Different organization and altered transcription of the mitochondrial *atp6* gene in the male-sterile cytoplasm of rapeseed (*Brassica napus* L.)

Hirokazu Handa and Kousuke Nakajima

Department of Cell Biology, National Institute of Agrobiological Resources, Tsukuba Science City, 305, Japan

Received July 25/October 2, 1991

Summary. The F_o-ATPase subunit 6 gene (atp6) of rapeseed mitochondria has been isolated from both pol malesterile and normal (fertile) cytoplasms in order to determine whether the rearrangements around the atp6 locus in *pol* male-sterile cytoplasm play a role in cytoplasmic male-sterility (cms). The pol cms and normal atp6 genes are identical and encode a 261-amino acid polypeptide. As a result of extensive rearrangement, a novel reading frame (pol-urf) was generated upstream of the atp6 gene only in *pol* cms mitochondria, which encoded 105 amino acids and might be co-transcribed with atp6. A 5'-portion of *pol-urf* shows sequence homology to the *Oenothera* ORFB gene associated with coxIII. A 5'-flanking region of the pol-urf also shows homology to that of ORF105 in Ogura cms radish mitochondria. These DNA rearrangements which give rise to *pol-urf* in the vicinity of the *atp6* locus may be responsible for cms in rapeseed.

Key words: Cytoplasmic male-sterility $-F_o$ -ATPase subunit 6 – Mitochondrial DNA – Rapeseed

Introduction

Cytoplasmic male sterility (cms) is a maternally inherited trait which prevents the production of functional pollen. Male-sterile cytoplasm has long been of interest for its usefulness in the production of hybrid seeds. In many crops, the cms phenotype is expressed in alloplasmic lines arising from interspecific or intergeneric crosses. It is thought to be due to incompatibility between the nucleus and the cytoplasm. Several lines of evidence suggest that the cms determinants reside on the mitochondrial (mt) genome (reviewed in Hanson and Conde 1985; Lonsdale 1987).

There are two types of cms cytoplasm, *nap* cms and *pol* cms, in rapeseed (*Brassica napus* L.). The first observations of cms were by Shiga and Baba (1971, 1973) and Thompson (1972), who detected male-sterile plants in F_2

progeny of intraspecific crosses (*nap* cms cytoplasm). Fu (1981) later observed that the cytoplasm of a Polish variety "Polima" induced male sterility (*pol* cms cytoplasm).

The mt genome of *Brassica* is the smallest and the best-characterized among higher plant mt genomes (Palmer 1988). A complete restriction map showed the genome size of rapeseed mitochondria to be 221 kb (Palmer and Herbon 1988). We previously classified the mt genome of rapeseed into two major types, each of which was further divided into three subtypes based on length differences in the restriction fragments of mtDNAs. The mt genome of *nap* cms cytoplasm was classified as type I, whereas the *pol* cms cytoplasm carried the type II mt genome (Handa et al. 1990).

To determine the molecular basis of cms in rapeseed, we examined the physical organization of several mitochondrial genes from *pol* cms cytoplasm, type II b mt genome, and its normal counterpart, type II a. Our findings implicate DNA rearrangements located around the mitochondrial *atp6* locus of *pol* cms cytoplasm as candidates for causing the male-sterile phenotype. We analyzed the *atp6* locus from both *pol* cms and normal cytoplasms to further examine its role in *pol* cms. Nucleotide sequences of the two *atp6* locus was found to be associated with a novel 105-amino acid ORF (*pol-urf*) which may be co-transcribed with *atp6*. The possible role of this alteration in *pol* cms cytoplasm is discussed.

Materials and methods

Plant materials. B. napus L. cv. Isuzu-natane was the source of the normal (fertile) cytoplasm. A cytoplasmic male-sterile line, which has the "Polima" cms cytoplasm under the nuclear background of Isuzu-natane, was also used.

Isolation of nucleic acids. Mitochondria were isolated from 8 weekold plants. Mitochondrial DNA (mtDNA) was extracted as described previously (Handa et al. 1990). Mitochondrial RNA (mtRNA) was isolated by the procedure of Stern and Newton (1986), except that no sucrose step gradient was used for the purification of mitochondria.

Southern and Northern blot analyses. Restricted mtDNAs separated by agarose gel electrophoresis were transferred to a nylon membrane filter (Hybond N, Amersham, UK) by a modification of Southern (1980). Isolated mtRNA (10 µg/lane) was electrophoresed in a 1.2% agarose gel containing 0.6 M formaldehyde, 20 mM MOPS (pH 7.0), 5 mM sodium acetate, 1 mM EDTA, and blotted to Hybond N filters with $20 \times SSC$. Eight mitochondrial-specific genes [pea atpA, coxII (Morikami and Nakamura 1987, and unpublished data), rice atp6, cob, coxI, nad3/rps12 (Kadowaki et al. 1989, 1990, and unpublished data), tomato atp9 (Kazama et al. 1990), and Aegilops columnaris coxIII (Ikeda et al. 1990)] were labeled by the random primer method incorporating digoxigenin-labeled dUTP (Feinberg and Vogelstein 1983) and were used as hybridization probes. Hybridizations were carried out at 68 °C (for DNA blots without formamide), and at 42°C (for RNA blots with 50% formamide) for 16 h after prehybridization. Following hybridization, the hybridized bands were detected by an enzyme-linked immunoassay using an anti-digoxigenin-alkaline phosphatase conjugate and a subsequent enzyme-catalyzed color reaction with Xphosphate and NBT, or else the chemiluminescent reaction with AMPPD, following the manufacturer's instructions (Boehringer Mannheim, Germany; Tropix, Inc., Mass., USA). Strand-specific RNA probes were synthesized from the 1.35 kb HindIII-EcoRI fragment cloned into pBluescriptII vector (Stratagene, CA, USA) (Fig. 2, probes F and F') using T3 or T7 RNA polymerase.

