

---

**The wheat mitochondrial gene for apocytochrome b: absence of a prokaryotic ribosome binding site**

---

Poppo H.Boer, John E.McIntosh, Michael W.Gray and Linda Bonen

---

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

---

Received 7 January 1985; Accepted 27 February 1985

---

**ABSTRACT**

The wheat mitochondrial gene for apocytochrome b (CYB) has been identified by its hybridization to a yeast CYB probe and its nucleotide sequence has been determined. The wheat CYB sequence predicts a cytochrome b apoprotein of 398 amino acids; it is almost identical to that of maize but has ten additional amino acids at the carboxy terminus. No introns are present in the wheat CYB gene, but an internal segment of the gene is repeated at another genomic location.

Transcript analysis reveals a single wheat CYB mRNA of approximately 2.4 kb with a long untranslated leader. Sequences upstream of the CYB coding region are very similar in wheat and maize but the stretch proposed to be a ribosome binding site in maize is not conserved in wheat. The corresponding leader regions of the wheat mitochondrial mRNAs for cytochrome oxidase subunits I and II also lack complementarity to the 3'-end of the small subunit rRNA. We conclude that alternative signals are involved in the initiation of translation in plant mitochondria.

**INTRODUCTION**

The mitochondrial (mt) genomes of higher plants are larger and more complex in form and structure than the mtDNAs of other organisms [1]. This has impeded their physical characterization, so that to date only a few genes have been identified. These include mt ribosomal RNA genes in wheat and maize [2-6], a wheat mt transfer RNA gene [7] and genes for respiratory chain components, namely cytochrome oxidase subunit II (COII) in maize, Oenothera, wheat and rice [8-11] and apocytochrome b (CYB) in maize and Oenothera [12-14]. An inter-species comparison of rRNA [3,6,15] and COII [10,11] genes suggests that higher plant mitochondrial genes undergo a very low rate of nucleotide sequence divergence compared to animal mt genes. For example, the wheat and maize COII genes share 99% sequence identity [10], a level much higher than that observed between COII genes in mammals even of the same genus (cf. Rattus rattus vs. Rattus norvegicus COII genes with 92% sequence identity [16]).

On the other hand, markedly different mtDNA restriction patterns among

---

closely-related plants [17,18] indicate that there can be extensive variation in mt genome arrangement and size. Such differences can affect structural genes and their expression. In the case of the COII genes, the coding and intron sequences are highly conserved among the three monocotyledons wheat, maize and rice, but the wheat and rice COII introns have a large insert relative to maize [10,11]. In addition, sequences upstream of the maize COII gene are unrelated to those in wheat or rice and interestingly, the wheat and maize transcript patterns are quite different, with that of wheat being much less complex [10].

Little is known about transcriptional or translational controls in higher plant mitochondria but comparative sequence and transcript analysis may reveal motifs important in gene expression. Several features of the higher plant mt translational system are distinctively prokaryotic in nature. Their ribosomes (unlike those of other mitochondria) possess a 5S rRNA species [2] and the 18S and 26S rRNAs [4-6,15] as well as the fmet-tRNA [7] show strong sequence similarity with their eubacterial and chloroplast counterparts. Recently, Dawson et al. [13] have proposed that maize mt mRNAs may possess a ribosome binding site analogous to that involved in the initiation of translation in prokaryotes and chloroplasts [19,20].

We have determined the sequence of the wheat mitochondrial gene for apocytochrome b and find that its derived amino acid sequence supports a recently proposed model for CYB protein structure and function in energy transducing membranes [21,22]. We compare CYB flanking sequences to those of wheat mt COII and COI genes [10; ms. in preparation], as well as those of maize mt genes [13], to examine the proposal of a ribosome binding site in higher plant mitochondrial mRNAs.

