The EMBO Journal vol.3 no.11 pp.2531-2536, 1984

The wheat cytochrome oxidase subunit II gene has an intron insert and three radical amino acid changes relative to maize

Linda Bonen, Poppo H. Boer and Michael W. Gray

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

Communicated by G.S.P. Groot

We have determined the sequence of the wheat mitochondrial gene for cytochrome oxidase subunit II (COII) and find that its derived protein sequence differs from that of maize at only three amino acid positions. Unexpectedly, all three replacements are non-conservative ones. The wheat COII gene has a highly-conserved intron at the same position as in maize, but the wheat intron is 1.5 times longer because of an insert relative to its maize counterpart. Hybridization analysis of mitochondrial DNA from rye, pea, broad bean and cucumber indicates strong sequence conservation of COII coding sequences among all these higher plants. However, only rye and maize mitochondrial DNA show homology with wheat COII intron sequences and rye alone with intron-insert sequences. We find that a sequence identical to the region of the 5′ exon corresponding to the transmembrane domain of the COII protein is present at a second genomic location in wheat mitochondria. These variations in COII gene structure and size, as well as the presence of repeated COII sequences, illustrate at the DNA sequence level, factors which contribute to higher plant mitochondrial DNA diversity and complexity.

Key words: cytochrome oxidase subunit II/introns/mitochondrial genes/plant mitochondrial DNA/wheat

Introduction

Subunit II of cytochrome c oxidase (COII) is a well-studied component of the respiratory chain complex, involved in the transfer of electrons from cytochrome c to oxygen (for review, see Capaldi *et al.*, 1983). It is located in the inner mitochondrial membrane and, like subunits I and III of this complex, its gene is encoded in the mitochondrial DNA. The sequence of the COII gene has been determined in a number of organisms ranging from yeast to human, and only in maize is it interrupted by an intron (Fox and Leaver, 1981). A second higher plant COII gene, that of evening primrose, *Oenothera berteriana*, is not split (Hiesel and Brennicke, 1983). This optional nature of mitochondrial introns in closely-related organisms has previously been observed among yeast strains and is a factor contributing to differences in their genome size (Hensgens *et al.*, 1983). Higher plant mitochondrial DNA (mtDNA) is considerably larger than that of yeast (for review, see Leaver and Gray, 1982) and it can show even more pronounced variation in size among members of the same family. For example, within the cucurbit family, mtDNA size ranges from 330 kb in watermelon to ~2400 kb in muskmelon (Ward *et al.*, 1981), with < 10% being attributed to repeated sequences.

As yet, the significance of the increased mitochondrial genome size in higher plants (some 15- to 150-fold larger than human mtDNA) and the extreme diversity in size among closely-related plants is largely unknown. It seems likely that non-coding sequences account for much of the difference, perhaps having arisen by an extensive rearrangement of mitochondrial sequences (discussed in Ward *et al.*, 1981) and/or the integration of foreign sequences (cf. the presence of chloroplast sequences in mtDNA, Stern and Lonsdale, 1982; Stern and Palmer, 1984). Such events, either recombinational or transpositional, would expand genome size and might influence individual gene structure. Although information is emerging on the physical organization of higher plant mitochondrial genomes (Lonsdale *et al.*, 1983; Palmer and Shields, 1984; Falconet *et al.*, 1984), very little is known about the evolution of plant mtDNA at the individual gene level, either in terms of intragenic organization or nucleotide sequence divergence. Taken in conjunction with the knowledge that animal mtDNA is evolving very quickly (Brown *et al.*, 1979) and that the COII gene is one of the most rapidly evolving mitochondrial genes in primates (Cann *et al.*, 1984), this prompted us to examine the COII gene in a number of higher plants.

We have determined the sequence of the wheat COII gene and have used exon- and intron-specific regions of this gene to examine by hybridization analysis the COII genes in rye, pea, broad bean and cucumber. We find that these genes are all closely-related, but that their length varies depending on the presence or absence of an intron (or intron insert). We also find that 5′ exon sequences are repeated in the mtDNA of the monocotyledons.

