EXPERT SYSTEM SOFTWARE FOR EVALUATING AUTOMATED STR GENOTYPING RESULTS

T. Matsumoto[1], R. Nakashige[1], and S. B. Lee[2, 3]
[1]Hitachi Software Engineering Co., Ltd
[2]Hitachi Software Engineering America, Ltd. MiraiBio Division
[3]Forensic Science Program, Justice Studies Department, San Jose State University

In forensic science, various improvements in experimental techniques have achieved microsatellite results with fewer artifacts. For example, the tetranucleotide repeat loci included in the CODIS core loci have small amounts of stutter, and a new PIG (Prostate Investigation Group) tailing protocol was developed to distinguish true allele peaks from artifacts. Algorithms for automated genotyping are also able to achieve high accuracy rates.

Advances in both hardware and software automation have led to the ability to rapidly genotype large numbers of samples. However, the larger numbers of samples requires time-consuming manual review and interpretation following automated genotyping. Here, we propose a displaying technique to facilitate reviewing allele calls.

Although software systems in forensic science provide plot windows for visual assessment, analysts are required to examine each electropherogram. Therefore manually examining all electropherograms is difficult because of the large amount of time required. Some software systems have algorithms to detect suspicious electropherograms, however, these may not detect all samples requiring review.

Our technique enables users to review many electropherograms at-a-glance by categorizing individuals based on their genotypes and viewing electropherograms according to fragment size and peak height. Because most electropherograms show the same pattern by our normalization, analysts can easily sort out the small number of unusual electropherograms. After that, the raw data of these samples can be reviewed in detail.

With our technique, analysts only have to examine specific sections of the electropherogram, similar to a rule-based error detection approach. Furthermore, manual reviewing of all data in

the compiled format display is expected to enable reducing or eliminating overlooking samples that require secondary review.