LAGOS FORENSIC SYMPOSIUM

# The Use Of The Combined DNA Index System (CODIS) In Human Identification and Criminal Investigations.

Michael Coble, PhD

Associate Professor and Executive Director, Center for Human Identification, University of North Texas Health Science Center

**6th Lagos Forensic Symposium (Hybrid)**
**October 19-20, 2022**

# Forensic DNA testing – Criminal Investigations

- Direct DNA Comparison – The Questioned sample is compared to a Known reference sample


Q

K

# An Example

blood at
crime scene

DNA
testing

Marker XYZ
8, 11

8, 11

Suspect matches
the evidence

How rare
is this match?

---

| Allele | Freq |
|---|---|
| 6 | 0.20 |
| 7 | 0.10 |
| 8 | 0.40 |
| 9 | 0.10 |
| 10 | 0.10 |
| 11 | 0.10 |

Use “2pq”

Freq 8 = 0.40
Freq 11 = 0.10

2(0.40)(0.10) = **0.08**

We would expect about 8% of the population (or about 1 in 12.5 people) to match the evidence.

# Forensic DNA testing – Missing Persons Investigations

- Direct DNA Comparison – The Questioned sample is compared to a Known reference sample

Bone sample

DNA testing →

Marker XYZ
8, 11

8, 11

Known reference matches the evidence

Q

K

# Difficulties encountered

- In Criminal investigations – there may not be a suspect (e.g. Sexual Assault).

- Criminal casework is often presented as complex mixtures of multiple individuals – challenging to make a Q to K comparison.

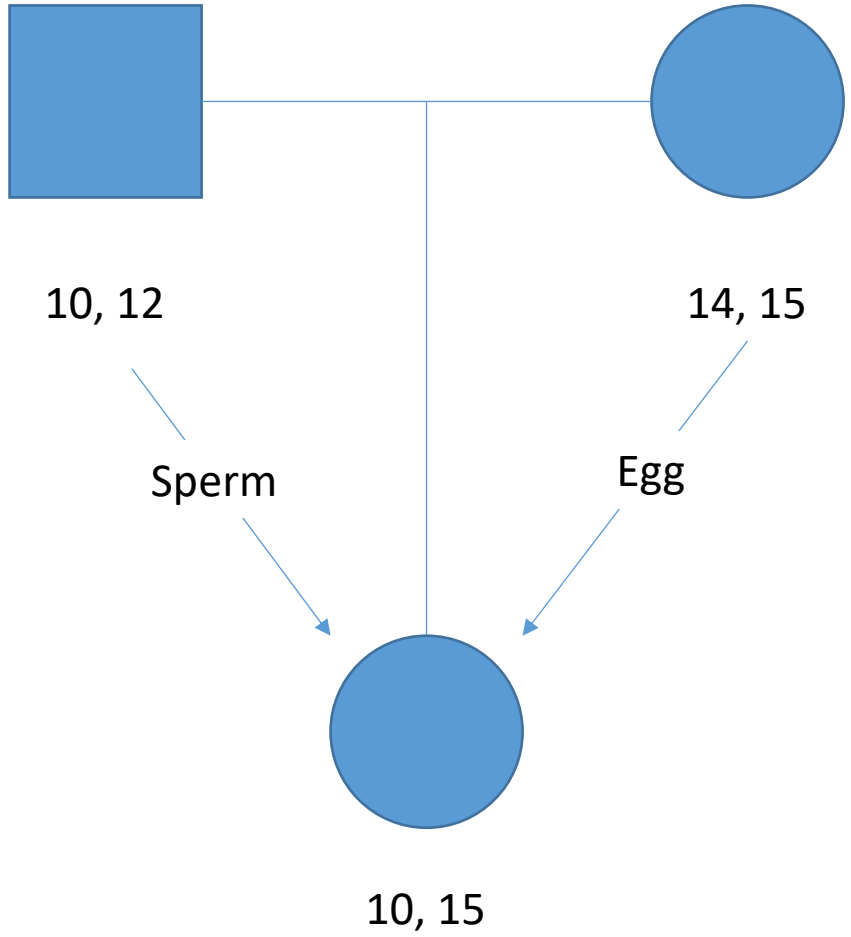
- In Missing Persons investigations, there may be a lack of an item of clothing or biological material that belonged to the missing individual.

# Forensic DNA testing – Criminal and Missing Persons Investigations

- Indirect DNA Comparisons – use of close family members are used to infer the relationship to the unidentified remains or evidence from the crime scene stain.

- Requires specialized knowledge of how to perform statistical kinship calculations.

# DNA Inheritance



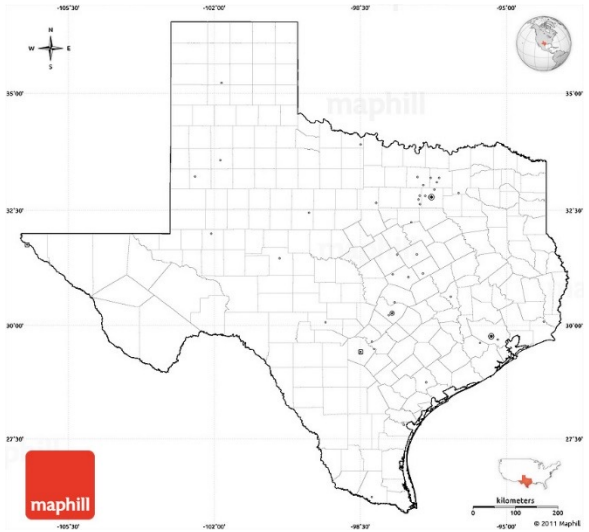

One allele (repeat) is obligate from each parent

# Criminal Investigations

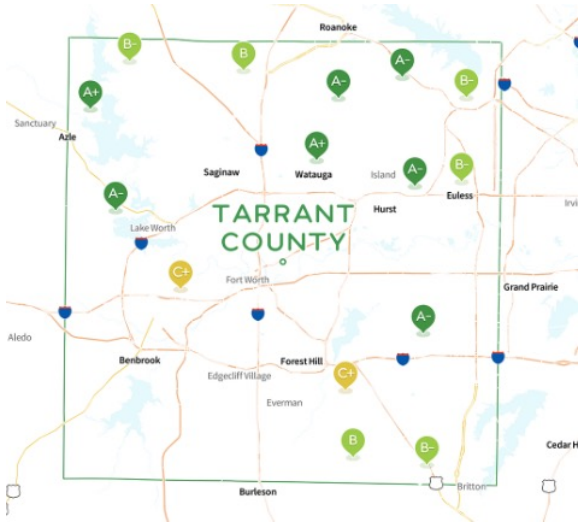
# CODIS – The Combined DNA Index System

- CODIS is a multilayered DNA database with several “indicies”


National level (NDIS)




State level (SDIS)




Local level (LDIS)

# CODIS Architecture

Convicted Offender or Arrestee Index – single source profiles entered into the state database given a qualifying event.

Forensic Index – profiles developed from forensic casework (evidence from a crime scene) – where no suspect has been determined.

Missing Person Index – profiles of skeletal remains and reference family profiles.

