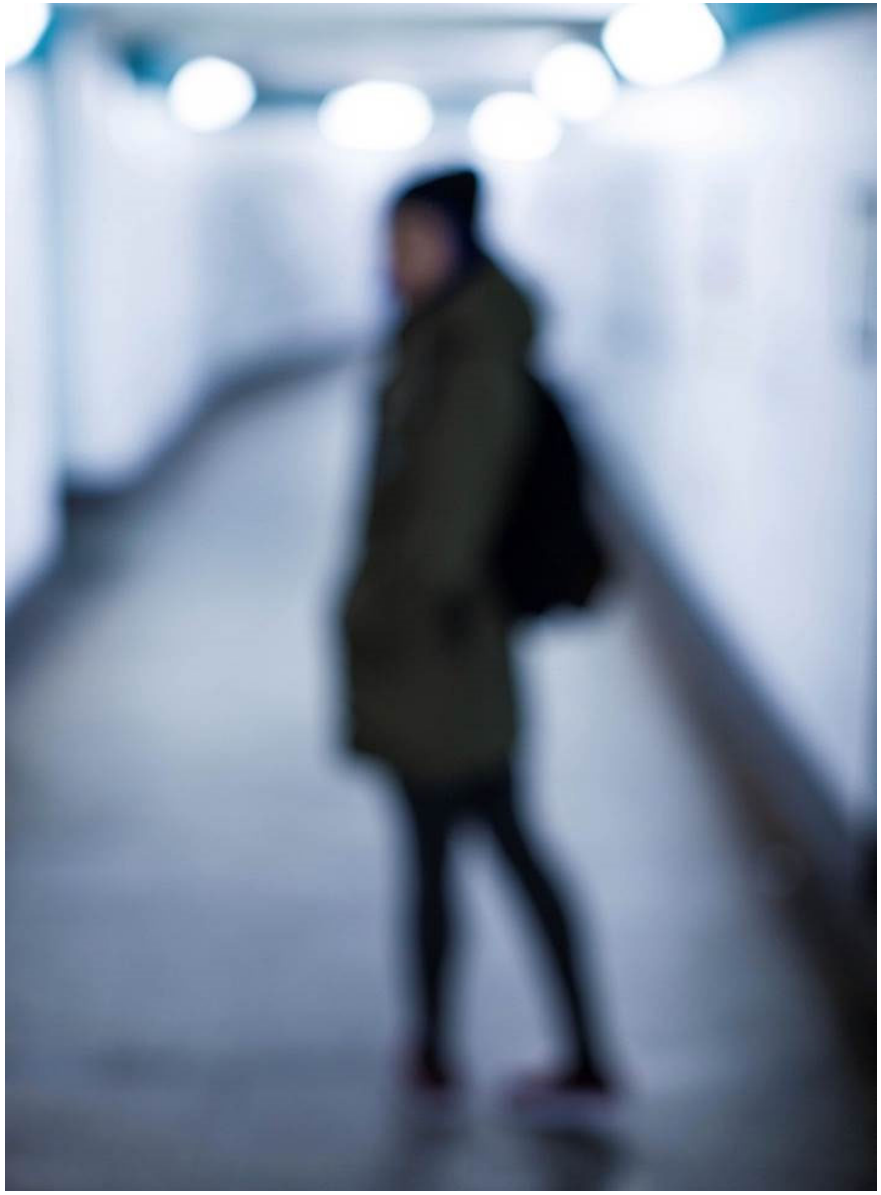
Interpretation of Y chromosome STRs for missing persons cases

Jianye Ge, Benjamin Crysup, Dixie Peters, Romy Franco, Meng Huang, Muyi Liu, Xuewen Wang, Bruce Budowle

Center for Human Identification

University of North Texas Health Science Center (UNTHSC)

Y-STR for forensic applications



Most violent crimes are committed by men

- 99% of the forcible rapes
- 88% of the robberies
- 85% of the burglaries
- 88.8% of the homicide offenders

Y chromosome inheritance

- Only inherit from father to sons
- Can be used to trace male lineage

Y chromosome tests provide information for

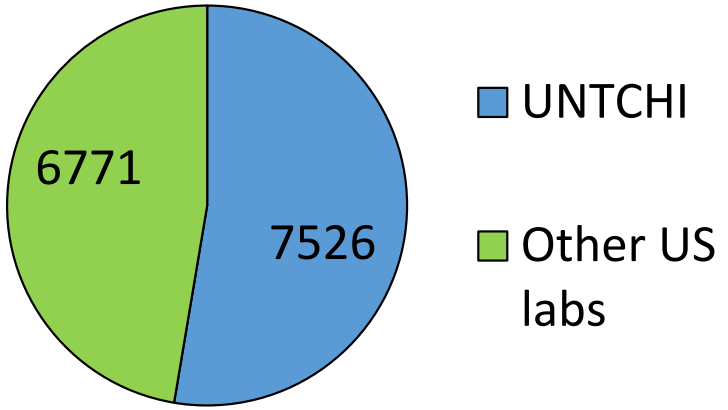
- Kinship analysis
- Missing persons identification
- Familial searching
- Mixture in sexual assault cases
- Ancestry inference
- Link multiple cases
- Y database searching

Snyder HN. Arrest in the United States, 1990-2010. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice; 2012.
Homicide Trends in the United States. 1980-2008. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice; 2011.