Sequence determination and analysis. The dideoxynucleotide chaintermination method of Sanger et al. (1977) was used for DNA sequencing. Restriction fragments were subcloned into pBluescriptII vectors. This sequencing strategy was supplemented with nested deletion methods using exonuclease III and subsequent mung bean nuclease digestions, following the manufacturer's procedures (Takara Shuzo, Japan). Sequence data were compiled and analyzed with the aid of GENETYX computer software programs.

Results and discussion

Southern blot analysis of restricted mtDNA

The mtDNAs from pol cms and normal cytoplasms were compared by restriction digests using EcoRI, HindIII, and KpnI (Fig. 1A). The mtDNAs from both cytoplasms showed different fragment patterns as reported previously (Handa et al. 1990). Southern blot analysis was used to determine the gene organization of mtDNA from pol cms and normal cytoplasms by probing with eight mitochondrial genes, atpA, atp6, atp9, cob, coxI, coxII, coxIII, and nad3/rps12. The mtDNA from pol cms and normal cytoplasms had the same hybridization signals when probes for the atpA, atp9, cob, coxI, coxII, coxIII, and nad3/ rps12 genes were used (data not shown). Only when the probe for the atp6 gene was used were different patterns observed (Fig. 1B). Restriction fragments encoding the atp6 gene of pol cms and of the normal mt genome were determined to be 4.9 kb and 6.5 kb, respectively, after restriction with EcoRI. These results suggested a different gene organization around the atp6 locus in the two mtDNAs. Recently, Witt et al. (1991) also detected a different gene organization around the atp6 loci in pol cms and normal mitochondrial genomes by Southern hybridization. Their results are coincident with ours, although their estimated fragment sizes are slightly larger.



Fig. 1A, B. Southern blot analysis of mtDNAs from *pol* cms (P) and normal (N) cytoplasms. A, agarose gel profile; B, hybridization patterns using the rice *atp6* gene as a probe



Fig. 2. Physical maps of the *atp6* loci from *pol* cms (top) and normal mitochondria (bottom). Clone pBNP3 was isolated from *pol* cms cytplasm, and clone pBNI54 from normal cytoplasm. Shaded and open boxes, respectively, represent the *atp6* coding region and *polurf. Horizontal arrows with a solid head* indicate the conserved sequences between the two clones. Arrow heads indicate the transcriptional direction of the *atp6* and *pol-urf* genes. Horizontal lines indicate the position of the probes (A, B, C, D, and E) used in this study. Horizontal arrows with open heads indicate the position of the probes F and F' and their directions. Restriction sites are indicated as follows: \bullet , *Eco*RI; \vartriangle , *Eco*RV; \bigstar , *Kpn*I; \square , *Pst*I; \bigcirc , *Sal*I; \blacksquare , *Xho*I

Cloning of rapeseed atp6 genes from pol cms and normal cytoplasms

The heterologous *atp6* gene from rice (Kadowaki et al. 1990) was used as a probe to isolate the rapeseed *atp6* gene. An 11 kb *Sal*I fragment of clone pBNP3 was isolated from *pol* cms, and a 6.5 kb *Eco*RI fragment of clone pBN154 was obtained from normal cytoplasm. Detailed physical mapping and further Southern analysis of plasmids pBNP3 and pBN154 were performed (Fig. 2) to better localize the difference between the two atp6 loci. The sizes of the isolated fragments were coincident with those expected based on the Southern blot analysis (Fig. 1 B) and the two clones shared a common 2.2 kb fragment including the *atp6* coding sequence (Fig. 2). However, the two fragments diverged upstream of the *atp6* coding sequence, indicating that a rearrangement existed in the

TTGAATGGG 1358

Fig. 3. Nucleotide and deduced amino-acid sequences of the *atp6* loci from the *pol* cms and the normal mt genome. The *pol* cms sequence (top row) is aligned with that from normal cytoplasm (bottom row). Where the two are identical, only the normal sequence is shown. The first base of the *atp6* initiation codon ATG is numbered as +1. Amino acids are shown above (*pol-urf*) and below (*atp6*) the nucleotide sequence and are numbered relative to the translation initiation codon. The homologous region to normal radish *atp6* locus is represented by *arrows* (below the sequence for normal cytoplasm). The truncated *trnfM* gene is *underlined with a dashed line*. The sequences nearly identical to the conserved plant mitochondrial promoter are *boxed*

5'-flanking region of the *pol* cms *atp6* gene. To further characterize the differences, both *atp6* genes and their flanking regions were sequenced.

Sequence analysis of the normal atp6 gene

The nucleotide sequence of the *atp6* locus on pBNI54 from normal cytoplasm was determined (Fig. 3). This sequence contains a 783 bp open reading frame corresponding to a 261-amino acid polypeptide that was identified as *atp6* by DNA sequence homology. The deduced amino acid sequence of the rapeseed ATP6 protein shows 100% homology to the ATP6 protein of radish (Makaroff et al. 1989). The deduced sequence also exhibits 81-86% identity with the conserved regions of the ATP6 proteins of maize, tobacco, *Oenothera*, and rice (Dewey et al. 1985; Bland et al. 1987; Schuster and Brennicke 1987; Kadowaki et al. 1990).

