Length heterogeneity of a conserved displacement-loop sequence in human mitochondrial DNA

William W. Hauswirth and David A. Clayton¹

Department of Immunology and Medical Microbiology, College of Medicine, University of Florida, Gainesville, FL 32610, and ¹Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA

Received 30 July 1985; Revised and Accepted 26 October 1985

ABSTRACT

Mitochondrial DNA from human tissue culture cells contains heterogeneous sequences located within a previously identified, evolutionally conserved region termed CSB-2. CSB-2 is located near the origin of heavy-strand mitochondrial DNA synthesis and the major transcriptional promoters for each strand of human mitochondrial DNA. Nucleotide sequence analysis of cloned mitochondrial DNA and electrophoretic analysis of appropriate small fragments from cellular mitochondrial DNA show that the variability is limited to a homopolymer sequence which can range in length from 6 to 12 residues. *In vitro* transcriptional analyses, using several of these cloned length polymorphs as templates and partially purified human mitochondrial RNA polymerase, demonstrate that the most common polymorphs will support accurate transcriptional initiation.

INTRODUCTION

Intracellular DNA sequence heterogeneity is a necessary prerequisite for gene evolution. For diploid nuclear genes in animals, heterozygosity is a common feature. This arises principally because the embryo receives a haploid set of genes from each parent. For mitochondrial genes, however, inheritance is strictly maternal (1-6) and intracellular heterogeneity (heteroplasmy) must arise either by inheritance of heterogeneous mitochondrial DNA (mtDNA) or by generation of heteroplasmy during female germ line development (3,7). A previous report has demonstrated and quantitated mtDNA sequence heterogeneity in mammalian tissue (9). In that work, bovine liver and brain mitochondria were shown to contain variable-length nucleotide sequences located within an evolutionary conserved sequence (Conserved Sequence Block 2, CSB-2) near the 5' end of the displacement-loop (D-loop) region. Cloned mtDNA and mtDNA from animal tissue contained a population of CSB-2 length polymorphs exhibiting from 9 to 19 cytosine residues within a homopolymer cytosine sequence in CSB-2. These observations demonstrated a state of significant intra-animal mtDNA sequence heterogeneity, which directly implied heteroplasmy. Additionally, there is a report of a single nucleotide insertion, from one variant clone, in the D loop of human placental mtDNA (10). To determine whether such implied heteroplasmy is a general feature of mammals, we have extended this analysis to cellular mtDNA and individual clones of

human KB cell mtDNA. We find CSB-2 length polymorphism in human tissue culture cell mtDNA, analogous to that found in bovine tissue (9). In contrast, several other similarly analyzed, noncoding regions of the human mitochondrial genome do not exhibit length heterogeneity.

This apparent universality of length variation within mammalian CSB-2s, coupled with its sequence conservation (11-13) and its proximity to the origin of heavy-strand (H-strand) DNA synthesis and transcriptional promoters (14-17), suggested that CSB-2 length heterogeneity might play a role in mitochondrial gene expression. To test one potential function, we have assayed several CSB-2 length polymorphs for their ability to serve as transcriptional templates in an in vitro system. The results indicate that the two most common length polymorphs are functional templates for correct in vitro transcription. Therefore, CSB-2 is probably not involved in accurate transcriptional initiation and any possible role in transcriptional efficiency is yet to be determined.

MATERIALS AND METHODS

DNA preparation, labeling and nucleotide sequence determination. mtDNA was prepared from human KB tissue culture cells grown in suspension. For determining length polymorphism within small restriction fragments, larger fragments containing the regions of interest were extracted by electroelution from 1% agarose gels using NA-45 DEAE paper. Each large fragment was subsequently digested with the indicated restriction enzymes and end labeled using the Klenow fragment of DNA polymerase I (9). To avoid incorporating a labeled nucleoside triphosphate at more than one position at each 3' end, all reactions were run for 3 min at room temperature and, according to the specific sequence (11), the penultimate 3' nucleotide was added as an unlabeled triphosphate to prevent the exonuclease activity of the polymerase from exposing additional template. Fragments were then separated by size on 6% acrylamide-8 M urea gels after denaturation in 50% formamide. Details for each specific fragment are given in the legend to Fig. 2.

