A109 Choosing Relatives for DNA Identification of Missing Person Identification

Jianye Ge, PhD*, University of North Texas Health Science Center, Institute of Investigative Genetics, 3500 Camp Bowie Boulevard, Fort Worth, TX 76107; Bruce Budowle, PhD, University of North Texas Health Science Center, Forensic & Investigative Genetics, 3500 Camp Bowie Boulevard, EAD 310, Fort Worth, TX 76107; and Ranajit Chakraborty, University of North Texas, Health Science Center, Institute of Investigative Genetics, 3500 Camp Bowie Boulevard, EAD 310, Fort Worth, TX 76137

After attending this presentation, attendees will be informed of the information content that can be useful with different combination of relatives, which relatives would be most informative, and how many relatives are sufficient for DNA missing person and other kinship analysis identifications.

This presentation will impact the forensic science community by comparing the information content of different relative combination scenarios for kinship analysis and providing practical guidance to select the most informative relatives for missing person identification, especially when resources are limited.

DNA-based analysis is integral to missing person identification cases. When direct references are not available, indirect relative references can be used to identify missing persons by kinship analysis. The DNA Commission of the International Society for Forensic Genetics (ISFG) suggested calculating posterior odds for identification and setting 99.9% as a degree of confidence. There are two ways to increase the power of identification: (1) type more markers; and, (2) type more relatives. In many cases, the quality and quantity of DNA is poor. Thus the number of markers that can be typed will be limited by the quality and quantity of DNA derived from remains. Increasing the number of reference relatives can increase the chances of identifying remains and particularly for challenged samples. However, typing all relatives of a large pedigree can be costly and may not be necessary to reach a defined threshold for identification. At times decisions may need to be made on which relatives to type. Since there are information and cost factors regarding the selection and number of relatives, some selection criteria should be considered to guide identity testers.

In this study, the 37 most common relative combination scenarios (e.g., both parents, three children, one child plus one parent plus spouse,

one parent plus three fullsibs, one child plus one parent, two unrelated uncles, etc.) in missing person identifications were selected and using the 13 CODIS STRs as genetic profile data, large numbers of pedigrees (e.g., 1,000,000) were simulated for each scenario. The distribution of LRs for each scenario was evaluated to first confirm the well-known single relative reference scenarios and second to determine the most informative combinations of relatives for identifying an unknown person. Thus guidance is given on which and how many relatives should be selected and typed for kinship analyses for identification so that efficiency can be optimized under the constraints of limited resources. The LRs were calculated by a software program MPKin, which jointly considers DNA profiles from all available family members and missing persons/remains.

Based on the simulation results, the following guidelines are recommended in choosing relatives for missing person identification.

Parents are the preferred relatives and both parents of the missing person should be typed when possible. The expected LRs of single parent and both parents are about 104 and 1010, respectively. If both parents are typed, all other relatives, including fullsibs, may not be necessary.

Children are the second preferred relatives. Type as many children as possible or until the unknown genotype of the missing person can be reconstructed. The expected LRs of two, three, or four children are about 1×10^7, 1×10^9, or 2.5×10^{10}, respectively. In absence of parents, if the missing person is male, sons are preferred because of the same Y- chromosome shared between father and sons; otherwise, sons and daughters are equivalent.

Even if a child is available, the spouse of the missing person (i.e., the father/mother of the child) should be considered for typing, if he/she is available. The LR can increase more than 100 fold, if the spouse of the missing person is further typed, for a scenario with a single child already available.

Fullsibs are the third preferred relatives. A single fullsib may not be informative, but typing three or four fullsibs with expected LRs 3.3×10^7 or 3.5×10^8 are sufficient for reliable identification. If the missing person is male, brothers are preferred compared to sisters, because the missing person and brothers can be tested for both Y-chromosome and mtDNA, and it is reasonable to type less relatives with the same discrimination power due to economical reasons in some conditions; otherwise, brothers and sisters are equivalent.

All other distant relatives, such as grandparents/grandchildren, halfsibs, uncles/aunts, and cousins, only provide limited identification capabilities based on autosomal markers, but their Y-chromosomes and mtDNA can be used to increase the LR or filter out false relationships.

Less genetic dependence between reference relatives provides a higher LR on average. This is practiced routinely for standard paternity cases where two unrelated parents are sought. But the concept can

be applied to extended pedigrees as well. For example, two biologically unrelated uncles (with an expected LR = 100) can be more informative than two related uncles (with an expected LR = 28).

With limited number of relatives, type as many markers as possible.

Kinship Analysis, Missing Person, Pedigree Likelihood Ratio