# CODIS Statistics

- The National DNA Index (NDIS) contains
- 14,836,490 offender profiles
- 4,513,955 arrestee profiles
- 1,144,255 forensic profiles.
- As of October 2021, CODIS has produced over 587,773 hits assisting in more than 574,343 investigations.

as of October 2021

# Searching in CODIS – 3 levels

High Stringency Searching

| Evidence Profile | CO Profile |
|---|---|
| 14, 20 | 14, 20 |
| 15, 17 | 15, 17 |
| 11, 11 | 11, 11 |

| Locus | Target<br>Unidentified Remains | Candidate<br>Convicted Offender |
|---|---|---|
| D3S1358 | 17 ,18 | [H] 17 ,18 |
| vWA | 14 | [H] 14 |
| FGA | 21 ,23 | [H] 21 ,23 |
| Amelogenin | X ,Y | _X ,Y |
| D8S1179 | 12 ,13 | [H] 12 ,13 |
| D21S11 | 29 | [H] 29 |
| D18S51 | 12 ,13 | [H] 12 ,13 |
| D5S818 | 11 | [H] 11 |
| D13S317 | 8 ,10 | [H] 8 ,10 |
| D7S820 | 10 ,11 | [H] 10 ,11 |
| D16S539 | 11 ,12 | [H] 11 ,12 |
| TH01 | 6 ,9.3 | [H] 6 ,9.3 |
| TPOX | 8 ,11 | [H] 8 ,11 |
| CSF1PO | 10 | [H] 10 |
| D2S1338 | 17 ,23 | [H] 17 ,23 |
| D19S433 | 13 ,14 | [H] 13 ,14 |

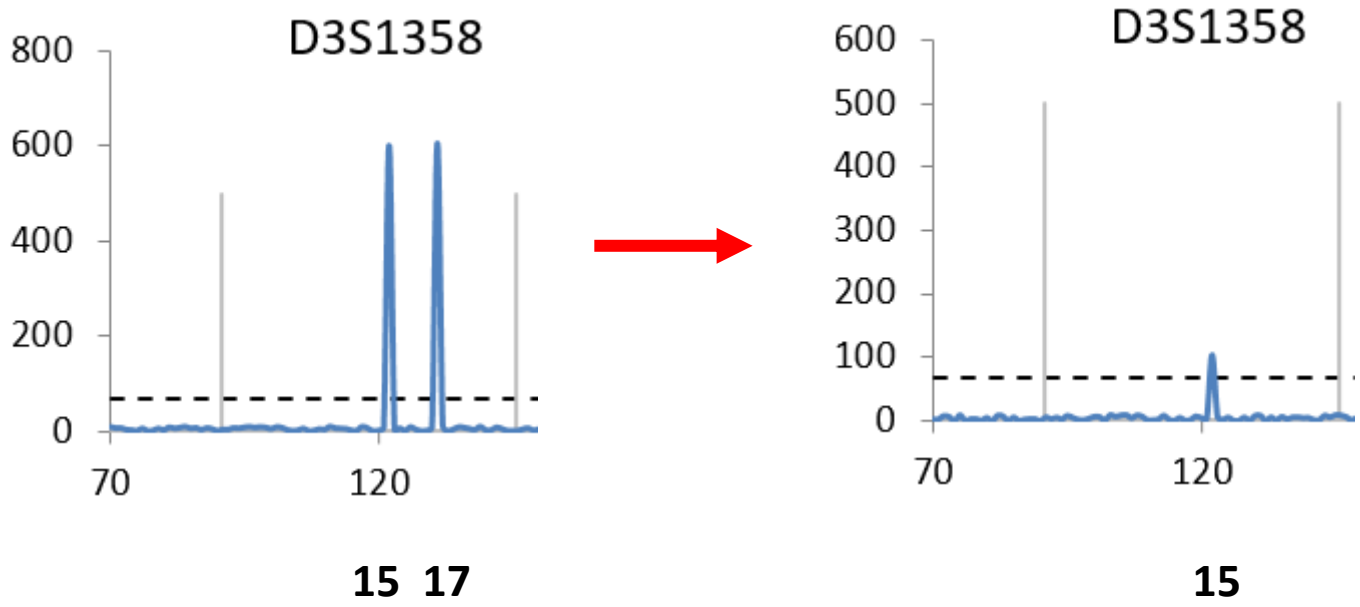
# Searching in CODIS – 3 levels

High Stringency Searching

| Evidence Profile | CO Profile |
|---|---|
| 14, 20 | 14, 20 |
| 15, 17 | 15, 17 |
| 11, 11 | 11, 11 |

| Evidence Profile | CO Profile |
|---|---|
| 14, 20 | 14, 20 |
| **15, 15** | 15, 17 |
| 11, 11 | 11, 11 |

Suppose – partial profile (drop-out) in evidence



# Searching in CODIS – 3 levels

Moderate Stringency Searching

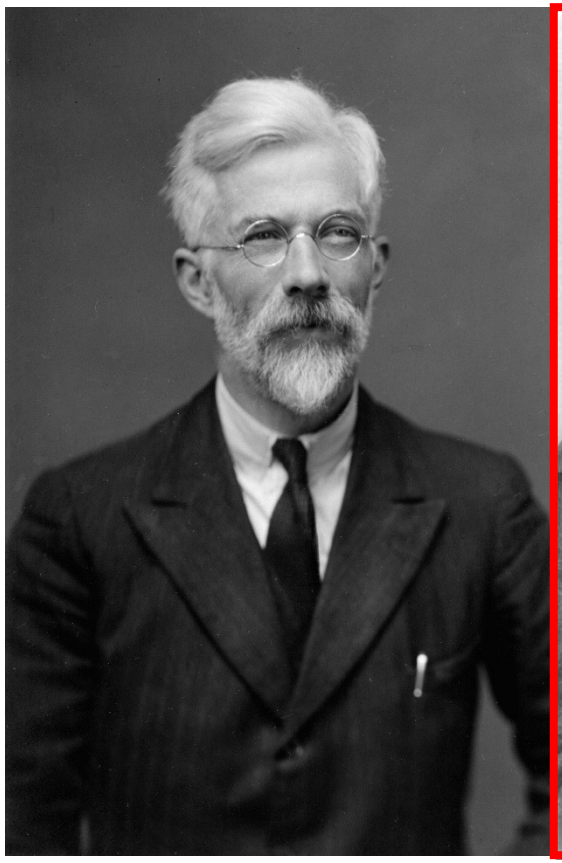
| Evidence Profile | CO Profile |
|---|---|
| 14, 20 | 14, 20 or 14, 14 or 20, 20 |
| 15, 17 | 15, 17 or 15, 15 or 17, 17 |
| 11, 11 | 11, X |

