PCR and Probe Based Enrichment Assays for Deep Sequencing mtDNA

Cassandra D. Calloway, Valarie McClain, George Sensabaugh, Henry Erlich
Children's Hospital Oakland Research Institute, Oakland, California, USA

Mitochondrial DNA (mtDNA) is routinely targeted for analysis of highly limited or degraded forensic samples. The most common approach taken by most forensic DNA laboratories uses PCR to amplify the most hypervariable regions (HVI/HVII) of the mitochondrial genome followed by Sanger sequencing using a capillary electrophoresis instrument. While this approach is successful in most cases, there are some limitations. Sanger sequencing is limited for detection and resolution of mixed or heteroplasmic samples and sequencing only the HVI/HVII has limited discrimination power. We have developed two assays using different enrichment strategies for targeting mtDNA followed by deep sequencing using the 454 GS Jr next-generation sequencing platform to address these limitations. We use a multiplex PCR to target the HVI/HVII regions of the mitochondrial genome and a solution based probe capture method for targeting the entire mitochondrial genome. Both assays allow for greater sensitivity for detecting and resolving mixtures and the sequence capture method allows for sequencing the entire mitochondrial genome for greatly increased discrimination power.

Next-Generation sequencing (NGS) technologies offer a high-throughput solution for parallel sequencing of thousands to millions of sequences and have the potential to revolutionize forensic DNA analysis. There are several NGS technologies available which are currently being explored for forensic applications. We have chosen to use the 454 GS Jr because of the longer read length, high accuracy and low instrument cost. The 454 Genome Sequencer is a scalable, highly parallel pyrosequencing system that uses emulsion PCR (emPCR) to amplify a single DNA sequence. The 'clonal sequencing' aspect of this technology enables separation of individual components of a mixture often encountered in forensic samples. The system can be used for direct sequencing of DNA products generated by PCR or de novo sequencing of small whole genomes by employing different strategies for generating the DNA library. We have developed two mtDNA 454 sequencing assays using different enrichment strategies which will be presented here.

We have developed a multiplex PCR assay for parallel sequencing of the HVI/HVII regions of the mitochondrial genome of up to 64 pooled samples using the 454 GS Jr platform. A duplex PCR assay is used to co-amplify the mtDNA HVI/HVII regions to conserve sample, which is essential for forensic applications whereby DNA is often limited. Multiplex identification (MID) tags are incorporated in the PCR primers and used to amplify individual samples which can then be pooled and sequenced in a single run. By using an amplification strategy for target enrichment, we were able to greatly reduce the amount of DNA required for sequencing. We show successful amplification and sequencing of ~1pg DNA or ~500 mtDNA copies. We also show greatly improved sensitivity for detection and quantification of minor sequences in a mixture compared to Sanger sequencing. Using the HVI/HVII 454 assay, we were able to detect the minor sequence component at ~1% in a two person DNA mixture with 500-1000 read depth, which is much more sensitive than Sanger sequencing (10-20%). This assay was also used for sensitive detection of heteroplasmy in hair and blood samples previously not detected by other analysis methods. We present here that the HVI/HVII 454 PCR and sequencing assay is a sensitive, robust assay which allows for increased detection and resolution of individual sequences in mixed and heteroplasmic samples.

We have also developed an assay which uses a solution based probe capture system for enrichment of mtDNA and the 454 GS technology for deep sequencing of the entire mitochondrial genome. We considered several probe capture methods and selected the

Nimblegen SeqCap EZ platform because of the extensive tiling design and ability to efficiently synthesize hundreds of thousands of probes. Several properties of mtDNA were taken into consideration in the probe design and included the circular nature of mtDNA, the high density and distribution of polymorphisms, and nuclear pseudogenes to maximize probe capture efficiency and specificity. We present here the sensitivity, specificity, and mixture detection limits of our mtDNA probe capture system. We successfully captured and sequenced 100% of the mitochondrial genome with ~500-1000 fold coverage on average and ~80% on target rate. We were also able to greatly reduce the recommended starting DNA amount from 500 ng to 1ng by modifying the provided protocol and are in the process of testing even lower DNA amounts. We also compared three day and one day hybridization times to determine if the hybridization time could be reduced to increase throughput. Reduction of the hybridization time from three days to one day did not affect the capture efficiency or specificity with no observed differences in the percent mtDNA covered (100%) or on target rates (~80%). Our results show we have developed a sensitive, robust capture system which can be used to target and sequence the entire mitochondrial genome for forensic applications requiring increased discrimination power. This assay also has the potential for capture and sequencing of highly degraded samples.