

### How much DNA is Identical by Descent (IBD)?

%	Avg cM	Range	Relationships	Note
50%	3600	Near 3600	Parent-Child	
50%	3600	3250 - 3900	Full Sibling at 23&Me	Counts fully identical regions (FIR) on each homolog
37%	2650	2300 - 2950	Full Sibling at Ancestry, FamilyTreeDNA and Gedmatch	Counts FIRs as if HIRs (HIR: half-identical region)
25%	1800	1300-2300	Half Sibling Aunt/Uncle to Niece/Nephew Double 1C Grandparent to Grandchild	
12.5%	900	575-1330	1C Half Aunt/Uncle/Niece/Nephew Great-Grandparent to Gr-GrChild Great-Aunt/Uncle/Niece/Nephew	1C = first cousin
6.25%	450	215-650	1C1R, H1C, Half-Aunt/Uncle/Niece/Nephew	1R = once removed H = half
3.125%	224	75-360	2C, 1C2R, H1C1R	
1.56%	112	30-215	2C1R, H2C, 1C3R, H1C2R	1 in 1000 won't match
0.78%	56	0-109	3C, 2C2R, H2C1R	10% of 3Cs won't match
0.4%	30	0-75	3C1R, H2C2R	25% of 3C1Rs won't match
0.2	13		4C	50% of 4Cs won't match

Adapted from Christa Stalcup, DNA Detectives Autosomal Statistics Chart, aka "The Green Chart." Sources - see also: Blaine Bettinger, <http://thegeneticgenealogist.com/2016/06/26/update-to-the-shared-cm-project/> and ISOGG, [http://isogg.org/wiki/Autosomal\\_DNA\\_statistics](http://isogg.org/wiki/Autosomal_DNA_statistics).

### Method for finding Bio-Parents after Test Results are in

1. Working from the best matches first (but ignoring descendants of the testee), use the Shared Matches tool to sort matches with trees into Genetic Kinship Networks, connecting each testee to the Common Ancestor (CA) or Common Ancestral Couple shared with another testee. Complete the immediate genealogies around each of the CA's to find connections to more testees. Pay attention as to whether the CA's had children by multiple spouses because this will significantly affect the amount of shared DNA. The adoptee descends from the CAs of the testees or from the recent ancestors of the CAs as indicated by the total shared centimorgans.
2. You will need to find least two Kinship Networks, one for each bio-parent. Often you will form more than two networks. In the case of say three networks, there should be a marriage or union connecting two of the people within two of the networks to collapse them into one, and this is where the adoptee fits in.
3. Determine the best fit branch for the bio-parent within the network as was shown during the lecture.
4. Often the best fit result will lead to an ambiguous result, such as, "bio-parent must be a one of these 5 siblings." Try to determine who was at the right place at the right time, but know that this is not necessarily conclusive.
5. When you must resort to matches without trees, try first contacting the match. Ask an easy to answer question such as "could you tell me your grandparents." Use public trees and censuses to find the grandparents relatives and work them into the network. Matches with full names can often be found in obits as survivors. Companies like Spokeo, BeenVerified and Pipl, if searched on an uncommon name and probable location, will often offer for free birthdates, and lists of relatives of the person. I try hard to identify all matches at or above the 200 cM level unless the case is solved without doing so.
6. If results remain inconclusive but narrowed down to a single set of siblings (or a single set of first cousins), see if any are willing to test so that the relationship can be further narrowed down.

## Testing “Recipe” for Adoptees seeking Bio-Parents

This is typically the most cost-effective approach.

1. Test first at Ancestry. Their testing pool is the largest and you cannot transfer in from elsewhere.
2. Transfer your raw data into FamilyTreeDNA, MyHeritage and GEDmatch. Cost is minimal or free.
3. If you are seeking bio-parents for a deceased parent, test yourself. If your other parent is living, have them test too, so that matches on the known side can easily be filtered out of consideration.
4. Test at 23andMe if you still cannot solve the case.