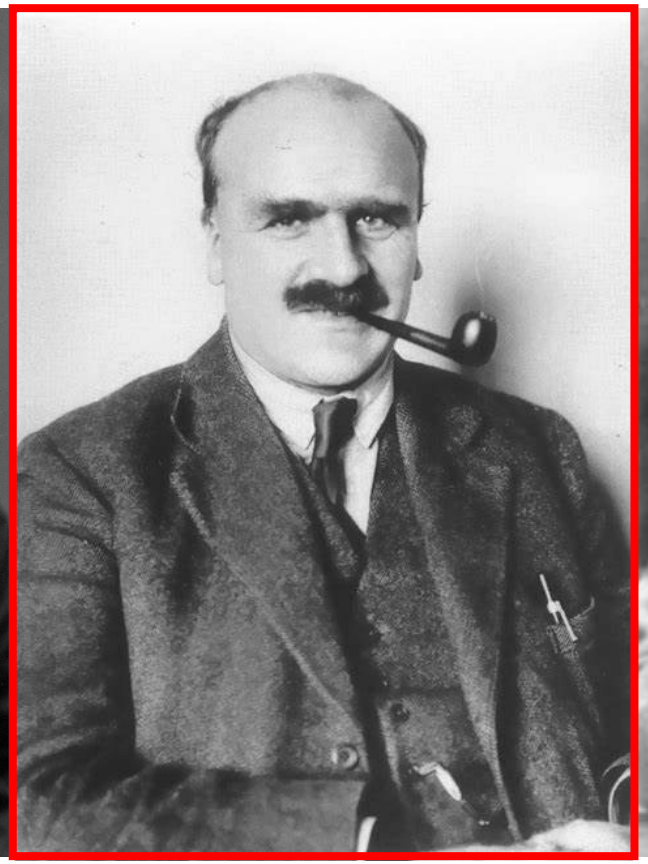
Low Stringency Searching

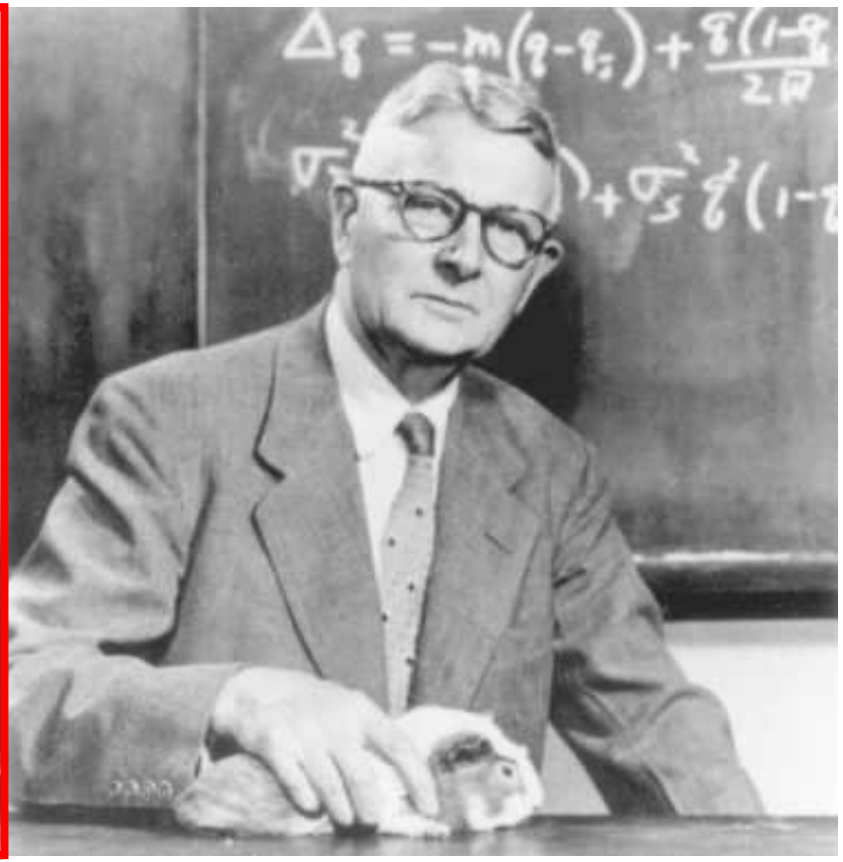
| Evidence Profile | CO Profile |
|---|---|
| 14, 20 | 14, X or 20, X |
| 15, 17 | 15, X or 17, X |
| 11, 11 | 11, X |

1 obligate allele is present

# What if… the offender is not in the database?







Ronald Fisher J.B.S. Haldane Sewall Wright

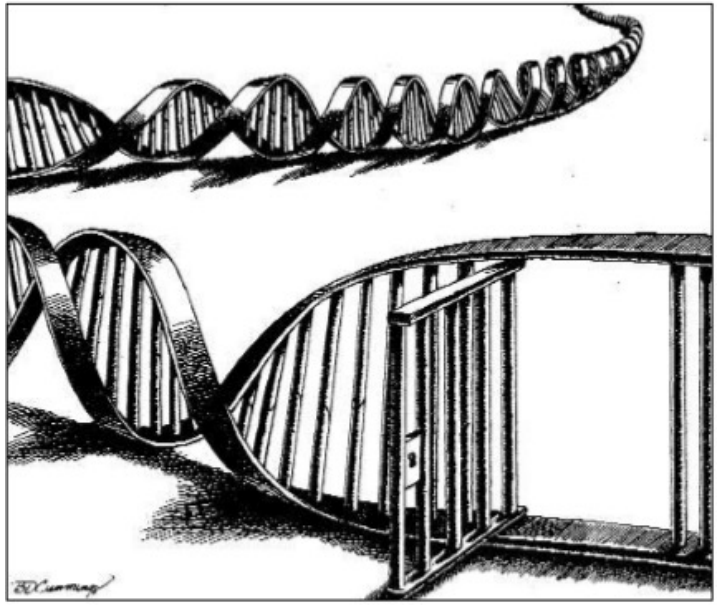
# Familial Searching

HUMAN GENETICS

## Finding Criminals Through DNA of Their Relatives

Frederick R. Bieber,[1*] Charles H. Brenner,[2] David Lazer[3]

