HIGH SENSITIVITY MULTIPLEX SHORT TANDEM REPEAT LOCI ANALYSES WITH MASSIVELY PARALLEL SEQUENCING

Xiangpei Zeng[1], Jonathan L. King[1], David H. Warshauer[1], Bobby L. LaRue[1], Antti Sajantila[1,2], Jaynish Patel[3], Doug Storts[3], Bruce Budowle[1,4]

[1]Institute of Applied Genetics, Department of Molecular and Medical Genetics, University of North Texas Health Science Center
[2]Department of Forensic Medicine, Hjelt Institute, University of Helsinki
[3]Promega Corporation
[4]Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University

Short tandem repeats (STRs) are the primary genetic markers used in forensic DNA human identification testing, due to their high discrimination power and relatively short amplicon size. Currently, capillary electrophoresis (CE) and laser optics are used to separate amplicons based on size and detect the fluorescently-labeled amplicons, respectively. However, CE-based methods have some limitations in the forensic field: for example, a limited number of STR loci (25-30 maximum) can be typed in a multiplex system; and output data cannot distinguish some stutter products, dye artifacts, and low-level alleles.

Massively parallel sequencing (MPS) technologies allow for a substantial increase in throughput and depth of coverage at a relatively-affordable price and thus now can be considered a viable technology for application-oriented forensic laboratories. With the advent of MPS, some of the CE-based limitations may be overcome. MPS generates both allele lengths and genetic sequences potentially increasing discrimination power and providing more information to distinguish stutter products from minor contributor alleles in mixtures. Because locus assignment is not based in part on size, as is required by CE methods, STR amplicons can be engineered to be smaller in size to be better suited for analyzing challenged samples, they can be more similar in size due to the lack of spectral limitations inherent with CE methods, and thus many more loci can be analyzed simultaneously.

In this study, a prototype multiplex STR System (Promega) containing 17 STR loci, was amplified by PCR. Library preparation was performed using the TruSeq DNA LT Sample Preparation Kit (Illumina); sequencing was carried out on the MiSeq (Illumina; 2×250bp); and STRait Razor was used for allele calling. The system was tested for amplicon size selection, required amounts of PCR product for library preparation, and sensitivity of detection (based on input DNA). MinElute PCR Purification Kit (Qiagen) was demonstrated to be the better size selection method compared with diluted bead mixtures. The library input sensitivity study showed that a wide range of amplicon product could be used and no notable differences in STR profile were observed. Results on 24 different individuals showed that a high depth of coverage and balanced heterozygote allele coverage ratios can be obtained with input DNA (≥250 pg) similar to that typically used in CE-based systems. Moreover, an input DNA sensitivity study showed that complete STR profiles could be generated from as little as 62 pg of input DNA and nearly complete (i.e., limited allele drop out) profiles were observed with as little as 16 pg of input DNA. These studies indicate that this STR multiplex and the Illumina MiSeq system can generate reliable DNA profiles at a sensitivity of detection level that rivals current CE-based approaches.