

SEQUENCING OF CYTOCHROME OXIDASE SUBUNIT II AND APOCYTOCHROME B GENES OF SUNFLOWER MITOCHONDRIAL DNA

L. Siculella; L.R. Ceci;^o A.S. Treglia;^o C. De Benedetto;^{oo} D. Pacoda;^o C. Ferrotta; R. Gallerani;^{ooo}
^oDipartimento di Biologia, Università di Lecce, Via Monteroni, 73100 Lecce; ^{oo}Sezione di Trani del Centro di Studio sui Mitochondri e Metabolismo Energetico, C.N.R. Via Corato, 70059 Trani; ^{ooo}Dipartimento di Biochimica e Biologia Molecolare, Università di Bari, Via Amendola 165/A, 70126 Bari Italia.

SUMMARY

Sunflower mitochondrial genes coding for apocytochrome b (CYB) and cytochrome oxidase subunit II (COXII) have been fully sequenced. The CYB gene is continuous being constituted by an uninterrupted reading frame of 1,194 bp; on the contrary COXII gene appears to be discontinuous. An intron interrupts two exons having 384 and 447 bp length respectively.

Nucleotide and amino acid sequences of coding regions of both genes are highly homologous with the corresponding sequenced so far in other higher plants.

The intron present in the COXII gene is unique; it is the first intron localized and sequenced in the COXII gene of a dicot plant and its structure results very similar to that of rice.

Comparative mapping studies of mitochondrial DNA (mtDNA) in fertile and male sterile line (CMS89) reveal a rearrangement in the region bearing the CYB gene.

INTRODUCTION

COXII and CYB genes have been sequenced so far in several different systems: maize (Fox and Leaver 1981), (Dawson et al. 1984); wheat (Bonen et al. 1984), (Boer et al. 1985); cenothena (Hiesel and Brennicke 1983), (Schuster and Brennicke 1985). In rice (Kao et al. 1984), pea (Moon et al. 1985) and soy bean (Grabau 1987) only the COXII gene has been sequenced. In this last system an initiator tRNA has been localized at a distance of 194 bp far from the 5' termini of COXII gene.

In monocot plants this gene is discontinuous for the presence of an intron of 794 bp in maize (Fox and Leaver 1981), 1,213 bp in wheat (Bonen et al. 1984) and 1,265 bp in rice (Kao et al. 1984) which interrupts the coding region exactly at the same position.

The CYB gene is a continuous reading frame of about 1,200 bp showing relevant nucleotide homology levels at its 5' adjacent non coding regions. In the comparison wheat/cenothena this homology has an extension of about 600 nucleotides. On the contrary in maize and cenothena a G/A rich stretch of 75 bp is present in the same region.

In these systems a putative ribosome binding site, fifteen nucleotides far from the initiation codons, has been identified. This ribosome binding site (an octanucleotide having the sequence 5'-AGT TGICA-3') is 68% complementary to a sequence located at the 3' termini of 18S ribosomal mitochondrial RNA of maize and cenothena (Dawson et al. 1984).

With the aim of better understanding the structure of both coding and regulatory regions of some mitochondrially coded protein genes in sunflower, we localized on different restriction fragments, COXII (Ferrotta et al. 1986) and CYB genes (Pacoda et al. 1987) and sequenced them and part of 5' and 3' adjacent non coding regions (Ceci et al. 1988), (Treglia et al. 1988).

Sequencing analysis of COXII gene confirmed our preliminar observation (Gallerani et al. 1987) demonstrating the presence of an intron having a length similar to those present in rice (Kao et al. 1984) and wheat (Bonen et al. 1984). In sunflower indeed a 1,331 bp long intron interrupts the coding region after a serine and before a tyrosine residue as in monocot plants, maize, wheat and rice. However the insertion sequence of 505 bp in sunflower (Ceci et al. 1988), corresponding to the equivalent of 461 bp in rice (Kao et al. 1984) which shows all the peculiar characteristics of a transposable element, does not show the same properties owing to relevant differences of primary structure at its 5' and 3' termini.

