# SHORT TANDEM REPEAT TYPING ON THE 454 PLATFORM: STRATEGIES AND CONSIDERATIONS FOR TARGETED SEQUENCING OF COMMON FORENSIC MARKERS

Melissa Scheible[1,2,*], Odile Loreille[1,2], Rebecca Just[1,2] and Jodi Irwin[1,2,3]
[1]American Registry of Pathology
[2]Armed Forces DNA Identification Laboratory, Armed Forces Medical Examiner System
[3]Present affiliation: Federal Bureau of Investigation, Quantico

The past several years have seen a dramatic advance in the methods, chemistries and detection platforms available for DNA sequence data generation. Next generation sequencing (NGS) technologies, which produce large volumes of sequence data at extremely low cost relative to current platforms, are being broadly applied to various questions in medical genetics, evolutionary biology, molecular anthropology, phylogeny, epidemiology and metagenomics. For many of these applications, NGS is being used to produce sequence data covering thousands of loci, or even entire organismal genomes in a single sequencing run. Given this capacity, it's not difficult to envision the potential implications of this technology for criminalistics, missing persons and disaster victim identification purposes. Historically, the recovery of large numbers of forensic markers in a single assay has been restricted by both the technical limitations of current, established capillary based sequencing genotyping technologies, as well as the quality and quantity of DNA originating from the damaged and degraded specimens regularly encountered in forensic casework. These limitations don't apply in quite the same way to NGS, however. As a result, the simultaneous recovery of the standard autosomal DNA, mitochondrial DNA, and X and Y-chromosomal markers regularly assayed in forensic genetics, along with additional markers of interest, may be possible with these new technologies [1].

To investigate the feasibility of next generation sequencing technology (NGS) for the multiplex detection and sequence production of short tandem repeats (STRs) from degraded and low DNA quantity samples, standard polymerase chain reaction amplification methods were used to enrich for commonly employed STR markers. Samples were amplified with two multiplexing strategies: a multiplex containing thirteen miniSTR markers and a series of multiplexes containing four miniSTR markers each. Each sample multiplex was barcoded with a sample-specific multiplex identifier (MID) for subsequent parallel tagged sequencing on the GS Junior System (454 Life Sciences, a Roche company, Branford, CT). The next generation sequencing results from these experiments, which encompassed over fifty DNA extracts representing both pristine samples and low-quality evidentiary specimens, reflected known genotypes and displayed consistency across multiple extracts and/or amplifications of the same sample. Furthermore, the NGS data revealed sequence variation not discernible with standard capillary electrophoresis-based detection alone, yielding information that is likely to be particularly valuable in missing persons and disaster victim identification cases for which only partial profiles are recovered and/or only distant kin are available as references. While our results clearly highlight the potential of NGS for sequencing STRs from evidentiary specimens, our data also reveal the complexities of NGS-based STR typing, both in terms of the laboratory assays themselves as well as the downstream data processing and analysis.

*The opinions and assertions contained herein are solely those of the authors and are not to be construed as official or as views of the US Department of Defense, its branches, the U.S. Army Medical Research and Materiel Command, the Armed Forces Medical Examiner System, the Federal Bureau of Investigation, or the U.S. Government.*

[1] B.L. Hancock-Hanser, A. Frey, M.S. Leslie, P.H. Dutton, F.I. Archer, P.A. Morin, Targeted multiplex next-generation sequencing: advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics, Mol.Ecol.Resour. 13 (2013) 254-268.