

Investigative Genetic Genealogy (IGG): current methods, knowledge and practice



Genetisti Forensi Italiani

XXVIII
CONGRESSO
NAZIONALE

**LA GENETICA
FORENSE:
QUALE FUTURO?**

10|11|12
Novembre 2022
GENOVA



Christopher Phillips

Forensic Genetics Unit,
University of Santiago de
Compostela, Spain



Buckskin Girl, a 37-year MP ID case - identified 11th April 2018

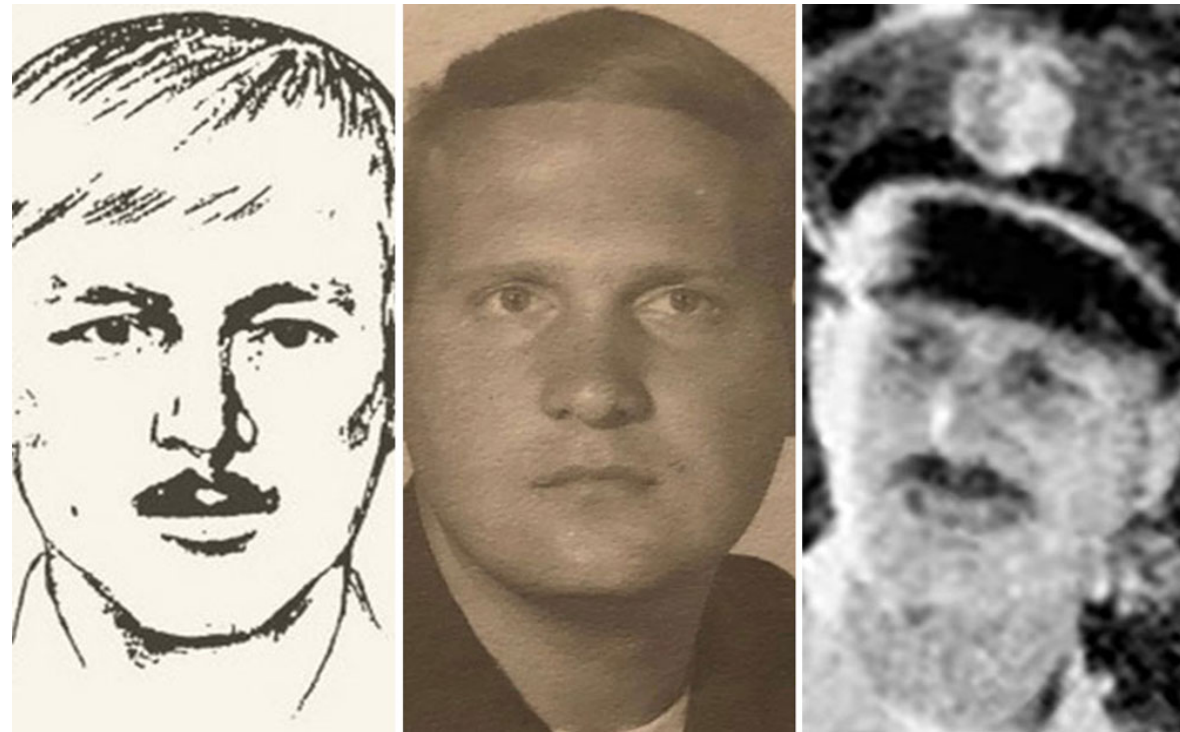


DNA Doe Project is pleased to announce our first success story. We have identified Buckskin Girl, as confirmed this week by the Miami County, Ohio Sheriff's Office. While all other forensic identification methods failed, our pioneering approach using genetic genealogy succeeded.

We participated in a press conference held in Troy, Ohio this afternoon where the news was made public by Chief Deputy Stephen Lord.

- Found in a roadside ditch in Troy, Ohio, 24th April 1981
- 2001 STR profile generated
- 2009 mtDNA profile to CODIS
- 2018 DNA Doe project involved
- New DNA extract made from highly degraded blood sample in storage
- Whole genome sequencing by *Full Genomes Corporation*
- 28 March 2018, DNA profile uploaded to GEDmatch community database
- Successful match with 1st cousin-once removed (equivalent to 2nd cousin)
- Full analysis took 4 hours
- 11th April 2018 press conference - identified as Marcia King, aged 21

The Golden State Killer Suspect - identified 24th April 2018



- 13 murders and 13 kidnappings linked to North California region
- Former police officer so had easy access
- He possibly interfered with evidence

UNKNOWN SUSPECT

East Area Rapist/Golden State Killer
California
1976 to 1986



REWARD

The FBI is offering a reward of up to \$50,000 for information leading to the arrest and conviction of the individual known as the East Area Rapist/Golden State Killer.

REMARKS

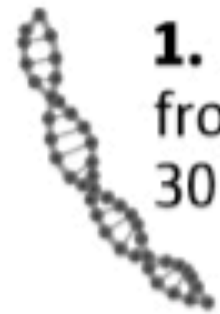
The East Area Rapist/Golden State Killer (EAR/GSK) is described as a White male, currently thought to be between the ages of 60 and 75 years old, and approximately 5'10" tall. He may have/had blond or light brown hair and an athletic build. He may have had an interest in the military, or had some military training, leaving him familiar and proficient with firearms.



- Directly descended from Italian immigrants to the US
- Likely to have premature baldness (but probable age was ~70Y)
- Blue eyes

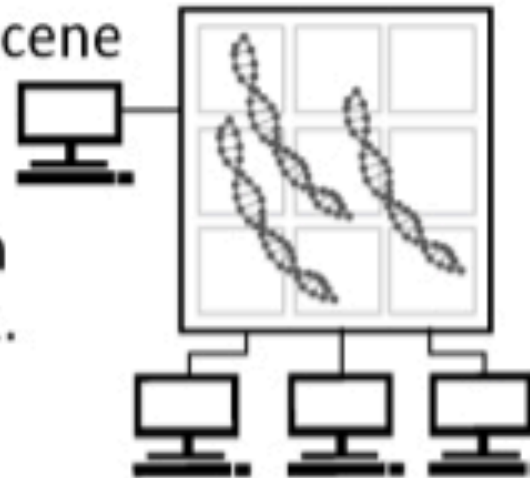
- 13 murders and 13 kidnappings linked to North California region
- Former police officer so had easy access
- He possibly interfered with evidence
- Four months to process rape-kit DNA
- It is likely, but not admitted, that a DTC 'buccal swab' was dosed with DNA and the SNP data uploaded to GEDmatch
- 10-20 relatives shared the same great-great-great grandparents
- Genealogist Barbara Rae Venter made multiple, extensive family trees
- Two suspects matched likely location, age, appearance, modus operandi data
- Collected DNA from cup/car door/tissue
- 24th April 2018 press conference - identified as Joseph James DeAngelo, aged 72: 8 plus 4 FD murder charges

Identifying the Golden State Killer - the critical steps



1. DNA was collected from the crime scene 30 years ago.

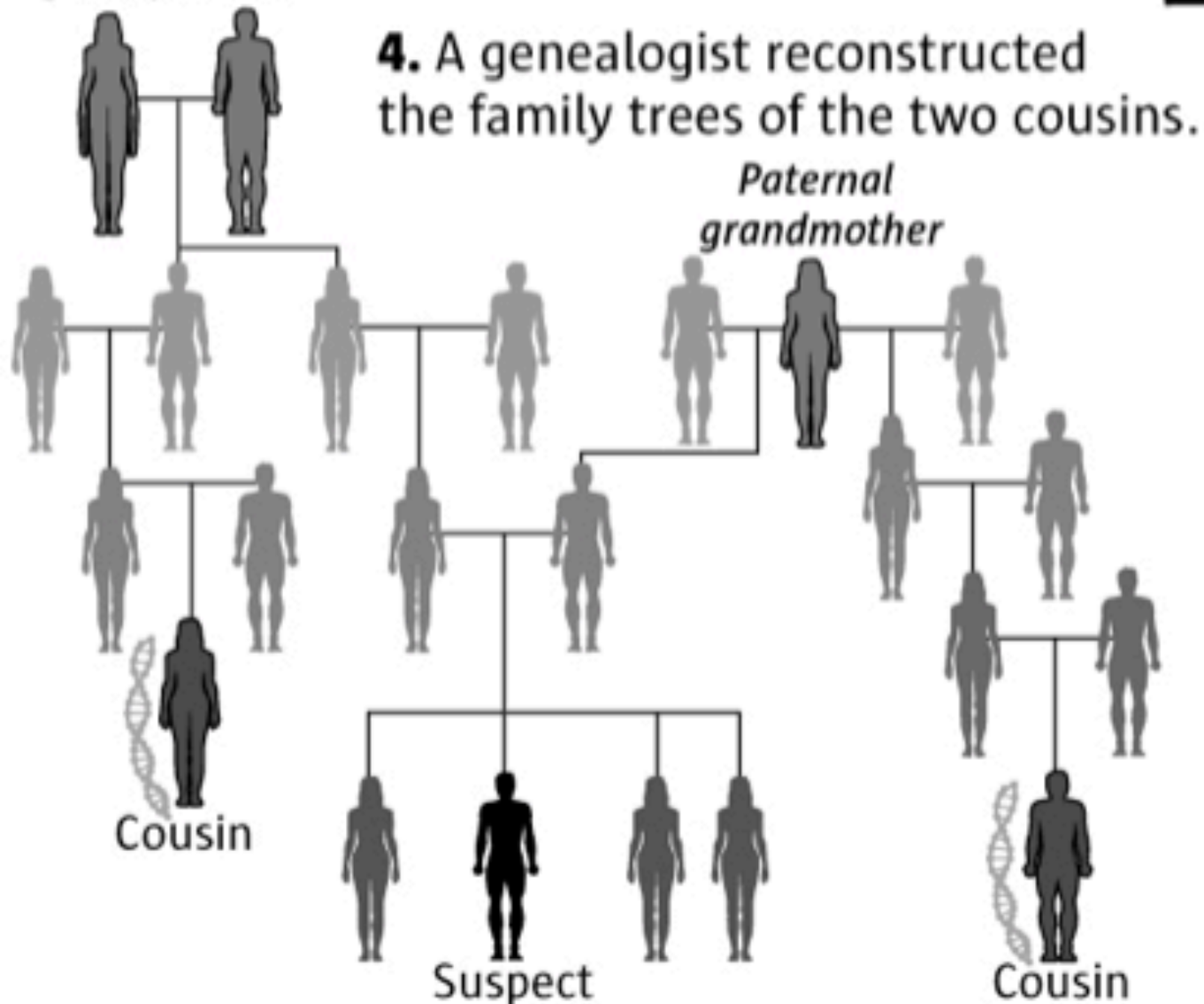
2. Crime-scene DNA run through GEDmatch data bank.



3. Results indicate shared DNA with two second cousins.



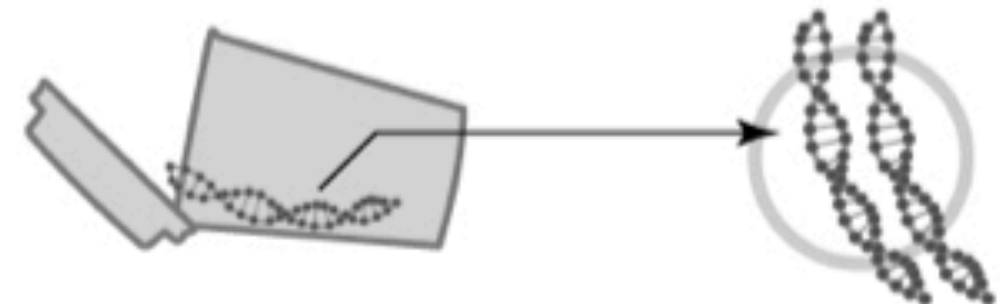
Maternal great grandparents



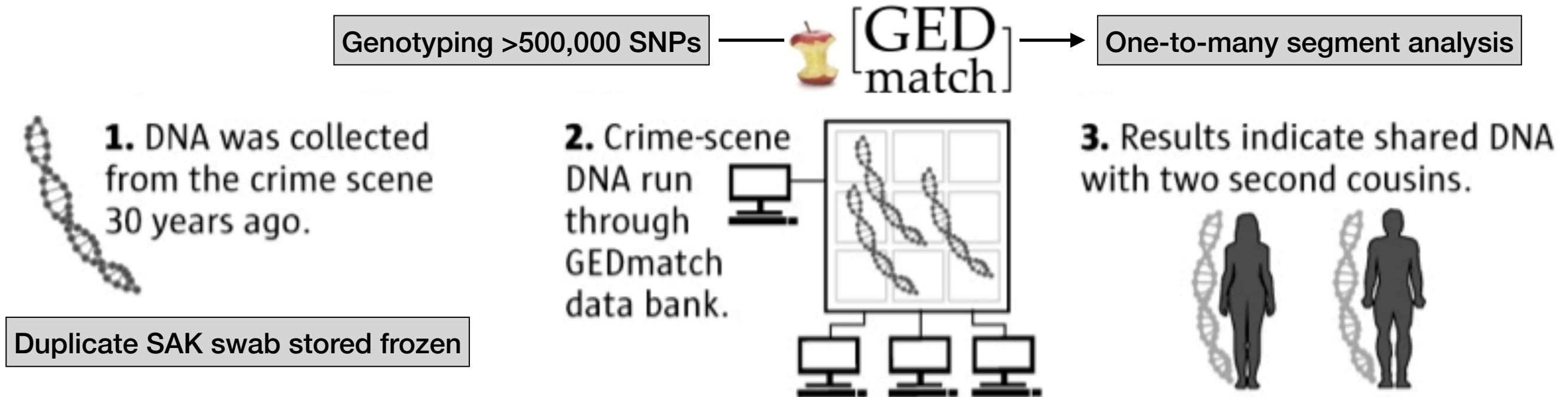
5. The reconstructed family trees lead to a suspect through common ancestors.



6. Detectives collect DNA from a discarded cup. It matches DNA from the crime scene.



Identifying the Golden State Killer - steps 2-3 were *DNA analysis*



Forensic Science International: Genetics 36 (2018) 186–188



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Commentary

The Golden State Killer investigation and the nascent field of forensic genealogy

Chris Phillips

Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain



Forensic Science International: Genetics 36 (2018) 203–204



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Forensic genealogy: Some serious concerns

Denise Syndercombe Court

King's Forensics, Faculty of Life Sciences and Medicine, King's College, London, UK



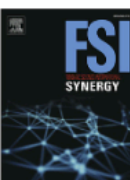
Forensic Science International: Synergy 1 (2019) 114–125



Contents lists available at ScienceDirect

Forensic Science International: Synergy

journal homepage: <https://www.journals.elsevier.com/forensic-science-international-synergy/>



Forensic genealogy, bioethics and the Golden State Killer case

Ray A. Wickenheiser

Laboratory System Director, New York State Police Crime Laboratory System, 1220 Washington Ave, Building 30, Albany, NY, 12226-3000, USA



Forensic Science International 301 (2019) 107–117



Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint



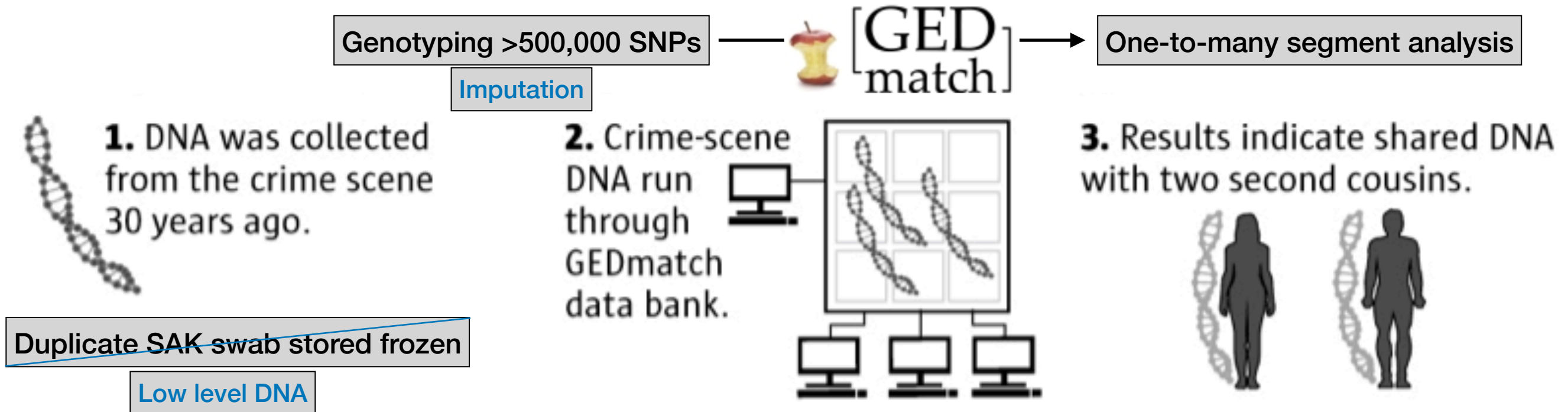
Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes

Debbie Kennett

Research Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, United Kingdom



Identifying the Golden State Killer - steps 2-3 were *DNA analysis*



Forensic Science International: Genetics 36 (2018) 186–188



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Commentary

The Golden State Killer investigation and the nascent field of forensic genealogy

Chris Phillips

Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain



Forensic Science International: Genetics 36 (2018) 203–204



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Forensic genealogy: Some serious concerns

Denise Syndercombe Court

King's Forensics, Faculty of Life Sciences and Medicine, King's College, London, UK



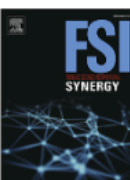
Forensic Science International: Synergy 1 (2019) 114–125



Contents lists available at ScienceDirect

Forensic Science International: Synergy

journal homepage: <https://www.journals.elsevier.com/forensic-science-international-synergy/>



Forensic genealogy, bioethics and the Golden State Killer case

Ray A. Wickenheiser

Laboratory System Director, New York State Police Crime Laboratory System, 1220 Washington Ave, Building 30, Albany, NY, 12226-3000, USA



Forensic Science International 301 (2019) 107–117



Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint



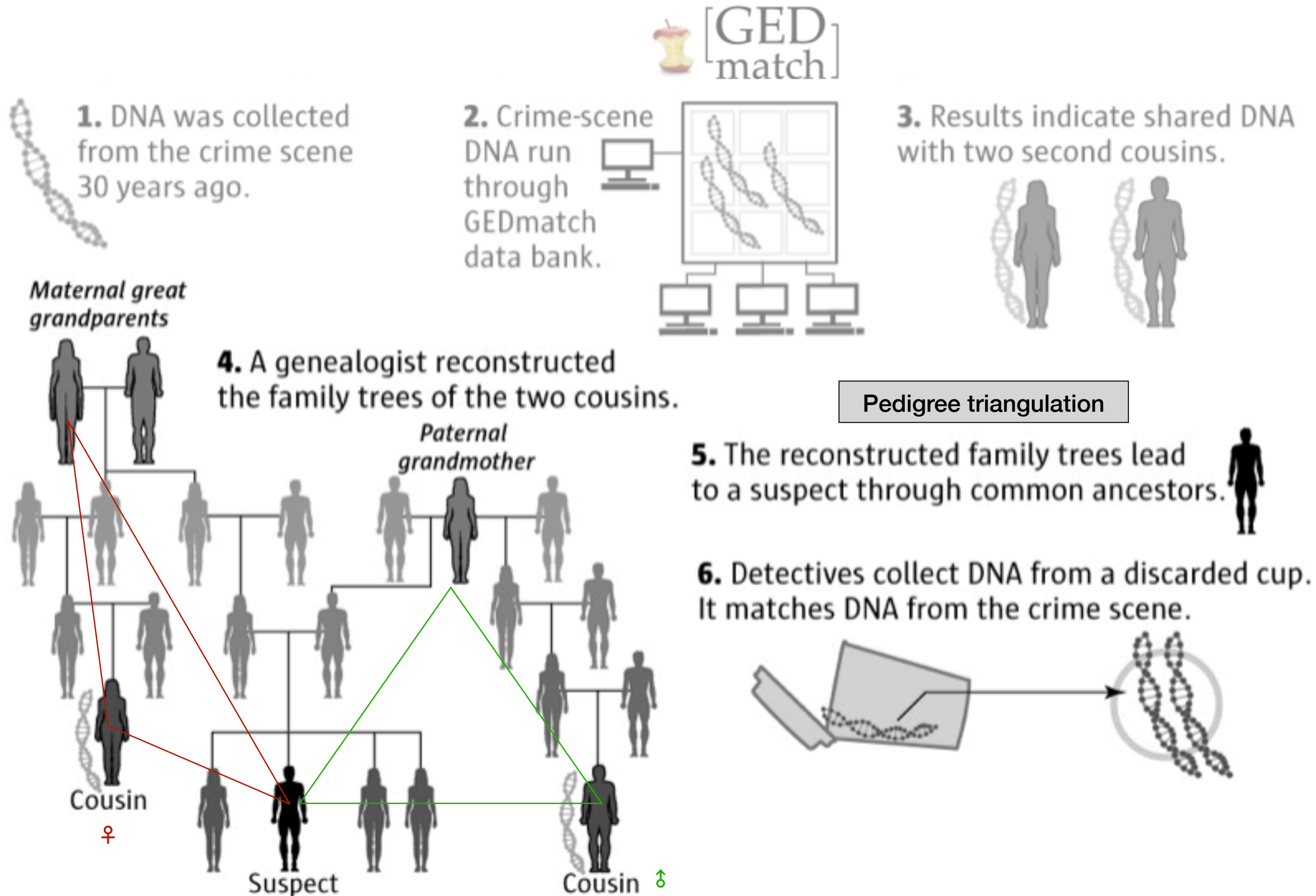
Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes

