Volume 16 Number 20 1988 Nucleic Acids Research

Mitochondrial DNA sequences from a 7000-year old brain

Svante Pääbo, John A.Gifford[1] and Allan C.Wilson

Department of Biochemistry, University of California, Berkeley, CA 94720 and [1]Department of Anthropology, University of Miami, PO Box 248106, Coral Gables, FL 33124, USA

Received May 5, 1988; Revised and Accepted September 2, 1988

ABSTRACT

Pieces of mitochondrial DNA from a 7000-year-old human brain were amplified by the polymerase chain reaction and sequenced. Albumin and high concentrations of polymerase were required to overcome a factor in the brain extract that inhibits amplification. For this and other sources of ancient DNA, we find an extreme inverse dependence of the amplification efficiency on the length of the sequence to be amplified. This property of ancient DNA distinguishes it from modern DNA and thus provides a new criterion of authenticity for use in research on ancient DNA. The brain is from an individual recently excavated from Little Salt Spring in southwestern Florida and the anthropologically informative sequences it yielded are the first obtained from archaeologically retrieved remains. The sequences show that this ancient individual belonged to a mitochondrial lineage that is rare in the Old World and not previously known to exist among Native Americans. Our finding brings to three the number of maternal lineages known to have been involved in the prehistoric colonization of the New World.

INTRODUCTION

In 1985, Pääbo (1) reported that pieces of repetitive nuclear DNA could be cloned and sequenced from ancient human remains. However, until now, no success has been achieved with such materials in routinely obtaining sequences from preselected regions of the genome known to harbor anthropologically informative sites. The success reported here was achieved with DNA from a 7000-year-old brain, recently excavated in Florida.

Ancient human brains survive in peat bogs and other wet sites at several places in Florida (2,3). Although human DNA is detectable in this material, it is highly modified (3,4). Lesions occur so often in this DNA that it breaks down into bits about 100-bp long upon treatment with alkali (3). Probably for this reason, no success was achieved before in

cloning informative DNA sequences from ancient brains preserved in the anaerobic peaty sediments typifying such sites. With this hypothesis in mind, we have modified the polymerase chain reaction so that it will amplify short segments of mitochondrial (mt) DNA from such a brain.

The brain studied here comes from Little Salt Spring (North Port, Florida), which is a large, flooded sinkhole that served prehistoric humans as a source of fresh water during periods of drought. Archaeological excavations around the spring have yielded evidence of Paleo-Indian as well as Archaic Indian activities, ranging in age from 12,000 to 9000 BP and from 6800 to 5200 BP, respectively (5). The environmental conditions at Little Salt Spring have allowed for the preservation of large quantities of artifacts and organic remains.

During an exploratory excavation in the southwest sector of the spring basin in April-May 1986 one relatively intact burial was discovered, the cranium of which contained macroscopically well-preserved neural tissue (J.A. Gifford et al., in preparation). An accelerator mass spectroscopy radiocarbon date of 6860±110 years BP (Beta-17208; corrected by ^{13}C for total isotope effect, see ref. 6 for method) was obtained from a small sample of the neural tissue, indicating a Middle Archaic age. We here report anthropologically informative mitochondrial DNA sequences from this individual, which demonstrate that he or she belonged to a maternal lineage not known before in Native Americans.

MATERIALS AND METHODS

Extraction and Purification of Ancient DNA

Samples of neural tissue, between 0.5 and 1 gram in size, were put into 8 ml of 10 mM Tris (pH 8.0), 2 mM EDTA and 10 mM NaCl. The tissue was broken up into small pieces by vortexing and 4 mg of collagenase (Sigma Chemical Co., St. Louis, MO, Cat. # C-0130) was added. The tubes were incubated at 37°C with slow agitation for 3 hours. 880 µl of a 10% stock solution of sodium dodecyl sulphate (SDS) was then added as well as 80 mg of dithiothreitol (DTT) (BRL, Gathersbury, MD) and 4 mg of proteinase K (Boehringer Mannheim GmbH, FRG). Incubation was continued for approximately 20 hours. At that time very little solid tissue remained. The solution was extracted twice with water-saturated phenol of neutral pH and once with chloroform/isoamylalcohol (24:1). At this point the bulk of the extract was stored at -70°C while an aliquot of 10% (500-700 µl) was diluted to a total volume of 2 ml with glass-distilled water and

concentrated on a Centricon 30 microconcentrator (Amicon Inc., Danvers, MA). This procedure avoids some of the losses of material encountered during ethanol precipitation. After addition of another 2 ml of distilled water and a second concentration, the extract was stored at 4°C. Sterile and disposable tubes and pipets were used throughout for the preparation of the solutions and the extractions. For control extractions either nothing or 10 µg of E.coli JM101 DNA was added to the extraction buffer.

