Sequence Analysis of Mitochondrial DNA Hypervariable Regions Using Infrared Fluorescence Detection

David L. Steffens¹ and Reena Roy^{1,2}

¹LI-COR, Inc. and ²Nebraska State Patrol Criminalistics Laboratory, Lincoln, NE, USA

BioTechniques 24:1044-1046 (June 1998)

ABSTRACT

The non-coding region of the mitochondrial genome provides an attractive target for human forensic identification studies. Two hypervariable (HV) regions, each approximately 250–350 bp in length, contain the majority of mitochondrial DNA (mtDNA) sequence variability among different individuals. Various approaches to determine mtDNA sequence were evaluated utilizing highly sensitive infrared (IR) fluorescence detection. HV regions were amplified either together or separately and cycle-sequenced using a Thermo Sequenase[®] protocol. An M13 universal primer sequence tail covalently attached to the 5' terminus of an amplification primer facilitated electrophoretic analysis and direct sequencing of the amplification products using IR detection.

INTRODUCTION

The mitochondrion is one of only two cellular organelles that possesses its own genome. The human mitochondrial genome contains 16 569 bp of DNA and has been completely sequenced (1). It primarily consists of densely organized genes encoding rRNAs, tRNAs and proteins involved in oxidative respiration, together with a non-coding region approximately 1100 bp in length containing the origin of replication. Sequence analysis of mitochondrial DNA (mtDNA) from different individuals reveals a higher degree of variability relative to nuclear genomic DNA, presumably because of a lower fidelity of the mtDNA polymerase (2). The majority of this sequence variation is found within the non-coding region sub-localized to two hypervariable (HV) regions (HV1 and HV2), each approximately 250–350 bases in length (6,15). This sequence variability provides an attractive target for forensic identification studies. In addition, human cells contain several

hundred copies of mtDNA, thus providing substantially more template for amplification than nuclear genomic DNA (8).

We have previously used polymerase chain reaction (PCR) amplification coupled with infrared (IR) fluorescence detection (7) for analysis of short tandem repeat (STR) polymorphisms (11,13), determination of gender (10) and typing of D1S80 alleles (9). Several potential strategies exist for amplification and sequence analysis of human HV1 and HV2. Both are sufficiently small (250–350 bp) to allow efficient amplification by PCR and can be co-amplified together in a single reaction. In addition, the regions are adjacently positioned, allowing amplification of both within a 1-kb PCR fragment.

Sequence analysis was performed using a Model 4200 Dual Dye Automated Sequencer (LI-COR, Lincoln, NE, USA). The dual-IR dye capability permits simultaneous generation of two sequence ladders in one reaction or loading of two separate reactions in one set of lanes. A gel length of 66 cm can provide sequence runs of 1000 bases or more, thus potentially permitting simultaneous sequence analysis of both HV regions in one PCR product. These various approaches towards forensic sequence analysis of human HV regions are evaluated in this report.

MATERIALS AND METHODS

Within the human mitochondrial genome, HV1 is located in the region of approximately between base numbers 16 024–16 365, while HV2 is approximately between bases 73 and 340 (17,18). We have primarily utilized amplification primer sequences recommended by Holland et al. (4). HV1 was amplified using primers F15990, R16239 and R16391 (F or R designates Forward or Reverse primer; the number specifies the position of the primer's 3' terminus on the mitochon-

drial genome). An F16004 primer (5'-AGCACCCAAAGC-TAAGATTC-3') was also evaluated. HV2 was amplified using primers F34, R266, R362, and R465. Both regions can be co-amplified within a 1-kb product using F15990 or F16004 together with R362 or R465.