A comparison of the flanking regions between *Brassi*ca family members revealed a number of differences. The 5'-flanking region of normal radish contains a gene for tRNA^{fMET} (trnfM) located on the same strand and situated between -232 bp and -159 bp upstream of the atp6 gene (Makaroff et al. 1989). A 98.3% sequence homology between the 5'-flanking region of normal rapeseed and radish atp6 genes was observed from -181 bp to the ATG initiation codon (Fig. 3), but upstream from -182 the two genes diverge extensively. Therefore, in normal rapeseed, the trnfM was truncated, leaving only 23 bp of the 3' end.

pol: GATCGCTACGCTGCCGGGCTGCCCCGCGGATCAAACTAT -1052

pol: CAATCTCATAAGAGAAGAAATCTCTATGCCCCCTGTTCTTGGTTTTCTCCCCATGCTTTTG -992

| <u>pol</u> : | TTG | GTC | AAC | AA | CCA | AC | CA | CAA | CTI | ΓTC | TA | TAC | ΠT | CTT n | CA | CT/ mal | CT L: | CC 1 GT 1 | TAG TAI - 8 | A GG GCG R 2 | CTT GTA | GAC ATG | GGA GCA | -932 |
|------------------------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|---------------|--------------------|------------|-----------------|------------|-----------------|-----------------|-----------------|-------------------|--------------------|-----------------|----------------------|-----------------|------------|
| | | | | | | | | | | | | | | | | | <u>p0</u> | <u>1-</u> ı | <u>irf</u> | : M | Р | Q | L | 4 |
| <u>pol</u> : ormal: | GTG TAA | AAG GCC | CTG AAA | TC: GC | rge Fac | GAG GAT | GG/ CA(| AT CCG | CA) GA(| CTT GTT | TG TA | TT(GG(| HAA CAA | ATC ATG | GG' | TT/ FA(| AT CGA | CT / GA (| AT SAA | CAT AAA | GCC GGT | TCA. TTT | ACT GAA | -872 |
| | D GGA | K Taa | F Att | CAC | r CTT | Y AT | F TT1 | S TTC |) Aca | l Lat | F TC | F TTC | W TG | L GTT | AT | C GCC | l TT | F TTC | F TTC | F CTT | T Tac | F TTT: | Y CTA | 24 |
| | ATG | CAT | CTT | CCO | GAA | AC | CAI | GAT | AA7 | rca | AC | CG/ | AG | GAG | AT: | rco | CC | TGC | CA A. | AAA | GGA | GAT | GGC | -812 |
| | I TAT TCT. | F TTT AGT | I CAT TTC |) AT(AT/ | C GCA ACT | N AT 'AA' | D GAI TCI | G IGG ITC | I Aga Gat |) ATG CT | G Ga TC | V GTA GA C | L CT | G TGG TCC | GA1 CCC | I FCA CAC | S .GC :GG | R Aga Agc | I AT GG | L CCT FAA | K AAA GGC | L ACT. ATC | W Atg CCA | 44 -752 |
| | N | Q | L | I | | S | Н | R | 0 | ; | K | T | L | L | 5 | 5 | K | G | R | L | G | K | N | 64 |
| | GAA AGG | CCA AGC | ACT GAA | GC3 GC0 | GGC | CA TC | CA(AA/ | CG NGG | GGG ATI | IGA IGG | ÅG. CA. | ACC ATC | CT AA | CCT GCT | GAC CAC | GCA GTC | AG TG | GG A GT C | AGC TG/ | GCT GGG | TGC. GGA | AAAA TGGI | AAA GCT | -692 |
| | R TCG: | S TAG | S TTC | I Aga |) \TT | S CA | S AG1 | R CG | F GTT | CG | E Agi | V GTA | S TC | A AGC | I GT 1 | : 166 | A CC | A GC C | H Cai | Y TA | F TT T | I FAT(| I CAT | 84 |
| | AGG | CTA | CTC | TCC | CC | AT. | ACA | TA | GGA | AG | AG | GGG | GA | ∖GT | CAA | GT | AC | CTA | GAI | CC. | AAG | ATC/ | ACG | -632 |
| | F TTTI | V CGTI | V GGT | F CCC |) CA A | K AA' | L TTO | G GG | Q ACA | .67 | F TT' | S PCT | T AC | L የተኛ | Y A'T A | (4.17.4 | Ĩ TA. | I ATT | F TT1 | ץ ידדי | ¥ ۲GT | C PTG1 | ₩ היירק | 104 |
| | ATA' | rcc -56 | GCT 5 | TTO | GT | 'GG' | TAA | TA: | rcc | GG | CA: | ΓAA | AA | GAA | AGO | CAC | CA | AAG | CA/ | AC. | AAAI | GTC | GAT | -572 |
| | GGGI | r ITG. | AAA | TGG | GG | GG | TAT | TAC | GGA | AA | rg, | AAA | TTI | GT | CAT | TT | CG | GCG | TCO | GA | CCA | GATO | GC | 100 |
| | GTTC | CCGI | 6 1 G. GCC: | | AG | AC: | CTG | GA1 | ICT | CA | | GAG | CGG | 300 100 | GA1 GCC | TCT | AG: TG(| CAT | tto Cti | TC: TT: | CTA) GTAC | CTCT | GA GA | -512 |
| | GGTU TGTI | GAG | GAG | GCT | GA | GAG ACT | JAG ICT | CAP | IGA | AT. | AGA CGC | GA | GC# AA7 | IGA. | AGC Tga | CTC | GC (TG: | CCG FAC | C G G G C G | IGC I | AAA(FCT) | GAC <i>i</i> Fago | IAG ICG | -452 |