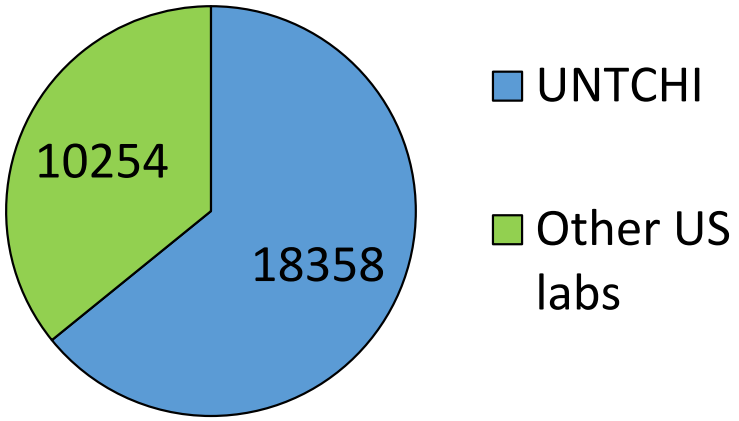
Missing Persons Unit at UNTCHI

- 650-1,000 remains samples processed/year
- 1,200-2,400 family reference samples/year
- Greater contribution of DNA profiles in CODIS/NDIS than all other MP lab

Remains at NDIS



References at NDIS



Current as of August 2021

SWGDAM guidelines

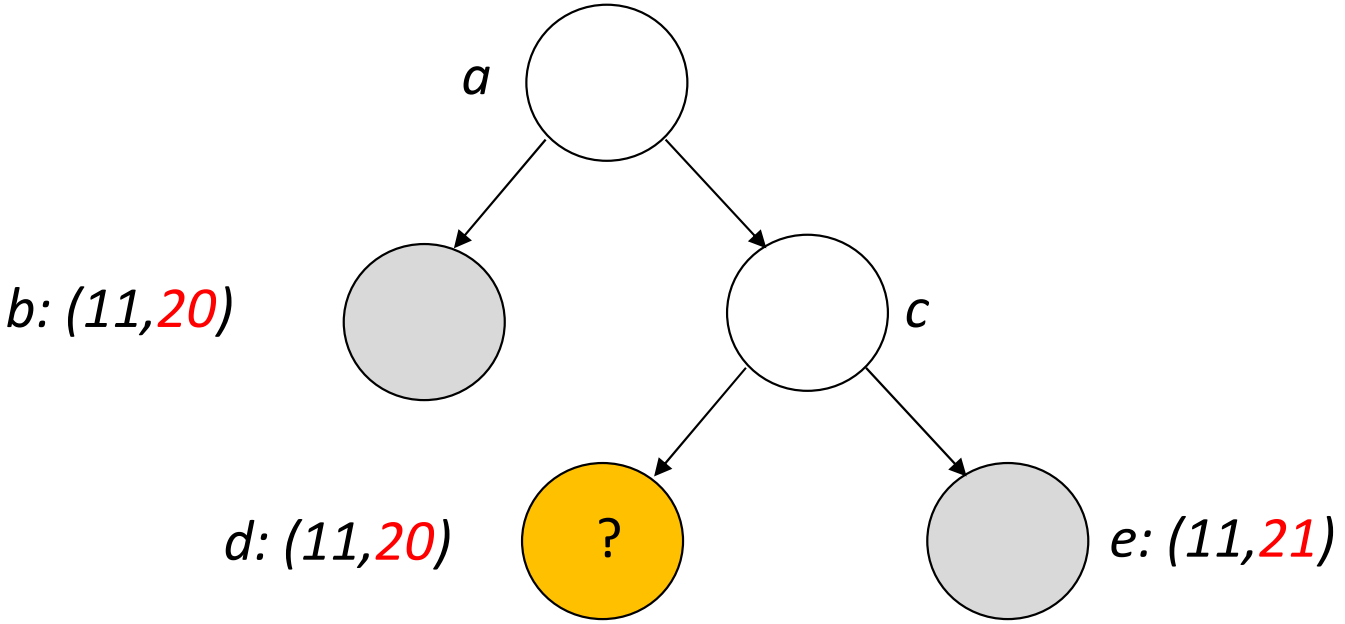
- **SWGDAM Guidelines for Missing Persons Casework**

  "For missing persons, relatives of missing persons and unidentified human remains samples, additional DNA methods other than autosomal STR typing (such as mtDNA or Y-STR typing) should always be considered, if relevant."

  "A Y-STR LR and an mtDNA LR are computed for each of those systems."

- **SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories**
  - Focuses on haplotype frequency estimation
  - No recommendation was given on how to calculate likelihood ratio for scenarios with multiple male references and/or mutations

A complex case - example



- A two-locus Y-STR pedigree including 2 typed references (*b* & *e*), 2 untyped individuals (*a* & *c*), and 1 missing person (*d*)
- The missing person is related to two references with different Y haplotypes

Interpretation methods for complex cases

- Pedigree likelihood ratio (PLR) for lineage markers, similar to PLR for autosomal markers

Int J Legal Med (2011) 125:519–525
DOI 10.1007/s00414-010-0514-9

ORIGINAL ARTICLE

Pedigree likelihood ratio for lineage markers

Jianye Ge · Arthur Eisenberg · Jiangwei Yan · Ranajit Chakraborty · Bruce Budowle

- A convenient and accurate method to determine if two Y-STR profiles are from the same lineage

Electrophoresis 2016, *37*, 1659–1668

Hai Liu[1]
Xiaoyang Li[1]
Julio Mulero[2]
Andrea Carbonaro[2]
Marc Short[2]
Jianye Ge[2]

[1]The Institute of Forensic Science and Technology, Henan Provincial Public Security Bureau, Zhengzhou, P. R. China
[2]Human Identification Division, Thermo Fisher Scientific, South San Francisco, CA, USA

Research Article

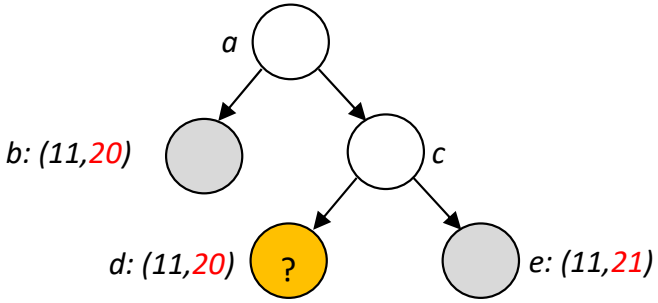
A convenient guideline to determine if two Y-STR profiles are from the same lineage

Y chromosome STR loci are used in forensics primarily for identification purposes by determining the male lineages. The Henan province in China has established a large Y-STR (>200 000 profiles) database for criminal investigations. A large proportion of the Y-STR profiles in the database were generated using either the Applied Biosystems Yfiler or Yfiler Plus PCR Amplification kits. The additional loci in the Yfiler Plus kit as compared to the Yfiler kit results in a concomitant cumulative mutation rate increase across the loci.