Modified Conditions for the Polymerase Chain Reaction

The polymerase chain reaction (PCR) was performed in a 50-µl reaction mixture containing 67 mM Tris/Cl, pH 8.8, 2 mM $MgCl_2$, 10 mM ß-mercaptoethanol, 250 µM each of dATP, dCTP, TTP and dGTP, 2 µg/ml bovine serum albumin (Sigma, MO, fraction V) and 12.5 units of Thermus aquaticus DNA polymerase (Perkin Elmer-Cetus, Norwalk, CT). The albumin as well as the high concentration of the enzyme was necessary to overcome an inhibitory activity of unknown origin which was present in the ancient extracts. Thermal cycling was performed in a programmable heating block (Perkin Elmer-Cetus, Norwalk, CT).

Primers A, B and C used for amplification and dideoxynucleotide sequencing of region V have been described (7). The sequence of primer M14725 is 5'-CGAAGCTTGATATGAAAAACCATCGTTG-3'-14724 and the sequence of primer cytb2 is 5'-AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA-3'-15149. This set of primers amplifies a 471 bp-long segment of the cytochrome b gene (c.f. ref. 8). With the above primers, forty PCR cycles were performed as follows: denaturation at 92°C for 30 seconds, cooling to 50°C over 50 seconds, annealing at 50°C for 1 minute, heating to 72°C over 1 minute and extension at 72°C for 1 minute. The primers used for amplification and sequencing of the region containing the polymorphic HincII site at position 13259 were: (M13232) 5'-CGCCCTTACACAAAATGACATCAA-3'-13232 and (M13286) 5'-GTGTGGTTGGTTGATGCCGA-3'-13286. The PCR cycle used for this pair of primers was as above except that the annealing was done at 60°C and extension carried out for 20 seconds.

Unbalanced Priming and Sequencing

To obtain high yields in subsequent sequencing reactions use was made of the unbalanced priming method of Gyllensten and Erlich (9), which generates a single-stranded product suitable for sequencing. Twenty µl of the first amplification mixture, containing double-stranded amplification products, was resolved in a 4% NuSieve agarose gel in Tris-acetate

buffer and stained with ethidium bromide. The relevant band was cut out of the gel and a thin slice of the band (volume 1-2 µl) was subjected to a second round of PCR. This reaction was performed as the first one except that one of the two primers used for the initial amplification was reduced 100-fold in concentration and only 30 PCR cycles were performed. The entire reaction mixture was then concentrated twice on Centricon 30 microconcentrators from 2 ml of distilled water. Approximately 10% of the concentrated reaction was used for sequencing by the dideoxynucleotide chain termination method (10) using the Sequenase enzyme (United States Biochemical Co., Cleveland, Ohio) (11). Electrophoresis in 8% polyacrylamide gels and autoradiography were by standard techniques (12).

RESULTS

Small Size of DNA from the Ancient Brain

Immediately upon its discovery the cranium was placed in spring water and refrigerated. On the next day, samples of neural tissue were removed through the foramen magnum and frozen in liquid nitrogen. Approximately one gram of tissue was used for each extraction. Control extractions using no tissue or 10 µg of bacterial DNA were performed in parallel with the brain extractions. Agarose gel electrophoresis of an aliquot of the extracted DNA showed that the majority of the DNA detectable with ethidium bromide was degraded to a relatively small molecular size of approximately 50-200 bp (Figure 1), while a fraction of the material was of a larger molecular size, mostly in the range from two to five kb. Such a reduction in size is typical of ancient DNA (3,4).

