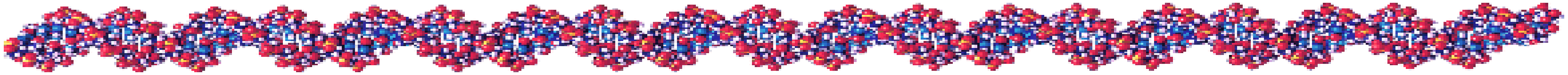
Identification of Distant Relatives Using Linked Autosomal STRs



Kristen E. Lewis, PhD
NRC Postdoctoral Fellow
NIST Applied Genetics Group

Melanie Caserza, Tom Walsh, and Mary-Claire King
University of Washington
Genome Sciences

DNA Identification

Unlinked Markers

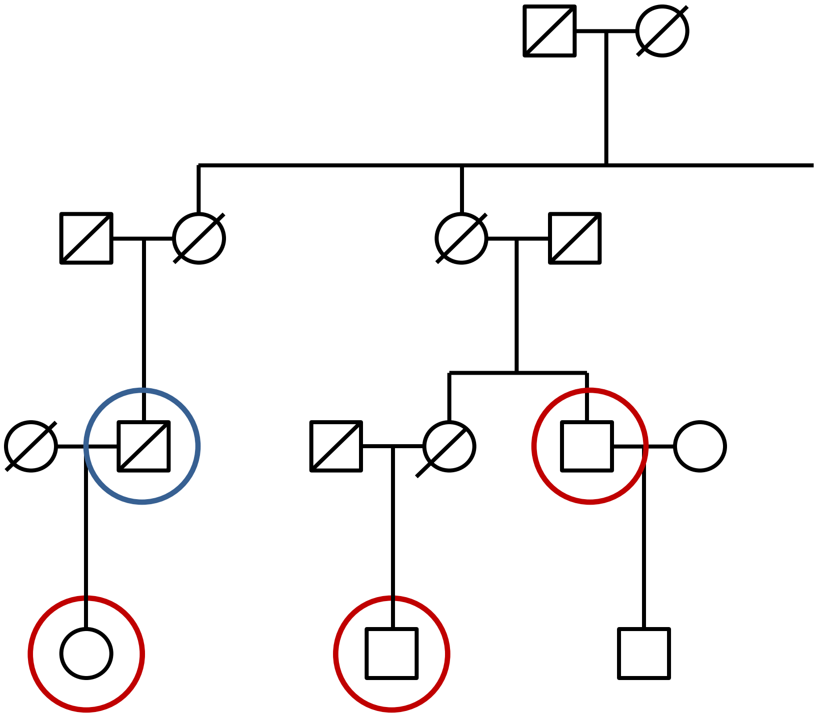
- Recombination
  - CODIS loci
- Direct matching
- Paternity analysis

Linked Markers

- No recombination
  - Y-STRs
  - Mitochondrial DNA
  - X-STRs
  - Autosomal SNP haploblocks
  - Linked autosomal STRs
- Complex kinship analysis
- Historical identification
- Remains identification

Remains Identification

- American soldier missing since WWII
- Remains discovered in ocean near Australia (2002)
- Identified through mtDNA sequencing