| | GGAI | GAI | TTC: | rrc | TA. | AG/ | ATA | CG/ | AC | TT | 300 | iGA | AGO | SAA' | TGA | AA | AGO | JTA' | TTC | GAO | CTAF | GAC | GG | -392 |
| | GGAC | ICA(CGG(| GAG GTA: | GAT FTC | CCI | GAA | CAA Aga | AAA TGC | ICT. | AGA GA(| GC CCT | GT AA | GAO GAO | CG(AT) | CGA AGA | .TA .GA | TC (CG / | GT G A G A. | C GG A CG | CG1 AGA | ICAA ACTA | IGC C | TC AA | -332 |
| | CTGC ATCA | GAT/ GAT | ATA IAT: | ATG FGG | AA. AG | ATC AGC | GGG GAA | AGO AGA | AC AG | AG/ CAC | AA TO | ICC ICA | TTC AGT | AG(AG/ | GAG Agt | CA AG | CT1 AGC | CA G | GGA CTT | TC1 GG1 | TTTC AGC | CGGC | AC CA | -272 |
| | CTTG AGGA | ATC GCC | GT(GAA/ | CTG | CG/ GC | AGA I'GC | ATT CTT | CGC GAC | GA. | AG/ ACC | GC GC | CA. | AGG Gaa | TG/ GCC | AAC CAA | GA CT | GA 1 CC 0 | CC STT | ETG CCT | AT (CC A | CT CT AGCT | 'TTC 'TCG | GC CG | -212 |
| | GA TAGA | GA/ | GC1 | TAG | GT' | FCC | TC | стт | 'GA' | TGI | -1 CC | 81 T <u>A</u> T | GGI | TCA | 188 | TC) | CTA | TC | r <u>cc</u> | GCA | <u>Ç</u> ta | AGT | AA | -152 |
| 1 | GGGT | TTC | ATT | CT | GC/ | \TC | AC. | ГСТ | CC | CCG | TC | GTI | CT | CGA | CC. | TC | GC A | AGO | FTT' | TTT | GAA | GCG | GC | -92 |
| (| CGAA | GCG | GGA | AG. | r G A | ACA | AT) | ACC | GC1 | ITT | TC | TTC | CAG | CAC | AT | TTI | ľGG | A ATO | L GAT' | TT G | AGC | GA A. | AA | -32 |
| (| CGGA | GTA | ĈĂĂ | AG | FT(| CAG | CC | ITT | AA(| GGA a | GG tp | CT/ 6: | ATG M | AAT N | CA) Q | AA1 1 | ľ á G | GGC G | TG L | GT G V | GCG A | CAG Q | ГС S | 29 10 |
| (| CCCA P | CTT L | GAC D | CAI Q | ATT F | TG | AG/ E | ATT I | GT (V | CCC P | AT | TG/ L | TT I | CCT P | AT(M | GA/ N | ATA I | TCE I | GA. G | AAC N | TTC F | TAT: Y | FT F | 89 30 |
| (| CTCA S | TTC F | ACA T | AA1 N | ГСС F | CAT | CT' S | rtg L | TTC F | CAT M | GC | TGC L | TA L | ACT T | CTO | GA (| TT S | TTI F | TC) F | CTA L | CTT L | CT G L | AT I | 149 50 |
| 1 | FCAT H | TTT F | ATT I | ACT T | raa K | AA [| AG(K | GGA G | GG/ G | AGG G | AA | ACJ N | ITA L | GTC V | CC/ P | AAA N | AT G | CTT A | 'GG W | CAA Q | TCC S | TTG L | GT V | 209 70 |
| ļ | AGAG E | CTT L | CTT L | TAT Y | rga D | TT) | TCO F | GTG V | CTO L | GA A N | СС | TGG L | TA. V | AAG K | GA/ E | ACA G | AA I | TAG I | GT (G | GGT G | CTT L | TCC(S | G G | 269 90 |
| ł | AAT N | GTG V | AAA K | CA/ Q | AT M | GT I | TTI F | TC F | CCI P | ttg C | CA | TCI I | TG L | GTC V | AC) T | CT1 F | TC | T T 7 L | TT' F | rtg L | TTA' L | TTT! F | rg C | 329 110 |
| 1 | FAAT N | CTT L | CAG Q | GG1 G | TA'I M | 'GA' | TAC I | CT P | TAT Y | CAG S | CT | ГСA F | CA T | GTG V | AC/ T | AA G S | TC. | ATT H | TT (F | CTC L | ATT. I | ACT: T | TT L | 389 130 |
| 0 | GCT(A | CTC L | TCA •S | TTI F | rtc S | TA | TTI I | TT. F | ATT I | GG G | CA | FTA I | CT. T | ATA I | GTO V | GGG G | AT | FTC F | AA/ Q | AGA R | CAT H | GGGG G | CT L | 449 150 |
| 1 | ICAT: H | ITT F | TTC F | AGC S | CTT F | TT | TAT L | TA L | CCC P | GC. A | AG | GAG G | TC) V | CCA P | CTE L | GC C P | GT | F A G L | CA (A | CT P | TTT F | ITA(L | ST V | 509 170 |
| A | LCTC | CTT L | GAG E | CTA L | AT I | TT | CTI S | AT: Y | rg1 C | TT F | rci I | GCG R | CA' A | FTA L | AGC S | CTT L | AG | GAA 3 | TAC I | CGT R | ΓΤΑ' L | FTTC F | EC A | 569 190 |
| τ | AAT/ N | ATG M | ATG M | GCC A | GG G | TC/ | ATA H | GT: S | TTA L | GT. V | 4.A.A 1 | \GA { | TT: I | lTA. | AGT S | GG G | GT. | rcg P | CTI A | 'GG/ W | ACTA T | ATGC M | T L | 629 210 |
| A | TGT/ C | ATG. M | AAT N | GAG E | AT I | TT: | FCT F | AT: Y | F F | 'AT. I | AGI (| GGG G | CT (A | L L | GGT G | CC P | TT: 1 | FAT L | TTA F | TAI I | GTTI V | CT T (L | C A | 689 230 |

Located 152 bp 5' of the *atp6* initiation ATG site was a 7-bp sequence, TAAGTAA (Fig. 3), which is nearly identical to part of the consensus sequence of the putative plant mitochondrial promoter (Young et al. 1986).

