NIST SPECIAL PUBLICATION 260-155

U. S. DEPARTMENT OF COMMERCE/Technology Administration
National Institute of Standards and Technology

Standard Reference Materials®

Human Mitochondrial DNA—Amplification and Sequencing Standard Reference Materials—SRM 2392 and SRM 2392-I

B. C. Levin, D. K. Hancock, K. A. Holland, H. Cheng and K. L. Richie

NIST
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

NIST Special Publication 260-155

Standard Reference Materials®

Human Mitochondrial DNA—Amplification and Sequencing - Standard Reference Materials—SRM 2392 and SRM 2392-I

B. C. Levin, D. K. Hancock, K. A. Holland, H. Cheng and K. L. Richie

Biotechnology Division
Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8311

U.S. DEPARTMENT OF COMMERCE, *Donald L. Evans, Secretary*
TECHNOLOGY ADMINISTRATION, *Phillip J. Bond, Under Secretary of Commerce for Technology*
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, *Arden L. Bement, Jr., Director*

Issued September 2003

Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

National Institute of Standards and Technology Special Publication 260-155
Natl. Inst. Stand. Technol. Spec. Publ. 260-155, 89 pages (September 2003)
CODEN: NSPUE2

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 2003

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov — Phone: (202) 512-1800 — Fax: (202) 512-2250
Mail: Stop SSOP, Washington, DC 20402-0001

Contents

NIST Special Publication 260-155

Standard Reference Materials:
Human Mitochondrial DNA – Amplification and Sequencing – SRM 2392 and SRM 2392-I

Barbara C. Levin, Diane K. Hancock, Koren A. Holland, Haiyan Cheng and Kristy L. Richie

Biotechnology Division
Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8311

Abstract

Standard Reference Materials SRM 2392 and 2392-I are intended to provide quality control when performing the polymerase chain reaction (PCR) and sequencing of human mitochondrial DNA (mtDNA) for forensic identification, medical diagnosis, or mutation detection. They may also serve as controls when amplifying (PCR) and sequencing any DNA. These SRMs can also be used for quality assurance when determining the sequence of in-house DNA controls. SRM 2392 is certified for the sequences of the entire human mtDNA (16,569 base pairs) from two lymphoblastoid cell culture lines (CHR and 9947A) from apparently normal individuals and the cloned HV1 region of CHR containing a C-stretch through which it is difficult to sequence. SRM 2392-I is certified for the mtDNA sequence from HL-60, a promyelocytic cell line prepared from the peripheral blood leukocytes from an individual with acute promyelocytic leukemia. The mtDNA sequence information (but not the DNA) of two additional DNA templates (GM03798 and GM10742A) that were amplified and sequenced in their entirety multiple times at NIST are provided in this Special Publication for comparison purposes. The sequences of fifty-eight unique primer sets that allow any area or the entire mtDNA (16,569 base pairs) to be amplified and sequenced are also provided. We found during the sequencing of these five mtDNA templates that some of the single nucleotide polymorphisms (SNPs) did result in amino acid changes when compared with the Cambridge Reference Sequence (Anderson *et al.,*1981; Andrews *et al.,*1999). Two interlaboratory evaluations, one for the amplification, sequencing, and data analysis of the CHR template and one for the HL-60 template were each conducted by three different laboratories and NIST. Corroboration of the results in SRMs 2392 and 2392-I will provide quality assurance that any unknown mtDNA is also being amplified and sequenced correctly.

Keywords: CHR; 9947A; HL-60; GM03798; GM10742A; Forensic Identification; Medical Diagnosis; Mitochondrial DNA sequence; Mutation Detection; Single Nucleotide Polymorphism (SNP); Standard Reference Material (SRM);

1. Introduction

Each human cell can have a few dozen to several thousand molecules of mtDNA (Bogenhagen and Clayton, 1974; King and Attardi, 1989). Sequence analysis of mtDNA is used by the forensic community for human identification especially in those cases where the genomic DNA is highly degraded or non-existent (Holland *et al.,* 1993; Holland *et al.,* 1995). Forensic analysis to distinguish between non-maternally related individuals is primarily based on the considerable sequence variation found in the two hypervariable regions (HV1, HV2) located in the non-coding displacement loop (D-loop). The medical community is also using sequence analysis of mtDNA for diagnoses of diseases associated with specific mutations and deletions in the coding regions (Wallace *et al.,* 1997). A third area of research which is largely unexplored and which needs sequence analysis is the examination of the mutagenic effects of chemical and physical agents on mtDNA (Grossman, 1995; Ballinger *et al.*, 1996).

Human mitochondrial DNA (mtDNA) has been completely sequenced and found to be circular double-stranded molecules containing 16,569 base pairs (Fig. 1) (Anderson *et al.,* 1981). The sequence determined by Anderson *et al.* (1981) is referred to as the Cambridge Reference Sequence and was a consensus sequence based primarily on the results from a human placenta, partly on the sequence from HeLa cells, and in five areas where the results were ambiguous, the authors inserted the results from the bovine mtDNA that they had also sequenced. In 1999, Andrews *et al.* resequenced the original placenta and found that the 1981 sequence contained a number of errors and rare polymorphisms that were specific for that placenta. In SRM 2392, which became available in 1999, we compared the mtDNA sequence of three apparently normal individuals with the 1981 Cambridge Reference Sequence. We also found a number of places that appeared to be errors and which were later confirmed by Andrews *et al.* (1999). In this NIST Special Publication on the development of the mtDNA amplification and sequencing SRM 2392 (Levin *et al*., 1999) and SRM 2392-I (Levin *et al*., 2003, in press), we compare the sequencing results from three apparently normal individuals and two individuals with diseases with both the 1981 and 1999 Cambridge Reference Sequences. Human mtDNA SRMs 2392 and 2392-I can be used for quality control in amplification, sequencing, forensic identifications, medical diagnostics, and mutation detection.

2. Materials and Methods

2.1. Origin of Extracted DNA

In SRM 2392, the DNA template designated CHR came from human white blood cells that were transformed with the Epstein-Barr virus and immortalized as a cell culture line (CHR cells) by the American Type Culture Collection (ATCC, Manassas, VA). After transformation, the cells were grown in Iscove's Modified Dulbecco's Media or RPMI 1640 media with L-glutamine, sodium bicarbonate, penicillin, streptomycin and 20% fetal calf serum (Life Technologies, Inc., Grand Island, NY). The cell cultures were grown at 37 ^{o}C in an humidified atmosphere containing 5% CO_2 and 95% air. The DNA was extracted from 2 x 10^8 CHR cells by the Qiagen Plasmid/Cosmid Purification Protocol (Qiagen, Inc., Chatsworth, CA). This procedure enhanced the concentration of mtDNA and reduced, but did not eliminate, nuclear DNA.

The CHR data presented in this paper were obtained primarily with the above mentioned immortalized CHR cell culture line. However, before starting production of the final SRM 2392, it was necessary to obtain fresh blood from CHR and to reestablish the cell line. This second CHR cell line was established by ATCC as above. The sequence of the mtDNA of this second CHR cell line was examined and found to be identical to that of the first CHR cell line, with the single exception that no heterplasmy was noted at nucleotide position (np) 6849[1]; the mtDNA of the second CHR cell line agreed with the Cambridge Reference Sequence at np 6849. It is the DNA from the second CHR cell line that is included in SRM 2392.

Also in SRM 2392, the DNA template 9947A was obtained from Life Technologies, Inc. (Grand Island, NY) who prepared it from a Epstein-Barr virus immortalized human lymphoid cell line. DNA from 9947A is also used in the PCR-based DNA profiling standard (SRM 2391[2]) designed for forensic and paternity testing, law enforcement training, and research.

SRM 2392-I contains the extracted DNA from HL-60; this DNA was prepared by the Professional Services Department of ATCC (Manassas, VA).

The entire mtDNA of two additional cell lines (GM03798 and GM10742A) were amplified and sequenced multiple times at NIST and these sequences are provided in this NIST Special Publication for information and comparison purposes. Although the extracted DNA from GM03798 and GM10742A are not part of SRM 2392 or 2392-I, the cell cultures can be obtained from the National Institute of General Medical Sciences (NIGMS) Human Genetic Mutant Cell Repository, Coriell Institute for Medical Research, Camden, NJ. GM03798, an apparently normal human lymphoblastoid cell culture, was grown in the same manner as the CHR cells and the DNA was extracted using DNA NOW™, a phenol-free DNA isolation reagent (BIOGENTEX, Seabrook, TX). The lymphoblast cell line GM10742A came from a patient with Leber Hereditary Optic Neuropathy (LHON), a disease that causes blindness in young adults and that has been associated with a number of mtDNA primary, intermediate and secondary mutations (Wallace *et al.*, 1997). This cell line (GM10742A) was purchased from NIGMS. The cells were grown at NIST in RPMI 1640 plus L-glutamine and sodium bicarbonate growth media (Sigma, St. Louis, MO), fetal calf serum (20%) (Sigma), and the antibiotics streptomycin and penicillin (final concentration: 100 U/mL) (Sigma). DNA was extracted from GM10742A using the QIAGEN Plasmid Kit following the Plasmid Mini Purification Protocol.

2.2. Isolation and Cloning of mtDNA Containing the C-Stretch in SRM 2392

Confluent CHR cells were harvested by centrifugation at 1500 rpm for 5 min. The mtDNA was isolated using the Qiagen Plasmid/Cosmid Purification Protocol (Qiagen Inc., Chatsworth, CA). Following isolation, the mtDNA was digested with restriction enzymes SacI and KpnI (New England Biolabs, Inc., Beverly, MA) into five fragments which were separated on a 0.7% low

[1]All nucleotide numbers referred to in this paper are based on the numbering system of the Cambridge Reference Sequence (Anderson *et al.,* 1981 and Andrews *et al.*, 1999).
[2]SRM 2391 may be obtained from the Standard Reference Material Program, NIST, Gaithersburg, MD 20899.

melting agarose gel. Bands of the size of the fragment containing the HV1 region were cut from the gel and melted at 65 °C. DNA was extracted with phenol twice and precipitated by adding sodium acetate (150 mmol/L) and two volumes of 100% ethanol. The final product was resuspended in 10 mmol/LTris-1mmol/L EDTA (TE) buffer. The cloning vector, M13mp18, was also digested with SacI and KpnI, treated with Calf Intestinal Alkaline Phosphatase (New England Biolabs, Inc., Beverly, MA) extracted with phenol and precipitated with sodium acetate and ethanol as described above. The vector was incubated with the mtDNA product and T_4 DNA ligase (Life Technologies, Inc., Grand Island, NY) at 4 °C overnight. An overnight culture of *E. coli* host TG-1 cells was diluted and grown at 37 °C in LB media (Sambrook *et al.,* 1989) until the OD_{650} reached 0.4 - 0.5. The cells were harvested by centrifugation at 1500 rpm for 5 min. The cell pellet was resuspended in 10 mL calcium chloride (50 mmol/L) and incubated for one hour on ice, centrifuged, and resuspended in 1 mL calcium chloride (50 mmol/L) and incubated for 30 min on ice. The treated TG-1 cell suspension (0.3 mL) was incubated for 30 min on ice with 20 µL of a ligation mixture containing the isolated mtDNA fragment, the cloning vector that had been treated overnight, T_4 Ligase and ligation buffer (Life Technologies, Inc., Gaithersburg, MD). This mixture was heat shocked at 42 °C for 2 min, mixed with 0.2 mL of untreated TG-1 cells (from the overnight culture), 4 µL of a 1mol/L isopropylthio-β-D-galactoside (IPTG), 40 µL of a 20 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and 3 mL of melted (55 °C) top agar and spread on the surface of freshly prepared LB agarose plates (Sambrook *et al.,* 1989). The plates were incubated at 37 °C overnight. Both colorless and blue plaques were visible in the morning. The colorless plaques indicate that insertion of the mtDNA fragment into the vector has occurred; whereas, the blue plaques have no insertion.

2.3. Bacteriophage DNA Isolation and Sequencing

Single, well isolated colorless plaques from the above LB plates were each placed in a sterile tube with 1.5 mL of a TG-1/ LB cell suspension which contained TG-1 cells that were grown overnight and diluted 1/100 in LB media and grown for 1 hour at 37 °C. The plaques and the TG-1/ LB cell suspension were grown at 37 °C for 5 hours. Cells were removed by centrifugation at 15000 rpm for 5 min. The supernatant containing the bacteriophage was incubated with 0.2 mL polyethylene glycol (20% PEG in 2.5 mol/L NaCl) overnight at 4 °C and the resultant precipitate containing the DNA phage particles was pelleted by centrifugation at 15000 rpm for 15 min. The bacteriophage DNA was isolated by phenol extraction and sodium acetate/ethanol precipitation as described above and then dissolved in 25 µL of TE buffer. The bacteriophage DNA was cycle sequenced with AmpliTaq DNA polymerase, FS and the -21M13 primer: 5'-TGTAAAACGACGGCCAGT-3' according to the protocol in the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (PerkinElmer, Foster city, CA). The cycle sequencing was conducted in a PerkinElmer Model 9600 thermocycler by first heating the DNA reaction mixture at 96 °C for 1 min and then subjecting the mixture to 25 cycles of 96 °C for 10 sec, 50 °C for 5 sec and 60 °C for 2 min. The cycle sequencing product was purified using a Centri-Sep spin column (Princeton Separations, Inc., Adelphia, NJ). The DNA pellet was rinsed with 70% ethanol, vacuum dried, resuspended in loading buffer prepared by combining deionized formamide and 25 mmol/L EDTA (pH 8.0) in a ratio of 5:1, loaded onto a 4.75% acrylamide gel and electrophoresed on an ABI 373A DNA sequencer. One of the clones containing the C-stretch sequence was used as the source of the cloned DNA for the SRM.

2.4. mtDNA Primers used in both SRM 2392 and 2392-I

Fifty-eight sets of unique primers (19-28 bp) for sequencing the entire human mtDNA (16,569 bp) were computer-designed using GENE RUNNER FOR WINDOWS (Hastings Software, Inc., Hastings, NY). The -21M13 primer was used to sequence the cloned HV1 region of the DNA from the CHR template. The sequences of all the primers are shown in Table 1 and are the same as those used for SRM 2392 (Levin *et al.*, 1999; Appendix C) with the exception of the reverse primer of set 51. During the course of the development of SRM 2392-I, this primer was changed since it contained a C at np 14368. Since this is the reverse primer, it would bind to a G at np 14368. Andrews *et al.* (1999) in their reevaluation of the placenta originally used to sequence human mtDNA in 1981 (Anderson *et al.*, 1981) found that np 14368 should have a C at that position. Therefore, the primer should have a G at np 14368. The new reverse primer 51 is 5'-TTAGCGATGGAGGTAGGATT**G**G-3' (np 14368 is in bold and underlined) which binds to the C at position 14368. The 5' end is np 14388 and the 3' end is 14367. The 58 sets of primers were custom-made by Bio-Synthesis, Inc. (Lewisville, TX); the new reverse primer 51 was obtained from Invitrogen (Carlsbad, CA).

2.5. Polymerase Chain Reaction (PCR) SRM 2392 completed in 1999

Extracted DNA was resuspended in TE buffer (pH 7.5) containing 10 mmol/L Tris and 1 mmol/L EDTA. The PCR reaction mixture contained: DNA (1μL), Taq DNA polymerase (0.5 μL or 2.5 units)((Boehringer Mannheim, Indianapolis IN) and 10x buffer (5 μL) (Boehringer Mannheim), dNTP's (0.2 mmol/L each) (Life Technologies, Inc., Gaithersburg, MD), forward and reverse primers (0.4 μmol/L each) plus H_2O to a final volume of 50 μL. The 10x buffer (pH 8.3) contained Tris-HCl (100 mmol/L), $MgCl_2$ (15 mmol/L), and KCl (500 mmol/L). Thermal cycling was conducted in a PerkinElmer Model 9600 thermocycler and consisted of 1 min at 96 ^{o}C, followed by 32 cycles of 15 sec at 94 ^{o}C (denaturation), 30 sec at 56 ^{o}C (annealing), and 15 sec at 72 ^{o}C (extension), and ending with a final extension of 7 min at 72 ^{o}C. A sample of the amplified DNA was electrophoresed in 0.7% agarose and stained with ethidium bromide to assess the purity and size of the PCR product. Before sequencing, extraneous materials were removed from the PCR product with a QIAquick PCR Purification Kit (QIAGEN, Inc., Chatsworth, CA).

2.6. Polymerase Chain Reaction (PCR) SRM 2392-I completed in 2002

The PCR mixture contained: DNA (1μL; 1.4 ng), AmpliTaq Gold® DNA polymerase, (0.5 μL; 2.5 units) (Applied Biosystems), 10x buffer (5 μL) containing 100 mmol/L Tris-HCl, pH 8.3, 500 mmol/L KCl, 15 mmol/L $MgCl_2$ and 0.01% (w/v) gelatin (Applied Biosystems), 10 mmol/L dNTP mix (1 μL) (Invitrogen), 10 μmol/L forward and reverse primers (1 or 2 μL), plus water to make a final volume of 50 μL. Thermal cycling was conducted in a PerkinElmer thermocycler Model 9700 and started with 10 minutes at 95 ºC, followed by 35 cycles of 20 sec at 94 ºC (denaturation), 10 sec at 56 ºC (annealing), 30 sec at 72 ºC (extension) and ended with a final extension of 7 min at 72 ºC. Amplified DNA was purified with a QIAquick PCR Purification Kit (Qiagen) and the purity and size of the PCR product was determined by electrophoresis in 2% agarose gels in 1x TBE buffer containing 0.5 μg/mL ethidium bromide (Sigma).

2.7. Sequencing SRM 2392 completed in 1999

Cycle sequencing using fluorescent dye-labeled terminators was performed with an ABI PRISM™ Dye Terminator Cycle Sequencing Kit with AmpliTaq® DNA Polymerase, FS (PerkinElmer). Thermal cycling was conducted in a PerkinElmer Model 9600 thermocycler and started with one minute at 96 °C. The reaction then underwent 25 cycles of 96 °C for 15 sec (denaturation), 50 °C for 5 sec (annealing), and 60 °C for 2 min (extension). The DNA products were purified by passage through a Centri-Sep spin column (Princeton Separations, Inc.). Electrophoresis and sequencing of the fluorescently-labeled purified DNA were performed with a 373A ABI Sequencer (PerkinElmer) using a 4.75% acrylamide gel. Data analysis was executed with the Sequence Navigator software package (PerkinElmer).

2.8. Sequencing SRM 2392-I completed in 2002

Cycle sequencing using fluorescent dye-labeled terminators was performed with an ABI PRISM® BigDye® Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (Applied Biosystems). Cycle sequencing reactions in both the forward and reverse modes were conducted with a 9700 PerkinElmer thermal cycler and started with one minute at 96 ºC. The reaction then underwent 25 cycles of 96 ºC for 15 sec (denaturation), 50 ºC for 5 sec (annealing), and 60 ºC for 2 min (extension). The DNA products were purified using Edge Gel Filtration Cartridges (Edge BioSystems, Gaithersburg, MD). Sequencing and data analysis of the purified DNA were performed using an Applied Biosystems PRISM® Model 310 Genetic Analyzer with POP-6™ polymer system and 47 cm x 50 µm capillaries (Applied Biosystems). Sequence data were analyzed with Sequencing Analysis Software 3.3, and comparisons to the Cambridge Reference Sequence (Anderson *et al.*,1981) were performed with Sequence Navigator Software 1.01.

2.9. Interlaboratory Evaluation for SRM 2392 completed in 1999

Three laboratories in addition to NIST participated in an interlaboratory evaluation of the CHR template. These laboratories were The Bode Technology Group, Inc., 21515 Ridgetop Circle, Suite 140, Sterling, VA 20166; IIT Research Institute, Virginia Technology Center, 8510 Cinderbed Road, Suite 300, P.O. Box 899, Newington, VA 22122; and Lark Technologies, Inc., 9545 Katy Fwy, Suite 465, Houston, TX 77024.

Each laboratory was sent:
1. Two tubes of DNA from the first CHR cell culture line. One tube contained extracted DNA ready for PCR amplification of the entire mtDNA. The other contained the cloned DNA ready for cycle sequencing of the HV1 region (this DNA did *not* need to be PCR amplified).

2. Fifty-eight sets of primers labeled with either F# (forward primer) or R# (reverse primer). Forward and reverse primers with the same number were paired and numbered from the 5' end. Primers were diluted to 10 µL and ready for use. Also enclosed was the -21M13 primer to do the sequencing of the cloned HV1 region of the CHR DNA, which covered base pairs 16133 through 16569 and 1 through 40.

3. The protocol used at NIST to amplify and sequence the DNA. The laboratories, however, were

free to use any protocol with which they were familiar and felt comfortable.

4. A form table to record the results. This table provided the number of the Primer Set, the region that each Primer Set amplified, and the length of the amplified region. We requested that the laboratory fill in the differences found when they compared the sequence that they determined for the mtDNA from CHR with that of the Cambridge Reference Sequence (Anderson *et al.,* 1981).

5. All the laboratories received the following cautionary note:

WARNING: The DNA and cells were derived from a cell culture line from an apparently healthy human subject. The cell culture line has been tested and found to be nonreactive for hepatitis B surface antigen and HIV. However, no test method can ensure that a product derived from human blood does not contain HIV, hepatitis or other infectious agents. HANDLE AS IF CAPABLE OF TRANSMITTING DISEASE. (The second CHR cell culture line generated from the same individual was not tested again for hepatitis or HIV. Normal precautions should be used.)

2.10. Differences in Methodology used by Laboratories in Interlaboratory Evaluation of SRM 2392, 1999.

The Bode Technology Group, Inc. essentially followed the NIST protocol except used a 6% acrylamide/8.3 mol/L urea gel for the sequencing electrophoresis instead of a 4.75% acrylamide.

IIT Research Institute also followed the NIST protocol except used Taq Gold (PerkinElmer) for the amplification reaction which was modified to include a hot start of 95 ^{o}C for 11 minutes. Microcon 100 microconcentrators (Amicon, Inc., Beverly, MA) were used to purify the PCR products. The quantities of DNA were determined by capillary electrophoresis (CE) with a Beckman P/ACE 5010 System (Beckman Instruments, Inc., Fullerton, CA) as follows: 1 µL of the amplified product was mixed with 25 µL of sterile deionized H_2O containing 0.52 ng/µL of a 200 bp internal standard (GenSura Laboratories, Inc., Del Mar, CA) and run on the CE. One determines the quantity of the amplified product from the ratio of the PCR product peak area to the internal standard peak area multiplied by a migration standard. A 6% acrylamide gel was used for the sequencing electrophoresis instead of a 4.75% acrylamide.

Lark Technologies, Inc., followed the NIST protocol with the following differences: Ampli-Taq DNA polymerase (PerkinElmer) was used to amplify the DNA; the dNTP's were purchased from Pharmacia Biotech, Inc. (Piscataway,NJ); the products were purified with Qiaquick PCR purification kit (Qiagen); in the sequencing reactions, the amount of PCR product used varied from 1 to 3 µL based on the concentration estimated from agarose gels; cycling conditions were 95 ^{o}C for one minute, followed by 25 cycles of 96 ^{o}C for 15 seconds, 50 ^{o}C for 15 seconds, and 60 ^{o}C for 4 minutes; the sequence reactions were cleaned up by ethanol precipitation; a 4.25% polyacrylamide gel was used for the sequencing electrophoresis instead of a 4.75% acrylamide; and a ABI 377 automated sequencer was used instead of the ABI 373. Electropherograms were printed for each reaction and the sequences were manually edited based on the electropherogram patterns. Printed electropherograms and a floppy disc with the sequence data were sent to NIST where the data were compared to the Cambridge Reference Sequence (Anderson *et al.,*1981).

2.11. Interlaboratory Evaluation for SRM 2392-I, 2002

Three laboratories, in addition to NIST, participated in an interlaboratory evaluation of the mtDNA sequence of HL-60. These laboratories included the FBI Laboratory, FBI Academy, Quantico, VA 22135; Armed Forces DNA Identification Laboratory (AFDIL), Armed Forces Institute of Pathology, Rockville, MD 20850; and the Georgia Bureau of Investigation (GBI), Decatur, GA 30034. Each laboratory was asked to amplify and sequence the entire mtDNA of HL-60. NIST provided: 1. A tube of DNA containing the extracted DNA ready for PCR amplification, 2. The 58 sets of primers labeled with either F# (forward primer) or R# (reverse primer), 3. A table to record the results, and 4. Any other supplies that were needed and requested by the participants. They were allowed to use any protocol for amplification or sequencing that they wished, but were requested to provide a copy of that protocol to NIST. NIST also requested copies of the electropherograms to enable us to resolve any discrepancies; although, as it turned out, there were no discrepancies.

2.12. Differences in Methodology Used by the Laboratories in the Interlaboratory Evaluation for SRM 2392-I

2.12.1. Armed Forces DNA Identification Laboratory (AFDIL), SRM 2392-I

AFDIL has developed a high-throughput, automated sequencing procedure using 12 primer sets that produce overlapping PCR products ranging from 825 to 1886 bp. The primers used to amplify the first 11 products are based on those published in Levin *et al.* (1999). They are designated: Amp01 – F361/R2216; Amp02 – F1993/R3557; Amp03 – F3441/R4983; Amp04 – F4797/R6526; Amp05 – F6426/R8311; Amp06 – F8164/R9848; Amp07 – F9754/R11600; Amp08 – F11403/R13123; Amp09 – F12793/R14388; Amp10 – F14189/R15396; Amp11 – F15260/R16084. The primers used to amplify the control region were developed at AFDIL and are Amp12 – F15878/R649 (F15878 is TTAACTCCACCATTAGCACC and R649 is TTTGTTTATGGGGTGATGTGA).

2.12.1.1. AFDIL PCR Amplification, SRM 2392-I:

The PCR mixture contained HL-60 DNA (1 µL), AmpliTaqGold® DNA polymerase (1 µL) (Applied Biosystems), 10x PCR buffer (5 µL) (Applied Biosystems), dNTP's (0.2 mmol/L) (Invitrogen), 2 µL of forward and reverse primers (10 µmol/L) (MWG Biotech, High Point, NC) plus dH_2O to a final volume of 50 µL. The 10x PCR buffer was the same as that used by NIST. Thermal cycling was conducted in a PerkinElmer 9700 thermocycler with the following conditions: 10 min at 96 ^oC (activation of AmpliTaq Gold®), plus 40 cycles of 94 ^oC for 15 sec, 56 ^oC for 30 sec, and 72 ^oC for 1 min. The purity and size of the PCR products were assessed by electrophoresis in a 0.7% agarose gel containing 0.3 µg/mL of ethidium bromide. The PCR products were purified with Shrimp Alkaline Phosphatase/Exonuclease I (Amersham Pharmacia, Piscataway, NJ). 5 µL of exonuclease I (10 U/µL) and 10 µL of Shrimp Alkaline Phosphatase (1 U/µL) was added to each tube containing PCR product. The tubes were heated at 37^oC for 15 min followed by 94^oC for 15 min in a PerkinElmer 9700 thermocycler.

2.12.1.2. AFDIL Sequencing, SRM 2392-I :

Cycle sequencing was performed with the ABI PRISM® BigDye® terminators (original version) cycle sequencing kit (Applied Biosystems). The sequencing mixture contained 9 µL dH_2O, 6 µL

BigDye® dilution buffer (400 mmol/L TRIS, 10 mmol/L $MgCl_2$, pH 9.0), 2 μL BigDye® terminator reaction mixture, 1 μL of forward or reverse primer (10 μmol/L each) and 2 μL of HL-60 PCR product for a total volume of 20 μL. A few of the sequencing primers (e.g., R649) required the use of the ABI PRISM® dGTP BigDye® terminator kit (Applied Biosystems). Thermal cycling was conducted in a PerkinElmer 9700 thermocycler at the following conditions: an initial 1 min denaturation at 96 °C, followed by 25 cycles of 94 °C for 15 sec, 50 °C for 5 sec, and 60 °C for 2 min. The DNA product was purified by filtration through a spin column matrix (Edge BioSystems, Gaithersburg, MD). Electrophoresis and sequencing were performed with an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems) using POP-6™ polymer (Applied Biosystems) with a 50 cm capillary. Data analysis was executed using Sequencher Plus 4.0.5b11 (Gene Codes, Ann Arbor, MI). The HL-60 sequence differences were identified by comparison to the Cambridge Reference Sequence as revised by Andrews *et al.* (1999). In most cases, sequence information was acquired for both the forward and reverse directions. In some regions, two separate reactions using the same primer were conducted (indicated by 2X in the following list of primers). A total of 95 sequencing reactions plus 1 pGEM reaction were conducted in a 96 well format. If the primer failed the first trial, the reaction was repeated. The finding of a heteroplasmy at np 12071 was also confirmed by an additional PCR and sequencing reaction. The following primers from Levin *et al.* (1999) were used to sequence the 12 PCR amplicons:

Amp01 (F361/R2216): F361, R921, F1234, R1425, F873, R2216, F1657, R1769
Amp02 (F1993/R3557): F1993, R2660, R2834, R3557, F2417, R3006, F3234
Amp03 (F3441/R4983): F3441, R3940, F3931 (2X), R4982, F4392, R4162
Amp04 (F4797/R6526): F4797 (2X), R6526, F5700 (2X), F5318, R5882, F6242
Amp05 (F6426/R8311): F6426 (2X), R7255, F7645 (2X), R8311, F7075, R7792
Amp06 (F8164/R9848): F8164 (2X), R9059, F8903, R9848, F8539, R9403, F9309
Amp07 (F9754/R11600): F9754 (2X), R10556, F11001 (2X), R11600, F10386, R11267
Amp08 (F11403/R13123): F11403 (2X), F12357, R13123, F11901 (2X), F12601, R12876 Amp09 (F12793/R14388): F12793, R13611, F13518 (2X), R14388, F13188, R13343, F13899, R13935
Amp10 (F14189/R15396): F14189 (2X), R15396, F14909, R14996, F14470
Amp11 (F15260/R16084): F15260, R16084, F15574, R15774
Amp12 (F15878/R649): F15971, R16175 (2X), F16450 (2X), R274, F314 (2X), R649 (2X), F16190, R16400
In Amp12, F15971 came from Levin *et al.* (1999). The other primers were designed by AFDIL and were as follows:
R16175: TGGATTGGGTTTTTATGTA
F16450: GCTCCGGGCCCATAACACTTG
R274: TGTGTGGAAAGTGGCTGTGC
F314: CCGCTTCTGGCCACAGCACT
R649: TTTGTTTATGGGGTGATGTGA
F16190: CCCCATGCTTACAAGCAAGT
R16400: GTCAAGGGACCCCTATCTGA

2.12.2. Georgia Bureau of Investigation (GBI), Interlaboratory Evaluation, SRM 2392-I

The Georgia Bureau of Investigation used the same protocol as that used by NIST with the

following exceptions: 1. A PerkinElmer model 9600 was used for thermal cycling; 2. Amplified DNA was electrophoresed in 2.75% agarose gels; 3. The cycle sequencing was performed with a PerkinElmer model 9600 using the program provided with the BigDye® kit (25 cycles of 96 °C for 10 sec, 50 °C for 5 sec, and 60 °C for 4 min); 4. Some samples that needed to be cycle sequenced again using more amplicon were not purified before precipitation; 5. The samples were precipitated using isopropanol precipitation as per the BigDye® instructions; and 6. Comparison of the sequence data was performed with Sequencer 3.1.1 software (Gene Codes, Ann Arbor, MI).