Results

Identification and sequence of the wheat COII gene

We identified wheat mtDNA restriction fragments that hybridized to the cloned maize COII gene (Fox and Leaver, 1981) and selected *Eco*RI and overlapping *Sal*I recombinant clones from our wheat mtDNA pUC9 clone bank for sequencing (Figure 1). The sequence of the wheat COII gene and its flanking regions is shown in Figure 2. The gene is 1997 nucleotides long with an intron at position 391, the same position as in maize. However, the wheat COII intron has an insert of 422 nucleotides at position 1062-1064 (dashed underline, Figure 2). (The precise position is ambiguous because it lies within a short repeat, as discussed below.) Both the COII coding regions and intron sequences are equally

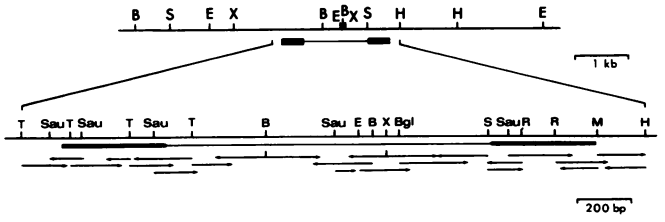

Fig. 1. Restriction map of the wheat mitochondrial COII gene and sequencing strategy. Regions of two *Eco*RI clones and an overlapping *Sal*I clone were subcloned into M13 vectors and sequenced as indicated by arrows. Only restriction sites used in sequencing are shown: B, *Bam*HI; Bgl, *Bgl*II; E, *Eco*RI; H, *Hind*III; M, *Msp*I; R, *Rsa*I; S, *Sal*I; Sau, *Sau*3A; T, *Taq*I; X, *Xho*I. Solid bars indicate exons.