Size Dependence of the Amplification Reaction

Aliquots of 2-4% of the extracts were used for enzymatic amplification (13) of various parts of the mitochondrial genome, using the heat-resistant Thermus aquaticus DNA polymerase (14). Initially, various sets of primers were used to determine the maximum size of the DNA fragments which could be amplified. It was found that pieces up to 120 bp in length could be successfully and reproducibly amplified whereas no amplification of over 200 bp was ever successful. This was in sharp contrast to contemporary human DNA extracted from brain tissue by the same procedure, from which pieces of up to 2 kb could be amplified (see Figure 2). From comparisons of the amounts of the specific amplification products obtained to dilution-series of

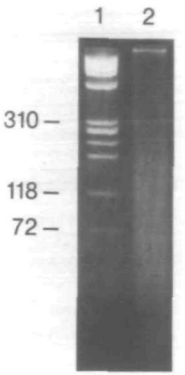

Figure 1. Agarose gel electrophoresis of the DNA extracted from the 7000-year-old brain. Lane 1 represents molecular size markers, some of which are indicated in number of base pairs, and lane 2 shows 10% of the DNA extracted from approximately one gram of brain tissue. DNA fragments over 2 kb in size are not separated on this 4% agarose gel.

contemporary human DNA extracted from brain, it was estimated that the amplification mixtures contained between 100 and 300 fg of human DNA, which could be replicated by the T. aquaticus polymerase. Based on

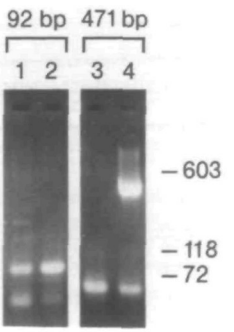

Figure 2. Agarose gel electrophoresis of the products of mtDNA amplification performed from the DNA of the 7000-year-old brain as well as contemporary human DNA. Lanes 1 and 3 show the amplification products from the Little Salt Spring brain whereas lanes 2 and 4 show the amplifications from 1.2 ng of contemporary human DNA. The 92-bp-long amplifications (lanes 1 and 2) were performed using primers A and C (7) and the 471-bp-long amplification using primers M14725 and cytb2. The amplification products of approximately 40 in lanes 1 and 2 and 60 bp in lanes 3 and 4 are dimers of primers. The migration positions of molecular size markers are indicated in number of base pairs.

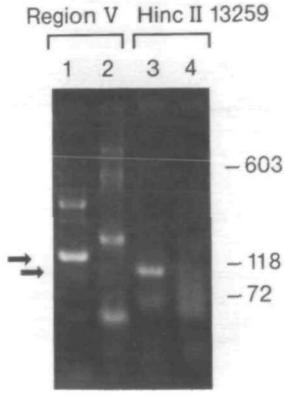

Figure 3. Agarose gel electrophoresis of the products of amplification reactions performed on mtDNA from a 7000-year-old brain. Lanes 1 and 3 represent DNA extracted from the Little Salt Spring brain that has been amplified with pairs of primers specific for region V and the HincII site at position 13259, respectively. Lanes 2 and 4 represent amplifications from control extract with the same primer pairs where no tissue was added. The migration positions of molecular size markers are indicated in number of base pairs. Arrows indicate the bands whose sequences appear in Figure 4.

the estimation that there exists 1000 copies of mtDNA in one mammalian cell, this would indicate that the amplifications started from approximately fifty copies of the relevant mitochondrial sequences. Two anthropologically informative regions of the mitochondrial genome were selected for further study.

Amplification of Informative DNA sequences

Region V is a small intergenic segment of the mitochondrial genome, located between the cytochrome oxidase II gene and the gene encoding lysyl tRNA. It carries two direct repeats of 9 bp, one of which has been shown to be deleted in some individuals of Asian origin (7, 15, 16) as well as in members of populations thought to originate from Asia, such as North American Indians (C. Beck, C. Orrego and A.C. Wilson, in preparation) and New Guineans (17). When the primers specific for region V were used and the PCR carried out for 40 cycles with aliquots from the Little Salt Spring extract as well as relevant control extracts, an amplification product of the expected size was detected only in the extract from the brain tissue (Figure 3). This band was cut out from the