2.12.3. Federal Bureau of Investigation (FBI), Interlaboratory Evaluation, SRM 2392-I

2.12.3.1. FBI Polymerase Chain Reaction (PCR):

The PCR mixture contained from 0.1 to 1.4 ng HL-60, AmpliTaq Gold® Polymerase (2.5 units) (Applied Biosystems), 10x PCR buffer (5 µL) (Applied Biosystems), GeneAmp® dNTPs (0.2 mmol/L each) (Applied Biosystems), forward and reverse primers (0.4 µmol/L each), plus dH_2O to a final volume of 50 µL. The 10x buffer (pH 8.3) was the same as used by NIST. Thermal cycling was conducted in a GeneAmp® PCR System 9700 (PerkinElmer) and consisted of 10 min at 95 °C followed by 35 cycles of 94 °C for 20 sec, annealing temperatures of 50 °C (primer set 49), 51 °C (primer sets 1, 7, 44, 57), 52 °C (primer sets 6, 8, 30, 45), 53 °C (primer set 53) and 56 °C (all other primer sets) for 10 sec, and 72 °C for 30 sec and ending with a final extension of 7 min at 72 °C. Amplified products were purified by treatment with Exo-SAP-IT (5 µL for every 25 µL of PCR product) (USB Corp., Cleveland, OH). Samples of the PCR products were electrophoresed in 1.2% agarose gels containing ethidium bromide to assess the purity, size, and quantity of the PCR products.

2.12.3.2. FBI Sequencing:

Cycle sequencing using fluorescent dye–labeled terminators was performed with an ABI PRISM® dRhodamine Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (Applied Biosystems). Thermal cycling was conducted in a GeneAmp® PCR System 9700 (Applied Biosystems) and started with 1 min at 96 °C followed by 25 cycles of 96 °C for 15 sec, 50 °C for 1 sec and 60 °C for 1 min and ended with a final incubation at 15 °C for 10 min. These products were purified with Centri-Sep™ spin columns (Princeton Separation, Inc., Adelphia, NJ) and electrophoresed on an ABI PRISM® 310 Genetic Analyzer (Applied Biosystems) using POP-6™ and either a 47 or 61 cm capillary. Comparison of the sequence data was performed with Sequencer 4.1 software (Gene Codes).

The FBI also discovered the problem with original primer set 51 (see Materials and Methods). The reverse primer incorporated a G at position 14368 [an error in the original Cambridge Reference Sequence (Anderson *et al.,* 1981) that was found and corrected by Andrews *et al.* (1999)]. However, sequencing with primer set 52 showed a C at position 14368. They prepared a new set of primers (primer set 51.5) to amplify approximately a 400 bp fragment that encompassed np 14368. The new primers were:
F14217: 5'-CTAATCAACGCCCATAATCATAC-3' and
R14620: 5'-GTTTTCTTCTAAGCCTTCTCC-3'.
The new primer set confirmed that the correct base at position 14368 was a C.

2.13. Permissions

The development of SRM 2392-I was deemed exempt from the policy of Part 27 of Title 15 of the Code of Federal Regulations by the NIST Institutional Review Board and the Director of the Chemical Science and Technology Laboratory. This work fit into the exemption category described in 15 CFR 27.101(b)(4) which states: "Research, involving the collection or study of existing data, documents, pathological specimens, or diagnostic specimens, if, these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects."

ATCC also waived condition 3(c) in their Material Transfer Agreement which states that the "purchaser shall not sell, lend, distribute or otherwise transfer the material or replicates to any others" for the use of HL-60 in the NIST mtDNA SRM. They stated that, in their view, "as a government agency, NIST will not be providing this material as a commercial product despite the collection of fees for the SRM."

2.14. Final Preparation of DNA for SRM 2392 and SRM 2392-I

2.14.1. CHR DNA Preparation and Confirmation SRM 2392:

CHR cells were grown in RPMI-1640 media supplemented with L-glutamine, sodium bicarbonate, penicillin, streptomycin and 10% fetal calf serum (Sigma Chemical Co.), centrifuged and resuspended in 10 mL of phosphate buffered saline (PBS) to a final cell count of 1.2×10^6 cells/mL. DNA was extracted using the Qiagen Maxi Kit according to the Qiagen protocol (Qiagen Inc.). The concentration of the DNA in the extract was determined with an UV spectrophotometer to be 82.3 ng/μL.

PCR amplification was performed on the extract using primer set 57 (HV1 region) and primer set 1 (HV2 region) in order to verify the sequence. The extract was diluted 1/10 (8.23 ng/μL), and 1 μL was used in PCR reaction. The PCR reaction mixture was the same as that described in the polymerase chain reaction section of Materials and Methods, with the exception that the forward and reverse primers were used at a concentration of 0.2 μM.

Thermal cycling was conducted in a PerkinElmer 9600 (HV1) or 9700 (HV2) and consisted of 30 sec at 95 ^{o}C, followed by 32 cycles of 20 sec at 94 ^{o}C (denaturation), 10 sec at 56 ^{o}C (annealing), and 30 sec at 72 ^{o}C (extension), and ending with a final extension of 7 minutes at 72 ^{o}C. A sample of the amplified DNA was electrophoresed and stained with ethidium bromide and purified as described in the polymerase chain reaction section of Materials and Methods except that a 2% agarose gel was used. Cycle sequencing, electrophoresis of the products, and the analysis of the sequence data were performed as described in the sequencing section of Materials and Methods. The polymorphisms of the HV1 and HV2 regions and the C-stretch (HV1 region) seen in the pre-production mode samples were confirmed in the production mode samples.

The DNA extract was diluted to 1 ng/μL with TE buffer and PCR amplified using Primer Set 1 (HV2) to verify that this concentration was appropriate to use for PCR. Sixty μL of the CHR DNA extract at a concentration of 1 ng/μL was aliquotted into 260 tubes.

2.14.2. CHR Cloned DNA Production, SRM 2392:

DNA from one of the C-stretch clones (containing 12 C's) prepared as described earlier was used to transform TG-1 cells. Following the transformation and growth of the clones, the bacteriophage DNA was isolated and sequenced to verify the presence of the 12 C region. Following this verification, a large scale transformation was performed for SRM production. The cloned DNA that was obtained following this transformation was quantitated by UV spectrophotometry and diluted to 100 ng/μL with TE buffer. The cloned DNA was sequenced and analyzed as described in section entitled Bacteriophage DNA Isolation and Sequencing in Materials and Methods to verify that it contained the C-stretch of 12 C's plus all the additional polymorphisms previously found in the HV1 region of the CHR cells. For the final SRM, 260 tubes containing 10 μL at a concentration of 100 ng/μL were prepared.

2.14.3. 9947A DNA Production, SRM 2392:

Cell culture line 9947A was grown by Life Technologies, Inc. (Gaithersburg, MD) who extracted the DNA and submitted a sample (at a concentration of 1 ng/μL) to NIST for verification of the sequence. The HV1 (primer set 57) and HV2 (primer set 1) regions as well as the region amplified by primer set 30 were checked by PCR and sequencing to verify that the DNA was the same as sequenced earlier. One μL of DNA (1 ng) was used as the template in the PCR reaction mixture, which was performed as described in the polymerase chain reaction section of the Materials and Methods section, with the exception that the forward and reverse primers were used at a concentration of 0.2 μmol/L. Thermal cycling was conducted in a PerkinElmer 9700 and consisted of 30 sec at 95°C, followed by 32 cycles of 20 sec at 94 °C (denaturation), 10 sec at 56 °C (annealing), and 30 sec at 72 °C (extension), and ending with a final extension of 7 minutes at 72 °C. A sample of the amplified DNA was electrophoresed and stained with ethidium bromide and purified as described in the section entitled polymerase chain reaction except that a 1% agarose gel was used. Cycle sequencing, electrophoresis of products, and analysis of sequence data were performed as described earlier. Examination of the results from primer sets 57 (HV1), 1 (HV2), and 30 indicated that the DNA supplied by Life Technologies was the same as that examined earlier in the pre-production mode.

After NIST verified that the sequence was okay, Life Technologies, Inc. prepared 250 tubes containing 60 μL of 9947A DNA at a concentration of 1 ng/μL.

2.14.4.HL-60 Production, SRM 2392-I:

DNA from the HL-60 cell culture was extracted, isolated, and quantified by the Professional Services Department of the American Type Culture Collection (ATCC, Manassas, VA). DNA was isolated from the HL-60 cell culture using the QIAamp DNA Blood Mini Kit (Qiagen, Inc.). Quantification was determined by the Quantiblot Human DNA Quantification Kit (Applied Biosystems). The final concentration, 1.4 ng/μL, is based on 12 replicate tests. The DNA purity (A260:A280 = 1.9) was determined spectrophotometrically. The integrity of DNA was determined electrophoretically using 0.4% agarose gels.

Sixty-five μL of HL-60 DNA were dispensed into 260 tubes. Three randomly chosen tubes were used to send to the laboratories that agreed to participate in the Interlaboratory Evaluation. Each of the three laboratories plus NIST used all 58 primer sets to sequence the entire 16,569 base pairs of the HL-60 mtDNA. All four laboratories found the identical sequence, thereby confirming the accuracy of the sequence determined at NIST.

3. Results And Discussion

3.1. SRM Templates:

Two DNA templates, CHR and 9947A, are included in the NIST human mtDNA amplification and sequencing SRM 2392. The CHR and 9947A DNA samples came from human cell culture lines that were developed from apparently normal individuals. The DNA template in SRM 2392-I, HL-60, came from a promyelocytic cell line prepared from the peripheral blood leukocytes from an individual with acute promyelocytic leukemia. The DNA extracted from 9947A and HL-60 is the total DNA including nuclear DNA. The DNA from CHR was isolated in a manner which enhanced the concentration of the mtDNA, but did not totally eliminate the nuclear DNA. However, both types of extracted DNA amplified and sequenced accurately. SRM 2392 also provides cloned DNA from the HV1 region of the CHR template which contains a C-stretch. In most people, the HV1 region has a string of cytosine (C) residues interrupted by a thymine (T) at np 16189. In some individuals, however, a transition occurs which changes the T to a C producing a long string of C's called the C-stretch. When this happens, sequencing beyond the C-stretch becomes very difficult, if not impossible. Clones of the HV1 region containing the C-stretch indicated that the number of C's differed among the different clones and the difficulty in sequencing was due to frameshifts that resulted from the simultaneous sequencing of templates with differing numbers of C's (Bendall and Sykes, 1995; Levin *et al.,* 1995; Levin *et al.,* 1997). We found, however, that one could sequence through the entire HV1 region including the C-stretch without problems if one used the clone of the area. Therefore, we included the cloned HV1 region of the CHR DNA template in SRM 2392.

In addition to templates CHR and 9947A in SRM 2392 and HL-60 in SRM 2392-I, we have included all the mtDNA sequence information from two additional templates in Table 2: 1. GM03798, a normal lymphoblastoid cell line and 2. GM10742A from an individual with the mtDNA disease, Leber Hereditary Optic Neuropathy (LHON), which causes blindness in early adulthood. Both of these cell lines were obtained from the NIGMS Human Genetic Mutant Cell Repository. The DNA from these cell lines are not part of SRMs 2392 or 2392-I and the data are included for information and comparison purposes only.

3.2. Primers:

The 58 sets of unique primers were designed to allow the amplification and sequencing of any region or the entire 16,569 base pairs that comprise human mtDNA. The sequence of both the forward and reverse primers that are in each set are shown in Table 1. The numbers indicate the 5' end of the primer. They are all between 19 and 28 base pairs long and the criteria that were used to choose these primers were primer melting temperatures (T_m) between 50 and 65 ^oC, primer length between 15 and 30 base pairs, and PCR product length between 220 and 1000 base pairs. In actuality, these primers gave PCR product lengths between 224 and 976 base pairs. The primers

were designed to produce sequences that overlapped with both the previous and following regions to allow those areas in the beginning and end of electropherograms which are difficult to sequence to become readable and to prevent any gaps in the sequence. The readable region is always smaller than that of the amplified region. In the process of developing SRM 2392-I, we found that reverse primer of primer set 51 needed to be changed (see explanation in section entitled "mtDNA primers" in the Materials and Methods section). The new reverse primer 51 is shown in Table 1. The change is in bold and underlined.

In addition to the designed primers, we also used the -21M13 primer (Table 1) to sequence the cloned DNA from the HV1 region of the CHR template which contained the C-stretch. The PCR products were produced under the same conditions for all 58 primer sets and generated single, distinct bands in all cases (Levin *et al.*, 1999).

3.3. Differences between the SRM DNAs and the Cambridge Reference Sequence:

Human mtDNA was completely sequenced in 1981 and is referred to as the Cambridge Reference Sequence (Anderson *et al.,* 1981). This sequence is a consensus sequence based primarily on a placenta, a few areas sequenced from HeLa cells, and five ambiguous areas where the sequence from bovine mtDNA was used. All investigators, who subsequently examined human mtDNA, have used the numbering system of the Cambridge Reference Sequence and have compared their sequence findings to those found in 1981. In 1999, Andrews *et al.* resequenced the original placenta used in 1981. They found that the 1981 sequence had a number of errors and that the placenta had a number of rare polymorphisms. At the present time, investigators compare their sequences to the 1999 revised sequence. However, the DNA from this placenta is not available for use as a positive control during actual experiments; whereas, NIST SRMs 2392 and 2392-I are available. Table 2 shows the mtDNA differences compared to the Cambridge Reference Sequence that were found at NIST with the three templates - CHR and 9947A (in SRM 2392) and HL-60 (in SRM 2392-I) and the two additional templates - GM03798, and GM10742A - that are included for information and comparison purposes. In all five templates, all 58 areas comprising the entire mtDNA were completely amplified and sequenced at least twice. There are 13, 9, 11, 4, and 12 differences in the non-coding regions of templates CHR, 9947A, HL-60, GM03798, and GM10742A, respectively and 33, 23, 33, 19, and 31 differences in the coding regions of templates CHR, 9947A, HL-60, GM03798, and GM10742A, respectively. All of the differences from the Cambridge Reference Sequence found in CHR, 9947A, HL-60, GM03798, and GM10742A are shown in Figure 1 along with many of the mutations associated with Leber Hereditary Optic Neuropathy (LHON) that have been noted in the literature (Wallace *et al.,* 1997). With one exception, none of the base pair changes found in the coding regions of the apparently normal templates (CHR, 9947A, GM03798) sequenced at NIST correlate with any of the changes found associated with the published disease states found in the MitoMap web site: http://www.mitomap.org The one exception was a secondary LHON mutation found in CHR. HL-60 and GM10742A, however, did have some mutations that have been associated with LHON.

3.4. Meaning of the Differences from Cambridge Reference Sequence:

Since three templates came from apparently normal individuals, one from an individual with acute promyelocytic leukemia, and one from an individual with LHON, it was of interest to determine if

the differences in the coding regions would actually cause amino acid changes in the resultant protein structures. The genetic code for human mtDNA is slightly different from the universal genetic code. Table 3 shows the differences in the universal genetic code that one needs to consider when determining the amino acid sequence designated by the three base pair codons in mtDNA. Many of the differences from the Cambridge Reference Sequence were in the third position wobble and did not affect the amino acid sequence. However, CHR, 9947A, HL-60, GM03798, and GM10742A had 11, 12, 15, 8, and 14 different base pairs, respectively, that would result in a different amino acid from that designated by the Cambridge Reference Sequence of 1981. However, comparison with the revised Cambridge Reference Sequence of 1999, indicate that only 3, 4, 7 plus 2 in tRNAs, 0, and 7 would result in amino acid changes in CHR, 9947A, HL-60, GM03798, and GM10742A, respectively. These structural changes, however, do not necessarily mean a functional change has occurred in the protein. To determine if a functional change has occurred, one still needs to decipher whether the amino acid change is in an active site on the protein. We have developed an interactive web site that will easily and quickly determine if the DNA change causes an amino acid change (Lee and Levin, 2002). The changes that result in amino acid changes are shown in Table 2.

3.5. The Interlaboratory Evaluation of the CHR and HL-60 Templates:

The interlaboratory evaluation of CHR was conducted by four laboratories including NIST. The interlaboratory evaluation of HL-60 was conducted by four laboratories including NIST. Any changes made by the various laboratories to the NIST protocol are listed in the section on Materials and Methods. Each laboratory was instructed to amplify and sequence the 58 areas designated by the 58 primer sets and those laboratories that sequenced CHR were also to sequence the cloned DNA from the HV1 region. Those sequencing CHR were designated Labs 1, 2, 3, and 4. Labs 1, 2 and 3 found essentially the same polymorphisms. Laboratory 4 had less experience with sequencing mtDNA and did find differences that the other laboratories did not observe. Interlaboratory data was considered ambiguous and excluded if the following problems were observed: 1. If the computer results were ambiguous as indicated by calling a peak "N" rather than A, C, G or T. 2. If the differences from the Cambridge Reference Sequence were not consistently seen within a laboratory, i.e., if the laboratory sequenced in both the forward and reverse directions and one direction agreed with the Cambridge Reference Sequence and the other direction did not agree, we assumed the results that agreed with the Cambridge Reference Sequence were correct. 3. If, within any one laboratory, the difference from Cambridge Reference Sequence was seen with one primer set, but not in the overlapping sequences seen in the previous or subsequent primer sets, the results agreeing with the Cambridge Reference Sequence were considered correct. Even with these exclusions, Laboratory 4 had many differences that were not seen by the other labs. One problem was that they did not provide data from primer sets 29, 39, and 41 and we were unable to check the overlapping sequences. Laboratory 2 was unable to sequence the clone, which was not a problem for the other laboratories. Laboratory 3 was missing data from primer sets 36 and 48. Laboratories 1 and 3 noted a heteroplasmy[3] at base number 6849 (the Cambridge Reference Sequence found an A at this site).

[3]In final preparation of SRM 2392, a new blood sample was obtained from CHR and a new cell culture line was established. The sequence analysis of the new CHR was identical to the first cell line except no heteroplasmy was found at bp 6849. Therefore, the cell line supplied with SRM 2392 does not have this heteroplasmy. The data on the first cell line are included in the text to indicate the agreement in the interlaboratory evaluation which was done with the first cell line.

Laboratory 1 found a G at this site, but closer examination of the electropherogram showed that an A peak existed under the G peak. Laboratory 3 also noted the A/G heteroplasmy at this site. Laboratory 4 did not note the heteroplasmy, but when their electropherograms were examined at NIST, the A/G heteroplasmy was noted. NIST did not have the electropherograms of Laboratory 2, but on questioning them, they agreed that the heteroplasmy was there, but that they had missed it. One of the problems with finding heteroplasmic sites is that if the computer call is the same as the Cambridge Reference Sequence, one would not necessarily examine that site more closely. If the computer call is different from the Cambridge Reference Sequence, one would look more closely at the electropherogram and then note the presence of a smaller peak under the main peak. The cell line CHR was reestablished with fresh blood and the entire 16,569 base pairs were amplified and sequenced again. The heteroplasmy at 6849 was not present in the new cell line. The new cell line agreed with the Cambridge Reference Sequence at this site. It is the DNA from the new cell line that is in SRM 2392.

With the exceptions of the differences noted here, the interlaboratory evaluation of CHR was successful in that most of the laboratories found the same results. The many differences noted by Laboratory 4, who was less experienced at sequencing mtDNA, confirms and emphasizes the need for Standard Reference Materials for sequencing mtDNA. If Laboratory 4 had the NIST mtDNA SRM 2392 or SRM 2392-I and had run it along side of their unknown sample, they would have realized that they were finding an undue number of differences and could have reexamined their procedures to try to determine the reason for these excessive differences.

The Interlaboratory Evaluation of the mtDNA sequence of HL-60 was conducted by four laboratories (see Materials and Methods section) including NIST. All the laboratories produced identical results).

4. Conclusions

Two NIST Standard Reference Materials (SRM 2392 and 2392-I) have been prepared. These SRMs allow one to amplify and sequence any region or the entire 16,569 base pairs which comprise human mtDNA. Fifty-eight pairs of unique primers have been designed, tested and shown to work well in the amplification and sequencing procedures (note: the reverse primer for primer set 51 has been changed from that originally listed in SRM 2392, in Levin *et al*., 1999, and in Appendix C). The three DNA templates (CHR and 9947A in SRM 2392, and HL-60 in SRM 2392-I) have characteristic polymorphisms throughout the non-coding and coding regions of the DNA and therefore, can serve as positive controls during PCR amplification and sequencing. With one exception, none of the polymorphisms found in CHR and 9947A correspond to any of the published base pair changes that have been correlated with specific diseases. The one exception is that CHR has one mutation that has been associated with a secondary mutation associated with LHON. HL-60, however, has 4 mutations that have been associated with the mitochondrial disease LHON.

Compared to the 1981 Cambridge Reference Sequence, CHR mtDNA had 13 differences in the non-coding regions and 33 differences in the coding regions, 9947A mtDNA had 9 differences in the non-coding regions and 23 differences in the coding regions, HL-60 had 11 differences in the non-coding regions and 33 differences in the coding regions. GM03798 and GM 10742A, whose data are included for comparison and information, had 4 and 12 differences in the non-coding regions,

respectively and 19 and 14 differences in the coding regions, respectively. All of these differences are shown in Fig. 1 and Table 2. These differences in the coding regions do result in some amino acid changes in the proteins coded for by mtDNA. Four laboratories participated in an interlaboratory evaluation of the CHR template and four laboratories participated in the interlaboratory evaluation of HL-60; some differences between laboratories were noted in the CHR evaluation, but, in general, agreement was good. No differences were noted in the HL-60 interlaboratory evaluation. The use of NIST SRM 2392 and SRM 2392-I will provide quality control to the scientific and medical communities when they amplify and sequence human mtDNA.

5. Acknowledgments

The authors are grateful to The Bode Technology Group, Inc., Sterling, VA; IIT Research Institute, Newington, VA; and Lark Technologies, Inc., Houston, TX for their participation in the Interlaboratory Evaluation of the CHR template and to the FBI Laboratory, FBI Academy, Quantico, VA; the Armed Forces DNA Identification Laboratory (AFDIL), Armed Forces Institute of Pathology, Rockville, MD; and the Georgia Bureau of Investigation (GBI), Decatur, GA for their participation in the Interlaboratory Evaluation of HL-60. The research to prepare SRM 2392-I containing HL-60 was supported by Interagency Agreement number 1999-IJ-R-094 from the National Institute of Justice. The authors also acknowledge the funding received from the National Institute of Justice through the NIST Office of Law Enforcement Standards and the support from the Advanced Technology Program at NIST to develop SRM 2392. Points of view are those of the authors and do not necessarily represent the position of the U.S. Department of Justice. The authors thank Prasad Reddy for his help and advice on cloning, Lois A. Tully (when serving as a post-doctoral research associate at NIST) and Julie T. Chen for checking selected sequence areas of DNA samples from CHR, the CHR clone, and 9947A before the final assembly of SRM 2392 that became available in 1999. We also thank MaryPat Jones for confirming the sequences of some areas of the second generation of the CHR cell line. The authors are grateful to Gettysburg College for the partial funding of Koren A. Holland during her sabbatical year at NIST. We also acknowledge and thank Mary Pat Jones and Kathryn Kapfer for help in sequencing GM10742A at NIST.

6. References

Anderson, S., Bankier, A. T., Barrell, B. G., deBrujin, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**: 457-465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N. (1999). Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA. *Nature Genetics* **23**:147.

Ballinger, S. W., Bouder, T. G., Davis, G. S., Judice, S. A., Nicklas, J. A., and Albertini, R. J. (1996). Mitochondrial genome damage associated with cigarette smoking. *Cancer Research* **56**: 5692-5697.

Bendall, K. E., and Sykes, B. C. (1995). Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am. J. Hum. Genet.* **57**: 248-256.

Bogenhagen, D., and Clayton, D. A. (1974). The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. *J. Biol. Chem.* **249**: 7991-7995.

Grossman, L. I. (1995). Mitochondrial mutations and human disease. *Env. & Mol. Mutag.* **25**: (supplement 26) 30-37.

Holland, M. M., Fisher, D. L., Mitchell, L. G., Rodriquez, W. C., Canik, J. J., Merril, C. R., and Weedn, V. W. (1993) Mitochondrial DNA sequence analysis of human skeletal remains: Identification of remains from the Vietnam War. *J. Forensic Sciences* **38**: 542-553.

Holland, M. M., Fisher, D. L., Roby, R. K., Ruderman, J., Bryson, C., and Weedn, V. W. (1995) Mitochondrial DNA sequence analysis of human remains. *Crime Laboratory Digest* **22**: 109-115.

Howell, N., McCullough, D. A., Kubacka, I., Halvorson, S., and Mackey, D. (1992). The sequence of human mtDNA: the question of errors versus polymorphisms. *Am. J. Hum. Genet.* **50**: 1333-1337.

King, M. P., and Attardi, G. (1989). Human cells lacking mtDNA: Repopulation with exogenous mitochondria by complementation. *Science* **246**: 500-503.

Lee, M.S. and Levin, B.C. (2002). MitoAnalyzer, a computer program and interactive web site to determine the effects of single nucleotide polymorphisms and mutations in human mitochondrial DNA. *Mitochondrion* **1:** 321-326.

Levin, B. C., Cheng, H., O'Connell, C., Reeder, D. J., and Holland, M. (1995). Evaluation of alternative enzymes and additives to improve sequencing of mitochondrial DNA. Proceedings of the Fifth International Symposium on Human Identification, Scottsdale, AZ, October 8-12, 1994, p. 170.

Levin, B. C., Cheng, H., Holland, M., and Reeder, D. J. (1997). Heteroplasmy responsible for difficulty experienced in sequencing the human mitochondrial DNA HV1 region containing the C-stretch. Proceedings of the Seventh International Symposium on Human Identification, Scottsdale, AZ, September 19-21, 1996, p. 166.

Levin, B. C., Cheng, H., and Reeder, D. J. (1999). A human mitochondrial DNA Standard Reference Material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics* **55**: 135-146.

Levin, B.C., Holland, K.A., Hancock, D.K., Coble, M., Parsons, T.J., Kienker, L.J., Williams, D.W., Jones, MP., and Richie, K.L., (2003). Comparison of the complete mtDNA genome sequences of human cell lines - HL-60 and GM10742A - from individuals with promyelocytic leukemia and leber hereditary optic neuropathy, respectively, and the inclusion of HL-60 in the NIST human mitochondrial DNA Standard Reference Material - SRM 2392-I. Mitochondrion (in press, 2003).

Marzuki, S., Lertrit, P., Noer, A. S., Kapsa, R. M. I., Sudoyo, H., Byrne, E., and Thyagarajan, D. (1992). Reply to Howell *et al*.: the need for a joint effort in the construction of a reference data base

for normal sequence variants of human mtDNA. *Am. J. Hum. Genet.* **50**: 1337-1340.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning, A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, NY.

Wallace, D. C. (1992). Mitochondrial genetics: a paradigm for aging and degenerative diseases. *Science* **256**: 628-632.

Wallace, D. C., Brown, M. D., and Lott, M. T. (1997). Mitochondrial genetics. In "Emery and Rimoin's Principles and Practice of Medical Genetics" (D. L. Rimoin, J. M. Connor, and R. E. Pyeritz RE, Eds.), Vol. 1, Third Edition, pp. 277-332, Churchill Livingstone, NY.

Table 1. Sequences for Primer Sets Used for PCR Amplification of Human mtDNA

Primer Set Number		Primer Sequence
1(HV2)	F15	CACCCTATTAACCACTCACG
	R484	TGAGATTAGTAGTATGGGAG
2	F361	ACAAAGAACCCTAACACCAGC
	R921	ACTTGGGTTAATCGTGTGACC
3	F756	CATCAAGCACGCAGCAATG
	R1425	AATCCACCTTCGACCCTTAAG
4	F873	GGTTGGTCAATTTCGTGCCAG
	R1425	AATCCACCTTCGACCCTTAAG
5	F1234	CTCACCACCTCTTGCTCAGC
	R1769	GCCAGGTTTCAATTTCTATCG
6	F1587	TGCACTTGGACGAACCAGAG
	R2216	TGTTGAGCTTGAACGCTTTC
7	F1657	CTTGACCGCTCTGAGCTAAAC
	R2216	TGTTGAGCTTGAACGCTTTC
8	F1993	AAACCTACCGAGCCTGGTG
	R2216	TGTTGAGCTTGAACGCTTTC
9	F2105	GAGGAACAGCTCTTTGGACAC
	R2660	AGAGACAGCTGAACCCTCGTG
10	F2417	CACTGTCAACCCAACACAGG
	R3006	ATGTCCTGATCCAACATCGAG
11	F2834	CCCAACCTCCGAGCAGTACATG
	R3557	AGAAGAGCGATGGTGAGAGC
12	F2972	ATAGGGTTTACGACCTCGATG
	R3557	AGAAGAGCGATGGTGAGAGC
13	F3234	AGATGGCAGAGCCCGGTAATC
	R3557	AGAAGAGCGATGGTGAGAGC
14	F3441	ACTACAACCCTTCGCTGACG
	R3940	TGAAGCCTGAGACTAGTTCGG
15	F3635	GCCTAGCCGTTTACTCAATCC
	R4162	TGAGTTGGTCGTAGCGGAATC
16	F3931	TCAGGCTTCAACATCGAATACG
	R4728	TTATGGTTCATTGTCCGGAGAG
17	F4183	TTTCTACCACTCACCCTAGCATTAC
	R4728	TTATGGTTCATTGTCCGGAGAG
18	F4392	CCCATCCTAAAGTAAGGTCAGC
	R4983	GGTTTAATCCACCTCAACTGCC
19	F4447	TTGGTTATACCCTTCCCGTAC
	R4982	GTTTAATCCACCTCAACTGCC
20	F4797	CCCTTTCACTTCTGAGTCCCAG
	R5553	AGGGCTTTGAAGGCTCTTG
21	F4976	ATTAAACCAGACCCAGCTACG
	R5553	AGGGCTTTGAAGGCTCTTG

22	F5318	CACCATCACCCTCCTTAACC
	R5882	GCTGAGTGAAGCATTGGACTG
23	F5700	TAAGCACCCTAATCAACTGGC
	R6262	GCCTCCACTATAGCAGATGCG
24	F5999	TCTAAGCCTCCTTATTCGAGC
	R6526	ATAGTGATGCCAGCAGCTAGG
25	F6242	CGCATCTGCTATAGTGGAGG
	R6526	ATAGTGATGCCAGCAGCTAGG
26	F6426	GCCATAACCCAATACCAAACG
	R7030	TGGGCTACAACGTAGTACGTG
27	F6744	GGCTTCCTAGGGTTTATCGTG
	R7255	TTTCATGTGGTGTATGCATCG
28	F7075	GAGGCTTCATTCACTGATTTCC
	R7792	GGGCAGGATAGTTCAGACGG
29	F7215	CGACGTTACTCGGACTACCC
	R7792	GGGCAGGATAGTTCAGACGG
30	F7645	TATCACCTTTCATGATCACGC
	R8215	GACGATGGGCATGAAACTG
31	F7901	TGAACCTACGAGTACACCGACTAC
	R8311	AAGTTAGCTTTACAGTGGGCTCTAG
32	F8164	CGGTCAATGCTCTGAAATCTGTG
	R8669	CATTGTTGGGTGGTGATTAGTCG
33	F8539	CTGTTCGCTTCATTCATTGCC
	R9059	GTGGCGCTTCCAATTAGGTG
34	F8903	CCCACTTCTTACCACAAGGC
	R9403	GTGCTTTCTCGTGTTACATCG
35	F9309	TTTCACTTCCACTCCATAACGC
	R9848	GAAAGTTGAGCCAATAATGACG
36	F9449	CGGGATAATCCTATTTATTACCTCAG
	R9995	AGAGTAAGACCCTCATCAATAGATGG
37	F9754	AGTCTCCCTTCACCATTTCCG
	R10275	AAAGGAGGGCAATTTCTAGATC
38	F10127	ACTACCACAACTCAACGGCTAC
	R10556	GGAGGATATGAGGTGTGAGCG
39	F10386	GGATTAGACTGAACCGAATTGG
	R11166	CATCGGGTGATGATAGCCAAG
40	F10704	GTCTCAATCTCCAACACATATGG
	R11267	TGTTGTGAGTGTAAATTAGTGCG
41	F11001	AACGCCACTTATCCAGTGAACC
	R11600	CTGTTTGTCGTAGGCAGATGG
42	F11403	GACTCCCTAAAGCCCATGTCG
	R11927	TTGATCAGGAGAACGTGGTTAC
43	F11760	ACGAACGCACTCACAGTCG
	R12189	AAGCCTCTGTTGTCAGATTCAC
44	F11901	TGCTAGTAACCACGTTCTGGTG
	R12876	GATATCGCCGATACGGTTG
45	F12357	AACCACCCTAACCCTGACTTCC

	R12876	GATATCGCCGATACGGTTG
46	F12601	TTCATCCCTGTAGCATTGTTCG
	R13123	AGCGGATGAGTAAGAAGATTCC
47	F12793	TTGCTCATCAGTTGATGATACG
	R13343	TTGAAGAAGGCGTGGGTACAG
48	F13188	CACTCTGTTCGCAGCAGTATG
	R13611	TCGAGTGCTATAGGCGCTTGTC
49	F13518	CATCATCGAAACCGCAAAC
	R13935	TGTGATGCTAGGGTAGAATCCG
50	F13715	GAAGCCTATTCGCAGGATTTC
	R14118	TGGGAAGAAGAAAGAGAGGAAG
51	F13899	TTTCTCCAACATACTCGGATTC
	R14388	TTAGCGATGGAGGTAGGATT**G**G (New Primer)
	R14388	TTAGCGATGGAGGTAGGATTCG (Old Primer)
52	F14189	ACAAACAATGGTCAACCAGTAAC
	R14926	TGAGGCGTCTGGTGAGTAGTGC
53	F14470	TCCAAAGACAACCATCATTCC
	R14996	CGTGAAGGTAGCGGATGATTC
54	F14909	TACTCACCAGACGCCTCAACCG
	R15396	TTATCGGAATGGGAGGTGATTC
55	F15260	AGTCCCACCCTCACACGATTC
	R15774	ACTGGTTGTCCTCCGATTCAGG
56	F15574	CGCCTACACAATTCTCCGATC
	R16084	CGGTTGTTGATGGGTGAGTC
57 (HV1)	F15971	TTAACTCCACCATTAGCACC
	R16451	GCGAGGAGAGTAGCACTCTTG
58	F16097	TACATTACTGCCAGCCACCATG
	R336	TTAAGTGCTGTGGCCAGAAG
-21M13	F	TGTAAAACGACGGCCAGT

HV2: Hypervariable region 2
HV1: Hypervariable region 1
F: forward primer
R: reverse primer
These are the same primers used for SRM 2392, Levin *et al.*, 1999, and Appendix C except the reverse primer of set 51 has been changed to: TTAGCGATGGAGGTAGGATT**G**G. The change (C to G) occurs at np 14368 and is in bold and underlined. Those using SRM 2392 or SRM 2392-I should use the new reverse primer 51.

