of Medicine

Y- Chromosome DNA (Y-DNA)

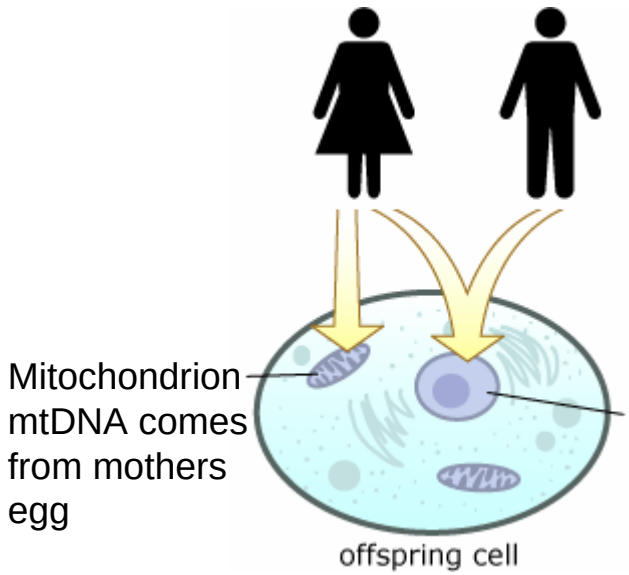
- **Y-DNA** is a type of DNA that is only carried by men, who **inherit it from their fathers**
- This means that males with a common paternal ancestor have similar Y-DNA
- Particularly useful for tracing one's direct paternal line because it changes slowly from generation to generation
- Surname of the father is also inherited by his sons



All males connected by blue line have common Y-DNA

Mitochondrial DNA (mtDNA)

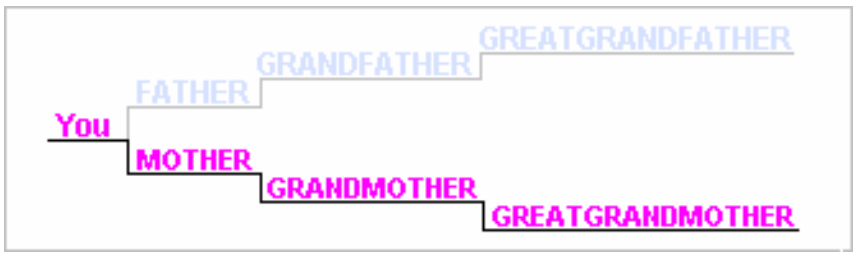
- Each cell has about 1700 mitochondria with many copies of the same loop of DNA
- It carries 37 genes in approximately 16,000 base pairs
- **MtDNA** is carried by both males and females, but is **only inherited from mother**

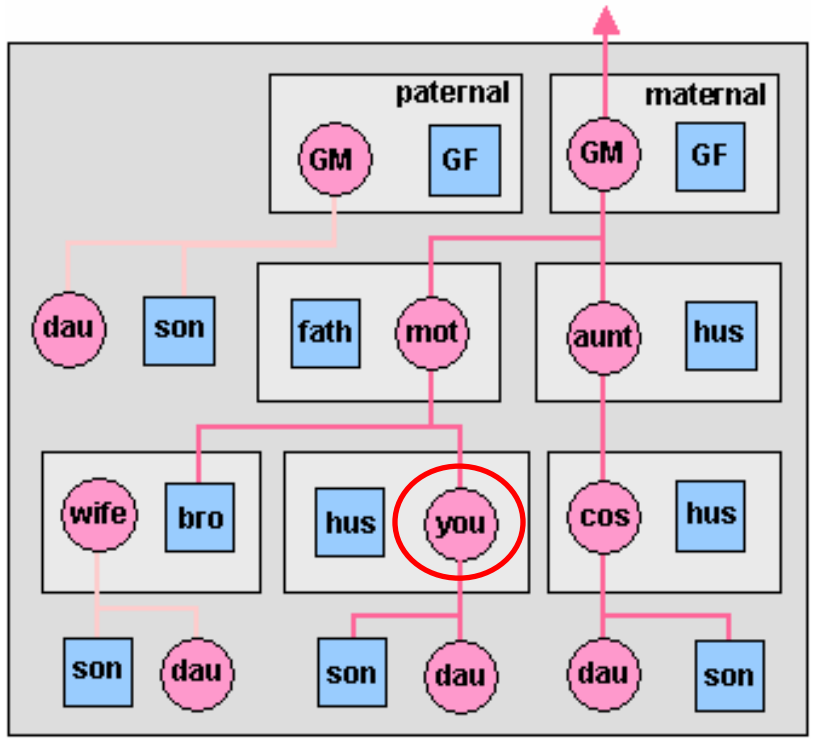


http://evolution.berkeley.edu/evolibrary/news/071101_genealogy

Mitochondrial DNA (mtDNA)

- mtDNA is particularly useful for tracing one's direct maternal line





All persons connected by pink lines have common mtDNA

Autosomal DNA

- **Autosomal DNA** is the type of DNA responsible for most physical characteristics (height, eye color, etc.)
- **Autosomal DNA is inherited** by sons and daughters **from both parents** (and from all four grandparents, etc).
- Currently, autosomal DNA is not used extensively for genealogy.