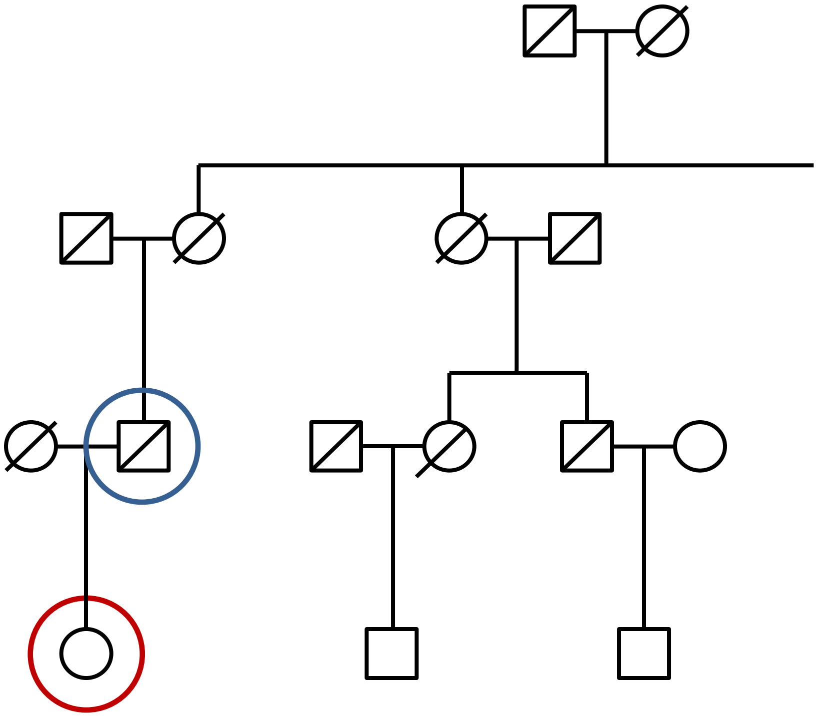


Humanitarian Identifications by the King Laboratory

- Children of the Disappeared (Argentina)
  - mtDNA sequencing
  - Kinship analysis
- Present case
  - Associate the alleged daughter with living members of the soldier’s family
  - Infer paternity of the alleged father

A Question of Kinship

- Australian woman raised as soldier’s posthumous daughter
- Baptismal certificate
- Confirm kinship for personal and legal reasons
- No autosomal DNA from remains
- No Y or mtDNA relationship