Table 2. Certified Human mtDNA Sequence Differences from the Cambridge Reference Sequence (CRS) Found in the Two Templates (CHR and 9947A) in NIST SRM 2392, One Template (HL-60) in NIST SRM 2392-I and in GM03798 and GM10742A.

Comparison with the Cambridge Reference Sequence (CRS)								
CRS								
# [a]	Base[b] 1981/1999	Template CHR[d]	Template 9947A[d]	Template HL-60	Template GM03798	Template GM10742A	Amino acid change	Region
73	A	G	-	G	-	G		HV2
93	A	-	G	-	-	-		HV2
150	C	-	-	T	-	-		HV2
152	T	-	-	C	-	-		HV2
185	G	-	-	-	-	A		HV2
195	T	C	C	-	-	-		HV2
204	T	C	-	-	-	-		HV2
207	G	A	-	-	-	-		HV2
214	A	-	G	-	-	-		HV2
228	G	-	-	-	-	A		HV2
263*R	A	G	G	G	G	G		HV2
295	C	-	-	T	-	T		HV2
303-309	-	C (ins)	CC (ins)	-	-	-		HV2
311-315*R	-	C (ins)	C (ins)	C (ins)	C (ins)	C (ins)		HV2
462	C	-	-	-	-	T		HV2

482	T	-	-	-	-	C		HV2
489	T	-	-	C	-	C		HV2
709	G	A	-	-	A	-		12sRNA
750 *R	A	G	G	G	G	G		12sRNA
1438*R	A	G	G	G	G	G		12sRNA
1719	G	A	-	-	-	-		16sRNA
2706	A	G	-	G	-	G		16sRNA
2841	T	-	-	-	-	A		16sRNA
3010	G	-	-	-	A	A		16sRNA
3106-3107*E	C/del	del C	del C	del C	del C	del C		16sRNA
3394	T	-	-	-	-	C	Tyr → His	ND1 LHON
3423*E	G/T	T	T	T	T	T	Silent	ND1
4135	T	-	C	-	-	-	Tyr → His	ND1
4216	T	-	-	C	-	C	Tyr → His	ND1 LHON
4769*R	A	G	G	G	G	G	Silent	ND2
4985*E	G/A	A	A	A	A	A	Silent	ND2
5186	A	G	-	-	-	-	Silent	ND2
5228	C	-	-	G	-	-	Silent	ND2
5633	C	-	-	T	-	-		tRNA Ala
6221	T	C	-	-	-	-	Silent	COI
6371	C	T	-	-	-	-	Silent	COI

6791	A	G	-	-	-	-	Silent	COI
6849[h]	A	G(0.3A)[h]	-	-	-	-	Thr → Ala[h]	COI
7028	C	T	-	T	-	T	Silent	COI
7476	C	-	-	T	-	-		tRNA Ser
7645	T	-	C	-	-	-	Silent	COII
7861	T	-	C	-	-	-	Silent	COII
8020	G	-	-	-	-	A	Silent	COII
8448	T	-	C	-	-	-	Met → Thr	ATPase 8
8503	T	C	-	-	-	-	Silent	ATPase 8
8860*R	A	G	G	G	G	nd	Thr → Ala	ATPase6
9315	T	-	C	-	-	-	Phe → Leu	COIII
9559*E	G/C	C	C	C	C	C	Arg → Pro	COIII
10172	G	-	-	A	-	-	Silent	ND3
10398	A	-	-	G	-	G	Thr → Ala	ND3
11251	A	-	-	G	-	G	Silent	ND4
11287	T	-	-	-	-	C	Silent	ND4
11335*E	T/C	C	C	C	C	C	Silent	ND4
11719	G	A	-	A	-	A	Silent	ND4
11778	G	-	-	-	-	A	Arg → His	ND4 LHON
11878	T	C	-	-	-	-	Silent	ND4
12071[het]	T	-	-	C/T[het]	-	-	Phe→Leu[het]	ND4
12612	A	G	-	G	-	G	Silent	ND5
12705	C	T	-	-	-	-	Silent	ND5
13572	T	-	C	-	-	-	Silent	ND5

13702*E	G/C	C	C	C	C	C	Gly → Arg	ND5
13708	G	A	-	A	-	A	Ala → Thr	ND5 LHON
13759	G	-	A	-	-	-	Ala → Thr	ND5
13966	A	G	-	-	-	-	Thr → Ala	ND5
14199*E	G/T	T	T	T	T	T	Pro → Thr	ND6
14272*E	G/C	C	C	C	C	C	Phe → Leu	ND6
14365*E	G/C	C	C	C	C	C	Silent	ND6
14368*E	G/C	C	C	C	C	C	Phe → Leu	ND6
14470	T	C	-	-	-	-	Silent	ND6
14569	G	-	-	A	-	-	Silent	ND6
14766*E	T/C	T	C	T	C	T	Ile → Thr	ND6
14798	T	-	-	-	-	C	Phe → Leu	CYT B
15257	G	-	-	A	-	-	Asp→ Asn	CYT B LHON
15326*R	A	G	G	G	G	G	Thr → Ala	CYT B
15452	C	-	-	A	-	A	Leu → Ile	CYT B
15646	C	-	-	-	T	-	Silent	CYT B
15812	G	-	-	A	-	-	Val → Met	CYT B LHON
16069	C	-	-	T	-	T		HV1
16126	T	-	-	-	-	C		HV1

16183	A	C	-	-	-	-	HV1
16184-93	-	C (ins)	-	-	-	-	HV1
16189	T	C	-	-	-	-	HV1
16193	C	-	-	T	-	-	HV1
16223	C	T	-	-	-	-	HV1
16278	C	T	-	T	-	-	HV1
16292	C	-	-	-	-	T	HV1
16311	T	-	C	-	-	-	HV1
16357	T	-	-	-	C	-	HV1
16362	T	-	-	C	-	-	HV1
16519	T	C	C	-	C	nd	HV1

a : Numbers correspond to Cambridge Reference Sequence (Anderson *et al*., 1981).
b: Base found in 1981 (Anderson *et al*., 1981)/ Base found in 1999 (Andrews *et al*., 1999).
d: SRM 2392 (Levin *et al*., 1999).
e: SRM 2392-I (Levin *et al.,* 2003 in press).
- : Base pair same as in 1981 Cambridge Reference Sequence (Anderson *et al*., 1981).
h : Possible heteroplasmic site. This heteroplasmy seen in the mtDNA from the first CHR cell culture line is not seen in the mtDNA from the second CHR cell culture line. The second CHR cell culture line agrees with the CRS at np 6849. It is DNA from the second CHR cell culture line that is supplied in NIST SRM 2392.
*R: Rare polymorphisms in Cambridge Reference Sequence discovered by reanalysis of original placenta (Andrews *et al.*, 1999).
*E: Error in Cambridge Reference Sequence discovered by reanalysis of original placenta (Andrews *et al.*, 1999).
del: Deletion
ins: Insertion
het: Heteroplasmy found in HL-60 at np 12071.
HV1: Non-coding region found from 16024 and 16569.
HV2: Non-coding region found from 1 and 576.
CHR DNA: Sequence based on two amplifications and cycle sequencing procedures with DNA from the first cell culture line and at least one amplification and cycle sequencing procedure with DNA from the second cell culture line.
9947A DNA: Sequence based on two amplifications and cycle sequencing procedures.

HL-60 DNA: Sequence based on two amplifications and cycle sequencing procedures in both the forward and reverse directions for a total of 4 sequences
ATPase 6: ATP synthase 6
ATPase 8: ATP synthase 8
CYTB: Cytochrome B
COI: Cytochrome C Oxidase I
COII: Cytochrome C Oxidase II
COIII: Cytochrome C Oxidase III
ND1: NADH dehydrogenase 1
ND2: NADH dehydrogenase 2
ND3: NADH dehydrogenase 3
ND4: NADH dehydrogenase 4
ND5: NADH dehydrogenase 5
ND6: NADH dehydrogenase 6

Table 3. Universal Genetic Code and (Human mtDNA Differences)

5' End	2nd Position				3' End
	T	C	A	G	
T	Phe	Ser	Tyr	Cys	T
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	TER[1]	TER[1] (Trp)	A
	Leu	Ser	TER[1]	Trp	G
C	Leu	Pro	His	Arg	T
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile (INT)[2]	Thr	Asn	Ser	T
	Ile (INT)[2]	Thr	Asn	Ser	C
	Ile (Met) (INT)[2]	Thr	Lys	Arg (TER)[1]	A
	Met (INT)[2]	Thr	Lys	Arg (TER)[1]	G
G	Val	Ala	Asp	Gly	T
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

1. TER: Termination codon
2. INT: Initiation codon

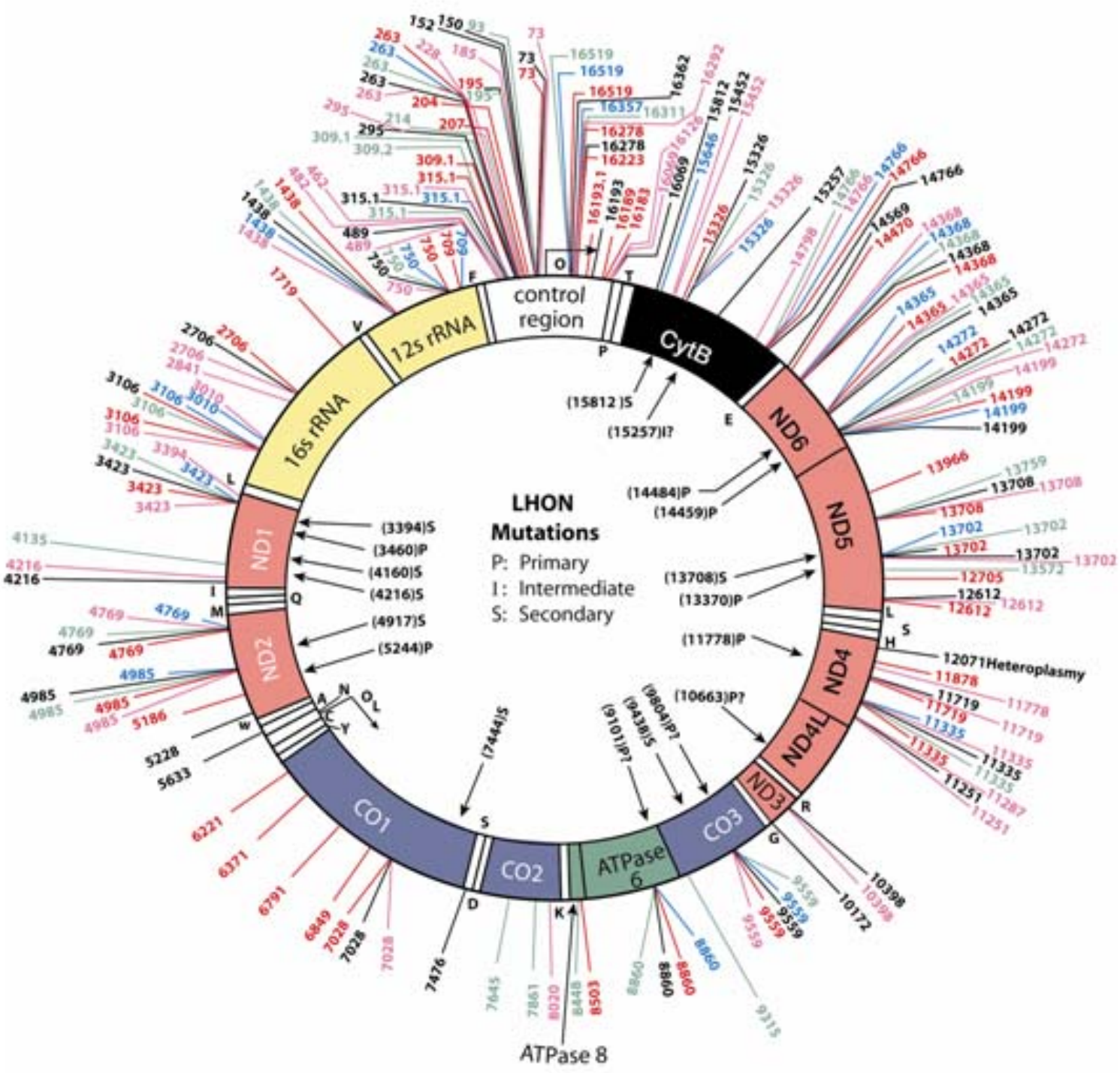

control region
12s rRNA
16s rRNA
CytB
ND6
ND5
ND4
ND4L
ND3
CO3
ATPase 6
CO2
CO1
ND2
ND1
ATPase 8
LHON Mutations
P: Primary
I: Intermediate
S: Secondary
(15812)S
(15257)I?
(14484)P
(14459)P
(13708)S
(13370)P
(11778)P
(10663)P?
(9804)P?
(9438)S
(9101)P?
(7444)S
(4917)S
(5244)P
(4160)S
(4216)S
(3460)P
(3394)S
12071Heteroplasmy

Figure 1. Schematic of human mtDNA showing its circular double-stranded DNA and all the differences from the Cambridge Reference Sequence (Anderson *et al.*, 1981) found in CHR (red), 9947A (green), HL-60 (black), GM03798 (blue), and GM10742A (purple) as numbers along the outside of the color-coded circle. Locations of the control region, rRNAs and genes (see legend to Table 2 for abbreviations) coded by human mtDNA are shown. The locations of the 22 tRNAs are noted by white areas in the circle and designated by their single letter amino acid code. Since a number of mutations found in GM10742A and HL-60 and one change in CHR have been associated with primary, intermediate or secondary mutations linked to the disease LHON, the position of these mutations plus other LHON mutations are shown on the inside of the circle (Wallace *et al.*, 1997). The question mark following the np of the LHON mutations indicates the assignment is not confirmed. One of the primary mutations that have been associated with LHON, G11778A, was found in GM10742A (Table 2) but not found in the other DNA templates examined in this research. (Modified from Levin *et al.*, 1999).

Appendix A

SRM 2392 Certificate

National Institute of Standards & Technology

Certificate of Analysis

Standard Reference Material® 2392

Mitochondrial DNA Sequencing (Human)

This Standard Reference Material (SRM) is intended to provide quality control when performing the polymerase chain reaction (PCR) and sequencing of human mitochondrial DNA (mtDNA) for forensic identifications, medical diagnosis, or mutation detection. It may also be used as a control when amplifying (PCR) and sequencing any DNA. This SRM can also be used for quality assurance when assigning values to in-house control materials. It is certified for the sequences of the entire human mtDNA (16 569 base pairs) from two lymphoblastoid cell culture lines (CHR and GM09947A) from apparently normal individuals, plus the cloned HV1 region of CHR containing a C-stretch which is difficult to sequence. The SRM is packaged in a single box containing three components: (1) extracted DNA from cell culture line CHR (tube contains 60 µL of DNA at a concentration of 1 ng/µL); (2) extracted DNA from cell culture line GM09947A (tube contains 60 µL of DNA at a concentration of 1 ng/µL); and (3) cloned DNA from the CHR HV1 region containing the C-stretch (tube contains 10 µL of DNA at a concentration of 100 ng/µL).

This SRM is composed of well-characterized extracted human DNA from CHR and GM09947A and cloned DNA from the HV1 region of CHR. Table 1 contains the certified sequence information of two entire mtDNA templates (CHR and GM09947A). Table 2 contains the reference sequences of 58 unique primer sets which were designed to amplify any portion or the entire human mtDNA. The sequence information of a third DNA template (GM03798) that was amplified and sequenced in its entirety three to four times at NIST is provided in reference 1. Although the extracted DNA from GM03798 is not provided, the cell culture line can be obtained from NIGMS Human Genetic Mutant Cell Repository, Coriell Institute for Medical Research, Camden, NJ.

Expiration of Certification: The certification of this SRM is valid until **31 May 2008**, provided the SRM is handled and stored in accordance with the instructions given in this certificate. This certification is nullified if the SRM is damaged, contaminated, or modified.

Maintenance of SRM Certification: NIST will monitor this SRM over the period of its certification. If substantive technical changes occur that affect the certification before the expiration of this certificate, NIST will notify the purchaser. Return of the attached registration card will facilitate notification.

The overall direction and coordination of the technical measurements leading to the certification were performed by B.C. Levin and D.J. Reeder of the NIST DNA Technologies Group, Biotechnology Division.

The analytical determination and technical measurements for the certification of this SRM were performed by B.C. Levin, H. Cheng, L.A. Tully, M.P. Jones, and D.J. Reeder of the NIST DNA Technologies Group, Biotechnology Division.

The support aspects involved in the preparation, certification, and issuance of this SRM were coordinated through the NIST Standard Reference Materials Program by J.C. Colbert and B.S. MacDonald of the NIST Measurement Services Division.

Vincent L. Viker, Chief
Biotechnology Division

Gaithersburg, MD 20899
Certificate Issue Date: 17 June 2003
See Certificate Revision History on Last Page

John Rumble, Jr., Chief
Measurement Services Division

NOTICE AND WARNING TO USER

Warning: SRM 2392 IS A HUMAN SOURCE MATERIAL. SINCE THERE IS NO CONSENSUS ON THE INFECTIOUS STATUS OF EXTRACTED DNA, HANDLE PRODUCT AS A BIOHAZARDOUS MATERIAL CAPABLE OF TRANSMITTING INFECTIOUS DISEASE.

Storage: Store frozen at a temperature of -20 °C. **DO NOT** use a self-defrosting freezer because periodic cycling of temperatures may cause shortened shelf life of this SRM.

INSTRUCTIONS FOR USE

It is recommended that once thawed, each SRM component should be used in its entirety. Repeated freezing and thawing is **NOT** recommended as this might shorten the shelf life of the SRM. If it is necessary to perform repeated analyses, thaw the SRM and divide the tube contents into aliquots that will be kept frozen until use. Thawing can be conducted at refrigerator temperatures, room temperature, or at 37 °C. Once thawed, the sample should be processed without delay.

SOURCE AND ANALYSIS[1]

Source of Material: CHR DNA, both extracted and cloned, was prepared in the NIST DNA Technologies Group, Biotechnology Division. DNA for GM09947A was prepared by Life Technologies, Inc., Gaithersburg, MD.

NIST Analysis: NIST extracted DNA from the CHR cell culture, PCR was used to amplify both the CHR DNA and GM09947A DNA with all 58 primer sets multiple times, and sequenced the PCR products with a Perkin-Elmer Applied Biosystems, Inc. (ABI) 373 automated sequencer or an ABI 310 sequencer. The cloned DNA was prepared at NIST as described in reference 1. The sequence of the CHR clone and of representative PCR products of the final CHR and GM09947A DNA included in SRM 2392 was reanalyzed to ensure sequence accuracy.

Interlaboratory Analysis: An interlaboratory evaluation of the amplification, sequencing, and data analysis of the CHR template was conducted by four laboratories, including NIST. These laboratories were: The Bode Technology Group, Inc., Sterling, VA; IIT Research Institute, Virginia Technology Center, Newington, VA; and Lark Technologies, Inc., Houston, TX. Description of the interlaboratory analyses is described in reference 1.

Description of Components: Three components are included in each unit; all components must be stored at -20 ºC.

- # 1 Extracted DNA from cell culture line CHR (tube contains 60 µL of DNA at a concentration of 1 ng/µL)
- # 2 Extracted DNA from cell culture line GM09947A (tube contains 60 µL of DNA at a concentration of 1 ng/µL)
- # 3 Cloned DNA from the CHR HV1 region containing the C-stretch (tube contains 10 µL of DNA at a concentration of 100 ng/µL)

NOTE: DNA concentrations given are nominal values and are not intended for use as concentration standards.

[1]Certain commercial equipment, instruments, materials, or companies are identified in this paper to specify the experimental procedure. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the materials or equipment identified are the best available for this purpose.

Table 1. Certified Human mtDNA Sequence Differences from the Cambridge Reference Sequence [2] Found in Two Templates at NIST

Primer Set	Amplified Region[a]	Length of Amplified Region	Comparison with Cambridge Reference Sequence (CRS)				Amino Acid
			CRS #	bp	Template CHR	Template GM09947A	Change
1 (HV2)	15 - 484	470	73 93 195 204 207 214 263 309.1 309.2 315.1	A A T T G A A	Start 39 G - C C A - G C(ins) - C(ins) End 436	Start 39 - G C - - G G C(ins) C(ins) C(ins) End 473	
2	361 - 921	561	709 750	G A	Start 429 A G End 891	Start 421 - G End 846	
3	756 - 1425	670	NONE		Start 778 End 1197	Start 778 End 1278	
4	873 - 1425	553	NONE		Start 931 End 1335	Start 928 End 1377	
5	1234 - 1769	536	1438 1719	A G	Start 1279 G A End 1738	Start 1275 G - End 1741	

Table 1. Continued

6	1587 - 2216	630	1719[d]	G	Start 1632 A End 2106	Start 1632 - End 2106	
7	1657 - 2216	560	1719[d]	G	Start 1691 A End 2170	Start 1686 - End 2173	
8	1993 - 2216	224	NONE		Start 2036 End 2213	Start 2018 End 2217	
9	2105 - 2660	556	NONE		Start 2157 End 2636	Start 2150 End 2586	
10	2417 - 3006	590	2706	A	Start 2465 G End 2920	Start 2458 - End 2956	
11	2834 - 3557	724	3106/3107	C	Start 2861 Del End 3350	Start 2869 Del End 3373	
12	2972 - 3557	586	3106/3107[d] 3423	C G	Start 2999 Del E End 3422	Start 2999 Del T End 3460	Silent
13	3234 - 3557	324	3423[d]	G	Start 3265 T End 3548	Start 3258 T End 3545	Silent[d]
14	3441 - 3940	500	NONE		Start 3487 End 3916	Start 3491 End 3920	
15	3635 - 4162	528	NONE		Start 3667 End 4126	Start 3662 End 4061	
16	3931 - 4728	798	4135	T	Start 3964 - End 4399	Start 3968 C End 4427	Try➔His

Table 1. Continued

17	4183 - 4728	546	NONE		Start 4208 End 4657	Start 4249 End 4657	
18	4392 - 4982	591	4769	A	Start 4449 G End 4860	Start 4453 G End 4935	Silent
19	4447 - 4982	536	4769[d]	A	Start 4492 G End 4958	Start 4492 G End 4921	Silent[d]
20	4797 - 5553	757	4985 5186	G A	Start 4838 A G End 5327	Start 4845 A - End 5324	Silent Silent
21	4976 - 5553	578	5186[d]	A	Start 5000 G End 5516	Start 5007 - End 5521	Silent[d]
22	5318 - 5882	565	NONE		Start 5361 End 5754	Start 5360 End 5758	
23	5700 - 6262	563	NONE		Start 5741 End 6149	Start 5744 End 6163	
24	5999 - 6526	528	6221 6371	T C	Start 6043 C T End 6442	Start 6058 - - End 6503	Silent Silent
25	6242 - 6526	285	6371[d]	C	Start 6271 T End 6520	Start 6302 - End 6520	Silent[d]
26	6426 - 7030	605	6791 6849*	A A	Start 6451 G G(0.3A)[h]* End 6916	Start 6474 - - End 6930	Silent Thr➔Ala*

Table 1. Continued

27	6744 - 7255	512	6849[d]* 7028	A C	Start 6775 G(0.3A)[h]* T End 7215	Start 6782 - - End 7221	Thr➔Ala[d]* Silent
28	7075 - 7792	718	NONE		Start 7123 End 7602	Start 7123 End 7601	
29	7215 - 7792	578	7645	T	Start 7263 - End 7722	Start 7280 C End 7769	Silent
30	7645 - 8215	571	7861	T	Start 7671 - End 8149	Start 7666 C End 8155	Silent
31	7901 - 8311	411	NONE		Start 7960 End 8289	Start 7959 End 8288	
32	8164 - 8669	506	8448 8503	T T	Start 8211 - C End 8646	Start 8212 C - End 8641	Met➔Thr Silent
33	8539 - 9059	521	8860	A	Start 8581 G End 9019	Start 8582 G End 8999	Thr➔Ala
34	8903 - 9403	501	9315	T	Start 8947 - End 9380	Start 8944 C End 9381	Phe➔Leu
35	9309 - 9848	540	9559	G	Start 9334 C End 9823	Start 9333 C End 9827	Arg➔Pro
36	9449 - 9995	547	9559[d]	G	Start 9476 C End 9964	Start 9485 C End 9940	Arg➔Pro[d]

Table 1. Continued

37	9754 - 10275	522	NONE		Start 9777 End 10225	Start 9781 End 10251	
38	10127 - 10556	430	NONE		Start 10168 End 10534	Start 10166 End 10536	
39	10386 - 11166	781	NONE		Start 10410 End 10899	Start 10416 End 10916	
40	10704 - 11267	564	NONE		Start 10734 End 11223	Start 10742 End 11197	
41	11001 - 11600	600	11335	T	Start 11026 C End 11461	Start 11040 C End 11517	Silent
42	11403 - 11927	525	11719	G	Start 11428 A End 11795	Start 11432 - End 11853	Silent
43	11760 - 12189	430	11878	T	Start 11784 C End 12159	Start 11802 - End 12164	Silent
44	11901 - 12876	976	NONE		Start 11926 End 12404	Start 11926 End 12443	
45	12357 - 12876	520	12612 12705	A C	Start 12404 G T End 12769	Start 12391 - - End 12849	Silent Silent
46	12601 - 13123	523	12705[d]	C	Start 12627 T End 13102	Start 12645 - End 13045	Silent[d]
47	12793 - 13343	551	NONE		Start 12817 End 13295	Start 12807 End 13307	

Table 1. Continued

48	13188 - 13611	424	13572	T	Start 13238 - End 13587	Start 13238 C End 13593	Silent
49	13518 - 13935	418	13572[d] 13702 13708 13759	T G G G	Start 13541 - C A - End 13910	Start 13541 C C - A End 13921	Silent[d] Gly➔Arg Ala➔Thr Ala➔Thr
50	13715 - 14118	404	13966	A	Start 13775 G End 14094	Start 13760 - End 14110	Thr➔Ala
51	13899 - 14388	490	13966[d] 14199 14272 14365	A G G G	Start 13926 G T C C End 14369	Start 13927 - T C C End 14374	Thr➔Ala[d] Pro➔Thr Phe➔Leu Silent
52	14189 - 14926	738	14272[d] 14365[d] 14368 14470 14766	G G G T T	Start 14216 C C C C E End 14699	Start 14216 C C C - C End 14806	Phe➔Leu[d] Silent[d] Phe➔Leu Silent Ile➔Thr
53	14470 - 14996	527	14766[d]	T	Start 14502 - End 14957	Start 14513 C End 14972	Ile➔hr[d]
54	14909 - 15396	488	15326	A	Start 14941 G End 15380	Start 14933 G End 15373	Thr➔Ala

Table 1. Continued

55	15260 - 15774	515	15326[d]	A	Start 15305 G End 15754	Start 15293 G End15950	Thr➔Ala[d]
56	15574 - 16084	511	NONE		Start 15637 End 16056	Start 15599 End 16058	
57 (HV1)	15971 - 16451	481	16183 16189 16311	A T T	Start 16014 C C E End 16193	Start 16011 - - C End 16430	
58	16097 - 336	809	16183[d] 16189[d] 16311[d] 16519	A T T T	Start 16125 C C E E End 16193	Start 16130 - - C C End 59	
-21M13[c] cloned DNA	16133 - 40	477	16183[d] 16189[d] 16193.1 16223 16278 16519[d]	A T C C T	Start 16131 C C C(ins) T T C End 40	ND	