Pedigree likelihood ratio (PLR) for Y-STRs

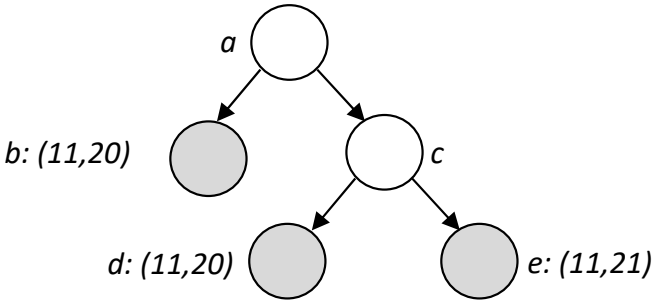
Pedigree likelihood ratio (PLR) with Y-STRs

- Multiple Y-STR references with different haplotypes

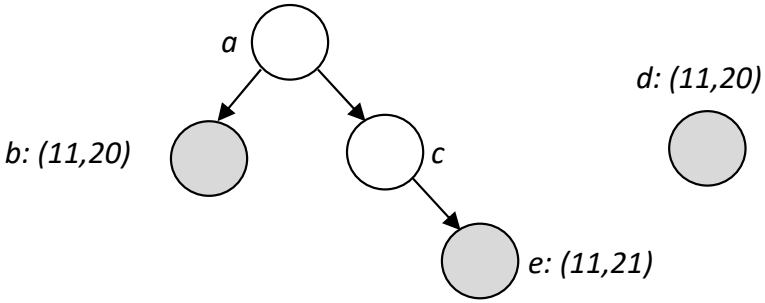


A two-locus Y-STR pedigree including 2 typed references (*b* & *e*), 2 untyped individuals (*a* & *c*), and 1 missing person (*d*)

- Two competing hypotheses



Kinship hypothesis (H_k)

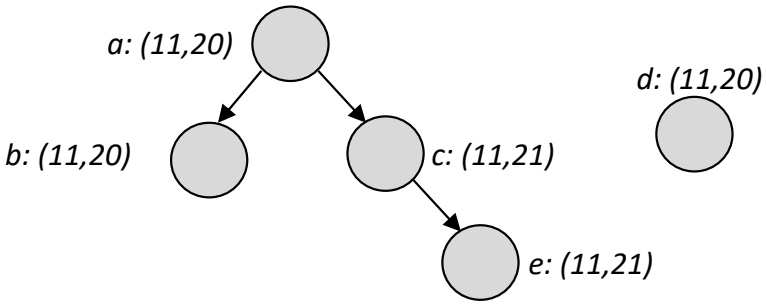


Non-Kinship hypothesis (H_{nk})

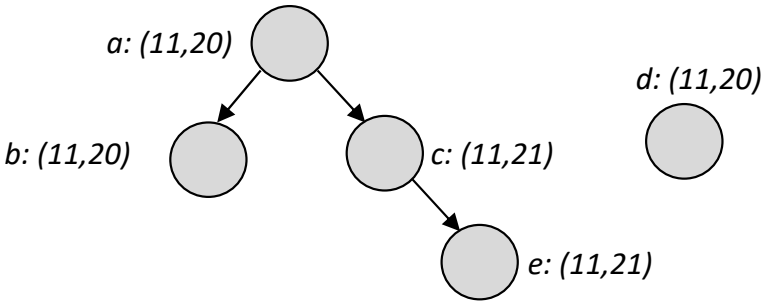
Pedigree likelihood (PL)

- PL = the cumulative haplotype frequency(ies) of the founder(s)
  × the cumulative transmission probabilities of all father-son pairs

- Predict possible haplotypes of the untyped individuals
  - *a* = (11,20), (11,21), etc.; *c* = (11,20), (11,21), etc.
  - Calculate likelihood for each haplotype combination (e.g., *a* = (11,20) & *c* = (11,21))



Pedigree likelihood (PL) – cont.

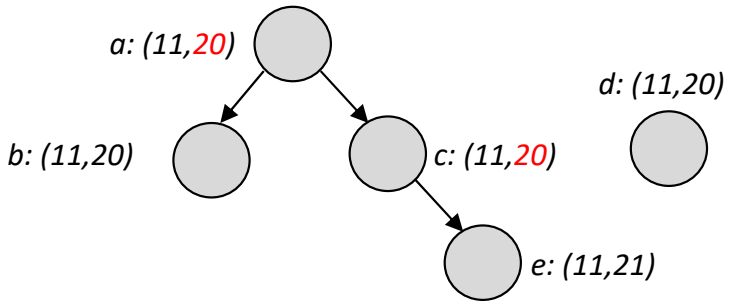


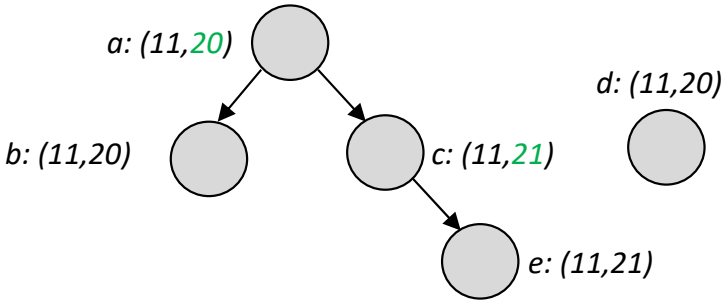
- Founders' haplotype frequencies = Pr(a)×Pr(d)
  - Search Y-STR database (e.g., YHRD) to obtain haplotype frequencies

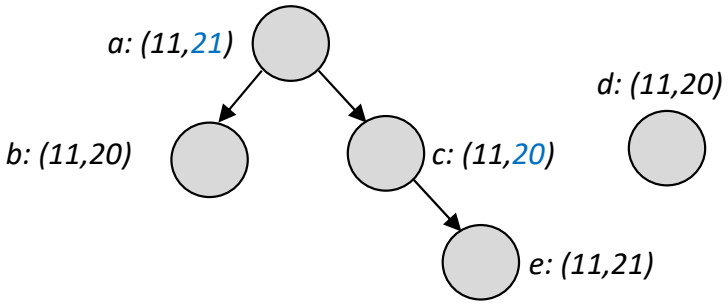
- Transmission probabilities = Pr(a→b)×Pr(a→c)×Pr(c→e)
  - Pr(c→e) = Pr(11→11)×Pr(20→21), assuming independent mutations between markers
  - Mutation rates and mutation model (e.g. Two Phase Model)