## DNA Test Companies

- Ancestry DNA: <http://dna.ancestry.com> largest database, no chromosome browser.
- My Heritage DNA: [www.myheritage.com/dna](http://www.myheritage.com/dna) accepts transfers, growing database, offers a chromosome browser, and shows *how much* matches match one another.
- Family Tree DNA: <http://www.familytreedna.com> medium database, good tools, Y & mito tests, Y Surname Projects.
- 23 and Me: <https://www.23andme.com> medium database, good tools & science
- Living DNA: <https://www.livingdna.com/> newer, British focus. Database is currently too small to recommend this site to US adoptees.
- GEDmatch (not a testing company, but accepts data uploads and offers the best toolkit)  
<https://www.gedmatch.com/>

## DNA Testing Companies Compared

more detail at [http://isogg.org/wiki/Autosomal\\_DNA\\_testing\\_comparison\\_chart](http://isogg.org/wiki/Autosomal_DNA_testing_comparison_chart)

Feature	Ancestry	Family Tree DNA	23 and Me	My Heritage
Medical Info	No	No	Yes (extra\$)	No
<b>Chromosome Browser</b>	<b>No</b>	Yes	Sharing testees only	Yes
<b>In Common With tool</b>	Yes	Yes	Sharing testees only	Yes
<b>How Matches Match</b>	No	No	Yes	Yes
List of Matches	Yes	Yes	Yes	Yes
Share by Default	Yes	Yes	New testees only	Yes
<b>Sharing Database Size*</b>	> 12 M	23% of Ancestry size	35% of Ancestry size	21% of Ancestry size
Correspondence	Thru site	Direct email	Thru site	Thru Site
Match Responsiveness	Poor	Better	Poor	Better
Hosts Family Trees	Yes	Yes	Yes**	Yes
Y & mt haplogroups	No	Available, extra \$\$	Yes	No
Y & mt genealogical relevant tests, Y surname projects	No	Available, extra \$\$	No	No
Accepts Transfers-In	No	Yes	No	Yes

GEDMatch database size is over 1.2 M, I estimate about 10% of the size of Ancestry.

All autosomal tests run about \$59-\$69 when on sale. FTDNA's mt and Y tests range from about \$150-450.

\* My 2017 **rough** estimate, based off info on 35 kits reported by 28 people on CAGGNI's Facebook page. My main point here is that the size of the database by itself is not interesting; it is the size of the *sharing* database that matters.

\*\*23andMe is the only company that does not allow uploading of Gedcoms. It builds empty trees connecting you to your matches in the most probable way and then allows you to fill in the connecting ancestors and make corrections to the tree as necessary.

**To Learn More**

- International Society of Genetic Genealogy - ISOGG: <http://www.isogg.org>. The ISOGG Wiki [http://www.isogg.org/wiki/Wiki\\_Welcome\\_Page](http://www.isogg.org/wiki/Wiki_Welcome_Page). ISOGG also has an active Facebook page.
- Join CeCe Moore's Facebook Group *DNA Detectives* (CeCe Moore, lead admin). Other excellent genetic genealogy Facebook groups include Blaine Bettinger's *Genetic Genealogy Tips and Tricks* and the ISOGG group.

**3rd Party Tools**

- DNAadoption: <http://www.dnaadoption.com/> the team here pioneered a formalized method for finding bio-parents and at one time offered a 20 hour class to learn the method.
- GedMatch: <https://www.gedmatch.com/> has the best toolkit but a smaller matching database.
- DNAGedCom: <http://dnagedcom.com> the ADSA tool, JWorks and KWorks help identify all matches on a segment
- Genetic Affairs <https://www.geneticaffairs.com/> Clustering Tool & updates on your matches from many of the companies. MyHeritage now has clustering built in; Ancestry has recently prevented use of Genetic Affairs.
- DNA Painter <https://dnapainter.com/> color in your chromosomal segments from multiple companies. Also popular here: the Shared cM and What Are The Odds (WATO) tools.

