Volume 16 Number 13 1988 Nucleic Acids Research

Conformational mutation in human mtDNA detected by direct sequencing of enzymatically amplified DNA

Linda Vigilant, Mark Stoneking and Allan C.Wilson

Department of Biochemistry, University of California, Berkeley, CA 94720, USA

Received March 21, 1988; Revised and Accepted June 2, 1988

ABSTRACT

Restriction enzyme analysis of 241 human mtDNAs revealed polymorphism in the electrophoretic mobility of a fragment corresponding to part of the ND4 gene. Enzymatic amplification and direct sequencing of this fragment demonstrates that a single T <-> C transition correlates with the faster mobility exhibited by the fragment in seven mtDNAs from Papua New Guinea. The enhanced mobility caused by this transition could result from disrupting an AT-rich region near the middle of the fragment that might make it curve. An analogous mutation at an adjacent position in a European mtDNA causes a similar but more pronounced alteration in mobility. The seven New Guineans with this substitution are all from the Eastern Highlands Province and constitute one clade in a genealogical tree based upon restriction analysis. The additional information provided by sequencing allows refinement of the genealogical tree but does not require modification of higher order branching structure.

INTRODUCTION

Restriction mapping has been used extensively to examine polymorphism in human mitochondrial DNA (mtDNA). Until recently, differences in the electrophoretic mobilities of restriction fragments were assumed to be due either to base substitutions within restriction sites or to length mutations (i.e., additions or deletions of bases) between sites. However, Singh et al. (1) described two cases of base substitutions between sites that caused changes in fragment mobility, probably by affecting the curvature of DNA. Therefore, some polymorphisms that had previously been attributed to length mutations may turn out to be the result of conformational substitutions.

High-resolution restriction mapping of 112 human mtDNAs identified nine regions of apparent length mutation (2). Each of these nine regions includes an intergenic junction usually containing a noncoding segment (Fig. 1) which could contain the length polymorphism. One of these regions was characterized by enzymatic amplification and direct sequencing and shown to contain a 9-bp deletion in a noncoding segment in some individuals, while in another individual there was an

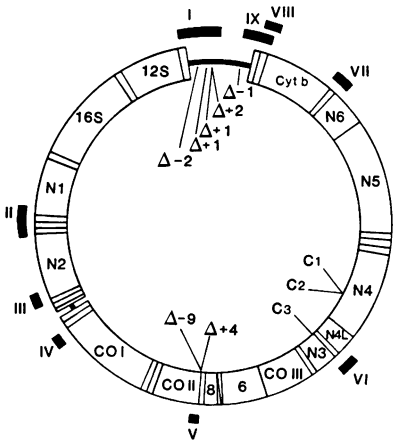

Figure 1 Length and conformational polymorphisms in the human mtDNA genome. Abbreviations for the genes are as follows: two ribosomal RNA genes (12S and 16S); seven genes for NADH dehydrogenase subunits (N1-N6, N4L); three genes for cytochrome oxidase subunits (COI-COIII); two genes for ATPase subunits (6 and 8); and the gene for cytochrome b (Cyt b). The 22 transfer RNA genes are represented by the hollow unlabelled bars. The solid bars labelled I-IX show the approximate locations of nine apparent length polymorphisms described by Cann and Wilson (2). Seven examples of authentic length mutations are now known to occur in noncoding portions of mtDNA, including five in the displacement loop (3,4). In contrast, the three conformational mutations (C) known are in coding DNA (1, this report). C1 is the T<->C transition at 11254 reported here, C2 is the T<-> C transition at 11253 (1), and C3 is an A<->G transition at 10398 (1).

addition of 4-bp (3). In addition, Greenberg et al. (4) found five length mutations in the major noncoding region by conventional cloning and sequencing (Fig. 1).

A restriction mapping survey of mtDNA from 129 additional individuals by Stoneking (5) revealed a tenth region of apparent length polymorphism mapping entirely to DNA coding for subunit 4 of NADH dehydrogenase (Fig. 1). Because this appeared to be a good candidate for a conformational polymorphism rather than a length polymorphism, we decided to sequence the variant mtDNAs. In this study, enzymatic amplification by the polymerase chain reaction and direct sequencing of the affected region proved to be a convenient way of showing that single-base substitutions occurring outside restriction sites can alter DNA mobility, apparently by affecting DNA curvature.