IR fluorescent visualization of amplification products is facilitated by covalently attaching a universal M13 primer sequence "tail" on the 5' end of one or both amplification primers, as described previously (11). M13 tails utilized in this work were M13 Forward (-29) (5'-CACGACGTTGTAAA-ACGAC-3') (14) and M13 Reverse (5'-GGATAACAATTTC-ACACAGG-3'). A small quantity of IR-labeled universal M13 primer included in the amplification reaction will be incorporated into product yielding IR-labeled PCR fragments, which can be conveniently analyzed on the LI-COR Model 4200 sequencer. Either amplification primer can be synthesized with either of the two M13 tails at its 5' terminus, or one primer can have one tail and the other primer have the other tail.

DNA for amplification was isolated from saliva samples using Chelex® extraction (16), yielding a final concentration of 5–10 ng/μL. Alternatively, DNA was isolated from liquid human blood samples by standard organic extraction (3). Briefly, samples were digested with proteinase K in a sodium dodecyl sulfate (SDS) solution followed by phenol-chloroform-isoamyl alcohol extraction. DNA was ethanol-precipitated and redissolved in 1× TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) at a concentration of approximately 20 ng/μL.

Amplifications utilized either a 50° or 60°C annealing temperature, depending upon the melting temperatures (T_m)s of the particular amplification primers, and were cycled with a GeneAmp® PCR System 2400 (Perkin-Elmer, Norwalk, CT, USA). Reactions were performed in a 50-μL volume contain-

ing PCR buffer, 0.2 mM dNTPs, 10 pmol of each amplification primer (one or both containing an M13 tail at the 5' terminus), 1 pmol IRD800-labeled universal M13 primer (LI-COR), 10–40 ng template DNA and 1.0 U of *Taq* DNA Polymerase (Boehringer Mannheim, Indianapolis, IN, USA). Amplification products were electrophoretically analyzed for purity using a 25-cm denaturing gel (6% Long Ranger™; FMC Bio-Products, Rockland, ME, USA) on the LI-COR Model 4200 Dual Dye Automated DNA Sequencer before sequencing.

Figure 1 demonstrates that the tailed primer labeling strategy is a convenient means of labeling amplified mtDNA HV regions. Lane 1 depicts a 460-base fragment encompassing the entire HV1 region, while lane 2 contains a 291-base fragment deriving from a portion of HV2. Although spurious bands of lesser intensity were occasionally visualized during electrophoretic analysis of amplified HV regions, these did not seem to adversely affect subsequent sequencing of the PCR products. Before sequence analysis, amplification products were purified using an Amicon® Microcon®-100 Ultrafiltration Microconcentrator (Millipore, Bedford, MA, USA) to remove residual primers, nucleotides and salts. Reaction (10–15 μL) was mixed with 500 μL 0.1× TE buffer in the filter cup and spun at 500× *g* in a microcentrifuge for 10–20 min until 5–10 μL of retentate remained.

Amplified HV regions were sequenced using a Thermo Sequenase™ Cycle Sequencing Kit (Amersham Pharmacia Biotech, Piscataway, NJ, USA). A pre-mixture composed of 1.0 μL universal M13 primer [1 pmol/μL M13 Forward (-29) or M13 Reverse (LI-COR) labeled with IRD700 or IRD800], 1.0 μL purified PCR product and 24.0 μL H₂O were prepared, and 6.0 μL were aliquoted into each of four reaction tubes, which contained 2.0 μL of the appropriate Thermo Sequenase reagent. The volume of some purified PCR products needed to be empirically adjusted (0.5–2.5 μL) to obtain optimal results. Sequencing reactions were thermal cycled using a pre-incubation of 2.5 min at 95°C, followed by 40 cycles at 95°C for 30 s, 60°C for 30 s and 72°C for 90 s. When complete, 6.0 μL of IR2 Stop/Loading Buffer (LI-COR; which contains an alternate dye in place of bromphenol blue to reduce sequence ladder distortion near the primer front) were added to each tube. Samples were denatured at 95°C, and 1–2 μL of each loaded onto a 41-cm length gel (6% Long Ranger) mounted on a Model 4200 Dual Dye Automated DNA Sequencer and electrophoresed according to the manufacturer's recommendations. For large PCR products (1 kb), a 66-cm length gel (4% Long Ranger) was utilized. Sequence ladders were visually inspected to assess data quality and analyzed using BaseImagIR™ software (LI-COR).