For sequence determination, a region of human mtDNA containing CSB-2 was cloned into pBR322 by similarly isolating the Hpa II 828-bp fragment extending from position 104 to position 931. This fragment was ligated into the Cla I site of pBR322 for subsequent transformation and amplification. The cloned mtDNA fragment was excised using Eco RI and Hind III, then digested with Fnu 4HI, and labeled at the 3' ends as above using α -³²P-dCTP and unlabeled dGTP. The labeled fragments were then redigested with Hae III, separated on a 6% acrylamide-8 M urea gel and the polymorphic band containing the CSB-2 of 61 to 63 bases was extracted from the gel by elution and sequenced according to the chemical sequencing technique of Maxam and Gilbert (18).

In vitro transcription. The in vitro template activity of individual CSB-2

polymorphs was determined according to the technique of Walberg and Clayton (19). KB cell mitochondrial RNA (mtRNA) polymerase was a gift of J.E. Hixson and J.N. Doda and was prepared from purified mitochondria (19). *In vitro* runoff reactions contained 10 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 1 mM dithiothreitol, 100 µg/ml BSA, unlabeled nucleotide triphosphates at 200 µM and α-³²P UTP at about 1 µM (20 µCi). Approximately 0.5 µg of each template DNA was added to the reaction. Each template had been digested with *Hind* III and *Eco* RI, extracted with phenol, and ethanol precipitated prior to addition to the reaction mix. Total reaction volumes were 45-50 µl. The reactions were carried out at 30°C for 60 min and then were stopped with 200 µl of 0.3 M sodium acetate, pH 7, 10 mM EDTA and 20 µg yeast tRNA. This was followed by two phenol:chloroform (1:1) extractions and ethanol precipitation. Each sample was taken up in 5 µl of loading buffer containing 8 M urea, heated at 70°C for 1 min and loaded onto a 6% polyacrylamide-8 M urea gel for sizing of the runoff products.

Figure 1. Map locations and lengths in bases of restriction fragment strands used for assessing length polymorphisms in human KB cell mtDNA. Genes are indicated with the single-letter amino acid code used for tRNA genes, and, as indicated, for protein-coding regions. Restriction enzyme site abbreviations are: Ha for *Hae* III, Dd for *Dde* I, Hp for *Hpa* I, Fn for *Fnu* 4HI, Ta for *Tag* I, and Mb for *Mbo* I. Panel A is the strategy for assessing length variation within the noncoding region at the 3' end of the COII gene; panel B is the analogous strategy for the origin of L-strand DNA replication (O_L) region; and panel C is the strategy for the 5' half of the D-loop region containing the origin of H-strand DNA replication (O_H). In panel C the numbered, filled boxes indicate the position and extent of CSB-1, CSB-2, and CSB-3, respectively. More detailed protocols for preparing the fragments for each region are described in the legend to Fig. 2.

Figure 2. Gel analyses of restriction fragment strands containing potential polymorphic regions in the human mitochondrial genome. Map positions for each are shown in Fig. 1. For all samples, 8% polyacrylamide-8 M urea gels were run at 1200 to 1500 V for 1.5 to 4.5 hr. Samples were denatured in 50% formamide immediately before gel electrophoresis. The indicated lengths were determined by nucleotide sequence ladders run on the same gel (data not shown). Lanes A and B are an autoradiogram of the 59-base L-strand fragment containing nucleotides 8251 to 8309, encompassing the 28-bp noncoding region between the COII gene and the lysine tRNA gene. Five times as much sample was loaded in Lane A as in Lane B. To prepare this fragment, a 451-bp *Mbo* I fragment (nucleotides 7859 to 8309) was isolated, digested with *Dde* I, 3' end labeled with α - 32 P-dCTP and 10 mM dATP, and digested with *Hae* III. Lane C contains the 98-base H-strand fragment (nucleotides 5694 to 5791) and the 47-base H-strand fragment (nucleotides 5694 to 5740) containing all or part of the origin of L-strand DNA replication, respectively. These strands were prepared by digesting the isolated, 3768-bp *Mbo* I fragment (nucleotides 3693 to 6460) with *Hpa* I, 3' end labeling with α - 32 P-dTTP and 10 mM dGTP, and then digesting with *Taq* I or *Fnu* 4HI. The samples were then combined for analysis. Lanes D, E, F, and G are autoradiograms of fragments containing the 5' 577 bases of the D-loop region (nucleotides 1 to 577). All fragments were made from an isolated *Mbo* I fragment (nucleotides 1 to 740). Lane D shows the 152 base L-strand fragment (nucleotides 528 to 679) prepared by digesting with *Dde* I, 3' end labeling with α - 32 P-dCTP, and then digesting with *Fnu* 4HI. Lanes E and F are different portions of the same lane (the 137-base fragment is shown on both) analyzing a 206-base H-strand fragment (nucleotides 324 to 528) and a 62-base L-strand fragment (nucleotides 262 to 323) prepared by digesting with *Hae* III, 3' end labeling with α - 32 P-dGTP and 10 mM dTTP, and then digesting with *Fnu* 4HI. Additionally, a 137-base H-strand fragment (nucleotides 3-139), prepared by 3' end labeling the original *Mbo* I fragment with α - 32 P-dGTP and 10 mM dTTP, followed by digestion with *Hinf* I, was analyzed in the same lane. Lane G shows the 125-base L-strand fragment (nucleotides 137 to 261) prepared by digestion with *Fnu* 4HI, 3' end labeling with α - 32 P-dCTP and 10 mM dGTP, and then digestion with *Hinf* I.