Table 1. Continued

[a]	Numbers correspond to Cambridge Reference Sequence [2].
E	Base pair change came after the readable sequence.
-	Base pair same as in Cambridge Reference Sequence [2].
h*	Possible heteroplasmic site
*	This heteroplasmy seen in the first CHR cell culture line was not seen with the second CHR cell culture line. It is the second CHR cell culture line that is supplied in NIST SRM 2392.
[c]	This primer is used for sequencing the cloned DNA of the HV1 region.
[d]	Change also seen in previous primer set.
Start	Start of readable sequence.
End	End of readable sequence.
CHR cells	Sequence based on two amplifications and cycle sequencing procedures in first cell culture line and at least one amplification and cycle sequencing procedure with the second cell culture line.
GM09947A cells	Sequence based on two amplifications and cycle sequencing procedures.
Ins	Insertion
Del	Deletion
ND	Not done

Table 2. Reference Sequences for Primer Sets Used for PCR Amplification of Human mtDNA

PRIMER SET NUMBER	PRIMER SEQUENCE	
1 (HV2)	F15	CACCCTATTAACCACTCACG
	R484	TGAGATTAGTAGTATGGGAG
2	F361	ACAAAGAACCCTAACACCAGC
	R921	ACTTGGGTTAATCGTGTGACC
3	F756	CATCAAGCACGCAGCAATG
	R1425	AATCCACCTTCGACCCTTAAG
4	F873	GGTTGGTCAATTTCGTGCCAG
	R1425	AATCCACCTTCGACCCTTAAG
5	F1234	CTCACCACCTCTTGCTCAGC
	R1769	GCCAGGTTTCAATTTCTATCG
6	F1587	TGCACTTGGACGAACCAGAG
	R2216	TGTTGAGCTTGAACGCTTTC
7	F1657	CTTGACCGCTCTGAGCTAAAC
	R2216	TGTTGAGCTTGAACGCTTTC
8	F1993	AAACCTACCGAGCCTGGTG
	R2216	TGTTGAGCTTGAACGCTTTC
9	F2105	GAGGAACAGCTCTTTGGACAC
	R2660	AGAGACAGCTGAACCCTCGTG
10	F2417	CACTGTCAACCCAACACAGG
	R3006	ATGTCCTGATCCAACATCGAG
11	F2834	CCCAACCTCCGAGCAGTACATG
	R3557	AGAAGAGCGATGGTGAGAGC
12	F2972	ATAGGGTTTACGACCTCGATG
	R3557	AGAAGAGCGATGGTGAGAGC
13	F3234	AGATGGCAGAGCCCGGTAATC
	R3557	AGAAGAGCGATGGTGAGAGC
14	F3441	ACTACAACCCTTCGCTGACG
	R3940	TGAAGCCTGAGACTAGTTCGG
15	F3635	GCCTAGCCGTTTACTCAATCC
	R4162	TGAGTTGGTCGTAGCGGAATC
16	F3931	TCAGGCTTCAACATCGAATACG
	R4728	TTATGGTTCATTGTCCGGAGAG

Table 2. Continued

17	F4183 R4728	TTTCTACCACTCACCCTAGCATTAC TTATGGTTCATTGTCCGGAGAG
18	F4392 R4983	CCCATCCTAAAGTAAGGTCAGC GGTTTAATCCACCTCAACTGCC
19	F4447 R4982	TTGGTTATACCCTTCCCGTAC GTTTAATCCACCTCAACTGCC
20	F4797 R5553	CCCTTTCACTTCTGAGTCCCAG AGGGCTTTGAAGGCTCTTG
21	F4976 R5553	ATTAAACCAGACCCAGCTACG AGGGCTTTGAAGGCTCTTG
22	F5318 R5882	CACCATCACCCTCCTTAACC GCTGAGTGAAGCATTGGACTG
23	F5700 R6262	TAAGCACCCTAATCAACTGGC GCCTCCACTATAGCAGATGCG
24	F5999 R6526	TCTAAGCCTCCTTATTCGAGC ATAGTGATGCCAGCAGCTAGG
25	F6242 R6526	CGCATCTGCTATAGTGGAGG ATAGTGATGCCAGCAGCTAGG
26	F6426 R7030	GCCATAACCCAATACCAAACG TGGGCTACAACGTAGTACGTG
27	F6744 R7255	GGCTTCCTAGGGTTTATCGTG TTTCATGTGGTGTATGCATCG
28	F7075 R7792	GAGGCTTCATTCACTGATTTCC GGGCAGGATAGTTCAGACGG
29	F7215 R7792	CGACGTTACTCGGACTACCC GGGCAGGATAGTTCAGACGG
30	F7645 R8215	TATCACCTTTCATGATCACGC GACGATGGGCATGAAACTG
31	F7901 R8311	TGAACCTACGAGTACACCGACTAC AAGTTAGCTTTACAGTGGGCTCTAG
32	F8164 R8669	CGGTCAATGCTCTGAAATCTGTG CATTGTTGGGTGGTGATTAGTCG
33	F8539 R9059	CTGTTCGCTTCATTCATTGCC GTGGCGCTTCCAATTAGGTG
34	F8903 R9403	CCCACTTCTTACCACAAGGC GTGCTTTCTCGTGTTACATCG

Table 2. Continued

35	F9309 R9848	TTTCACTTCCACTCCATAACGC GAAAGTTGAGCCAATAATGACG
36	F9449 R9995	CGGGATAATCCTATTTATTACCTCAG AGAGTAAGACCCTCATCAATAGATGG
37	F9754 R10275	AGTCTCCCTTCACCATTTCCG AAAGGAGGGCAATTTCTAGATC
38	F10127 R10556	ACTACCACAACTCAACGGCTAC GGAGGATATGAGGTGTGAGCG
39	F10386 R11166	GGATTAGACTGAACCGAATTGG CATCGGGTGATGATAGCCAAG
40	F10704 R11267	GTCTCAATCTCCAACACATATGG TGTTGTGAGTGTAAATTAGTGCG
41	F11001 R11600	AACGCCACTTATCCAGTGAACC CTGTTTGTCGTAGGCAGATGG
42	F11403 R11927	GACTCCCTAAAGCCCATGTCG TTGATCAGGAGAACGTGGTTAC
43	F11760 R12189	ACGAACGCACTCACAGTCG AAGCCTCTGTTGTCAGATTCAC
44	F11901 R12876	TGCTAGTAACCACGTTCTGGTG GATATCGCCGATACGGTTG
45	F12357 R12876	AACCACCCTAACCCTGACTTCC GATATCGCCGATACGGTTG
46	F12601 R13123	TTCATCCCTGTAGCATTGTTCG AGCGGATGAGTAAGAAGATTCC
47	F12793 R13343	TTGCTCATCAGTTGATGATACG TTGAAGAAGGCGTGGGTACAG
48	F13188 R13611	CACTCTGTTCGCAGCAGTATG TCGAGTGCTATAGGCGCTTGTC
49	F13518 R13935	CATCATCGAAACCGCAAAC TGTGATGCTAGGGTAGAATCCG
50	F13715 R14118	GAAGCCTATTCGCAGGATTTC TGGGAAGAAGAAAGAGAGGAAG
51[a]	F13899 R14388 R14388	TTTCTCCAACATACTCGGATTC TTAGCGATGGAGGTAGGATTCG (Old Primer) TTAGCGATGGAGGTAGGATT**G**G (New Primer)
52	F14189 R14926	ACAAACAATGGTCAACCAGTAAC TGAGGCGTCTGGTGAGTAGTGC

Table 2. Continued

53	F14470 R14996	TCCAAAGACAACCATCATTCC CGTGAAGGTAGCGGATGATTC
54	F14909 R15396	TACTCACCAGACGCCTCAACCG TTATCGGAATGGGAGGTGATTC
55	F15260 R15774	AGTCCCACCCTCACACGATTC ACTGGTTGTCCTCCGATTCAGG
56	F15574 R16084	CGCCTACACAATTCTCCGATC CGGTTGTTGATGGGTGAGTC
57 (HV1)	F15971 R16451	TTAACTCCACCATTAGCACC GCGAGGAGAGTAGCACTCTTG
58	F16097 R336	TACATTACTGCCAGCCACCATG TTAAGTGCTGTGGCCAGAAG
-21M13	F	TGTAAAACGACGGCCAGT

[a] The reverse primer of set 51 has been changed. One should use the new primer.

REFERENCES

[1] Levin, B.C.; Cheng, H.; Reeder, D.J.; *Human Mitochondrial DNA Standard Reference Material for Quality Control in Forensic Identification, Medical Diagnosis, and Mutation Detection*; Genomics, Vol. 55, pp. 135-146 (1999).

[2] Anderson, S.; Bankier, A.T.; Barrell, B.G.; deBrujin, M.H.L.; Coulson, A.R.; Drouin, J.; Eperon, I.C.; Nierlich, D.P.; Roe, B.A.; Sanger, F.; Schreier, P.H.; Smith, A.J.H.; Staden, R.; Young, I.G.; *Sequence and Organization of the Human Mitochondrial Genome*; Nature, Vol. 290, pp. 457-465 (1981).

Certificate Revision History: 17 June 2003 (This revision reports an extension in expiration date and replacement of reverse primer 51); 29 December 1999 (Original certificate date).

Users of this SRM should ensure that the certificate in their possession is current. This can be accomplished by contacting the SRM Program at: telephone (301) 975-6776; fax (301) 926-4751; e-mail srminfo@nist.gov; or via the Internet http://www.nist.gov/srm.

Appendix B

SRM 2392-I Certificate

National Institute of Standards & Technology

Certificate of Analysis

Standard Reference Material® 2392-I

Mitochondrial DNA Sequencing
(Human HL-60 DNA)

This Standard Reference Material (SRM) is intended to provide quality control when performing the polymerase chain reaction (PCR) and sequencing of human mitochondrial DNA (mtDNA) for forensic identification, medical diagnosis, or mutation detection. It may also serve as a control when amplifying (PCR) and sequencing any DNA. This SRM can also be used for quality assurance when assigning values to in-house control materials. It is certified for the sequences of the entire human mtDNA (16 569 base pairs) from a promyelocytic cell line (HL-60) prepared from the peripheral blood leukocytes from an individual with acute promyelocytic leukemia. A unit of SRM 2392-I consists of 65 µL of extracted DNA from cell culture line HL-60 at a nominal concentration of 1.4 ng/µL, which is contained in a vial packaged in a protective plastic box.

Certified Values: The certified sequence information of extracted human DNA from HL-60 is provided in Table 1. Also provided in Table 1 is the certified sequence information for two additional entire mtDNA templates, CHR and GM09947A, which are provided in SRM 2392. SRM 2392-I only contains the HL-60 template. Table 2 contains the sequences of 58 unique primer sets that were designed to amplify any portion or the entire human mtDNA [1].

Supplemental Information: The sequence information of an additional two DNA templates, (GM03798 [1] and GM10742A [2]), that were amplified and sequenced in their entirety multiple times at NIST are provided in references 1 or 2. Although the extracted DNA from GM03798 and GM10742A are not provided, the cell cultures can be obtained from NIGMS Human Genetic Mutant Cell Repository, Coriell Institute for Medical Research, Camden, NJ. A schematic of the differences from the Cambridge Reference Sequence [3] found in the mtDNA from all five templates is shown in Figure 1.

Expiration of Certification: The certification of this SRM is valid until **31 March 2008**, provided the SRM is handled and stored in accordance with the instructions given in this certificate. This certification is nullified if the SRM is damaged, contaminated, or modified.

Maintenance of SRM Certification: NIST will monitor this SRM over the period of its certification. If substantive technical changes occur that affect the certification before the expiration of this certificate, NIST will notify the purchaser. Return of the attached registration card will facilitate notification.

The analytical determination, technical measurements and analysis of data for the certification of this SRM were performed by D.K. Hancock, K.L. Richie, K.A. Holland (on sabbatical from Gettysburg College, Gettysburg, PA), and B.C. Levin of the NIST DNA Technologies Group, Biotechnology Division.

The overall direction and coordination of the technical measurements leading to the certification was performed by B.C. Levin of the NIST DNA Technologies Group, Biotechnology Division.

The support aspects involved in the issuance of this SRM were coordinated through the NIST Standard Reference Materials Program by B.S. MacDonald of the NIST Measurement Services Division.

Vincent L. Vilker, Chief
Biotechnology Division

Gaithersburg, MD 20899
Certificate Issue Date: 13 June 2003

John Rumble, Jr., Chief
Measurement Services Division

Support for the preparation and certification of this SRM was provided by the National Institute of Justice through the NIST Office of Law Enforcement Standards.

NOTICE AND WARNINGS TO USER

Warning: SRM 2392-I IS A HUMAN SOURCE MATERIAL. SINCE THERE IS NO CONSENSUS ON THE INFECTIOUS STATUS OF EXTRACTED DNA, HANDLE PRODUCT AS A BIOHAZARDOUS MATERIAL CAPABLE OF TRANSMITTING INFECTIOUS DISEASE.

Permissions: The research to use HL-60 DNA in SRM 2392-I was deemed exempt from the policy of Part 27 of Title 15 of the Code of Federal Regulations by the NIST Institutional Review Board and the Director of the Chemical Science and Technology Laboratory. This work fit into the exemption category described in 15 CFR 27.101(b)(4) which states: "Research, involving the collection or study of existing data, documents, pathological specimens, or diagnostic specimens, if, these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects."

ATCC also waived condition 3(c) in their Material Transfer Agreement which states that the "purchaser shall not sell, lend, distribute or otherwise transfer the material or replicates to any others" for the use of HL-60 in the NIST mitochondrial DNA SRM. They stated that, in their view, "as a government agency, NIST will not be providing this material as a commercial product despite the collection of fees for the SRM."

Storage: Store frozen at a temperature of –20 ^{o}C. **DO NOT** use a self-defrosting freezer because of periodic cycling of temperatures may shorten the shelf life of this SRM.

INSTRUCTIONS FOR USE

It is recommended that once thawed, each SRM component should be used in its entirety. Repeated freezing and thawing is **NOT** recommended as this might shorten the shelf life of the SRM. If it is necessary to perform repeated analyses, thaw the SRM and divide the tube contents into aliquots that will be kept frozen until use. Thawing can be conducted at refrigerator temperatures, room temperature, or at 37 ^{o}C. Once thawed, the sample should be processed without delay. DNA concentrations given are nominal values and are **NOT** intended for use as concentration standards.

SOURCE AND ANALYSIS[1]

Source of Material: DNA from HL-60 was prepared by the Professional Services Department of the American Type Culture Collection (ATCC), Manassas, VA. This material was subsequently vialed at NIST into 65 µL portions (nominal DNA concentration of 1.4 ng/µL) and labeled SRM 2392-I Component D (Components A, B, and C are available in SRM 2392).

NIST Analysis: PCR was used to amplify the HL-60 DNA in its entirety multiple times using all 58 primer sets. The PCR products were sequenced with an Applied Biosystems, Inc. 310 automated sequencer. The sequences of representative PCR products of the final HL-60 DNA included in SRM 2392-I were reanalyzed to ensure sequence accuracy.

Interlaboratory Analyses: An interlaboratory evaluation of the amplification, sequencing and data analysis of the HL-60 template was conducted by four laboratories, including NIST. These laboratories were: The Armed Forces DNA Identification Laboratory (AFDIL), Rockville, MD; Federal Bureau of Investigation Laboratory (FBI), Quantico, VA; and The Georgia Bureau of Investigation (GBI), Decatur, GA. The sequences obtained by all of the laboratories were identical. Description of the interlaboratory analysis of HL-60 is described in reference 2.

[1]Certain commercial equipment, instruments, or materials are identified in this certificate in order to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Table 1. Certified Human mtDNA Sequence Differences from the Cambridge Reference Sequence (CRS) [3,4] Found in the Two Templates (CHR and GM09947A) in NIST SRM 2392 and One Template (HL-60) in NIST SRM 2392-I

Comparison with the Cambridge Reference Sequence (CRS)						
CRS						
# [a]	Base[b] 1981/ 1999	Template CHR[d]	Template 9947A[d]	Template HL-60[e]	Amino acid change	Region
73	A	G	-	G		HV2
93	A	-	G	-		HV2
150	C	-	-	T		HV2
152	T	-	-	C		HV2
195	T	C	C	-		HV2
204	T	C	-	-		HV2
207	G	A	-	-		HV2
214	A	-	G	-		HV2
263*R	A	G	G	G		HV2
295	C	-	-	T		HV2
303-309	-	C (ins)	CC (ins)	-		HV2
311-315*R	-	C (ins)	C (ins)	C (ins)		HV2
489	T	-	-	C		HV2
709	G	A	-	-		12sRNA
750 *R	A	G	G	G		12sRNA
1438*R	A	G	G	G		12sRNA
1719	G	A	-	-		16sRNA
2706	A	G	-	G		16sRNA
3106-3107*E	CC/del	del C	del C	del C		16sRNA
3423*E	G/T	T	T	T	Silent	ND1
4135	T	-	C	-	Tyr → His	ND1
4216	T	-	-	C	Tyr → His	ND1 LHON
4769*R	A	G	G	G	Silent	ND2
4985*E	G/A	A	A	A	Silent	ND2
5186	A	G	-	-	Silent	ND2
5228	C	-	-	G	Silent	ND2
5633	C	-	-	T		tRNA Ala
6221	T	C	-	-	Silent	COI
6371	C	T	-	-	Silent	COI

Comparison with the Cambridge Reference Sequence (CRS)						
CRS						
# [a]	Base[b] 1981/ 1999	Template CHR[d]	Template 9947A[d]	Template HL-60[e]	Amino acid change	Region
6791	A	G	-	-	Silent	COI
6849[h]	A	G(0.3A)[h]	-	-	Thr → Ala[h]	COI
7028	C	T	-	T	Silent	COI
7476	C	-	-	T		tRNA Ser
7645	T	-	C	-	Silent	COII
7861	T	-	C	-	Silent	COII
8448	T	-	C	-	Met → Thr	ATPase 8
8503	T	C	-	-	Silent	ATPase 8
8860*R	A	G	G	G	Thr → Ala	ATPase 6
9315	T	-	C	-	Phe → Leu	COIII
9559*E	G/C	C	C	C	Arg → Pro	COIII
10172	G	-	-	A	Silent	ND3
10398	A	-	-	G	Thr → Ala	ND3
11251	A	-	-	G	Silent	ND4
11335*E	T/C	C	C	C	Silent	ND4
11719	G	A	-	A	Silent	ND4
11878	T	C	-	-	Silent	ND4
12071[het]	T	-	-	C/T[het]	Phe→Leu[het]	ND4
12612	A	G	-	G	Silent	ND5
12705	C	T	-	-	Silent	ND5
13572	T	-	C	-	Silent	ND5
13702*E	G/C	C	C	C	Gly → Arg	ND5
13708	G	A	-	A	Ala → Thr	ND5 LHON
13759	G	-	A	-	Ala → Thr	ND5
13966	A	G	-	-	Thr → Ala	ND5
14199*E	G/T	T	T	T	Pro → Thr	ND6
14272*E	G/C	C	C	C	Phe → Leu	ND6
14365*E	G/C	C	C	C	Silent	ND6
14368*E	G/C	C	C	C	Phe → Leu	ND6
14470	T	C	-	-	Silent	ND6
14569	G	-	-	A	Silent	ND6
14766*E	T/C	T	C	T	Ile → Thr	ND6
15257	G	-	-	A	Asp→ Asn	CYT B LHON

Comparison with the Cambridge Reference Sequence (CRS)						
CRS						
# [a]	Base[b] 1981/ 1999	Template CHR[d]	Template 9947A[d]	Template HL-60[e]	Amino acid change	Region
15326*R	A	G	G	G	Thr → Ala	CYT B
15452	C	-	-	A	Leu → Ile	CYT B
15812	G	-	-	A	Val → Met	CYT B LHON
16069	C	-	-	T		HV1
16183	A	C	-	-		HV1
16184-93	-	C (ins)	-	-		HV1
16189	T	C	-	-		HV1
16193	C	-	-	T		HV1
16223	C	T	-	-		HV1
16278	C	T	-	T		HV1
16311	T	-	C	-		HV1
16362	T	-	-	C		HV1
16519	T	C	C	-		HV1

[a] Numbers correspond to Cambridge Reference Sequence [3]
[b] Base found in 1981[3] Base found in 1999 [4]
[d] SRM 2392, reference 1
[e] SRM 2392-I, reference 2
- Base pair same as in 1981 Cambridge Reference Sequence [3]
[h] Possible heteroplasmic site. This heteroplasmy seen in the mtDNA from the first CHR cell culture line is not seen in the mtDNA from the second CHR cell culture line. The second CHR cell culture line agrees with the CRS at np 6849. It is DNA from the second CHR cell culture line that is supplied in NIST SRM 2392.
*R Rare polymorphisms in Cambridge Reference Sequence discovered by reanalysis of original placenta by Andrews et al., 1999 [4].
*E Error in Cambridge Reference Sequence discovered by reanalysis of original placenta by Andrews et al., 1999 [4].
del Deletion
ins Insertion
het Heteroplasmy found in HL-60 at np 12071
HV1 Non-coding region found from 16024 and 16569
HV2 Non-coding region found from 1 and 576
CHR DNA Sequence based on two amplifications and cycle sequencing procedures with DNA from the first cell culture line and at least one amplification and cycle sequencing procedure with DNA from the second cell culture line.
GM09947A DNA Sequence based on two amplifications and cycle sequencing procedures.
HL-60 DNA Sequence based on two amplifications and cycle sequencing procedures in both the forward and reverse directions for a total of 4 sequences.
ATPase 6 ATP synthase 6
ATPase 8 ATP synthase 8
CYTB Cytochrome B
COI Cytochrome C Oxidase I
COII Cytochrome C Oxidase II
COIII Cytochrome C Oxidase III
LHON Leber Hereditary Optic Neuropathy
ND1 NADH dehydrogenase 1
ND2 NADH dehydrogenase 2
ND3 NADH dehydrogenase 3
ND4 NADH dehydrogenase 4
ND5 NADH dehydrogenase 5
ND6 NADH dehydrogenase 6

Table 2. Reference Sequences for Primer Sets Used for PCR Amplification of Human mtDNA

Primer Set Number		Primer Sequence
1(HV2)	F15	CACCCTATTAACCACTCACG
	R484	TGAGATTAGTAGTATGGGAG
2	F361	ACAAAGAACCCTAACACCAGC
	R921	ACTTGGGTTAATCGTGTGACC
3	F756	CATCAAGCACGCAGCAATG
	R1425	AATCCACCTTCGACCCTTAAG
4	F873	GGTTGGTCAATTTCGTGCCAG
	R1425	AATCCACCTTCGACCCTTAAG
5	F1234	CTCACCACCTCTTGCTCAGC
	R1769	GCCAGGTTTCAATTTCTATCG
6	F1587	TGCACTTGGACGAACCAGAG
	R2216	TGTTGAGCTTGAACGCTTTC
7	F1657	CTTGACCGCTCTGAGCTAAAC
	R2216	TGTTGAGCTTGAACGCTTTC
8	F1993	AAACCTACCGAGCCTGGTG
	R2216	TGTTGAGCTTGAACGCTTTC
9	F2105	GAGGAACAGCTCTTTGGACAC
	R2660	AGAGACAGCTGAACCCTCGTG
10	F2417	CACTGTCAACCCAACACAGG
	R3006	ATGTCCTGATCCAACATCGAG
11	F2834	CCCAACCTCCGAGCAGTACATG
	R3557	AGAAGAGCGATGGTGAGAGC
12	F2972	ATAGGGTTTACGACCTCGATG
	R3557	AGAAGAGCGATGGTGAGAGC
13	F3234	AGATGGCAGAGCCCGGTAATC
	R3557	AGAAGAGCGATGGTGAGAGC
14	F3441	ACTACAACCCTTCGCTGACG
	R3940	TGAAGCCTGAGACTAGTTCGG
15	F3635	GCCTAGCCGTTTACTCAATCC
	R4162	TGAGTTGGTCGTAGCGGAATC
16	F3931	TCAGGCTTCAACATCGAATACG
	R4728	TTATGGTTCATTGTCCGGAGAG
17	F4183	TTTCTACCACTCACCCTAGCATTAC
	R4728	TTATGGTTCATTGTCCGGAGAG
18	F4392	CCCATCCTAAAGTAAGGTCAGC
	R4983	GGTTTAATCCACCTCAACTGCC
19	F4447	TTGGTTATACCCTTCCCGTAC
	R4982	GTTTAATCCACCTCAACTGCC
20	F4797	CCCTTTCACTTCTGAGTCCCAG
	R5553	AGGGCTTTGAAGGCTCTTG
21	F4976	ATTAAACCAGACCCAGCTACG
	R5553	AGGGCTTTGAAGGCTCTTG
22	F5318	CACCATCACCCTCCTTAACC
	R5882	GCTGAGTGAAGCATTGGACTG
23	F5700	TAAGCACCCTAATCAACTGGC
	R6262	GCCTCCACTATAGCAGATGCG
24	F5999	TCTAAGCCTCCTTATTCGAGC
	R6526	ATAGTGATGCCAGCAGCTAGG
25	F6242	CGCATCTGCTATAGTGGAGG
	R6526	ATAGTGATGCCAGCAGCTAGG
26	F6426	GCCATAACCCAATACCAAACG
	R7030	TGGGCTACAACGTAGTACGTG
27	F6744	GGCTTCCTAGGGTTTATCGTG
	R7255	TTTCATGTGGTGTATGCATCG

Primer Set Number		Primer Sequence
28	F7075	GAGGCTTCATTCACTGATTTCC
	R7792	GGGCAGGATAGTTCAGACGG
29	F7215	CGACGTTACTCGGACTACCC
	R7792	GGGCAGGATAGTTCAGACGG
30	F7645	TATCACCTTTCATGATCACGC
	R8215	GACGATGGGCATGAAACTG
31	F7901	TGAACCTACGAGTACACCGACTAC
	R8311	AAGTTAGCTTTACAGTGGGCTCTAG
32	F8164	CGGTCAATGCTCTGAAATCTGTG
	R8669	CATTGTTGGGTGGTGATTAGTCG
33	F8539	CTGTTCGCTTCATTCATTGCC
	R9059	GTGGCGCTTCCAATTAGGTG
34	F8903	CCCACTTCTTACCACAAGGC
	R9403	GTGCTTTCTCGTGTTACATCG
35	F9309	TTTCACTTCCACTCCATAACGC
	R9848	GAAAGTTGAGCCAATAATGACG
36	F9449	CGGGATAATCCTATTTATTACCTCAG
	R9995	AGAGTAAGACCCTCATCAATAGATGG
37	F9754	AGTCTCCCTTCACCATTTCCG
	R10275	AAAGGAGGGCAATTTCTAGATC
38	F10127	ACTACCACAACTCAACGGCTAC
	R10556	GGAGGATATGAGGTGTGAGCG
39	F10386	GGATTAGACTGAACCGAATTGG
	R11166	CATCGGGTGATGATAGCCAAG
40	F10704	GTCTCAATCTCCAACACATATGG
	R11267	TGTTGTGAGTGTAAATTAGTGCG
41	F11001	AACGCCACTTATCCAGTGAACC
	R11600	CTGTTTGTCGTAGGCAGATGG
42	F11403	GACTCCCTAAAGCCCATGTCG
	R11927	TTGATCAGGAGAACGTGGTTAC
43	F11760	ACGAACGCACTCACAGTCG
	R12189	AAGCCTCTGTTGTCAGATTCAC
44	F11901	TGCTAGTAACCACGTTCTGGTG
	R12876	GATATCGCCGATACGGTTG
45	F12357	AACCACCCTAACCCTGACTTCC
	R12876	GATATCGCCGATACGGTTG
46	F12601	TTCATCCCTGTAGCATTGTTCG
	R13123	AGCGGATGAGTAAGAAGATTCC
47	F12793	TTGCTCATCAGTTGATGATACG
	R13343	TTGAAGAAGGCGTGGGTACAG
48	F13188	CACTCTGTTCGCAGCAGTATG
	R13611	TCGAGTGCTATAGGCGCTTGTC
49	F13518	CATCATCGAAACCGCAAAC
	R13935	TGTGATGCTAGGGTAGAATCCG
50	F13715	GAAGCCTATTCGCAGGATTTC
	R14118	TGGGAAGAAGAAAGAGAGGAAG
51	F13899	TTTCTCCAACATACTCGGATTC
	R14388	TTAGCGATGGAGGTAGGATTGG (New Primer)
	R14388	TTAGCGATGGAGGTAGGATTCG (Old Primer)
52	F14189	ACAAACAATGGTCAACCAGTAAC
	R14926	TGAGGCGTCTGGTGAGTAGTGC
53	F14470	TCCAAAGACAACCATCATTCC
	R14996	CGTGAAGGTAGCGGATGATTC
54	F14909	TACTCACCAGACGCCTCAACCG
	R15396	TTATCGGAATGGGAGGTGATTC
55	F15260	AGTCCCACCCTCACACGATTC
	R15774	ACTGGTTGTCCTCCGATTCAGG

Primer Set Number		Primer Sequence
56	F15574	CGCCTACACAATTCTCCGATC
	R16084	CGGTTGTTGATGGGTGAGTC
57 (HV1)	F15971	TTAACTCCACCATTAGCACC
	R16451	GCGAGGAGAGTAGCACTCTTG
58	F16097	TACATTACTGCCAGCCACCATG
	R336	TTAAGTGCTGTGGCCAGAAG
-21M13	F	TGTAAAACGACGGCCAGT

HV2: Hypervariable region 2
HV1: Hypervariable region 1
F: forward primer
R: reverse primer

These are the same primers used for SRM 2392 and reference 1 except the reverse primer of set 51 has been changed to: TTAGCGATGGAGGTAGGATT**G**G. The change (C to G) occurs at np 14368 and is in bold and underlined. Those using SRM 2392 should also use the new reverse primer 51.

REFERENCES

[1] Levin, B.C.; Cheng, H.; Reeder, D.J.; *A Human Mitochondrial DNA Standard Reference Material for Quality Control in Forensic Identification, Medical Diagnosis, and Mutation Detection*; Genomics, Vol. 55, pp. 135-146 (1999).

[2] Levin, B.C.; Holland, K.A.; Hancock, D.K.; Coble, M.; Parsons, T.J.; Kienker, L.J.; Williams, D.W.; Jones, MP.; Richie, K.L.; *Comparison of the Complete mtDNA Genome Sequences of Human Cell Lines - HL-60 and GM10742A - from Individuals with Promyelocytic Leukemia and Leber Hereditary Optic Neuropathy, Respectively, and the Inclusion of HL-60 in the NIST Human Mitochondrial DNA Standard Reference Material - SRM 2392-I*; Mitochondrion, Vol. 2, pp. 386-399 (2003).

[3] Anderson, S.; Bankier, A.T.; Barrell, B.G.; deBrujin, M.H.L.; Coulson, A.R.; Drouin, J.; Eperon, I.C.; Nierlich, D.P.; Roe, B.A.; Sanger, F.; Schreier, P.H.; Smith, A.J.H.; Staden, R.; Young, I.G.; *Sequence and Organization of the Human Mitochondrial Genome*; Nature, Vol. 290, pp. 457-465 (1981).

[4] Andrews, R.M.; Kubacka, I.; Chinnery, P.F.; Lightowlers, R.N.; Turnbull, D.M.; Howell, N.; *Reanalysis and Revision of the Cambridge Reference Sequence for Human Mitochondrial DNA*; Nature Genetics, Vol. 23, p. 147 (1999).

Users of this SRM should ensure that the certificate in their possession is current. This can be accomplished by contacting the SRM Program at: telephone (301) 975-6776; fax (301) 926-4751; e-mail srminfo@nist.gov; or via the Internet http://www.nist.gov/srm.

