Intro to Genetic Genealogy

Craig Rhombs

Welcome

- Who am I?
- Why am I doing this?
- Definition of genealogy = “an account of the descent of a person, family, or group from an ancestor or from older forms”
- Basic objectives
  - Present the basics from biology
  - Explain the practical application to genealogy

Where Can I Find Class Materials?

https://281connections.us

Email: 281connections@gmail.com

Topic Outline

- Biological Groundings
  - The Cell and DNA
  - Nuclear DNA
    - Autosomal DNA
    - Sex chromosomes
    - Statistical processes
  - Mitochondrial DNA
- Genetic Testing for Genealogy
  - What you get
  - Interpreting autosomal match data
  - Interpreting yDNA and mtDNA match data
  - Examples from familytreedna.com and ancestry.com
  - Rules of Thumb and Cautions

Attitudes & Respect

- "I don't know who my grandfather was, I am much more concerned to know what his grandson will be." -- Abraham Lincoln
- "When a society or a civilization perishes, one condition can always be found. They forgot where they came from." -- Carl Sandburg
- "We are the children of many sires, and every drop of blood in us in its turn ... betrays its ancestor." -- Ralph Waldo Emerson

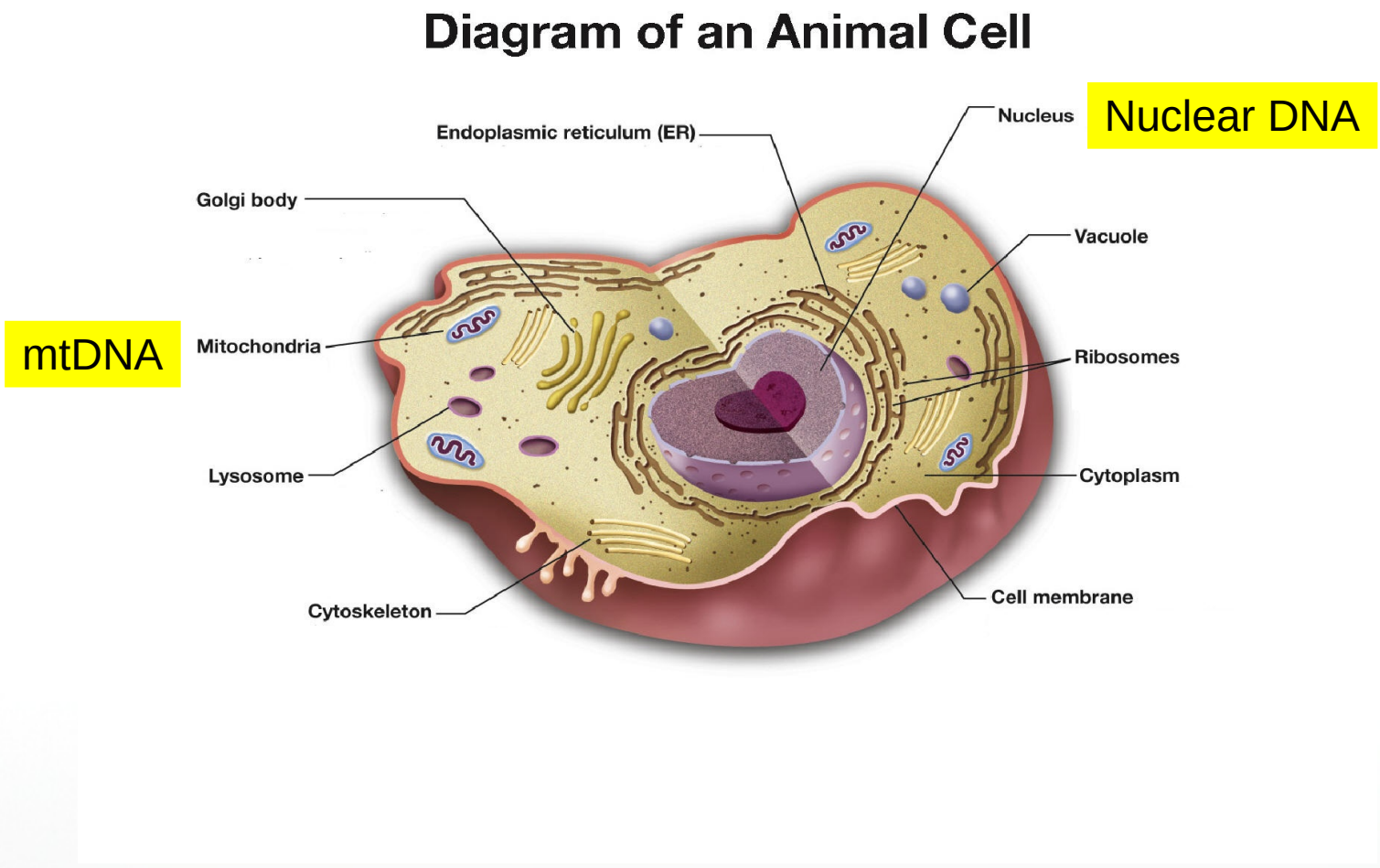
Review Expectations

- This type of analysis may be difficult.
- It will cost $.
- It may reveal stuff that you are uncomfortable with.
- It will be most useful if you have already done significant tradition research.
- It may also help you confirm suspicions or unclear results from traditional research.
- Some analytic results will vary for good reasons.
- There is a lot of marketing hype about what to expect.

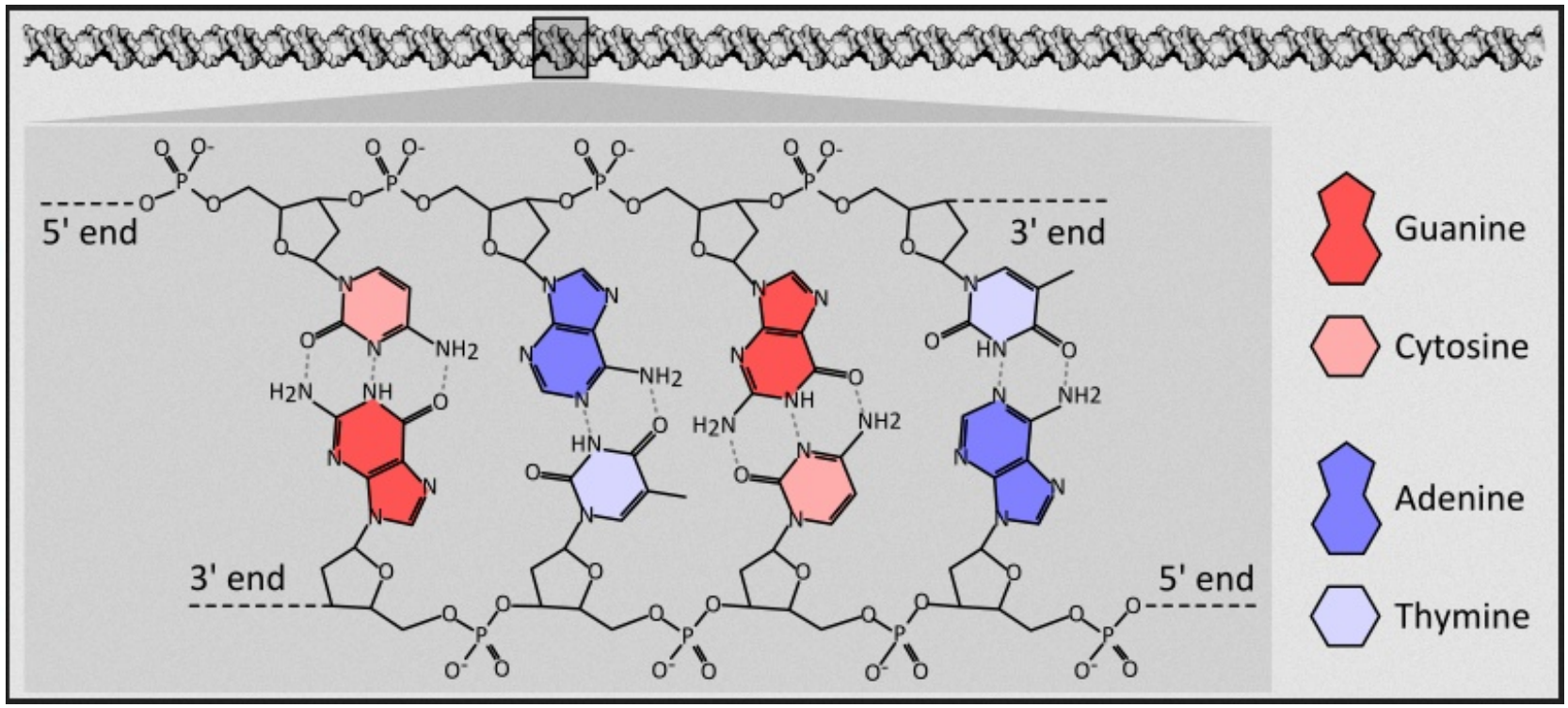
More on Expectations

- Find cousins you did not know you had.
- Biology does not obfuscate.
- There are practical limits to how far back in time you can detect a genetic “signal”.
- Understand the math.
- https://en.wikipedia.org/wiki/Introduction_to_genetics

Cell Components – 2 kinds of DNA



DNA at the Molecular Level



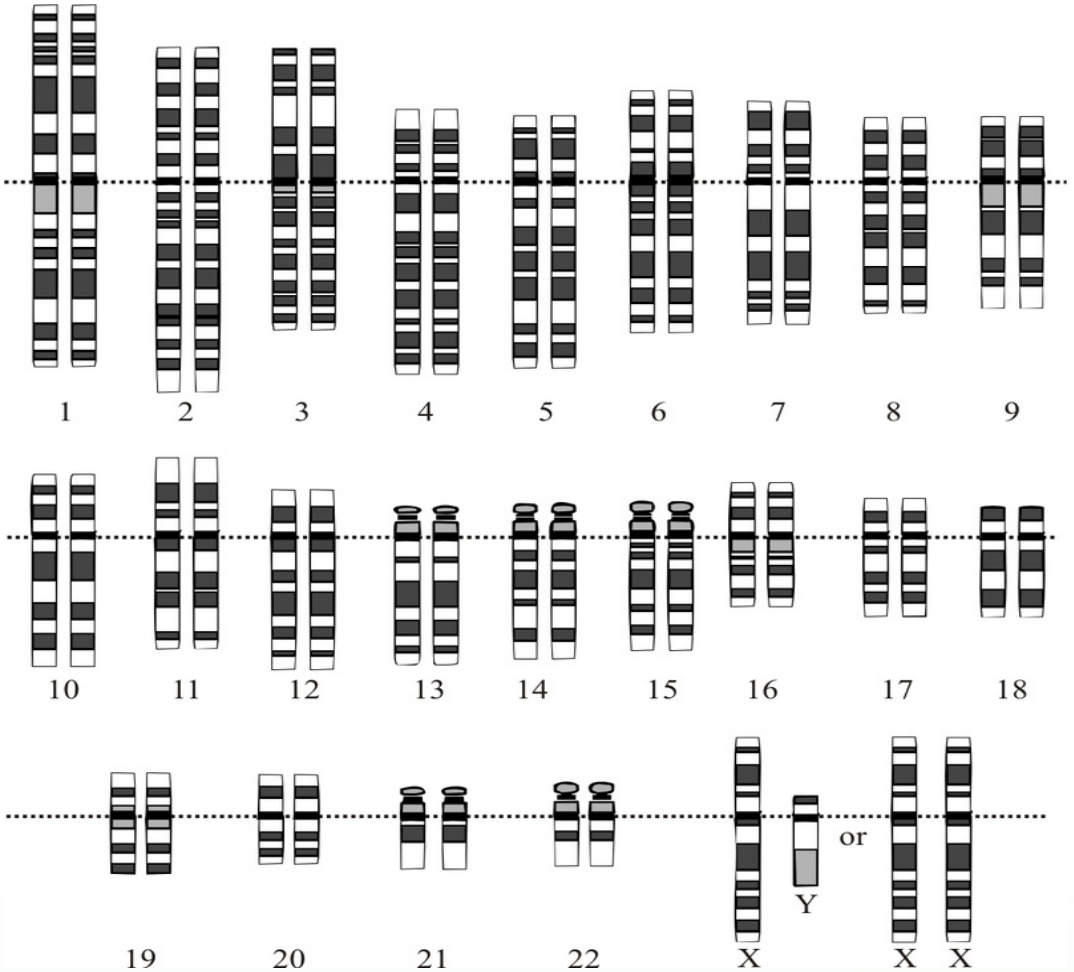
The chemical structure of a four base pair fragment of a DNA double helix. The sugar-phosphate backbone chains run in opposite directions with the bases pointing inwards, base-pairing A to T and C to G with hydrogen bonds.

Nuclear DNA

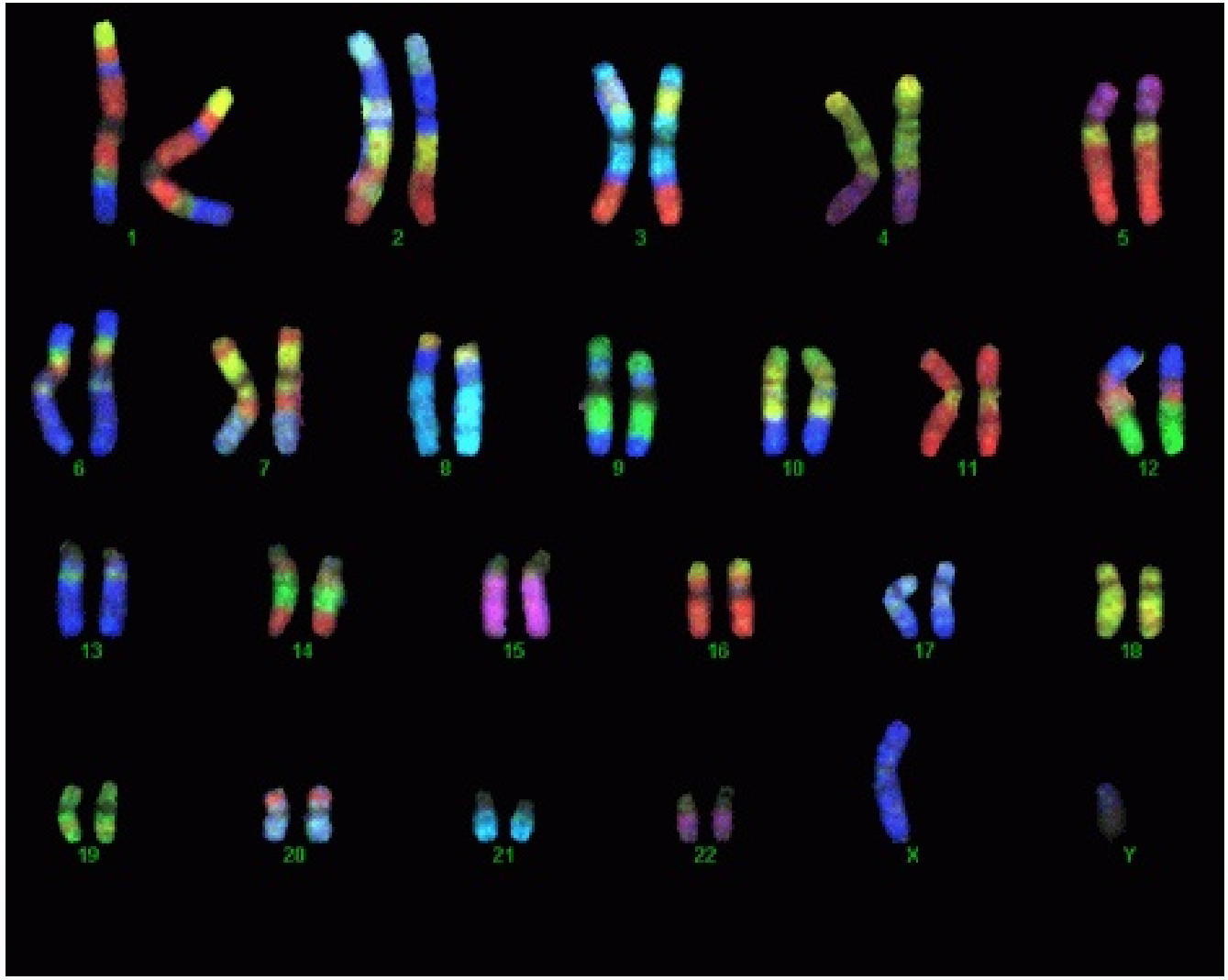
- 23 chromosome pairs (diploid cells)
- 23 single chromosomes (haploid cells)
- Meiosis (creating gametes)
- Genes (parts of chromosomes; alleles)
- Parental contributions
- Special nature of the sex chromosome pair