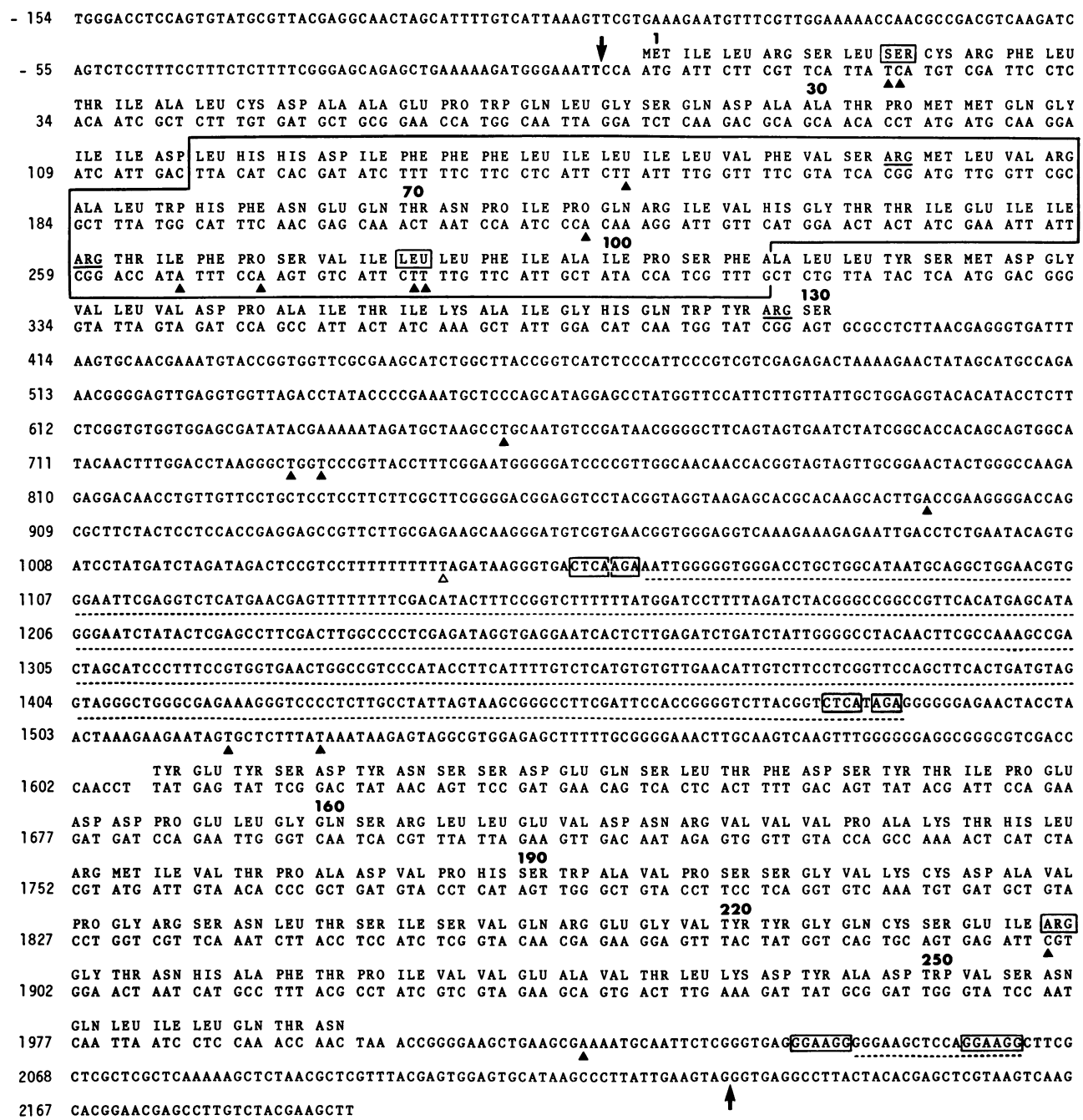

Fig. 2. Nucleotide sequence of the wheat mitochondrial gene for cytochrome oxidase subunit II and its predicted amino acid sequence. CGG codons at amino acid positions 57, 87 and 129 (underlined) likely specify Trp rather than Arg (see text, and Fox and Leaver, 1981; Hiesel and Brennicke, 1983). Single nucleotide substitutions relative to the maize sequence (a total of 15 within the gene and one downstream; comprised of nine transitions and seven transversions and one addition within the intron) are indicated by solid and open triangles, respectively. The three predicted amino acid replacements relative to maize are boxed. Inserts in the intron and downstream region are underlined (dashed) and flanking repeats at the insert sites are boxed. The downstream arrow indicates the end of the maize sequence reported in Fox and Leaver (1981) and the upstream arrow indicates the start of obvious sequence similarity between the wheat and maize COII sequences. The wheat sequence begins at a *TaqI* site immediately preceding position −154. The 193 nucleotide region of the 5′ exon which is found at a second genomic location is blocked.

well-conserved (98.9% and 99.3% nucleotide identity, respectively) between wheat and maize.

Within the coding regions, only nine nucleotide changes were found, five of which are silent substitutions. The other four differences result in codon changes that predict three amino acid replacements (Figure 3). Surprisingly, all three substitutions are non-conservative ones, namely Glu→Ser at position 7, Pro→Leu at position 95, and Cys→Arg at position 228 (see Discussion).

The region immediately downstream of the wheat and maize COII coding sequences is also highly conserved (98.3% nucleotide identity over 116 nucleotides), with one insert of 17 nucleotides in wheat (dashed underline, Figure 2). In contrast, the upstream sequences show only 27% nucleotide identity (in the 105 nucleotides of comparable region sequenced in maize; Fox and Leaver, 1981). Obvious nucleotide sequence similarity begins three nucleotides before the AUG that is likely to be the initiation codon. This is the second of two in-frame AUG codons seen in maize.

A comparison of the wheat and *Oenothera* (Hiesel and

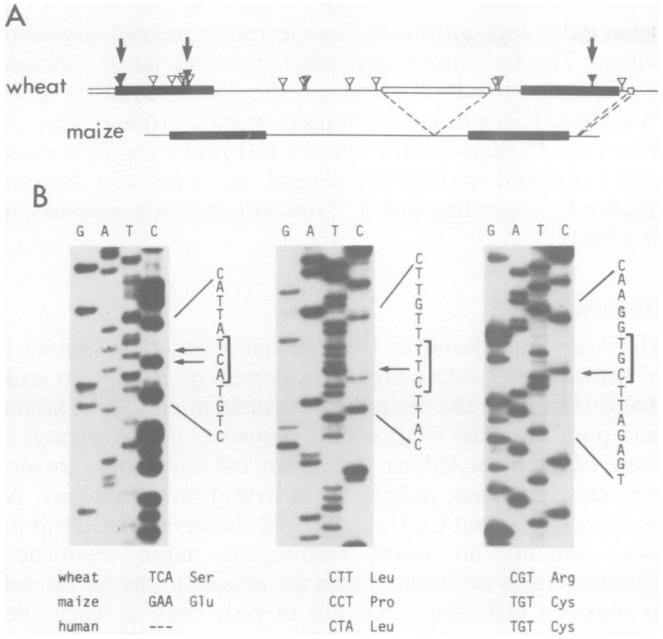

Fig. 3. **(A)** Schematic comparison of the wheat and maize COII genes, with exons being shown by black bars and sequences not held in common by open bars. Single nucleotide differences are indicated by triangles with the ones predicting amino acid substitutions indicated by solid triangles and arrows. **(B)** Sequencing gels showing nucleotide positions (arrows) which predict amino acid differences between wheat and maize. The first panel is shown mirrored to facilitate direct reading of the sense strand sequence. The relevant codons are bracketed and compared below with those of maize (Fox and Leaver, 1981) and human (Anderson *et al.*, 1981). The human COII protein lacks the amino-terminal region in which the first wheat/maize difference is seen.

Brennicke, 1983) COII genes shows 88.0% nucleotide identity (and 84.6% amino acid identity) with no obvious sequence similarity outside the coding sequence. As in maize and *Oenothera*, wheat mitochondrial codon usage appears to deviate from the universal genetic code in that CGG likely specifies Trp rather than Arg (Fox and Leaver, 1981; Hiesel and Brennicke, 1983).