RESULTS

Because length variation in coding regions is expected, in most cases, to destroy function, all genome regions examined for length polymorphism contain no known coding function. Three regions were considered potential candidates (Fig. 1). The first is a portion of the D-loop region analogous to that exhibiting length variation in the cow. The other two include the origin of light-strand (L-strand) replication located between nucleotides 5730 and 5760, and a 28-bp region immediately 3' to the cytochrome oxidase subunit II (COII) gene and 5' to the lysine tRNA gene for which no coding function is apparent (11). The strategy for examining each for length polymorphism is identical. A small restriction fragment containing the region of interest is labeled at one end and the labeled strand is sized by gel electrophoresis. If length polymorphs exist, they can be distinguished even if they differ by only a single nucleotide.

A 59-base nucleotide fragment containing the 28-bp noncoding region 3' to the COII gene, so analyzed, exhibited no polymorphism (Fig. 2, lanes A and B). Two overlapping fragments of 47 bp or 98 bp containing either part or all of the origin of L-strand replication similarly exhibited no length polymorphism (Fig. 2, lane C). A series of fragments, as indicated in Fig. 1, spanning 577 nucleotides and encompassing the 5' half of the D loop, were also analyzed. Of the five species (Fig. 2, lanes D,E,F, and G) only the 62-bp Fnu 4HI/Hae III fragment mapping between nucleotides 262 and 323 exhibited length variation (Fig. 2, lane F). This fragment contains the human CSB-2 region and therefore human mtDNA, in addition to bovine mtDNA (9), could contain CSB-2 length polymorphism. By examination of the size range of the polymorphic fragment it appears that individual molecules vary by as many as six nucleotides (59 to 65 bases) with each intermediate nucleotide length being represented. The most common length is 62 nucleotides and this species constitutes approximately 60% of the total.

In order to determine unequivocally the nature of this length polymorphism, the Hpa II fragment of KB cell mtDNA mapping from nucleotide 104 to nucleotide 931 was cloned into pBR322. Recombinant plasmid DNAs from 20 independent clones were isolated and analyzed. The distribution of lengths of the Fnu 4HI/Hae III fragments that contain CSB-2 is shown in Table 1 and indicates that, as found with cellular DNA, the 62-base length predominates. The nucleotide sequences for three of these cloned fragments were determined (Fig. 3). All variation in fragment lengths can be accounted for by variation within the CSB-2 region. Although the human CSB-2 contains two guanine homopolymers separated by a single adenine residue (11), all variation is found within the larger homopolymer.

