

## **EXTRACTING INTELLIGENCE FROM OTHERWISE UNINTERPRETABLE DNA MIXTURES**

Karen E. Olson, Matthew J. Bohn, Jason J. LeBlanc, Michael J. Salyards, and Richard E. Tontarski Jr., Defense Forensic Science Center

DNA mixture interpretation continues to be a challenge with traditional, capillary electrophoresis (CE) based DNA analysis. Stochastic effects including heterozygote imbalance, drop-out, drop-in, stutter, and variation between replicate samples are compounded in mixtures due to contributor ratios, quantity and integrity of total DNA in the mixture and from each contributor, as well as the amount of allele sharing between contributors. The need to accurately interpret DNA mixtures has spawned various deconvolution methods and software to assist with the interpretation and prediction of genotypes. However, many mixtures generated from evidence items are unusable because they are too complex to conclusively identify one or more profiles. Although it may never be possible to fully deconvolve these complex mixtures with current short tandem repeat (STR) typing technologies, it might still be possible to leverage information and develop investigative leads based on the electropherogram data. We are developing a principal component analysis (PCA) algorithm that clusters DNA mixture profiles based on the presence of a single common contributor together with a K-means clustering algorithm to separate them from mixtures that do not share that same individual's profile. To utilize this analysis process, STR DNA profiles are generated following traditional, CE-based methods. The allele calls, including peak-heights, are imported into the clustering algorithm which assigns a weight based on the relative frequency and peak height of the allele. The algorithm then maximizes the differences between the composite, weighted mixture profiles, causing the mixtures sharing a common contributor to become grouped. Those mixtures are separated from unique, random mixtures that do not contain that individual. Success has been observed with clustering 3-person mixtures at various contributor ratios. We are currently working to determine the maximum number of contributors and common contributor ratio limits that will still allow accurate clustering. While this type of analysis cannot determine the identity of the common contributor, it may help investigators determine if crimes or illicit activities committed in the same area or in similar manners, for example burglaries, gang violence, or improvised explosive device construction, may have been committed or assisted by the same individual, which can shape investigations and prosecution strategies.

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense. Names of commercial manufacturers or products included are incidental only, and inclusion does not imply endorsement by the authors, DFSC, OPMG, DA or DoD.