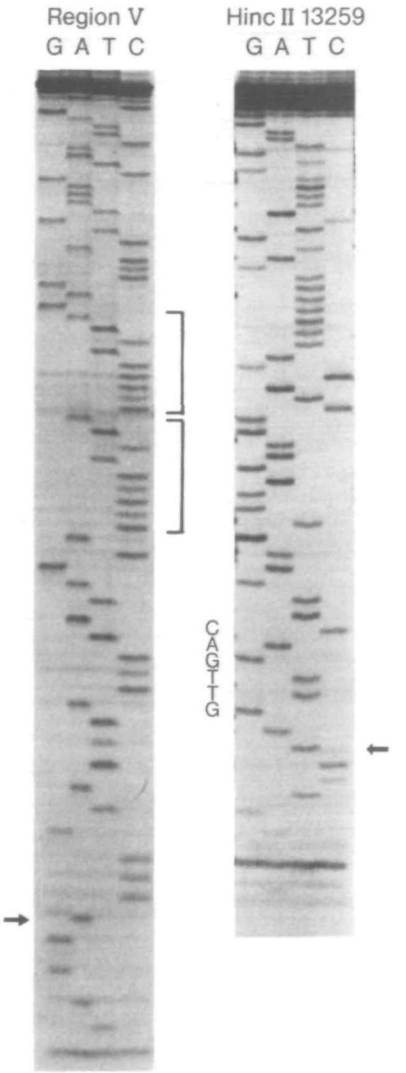

Figure 4. Direct sequencing of amplification products from the 7000-year-old brain. The sequence of region V was determined with primer C (ref. 7). Brackets indicate the two direct repeats and the arrow points to the substitution at position 8251. The sequence of the amplification product containing the HincII site starting at position 13259 was determined using primer M13286. The sequence of the restriction site is indicated and the arrow points to the double sequence at position 13266.

agarose gel along with the same regions from the control lanes where no specific products were discernible. Slices of these agarose pieces were subjected to a further thirty cycles of PCR, with one of the primers being present in a 100:1 excess over the other. A representative sequence from the brain extract appears in Figure 4. It is evident that this individual carries both of the repeats in region V. The presence of the two repeats in region V was confirmed in two additional extracts prepared from the same brain.

No specific sequence was obtained even after long exposures of the autoradiograms from the control experiments. Some other bands which are of higher molecular weight can be seen in Figure 3. These bands do not contain the relevant region V sequence since they cannot be sequenced using an internal sequencing primer (c.f. ref. 7). Such nonspecific amplification products are sometimes obtained even in control extracts where no template DNA should be present when large amounts of T. aquaticus DNA polymerase are used. They presumably stem from weak priming sites occurring in non-human DNA molecules, which are introduced into the reaction in minute amounts with the reagents used.

An additional informative site is present in the sequence obtained from region V. At position 8251 the sequence from the brain contains an A whereas that position in the published sequence of the human mitochondrial genome (8) carries a G. This represents a silent transition which results in a loss of a HaeIII site at position 8250 and the gain of an AvaII site at position 8249 in the gene for cytochrome oxidase II. Among the 241 individuals examined in a world-wide survey of human mtDNA (18,19), this restriction site variant occurs in only three individuals, originating from different geographical regions. Furthermore, in these three cases the variant could be due to a change at position 8251 or 8252. Therefore, this is a rare sequence. It might thus be of significance that the HaeIII loss in conjunction with the AvaII gain occur in five mtDNA types forming a clade in a phylogenetic tree of 62 mtDNA types in the Japanese population (15) and that in all those cases the same nucleotide change is responsible for the change in restriction patterns (S. Pääbo, recent observation).

An anthropologically informative HincII site has been described at position 13259 in the published human mitochondrial genome (8). This site is absent in 40% of Amerindians living in southwestern United

States (20), whereas it is absent in only one out of 55 Orientals analyzed (21). The site has also been shown to be absent in another (Apache) Indian included in the survey of human mitochondrial types (18; M. Stoneking, pers. comm.). An attempt was made to amplify a 97-bp region encompassing position 13259 from the Little Salt Spring brain. A specific band of the expected size was detected in the amplification mixture after 40 cycles of PCR, whereas no band was seen in the controls (Figure 3). Single-stranded template was generated and the product along with corresponding controls were sequenced (Figure 4). It can be seen that the sequence from the brain sample carries the intact HincII site at 13259. At position 13266 a weak band can be seen in the C reaction at this position as well as a strong band in the T reaction (see Figure 4). An analogous situation was present when the other strand was sequenced (not shown). In two independent extracts from the same brain a similar double-sequence was observed. This may represent a contamination of the archaeological sample with extraneous DNA or a heteroplasmy of the tissue, which might be speculated to occur in tissues such as brain, where little or no cell-division occurs after birth (22).