The recent availability of an *in vitro* transcription system from human mito-

Table 1

CSB-2 Configuration					
	G_8AG_6	G_9AG_6	$G_{10}AG_6$	$G_{11}AG_6$	$G_{12}AG_6$
Number of clones	2	13	4	0	1

chondria allowed us to test preliminarily one potential effect of CSB-2 polymorphism. Although the presence of the most common CSB-2 is not required for mitochondrial transcription initiation in human cells (15), it is unclear what effect other CSB-2 polymorphs might have on the fidelity of transcription initiation. Accordingly, cloned human CSB-2 length polymorphs were used as templates in an in vitro transcription runoff assay using partially purified human mtRNA polymerase (15,19) (Fig. 4). Independent clones containing CSB-2 regions of either 9 or 10 guanine residues within the variable homopolymer region were analyzed. The major runoff products were consistent with L-strand transcription initiation at the known L-strand promoter (LSP). These results indicate that both polymorphs are capable of supporting accurate in vitro transcription; thus, both polymorphs are functional. We specifically note that any effect on L-strand transcriptional efficiency cannot be determined until a complete series of quantitative experiments is assayed for transcriptional activity.

DISCUSSION

The observations reported here document the existence of mtDNA length heterogeneity within cells derived from a human tissue culture line. The variable region is located within an evolutionally conserved mtDNA sequence block, CSB-2, near the 5' end of the D-loop region. Variability is confined to one guanine homopolymer region within CSB-2, varying in length from <8 nucleotides to more than 10 nucleotides. As such, it is the second reported case of polymorphism within mammalian mitochondria. The most significant aspects of this report are the inferences that all mammalian mitochondria may contain similarly variable CSB-2 regions and that other regions of mtDNA, of similar sequence, do not exhibit a parallel heterogeneity.

Existence and role of CSB-2 sequence variability. Previous work on CSB-2 sequences from various tissue sources has demonstrated intraspecific variation in human mtDNA. Both KB and HeLa tissue culture cells are reported to contain the same G₉AG₆ CSB-2 configuration (20,21) which we find with the highest frequency. From different sources of human placental tissue, configurations of G₇AG₅, G₇AG₆, and

Figure 3. Autoradiograms of nucleotide sequencing gels for three CSB-2 polymorphs (Panels A,B,C). Guanine and adenine (G and A) and guanine-only (G) reactions are shown for each sample. Numbers along the side of each panel refer to the length of each iteration of guanines on the H strand within CSB-2.

G_8AG_6 have been reported (10,11). In each case most of these data appear to be derived from a single cloned DNA fragment and it is therefore not possible to assess whether the reported sequence represents the most common CSB-2 in the tissue source.

Figure 4. Runoff transcription analyses of cloned mtDNA fragments containing L-strand and H-strand promoters (LSP and HSP, respectively) using as templates different CSB-2 polymorphs. Panel A is an ethidium-bromide stained agarose gel showing the amount of templates 9-6b and 10-6 used in the reactions. The 856-bp species contains the polymorphic CSB-2 region and both promoters, as shown in the diagram below. Panel B is an autoradiogram of the *in vitro* runoff products using human mtRNA polymerase and the amount of template shown in Panel A. Lane 9-6a is another template of the 9-6 CSB-2 configuration (see text). The runoff products were sized by electrophoresis on a 6% polyacrylamide-8 M urea gel relative to denatured *Hpa* II fragments of pBR322 DNA. Also shown is a diagram of the expected lengths of accurately initiated transcripts. Restriction enzyme site abbreviations are: Bg for *Bgl* II, Ec for *Eco* RI, Ha for *Hae* III, and Mb for *Mbo* I.

However, intraspecific CSB-2 variability is clear and is entirely consistent with this report of intratissue variability, except that we have yet to detect variation in the second, shorter guanine homopolymer of CSB-2.

We are therefore faced with an interesting apparent contradiction: a D-loop sequence block, conserved in all mammals (and probably among all vertebrates because *Xenopus laevis* mtDNA also contains CSB-2 (22)), appears to be variable in length within the tissues or cells of every individual. Because it is normally supposed that evolutionally conserved regions reflect function at some level of gene expression, one implication of this conservation may be that CSB-2 plays a regulatory role in mitochondrial nucleic acid synthesis. If so, then CSB-2 length variation may relate to that role. We have attempted to test experimentally one potential role for such variability. Since CSB-2 is within a few hundred nucleotides of both transcriptional promoters (15), *in vitro* runoff transcription experiments using different polymorphs as templates were carried out in order to determine whether polymorphs vary in their ability to support accurate transcription initiation. The results suggest that different polymorphs can serve as templates for correct initiation. A more complete analysis is necessary to determine whether different length polymorphs vary in the efficiency with which they support *in vitro* transcription.

