A165 A Study of the Effects of Database Size and Loci Number on the Potential for Successful Familial Searches in Databases of STR Profiles

Rick W. Staub, PhD, Orchid Cellmark Incorporated, 13988 Diplomat Drive, Suite 100, Dallas, TX 75234; Dmitry Maschenski, MS, Orchid Cellmark Incorporated, 13988 Diplomat Drive, Suite 100, Dallas, TX 75234; Carlos Baena, BS, Orchid Cellmark Incorporated, 13988 Diplomat Drive, Suite 100, Dallas, TX 75234; Pedro Reinoso, BS, Orchid Cellmark Incorporated, 13988 Diplomat Drive, Suite 100, Dallas, TX 75234; Sid Sinha, MBA, Orchid Cellmark Incorporated, 13988 Diplomat Drive, Suite 100, Dallas, TX 75234*

After attending this presentation, attendees will understand the logic used to perform familial searches and how results are influenced by methodologies, database size, and genetic loci utilized.

This presentation will impact the forensic science community by clarifying potential problems involved with familial searches and how to best overcome them through the use of optimal scientific methods.

Over the past several years, the utilization of familial searching tools by law enforcement to locate relatives of the perpetrators of violent crimes has received a good deal of publicity in the media. This has initiated a good deal of public debate over potential violation of civil rights of those individuals in DNA databases who are identified by familial searches as "persons of interest." There is a legitimate potential for several individuals completely unrelated to perpetrators of violent crimes to arise as persons of interest when database searches are performed, particularly if the searching technique used does not have a strong discrimination power for locating relatives. Even when optimal methodologies are used to carry out familial searches, the number of false positive and false negative results in familial searches will be affected by the number of loci used and the number of profiles in searched databases.

A generally accepted methodology for assessing putative kinship is via the utilization of a Kinship Index (KI), which is a likelihood ratio comparing the likelihood of obtaining genetic results from individuals that are biologically related to the likelihood of obtaining the same genetic results from unrelated individuals. A less robust method of assessing kinship is via allele sharing (identical by descent in biological relatives vs. identical by chance in non-related individuals). In this study we compute KIs and number of shared alleles in comparisons of STR profiles from simulated evidence items to databases of varying size that also contain profiles of known relatives of the donor of the simulated evidence profile. In making these comparisons, the number of loci in the evidence profile is also varied to obtain data on the likelihood of successful searches based on loci used in the search. These data provide the practitioner of familial searches with a reasonable framework within which to establish guidelines to optimize the chances for carrying out successful searches in the future. Additionally, they illuminate the impact of database size on the frequency of false positive hits when performing familial searches. This presentation will examine possible means to reduce false negative and false positive hits that arise, such as the comparison of Y-STR haplotypes for male samples in databases, or comparison of most likely race of evidence and database profiles.

All databases used in these studies were generated from simulated profiles based on an algorithm that utilizes input STR allele frequencies from published data. Profiles of biologically related individuals for search purposes were either obtained from families of known racial/ethnic origin or generated from profiles in databases using rules of Mendelian inheritance.

Familial Search, DNA Profile Database, Y-STR Haplotype