

2-22-2007

Empirical Analysis of the STR Profiles Resulting from Conceptual Mixtures

Dan E. Krane

Wright State University - Main Campus, dan.krane@wright.edu

Follow this and additional works at: <https://corescholar.libraries.wright.edu/biology>



Part of the [Biology Commons](#), [Medical Sciences Commons](#), and the [Systems Biology Commons](#)

Repository Citation

Krane, D. E. (2007). Empirical Analysis of the STR Profiles Resulting from Conceptual Mixtures. .
<https://corescholar.libraries.wright.edu/biology/257>

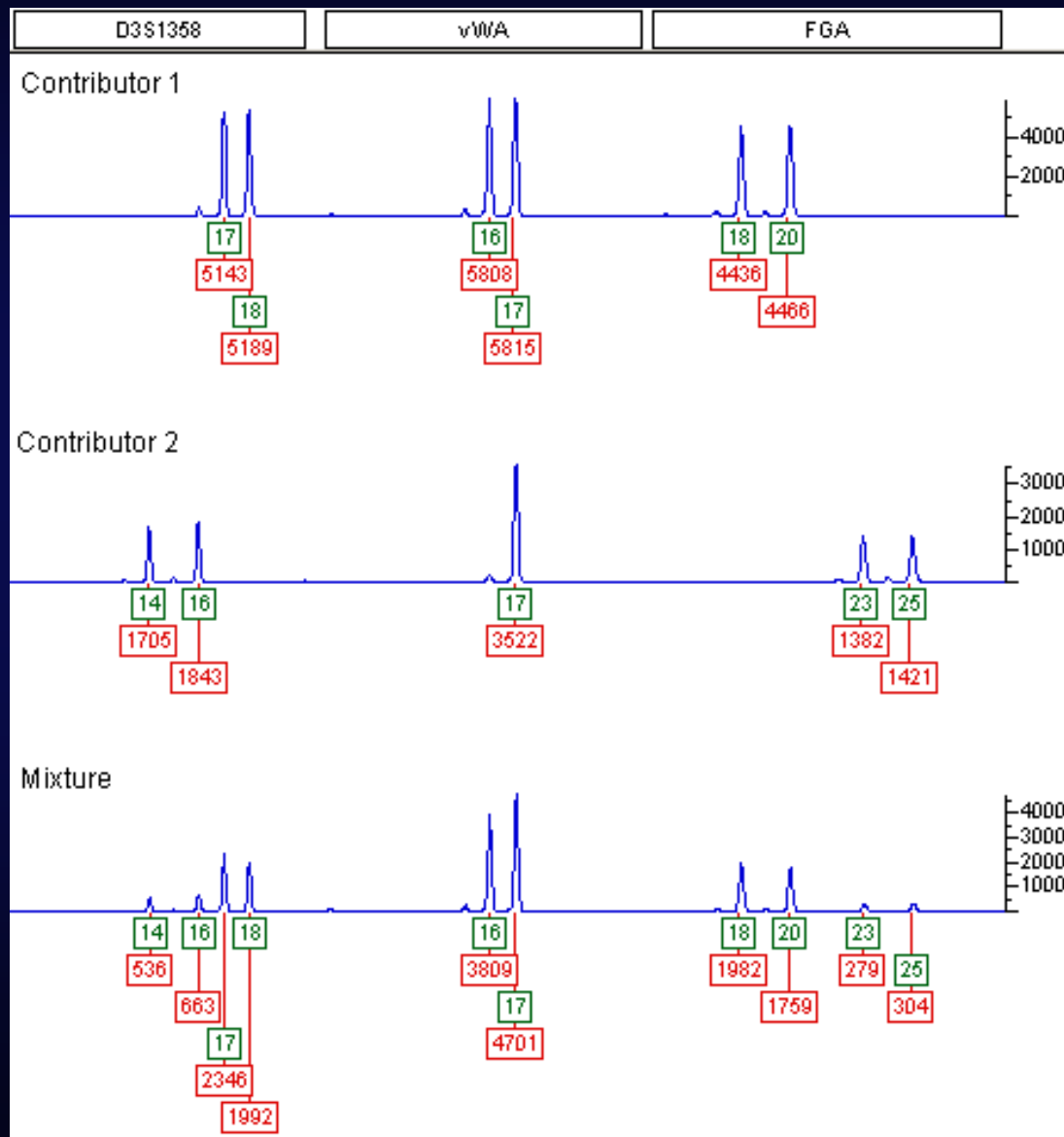
This Presentation is brought to you for free and open access by the Biological Sciences at CORE Scholar. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