Debbie Kennett

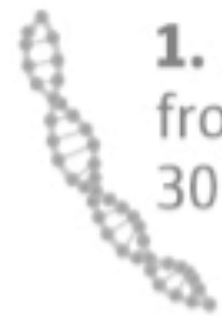
Research Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, United Kingdom



Identifying the Golden State Killer - steps 4-5 were *genealogy*

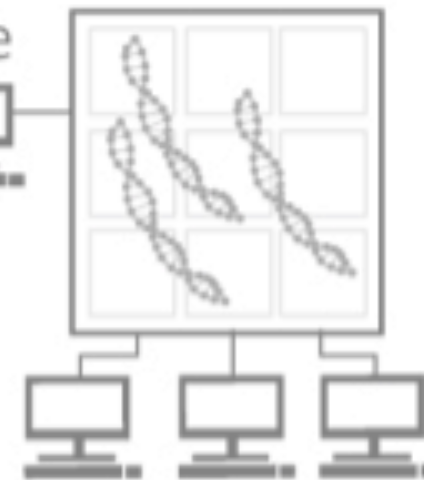


Identifying the Golden State Killer - steps 4-5 were *genealogy*



1. DNA was collected from the crime scene 30 years ago.

2. Crime-scene DNA run through GEDmatch data bank.



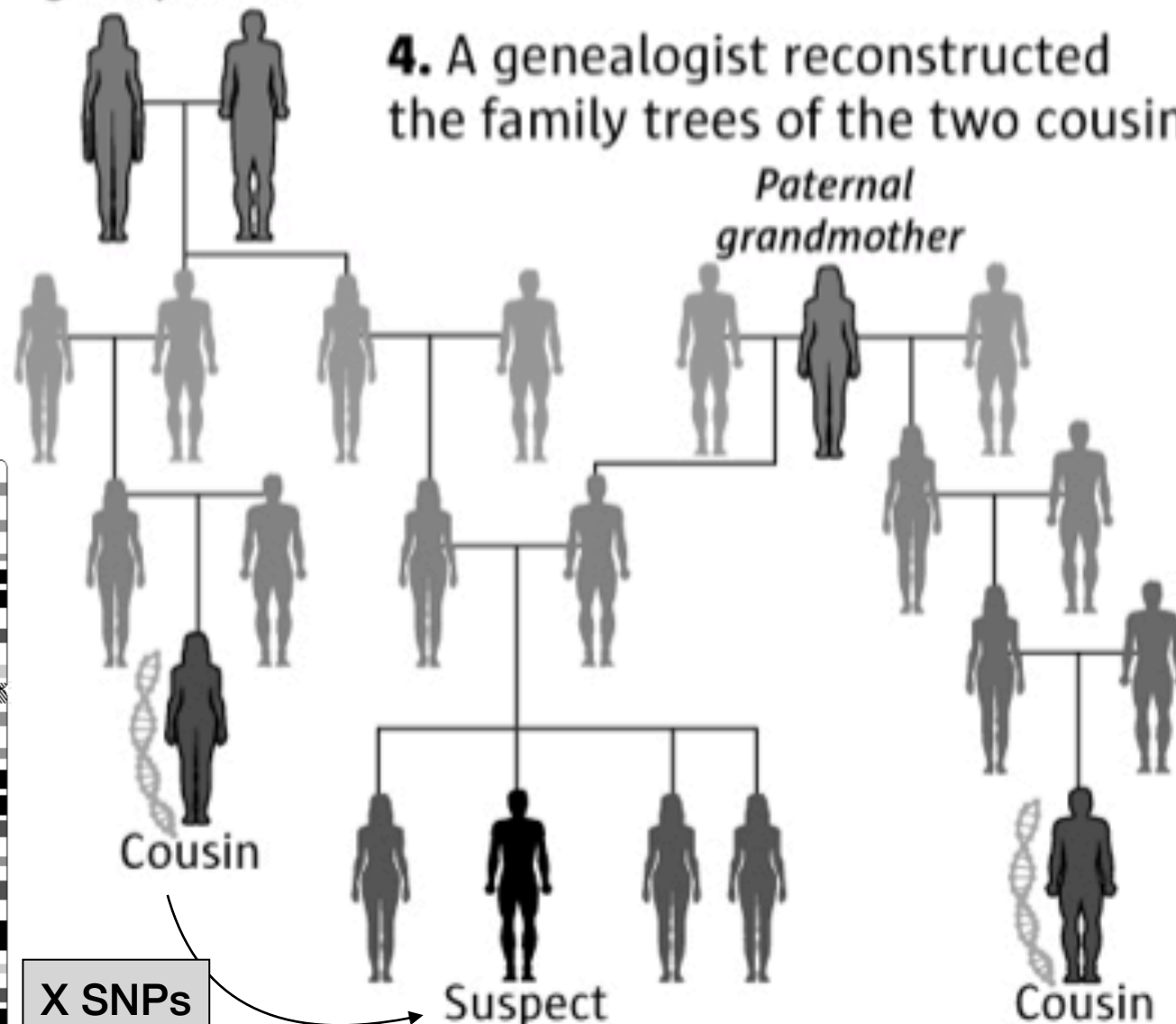
3. Results indicate shared DNA with two second cousins.



Maternal great grandparents

4. A genealogist reconstructed the family trees of the two cousins.

Paternal grandmother

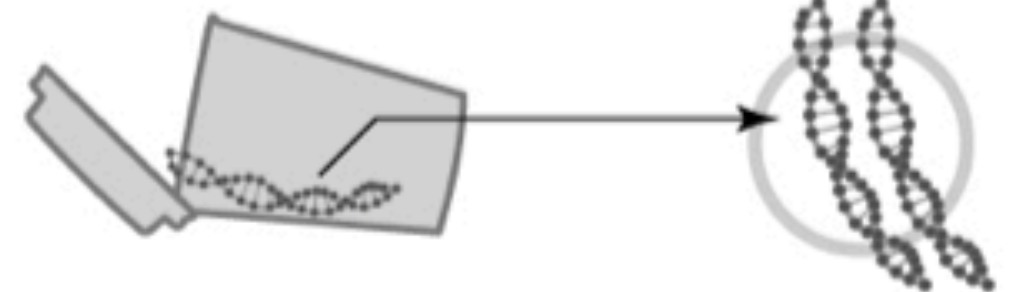


Pedigree triangulation

5. The reconstructed family trees lead to a suspect through common ancestors.



6. Detectives collect DNA from a discarded cup. It matches DNA from the crime scene.



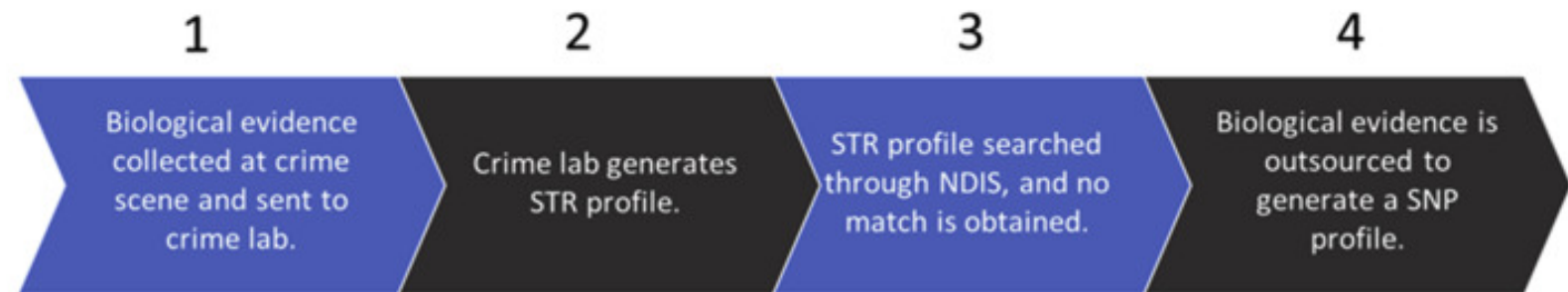
The confirmation of identity applies the principle of 'CODIS first and last'

The IGG workflow in more detail



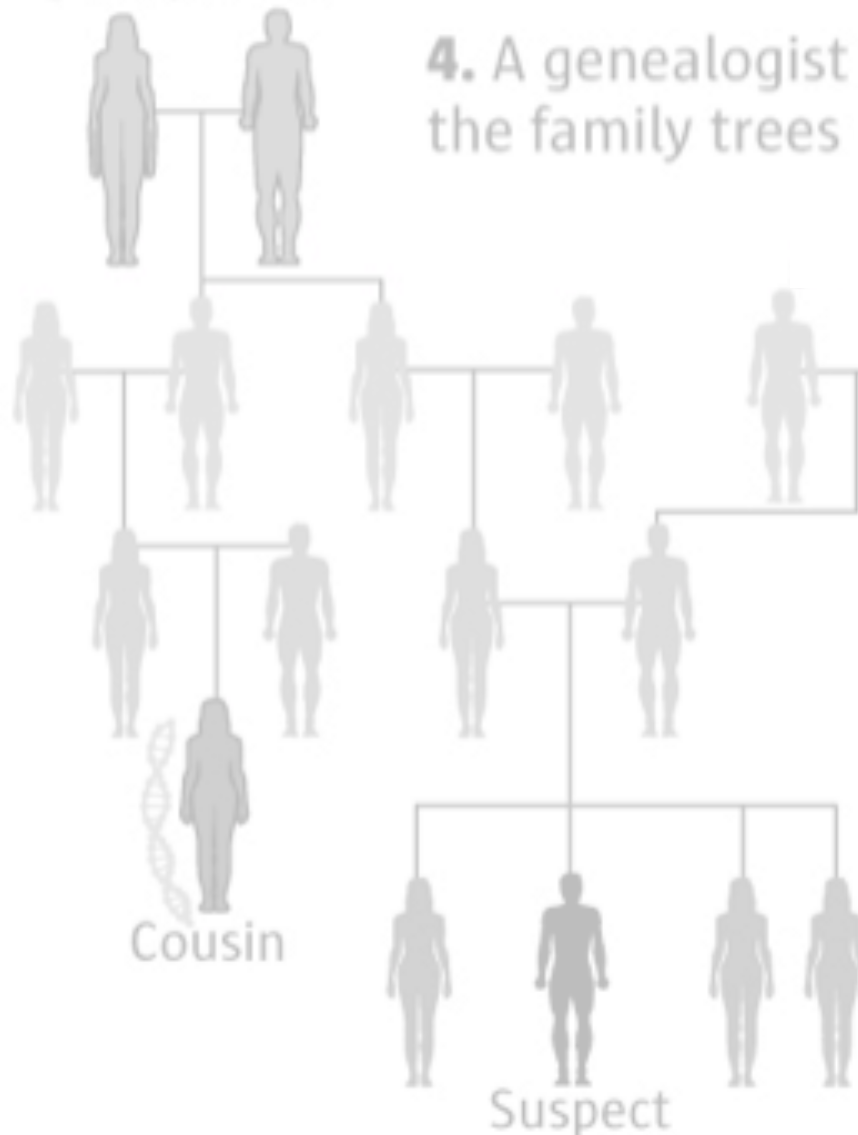
1. DNA was collected from the crime scene 30 years ago.

Phase 1: DNA Analysis

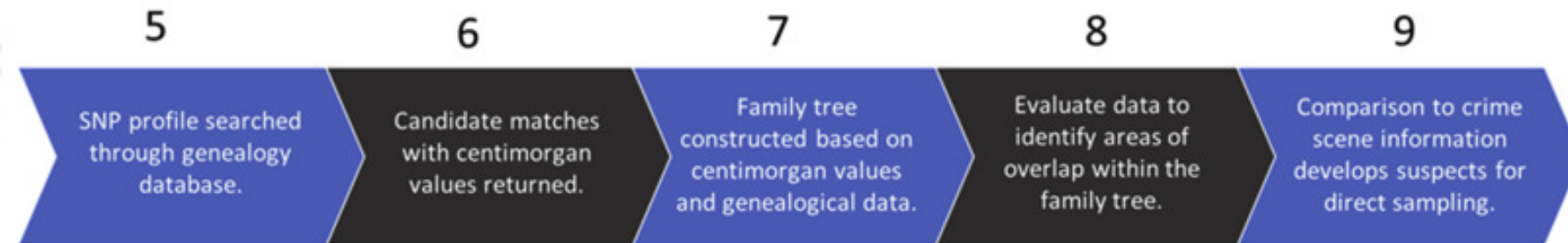


CODIS first

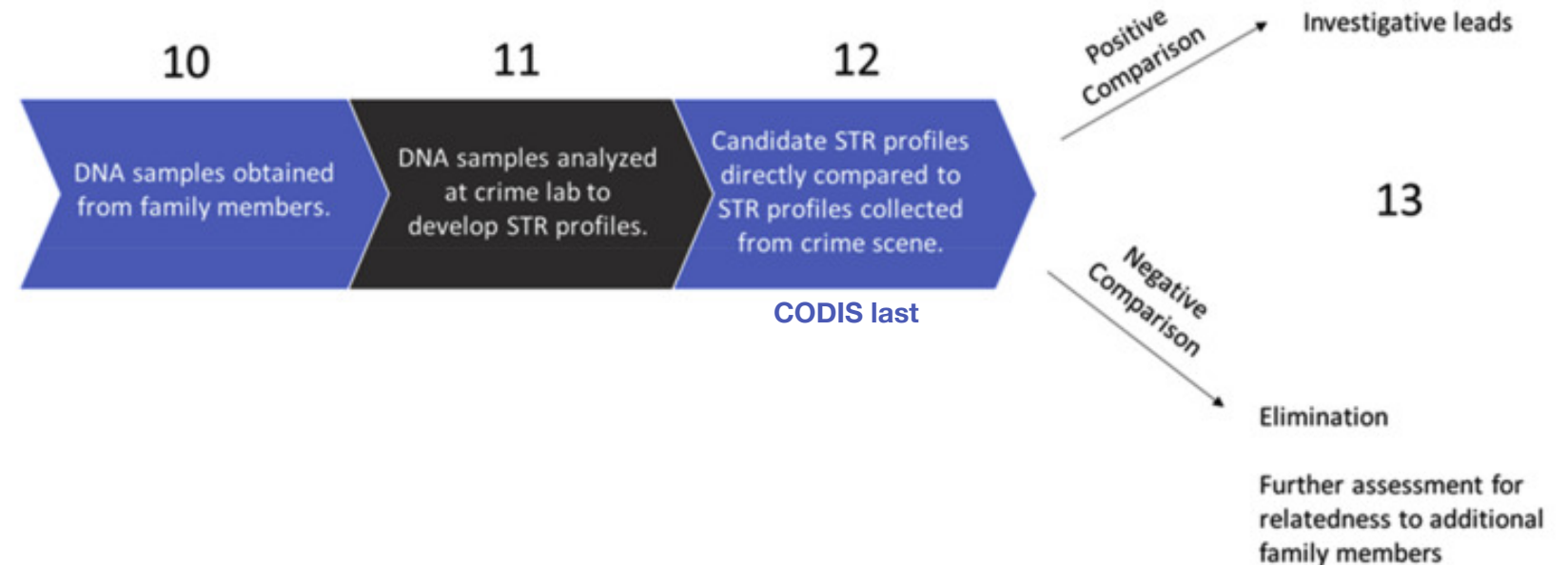
Maternal great grandparents



Phase 2: Genealogy

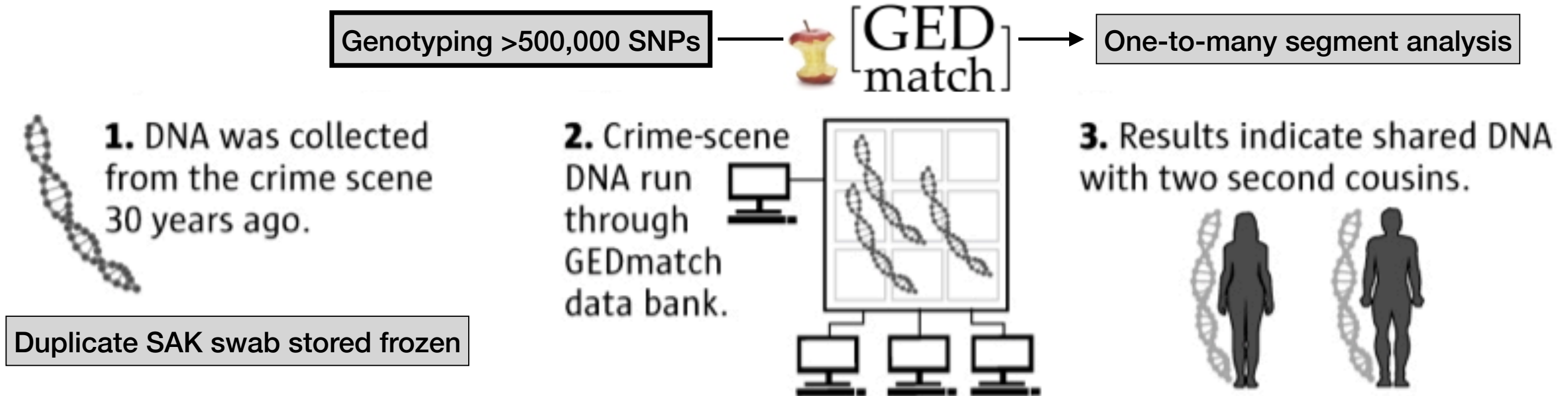


Phase 3: Investigation



CODIS last

Identifying the Golden State Killer - genotyping high numbers of SNPs

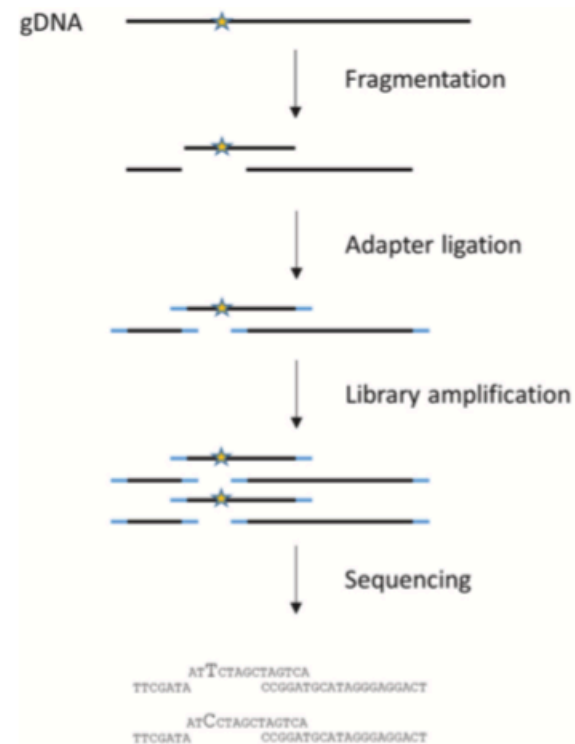


Whole-genome sequencing

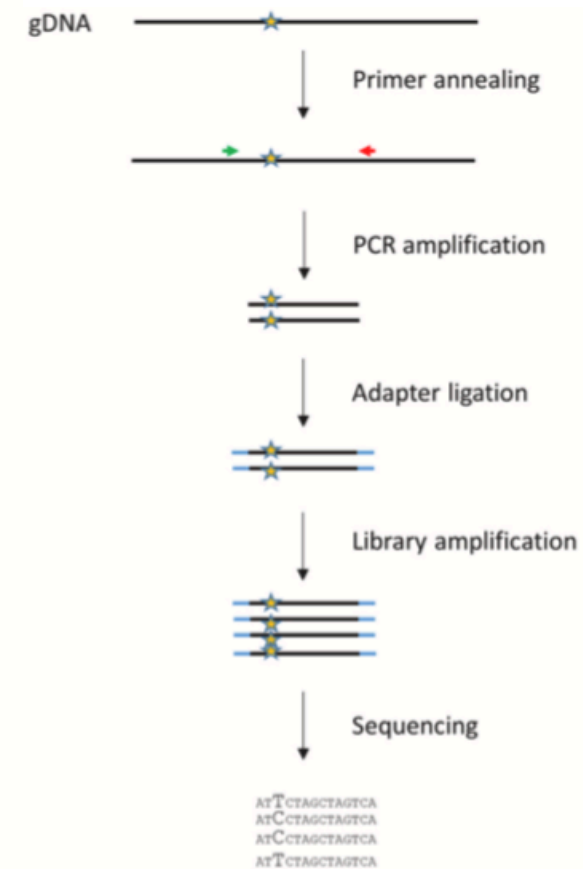
Sonication creates a pool of random short fragments