Transcripts of the wheat COII gene

The wheat COII mRNA was characterized by Northern hybridization analysis using M13-cloned probes specific to the 5′ exon, intron and 3′ exon (Figure 4). An abundant transcript of ~1.5 kb was detected with exon probes and a minor transcript of 1.2 kb with the intron probe. All probes revealed minor levels of a large transcript ~2.7 kb in length, which is the expected size of an intron-containing precursor.

Additional wheat COII 5′ exon-specific sequences

To determine whether any additional copies of COII gene-specific sequences are present in total wheat mtDNA, we conducted Southern hybridization experiments using the same exon- and intron-specific probes as in the transcript analysis. The 3′ exon and intron probes showed hybridization patterns consistent with those expected from restriction mapping and sequence analysis of the COII gene (data not shown). However, the 5′ exon-specific probe (a 224-bp *Taq*I fragment, Figure 1), showed intense hybridization to a second restriction fragment (square, Figure 5) in addition to the one containing the sequenced COII gene (triangle, Figure 5). We determined the sequence of the additional hybridizing region

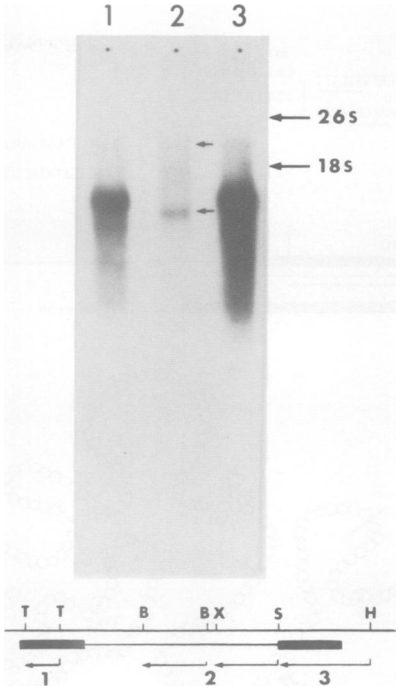

Fig. 4. Transcript pattern of the wheat COII gene. Northern blots of wheat mtRNA were hybridized with wheat COII probes specific to the 5′ exon (**lane 1**), intron (**lane 2**), and 3′ exon (**lane 3**) as shown below. The intron probe was a mixture of two M13 clones. Intron and precursor bands in **lane 2** are indicated with arrows. 26S and 18S rRNA size markers were determined by methylene blue staining of the blots (Maniatis *et al.*, 1982). Restriction site abbreviations are as in Figure 1.

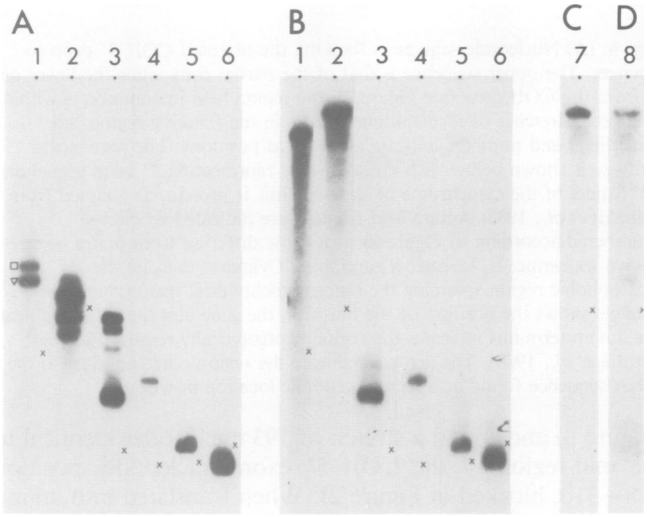

Fig. 5. Detection of sequences homologous to the wheat COII gene in other higher plant mtDNAs. Restriction digests of: **lane 1**, wheat (*Bam*HI); **lanes 2, 7, 8**, rye (*Bam*HI); **lane 3**, maize (*Eco*RI); **lane 4**, cucumber (*Eco*RI); **lane 5**, broad bean (*Eco*RI); and **lane 6**, pea (*Eco*RI) were electrophoresed on agarose gels, blotted to nitrocellulose and individual strips were hybridized with probes from **(A)** 5′ exon, **(B)** 3′ exon, **(C)** intron-insert and **(D)** intron regions. Probes in **(A)**, **(B)** and **(D)** are as in Figure 4 and in **(C)** an 80-bp *Taq*I-*Xho*I clone (positions 1139–1218, Figure 2) was used. Less DNA is present in **lane 1** (wheat mtDNA). A size marker of ~2 kb is indicated by the cross (x) in each track. In **lane 1**, a triangle indicates the *Bam*HI fragment containing the 5′ region of the wheat COII gene (see Figure 1) and the square indicates the fragment containing the repeated copy of 5′ exon sequences.