Empirical analysis of the STR profiles resulting from conceptual mixtures

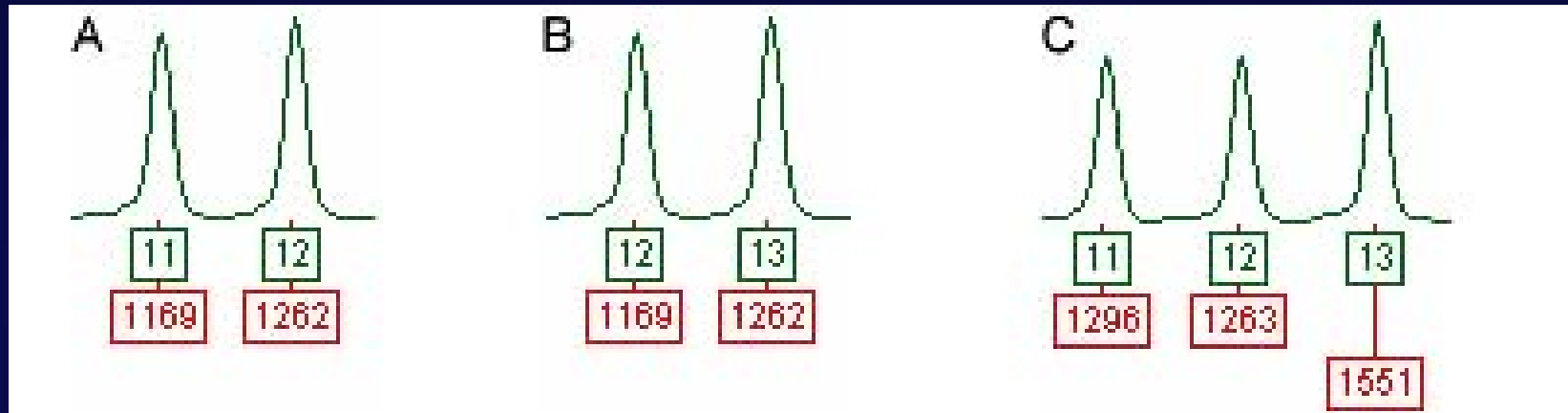
Dan E. Krane, Wright State University, Dayton, OH

Forensic Bioinformatics
(www.bioforensics.com)

Mixed DNA samples



Mixed DNA samples



Suitable profiles for empirical mixing

- 959 complete 13-locus (CODIS-loci) STR genotypes used by the FBI for the purpose of allele frequency databases
- Includes: Bahamians (153); Trinidadians (76); US African Americans (177); Southwest Hispanics (202); Jamaicans (157); and US Caucasians (194)
- Available on-line at:
<http://www.fbi.gov/hq/lab/fsc/backissu/july1999/dnaloci>
- Analyzed for Hardy-Weinberg equilibrium but no mention of possibility of relatives

How many contributors to a mixture? if analysts can discard a locus?

Maximum # of alleles observed in a 3-person mixture	# of occurrences	Percent of cases
2	0	0.00
3	3,398 78	0.00 0.00
4	26,788,540 4,967,034	18.28 3.39
5	112,469,398 93,037,010	76.75 63.49
6	7,274,823 48,532,037	4.96 33.12

There are 146,536,159 possible different 3-person mixtures of the 959 individuals in the FBI database (Paoletti et al., November 2005 *JFS*).

How many contributors to a mixture? If analysts can discard a locus?

Maximum # of alleles observed in a 3-person mixture	# of occurrences	Percent of cases
2	0	0.00
3	8,151 310	0.02 0.00
4	11,526,219 2,498,139	25.53 5.53
5	32,078,976 29,938,777	71.07 66.32
6	1,526,550 12,702,670	3.38 28.14