MATERIALS AND METHODS

Preparation and Electrophoresis of DNA

Highly purified mtDNAs from nine placentas and three cell lines (HT1080,

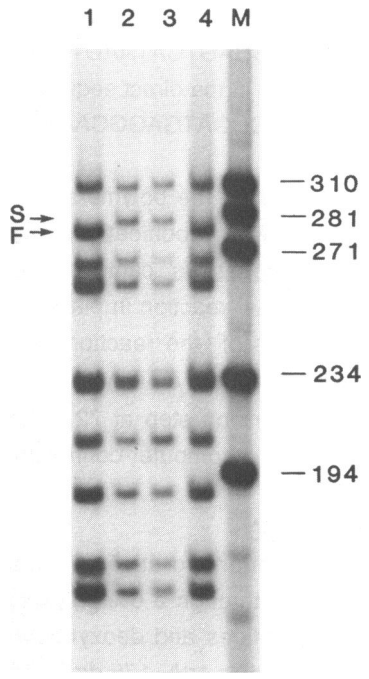

Figure 2 Electrophoretic demonstration of genetic polymorphism in the mobility of a TaqI fragment of human mtDNA. This autoradiograph of a 6% native polyacrylamide gel contains some of the end-labelled fragments resulting from TaqI digestion of purified mtDNA. Lanes 1-4 contain mtDNA from individuals 5,2,3, and 4 as listed in Table 1. Arrows point to the polymorphic 297-bp fragment, whose mobility is described as "fast" (F) for individuals 5 and 4 and "slow" (S) for individuals 2 and 3. M refers to the size standard (ΦX174 DNA cut with HaeIII).

GM3043, and HeLa) were prepared by differential ultracentrifugation through cesium chloride density gradients as previously described (5, 6). Restriction digests were performed according to the recommendations of the supplier (New England Biolabs). DNA was end-labelled with ^{32}P- or ^{35}S-nucleotides by standard methods (7). DNA was electrophoresed at room temperature for 15-20 hours at 5mA (110V) through 40-cm long 6% nondenaturing polyacrylamide gels in a Tris-borate buffer described by Maniatis et al. (8).

Primers

Two 20-base oligonucleotide primers made by a Biosearch DNA synthesizer were used for the amplification of a 325-bp region from bases 11091 to 11415 of the published sequence (9). These primers were also used for direct sequencing, and have the following sequences:

Primer 1 11091-5'-CAGCCACAGAACTAATCATA-3'-11110

Primer 2 11415-5'-GCTTTAGGGAGTCATAAGTG-3'-11396

A third primer of 19 bases used only in the direct sequencing is:

Primer 3 11154-5'-TCATCACCCGATGAGGCAA-3'-11172

Polymerase Chain Reaction

Enzymatic amplification based on the polymerase chain reaction followed the method of Wrischnik et al. (3) with modifications. DNA polymerase from the thermophilic bacterium Thermus aquaticus came from Cetus Corporation and was used at a concentration of 5 units per reaction in place of E. coli DNA Polymerase I (Klenow fragment). The 25 cycles of the reaction each consisted of a template denaturation step at 94 C for one minute, an oligonucleotide annealing step at 50 C for one minute and a product extension step at 72 C for two minutes; the entire amplification process took place in a computer-controlled water bath designed and built by T. Kocher.

Direct Sequencing of Amplified DNA

Sequencing reactions were performed at room temperature and the products were separated by electrophoresis through a 6% polyacrylamide gel. The reaction mixes contained dideoxyribonucleotides and deoxyribonucleotides in the following molar ratios: 75 ddG : 1 dG, 160 ddA : 1dA, 175 ddT : 1dT and 60 ddC : 1dC. All other details of the sequencing procedure were as described by Wrischnik et al. (3).

RESULTS

Electrophoretic Mobilities of Fragments

In the case of 234 individuals a 297-bp fragment produced from the ND4 gene by Taq I appeared to move at the conventional rate during electrophoresis through polyacrylamide gels (2, 5); however, seven individuals from Papua New Guinea were exceptional in possessing a fragment of faster mobility (Fig. 2). A 420-bp Hinf I fragment showed a similar shift in mobility and has a 267-bp zone of overlap with the Taq I fragment. We assumed, therefore, that the basis for the mobility difference resides within this 267-bp zone, which spans from bases 11125 to 11391 in the published sequence.