The sunflower CYB gene shows as initiation codon the unusual triplet GUG. Sequence homologies both

of nucleotides and derived amino acids of sequenced region on sunflower mtDNA with those of other CYB genes sequenced so far in higher plants, confirmed its identity with that of apocytochrome b gene. Further support to this conclusion was given by high similarities between hydrophathy profile of deduced sunflower protein and that of beef apocytochrome b (Saraste 1984).

Restriction analysis of a region of the master chromosome of sunflower mtDNA bearing the CYB gene reveals in the male sterile line CMS89 a 12 kbp inversion mapping close to the gene. This rearrangement however does not affect its transcriptional pattern.

MATERIALS AND METHODS

The isolation of mitochondria and the extraction of mtDNA is reported elsewhere (Ceci et al. 1988). Restriction and hybridization analysis were carried out by using maize COXII and CYB probes. Hybridization conditions were: 5xSSC, 0.1% SDS, 2xDenhardt's, 0.2 mg/ml carrier DNA, 65 °C. Restriction enzyme digestions were carried out by using the conditions recommended by the suppliers (Boehringer Mannheim). Sequencing of the genes was as reported by Maxam and Gilbert (Maxam and Gilbert 1977) and Korneluk (Korneluk et al. 1985).

RESULTS

Apocytochrome b gene

The region of sunflower mtDNA containing the CYB gene was originally identified on two different SalI and HindIII fragments by hybridization and restriction analysis (Pacoda et al. 1987). The precise position of the gene on these fragments was determined by alignment of their sequences with those of other CYB genes reported so far. A continuous reading frame of 1,194 nucleotides was identified, having very high homology (about 95%, see table I) with corresponding genes sequenced so far in maize, wheat and oenothera.

Table I

Sequence homologies of coding and 5' adjacent non coding regions of CYB genes of sunflower wheat oenothera and maize mitochondria.

	5' non coding regions	coding regions
sunflower/wheat	45	94
sunflower/oenothera	49	95
sunflower/maize	47	95

The homology values of coding regions refer to sequences of the entire gene.

The observed high homology levels of coding regions led also to the identification of the initiation codon (table II). As it is possible to see in table I the comparison of sequenced regions on

Table II

Alignment of 5' termini and adjacent non coding regions of CYB genes of sunflower wheat oenothera and maize mitochondria.

	-40	-30	-20	-10	0	
sunflower	GIGAGG	GAACCTGAGT	TCCTCTTCTA	AAAAATCAAA	ATAAAAATAA	GIG --- ACT ATA AGG AAC CAA ...
wheat	AA G	GAG	G AACGA	G GAG	A	---
oenothera	G	AG	G ACGA	G AAGG	A	GCG
maize	AA G	GAG	G ACGA	G GAG	A	---

In this table 45 nucleotides upstream the 5' termini of four different CYB genes and the first 18 (21 in oenothera) of coding regions are reported. Only those nucleotides which differ from sunflower gene are shown for wheat, oenothera and maize. Dashes indicate gaps necessary to align the sequences.

the 5' site of the genes allowed the estimation of a much lower homology level (45% vs. 95%). Furthermore in sunflower, just before the initiation codon, a decanucleotide insertion was detected. It contains two TAA stop codons separated by an AAA triplet, in register with the coding region. The only part of the gene having slight differences among those sequenced so far, concerns its 3' termini. The deduced protein results indeed, regarding to the shortest (maize, 388 residues) (Dawson et al. 1984), nine, eleven and six amino acids longer in sunflower (Treglia et al. 1988), wheat (Boer et al. 1985) and oenothera (Schuster and Brennicke 1985) respectively.

Cytochrome oxidase subunit II gene

Sequencing analysis of HindIII insert of the recombinant plasmid pCP 798, part of a Sall and HindIII library of sunflower mtDNA restriction fragments (Perrotta et al. 1986) led to the identification of 2,165 nucleotides corresponding to the gene of cytochrome oxidase subunit II of sunflower mitochondria (Ceci et al. 1988). The structure of this gene, compared to the most similar (rice) among those sequenced so far in higher plants, is reported in figure 1.