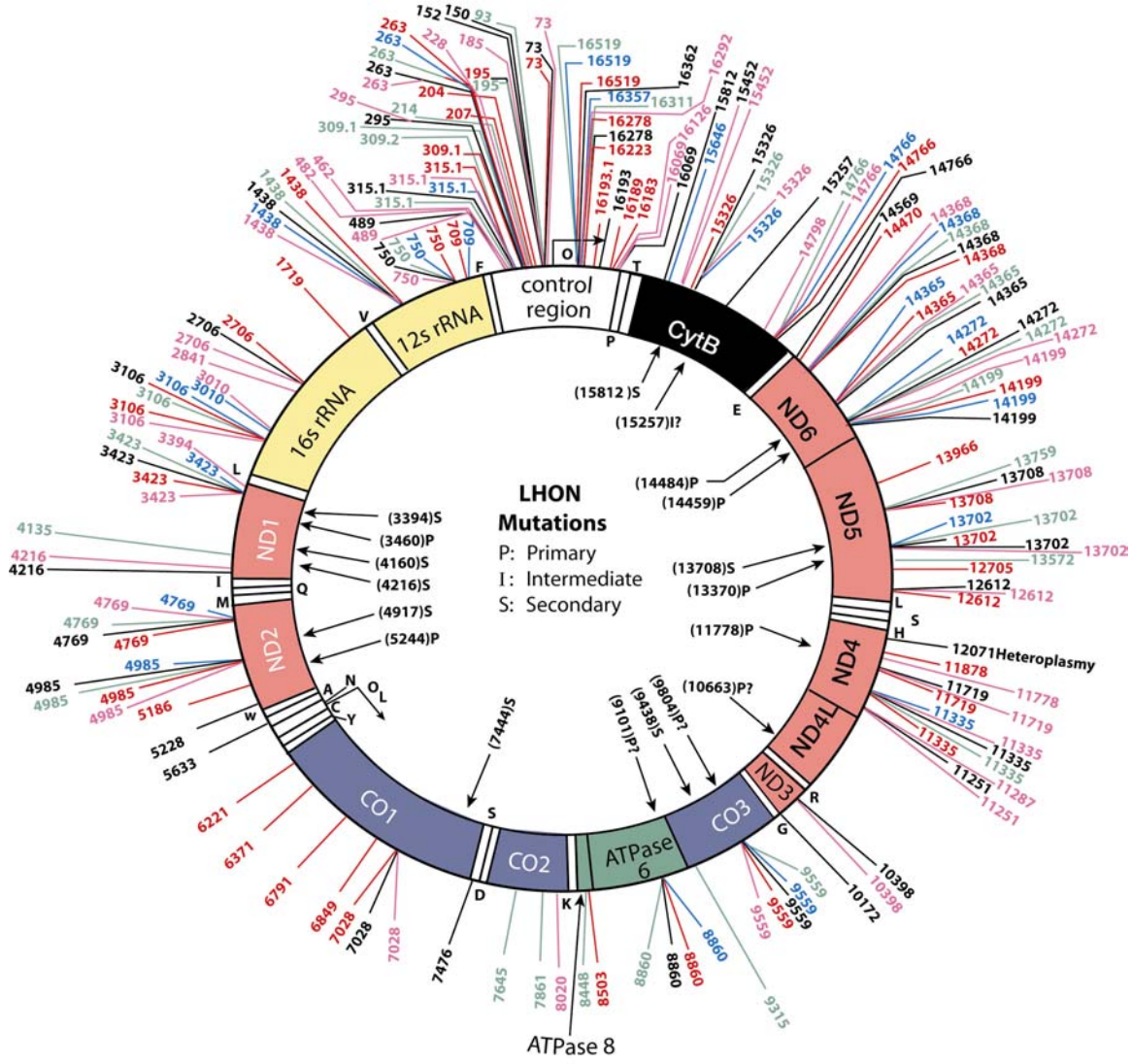


Figure 1. Schematic of human mtDNA showing its circular double-stranded DNA and all the differences from Cambridge Reference Sequence (1981) found in CHR (red), 9947A (green), HL-60 (black), GM03798 (blue), and GM10742A (purple) as numbers along the outside of the color-coded circle. Locations of the control region, rRNAs and genes coded by human mtDNA are shown. The locations of the 22 tRNAs are noted by white areas in the circle and designated by their single letter amino acid code. Since a number of mutations found in GM10742A and HL-60 and one change in CHR have been associated with primary, intermediate or secondary mutations linked to the disease Leber Hereditary Optic Neuropathy (LHON), the position of these mutations plus other LHON mutations are shown on the inside of the circle (Wallace et al., 1997). The question mark following the np of the LHON mutations indicates the assignment is not confirmed. One of the primary mutations that have been associated with LHON, G11778A, was found in GM10742A [2] but not found in the other DNA templates examined in this research. (Modified from Levin et al., 1999).

Appendix C

Journal Article:

B. C. Levin, H. Cheng, and D. J. Reeder, "A Human Mitochondrial DNA Standard Reference Material for Quality Control in Forensic Identification, Medical Diagnosis, and Mutation Detection," Genomics **55**, 135-146 (1999).

(See end of Appendix C for color representation of Figure 2.)

Genomics 55, 135–146 (1999)
Article ID geno.1998.5513, available online at http://www.idealibrary.com on IDEAL®

A Human Mitochondrial DNA Standard Reference Material for Quality Control in Forensic Identification, Medical Diagnosis, and Mutation Detection

Barbara C. Levin,*,[1] Haiyan Cheng,† and Dennis J. Reeder*

*Biotechnology Division, Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899; and †GEO-CENTERS, Inc., Newton Centre, Massachusetts 02159

Received May 8, 1998; accepted August 4, 1998

A human mitochondrial DNA (mtDNA) standard reference material (SRM 2392) will provide quality control when mtDNA is sequenced for forensic identifications, medical diagnosis, or mutation detection. SRM 2392 includes DNA from two lymphoblast cell cultures (CHR and 9947A) and cloned DNA from the CHR HV1 region, which contains a C stretch and is difficult to sequence. The mtDNA sequence (but not the DNA) of a third human template GM03798 is provided for comparison. Fifty-eight unique primer sets allow any area or the entire mtDNA (16,569 bp) to be amplified and sequenced. While none of the differences in these three templates correspond to published mutations associated with specific diseases, some of these differences did result in animo acid changes compared with that published by S. Anderson *et al.* (1981, *Nature* 290: 457–465). An interlaboratory evaluation of the amplification, sequencing, and data analysis of the CHR template was conducted by four laboratories. Corroboration of the SRM results will provide quality assurance that any unknown mtDNA is also being amplified and sequenced correctly.

 Some of these data have been presented at the following meetings: Environmental Mutagen Society 27th Annual Meeting, Victoria, British Columbia, Canada, March 1996; Seventh International Symposium on Human Identification, Scottsdale, Arizona, September 1996; Human Identification User's Meeting, Rockville, Maryland, November 1996; The American Association of Clinical Chemistry, Inc., San Diego, California, November 1996; The Society of Toxicology Annual Meeting, Cincinnati, Ohio, March 1997; Genetic Toxicology Gordon Conference, New London, New Hampshire, June 1997; and the 47th Annual Meeting of The American Society of Human Genetics, Baltimore, Maryland, October 1997.

[1] To whom correspondence should be addressed at 100 Bureau Drive, Stop 8311, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8311. Telephone: (301) 975-6682. Fax: (301) 330-3447. E-mail: barbara.levin@nist.gov.

INTRODUCTION

Human mitochondrial DNA (mtDNA) has been completely sequenced and found to be circular double-stranded molecules containing 16,569 bp (Anderson *et al.*, 1981). Each human cell can have a few dozen to several thousand molecules of mtDNA (Bogenhagen and Clayton, 1974; King and Attardi, 1989). Sequence analysis of mtDNA is being used by the forensic community for human identification, especially in those cases in which genomic DNA is highly degraded or nonexistent (Holland *et al.*, 1993, 1995). Forensic analysis to determine the distinction between individuals is primarily based on the considerable sequence variation found in the two hypervariable regions (HV1, HV2) located in the noncoding displacement loop (D loop). The medical community is also using sequence analysis of mtDNA for diagnoses of diseases associated with specific mutations and deletions (Wallace *et al.*, 1997). A third area of research, which is largely unexplored and which needs sequence analysis, is the examination of the mutagenic effects of chemical and physical agents on mtDNA (Grossman, 1995; Ballinger *et al.*, 1996). The objective of this research was to develop a human mtDNA standard reference material (SRM) for quality control in sequencing, forensic identifications, medical diagnostics, and mutation detection.

MATERIALS AND METHODS

Origin of extracted DNA. The DNA template designated CHR came from human white blood cells that were transformed with the Epstein–Barr virus and immortalized as a cell culture line (CHR cells) by the American Type Culture Collection (ATCC; Rockville, MD). After transformation, the cells were grown in Iscove's modified Dulbecco's media or RPMI 1640 media with L-glutamine, sodium bicarbonate, penicillin, streptomycin, and 20% fetal calf serum (Life Technologies, Inc., Grand Island, NY). The cell cultures were grown at 37°C in humidified atmosphere containing 5% CO_2 and 95% air. The DNA was extracted from 2×10^8 CHR cells by the Qiagen Plasmid/Cosmid Purification Protocol (Qiagen, Inc., Chatsworth, CA). This procedure enhanced the concentration of mtDNA and reduced, but did not eliminate, nuclear DNA.

The CHR data presented in this paper were obtained primarily with the above-mentioned immortalized CHR cell culture line. How-

0888-7543/99 $30.00

ever, before production of the final SRM began, it was necessary to obtain fresh blood from CHR and to reestablish the cell line. This second CHR cell line was established by the ATCC as above. The sequence of this second CHR cell line was examined and found to be identical to that of the first CHR cell line, with the single exception that no heteroplasmy was noted at bp 6849; the second CHR cell line agreed with Anderson at bp 6849. It is the second CHR cell line that is included in SRM 2392.

The DNA template 9947A was obtained from Life Technologies, Inc., who prepared it from a Epstein–Barr virus-immortalized human lymphoid cell line. DNA from 9947A is also used in the PCR-based DNA profiling standard (SRM 2391)[2] designed for forensic and paternity testing, law enforcement training, and research.

A third DNA template was extracted from an apparently normal human lymphoblastoid cell culture (GM03798) obtained from NIGMS[3] and grown in the same manner as the CHR cells. The DNA was extracted using DNA NOW, a phenol-free DNA isolation reagent (BIOGENTEX, Seabrook, TX). The information on this template is included for informational purposes only; the DNA is not included as part of this SRM.

Isolation and cloning of mtDNA containing the C stretch. Confluent CHR cells were harvested by centrifugation at 1500 rpm for 5 min. The mtDNA was isolated using the Qiagen Plasmid/Cosmid Purification Protocol (Qiagen, Inc.). Following isolation, the mtDNA was digested with restriction enzymes *Sac*I and *Kpn*I (New England Biolabs, Inc., Beverly, MA) into five fragments which were separated on a 0.7% low-melting-agarose gel. Bands of the size of the fragment containing the HV1 region were cut from the gel and melted at 65°C. DNA was extracted with phenol twice and precipitated by adding sodium chloride (150 mM) and 2 vol of 100% ethanol. The final product was resuspended in Tris–EDTA (TE) buffer.

The cloning vector, M13mp18, was also digested with *Sac*I and *Kpn*I, treated with calf intestinal alkaline phosphatase (New England Biolabs, Inc.), extracted with phenol, and precipitated with NaCl and ethanol as described above. The vector was incubated with the mtDNA product and T_4 DNA ligase (Life Technologies, Inc.) at 4°C overnight. An overnight culture of *Escherichia coli* host TG-1 cells was diluted and grown at 37°C in LB media (Sambrook *et al.*, 1989) until the OD_{650} reached 0.4–0.5. The cells were harvested by centrifugation at 1500 rpm for 5 min. The cell pellet was resuspended in 10 ml calcium chloride (50 mM) and incubated for 1 h on ice, centrifuged, and resuspended in 1 ml calcium chloride (50 mM) and incubated for 30 min on ice. The treated TG-1 cell suspension (0.3 ml) was incubated for 30 min on ice with 20 μl of a ligation mixture containing the isolated mtDNA fragment, the cloning vector that had been treated overnight, T_4 ligase, and ligation buffer (Life Technologies, Inc.). This mixture was then exposed to a heat shock of 42°C for 2 min; mixed with 0.2 ml of untreated TG-1 cells (from the overnight culture), 4 μl of 1 M isopropylthio-β-D-galactoside, 40 μl of 20 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, and 3 ml of melted (55°C) top agar; and spread on the surface of freshly prepared LB agarose plates (Sambrook *et al.*, 1989). The plates were incubated at 37°C overnight. Both colorless and blue plaques were visible in the morning. The colorless plaques indicate that insertion of the vector has occurred, whereas the blue plaques have no insertion of the vector.

Bacteriophage DNA isolation and sequencing. Single, well-isolated colorless plaques from the above LB plates were each placed in a sterile tube with 1.5 ml of a TG-1/LB cell suspension that contained TG-1 cells that were grown overnight and diluted 1/100 in LB media and grown for 1 h at 37°C. The plaques and the TG-1/LB cell suspension were grown at 37°C for 5 h. Cell debris was removed by centrifugation at 15,000 rpm for 5 min. The supernatant containing the bacteriophage was incubated with 0.2 ml polyethylene glycol (20% PEG in 2.5 M NaCl) overnight at 4°C, and the resultant precipitate containing the DNA was pelleted by centrifugation at 15,000 rpm for 15 min. The bacteriophage DNA was isolated by phenol extraction and NaCl/ethanol precipitation as described above and then dissolved in 25 μl of TE buffer. The bacteriophage DNA was cycle sequenced with AmpliTaq DNA polymerase, FS, and the −21 M13 primer: 5′-TGTAAAACGACGGCCAGT-3′ according to the protocol in the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin–Elmer, Foster City, CA). The cycle sequencing was conducted in a Perkin–Elmer Model 9600 thermocycler by first heating the DNA reaction mixture at 96°C for 1 min and then subjecting the mixture to 25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 2 min. The cycle sequencing product was purified using a Centri-Sep spin column (Princeton Separations, Inc., Adelphia, NJ). The DNA pellet was rinsed with 70% ethanol, vacuum dried, resuspended in loading buffer prepared by combining deionized formamide and 25 mM EDTA (pH 8.0) in a ratio of 5:1, loaded onto a 4.75% acrylamide gel, and electrophoresed on an ABI 373 DNA sequencer. One of the clones containing the C stretch sequence was used as the source of the cloned DNA for the SRM.

mtDNA primers. Fifty-eight sets of unique primers (19–28 bp) for sequencing the entire mtDNA (16,569 bp) were computer-designed using Gene Runner for Windows (Hastings Software, Inc., Hastings, NY) and custom made by Bio-Synthesis, Inc. (Lewisville, TX). The −21M13 primer was used to sequence the cloned HV1 region of the DNA from the CHR template. The sequences of all the primers are shown in Table 1.

Polymerase chain reaction (PCR). Extracted DNA was resuspended in TE buffer (pH 7.5) containing 10 mM Tris and 1 mM EDTA. The PCR mixture contained DNA (1 μl), *Taq* DNA polymerase (0.5 μl or 2.5 units) (Boehringer Mannheim), and 10× buffer (5 μl) (Boehringer Mannheim), dNTP's (0.2 mM each) (Life Technologies, Inc.), forward and reverse primers (0.4 μM each), plus H_2O to a final volume of 50 μl. The 10× buffer (pH 8.3) contained Tris–HCl (100 mM), $MgCl_2$ (15 mM), and KCl (500 mM).

Thermal cycling was conducted in a Perkin–Elmer Model 9600 thermocycler and consisted of 1 min at 96°C; followed by 32 cycles of 15 s at 94°C (denaturation), 30 s at 56°C (annealing), and 15 s at 72°C (extension); and ending with a final extension of 7 min at 72°C.

A sample of the amplified DNA was electrophoresed in 0.7% agarose and stained with ethidium bromide to assess the purity and size of the PCR product. Before sequencing, extraneous materials were removed from the PCR product with a QIAquick PCR Purification Kit (QIAGEN, Inc.).

Sequencing. Cycle sequencing using fluorescent dye-labeled terminators was performed with an ABI PRISM Dye Terminator Cycle Sequencing Kit with AmpliTaq DNA polymerase, FS (Perkin–Elmer, Foster City, CA). Thermal cycling was conducted in a Perkin–Elmer Model 9600 thermocycler and started with 1 min at 96°C. The reaction then underwent 25 cycles of 96°C for 15 s (denaturation), 50°C for 5 s (annealing), and 60°C for 2 min (extension). The DNA product was purified by passage through a Centri-Sep spin column (Princeton Separations, Inc.).

Electrophoresis and sequencing of the fluorescently labeled purified DNA were performed with a 373 ABI Sequencer (Perkin–Elmer) using a 4.75% acrylamide gel. Data analysis was executed with the Sequence Navigator software package (Perkin–Elmer).

Interlaboratory evaluation. Three laboratories in addition to NIST participated in an interlaboratory evaluation of the CHR template. These laboratories were The Bode Technology Group, Inc. (21515 Ridgetop Circle, Suite 140, Sterling, VA 20166), IIT Research Institute (Virginia Technology Center, 8510 Cinderbed Road, Suite 300, P.O. Box 899, Newington, VA 22122), and Lark Technologies, Inc. (9545 Katy Fwy, Suite 465, Houston, TX 77024).

Each laboratory was sent:

1. Two tubes of DNA from the first CHR cell culture line. One tube contained extracted DNA ready for PCR amplification of the entire mtDNA. The other contained the cloned DNA ready for cycle sequencing of the HV1 region (this DNA did *not* need to be PCR amplified).

[2] SRM 2391 may be obtained from the Standard Reference Material Program, NIST, Gaithersburg, Maryland 20899.

[3] NIGMS Human Genetic Mutant Cell Repository, Coriell Institute for Medical Research, 401 Haddon Avenue, Camden, New Jersey 08103.

2. Fifty-eight sets of primers labeled with either F# (forward primer) or R# (reverse primer). Forward and reverse primers with the same number were paired and numbered from the 5′ end. Primers were diluted to 10 μl and ready for use. Also enclosed was the −21M13 primer for the sequencing of the cloned HV1 region of the CHR DNA, which covered basepairs 16133 to 40.

3. The protocol used at NIST to amplify and sequence the DNA. The laboratories, however, were free to use any protocol with which they were familiar and felt comfortable.

4. A form table to record the results. This table provided the number of the primer set, the region that each primer set amplified, and the length of the amplified region. We requested that the laboratory fill in the differences found when they compared the sequence that they determined for the SRM with that of Anderson *et al.* (1981).

5. The following cautionary note:

> WARNING: The DNA and cells were derived from a cell culture line from an apparently healthy human subject. The cell culture line has been tested and found to be nonreactive for hepatitis B surface antigen and HIV. However, no test method can ensure that a product derived from human blood does not contain HIV, hepatitis or other infectious agents. HANDLE AS IF CAPABLE OF TRANSMITTING DISEASE. (The second CHR cell culture line generated from the same individual was not tested again for hepatitis or HIV. Normal precautions should be used.)

Differences in methodology used by laboratories in interlaboratory evaluation. The Bode Technology Group, Inc., essentially followed the NIST protocol except that they used a 6% acrylamide/8.3 M urea gel for the sequencing electrophoresis instead of a 4.75% acrylamide.

IIT Research Institute also followed the NIST protocol except that they used *Taq* Gold (Perkin–Elmer) for the amplification reaction, which was modified to include a hot start of 95°C for 11 min. Microcon 100 microconcentrators (Amicon, Inc., Beverly, MA) were used to purify the PCR products. The quantities of DNA were detemined by capillary electrophoresis (CE) with a Beckman P/ACE 5010 System (Beckman Instruments, Inc., Fullerton, CA) as follows: 1 μl of the amplified product was mixed with 25 μl of sterile deionized H_2O containing 0.52 ng/μl of a 200-bp internal standard (GenSura Laboratories, Inc., Del Mar, CA) and run on the CE. One determines the quantity of the amplified product from the ratio of the PCR product peak area to the internal standard peak area multiplied by a migration standard. A 6% acrylamide gel was used for the sequencing electrophoresis instead of a 4.75% acrylamide.

Lark Technologies, Inc., followed the NIST protocol with the following differences: AmpliTaq DNA polymerase (Perkin–Elmer) was used to amplify the DNA; the dNTP's were purchased from Pharmacia Biotech, Inc. (Piscataway, NJ); the products were purified with Qiaquick PCR purification kit (Qiagen); in the sequencing reactions, the amount of PCR product used varied from 1 to 3 μl based on the concentration estimated from agarose gels; cycling conditions were 95°C for 1 min followed by 25 cycles of 96°C for 15 s, 50°C for 15 s, and 60°C for 4 min; the sequence reactions were cleaned up by ethanol precipitation; a 4.25% polyacrylamide gel was used for the sequencing electrophoresis instead of 4.75% acrylamide; and an ABI 377 was used instead of the ABI 373. Electropherograms were printed for each reaction, and the sequences were manually edited based on the electropherogram patterns. Printed electropherograms and a floppy disk with the sequence data were sent to NIST where the data were compared to the Anderson sequence.

RESULTS AND DISCUSSION

SRM templates. Two DNA templates, CHR and 9947A, are included in the NIST human mtDNA sequencing SRM 2392. Both of these DNA samples come from human cell culture lines that were developed from apparently normal individuals. The DNA from 9947A is the total extracted DNA, which also includes nuclear DNA. The DNA from CHR was isolated in a manner that enhanced the concentration of the mtDNA, but did not totally eliminate the nuclear DNA. The SRM also provides cloned DNA from the HV1 region of the CHR template, which contains a C stretch. In most people, the HV1 region has a string of cytosine (C) residues interrupted by a thymine (T) at nucleotide position 16189.[4] In some individuals, however, a transition that changes the T to a C occurs, producing a long string of Cs called the C stretch. When this happens, sequencing beyond the C stretch becomes very difficult, if not impossible. Clones of the HV1 region containing the C stretch indicated that the number of Cs differed among the different clones and the difficulty in sequencing was due to frameshifts that resulted from the simultaneous sequencing of templates with differing numbers of Cs (Bendall and Sykes, 1995; Levin *et al.,* 1995, 1997). We found, however, that one could sequence through the entire HV1 region including the C stretch without problems if one used the clone of the area. Therefore, we have included the cloned HV1 region of the CHR DNA template in the SRM.

In addition to templates CHR and 9947A, we have included all the information regarding a third template from a lymphoblastoid cell line (GM03798) that was obtained from the NIGMS Human Genetic Mutant Cell Repository and that was completely sequenced at NIST three times. The DNA from this cell line is not part of the SRM, and the data are included for information only.

Primers. The 58 sets of unique primers were designed to allow the amplification and sequencing of any region or the entire 16,569 bp that comprise human mitochondrial DNA. The sequences of both the forward and the reverse primers that are in each set are shown in Table 1. The numbers indicate the 5′ end of the primer. They are all between 19 and 28 bp long, and the criteria that were used to choose these primers were primer T_m = 50–65°C, primer length = 15–30 bp, and PCR product length = 400–850 bp. The primers were designed to produce sequences that overlapped with both the previous and the following regions to allow those areas in the beginning and end of electropherograms, which are difficult to sequence, to become readable. Table 2 shows the number of the base where the readable sequence begins and ends (indicated in the table as "start" and "end"). The readable region is always smaller than the amplified region.

In addition to the designed primers, we also used the −21M13 primer (Table 1) to sequence the cloned DNA from the HV1 region of the CHR template that contained the C stretch. The PCR products produced single, distinct bands for all 58 primer sets (Figs. 1A and 1B).

Differences between the SRM templates and the Anderson sequence. Anderson and his co-workers completely sequenced human mtDNA in 1981 (Anderson *et al.,* 1981). All investigators who subsequently examined human mtDNA have used the numbering system of Ander-

[4] All nucleotide numbers referred to in this paper are based on the numbering system of Anderson *et al.* (1981).

TABLE 1

Primer Sets Used for PCR Amplification of Human mtDNA

Primer set		Primer sequence	Primer set		Primer sequence
1 (HV2)	F15	CACCCTATTAACCACTCACG		R8215	GACGATGGGCATGAAACTG
	R484	TGAGATTAGTAGTATGGGAG	31	F7901	TGAACCTACGAGTACACCGACTAC
2	F361	ACAAAGAACCCTAACACCAGC		R8311	AAGTTAGCTTTACAGTGGGCTCTAG
	R921	ACTTGGGTTAATCGTGTGACC	32	F8164	CGGTCAATGCTCTGAAATCTGTG
3	F756	CATCAAGCACGCAGCAATG		R8669	CATTGTTGGGTGGTGATTAGTCG
	R1425	AATCCACCTTCGACCCTTAAG	33	F8539	CTGTTCGCTTCATTCATTGCC
4	F873	GGTTGGTCAATTTCGTGCCAG		R9059	GTGGCGCTTCCAATTAGGTG
	R1425	AATCCACCTTCGACCCTTAAG	34	F8903	CCCACTTCTTACCACAAGGC
5	F1234	CTCACCACCTCTTGCTCAGC		R9403	GTGCTTTCTCGTGTTACATCG
	R1769	GCCAGGTTTCAATTTCTATCG	35	F9309	TTTCACTTCCACTCCATAACGC
6	F1587	TGCACTTGGACGAACCAGAG		R9848	GAAAGTTGAGCCAATAATGACG
	R2216	TGTTGAGCTTGAACGCTTTC	36	F9449	CGGGATAATCCTATTTATTACCTCAG
7	F1657	CTTGACCGCTCTGAGCTAAAC		R9995	AGAGTAAGACCCTCATCAATAGATGG
	R2216	TGTTGAGCTTGAACGCTTTC	37	F9754	AGTCTCCCTTCACCATTTCCG
8	F1993	AAACCTACCGAGCCTGGTG		R10275	AAAGGAGGGCAATTTCTAGATC
	R2216	TGTTGAGCTTGAACGCTTTC	38	F10127	ACTACCACAACTCAACGGCTAC
9	F2105	GAGGAACAGCTCTTTGGACAC		R10556	GGAGGATATGAGGTGTGAGCG
	R2660	AGAGACAGCTGAACCCTCGTG	39	F10386	GGATTAGACTGAACCGAATTGG
10	F2417	CACTGTCAACCCAACACAGG		R11166	CATCGGGTGATGATAGCCAAG
	R3006	ATGTCCTGATCCAACATCGAG	40	F10704	GTCTCAATCTCCAACACATATGG
11	F2834	CCCAACCTCCGAGCAGTACATG		R11267	TGTTGTGAGTGTAAATTAGTGCG
	R3557	AGAAGAGCGATGGTGAGAGC	41	F11001	AACGCCACTTATCCAGTGAACC
12	F2972	ATAGGGTTTACGACCTCGATG		R11600	CTGTTTGTCGTAGGCAGATGG
	R3557	AGAAGAGCGATGGTGAGAGC	42	F11403	GACTCCCTAAAGCCCATGTCG
13	F3234	AGATGGCAGAGCCCGGTAATC		R11927	TTGATCAGGAGAACGTGGTTAC
	R3557	AGAAGAGCGATGGTGAGAGC	43	F11760	ACGAACGCACTCACAGTCG
14	F3441	ACTACAACCCTTCGCTGACG		R12189	AAGCCTCTGTTGTCAGATTCAC
	R3940	TGAAGCCTGAGACTAGTTCGG	44	F11901	TGCTAGTAACCACGTTCTGGTG
15	F3635	GCCTAGCCGTTTACTCAATCC		R12876	GATATCGCCGATACGGTTG
	R4162	TGAGTTGGTCGTAGCGGAATC	45	F12357	AACCACCCTAACCCTGACTTCC
16	F3931	TCAGGCTTCAACATCGAATACG		R12876	GATATCGCCGATACGGTTG
	R4728	TTATGGTTCATTGTCCGGAGAG	46	F12601	TTCATCCCTGTAGCATTGTTCG
17	F4183	TTTCTACCACTCACCCTAGCATTAC		R13123	AGCGGATGAGTAAGAAGATTCC
	R4728	TTATGGTTCATTGTCCGGAGAG	47	F12793	TTGCTCATCAGTTGATGATACG
18	F4392	CCCATCCTAAAGTAAGGTCAGC		R13343	TTGAAGAAGGCGTGGGTACAG
	R4983	GGTTTAATCCACCTCAACTGCC	48	F13188	CACTCTGTTCGCAGCAGTATG
19	F4447	TTGGTTATACCCTTCCCGTAC		R13611	TCGAGTGCTATAGGCGCTTGTC
	R4982	GTTTAATCCACCTCAACTGCC	49	F13518	CATCATCGAAACCGCAAAC
20	F4797	CCCTTTCACTTCTGAGTCCCAG		R13935	TGTGATGCTAGGGTAGAATCCG
	R5553	AGGGCTTTGAAGGCTCTTG	50	F13715	GAAGCCTATTCGCAGGATTTC
21	F4976	ATTAAACCAGACCCAGCTACG		R14118	TGGGAAGAAGAAAGAGAGGAAG
	R5553	AGGGCTTTGAAGGCTCTTG	51	F13899	TTTCTCCAACATACTCGGATTC
22	F5318	CACCATCACCCTCCTTAACC		R14388	TTAGCGATGGAGGTAGGATTCG
	R5882	GCTGAGTGAAGCATTGGACTG	52	F14189	ACAAACAATGGTCAACCAGTAAC
23	F5700	TAAGCACCCTAATCAACTGGC		R14926	TGAGGCGTCTGGTGAGTAGTGC
	R6262	GCCTCCACTATAGCAGATGCG	53	F14470	TCCAAAGACAACCATCATTCC
24	F5999	TCTAAGCCTCCTTATTCGAGC		R14996	CGTGAAGGTAGCGGATGATTC
	R6526	ATAGTGATGCCAGCAGCTAGG	54	F14909	TACTCACCAGACGCCTCAACCG
25	F6242	CGCATCTGCTATAGTGGAGG		R15396	TTATCGGAATGGGAGGTGATTC
	R6526	ATAGTGATGCCAGCAGCTAGG	55	F15260	AGTCCCACCCTCACACGATTC
26	F6426	GCCATAACCCAATACCAAACG		R15774	ACTGGTTGTCCTCCGATTCAGG
	R7030	TGGGCTACAACGTAGTACGTG	56	F15574	CGCCTACACAATTCTCCGATC
27	F6744	GGCTTCCTAGGGTTTATCGTG		R16084	CGGTTGTTGATGGGTGAGTC
	R7255	TTTCATGTGGTGTATGCATCG	57 (HV1)	F15971	TTAACTCCACCATTAGCACC
28	F7075	GAGGCTTCATTCACTGATTTCC		R16451	GCGAGGAGAGTAGCACTCTTG
	R7792	GGGCAGGATAGTTCAGACGG	58	F16097	TACATTACTGCCAGCCACCATG
29	F7215	CGACGTTACTCGGACTACCC		R336	TTAAGTGCTGTGGCCAGAAG
	R7792	GGGCAGGATAGTTCAGACGG	−21M13	F	TGTAAAACGACGGCCAGT
30	F7645	TATCACCTTTCATGATCACGC			

son and have compared their sequence findings to those described by Anderson. However, the DNA sequenced by Anderson is not available for use as a positive control during actual experiments, whereas NIST SRM 2392 would be available. Table 2 shows the mtDNA differences compared to the Anderson sequence that were found at NIST with all three templates—CHR, 9947A, and GM03798. In all three templates, all 58 areas comprising