DISCUSSION

The Polymerase Chain Reaction in Molecular Archaeology

The study of ancient DNA poses numerous technical problems. For example, the DNA sequences of interest can be expected to exist in low amounts in the presence of an excess of other DNA, often of bacterial or fungal origin. This makes the screening of genomic libraries prepared from ancient DNA a tedious process. Also, the vast majority of the DNA molecules that are extracted from archaeological remains have suffered physical damage that precludes replication (S. Pääbo, in preparation), which drastically reduces the efficiency of conventional cloning procedures. The polymerase chain reaction (13) in conjunction with the heat-resistant DNA-polymerase of T. aquaticus (14) can be expected to overcome these problems since this technique makes possible the specific in vitro amplification of extremely small numbers of a relevant DNA sequence up to amounts which allow for its study by conventional sequencing techniques (23). However, DNA extracts from ancient tissue remains often contain an unidentified component which inhibits the activity of T. aquaticus polymerase. Here, we show that this inhibition

can be overcome by the addition of bovine serum albumin and increased amounts of enzyme. Under these conditions, the PCR is able to detect what corresponds to 100 fg of intact human DNA in a vast excess of extraneous or damaged DNA in extracts prepared from a 7000-year-old brain, excavated from an Archaic Indian site in Florida. Since this approach may be fruitfully applied also to the study of both desiccated (S. Pääbo, in preparation) and frozen (R.G. Higuchi and A.C. Wilson, in preparation) ancient remains, the polymerase chain reaction will undoubtedly prove to be the method of choice for the study of DNA retrieved from archaeological samples.

Authenticity of Amplified DNA

When DNA from an extinct species, such as the quagga (24, 25, 26), is studied, the authenticity of the DNA fragments obtained can be verified by their degree of similarity to orthologous fragments in closely related, extant species. Unfortunately, this phylogenetic criterion of authenticity cannot be applied when ancient human DNA is studied, since the most likely source of contamination is DNA from other humans. Such contaminations may stem from a variety of sources, such as ancient blood-stains on a specimen, handling during excavation or removal of samples and contamination of reagents and glassware used during the extraction and analysis procedures.

The laboratory contaminations can be controlled for by doing control extracts and amplifications and discarding all reagents which give rise to specific amplification products. Contaminations derived from ancient or contemporary handling of, e.g., a body can be controlled for by performing several independent extracts from the same individual, preferably from different organs.

An additional criterion of authenticity is made possible by our finding that for ancient DNA there is a strong inverse dependence of the amplification efficiency on the size of the segment to be amplified (Figure 2). The small molecular size of amplifiable template DNA of ancient origin is a feature that we have found to be a consistent characteristic of archaeologically retrieved DNA. By contrast, for modern DNA, the amplification efficiency for a 500-bp segment is nearly the same as for a 100-bp segment. This property of old DNA can often be used to detect and discriminate against relatively recent contaminations. For example, in extracts from the remains of a quagga that has been used previously to clone mitochondrial DNA fragments (24), the polymerase

chain reaction routinely amplifies small DNA fragments (≤140 bp) that consistently yield quagga sequences, whereas longer amplification products, which can sporadically be obtained, invariably yield human sequences (26).

Thus, for ancient human remains, the small size of an amplifiable sequence in conjunction with the reproducibility of the results within and between extracts can be taken as indicating that a particular DNA sequence is of ancient origin. Furthermore, in the present case, the sequence found is uncommon. If it had originated from laboratory personnel, a G would have been expected at position 8251, not an A. Nevertheless, indisputable proof that the amplified human sequences reported here are of ancient origin must await more extended work involving numerous individuals from ancient populations, which will allow for a phylogenetic analysis of sequence polymorphisms. Such work is in progress.