### MATERIALS AND METHODS

Wheat mtDNA was digested with either BamHI or HindIII restriction endonuclease and ligated into a phosphatase-treated plasmid vector pUC9 [23] for transformation of E. coli JM83. Recombinant clones containing the wheat CYB gene were identified by colony hybridization using a cytochrome b gene probe from the yeast Kluveromyces lactis (kindly provided by L.A. Grivell and M. de Haan, University of Amsterdam). The yeast CYB insert was separated from vector by gel electrophoresis in low melting point agarose. Heterologous hybridization experiments were conducted at 50°C and 5 x SSC for at least 36 h. Similar Southern hybridization experiments were performed with total wheat mtDNA. Homologous hybridization experiments were followed by stringent

washes with 0.1 x SSC at 65°C.

Recombinant plasmid DNA minipreps were prepared by the boiling method [24] followed by polyethylene glycol precipitation [25]. Cloned fragments were characterized by restriction enzyme and Southern hybridization analysis and appropriate fragments were subcloned into M13 vectors. Sequencing was carried out by the M13 dideoxy chain termination method [26]. The region containing the 3' terminus of the gene (clone pH410, see Fig.2) could only be obtained using the *E. coli* TBl host, an *E. coli* JM83 derivative that lacks the K12 restriction system (kindly provided by T. Baldwin, Texas A&M). This clone was detected by screening a HindIII pUC9 clone bank with the HindIII-BamHI 718 bp fragment isolated from pB376 (see Fig. 2).

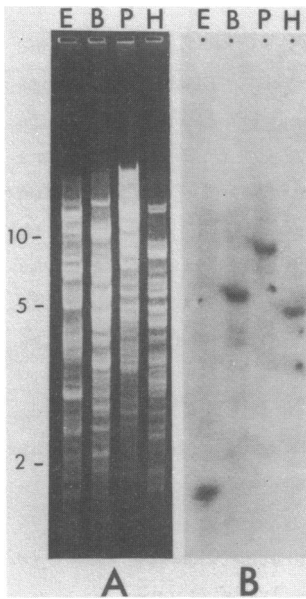
For transcript analysis, wheat mt RNA was isolated from 24 h germinating embryos [27] and the fraction that was insoluble in 1M NaCl was separated by size on 2.2M formaldehyde, 1.25% agarose gels [28] and transferred to nitrocellulose. Northern hybridizations were conducted at 42°C in 50% formamide, 5 x SSC using either M13 single-stranded probes labeled with <sup>32</sup>P by second-strand synthesis or plasmids labeled by nick translation.

## RESULTS

### Identification and sequence of the wheat mitochondrial CYB gene

Wheat mtDNA which had been digested with various restriction endonucleases was hybridized with the *K. lactis* CYB gene under conditions of low stringency. In each case, a strongly hybridizing fragment was observed (Fig. 1). Clones containing the major hybridizing 4.9 kb HindIII and 5.6 kb BamHI fragments (pH354 and pB376) were selected from wheat mtDNA pUC9 clone banks by colony hybridization and the CYB gene was localized on the restriction map (Fig. 2) by Southern analysis. The strongly hybridizing 1.75 kb EcoRI fragment is internal to the BamHI fragment (clone pB376).

The nucleotide sequence of the wheat mt CYB gene was determined as shown by the sequencing strategy in Fig. 2 and the sequence with flanking regions is shown in Fig. 3. The structural gene is 1194 nucleotides long but unlike many fungal CYB genes it has no introns. The wheat CYB gene predicts a protein of 398 amino acids which shows 98.5% amino acid sequence identity with maize [13], 94% with *Oenothera* [14], 49% with human [29] and 49% with yeast [30]. Only six amino acid residues differ from those of maize, namely Ser<sub>40</sub> → Cys (in maize), His<sub>60</sub> → Tyr, Val<sub>197</sub> → Ala, Pro<sub>361</sub> → Ser, Val<sub>363</sub> → Phe and Asp<sub>388</sub> → Glu (Fig. 3). Values for the degree of chemical difference [31] between these amino acid pairs are 112, 83, 64, 74, 50 and 45,



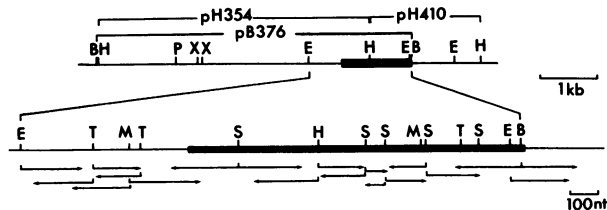
**Figure 1.** Identification of wheat mtDNA restriction fragments having homology with the apocytochrome b gene from *K. lactis*.