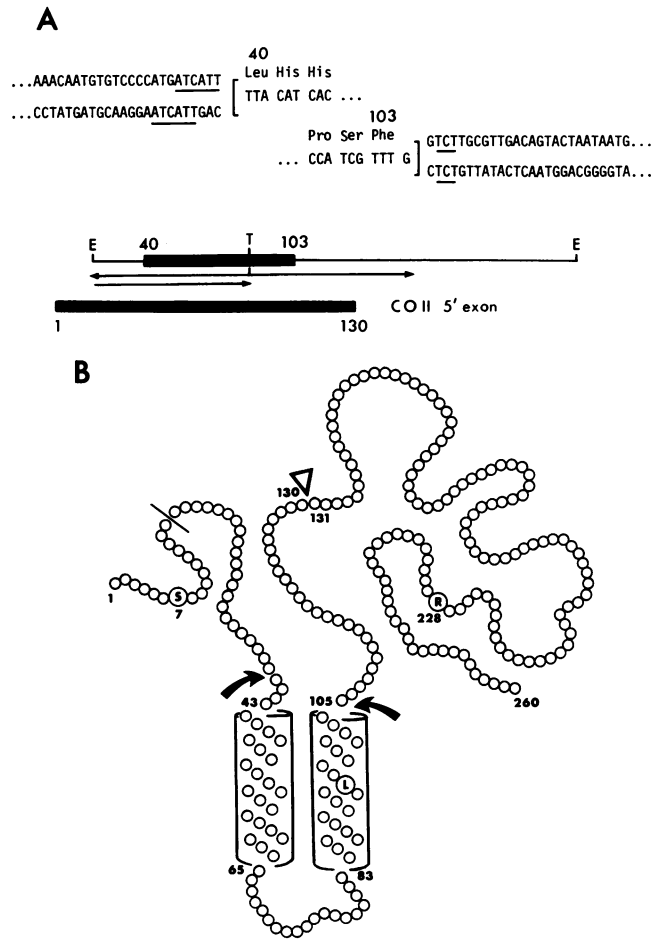

Fig. 6. (A) Nucleotide sequences flanking the repeated COII 5′ exon sequences. The upper sequence is that of the partial copy while the lower one is from the COII gene (see Figure 2); the region held in common is within brackets. Stretches of identical nucleotides in the flanking regions are underlined and numbers indicate amino acid positions. The sequencing strategy is shown below with the solid bars representing 5′ exon sequences. (B) Model of the cytochrome oxidase subunit II protein, as adapted from Capaldi *et al.*, 1983. Amino acid residues are indicated by circles (numbered according to Figure 2) with those differing from maize being shown (S, serine; L, leucine; R, arginine). Cylinders indicate the hydrophobic region spanning the inner mitochondrial membrane. The triangle shows the position of the intron in the gene and the solid line near the amino terminus indicates the region proteolytically removed in yeast (Pratje *et al.*, 1983). The arrows delineate the region corresponding to the DNA sequence found at a second genomic location in wheat.

(Figure 6) and found a stretch of 193 nucleotides identical to the mid-region of the COII 5′ exon (nucleotides position 118–310, blocked in Figure 2). When translated into amino acids, this region corresponds to the transmembrane domain of the COII protein in the model presented by Capaldi *et al.* (1983) (Figure 6, arrows). It too contains the Pro→Leu substitution relative to the maize COII gene. In fact, six of the nine nucleotide differences seen between the wheat and maize COII coding sequences are clustered within this 193 nucleotide stretch.

COII genes in other higher plants

We searched for sequences homologous to wheat COII exon and intron regions in the mtDNA of rye, pea, broad bean and cucumber, again using Southern hybridization analysis. In the case of pea, bean, and cucumber, the same single restriction fragment hybridized to both 5′ exon and 3′ exon wheat COII probes (Figure 5). However, no hybridization was detected with the intron probe, suggesting that these COII genes either lack introns or have introns unrelated to that of wheat. The two monocotyledons, rye and maize, showed hybridization to wheat intron-specific probes, while rye alone hybridized to a wheat intron-insert-specific probe (Figure 5). When the 5′ exon-specific wheat COII probe was hybridized with maize and rye mtDNA, several fragments were detected (Figure 5), suggesting that 5′ exon sequences are repeated, as in wheat.