Adaptors are ligated that identify each sample and aid library amplification

Whole sequences are stitched together from segment overlaps



MPS using conventional forensic PCR

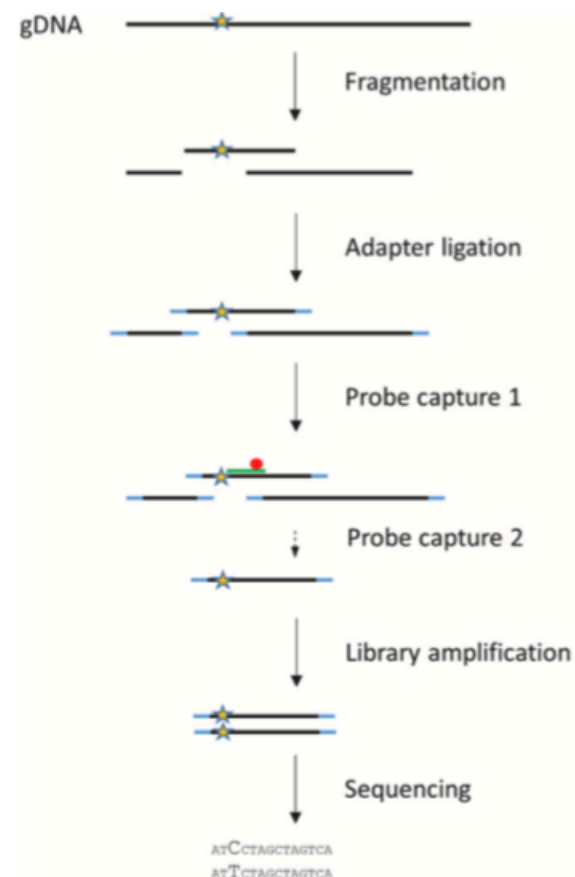


Hybridisation capture methods

DNA fragments with ligated adaptors are captured by probe hybridisation (often termed 'baits')

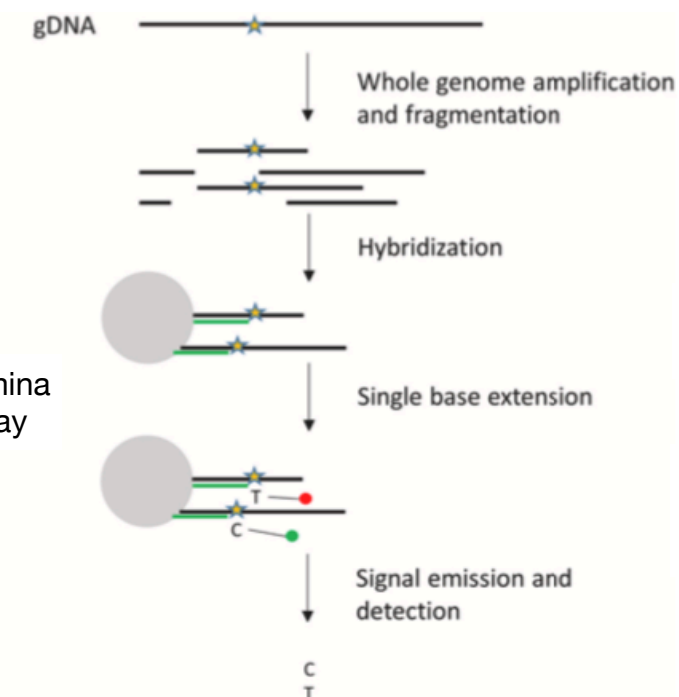
Probes removed and the enriched DNA pool provides efficient library amplification

Sequencing can use established platforms at varied scales: Illumina NextSeq/ NovaSeq, Thermo Fisher Ion GeneStudio



Whole-genome SNP arrays

e.g. Illumina BeadArray



SBE principle is identical to SNaPshot

Tend to use Cy5-Cy3 dyes so only binary SNPs detected

Large-scale SNP genotyping

Whole-genome sequencing

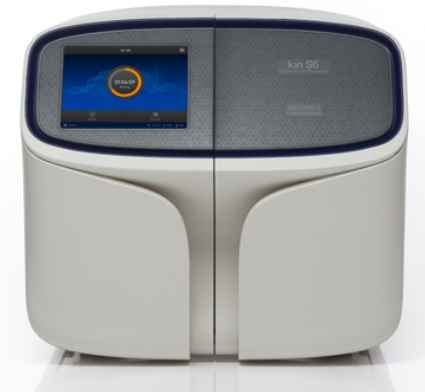
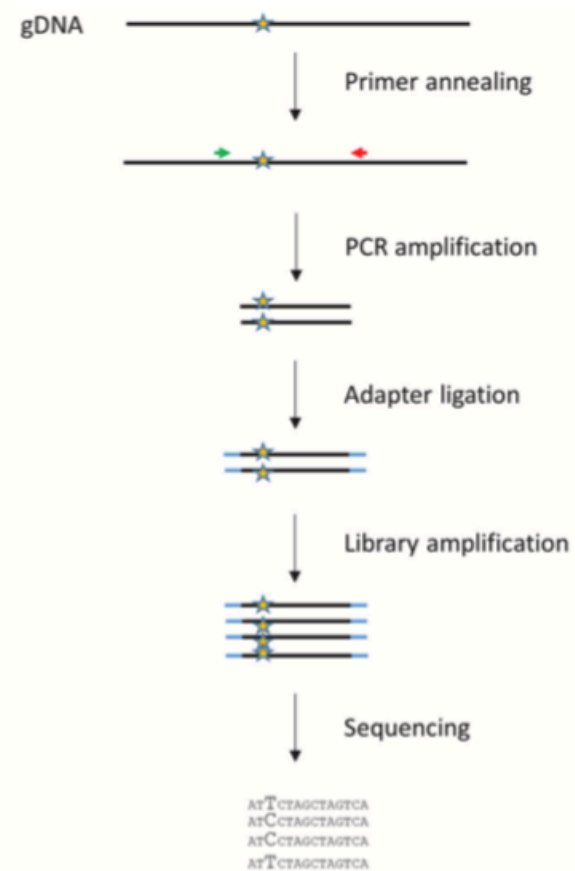
Sonication creates a pool of random short fragments

Adaptors are ligated that identify each sample and aid library amplification

Whole sequences are stitched together from segment overlaps



MPS using conventional forensic PCR

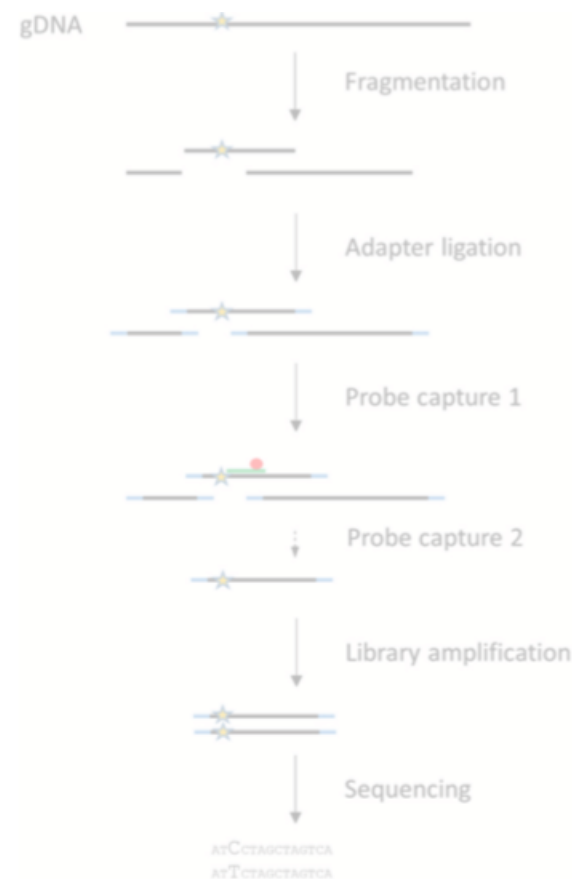


Hybridisation capture methods

DNA fragments with ligated adaptors are captured by probe hybridisation (often termed 'baits')

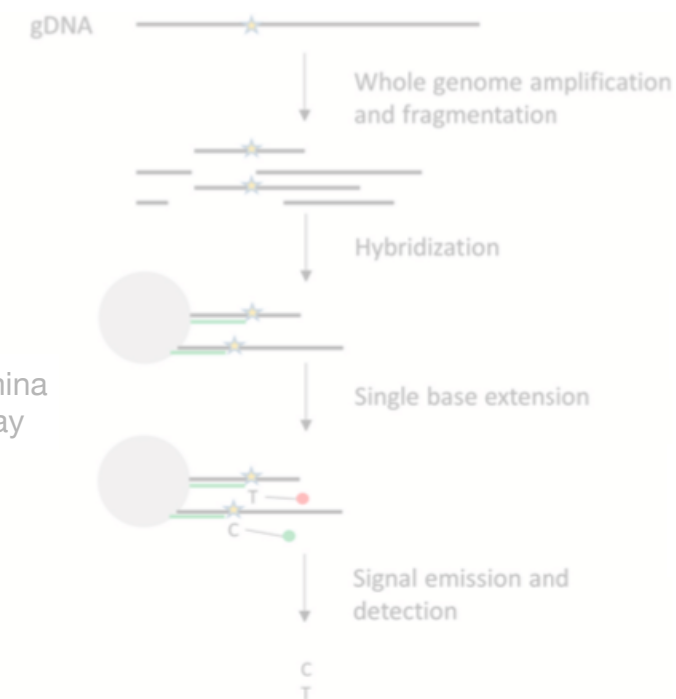
Probes removed and the enriched DNA pool provides efficient library amplification

Sequencing can use established platforms at varied scales: Illumina NextSeq/ NovaSeq, Thermo Fisher Ion GeneStudio



Whole-genome SNP arrays

e.g. Illumina BeadArray



SBE principle is identical to SNaPshot

Tend to use Cy5-Cy3 dyes so only binary SNPs detected

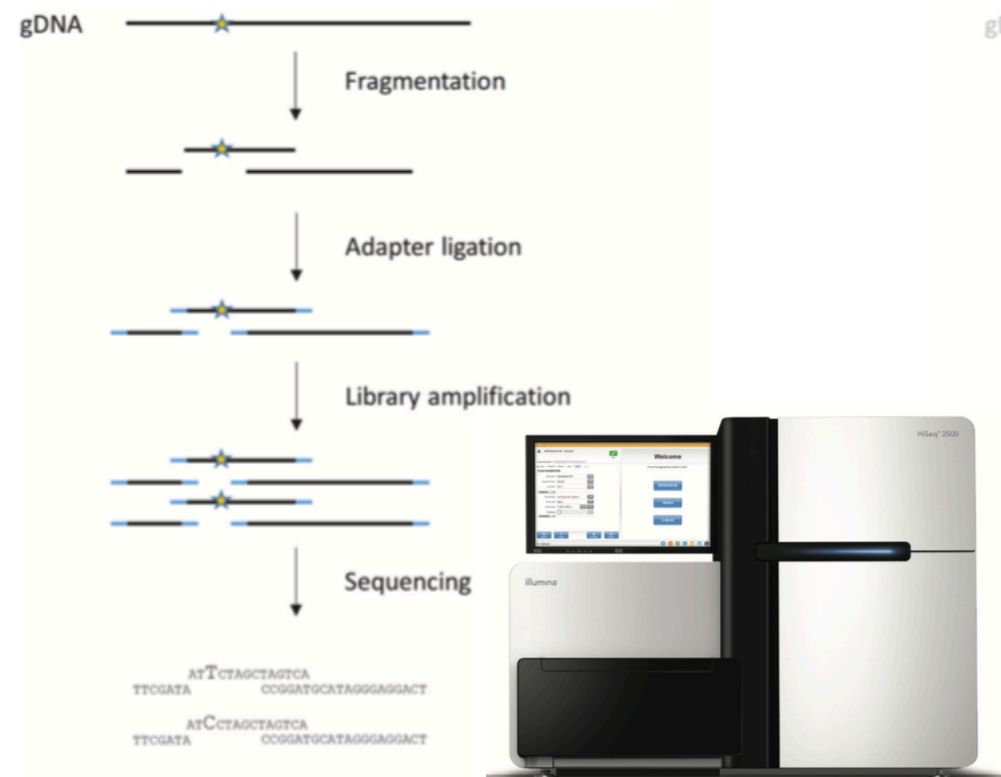
Massively Parallel Sequencing

Whole-genome sequencing

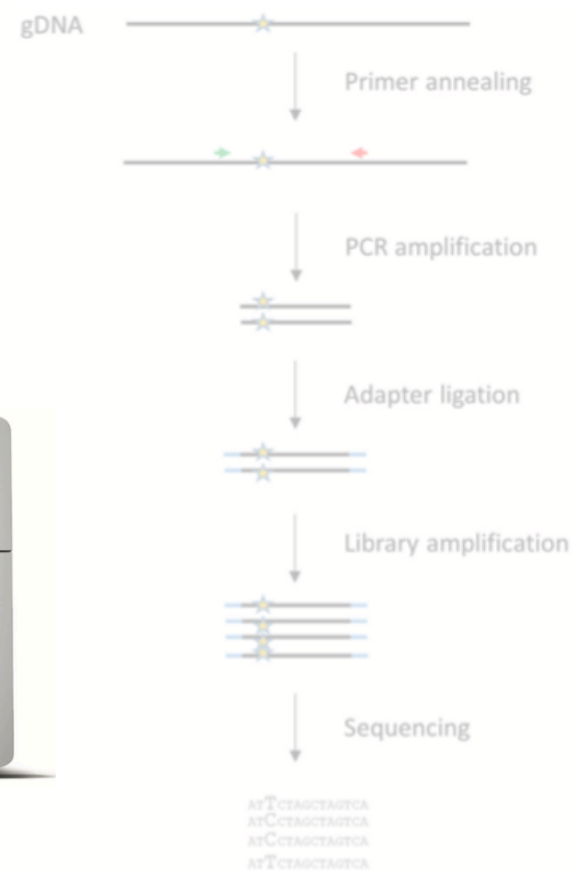
Sonication creates a pool of random short fragments

Adaptors are ligated that identify each sample and aid library amplification

Whole sequences are stitched together from segment overlaps



MPS using conventional forensic PCR



Hybridisation capture methods

DNA fragments with ligated adaptors are captured by probe hybridisation (often termed 'baits')

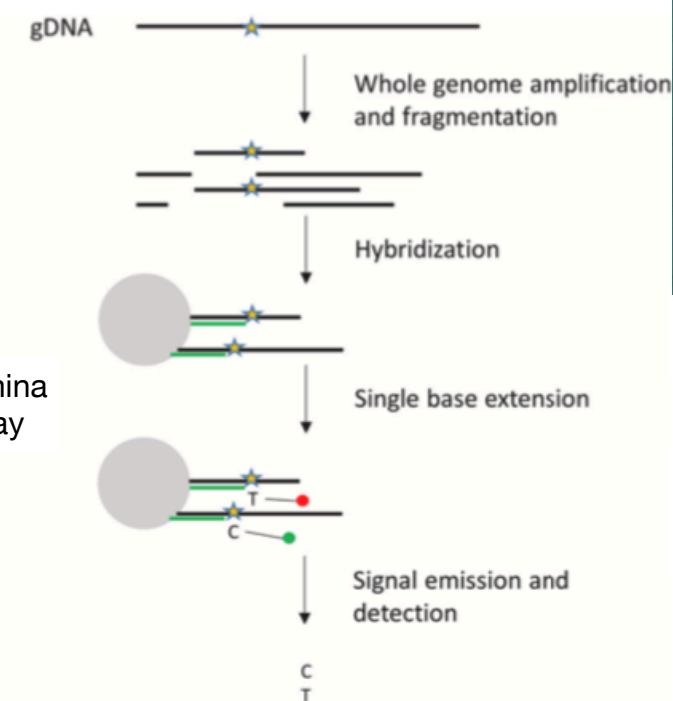
Probes removed and the enriched DNA pool provides efficient library amplification

Sequencing can use established platforms at varied scales: Illumina NextSeq/ NovaSeq, Thermo Fisher Ion GeneStudio



Whole-genome SNP arrays

e.g. Illumina BeadArray



SBE principle is identical to SNaPshot

Tend to use Cy5-Cy3 dyes so only binary SNPs detected

Current genealogy mainstays for SNPs

Whole-genome sequencing

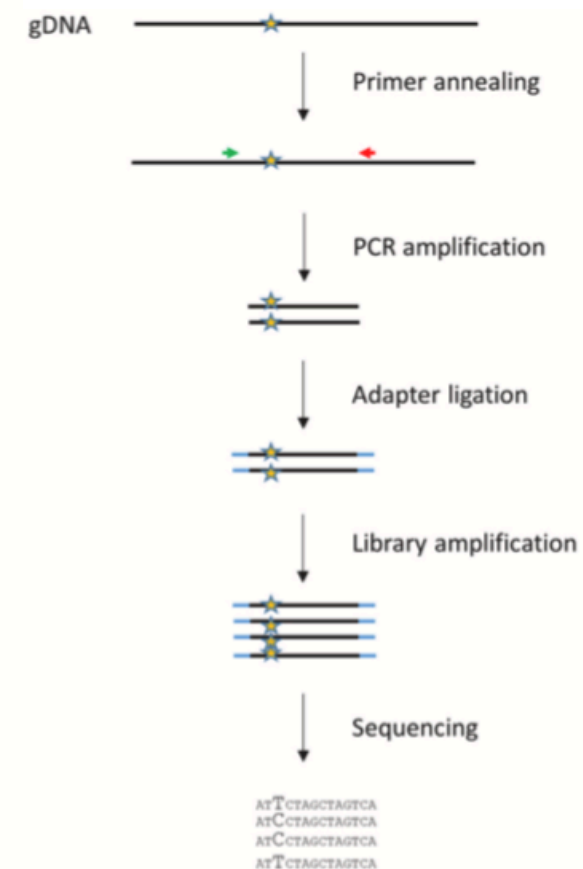
Sonication creates a pool of random short fragments

Adaptors are ligated that identify each sample and aid library amplification

Whole sequences are stitched together from segment overlaps



MPS using conventional forensic PCR

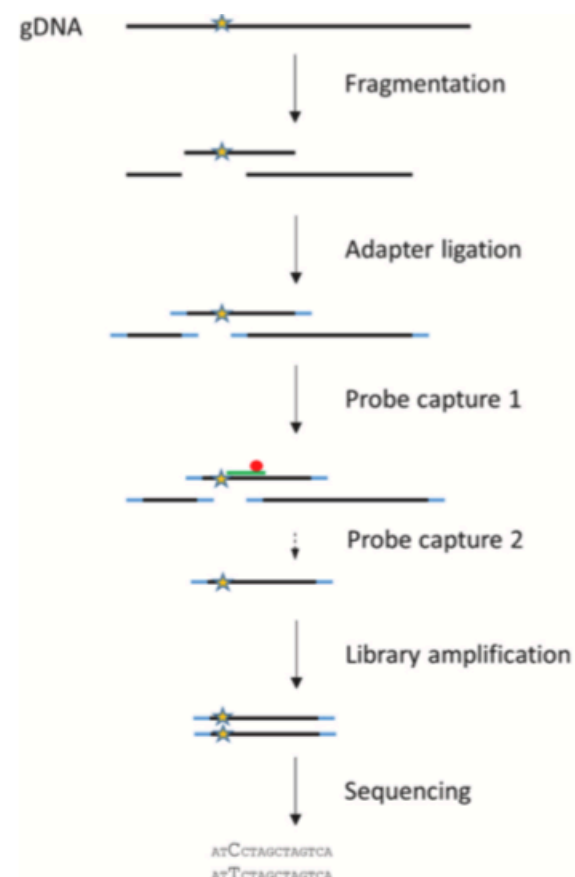


Hybridisation capture methods

DNA fragments with ligated adaptors are captured by probe hybridisation (often termed 'baits')