TABLE 2

Primer Sets Used for PCR Amplification of Human mtDNA and Differences with the Anderson Sequence Found in Three Templates at NIST

Primer set	Amplified region[a]	Length of amplified region	Comparison with Anderson					Amino acid change
			Anderson No.	Anderson bp	Template CHR	Template 9947A	Template GM03798	
1 (HV2)	15–484	470			Start 39	Start 39	Start 55	
			73	A	G	—	—	
			93	A	—	G	—	
			195	T	C	C	—	
			204	T	C	—	—	
			207	G	A	—	—	
			214	A	—	G	—	
			263	A	G	G	G	
			309.1		C(ins)	C(ins)	—	
			309.2		—	C(ins)	—	
			315.1		C(ins)	C(ins)	C(ins)	
					End 436	End 473	End 454	
2	361–921	561			Start 429	Start 421	Start 415	
			709	G	A	—	A	
			750	A	G	G	G	
					End 891	End 846	End 834	
3	756–1425	670	None		Start 778	Start 778	Start 818	
					End 1197	End 1278	End 1146	
4	873–1425	553	None		Start 931	Start 928	Start 938	
					End 1335	End 1377	End 1323	
5	1234–1769	536			Start 1279	Start 1275	Start 1295	
			1438	A	G	G	G	
			1719	G	A	—	E	
					End 1738	End 1741	End 1654	
6	1587–2216	630			Start 1632	Start 1632	Start 1649	
			1719[b]	G	A	—	—	
					End 2106	End 2106	End 2031	
7	1657–2216	560			Start 1691	Start 1686	Start 1715	
			1719[b]	G	A	—	—	
					End 2170	End 2173	End 2097	
8	1993–2216	224	None		Start 2036	Start 2018	Start 2069	
					End 2213	End 2217	End 2212	
9	2105–2660	556	None		Start 2157	Start 2150	Start 2161	
					End 2636	End 2586	End 2560	
10	2417–3006	590			Start 2465	Start 2458	Start 2483	
			2706	A	G	—	—	
					End 2920	End 2956	End 2915	
11	2834–3557	724			Start 2861	Start 2869	Start 2888	
			3010	G	—	—	A	
			3106/3107	C	Del	Del	Del	
					End 3350	End 3373	End 3243	
12	2972–3557	586			Start 2999	Start 2999	Start 3031	
			3106/3107[b]	C	Del	Del	Del	
			3423	G	E	T	T	Silent
					End 3422	End 3460	End 3425	
13	3234–3557	324			Start 3265	Start 3258	Start 3292	
			3423[b]	G	T	T	T	Silent[b]
					End 3548	End 3545	End 3541	
14	3441–3940	500	None		Start 3487	Start 3491	Start 3499	
					End 3916	End 3920	End 3847	
15	3635–4162	528	None		Start 3667	Start 3662	Start 3725	
					End 4126	End 4061	End 4044	
16	3931–4728	798			Start 3964	Start 3968	Start 3987	
			4135	T	—	C	—	Try → His
					End 4399	End 4427	End 4436	
17	4183–4728	546	None		Start 4208	Start 4249	Start 4208	
					End 4657	End 4657	End 4642	
18	4392–4982	591			Start 4449	Start 4453	Start 4440	
			4769	A	G	G	G	Silent
					End 4860	End 4935	End 4877	
19	4447–4982	536			Start 4492	Start 4492	Start 4492	
			4769[b]	A	G	G	G	Silent[b]

TABLE 2—*Continued*

Primer set	Amplified region[a]	Length of amplified region	Comparison with Anderson					Amino acid change
			Anderson No.	Anderson bp	Template CHR	Template 9947A	Template GM03798	
					End 4958	End 4921	End 4931	
20	4797–5553	757			Start 4838	Start 4845	Start 4838	
			4985	G	A	A	A	Silent
			5186	A	G	—	—	Silent
					End 5327	End 5324	End 5215	
21	4976–5553	578			Start 5000	Start 5007	Start 5016	
			5186[b]	A	G	—	—	Silent[b]
					End 5516	End 5521	End 5400	
22	5318–5882	565	None		Start 5361	Start 5360	Start 5371	
					End 5754	End 5758	End 5800	
23	5700–6262	563	None		Start 5741	Start 5744	Start 5754	
					End 6149	End 6163	End 6136	
24	5999–6526	528			Start 6043	Start 6058	Start 6047	
			6221	T	C	—	—	Silent
			6371	C	T	—	—	Silent
					End 6442	End 6503	End 6456	
25	6242–6526	285			Start 6271	Start 6302	Start 6293	
			6371[b]	C	T	—	—	Silent[b]
					End 6520	End 6520	End 6520	
26	6426–7030	605			Start 6451	Start 6474	Start 6487	
			6791	A	G	—	—	Silent
			6849*	A	G (0.3A)[h]*	—	—	Thr → Ala*
					End 6916	End 6930	End 6885	
27	6744–7255	512			Start 6775	Start 6782	Start 6801	
			6849[b,]*	A	G (0.3A)[h]*	—	—	Thr → Ala[b,]*
			7028	C	T	—	—	Silent
					End 7215	End 7221	End 7177	
28	7075–7792	718	None		Start 7123	Start 7123	Start 7130	
					End 7602	End 7601	End 7547	
29	7215–7792	578			Start 7263	Start 7280	Start 7273	
			7645	T	—	C	—	Silent
					End 7722	End 7769	End 7706	
30	7645–8215	571			Start 7671	Start 7666	Start 7701	
			7861	T	—	C	—	Silent
					End 8149	End 8155	End 8156	
31	7901–8311	411	None		Start 7960	Start 7959	Start 7960	
					End 8289	End 8288	End 8258	
32	8164–8669	506			Start 8211	Start 8212	Start 8230	
			8448	T	—	C	—	Met → Thr
			8503	T	C	—	—	Silent
					End 8646	End 8641	End 8637	
33	8539–9059	521			Start 8581	Start 8582	Start 8581	
			8860	A	G	G	G	Thr → Ala
					End 9019	End 8999	End 8991	
34	8903–9403	501			Start 8947	Start 8944	Start 8951	
			9315	T	—	C	—	Phe → Leu
					End 9380	End 9381	End 9370	
35	9309–9848	540			Start 9334	Start 9333	Start 9333	
			9559	G	C	C	C	Arg → Pro
					End 9823	End 9827	End 9800	
36	9449–9995	547			Start 9476	Start 9485	Start 9479	
			9559[b]	G	C	C	C	Arg → Pro[b]
					End 9964	End 9940	End 9911	
37	9754–10275	522	None		Start 9777	Start 9781	Start 9808	
					End 10225	End 10251	End 10184	
38	10127–10556	430	None		Start 10168	Start 10166	Start 10180	
					End 10534	End 10536	End 10524	
39	10386–11166	781	None		Start 10410	Start 10416	Start 10439	
					End 10899	End 10916	End 10865	
40	10704–11267	564	None		Start 10734	Start 10742	Start 10758	
					End 11223	End 11197	End 11167	
41	11001–11600	600			Start 11026	Start 11040	Start 11059	
			11335	T	C	C	C	Silent
					End 11461	End 11517	End 11497	

TABLE 2—*Continued*

Primer set	Amplified region[a]	Length of amplified region	Comparison with Anderson					Amino acid change
			Anderson No.	Anderson bp	Template CHR	Template 9947A	Template GM03798	
42	11403–11927	525			Start 11428	Start 11432	Start 11456	
			11719	G	A	—	—	Silent
					End 11795	End 11853	End 11855	
43	11760–12189	430			Start 11784	Start 11802	Start 11802	
			11878	T	C	—	—	Silent
					End 12159	End 12164	End 12163	
44	11901–12876	976	None		Start 11926	Start 11926	Start 11961	
					End 12404	End 12443	End 12397	
45	12357–12876	520			Start 12404	Start 12391	Start 12391	
			12612	A	G	—	—	Silent
			12705	C	T	—	—	Silent
					End 12769	End 12849	End 12775	
46	12601–13123	523			Start 12627	Start 12645	Start 12643	
			12705[b]	C	T	—	—	Silent[b]
					End 13102	End 13045	End 13024	
47	12793–13343	551	None		Start 12817	Start 12807	Start 12816	
					End 13295	End 13307	End 13266	
48	13188–13611	424			Start 13238	Start 13238	Start 13244	
			13572	T	—	C	—	Silent
					End 13587	End 13593	End 13590	
49	13518–13935	418			Start 13541	Start 13541	Start 13571	
			13572[b]	T	—	C	—	Silent[b]
			13702	G	C	C	C	Gly → Arg
			13708	G	A	—	—	Ala → Thr
			13759	G	—	A	—	Ala → Thr
					End 13910	End 13921	End 13900	
50	13715–14118	404			Start 13775	Start 13760	Start 13760	
			13966	A	G	—	—	Thr → Ala
					End 14094	End 14110	End 14104	
51	13899–14388	490			Start 13926	Start 13927	Start 13961	
			13966[b]	A	G	—	—	Thr → Ala[b]
			14199	G	T	T	T	Pro → Thr
			14272	G	C	C	C	Phe → Leu
			14365	G	C	C	E	Silent
					End 14369	End 14374	End 14342	
52	14189–14926	738			Start 14216	Start 14216	Start 14240	
			14272[b]	G	C	C	C	Phe → Leu[b]
			14365[b]	G	C	C	C	Silent[b]
			14368	G	C	C	C	Phe → Leu
			14470	T	C	—	—	Silent
			14766	T	E	C	E	Ile → Thr
					End 14699	End 14806	End 14698	
53	14470–14996	527			Start 14502	Start 14513	Start 14527	
			14766[b]	T	—	C	C	Ile → Thr[b]
					End 14957	End 14972	End 14956	
54	14909–15396	488			Start 14941	Start 14933	Start 14950	
			15326	A	G	G	G	Thr → Ala
					End 15380	End 15373	End 15359	
55	15260–15774	515			Start 15305	Start 15293	Start 15287	
			15326[b]	A	G	G	G	Thr → Ala[b]
			15646	C	—	—	T	Silent
					End 15754	End 15950	End 15723	
56	15574–16084	511			Start 15637	Start 15599	Start 15601	
			15646[b]	C	—	—	T	Silent[b]
					End 16056	End 16058	End 16030	
57 (HV1)	15971–16451	481			Start 16014	Start 16011	Start 16004	
			16183	A	C	—	—	
			16189	T	C	—	—	
			16311	T	E	C	—	
			16357	T	E	—	C	
					End 16193	End 16430	End 16403	
58	16097–336	809			Start 16125	Start 16130	Start 16151	
			16183[b]	A	C	—	—	
			16189[b]	T	C	—	—	

TABLE 2—*Continued*

Primer set	Amplified region[a]	Length of amplified region	Comparison with Anderson: Anderson No.	Anderson bp	Template CHR	Template 9947A	Template GM03798	Amino acid change
			16311[b]	T	E	C	—	
			16357[b]	T	E	—	C	
			16519	T	E	C	C	
					End 16193	End 59	End 103	
−21M13[c] cloned DNA	16133–40	477			Start 16131			
			16183[b]	A	C	ND	ND	
			16189[b]	T	C			
			16193.1		C(ins)			
			16223	C	T			
			16278	C	T			
			16519[b]	T	C			
					End 40			

Note. B, basepair change came before the readable sequence; E, basepair change came after the readable sequence; —, basepair same as in Anderson sequence; h*, possible heteroplasmic site. * This heteroplasmy seen in the first CHR cell culture line was not seen with the second CHR cell culture line. It is the second CHR cell culture line that is supplied in NIST SRM 2392; Start, start of readable sequence; End, end of readable sequence; CHR cells, sequence based on two amplifications and cycle sequencing procedures in first cell culture line and at least one amplification and cycle sequencing procedure with the second cell culture line; 9947A cells, sequence based on two amplifications and cycle sequencing procedures; GM03798, sequence based on three to four amplifications and cycle sequencing procedures; ins, insertion; Del, deletion; ND, not done.

[a] Numbers correspond to Anderson sequence (Anderson *et al.,* 1981).

[b] Change also seen in previous primer set.

[c] This primer is used for sequencing the cloned DNA of the HV1 region.

the entire mtDNA were completely amplified and sequenced at least twice (GM03798 was done three times). There were 13, 9, and 4 differences in the noncoding regions of templates CHR, 9947A, and GM03798, respectively, and 33, 23, and 19 differences in the coding regions of templates CHR, 9947A, and GM03798, respectively. All of the differences from Anderson found in these three templates are shown in Fig. 2 along with many of the diseases that have been noted in the literature (Wallace *et al.,* 1997). None of the basepair changes found in the coding regions of the three templates sequenced at NIST correlate with any of the changes found associated with these published disease states.

Meaning of the differences from Anderson. Since all three templates had come from apparently normal individuals, it was of interest to determine if the differences in the coding regions would actually cause amino acid changes in the resultant protein structures. The genetic code for human mtDNA is slightly different from the universal genetic code (Anderson *et al.,* 1981). One needs to consider these differences in the universal genetic code when determining the amino acid sequence designated by the 3-bp codons in mtDNA. Many of the differences from the Anderson sequence were in the third position wobble and did not affect the amino acid sequence (silent changes) (Table 2). However, CHR, 9947A, and GM03798 had 10 (9 without the heteroplasmy at bp 6849), 12, and 8 different basepairs, respectively, that would result in a different amino acid from that designated by the Anderson sequence (Table 2). However, data from the literature indicate that perhaps 14 basepair designations in the consensus sequence of Anderson may not be the sequence found in the majority of normal individuals (Howell *et al.,* 1992; Marzuki *et al.,* 1992). Our results agree with 11 of these 14 new designations (Table 3). Examination of our results using these new designations indicates that only 4 (3 without the heteroplasmy at bp 6849), 6, and 2 differences for CHR, 9947A, and GM03798, respectively, would result in amino acid changes. These structural changes, however, do not necessarily mean a functional change has occurred in the protein. To determine if a functional change has occurred, one still needs to decipher whether the amino acid change is in an active site on the protein.

The interlaboratory evaluation of the CHR template. An interlaboratory evaluation was conducted by four laboratories including NIST. All of the laboratories essentially followed the NIST protocol sent with the DNA from the first CHR cell culture line and the primers. Any changes to the protocol are listed under Materials and Methods. Each laboratory was instructed to amplify and sequence the 58 areas designated by the 58 primer sets and also to sequence the cloned DNA for the HV1 region. Laboratory 1 amplified and sequenced each area at least twice. The other labs amplified and sequenced the areas from one to six times. Laboratories 1, 2, and 3 found essentially the same polymorphisms. Laboratory 4 had less experience with sequencing mtDNA and did find differences that the other laboratories did not observe. Data were excluded

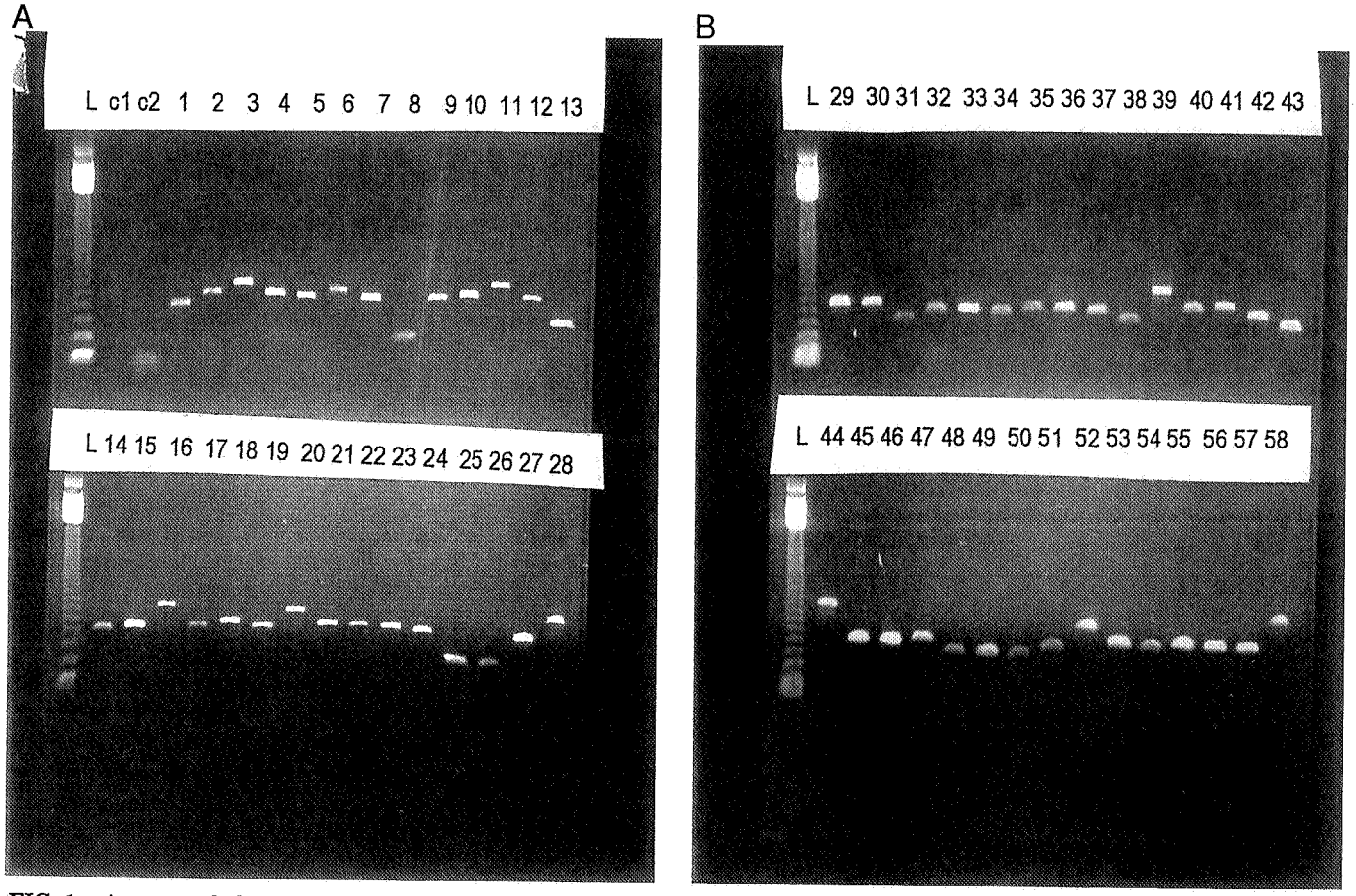


FIG. 1. Agarose gel electrophoresis of PCR products from 58 primer sets designed for human mitochondrial DNA. (**A**) Lane **L,** 123-bp ladder from Gibco BRL; lane **c1,** negative control with no primers; lane **c2,** negative control with no DNA; lanes **1–28,** PCR products from primer sets 1–28. (**B**) Lane **L,** 123-bp ladder from Gibco BRL; lanes **29–58,** PCR products from primer sets 29–58.

from the analysis of this interlaboratoy evaluation if the following conditions were observed: (1) The computer results were ambiguous as indicated by calling a peak "N" rather than A, C, G, or T. (2) The differences from Anderson were not consistently seen within a laboratory, i.e., if the laboratory sequenced in both the forward and the reverse directions and one direction agreed with Anderson and the other direction did not agree, we assumed the results that agreed with Anderson were correct. (3) Within any one laboratory, the difference from Anderson was seen with one primer set, but not in the overlapping sequences seen in the previous or subsequent primer sets. Even with these exclusions, Laboratory 4 had many differences that were not seen by the other labs. One problem was that they did not provide data from primer sets 29, 39, and 41, and we were unable to check those overlapping sequences. Laboratory 2 was unable to sequence the clone, which was not a problem for the other laboratories. Laboratory 3 was missing data from primer sets 36 and 48. Laboratories 1 and 3 noted a heteroplasmy[5] at base number 6849 (Anderson found an A at this site). Laboratory 1 found a G at this site, but closer examination of the electropherogram showed that an A peak existed under the G peak. Laboratory 3 also noted the A/G heteroplasmy at this site. Laboratory 4 did not note the heteroplasmy, but when their electropherograms were examined at NIST, the A/G heteroplasmy was noted. NIST did not have the electropherograms of Laboratory 2, but on questioning them, they agreed that the heteroplasmy was there, but that they had missed it. One of the problems with finding heteroplasmic sites is that if the computer call is the same as Anderson, one would not necessarily examine that site more closely. If the computer call is different from Anderson, one would look more closely at the electropherogram and then note the presence of a smaller peak under the main peak.

With the exceptions of the differences noted here, the interlaboratory evaluation was successful in that most of the laboratories found the same results. The many differences noted by Laboratory 4, which was less experienced at sequencing mtDNA, confirm and emphasize the need for a standard reference mate-

[5] In final preparation of this SRM, a new blood sample was obtained from CHR and a new cell culture line was established. The sequence analysis of the new CHR was identical to the first cell line except no heteroplasmy was found at bp 6849. Therefore, the cell line supplied with this SRM does not have this heteroplasmy. The data on the first cell line are included in the text to indicate the agreement in the interlaboratory evaluation that was done with the first cell line.

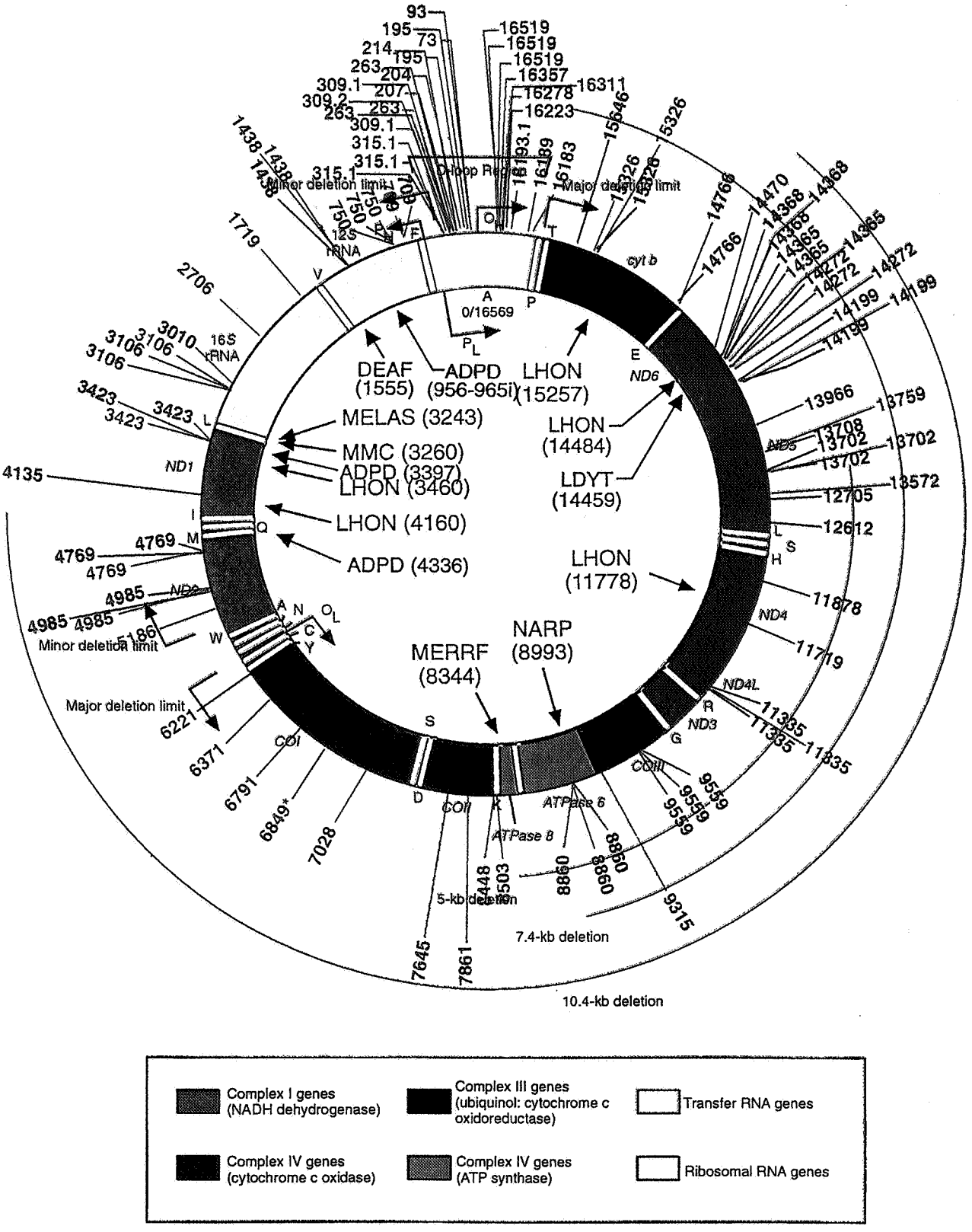


FIG. 2. Human mitochondrial DNA (color-coded circle) showing positions of genes, diseases, and deletions (Wallace, 1992; Wallace *et al.*, 1997) and the polymorphisms found in three apparently normal individuals examined for this study. The pink (CHR), green (9947A), and blue (GM03798) numbers and lines along the outside of the color-coded circle indicate the locations of differences from Anderson found in this study. Locations of disease base substitutions are shown on the inside of the circle. Abbreviations: ADPD, Alzheimer disease and/or Parkinson disease; DEAF, neurosensory hearing loss; LDYT, LHON plus dystonia; LHON, Leber hereditary optic neuropathy; MELAS, mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes; MERRF, myoclonic epilepsy and ragged-red fiber disease; MMC, myopathy and cardiomyopathy; NARP, neurogenic muscle weakness, ataxia, and retinitus pigmentosum. Large deletions are shown by the concentric black semicircles along the outside of the colored circle. tRNA genes (white areas in color-coded circle) are indicated by their amino acid single-letter abbreviation. Modified with permission of the publisher from Wallace (1992) where additional information on this figure may be found. 6849*, see legend to Table 2.

TABLE 3

Errors vs Polymorphisms in mtDNA Sequence Determined by Anderson

Basepair	Anderson designation → literature designation (No. found/No. examined)	Anderson designation → NIST mt SRM designation (No. found/No. examined)	Change
1438	A → G[a]	A→G (3/3)	12 s rRNA
3423	G → T (87/87)[b]	G→T (3/3)	SILENT
4769	A → G (28/30)[b]	A→G (3/3)	SILENT
4985	G → A (9/9)[b]	G→A (3/3)	SILENT
8860	A → G[a]	A→G (3/3)	Thr → Ala
11335	T → C (8/8)[b]	T→C (3/3)	SILENT
11719	G → A (26/37)[b]	G→A (1/3)	SILENT
12308	A → G (3/9)[b]	Change not found	$tRNA^{leu}$
13702	G → C (105/105)[b]	G→C (3/3)	Gly → Arg
14199	G → T (9/9)[b]	G→T (3/3)	Pro → Thr
14272	G → C (9/9)[b]	G→C (3/3)	Phe → Leu
14365	G → C (9/9)[b]	G→C (3/3)	SILENT
14368	G → C (9/9)[b]	G→C (3/3)	Phe → Leu
15326	A → G (6/6)[b]	A→G (3/3)	Thr → Ala

Note. Anderson *et al.* (1981) sequenced mtDNA mainly from a single human placenta, although some regions were from HeLa cells. Five ambiguous regions (bp 10, 934, 935, 14272, 14365) were assumed by Anderson to be same as that found in bovine mtDNA.

[a] Marzuki *et al.*, 1992.

[b] Howell *et al.*, 1992.

rial for sequencing mtDNA. If Laboratory 4 had the NIST mtDNA SRM 2392 and had run it alongside their unknown sample, they would have realized that they were finding an undue number of differences and could have reexamined their procedures to try to determine the reason for these excessive changes.

CONCLUSIONS

A NIST standard reference material (SRM 2392) that allows one to sequence any region or the entire 16,569 bp that comprise human mtDNA has been prepared. Fifty-eight pairs of unique primers have been designed, tested, and shown to work well in the amplification and sequencing procedures. The two DNA templates (CHR and 9947A) included in the SRM have characteristic polymorphisms throughout the noncoding and coding regions of the DNA and, therefore, can serve as positive controls during PCR amplification and sequencing. None of these polymorphisms correspond to any of the published basepair changes that have been correlated with specific diseases.

Compared to the Anderson sequence, CHR mtDNA had 13 differences in the noncoding regions and 33 differences in the coding regions, and the 9947A mtDNA had 9 differences in the noncoding regions and 23 differences in the coding regions. GM03798, whose data are included for comparison and information, had 4 differences in the noncoding regions and 19 differences in the coding regions. These differences in the coding regions do result in some amino acid changes in the proteins coded for by mtDNA. Four laboratories participated in an interlaboratory evaluation of the CHR template; some differences between laboratories were noted, but, in general, agreement was good. The use of NIST SRM 2392 will provide quality control to the scientific and medical communities when they sequence human mtDNA.

ACKNOWLEDGMENTS

The authors are grateful to The Bode Technology Group, Inc. (Sterling, VA), IIT Research Institute (Newington, VA), and Lark Technologies, Inc. (Houston, TX) for their participation in the interlaboratory evaluation. The authors also acknowledge the funding received from the National Institute of Justice through the NIST Office of Law Enforcement Standards and the support from the Advanced Technology Program at NIST. The authors thank Prasad Reddy for his help and advice on cloning, MaryPat Jones for confirming the sequences of some areas of the second CHR cells, and Catherine O'Connell for the use of her computer with the Gene Runner Program and her advice in the use of the program.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., deBrujin, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290:** 457–465.

Ballinger, S. W., Bouder, T. G., Davis, G. S., Judice, S. A., Nicklas, J. A., and Albertini, R. J. (1996). Mitochondrial genome damage associated with cigarette smoking. *Cancer Res.* **56:** 5692–5697.

Bendall, K. E., and Sykes, B. C. (1995). Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am. J. Hum. Genet.* **57:** 248–256.

Bogenhagen, D., and Clayton, D. A. (1974). The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. *J. Biol. Chem.* **249:** 7991–7995.

Grossman, L. I. (1995). Mitochondrial mutations and human disease. *Environ. Mol. Mutagen.* **25**(Suppl. 26): 30–37.

Holland, M. M., Fisher, D. L., Mitchell, L. G., Rodriquez, W. C., Canik, J. J., Merril, C. R., and Weedn, V. W. (1993). Mitochondrial DNA sequence analysis of human skeletal remains: Identification of remains from the Vietnam War. *J. Forensic Sci.* **38:** 542–553.

Holland, M. M., Fisher, D. L., Roby, R. K., Ruderman, J., Bryson, C., and Weedn, V. W. (1995). Mitochondrial DNA sequence analysis of human remains. *Crime Lab. Dig.* **22:** 109–115.

Howell, N., McCullough, D. A., Kubacka, I., Halvorson, S., and Mackey, D. (1992). The sequence of human mtDNA: The question of errors versus polymorphisms. *Am. J. Hum. Genet.* **50:** 1333–1337.

King, M. P., and Attardi, G. (1989). Human cells lacking mtDNA: Repopulation with exogenous mitochondria by complementation. *Science* **246:** 500–503.