A 325-bp region including this zone of overlap was amplified enzymatically using the Taq DNA polymerase. The mobility of the amplified fragment was greater for the seven New Guineans than for four other individuals but slightly less than the mobility of the corresponding fragment from the HT1080 cell line (Fig. 3).

Sequencing the Amplified DNA

As is evident from the sequencing gel autoradiograph (Fig. 4) and Table 1, the electrophoretic mobility alteration observed is not due to a length mutation but is

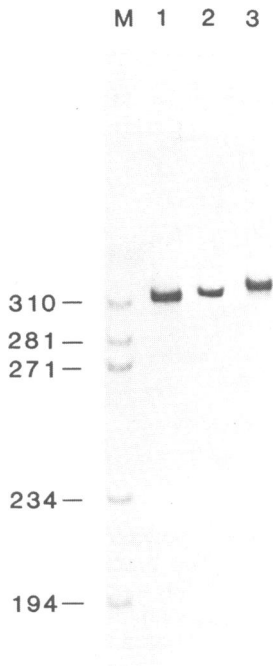

Figure 3 Autoradiograph of amplification products from mtDNAs of three humans {HT1080 (lane 1), individual 9 (lane 2), and individual 1 (lane 3)} separated in a 6% native polyacrylamide gel. Each lane contains approximately 50 ng of DNA end-labelled with ^{35}S-dGTP and ^{35}S-dCTP. Lane M contains size marker (ΦX174 cut with HaeIII). The electrophoretic mobility difference seen in the restriction fragments persists in the DNA fragments resulting from the amplification process. The DNA in lane 3 has the slowest mobility relative to the "fast" DNA in lane 2 and the "faster" DNA in lane 1. Note the absence of any bands other than the expected 325-bp product of the enzymatic amplification procedure.

correlated with a single T<->C transition . Of twelve mtDNAs surveyed, the four that yielded Taq I and Hinf I fragments of conventional mobility exhibited a T at position 11254, as does the reference sequence. However, the seven New Guinean samples with fragments of faster mobility were found to possess a C at 11254. In addition, our enzymatic amplification and direct sequencing of mtDNA from HT1080 verified the previously reported presence of a C at position 11253 (Table 1). One other substitution, a silent T<->C transition at position 11335, was also detected but appeared in all twelve mtDNAs and so could not account for the mobility shift.

Mobilities of Smaller Fragments

The DNA region responsible for the mobility shift was deduced by comparing the mobilities of smaller fragments having either T or C at position 11254. These smaller

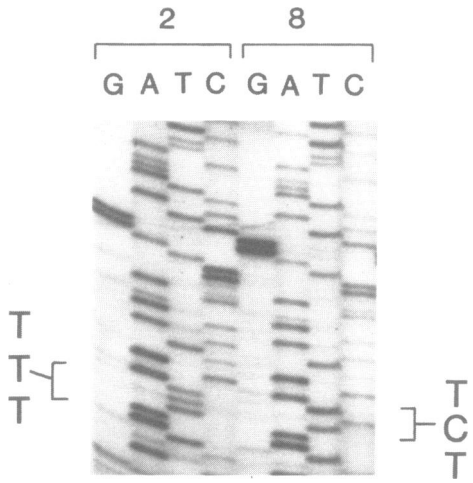

Figure 4 Autoradiograph of a 6% polyacrylamide sequencing gel of amplified mtDNA from individuals 2 and 8. The brackets indicate the single base change, T<->C at 11254, that correlates with the electrophoretic mobility difference seen in restriction fragments in native polyacrylamide gels. Individual 2 has a T at 11254 and a slow fragment, while individual 8 has a C and a faster fragment.

fragments were produced by cleaving the amplified 325-bp segment with restriction enzymes. A plot of relative gel mobility vs. cleavage position indicates that the mobility difference is minimized when the polymorphic base is located towards the end of a fragment (Fig. 5). The position of the minimum in the plot implicates a region near the middle of the 325-bp fragment and inclusive of positions 11253 and 11254 as the center of curving (Fig. 5).