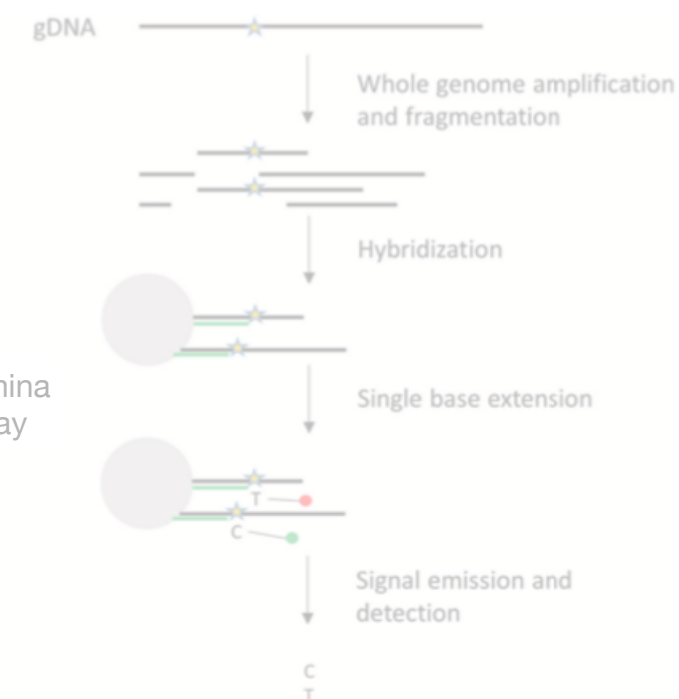
Probes removed and the enriched DNA pool provides efficient library amplification

Sequencing can use established platforms at varied scales: Illumina NextSeq/ NovaSeq, Thermo Fisher Ion GeneStudio



Whole-genome SNP arrays

e.g. Illumina BeadArray





SBE principle is identical to SNaPshot

Tend to use Cy5-Cy3 dyes so only binary SNPs detected

Medium-scale SNP typing - kinship *not* IGG

Article

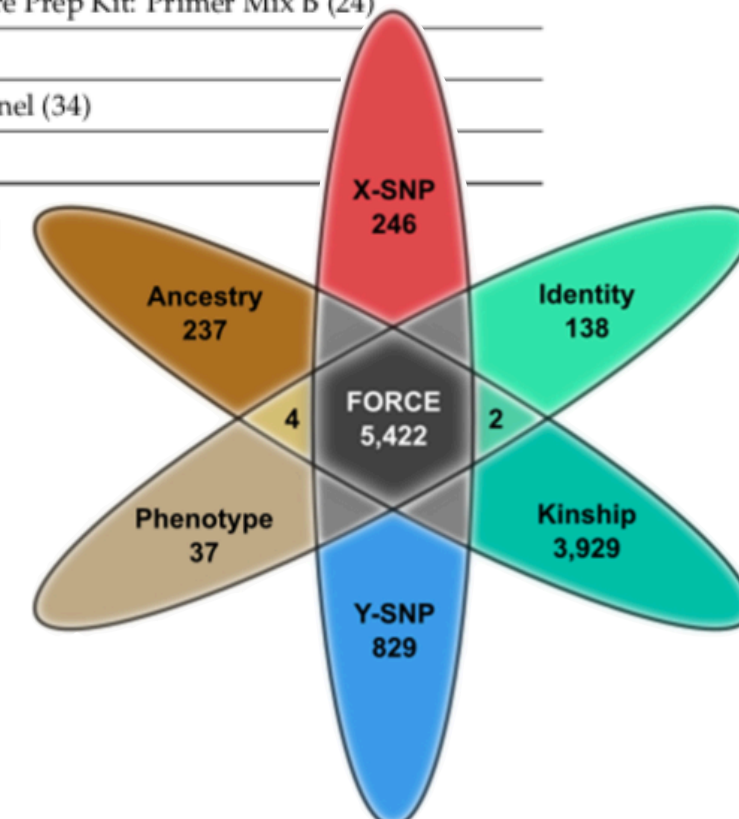
The FORCE Panel: An All-in-One SNP Marker Set for Confirming Investigative Genetic Genealogy Leads and for General Forensic Applications

Andreas Tillmar^{1,2,*}, Kimberly Sturk-Andreaggi^{3,4,5} , Jennifer Daniels-Higginbotham^{3,4}, Jacqueline Tyler Thomas^{3,4} and Charla Marshall^{3,4,6,*} 

Marker Type	SNP Panels (# SNPs)
Kinship SNP/ X-SNP	Infinium Global Screening Array (654,027), Infinium Omni Express (710,000), Infinium CytoSNP-850K (850,000)
iiSNP	ForenSeq DNA Signature Prep Kit: Primer Mix A (94) Precision ID Identity Panel (90) QIAseq Investigator 140 SNP panel (140)
aiSNP	ForenSeq DNA Signature Prep Kit: Primer Mix B (56) Precision ID Ancestry Panel (165) VISAGE panel (115)
piSNP	ForenSeq DNA Signature Prep Kit: Primer Mix B (24) VISAGE panel (41)
Y-SNP	Precision ID Identity Panel (34) AmpliSeq (884)

5422 SNPs in total

The FORCE panel



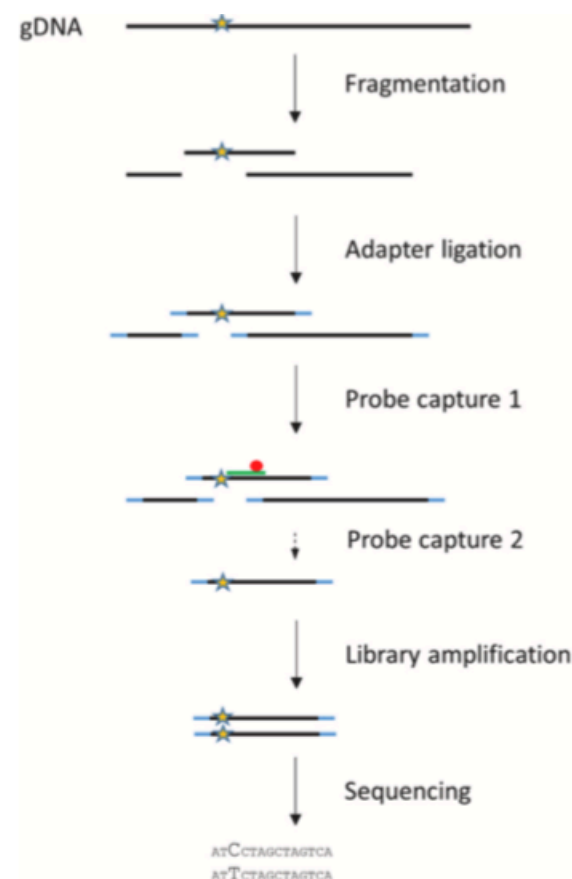
Case	WWII Context	Sample	Sample Type	Relationship to Service Member (DOR)
A	USS Oklahoma	1	Buccal swab	Grandniece (3)
		2	Buccal swab	Nephew (2)
		3	Left femur	Self
B	USS Oklahoma	4	Buccal swab	Nephew (2)
		5	Buccal swab	Great grandnephew (4)
		6	Buccal swab	Great grandniece (4)
		7.1	Left femur	Self
		7.2		
C	Austria	8	Buccal swab	First cousin twice removed, male (5)
		9	Buccal swab	Nephew (2)
		10	Buccal swab	Daughter (1)
		11.1	Long bone	Self
		11.2		
D	Italy	12	Buccal swab	Sister (1)
		13	Buccal swab	Grandniece (3)
		14.1	Right parietal	Self
		14.2		
		14.3		
E	Tarawa	15	Buccal swab	Son (1)
		16	Buccal swab	Daughter (1)
		17.1	Right tibia	Self
		17.2		

Hybridisation capture methods

DNA fragments with ligated adaptors are captured by probe hybridisation (often termed 'baits')



Probes removed and the enriched DNA pool provides efficient library amplification

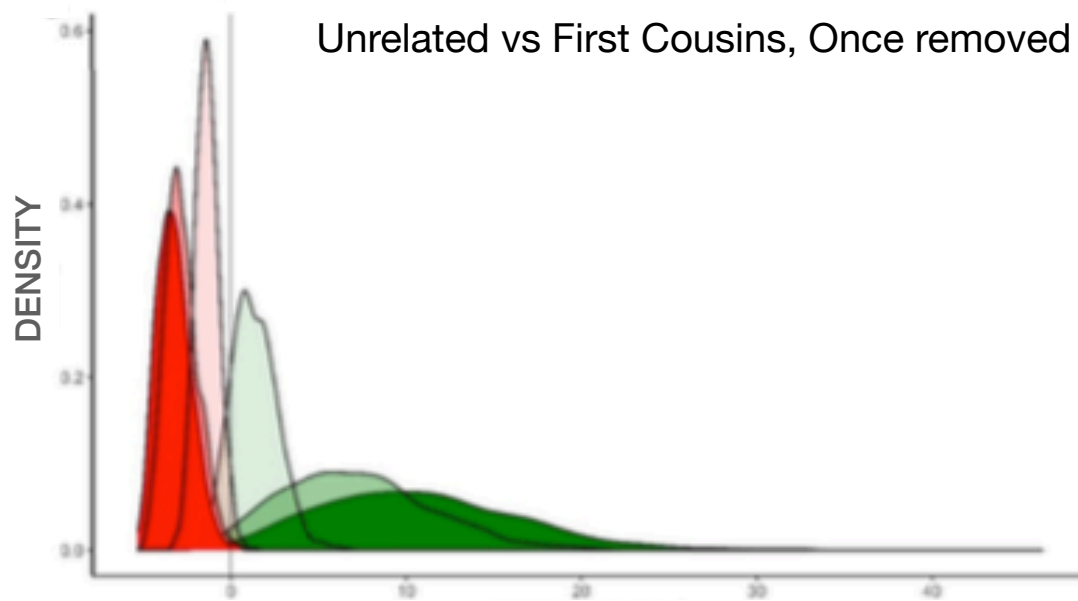
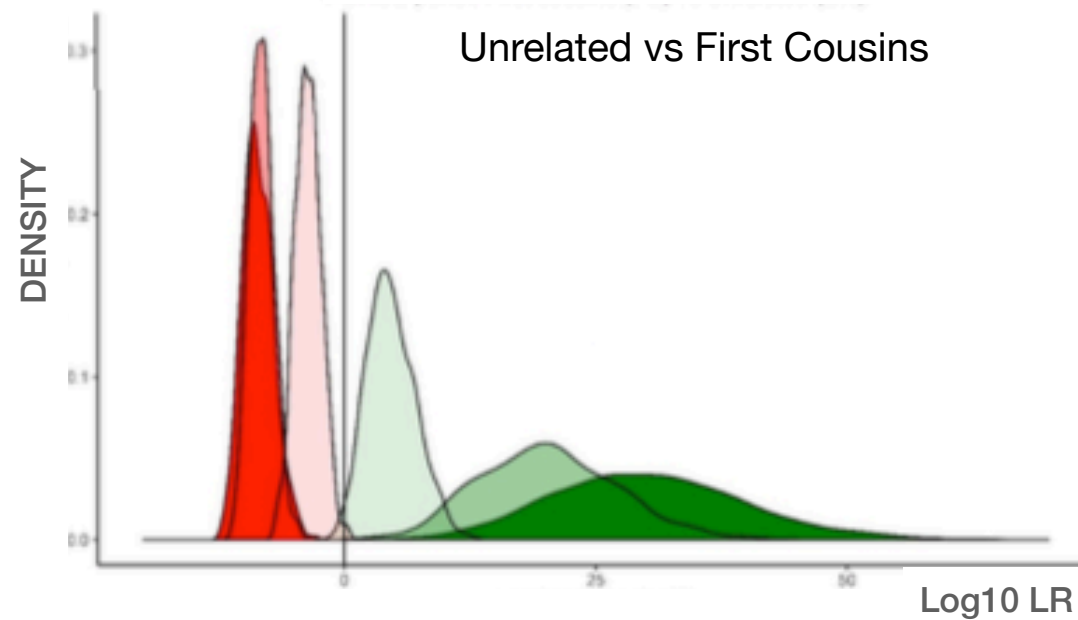
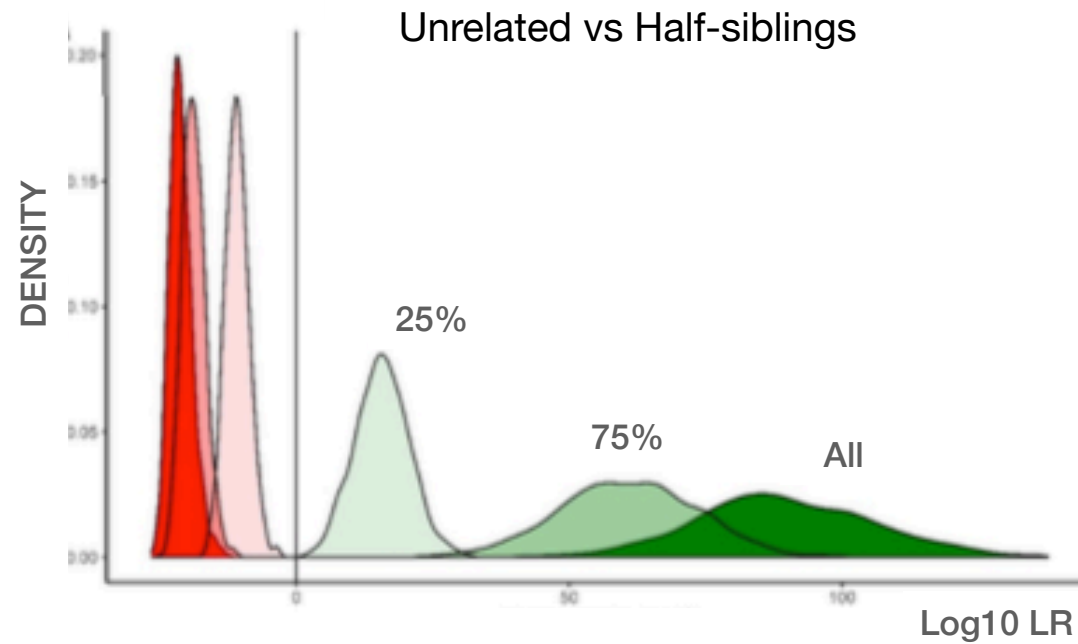
Sequencing can use established platforms at varied scales: Illumina NextSeq/ NovaSeq, Thermo Fisher Ion GeneStudio



Article

The FORCE Panel: An All-in-One SNP Marker Set for Confirming Investigative Genetic Genealogy Leads and for General Forensic Applications

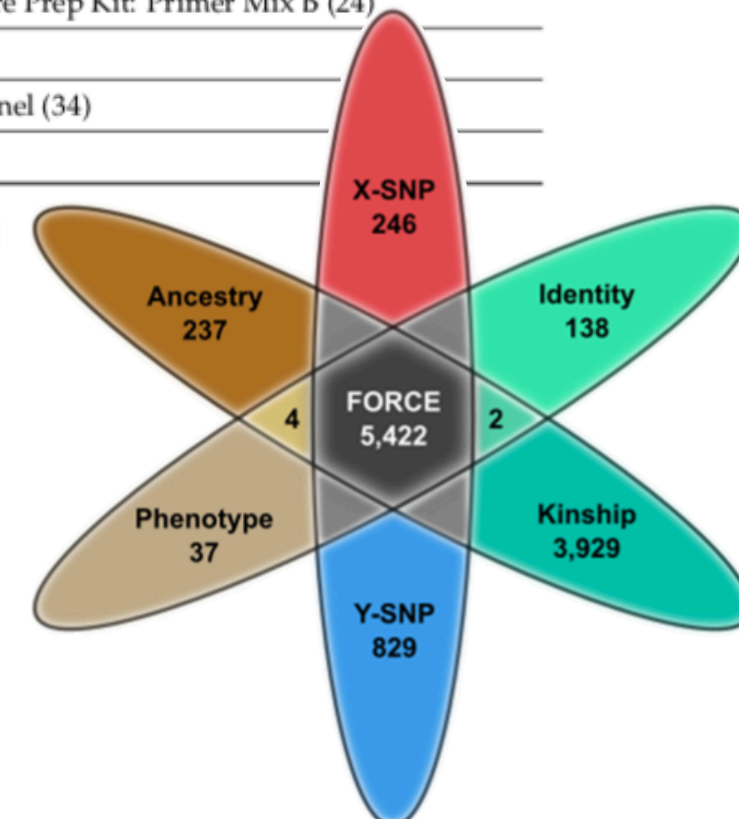
Andreas Tillmar^{1,2,*}, Kimberly Sturk-Andreaggi^{3,4,5} , Jennifer Daniels-Higginbotham^{3,4},
Jacqueline Tyler Thomas^{3,4} and Charla Marshall^{3,4,6,*} 



Marker Type	SNP Panels (# SNPs)
Kinship SNP/ X-SNP	Infinium Global Screening Array (654,027), Infinium Omni Express (710,000), Infinium CytoSNP-850K (850,000)
iiSNP	ForenSeq DNA Signature Prep Kit: Primer Mix A (94) Precision ID Identity Panel (90) QIAseq Investigator 140 SNP panel (140)
aiSNP	ForenSeq DNA Signature Prep Kit: Primer Mix B (56) Precision ID Ancestry Panel (165) VISAGE panel (115)
piSNP	ForenSeq DNA Signature Prep Kit: Primer Mix B (24) VISAGE panel (41)
Y-SNP	Precision ID Identity Panel (34) AmpliSeq (884)

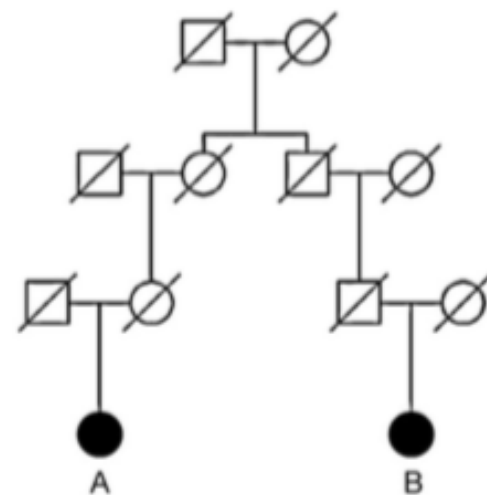
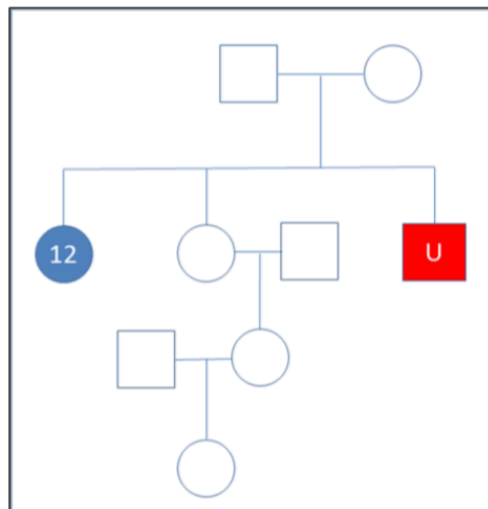
5422 SNPs in total

