PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Molecular instability in the *COII–tRNA^{Lys}* intergenic region of the human mitochondrial genome: multiple origins of the 9 –bp deletion and heteroplasmy for expanded repeats

Mark G. Thomas, Charles E. Cook, Kevin W. P. Miller, Michael J. Warin and Erika Hagelberg

Phil. Trans. R. Soc. Lond. B 1998 **353**, 955-965
doi: 10.1098/rstb.1998.0260

The Royal Society

Molecular instability in the *COII–tRNA^{Lys}* intergenic region of the human mitochondrial genome: multiple origins of the 9-bp deletion and heteroplasmy for expanded repeats

Mark G. Thomas[1*], Charles E. Cook[1†], Kevin W. P. Miller[1††], Michael J. Waring[2] and Erika Hagelberg[1‡]

[1]*Department of Biological Anthropology, University of Cambridge, Downing Street, Cambridge CB2 3DZ, UK*
[2]*Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, UK*

We have identified two individuals from Glasgow in Scotland who have a deletion of one of two copies of the intergenic 9-bp sequence motif CCCCCTCTA, located between the cytochrome oxidase II (*COII*) and lysine tRNA (*tRNA^{Lys}*) genes of the human mitochondrial genome. Although this polymorphism is common in Africa and Asia, it has not been reported in Northern Europe. Analysis of the mitochondrial DNA control region sequences of these two individuals suggests that they belong to a lineage that originated independently of the previously characterized African and Asian 9-bp deleted lineages. Among the Scottish population we have also identified a maternal lineage of three generations exhibiting heteroplasmy for two, three and four copies of the CCCCCTCTA motif. Polymerase chain reaction amplification across the *COII–tRNA^{Lys}* intergenic region of these individuals gives different ratios of the three product lengths that are dependent on the concentration of the DNA-binding dye crystal violet. To investigate whether changes in repeat number were generated *de novo*, we constructed clones containing known numbers of the CCCCCTCTA motif. In the presence of high concentrations of crystal violet we obtained two, three and four copies of this motif when the amplification template contained only four copies. Various DNA-binding drugs are known to stabilize bulged structures in DNA and contribute to the process of slipped-strand mispairing during DNA replication. These results suggest that the *COII–tRNA^{Lys}* intergenic region is unstable owing to slipped-strand mispairing. Although sequences containing four copies of the CCCCCTCTA motif are less stable *in vitro*, we observed an increase in the proportion of mitochondrial genomes with four repeats between a mother and a daughter in the heteroplasmic lineage. From this we conclude that drift in the germ-line lineage is a main factor in the maintenance or loss of heteroplasmy.

Keywords: human; mtDNA; 9-bp deletion; heteroplasmy; DNA-binding drugs; slipped-strand mispairing

1. INTRODUCTION

The *COII–tRNA^{Lys}* intergenic 9-bp deletion (Wrischnik *et al.* 1987) is one of the most commonly studied human mitochondrial DNA (mtDNA) polymorphisms. It consists of the loss of one of two tandemly repeated copies of the sequence CCCCCTCTA from a non-coding region located between the cytochrome oxidase II (*COII*) and lysine tRNA genes *tRNA^{Lys}*. The 9-bp deletion was originally thought to be specific to individuals of Asian origin (Ballinger *et al.* 1992; Harihara *et al.* 1992). Its occurrence in ancient (Hagelberg & Clegg 1993; Hagelberg *et al.* 1994) and modern (Hertzberg *et al.* 1989; Lum *et al.* 1994; Redd *et al.* 1995; Melton *et al.* 1995; Sykes *et al.* 1995) oceanic populations, the Tharus of Southern Nepal (Passarino *et al.* 1993), Malagasy (Soodyall *et al.* 1995) and some Amerind populations (Schurr *et al.* 1990; Horai *et al.* 1993; Torroni *et al.* 1992, 1993*a*; Bailliet *et al.* 1994; Merriwether *et al.* 1994) has been used to argue that these populations have a common origin in Asia (Ballinger *et al.* 1992; Wallace & Torroni 1992; Horai *et al.* 1993; Torroni *et al.* 1993*b*, 1995; Bailliet *et al.* 1994; Cann 1994; Monsalve *et al.* 1994). However, phylogenetic analyses of 9-bp deleted mtDNA haplotypes from Africa (Vigilant 1990; Redd *et al.* 1995; Soodyall *et al.* 1995, 1996) and Europe (Torroni *et al.* 1995), using control region (CR) sequences and high resolution restriction site analysis respectively, indicate that the deletion occurred more than once during human history.

*Author and address for correspondence: The Centre for Genetic Anthropology, Department of Biology, University College, Gower Street, London WC1E 6BT, UK (m.thomas@ucl.ac.uk).
†Present address: University Museum of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.
††Present address: North Louisana Criminalistics Laboratory, DNA Section, 1115 Brooks Street, Shreveport, LA 71101, USA.
‡Present address: Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

Received 14 February 1997 *Accepted* 4 June 1997

As well as a deletion of one of the 9-bp motifs, a triplication has been observed in a Chuckchi from Siberia (Shields *et al.* 1992) and a Tharu from Nepal (Passarino *et al.* 1993). The occurrence of these and other length polymorphisms (Wrischnik *et al.* 1987; Ballinger *et al.* 1992; Redd *et al.* 1995) suggests the existence of sequence length instability in this mtDNA region. Redd *et al.* (1995) proposed that the instability could be the result of slipped-strand mispairing (SSM) during DNA replication (Levinson & Gutman 1987; Richards & Sutherland 1994), although this has not been shown experimentally.

We report the occurrence of the 9-bp deletion in two Scottish individuals from Glasgow. In addition, we have observed heteroplasmy for two, three, and four copies of the 9-bp motif in three generations of a Scottish family. We present the first experimental evidence that the instability found in this region of the mitochondrial genome is caused by SSM during replication.

2. MATERIALS AND METHODS

(a) *Sample preparation*

DNA was extracted from buccal epithelial cells as follows: 15 ml of a 4% sucrose solution was used as a mouthwash for 15 s, then centrifuged at 1000 *g* for 10 min. The cell pellet was washed in 500 µl of 10 mM NaCl–10 mM EDTA and centrifuged at 1000 *g* for 5 min. The cell pellet was resuspended in 500 µl of 50 mM NaOH and the tubes incubated at 95 °C for 20 min to lyse the cells. Finally, 100 µl of 1 M Tris-HCl (pH 7.5) were added and the samples stored at 4 °C.

(b) *PCR amplification*

The presence–absence of the 9-bp deletion was determined using primers RVM1 (5′-AGGGCCCGTATTTACCCTATAG-3′) and RVM2 (5′-ATTTAGTTGGGGCATTTCACTG-3′) which amplify a 133-bp or a 124-bp product for non-deleted and deleted samples, respectively. Amplifications were done in 25 µl reaction volumes containing 0.075 µg of each primer, 200 µM dNTPs, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X-100, 0.02% gelatin, 0.15 units of *Taq* DNA polymerase (HT Biotechnology, Cambridge, UK) and 1 µl of DNA template. Polymerase chain reaction (PCR) parameters were: 35 cycles consisting of 94 °C for 1 min, 57 °C for 30 s and 72 °C for 30 s. Then 5 µl aliquots of the PCR products were electrophoresed on 3% agarose gels, stained with ethidium bromide, and visualized by ultraviolet transillumination. High-resolution analysis of length variation in the *COII–tRNA*^{Lys} intergenic region was done using (^{33}P-)end-labelled primer RVM1. PCR conditions were the same as before except that the reactions were done in 20 µl volumes containing 0.04 µg of each primer.