Discussion

The wheat mitochondrial DNA sequence encoding subunit II of cytochrome oxidase predicts a protein of 260 amino acids that differs from the maize COII protein at only three amino acid positions. The COII coding sequences of rye, as well as those of the dicotyledons, pea, bean and cucumber, are also very closely-related, judging from hybridization analysis. As in maize, the wheat COII gene is split; however, the wheat intron contains an insert relative to maize. Sequences homologous to the wheat intron are present in the rye (as well as maize) COII genes, but not in pea, bean or cucumber mtDNA, either within the COII gene or elsewhere in their genomes. The only mtDNA among these plants that hybridized with the wheat COII intron-insert region was that of rye. These observations imply that the COII intron (and intron insert) found in monocotyledons did not arise by rearrangement events translocating DNA sequences which are extragenic in dicotyledons to intragenic positions in monocotyledons.

Although the precise evolutionary distances among these plants are not yet known, our findings point to a slower rate of evolution in higher plant mtDNA than that seen in animal mtDNA. We see 99% nucleotide identity between the wheat and maize COII genes, whereas 92% was observed between those of two closely-related rat species, *Rattus rattus* and *Rattus norvegicus* (Brown and Simpson, 1982). In primate mtDNA, the rate of mutation fixation in COII genes is even higher (Cann *et al.*, 1984); 'radical' amino acid replacements were predicted even among individual humans ('radical' as defined by pronounced differences in amino acid polarity, composition, and molecular volume; Grantham, 1974; Cann *et al.*, 1984). The fact that we *do* observe several non-conservative changes between wheat and maize may provide insight into constraints on protein structure. The Glu→Ser substitution near the amino terminus (Figure 3) is in a region of low amino acid homology between maize and yeast (Fox and Leaver, 1981), and in yeast this stretch is removed proteolytically from the precursor COII protein (Pratje *et al.*, 1983). The Pro→Leu replacement occurs in the transmembrane domain of the protein (Capaldi *et al.*, 1983) and presumably affects α-helical structure. Almost all organisms, including wheat, have Leu at this position. The third difference, Cys→Arg, is in the region which contains several His and Cys residues implicated in copper binding. From comparative sequence analysis, five conserved His and Cys residues are candidates as copper ligands (Capaldi *et al.*, 1983). However, in wheat we find that Cys_{228} (position 196 in the beef protein, Steffens and Buse, 1979) is predicted to be replaced by Arg (Figure 3). This would reduce the total number of conserved His and Cys residues to four; namely, His_{24}, His_{161}, His_{204} and Cys_{200} (numbered according to the beef protein). The possibility that CGU encodes Cys rather than Arg in wheat (analogous to the CGG Trp *versus* Arg

deviation in the genetic code mentioned above) seems unlikely in view of the number of CGU codons at conserved Arg positions. In addition, we do not find 3′ exon sequences repeated elsewhere in the wheat mitochondrial genome (as judged by the exclusive hybridization of the 3′ exon probe to single bands in digests and double digests of wheat mtDNA).

The wheat and maize intron sequences are as well-conserved as the COII coding regions, suggesting that there are equally strong constraints on them. This differs from the situation in eukaryotic nuclear genes, where intron nucleotide conservation is low (cf. mouse *versus* rabbit β-globin exon/intron: 88%/40% identity, respectively; van Ooyen *et al.*, 1979) and is consistent with the view that folding of the mitochondrial intron RNA is important for splicing of the precursor transcript. When the sequence of the wheat COII intron is folded according to the secondary structure model proposed for maize (Michel and Dujon, 1983), only one nucleotide change in the core structure is seen (resulting in an A-U pair *versus* G-U in maize). However, the insert in the wheat COII intron makes it 1.5 times longer than its maize counterpart. The insert occurs in a non-conserved region of the secondary structure model, a region that is variable in length among members of this class of introns (as defined in Michel and Dujon, 1983). We find an open reading frame of 138 amino acids in the wheat COII intron insert, but it shows no detectable amino acid homology with those in yeast mitochondrial introns (Bonitz *et al.*, 1980) which have related secondary structure (Michel and Dujon, 1983). The intron insert is located in a stretch of purine residues and is flanked by a direct repeat of eight nucleotides (with one mismatch) (boxed region, Figure 2). This resembles the duplication of target site DNA seen in transpositional events (Calos and Miller, 1980); however, the insert is present only at this location on the wheat mtDNA as determined by hybridization analysis. The second insert relative to maize, a 17-nucleotide stretch in the downstream region, also occurs in a purine-rich stretch and is flanked by a direct repeat of six nucleotides (boxed, Figure 2). Aside from this insert, the downstream sequence (of 116 nucleotides) is as highly conserved as the COII gene.