The FORCE panel



IGG looks beyond the pairwise comparisons of kinship tests

Kinship testing



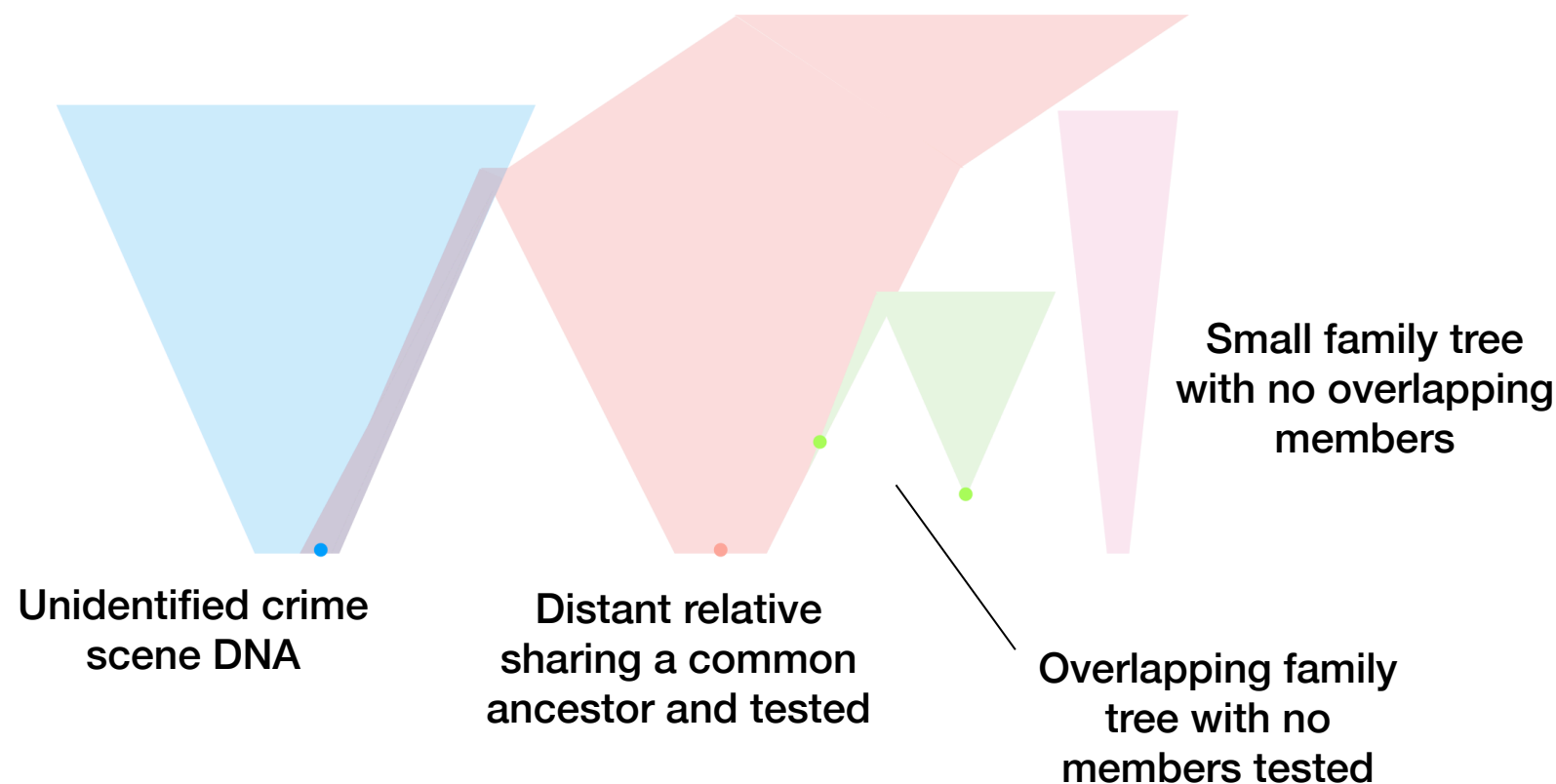
LR $\frac{H1: \text{related as claimed}}{H2: \text{unrelated}}$

STR-based familial searching

- 'near matching' one unknown to one or two first degree relatives

SNP-based (long range) familial searching

- matching one unknown to many potential relatives



Small family tree with no overlapping members

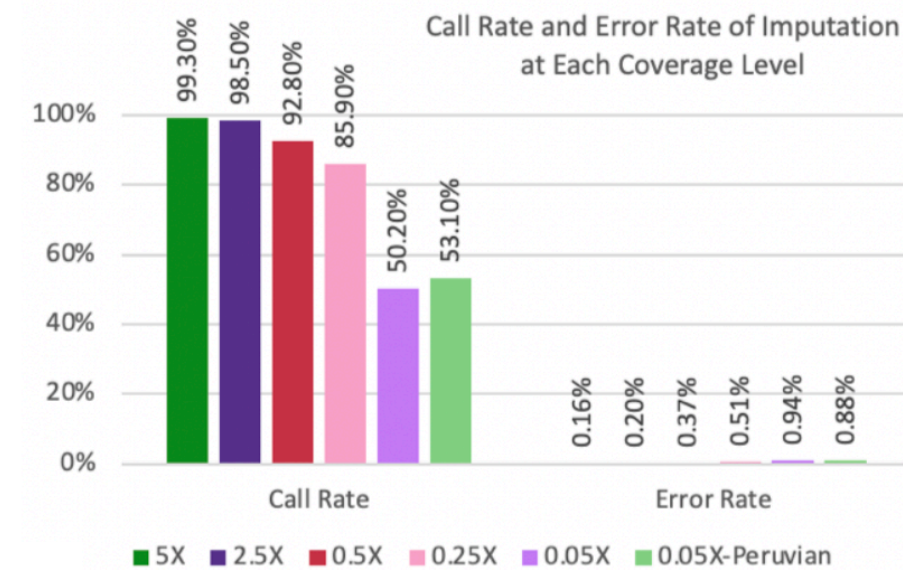
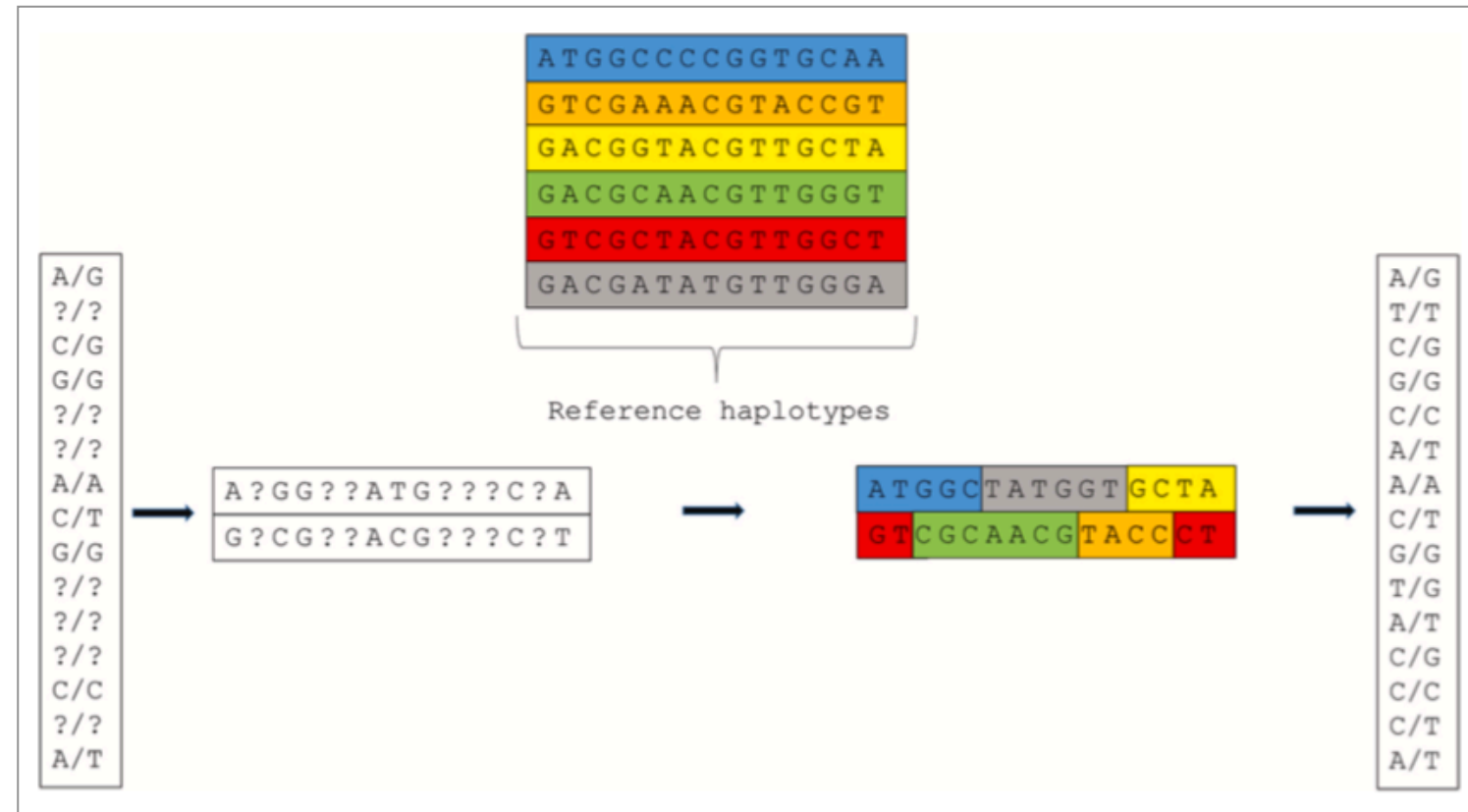
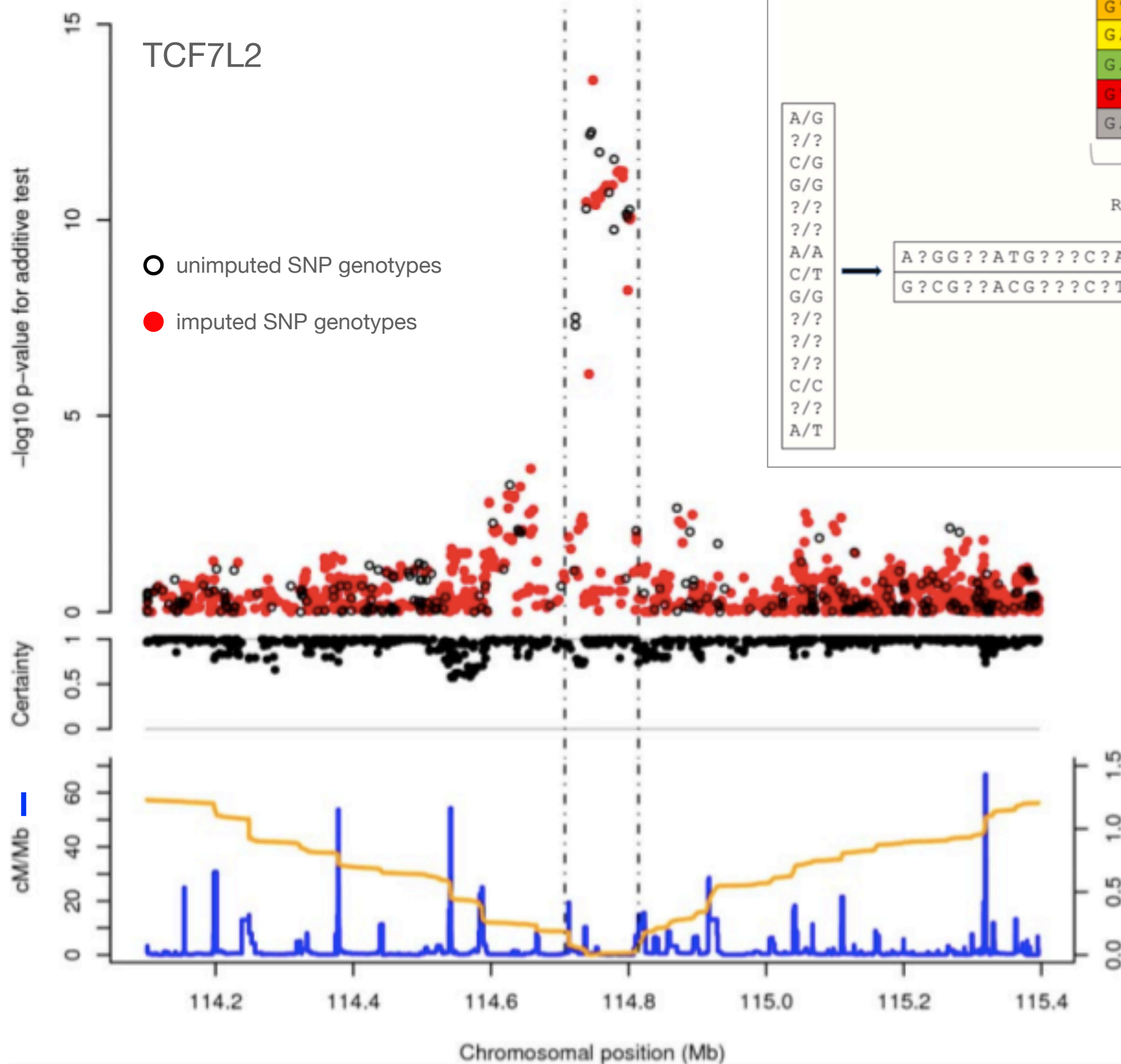
Only 12/50 US states allow CODIS familial searching

Severe sexual assault kit backlog

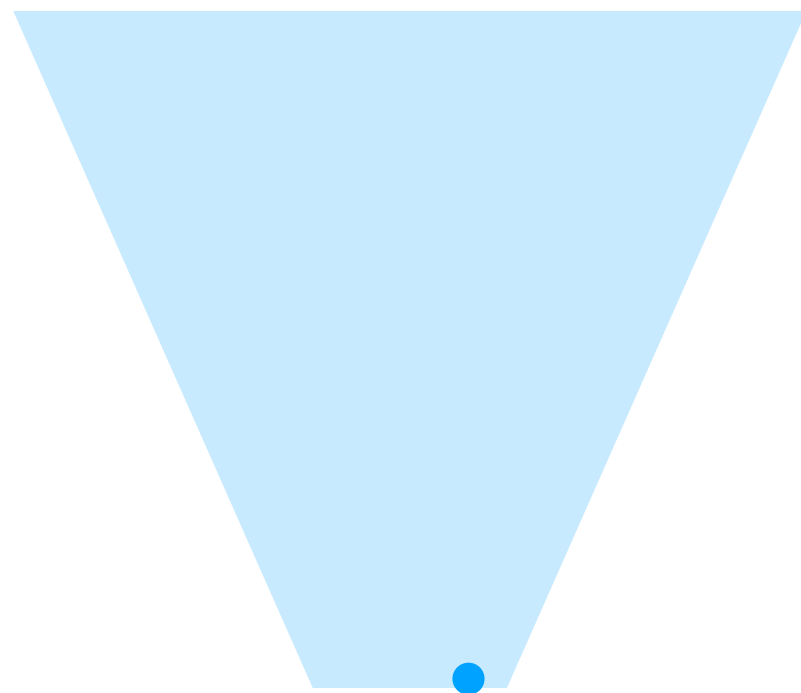
Principle of imputation assumes no recombination in closely sited SNPs - so missing genotypes are inferred from haplotype patterns

TCF7L2

○ unimputed SNP genotypes
● imputed SNP genotypes



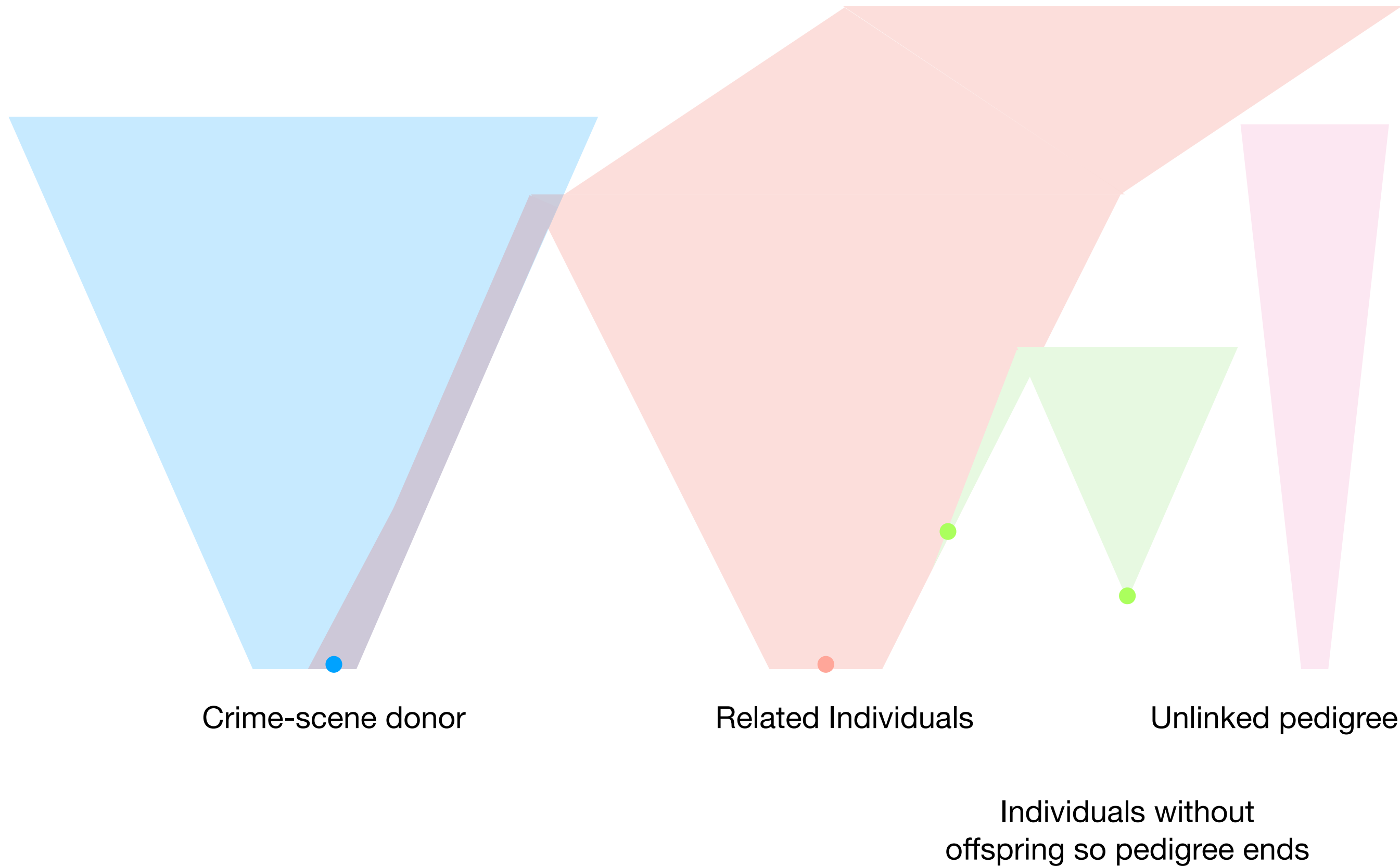
STR-based familial searches finds a group of close relatives



Crime-scene donor
(not in a DNA database)

5-10 generations

Most pedigrees are unlinked until they go back > 10-20 generations



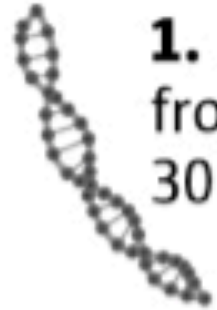
Identifying the Golden State Killer - segment analysis

Genotyping >500,000 SNPs



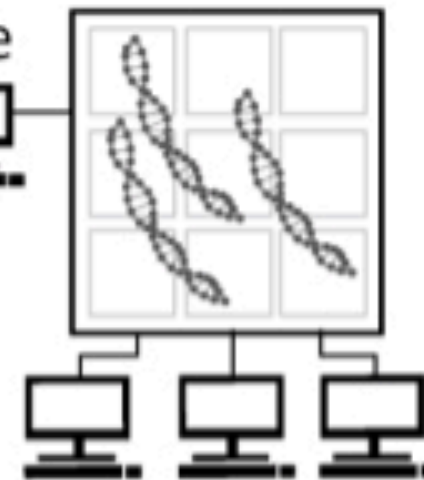
[GED
match]