The sunflower COXII gene has a discontinuous organization revealed so far only by the equivalent genes in monocot plants. In the case of rice (Kao et al. 1984) and sunflower in particular, in the non coding part of the genes three different regions can be detected. The first, flanking the 3' termini of first exon and the 5' termini of second exon, is highly homologous to equivalent regions present also in wheat (Bonen et al. 1984) and in maize (Fox and Leaver 1981) (dotted regions in figure 1). The second, common to rice, wheat and sunflower, shows in the first two systems the main characteristics of a transposable element (dashed regions). The third is a small insertion sequence present only in rice and sunflower intron having in both systems the same length (52 bp, black regions).

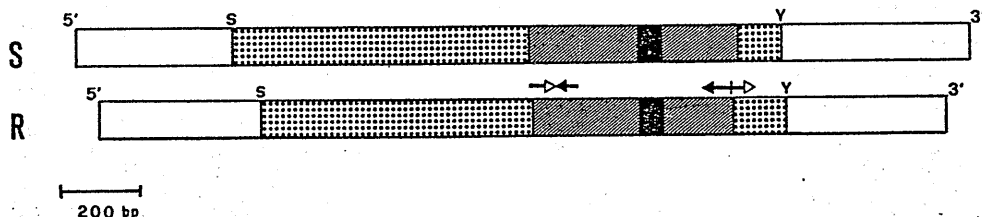


Figure 1. Organization of sunflower and rice mitochondrial COXII genes.

White regions correspond to the exons. The arrows indicate the relative positions of direct (→) and inverted repeats (⇌). S (serine) and Y (tyrosine) indicate the last amino acid of first exon and the first of second exon.

As far as the coding regions are concerned, the most relevant domains identified in the beef cytochrome oxidase subunit II, according to the models proposed by Capaldi (Capaldi et al. 1983) and Holm (Holm et al. 1987), are almost completely confirmed in the sunflower protein (Ceci et al. 1988)

Restriction analysis of CYB gene region in male sterile lines

In figure 2 the comparative restriction analysis of a 20 kbp region of sunflower mtDNA, both from fertile and sterile lines is reported. The results clearly show that the map is coincident for the first 8 kbp whilst it results inverted for the remaining 12 kbp. The end point of the inversion maps 2 kbp far from the CYB gene.

DISCUSSION

The relevant differences in the organization of 5' non coding region of mitochondrial CYB gene in

sunflower, compared with the equivalent of other homologous genes sequenced so far led us to consider other elements which could demonstrate that the sequence we obtained corresponded to the *CYB* gene. The most relevant observations are listed below. The high homology levels both for nucleotides and deduced amino acids of the region sequenced on sunflower mtDNA restriction fragments (table I) (Treglia et al. 1988) support the hypothesis that the sequences we obtained correspond to those of sunflower *CYB* gene. On the other hand the detection of two stop codons localized at the 5' site of GUG triplet and in register with it rules out the possibility that any other AUG upstream the deca nucleotide 5'-ATAAAAATAA-3' could codify the initiator methionine. This observation also confirms the identification of GUG as the initiator codon. Finally, the hydropathy profile obtained by using the Kyte and Doolittle procedure (Kyte and Doolittle 1982) on deduced amino acid sequence of sunflower protein resulted identical to that obtained in the case of beef (Saraste 1984) and maize apocytochrome b (Dawson et al. 1984). As in the case of these proteins the sunflower hydropathy profile (Treglia et al. 1988) shows nine hydrophobic transmembrane domains corresponding to α -elix structures of the protein which should lie in the inner mitochondrial membrane. These results all together confirm that the region we sequenced is indeed the mitochondrial gene for the sunflower apocytochrome b.

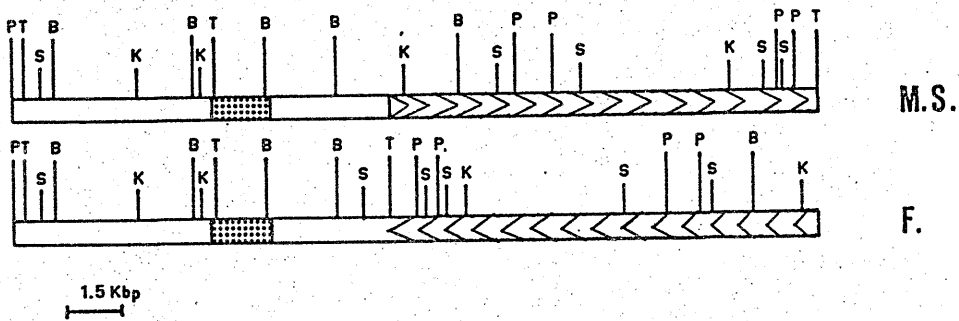


Figure 2. Restriction map of a 20kbp region of sunflower mtDNA of fertile and CMS89 male sterile line bearing the *CYB* gene.