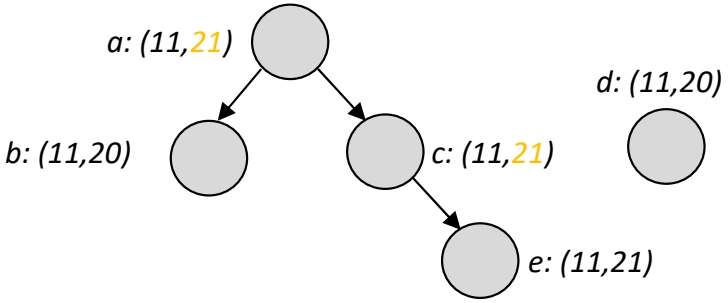
Pedigree likelihood (PL) – cont.

- Sum the likelihoods for all possible haplotype combinations for the untyped individuals
- PL = L(G|*a* = (11,20) & *c* = (11,20)) + L(G|*a* = (11,20) & *c* = (11,21)) +
  L(G|*a* = (11,21) & *c* = (11,20)) + L(G|*a* = (11,21) & *c* = (11,21))









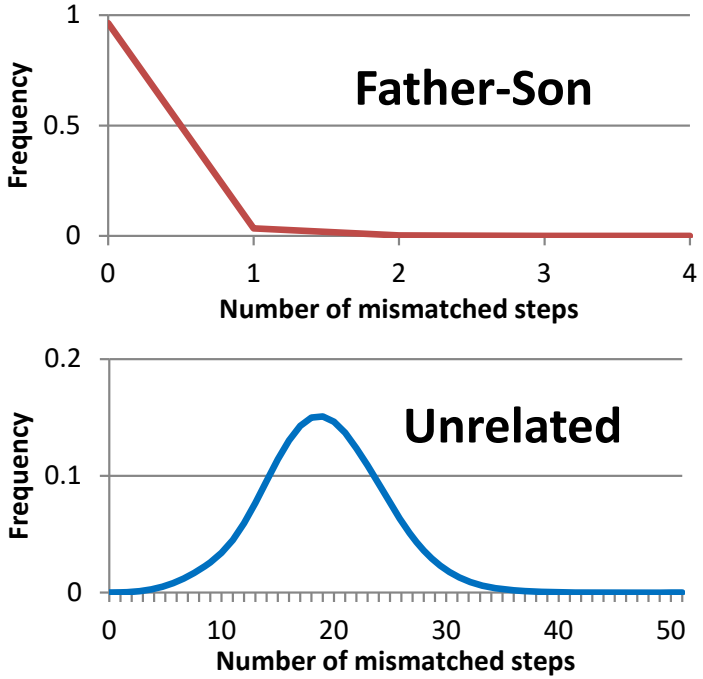
A convenient method to assess if two Y-STR profiles are from the same lineage

A convenient method to assess if two Y-STR profiles are from the same lineage

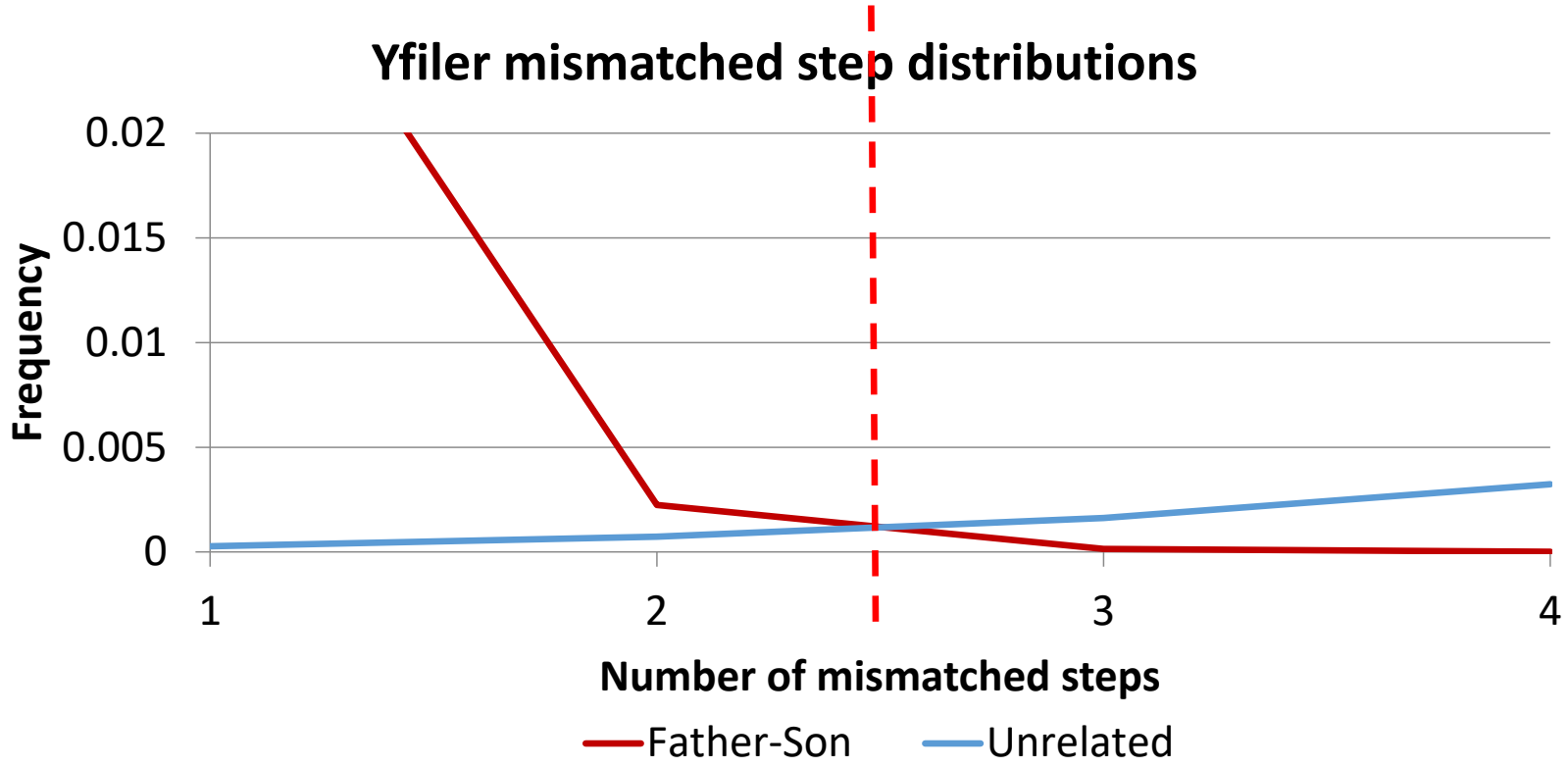
- PLR approach requires pedigree structure is defined, which may not be available in some cases
- Analysts just need to know if two Y-STR profiles are related or not
- A simple solution – use the number of mismatched loci/steps
  - More distant relationship → More mismatched loci/steps