For the analysis of the mtDNA control region, a 1239-bp fragment of the mitochondrial genome was amplified using a biotinylated version of primer HVM1 (5′-biotin-CTAACCTGAATCGGAGGACAAC-3′) and primer HVM4 (5′-GCATACCGCCAAAAGATAAAA-3′). The PCR was done as before except that 0.3 µg of each primer and 1 unit of *Taq* polymerase were used in 100 µl reactions. Cycling parameters were: pre-incubation at 95 °C for 60 s followed by 35 cycles of 94 °C for 60 s, 56 °C for 60 s, 72 °C for 90 s, with an additional 1 s per round auto-extension; then a final incubation step of 72 °C for 10 min. PCR products were stored at 4 °C. Before sequencing, the products were extracted once with 100 µl of chloroform to remove mineral oil. Then 5 µl aliquots of each reaction were electrophoresed on 0.7% agarose gels and visualized as before.

(c) *DNA sequencing*

Single-stranded sequencing templates were prepared from PCR products by using streptavidin-coated magnetic beads (Dynal, Merseyside UK) to bind the biotinylated strand of the PCR product, then both strands were recovered (Thomas *et al.* 1996). The Dynabead–DNA complex was resuspended in 21 µl of TE (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) and 7 µl used for each of three sequencing reactions. The unbiotinylated strand was resuspended in 7 µl of TE and used for one sequencing reaction. A stringency wash at 55 °C was done on biotinylated templates (Thomas *et al.* 1994). Sequencing primers were L15996, H16498 (Kocher *et al.* 1989), MT1 (5′-TGGTTAGGCTGGTGTTA-3′), and PMT1 (5′-biotin-CTCACCCATCAACAACCGCTAT-3′). Sequencing was done using the Sequenase kit (USB, Cleveland, Ohio) except that each reaction contained 0.02 µl of labelling mix, 0.083 µl of 10 µCi–µl ^{35}S-dATP (Amersham, UK), 0.0625 µl of pyrophosphatase and 0.125 µl of Sequenase. Sequencing products were run on 6% acrylamide gels (National Diagnostics, Hull, UK), fixed in 7% acetic acid, 7% methanol, dried onto filter paper and exposed to X-ray film (Kodak X-Omat LS) for 3–30 d.

Sequences were recorded by using the UWGCG program SEQED and aligned using the programs PILEUP and PRETTY (Program Manual for the Wisconsin Package 1994). Phylogenetic analysis was performed using the PHYLIP package (Felsenstein 1995). A distance tree was constructed using the DNADIST and NEIGHBOR programs. Distances were calculated according to the two-parameter method of Kimura (1980), assuming a transition–transversion ratio of ten, and the phylogenetic tree constructed using the neighbour-joining algorithm with the chimpanzee control region (CR) as an outgroup (Horai *et al.* 1995). The reliability of the interior nodes was estimated by bootstrap resampling 100 times (Felsenstein 1975). Within and between population nucleotide diversities were calculated using the program IWAVE (S. Sherry, Department of Anthropology, Pensylvania State University). Corrected intermatch distance was calculated according to Nei & Li (1979).

(d) *Cloning of PCR products*

PCR products from amplification of the *COII–tRNA*^{Lys} intergenic region were blunt-end cloned into a plasmid vector (pBSSKII+, Stratagene) and transformed into *Escherichia coli* DH5α host cells by electroporation. Recombinant plasmid clones were screened for inserts of the appropriate size by PCR amplification with the T3 and T7 primers that flank the multiple cloning site of the vector. Clones containing no insert produced a PCR product of 197 bp whereas clones with inserts produced PCR products of 332–350 bp.

PCR products from clones containing inserts were sequenced by dye terminator cycle sequencing with *Taq*

polymerase and run on an ABI automated sequencer. All clones were sequenced in both directions with the T3 and T7 primers. At least six clones from each PCR product were sequenced.

3. RESULTS

(a) *Samples*

A total of 115 individuals from Scotland, including 76 from the Orkney islands and 39 from Glasgow, were surveyed for length variation in the *COII–tRNA*^{Lys} intergenic region. We detected two individuals from Glasgow (numbers 19 and 34) with the 9-bp deletion, and an additional individual (number 39), also from Glasgow, who appeared to carry an insertion in the *COII–tRNA*^{Lys} intergenic region. Individuals 19 and 34 were of Scottish descent on the maternal line for at least two generations while individual 39 could trace her maternal Scottish ancestry for five generations. Analysis of DNA samples from the mother, maternal grandmother, and maternal aunt of individual 39 revealed that they also had an insertion in the *COII–tRNA*^{Lys} intergenic region.

To characterize this insertion, DNA of individual 39 and her maternal relatives was amplified with a (^{33}P-)labelled primer and the products were run on 6% polyacrylamide sequencing gels (see figure 1*a*). All the maternal relatives of individual 39 showed a multiple banding pattern with fragment sizes of 133, 142 and 151 bp. There were also lower intensity bands of intermediate sizes. Although the ratios of the three main fragments in the mother, grandmother and aunt were similar to one another, they differed from that of individual 39 (figure 2). One 9-bp deleted and one non-deleted sample (individuals 19 and 20) were run for comparison, and each produced a single band. The occurrence of multiple bands suggested that individual 39 and her maternal relatives were heteroplasmic. However, a similar banding pattern could also be the consequence of repeat number instability during amplification (Behnkrappa & Doerfler 1994).

DNA sequencing of the PCR products from individual 39 and her maternal relatives gave poor results (not shown) owing to the multiple banding pattern observed. However, the sequencing ladder did suggest that the 142-bp and 151-bp fragments differed from the 133-bp fragment by the insertion of extra copies of the 9-bp motif. The PCR products were then cloned into a plasmid vector. These clones yielded unambiguous sequence information and showed that the three different PCR bands obtained from individual 39 were the result of amplification of templates with two, three and four copies of the CCCCCTCTA motif.

To investigate the possibility that the length differences in the PCR products from genomic DNA might be resulting from errors in PCR and not from heteroplasmy, we amplified the *COII–tRNA*^{Lys} intergenic region in plasmids with two, three, and four copies of the CCCCCTCTA motif by using five different concentrations of plasmid DNA as PCR templates (see figure 1). Assuming that the likelihood of a replication error of the 9-bp motif is constant for every replication event, the ratio of erroneous products to accurate PCR products will depend on the average number of replication events undergone by each molecule. We predict that lower initial template concentrations increase the number of PCR rounds necessary to reach the plateau stage, increasing the average number of replication events, and therefore increasing the ratio of erroneous products to accurate products.

Whereas PCR amplification of clones containing two and three repeats of the CCCCCTCTA motif gave rise to single-sized products, the amplification products from diluted plasmid DNA containing four repeats of the CCCCCTCTA motif were of several sizes (figure 1*b*). This multiple banding was less pronounced than if PCR was done directly on the genomic DNA of individual 39. Thus, although PCR amplification of DNA sequences with four copies of the CCCCCTCTA motif gives rise to products of more than one size, this does not account for the banding pattern seen with individual 39 and her maternal relatives. We conclude that these individuals are heteroplasmic for two, three and four copies of the CCCCCTCTA motif.