[A] UV fluorescence pattern of wheat mtDNA digested with *EcoRI* (E), *BamHI* (B), *PstI* (P) or *HindIII* (H) and fractionated on 1% agarose gels.

[B] Southern blot analysis using the *K. lactis* CYB probe. Arrows indicate molecular weight sizes of 10, 5 and 2 kb.

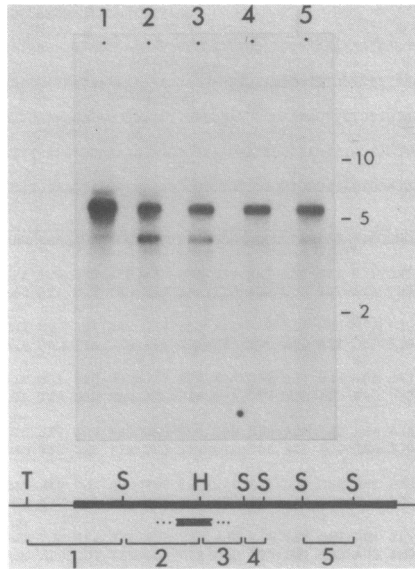
respectively, where a value greater than 73 is regarded as non-conservative [cf. ref.32]. Sequence comparison predicts that the wheat CYB protein is 10 amino acids longer than that of maize because of a first position codon substitution (TAG → GAG) at the position of the maize termination codon.

The wheat and maize CYB genes show 98.8% nucleotide sequence identity, with a total of 14 nucleotide differences scattered along the length of the gene. Transversions and transitions are equally represented and seven changes are third position codon substitutions. A first position change at amino acid position 120 (TGG → CGG) appears to be a silent codon substitution for Trp,



**Figure 2.** Restriction map of the wheat mt CYB gene region and sequencing strategy. The black bar indicates the CYB coding region and the arrows show the direction and extent of sequence determined. Only restriction sites used in the analysis are shown: *BamHI* (B), *EcoRI* (E), *HindIII* (H), *MspI* (M), *PstI* (P), *Sau3A* (S), *TaqI* (T), *XbaI* (X).

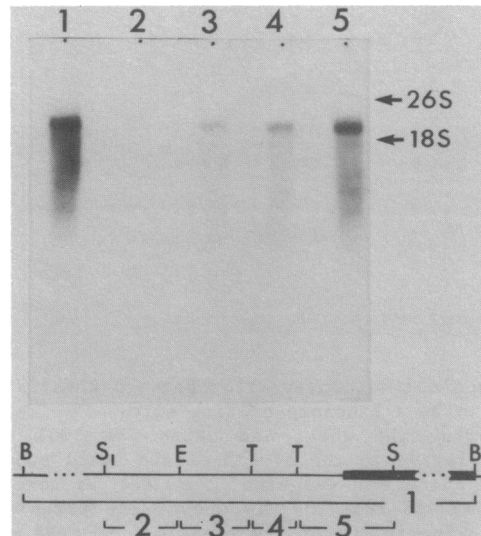




**Figure 4.** Southern blot analysis of subregions of the wheat CYB gene region. Wheat mtDNA digested with BamHI was electrophoresed on 1% agarose gels and transferred to nitrocellulose. Strips were probed with M13 clones from the CYB gene region as shown below. The long black bar indicates the coding region and the short black bar the region hybridizing to an additional BamHI fragment. Restriction sites are abbreviated as in Fig. 2.

since it is generally accepted that in higher plant mitochondria CGG encodes Trp rather than the 'universal' Arg [8,9]. Of the 14 Trp residues in the derived wheat CYB protein sequence, 12 are encoded by TGG, 2 by CGG, but none by TGA, which specifies Trp in most fungal and animal mt genes.