Restriction Analysis of HT1080 mtDNA

Twelve restriction enzymes recognizing approximately 370 restriction sites were used to map HT1080 mtDNA by the sequence comparison method (11). This survey revealed three restriction site polymorphisms that could be used to place HT1080 mtDNA on the phylogenetic tree of 147 human mtDNA types (3): an Hha I site absent at position 9053, a Hae III site absent at position 16517, and a Dde I site present at position 16478. HT1080 mtDNA is most closely related to type 56, differing from it by 0.17%.

Refinement of the Phylogenetic Tree

A phylogenetic tree relating the eight mtDNA types represented among the thirteen individuals studied appears in Figure 6. In tree I, based solely on restriction site polymorphisms, the two New Guinean mtDNA types with the conformational mutation at position 11254 (types 107 and 108) are not phylogenetically

Table 1. Sequence and Relative Mobility Information for 13 Human MtDNAs

Individual	Sequence[a] (11250-11260)	Relative Mobility	Geographic Origin	Type of MtDNA[b]
Standard[c]	TAAT TTACACT	unknown	Europe	167
HeLa	. . . . .	slow	Africa	61
1[d]	. . . . .	slow	South Africa	1
2	. . . . .	slow	S.Highlands[e]	109
3	. . . . .	slow	E.Highlands[f]	106
4	. . . . C	fast	E.Highlands	107
5	. . . . C	fast	E.Highlands	107
6	. . . . C	fast	E.Highlands	107
7	. . . . C	fast	E.Highlands	107
8	. . . . C	fast	E.Highlands	107
9	. . . . C	fast	E.Highlands	107
10	. . . . C	fast	E.Highlands	108
HT1080[g]	. . . C	faster	Europe[h]	---

[a] Dots indicate identity with the standard sequence.

[b] MtDNA type is numbered according to Stoneking (5), which differs from previous numbering systems (3,11,17).

[c] Standard is the published sequence (9).

[d] Individual 1 is the GM3043 cell line.

[e] Southern Highlands Province of Papua New Guinea.

[f] Eastern Highlands Province of Papua New Guinea.

[g] Although we confirmed the sequence reported earlier by Singh et al. (1) for HT1080, we did not confirm their report that mtDNA from GM6224 also had a C at position 11253 and an enhanced mobility; in our hands GM6224 mtDNA was identical to the slow type of sequence, as in individuals 1-3 and HeLa.

[h] European origin inferred because the cell line came from a Caucasian North American (10).

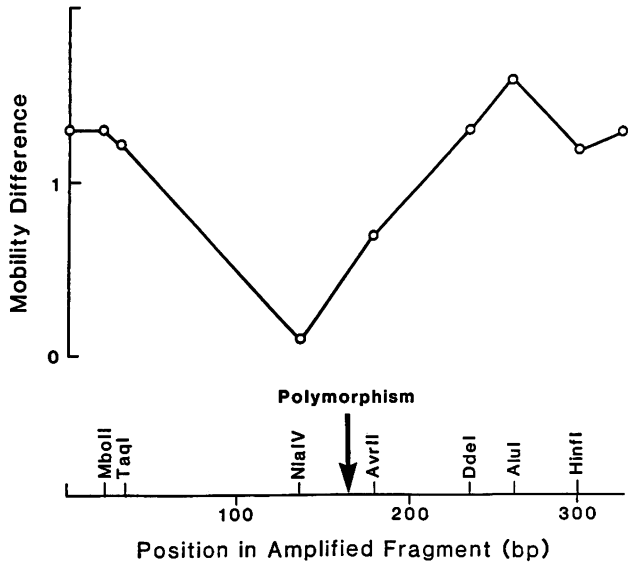

Figure 5 Mapping the locus of bending. Amplified 325-bp fragments from both normal (curved) mtDNA and variant (less curved) mtDNA were cut with 7 restriction enzymes and electrophoresed in native 6% polyacrylamide gels. Each enzyme cut once at the position indicated. Mobility difference (d) refers to the bigger resulting fragment and is defined by the equation d =100 (S - B) / B, where B is the distance migrated by the curved DNA and S is the corresponding value for the less curved DNA. The polymorphic base causing the mobility difference is at position 163 in the 325-bp fragment.

distinguishable from a New Guinean mtDNA type lacking this mutation (type 109). Incorporating the conformational mutation into the phylogenetic analysis allows this trifurcation to be resolved without introducing any other changes in the order of branching in the tree (Fig. 6, tree II).