In the 3'-flanking region, the 104 bps (from position +784 to +887) after the termination codon are highly conserved (98% identical) between the normal rapeseed and normal radish atp6 genes. Downstream from + 888 there is no further sequence homology. This lack of homology differed from what was observed in comparison with the *atp6* gene from Ogura cms radish. Compared to the Ogura cms atp6 gene, a 575 bp sequence beginning from the termination codon in normal rapeseed atp6 (from position +784 to +1358) has 99.6% homology. In the normal radish mt genome, the homologous sequence is located on another part of the genome 60 kb from the atp6 locus (Makaroff et al. 1989). These results indicate that extensive rearrangements have occurred at the atp6 loci of Brassica. Unexpectedly, the mt genome of Ogura cms radish might be more closely related to that of normal rapeseed than that of normal radish, because it is more likely that the 575 bp sequence shared by Ogura cms radish and rapeseed in the 3'-flanking region of the atp6 locus was present in the common ancestor of the genus and was subsequently rearranged in normal radish, rather than the same rearrangements occurring independently in rapeseed and Ogura cms radish. These results also support the findings that plant mtDNA evolves rapidly in structure, but slowly in sequence, as described previously by Palmer and Herbon (1988).

Sequence analysis of the pol cms atp6 gene and an associated 105-amino acid ORF

Nucleotide sequence analysis of pBNP3 revealed a complete *atp6* gene having an ORF of 783 nucleotides which could encode a polypeptide of 261 amino acids (Fig. 3). The *atp6* coding region (783 bp), 3'-flanking sequence (575 bp), and part of the 5'-flanking sequence (211 bp) are highly conserved (99.7% identical) between the *pol* cms and normal *atp6* genes. These results indicated that the *pol* cms *atp6* gene is intact and probably functional.

The differences in the *atp6* loci noted in the physical mapping of the two genomic fragments were detected in the upstream region beginning from nucleotide -212, after which the sequences diverge completely. The region spanning from -882 to -565 in *pol* cms *atp6* harbors an open reading frame that has the capacity to encode an 105-amino acid, 12270-Da, polypeptide. This ORF was named pol-urf (unidentified reading frame of pol cms mitochondria). This novel ORF showed a partial sequence homology to the ORFB gene associated with the Oenothera mitochondrial coxIII gene (Hiesel et al. 1987), beginning in *pol-urf* at position -118 in the 5'-flanking region and ending at position +175 bp in the open reading frame. However, the following downstream coding sequence showed no homology with the Oenothera ORFB gene, or any other reported sequence. Neither ORFB nor *pol-urf* show a preference for U in the third position of their codons, a common feature of plant mitochondrial genes; U is found in 33% of all codons in the third position of the *pol-urf* gene (40.5% in *atp6*). The absence of a typical plant mitochondrial codon usage suggested that ORFB and *pol-urf* are not originally mitochondrial genes despite their mitochondrial location.

The 177 bp sequence block adjacent to the 5'-end of the *pol-urf* also has sequence homology (91%) with the 5'-flanking region of the radish ORF105 sequence which is associated with the atp6 gene in Ogura cms radish (Makaroff et al. 1989). The 14 bp sequence AATCT-CATAAGAGA located at the 5'-end of this sequence block (from -1050 to -1037, Fig. 4) is nearly identical to the consensus sequence of the putative plant mitochondrial promoter (Young et al. 1986). If this 14 bp sequence works as a promoter, then ORF105 and *pol-urf* may both be transcribed in the same manner, for example specifically in male-sterile cytoplasm. Much (118 bp) of this sequence block is homologous to the 5'-flanking region of Oenothera ORFB (described above). Further work is necessary to determine what role this 177 bp sequence block plays in plant mtRNA structure and/or function. The coding sequence showed no homology with the radish ORF105 gene (Fig. 4), and the structure of a truncated gene fused to an unidentified sequence suggested that *pol-urf* might have been generated through a series of duplication and rearrangement events.

To determine where the *pol-urf* sequence resides on the normal mitochondrial genome, Southern blot analysis was performed. Two DNA fragments containing the 5'half or 3'-half of the *pol-urf* sequence (probes A, a 219 bp HincII-EcoRI fragment, and B, a 415 bp EcoRI-EcoRV fragment, respectively, Fig. 2) were used to probe Southern blots containing EcoRI, HindIII, and KpnI digests of pol cms and normal mtDNA (Fig. 5). The 3'-half, probe B, hybridized only to the same fragments containing the *atp6* gene from both *pol* cms and normal mtDNA, indicating that this sequence is not repeated in the mt genome. However, hybridization experiments with the 5'-half (probe A) showed more complicated patterns for both pol cms and normal mtDNA. The homologous sequence to probe A is present about three times in *pol* cms mtDNA and twice in normal rapeseed mtDNA. Because plant mitochondrial genomes are characterized by frequent homologous recombination events (Lonsdale et al. 1984; Palmer and Shields 1984), this dispersed sequence may be involved in such events, and may explain how the pol-urf gene originated.

Transcriptional analysis of the atp6 locus

To determine whether, and to what levels, the *atp6* and *pol-urf* genes are transcribed, total mtRNAs extracted from green leaves were probed with two different DNA fragments from the *pol* cms *atp6* locus (probes B and C, Fig. 6). Probe C, a 897 bp EcoRV/HindIII fragment, contains the 5'-coding region of the *atp6* gene (546 bp) and 351 bp of the 5'-flanking sequence. Probe B is as described above.