Example:
10 and 12 → 1 mismatched locus or 2 steps

- Y-STR profiles from the **same lineage** usually have **no or a small number of mismatched** loci/steps
- Y-STR profiles from **different lineages** usually have **a relatively large number of mismatched** loci/steps



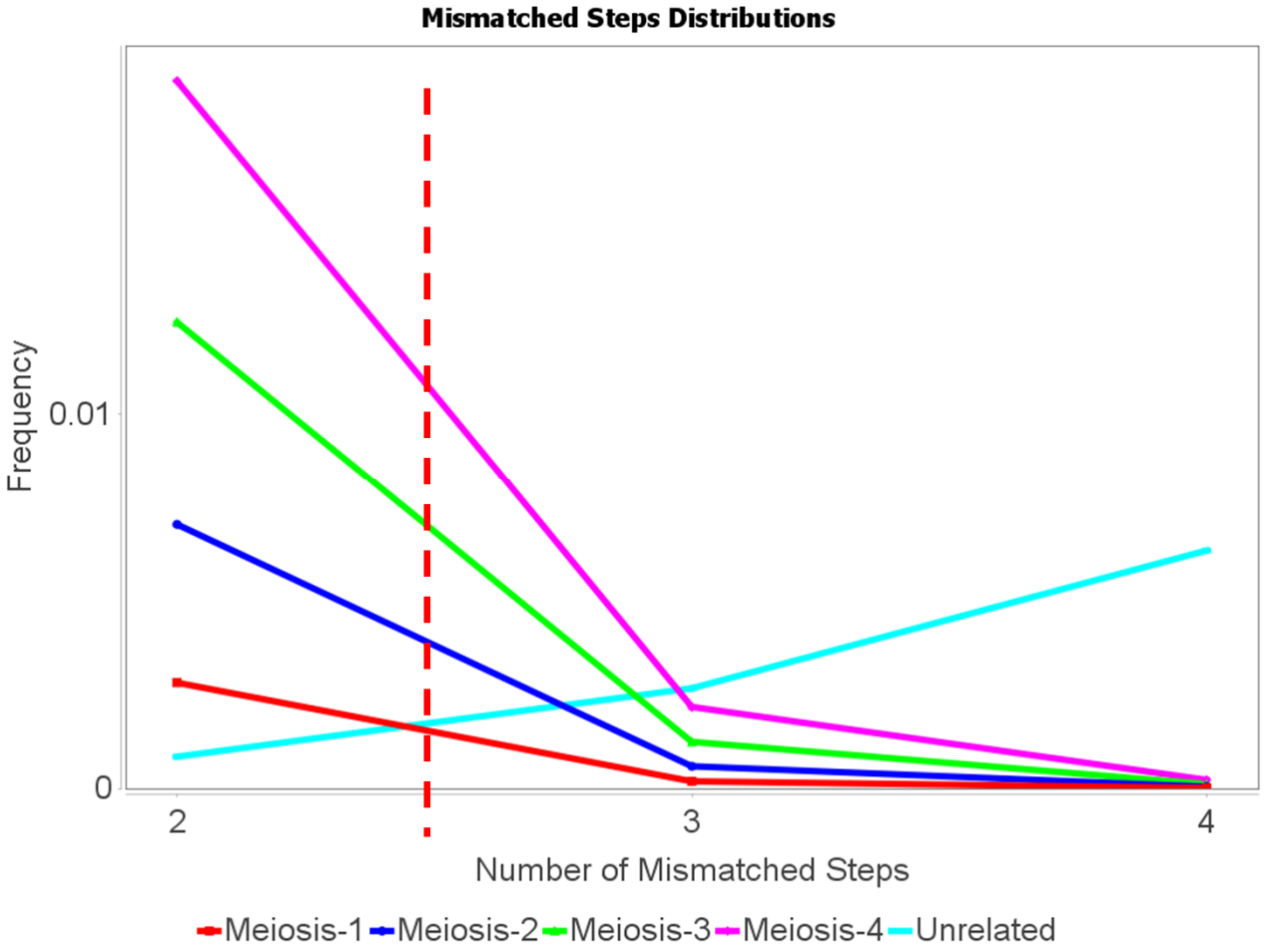
Merge distributions and select a threshold