When total wheat mtDNA that had been digested by BamHI endonuclease was hybridized under stringent conditions with clones containing segments of the wheat CYB gene (used for DNA sequence analysis), hybridization was consistent with the presence of a single copy of the gene on the cloned BamHI 5.6 kb fragment (Fig. 4). However, two subprobes of the coding region hybridized strongly to an additional wheat mtDNA BamHI fragment of 4.3 kb implying that a segment of the CYB gene is present at another genomic location (Fig. 4, lanes 2 and 3). Minor additional hybridization of BamHI fragments to all CYB probes may reflect very low levels of other genomic arrangements of the CYB gene in wheat mtDNA. Similar hybridization patterns (data not shown) were obtained using HindIII digested wheat mtDNA.



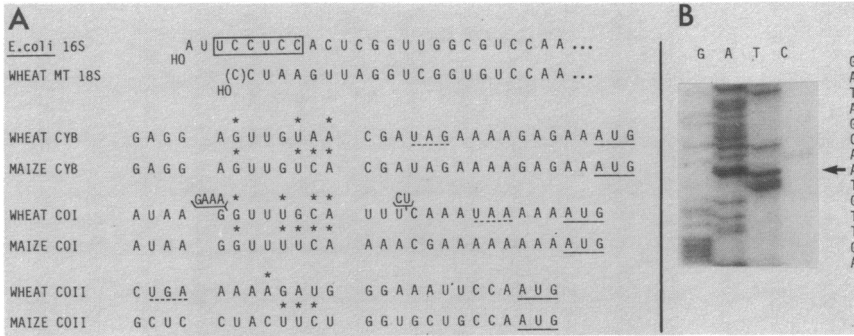
**Figure 5.** Northern blot analysis of the wheat CYB gene region. Wheat mtRNA fractionated on formaldehyde/agarose gels was transferred to nitrocellulose and probed with fragments from the wheat CYB gene region, labeled by nick-translation (lane 1) or M13 second-strand synthesis (lanes 2-5). Ribosomal RNA positions are indicated by arrows. The black bar represents the CYB coding region and restriction site abbreviations are as in Fig. 2, with  $S_{1I}$  =  $S_I$ .

#### Transcript analysis of the wheat CYB gene

Northern hybridization analysis of wheat mt RNA using a number of DNA probes from the wheat CYB gene region revealed a single major transcript of approximately 2.4 kb (Fig. 5). Its size indicates that the wheat CYB mRNA exceeds the coding length by approximately 1.2 kb. A probe extending from position -346 to -615 (TaqI-EcoRI) hybridized to the 2.4 kb transcript (Fig. 5 lane 3) while ones from regions further upstream did not (Fig. 5, lane 2). This suggests that the wheat mt CYB mRNA has a leader of about 400-650 nucleotides. Essentially the same transcript pattern was seen when the total 5.6 kb BamHI fragment was used in RNA blot analysis (Fig. 5, lane 1).

#### Analysis of regions upstream of wheat mt protein coding genes

When the region preceding the CYB coding sequence was compared to the analogous region in maize [12], strong conservation (93% nucleotide sequence identity) was observed until position -530. At that point, homology abruptly ends (arrow, Fig. 3). Interestingly, this position is in the vicinity of the predicted 5' terminus of the wheat CYB mRNA, that is, 400-650 nucleotides upstream of the initiation codon.



**Figure 6.** Lack of nucleotide conservation between wheat and maize mt mRNAs in the region proposed to be a ribosome binding site.

[A] Sequences preceding the wheat and maize CYB, COI, and COII initiation codons are aligned relative to the 3'-end of the small subunit rRNA as in ref. [13]. The *E. coli* 'Shine and Dalgarno' sequence is boxed and the analogous region of the wheat mt 18S rRNA [34] is shown. The parenthesis indicates that this C residue is lacking in 20% of the wheat mt 18S rRNA molecules. Initiation codons and in-phase termination codons are shown by solid and dashed underlines, respectively. Nucleotides which are complementary to the 3'-end of the 18S rRNA are shown with asterisks.