Altered transcriptional patterns were observed when the *atp6* gene was used as a probe between *pol* cms and

| AA) | TC | T | TA | TG | с.и | СС | CCI | IGT | TC | FT (| GGI | TT | TC: | rce | CCA | TGO | С.Т | TTI | [G] | TTG | GTI | CAA | CAA | CCAA | 9' |
|----------|----------|-----|----------|----------|-------------|----------|----------|----------|----------|-----------|----------|----------|------------|------------------------|------------|----------|---------------|--------------|-----------|-----------|----------|---------------|------------|----------|-----|
| G | -Ge | -1 | '-G | | -A' | ľ | AA/ | l−G | | ' | ΓT- | -CC | CA- | -G | -A- | | -A- | -00 |] | | | | | C | |
| | | | | | | | | | | | | | | | | | ••• | | | | | | | | |
| CCA | IC/ | A | TT | TC | FA' | r. | . A(| STT | CT: | rc, | AC1 | CAC | TC | CTA | AGA | GGC | CTT | GAO | CGC | GAG | TG | AAG | CT G. | CTG | ~ 9 |
| AG0 | | Т- | A | _ | A | -A. | AG- | -A | | | | | | <u>-</u> | -0- | | | .т- | # -C1 | ₹ Г-А | - A · | | | | |
| | | | | | | | | | | | | | | | -88 | 2 | | | | | | | | | |
| GÅ | 360 | AA | TC | AT' | TT: | ſG | TT(| AA | AT(| CAI | ATI | 4A] ת | TC: | ['A <i>i</i> - 11 1 | ATC ef: | ATC M | '33 ק | TCA r | 1A () | TG | GA' D | E A A K | ATT(| CACT | -8 |
| | | | -T | | | ' | | -T | C | | | | -~. | . G | GA | | · | | | | | | '] | 1 | |
| | | | | | | | | | | | | | (| DRI | B: | * | * | * | : | * | * | * | * | * | |
| | | • • | ••• | . A | ~ | | | | | | | | (0RI |) 711 | A)5: | * | R | A-U F | i−8 } | τ Γ | TT- F | -Մ- ն | * | D D | |
| | | | | | | | | | | | | | • • • • | | | | | | | î | î | | | - | |
| TAT | TT' | Τĩ | CA | CA | AT? | ['C' | rto | TG | GTI | ľA: | rg(| CT | TT1 | C. | TC | TTI | 'QA' | TT I | CJ. | 'AT | AT' | TT 1 | CATA | TGC | -7 |
| | r | CA | ة | | | | 1 | | ر) | ; ; | ι. | | | (| r] | r | 1 | P | | 1 | 1 | r -CC- | | ы. | |
| * | * | | T | * | 2 | k | * | * | ŝ | 3 | * | * | 4 | * | L | * | * | * | t | * | * | P | * | * | |
| G∽/ | AC. T | -A | .GC | -T' | <i>[</i> ۸] | ١T٠ ، | -CA | A-' | T-(| C/ | ۲-۱ م | 'GA' | -A(| CA- | -AT | -C- | Gi | ۹۹C ۳ | TG | CG | -G- | -CA. | ATCO | ACG | |
| ъ | 1 | | ~ | Ľ | 1 | ` | 0 | 0 | | , | U | D | | | 1 | 0 | п | - | | л | 0 | ય | 5 | 1 | |
| AAJ | 'GA | TG | GA | GA' | rgo | GA | GTA | CT | TGO | GA | AT(| AG | CAC | GA/ | ١T | CTA | AA | ACT | ΆT | GG | AA(| CCA | ACTO | CTT | -73 |
| N | D | | G | D | (| 3 | V | L | (| ì | Ι | S | F | } | I | L | , K | ľ | | ¥ | N | Q | L | L | |
| * | * | | * | * | 1 | ķ | * | * | × | k | * | * | 4 | ŧ | * | * | , * | * | | R | * | * | * | * | |
| -00 | AC | -A | TT | A-(| C-4 | ۲ï | r a 1 | AG | -T] | ! | AA | TC. | A | -T1 | 1CC | AGΊ | 'G-1 | r-c | | AT | TCI | r ag | -ACT | GCG | |
| Т | T | | I | N | I |) | Y | S | I | | K | * -7 | 3 N R N | 3 | S | S | D | P |) | Y | S | R | T | A | |
| TCA | CA | СС | GG | GG(| GA/ | lG/ | ACC | CT | CCI | 'G/ | ١GC | AA | GGC | ia A | GG | CTT | GG | \ A A | AA | AT | CGI | 'AG' | FTCA | GAT | -67 |
| S | H | | R | G | } | (| T | Ľ | I | | S | K | 0 | 1 | R | L | G | K | | N | R | S | S | D | |
| * | * | | * | * |) * | ; | -A N | A T | * | i UU : | ; R | * | я Т | ili I | CA. P | AAC N | A-1 S | ITT L | GG | i-A' R | jA/ R | ICTI L | U∼Tt I. | R | |
| AGI | | AT | CC | AC' | [| C | rt'i | TC | TAC | STO | 3-1 | C | AC- | T | | GAA | TAC | C-G | GA | TG | - T- | T | GCT- | -TG | |
| * | Q | | S | Т |) | l | F | S | 5 | 5 | G | Q | ł | 2 | Р | E | Y | R | | M | ľ | M | L | V | |
| TCA | AG | TC | GG | TTO | CGA | G | GTA | TC | AGC | GI | TG | GC | CGC | :00 | AT | ŤAŦ | TT | TAT | ĊA | TT | rto | GT | GTC | CCA | -61 |
| S | S | | R | F | F | | V | 5 | A | | L | A | A | | H | Y | F | I | | I | F | V | V | Р | |
| AA- K | ئا G | -1 | TT. F | AG- S |) A- ד | .C- | -GI G | GT· V | -т- s | -C- | •AT V | AT(M | STA V | 1-1 | :C~. S | AG- S | <i> </i> L | -T۲ F | -G | iagi E | i−A V | CTC(S | λΑΊ); Ω | TGG W | |
| GTC | -A | G- | - -T | ссс | ;-1 | T | c | , AG(| GAT | -0 | T | ATI | GCT | AC | TG | GTC | AAC | G | TC | ĊG | -A | .C-(| CAG- | ATG | |
| ¥ | K | | R | P | ١ | 1 | L | R | M | | * | M | Ι | | V | V | K | R | | P | ¥ | Ľ | S | M | |
| AAA | TT | GG | GA | CAG | TT | T | ICT | AC | AT7 | AT | 'AT | AT) | AAT | TT | TT" | TTT | GTI | TG | TT | GG | - 361 | - 26 TG/ |) AAAT | GGG | -55 |
| K | L | | G | Q | F | | S | T | I | | Y | Ι | I | - | F | F | V | C | | W | G | ++- | + | | 54 |
| TGT | AA V | | CT | GT(| CGA | C- | TA | TTO | 366 7 | -A | -A | G(| β−A γ | GA | AA. | A-C | AC- | T | GA | TA: T | C-01 | 1 C | TTC F | A | |
| CTT | л А- | -C | н Т-1 | GT- | GG | A- | -T | TT | u AA | GA | а | CA- | n TC- | CA | л GG· | -c- | -AA | <u>م</u> -۲ | c- | A CO | | | GCG | A-T | |



