

## Characterization of the gene *urf13-T* and an unidentified reading frame, ORF 25, in maize and tobacco mitochondria

S. E. Stamper, R. E. Dewey, M. M. Bland, and C. S. Levings, III

Department of Genetics, Box 7614, North Carolina State University, Raleigh, NC 27695-7614, USA

**Summary.** We have previously identified two large open reading frames, designated ORF13 and ORF25, in the Texas male-sterile cytoplasm (*cms-T*) of maize mitochondrial DNA (mtDNA). ORF13 is a single copy gene of chimeric origin that is uniquely transcribed and translated in the mitochondria of *cms-T* maize, where it produces a polypeptide of approximately 13,000 Mr. The ORF13 reading frame does not occur in the maize N, C or S cytoplasms or *Nicotiana tabacum*. ORF25 exists as a single copy and is transcribed in the four major maize cytoplasms (N, T, C and S) and *N. tabacum*. The predicted ORF25 polypeptide has a molecular weight of 24,374 in normal maize and 22,439 in tobacco. Several nucleotide and predicted amino acid changes have occurred in the ORF25 gene among the four maize cytoplasms and *N. tabacum*. Properties such as transcription and conservation of the sequence between two diverse species suggests that ORF25 encodes a functional plant mitochondrial gene. The ORF25 sequence of maize contains a chloroplast DNA insert homologous to a tRNA-Arg gene; this chloroplast DNA insert is absent in the tobacco ORF25 sequence. Comparison of the ORF25 and ORF13 sequences in restored and non-restored *cms-T* indicates no differences in their nucleotide sequences. Thus fertility restoration does not alter the primary sequences of ORF13 or ORF25.

**Key words:** Nucleotide sequence – Male-sterile cytoplasms – Chloroplast insert

### Introduction

A mtDNA sequence associated with the T type of cytoplasmic male sterility (*cms-T*) has been reported to contain two open reading frames, designated ORF13 and ORF25, which are separated by an intergenic region of 77 base pairs (Dewey et al. 1986). ORF13 is a novel reading frame of chimeric origin that is associated with the T type of male sterility. The ORF13 coding region contains sequences homologous to portions of the coding and 3' flanking regions of the maize mitochondrial 26S ribosomal gene. In addition, the ORF13 5' flanking region is homologous to a portion of the *atp6* 5' flanking region that presumably includes the *atp6* promoter. It is thought that this *atp6* promoter sequence is responsible for the transcription of ORF13 and ORF25 in the T cytoplasm. Recently, it has been shown that ORF13 encodes a 13,000 Mr polypeptide and the gene symbol, *urf13-T*, has been assigned (Dewey et al. 1987).

Although ORF25-like sequences are not detected in animal or fungal mitochondria, sequences homologous to ORF25 are apparently common to higher plant mitochondria. Southern blot analyses have revealed that homologous sequences occur in bean, wheat, pea and rice. In maize, northern blot analysis has indicated that ORF25 is transcribed, but as yet a polypeptide gene product has not been identified. The ORF25 of *cms-T* maize is terminated by a sequence homologous to a chloroplast tRNA-Arg gene (Dewey et al. 1986).

Since previous studies have suggested that the *urf13-T* sequence is only found in *cms-T* and that ORF25 is present in maize and other mitochondrial genomes, we have investigated the organization of the mtDNA region containing *urf13-T* and ORF25 in four maize cytoplasms and in tobacco. In addition, we have compared the flanking and coding sequences of *urf13-T* and ORF25 in *cms-T* fertile and sterile lines to determine

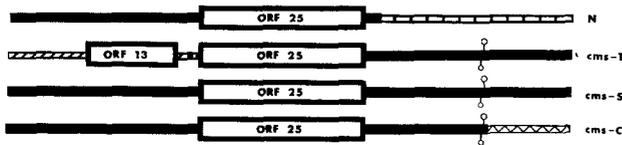


Fig. 1. Differences in the flanking regions of ORF25 among the four maize cytoplasms. *Boxed regions* indicate open reading frames. *Solid lines* represent homologous sequences. Potential 3' RNA stem-loop structures are illustrated by *lollipops*

ORF 25 5' FLANKING REGION

A

N -91	TTCGAATGTT	CTTTTTTT--	-----GGA	AAAAACCAAC	CACAAAAAAA	-51
C -96	TTCGAATGTT	CTTTTTTTTCG	TTGGGTGGA	AAAAACCAAC	CACAAAAAAA	-47
S -102	TTCGAATGTT	CTTTTTTTTCG	TTGGGTGGA	-----CCAAC	CACAAAAAAA	-59
	*****	*****	**	*****	*****	