Initial attempts at HV sequence analysis focused upon shorter PCR products (200–300 bp). Optimization of reaction conditions produced readable sequence ladders corresponding to the previously published sequence. Amplification primer sequences were present at the beginning and end of the sequence ladder. Subsequent analysis attempts utilized longer products containing an entire HV region or both HV1 and HV2. Sequencing of the latter 1-kb products yielded multiple strong stops in all four lanes, primarily within the first 200 bases, producing ambiguities at these positions. Interestingly, long PCR products generated using alternate primers produced strong stops at different positions, which remained constant when other sequencing enzymes were used (e.g., SequiTherm™ and Excel™ II; Epicentre Technologies, Madison, WI, USA). It is hypothesized that these stops result from

Figure 1. Amplification of mitochondrial HV regions using IR fluorescent detection. Human mtDNA amplification products were electrophoretically separated using a 25-cm length gel on a Model 4200 Dual Dye Automated DNA Sequencer. Lane 1 is a 460-bp product obtained using primers F15990 with an M13 Forward (-29) tail and R16391, which amplify the entire HV1 region. Lane 2 is a 291-bp product obtained using primers F34 with an M13 Forward (-29) tail and R266 to amplify a portion of HV2. M designates a molecular weight marker with sizes listed to the left of the figure.

Figure 2. Sequencing of mtDNA HV1 and HV2 regions. HV1 was amplified using primers F15990 with an M13 Forward (-29) tail and R16391, and HV2 was amplified using F34 with an M13 Reverse tail and R362. The figure shows two images collected simultaneously in two different detection channels during one electrophoresis run. Reaction set 1 is HV1-sequenced using M13 Forward (-29) labeled with IRD700, and reaction set 2 is HV2-sequenced using M13 Reverse labeled with IRD800. Reaction set 3 consists of a mixture of the first two, which were mixed together before gel loading.

secondary structure within the PCR product and that alternate fragments contain alternate secondary structures producing alternate sequence ladder stops. Shorter PCR products containing a single HV region produced substantially fewer sequencing stops.

Separate amplifications of HV1 using F15990 with an M13 Forward (-29) tail and R16391, and HV2 using F34 with an M13 Reverse tail and R362 were also performed. When sequenced separately using the appropriate IR-labeled primer, high-quality sequence ladders with few artifacts are produced, as shown in Figure 2 (reaction sets 1 and 2). These sets can also be mixed together before loading, and electrophoresed together in the same lanes (reaction set 3). These four primers were also utilized to simultaneously amplify both HV-region products in a single PCR. However, when this template mixture was sequenced using M13 Forward (-29) or M13 Reverse, or a combination of both, a number of sequencing artifacts and strong stops were produced. During these studies, it became apparent that organically extracted template yielded higher quality sequence ladders than did Chelex-extracted DNA. Presumably, high-molecular-weight impurities present in the Chelex extract are carried over and interfere with sequencing reactions.

CONCLUSIONS

Sequence data produced during these experiments corresponded to published human mtDNA sequence (1). Several different individuals were used during this study and definite sequence variations were easily distinguishable. Sequencing of HV regions separately yielded the most accurate sequence data by reducing the required read length. In general, performing each step separately rather than combining them produces the most accurate DNA sequence data. Higher throughput can be obtained by utilizing different IR fluorescent dyes and electrophoresing them together in the same lanes on the Model 4200 sequencer. The availability of two IR dyes and

multiple-tailed primers provide a variety of possibilities for amplification and sequencing of targets of interest. This versatility has not been previously examined, although an earlier report explored different mtDNA sequencing strategies involving alternate sequencing chemistries (5). IR fluorescence is more sensitive than visible fluorescence (12), which may prove useful in analysis of forensic samples that are degraded or limited in quantity.