In contrast, the upstream sequences show no sequence similarity between wheat and maize until three nucleotides before the presumptive initiation codon. This raises questions about signals for the initiation of translation, which would be expected to be conserved between such closely-related genes. Presumably the signals differ from those in prokaryotes, since the mature small subunit rRNA in wheat mitochondria lacks a classical Shine and Dalgarno sequence (Schnare and Gray, 1982). In addition, the presence of an out-of-frame AUG closely preceding the wheat initiation codon (positions −11 to −13, Figure 2) does not support a eukaryotic-type scanning model for initiation (Kozak, 1978). One prominent feature of the upstream region is a stretch rich in purines preceded by a string of pyrimidines; a similar composition is seen preceding the *Oenothera* COII gene (Hiesel and Brennicke, 1983).

Our Northern hybridization analysis demonstrates the presence of a wheat COII mRNA of ~1.5 kb (about twice the length of the coding sequence), present at levels comparable with that of the wheat COI and cytochrome b mRNAs (unpublished observations). We also see an intron-specific transcript of 1.2 kb, corresponding to a full-length excised intron RNA. Although its physical conformation is not known, this transcript is present at relatively low levels compared with those of yeast mitochondrial circular introns (Hensgens *et al.*, 1983). The wheat COII transcript pattern is much simpler than that seen in maize, where numerous transcripts were observed, promoting the suggestion that there may be several COII mRNAs (Fox and Leaver, 1981). The differences in transcript patterns may be related to differences in upstream sequences between the two plant COII genes.

Our analysis has also uncovered a second copy of 193 nucleotides identical to the part of the 5′ exon that corresponds to the transmembrane domain of the wheat COII protein. We are currently investigating the functional significance, if any, of this extra copy. It may be a pseudogene resulting from recombinational or transpositional events, but other possibilities include a multiply-split second COII gene copy, or even a domain of another membrane protein. Because probes from the 3′ exon and intron regions appear not be represented elsewhere on the wheat mtDNA, this implies that if the second 5′ exon copy is part of a functional COII gene, its expression must involve alternative splicing to the unique 3′ exon sequences.

Although this second copy is readily apparent in hybridization experiments using highly specific probes (Figure 5), it was detected only as a minor signal when longer probes, such as the 2.5-kb *Eco*RI fragment (Figure 1) were used. Consequently, the presence of such short repeats could well be a common feature of higher plant mtDNA (cf. Ward *et al.*, 1981), but one that escapes detection in hybridization and kinetic renaturation analyses.

After these studies were completed, we learned that the rice COII gene also has an intron insert at the same position and exhibits strong homology with the one that we have characterized in wheat (T.-H. Kao, E. Moon and R. Wu, personal communication).

Materials and methods

Identification, cloning and sequence analysis of the wheat COII gene

MtDNA, isolated from wheat embryos (*Triticum aestivum* L.) which were germinated for 24 h (Bonen and Gray, 1980) was used to construct *Eco*RI, *Sal*I and *Bam*HI clone banks in the plasmid vector pUC9 (Vieira and Messing, 1982), using standard recombinant DNA techniques (Maniatis *et al.*, 1982). Recombinant clones, as well as restricted total wheat mtDNA, were screened for sequences homologous to the maize COII gene using the clone pZmE1 (Fox and Leaver, 1981) labelled by nick translation. Hybridizations were conducted at 5 x SSC, 0.1% SDS, 2 x Denhardt's, 20 μg/ml sheared, denatured salmon sperm DNA, 65°C overnight. In initial experiments, 60°C conditions were used. Two *Eco*RI clones and an overlapping *Sal*I clone were used for sequencing the wheat COII gene by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) after subcloning fragments into M13 vectors. The 5′ exon partial copy was sequenced from an *Eco*RI subclone of a *Bam*HI clone containing the fragment indicated by a square in Figure 5.

Northern hybridization analysis

Wheat mtRNA isolated from embryos germinated for 24 h (Bonen and Gray, 1980) was fractionated by precipitation from 1 M NaCl and the insoluble fraction was electrophoresed (1−2 μg/lane) on 1.5% agarose, 2.2 M formaldehyde gels (Lehrach *et al.*, 1977). Nitrocellulose blots (prepared and prehybridized according to Thomas, 1980) were hybridized with wheat COII probes (labelled by M13 primer extension) in 50% deionized formamide, 5 x SSC, 2 x Denhardt's, and 20 μg/ml denatured, sheared salmon sperm DNA at 42°C, overnight.