Fig. 4. A Alignment of the nucleotide sequences around the pol-urf gene. The nucleotide sequence around the pol-urf gene (top) is shown aligned with the Oenothera ORFB (middle) and Ogura radish ORF105 (bottom) sequences. Nucleotides of pol-urf are numbered as in Fig. 3; identical bases are shown as a -. Gaps (indicated by \bullet) have been introduced only in the 5'-flanking region of the *pol-urf* gene in order to achieve the best alignment. The deduced amino acids are shown below each respective nucleotide sequence; identical amino acids are represented by an asterisk. The 14-nucleotide sequence nearly identical to the consensus sequence of the putative plant mitochondrial promoter is boxed. B A schematic drawing to indicate the alignment of nucleotide sequences around the *pol-urf* gene. Open reading frames are represented by a box; open boxes represent homologous regions between pol-urf and ORFB, other boxes represent non-homologous regions. Homologous 5'flanking regions are shown as a *filled bar*



Fig. 5A, B. Detection of homologies with the *pol-urf* gene in *pol* cms (P) and normal (N) mtDNA. The probes used for hybridizations to the filters shown in *panels A and B* were probe A and probe B (Fig. 2), respectively



Fig. 6A–D. Transcriptional patterns of the *atp6* and *pol-urf* genes. Transcriptional patterns observed when mtRNAs from *pol* cms (P) and normal (N) cytoplasm were probed with probes B and C (Fig. 2), in *panels A and B*, respectively. Filters were probed with strand-specific probe F', corresponding to the anti-sense RNA (*panel C*) and probe F, the sense RNA, which is the complement of the probe in panel C (*panel D*). *Panel D* was obtained by a five times longer exposure than *panel C*

normal cytoplasms (probe C, Fig. 6B). Transcriptional alterations of the *atp6* gene were also detected in the Ogura male-sterile radish mitochondria (Makaroff and Palmer 1988). In rapeseed, one highly abundant transcript [1100 nucleotides (nt)] was observed in mtRNA from normal cytoplasm, while two abundant transcripts (2100 and 1100 nt) were evident in *pol* cms cytoplasm. These results of Northern hybridizations parallel those of

Witt et al. (1991), although several minor transcripts were detected in their results in addition to the major transcripts. Witt et al. (1991) carried out the Northern analysis of *pol* cms and normal mtRNA using 12 mitochondrial gene probes, and showed that only hybridization to *atp6* led to different patterns between *pol* cms and normal lines. They also reported that the large major transcript (2100 nt in our data, 2000 nt in theirs) is influenced by the nuclear background, indicating the presence of restorer genes. When the 3'-portion of the *pol-urf* (probe B) was used as a probe, only one abundant transcript (2100 nt) was detected and only from *pol* cms cytoplasm (Fig. 6A).

Probe F', corresponding to the anti-sense RNA, revealed the same transcript patterns as did probe C (Fig. 6C). A 1100 nt transcript was detected in both *pol* cms and normal mtRNAs. A 2100 nt transcript was present in *pol* cms mtRNA, but not in normal mtRNA. When the sense RNA from same region was used as a probe in Northern hybridizations (probe F), no hybridization signals were observed for any of *pol* cms and normal mtRNAs even upon prolonged exposure (Fig. 6D).

No hybridization signals were detected when probing *pol* and normal mtRNA blots with a *XbaI/Sal*I fragment covering a further upstream sequence of *pol-urf* (Fig. 2, probe D) and a *Hin*dIII/*Sac*I fragment covering a further downstream sequence of *atp6* (Fig. 2, probe E) (data not shown). These results, and the presence of sequences similar to the putative promoter sequence of other plant mitochondrial genes, indicated that the abundant 1100 nt RNA represents the transcript of the *atp6* gene, while the 2100 nt transcript from *pol* cms mitochondria may be a co-transcribed RNA consisting of both *pol-urf* and *atp6* sequences.