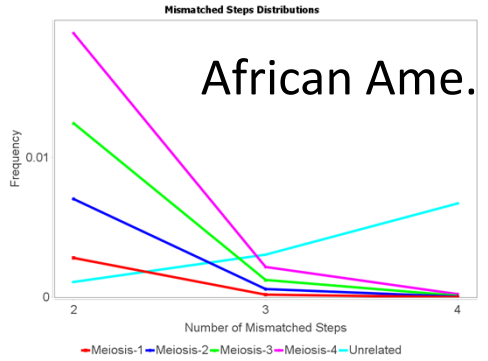


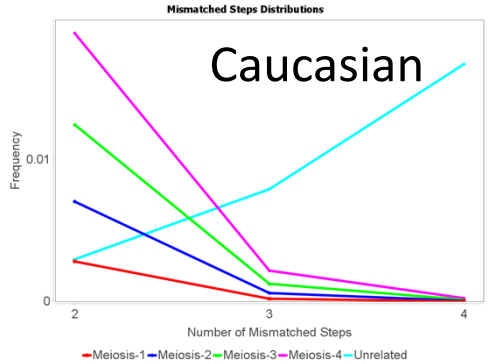
≤2: More likely they are father-son

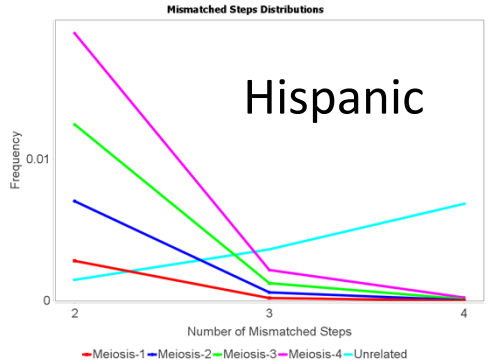
≥3: More likely they are unrelated

Yfiler









- Yfiler profiles: Budowle et al. (2010)
- Mutation rates: YHRD, https://yhrd.org/pages/resources/mutation_rates
- Meiosis-1 = father-son; Meiosis-2 = full-sibling; Meiosis-3 = uncle-nephew; Meiosis-4 = first-cousin.

Budowle B, Ge J, Aranda XG, Planz JV, Eisenberg AJ, Chakraborty R. Texas population substructure and its impact on estimating the rarity of Y STR haplotypes from DNA evidence. Journal of forensic sciences. 2009 Sep;54(5):1016-21.

How reliable is this method?

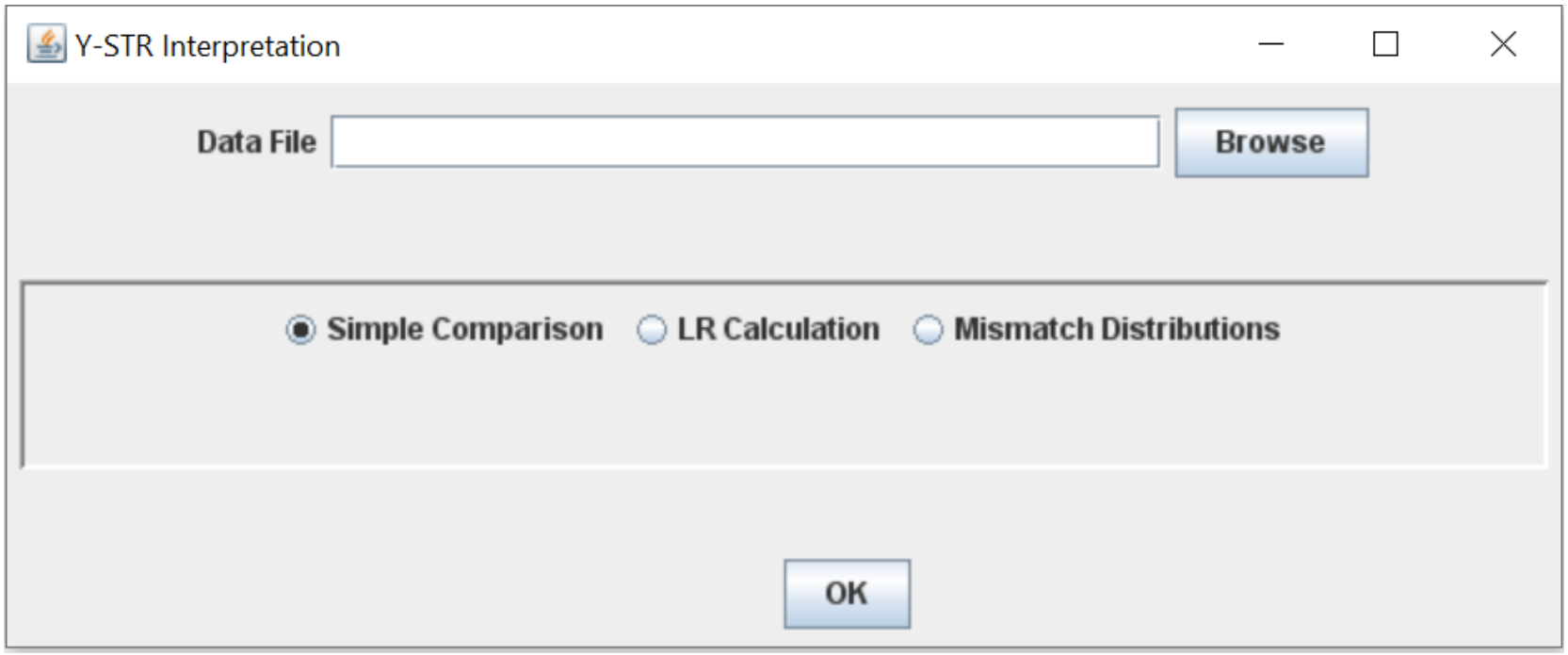
- A study by Liu et al. (2016)
  - 7,405 Chinese unrelated Yfiler Plus profiles
  - Define a lineage as up to 4 meioses relationships (3rd degree or first-cousin)
  - ≤5: more likely from the same lineage
- The chance of being incorrect is extremely low
  - Two unrelated Yfiler Plus profiles, in most cases (i.e., >99.975%), will have more than 5 mismatched steps
  - For almost all close relatives, the mismatched steps are ≤5

Relationship	Chance of mismatched steps ≤5
Father-son	99.999997%
Grandfather-Grandson; brother	99.999984%
Uncle-nephew	99.999950%
First-Cousin	99.999878%