Human Genome (23 pairs)

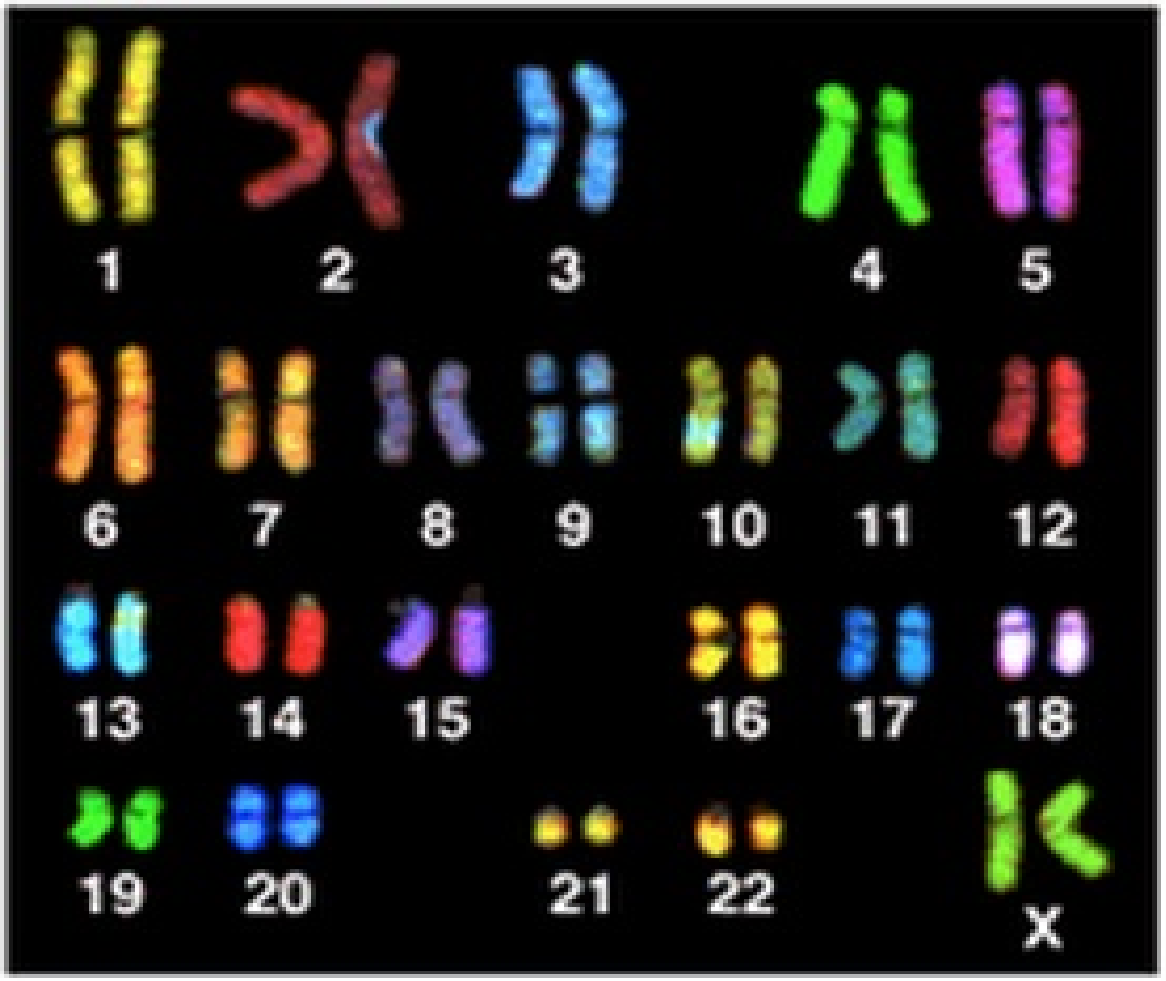
https://en.wikipedia.org/wiki/Human_genome



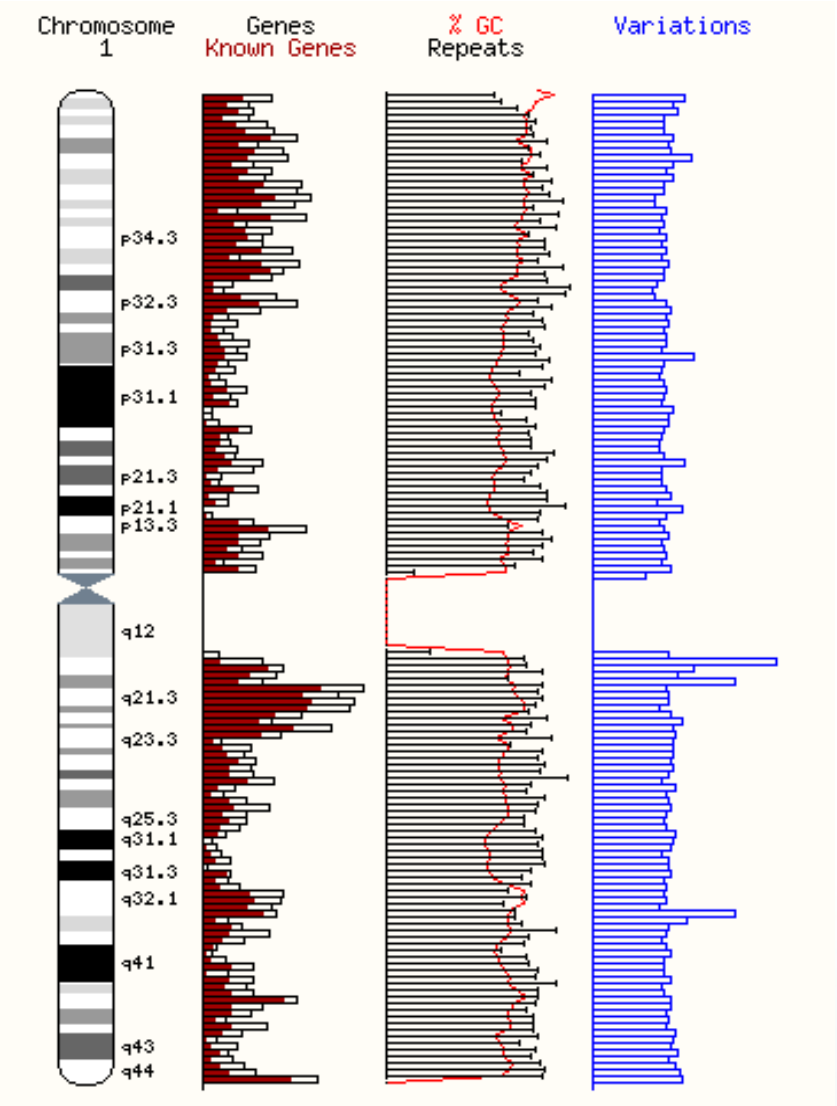
Human Male Karyotype Example



Human Female Karyotype Example



Genes


Chromosome
1
Genes
Known Genes
% GC
Repeats
Variations
p34.3
p32.3
p31.3
p31.1
p21.3
p21.1
p13.3
q12
q21.3
q23.3
q25.3
q31.1
q31.3
q32.1
q41
q43
q44

Autosomal DNA

- Chromosome Pairs 1-22
- A true balanced mix of mother and father
- Each pair composed of one from father and one from mother.

Theoretical Composition of Autosomal DNA

Gen 0	Gen 1	Gen 2	Gen 3	Gen 4	Gen 5
1	0.5	0.25	0.125	0.0625	0.03125
YBP 0	YBP 30	YBP 60	YBP 90	YBP 120	YBP 150
you	1-1 Dad	2-1 Pgd	3-1 G	4-1 GG	5-1 GGG
					5-2 GGG
				4-2 GG	5-3
					5-4
			3-2 G	4-3 GG	5-5
					5-6
				4-4 GG	5-7
					5-8
		2-2 Pgm	3-3 G	4-5 GG	5-9
					5-10
				4-6 GG	5-11
					5-12
			3-4 G	4-7 GG	5-13
					5-14
				4-8 GG	5-15
					5-16
	1-2 Mom	2-3 Mgd	3-5 G	4-9 GG	5-17
					5-18
				4-10 GG	5-19
					5-20
			3-6 G	4-11 GG	5-21
					5-22
				4-12 GG	5-23
					5-24
		2-4 Mgm	3-7 G	4-13 GG	5-25
					5-26
				4-14 GG	5-27
					5-28
			3-8 G	4-15 GG	5-29
					5-30
				4-16 GG	5-31
					5-32

Generation number.

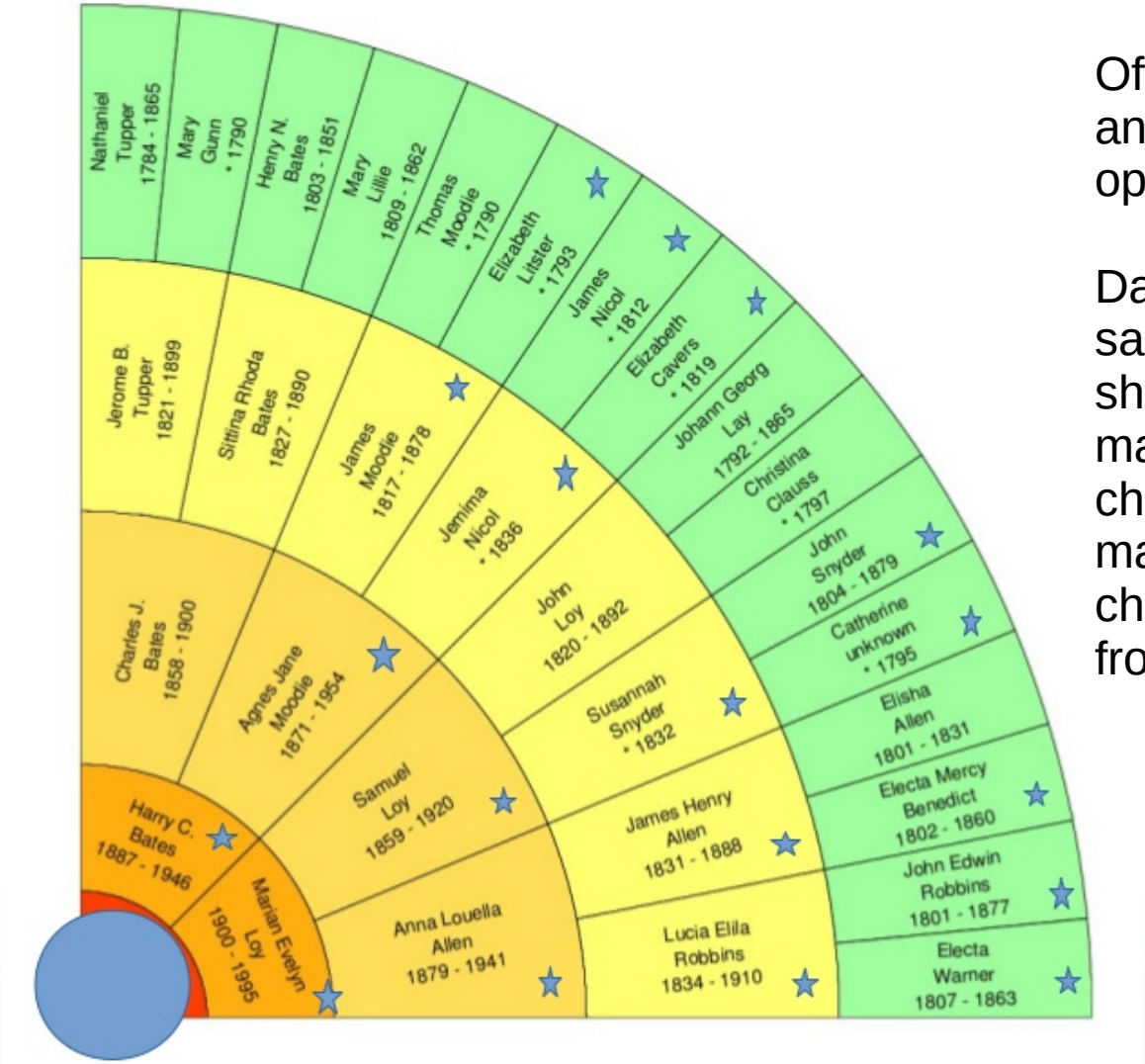
Fractional contribution to you.

Years before present assuming 30 years per generation.

23rd Pair (XX, XY)

- Y Chromosome
  - Only from father to son
  - Changes faster than mtDNA
  - Only for tracing paternal line
- X Chromosome
  - Males get one chromosome from mother
  - Females get one from each parent (only one is active in a given cell)
  - Presents some interesting analytic opportunities

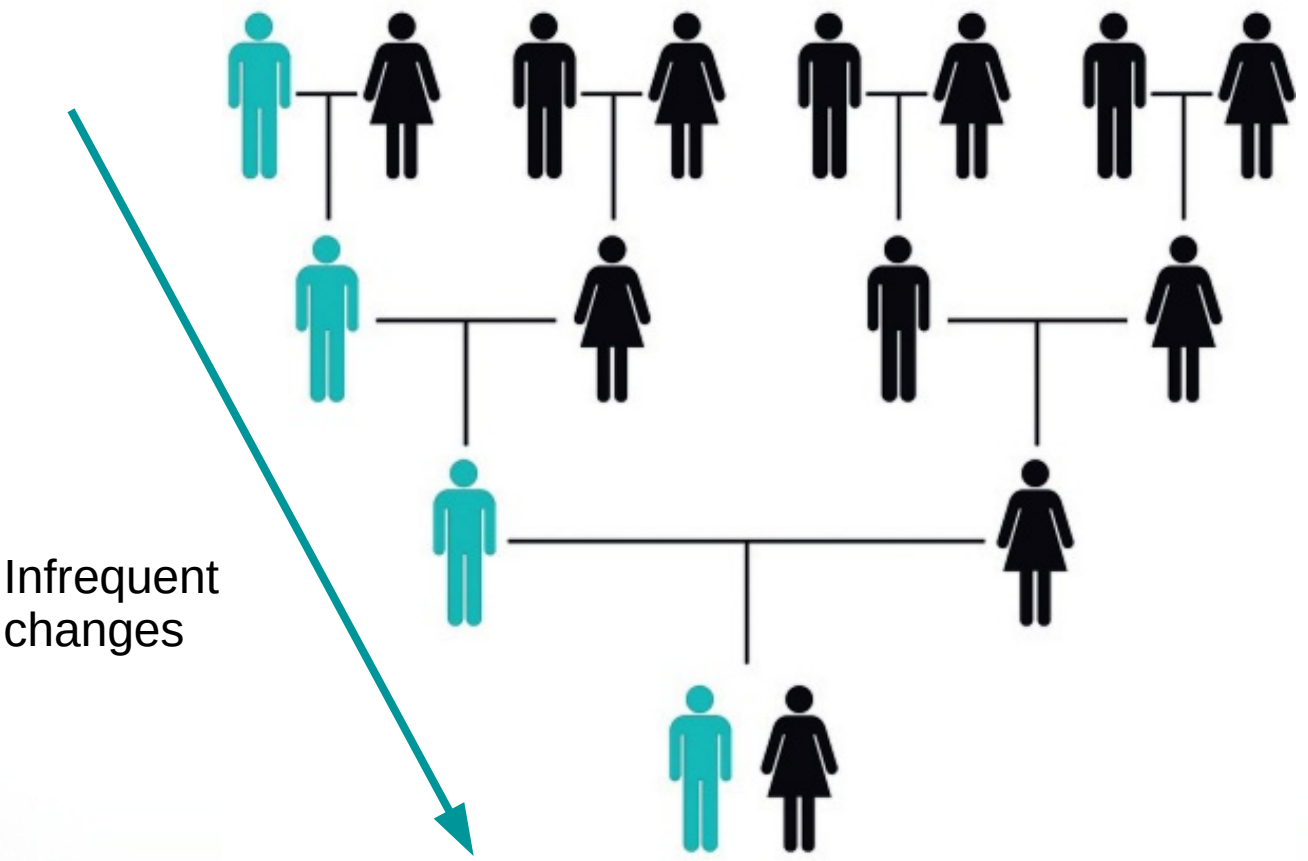
X Inheritance Example



Offers special analytic opportunities -

Daughters of the same father should have matching x chromosomes; males inherit x chromosome only from mother