(b) *Effects of DNA-binding drugs on copy-number instability of the CCCCCTCTA motif during the PCR*

It has been proposed that repeat number instability in homopolymeric regions (Bendall & Sykes 1995; Greenblatt *et al.* 1996), simple repetitive DNA (SR-DNA) sequences (Tran *et al.* 1995), longer repeats (Wilkinson & Chapman 1991), and regions containing hairpins and other secondary structures (Rosche *et al.* 1995), are due to the phenomenon of slipped-strand mispairing (Levinson & Gutman 1987; Richards & Sutherland 1994). The formation of bulges of single-stranded DNA, hairpin loops, or triplex structures during DNA replication is an essential intermediate in the SSM mechanism. Previous studies have shown that intercalators can sometimes stabilize such structures (Rentzeperis *et al.* 1995; Tuite & Norden 1995; Zhou *et al.* 1995; Caceresortes & Wang 1996) and contribute to repeat number instability (Demarini *et al.* 1992), and it has long been suspected that such processes might underlie the phenomenon of drug-induced frameshift mutagenesis (Streisinger *et al.* 1966; Imada *et al.* 1970; Gale *et al.* 1981).

To gain an insight into the possible mechanism of repeat number instability of the CCCCCTCTA motif during the PCR, we investigated the effects of different concentrations of the well-known intercalator ethidium bromide (EtBr) (Rentzeperis *et al.* 1995) on the PCR amplification of DNA templates containing different numbers of repeats from individual 39, and of templates containing known numbers of repeats (clones). We elected also to investigate the effects of crystal violet (CV), another DNA-binding drug which interacts with nucleic acids by diverse mechanisms but not by classical intercalation (Wakelin *et al.* 1981). Figure 1*c* shows amplification products of template DNA from individuals 33, 34 and 39 in the presence of various concentrations of the drugs. Both EtBr and CV increase the ratio of PCR products with two repeats of the CCCCCTCTA motif to those containing three and four repeats. Inhibition of PCR occurred at between 20 µM and 40 µM EtBr and between 250 µM and 500 µM CV. A more pronounced effect on the ratios of different-sized products was obtained when using sub-inhibitory concentrations of CV than with sub-inhibitory concentrations of EtBr. This was possibly because CV was less inhibitory to

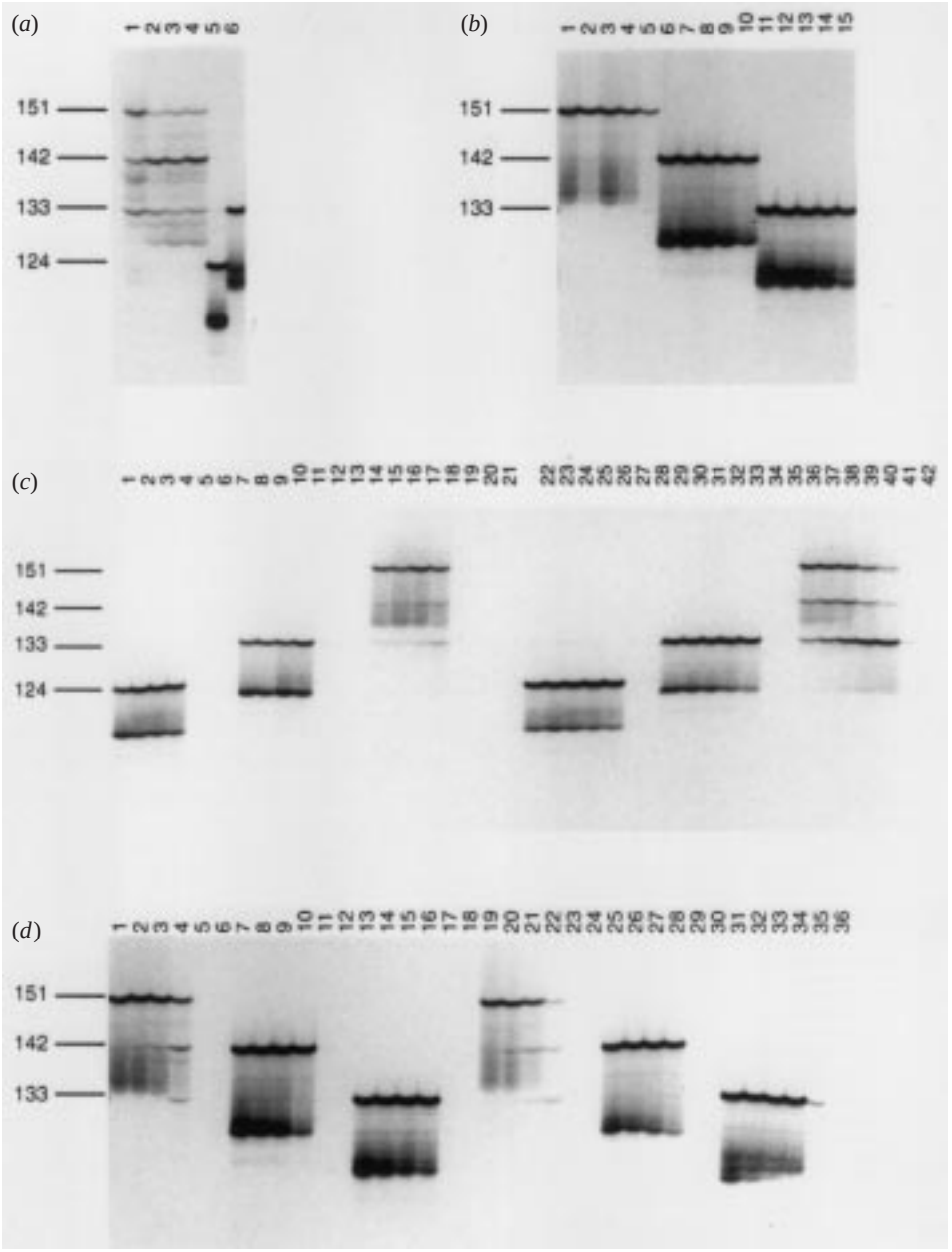

Figure 1. PCR products amplified by using primers RVM1 and RVM2 with different template DNAs. Primer RVM1 was end-labelled with ^{33}P and all products were run on 50-cm long 6% denaturing polyacrylamide gels. Lengths (bp) of PCR products are indicated to the left of each photograph and were estimated by using a sequencing ladder as a molecular weight marker (not shown). The occurrence of extra bands of smaller apparent size to the correct size band was also observed, particularly in those reactions for which the main PCR products were either 124 or 133 bp in length. These bands are probably caused by re-annealed, double stranded DNA because in each case, the apparent size is proportional to the size of the main (denatured) PCR product. (*a*) A total of six templates from six different individuals: lanes 1–6 are, in order, individual 39, her mother, her aunt, her grandmother, individual 19, and individual 20. (*b*) PCR amplification of a dilution series of cloned plasmid DNA containing four (lanes 1–5), three (lanes 6–10), and two (lanes 11–15) copies of the CCCCCTCTA motif. Plasmid DNA was prepared by alkaline lysis and centrifugation of cellular debris to a concentration of approximately 0.05 μg μl^{-1}. Of the following dilutions 1 μl was used in each of five 25 μl PCR reactions; 1:10, 1:100, 1:1000, 1:10 000, 1:100 000, for each of the three templates. (*c*) PCR amplification of DNA extracted directly from individuals 19 (lanes 1–7, 22–28), 20 (lanes 8–14, 29–35), and 39 (lanes 15–21, 36–42) in the presence of EtBr (lanes 1–21) and crystal violet (lanes 22–42). Each set of seven PCR reactions used the following concentration series: for EtBr 0, 5, 10, 20, 40, 80, 160 μM and for crystal violet 0, 31.25, 62.5, 125, 250, 500, 1000 μM. (*d*) PCR amplifications of DNA extracted from plasmid clones containing four (lanes 1–6, 19–24), three (lanes 7–12, 25–30), and two (lanes 13–18, 31–36) copies of the CCCCCTCTA motif. In total, two sets of dilutions of template DNA were used. Lanes 1–18 used 1 μl of a 1:10 dilution of plasmid DNA in each PCR reaction. Lanes 19–36 used 1 μl of a 1:1000 dilution as the PCR template. Each set of six PCR reactions used the following concentrations of crystal violet: 0, 62.5, 125, 250, 500, 1000 μM.