[B] Sequencing gel showing the region immediately upstream of the wheat CYB gene. The arrow indicates the nucleotide change relative to maize.

We also compared the sequence immediately upstream of the wheat CYB initiation codon to regions preceding the wheat COII and COI genes [10, ms. in preparation], as well as those of maize [13], to search for conserved sequences that might be involved in expression. It has been proposed [13] that the sequence AGUUGUCA may be a ribosome binding site in the CYB mRNA of maize. However, we observed a nucleotide change in wheat (C → A) (Fig. 6) which disrupts complementarity with the 3'-end of the small subunit rRNA. A summary of similar differences between the comparable regions of other wheat and maize mt mRNAs, aligned as in ref. [13], is given in Fig. 6. There is no conserved stretch that can base pair with the 3'-end of the small subunit rRNA. Translation is assumed to begin at the indicated AUG codon (underlined in Fig. 6) because there are in-frame termination codons a short distance upstream in each case (dashed underline, Fig. 6).

**DISCUSSION**

The wheat mitochondrial CYB gene is 1194 nucleotides in length and encodes the longest apocytochrome b protein yet reported. Extensions at both the amino and carboxy termini result in it being 18 amino acids longer than its human counterpart. The wheat CYB protein is almost identical to that of



maize [13] although a nucleotide substitution at the position of the maize termination codon results in it being 10 amino acids longer. The Oenothera CYB sequence is identical to that of wheat at the position of the maize stop codon but the former terminates after encoding five additional amino acids [14]. The wheat CYB sequence conforms to a recently proposed model [21,22] in which the protein is folded into nine hydrophobic transmembrane domains. Four conserved histidines in transmembrane domains II and V (positions 88, 102, 189 and 203 in wheat) are predicted to bind two heme prosthetic groups. None of the six amino acid differences between the wheat and maize CYB proteins affect the model to any extent. The three non-conservative replacements and the carboxy terminal extension of the wheat CYB protein (if the latter is present in the mature protein) presumably reflect low functional constraint in these regions of the protein.

Hybridization analysis indicates that one region within the wheat CYB gene is present at another genomic location and hybridization intensities suggest that this region corresponds approximately to transmembrane domains III and IV in the protein model [21,22]. The presence of repeated segments of higher plant mt genes may be a general feature of these complex genomes, since a similar observation has been made for the wheat COII gene, where an extra copy corresponding to the single transmembrane domain of the COII protein was found [10].

The strong sequence conservation between the wheat and maize CYB genes reinforces the view that higher plant mt genes are evolving slowly relative to mammalian mt genes. The 99% nucleotide sequence identity seen between the wheat and maize CYB genes is much higher than that (93%) observed between rat and mouse CYB genes [33]. The two plants are believed to have diverged some 50-70 million years ago compared to a divergence time of approximately 25-30 million years for the two rodents [cf. ref.6]. As for the non-coding regions of plant mt genes, the level of conservation between wheat and maize is somewhat lower (93% identity) in the CYB upstream region than it is within the COII intron (99%) [10], suggesting that there are different constraints on various untranslated regions of plant mt transcripts.

Northern hybridization analysis identifies a major CYB transcript of approximately 2.4 kb, with the CYB mRNA having a leader of 400-650 nucleotides. The wheat CYB transcript pattern is simpler than that of maize, where a prominent 4.2 kb transcript upstream and overlapping the maize CYB gene was seen in addition to the 2.25 kb CYB mRNA. The wheat COII gene also shows a less complex transcript pattern than its maize counterpart [10];

in that case, sequence similarity abruptly ends three nucleotides upstream of the ATG initiation codon. Differences in transcript patterns between the two plants may well be related to differences in upstream sequences and these observations demonstrate that different genomic arrangements among closely related organisms can have an effect on gene expression. We do not observe any pronounced sequence similarities among the upstream regions of the three wheat mt genes but a detailed search for transcriptional or processing motifs must await precise mapping of the 5' ends of the transcripts.