Y Inheritance (paternal line)

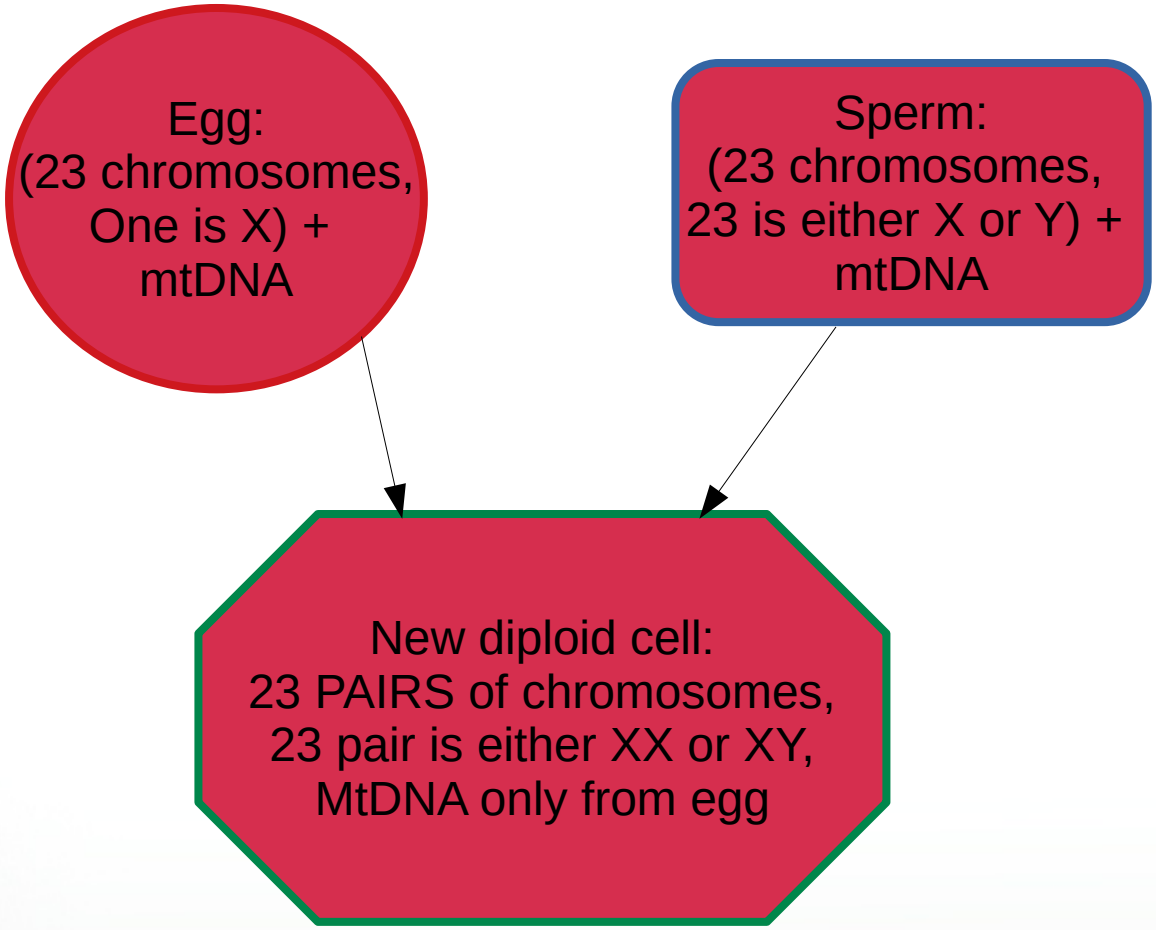


from Jordi Picart, copyleft CC-BY-SA
https://creativecommons.org/licenses/by-sa/4.0/deed.en

Meiosis

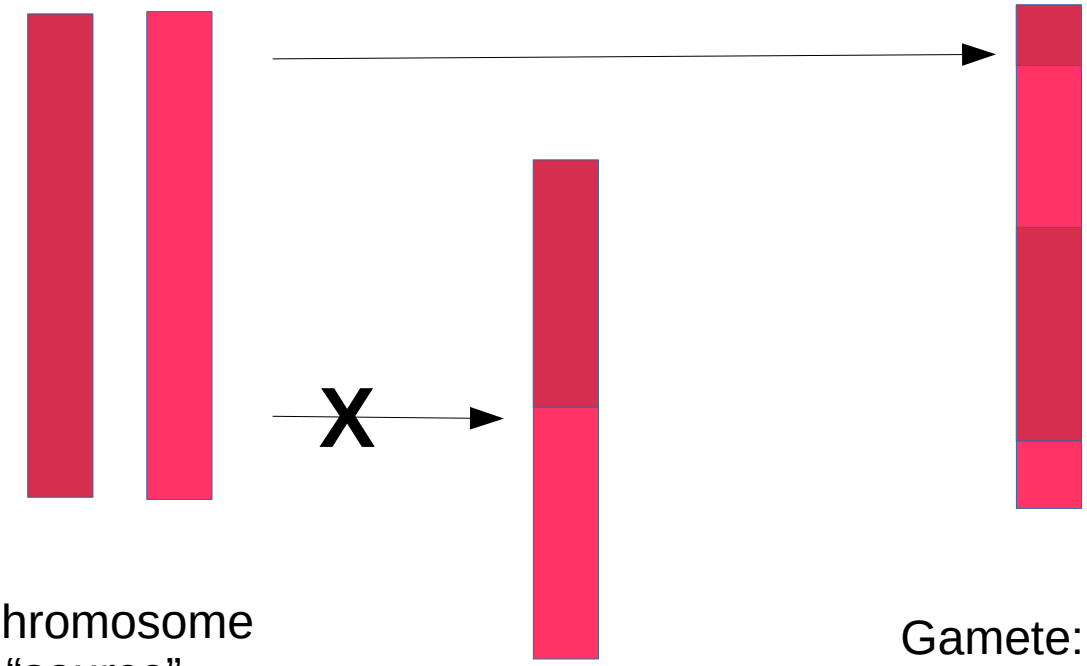
- Meiosis – one diploid cell creates 4 gametes
- Duplication of chromosome pairs to produce sister pairs
- Production of haploid cells
- Crossover (mixing mother and father traits) yields uncertainty in inheritance. (statistical processes)
- # of possible combinations is huge !
- This is why children of one pair of parents can be so different.

Haploid Cells (Gametes) and Diploid Cells



Crossover During Meiosis

(Homologous Recombination)



One Chromosome pair in “source” cell -

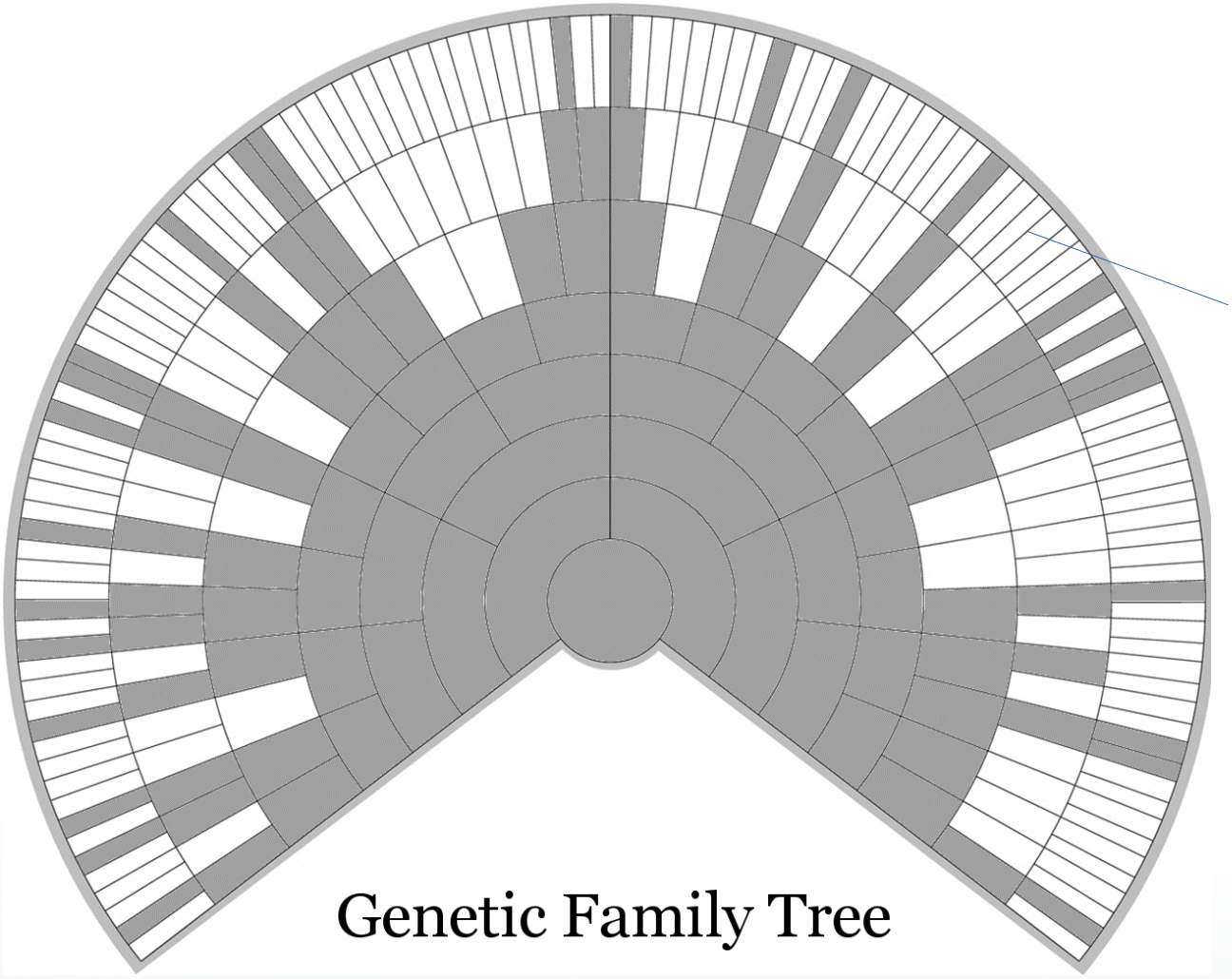
One from mother.
One from father.

Gamete: single chromosome if crossover did not have random aspects.

Gamete: single chromosome illustrating randomness. (4x, each different)

Inheritance is a Statistical Process

adapted from a graphic by Blaine Bettinger



A “signal” can be lost.

Actual Composition of Autosomal DNA

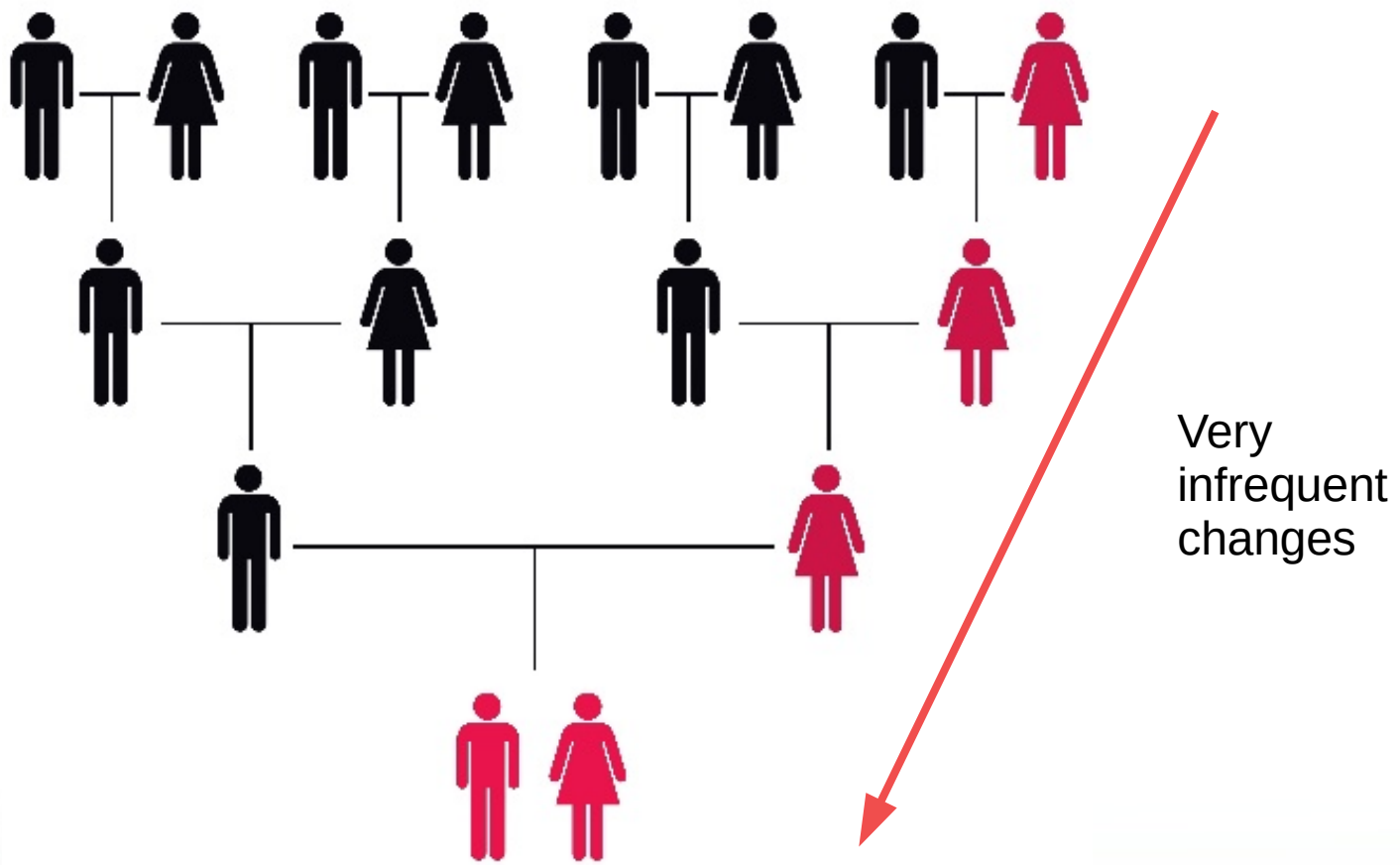
Gen 0	Gen 1	Gen 2	Gen 3	Gen 4	Gen 5
1	0.5	0.25	0.125	0.0625	0.03125
YBP 0	YBP 30	YBP 60	YBP 90	YBP 120	YBP 150
you	1-1 Dad	2-1 Pgd	3-1 G	4-1 GG	5-1 GGG
					5-2 GGG
				4-2 GG	5-3
					5-4
			3-2 G	4-3 GG	5-5
					5-6
				4-4 GG	5-7
					5-8
		2-2 Pgm	3-3 G	4-5 GG	5-9
					5-10
				4-6 GG	5-11
					5-12
			3-4 G	4-7 GG	5-13
					5-14
				4-8 GG	5-15
					5-16
	1-2 Mom	2-3 Mgd	3-5 G	4-9 GG	5-17
					5-18
				4-10 GG	5-19
					5-20
			3-6 G	4-11 GG	5-21
					5-22
				4-12 GG	5-23
					5-24
		2-4 Mgm	3-7 G	4-13 GG	5-25
					5-26
				4-14 GG	5-27
					5-28
			3-8 G	4-15 GG	5-29
					5-30
				4-16 GG	5-31
					5-32

Fractional contribution to your DNA. These are theoretical values. Actual values will differ because of the statistical effects induced by crossover during meiosis.