N -50	GCTCCCTTT	CTCT-----	-TGGAGCAGA	GCTTCATCAT	AAAAG-TGGA	-9
C -46	GCTCCCTTT	CTCT-----	-TGGAGCAGA	GCTTCATCAT	AAAAGGTGA	-4
S -58	GCTCCCTTT	CTTTTTTTTT	TGGAGCAGA	GCTTCATCAT	AAAAG-TGA	-9
	*****	****	*****	*****	*****	**

N -8	GAGTCACA	<u>ATG</u>
C -3	GAG-----	<u>ATG</u>
S -8	GAGTCACA	<u>ATG</u>
	***	

B

N -82	TCFTTTTTTG	GAAAAACCA	ACCACAAAA	AAGTCTCCCT	TTCTCTTGA	-33
T -100	CCGATAGCGT	ACAAGTACCG	TGAGGACCA	AAGTCTCCCT	TTCTCTTTTG	-51
	* * *	** ***	* * *	*****	*****	

N -32	GCA-----	-----	-GAGCTTCAT	CATAAAAGTG	GAGAGTCACA	<u>ATG</u>
T -50	GGTGGGGGC	GGAGCTGAAT	AAATTGTATT	TTATAAGTIT	TAGAGTCACA	<u>ATG</u>
	*****	*****	*****	*****	*****	

Fig. 2A, B. 5' flanking regions of ORF25 in maize. Homologous and absent nucleotides are indicated with *asterisks* and *dashes*, respectively. A Comparison of 5' flanking regions of N, C and S cytoplasms. B Comparison of N and T 5' flanking sequences. The ORF13 termination codon and the ORF25 putative initiation codon are *underlined*

if restorer genes could act by modifying the DNA sequence.

## Materials and methods

**Nucleic acid isolations.** Mitochondrial DNA (mtDNA) and RNA (mtRNA) were isolated from 5 to 7 day old etiolated coleoptiles of *Zea mays* L., as previously described (Pring and Levings 1978; Schuster et al. 1983). Nucleic acids of male-fertile and male-sterile maize were obtained from: B73 (normal, male fertile), B73 (*cms-T*, *rfl Rf2*), Mo17 (*cms-C*), B73 (*cms-Vg*), B73 x Ky21 (*cms-T*, *Rfl Rf2*, male fertile). *Cms-Vg* is a member of the S group of male-sterile cytoplasms (Beckett 1971). mtDNA and mtRNA from *Nicotiana tabacum* L. cv. Coker 139, were prepared as previously described (Bland et al. 1985).

**RNA electrophoresis, transfer and hybridization.** Maize and tobacco mRNAs were heat denatured and fractionated in 1.2% agarose gels containing 6% formaldehyde (Maniatis et al. 1982). Gels were transferred to nitrocellulose as described by

Thomas (1980). DNA fragments specific to ORF25 were labeled by nick translation (Maniatis et al. 1982) using [<sup>32</sup>P]-dATP (New England Nuclear, 3,200 Ci/mmol) and hybridized to the nitrocellulose blots as described (Maniatis et al. 1982).

**Construction and screening of mtDNA libraries.** DNA libraries were constructed from BamHI digests and EcoRI digests of total maize mtDNA and BamHI digests of tobacco mtDNA cloned into the plasmid vectors pUC8 or 9 (Vieira and Messing 1982) and transformed into the *E. coli* hosts JM109 or JM83. Ampicillin resistant, lacZ<sup>-</sup> colonies were selected, transferred to nitrocellulose filters (Maniatis et al. 1982) and screened with a nick translated *cms-T* mtDNA fragment which extends from a BglII site (position 1722) to a SmaI site (position 2272) (Dewey et al. 1986). DNA hybridization conditions were as described (Maniatis et al. 1982).

**DNA sequence analysis.** DNA fragments were subcloned into M13 bacteriophage vectors mp18 and mp19 (Yanisch-Perron et al. 1985). DNA sequences were determined by the dideoxynucleotide chain-termination method (Sanger et al. 1977). Sequence analyses, polypeptide molecular weight estimations and predicted RNA secondary structures were performed with computer programs furnished by BIONET.