What signals are involved in the translation of these plant mt mRNAs? Higher plant mt ribosomes are prokaryotic-like in that they possess a 5S rRNA [2] and their ribosomal RNAs [4-6] and initiator methionine tRNA [7] show strong sequence similarity with eubacterial counterparts. At present it is not known whether other features of the plant mt translational system also resemble those of bacteria, but Dawson et al. [13] have recently proposed that plant mt mRNAs may possess a ribosome binding site analogous to that observed in prokaryotes. In *E. coli* and chloroplasts, the ability of the leaders of mRNAs to base pair with the sequence CCUCCU at the 3'-end of the small subunit rRNA has been shown to be important in the initiation of translation [19,20]. However, there are several differences in plant mitochondria. Although the DNA sequence of an extended region at the 3'-end of both wheat and maize mt 18S rRNA genes is almost identical to that of *E. coli*, the 'Shine and Dalgarno' equivalent CCTCCT is replaced by AATCCT. It should also be noted that in the alignment of Dawson et al. [13], the putative 'Shine and Dalgarno' region in the maize 18S rRNA includes a U residue at the extreme 3'-end which by direct RNA sequencing has been shown [34] to be lacking in the mature 18S rRNA in wheat mitochondria (and in fact 20% of the rRNA molecules also lack the penultimate C residue).

In a comparison of the upstream regions of the wheat CYB, COII, and COI genes with each other and with their maize homologues, we found no conserved sequence able to base pair with the 3'-end of the small subunit rRNA. The putative CYB ribosome binding site of maize [13] has a nucleotide change in wheat that disrupts base pairing. Similarly, the analogous region upstream of the wheat COI gene shows a nucleotide change and an insert. Stretches of several nucleotides complementary to the 3'-end of the small subunit rRNA can be found upstream of the wheat COI and COII initiation codons in the region analogous to the prokaryotic position, that is positions -5 to -9, but none is conserved in maize. Since it is unlikely that such closely-related genes would differ in fundamental mechanisms of expression, we conclude that

---

alternative signals must be involved in the initiation of higher plant mt translation.

It is not yet clear what the nature of this control might be. The presence of a long leader in the wheat CYB mRNA (400-650 nucleotides) contrasts with the situation in human mitochondria where mRNAs lack leaders [35]; however it resembles the case of the yeast CYB mRNA which has a leader of approximately 1 kb [36]. The presence of numerous out-of-frame AUG codons in the wheat CYB leader does not support a scanning model such as that proposed for the initiation of eukaryotic translation [20]. The only striking features of the three wheat mt mRNAs are long stretches of purines preceding the initiation codons. For example, within the first 40 nucleotides of the wheat CYB leader, the A+G content is 85%. The initiation codons are flanked by A residues, and as noted [13], this would allow a four nucleotide interaction (AUGA) with the anticodon loop of the fmet-tRNA [7] in a fashion analogous to that for bacterial initiation [37]. Additional sequence information on other plant mitochondrial genes as well as in vitro studies on ribosome binding and translation initiation are needed to resolve this issue.

#### ACKNOWLEDGEMENTS

We thank L.A. Grivell and M. de Haan for the K. lactis CYB clone, T. Baldwin for the E. coli TBl strain and P. Isaac, A. Dawson, C. Leaver and A. Brennicke for communicating sequence data prior to publication. Provision of a Dalhousie Medical Research Fellowship (L.B.) is gratefully acknowledged. This work was supported by a grant (MT-4124) from the Medical Research Council of Canada to M.W.G.