How Can DNA Identify Us?

- Mutations and the shuffle of maternal and paternal genes through sexual reproduction, ensure that each member of a species (except identical twins) has a unique DNA sequence

- The ideal way to distinguish an individual from all the other people on Earth would be to describe the entire sequence of nucleotides in his or her DNA

- However, since each human genome (all the DNA in a person's chromosomes) is made up of **more than 3 billion nucleotide basepairs**, describing an individual's complete DNA would be far to complicated and expensive to be practical!

How Can DNA Identify Us?

- **Only 5% of human DNA codes for physical traits**
- The rest is made of long stretches of nucleotide basepairs with no apparent function ("junk" DNA)
- Within these nonfunctional stretches are short, moderately repetitive base pair sequences. The number of repeats is inherited and is easily detectable making them ideal identifying markers

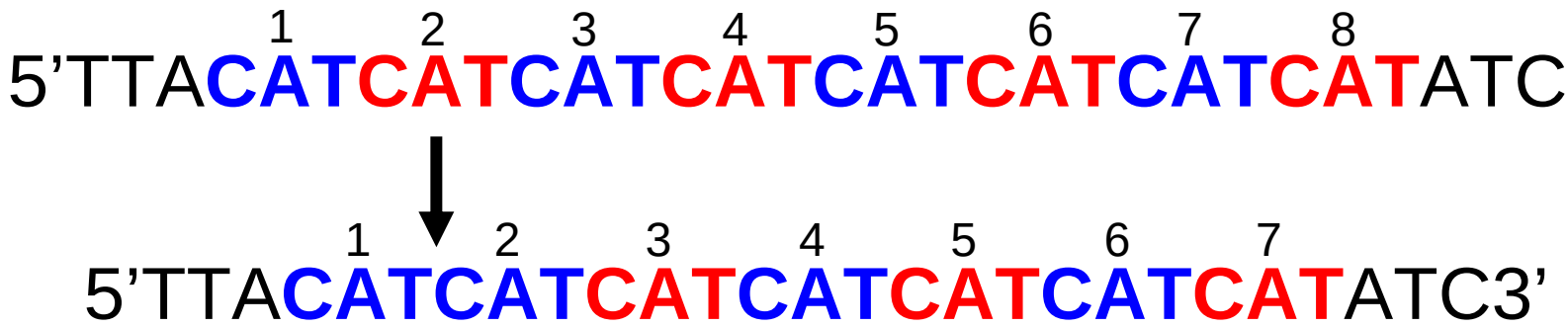
1 2 3 4 5 6 7 8

TTA**CATCATCATCATCATCATCATCAT**ATC

- In this example the nucleotides “CAT” are repeated 8 times

How Can DNA Identify Us?

- The number of repeating units can occasionally change during evolution and descent
- Example:



- They are thus useful markers for familial relationships and have been used in paternity testing, forensic science and in the identification of human remains

How Can DNA Identify Us?

There are two types of repetitive DNA sequences

- **VNTRs** (variable number tandem repeats) are repeated sequences that typically range from 10 to 80 bps. These occur fairly frequently in the human genome but there are relatively few different types

- **STRs** (Short tandem repeats, a.k.a. microsatellites) are much shorter (2-10 bps) and may be repeated as many as 100 times at a given location on a chromosome
  - The human genome contains hundreds of thousands of these STRs all evenly distributed on all the chromosomes

How Can DNA Identify Us?

In 1997, the FBI announced the selection of 13 STR markers to be used in forensic investigations

If any two samples of DNA obtained from different sources (say a crime scene and a suspect) have matching numbers of repeats at all 13 markers, it is virtually certain they are from the same person

How Is This Used for Genealogy?

- The location of these STRs is known and the number of repeats can be analyzed
- These unique locations are called "markers“
- Markers are identified with specific nomenclature
  - Chromosomal markers are described by
    - D (number of chromosome or X or Y)
    - S (Numbers the signify a place on the chromosome).
    - Example: DYS392 or D21S458
  - Mitochondrial DNA (mtDNA) is a continuous circle of 16569 DNA bases. Locations are designated by number from 00001 to 16569

How Is This Used for Genealogy?