*Science* (2006) **312:** 1315-1316

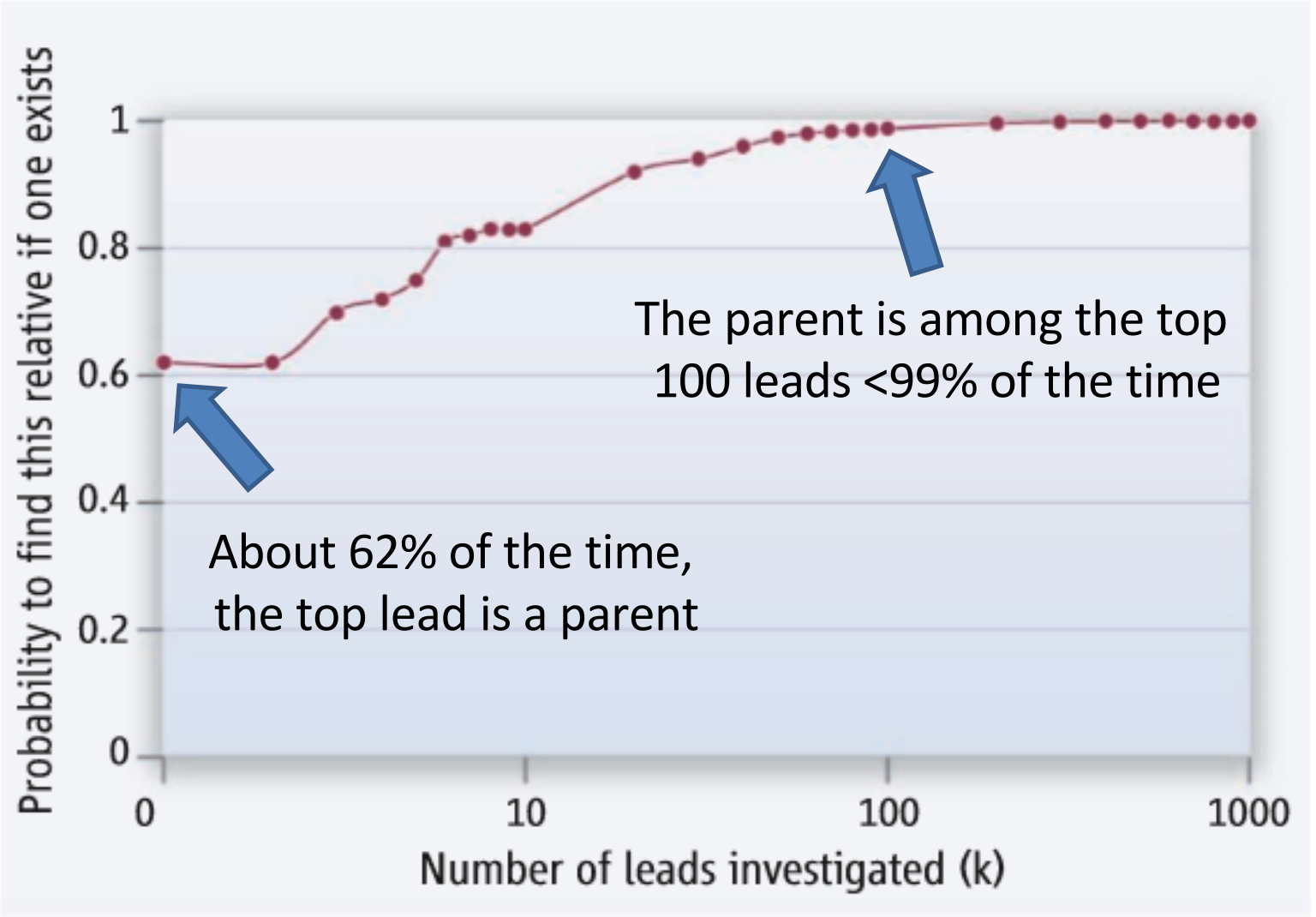


# Why Search for Relatives?

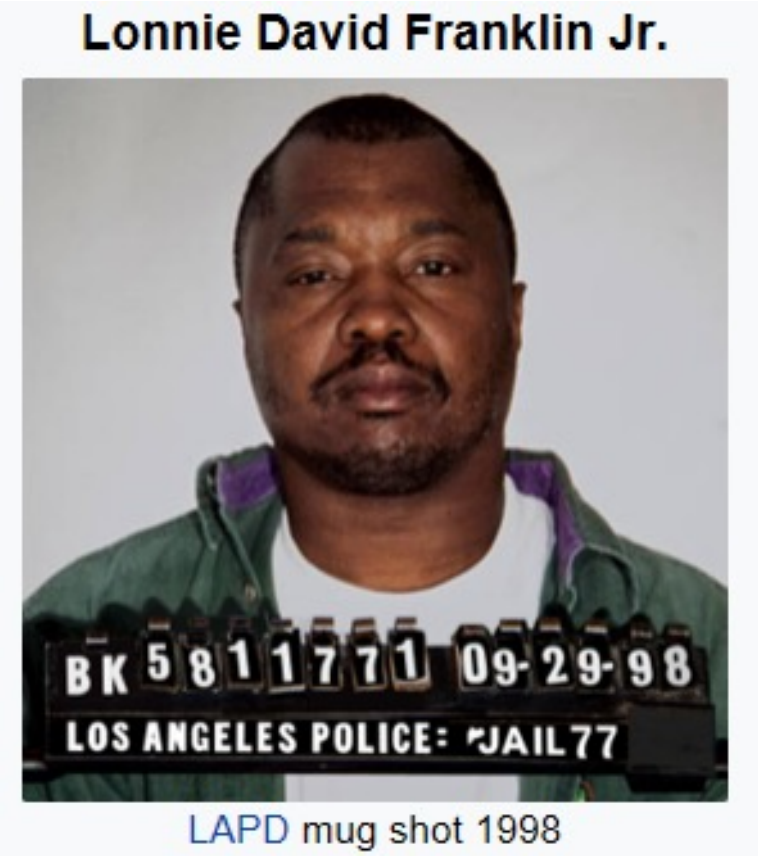
- Statistics show that (in the US) there is a strong probabilistic dependency between the conviction of a parent and their children.

- Surveys have shown that about 46% of prisoners have indicated that they have a close relative also in prison.

# Bieber et al. 2006



# Success Story



LAPD mug shot 1998

## "The Grim Sleeper"

12 Victims (11 homicides) from 1985 – 2007

Nearly all the victims were prostitutes.

California – passed an arrestee DNA collection law in 2004

Franklin was arrested in 2003, but his sample was never collected since he was on parole.

His son, Christopher, was arrested for gun possession and uploaded into CODIS in 2008.