#### REFERENCES

1. Leaver, C.J. and Gray, M.W. (1982) *Annu. Rev. Plant Physiol.* 33, 373-402.
2. Spencer, D.F., Bonen, L. and Gray, M.W. (1981) *Biochemistry* 20, 4022-4029.
3. Chao, S., Sederoff, R.R. and Levings III, C.J. (1983) *Plant Physiol.* 71, 190-193.
4. Spencer, D.F., Schnare, M.N. and Gray, M.W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 493-497.
5. Dale, R.M.K., Mendu, N., Ginsburg, H. and Kridl, J.C. (1984) *Plasmid* 11, 141-150.
6. Chao, S., Sederoff, R. and Levings III, C.S. (1984) *Nucl. Acids Res.* 12, 6629-6644.
7. Gray, M.W. and Spencer, D.F. (1983) *FEBS Lett.* 161, 323-327.
8. Fox, T.D. and Leaver, C.J. (1981) *Cell* 26, 315-323.
9. Hiesel, R. and Brennicke, A. (1983) *EMBO J.* 2, 2173-2178.
10. Bonen, L., Boer, P.H. and Gray, M.W. (1984) *EMBO J.* 3, 2531-2536.
11. Kao, T.-H., Moon, E. and Wu, R. (1984) *Nucl. Acids Res.* 12, 7305-7315.

12. Leaver, C.J., Hack, E., Dawson, A.J., Isaac, P.G. and Jones. V. (1983) in *Mitochondria 1983*, Schweyen, R.J., Wolf, K. and Kaudewitz, F. Eds., pp.269-283, de Gruyter, Berlin.
13. Dawson, A.J., Jones, V.P. and Leaver, C.J. (1984) *EMBO J.* 3, 2107-2113.
14. Schuster, W. and Brennicke, A. (1984) *Curr. Genetics*. In press.
15. Gray, M.W., Sankoff, D. and Cedergren, R.J. (1984) *Nucl. Acids Res.* 12, 5837-5852.
16. Brown, G.G. and Simpson, M.V. (1982) *Proc. Natl. Acad. Sci. USA* 79, 3246-3250 .
17. Sederoff, R.R., Levings III, C.S., Timothy, D.H. and Hu, W.W.L. (1981) *Proc. Natl. Acad. Sci. USA* 78, 5953-5957.
18. Ward, B.L., Anderson, R.S. and Bendich, A.J. (1981) *Cell* 25, 793-803.
19. Shine, J. and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* 71, 1342-1346.
20. Kozak, M. (1983) *Microbiol. Rev.* 47, 1-45.
21. Widger, W.R., Cramer, W.A., Herrmann, R.G. and Trebst, A. (1984) *Proc. Natl. Acad. Sci. USA* 81, 674-678.
22. Saraste, M. (1984) *FEBS Lett.* 166, 367-372.
23. Vieira, J. and Messing J. (1982) *Gene* 19, 259-268.
24. Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.* 114, 193-197.
25. Rambach, A. and Hogness, D. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5041-5045.
26. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
27. Bonen, L. and Gray, M.W. (1980) *Nucl. Acids Res.* 8, 319-335.
28. Lehrach, H. Diamond, D., Wozney, J.M. and Boedtker, H. (1977) *Biochemistry* 16, 4743-4751.
29. 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.
30. Nobrega, F.G. and Tzagoloff, A. (1980) *J. Biol. Chem.* 255, 9828-9837.
31. Grantham, R. (1974) *Science* 185, 862-864.
32. Cann, R.L., Brown, W.M. and Wilson, A.C. (1984) *Genetics* 106, 479-499.
33. Koike, K., Kobayashi, M., Yaginuma, K., Taira, M., Yoshida, E. and Imai, M. (1982) *Gene* 20, 177-185.
34. Schnare, M.N. and Gray, M.W. (1982) *Nucl. Acids Res.* 10, 3921-3932.
35. Ojala, D., Montoya, J. and Attardi, G. (1981) *Nature* 290, 470-474.
36. Grivell, L.A., Hensgens, L.A.M., Osinga, K.A., Tabak, H.F., Boer, P.H., Crusius, J.B.A., van der Laan, J.C., de Haan, M., van der Horst, G., Evers, R.F. and Arnberg, A.C. (1982) in *Mitochondrial Genes*, Slonimski, P., Borst, P. and Attardi, G. Eds., pp. 225-239, Cold Spring Harbor Laboratory, New York.
37. Taniguchi, T. and Weissman, C. (1978) *J. Mol. Biol.* 118, 533-565.