OPTIMIZATION OF mtDNA CONTROL REGION SEQUENCING USING AN ABI 3100® GENETIC ANALYZER 50cm ARRAY AND FULL AND HALF REACTION BIGDYE® VERSION 1.0

Ryan Vachon BS; Suzanne M Barritt, MS, Col. Brion C Smith, Timothy P McMahon, PhD
Armed Forces DNA Identification Laboratory, Armed Forces Institute of Pathology Rockville, MD

The primary mission of the mtDNA section of the Armed Forces DNA Identification Laboratory (AFDIL) is to aid the Central Identification Laboratory Hawaii (CILHI), in the identification of human skeletal remains recovered from World War II, the Korean War, and the Vietnam Conflict. For blood specimens, AFDIL currently sequences mtDNA control region (nt15971-nt599) amplicons with ¼ reaction (in relation to the manufacturer's recommended quantity) Big Dye v. 1.0 (ABI, Foster City, CA) and seven primers in individual reactions. Sequencing products are analyzed on the ABI 3100® Genetic Analyzer using POP6 and a 36cm capillary array. However, since the ABI 3100® Genetic Analyzers and quarter BigDye sequencing reactions were validated, several problems have arisen in generating full control region sequences. These include the necessity of eight or more primers to cover the entire control region (CR), the use of alternative primers to cover difficult HV1 and HV2 C-stretches, and the loss of peak definition due to shouldering caused by a high degree of secondary structure within parts of the control region. The experiments to validate the use of full strength or half strength BigDye reactions and 50cm capillary arrays for analyzing mtDNA family references were conducted in an effort to standardize sequencing reactions and to reduce the amount of resequencing needed to confirm control regions, thus saving time and money.

First, a sensitivity study in which 300pg control region (nt15971-599) amplicons were sequenced with F15, F29, F34, R16410 and F15971 at either full or half reaction BigDye v. 1.0 and analyzed using a 50 cm array was performed. The sequence generated from these primers produced 500 or more well-defined bases, with half sequencing reactions having to be cut 20-50 base pairs sooner than full reactions. However, this did not affect the ability to confirm sequencing regions since an overlapping primer scheme was used. Further studies involved using only ½ BigDye reactions and demonstrated that the sequences obtained from a 50cm array could accurately generate full control region data using four to five sequencing primers.

Next, eight case samples previously typed and reported using ¼ BigDye reactions on the 36cm array were amplified, sequenced and typed on the 3100 using a 50cm array to demonstrate accurate reproducibility of DNA sequences. The eight case samples were randomly selected. The case samples were sequenced using both full and half BigDye reactions and the following primers (F15971, R16410, F16190, F16450, R285, F15, F314 and R599). Sequence generated on the 50 cm array was of similar if not better quality than the same sample sequences previously reported using the 36 cm array. As seen above, full BigDye reactions generated slightly longer sequence reads than the half BigDye reactions, though both confirmed the control region with fewer than the eight primers needed when using ¼ BigDye reactions and a 36cm array. A possible future experiment comparing data produced using full BigDye reactions or half BigDye reactions on the 50cm array and full BigDye reactions on the 36cm array could be useful to further explore the advantages and disadvantages of the 50cm capillary array.

Overall, the results described in these experiments demonstrated that analyzing sequencing reactions on the ABI 3100® Genetic Analyzer and a 50cm capillary array generate longer confirmed sequencing ranges of higher quality than the 36cm equivalents. Likewise, there was little difference observed in the quality of data produced with either full BigDye reactions or half BigDye reactions, although both are better than ¼ BigDye reactions. The two-fold increase in sequencing time for the 50cm array is an acceptable change in time as the 3100 can be run overnight and weekends with little or no human interaction. This not only frees the analysts up for data analysis, but also saves time and money by reducing the necessity for resequencing.