Journal of Forensic Research
ISSN: 2157-7145

Ge, J Forensic Res 2012, 3:8
DOI: 10.4172/2157-7145.1000e112

Editorial Open Access

DNA Based Familial Searching and Related Statistical Issues

Jianye Ge*

Department of Forensic and Investigative Genetics, University of North Texas Health Science Center at Fort Worth, 3500 Camp Bowie Blvd, Fort Worth, TX 76107, USA

DNA databases have become essential tools for forensic investigations. The typical DNA database search seeks direct matches between a forensic profile from crime scene evidence and an offender profile in the database. However, direct matches may not be obtained because the database does not contain all people in the population. As of January 2012, there are about 10.5 million offender profiles in the US National DNA Index (NDIS) [1], which only covers about 3% of the US population (i.e., there are more than 313 million people in the US [2]). To extend the investigative lead value of current databases, an additional approach, familial searching, has been used to determine the source of a forensic sample by searching the database for possible relatives (e.g., parents, children, full-sibs) of the true source of the sample. It also is an effective way to expand the utility of DNA databases with relatively low costs. This approach, when performed under stringent criteria, has been used to develop strong investigative leads and to successfully identify perpetrators of crimes [3-5].

There are a number of strategies to perform familial searching and varying methods to assess the strength of candidate associations from such searches. Generally, familial searching is carried out based on either an Identity-by-State (IBS) approach, which simply compares the number of shared alleles or loci between the forensic profile and the offender profiles, or likelihood ratio (LR) approach, which compares the probabilities of the forensic and offender profiles given that the donors are related (e.g., parent-child or full-sib) versus being unrelated. Studies have been carried out to evaluate the performance of familial searching strategies under ideal conditions [6,7]. More realistic issues in familial searching practices, such as profiles with mixtures, distant relatives, and multiple possible ancestry populations, need to be investigated further.

Mixture profiles are commonly observed in forensic samples. Chung et al. [8,9] described a method which can search the relatives of the unknown contributor in a two-person mixture with one contributor known. As expected, in this particular scenario, the accuracy of familial searching is lower than but close to that of single source profile searching, since the known contributor limits the possible alleles of the other contributor in a two-person mixture. This method can be extended further to allow for more unknown contributors.

Attempting to address the LR variation among reference populations, SWGDAM's [10] recommendation suggested that the maximum and minimum EKRs among Caucasians, African Americans, southwestern Hispanics, and southeastern Hispanics should be greater than 1 and 0.1, respectively. Ge et al. [6,11] estimated the LR variation distributions among the populations and compared the false negative/positive rates of the searching strategies by (1) The minimum LR among the populations or (2) SWGDAM strategy, and suggested that SWGDAM strategy might not be the best strategy for familial searching.

Because of different relationships and varying number of loci in profiles that will be searched, at least two candidate lists could be generated with multiple KI measures, i.e., KIs for parent-child or full-sib with 13 or 15 loci (or 10 loci if bounded by the minimum number of loci for an evidence profile). However, the KI values from different relationships or different sets of markers cannot be compared because they are calculated under different frameworks [11]. The profile-dependent or false rates based familial searching strategy proposed by Slooten and Meester [12] could be a good approach to merge the candidate lists from different frameworks, although more sophisticated approaches need to be developed.

When lineage markers (e.g., Y-STR and mtDNA sequences) are available, the efficiency of familial searching can be substantially improved [6,7,13] and distant relationships (e.g., half-sib, cousin) may be identified with relatively high confidence. Ge et al. [13] also suggested including Y-STRs into the core loci of CODIS to provide a higher capability of kinship analysis. Generally, more studies need to be done to estimate the statistical power of identifying relationships with varying numbers of Y-STR loci and/or varying regions of mtDNA sequences. Additionally, the assumption of independence among the autosomal STRs, Y-STR, and mtDNA still needs additional investigation, although the general belief is that these marker groups can be treated as independent for large populations, but may not be so for small, isolated populations.

To facilitate familial searching in casework, more work needs to be done to design more efficient strategies, with more realistic and complicated factors, that will enhance the capabilities of familial searching software and eventually provide forensic scientists a robust set of guidelines on familial searching and accompanying software tools. With new developments we look forward to assisting investigations with this very powerful tool of DNA typing and familial searching.

References

1. http://www.fbi.gov/about-us/lab/codis/ndis-statistics
2. http://www.census.gov/main/www/popclock.html
3. Bhattacharya S (2004) Killer Convicted Thanks to Relative's DNA.
4. Grim Sleeper case.
5. http://www.denverda.org/DNA/Familial_DNA_Database_Searches.htm
6. Ge J, Chakraborty R, Eisenberg A, Budowle B (2011) Comparisons of familial DNA database searching strategies. J Forensic Sci 56: 1448-1456.
7. Myers SP, Timken MD, Piucci ML, Sims GA, Greenwald MA, et al. (2011)

***Corresponding author:** Jianye Ge, Institute of Investigative Genetics, Department of Forensic and Investigative Genetics, University of North Texas Health Science Center at Fort Worth, 3500 Camp Bowie Blvd, Fort Worth, TX 76107, USA, Tel: 817 735 2773; E-mail: Jianye.Ge@unthsc.edu

Received August 20, 2012; **Accepted** August 21, 2012; **Published** August 24, 2012

Citation: Ge J (2012) DNA Based Familial Searching and Related Statistical Issues. J Forensic Res 3:e112. doi:10.4172/2157-7145.1000e112

Searching for first-degree familial relationships in California's offender DNA database: Validation of a likelihood ratio-based approach. Forensic Sci Int Genet 5: 493-500.

8. Chung YK, Fung WK, Hu YQ (2010) Familial database search on two-person mixture. Computational Statistics & Data Analysis 54: 2046-2051.
9. Chung YK, Fung WK (2012) Identifying contributors of two-person DNA mixtures by familial database search. Int J Legal Med.
10. Scientific Working Group on DNA Analysis Methods Ad Hoc Committee on Partial Matches (2009) SWGDAM Recommendations to the FBI Director on the "Interim Plan for the Release of Information in the Event of a 'Partial Match' at NDIS", Forensic Sci Comm 11.
11. Ge J, Budowle B (2012) Kinship index variations among populations and thresholds for familial searching. PLoS ONE 7: e37474.
12. Slooten K, Meester R (2012) (under review) Database likelihood ratios and familial DNA searching, arXiv: 1201.4261v2.
13. Ge J, Eisenberg A, Budowle B (2012) Developing criteria and data to determine best options for expanding the core CODIS loci. Investig Genet 3: 1.