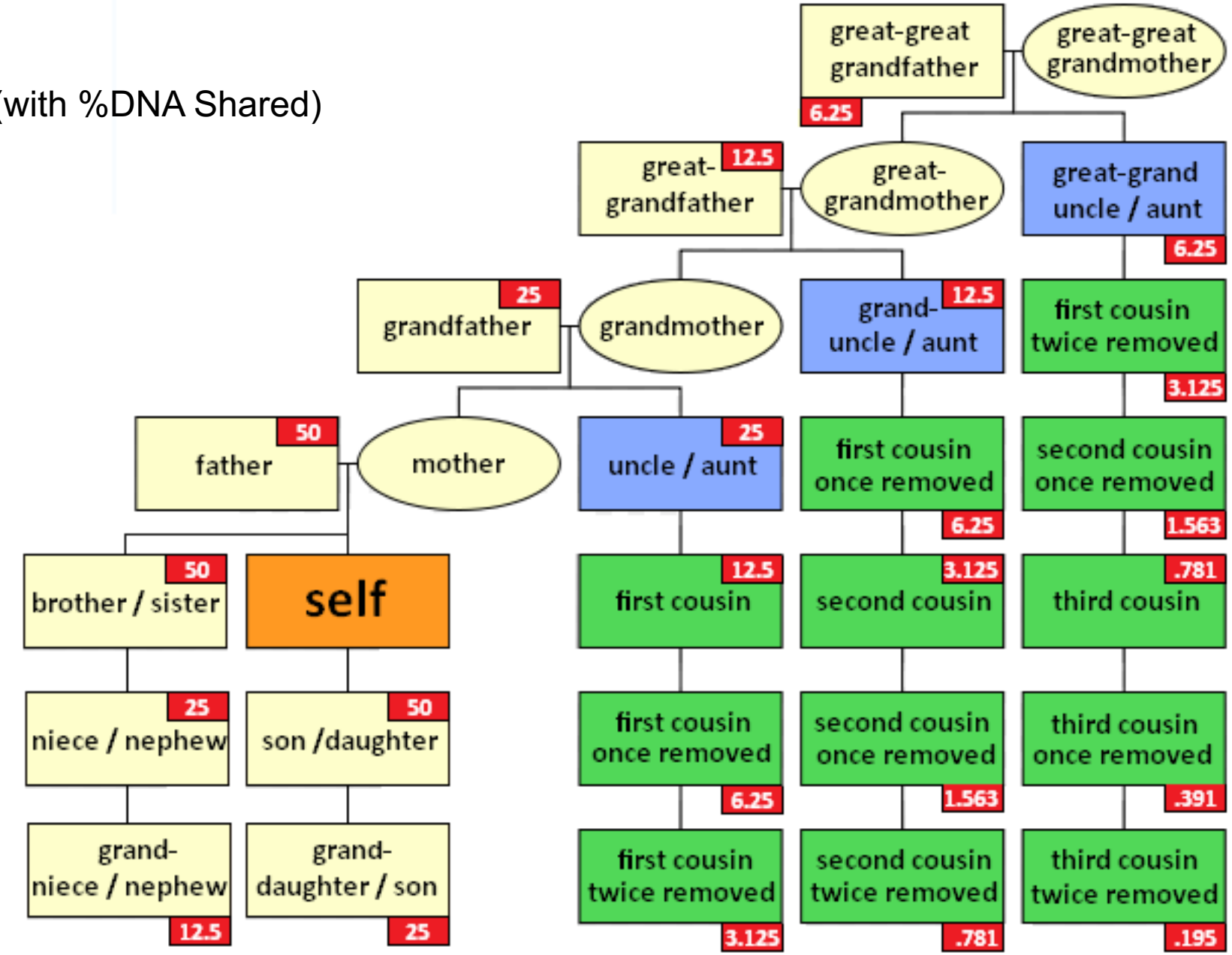
# Limitations of Familial Searching

- Not performed at a national level (only states)
- Some states prohibit familial testing in the U.S.
- False inclusions (or fortuitous matching)
- Females in the database – No Y-STR testing!!

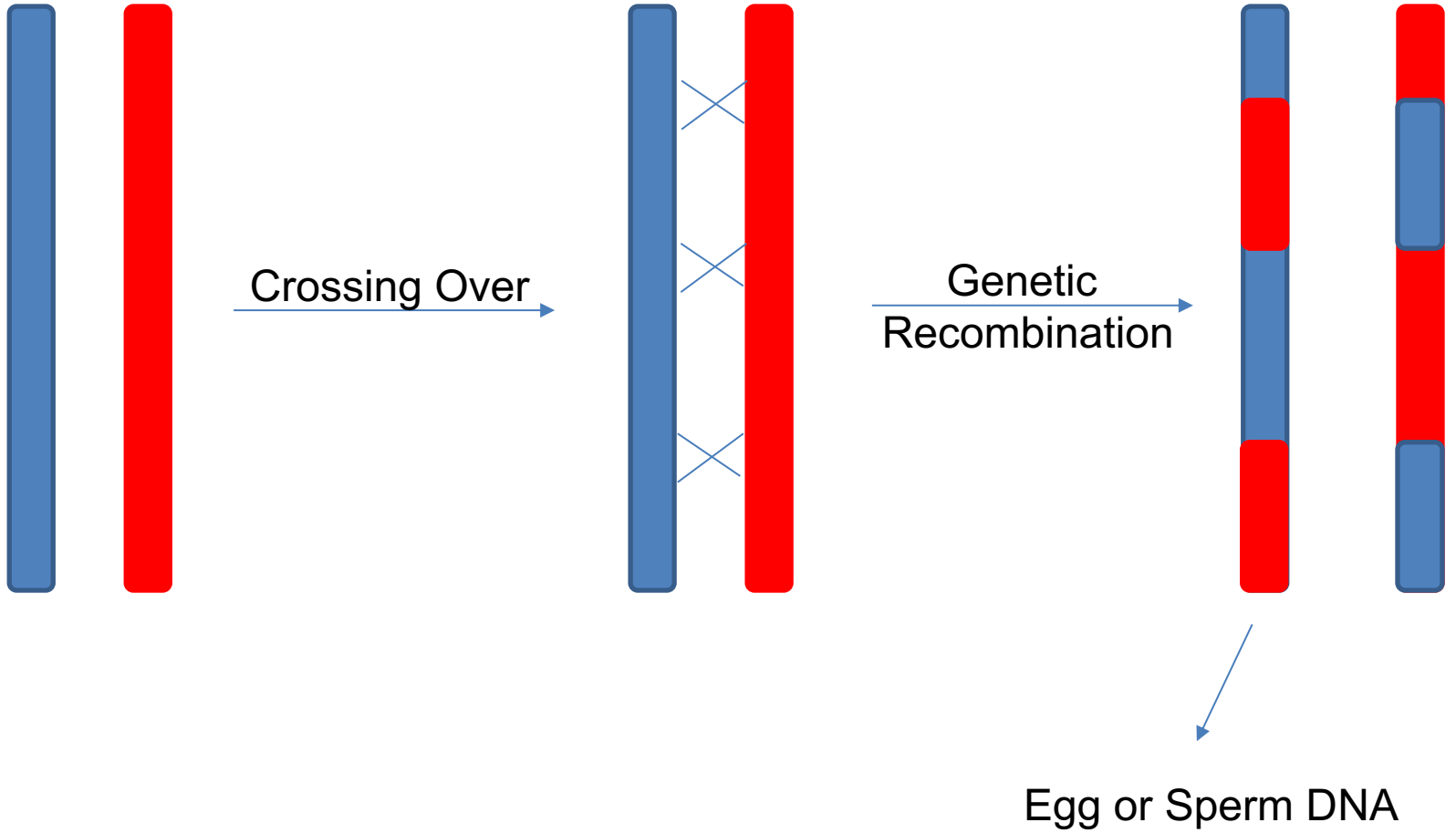
# Forensic Genetic Genealogy (Ray's lecture)

- Great investigative tool for cold cases and unidentified human remains
- About 1.2 million forensic samples in CODIS (no suspects)
- About 35,000 missing and unidentified persons in NamUS

Relationships (with %DNA Shared)