One-to-many segment analysis



1. DNA was collected from the crime scene 30 years ago.

2. Crime-scene DNA run through GEDmatch data bank.



3. Results indicate shared DNA with two second cousins.



Duplicate ♀ SAK swab stored frozen

All commercial testers have adapted Illumina SNP arrays

GEDmatch compares data in the intersect

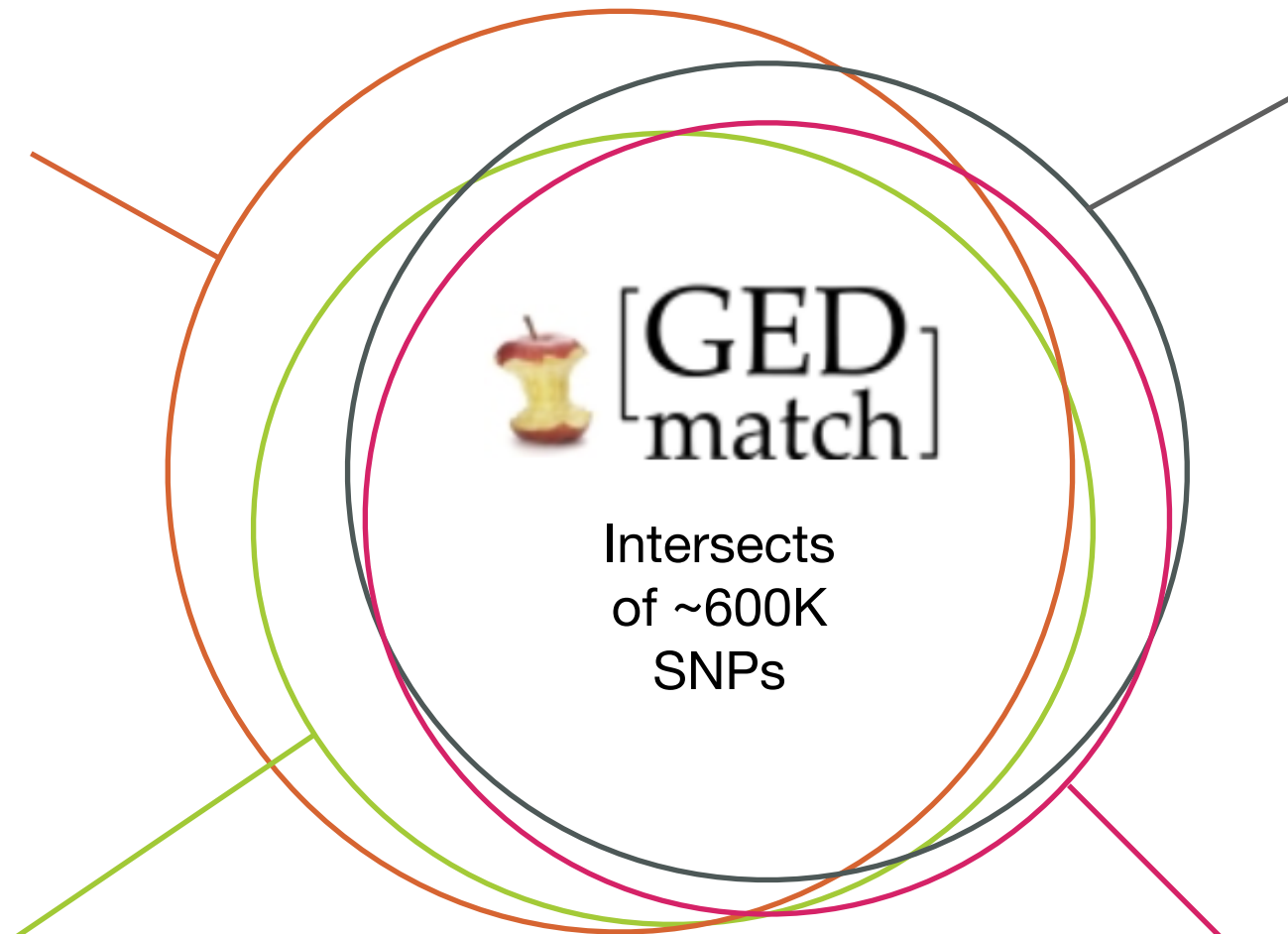


M -
A 702,442
X 17,889
Y 482

M 179
A 612,272
X 16,271
Y -

M 263
A 637,639
X 28,892
Y 1,691

M 4,318
A 630,132
X 16,530
Y 3,733





GEDmatch® One-to-Many

Useful YouTube video on how to use One to Many. [WATCH VIDEO](#) ▶

One-to-many DNA comparison for: Kit

Note: Additional facilities in Tier 1 One-to-many version.

Filter By:

- ☒ Autosomal
☐ X

With this offset

With this limit

[Prev](#) [Next](#)

cM size

Tag Groups

- ☒ None
☐ All
☐ One

Overlap Cutoff

SUBMIT

Tips ⓘ

Select all

GEDmatch is an open database

NEW MATCH NOTIFICATIONS

NEW FEATURE

2021

We are pleased to announce a new feature at GEDmatch: New Match Notification emails!

If you are a free user, you can receive weekly email notifications of new matches that are 100 cM or greater.

If you are a Tier 1 user, you can receive weekly OR daily email notifications, and can choose 30 cM, 70 cM, 100cM or 200 cM as the threshold for new match notifications.

VISUALIZATION OPTIONS									Haplogroup		Autosomal			X-DNA		
Select ▾	Kit ▾	Name (* => alias) ▾	Email ▾	GED WikiTree ▾	Age(days) ▾	Type ▾	Sex ▾		Mt ▾	Y ▾	Total cM ▾	Largest ▾	Gen ▾	Total cM ▾	Largest ▾	Source ▾
<input type="checkbox"/>	A492222	*Zobiana	zobiana@hotmail.co.uk		1803	2	F				235.7	48.2 Q	2.96	0	0	Migration - F2 - A
<input type="checkbox"/>	JQ8223445	Kate Wilson	macfearsome@gmail.com		142	2	F				51.9	12.5 Q	4.05	0	0	FTDNA
<input type="checkbox"/>	MX9810637	Simon Eriksson	simon.frewins@gmail.com		401	2	M				49.5	25.2 Q	4.09	0	0	23andMe
<input type="checkbox"/>	M714230	Jacqueline Gordier	Jacquigordier@hotmail.com		1744	2	F	H1c3			48.3	38.8 Q	4.11	0	0	Migration - V4 - M
<input type="checkbox"/>	YF5534989	Margaret Stephens	stracey@traceylawfirm.com		339	2	F				48.1	10.4 Q	4.11	0	0	-
<input type="checkbox"/>	RJ5853300	Justin Kincaid	kincaid4christ@yahoo.com		175	2	M		R-M269		43.4	13.3 Q	4.18	0	0	FTDNA
<input type="checkbox"/>	A373001	Joe Harper	jmharper@frontiernet.net		1204	2	M	C1c	M269		38.2	26.5 Q	4.28	0	0	Migration - F2 - A

GEDmatch applies IBD chromosome segment analysis to match closest relatives

Chr	B37 Start Pos'n	B37 End Pos'n	Centimorgans (cM)	SNPs
1	182,392,412	204,081,196	19.9	772

Chr 1

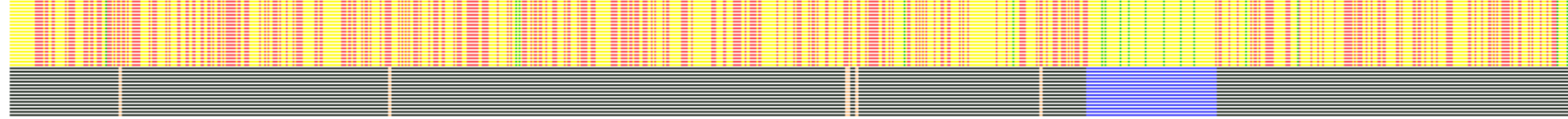
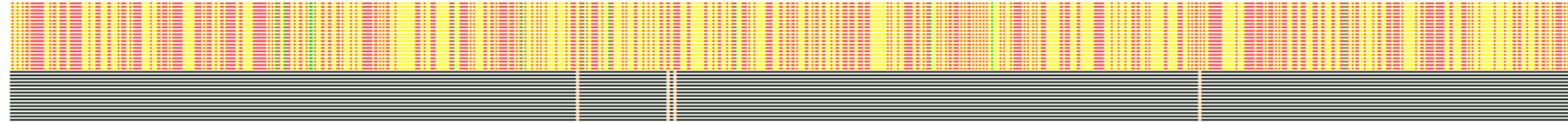


Image size reduction: 1/10

Chr 2



Chr	B37 Start Pos'n	B37 End Pos'n	Centimorgans (cM)	SNPs
3	10,275,475	16,110,666	9	322

Chr 3

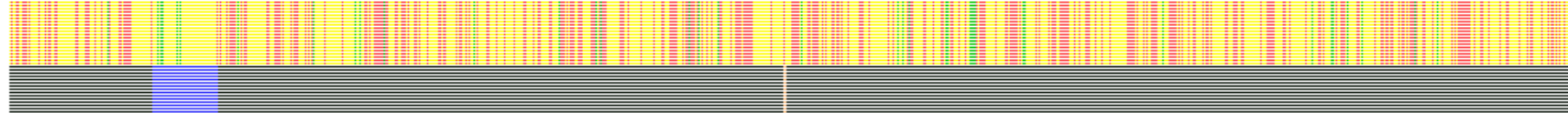
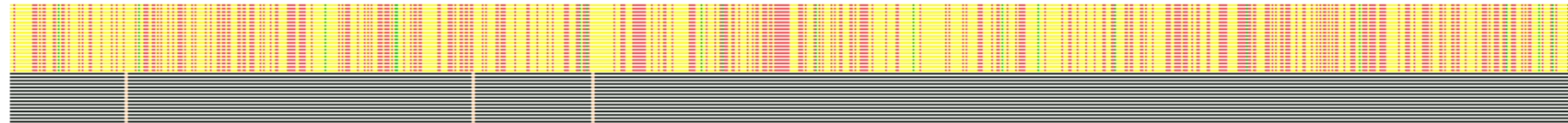


Image size reduction: 1/8

Chr 4



Chr	B37 Start Pos'n	B37 End Pos'n	Centimorgans (cM)	SNPs
5	10,148,758	28,028,745	20.8	636

Chr 5

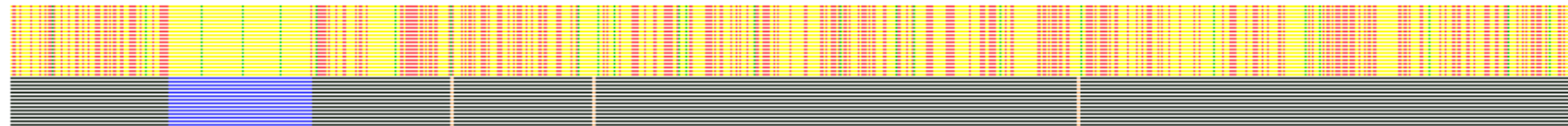


Image size reduction: 1/8

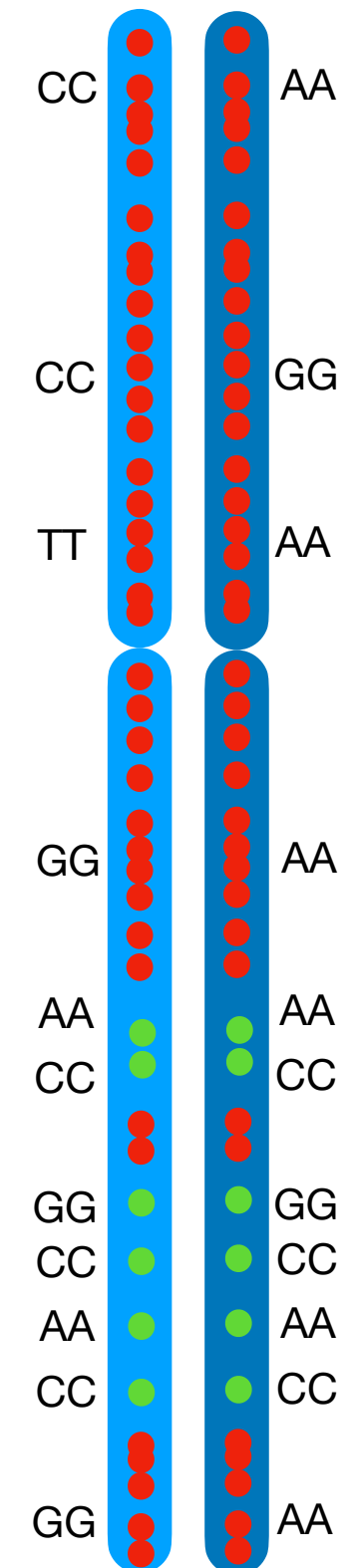
Individual marker indications:

Base Pairs with Full Match	
Base Pairs with Half Match	
Match with Phased data	
Base Pairs with No Match	

Validity of segments:

Significant	
Moderate	
Low	
Insignificant	
Large gap between adjacent SNPs	
No Match	

Defines IBD segments by identifying runs of homozygosity



True phase of a string of forty SNPs (35 give half match)

♂

♀

♂

♀

G A G G A G A G G G C T A G G G A G A A G A A G C C C C G G G G A A G C A A A T

A A G G A G A G A G C T A G G G G G A A G A G G T T C A G G G G C A G C A G A T

A G G G A G A G G G C T G G A G G G A A A A G G T T A A G G G G C A A A A A G T

G G A G A G A G G G C T A G G G A G A A G A A G C C C C G G G G A A G C A A G T

Forensic DNA

Putative relative

Unphased SNP array data output (i.e. it is alphabetic: both AG and GA = AG) in 10-SNP panels or windows

A A G G A G A G A G

C T A G G G A G A A

G A A G C C C A G G

G G A A G C A A A T

G A G G A G A G G G

C T A G G G G G A A

G A G G T T C C G G

G G C A G C A G A T

A G A G A G A G G G

C T A G A G A G A A

A A A G C C A A G G

G G A A A A A A G T

G G G G A G A G G G

C T G G G G G G A A

G A G G T T C C G G

G G C A G C A A G T

1

2

3

4

Phase windows

Inferred phase made for each window (six SNPs incorrectly phased)

G A G G A G A G G G

C T A G G G A G A A

G A A G C C C C G G

G G C A G C A G A T

A A G G A G A G A G

C T A G G G G G A A

G A G G T T C A G G

G G A A G C A A A T

A G G G A G A G G G

C T G G A G G G A A

A A G G T T A A G G

G G C A A A A A G T

G G A G A G A G G G

C T A G G G A G A A

G A A G C C C C G G

G G A A G C A A G T

Two 'seed windows' with full SNP allele agreement are extended to the first opposite-homozygote incompatibilities detected each side

G A G G A G A G G G

C T A G G G A G A A

G A A G C C C C G G

G G C A G C A G A T

A A G G A G A G A G

C T A G G G G G A A

G A G G T T C A G G

G G A A G C A A A T

A G G G A G A G G G

C T G G A G G G A A

A A G G T T A A G G

G G C A A A A A G T

G G A G A G A G G G

C T A G G G A G A A

G A A G C C C C G G

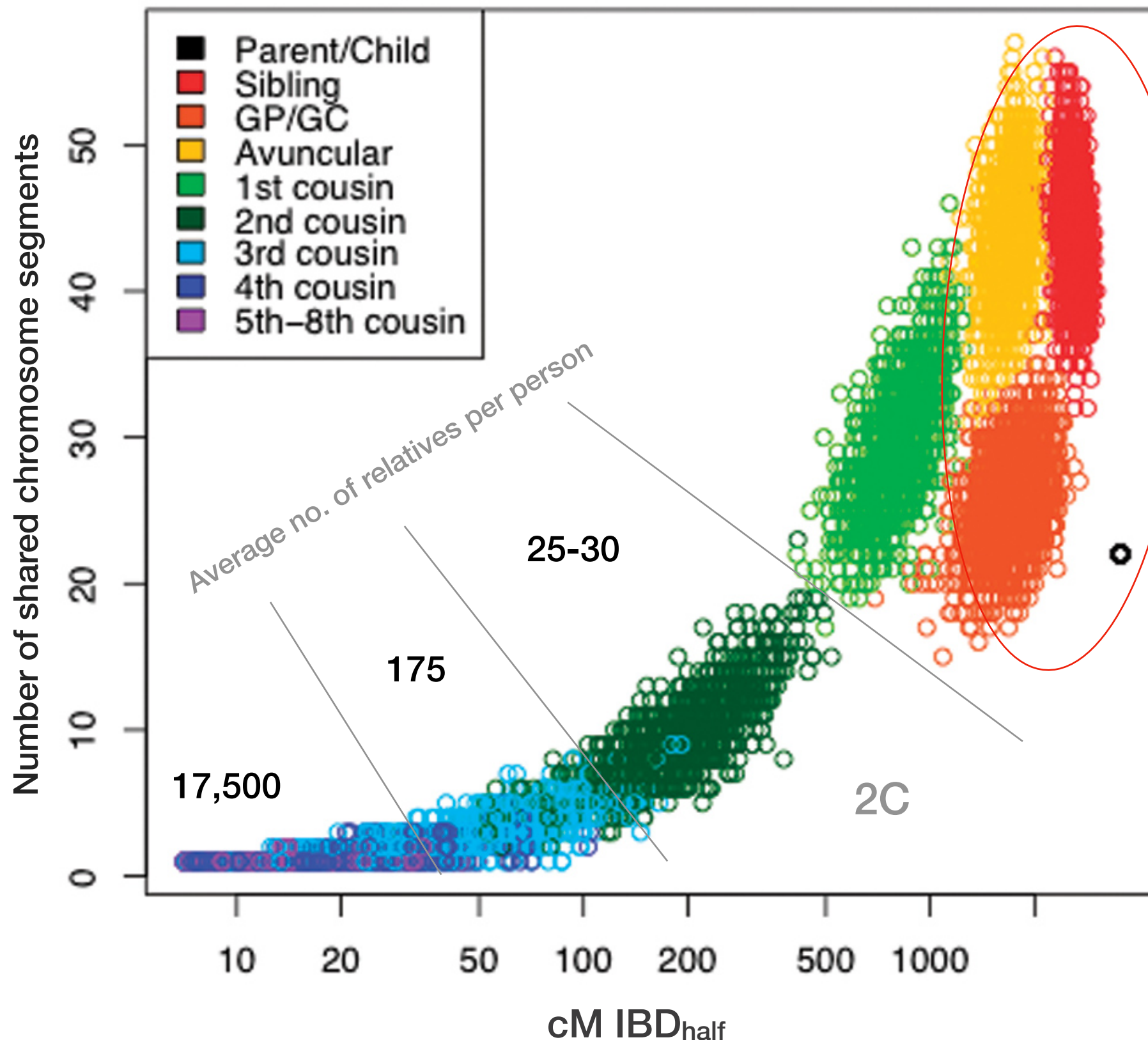
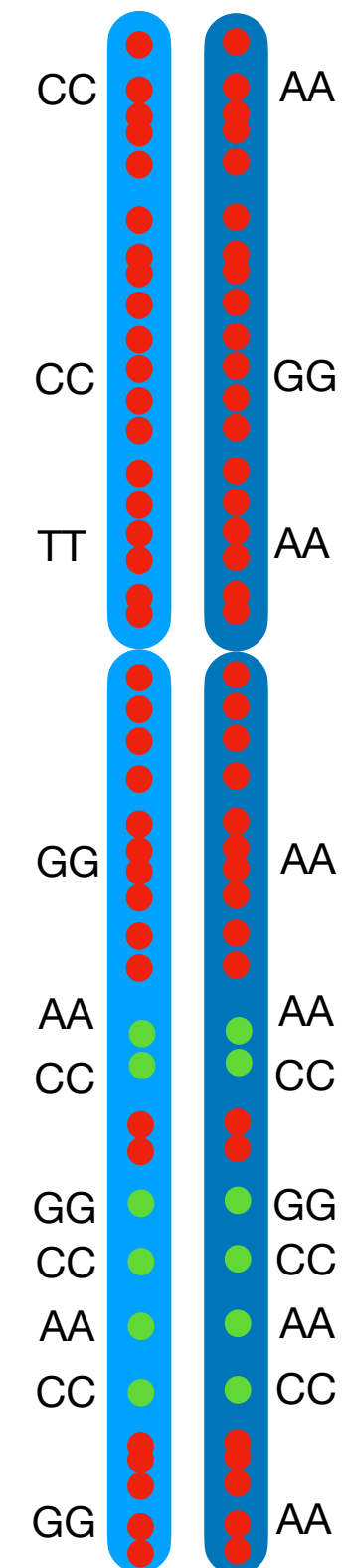
G G A A G C A A G T

5'

3'

Inferred IBD segment match with an undetected SNP allele incompatibility

IBD segments can identify 2nd to 3rd cousins with reasonably good likelihoods, less reliably beyond these relationships



The first crowdsourced genome study was the Shared cM Project

August 2017

Blaine T. Bettinger

www.thegeneticgenealogist.com

[More about this project](#)

[CC 4.0 Attribution License](#)

Interactive version by [Jonny Perl](#) at [DNA Painter](#)

Last updated 7th Oct 2017

[Source for this version](#)

Filter

Enter the total number of cM for your match here:

reset

[show %](#)

How to read this chart

Relationship

Average

Range

(low to high)

(99th percentile)

Then any relationships that fit will stand out below

[Click here for a sharable link to the cM amount above](#)

