DEVELOPMENT OF A NUCLEAR SNP PROBE CAPTURE ASSAY FOR MASSIVELY PARALLEL SEQUENCING OF DEGRADED AND MIXED DNA SAMPLES

Nikhil Bose[a,b], Katie Carlberg (MD)[c], George Sensabaugh (D.Crim)[b,d], Henry Erlich (PhD)[a], Cassandra Calloway (PhD)[a,b]
[a]Children's Hospital Oakland Research Institute (CHORI
[b]Forensic Science Graduate Program
[c]Children's Hospital Oakland
[d]University of California, Berkeley, School of Public Health

DNA extracted from biological forensic evidence can be extremely fragmented and in very low quantity. In such cases, chances of allelic dropouts are high in STR analysis because STR analysis requires large DNA fragments with intact primer binding sites. One approach to analyzing degraded DNA is SNP typing since only a single base variation is analyzed. However, the electrophoretic technologies previously used for SNP typing targeted a limited number of SNPs. New technologies such as massively parallel sequencing (MPS) now allows for analysis of many SNP regions across multiple samples simultaneously. Also, probe capture enrichment with MPS has already shown success in targeting and sequencing degraded mitochondrial DNA. The goal of this project was to design and test a probe capture assay targeting forensically relevant nuclear DNA SNPs for MPS of degraded and limited DNA samples as well as mixed DNA samples.

In order to design and develop the custom SNP probe capture assay, 451 SNPs were selected for individualization, analyzing mixtures, and determining phenotype, ancestry, and lineage. These SNPs include 136 Identity Informative SNPs, 41 Ancestry Informative SNPs, 24 Phenotypically Informative SNPs, 25 X chromosome SNPs, 81 Y chromosome SNPs, 31 Tri-allelic SNPs, 39 Tetra-allelic SNPs, and 36 Micro-haplotypes. The specificity of the probes to the regions of interest expressed as a proportion of on-target sequence reads was 95.37% for 16 samples at 25ng. The percent of SNPs with a minimum coverage of 10X for these 16 samples was 99.33%. Also, the sensitivity of this hybrid capture system was tested by varying the sample DNA amount from 50ng to 50pg. For amounts greater than 5ng, the percent of SNPs covered (measure of sensitivity) was 100% for all SNPs and the correct SNP genotype assignment was at least 99.78%. The percent of SNPs covered and correct SNP genotype assignment decreased as the sample amounts reduced from 1ng to 0.05ng. Next, the performance of the system was tested with size selected and mock degraded DNA. 0.5ng of size selected sample DNA ≤ 75bp obtained 96.65% coverage of all SNPs with 95.09% correct SNP genotype assignment. Mock degraded samples at 10ng, 1ng, and 0.5ng with an average size of 150bp were sequenced. The coverage results for the 10ng, 1ng, 0.5ng samples were 100% for 10ng, 94.87% for 1ng, and 75.67% for 0.5ng. Finally, a two-person male-male mixture was tested in contributor ratios ranging from 60:40 to 97.5:2.5. X SNPs, Y SNPs, Tri-allelic SNPs, Tetra-allelic SNPs, and Microhaplotypes helped in detecting 90% of the minor contributor SNPs for ratios between 60:40 and 90:10.

In summary, the custom SNP probe capture assay was highly successful in analyzing degraded DNA and DNA mixtures. Sample DNA ≤ 75bp at 0.5ng obtained 96.65% coverage of all SNPs with 95.09% correct SNP genotype assignment. The minor contributor SNPs can clearly be detected in ratios as low as 97.5:2.5 and 90% of minor contributor SNPs can be detected for 90:10 mixtures. Based on these results, we expect that this system can be successfully applied to analyze highly degraded DNA obtained from mass disasters and missing person cases as well as mixed DNA samples.