# Genetic Recombination



# Thomas Hunt Morgan

1933 Nobel Prize in Physiology or Medicine for discoveries elucidating the role that the chromosome plays in heredity.

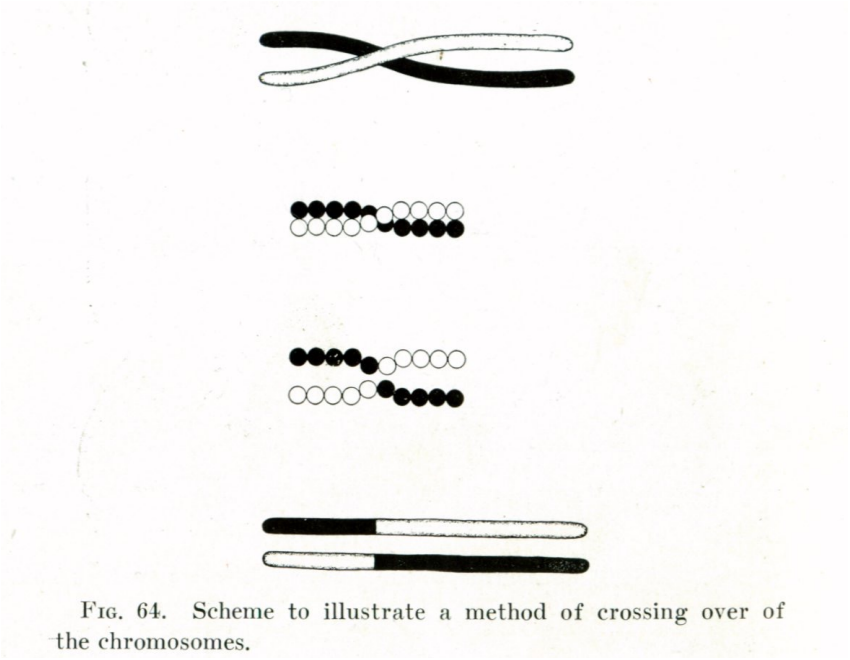


FIG. 64. Scheme to illustrate a method of crossing over of the chromosomes.

# Centi-Morgan (cM)

- A unit for measuring genetic linkage along chromosomes.
- 1 cM = 1 million bases of DNA (AKA “MegaBase”)
- Typically if two SNPs (or STRs) are 50 cM apart, they are considered unlinked and behave as if they are on separate chromosomes.

# My Cousins

Family Finder Matches

Help

All Matches | Detail View | Table View | Exact Search | Search | All

All (9494) | Paternal (0) | Maternal (0) | Both (0) | Filter | Sort by | Export CSV

Y-DNA37

| Ancestral Surnames | Haplogroup | Relationship Range | Shared DNA | Longest Block | X Match |
|---|---|---|---|---|---|
| Not Provided | Y-DNA: I-M253<br>mtDNA: N/A | 2nd Cousin - 4th Cousin<br>Link on Family Tree | 108 cM | 37 cM | No Match<br>Match date: May 10 2022 |

| Ancestral Surnames | Haplogroup | Relationship Range | Shared DNA | Longest Block | X Match |
|---|---|---|---|---|---|
| Not Provided | mtDNA: N/A | 2nd Cousin - 4th Cousin<br>Link on Family Tree | 81 cM | 81 cM | No Match<br>Match date: May 10 2022 |

FamilyTreeDNA

NEWS

# Former California cop arrested as Golden State Killer

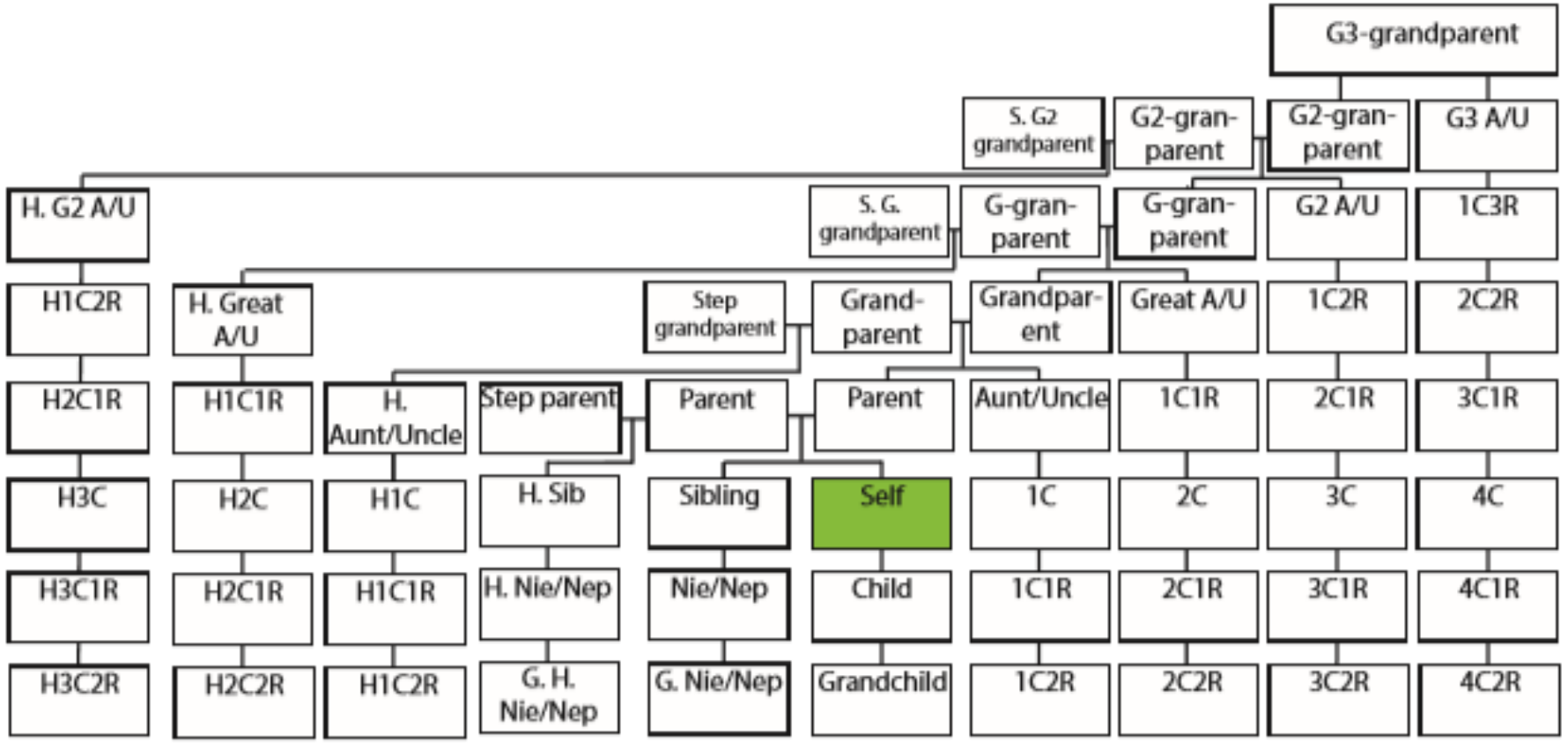
By Natalie O'Neill

April 25, 2018 | 5:20pm | Updated

Joseph James DeAngelo

Reuters

# Great, Great, Great Grandparent



# How Effective is FGG?

Science

REPORTS

Cite as: Y. Erlich *et al.*, *Science*
10.1126/science.aau4832 (2018).

## Identity inference of genomic data using long-range familial searches

**Yaniv Erlich[1,2,3,4*], Tal Shor[1], Itsik Pe'er[2,3], Shai Carmi[5]**

[1]MyHeritage, Or Yehuda 6037606, Israel. [2]Department of Computer Science, Fu Foundation School of Engineering, Columbia University, New York, NY, USA. [3]Center for Computational Biology and Bioinformatics (C2B2), Department of Systems Biology, Columbia University, New York, NY, USA. [4]New York Genome Center, New York, NY, USA. [5]Braun School of Public Health and Community Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

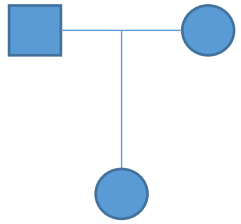
***Corresponding author. Email: erlichya@gmail.com**

With ~ I million people in GEDmatch, about 60% of the US European population can be identified (using 3rd cousins).