**Ethics**

- <https://petrieflom.law.harvard.edu/events/details/2019-petrie-flom-center-annual-conference> see esp. Kif Augustine-Adams and Natalie Ram
- Humor, but so true! <https://www.mcsweeney.net/articles/jon-snow-emails-23andme-about-his-dna-results>

**Autosomal DNA-Related Terms of relevance to unknown parentage searches**

CA	Common Ancestor. The person or persons (CAs), often an ancestral couple, from whom two or more matches descend.
cM, total shared cM	The unit of recombinant frequency used to measure DNA segments. Total shared cM is used to estimate genealogical relationships. Long matching segments are indicative of shared ancestry within the genealogical timeframe. See table in this handout.
IBD	Identical by Descent. Matching segments 10 cM or longer are 99% likely to be IBD.
IBS	Identical by State. Matching segments of 7 cM are 70% likely to be IBS; the match is characteristic of the population being sampled and so well outside the genealogical timeframe or is a false match due to a phasing error. See table in this handout.
ICW, Shared Matches, Relatives in Common	In Common With. When two of your matches are in each other's match lists, they are ICW. If they are ICW and match you at the same chromosomal location for 9+ cM, all three of you almost certainly share a recent genealogical ancestor. Ancestry uses the term "Shared Matches", 23 and Me uses the term "Relatives in Common."
NPE, misattribution	Non-paternity event or informally "not the parent expected." Our genetics do not always match our expectations...
Segment	A long series of matching SNPs (single nucleotide polymorphisms) found when comparing the raw DNA. A match must have sufficiently long segment(s) to qualify as a match.
Targeted Testing	Testing a suspected closer relation to further narrow down or confirm parent candidates.
Triangulation	The process of using three or more ICW matches, together with a 1:1 verified shared segments and genealogical paper trails, to locate the ancestor that contributed a particular segment of DNA. See Blaine Bettinger's "Triangulation Intervention" blog entry for details. Note that the triangulation process is not required when working unknown parentage cases – you and two bona fide third cousins may all share different segments from the Common Ancestors.

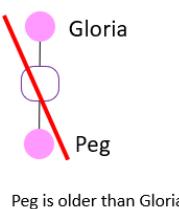
## Networks shown in the lecture

## 1. How are Peg and Gloria Related?

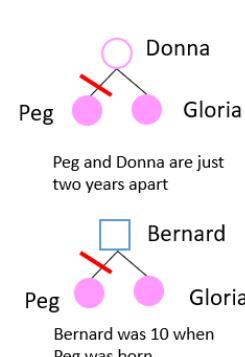
Peg and Gloria Share 1698 cM.  
Which relationship is it?

- Grandparent / Grandchild
- Half-Siblings
- Aunt / Niece

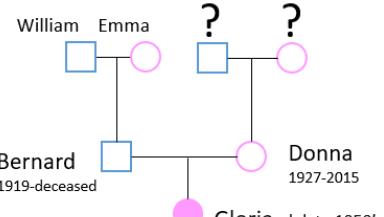
## Grandparent / Grandchild



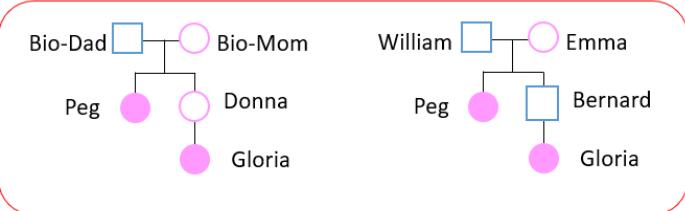
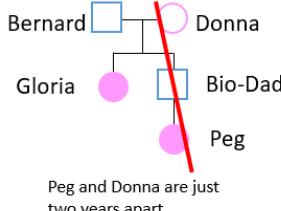
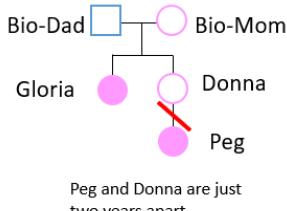
## Half Sibs