DISCUSSION

Information for Bending DNA

The existence of genetic polymorphisms affecting the state of curvature of DNA can be a valuable aid to testing ideas about the structural requirements for such curvature. MtDNA provides the first cases of such polymorphisms. Our results are in agreement with postulates of the existence of a curvature in human mtDNA (1), with the substitution at position 11254 causing a less curved fragment that moves faster in nondenaturing polyacrylamide gels. The 325-bp segment of mtDNA of this report, while AT-rich at the apparent curving locus, does not contain any obvious AT sequence repeats of the kind usually correlated with curving (12, 13). The assertion

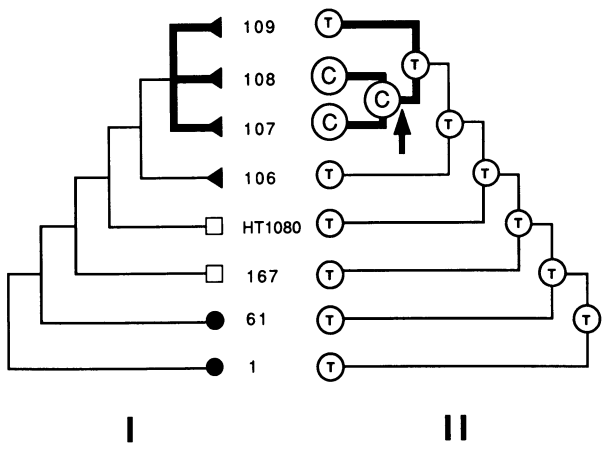

Figure 6 Two views of the phylogenetic relationships among eight types of mtDNA (identified in Table 1). Tree I is based only on the presence and absence of restriction sites (5, 11) while Tree II has a slightly different branching order because it incorporates sequence information concerning the conformational polymorphism at position 11254 (indicated by circled letters). Thick lines draw attention to the difference in branching order between the two trees. The arrow indicates the lineage on which the T -> C mutation at position 11254 is postulated to have occurred. Symbols indicate geographic origin: ▲- New Guinea; □- Europe; ●- Africa.

that curving occurs in the absence of sequence repeats is supported by a case recently described by Coll et al. (14) in which poly(dA)poly(dT) segments may stabilize curves in DNA by adopting a modified B-DNA conformation. In addition, random, single base pair mutagenesis of adenine tracts causing curves in SV40 DNA has shown that each base does not contribute equally to the curve (15). These observations and the recent description of an adenine tract that does not cause DNA to curve as predicted indicate that a general theory to explain DNA bending is still lacking (16).

Polarity of the Conformational Mutations

Since a phylogenetic tree of the 182 mtDNA types found in a survey of 241 individuals has been deduced from mapping approximately 370 restriction sites per genome (5, 11), we can determine the polarity of the T<->C mutation at position 11254 responsible for the mobility difference. The seven individuals bearing a C at this position are each other's closest relatives in the genealogical tree (Fig. 6), which is consistent with a single T -> C mutation. If one were to attempt to explain the distribution of C and T at 11254 shown in Figure 6 by C -> T mutations instead, five additional mutations would be required. On the grounds of parsimony, one is obliged to accept the simpler explanation, which is that the ancestral state is T and

the descendant state in the seven New Guineans is C. Similar reasoning applied to HT1080 type mtDNA implies that the ancestral state at 11253 is T. Hence the ancestral and prevalent form of mtDNA in this part of the ND4 gene is curved, while the mutant or descendant forms found in the seven New Guineans and HT1080 are straighter.

Ages of the Conformational Mutations

All seven of the New Guineans with the faster mobility form are from the Eastern Highlands Province, implying that the faster, less curved form arose in the Eastern Highlands from the conventional, curved form. Two mtDNA types (107 and 108) are represented among these seven individuals; the amount of sequence divergence (based on restriction mapping) among these individuals and between them and their nearest relative, type 109 (Fig. 6), can be used to estimate when the conformational mutation arose. The average sequence divergence is 0.01% among the seven individuals and 0.07% between them and type 109. The rate of human mtDNA sequence divergence has been estimated to be approximately 2-4% per million years (17). Thus, the conformational mutation arose at least 2500 - 5000 years ago, but less than 17500 - 35000 years ago, which is within archaeological estimates for human occupation of the New Guinea highlands (18). With a similar approach, we estimate that the T -> C change at position 11253 occurred no longer than 43000 to 86000 years ago in the lineage leading to HT1080 mtDNA.