# Missing Persons Investigations

# The Missing Persons Index in CODIS

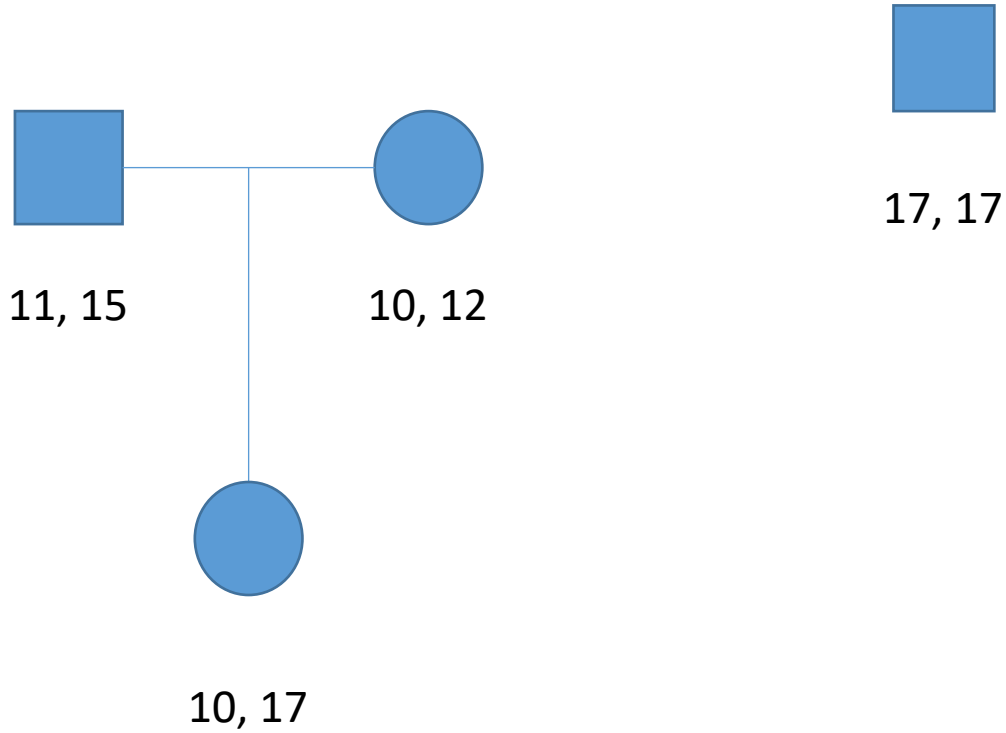
- Searching for missing persons in CODIS is driven by pedigree searches.
- The Unidentified Human Remains (UHRs) are entered into the MP Index of CODIS
- Family references are entered into the Index using a pedigree



- The software will then search each UHR against every pedigree

# Example of using a Low Stringency Search for Missing Persons Investigations

Suppose…



# Kinship Statistics

- The Kinship Index (KI) is a Likelihood Ratio (LR) that is used to measure the weight of the evidence given two alternative hypothesis.

$$LR = \frac{\Pr\left(E \mid H_p, I\right)}{\Pr\left(E \mid H_d, I\right)}$$

This person is a true child of these two parents

This is some random person in the population and is not the true child of these two parents

LRs<1 give more weight to the hypothesis that this person is not a child

LRs>1 give more weight to the hypothesis that this person is a child

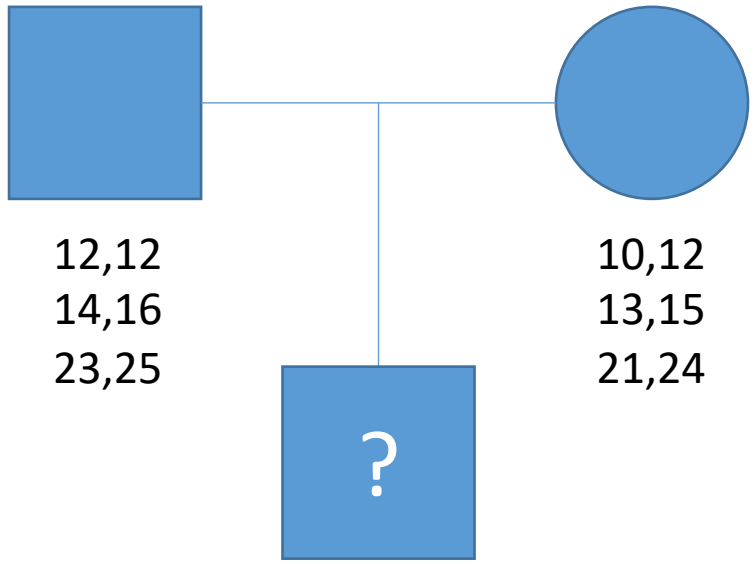
LRs=1 supports neither hypothesis (uninformative)