## Gloria's Tree

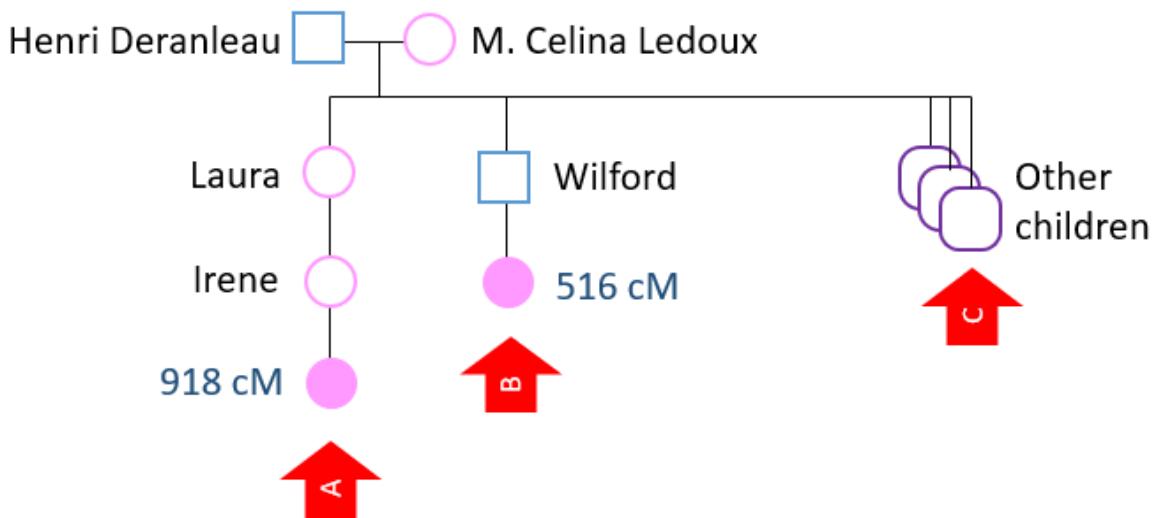


## Aunt / Niece



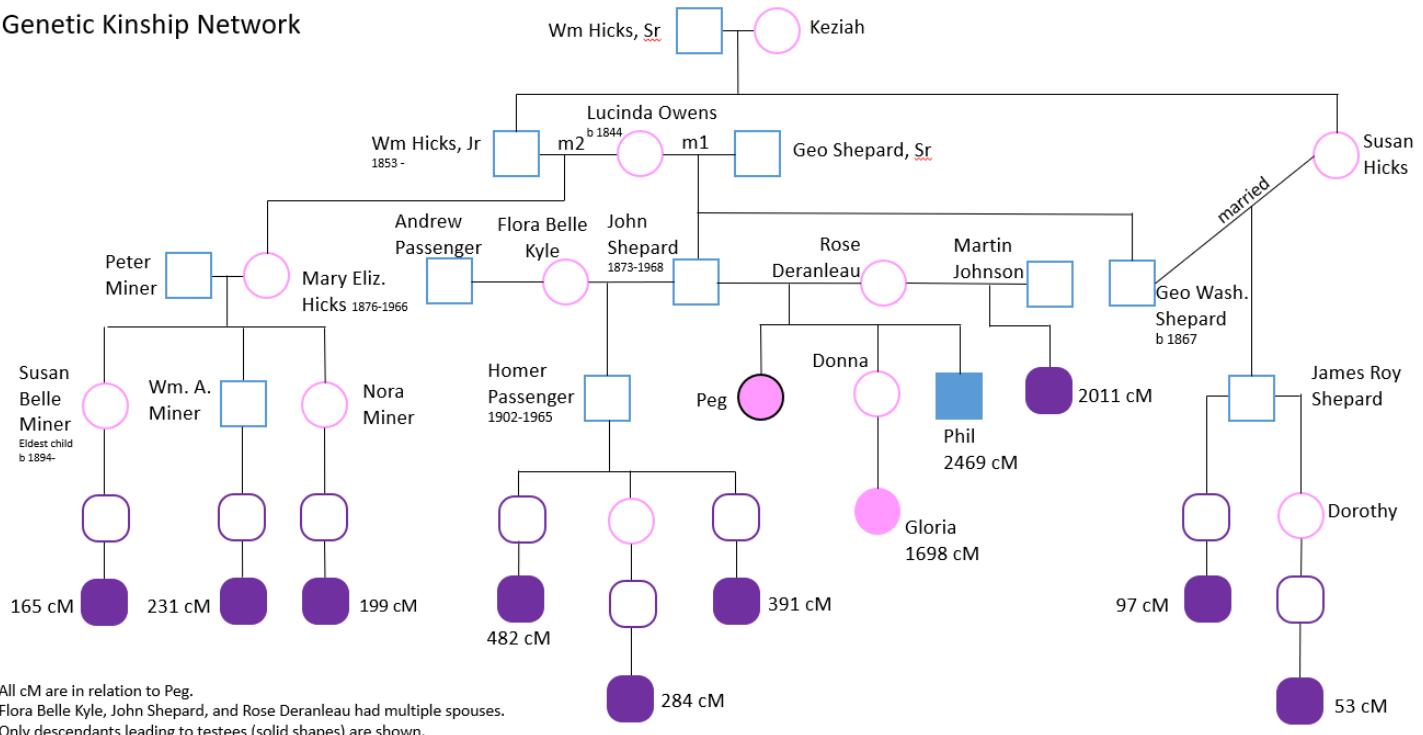
Showing how ages alone rule out all but two possible relationships. Shared matches (not shown here) show the connection is thru Donna.