Levin, B. C., Cheng, H., O'Connell, C., Reeder, D. J., and Holland, M. (1995). Evaluation of alternative enzymes and additives to improve sequencing of mitochondrial DNA. Proc. Fifth Int. Symp. on Human Identification, Scottsdale, Arizona, October 8–12, 1994, p. 170.

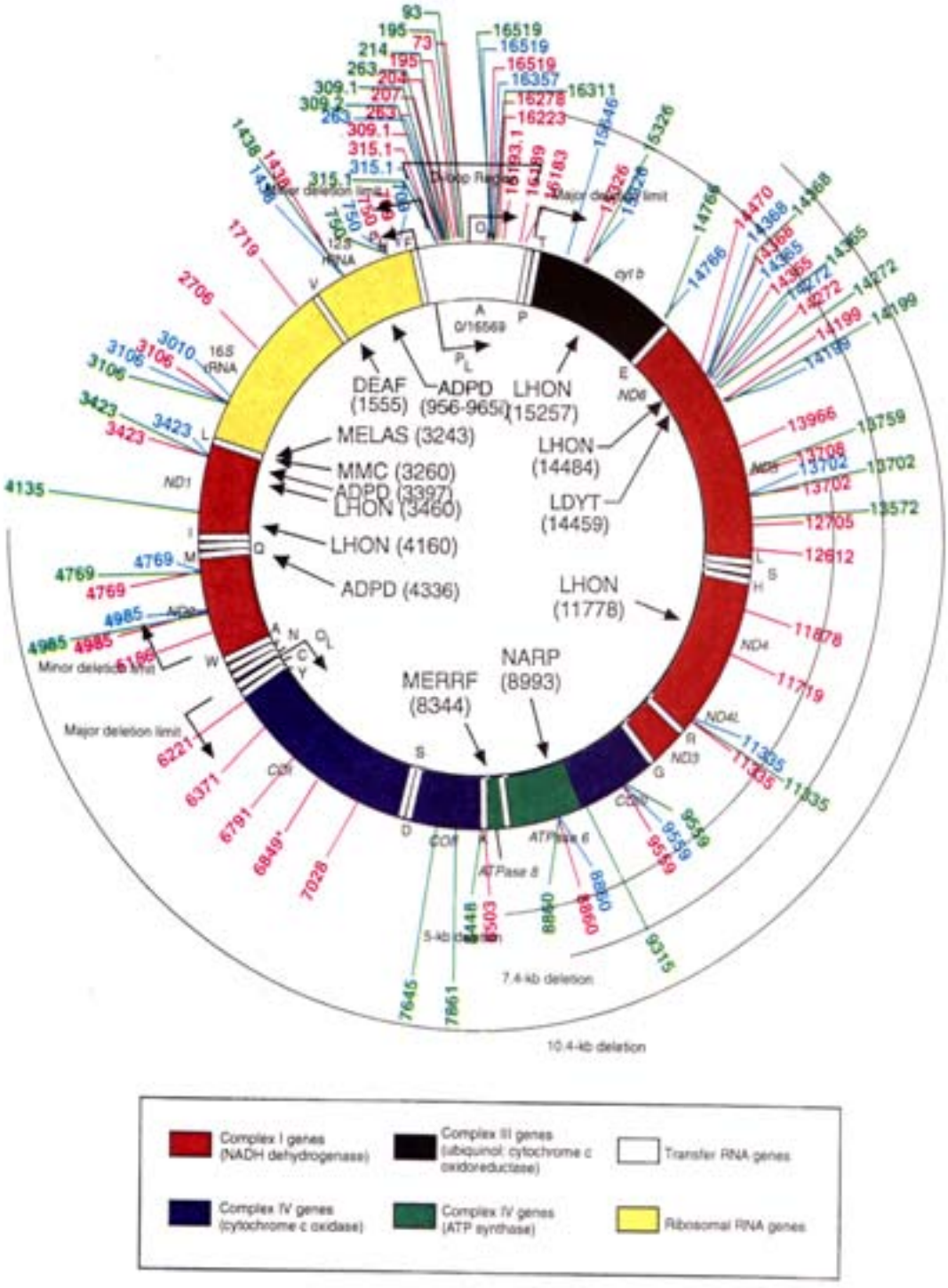
Levin, B. C., Cheng, H., Holland, M., and Reeder, D. J. (1997). Heteroplasmy responsible for difficulty experienced in sequencing the human mitochondrial DNA HV1 region containing the C-stretch. Proc. Seventh Int. Symp. on Human Identification, Scottsdale, Arizona, September 19–21, 1996, p. 166.

Marzuki, S., Lertrit, P., Noer, A. S., Kapsa, R. M. I., Sudoyo, H., Byrne, E., and Thyagarajan, D. (1992). Reply to Howell *et al.*: The need for a joint effort in the construction of a reference data base for normal sequence variants of human mtDNA. *Am. J. Hum. Genet.* **50:** 1337–1340.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Wallace, D. C. (1992). Mitochondrial genetics: A paradigm for aging and degenerative diseases. *Science* **256:** 628–632.

Wallace, D. C., Brown, M. D., and Lott, M. T. (1997). Mitochondrial genetics. *In* "Emery and Rimoin's Principles and Practice of Medical Genetics" (D. L. Rimoin, J. M. Connor, and R. E. Pyeritz, Eds.), 3rd ed., Vol. 1, pp. 277–332, Churchill Livingstone, New York.


DEAF (1555)
ADPD (956-965)
LHON (15257)
MELAS (3243)
MMC (3260)
ADPD (3397)
LHON (3460)
LHON (4160)
ADPD (4336)
LHON (14484)
LDYT (14459)
LHON (11778)
NARP (8993)
MERRF (8344)
Major deletion limit
Minor deletion limit
5-kb deletion
7.4-kb deletion
10.4-kb deletion
Complex I genes (NADH dehydrogenase)
Complex III genes (ubiquinol cytochrome c oxidoreductase)
Transfer RNA genes
Complex IV genes (cytochrome c oxidase)
Complex V genes (ATP synthase)
Ribosomal RNA genes

Appendix D

Journal Article:

B. C. Levin, K. A. Holland, D. K. Hancock, M. Coble, T. J. Parsons, L. J. Kienker, D. W. Williams, MP. Jones, and K. L. Richie, "Comparison of the Complete mtDNA Genome Sequences of Human Cell Lines - HL-60 and GM10742A - From Individuals with Pro-Myelocytic Leukemia and Leber Hereditary Optic Neuropathy, Respectively, and the Inclusion of HL-60 in the NIST Human Mitochondrial DNA Standard Reference Material - SRM 2392-I," Mitochondrion **2**, 387-400, (2003).

(See end of Appendix D for color representation of Figure 1.)

ELSEVIER

Mitochondrion 2 (2003) 387–400

Mitochondrion

www.elsevier.com/locate/mito

Comparison of the complete mtDNA genome sequences of human cell lines – HL-60 and GM10742A – from individuals with pro-myelocytic leukemia and leber hereditary optic neuropathy, respectively, and the inclusion of HL-60 in the NIST human mitochondrial DNA standard reference material – SRM 2392-I☆☆☆★

Barbara C. Levin[a,*], Koren A: Holland[a,b], Diane K. Hancock[a], Michael Coble[c,d], Thomas J. Parsons[c], Laura J. Kienker[e], Diana W. Williams[f], MaryPat Jones[g], Kristy L. Richie[a]

[a]*National Institute of Standards and Technology, 100 Bureau Drive, Mail Stop 8311, Gaithersburg, MD 20899-8311, USA*
[b]*Gettysburg College, Gettysburg, PA, USA*
[c]*Armed Forces DNA Identification Laboratory, Rockville, MD, USA*
[d]*The George Washington University, Washington, DC, USA*
[e]*Federal Bureau of Investigation Laboratory, Quantico, VA, USA*
[f]*Georgia Bureau of Investigation, Decatur, GA, USA*
[g]*National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA*

Received 12 December 2002; received in revised form 26 January 2003; accepted 28 January 2003

☆ This paper is a contribution of the U.S. National Institute of Standards and Technology (NIST), Armed Forces DNA Identification Laboratory (AFDIL), Federal Bureau of Investigation (FBI), the Georgia Bureau of Investigation (GBI) and the National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH) and is not subject to copyright. Certain commercial equipment, instruments, materials, or companies may be identified in this paper to specify the experimental procedure. Such identification does not imply recommendation or endorsement by the authors, nor does it imply that the materials or equipment identified are the best available for this purpose.

☆☆ Orders and requests for information concerning this SRM should be directed to the Standard Reference Materials Program, National Institute of Standards and Technology, 100 Bureau Drive, Stop 3222, Gaithersburg, MD 20899-2322, USA. Tel.: +1-301-975-6776; fax: +1-301-948-3730. E-mail: srminfo@nist.gov, web address http://www.nist.gov/srm.

★ Some of the data in this paper has been presented at the following meetings: The 12th International Symposium on Human Identification, Biloxi, MS, October, 2001; the American Academy of Forensic Sciences Annual Meeting, Atlanta, GA, February, 2002; the Third Annual Biotechnology Retreat, Shepherdstown, WV, May, 2002; the United Mitochondrial Disease Foundation Mito Dallas 2002 Symposium and Mitochondrial Standards Workshop, Dallas, TX, June, 2002; the Third Annual DNA Grantees' Workshop, National Institute of Justice, June, 2002; and the 13th International Symposium on Human Identification, Phoenix, AZ, October, 2002.

* Corresponding author. Tel.: +1-301-975-6682; fax: +1-301-975-8505.
E-mail address: barbara.levin@nist.gov (B.C. Levin).

doi:10.1016/S1567-7249(03)00010-2

Abstract

Forensic and clinical laboratories benefit from DNA standard reference materials (SRMs) that provide the quality control and assurance that their results from sequencing unknown samples are correct. Therefore, the mitochondrial DNA (mtDNA) genome of HL-60, a promyelocytic leukemia cell line, has been completely sequenced by four laboratories and will be available to the forensic and medical communities in the spring of 2003; it will be called National Institute of Standards and Technology (NIST) SRM 2392-I. NIST human mtDNA SRM 2392 will continue to be available and includes the DNA from two apparently healthy individuals. Both SRM 2392 and 2392-I contain all the information (e.g. the sequences of 58 unique primer sets) needed to use these SRMs as positive controls for the amplification and sequencing any DNA. Compared to the templates in SRM 2392, the HL-60 mtDNA in SRM 2392-I has two tRNA differences and more polymorphisms resulting in amino acid changes. Four of these HL-60 mtDNA polymorphisms have been associated with Leber Hereditary Optic Neuropathy (LHON), one as an intermediate mutation and three as secondary mutations. The mtDNA from a cell line (GM10742A) from an individual with LHON was also completely sequenced for comparison and contained some of the same LHON mutations. The combination of these particular LHON associated mutations is also found in phylogenetic haplogroup J and its subset, J_2, and may only be indicative that HL-60 belongs to haplogroup J, one of nine haplogroups that characterize Caucasian individuals of European descent or may mean that haplogroup J is more prone to LHON. Both these mtDNA SRMs will provide enhanced quality control in forensic identification, medical diagnosis, and single nucleotide polymorphism detection.

Keywords: Forensic identification; GM10742A; Haplogroup J; HL-60; Leber hereditary optic neuropathy (LHON); Medical diagnosis; Mitochondrial DNA sequence; Single nucleotide polymorphism (SNP); Standard reference material (SRM)

1. Introduction

On July 15, 1998, the Federal Bureau of Investigation (FBI) Director signed Standard 9.5, which stated "The laboratory shall check its DNA procedures annually or whenever substantial changes are made to the protocol(s) against an appropriate and available National Institute of Standards and Technology (NIST) standard reference material or standard traceable to a NIST standard". At the present time, there are a number of human DNA standard reference materials (SRMs) available from NIST (Levin et al., 2001). One of these SRMs on mitochondrial DNA (mtDNA SRM 2392) is used by the forensic community for human identification and by the medical community for diagnoses of a number of human diseases now known to be associated with specific mutations, insertions and deletions of mtDNA (Levin et al., 1999). This SRM was prepared by NIST to provide quality control to the scientific community when they amplify and sequence human mtDNA or any DNA and includes DNA from two apparently healthy individuals (CHR and 9947A) plus cloned DNA of an area in CHR containing a C-Stretch beyond which sequencing becomes difficult. All the information necessary to successfully conduct the polymerase chain reaction (PCR) amplification process, cycle sequencing steps, gel separation, and data analysis to obtain the final DNA sequence is provided as well as the information on the sequence of 58 unique primer sets that allow the sequencing of any specific portion or the entire mtDNA (16,569 bp) without any gaps. Following an interlaboratory evaluation, SRM 2392 became available to the public in December 1999.

The FBI acknowledges the utility of DNA SRMs to provide the quality control and quality assurance that the results from forensic laboratories that are sequencing unknown samples are correct. One of the FBI's Combined DNA Index Systems now includes mtDNA from unidentified remains, as well as from relatives of missing persons. In order for authorized laboratories to contribute to these indices, certain quality standards must be met. These include the use of DNA from the human cell line HL-60 as a positive control to be run in conjunction with the unknown samples. HL-60 was chosen as the positive control because of several features present in HL-60 but not in the cell lines currently available in NIST SRM 2392. Some of the advantageous features of HL-60 are well-spaced polymorphisms throughout the non-coding hypervariable regions, HV1 and HV2, of the mtDNA control

region, and no insertions at the HV2 C-stretch area (positions 303–309). One of the current templates in SRM 2392, the CHR DNA has a C-stretch in the HV1 region caused by a T to C change at position 16189 that produces a length heteroplasmy (Bendall and Sykes, 1995; Butler and Levin, 1998). Sequencing through this C-stretch is difficult and time-consuming and results in the need to perform additional sequencing reactions to resolve this region and the area following the C-stretch. The CHR template was chosen by NIST specifically for the C-stretch region since some laboratories wanted the opportunity to address this difficult sequencing problem and try to resolve it. The DNA from the other current SRM DNA template, 9947A, has only two polymorphisms in the HV1 region with respect to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999) and those polymorphisms are at common sites. For the work that the forensic laboratories are doing on human identification, several evenly spaced polymorphisms within the HV1 region are more desirable to differentiate the positive control from the test sample. Thus, the FBI suggested to NIST that the development of a SRM containing the HL-60 DNA template would be of great utility to the forensic community. All of the research and the interlaboratory evaluation of the HL-60 template necessary to produce this new SRM containing HL-60 have been completed and are described in this paper. Both SRM 2392 (containing two DNA templates) and SRM 2392-I (containing one DNA template) should fulfill the needs of forensic laboratories by providing additional quality control when sequencing human mtDNA. Corroboration of the SRM results will provide assurance that the techniques being used for amplification and sequencing unknown DNA are being conducted correctly.

The DNA templates in SRM 2392 came from apparently healthy individuals; however, HL-60 in SRM 2392-I is from a promyelocytic cell line from peripheral blood leukocytes provided by a 36-year-old Caucasian female with acute promyelocytic leukemia. Therefore, it was of interest to determine if the HL-60 DNA contained specific mutations that may be characteristic of this disease. Although more DNA samples from leukemic patients would have to be examined to associate specific mutations with this disease, we did find four polymorphisms that are considered either intermediate or secondary mutations associated with leber hereditary optic neuropathy (LHON) (Wallace et al., 1997). We, therefore, sequenced the mtDNA from cell line GM10742A that was prepared from a patient with LHON for comparison with the HL-60 mtDNA sequence. Two of the four LHON-associated mutations in HL-60 were also found in GM10742A. Two additional mutations associated with LHON were found in GM10742A, one primary and one secondary mutation. Many of the polymorphisms common to both HL-60 and GM10742A (some of which are associated with LHON) characterize the phylogenetic haplogroup J and sub-haplogroup J_2, a subset of J (Finnilä et al., 2001; Helgason et al., 2001; Herrnstadt et al., 2002). Haplogroup J is one of nine haplogroups that characterize Caucasian populations of European decent. A number of papers in the literature indicate that individuals with LHON usually belong to haplogroup J and suggest that those in haplogroup J may therefore be more prone to this mtDNA disease than those in other haplogroups (Brown et al., 2002; Herrnstadt et al., 2002; Torroni et al., 1997).

2. Materials and methods

2.1. HL-60 DNA

DNA from the HL-60 cell culture was extracted, isolated, and quantified by the Professional Services Department of the American Type Culture Collection (ATCC, Manassas, VA).

DNA was isolated from the HL-60 cell culture using the QIAamp DNA Blood Mini Kit (Qiagen, Inc., Valencia, CA). Quantification was determined by the Quantiblot Human DNA Quantification Kit (Applied Biosystems, Foster City, CA). The final concentration, 1.4 ng/μl, is based on 12 replicate tests.

The DNA purity (A260:A280 = 1.9) was determined spectrophotometrically. The integrity of DNA was determined electrophoretically using 0.4% agarose gels.

2.1.1. GM10742A DNA

The lymphoblast cell line GM10742A was purchased from NIGMS Human Genetic Mutant Cell Repository, Coriell Institute of Medical Research,

Camden, NJ. The cells were grown at NIST in RPMI 1640 plus L-glutamine and sodium bicarbonate growth media (Sigma, St. Louis, MO) plus fetal calf serum (20%) (Sigma) plus the antibiotics Streptomycin and Penicillin (final concentration: 100 U/ml) (Sigma). DNA was extracted using the QIAGEN Plasmid Kit following the Plasmid Mini Purification Protocol.

2.1.2. mtDNA primers

Fifty-eight sets of unique primers (19–28 bp) for sequencing any portion or the entire human mtDNA (16,569 bp), including both the HV1 and HV2 regions, were computer-designed using GENE RUNNER FOR WINDOWS (Hastings Software, Inc., Hastings, NY) and were the same as those used for SRM 2392 (Levin et al., 1999) with the exception of the reverse primer of set 51. During the course of this study, this primer was changed since it contained a C at nucleotide position (np) 14,368. Since this is the reverse primer, it would bind to a G at np 14,368. Andrews et al. (1999) in their reevaluation of the placenta originally used to sequence human mtDNA in 1981 (Anderson et al., 1981) found that np 14,368 should have a C at that position. The new reverse primer 51 is 5′-TTAGCGATGGAGGTAGGATT**G**G-3′ (np 14,368 is in bold and underlined) which binds to the C at position 14,368. The 5′ end is np 14,388 and the 3′ end is 14,367. The 58 sets of primers were custom-made by Bio-Synthesis, Inc. (Lewisville, TX); the new reverse primer 51 was obtained from Invitrogen (Carlsbad, CA). Those laboratories using either SRM 2392 or SRM 2392-I should use the new reverse primer 51.

2.1.3. Polymerase chain reaction (PCR) at NIST

The PCR mixture contained: HL-60 DNA (1 μl; 1.4 ng), AmpliTaq Gold® DNA polymerase, (0.5 μl; 2.5 units) (Applied Biosystems), 10 × buffer (5 μl) containing 100 mmol/l Tris–HCl, pH 8.3, 500 mmol/l KCl, 15 mmol/l $MgCl_2$ and 0.01% (w/v) gelatin (Applied Biosystems), 10 mmol/l dNTP mix (1 μl) (Invitrogen), 10 μmol/l forward and reverse primers (1 or 2 μl), plus water to make a final volume of 50 μl.

Thermal cycling was conducted in a PerkinElmer thermocycler Model 9700 and started with 10 min at 95°C, followed by 35 cycles of 20 s at 94°C (denaturation), 10 s at 56°C (annealing), 30 s at 72°C (extension) and ended with a final extension of 7 min at 72°C.

Amplified DNA was purified with a QIAquick PCR Purification Kit (Qiagen) and the purity and size of the PCR product was determined by electrophoresis in 2% agarose gels in 1 × TBE buffer containing 0.5 μg/ml ethidium bromide (Sigma).

2.1.4. Sequencing at NIST

Cycle sequencing using fluorescent dye-labeled terminators was performed with an ABI PRISM® BigDye® Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (Applied Biosystems).

Cycle sequencing reactions in both the forward and reverse modes were conducted with a 9700 PerkinElmer thermal cycler and started with 1 min at 96°C. The reaction then underwent 25 cycles of 96°C for 15 s (denaturation), 50°C for 5 s (annealing), and 60°C for 2 min (extension). The DNA products were purified using Edge Gel Filtration Cartridges (Edge BioSystems, Gaithersburg, MD).

Sequencing and data analysis of the purified DNA were performed using an Applied Biosystems PRISM® Model 310 Genetic Analyzer with POP-6™ polymer system and 47 cm × 50 μm capillaries (Applied Biosystems). Sequence data were analyzed with Sequencing Analysis Software 3.3, and comparisons to the Cambridge Reference Sequence (1981) were performed with Sequence Navigator Software 1.01.

2.2. Interlaboratory evaluation of HL-60

Three laboratories, in addition to NIST, participated in an interlaboratory evaluation of the sequence of HL-60 DNA. These laboratories included the FBI Laboratory, FBI Academy, Quantico, VA 22135; Armed Forces DNA Identification Laboratory (AFDIL), Armed Forces Institute of Pathology, Rockville, MD 20850; and the Georgia Bureau of Investigation (GBI), Decatur, GA 30034. Each laboratory was asked to amplify and sequence the entire mtDNA of HL-60. NIST provided: (1) a tube of DNA containing the extracted DNA ready for PCR amplification; (2) the 58 sets of primers labeled with either F# (forward primer) or R# (reverse primer); (3) a table to record the results; and (4) any other supplies

that were needed and requested by the participants. They were allowed to use any protocol for amplification or sequencing that they wished, but were requested to provide a copy of that protocol to NIST. NIST also requested copies of the electropherograms to enable us to resolve any discrepancies; although as it turned out, there were no discrepancies.

2.2.1. Differences in methodology used by the laboratories in the interlaboratory evaluation

2.2.1.1. Armed forces DNA identification laboratory (AFDIL). AFDIL has developed a high-throughput, automated sequencing procedure using 12 primer sets that produce overlapping PCR products ranging from 825 to 1886 bp. The primers used to amplify the first 11 products were based on those published in Levin et al. (1999) and are: Amp01 – F361/R2216; Amp02 – F1993/R3557; Amp03 – F3441/R4983; Amp04 – F4797/R6526; Amp05 – F6426/R8311; Amp06 – F8164/R9848; Amp07 – F9754/R11600; Amp08 – F11403/R13123; Amp09 – F12793/R14388; Amp10 – F14189/R15396; Amp11 – F15260/R16084. The primers used to amplify the control region were developed at AFDIL and are Amp12 – F15878/R649 (F15878 is TTAACTCCACCATTAGCACC and R649 is TTTGTTTATGGGGTGATGTGA).

2.2.1.1.1. AFDIL PCR amplification. The PCR mixture contained HL-60 DNA (1 μl), AmpliTaq-Gold® DNA polymerase (1 μl) (Applied Biosystems), 10 × PCR buffer (5 μl) (Applied Biosystems), dNTP's (0.2 mmol/l) (Invitrogen), 2 μl of forward and reverse primers (10 μmol/l) (MWG Biotech, High Point, NC) plus dH_2O to a final volume of 50 μl. The 10 × PCR buffer was the same as that used by NIST. Thermal cycling was conducted in a PerkinElmer 9700 thermocycler with the following conditions: 10 min at 96°C (activation of AmpliTaq Gold®), plus 40 cycles of 94°C for 15 s, 56°C for 30 s, and 72°C for 1 min. The purity and size of the PCR products were assessed by electrophoresis in a 0.7% agarose gel containing 0.3 μg/ml of ethidium bromide. The PCR products were purified with Shrimp Alkaline Phosphatase/Exonuclease I (Amersham Pharmacia, Piscataway, NJ). Five μl of exonuclease I (10 U/μl) and 10 μl of Shrimp Alkaline Phosphatase (1 U/μl) was added to each tube containing PCR product. The tubes were heated at 37°C for 15 min followed by 94°C for 15 min in a PerkinElmer 9700 thermocycler.

2.2.1.1.2. AFDIL sequencing. Cycle sequencing was performed with the ABI PRISM® BigDye® terminators (original version) cycle sequencing kit (Applied Biosystems). The sequencing mixture contained 9 μl dH_2O, 6 μl BigDye® dilution buffer (400 mmol/l TRIS, 10 mmol/l $MgCl_2$, pH 9.0), 2 μl BigDye® terminator reaction mixture, 1 μl of forward or reverse primer (10 μmol/l each) and 2 μl of HL-60 PCR product for a total vol of 20 μl. A few of the sequencing primers (e.g. R649) required the use of the ABI PRISM® dGTP BigDye® terminator kit (Applied Biosystems). Thermal cycling was conducted in a PerkinElmer 9700 thermocycler at the following conditions: an initial 1 min denaturation at 96°C, followed by 25 cycles of 94°C for 15 s, 50°C for 5 s, and 60°C for 2 min. The DNA product was purified by filtration through a spin column matrix (Edge BioSystems, Gaithersburg, MD). Electrophoresis and sequencing were performed with an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems) using POP-6™ polymer (Applied Biosystems) with a 50 cm capillary. Data analysis was executed using Sequencher Plus 4.0.5b11 (Gene Codes, Ann Arbor, MI). The HL-60 sequence differences were identified by comparison to the Cambridge Reference Sequence as revised by Andrews et al. (1999). In most cases, sequence information was acquired for both the forward and reverse directions. In some regions, two separate reactions using the same primer were routinely conducted (indicated by '2 ×' in the following list of primers) to achieve full sequence confirmation. A total of 95 sequencing reactions plus one pGEM reaction were conducted in a 96 well format. If the primer failed the first trial, the reaction was repeated. The finding of a heteroplasmy at nucleotide position (np) 12,071 was also confirmed by an additional PCR and sequencing reaction.

The following primers from Levin et al. (1999) were used to sequence the 12 PCR amplicons:

Amp01 (F361/R2216): F361, R921, F1234, R1425, F873, R2216, F1657, R1769
Amp02 (F1993/R3557): F1993, R2660, R2834, R3557, F2417, R3006, F3234
Amp03 (F3441/R4983): F3441, R3940, F3931 (2 ×), R4982, F4392, R4162

Amp04 (F4797/R6526): F4797 (2 ×), R6526, F5700 (2 ×), F5318, R5882, F6242
Amp05 (F6426/R8311): F6426 (2 ×), R7255, F7645 (2 ×), R8311, F7075, R7792
Amp06 (F8164/R9848): F8164 (2 ×), R9059, F8903, R9848, F8539, R9403, F9309
Amp07 (F9754/R11600): F9754 (2 ×), R10556, F11001 (2 ×), R11600, F10386, R11267
Amp08 (F11403/R13123): F11403 (2 ×), F12357, R13123, F11901 (2 ×), F12601, R12876
Amp09 (F12793/R14388): F12793, R13611, F13518 (2 ×), R14388, F13188, R13343, F13899, R13935
Amp10 (F14189/R15396): F14189 (2 ×), R15396, F14909, R14996, F14470
Amp11 (F15260/R16084): F15260, R16084, F15574, R15774
Amp12 (F15878/R649): F15971, R16175 (2 ×), F16450 (2 ×), R274, F314 (2 ×), R649 (2 ×), F16190, R16400.

In Amp12, F15971 came from Levin et al. (1999). The other primers were designed by AFDIL and were as follows:

R16175: TGGATTGGGTTTTTATGTA
F16450: GCTCCGGGCCCATAACACTTG
R274: TGTGTGGAAAGTGGCTGTGC
F314: CCGCTTCTGGCCACAGCACT
R649: TTTGTTTATGGGGTGATGTGA
F16190: CCCCATGCTTACAAGCAAGT
R16400: GTCAAGGGACCCCTATCTGA.

2.2.1.2. Georgia Bureau of Investigation (GBI). The GBI used the same protocol as that used by NIST with the following exceptions: (1) a PerkinElmer model 9600 was used for thermal cycling; (2) amplified DNA was electrophoresed in 2.75% agarose gels; (3) the cycle sequencing was performed with a Perkin-Elmer model 9600 using the program provided with the BigDye® kit (25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min); (4) some samples that needed to be cycle sequenced again using more amplicon were not purified before precipitation; (5) the samples were precipitated using isopropanol precipitation as per the BigDye® instructions; and (6) comparison of the sequence data was performed with Sequencer 3.1.1 software (Gene Codes, Ann Arbor, MI).

2.2.1.3. Federal Bureau of Investigation (FBI)

2.2.1.3.1. FBI Polymerase chain reaction (PCR). The PCR mixture contained from 0.1 to 1.4 ng HL-60, AmpliTaq Gold® Polymerase (2.5 units) (Applied Biosystems), 10 × PCR buffer (5 μl) (Applied Biosystems), GeneAmp® dNTPs (0.2 mmol/l each) (Applied Biosystems), forward and reverse primers (0.4 μmol/l each), plus dH_2O to a final volume of 50 μl. The 10 × buffer (pH 8.3) was the same as used by NIST. Thermal cycling was conducted in a GeneAmp® PCR System 9700 (PerkinElmer) and consisted of 10 min at 95°C followed by 35 cycles of 94°C for 20 s, annealing temperatures of 50°C (primer set 49), 51°C (primer sets 1, 7, 44, 57), 52°C (primer sets 6, 8, 30, 45), 53°C (primer set 53) and 56°C (all other primer sets) for 10 s, and 72°C for 30 s and ending with a final extension of 7 min at 72°C. Amplified products were purified by treatment with Exo-SAP-IT (5 μl for every 25 μl of PCR product) (USB Corp., Cleveland, OH). Samples of the PCR products were electrophoresed in 1.2% agarose gels containing ethidium bromide to assess the purity, size, and quantity of the PCR products.

2.2.1.3.2. FBI sequencing. Cycle sequencing using fluorescent dye–labeled terminators was performed with an ABI PRISM® dRhodamine Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (Applied Biosystems). Thermal cycling was conducted in a GeneAmp® PCR System 9700 (Applied Biosystems) and started with 1 min at 96°C followed by 25 cycles of 96°C for 15 s, 50°C for 1 s and 60°C for 1 min and ended with a final incubation at 15°C for 10 min. These products were purified with Centri-Sep™ spin columns (Princeton Separation, Inc., Adelphia, NJ) and electrophoresed on an ABI PRISM® 310 Genetic Analyzer (Applied Biosystems) using POP-6™ and either a 47 or 61 cm capillary. Comparison of the sequence data was performed with Sequencer 4.1 software (Gene Codes).

The FBI also discovered the problem with original primer set 51 (see Section 2). The reverse primer incorporated a G at position 14,368 [an error in the original Anderson et al. (1981) sequence that was found and corrected by Andrews et al. (1999)].

However, sequencing with primer set 52 showed a C at position 14,368. They prepared a new set of primers (primer set 51.5) to amplify approximately a 400 bp fragment that encompassed np 14,368. The new primers were: F14217: 5′-CTAATCAACGCCCA-TAATCATAC-3′ and R14620: 5′-GTTTTCTTC-TAAGCCTTCTCC-3′. The new primer set confirmed that the correct base at position 14,368 was a C.

2.3. Permissions

The research to prepare SRM 2392-I containing HL-60 DNA was deemed exempt from the policy of Part 27 of Title 15 of the Code of Federal Regulations by the NIST Institutional Review Board and the Director of the Chemical Science and Technology Laboratory. This work fit into the exemption category described in 15 CFR 27.101(b)(4) which states: "Research, involving the collection or study of existing data, documents, pathological specimens, or diagnostic specimens, if, these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects". ATCC also waived condition 3(c) in their Material Transfer Agreement which states that the "purchaser shall not sell, lend, distribute or otherwise transfer the material or replicates to any others" for the use of HL-60 in the NIST mitochondrial DNA SRM 2392-I. They stated that, in their view, "as a government agency, NIST will not be providing this material as a commercial product despite the collection of fees for the SRM".