The cms trait has been associated with a number of mitochondrial chimeric genes and pseudogenes, including maize *urf-13T* (Dewey et al. 1986), sorghum *coxI* (Bailey-Serres et al. 1986), petunia *Pcf* (Young and Hanson 1987), and rice *urf-rmc* (Kadowaki et al. 1990). The *pol-urf* reading frame also has the structure of a chimeric gene (Fig. 4). Moreover, it is associated with one of only three or four unique restriction fragments noted when comparing the *pol* cms mtDNA with that of normal rapeseed (Fig. 1 A), and its postulated transcript is altered by nuclear restorer genes (Witt et al. 1991).

Therefore, *pol-urf* is a candidate for contributing to the defect in male fertility. Its possible translation product and observed transcription are consistent with two postulated models for explaining how chimeric genes cause the cms trait. First, though no translation product for *pol-urf* has yet been identified, it would be expected to have an amino-terminus homologous to the Oenothera ORFB protein, since the first 53 amino acids of the two peptides are 91% homologous (Fig. 4). Because ORFB is well conserved among sunflower and several other plant mitochondrial genomes (Quagliariello et al. 1990), it would appear to play a vital role in mitochondrial function. The possible *pol-urf* protein, a fusion of a homologue to ORFB and an unknown peptide, might thus act as an antagonist to the functional ORFB in the mitochondria. Second, the presence of the 2100 nt transcript

may interfere with normal atp6 mRNA function either by limiting the availability of the 1100 nt atp6 mRNA by necessitating a processing step or interfering with normal transcription. Makaroff and Palmer (1988) also reported transcriptional differences for the atpA locus between sterile and restored Ogura radish. But their further work demonstrated that transcriptional alteration of atpA was not associated with sterile or restored states (Makaroff et al. 1990). Therefore, additional work is required to better characterize the structure of the 2100 nt transcript and the relationship between the cms trait and the presence of this transcript. These and other studies of *pol-urf* could help in understanding the difference betweeen *pol* cms and normal cytoplasm, and so provide one of clues necessary to clarify the mechanism of cms.

Acknowledgements. We are grateful to Dr. D.B. Wing and Dr. S. Kikuchi, NIAR, for critical reading of the manuscript and many helpful suggestions. We also thank Dr. K. Kadowaki, NIAR, Dr. A. Morikami, Nagoya University, and Dr. TM. Ikeda, Kyoto University, for providing the mitochondrial gene probes. The nucleotide sequence data reported here will appear in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under the following accession numbers: X58276, Brassica napus L. mitochondrial atp6 gene (clone pBNI54); X58277, B. napus mitochondrial pol-urf gene (clone pBNP3).

References

- Bailey-Serres J, Hanson DK, Fox TD, Leaver CJ (1986) Cell 47:567-576
- Bland MM, Levings CS III, Matzinger DF (1987) Curr Genet 12:475-481
- Dewey RE, Levings CS III, Timothy DH (1985) Plant Physiol 79:914-919
- Dewey RE, Levings CS III, Timothy DH (1986) Cell 44:439-449
- Feinberg AP, Vogelstein B (1983) Anal Biochem 132:6-13
- Fu TD (1981) Eucarp Crucif Newslett 6:6-7
- Handa H, Ohkawa Y, Nakajima K (1990) Jpn J Genet 65:17-24
- Hanson MR, Conde MF (1985) Int Rev Cytol 94:213-267
- Hiesel R, Schobel W, Schuster W, Brennicke A (1987) EMBO J 6:29-34
- Ikeda TM, Ishii K, Terachi T, Tsunewaki K (1990) Jpn J Genet 65:435
- Kadowaki K, Suzuki T, Kazama S, Oh-fuchi T, Sakamoto W (1989) Nucleic Acids Res 17:7519
- Kadowaki K, Suzuki T, Kazama S (1990) Mol Gen Genet 224:10-16
- Kazama S, Suzuki T, Kadowaki K, Akihama T (1990) Nucleic Acids Res 18:5879
- Lonsdale DM (1987) Plant Physiol Biochem 25:265-271
- Lansdale DM, Hodge TP, Fauron CMR (1984) Nucleic Acids Res 12:9249-9261
- Makaroff CA, Palmer JD (1988) Mol Cell Biol 8:1474-1480
- Makaroff CA, Apel IJ, Palmer JD (1989) J Biol Chem 264:11706-11713
- Makaroff CA, Apel IJ, Palmer JD (1990) Plant Mol Biol 15:735~746
- Morikami A, Nakamura K (1987) J Biochem 101:967–976
- Palmer JD (1988) Genetics 118:341-351
- Palmer JD, Herbon A (1988) J Mol Evol 28:87-97
- Palmer JD, Shields CR (1984) Nature 307:437-440
- Quagliariello C, Saiardi A, Gallerani R (1990) Curr Genet 18:355– 363
- Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci USA 74: 5463–5467
- Schuster W, Brennicke A (1987) Nucleic Acids Res 15:9092

Shiga T, Baba S (1971) Jpn J Breed 21(suppl 2): 16-17 (in Japanese) Shiga T, Baba S (1973) Jpn J Breed 23:187-197 (in Japanese with English summary)

- Southern E (1980) Methods Enzymol 68:152-176
- Stern DB, Newton KJ (1986) Methods Enzymol 118:488-496
- Thompson KF (1972) Heredity 29:253-257
- Witt U, Hansen M, Albaum M, Abel WO (1991) Curr Genet 19:323-327
- Young EG, Hanson MR (1987) Cell 50:41-49 Young EG, Hanson MR, Dierks PM (1986) Nucleic Acids Res 14:7995-8006

Communicated by C.W. Birky, Jr.