Half GG-Aunt/Uncle 187 12 – 383	Great-Grandparent 881 464 – 1486						Great-Great Aunt/Uncle 427 191 – 885	1C3R 123 0 – 283	2C3R 57 0 – 139	Other Relationships	
Half 1C2R 145 37 – 360	Half-Great Aunt / Uncle 432 125 – 765	Grandparent 1766 1156 – 2311				Great Aunt/Uncle 914 251 – 2108	1C2R 229 43 – 531	2C2R 74 0 – 261	3C2R 35 0 – 116	6C 21 0 – 86	
Half 2C1R 73 0 – 341	Half 1C1R 226 57 – 530	Half Aunt/Uncle 891 500 – 1446		Parent 3487 3330 – 3720	Aunt/Uncle 1750 1349 – 2175	1C1R 439 141 – 851	2C1R 123 0 – 316	3C1R 48 0 – 173	4C1R 28 0 – 117	6C1R 16 0 – 72	
Half 3C 61 0 – 178	Half 2C 117 9 – 397	Half 1C 457 137 – 856	Half-Sibling 1783 1317 – 2312	Sibling 2629 2209 – 3384	SELF	1C 874 553 – 1225	2C 233 46 – 515	3C 74 0 – 217	4C 35 0 – 127	5C 25 0 – 94	6C2R 17 0 – 75
Half 3C1R 42 0 – 165	Half 2C1R 73 0 – 341	Half 1C1R 226 57 – 530	Half-Niece / Nephew 891 500 – 1446	Niece / Nephew 1750 1349 – 2175	Child 3487 3330 – 3720	1C1R 439 141 – 851	2C1R 123 0 – 316	3C1R 48 0 – 173	4C1R 28 0 – 117	5C1R 21 0 – 79	7C 13 0 – 57
Half 3C2R 34 0 – 96	Half 2C2R 61 0 – 353	Half 1C2R 145 37 – 360	Half Great-Niece / Nephew 432 125 – 765	Great Niece / Nephew 910 251 – 2108	Grandchild 1766 1156 – 2311	1C2R 229 43 – 531	2C2R 74 0 – 261	3C2R 35 0 – 116	4C2R 22 0 – 109	5C2R 17 0 – 43	7C1R 13 0 – 53
Half 3C3R	Half 2C3R	Half 1C3R 87 0 – 191	Half GG Niece / Nephew 187 12 – 383	Great Great Niece / Nephew 427 191 – 885	Great-Grandchild 881 464 – 1486	1C3R 123 0 – 283	2C3R 57 0 – 139	3C3R 22 0 – 69	4C3R 29 0 – 82	5C3R 11 0 – 44	8C 12 0 – 50

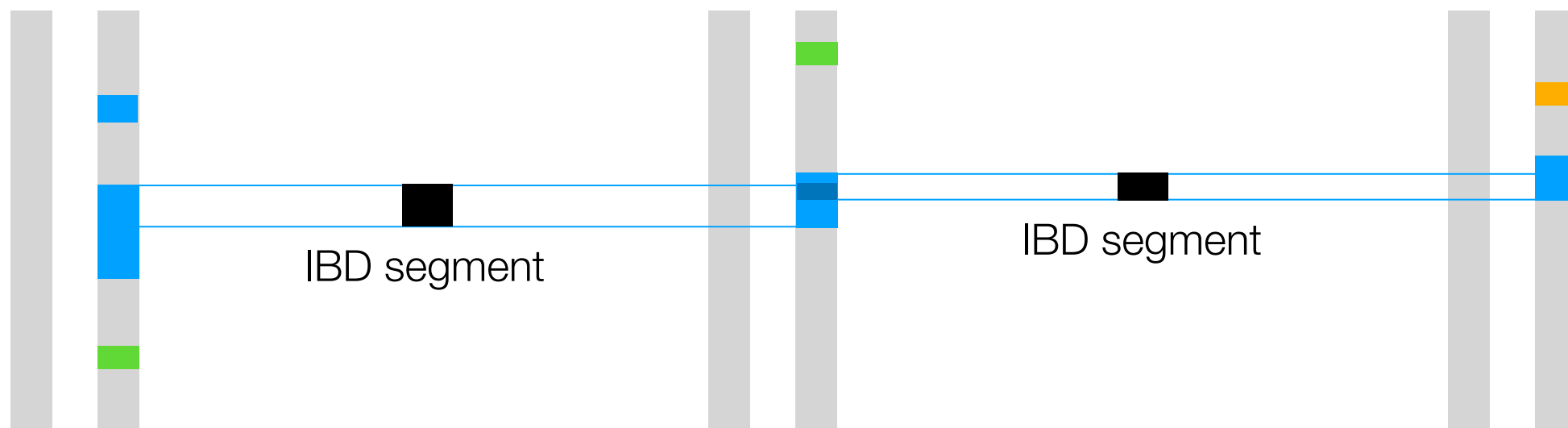
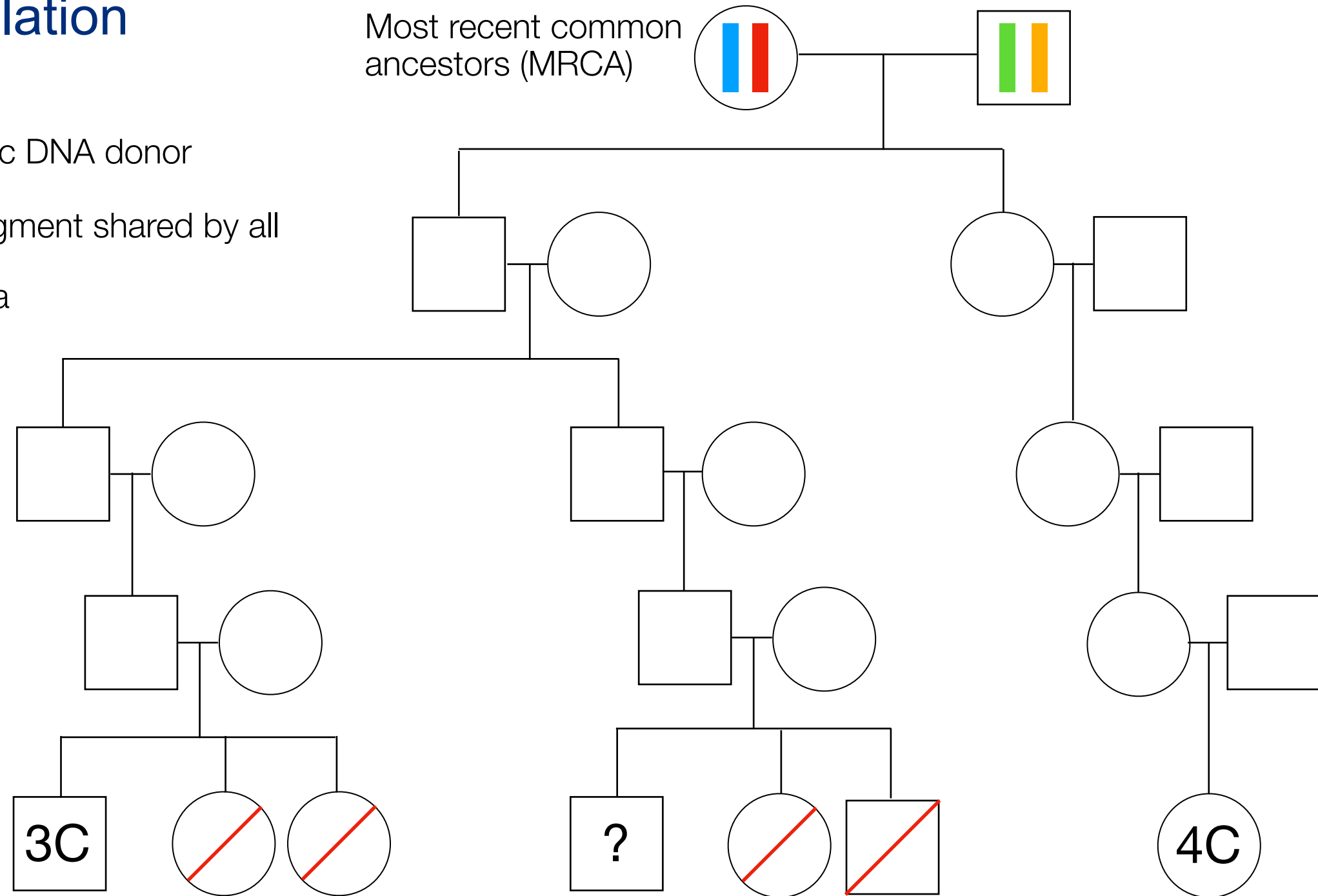
Triangulation

? Forensic DNA donor

■ IBD segment shared by all

— No data

Most recent common
ancestors (MRCA)



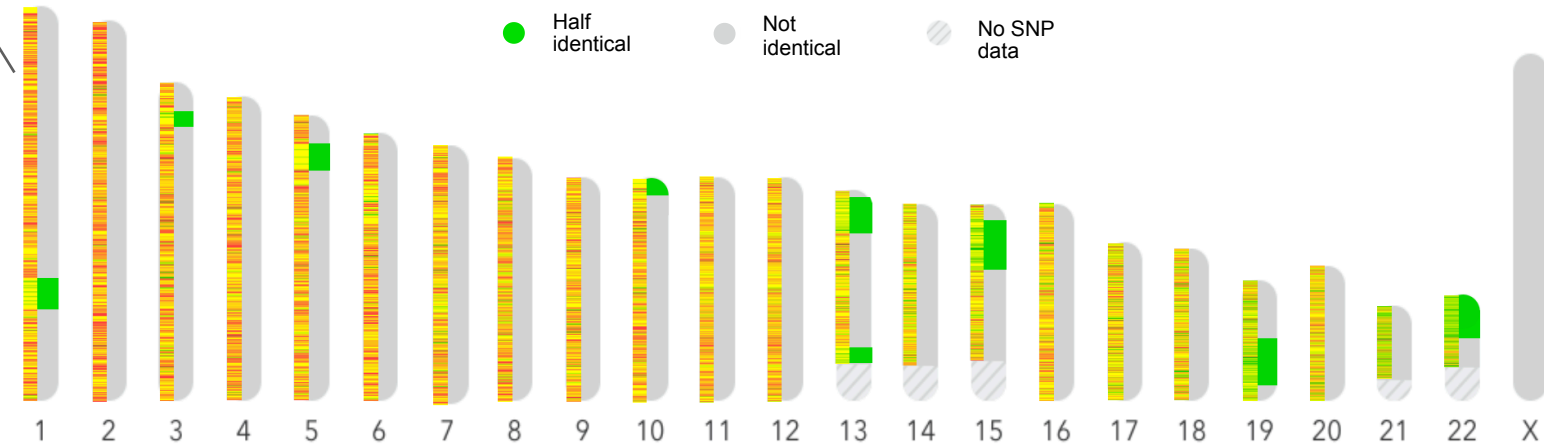
Exploring segments - two different matches in different databases can initiate triangulation studies



ZO

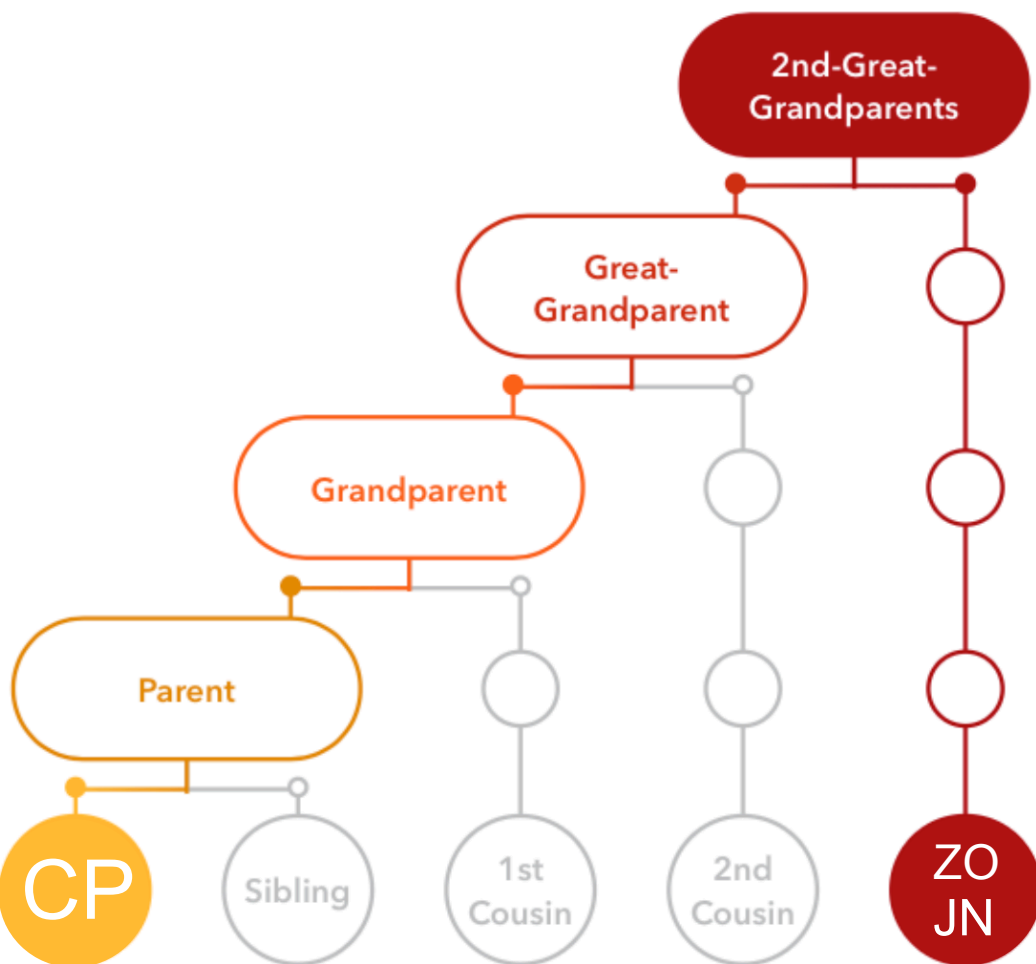
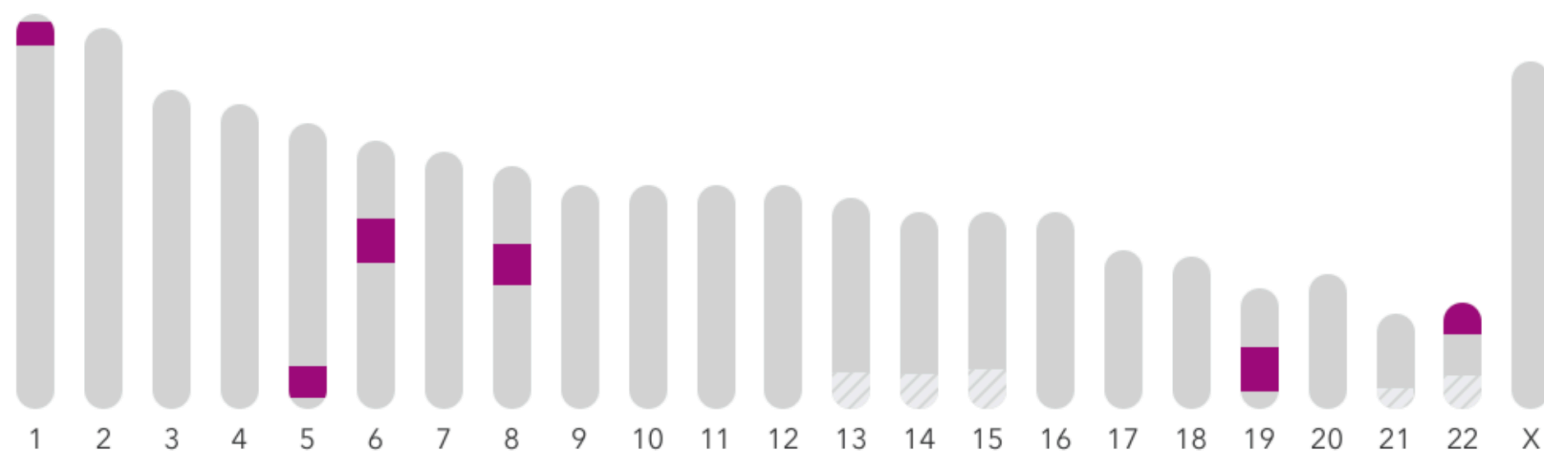
- Half-match (2 phased chromosomes)
- Full match (4 chromosomes)
- No match