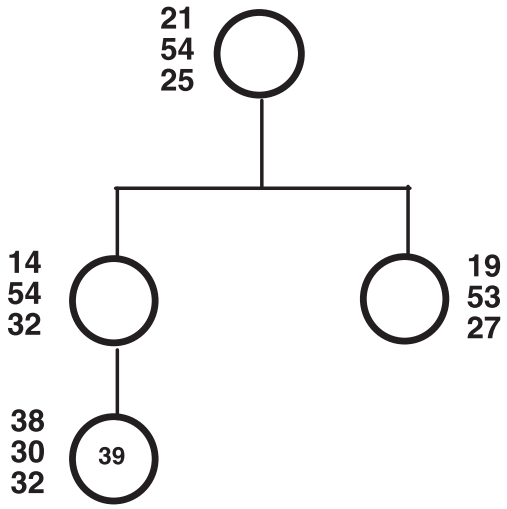

Figure 2. Pedigree showing percentage of PCR products with four, three, and two copies (from top to bottom) of the CCCCCTCTA motif in individual 39, her mother, her aunt, and her grandmother. The figures were calculated by comparing the relative densities of the bands for each product on figure 1*a* lanes 1, 2, 3, and 4 by using the program NIH image (available from `http://rsb.info.nih.gov/nih-image/`).

the polymerase. We observed inhibition of amplification between 250 µM and 500 µM CV using DNA extracted from clones containing two, three and four tandem copies of the motif (figure 1*d*). At sub-inhibitory concentrations of CV, deletions of the four repeated sequences occurred in a dose-dependent manner, although no deletions were observed by using template DNA containing three and two repeats of the CCCCCTCTA motif.

(c) *MtDNA control region sequences associated with the 9-bp deletion*

In total, two regions of the mtDNA-CR (hypervariable regions I and II) have been widely used to study relationships among human populations. Accordingly, CR sequences were obtained for individuals 19, 34 and 39 between mtDNA positions 16 020 and 16 478 (hypervariable region I), and between positions 84 and 327 (hypervariable region II) (Anderson *et al.* 1981). These were compared with 30 CR sequences from individuals with the 'Asian' 9-bp deletion (Redd *et al.* 1995), 22 sequences of Africans with the 9-bp deletion (Soodyall *et al.* 1996), 81 sequences of unrelated British individuals (Piercy *et al.* 1993) and the Cambridge reference sequence (Anderson *et al.* 1981). Sequences between positions 16 024 and 16 371, and between positions 53 and 302 were used for phylogenetic analysis. The use of these two regions maximized the number of variable sites used in the phylogenetic analysis while minimizing the number of sites for which data were missing. Figure 3 shows the sequence for this region from all 9-bp deleted individuals and from Glasgow individual 39.

The two individuals from Glasgow with the 9-bp deletion (19, 34) shared transversions at positions 16 306 and 16 332 when compared to the reference sequence. Transitions were also observed at positions 16 327 and 152 in individual 19, and at positions 16 126, 16 189, 16 294, 16 296, 16 298 and 195 in individual 34. The two transversions common to 19 and 34 were not detected in any of the other 9-bp deleted sequences shown in figure 3, nor were they found in 100 British individuals (Piercy *et al.* 1993) or in 920 human CR sequences in the GenBank–EMBL sequence database. The sharing of the two rare transversions therefore suggest a common origin for the 9-bp deletion in these two individuals. Assuming a CR mutation rate of 1.142×10^{-7} per site per year per mtDNA lineage (Stoneking *et al.* 1992), the eight site differences observed between individuals 19 and 34 would date their most common ancestor to about 50 000 years before present. Although estimates of coalescence based on only two sequences are susceptible to large errors, it is interesting to note that the estimated date is earlier than the arrival of modern humans in Europe. However, the large number of differences between these two sequences could also occur if the three adjacent changes at postions 16 294, 16 296 and 16 298 were linked by a single mutational event.

Table 1. *Mismatch, intermatch and corrected intermatch distances within and between populations*

(Mean within- (diagonal) and between- (above diagonal) population nucleotide diversities, together with the corrected between-population (below diagonal) distances for three populations of 9-bp deleted individuals only (African, Asian and Glasgow) and a British population. The corrected between-population distance for two populations was calculated by subtracting the mean of the nucleotide diversity within each population from the between-population nucleotide diversity (Nei & Li 1979).)

	African ($n=22$)	Asian ($n=30$)	British ($n=81$)	Glasgow ($n=2$)
African	**8.69**	18.88	18.36	19.66
Asian	11.88	**5.31**	9.52	11.45
British	10.07	2.92	**7.89**	8.98
Glasgow	11.31	4.80	1.04	**8.00**

(d) *Phylogenetic, mismatch and intermatch analysis*

Phylogenetic, mismatch, and intermatch analyses were done to determine whether the CR sequences of the two individuals from Glasgow with the 9-bp deletion were more closely related to other British sequences or to other 9-bp deleted sequences. Figure 4 shows a neighbour-joining tree of the CR sequences of 9-bp deleted African and Asian individuals, the two 9-bp deleted individuals from Glasgow, individual 39 (with four copies of the 9-bp motif) and 81 other British sequences (Piercy *et al.* 1993). The tree shows that the Asian and African 9-bp deleted sequences form distinct clusters, and the two Glasgow sequences together are part of a broader group containing only British sequences. The African sequences form three clades with strong statistical support as indicated by the bootstrap test: cluster 1, consisting of African sequences 1–18 (92%); cluster 2, with African sequences 19 and 20 (78%); and cluster 3, with African sequences 21 and 22 (86%). The bootstrap estimates are in good agreement with those of Soodyall *et al.* (1996). The Asian 9-bp deleted sequences form a single clade within a larger group of British sequences, but there is no statistical support for this grouping (7% bootstrap value). Support for the clade