Mitochondrial DNA

- Small but important
- Outside the nucleus
- Comes from mother so can be used to trace maternal lineage
- Stable over many generations

mtDNA Inheritance



from Jordi Picart, copyleft CC-BY-SA
https://creativecommons.org/licenses/by-sa/4.0/deed.en

Summary

- Nuclear DNA
  - 23 pairs of chromosomes
  - One in each pair from father, the other from mother
  - 22 autosomal, 1 sex pair (X & Y)
  - Y passed down from father to sons (fairly stable)
  - Statistical processes for X and autosomal
- Mitochondrial DNA
  - Passed down from mother to all offspring
  - Very stable over time


BREAK

Genetic Genealogy in Practice

- So what? How can DNA help me with Genealogy?
- Find cousins you did not know you had.
- Use to clarify/confirm ancestry.
- Biology does not obfuscate.
- There are practical limits to how far back in time you can detect a genetic “signal”.
- Understand the math.
- Have reasonable expectations
- This costs $.
- https://en.wikipedia.org/wiki/Introduction_to_genetics

Genetic Testing

- Example Companies:
- www.23andme.com
- www.familytreedna.com
- www.africanancestry.com
- www.ancestry.com
- https://www.myheritage.com
- https://www.livingdna.com/en-us (new, emphasis on GB)
- How? $? How much time?
- Comparison:
  https://www.smarterhobby.com/genealogy/best-dna-test/

Transferring Test Results

- In some cases, you can test with one company and transfer raw autosomal data to another for matching.
- May cost less than testing at multiple places.
- Expand the universe of potential matches.
- See https://www.yourdnaguide.com/transferring
- Will receive raw data from elsewhere: familytreedna.com, livingdna.com, myheritage.com, gedmatch.org, geneanet.org, . . .

What You Get

- List of Matches with others (autosomal, X, Y, mtDNA)
  - Ancestry.com – autosomal matches only
  - Familytreedna.com – matches based on all types of dna
  - 23andme.com – autosomal matches only, but reports some info about X, Y haplogroup, mt haplogroup
- Chromosome browsers (autosomal)
- Ancestral admixture estimates
  - Be careful! Understand the shortcomings.
  - https://familyhistorydaily.com/genealogy-help-and-how-to/understanding-dna-results/
- Connecting with cousins
- The value of multiple participants (projects)
- Some companies will allow you to upload raw data from other companies. (e.g., FTDNA, MyHeritage)

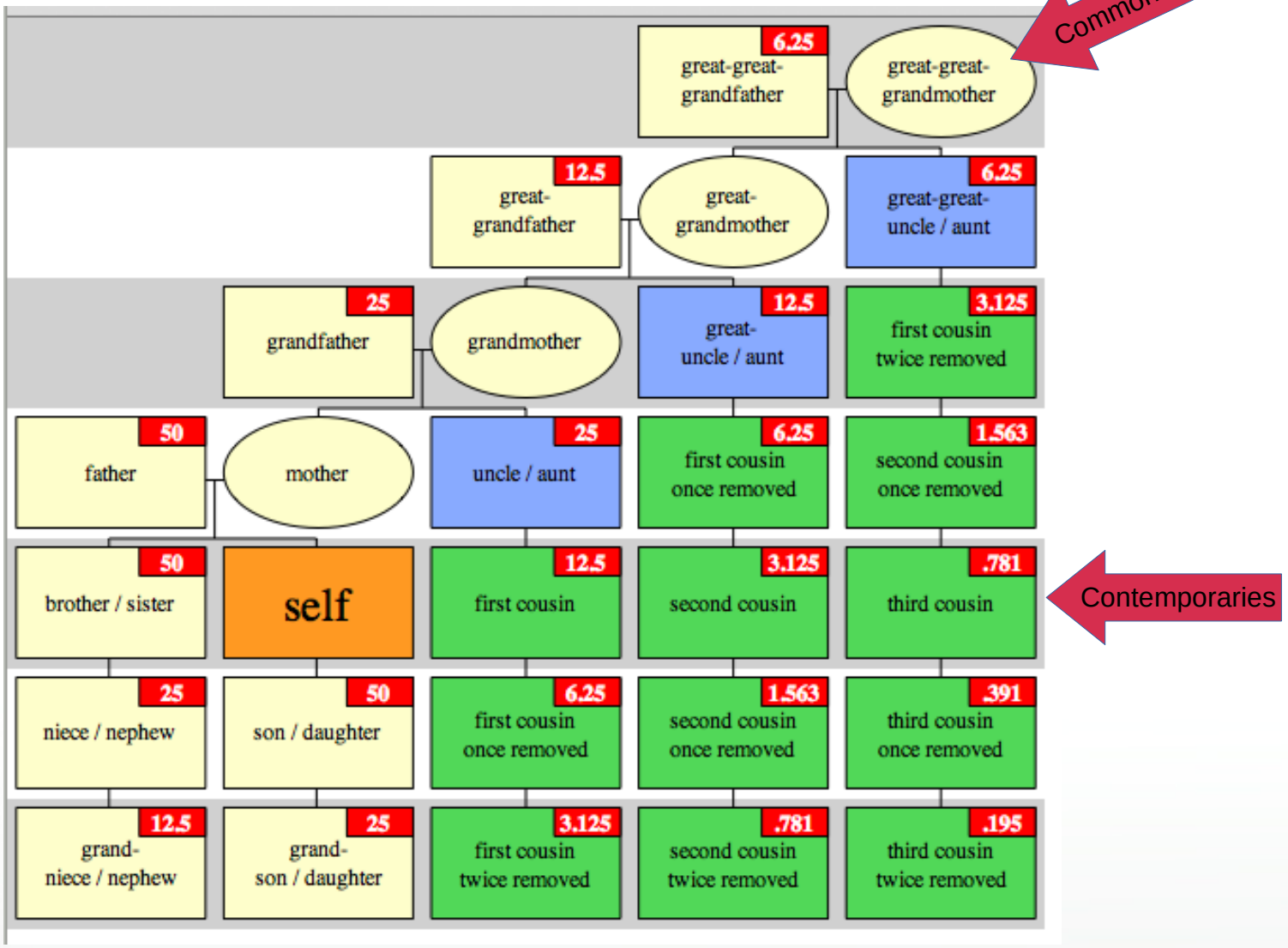
Cautions

- Health/disease info not initiated by your doctor
  - Your health is more complicated than just your DNA
  - Your doc's eyes will roll
- Privacy concerns
  - Read the privacy agreements
  - You are probably protected against use by insurance companies.
  - Workplace testing has failed in the courts already.
  - Law enforcement uses.
- CBC Video: https://youtu.be/Isa5c1p6aC0
- "Science Versus" Podcasts (at gimletmedia.com):
  - https://gimletmedia.com/shows/science-vs/8wh2mk/dna-kits-can-you-trust-them
  - https://gimletmedia.com/shows/science-vs/6nhgxk/race-can-we-see-it-in-our-dna

Autosomal Matches (atDNA)

- Shown as % or cM (centiMorgans).
- There are 68 cM per shared %
- Concept of “distance” to a common ancestor. How many generations might have been required to yield the % match seen?
- Most recent common ancestor (MRCA)

Theoretical Autosomal Shared %



Autosomal Matches - examples

Ancestry.com

Parent/Child
Shared DNA: 3,449 cM across 80 segments

Close Family–1st Cousin
Shared DNA: 1,797 cM across 78 segments

1st–2nd Cousin
Shared DNA: 850 cM across 43 segments

1st–2nd Cousin
Shared DNA: 669 cM across 33 segments

Relationship
Grandparent
Grandchild
Half sibling
Aunt/uncle
Niece/nephew

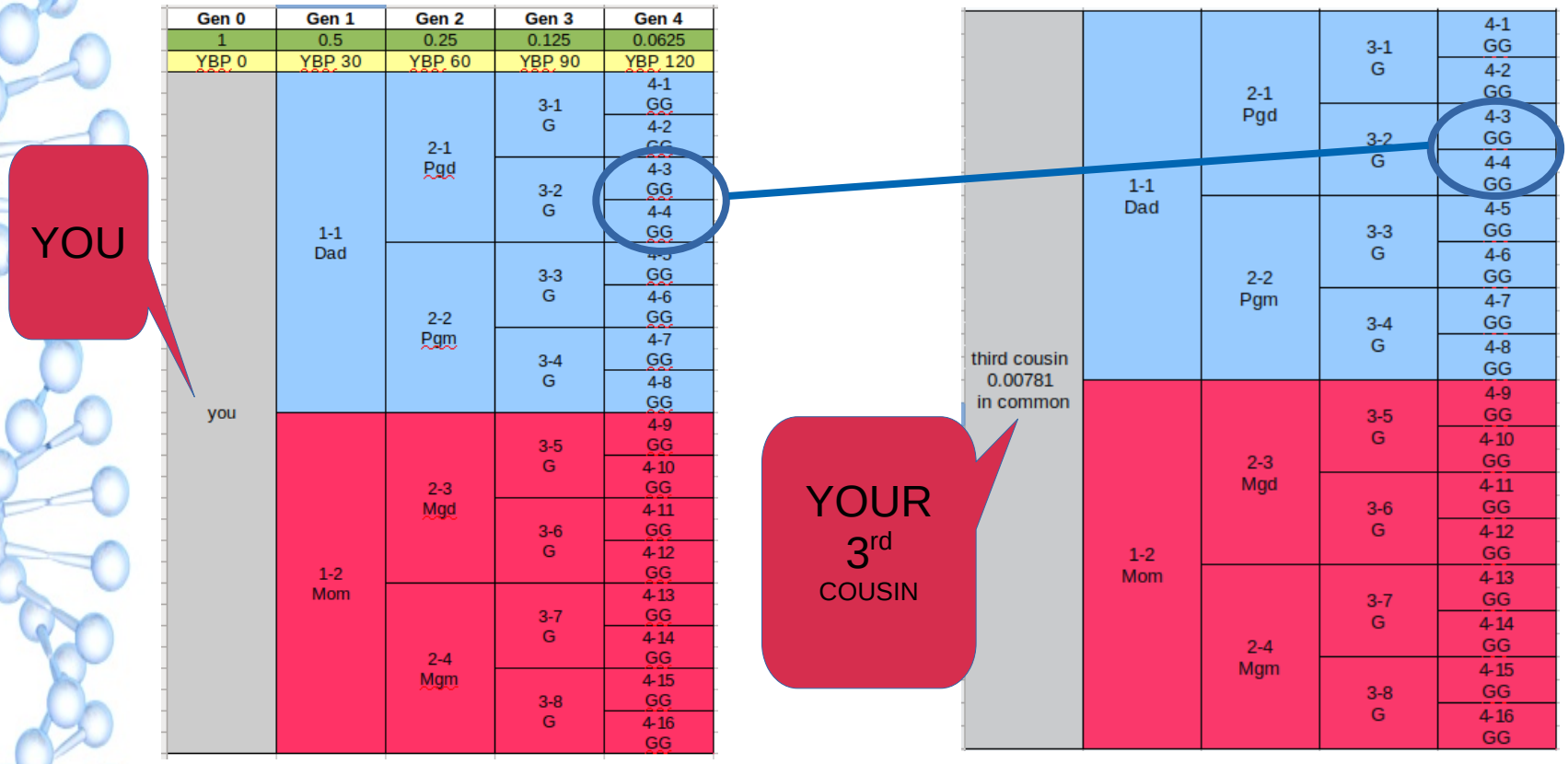
Match Date	Relationship Range	Shared Centimorgans	Longest Block	X-Match
11/21/2012	Father/ Son	3,380	267	X-Match
05/15/2017	1st Cousin - 2nd Cousin	530	56	
03/28/2018	1st Cousin - 3rd Cousin	282	54	X-Match
01/20/2017	2nd Cousin - 3rd Cousin	236	43	

Familytreedna.com

Most Recent Common Ancestor (MRCA)

- You must find a person/couple on your pedigree chart in common with the person with whom you share DNA (your genetic cousin).
- Assume that this genetic cousin is a contemporary. i.e., that he/she is the same number of generations distant from the common ancestor.

MRCA - Third Cousin Example



All people <u>not</u> circled are different. The circled family is common between you and your 3rd cousin. Circled people in theory contribute 0.781% (0.00781) common DNA to you and your cousin. You and your cousin (each) probably had to wade through information about at least 30 people to find the commonality.

MRCA – estimating formula

Cousin degree	common ancestor relation	predicted shared dna fraction with cousin	percent shared dna with cousin	low percent	high percent	generations to common ancestor	est. years to common ancestor	est. birth year for ancestor assuming a start of 1950
1	grandparent	0.12500	12.50000	10.50000	13.50000	2	60	1890
2	g grandparent	0.03125	3.12500	1.62500	4.12500	3	90	1860
3	2xg grandparent	0.00781	0.78125	0.00000	1.28125	4	120	1830
4	3xg grandparent	0.00195	0.19531	0.00000	0.69531	5	150	1800
5	4xg grandparent	0.00049	0.04883	0.00000	0.54883	6	180	1770
6	5xg grandparent	0.00012	0.01221	0.00000	0.26221	7	210	1740
7	6xg grandparent	0.00003	0.00305	0.00000	0.25305	8	240	1710
8	7xg grandparent	0.00001	0.00076	0.00000	0.20076	9	270	1680

NOTE: High and low estimates are only for illustration of the principle that actual values will lie in a range around the predicted values.

MATH NOTE: generations = 0.5 – 0.7213*ln(shared fraction)

MRCA – Calculation Example

You have a match at the 0.4% (27 cM) level with someone. How far back in time did your most recent common ancestor live?

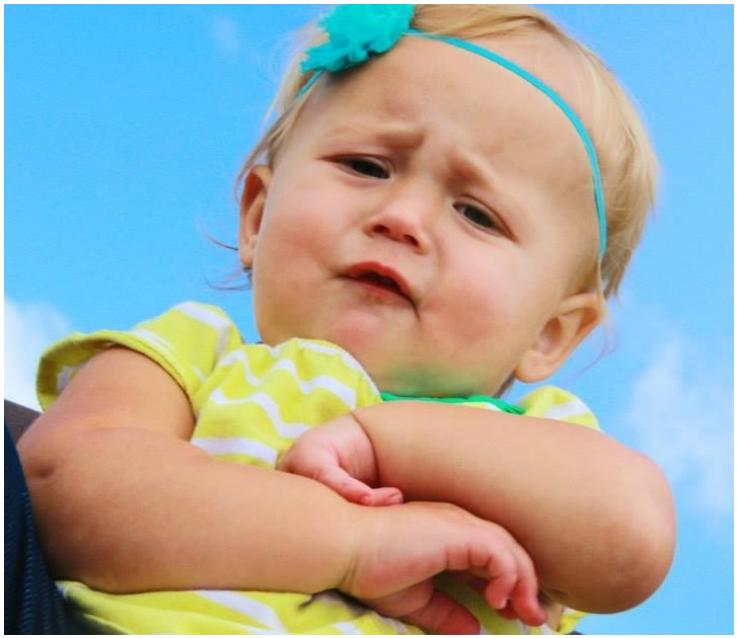
Hints:
Remember: generations = 0.5 – 0.7213*ln(shared fraction)

Solution:
0.4% is the same as a shared fraction of 0.004.
Using the equation above,
Generations = 0.5-0.7213*ln(0.004) = 4.48

So, your most recent common ancestor was born about 4-5 generations before you. That is, it could be a gg-grandparent or a ggg-grandparent. If you were born in 1950 and you assume a generational duration of 30 years, then you are looking for someone born *around* 1800-1830.

Does your family tree go back that far, and do you have confidence in it at that point in time?

MRCA – but wait!