REFERENCES

1. Anderson, S., A.T. Bankier, G.B. Barrell, M.H.L. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465.
2. Cann, R.L., M. Stoneking and A.C. Wilson. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31-36.
3. Comey, C.T., B.W. Koons, K.W. Presley, J.B. Smerick, C.A. Sobieralski, D.M. Stanley and F.S. Baechtel. 1994. DNA extraction strategies for amplified fragment length polymorphism analysis. *J. Forensic Sci.* 39:1254-1259.
4. Holland, M.M., D.L. Fisher, R.K. Roby, J. Ruderman, C. Bryson and V.W. Weedn. 1995. Mitochondrial DNA sequence analysis of human remains. *Crime Lab. Digest* 22:109-115.
5. Hoggood, R., K.M. Sullivan and P. Gill. 1992. Strategies for automated sequencing of human mitochondrial DNA directly from PCR products. *BioTechniques* 13:82-92.
6. Horai, K., Y. Nakamura and R. White. 1990. Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am. J. Hum. Genet.* 46:828-842.
7. Middendorf, L.R., J.C. Bruce, R.C. Bruce, R.D. Eckles, D.L. Grone, S.C. Roemer, G.S. Sloniker, D.L. Steffens, S.L. Sutter, J.A. Brumbaugh and G. Patonay. 1992. Continuous, on-line DNA sequencing using a versatile infrared laser scanner/electrophoresis apparatus. *Electrophoresis* 13:487-494.
8. Robin, E.D. and R. Wong. 1988. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *J. Cell. Physiol.* 136:507-513.
9. Roy, R. Infrared fluorescent detection of D1S80 alleles. 1997. *Forensic Sci. Int.* 87:63-71.
10. Roy, R. and D.L. Steffens. 1997. Infrared detection of PCR amplified gender identifying alleles. *J. Forensic Sci.* 42:452-460.
11. Roy, R., D.L. Steffens, B. Gartside, G.Y. Jang and J.A. Brumbaugh. 1996. Producing STR locus patterns from bloodstains and other forensic samples using an infrared fluorescent DNA sequencer. *J. Forensic Sci.* 41:418-424.
12. Steffens, D.L., G.Y. Jang, S.L. Sutter, J.A. Brumbaugh, L.R. Middendorf, K. Mühlegger, E.R. Mardis, L.A. Weinstock and R.K. Wilson. 1995. An infrared fluorescent dATP for labeling DNA. *Genome Res.* 5:393-399.
13. Steffens, D.L., R. Roy and J.A. Brumbaugh. 1997. Multiplex amplification of STR loci with gender alleles using infrared fluorescence detection. *Forensic Sci. Int.* 85:225-232.
14. Steffens, D.L., S.L. Sutter and S.C. Roemer. 1993. An alternate universal forward primer for improved automated DNA sequencing of M13. *BioTechniques* 15:580-582.
15. Stoneking, M., D. Hedgecock, R.G. Higuchi, L. Vigilant and H.A. Erlich. 1991. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *Am. J. Hum. Genet.* 48:370-382.
16. Walsh, P.S., D.A. Metzger and R. Higuchi. 1991. Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* 10:506-513.
17. Wilson, M.R., M.M. Holland, M. Stoneking, J.A. DiZinno and B. Budowle. 1993. Guidelines for the use of mitochondrial DNA sequencing in forensic science. *Crime Lab. Digest* 20:68-77.
18. Wilson, M.R., D. Polanskey, J. Butler, J.A. DiZinno, J. Replogle and B. Budowle. 1995. Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. *BioTechniques* 18:662-669.

Address correspondence to Dr. David L. Steffens, LI-COR Biotechnology Division, 4308 Progressive Ave., P.O. Box 4000, Lincoln, NE 68504, USA. Internet: dsteffens@bio.licor.com