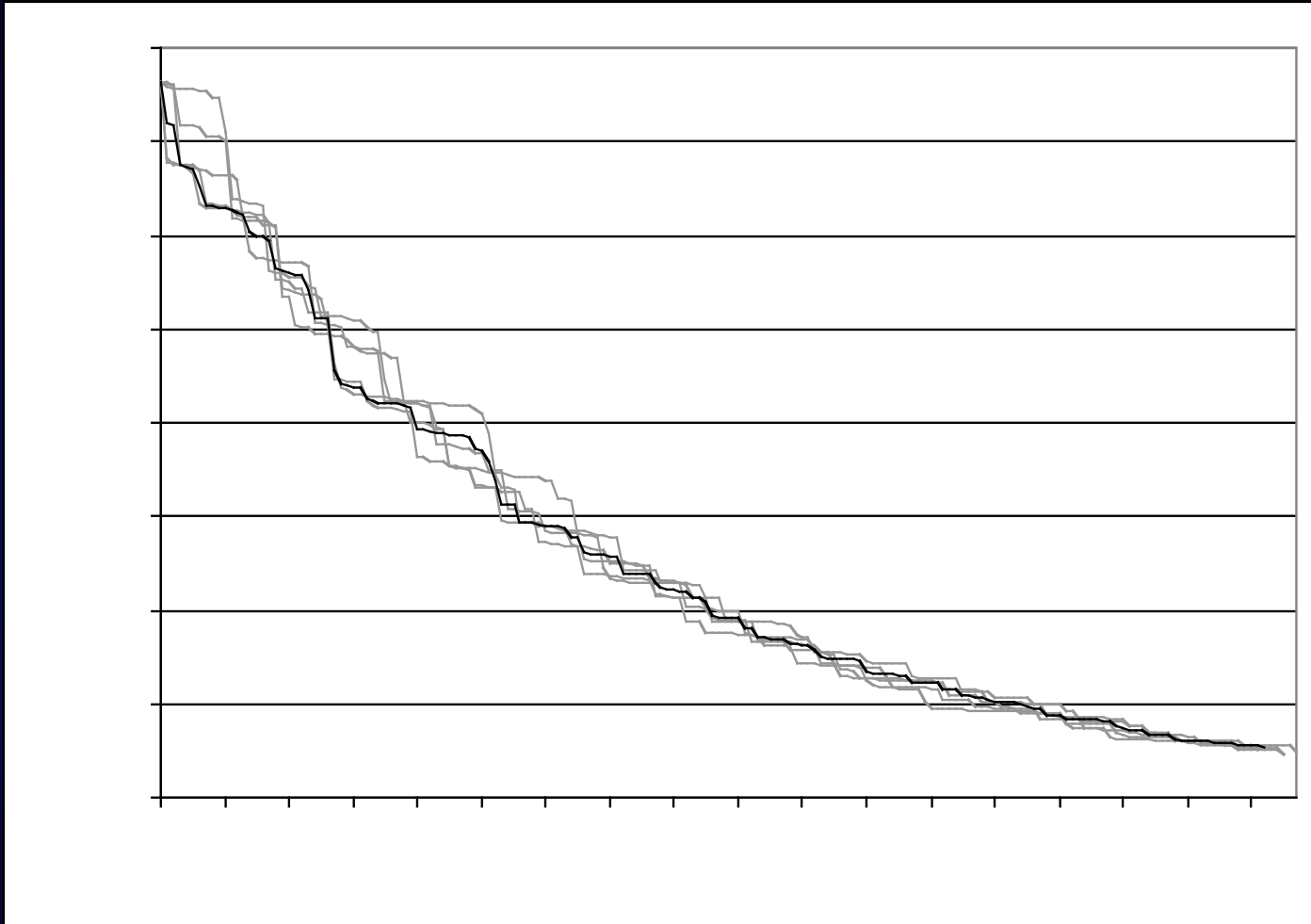
There are 45,139,896 possible different 3-person mixtures of the 648 individuals in the MN BCI database (genotyped at only 12 loci).

How many contributors to a mixture?

Maximum # of alleles observed in a 4-person mixture	# of occurrences	Percent of cases
4	13,480	0.02
5	8,596,320	15.03
6	35,068,040	61.30
7	12,637,101	22.09
8	896,435	1.57

There are 57,211,376 possible different 4-way mixtures of the 194 individuals in the FBI Caucasian database (Paoletti et al., November 2005 *JFS*). (35,022,142,001 4-person mixtures with 959 individuals.)

Does testing more loci help?