Southern hybridization analysis of rye, pea, bean, cucumber and maize mtDNA

MtDNA isolated from maize (*Zea mays* L., cv. Golden Bantam Early), rye (*Secale cereale* L., v. Kodiak), pea (*Pisum sativum* L., cv. Laxton Progress), broad bean (*Vicia faba* L., cv. Broad Windsor) and cucumber (*Cucumis sativus* L., cv. Long Green) as described by Huh and Gray (1982) was restricted, electrophoresed on separate gels, and blotted to nitrocellulose (Southern, 1975). Individual strips were hybridized with wheat COII probes (labelled by M13 primer extension) at 5 x SSC, 0.1% SDS, 2 x Denhardt's,

20 μg/ml sheared, denatured salmon sperm DNA at 65°C, overnight, with washes to 0.5 x SSC. Representative u.v. fluorescence patterns of such restricted mtDNAs are shown in Huh and Gray (1982).

Acknowledgements

We thank C.J. Leaver for the maize COII gene clone, T.Y. Huh for preparing the Southern blots and T.-H.Kao, E.Moon and R.Wu for communicating their unpublished results. Financial support from the Medical Research Council of Canada and the Dalhousie Medical Research Foundation is gratefully acknowledged.

References

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.

Bonen,L. and Gray,M.W. (1980) *Nucleic Acids Res.,* **8**, 319-335.

Bonitz,S.G., Coruzzi,G., Thalenfeld,B.E., Tzagoloff,A. and Macino,G. (1980) *J. Biol. Chem.,* **255**, 11927-11941.

Brown,G.G. and Simpson,M.V. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 3246-3250.

Brown,W.M., George,M. and Wilson,A.C. (1979) *Proc. Natl. Acad. Sci. USA,* **76**, 1967-1971.

Calos,M.P. and Miller,J.H. (1980) *Cell,* **20**, 579-595.

Cann,R.L., Brown,W.M. and Wilson,A.C. (1984) *Genetics,* **106**, 479-499.

Capaldi,R.A., Malatesta,F. and Darley-Usmar,V.M. (1983) *Biochim. Biophys. Acta,* **726**, 135-148.

Falconet,D., Lejeune,B., Quetier,F. and Gray,M.W. (1984) *EMBO J.,* **3**, 297-302.

Fox,T.D. and Leaver,C.J. (1981) *Cell,* **26**, 315-323.

Grantham,R. (1974) *Science (Wash.),* **185**, 862-864.

Hensgens,L.A.M., Arnberg,A.C., Roosendaal,E., van der Horst,G., van der Veen,R., van Ommen,G.J.B. and Grivell,L.A. (1983) *J. Mol. Biol.,* **164**, 35-58.

Hiesel,R. and Brennicke,A. (1983) *EMBO J.,* **2**, 2173-2178.

Huh,T.Y. and Gray,M.W. (1982) *Plant Mol. Biol.,* **1**, 245-249.

Kozak,M. (1978) *Cell,* **15**, 1109-1123.

Leaver,C.J. and Gray,M.W. (1982) *Annu. Rev. Plant Physiol.,* **33**, 373-402.

Lehrach,H., Diamond,D., Wozney,J.M. and Boedtker,H. (1977) *Biochemistry (Wash.),* **16**, 4743-4751.

Lonsdale,D.M., Hodge,T.P., Fauron,C.M.-R. and Flavell,R.B. (1983) in Goldberg,R.B. (ed.), *UCLA Symposia on Molecular and Cellular Biology,* New Series, Vol. **12**, Alan R. Liss, Inc., NY, pp. 445-456.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning*, published by Cold Spring Harbor Laboratory Press, NY.

Michel,F. and Dujon,B. (1983) *EMBO J.,* **2**, 33-38.

Palmer,J.D. and Shields,C.R. (1984) *Nature,* **307**, 437-440.

Pratje,E., Mannhaupt,G., Michaelis,G. and Beyreuther,K. (1983) *EMBO J.,* **2**, 1049-1054.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA,* **74**, 5463-5467.

Schnare,M.N. and Gray,M.W. (1982) *Nucleic Acids Res.,* **10**, 3921-3932.

Southern,E.M. (1975) *J. Mol. Biol.,* **98**, 503-517.

Steffens,G.J. and Buse,G. (1979) *Hoppe-Seyler's Z. Physiol. Chem.,* **360**, 613-619.

Stern,D.B. and Lonsdale,D.M. (1982) *Nature,* **299**, 698-702.

Stern,D.B. and Palmer,J.D. (1984) *Proc. Natl. Acad. Sci. USA,* **81**, 1946-1950.

Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 5201-5205.

van Ooyen,A., van den Berg,J., Mantei,N. and Weissmann,C. (1979) *Science (Wash.),* **206**, 337-344.

Vieira,J. and Messing,J. (1982) *Gene,* **19**, 259-268.

Ward,B.L., Anderson,R.S. and Bendich,A.J. (1981) *Cell,* **25**, 793-803.

Received on 23 July 1984