B199 A Statistical Approach to the Identification of Criminals and Mass Disaster Victims Through Kinship DNA Database Searches

Kristen E. Lewis, MS, University of Washington, Department of Genome Sciences, Seattle, WA 98195-7730; Bruce S. Weir, PhD, University of Washington, Department of Biostatistics, Seattle, WA 98195-7232; and Mary-Claire King, PhD, University of Washington, Department of Genome Sciences, Seattle, WA 98195-7720*

After attending this presentation, attendees will learn how to use familial searching/genetic kinship analysis to generate investigative leads that would otherwise go undetected by the current database searching methods.

This presentation will impact the forensic community and/or humanity by providing new methodology to aid in the identification of suspects in unsolved cases and victims of mass disasters or human rights abuses.

The typical use of state and federal forensic DNA databases (i.e., Combined DNA Index System or CODIS) for comparison of unknown biological evidence profiles with known offender profiles has increased dramatically in the past several years as the number of database profiles grows and as it becomes legal to include misdemeanor offenders and arrestees in these databases. However, in certain situations encountered by forensic scientists and law enforcement, a direct search of local, state, or national CODIS databases will not produce a hit if the contributor of the evidence either has not been convicted of a felony or has not been arrested in a state that allows arrestee testing. Without any other evidence in these cases, the investigation and prosecution will stall.

One way to generate investigative leads in these situations is to perform indirect searches of the CODIS database. In this process, the unknown evidence profile is searched against the CODIS database to identify possible close relatives of the true offender, as these individuals would share more alleles with the true criminal than an unrelated person would. This type of search, called familial searching or genetic kinship analysis, has been routinely used to identify victims of mass disasters, human rights violations, and missing persons cases. Most states do not have legislation that explicitly addresses whether or not evidence profiles can be searched against the database to find partial profile matches with individuals who may be close relatives of the person from which the evidence came.

Likelihood ratio calculations can be used to assess the probability of identifying a close relative (i.e., parent-offspring, full or half-siblings, or uncle-nephew) depending on the hypothesized relationship by comparing the frequency and number of alleles shared between the evidence and each database profile (Brenner and Weir, 2003; Bieber et al., 2006). The largest likelihood ratios are typically the true relative, but unrelated individuals may generate high likelihood ratios by chance.

This presentation will describe a model-based approach that simulates a familial search of forensic databases by using FBI data composed of 13 CODIS STR genotypes from six populations (Budowle and Moretti, 1999). Each population database contains approximately 200 individuals. The allele frequencies at each CODIS locus are calculated for each population database. In the parent-offspring scenario, the model simulates a search of an evidence sample profile ("child") against all persons in the database, which includes the child's true "parent." To do this, two individuals from one database are randomly chosen to be hypothetical parents. Their profiles are used to generate a child's profile according to Mendelian rules of inheritance. This child's genotype is searched against every offender's profile in the database including his parents', and the likelihood that the two compared individuals have a parent-child relationship versus that they are unrelated is calculated using the allele frequency data. The likelihood ratios are rank ordered and the position of one of the true parents (randomly selected) is recorded. This process is repeated to generate 1,000 likelihood ratios and corresponding ranks of one of the true parents. This information can be used to identify the probability of finding a true parent in a forensic database based on the use of likelihood ratios. A similar methodology will be presented for the far more complex sibling and uncle-nephew cases.

Familial Search, CODIS Database, Likelihood Ratio