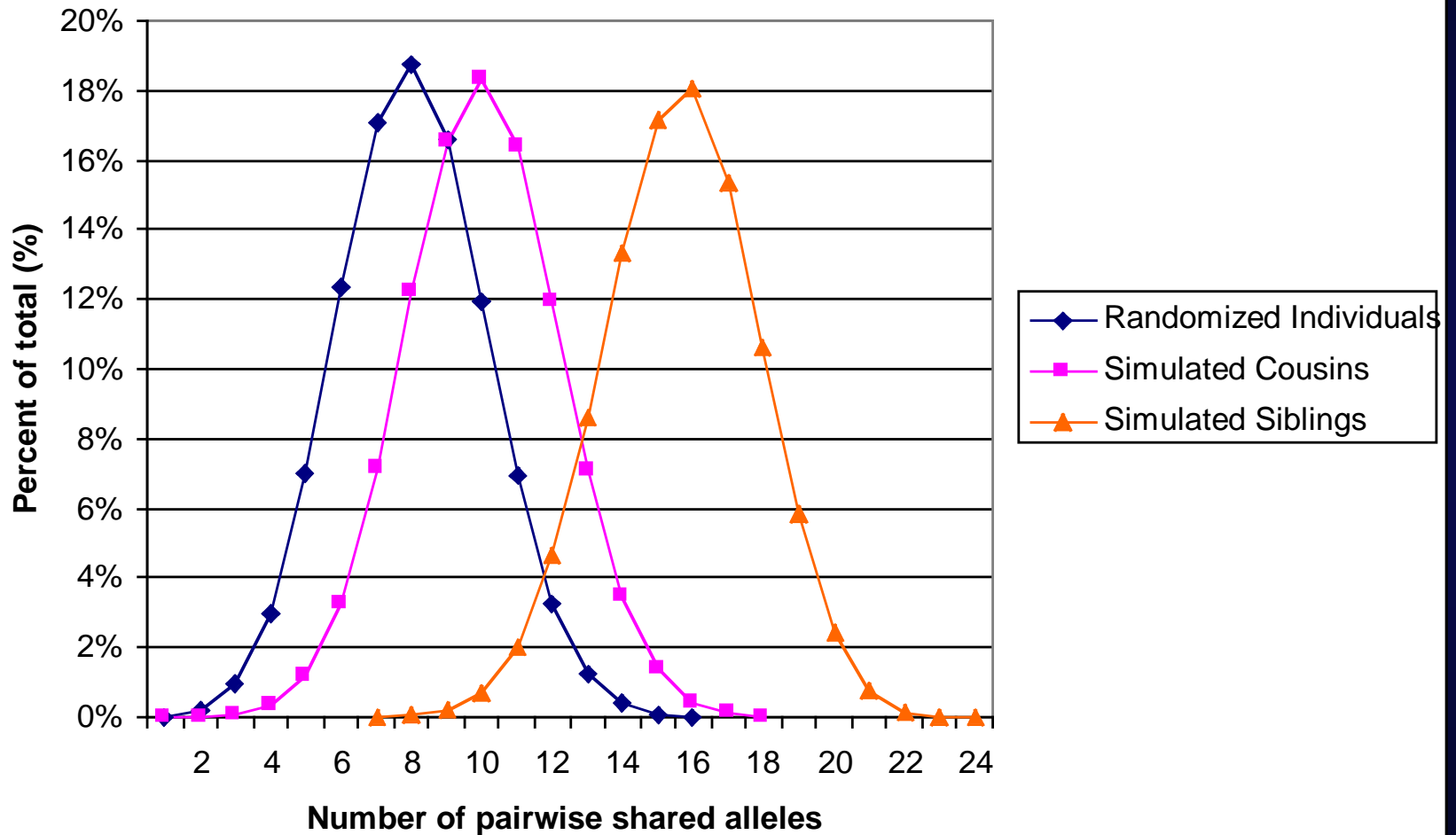


Five simulations are shown with each data point representing 57,211,376 4-person mixtures (average shown in black). (Paoletti et al., November 2005 *JFS*). Mischaracterization rate of 76.34% for original 13 loci.

What contributes to overlapping alleles between individuals?

- Identity by state
 - many loci have a small number of detectable alleles (only 6 for TPOX and 7 for D13, D5, D3 and TH01)
 - some alleles at some loci are relatively common
- Identity by descent
 - relatives are more likely to share alleles than unrelated individuals
 - perfect 13 locus matches between siblings occur at an average rate of 3.0 per 459,361 sibling pairs

Allele sharing between individuals



Allele sharing in databases

- Original FBI dataset's mischaracterization rate for 3-person mixtures (3.39%) is more than two σ above the average observed in five sets of randomized individuals
- Original FBI dataset has more shared allele counts above 19 than five sets of randomized individuals (3 vs. an average of 1.4)

Conclusions

- Maximum allele count by itself is not a reliable predictor of the number of contributors to mixed forensic DNA samples.
- Simply reporting that a sample “arises from two or more individuals” is reasonable and appropriate.
- Analysts should exercise great caution when invoking discretion.
- Excess allele sharing observed in the FBI allele frequency database is most easily explained by the presence of relatives in that database.

Resources

- Internet
 - **Forensic Bioinformatics Website:**
<http://www.bioforensics.com/>
 - **FBI's allele frequency database of genotypes:**
<http://www.fbi.gov/hq/lab/fsc/backissu/july1999/dnaloci>
 - **Additional figures and simulations:** birg.cs.wright.edu/~dpaolett/JFS-materials.zip
- Scientists
 - **Jason Gilder (Forensic Bioinformatics)**
 - **Carrie Rowland (Forensic Bioinformatics)**
 - **Keith Inman (Forensic Analytical, Haywood, CA)**
- *Journal of Forensic Sciences*
 - Paoletti, D., Doom, T., Krane, C., Raymer, M. and Krane, D. 2005. Empirical analysis of the STR profiles resulting from conceptual mixtures. *JFS*, 50:1361-1366.