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

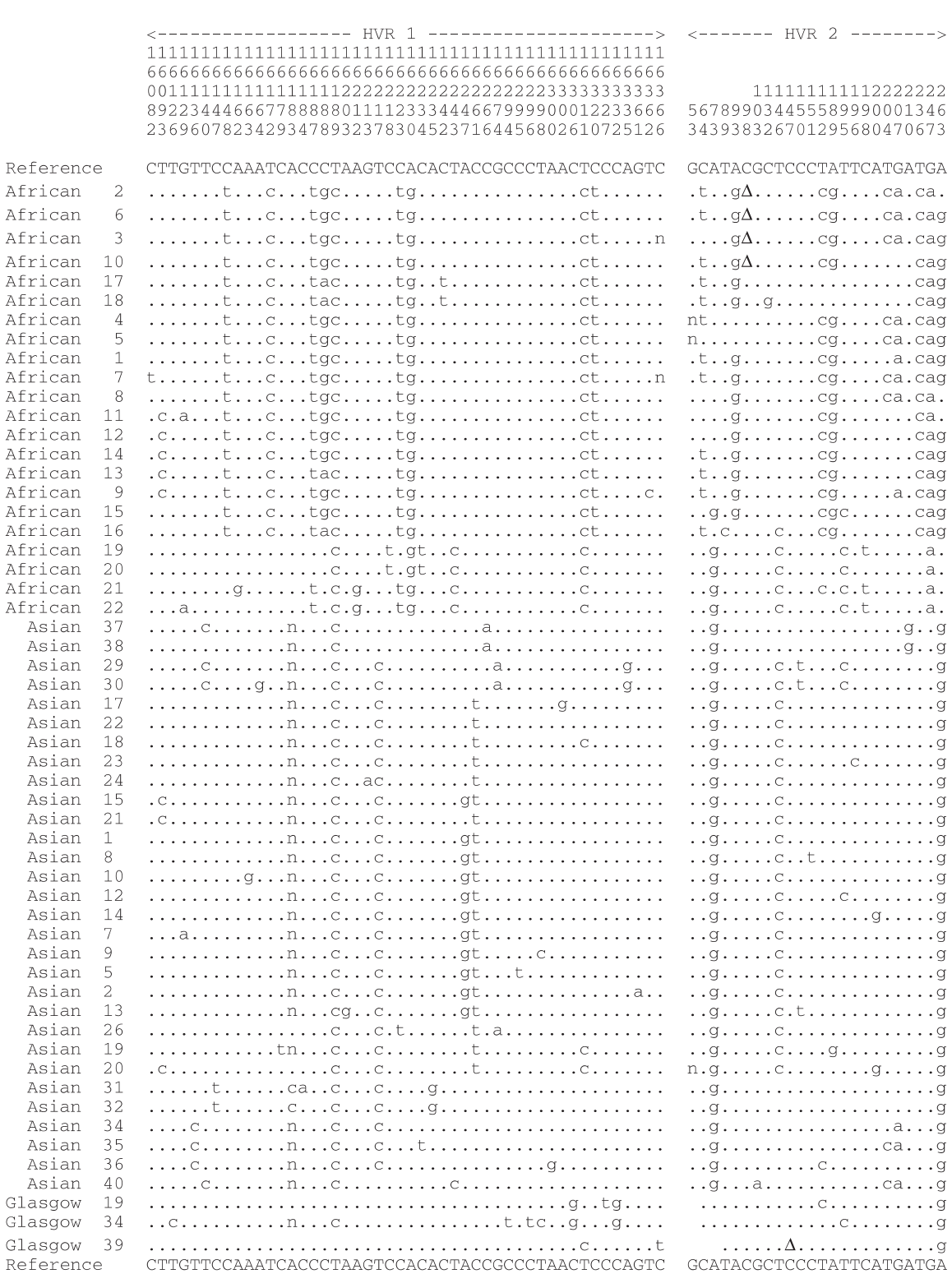

Figure 3. Control region sequence differences among mtDNAs with either repeat number heteroplasmy in the *COII–tRNA*^{Lys} intergenic region (Glasgow 39) or a 9-bp deletion (all others). Positions are numbered according to the reference sequence (Anderson *et al.* 1981). Variable sites only are shown; dots denote identity with the reference sequence; a 'Δ' indicates a deletion at that site; no dot indicates no sequence data; *n* indicates ambiguous sites. African sequences are those described by Soodyall *et al.* (1996) with numbering according to table 2 of that paper. Asian sequences are those described in Redd *et al.* (1995) with numbering according to figure 4. HVR1 and HVR2 refer to hypervariable regions 1 and 2.

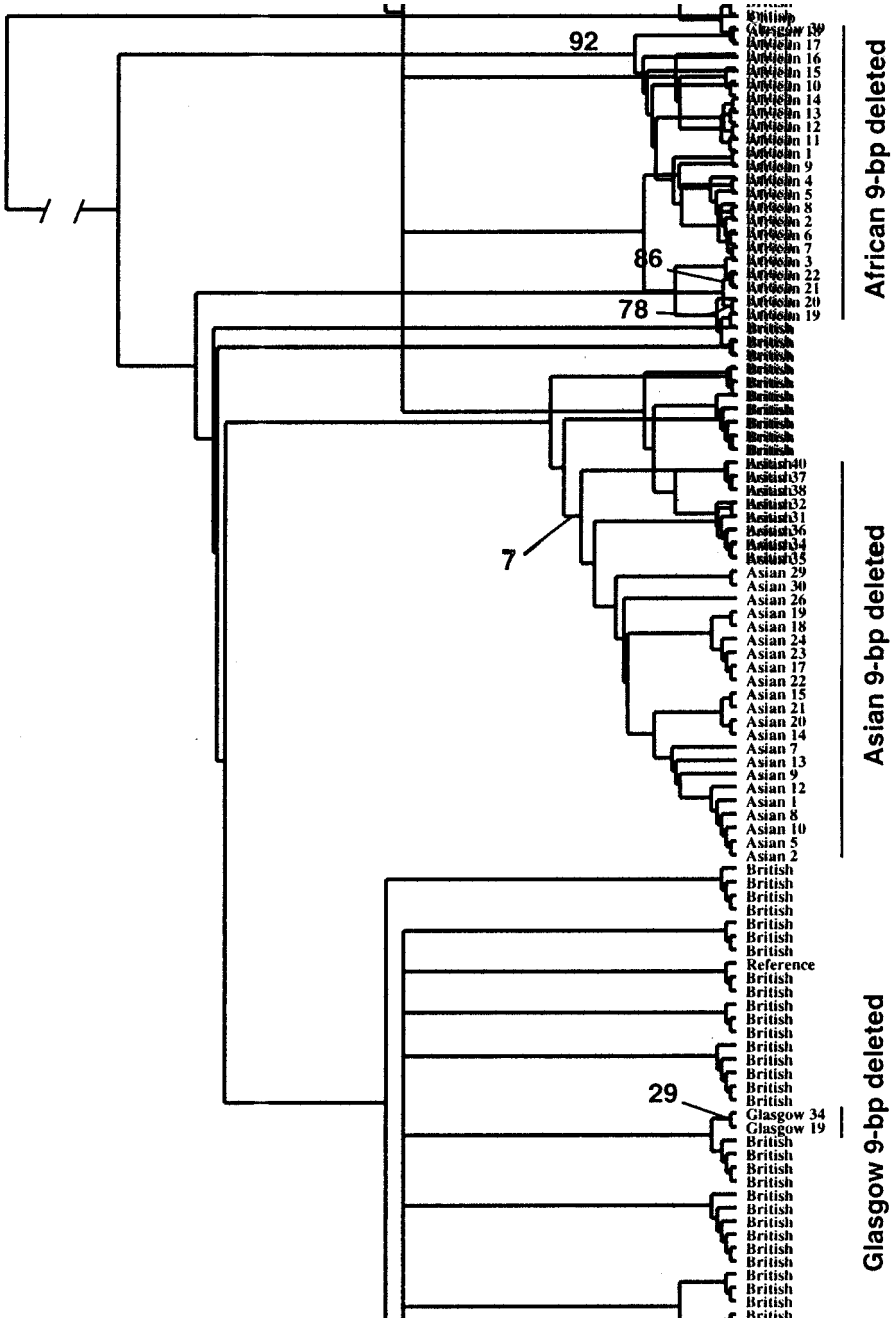

Figure 4. Consensus neighbour-joining tree of control region sequences from 22 African (Soodyall *et al*. 1996), 30 Asian (Redd *et al*. 1995), 81 British (Piercy *et al*. 1993), three Scottish individuals, and the Cambridge reference sequence (Anderson *et al*. 1981). Distances were calculated as described in the text and the consensus tree constructed by using the programs CONSENSE and DRAWGRAM (Felsenstein 1995). Numbers on nodes represent number of times that node was supported in 100 bootstrap replicates of the data.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

containing only the two Glasgow samples is also weak (29%). The low degree of support for these clades may be due to the relatively small distances between the sequences, and particularly between the British and the Asian sequences. This is reflected in the relatively low mean mismatch and intermatch distances between each group of sequences (table 1). Whereas the African sequences give large intermatch and corrected intermatch (Nei & Li 1979) distances with all other groups, the intermatch, and particularly the corrected intermatch distances between the British group and both the Asian and the Glasgow 9-bp deleted groups are low (2.92 and 1.04, respectively). However, the corrected intermatch distance between the Asian and Glasgow 9-bp deleted groups is somewhat higher (4.80), and suggests separate origins for the Asian and Glasgow 9-bp deleted lineages.