Recent evidence demonstrates that the transitions of primer RNA synthesis to actual H-strand DNA synthesis occur near the CSB-2 sequences of human (23) and mouse (24) mtDNA. Priming of H-strand replication begins at the LSP and the eventual primer RNA is most likely generated by a sequence-specific mitochondrial endonuclease (24, D.D. Chang and D.A. Clayton, in preparation). Because a portion of the RNA/DNA junctions has been mapped to CSB-2 (23,24), it is possible that CSB-2 plays a role in determining whether RNA synthesis or DNA synthesis occurs at this site. In this regard it will be of interest to learn whether one or another length polymorph of CSB-2 can be predicted to direct the synthesis of a more or less efficiently cleaved primer RNA.

In order to understand better mtDNA length variation, we have also surveyed other regions of the human mitochondrial genome which could, in theory, sustain length polymorphism without destroying biological function. In past studies, except for one clonal variation in the rat (25), all reported length variation outside the D-loop region in mammals involved differences between individual animals and not variation within a single cell or tissue (26,27). However, a general conclusion from these works was that noncoding regions were the ones most likely to undergo intraspecific length variation. The question of tissue variation in regions outside the D loop remained open. We therefore chose to examine both the origin of L-strand DNA replication, which encompasses a 31-base pair noncoding region flanked by the asparagine and cysteine tRNA genes, and the 24 nucleotides immediately 3' to the COII gene which serve no

apparent coding function (11). Besides the D loop, these are the only two noncoding regions of significant length in the genome. Neither of these regions were found to exhibit length polymorphism. We conclude, therefore, that length variability within noncoding regions is not a common phenomenon within an animal. Therefore some other aspect of CSB-2 must contribute to its particularly variable nature.

The fact that length variation may be limited to CSB-2 suggests that such heterogeneity is not simply tolerated, but may be advantageous to the system. A basis for such an advantage is not apparent; however, one can postulate possible models for up or down regulation of either transcription or replication (or both) depending on the CSB-2 variant molecule. Any need for diversity within a cell of any putative mtDNA regulatory sequence has not been established. In a broader context, these data also suggest that many independent intracellular lineages of mtDNA molecules exist within tissues of an animal and that uniparental inheritance of a subset of this population may be a key process leading to rapid mtDNA evolution.

How is CSB-2 variability generated? A general mechanism for generating CSB-2 variability would be replication slippage, as suggested earlier (9). Mismatching during replication of repeated sequences, including homopolymers, has been commonly suggested to explain length variation in DNA of prokaryotic, eukaryotic, and cytoplasmic origin (28-31). However, if merely the existence of a homopolymer sequence in the mitochondrial genome was sufficient to trigger length variation, many other variable regions should exist. Contradicting this view, the immediately adjacent homopolymer of six guanines in CSB-2 appears to be invariant. Additionally, analysis of all homopolymer sequences in the human mitochondrial genome reveals three other homopolymers of six or more base pairs within noncoding regions. Two of these have been examined in this study. One is a sequence of six guanines within the 137-base fragment analyzed in Fig. 2 and it shows no variation. The second is a homopolymer of six adenines within the 62-base fragment that also contains the CSB-2 sequence. All variability within this fragment is accounted for by CSB-2 sequences because all clones examined by sequencing contain the homopolymer adenine as a six-base entity (data not shown). In the human mitochondrial genome, except for CSB-2, all longer homopolymers exist only within coding regions. In the bovine genome, the only homopolymer sequence >7 nucleotides (except CSB-2) occurs at the origin of L-strand DNA replication and this sequence is also not polymorphic (unpublished data). Therefore there appears to be no obvious structural feature of homopolymer sequences in CSB-2 which might account for length variation.

A more interesting possibility is that CSB-2 is susceptible to sequence changes because it is replicated differently from the bulk of the genome. Because priming of H-strand synthesis occurs at the LSP and the transition to DNA synthesis is downstream of

CSB-2 (23,24), CSB-2 is first copied by RNA polymerase, not DNA polymerase. The nature of primer excision in mitochondria is unknown, but DNA synthesis at CSB-2 could be linked to the primer removal event. Incomplete removal of the primer RNA sequence (14) would result in an error frequency characteristic of mtRNA polymerase rather than mtDNA polymerase. Alternatively, CSB-2 could be replicated as a gap-filling phenomenon involving a different type of fidelity than that associated with the major amount of genomic synthesis.