A New Native American Mitochondrial Lineage

Our demonstration that the Middle Archaic Indian excavated at Little Salt Spring carries two direct repeats in region V is of anthropological importance. This individual does not belong to the mitochondrial lineage bearing the 9-bp deletion that occurs specifically in Asian populations (7,17) and is known to be present, in conjunction with an intact HincII site at position 13259, among American Indians (Beck et al., in preparation). Also, since it carries an intact HincII site at position 13259, the ancient mtDNA does not belong to the lineage that has been shown to be present in 40% of present-day Amerindians in the southwestern USA (20, 21). Thus, the ancient mitochondrial lineage is previously not described in the New World. The prehistoric colonisation of the New World therefore involved a minimum of three distinct maternal lineages.

The transition at position 8251 suggests that this ancient individual may stem from a small group of mitochondrial lineages known to be present at low frequency in Asia (15). Only further investigations can tell whether this is indeed the case. Furthermore, future studies of brains from Florida peat bogs will address the question whether the mitochondrial genotype of this individual was a frequent one among the Archaic Indians of Florida and if they are maternally related to present day Indians in this or other regions of the New World.

ACKNOWLEDGEMENTS

We are grateful to D. Hillis for fruitful collaboration and discussions, to T. Kocher for continuous constructive criticism, and to E. Prager and E. Zimmer for comments on the manuscript. C. Beck, C. Orrego, M. Stoneking and U. Gyllensten kindly made their unpublished results available to us. S.P. is supported by a long-term fellowship (ALTF 76-1987) from the European Molecular Biology Organization.

REFERENCES

1. Pääbo, S. (1985) Nature **314**, 644-645.
2. Royal, W. and Clark, E. (1960) Am. Antiq. **26**, 285-287.
3. Doran, G.H., Dickel, D.N., Ballinger Jr., W.E., Agee, O.F., Laipis, P.J. and Hauswirth, W.W. (1986) Nature **323**, 803-806.
4. Pääbo, S. (1986) Cold Spring Harbor Sympos. Quantit. Biol. **51**, 441-446.
5. Clausen, C.J., Cohen, A.D., Emiliani, C., Holman, J.A. and Stipp, J.J. (1979) Science **203**, 609-614.
6. Gowlett, J. (1987) J. World Prehist. **1** (2), 127-170.
7. Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. and Wilson, A.C. (1987) Nucl. Acids Res. **15**, 529-542.
8. Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, 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) Nature **290**, 457-465.
9. Gyllensten, U. and Erlich, H.A. (1988) Proc. Natl. Acad. Sci. USA, in press.
10. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467.
11. Tabor, S. and Richardson, C.C. (1987) Proc. Natl. Acad. Sci. USA **84**, 4767-4771.
12. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).
13. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Science **230**, 1350-1354.
14. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R.G., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science **239**, 487-494.
15. Horai, S. and Matsunaga, E. (1986) Human Genet. **72**, 105-117.
16. Horai, S., Gojobori, T. and Matsunaga, E. (1987) In Human Genetics, Proceedings of the Seventh International Congress, Berlin 1986 (Eds. F. Vogel and K. Sperling). Springer-Verlag, Berlin, pp. 177-181.
17. Stoneking, M. and Wilson, A.C. (1988) In The Colonisation of the Pacific: A Genetic Trail (Eds. A. Hill and S. Serjeantson), in press. Oxford University Press.

18. Cann, R.L., Stoneking, M. and Wilson, A.C. (1987) Nature **325**, 31-36.
19. Stoneking, M. Ph.D. Thesis, Univ. California, Berkeley, California (1986).
20. Wallace, D.C., Garrison, K. and Knowler, W.C. (1985) Am. J. Phys. Anthropol. **68**, 149-155.
21. Blanc, H., Chen, K.-H., D'Amore, M.A. and Wallace, D.C. (1983) Am. J. Hum. Genet. **35**, 167-176.
22. Holt, I.J., Harding, A.E. and Morgan-Hughes, J.A. (1988) Nature **331**, 717-719.
23. Higuchi, R., von Beroldingen, C.H., Sensabaugh, G.F. and Erlich, H.A. (1988) Nature **332**, 543-546.
24. Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A. and Wilson, A.C. (1984) Nature **312**, 282-284.
25. Higuchi, R.G., Wrischnik, L.A., Oakes, E., George, M., Tong, B. and Wilson, A.C. (1987) J. Mol. Evol. **25**, 283-287.
26. Pääbo, S. and Wilson, A.C. (1988) Nature **334**, 387-388.