## 2. Which branch does the bio-parent fit into? (Must be within the "A" branch; discovered thru records that bio-mother was Rose Deranleau, a sister to Irene).



## 3. Final network revealing the Bio-Parents for Peg and Donna

Genetic Kinship Network



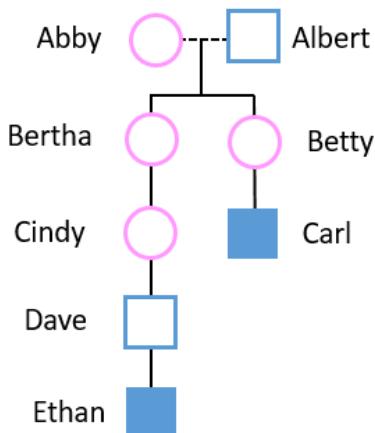
## Answer to Exercise on Page 6:

Carl and Ethan are 1C2R, sharing 1/32 or 213 cM (formula:  $2 \times (1/2^6) = 2 \times 1/64$ ).

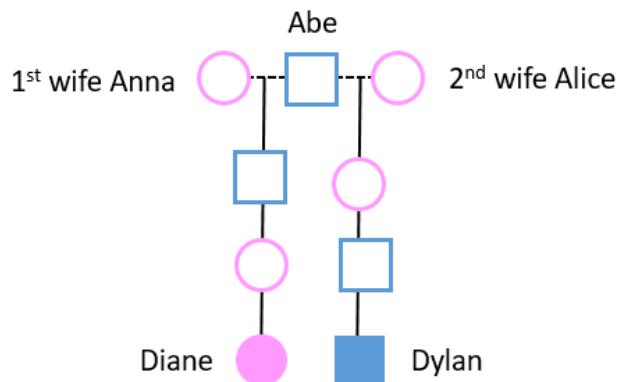
Diane and Dylan are H2C, sharing 1/64 or 1.56% (formula  $1/2^6$ )

Try your hand at calculating Average Shared DNA (average amount Identical by Descent, "IBD"):

**1. Carl and Ethan**



**2. Diane and Dylan**



What is the relationship between Carl and Ethan \_\_\_\_\_

between Diane and Dylan \_\_\_\_\_

How much DNA is shared by descent, on average\* \_\_\_\_\_

\_\_\_\_\_

Answers on Page 5.

**Method** – for each step apart thru a *single* CA multiply by  $\frac{1}{2}$ , for example, five steps apart is  $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = 1/32$ . Repeat for *each* common ancestor and add the results.

Steps Apart	DNA shared	% shared	cM (x72)
1	1/2	50.00%	3600
2	1/4	25.00%	1800
3	1/8	12.50%	900
4	1/16	6.25%	450
5	1/32	3.13%	225
6	1/64	1.56%	112
7	1/128	0.78%	56
8	1/256	0.39%	28
9	1/512	0.20%	14

\* In practice there will be a range of possible cM. See the table on page 1.