

DEVELOPMENT OF AN EFFICIENT, HIGH-THROUGHPUT STRATEGY FOR SEQUENCE ANALYSIS OF THE ENTIRE HUMAN MITOCHONDRIAL DNA CONTROL REGION

Susan W. Jones, Ph.D.¹, Jodi A. Irwin-Ross, M.S.¹, Faith A. Love, B.S.², Amy B. Welsh, M.F.S.¹, Mitchell M. Holland, Ph.D.¹, and Thomas J. Parsons, Ph.D.¹

¹*Armed Forces DNA Identification Laboratory*

²*George Washington University, Dept. of Forensic Sciences*



Currently, human mtDNA forensic analysis most often involves complete sequence determination from hypervariable regions 1 and 2 (HV1 and HV2) of the control region (CR). In a majority of cases, the variation present in these regions is sufficient to discriminate between individuals of different mtDNA lineages. Nonetheless, there are substantial limitations on the power of mtDNA discrimination caused by the occurrence of certain mtDNA types that are common in the population (e.g., the most common HV1/HV2 type within Caucasians occurs in ~5% of the population). Recent work in our laboratory (Anderson *et al.*, 1999), and elsewhere, has demonstrated that significant sequence variation in the CR resides outside of HV1 and HV2. Targeting this variation can substantially improve the discriminatory potential of mtDNA. For this reason we plan to routinely sequence the entire CR for reference samples and population databasing. We will present our development of an effective strategy for high-throughput sequencing of the entire CR.

We designed a new PCR primer set that robustly amplifies a 1.3 kb mtDNA fragment containing the complete CR, and demonstrated amplification sensitivity down to 200 pg of total DNA. Our next step was to evaluate a wide variety of internal sequencing primers, and to develop and test a strategy that would robustly provide reportable sequence data over the entire region. AFDIL reporting criteria require confirmed sequence from both DNA strands or confirmation from two different amplifications off of the same strand. These criteria directed our strategy for identifying a minimum number of sequencing reactions from two independent amplifications that would yield complete reporting. Unique properties of the control region, such as length-heteroplasmic polycytosine stretches in HV1 and HV2, also are important in determining an optimal approach. Recognizing that in a high-throughput laboratory sample tracking and analysis are primary challenges, our goal has been to minimize the frequency with which samples need to be re-sequenced. In addition to *de novo* CR sequencing, we have developed also an efficient strategy for obtaining the remainder of the CR sequence from individuals whose HV1 and HV2 sequence has previously been determined.

The need for rapid and inexpensive sequencing of the entire control region will continue to grow for the foreseeable future. Population databases need to be established for the entire control region, and it will be highly desirable to "retro-database" the current HV1/HV2 forensic databases to include the entire CR. Therefore, our sequencing strategies have been developed with an eye toward the eventual implementation of robotics. We have evaluated and validated PCR product clean-up and post-cycle sequencing clean up steps that are performed in a 96-well microtitre tray format.