3. Results and discussion

HL-60 is a promyelocytic cell line from the peripheral blood leukocytes of a Caucasian female (age 36) with acute promyelocytic leukemia. HL-60 DNA will be available in SRM 2392-I. The two DNA templates available in SRM 2392 are from apparently healthy individuals (both the DNA from CHR and 9947A are from Caucasian females in their twenties and thirties). The entire mtDNA of GM03798 (a 10-year-old apparently healthy Caucasian male) was amplified and sequenced at NIST, discussed in a previous publication (Levin et al., 1999), and is shown in Fig. 1 in this paper; however, that DNA is not included in SRM 2392 or SRM 2392-I. Another mtDNA template (GM10742A) from a 30-year-old Caucasian male patient with LHON was also sequenced at NIST; that sequence is shown in Table 1 and Fig. 1, but the DNA is not included in SRM 2392 or 2392-I. Both GM03798 and GM10742A can be obtained from the National Institute of General Medical Sciences Human Genetic Mutant Cell Repository, Coriell Institute for Medical Research, Camden, NJ. At NIST, 58 primer sets (Levin et al., 1999) were used to amplify the entire HL-60 mtDNA and GM10742A at least twice. In the case of HL-60, all the PCR products were sequenced in both the forward and reverse directions for a total of four sequences for each amplicon. During this study, reverse primer 51, which contained np 14,368, was changed (see Section 2) to accommodate the difference found at that position by Andrews et al. (1999) when they resequenced the original placenta used by Anderson et al. (1981). Any sequence ambiguities between experiments at NIST were resolved by additional PCR reactions and sequencing.

The NIST sequencing results from the mtDNA of HL-60, CHR, 9947A and GM10742A are shown in Table 1; all of these sequences plus GM03798 are shown in Fig. 1. Table 1 also compares the CHR, 9947A, HL-60 and GM10742A templates to the original Cambridge Reference Sequence (Anderson et al., 1981) and the revised sequence (Andrews et al., 1999). In Fig. 1, the HL-60 results are shown in black numbers surrounding the circular depiction of human mtDNA; the CHR results are shown in red; 9947A results are shown in green; GM03798 results are shown in blue; and GM10742A are in purple. At the present time, GM03798 and GM10742A are not part of SRM 2392 or SRM 2392-I, and the results are presented for information and comparison only.

The numbering system in Table 1 and Fig. 1 is that of the original Cambridge Reference Sequence (1981). The Cambridge Reference Sequence (1981) was based on a consensus analysis of a placenta, the HeLa cell line and the bovine sequence (the bovine sequence was used in five ambiguous human sites). The 1981 results were reexamined in 1999 by Andrews et al. who resequenced the original placenta. The original Cambridge Reference Sequence was

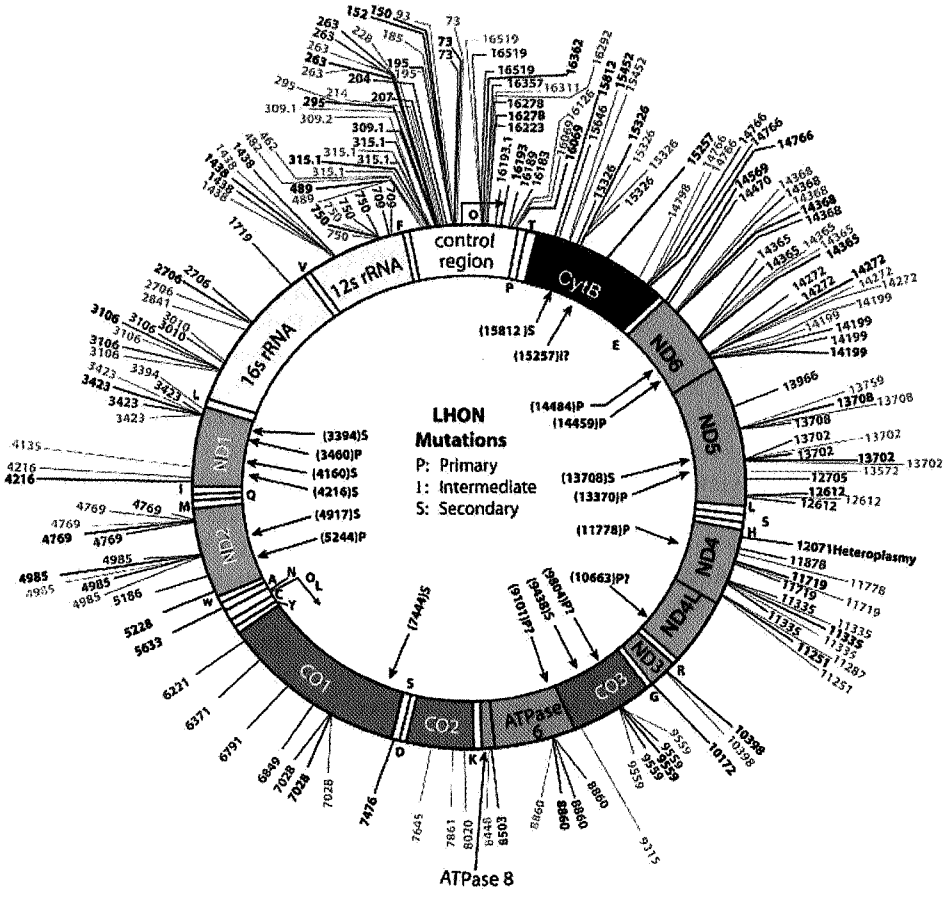


Fig. 1. HL-60 DNA polymorphic differences from the Cambridge Reference Sequence and positions of LHON mutations are added to the figure from Levin et al., 1999. This figure is a schematic of human mtDNA showing its circular double-stranded DNA and all the differences from Cambridge Reference Sequence found in CHR (red), 9947A (green), GM03798 (blue), HL-60 (black), and GM10742A (purple) as numbers along the outside of the color-coded circle. Locations of the control region, rRNAs and genes (see footnote to Table 1 for abbreviations) coded by human mtDNA are shown. The locations of the 22 tRNAs are noted by white areas in the circle and designated by their single letter amino acid code. (Modified from Levin et al., 1999).

found to have a number of rare polymorphisms and errors, which are noted in Table 1 with an asterisk.

Table 1 also shows the unique (U) coding region polymorphisms (either silent with no amino acid change or resulting in an amino acid change) in CHR, 9947A, HL-60 and GM10742A that were not found in the MITOMAP web site database (http://www.mitomap.org). This is a very extensive internet database on the mitochondrial DNA polymorphisms found in the literature. CHR has five polymorphisms (one produces an amino acid change), 9947A has four polymorphisms (two producing amino acid changes), HL-60 has two polymorphisms (one producing an amino acid change and is heteroplasmic), and GM10742A has two (one in the 16sRNA and one that is silent) that were not found in the MITOMAP database.

Table 2 shows the number of differences compared to the Cambridge Reference Sequence found in CHR, 9947A, HL-60 and GM10742A. There are 13, 9, 11 and 12 differences in the non-coding regions of CHR, 9947A, HL-60, and GM10742A, respectively. The non-coding region is the primary area used by the forensic community for human identification. There

are 33, 23, 33 and 31 differences in the coding regions of CHR, 9947A, HL-60, and GM10742A, respectively. These differences would have resulted in 11, 12, 15 and 14 amino acid changes in their respective proteins if one compares the results to the original Cambridge Reference Sequence published by Anderson et al. (1981). If compared to the revised Cambridge Reference Sequence (Andrews et al., 1999), the actual amino acid changes would be reduced to 3 in CHR, 4 in 9947A, and 7 in both HL-60 and GM10472A (the non-normal cell lines sequenced). HL-60 also has a change in the tRNA for alanine and the tRNA for serine. It would be of interest to examine other cell lines from patients with acute promyelocytic leukemia or LHON to see if they have any of these same mutations.

Table 3 shows the HL-60 and GM10742A polymorphisms in the coding regions that (with the exception of one CHR mutation included because of its association with LHON) were not found in the other templates in SRM 2392. This table also indicates if the mutation has been associated with a disease according to the MITOMAP web site database. At least four HL-60 mutations (T4216C in NADH dehydrogenase 1, G13708A in NADH dehydrogenase 5, and G15257A and G15812A in Cytochrome B) were found to be associated with LHON (Wallace et al., 1997). Four mutations associated with LHON were also found in GM10742A; these are T3394C and T4216C in NADH dehydrogenase 1, G11778A in NADH dehydrogenase 4, and G13708A in NADH dehydrogenase 5. Two of these mutations, T4216C and G13708A, were found in both HL-60 and GM10742A. The mutation G13708A was also seen in the CHR template. LHON causes central vision loss in patients in their twenties or thirties although the onset of symptoms can occur both earlier and later. The disease is inherited like other mitochondrial DNA diseases through the maternal lineage. LHON was first associated with a G11778A mtDNA mutation (Wallace et al., 1988) that results in an arginine to histidine change in NADH dehydrogenase 4 at position 340 in a protein that contains 460 amino acids (Lee and Levin, 2002). G11778A is considered a primary mutation causing LHON and was seen in GM10742A, a cell line from a patient who experienced the sudden onset of blindness at the age of 24. Since 1988, 27 missense mutations and one mutation in tRNA for leucine have been associated with LHON (http://www.mitomap.org). Some of these mutations are shown inside the inner circle of Fig. 1. The four mutations considered 'primary' in causing LHON are G14459A, G11778A, G3460A, and T14484C and are presented in order of decreasing severity (Wallace et al., 1997). The mutations associated with LHON found in this study are considered 'primary' (G11778A is in GM10742A), 'intermediate' (G15257A is in HL-60) or 'secondary' (T3394C in GM10742A; T4216C in HL-60 and GM10742A; G13708A in HL-60, CHR and GM10742A; G15812A in HL-60). According to Wallace et al. (1997), the intermediate or secondary mutations may increase the probability of having LHON or may be linked to one of the primary mutations. It is interesting that GM10742A has two of the secondary mutations also found in HL-60, namely T4216C and G13708A. CHR also has the G13708A mutation. Torroni et al. (1997) found that the combination of the np 4216 and 13,708 mutations were more frequent among the LHON patients containing the mutation at 11,778 than among the controls.

Characteristic sets of polymorphisms in human mtDNA are being used to distinguish various groups and to trace their maternal genealogy (Macaulay et al., 1999). Haplogroup J and sub-haplogroup J_2 have specific polymorphisms that distinguish this group of Caucasians with European ancestry from the other eight haplogroups (H, I, K, T, U, V, W and X) that have also been found to characterize those with a European background (Torroni et al., 1997). Both HL-60 and GM10742A have a number of identical polymorphisms that place them in Haplogroup J and Sub-haplogroup J_2 (Table 4).

In the interlaboratory evaluation of the entire sequence of HL-60 that was conducted by three laboratories plus NIST, all four laboratories found the same sequence with the exception of the Georgia Bureau of Investigation who had the problem noted in Section 2 with primer set 51 and was unable to examine that amplicon and therefore, did not find the polymorphism at 14,199.

In conclusion, NIST has sequenced the entire mtDNA (16,569 bp) from the HL-60 cell line multiple times and compared these results with those from AFDIL, FBI, and GBI who participated in the interlaboratory evaluation. All four laboratories

Table 1
Cambridge Reference Sequence nucleotide differences (some unique) found at NIST in the two DNA templates, CHR and 9947A included in SRM 2392; HL-60 included in SRM 2392-I and GM10742A DNA[a]

Comparison with Cambridge Reference Sequence (CRS)							
CRS #[d]	Base 1981/1999[e]	Template CHR[b]	Template 9947A[b]	Template HL-60[c]	Template GM10472A	Amino acid change	Region
73	A	G	–[f]	G	G		HV2
93	A	–	G	–	–		HV2
150	C	–	–	T	–		HV2
152	T	–	–	C	–		HV2
185	G	–	–	–	A		HV2
195	T	C	C	–	–		HV2
204	T	C	–	–	–		HV2
207	G	A	–	–	–		HV2
214	A	–	G	–	–		HV2
228	G	–	–	–	A		HV2
263*R[g]	A	G	G	G	G		HV2
295	C	–	–	T	T		HV2
303–309	–	C (ins)	CC (ins)	–	–		HV2
311–315*R	–	C (ins)	C (ins)	C (ins)	C (ins)		HV2
462	C	–	–	–	T		HV2
482	T	–	–	–	C		HV2
489	T	–	–	C	C		HV2
709	G	A	–	–	–		12sRNA
750*R	A	G	G	G	G		12sRNA
1438*R	A	G	G	G	G		12sRNA
1719	G	A	–	–	–		16sRNA
2706	A	G	–	G	G		16sRNA
2841	T	–	–	–	A (U)		16sRNA
3010	G	–	–	–	A		16sRNA
3106–3107*E[h]	C/del	del	del	del	del		16sRNA
3394	T	–	–	–	C	Tyr → His	ND1 LHON
3423*E	G/T	T	T	T	T	Silent	ND1
4135	T	–	C (U)	–	–	Tyr → His	ND1
4216	T	–	–	C	C	Tyr → His	ND1 LHON
4769*R	A	G	G	G	G	Silent	ND2
4985*E	G/A	A	A	A	A	Silent	ND2
5186	A	G (U)	–	–	–	Silent	ND2
5228	C	–	–	G (U)	–	Silent	ND2
5633	C	–	–	T	–		tRNA Ala
6221	T	C	–	–	–	Silent	COI
6371	C	T (U)	–	–	–	Silent	COI
6791	A	G	–	–	–	Silent	COI
6849[i]	A	G(0.3A)[i]	–	–	–	Thr → Ala[i]	COI
7028	C	T	–	T	T	Silent	COI
7476	C	–	–	T	–		tRNA Ser
7645	T	–	C	–	–	Silent	COII
7861	T	–	C (U)	–	–	Silent	COII
8020	G	–	–	–	A	Silent	COII
8448	T	–	C	–	–	Met → Thr	ATPase 8
8503	T	C	–	–	–	Silent	ATPase 8
8860*R	A	G	G	G	nd	Thr → Ala	ATPase 6
9315	T	–	C (U)	–	–	Phe → Leu	COIII
9559*E	G/C	C	C	C	C	Arg → Pro	COIII
10,172	G	–	–	A	–	Silent	ND3
10,398	A	–	–	G	G	Thr → Ala	ND3
11,251	A	–	–	G	G	Silent	ND4
11,287	T	–	–	–	C (U)	Silent	ND4

Table 1 (*continued*)

Comparison with Cambridge Reference Sequence (CRS)							
11,335*E	T/C	C	C	C	C	Silent	ND4
11,719	G	A	–	A	A	Silent	ND4
11,778	G	–	–	–	A	Arg → His	ND4 LHON
11,878	T	C (U)	–	–	–	Silent	ND4
12,071[het]	T	–	–	C/T[het] (U)	–	Phe → Leu[het]	ND4
12,612	A	G	–	G	G	Silent	ND5
12,705	C	T	–	–	–	Silent	ND5
13,572	T	–	C (U)	–	–	Silent	ND5
13,702*E	G/C	C	C	C	C	Gly → Arg	ND5
13,708	G	A	–	A	A	Ala → Thr	ND5 LHON
13,759	G	–	A	–	–	Ala → Thr	ND5
13,966	A	G	–	–	–	Thr → Ala	ND5
14,199*E	G/T	T	T	T	T	Pro → Thr	ND6
14,272*E	G/C	C	C	C	C	Phe → Leu	ND6
14,365*E	G/C	C	C	C	C	Silent	ND6
14,368*E	G/C	C	C	C	C	Phe → Leu	ND6
14,470	T	C	–	–	–	Silent	ND6
14,569	G	–	–	A	–	Silent	ND6
14,766*E	T/C	T	C	T	T	Ile → Thr	ND6
14,798	T	–	–	–	C	Phe → Leu	CYT B
15,257	G	–	–	A	–	Asp → Asn	CYT B LHON
15,326*R	A	G	G	G	G	Thr → Ala	CYT B
15,452	C	–	–	A	A	Leu → Ile	CYT B
15,812	G	–	–	A	–	Val → Met	CYT B LHON
16,069	C	–	–	T	T		HV1
16,126	T	–	–	–	C		HV1
16,183	A	C	–	–	–		HV1
16,184–93	–	C (ins)	–	–	–		HV1
16,189	T	C	–	–	–		HV1
16,193	C	–	–	T	–		HV1
16,223	C	T	–	–	–		HV1
16,278	C	T	–	T	–		HV1
16,292	C	–	–	–	T		HV1
16,311	T	–	C	–	–		HV1
16,362	T	–	–	C	–		HV1
16,519	T	C	C	–	nd		HV1

[a] ATPase 6, ATP synthase 6; ATPase 8, ATP synthase 8; CYTB, Cytochrome B; COI, Cytochrome C Oxidase I; COII, Cytochrome C Oxidase II; COIII, Cytochrome C Oxidase III; del, deletion; het, heteroplasmy found in HL-60 at np 12,071; HV1, non-coding region found from 16,024 and 16,569; HV2, non-coding region found from 1 and 576; ins, insertion; nd, area not amplified or sequenced; ND1, NADH dehydrogenase 1; ND2, NADH dehydrogenase 2; ND3, NADH dehydrogenase 3; ND4, NADH dehydrogenase 4; ND5, NADH dehydrogenase 5; ND6, NADH dehydrogenase 6; (U), Unique polymorphisms found in coding regions of CHR, 9947A, HL-60 and GM10742A determined by comparison with the MITOMAP database (http://www.mitomap.org).

[b] Levin et al. (1999). CHR DNA: Sequence based on two amplifications and cycle sequencing procedures with DNA from the first cell culture line and at least one amplification and cycle sequencing procedure with DNA from the second cell culture line. 9947A DNA: Sequence based on two amplifications and cycle sequencing procedures.

[c] HL-60 DNA: Sequence based on two amplifications and cycle sequencing procedures in both the forward and reverse directions for a total of four sequences.

[d] Numbers correspond to Cambridge Reference Sequence (Anderson et al., 1981).

[e] Base found in 1981 (Anderson et al., 1981)/base found in 1999 (Andrews et al., 1999).

[f] Base pair same as in 1981 Cambridge Reference Sequence.

[g] *R: Rare polymorphisms in Cambridge Reference Sequence discovered by reanalysis of original placenta by Andrews et al. (1999).

[h] *E: Error in Cambridge Reference Sequence discovered by reanalysis of original placenta by Andrews et al. (1999).

[i] Possible heteroplasmic site. This heteroplasmy seen in the mtDNA from the first CHR cell culture line is not seen in the mtDNA from the second CHR cell culture line. It is DNA from the second CHR cell culture line that is supplied in NIST SRM 2392.

Table 2
Number of differences from the Cambridge Reference Sequence

DNA template	Non-coding regions[a]	Coding region	Amino acid changes[b]	Real change[c]
CHR[d]	13	33	11	3
9947A[d]	9	23	12	4
HL-60	11	33	15	7 + 2 tRNA
GM10742A	12	31	14	7

[a] The non-coding regions cover the areas from 16,024 to 16,569 and 1 to 576.
[b] Amino acid changes compared to original 1981 Cambridge Reference Sequence.
[c] Amino acid and tRNA changes compared to revised 1999 Cambridge Reference Sequence and not counting those changes considered a rare polymorphism.
[d] Levin et al. (1999).

found identical results in all amplified regions. SRM 2392-I containing HL-60 will be available in the spring of 2003. SRM 2392 that currently includes the CHR and 9947A templates also will continue to be available. The CHR and 9947A DNAs come from apparently normal individuals, but the HL-60 DNA comes from an individual who has acute promyelocytic leukemia. Since this is the first time that the entire mtDNA from HL-60 has been sequenced, it would be interesting to determine if the differences found in HL-60 are also found in other leukemia patients. Another interesting finding is the four differences in HL-60 that have been associated with LHON disease. For comparison, we also sequenced GM10472A, a cell line from a patient with LHON and found many of the same polymorphisms and mutations. Analysis of these polymorphisms places both HL-60 and GM10742A in Haplogroup J. These results support the premise that individuals in Haplogroup J may be more prone to LHON. At this time,

Table 3
HL-60 and GM10472A Polymorphisms in Coding Regions and not found in original SRM 2392[a]

Nucleotide position	Nucleotide change	Region[b]	Amino acid difference	Associated disease	Reference
2841^{GM}	T → A	16sRNA			MA
3010^{GM}	G → A	16sRNA			MM, MA
3394^{GM}	T → C	ND1	Tyr → His	LHON, NIDDM	MM, MA
$4216^{HL60;GM}$	T → C	ND1	Tyr → His	LHON	MM, MA
5228^{HL60}	C → G	ND2	Silent		MA
5633^{HL60}	C → T	tRNA ala			MM, MA
7476^{HL60}	C → T	tRNA ser			MM, MA
8020^{GM}	G → A	COII	Silent		MM, MA
$10,172^{HL60}$	G → A	ND3	Silent		MM, MA
$10,398^{HL60;GM}$	A → G	ND3	Thr → Ala		MM, MA, Mt-1
$11,251^{HL60;GM}$	A → G	ND4	Silent		MM, MA
$11,287^{GM}$	T → C	ND4	Silent		MA
$11,778^{GM}$	G → A	ND4	Arg → His	LHON	MM, MA
$12,071^{HL60}$	T → C/T	ND4	Phe → Leu		MA
$13,708^{HL60;GMc}$	G → A	ND5	Ala → Thr	LHON	MM, MA
$14,569^{HL60}$	G → A	ND6	Silent		MM, MA
$14,798^{GM}$	T → C	CYT B	Phe → Leu		MM, MA
$15,257^{HL60}$	G → A	CYT B	Asp → Asn	LHON	MM, MA
$15,452^{HL60;GM}$	C → A	CYT B	Leu → Ile		MM, MA
$15,812^{HL60}$	G → A	CYT B	Val → Met	LHON	MM, MA

[a] GM, GM10742A; NIDDM, non-insulin dependent diabetes mellitus; MM, http://www.mitomap.org; MA, http://www.cstl.nist.gov/biotech/strbase/mitoanalyzer.html; Mt-1, Levin, BC, Sekiguchi, K, Tully, LA, Chen, JT, and Gropman, A. A patient with chronic progressive external ophthalmoplegia reexamined 30 years later. (Manuscript In preparation, 2003).
[b] See footnote to Table 1.
[c] Also seen in CHR as well as HL-60 and GM10742A.

Table 4
Polymorphisms common to haplogroup J and sub-haplogroup J_2 found in the mtDNA of HL-60 and GM10742A[a]

Polymorphisms characteristic of haplogroup J	Polymorphisms characteristic of sub-haplogroup J_2	Found in HL-60	Found in GM10742A	Also found in	Disease associated
A73G		Yes	Yes	CHR	
C150T		Yes	No		
T152C		Yes	No		
G185A		No	Yes		
T195C		No	No	CHR 9947A	
G228A		No	Yes		
C462T		No	Yes		
T489C		Yes	Yes		
G3010A		No	Yes		
T4216C		Yes	Yes		LHON (secondary)
A10398G		Yes	Yes		
A11251G		Yes	Yes		
G11719A		Yes	Yes	CHR	
A12612G		Yes	Yes	CHR	
G13708A		Yes	Yes	CHR	LHON (secondary)
T14798C		No	Yes		
C15452A		Yes	Yes		
C16069T		Yes	Yes		
T16126C		No	Yes		
T16189C		No	No	CHR	
C16278T		Yes	No	CHR	
T16311C		No	No	9947A	
	C295T	Yes	Yes		
	C5633T	Yes	No		
	C7476T	Yes	No		
	G10172A	Yes	No		
	G15257A	Yes	No		LHON (intermediate)
	G15812A	Yes	No		LHON (secondary)
	C16193T	Yes	No		

[a] References: Torroni et al. (1997); and Finnilä et al. (2001).

since the linkage of HL-60 to the actual donor has been broken, there is no way to determine if this patient actually had LHON.

Acknowledgements

This research was supported by Interagency Agreement number 1999-IJ-R-094 from the National Institute of Justice. Points of view are those of the authors and do not necessarily represent the position of the U.S. Department of Justice. The authors are grateful to Gettysburg College for the partial funding of Koren A. Holland during her sabbatical year at NIST. We also acknowledge and thank Kathryn Kapfer for help in sequencing GM10742A at NIST.

References

Anderson, S., Bankier, A.T., Barrell, B.G., et al., 1981. Sequence and organization of the human mitochondrial genome. Nature 290, 457–465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N., 1999. Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA. Nat. Genet. 23, 147.

Bendall, K.E., Sykes, B.C., 1995. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am. J. Hum. Genet. 57, 248–256.

Brown, M.D., Starikovskaya, E., Derbeneva, O., et al., 2002. The role of mtDNA background in disease expression: a new primary LHON mutation associated with Western Eurasian haplogroup. J. Hum. Genet. 110, 130–138.

Butler, J.M., Levin, B.C., 1998. Forensic applications of mitochondrial DNA. Trends Biotechnol. 16, 158–162.

Finnilä, A., Lehtonen, M.S., Majamaa, K., 2001. Phylogenetic network for European mtDNA. Am. J. Hum. Genet. 68, 1475–1484.

Helgason, A., Hickey, E., Goodacre, S., et al., 2001. MtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic Ancestry. Am. J. Hum. Genet. 68, 723–737.

Herrnstadt, C., Elson, J.L., Fahy, E., et al., 2002. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. Am. J. Hum. Genet. 70, 1152–1171.

Lee, M.S., Levin, B.C., 2002. MitoAnalyzer, a computer program and interactive web site to determine the effects of single nucleotide polymorphisms and mutations in human mitochondrial DNA. Mitochondrion 1, 321–326.

Levin, B.C., Cheng, H., Reeder, D.J., 1999. A human mitochondrial DNA Standard Reference Material for quality control in forensic identification, medical diagnosis, and mutation detection. Genomics 55, 135–146.

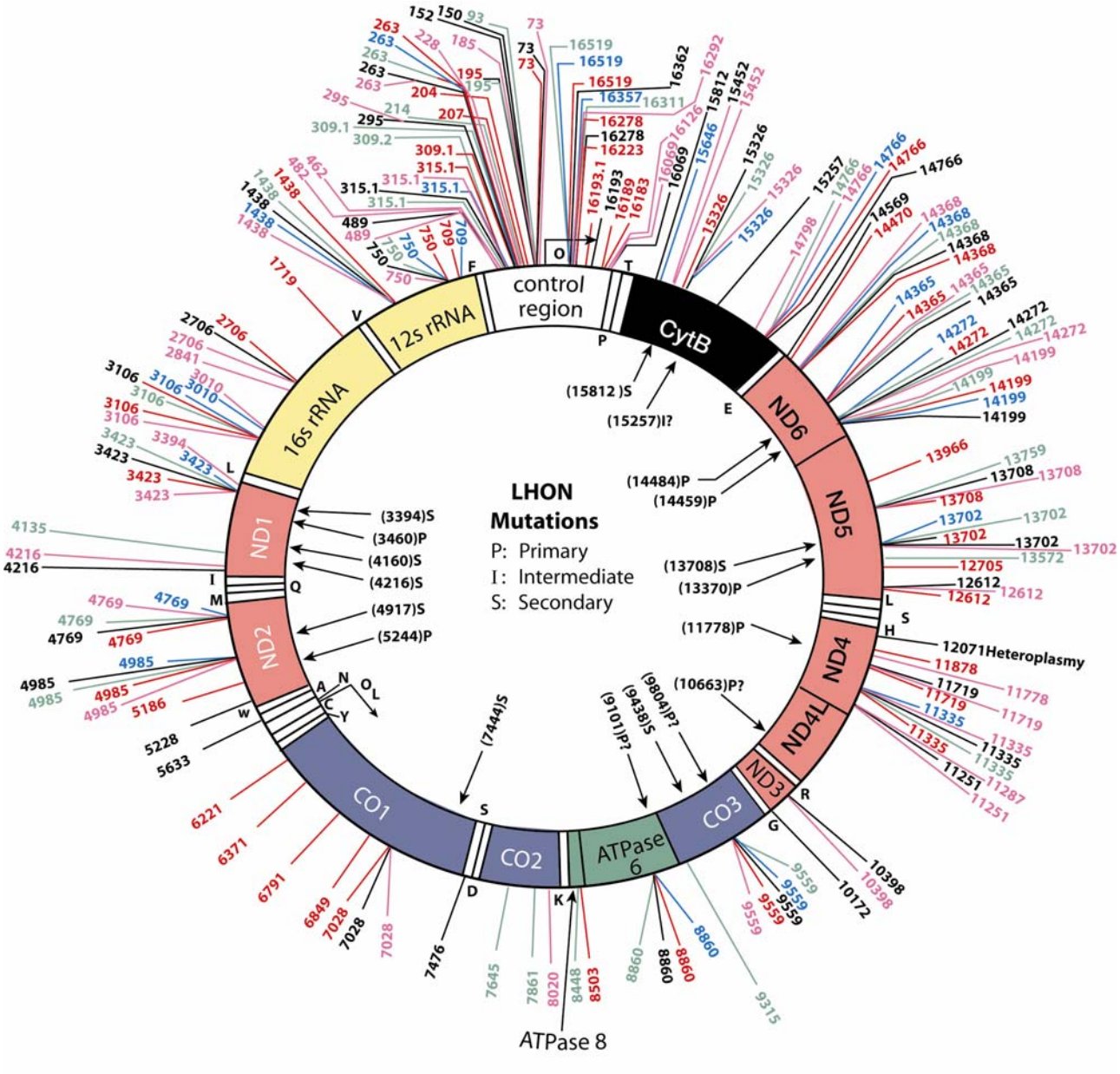
Levin, B.C., Cheng, H., Kline, M.C., Redman, J.W., Richie, K.L., 2001. A review of the DNA standard reference materials developed by the National Institute of Standards and Technology. Fresenius J. Anal. Chem. 370, 213–219.

Macaulay, V., Richards, M., Hickey, E., et al., 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control region sequences and RFLPs. Am. J. Hum. Genet. 64, 232–249.

Torroni, A., Petrozzi, M., D'Urbano, L., et al., 1997. Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber Hereditary Optic Neuropathy by increasing the penetrance of the primary mutations 11,778 and 14,484. Am. J. Hum. Genet. 60, 1107–1121.

Wallace, D.C., Singh, G., Lott, M.T., et al., 1988. Mitochondrial DNA mutation associated with Leber Hereditary Optic Neuropathy. Science 242, 1427–1430.

Wallace, D.C., Brown, M.D., Lott, M.T., 1997. Mitochondrial genetics. In: Rimoin, D.L., Connor, J.M., Pyeritz, R.E. (Eds.), Chapter 13 in Emery and Rimoin's Principles and Practice of Medical Genetics, Churchill Livingstone, New York, Edinburgh, London, pp. 277–332.


LHON
Mutations
P: Primary
I: Intermediate
S: Secondary
control region
12s rRNA
16s rRNA
ND1
ND2
CO1
CO2
ATPase 8
ATPase 6
CO3
ND3
ND4L
ND4
ND5
ND6
CytB
(15812)S
(15257)I?
(14484)P
(14459)P
(13708)S
(13370)P
(11778)P
(10663)P?
(9804)P?
(9438)S
(9101)P?
(7444)S
(5244)P
(4917)S
(4216)S
(4160)S
(3460)P
(3394)S
12071Heteroplasmy