- Half identical
- Not identical
- No SNP data



JN

- Half identical
- Not identical
- No SNP data



Multiple individuals with matching segments allows triangulation

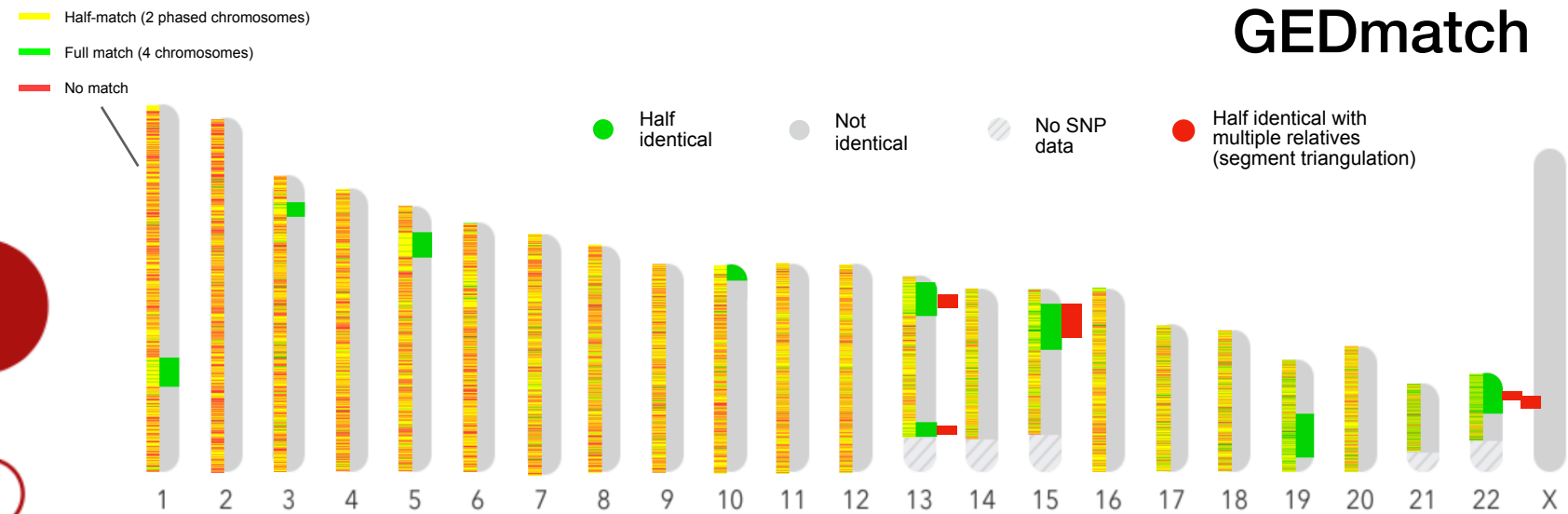
Segment matching

Segment triangulation

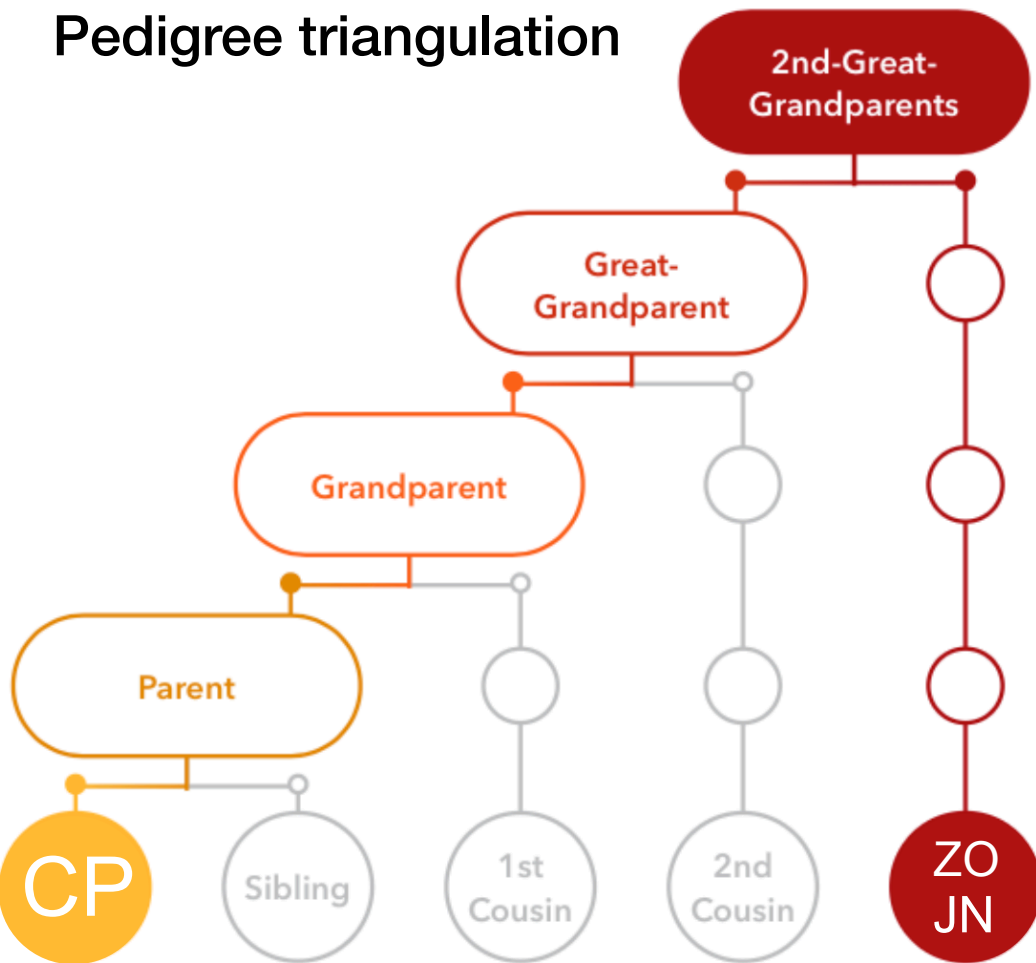
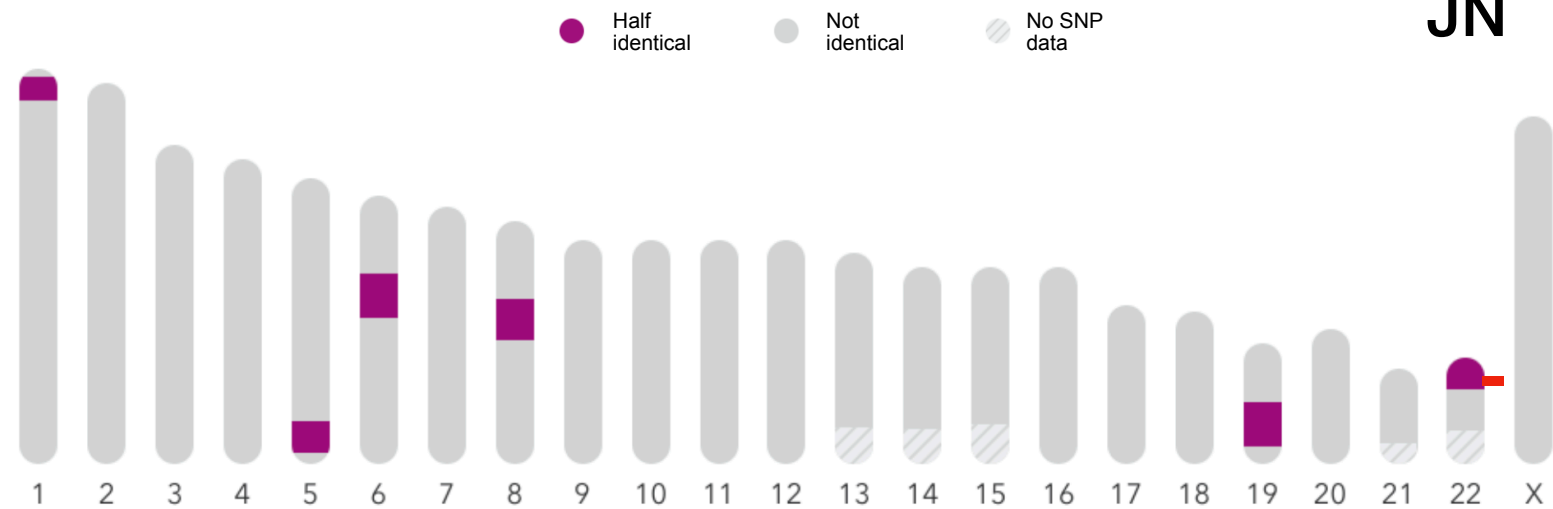
Pedigree triangulation



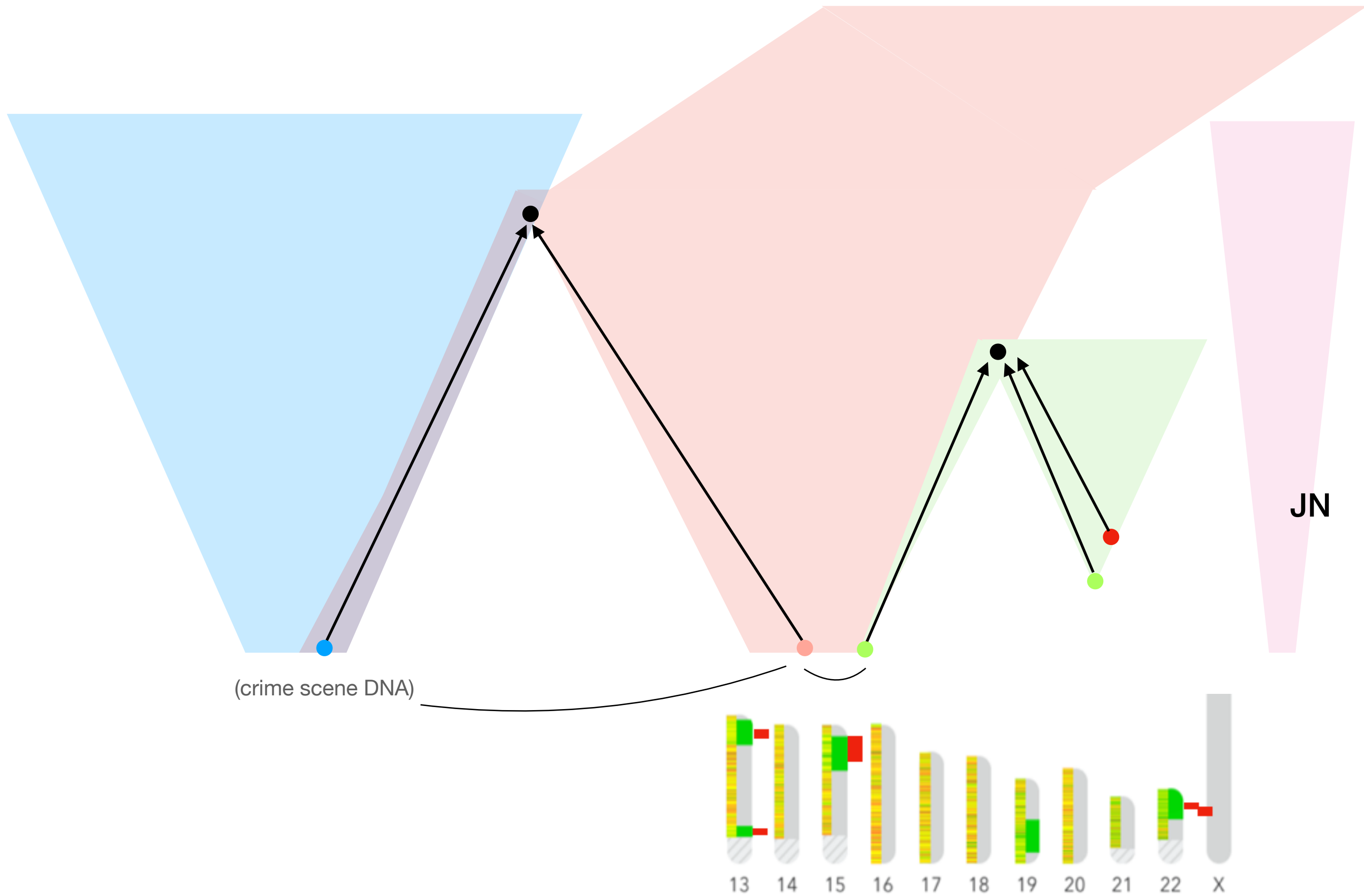
ZO+others in
GEDmatch



JN



Segment triangulation indicates shared ancestors amongst kinships





Whole-genome sequencing of human remains to enable genealogy DNA database searches – A case report

Andreas Tillmar^{a,b,*}, Peter Sjölund^c, Bo Lundqvist^d, Therese Klippmark^a, Cajsa Älgenäs^a, Henrik Green^{a,e}

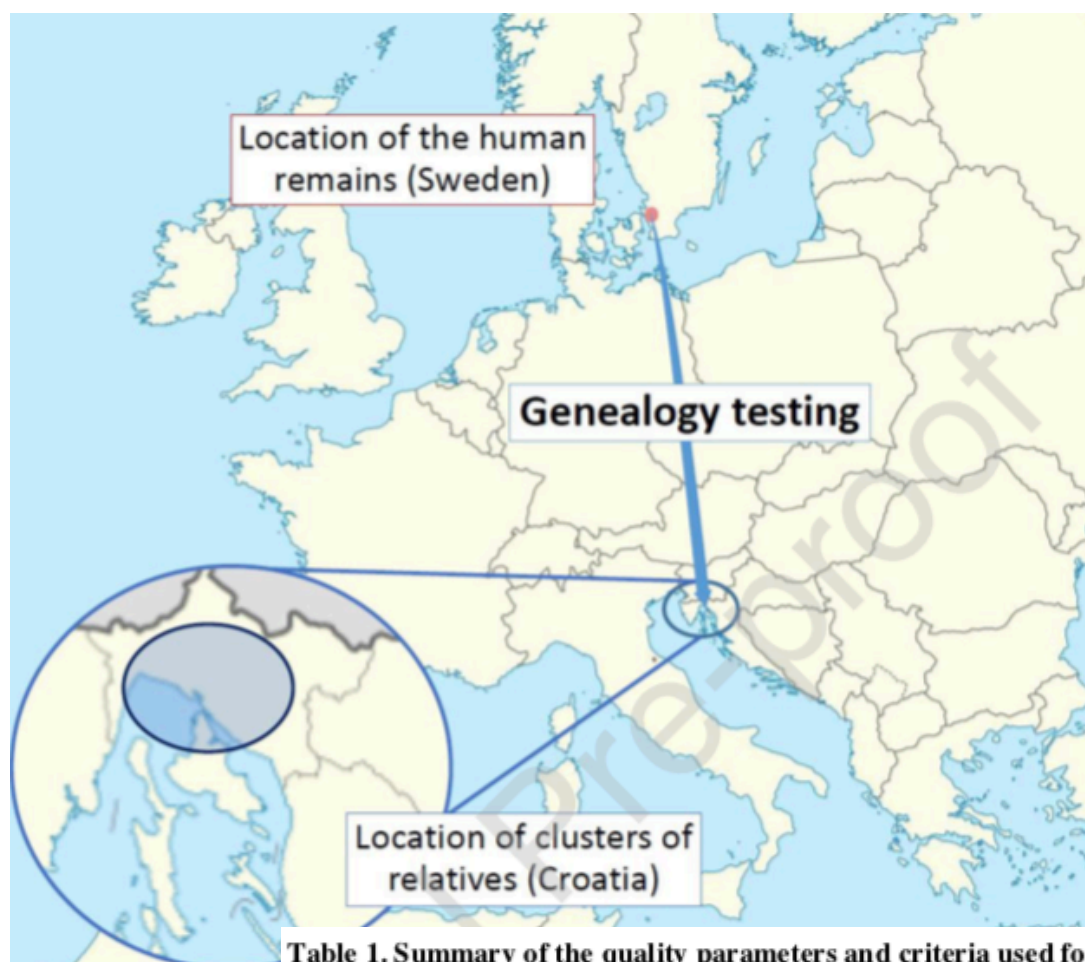


Table 1. Summary of the quality parameters and criteria used for the genotype calling

Parameter	Description	Threshold for genotype calling
Coverage	Number of individual reads that has a certain nucleotide at a given position.	≥ 10 (homozygote) ≥ 5 per allele (heterozygote)
Allelic balance*	[coverage for the most common nucleotide]/[coverage for the most common nucleotide + coverage for the second most common nucleotide].	$= 1$ (homozygote) $0.5 \leq x \leq 0.7$ (heterozygote)
Q score	Mean quality score (nucleotide Q).	> 25
Forward/Reverse-ratio	Forward and reverse read ratio for the nucleotide(s) comprising the genotype.	> 0.2

Cases in National Forensic Labs now being reported

- 3 x 3 ng DNA from a femur of murder victim generated 3 libraries for HiSeq^x
- Fragments of av. 400 bp size gave av. 150 bp read lengths as basis for 1,378,481 SNPs for an optimum GEDmatch intersect
- DNA quality and SNP call reliability were checked with STRs and the Qiagen 140-SNP ID panel respectively
- 3 billion reads gave 1,035,274 SNPs using Table 1 thresholds
- GEDmatch search made before terms and conditions changed
- A cluster of relatives was found within a 40km radius in NW Croatia - likely from a single pair of victim's grandparents



Research paper

Getting the conclusive lead with investigative genetic genealogy – A successful case study of a 16 year old double murder in Sweden

Andreas Tillmar^{a,b,*}, Siri Aili Fagerholm^c, Jan Staaf^d, Peter Sjölund^e, Ricky Ansell^{c,f,**}^a Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden^b Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Linköping, Sweden^c National Forensic Centre, Swedish Police Authority, Linköping, Sweden^d Polisregion Öst, Swedish Police Authority, Linköping, Sweden^e Peter Sjölund AB, Hårnösand, Sweden^f Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden**Table 2**

Summary of the established dataset and database searches.

Sample/ library preparation	Dataset	Database	Number of genotypes ^a (approx.)	Total shared segment length for top hits
DNA library 1.1	WGS analysis	GEDmatch	155,000–269,000	Less than 30 cM
DNA library 1.1	WGS analysis and genotype imputation	GEDmatch	864,000–1026,000	Less than 30 cM
DNA library 1.1 + DNA library 1.2	WGS analysis and genotype imputation	GEDmatch	908,000–1050,000	Less than 30 cM
DNA library 2.1	WGS analysis	GEDmatch	1279,000	Less than 30 cM
		FTDNA	1861,000	~350 cM, ~100 cM, ~60 cM and decreasing

Cases in National Forensic Labs now being reported

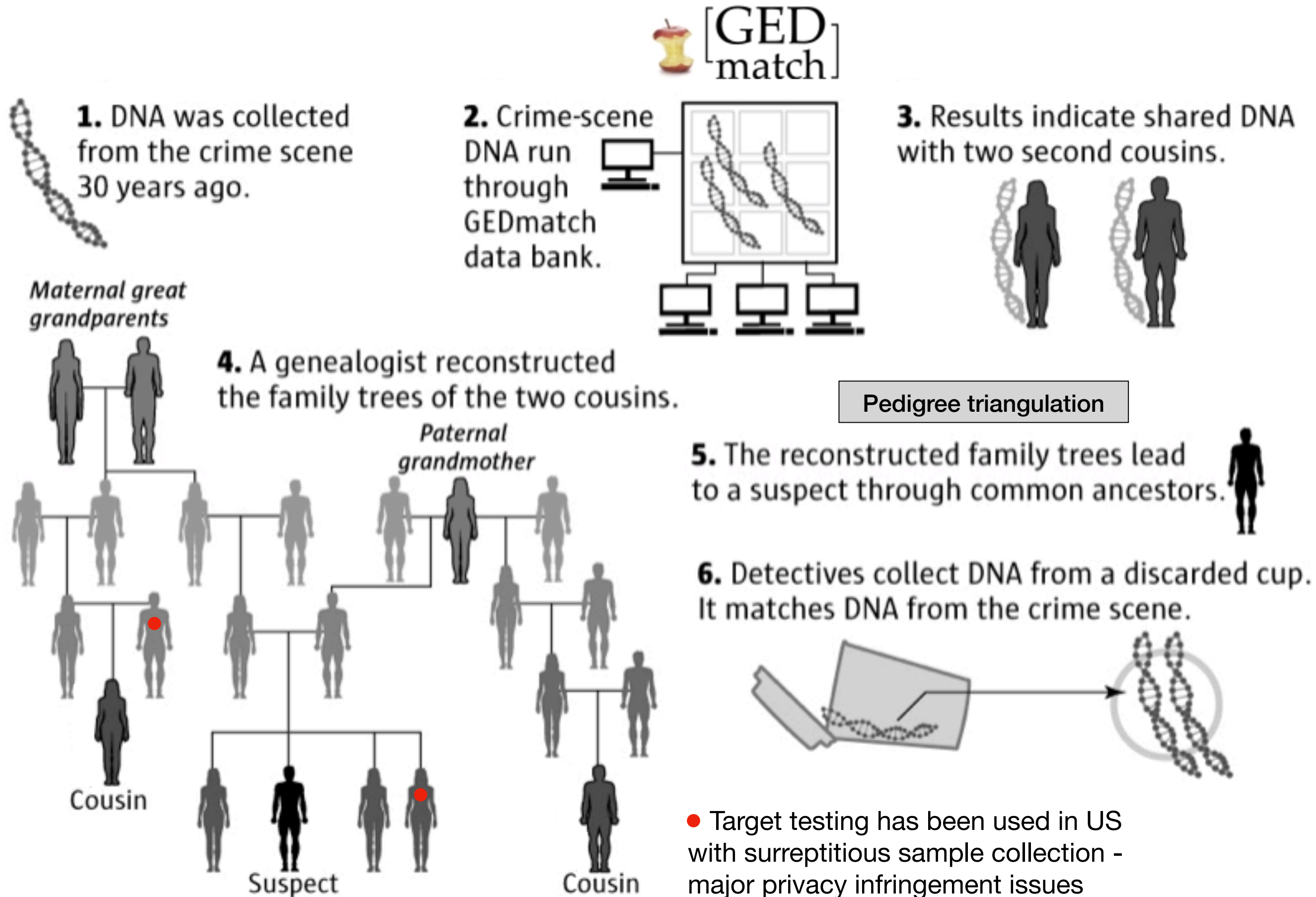
- Mother and child stabbed to death in Linköping in 2004. Mass screen of 6000 men to try to match DNA left at scene
- 3 separate whole-genome-sequence runs from 20 ng pooled DNA with 150 bp paired-end sequencing (NovaSeq 6000). Targeted 1,378,481 SNPs
- SNP genotypes were increased with imputation and a 1.3 million and 1.9 million dataset was applied to GEDmatch and FT DNA

Table 1

Summary statistics from the WGS runs.

Parameter	WGS run 1 (DNA library 1.1)	WGS run 2 (DNA library 1.2)	WGS run 3 (DNA library 2.1)
Average Coverage	13X	10X	60X
Duplication rate	~ 70%	~ 60%	< 10%
Median insert size	~ 60 bp	~ 60 bp	~ 180 bp

Investigators may need to extend to target testing to fill pedigree



- The application of IGG is most actively pursued in the US, but has been successfully used elsewhere - notably Sweden

	Country	Percentage
1	United States	65%
2	United Kingdom	9%
3	Canada	6%
4	Australia	4%
5	France	2%
6	Germany	1%
7	Sweden	1%
8	Ireland	1%
9	New Zealand	1%
10	Netherlands	1%

Uploads to GEDmatch

- The application of IGG is most actively pursued in the US, but has been successfully used elsewhere - notably Sweden
- Ethical considerations about individual privacy have been ignored in some cases and this has caused concern



Should we be making use of genetic genealogy to assist in solving crime?

A report on the feasibility of such methods in the UK

September 2020

The Biometrics and Forensics Ethics Group

UK and Sweden currently consider ethical/privacy concerns outweigh operational benefits of IGG and have not advocated its use - but such policies also relate to the success of STR-based familial searching in the UK compared to the US

- The application of IGG is most actively pursued in the US, but has been successfully used elsewhere - notably Sweden
- Ethical considerations about individual privacy have been ignored in some cases and this has caused concern
- A move towards much larger MPS SNP multiplexes will open up IGG to more labs - but extending familial search to 2nd-3rd cousin - e.g. Kintelligence (10,230 SNPs) + GEDmatch Pro



ForenSeq Kintelligence Kit

The only forensic genetic genealogy assay that targets the relevant markers in degraded and low-input DNA samples



Thank you



Genetisti Forensi Italiani

XXVIII
CONGRESSO
NAZIONALE

**LA GENETICA
FORENSE:
QUALE FUTURO?**

10|11|12
Novembre 2022
GENOVA



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Review article

Investigative genetic genealogy: Current methods, knowledge and practice

Daniel Kling^{a, b, *}, Christopher Phillips^{c, **}, Debbie Kennett^d, Andreas Tillmar^{a, e}

^a Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden

^b Department of Forensic Sciences, Oslo University Hospital, Oslo, Norway

^c Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain

^d Research Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, United Kingdom

^e Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Linköping, Sweden

