



E34 The Time Has Come to Analyze DNA Profile Databases

Dan Krane, PhD, 3640 Colonel Glenn Highway, Dept Bio Sci, Dayton, OH 45435; and Roger G. Koppl, PhD*, Syracuse University, Dept of Finance, 721 University Avenue, Syracuse, NY 13244*

After attending this presentation, attendees will have a better understanding of the need for disclosure of the National DNA Index (NDIS) database.

This presentation will impact the forensic science community by increasing awareness of the questions that could be addressed by an independent, transparent analysis of the profiles contained in the NDIS database.

The Federal Bureau of Investigation (FBI) has controlled the NDIS since it was established in 1994. Even though the federal legislation that established NDIS explicitly anticipated that records within the database would be available for purposes of research and quality control, the FBI has published no research derived from NDIS and has vigorously declined to disclose these records to academic scholars for the past 20 years. The time has come for the FBI to release anonymized NDIS profiles to academic scientists.

Analyses of NDIS profiles would allow independent evaluation of some of the population-genetic assumptions used to generate statistical weights for DNA profile matches with samples large enough to allow sensitive evaluation of population structure. Databases that have been made available for such analyses at the present time are relatively small collections of hundreds of individuals while NDIS presently contains more than 10,000,000 complete 13-loci STR DNA profiles. The relative utility of local, state, and nationwide allele frequency databases could also be assessed for the very first time.

Other analyses that thus far have only been possible with simulations could finally be performed with real-world data. Those analyses include (but are not limited to) determination of: the frequency with which three-person mixtures actually appear to be two-person mixtures; the effect of close relatives in a database upon match probabilities, and; the extent to which identity by descent causes clustering of DNA profiles. In a different vein, quality assurance analyses of a government database from Victoria, Australia, containing 15,021 9-loci STR DNA profiles found an error rate of approximately 1 in 300 and raised concerns about missed opportunities to develop investigative leads that could easily apply to the NDIS database as well.

Disclosure of the 13-loci STR DNA profiles in NDIS would not trespass on privacy interests in any meaningful way. The U.S. Supreme Court's recent ruling allowing for the inclusion of arrestees DNA profiles to be entered into databases like NDIS in part relies upon the fact that the 13 STR loci that are genotyped are non-coding and, therefore, cannot reveal sensitive information. Copying of the NDIS database itself should not be overly burdensome and would probably be accomplished in a matter of minutes.

Open access to data is a fundamental tenet of science. The need for openness in forensic science in particular was reinforced by a National Research Council report five years ago that decried the insularity of forensic science and called for greater involvement of the academic community in assessment, validation, and improvement of forensic science methods. The FBI should honor the norms of science and open the NDIS database to independent scientific scrutiny. Doing so poses no meaningful risk and can only strengthen the quality of forensic DNA analysis.

Database, DNA, Quality Control