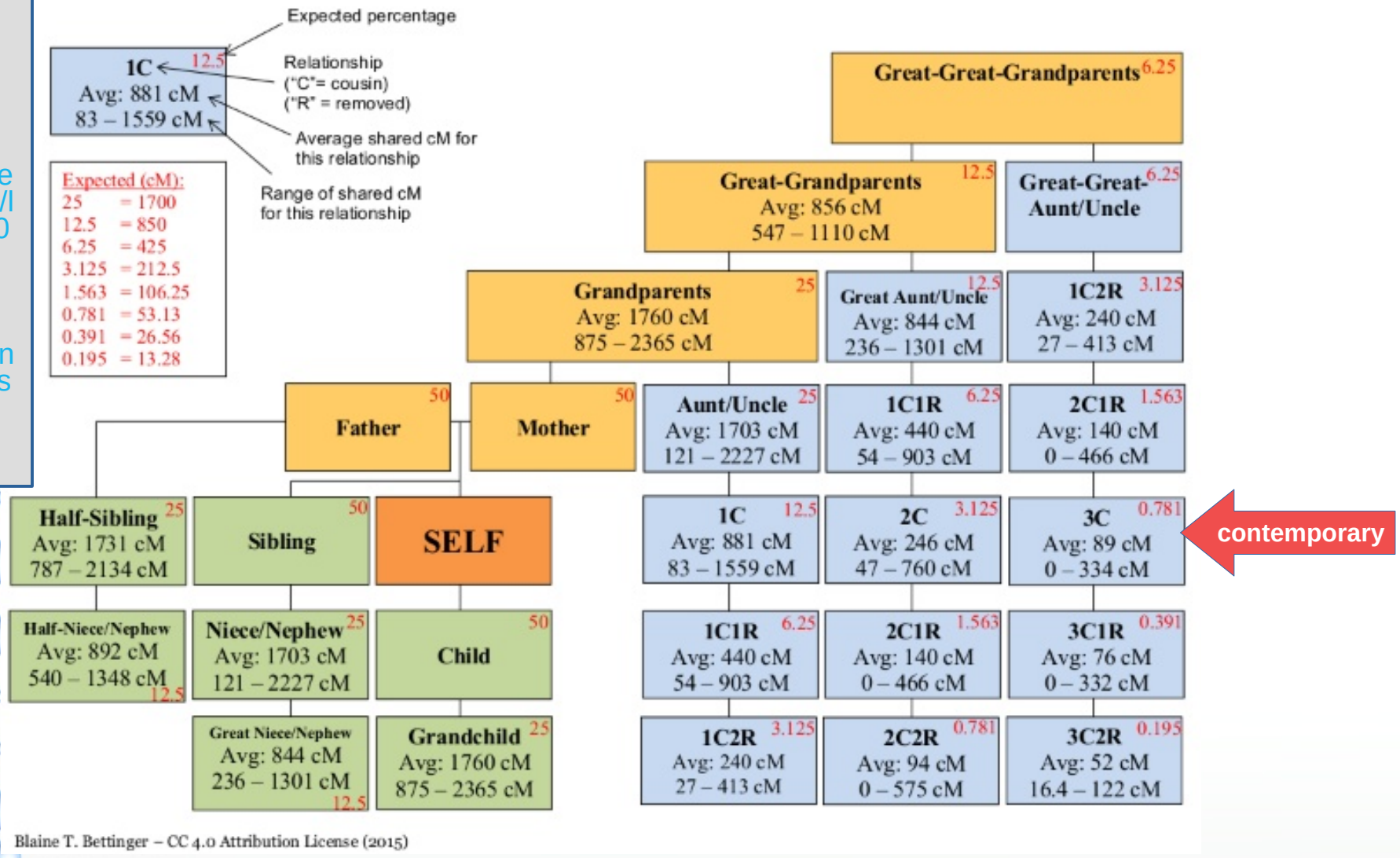
What about the effects due to the statistical aspects of inheritance?

What about cousin marriages?

Shared Autosomal DNA (%)

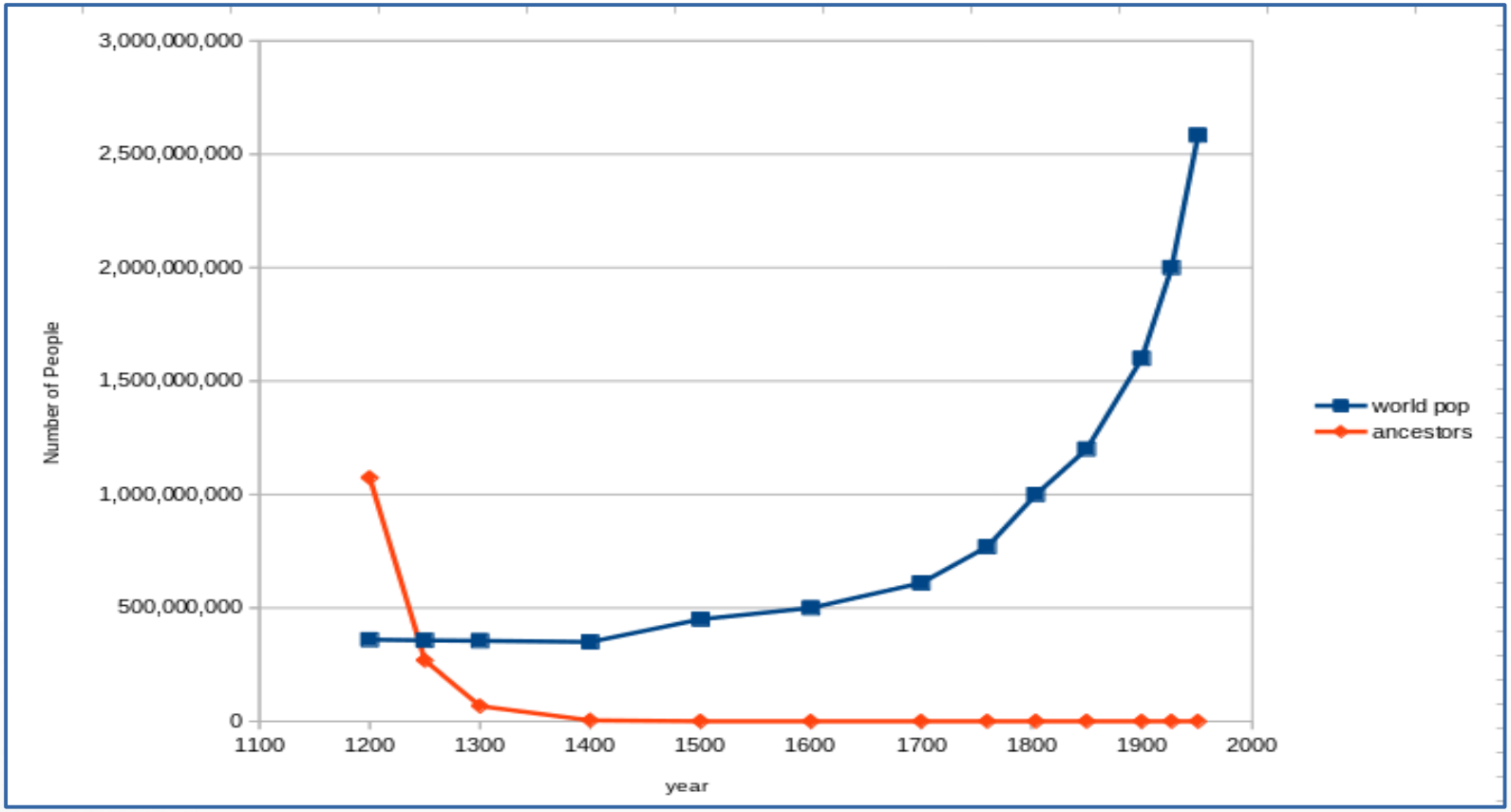
From a diagram by Blaine T. Bettinger, copyleft CC-BY https://creativecommons.org/licenses/by/4.0/deed.en

See also: https://dnapainter.com/tools/sharedcmv4



Cousins Marrying Cousins!?

- If you blindly assume that every generation of your pedigree consists of unrelated individuals, you are sure to be wrong at some point.
- Recent *Scientific American* article: link



Rules of Thumb

- The greater the amount in common, the greater the number of common segments, and the longer the common segments the better.
- Start by paying attention to matches with greater than about 0.4% (27cM) DNA in common.
- Common segments with lengths greater than 7-8 cM are the most likely to not be chance matches. (Again, you are playing a game of statistics)
- Have parents, 1st & 2nd cousins test as well to simplify analysis (if possible). Even sibling tests can help.
- Be ware of what you might find. Be considerate and respectful of privacy.

What about Y and mtDNA?

- These are not mixes like autosomal DNA!
- Both change with time.
- mtDNA changes very slowly & so can be used to reach farther back in time to a common maternal ancestor.
- Y DNA changes more quickly and is more fragile. Can be more easily used than autosomal DNA to trace a male line farther back in time.
- Testing can be useful in some situations.

y-DNA Matches Example

Concept of “distance” between two individuals. Familytreedna helps you interpret the matches with “TiP” reports. The example is for a distance of just 3 between testers. (I would be pretty sure to find a common ancestor IF our trees went back to the 15th century. OOF!)

Y-DNA TiP Report

In comparing Y-DNA 37 marker results, the probability that [redacted] and **Craig Rhombs** shared a common ancestor within the last...

COMPARISON CHART	
Generations	**Percentage**
4	13.35%
8	50.31%
12	78.93%
16	92.6%
20	97.7%
24	99.35%

My tree goes back 9 generations (to a person born in the 1600s).

mtDNA Matches

- Guidance from familytreedna . . . https://www.familytreedna.com/learn/mtdna-testing/tell-closeness-relationship/
- Based on analysis of “hypervariable” and “coding” regions.
- If you have a match, your common ancestor could have existed over a large time span!

Testing Level	Matching Level	Generations to Common Ancestor	
		50% Confidence Interval	95% Confidence Interval
mtDNA	HVR1	52 (about 1,300 years)	— *NA**
mtDNAPlus	HVR1 & HVR2	28 (about 700 years)	— *NA**
mtFullSequence	**HVR1, HVR2, & Coding Region**	**5 (about 125 years)**	**22 (about 550 years)**

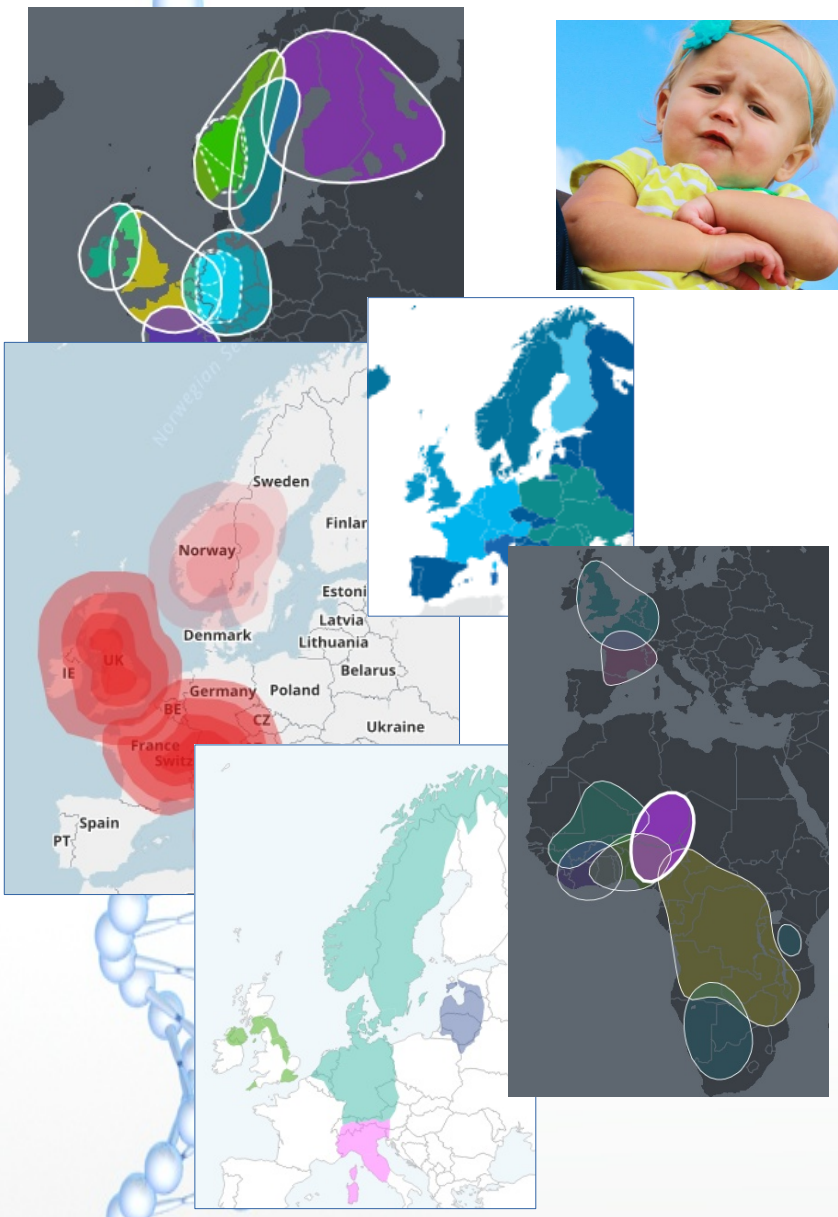
*** The range of generations to a common ancestor at this level is too broad to calculate a 95% confidence period.**

Examples

www.Familytreedna.com

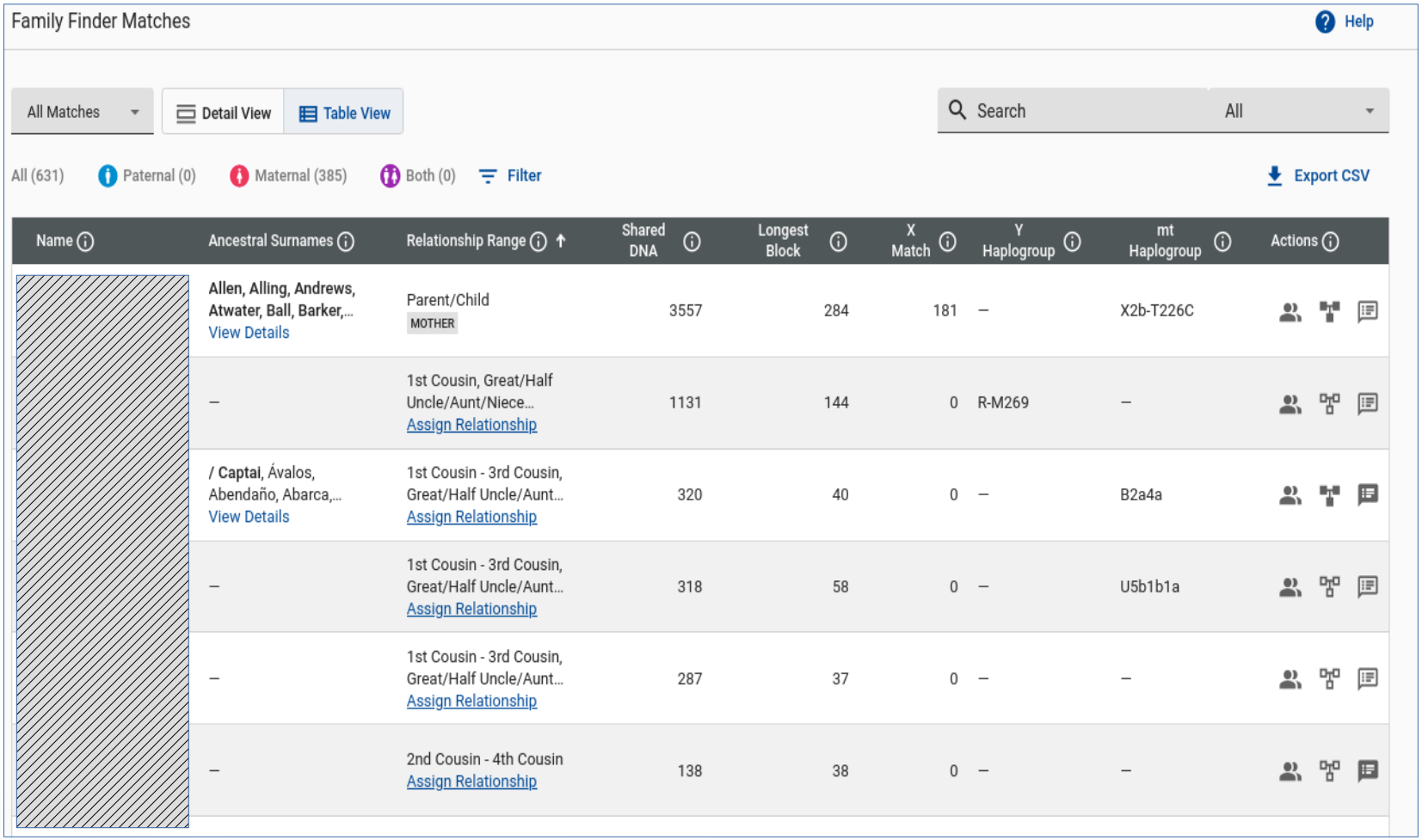
- www.Ancestry.com
- Features to look for . . .
  - Autosomal Genetic matches list
  - Common matches
  - Chromosome browser
  - Note the power of an analysis that combines DNA matching and a family tree!
  - Origins (admixture) estimate
  - Y matches
  - X matches
  - Joining groups

Admixture Estimates = Confusing!??