- Since each marker represents a specific location on a chromosome we can use this to report the number of repeats at a specific STR.

1 2 3 4 5 6 7 8

TTA**CATCATCATCATCATCATCATCAT**ATC

- Note there are eight repeats of the segment CAT. The number of repeats is the "value" that is shown on a DNA test report for the marker.

- In this example, a lab report would show DYS391=8

Haplotype

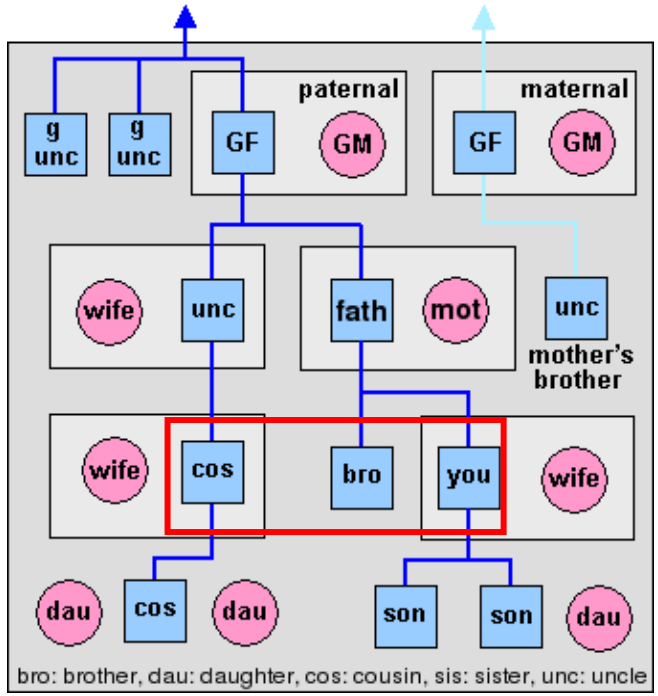
- DNA test reports typically show a series of markers and their corresponding values.
- These results are referred to as a "haplotype."
- The example below would be a 12-marker haplotype.

DYS391 **12**	DYS456 **15**	DYS876 **13**	DYS124 **12**
DYS234 **29**	DYS654 **22**	DYS938 **10**	DYS336 **11**
DYS765 **12**	DYS243 **16**	DYS569 **11**	DYS700 **15**

How Is This Used for Genealogy?

- DNA marker values are like telephone number
  - Consider that the same seven-digit telephone number might appear in both Boston and Miami. However, adding more numbers ("area codes") allows us to distinguish between regions.
  - The same thing is true about DNA marker values
- Comparing limited numbers of DNA markers makes it for two individuals to have the same marker values and not be closely related.
  - The FBI uses 13 markers to identify an individual most molecular genealogy recommend using 36 markers for genealogy work.
- Testing for more markers helps avoid ambiguity. In general, the more markers tested, the easier it is to distinguish individuals and family tree branches.

How Is This Used for Genealogy?



Using our example, let's look at "you", bro", and "cos", who have the following Y-DNA 18-marker haplotype:

YOU: 11 14 12 13 29 13 12 15 12 12 13 12 12 12 14 25 19 **30**
BRO: 11 14 12 13 29 13 12 15 12 12 13 12 12 12 14 25 19 **30**
COS: 11 14 12 13 29 13 12 15 12 12 13 12 12 12 14 25 19 31

- Note that YOU and BRO have the same haplotype, but COS has a difference of one marker value (31 instead of 30).
- That difference would have been due to a mutation that occurred in his Y-DNA (or his father's), but not in the other cousins.
- In general, the greater the number of mutations we find between two individuals, the further in the past their common ancestor lived.

Review Slide

- Males who share a common paternal ancestor will have virtually the same Y-chromosome DNA.
- Everyone will have virtually the same mtDNA as their maternal ancestors.
- We use the word "virtually" since occasionally there are small changes or "copy errors" that might occur with each descendant.
- Those copy errors are called "mutations" and are generally harmless, but are useful for tracing one's direct paternal or maternal line.

What is Currently Available for Genetic Genealogy?