4. DISCUSSION

The data presented here suggest that the CR sequences of the two 9-bp deleted individuals from Glasgow are not typical of those sequences found in 9-bp deleted individuals from Asia and Africa. According to Soodyall *et al.* (1996), several lines of evidence can be used to argue for an independent origin of this lineage: first, the two deleted individuals share two C→G transversions (at positions 16 306 and 16 332) not seen before in published CR sequences. The sharing of these rare polymorphisms suggests that the 9-bp deletion in these two individuals stems from a single mutation event. The CR sequences of these individuals differ by eight nucleotides, suggesting that their common ancestor was not recent and probably lived before the arrival of modern humans in Europe. Second, phylogenetic analysis indicates that individuals 19 and 34 group more closely with other British than with either Asians or Africans with the 9-bp deletion. Although bootstrap support for the all-African clusters are high (92%, 86% and 72%), support for an all-Asian cluster and an all-Glasgow-9-bp deleted cluster is low. This may be because of a number of factors: (i) the inclusion of a large number of diverse British sequences in the analysis; (ii) the limited variability in human CR sequences; and (iii) the tendency for the bootstrap method to underestimate statistical support for groupings (Zharkikh & Li 1992).

Pairwise sequence differences within (mismatch) and between (intermatch) groups provide a third line of evidence supporting the independent origin of the 9-bp deleted lineage in Glasgow. The mean intermatch distances between Africans and Asians, and between Africans and the Glasgow group, are more than twice the magnitude of the mean mismatch distance within any group, and are highest between the African and Glasgow groups. This indicates that the Glasgow–African difference is at least as great as the Asian–African difference and strongly supports an independent origin for the African group (Soodyall *et al.* 1996). The mean Glasgow–Asian intermatch distance, although only 1.44 times the mean mismatch distance within each group, is still greater than the mean intermatch distance between each population and the British sample. Thus, against a background of diverse British CR sequences, the Glasgow sequences are even less like the Asian sequences than expected.

Although the evidence for the independent origin of the 9-bp deleted lineage in Glasgow is compelling, it is not conclusive. Given the close similarity between all of the European and Asian CR sequences included in our analysis it is difficult to envisage a conclusive proof of independent origins by using CR sequences alone. The 9-bp deleted sequences from Glasgow are typically British, as indicated by the low corrected intermatch distance (Nei & Li 1979) between the two groups (1.04). If, as our data suggest, the deletion event occurred before the arrival of modern humans in Europe, then this polymorphism should be present elsewhere in Europe and we suggest that it would be rewarding to do additional surveys for the deletion.

The occurrence of repeat number instability in the 9-bp motif is suggested by three lines of evidence: (i) multiple independent deletion of one of the two copies of the motif in different populations (Vigilant 1990; Redd *et al.* 1995; Soodyall *et al.* 1995, 1996; Torroni *et al.* 1995; present study); (ii) the occurrence of triplications and quadruplications of the motif (Shields *et al.* 1992; Passarino *et al.* 1993; present study); and (iii) insertions other than that of the entire CCCCCTCTA motif in the *COII–tRNA*^{Lys} intergenic region (Wrischnik *et al.* 1987; Ballinger *et al.* 1992; Redd *et al.* 1995). We have also shown that repeat number instability can be demonstrated *in vitro* using *Taq* polymerase as a replication enzyme in the presence of DNA-binding drugs.

While it can be argued that the presence of ligands introduces a bias in the ability to amplify shorter repeats from DNA templates of mixed lengths, it is doubtful whether this alone can explain the multiple banding pattern observed by using a DNA template of known size derived from a plasmid clone. The simplest explanation of how intercalators cause repeat number instability is that they stabilize the formation of bulges of single-stranded DNA during DNA replication and thus increase the rate of SSM. The induced deletions occurring during the PCR amplification of a cloned sequence containing four repeats of the CCCCCTCTA motif appear to be of discrete 9-bp blocks. We do not, however, observe PCR products of such length as to indicate that they contain only one repeat of the CCCCCTCTA motif, regardless of how many repeats the template DNA contains. It may be that the background signal, probably caused by re-annealed double-stranded DNA (figure 1), obscures the one-repeat band on the gels. Alternatively, this could reflect the relative mutation rate of four repeats to three and two repeats above the rate at which four, three and two repeats mutate to one repeat. A third explanation is that DNA-binding drugs not only increase the occurrence of SSM but affect the form SSM takes by preferentially stabilizing different bulged structures during polymerization. This may explain why the phenomenon appears more marked with CV than EtBr, for CV undoubtedly engages in 'outside binding' to DNA more readily than does the classical intercalator EtBr (Wakelin *et al.* 1981).

We have shown that the multiple banding pattern obtained from individual 39 and her maternal relatives is owing, at least in part, to heteroplasmy. It is interesting to note that the ratio of mitochondrial genomes

containing four and two repeat elements over those containing three repeats is higher in individual 39 than in her mother or grandmother. It is difficult to account for this observation by the *de novo* generation of mitochondria containing two and four repeat elements from those containing three repeat elements because our experiments suggest that it is chiefly the four-repeats sequence that is unstable during replication. Although it should be remembered that our experiments using DNA-binding ligands only mimic general repeat number instability, and may not model how repeat number instability occurs *in vivo*, the most probable explanation for the observed ratios in individual 39 is drift in the germ line lineage. Ashley *et al.* (1989) have demonstrated rapid segregation in Holstein cows (sometimes within two or three generations) of mitochondria heteroplasmic for a point mutation. Rapid segregation of point mutation heteroplasmy has also been recently demonstrated in mice (Jenuth *et al.* 1996) and humans (Bendall *et al.* 1996; Ivanov *et al.* 1996). However, in general, length heteroplasmies have not been shown to segregate in this manner. Rabbits (Casane *et al.* 1994), carnivores (Hoelzel *et al.* 1994), shrews (Fumagalli *et al.* 1996) and crickets (Harrison *et al.* 1985) have been shown to maintain extensive length heteroplasmy from one generation to the next and the proportions of length variants in a homopolymeric region of the human CR are thought to be maintained, even in distantly related individuals (Bendall & Sykes 1995). A notable exception to this rule comes from the observation that nucleotide repeat number differences in the homopolymeric tract around position 310 of the human mtDNA sequence differ significantly between different mature oocytes from the same individual (Marchington *et al.* 1997).

Our results show a marked change in the proportion of length variants over a single generation of heteroplasmic individuals, although loss of heteroplasmy was not observed. There are three mechanisms that are likely to act simultaneously to determine the maintenance or rate of loss of mtDNA heteroplasmy (Casane *et al.* 1994): (i) genetic drift in the germ line; (ii) the mutation rate for a particular polymorphism; and (iii) any selective constraints on mutations that are either deleterious or cause replication inefficiency. The rate of genetic drift will be determined solely by the effective population size (N_e) of mtDNA molecules in the germ line lineage. Previous studies (Piko & Matsumoto 1976) have shown that mouse eggs contain an average of $92\,500 \pm 7000$ mitochondria. If this were to represent the effective population size of mtDNA molecules in the germ line lineage in other mammals, it would be difficult to account for the rapid segregation of point mutation heteroplasmy observed (Ashley *et al.* 1989; Bendall *et al.* 1996; Ivanov *et al.* 1996; Jenuth *et al.* 1996) without invoking extreme selection. Given that most of the polymorphisms studied are common and have no obvious effects on fitness, it seems likely that the bottleneck in the germ line mtDNA population does not lie at the egg cell stage. Jenuth *et al.* (1996) have recently demonstrated that this bottleneck does indeed occur before oocytes are formed. In experiments on mice with heteroplasmic mitochondrial genomes they showed that variation in the proportion of different mtDNA types is high between different primary and mature oocytes, but that it is low between primordial germ cells from the same individual. They concluded that the mtDNA bottleneck occurs primarily in the oogonia as these cells differentiate into primary oocytes, and estimate that there are about 200 segregating mtDNA units in mice.