ACKNOWLEDGMENTS

This work was supported by grant NP-9J from the American Cancer Society, Inc. and grant GM-33088-14 from the National Institute of General Medical Sciences. We thank D.D. Chang, R.P. Fisher and J.E. Hixson for review of the manuscript.

REFERENCES

1. Dawid, I.B. and Blackler, A.W. (1972) *Dev. Biol.* **29**:152-161.
2. Hutchinson, C.A., III, Newbold, J.E., Potter, S.S. and Edgell, M.H. (1974) *Nature (London)* **251**:536-538.
3. Kroon, A.M., Pepe, G., Bakker, H., Holtrop, M., Bollen, J.E., van Bruggen, E.F., Cantatore, P., Terpstra, P. and Saccone, C. (1977) *Biochem. Biophys. Acta* **478**:128-145.
4. Hayashi, J.I., Yonekawa, H., Gotoh, O., Tagshira, Y., Moriwaki, K. and Yoshida, T.H. (1979) *Biochem. Biophys. Res. Comm.* **81**:871-877.
5. Avise, J.C., Lansman, R.A. and Shade, R.O. (1979) *Genetics* **92**:279-295.
6. Giles, R.E., Blanc, H., Cann, H.M. and Wallace, D.C. (1980) *Proc. Natl. Acad. Sci. USA* **77**:6715-6719.
7. Hauswirth, W.W. and Laipis, P.J. (1982) *Proc. Natl. Acad. Sci. USA* **79**:4686-4690.
8. Olivo, P.D., Van De Walle, M.J., Laipis, P.J. and Hauswirth, W.W. (1983) *Nature* **306**:400-402.
9. Hauswirth, W.W., Van De Walle, M.J., Laipis, P.J. and Olivo, P.D. (1984) *Cell* **37**:1001-1007.
10. Greenberg, B.D., Newbold, J.E. and Sugino, A. (1983) *Gene* **21**:33-49.
11. Anderson, S., Bankier, 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* **296**:457-465.
12. Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) *Cell* **26**:167-180.
13. Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. (1982) *J. Mol. Biol.* **156**:683-717.
14. Clayton, D.A. (1982) *Cell* **28**:693-705.
15. Chang, D.D. and Clayton, D.A. (1984) *Cell* **36**:635-643.
16. Bogenhagen, D.F., Applegate, E.F. and Yoza, B.K. (1984) *Cell* **36**:1105-1113.
17. Clayton, D.A. (1984) *Ann. Rev. Biochem.* **53**:573-594.
18. Maxam, A.M. and Gilbert, W. (1980) *Meth. Enzymol.* **65**:499-560.
19. Walberg, M.W. and Clayton, D.A. (1983) *J. Biol. Chem.* **258**:1268-1275.
20. Walberg, M.W. and Clayton, D.A. (1981) *Nucl. Acids Res.* **9**:5411-5421.
21. Ojala, D., Crews, S., Gelfand, R. and Attardi, G. (1981) *J. Mol. Biol.* **150**:303-314.
22. Wong, F.P., Mo, D.P., Wilson, R.E. and Roe, B.A. (1983) *Nucl. Acids Res.* **11**:4977-4995.
23. Chang, D.D. and Clayton, D.A. (1985) *Proc. Natl. Acad. Sci. USA* **82**:351-355.
24. Chang, D.D., Hauswirth, W.W. and Clayton, D.A. (1985) *EMBO J.* **4**:1559-1567.
25. Brown, G.G. and DesRosiers, L.J. (1983) *Nucl. Acids. Res.* **11**:6699-6708.

26. Taira, M., Yoshida, E., Kobayashi, M., Yaginuma, K. and Koike, K. (1983) Nucl. Acids Res. 11:1635-1643.
27. Cann, R.L. and Wilson, A.C. (1983) J. Mol. Evol. 19:87-91.
28. Streisinger, G., Okada, Y., Emrech, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31:77-84.
29. Farabaugh, P.J., Schmeissner, U., Hofer, M. and Miller, J.H. (1978) Mol. Biol. 126:847-863.
30. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21:653-668.
31. Jones, C.W. and Kafatos, F.C. (1982) J. Mol. Evol. 19:87-103.