Robust detection of individual forensic profiles in DNA mixtures Notes

J. Isaacson, E. Shwoebel, A. Scherbina, et al.

March 1, 2020

1 Introduction

There exists two main methods of forensic analysis for comparing uncharacterized human DNA profiles against reference DNA profiles for identification.

1. Short tandem repeats (STRs) are amplified, allele sizes are determined, and the results are used to match against databases. A suspect's DNA profile is compared against a sample to look for STR length similarity.
2. Mitochondrial DNA, as a stable maternal line of DNA is used to designate a subject as excluded, inconclusive, or cannot exclude.

When using these methods to identify people, the likelihood of falsely matching an innocent suspect, or the probabiliy of a random man not excluded $P(RMNE)$ is an important factor that needs to be better quantified, as current statistics place this number to be one in a billion to one in a thousand, making it an unreliable estimator for forensic use.

The purpose of this study was to estimate $P(RMNE)$ for mixtures ranging from 2 to 10 contributing individuals, sampling from a large collection of simulated DNA profiles. For each study, $P(RMNE)$ was calculated as a function of population-averaged minor-allele frequences (mAFs). This paper demonstrates a confidence interval of 99% for a mAF of 0.075 for all $P(RMNE)$ calculations. This study considers the following on $P(RMNE)$ calculations:

1. SNP panel size
2. Multiple relatives' DNA present in a mixutre
3. Population-specific SNP bias
4. Fixed genotyping error, binomially distributed among alleles

10^{-14} to 10^{-11} error may not affect a jury's interpretation of a crime, but 10^{-9} to 10^{-6} may affect the admissibility of DNA evidence.

2 Methodology

A panel of 480 SNPs to analyze complex mixtures is created.
Lots of science stuff I don't understand.

3 Results

The $P(RMNE)$ was evaluated for a panel of 480 SNPs. At each location, a minor allele ratio is calculated as the number of observed minor allele calls divided by the total number of calls. Only those with a minor

allele ratio of 0.01 are used for the $P(RMNE)$ calculation. This group was able to achieve a $P(RMNE)$ within the courts admissability guidelines for between 0 and 10 allele mismatches in a 3-5 person mixture. Outisde of these values, whether by adding more people to the mixture or increasing the number of allowed allele mismatches, the value for $P(RMNE)$ became too large to be useful.

4 Conclusion

This group developed a method for evaluating the exact probability of random man not excluded for SNP mixture analysis by including allele specific population frequences into the probability calculation. Still, with current methods, 5 of the 9 sequences tested fell below the one in a billion requirement for DNA admissability.