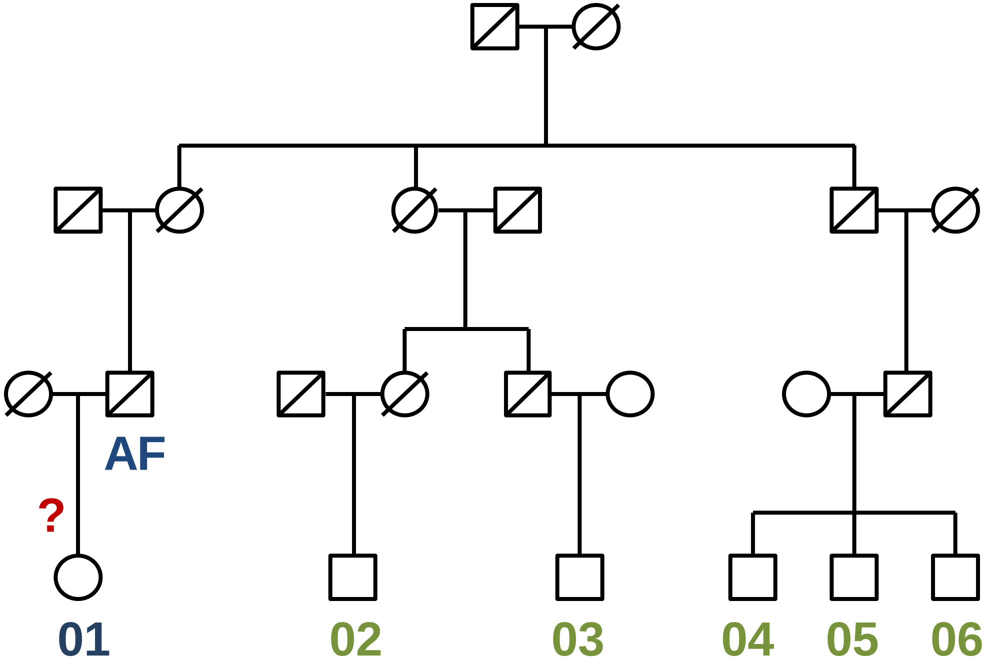
Approach

- Evaluate relatedness between alleged daughter and putative second cousins

- Likelihood ratio analysis using forensic autosomal STR loci

- Haplotype analysis using linked autosomal STR loci

Likelihood Ratio Analysis

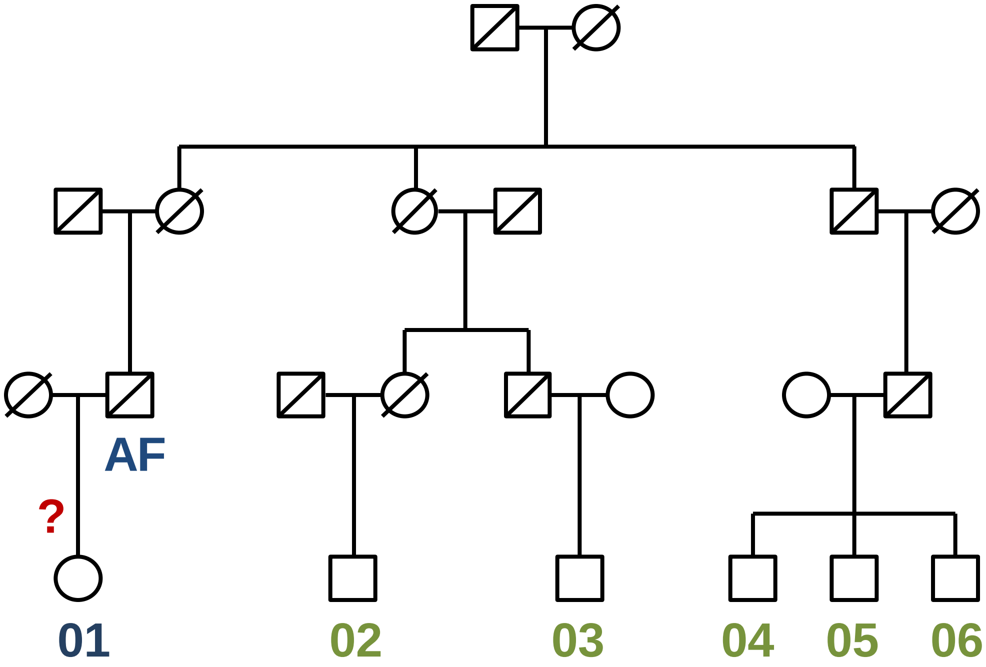


40 autosomal STRs
- Identifiler (15)
- NIST 26plex (25)

$$LR = \frac{\text{Probability of genotypes if 01 and putative } 2^{nd} \text{ cousin are } \mathbf{2^{nd}\ cousins}}{\text{Probability of genotypes if 01 and putative } 2^{nd} \text{ cousin are } \mathbf{unrelated}}$$

Hill, C.R., Butler, J.M., Vallone, P.M. (2009) A 26plex autosomal STR assay to aid human identity testing. *J. Forensic Sci.* 54(5): 1008-1015