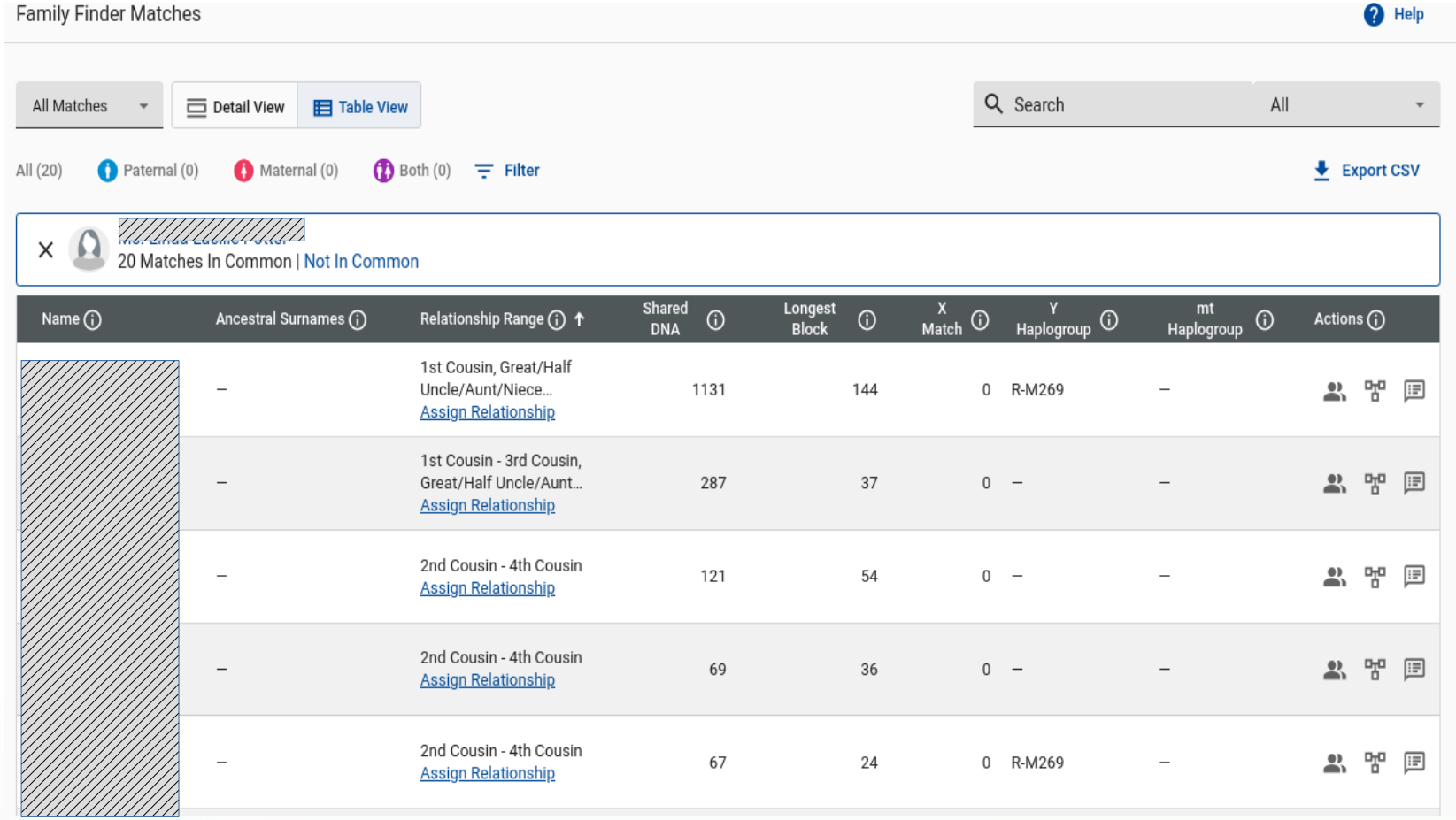


- Marketing and sales expectations
- Looking back past where your tree probably reaches
- Biology is more complicated than locations in your tree
- Your mixture will differ from that of your sibling's!
- Your ancestors may have been from areas where there was "high traffic"
- Calculations and sample groups vary according to testing company
- See https://youtu.be/Isa5c1p6aC0 and https://www.newshub.co.nz/home/opinion/2021/05/dna-ancestry-tests-are-they-telling-the-truth-that-s-questionable-genetic-anthropologist.html

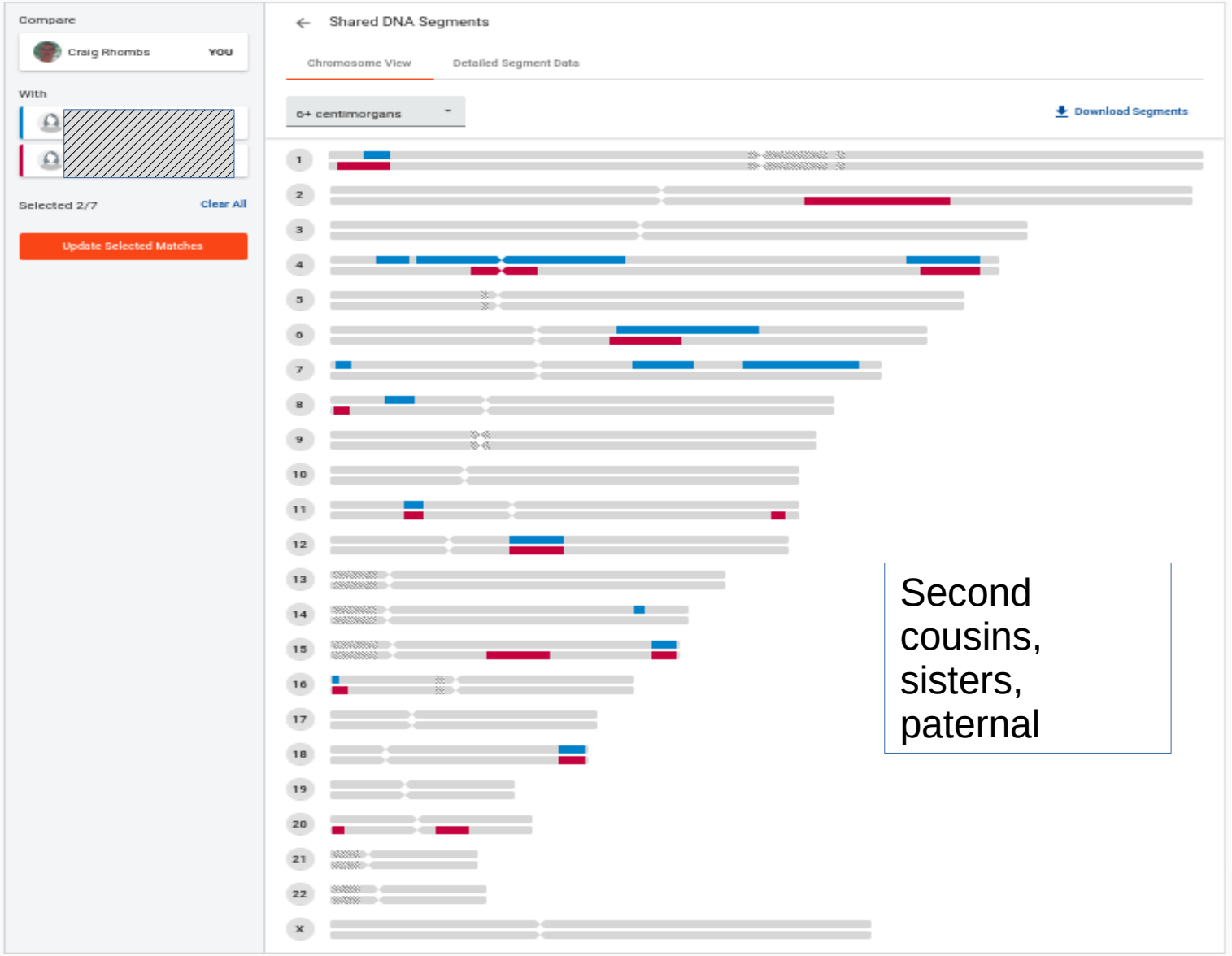
Family Tree DNA – Family Finder Matches



Family Tree DNA – common matches

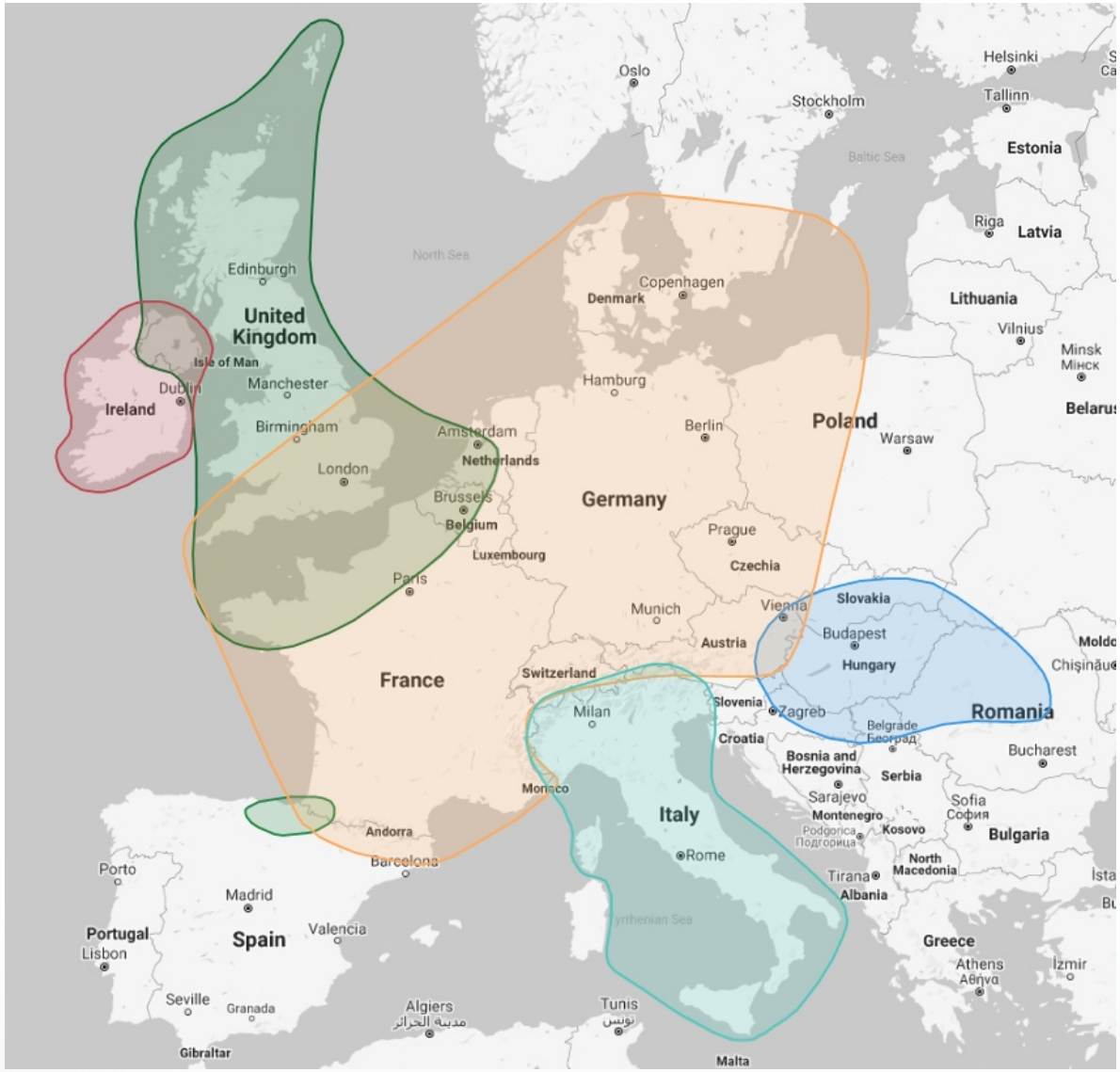


Family Tree DNA – Chromosome Browser



Second cousins, sisters, paternal

Family Tree DNA – Family Finder “origins”

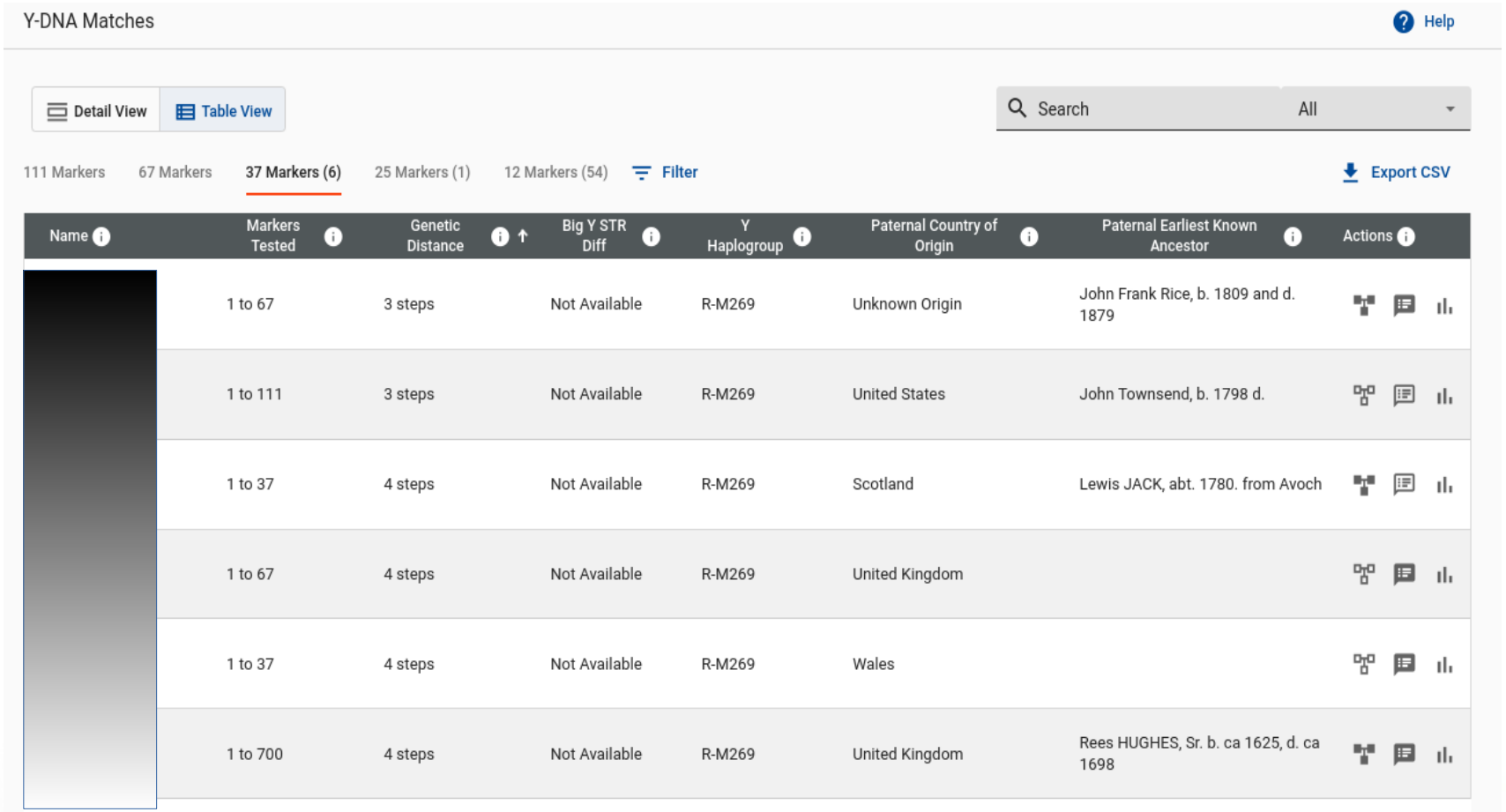


WORTHLESS!