Dotted areas correspond to the *CYB* gene. The opposite orientation of 12kbp regions is demonstrated by the inverted relative positions of restriction sites. A more detailed map of the same region is reported elsewhere (Siculella and Palmer 1988).

The analysis of sequencing data regarding the sunflower COXII gene revealed a characteristic feature of this gene. It has, as reported in figure 1, a discontinuous organization very similar to that of rice. As it was mentioned before, the intron present in this gene contains a 461 bp insertion element flanked by direct and inverted repeats and hence it has the characteristics of a transposable element. In sunflower the regions corresponding to the inverted and direct repeats of rice intron appear altered by nucleotide insertions or deletions of different length. In table III the comparison of primary structures of the intron of the above mentioned regions are reported. On the basis of these observations we conclude that in sunflower an insertion sequence highly homologous to the equivalent region of rice intron is present. However in sunflower this element has not the same characteristics of mobility as in rice.

The observation that in the sunflower male sterile line CMS89 a wide region flanking the *CYB* gene is inverted, suggested the possibility that this inversion could affect somehow the gene transcription. Transcript analysis (Siculella and Palmer 1988) carried out by using specific probes containing the 12kbp inversion or part of it and/or the *CYB* gene showed for this region an identical tran-

Table III

Alignment of sequences of sunflower COXII gene intron corresponding to regions of direct and inverted repeats in rice.

		1050	1060		1070
rice	-----	AAGGGT	GACTCAA	AAGAAA	-----TTGGGGGTGGGA--CC
	670	680	690	700	710
sunflower	TCTTTTTCTGAAAAA	AAAAAGAAGAAGGGT	GACTCAA	-----CTTCTCAGCTAGAGTTGGGGGTGGGATCC	
		1520	1530	1540	1550
rice	CACC-----	GGGGTCTTACGGT	TTCATAGAGGGGGGAGA	ACTACCTAACTAAAG	
	1200	1210	1220	1230	
sunflower	CCCCTTACTCATAAAGGGGTCTTACGGTC	-----	GGTA-CTAACTAAAG		

The numbering of sunflower gene is as reported by Ceci et al. (Ceci et al. 1988); that of rice gene is according to Kao et al. (Kao et al. 1984).

scriptional pattern both in male sterile and fertile lines. These observations demonstrate the 12 kbp inversion does not affect the transcription of CYB gene.

CONCLUSIONS

Sequencing analysis of CYB and COXII genes of sunflower mtDNA revealed interesting features for both of them. The sunflower CYB gene shows a characteristic structure at its 5' non coding region and the unusual GUG as initiator codon.

The sequencing of COXII gene reveals for the first time the presence of an intron in this gene of a dicot plant. The intron structure appears very similar to that of monocot rice. However in sunflower, part of the intron, equivalent to the transposable element of rice gene, has not the same characteristics of mobility.

A 12kbp inversion characterizes the organization of a region flanking the CYB gene in a sunflower male sterile line. This inversion however does not affect the transcriptional pattern of CYB gene.

ACKNOWLEDGMENTS

We thank C. J. Leaver for the maize cytochrome oxidase subunit II and apocytochrome b clones. This work was supported in part by 40% M.P.I. funds.