Likelihood Ratio Analysis

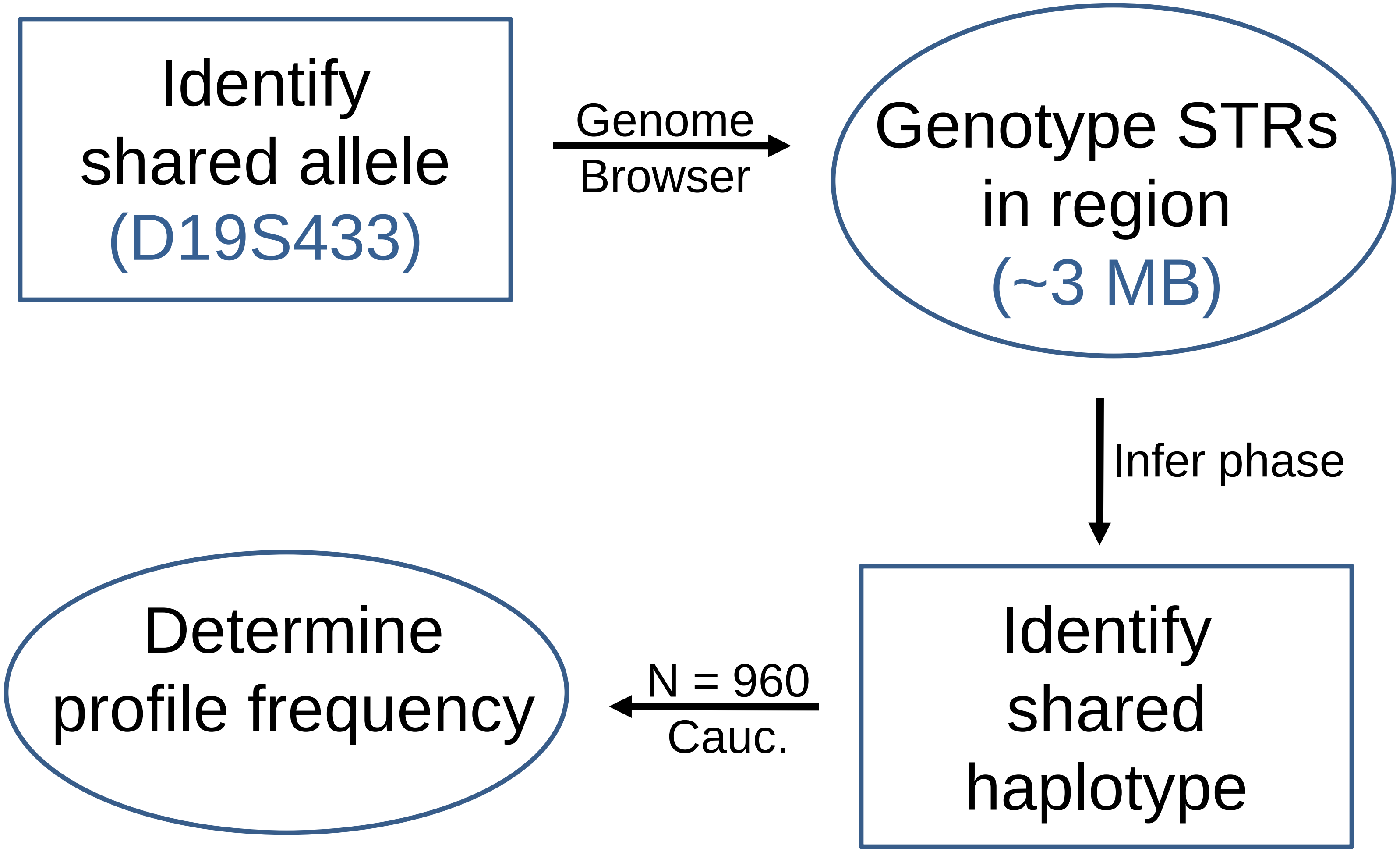


40 autosomal STRs
- Identifiler (15)
- NIST 26plex (25)

Likelihood ratios for comparisons of:
- Putative 2nd cousins ranged from **0.5 to 0.9**
- Undisputed 2nd cousins ranged from **0.5 to 3.9**

Probability of sharing one allele identical by descent = 1/16

Autosomal Haplotype Strategy



Identify shared allele (D19S433)