Family Tree DNA – Family Finder “ancient origins”

Entertaining!

Family Tree DNA – Y DNA info.



Y-DNA Matches

Help

Detail View | Table View

Search | All

111 Markers | 67 Markers | 37 Markers (6) | 25 Markers (1) | 12 Markers (54) | Filter | Export CSV

Name	Markers Tested	Genetic Distance	Big Y STR Diff	Y Haplogroup	Paternal Country of Origin	Paternal Earliest Known Ancestor	Actions
[redacted]	1 to 67	3 steps	Not Available	R-M269	Unknown Origin	John Frank Rice, b. 1809 and d. 1879	
	1 to 111	3 steps	Not Available	R-M269	United States	John Townsend, b. 1798 d.	
	1 to 37	4 steps	Not Available	R-M269	Scotland	Lewis JACK, abt. 1780. from Avoch	
	1 to 67	4 steps	Not Available	R-M269	United Kingdom		
	1 to 37	4 steps	Not Available	R-M269	Wales		
	1 to 700	4 steps	Not Available	R-M269	United Kingdom	Rees HUGHES, Sr. b. ca 1625, d. ca 1698	

Family Tree DNA – mtDNA info.

mtDNA Matches

Help

FILTER MATCHES

Show Matches For: The Entire Database | Regions: HVR1, HVR2, Coding Regions | Matches Per Page: 25

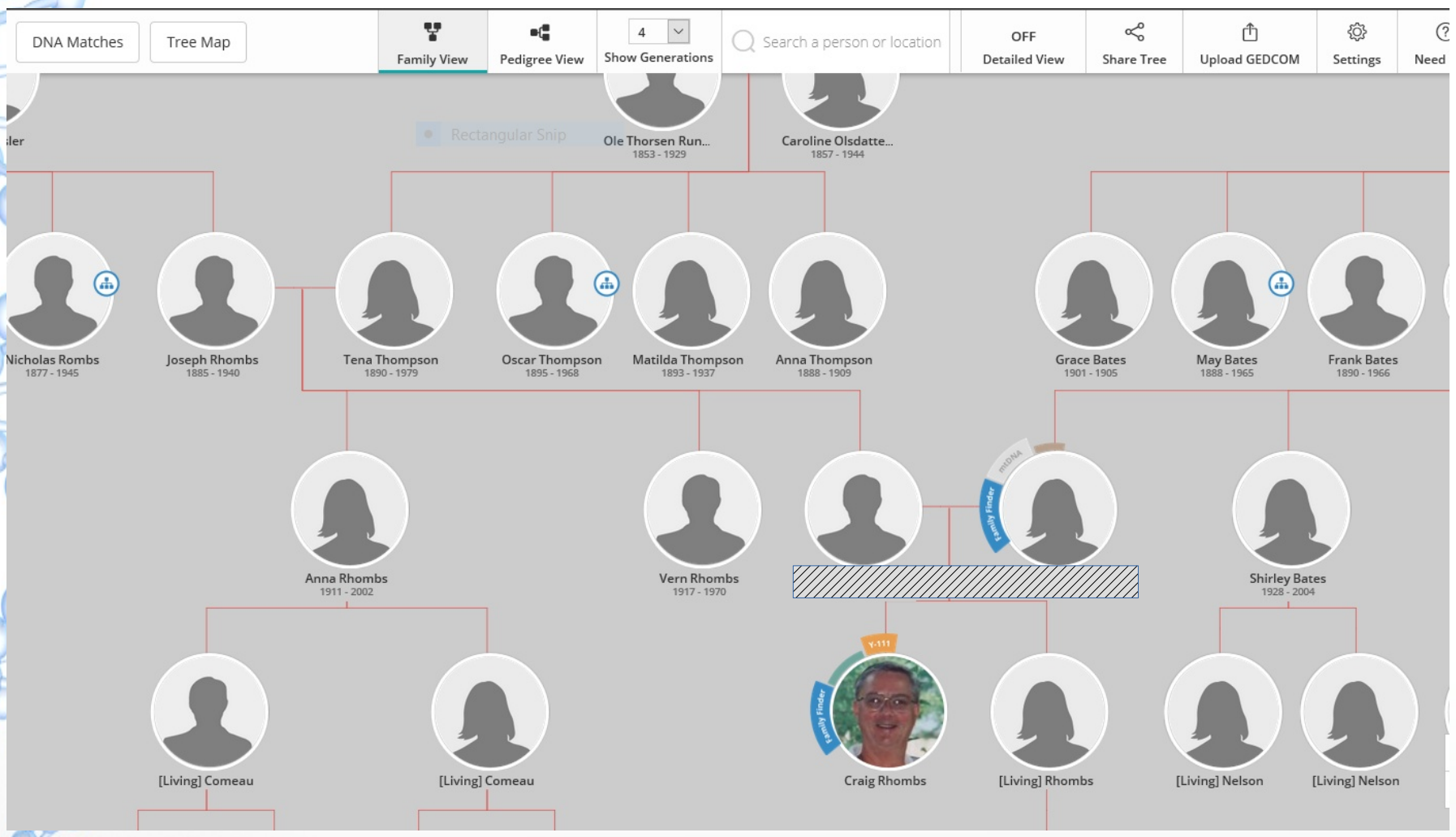
Last Name Starts With: | (Optional) New Since: | Run Report

HVR1, HVR2, CODING REGIONS - 197 MATCHES

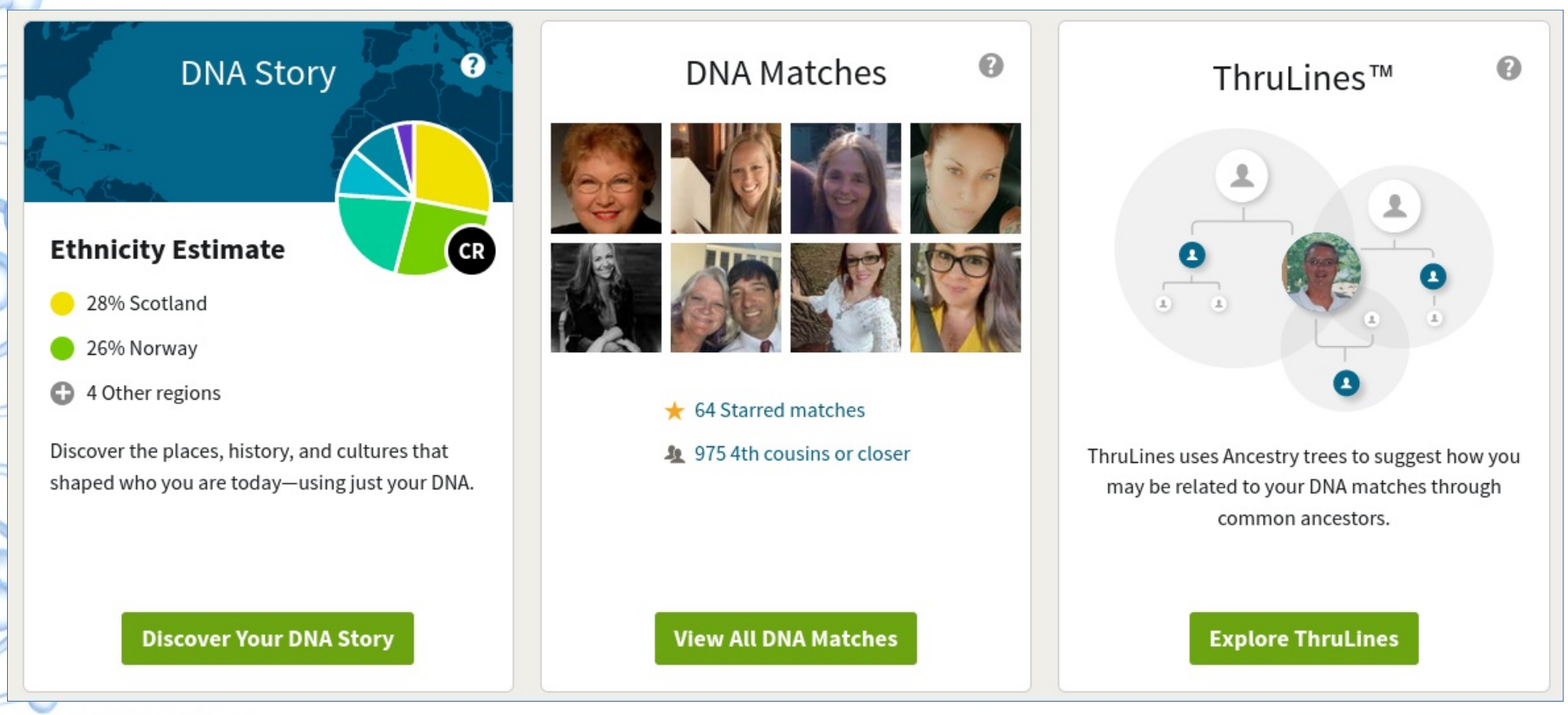
Page: 1 2 3 4 5 6 7 8 of

Genetic Distance	Name		Earliest Known Ancestor	mtDNA Haplogroup	Match Date
1	[redacted]	FMS		X2b-T226C	8/10/2020
1		FMS		X2b-T226C	12/13/2019
1		FMS		X2b-T226C	10/4/2019
1		FMS FF		X2b-T226C	4/25/2019
1		FMS FF	Mary Sanford, b. 1723	X2b-T226C	3/5/2019
1		FMS FF	Dorothy Cook	X2b-T226C	6/29/2018
1		FMS		X2b-T226C	5/11/2018
1		FMS FF		X2b-T226C	3/22/2018
1		FMS FF	Clarinda Clark	X2b-T226C	10/9/2017
1		FMS FF		X2b-T226C	5/22/2017
1		FMS	Bertha Mae Stewart b 1901 d 1979	X2b-T226C	5/12/2017
1		FMS FF	Catherine Wells Hall, b. 1801, d. July 5, 1872	X2b-T226C	12/13/2016

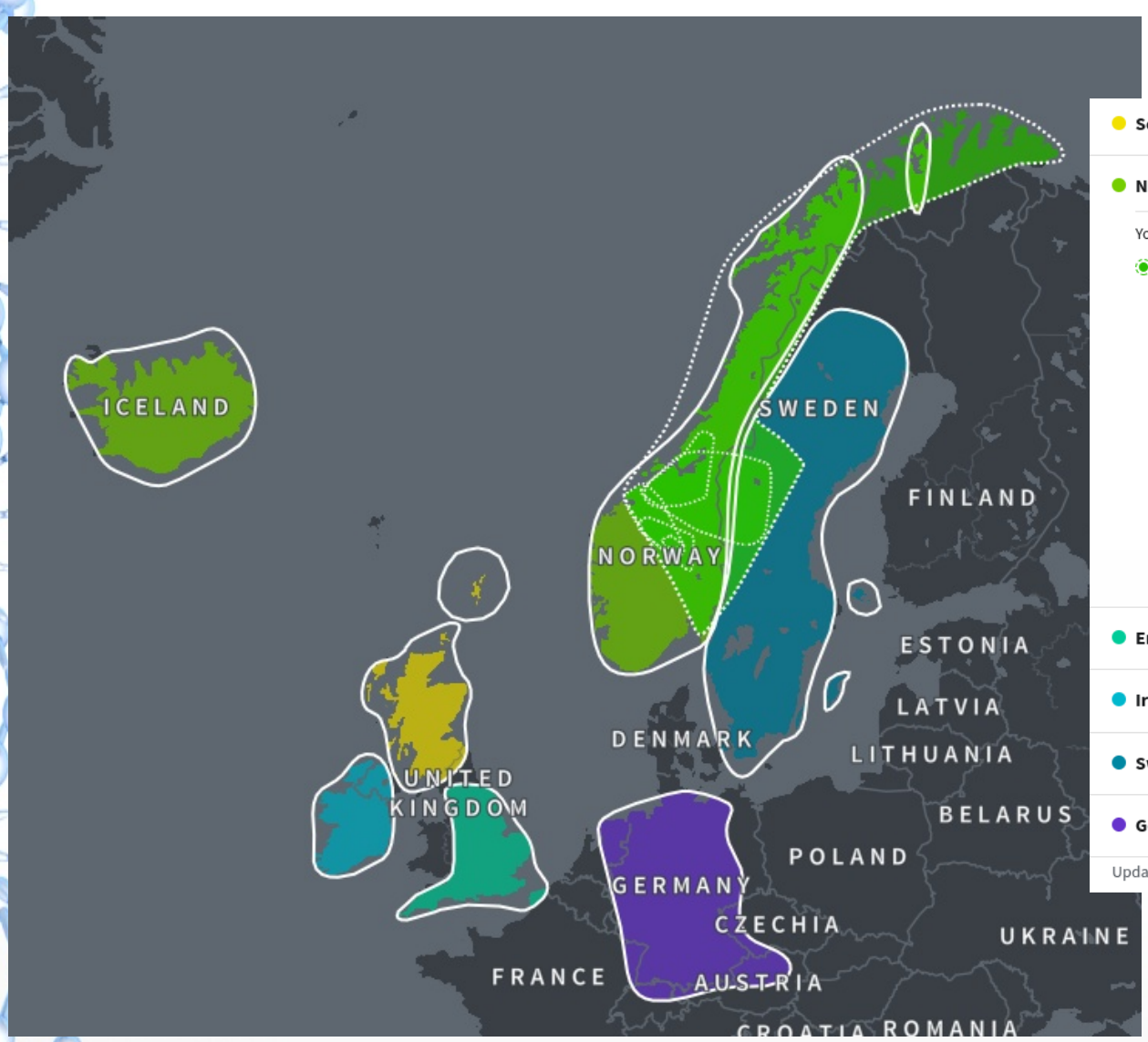
Family Tree DNA – Tree and test Integration



Ancestry.com Autosomal DNA



Ancestry.com – “ethnicity estimate”