- There are many commercial DNA testing companies that will collect your DNA and analyze it
- However, **there are few databases** that will allow you to compare your DNA to that of other people
- The largest is run by the **Sorenson Molecular Genealogy Foundation (www.smgf.org)**
- The SMGF have both a Y-DNA and mt-DNA database open to the public

DNA Databases

- Most databases have a strong focus on research
  - They are used for genealogic work
  - Main focus is to use the information collected for population genetics research
- Population genetics is a field of research that relies on DNA analysis to determine how groups of people are related, where they came from and how they have migrated over centuries

DNA Databases

- Because of their research emphasis, most databases welcome individual submissions of DNA and a pedigree chart

- At the **Sorenson Molecular Genealogy Foundation** you submit a DNA sample and a four generation pedigree chart

- **This is all free!** but will not help you directly with your genealogy because you will not receive the haplotype of your DNA

DNA Databases

- All databases have a privacy policy like this one:

  "To protect privacy, we do not reveal any personal information of our participants, or the names of any ancestors born after 1906. You may find references to ancestors in SMGF pedigree charts on message boards and in collections of family trees"

- You can still use the pedigree charts that others have submitted to augment your own.

DNA Databases

- To take full advantage of the database, you must know your haplotype
- You must pay to determine your haplotype
- There are many commercial labs that do this at prices ranging from $100-$300
  - (see additional info at the end of this presentation)
- Once you have your haplotype you may use the databases to **search by surname and haplotype**

Benefits of Molecular Genealogy

- A means to check or supplement the historical record with information from genetic data
- A positive test match with another individual may:
  - provide locations for further genealogical research
  - help determine ancestral homeland
  - discover living relatives
  - validate existing research
  - confirm or deny suspected connections between families
  - prove or disprove theories regarding ancestry

Genetic genealogy. (2009, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 20:06, March 25, 2009, from http://en.wikipedia.org/w/index.php?title=Genetic_genealogy&oldid=279634297

Drawbacks or Concerns

- Cost
- Quality of testing
- Privacy issues
- Loss of ethnic identity
- Finally, Y-DNA and mtDNA tests each only trace a single lineage
  - At 10 generations back, an individual has up to 1024 unique ancestors (fewer if ancestor cousins interbred)
  - However, most genealogists maintain contact with many cousins (1st, 2nd, 3rd, etc., with different surnames) whose Y-DNA and mtDNA are different, and thus can be encouraged to be tested to find additional ancestral DNA lineages.

Genetic genealogy. (2009, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 20:06, March 25, 2009, from

http://en.wikipedia.org/w/index.php?title=Genetic_genealogy&oldid=279634297

Not Perfect, But Promising Possibilities

- Genetic genealogy is based on probabilities, and like forecasting the weather, is not an exact science. It can provide important clues for family history research but traditional genealogy methods continue to be an important part of molecular genealogy

- As databases become larger Molecular Genealogy will become a more powerful tool in genealogy work.

Recent Developments: Social Networking and Genealogy

- Imagine the power of the internet harnessed not only to **RESEARCH** your ancestors but also to **CONNECT** with newly discovered relatives (and their ancestors)!
- GeneTree.com has exclusive rights to the SMGF database (including about 6 million ancestral links)
- Incorporates digital audio/video compression and encoding software
- Users may choose to have their haplotype analysis done
- When results come, and through their social network, you have the ability to connect and make “friends” with members who are genetically related to you!

www.genetree.com

Links to Check Out

DNA Testing Companies

African Ancestry
Ancestry.com
deCODEme
DNA Heritage
Ethnoancestry
Family Tree DNA
Gene Tree
Genetic Ancestor (mtDNA only)
GeoGene
NGS Genographic Project
Oxford Ancestors
Trace Genetics
23andMe

DNA Databases

SMGF
Ancestry.com
Gene Tree
ysearch
yhrd
ybase
mitosearch
DNA-Fingerprint
Oxford Ancestors

DNA Surname Projects

Through DNA Heritage
Through Family Tree DNA

Further reading: Books on Molecular Genealogy

Trace Your Roots with DNA : Using Genetic Tests to Explore Your Family Tree by Ann Turner and Megan Smolenyak

DNA and Family History: How Genetic Testing Can Advance Your Genealogical Research by Steve Jones (Foreword) and Chris Pomery

Forensic Genealogy by Colleen Fitzpatrick

Unlocking Your Genetic History by Thomas H. Shawker

The Journey of Man : A Genetic Odyssey by Spencer Wells

The Seven Daughters of Eve by Bryan Sykes

Mapping Human History : Genes, Race, and Our Common Origins by Steve Olson

The Real Eve: Modern Man's Journey Out of Africa by Stephen Oppenheimer

Deep Ancestry by Spencer Wells

Family History in the Genes: Trace Your DNA and Grow Your Family Tree by Chris Pomery

DNA & Genealogy by Colleen Fitzpatrick

Adam's Curse by Bryan Sykes