Refinement of the Genealogical Tree

Singh et al. (1) suggest that conformational mutations can be mistaken for restriction site polymorphisms, thereby introducing errors into phylogenetic analysis and estimates of sequence divergence. They base this on their failure to confirm by sequencing four of the six restriction site differences reported by Cann et al. (11) between HeLa mtDNA and the published sequence (9). However, these unconfirmed site differences reflect mapping errors due to incorrect semisite choices, and hence do not alter estimates of sequence divergence or phylogeny. Furthermore, it is unlikely that a conformational mutation would be confused with a restriction site polymorphism, for the following reasons. Each restriction site polymorphism involves mobility alterations in three fragments resulting from digestion with a single enzyme. For example, a site loss causes two bands to disappear and a new band to appear, while a site gain causes one band to disappear and two new bands to appear. By contrast, conformational mutations, like length mutations, involve mobility alterations in just one fragment and are usually detected by more than one restriction enzyme (2). In fact, the conformational mutations can refine rather than discredit a genealogical analysis (Fig. 6).

A more likely source of error would be to include conformational mutations in a

phylogenetic analysis without first sequencing the mutant regions. If, for instance, we were to include the altered mobility of the ND4 fragment in a phylogenetic analysis, there would be a tendency to associate the HT1080 mtDNA with New Guinean types 107 and 108. By confining attention to restriction site polymorphisms and sequence information this potential source of error is avoided. For these reasons, we recommend sequencing of regions that harbor length or conformational mutations before including them in a tree analysis (3).

ACKNOWLEDGEMENTS

We thank K. Bhatia for supplying placental tissue and T. Kocher, R. Higuchi and D. Irwin for helpful discussions and E. Prager for editorial assistance. This research received support from grants from the National Science Foundation and the National Institutes of Health, an NSF graduate fellowship to L.V. and an NIH postdoctoral fellowship to M.S.

REFERENCES

1. Singh, G., Neckelmann, N. and Wallace, D.C. (1987) *Nature* **329**, 270-272.
2. Cann, R.L. and Wilson, A.C. (1983) *Genetics* **104**, 699-711.
3. Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. and Wilson, A.C. (1987) *Nucl. Acids Res.* **15**, 529-542.
4. Greenberg, B.D., Newbold, J.E. and Sugino, A. (1983) *Gene* **21**, 33-49.
5. Stoneking, M. (1986) *Human Mitochondrial DNA Evolution in Papua New Guinea.* Ph.D. Thesis. University of California, Berkeley.
6. Brown, W.M., George, M. and Wilson, A.C. (1979) *Proc. Natl. Acad. Sci.* USA **76**, 1967-1971.
7. Brown, W.M. (1980) *Proc. Natl. Acad. Sci.* USA **77**, 3605-3609.
8. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual.* pp. 156,174. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
9. 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.
10. Rasheed, S., Nelson-Rees, W.A., Toth, E.M., Arnstein, P. and Gardner, M.B. (1974) *Cancer* **33**, 1027-1033.
11. Cann, R.L., Stoneking, M. and Wilson, A.C. (1987) *Nature* **325**, 31-36.
12. Koo, H.S., Wu, H.M. and Crothers, D.M. (1986) *Nature* **320**, 501-506.
13. Anderson, J.N. (1986) *Nucl. Acids Res.* **14**, 8513-8533.
14. Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) *Proc. Natl. Acad. Sci.* USA **84**, 8385-8389.
15. Milton, D.L. and Gesteland, R.F. (1988) *Nucl. Acids Res.* **16**, 3931-3949.
16. Burkhoff, A.M. and Tullius, T.D. (1988) *Nature* **331**, 455-457.
17. Stoneking, M., Bhatia, K. and Wilson, A.C. (1986) *Cold Spring Harbor Sympos. Quant. Biol.* **51**, 433-439.
18. White, J.P. and O'Connell, J.F. (1982) *A Prehistory of Australia, New Guinea and Sahul.* Academic Press, New York.