Genome Browser →

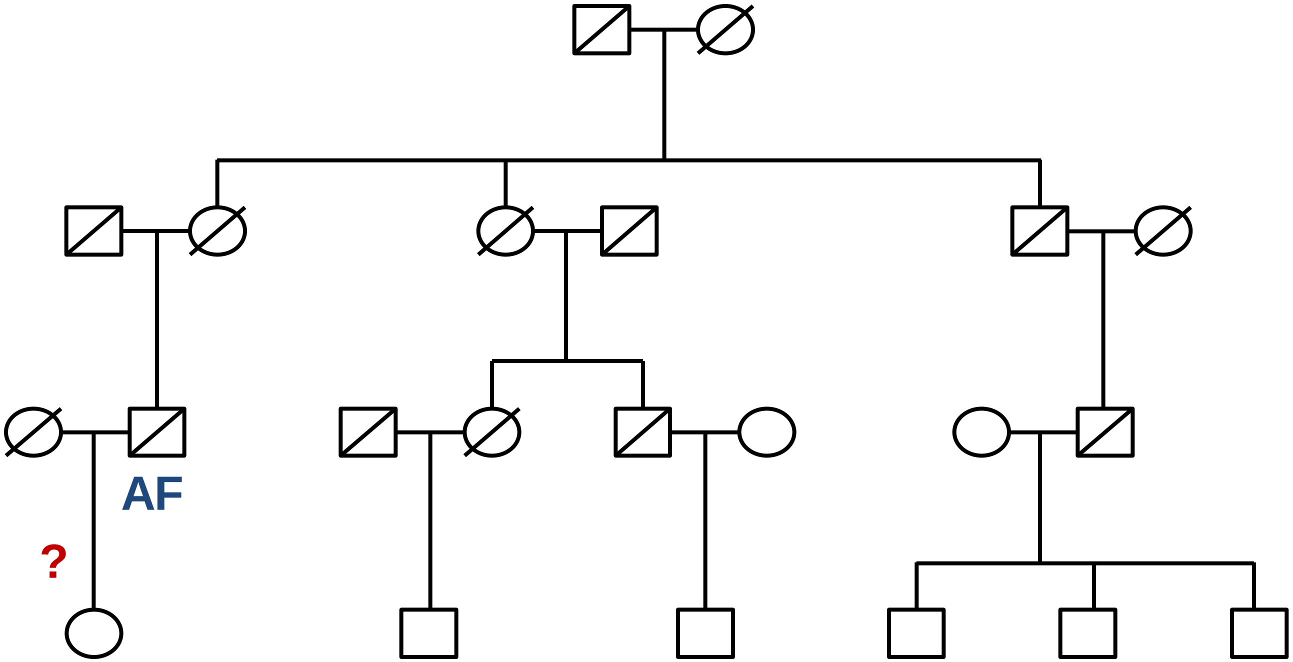
Genotype STRs in region (~3 MB)