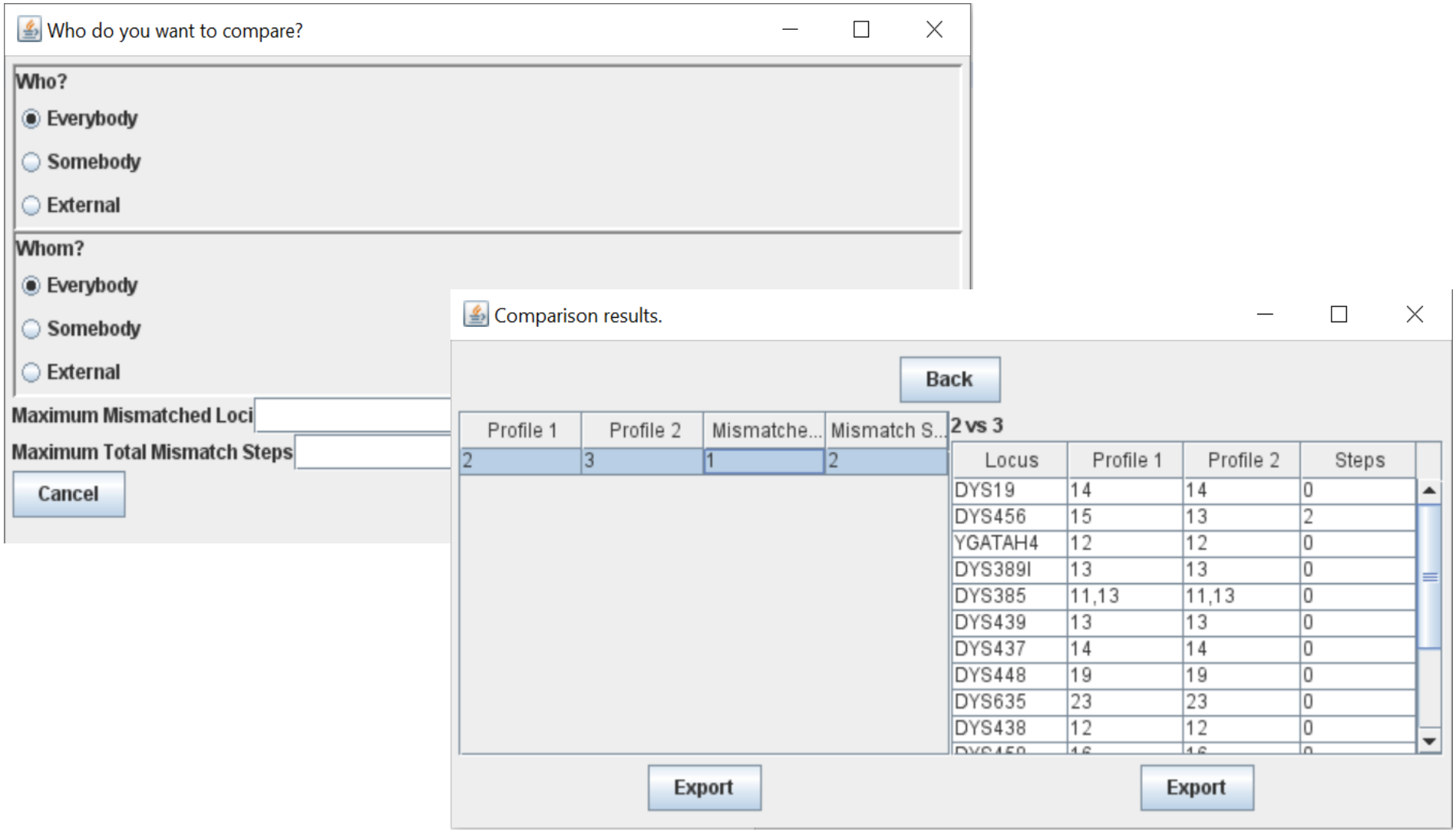
  - Genotyping error (i.e., ~0.1%) is greater than the error associated with this method
  - Accuracy will be lower if the lineage definition includes more distant relatives

- Liu H, Li X, Mulero J, Carbonaro A, Short M, Ge J. A convenient guideline to determine if two Y-STR profiles are from the same lineage. Electrophoresis. 2016 Jul 1;37(12):1659-68.
- Lackey A., https://www.thermofisher.com/blog/behindthebench/how-to-determine-if-two-y-str-haplotypes-are-from-the-same-male-lineage/, (2018).