Scotland	28%
Norway	26%

Your community with a connection to this ethnicity region

Southeastern, Central & Northern Norway

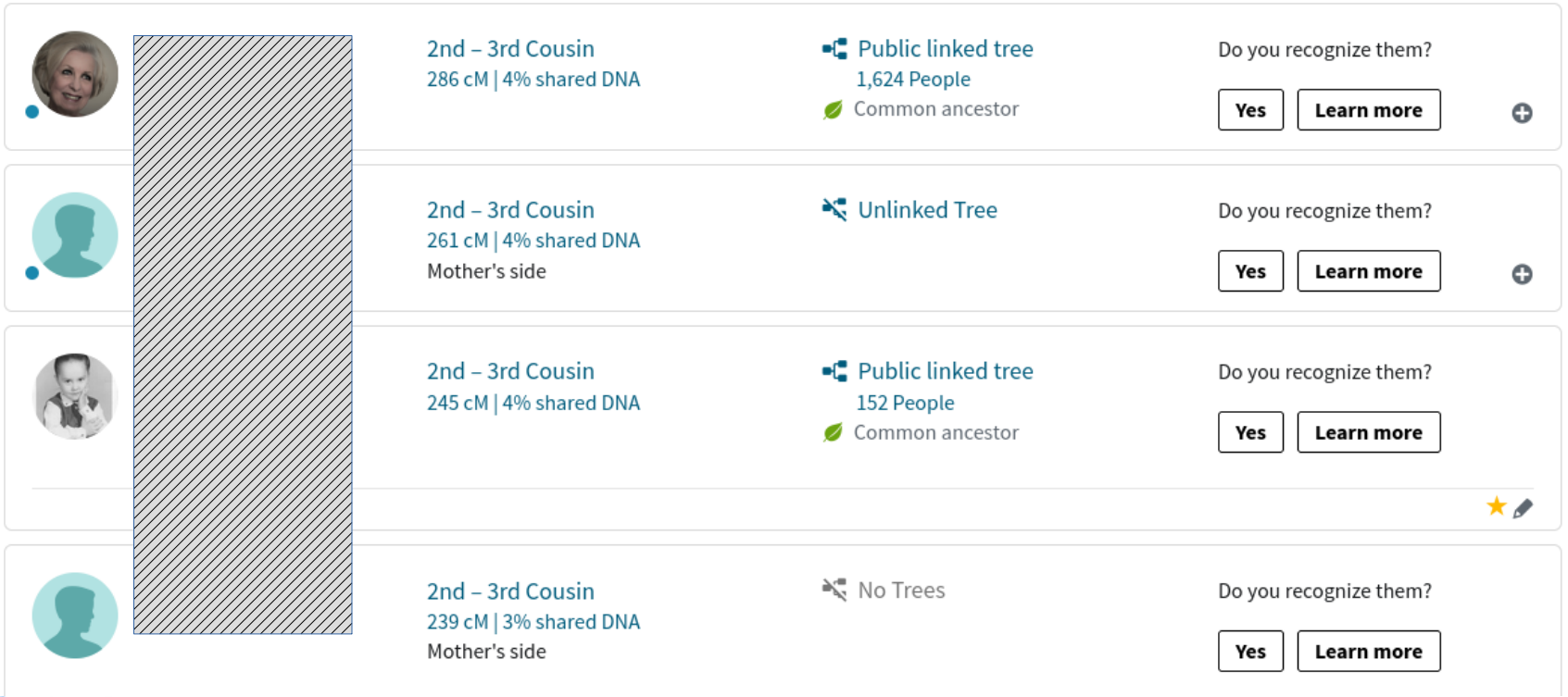
1 Possible Ancestor Story Found

- Northwestern Oppland
- Northern Møre og Romsdal, Oppland & Northern Hedmark
- Central Oppland
- Trøndelag, Northern Hedmark & Jämtlands län
- Central Norway

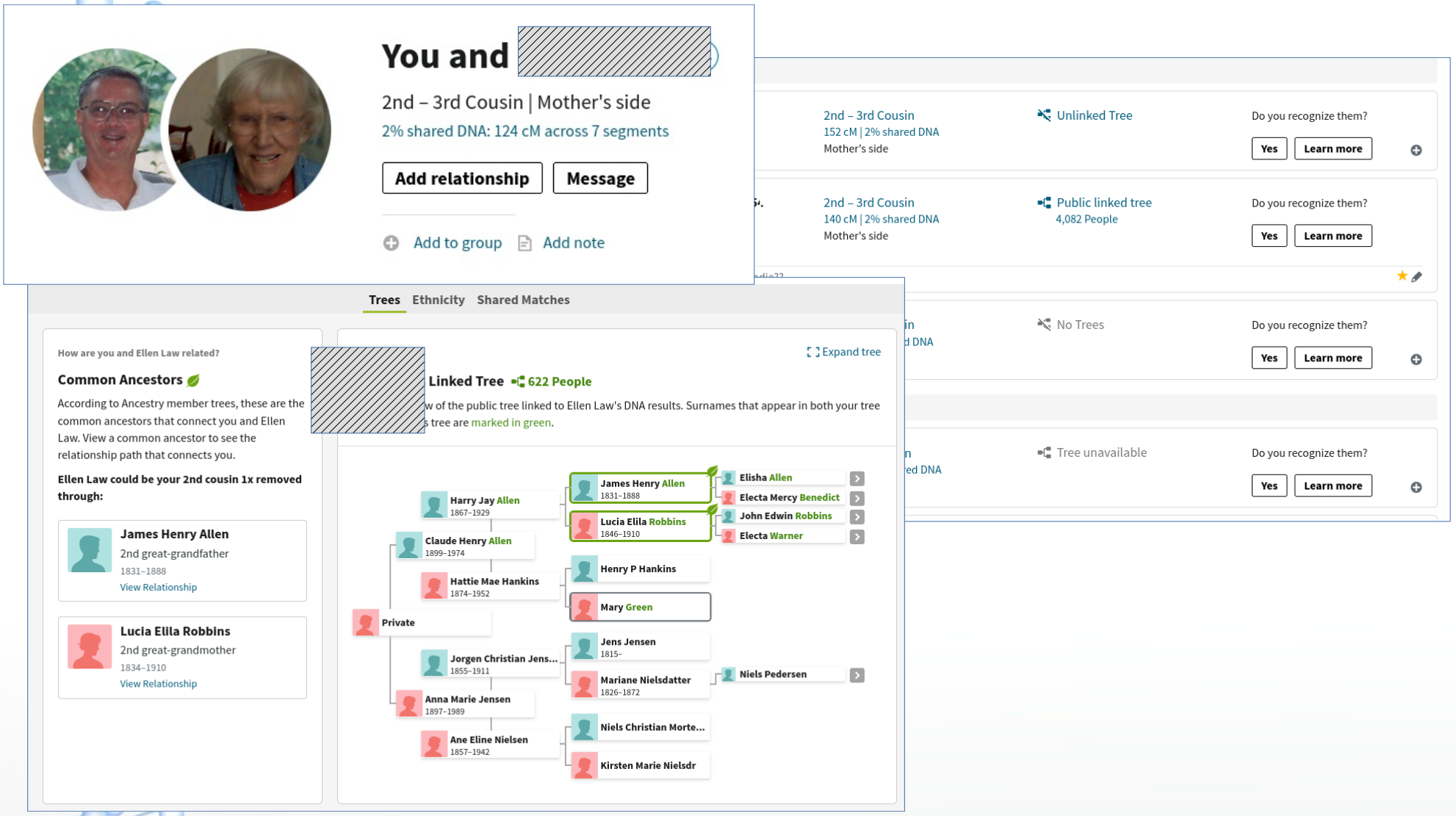
England & Northwestern Europe	22%
Ireland	10%
Sweden	10%
Germanic Europe	4%

Updated July 2020

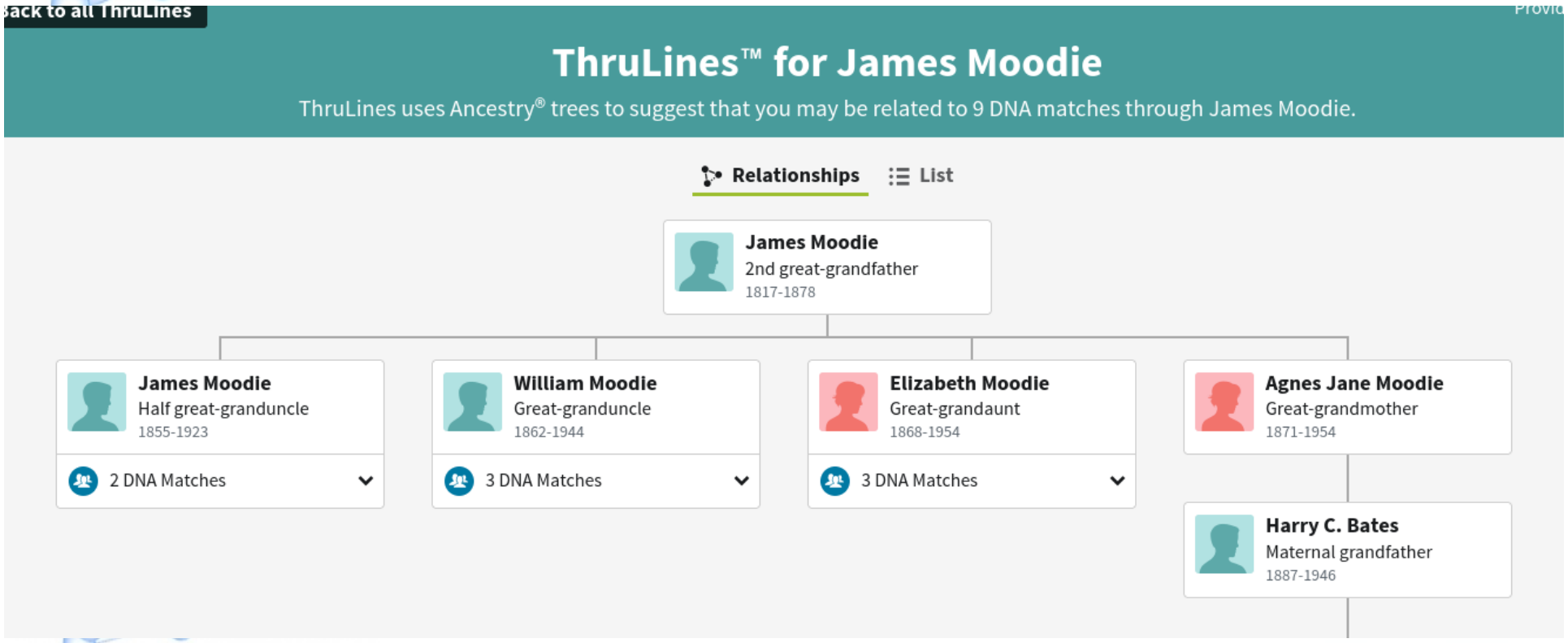
Ancestry.com – dna matches



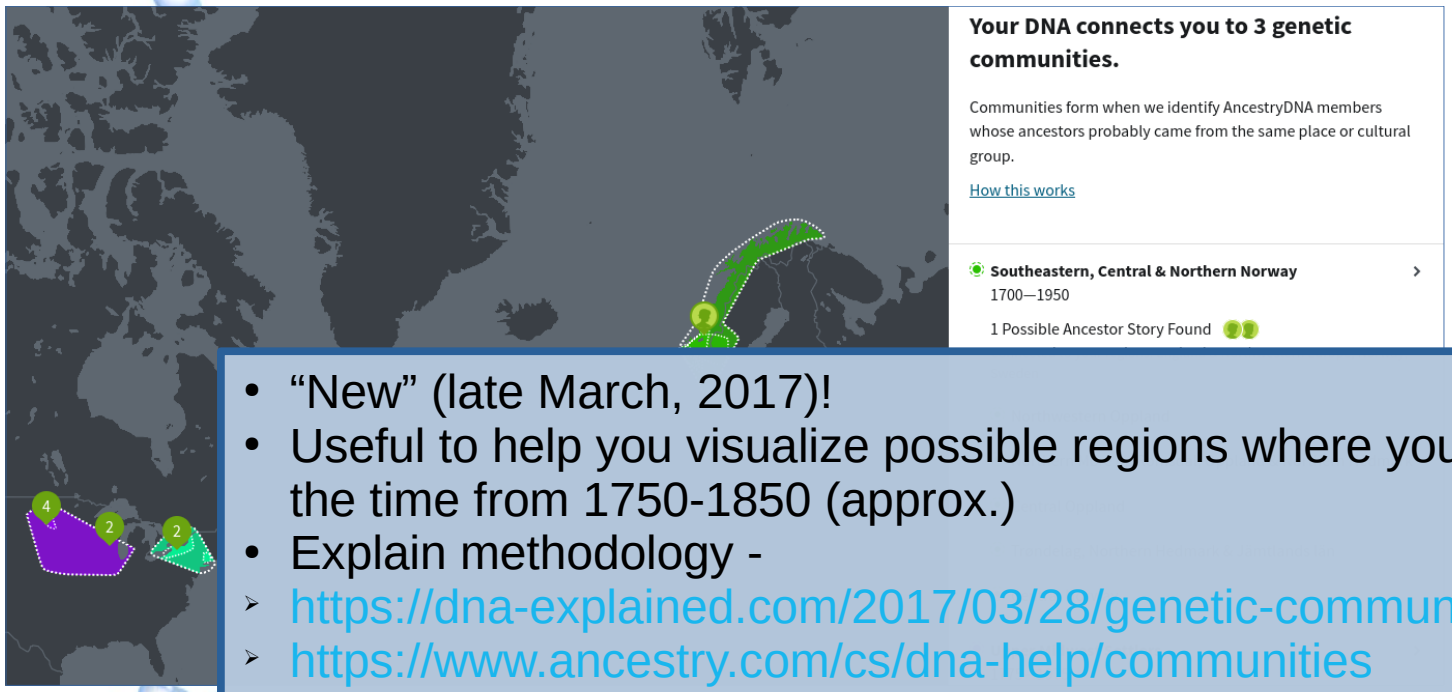
Ancestry.com – details



Ancestry.com – ThruLines



Ancestry.com – genetic communities



- "New" (late March, 2017)!
- Useful to help you visualize possible regions where your ancestors lived in the time from 1750-1850 (approx.)
- Explain methodology -
  - https://dna-explained.com/2017/03/28/genetic-communities/
  - https://www.ancestry.com/cs/dna-help/communities
  - https://www.ancestry.com/cs/dna-help/communities/whitepaper
- Formed from study of a couple million individuals with dna results and pedigree charts. Found clusters (about 300), created reference dna characteristics summary for each, then analyzed all ancestry members with dna tests to see what clusters they might be associated with.
- Uses: hints for passing through brick walls, confirmation of your research, etc. (use with caution and expect the analysis to change with time.)

Resources

- Local interest groups like . . .https://www.mngs.org/Interest-Groups
- ISOGG . . .https://isogg.org/wiki/Autosomal_DNA_testing_comparison_chart and

  https://isogg.org/wiki/Wiki_Welcome_Page
- If you get really “cranked up” on this kind of analysis, consider uploading raw data to GEDMATCH . . .https://www.gedmatch.com/login1.php
- DNA Painter (graphical tool & related video)

  https://dnapainter.com/

  https://youtu.be/wyjcJxywTZI
- Genome Mate Pro (compare match data from multiple sources)

  https://www.getgmp.com/
- Phasing; reconstruct genome of untested relative

  https://www.maps-phasing.com/

  https://www.borlandgenetics.com/

More Resources

- The Leeds manual analysis method for making sense of matches by forming clusters of related individuals.

  https://www.danaleeds.com/the-leeds-method/

- Take-offs on this type of analysis ($). Sometimes referred to as auto-clustering. Must upload matches (from myheritage, 23andme, or familytreedna).
  - https://www.myheritage.com/
  - https://www.dnagedcom.com/
  - https://geneticaffairs.com/
- CeCe Moore
  - https://thednadetectives.com/
- The McGuire Method

  https://thegeneticgenealogist.com/2017/03/19/guest-post-the-mcguire-method-simplified-visual-dna-comparisons/

References

- Really good overview from American Ancestors: https://www.americanancestors.org/education/learning-resources/read/dna-and-genealogy?_hsmi=122491487&_hsenc=p2ANqtz--mDBBopQ2yQ6n_sLBRA7wO2gxSB-GmVVItvVt6winLpMhq7esIXLGRUEfuU4HRP5lo1ixtz82YNIv7ewjDdXfQUxslRg
- Trace Your Roots with DNA, Smolenyak and Turner.
  - https://thegeneticgenealogist.com/
- The Family Tree Guide to DNA testing and Genetic Genealogy, Blaine T. Bettinger, c.2016, ISBN-13: 978-1440345326.
- The Invisible History of the Human Race, Christine Kenneally, c.2014, ISBN 978-0670-02555-8.
- A long story of a quest: https://maisonneuve.org/article/2021/04/12/genetic-mapping/?src=longreads&mc_cid=7aa42afeba&mc_eid=80b7342473&_hsmi=122491487&_hsenc=p2ANqtz-8yLR13w1wOuT6Zu6CR413oVAjh0wlVACui7WIFsmTS6l7Si-VWY25FSINQfJpjYhv3rKxR0fD2nMSD7FLLZlXiATYtUA
- Essentials of Genetics, Klug and Cummings
- The Lost Family - How DNA Testing Is Upending Who We Are, by Copeland, Libby (https://hclib.bibliocommons.com/v2/record/S109C5996780)

A Final Thought . . .

from Christine Kennealy in The Invisible History of the Human race . . .

"DNA tells us that we are creatures of chance and fate and that no one has quite the same mix of the two in his or her life. We think of ourselves as essentially whole, but when we look at our genome, we see that we are composed of many fragments stuck together. Many of our bits have different histories, and they each bring different probabilities into our lives.

…

Your genome is just the first hand that life deals you. How you play it is up to you."