REFERENCES

- Boer P.H., Mc Intosh J.E., Gray M.W., Bonen L., 1985, The wheat mitochondrial gene for apocytochrome b: absence of a procaryotic ribosome binding site, *Nucleic Acids Research* 13: 2281-2292.
- Bonen L., Boer P.H., Gray M.W., 1984, The wheat cytochrome oxidase subunit II gene has an intron insert and three radical amino acid changes relative to maize, *The EMBO Journal* 3: 2531-2536.
- Capaldi R.A., Malatesta F., Darley-Usmar V.M., 1983, Structure of cytochrome c oxidase, *Biochimica et Biophysica Acta* 726: 135-148.
- Ceci L.R., De Benedetto C., Siculella L., Perrotta C., Gallerani R., 1988, The discontinuous cytochrome oxidase subunit II gene of dicot plant *Helianthus annuus*, *Current Genetics*: submitted for publication.
- Dawson A.J., Jones V.P., Leaver C.J., 1984, The apocytochrome b gene in maize mitochondria does not contain introns and is preceded by a potential ribosome binding site, *The EMBO Journal* 3: 2107-2113.
- Fox T.D., Leaver C.J., 1981, The *Zea Mays* mitochondrial gene coding cytochrome oxidase subunit II has an intervening sequence and does not contain TGA codons, *Cell* 26: 315-323.
- Gallerani R., De Benedetto C., Siculella L., Perrotta C., Ceci L.R., 1987, A novel feature of mitochondrial COXII gene in higher plants: discontinuity in the gene of sunflower, in: *Quaglia-*

- riello E., Bernardi G., Ullmann A. Eds. From enzyme adaptation to natural philosophy: heritage from Jacques Monod. Elsevier Science Publishers: 149-158.
- Grabau E.A., 1987 Cytochrome oxidase subunit II gene is adjacent to an initiator methionine tRNA gene in soybean mitochondrial DNA, *Current Genetics* 11: 287-293.
- Hiesel R., Brennicke A., 1983 Cytochrome oxidase subunit II gene in mitochondria of *Oenothera* has no intron, *The EMBO Journal* 2: 2173-2178.
- Holm L., Saraste M., Wikstrom M., 1987, Structural models of the redox centres in cytochrome oxidase, *The EMBO Journal* 6: 2819-2823.
- Kao T., Moon E., Wu R., 1984, Cytochrome oxidase subunit II gene of rice has an insertion sequence within the intron, *Nucleic Acids Research* 12: 7305-7315.
- Korneluk R.G., Quan F., Gravel R.A., 1985, Rapid and reliable dideoxy sequencing of double stranded DNA, *Gene* 40: 317-323.
- Kyte J., Doolittle R.F., 1982, A simple method for displaying the hydropathic character of a protein, *Journal of Molecular Biology* 157: 105-132.
- Maxam A.M., Gilbert W., 1977, A new method for sequencing DNA, *Proceedings of the National Academy of Sciences* 74: 560-564.
- Moon E., Kao T., Wu R., 1985, Pea cytochrome oxidase subunit II gene has no intron and generates two mRNA transcripts with different 5'-termini, *Nucleic Acids Research* 13: 3195-3212.
- Pacoda D., Treglia A.S., Siculella L., Perrotta C., Gallerani R., 1987, Localization and partial sequencing analysis of mitochondrial gene for apocytochrome b in sunflower, in: Papa S., Chance B., Ernster L., Jaz J. Eds. *Cytochrome systems*, Plenum Publishing Co.: 133-134.
- Perrotta C., Siculella L., Ceci L.R., De Benedetto C., De Paolis A., Gallerani R., 1986, Genes for rRNA and proteins on restricted and cloned fragments of sunflower mitochondrial DNA, *Italian Journal of Biochemistry* 35: 177A-179A.
- Saraste M., 1984, Location of haem-binding sites in the mitochondrial cytochrome b, *FEBS letters* 166: 367-372.
- Schuster W., Brennicke A., 1985, TGA-termination codon in the apocytochrome b gene from *Oenothera* mitochondria, *Current Genetics* 9: 157-163.
- Siculella L., Palmer J.D., 1988, Physical and gene organization of mitochondrial DNA in fertile and male sterile sunflower. CMS-associated alterations in structure and transcription of the *atpA* gene, *Nucleic Acids Research*: submitted for publication.
- Treglia A.S., Pacoda D., Siculella L., Ceci L.R., Gallerani R., 1988, GUG as initiator codon in the sunflower mitochondrial gene of apocytochrome b, *Current Genetics* submitted for publication.