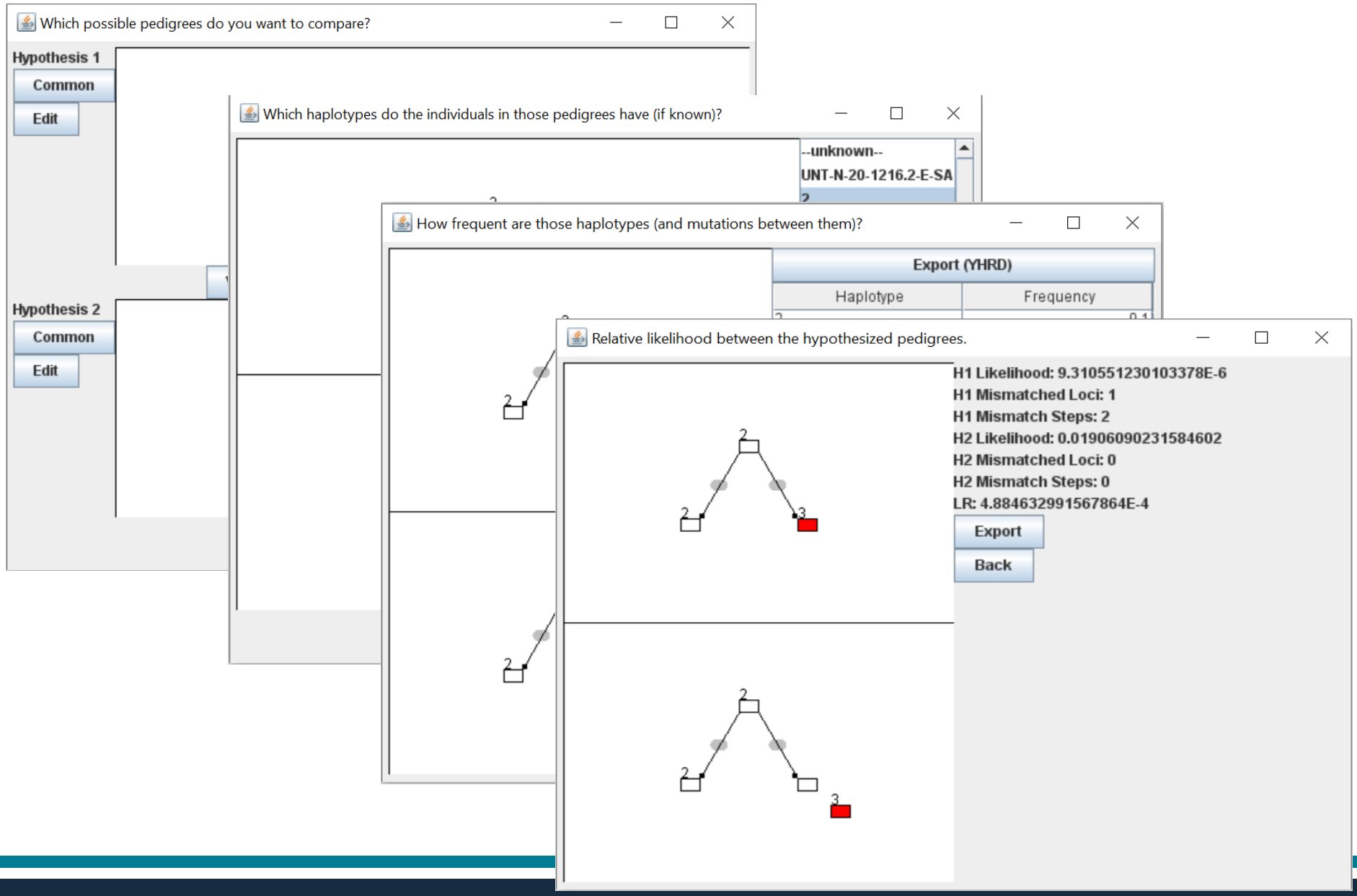
Interpretation software



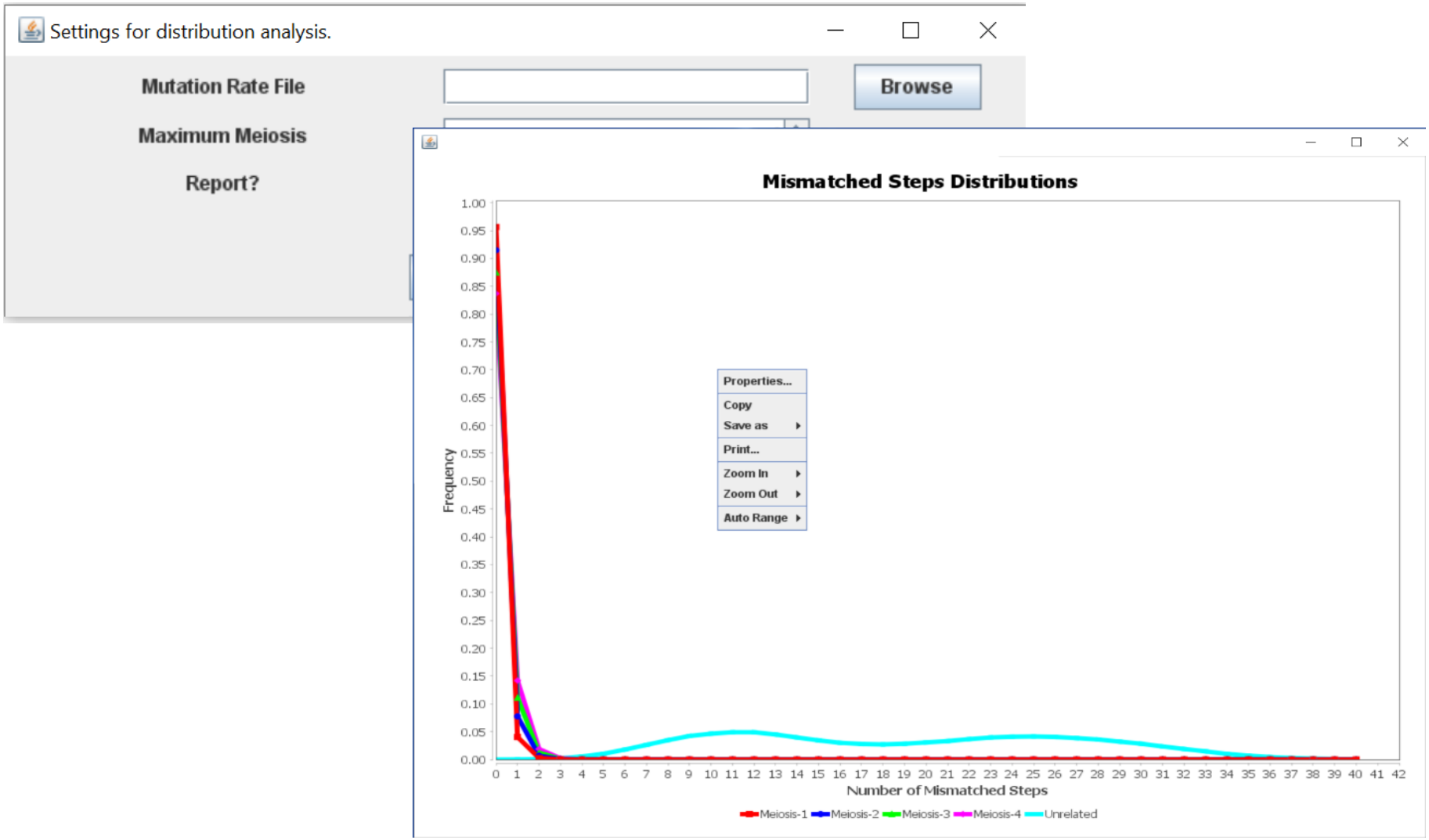
Simple comparisons



LR calculation



Mismatch distributions



Thank you!

This study was supported in part by award 2020-DQ-BX-0018 (Interpretation of Y chromosome STRs for missing persons cases), awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the U.S. Department of Justice.