Regardless of when a bottleneck occurs, the process of drift should be equal for point mutations and length polymorphisms alike. In the absence of any obvious selection effects, maintenance of length heteroplasmy is most probably caused by high mutation rates associated with repetitive DNA (Hoelzel *et al.* 1991).

In this study, we have not addressed the effect of different repeat numbers on general fitness or replication efficiency. We suggest that studies of the effects of DNA-binding drugs on mtDNA *in vivo* would aid our understanding both of the mechanisms of repeat number instability and of the consequence of different numbers of repeats on fitness. Such work, in conjunction with the recent estimates of effective population size of mtDNA molecules (Bendall *et al.* 1996; Jenuth *et al.* 1996) in the germ line lineage might allow the construction of a model for the maintenance or loss of heteroplasmy.

We thank C. Chandmal for help in labelling primers and L. Dickie and D. Smyth for help in collecting samples. M.G.T. is supported by Wellcome Trust grant 038052Z–93–Z and C.E.C. is supported by NERC–ABI grant GST–02–825, both awarded to E.H. Wellcome Trust grant 040049–Z–93 supported M.J.W.

REFERENCES

Anderson, S. (and 13 others) 1981 Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Ashley, M. V., Laipis, P. J. & Hauswirth, W. W. 1989 Rapid segregation of heteroplasmic bovine mitochondria. *Nucl. Acids Res.* **17**, 7325–7331.

Bailliet, G., Rothhammer, F., Carnese, F. R., Bravi, C. M. & Bianchi, N. O. 1994 Founder mitochondrial haplotypes in Amerindian population. *Am. J. Hum. Genet.* **55**, 27–33.

Ballinger, S. W., Schurr, T. G., Torroni, A., Gan, Y. Y., Hodge, J. A., Hassan, K., Chen, K. H. & Wallace, D. C. 1992 Southeast-Asian mitochondrial-DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics* **130**, 139–152.

Behnkrappa, A. & Doerfler, W. 1994 Enzymatic amplification of synthetic oligodeoxyribonucleotides—implications for triplet repeat expansions in the human genome. *Hum. Mutat.* **3**, 19–24.

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.

Bendall, K. E., Macaulay, V. A., Baker, J. R. & Sykes, B. C. 1996 Heteroplasmic point mutations in the human mtDNA control region. *Am. J. Hum. Genet.* **59**, 1276–1287.

Cacerescortes, J. & Wang, A. H. J. 1996 Binding of the antitumor drug nogalamycin to bulged DNA structures. *Biochemistry* **35**, 616–625.

Cann, R. L. 1994 MtDNA and native-Americans—a southern perspective. *Am. J. Hum. Genet.* **55**, 7–11.

Casane, D., Dennebouy, N., de Rochambeau, H., Mounolou, J. C. & Monnerot, M. 1994 Genetic analysis of systematic mitochondrial heteroplasmy in rabbits. *Genetics* **138**, 471–480.

Demarini, D. M., Abushakra, A., Gupta, R., Hendee, L. J. & Levine, J. G. 1992 Molecular analysis of mutations induced by the intercalating agent ellipticine at the hisd3052 allele of *Salmonella–Typhimurium* ta98. *Environ. Molec. Mutag.* **20**, 12–18.

Felsenstein, J. 1975 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Felsenstein, J. 1995 *PHYLIP: Phylogenetic Inference Package, version 3.57c*. Distributed by author. Department of Genetics, University of Washington, Seattle, WA, USA.

Fumagalli, L., Taberlet, P. & Favre, L. H. J. 1996 Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. *Molec. Biol. Evol.* **13**, 31–46.

Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H. & Waring, M. J. 1981 *The molecular basis of antibiotic action*, 2nd edn. London: Wiley.

Genetics Computer Group 1994 *Program manual for the Wisconsin Package*. Madison, WI: Genetics Computer Group.

Greenblatt, M. S., Grollman, A. P. & Harris, C. C. 1996 Deletions and insertions in the p53 tumor-suppressor gene in human cancers—confirmation of the DNA-polymerase slippage misalignment model. *Cancer Res.* **56**, 2130–2136.

Hagelberg, E. & Clegg, J. B. 1993 Genetic polymorphisms in prehistoric Pacific Islanders determined by analysis of ancient bone DNA. *Proc. R. Soc. Lond.* B **252**, 163–170.

Hagelberg, E., Quevedo, S., Turbon, D. & Clegg, J. B. 1994 DNA from ancient Easter-islanders. *Nature* **369**, 25–26.

Harihara, S., Hirai, M., Suutou, Y., Shimizu, K. & Omoto, K. 1992 Frequency of a 9-bp deletion in the mitochondrial DNA among Asian populations. *Hum. Biol.* **64**, 161–166.

Harrison, R. G., Rand, D. M. & Wheeler, W. C. 1985 Mitochondrial DNA size variation within individual crickets. *Science* **228**, 1446–1448.

Hertzberg, M., Mickleson, K. N. P., Serjeantson, S. W., Prior, J. F. & Trent, R. J. 1989 An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. *Am. J. Hum. Genet.* **44**, 504–510.

Hoelzel, A. R., Hancock, J. M. & Dover, G. A. 1991 Evolution of the cetacean mitochondrial D-loop region. *Molec. Biol. Evol.* **8**, 475–493.

Hoelzel, A. R., Lopez, J. V., Dover, G. A. & O'Brien, S. J. 1994 Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial-DNA control region of carnivores. *J. Molec. Evol.* **39**, 191–199.

Horai, S., Kondo, R., Nakagawahattori, Y., Hayashi, S., Sonoda, S. & Tajima, K. 1993 Peopling of the America, founded by 4 major lineages of mitochondrial-DNA. *Molec. Biol. Evol.* **10**, 23–47.

Horai, S., Hayasaka, K., Kondo, R., Tsugane, K. & Takahata, N. 1995 Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc. Natn. Acad. Sci. USA* **92**, 532–536.

Imada, M., Inouye, M., Eda, M. & Tsugita, A. 1970 Frameshift mutation in the lysozyme gene of bacteriophage T4: demonstration of the insertion of four bases and the preferential occurrence of base addition in acridine mutagenesis. *J. Molec. Biol.* **54**, 199–217.

Ivanov, P. L., Wadhams, M. J., Roby, R. K., Holland, M. M., Weedn, V. W. & Parsons, T. J. 1996 Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat. Genet.* **12**, 417–420.

Jenuth, J. P., Peterson, A. C., Fu, K. & Shoubridge, E. A. 1996 Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. *Nat. Genet.* **14**, 146–151.

Kimura, M. 1980 A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequence. *J. Molec. Evol.* **16**, 111–120.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X. & Wilson, A. C. 1989 Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natn. Acad. Sci. USA* **86**, 6196–6200.

Levinson, G. & Gutman, G. A. 1987 Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molec. Biol. Evol.* **4**, 203–221.

Lum, K. J., Rickards, O., Ching, C. & Cann, R. L. 1994 Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Hum. Biol.* **66**, 567–590.