# CODIS Architecture

## CODIS-MP
## (Missing Persons)

| UHR-1 | UHR-2 | UHR-3 |
|---|---|---|
| 13,14 | 12,12 | 10,11 |
| 14,16 | 14,15 | 14,16 |
| 22,23 | 23,24 | 23,25 |

## Pedigree-driven searches

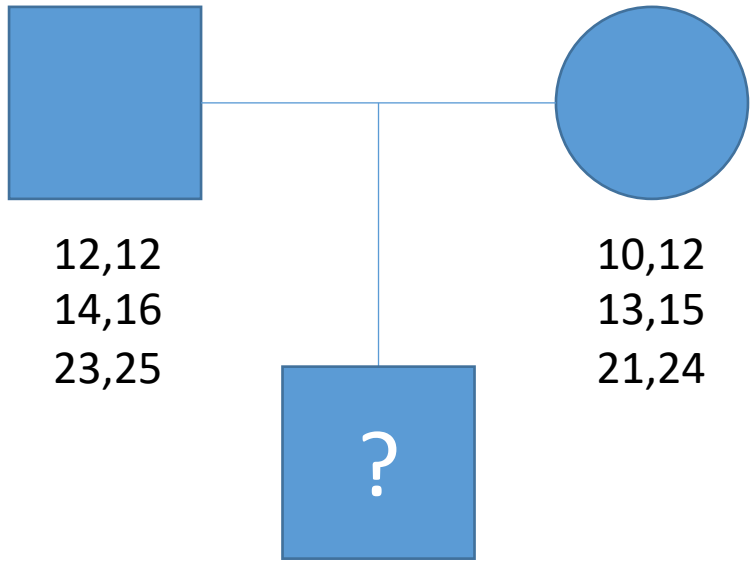


# CODIS Architecture

## CODIS-MP
## (Missing Persons)

| UHR-1 | UHR-2 | UHR-3 |
|---|---|---|
| 13,14 | 12,12 | 10,11 |
| 14,16 | 14,15 | 14,16 |
| 22,23 | 23,24 | 23,25 |

## Pedigree-driven searches

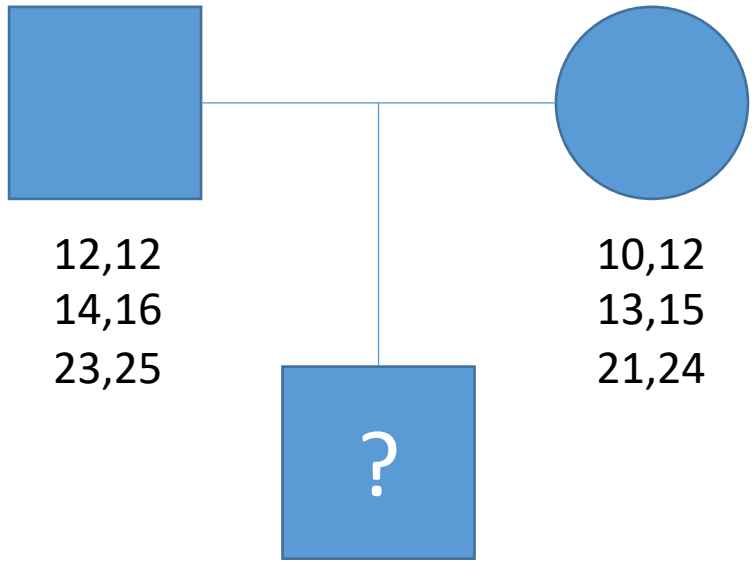


# CODIS Architecture

## CODIS-MP
## (Missing Persons)

| UHR-2 | UHR-3 |
|---|---|
| 12,12 | 10,11 |
| 14,15 | 14,16 |
| 23,24 | 23,25 |

## Pedigree-driven searches




12,12
14,16
23,25

10,12
13,15
21,24

UHR-1
13,14
14,16
22,23

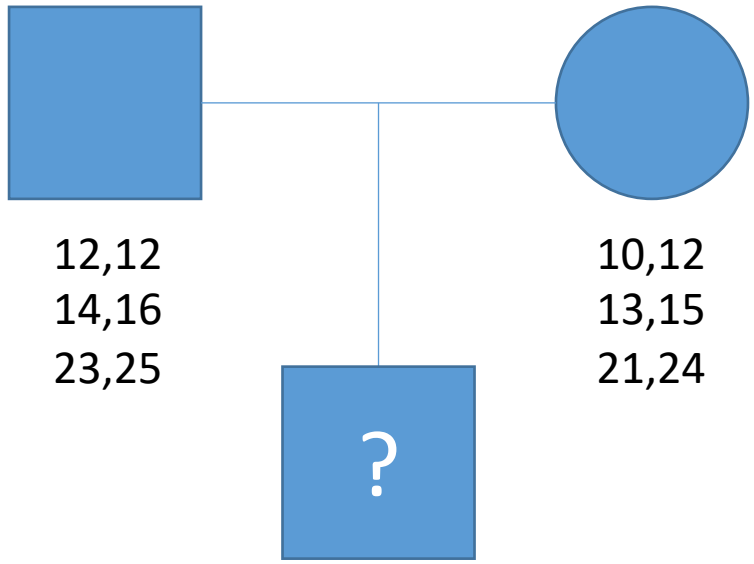
LR = 0
NO!

# CODIS Architecture

## CODIS-MP
## (Missing Persons)

| UHR-1 | UHR-2 | UHR-3 |
|---|---|---|
| 13,14 | 12,12 | 10,11 |
| 14,16 | 14,15 | 14,16 |
| 22,23 | 23,24 | 23,25 |

## Pedigree-driven searches



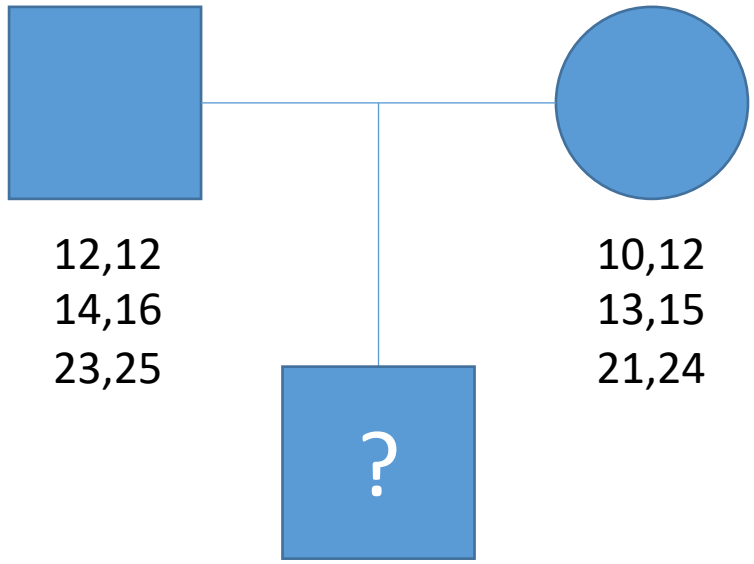
# CODIS Architecture

## CODIS-MP (Missing Persons)

UHR-1
13,14
14,16
22,23

UHR-3
10,11
14,16
23,25

## Pedigree-driven searches




12,12
14,16
23,25

10,12
13,15
21,24

UHR-2
12,12
14,15
23,24

LR = 100,000
YES

# CODIS Architecture

## CODIS-MP (Missing Persons)

| UHR-1 | UHR-2 | UHR-3 |
|---|---|---|
| 13,14 | 12,12 | 10,11 |
| 14,16 | 14,15 | 14,16 |
| 22,23 | 23,24 | 23,25 |

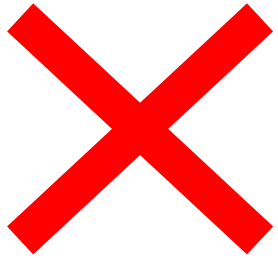
## CODIS-CO/A (Offenders)

| CO-1 | CO-2 | CO-3 |
|---|---|---|
| 15,16 | 12,12 | 11,11 |
| 12,17 | 14,15 | 15,16 |
| 25,25 | 23,24 | 24,27 |

# CODIS Architecture

## CODIS-MP (References)

| REF-1 | REF-2 | REF-3 |
|---|---|---|
| 13,14 | 12,12 | 10,11 |
| 14,16 | 14,15 | 14,16 |
| 22,23 | 23,24 | 23,25 |



## CODIS-FI (Forensic Evidence)

| FE-1 | FE-2 | FE-3 |
|---|---|---|
| 15,16 | 12,12 | 11,11 |
| 12,17 | 14,15 | 15,16 |
| 25,25 | 23,24 | 24,27 |

NO!!!!

# Conclusions

- DNA databases, especially those like CODIS, are critical for generating investigative leads.
- Familial Searching can sometimes lead to a close relative of the POI.
- Non-public databases are becoming equally as critical in criminal and missing persons cases, but may not be useful outside of W. European ancestry.

# Thank You!




- Dr. Richard Somiari (ITIS, LLC)
- The organizing committee of the Lagos Forensic Symposium
- Colleagues at the UNTHSC Center for Human Identification

michael.coble@unthsc.edu