↓ Infer phase

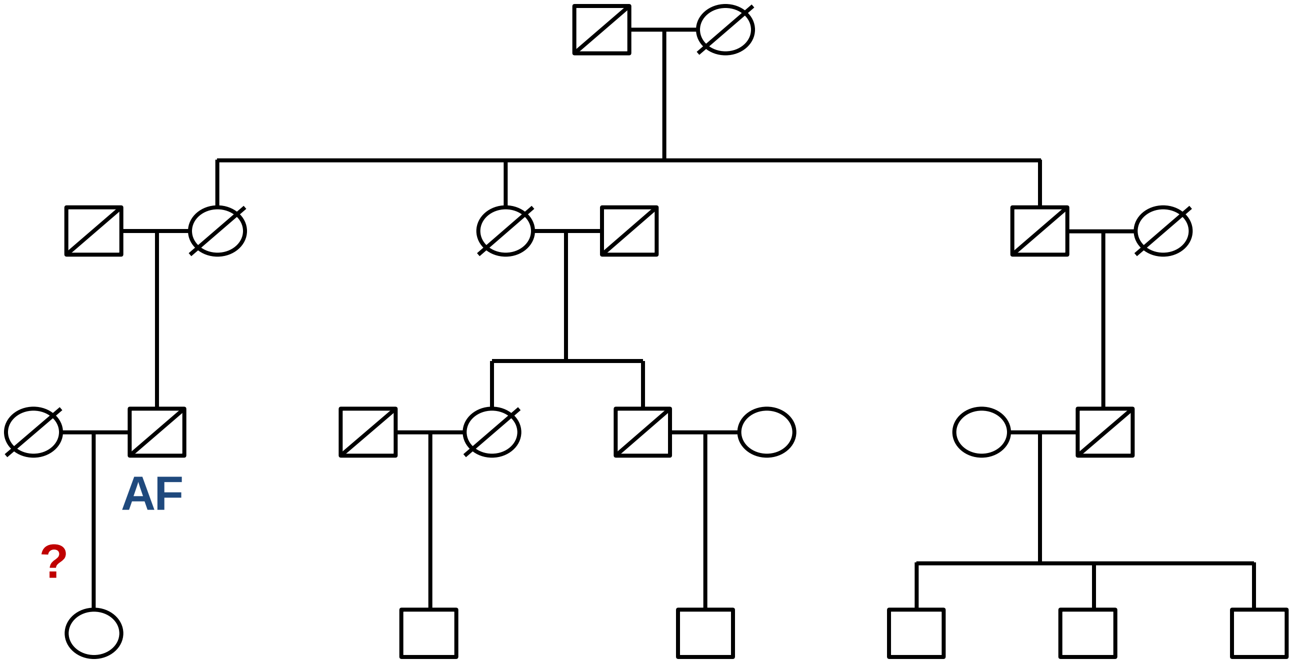
Identify shared haplotype

← N = 960 Cauc.

Determine profile frequency



Chr 19 (Mb)	01		02		03		04		05		06	
33.588	7	5	5	6	4	5	4	3	3	5	3	5
33.972	2	2	5	7	4	4	2	6	6	2	6	2
34.077	2	2	2	2	2	14	2	6	6	2	6	2
34.238	3	9	8	10	9	16	8	10	12	9	10	9
34.489	5	13	13	13	10	14	4	4	4	13	4	13
34.777	3	3	2	4	3	10	3	3	4	3	3	3
D19S433	15.2	14	14	14	13	15	12	14	15.2	14	14	14
35.477	2	7	7	7	7	7	7	5	5	7	5	7
35.549	3	3	2	3	2	3	1	4	2	3	4	3
36.506	6	10	8	8	1	7	9	2	9	10	2	10



Chr 19 (Mb)	01	02	03	04	05	06
33.588	7 **5**	5 6	4 5	4 3	3 **5**	3 **5**
33.972	2 **2**	5 7	4 4	2 6	6 **2**	6 **2**
34.077	2 **2**	2 2	2 14	2 6	6 **2**	6 **2**
34.238	3 **9**	8 10	9 16	8 10	12 **9**	10 **9**
34.489	5 **13**	13 13	10 14	4 4	4 **13**	4 **13**
34.777	3 **3**	2 4	3 10	3 3	4 **3**	3 **3**
D19S433	15.2 **14**	14 14	13 15	12 14	15.2 **14**	14 **14**
35.477	2 **7**	7 7	7 7	7 5	5 **7**	5 **7**
35.549	3 **3**	2 3	2 3	1 4	2 **3**	4 **3**
36.506	6 **10**	8 8	1 7	9 2	9 **10**	2 **10**

Profile Frequency

0 out of 960 controls could have a genotype consistent with the observed haplotype

95% confidence limit from zero proportion

$= 1 – \alpha^{(1/N)}$

where $\alpha = 0.05$ and $N = 1920$ chromosomes

$= 0.0016$

Holland and Parsons (1999)

Likelihood Ratio

$$LR = \frac{\text{Probability of profiles if 01 and 05/06 are } \textbf{related}}{\text{Probability of profiles if 01 and 05/06 are } \textbf{unrelated}}$$

$$LR = \frac{1}{0.0016} = 641$$

641 times more likely to observe the shared haplotype if the woman is related to this family than unrelated

Summary

- Defined a region of linked autosomal alleles
  - Constituted a haplotype in the known brothers
  - Shared by alleged daughter and brothers (putative 2nd cousins)

- Haplotype is rare in ancestry-matched population
  - Suggestive of identity by descent and not identity by state

- Region has not been broken up by recombination in the generations separating the tested relatives

- As many independent chromosomes as necessary could be evaluated

Linked Autosomal Markers

- Relatedness of persons separated by:
  - Multiple historical generations
  - War
  - Natural disaster

- Useful for particular identification or kinship cases
  - Not for direct matching

- Use when lineage markers are uninformative

Acknowledgments

University of Washington

Mary-Claire King
Tom Walsh
Melanie Caserza

NIST Applied Genetics Group
John Butler
Pete Vallone
Becky Hill

Questions

Kristen Lewis

kristen.lewis@nist.gov