

## **Adaptive DNA Databank Capacity using an Expert System Process**

Peter Wistort<sup>1</sup>, Thomas Leach<sup>1</sup>, Mark Perlin, MD, PhD<sup>2</sup>, and Barry Duceman, PhD<sup>1</sup>

<sup>1</sup>New York State Police, Forensic Identification Center, Albany, New York

<sup>2</sup>Cybergenetics Inc., Pittsburgh, PA

DNA databank laboratories handle an uncertain number of samples each year. In routine operation, the New York State Police (NYSP) Forensic Identification Center (FIC) processes around 5,000 samples each month within 60 days of arrival. However, changes in DNA law can create new backlogs. When the NYS legislature recently introduced a stronger DNA databank requirement for certain felonies and misdemeanors, 60,000 additional DNA samples had to be added within one year.

To cope with uncertain DNA demand, the NYSP FIC has developed an adaptive capacity DNA databank process that uses an expert system for rapid and accurate data review. The predictable DNA samples undergo biological STR analysis in the state DNA databank laboratory. The variable samples (e.g., from new legislation) are sent to an outside vendor laboratory, often subsidized by federal NIJ backlog funding. The DNA sequencer data from both laboratories are then reviewed in the NYSP databank unit using an NDIS-approved expert system.

The NYSP workflow begins with the arrival of an offender DNA sample at the FIC. These samples are logged into the NYSP laboratory database, and are routed for in-house or vendor processing. In-house samples undergo DNA extraction, quantitation, amplification and size separation in our laboratory. In the vendor pathway, DNA samples are sent out, and DNA sequencer files along with the vendor genotype results are returned.

The generated DNA sequencer .fsa files (from the FIC or vendor lab) are packaged into DataDisks for TrueAllele<sup>®</sup> Databank expert system review. The TrueAllele analysis process transforms the original .fsa files into quality checked data, providing a quality assessment of the size standards, allelic ladders, positive and negative controls. An analyst spends 1-2 minutes per plate (of 96 samples and controls) reviewing these quality results. Then, the TrueAllele Interpretation process quantifies every STR peak, designates the genotypes, and applies 25 quality check rules to the locus data of each sample. An analyst reviews the potentially problematic data, spending 6-8 minutes per 96-well sample plate.

Accepted DNA profiles are then formatted by TrueAllele into a computer file for NDIS upload. Rejected DNA sample information is relayed back to the DNA laboratory (in-house or vendor), along with computer generated rework instructions for reprocessing. The vendor-provided allele calls are compared with the expert system results using the TrueAllele AutoValidate<sup>™</sup> module, providing additional quality check information. Vendor samples that are not accepted by the expert system are manually reviewed using conventional genotyping software.

The NYSP expert system process provides adaptive capacity for high quality DNA databank review. Two FIC DNA analysts typically review databank data for several hours each day, using the TrueAllele expert system. The two analysts can review up to 1,000 DNA samples per day, from both in-house and vendor data. Internal checks on 10% of the samples show no discernable error in the FIC review process. These productivity improvements let the databank unit focus on process quality, rather than on data review.