Marchington, D. R., Hartshorne, G. M., Barlow, D. & Poulton, J. 1997 Homopolymeric tract heteroplasmy in mtDNA from tissues and single oocytes: support for a genetic bottleneck. *Am. J. Hum. Genet.* **60**, 408–416.

Melton, T., Peterson, R., Redd, A. J., Saha, N., Sofro, A. S. M., Martinson, J. & Stoneking, M. 1995 Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am. J. Hum. Genet.* **57**, 403–414.

Merriwether, D. A., Rothhammer, F. & Ferrell, R. E. 1994 Genetic-variation in the new-world—ancient teeth, bone, and tissue as sources of DNA. *Experientia* **50**, 592–601.

Monsalve, M. V., Derestrepo, H. G., Espinel, A., Correal, G. & Devine, D. V. 1994 Evidence of mitochondrial-DNA diversity in South-American aboriginals. *Ann. Hum. Genet.* **58**, 265–273.

Nei, M. & Li, W.-H. 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natn. Acad. Sci. USA* **76**, 5269–5273.

Passarino, G., Semino, O., Modiano, G. & Santachiarabenerecetti, A. S. 1993 COII transfer RNA(lys) intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (southern Nepal) have oriental affinities. *Am. J. Hum. Genet.* **53**, 609–618.

Piercy, R., Sullivan, K. M., Benson, N. & Gill, P. 1993 The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Legal Med.* **106**, 85–90.

Piko, L. & Matsumoto, L. 1976 Number of mitochondria and some properties of mitochondrial DNA in the mouse egg. *Dev. Biol.* **49**, 1–10.

Redd, A. J., Takezaki, N., Sherry, S. T., Mcgarvey, S. T., Sofro, A. & Stoneking, M. 1995 Evolutionary history of the COII–tRNA(lys) intergenic 9-base-pair deletion in human mitochondrial DNAs from the Pacific. *Molec. Biol. Evol.* **12**, 604–615.

Rentzeperis, D., Medero, M. & Marky, L. A. 1995 Thermodynamic investigation of the association of ethidium, propidium and bis-ethidium to DNA hairpins. *Bioorg. Med. Chem.* **3**, 751–759.

Richards, R. I. & Sutherland, G. R. 1994 Simple repeat DNA is not replicated simply. *Nat. Genet.* **6**, 14–16.

Rosche, W. A., Trinh, T. Q. & Sinden, R. R. 1995 Differential DNA secondary structure-mediated deletion mutation in the leading and lagging strands. *J. Bacteriol.* **177**, 4385–4391.

Schurr, T. G., Ballinger, S. W., Gan, Y. Y., Hodge, J. A., Merriwether, D. A., Lawrence, D. N., Knowler, W. C., Weiss, K. M. & Wallace, D. C. 1990 Amerindian mitochondrial DNAs have rare Asian mutations at high-frequencies, suggesting they derived from 4 primary maternal lineages. *Am. J. Hum. Genet.* **46**, 613–623.

Shields, G. F., Hecker, K., Voevoda, M. I. & Reed, J. K. 1992 Absence of the Asian-specific region-V mitochondrial marker in native Beringians. *Am. J. Hum. Genet.* **50**, 758–765.

Soodyall, H., Jenkins, T. & Stoneking, M. 1995 Polynesian mtDNA in the Malagasy. *Nat. Genet.* **10**, 377–378.

Soodyall, H., Vigilant, L., Hill, A. V., Stoneking, M. & Jenkins, T. 1996 MtDNA control-region sequence variation suggests multiple independent origins of an Asian-specific 9-bp deletion in sub-Saharan Africans. *Am. J. Hum. Genet.* **58**, 595–608.

Stoneking, M., Sherry, S. T., Redd, A. J. & Vigilant, L. 1992 New approaches to dating suggest a recent age for the human mtDNA ancestor. *Phil. Trans. R. Soc. Lond.* B **337**, 167–175.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, J., Terzaghi, E. & Inouye, M. 1966 Frameshift mutations and the genetic code. *Cold Spring Harb. Symp. Quant. Biol.* **31**, 77–84.

Sykes, B., Leiboff, A., Lowbeer, J., Tetzner, S. & Richards, M. 1995 The origins of the Polynesians—an interpretation from mitochondrial lineage analysis. *Am. J. Hum. Genet.* **57**, 1463–1475.

Thomas, M. G., Miller, K., Cook, C. E. & Hagelberg, E. 1994 A method for avoiding mis-priming when sequencing with dynabeads. *Nucl. Acids Res.* **22**, 3243–3244.

Thomas, M. G., Hesse, S. A., Foss, Y. J. & Farzaneh, F. 1996 The use of PCR for differential screening of cDNA libraries. In *Preparative PCR: protocols for molecular cloning and genetic engineering* (ed. B. White), ch. 39 pp. 405–418. Totowa, NJ: Humana.

Torroni, A. (and 10 others) 1992 Native-American mitochondrial-DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics* **130**, 153–162.

Torroni, A., Schurr, T. G., Cabell, M. F., Brown, M. D., Neel, J.V., Larsen, M., Smith, D. G., Vullo, C. M. & Wallace, D. C. 1993*a* Asian affinities and continental radiation of the 4 founding Native-American mtDNAs. *Am. J. Hum. Genet.* **53**, 563–590.

Torroni, A., Sukernik, R. I., Schurr, T. G., Starikovskaya, Y. B., Cabell, M. F., Crawford, M. H., Comuzzie, A. G. & Wallace, D. C. 1993*b* MtDNA variation of aboriginal Siberians reveals distinct genetic affinities with native-Americans. *Am. J. Hum. Genet.* **53**, 591–608.

Torroni, A., Petrozzi, M., Santolamazza, P., Sellitto, D., Cruciani, F. & Scozzari, R. 1995 About the 'Asian'-specific 9-bp deletion of mtDNA. *Am. J. Hum. Genet.* **57**, 507–508.

Tran, H. T., Degtyareva, N. P., Koloteva, N. N., Sugino, A., Masumoto, H., Gordenin, D. A. & Resnick, M. A. 1995 Replication slippage between distant short repeats in *Saccharomyces-cerevisiae* depends on the direction of replication and the rad50 and rad52 genes. *Molec. Cell Biol.* **15**, 5607–5617.

Tuite, E. & Norden, B. 1995 Intercalative interactions of ethidium dyes with triplex structures. *Bioorg. Med. Chem.* **3**, 701–711.

Vigilant, L. 1990 *Control region sequences from African populations and the evolution of human mitochondrial DNA*. PhD thesis, University of California, Berkeley, CA, USA.

Wakelin, L. P. G., Adams, A., Hunter, C. & Waring, M. J. 1981 Interaction of crystal violet with nucleic acids. *Biochemistry* **20**, 5779–5787.

Wallace, D. C. & Torroni, A. 1992 American-Indian prehistory as written in the mitochondrial-DNA—a review. *Hum. Biol.* **64**, 403–416.

Wilkinson, G. S. & Chapman, A. M. 1991 Length and sequence variation in evening bat D-loop mtDNA. *Genetics* **128**, 607–617.

Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H., Arnheim, N. & Wilson, A. C. 1987 Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucl. Acids Res.* **15**, 529–542.

Zharkikh, A. & Li, W. H. 1992 Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide-sequences. 1. 4 taxa with a molecular clock. *Molec. Biol. Evol.* **9**, 1119–1147.

Zhou, B. W., Puga, E., Sun, J. S., Garestier, T. & Helene, C. 1995 Stable triple helices formed by acridine-containing oligonucleotides with oligopurine tracts of DNA interrupted by one or 2 pyrimidines. *J. Am. Chem. Soc.* **117**, 10 425–10 428.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES