

# National Institute of Justice

**Award Title:** Development of Reference Sample DNA Profiling for Databases Using Next Generation Sequencing Technologies

**Award Description:**

The potential of DNA typing for developing investigative leads and for solving future crimes came to fruition with the development of DNA databases, such as CODIS. The number of reference samples from convicted felons, arrestees, detainees, and missing persons continues to increase and there is no indication of the demand subsiding. To meet the needs of forensic DNA typing for developing investigative leads and its infrastructure new technologies are sought. One challenge to resolve is the selection of markers that should be used routinely by forensic laboratories. A core set of STRs was selected fourteen years ago. Recently, the FBI recommended that the core loci for CODIS should be change and augmented. This strategy has limitations because loci selected should be driven by the demands of casework, i.e., loci should be selected based on performance with degraded and inhibited samples or that the technology can be more versatile to enable a variety of search strategies. However, these concerns can be rendered moot with the advent of next generation sequencing (NGS). NGS provides sequencing data with unprecedented capacity and speed at a reduced cost on a per nucleotide basis. The technology is evolving and currently does not offer the sensitivity of detection to analyze low quantity and quality DNA samples. However, NGS is sufficiently robust to type reference samples for loading DNA profiles into CODIS. A large battery of markers can be analyzed simultaneously, far exceeding the current capacity of 15-21 STRs in commercial kits. It is conceivable that all forensically-relevant identified autosomal STRs, such as the 24 STR loci selected by the FBI and beyond, a set of Y STRs and X STRs, and human identity SNPs (400-500 markers) can be typed simultaneously. Many different samples can be typed simultaneously with barcoding. Initially, mtDNA would have to be typed separately, because of its sheer number of copies compared with nuclear DNA. Data from evidence samples can be compared among the majority (if not all) of the reference. The inclusion of a more comprehensive set of markers for reference samples will overlap all current databases (i.e., legacy data) and foster investigations. The basis of this proposal is that NGS, with its economies of scale, can provide a system such that reference samples can be typed economically for a large battery of markers and eventually, if commercialized, could exceed a cost benefit compared with current costs for typing a modicum of autosomal STRs. The primary goal of this project is to develop the capability of typing reference samples for a large battery of autosomal, Y chromosome, and X chromosome STRs and human identity SNPs in a single multiplex analysis. In addition mtDNA sequencing will be converted to that of a NGS platform approach. The methodologies shall provide sufficiently robust data so that profiles can meet criteria for uploading to CODIS. Two NGS platforms will be used: Genetic Analyzer IIx and Ion Torrent Personal Genome Machine. DNA samples will be probed-captured, libraries generate and sequenced on one or both NGS platforms. The analyses will be run at different depths of coverage. The genetic data will be collected and assessed for the following: 1) completeness of typing the full batter of markers; 2) consistency with existing typing strategies, predominately autosomal and Y STRs and mtDNA; 3) read length and impact on successful typing or failures; 4) depth of coverage required for a reliable result; 5) subset of reads that span an entire STR allele and minimum reads required for accuracy; 6) software requirements; 7) labor; 8) workflow; and 9) cost per sample. ca/ncf

<b>Awardee Name:</b> University of North Texas Health Science Center	<b>Award Number:</b> 2012-DN-BX-K033
<b>Solicitation Title:</b> NIJ FY 12 Applied Research and Development in Forensic Science for Criminal Justice Purposes	<b>Fiscal Year:</b> 2012
<b>Amount:</b> \$747,797.00	<b>Earmark:</b> No
<b>Recovery Act:</b> No	<b>State/Territory:</b> TX
<b>County:</b> Tarrant	<b>Congressional District:</b> 12
<b>Award Status:</b> Open	