## Results

To study the organization around the ORF25 sequence of the various maize cytoplasms and tobacco, we prepared mtDNA libraries. Using an ORF25-specific hybridization probe (*cms-T*), BamHI clones containing homologous sequences were obtained from mtDNA libraries of maize C and T-restored cytoplasms, and *N. tabacum* (8.0, 9.0, and 1.1 kilobases (kb), respectively). EcoRI clones from N (fertile) and S maize cytoplasms (3.7 kb and 4.1 kb, respectively) were similarly derived. The ORF25 reading frames and immediate flanking regions were isolated from each of these clones and the DNA nucleotide sequences determined. Sequence comparisons were performed using normal maize as the standard. Based on the putative initiation and termination sites, ORF25 has a predicted amino acid length of 219 in N, 221 in T, C and S, and 198 in tobacco (Fig. 3). The putative polypeptides have molecular weights of 24,374 in normal maize and 22,439 in tobacco.

*Urf13-T* is absent in the maize N, C and S cytoplasms (Fig. 1). Comparison of the 5' flanking regions of ORF25 from the N, C and S cytoplasms reveals several short deletions/insertions as well as a single point mutation at position -11 (Fig. 2A). No other nucleotide changes are found between positions -74 and -511 in these three cytoplasms (data not shown). In-frame termination codons occur 18 bp, 6 bp and 138 bp upstream from the ORF25 ATG putative initiator codons in N, C and S, respectively. No other ATG codons occur between the putative initiator codon at position 1 and the upstream termination codons in any

## ORF 25

1	N	MET	ARG	PHE	SER	GLY	MET	ASP	MET	LYS	GLY	ILE	ASN	MET	LEU	PHE	ALA	ALA	ILE	PRO	SER	ILE	CYS	ALA	SER	SER	PRO	26
	T												LYS	ARG	LYS													
	C,S																											
	TBC			LEU		SER	THR	ASN		GLN	ALA	ARG	LYS							LEU						SER		
27	N	LYS	LYS	ILE	SER	ILE	TYR	ASN	GLU	GLU	MET	ILE	VAL	ALA	ARG	CYS	PHE	ILE	GLY	PHE	LEU	ILE	LEU	SER	TRP	LYS	SER	52
	T																											
	C,S																											
	TBC														LEU						ILE		PHE					
53	N	LEU	GLY	LYS	THR	PHE	LYS	GLU	THR	LEU	ASP	GLY	ARG	ILE	GLU	SER	ILE	GLN	GLU	SER	LEU	GLN	GLN	PHE	PHE	ASN	PRO	78
	T																									CYS		
	C,S																											
	TBC						VAL								GLN	ALA				GLU	SER					PRO		
79	N	ASN	GLU	VAL	ILE	LEU	GLU	GLU	SER	ASN	GLU	GLN	GLN	ARG	LEU	LEU	ASN	LEU	TRP	ILE	SER	LEU	ARG	ILE	CYS	SER	THR	104
	T																			GLN								
	C,S																											
	TBC				VAL	PRO	PRO												---	---	ARG					GLY		
105	N	VAL	LYS	VAL	VAL	GLU	SER	LEU	PRO	ALA	ALA	ARG	CYS	ALA	PRO	LYS	CYS	GLU	LYS	THR	VAL	GLN	ALA	LEU	LEU	CYS	ARG	130
	T																											
	C,S																											
	TBC	---	---																									
131	N	ASN	LEU	ASN	VAL	LYS	SER	ALA	THR	LEU	LEU	ASN	ALA	THR	SER	SER	ARG	ARG	ILE	ARG	LEU	GLN	ASP	ASP	ILE	VAL	THR	156
	T																											
	C,S																											
	TBC													PRO												ALA	ILE	
157	N	GLY	PHE	HIS	PHE	SER	VAL	SER	GLU	ARG	LEU	VAL	SER	GLY	SER	THR	THR	LEU	VAL	GLU	ALA	SER	THR	VAL	GLU	GLN	ILE	182
	T																											
	C,S																											
	TBC	LYS	MET		VAL	LEU		GLY	LYS		PHE	CYS	PRO	TRP	CYS	SER	SER	LYS	ALA		ARG	---	---			PHE		
183	N	ARG	GLU	ALA	PHE	LEU	LEU	GLU	PRO	ARG	ASP	LEU	ILE	ARG	GLU	GLY	PHE	ILE	VAL	LEU	ARG	LYS	VAL	ARG	VAL	GLY	GLY	208
	T																											
	C,S																											
	TBC																				SER	LEU	VAL		MET	TRP	ASP	
209	N	ILE	PRO	GLY	---	---	THR	CYS	GLY	ASP	GLY	VAL	GLY	LEU	TER												219	
	T						LYS	ARG	SER						TER													
	C,S						LYS	ARG	SER						TER													
	TBC	SER	LEU	LYS	---	---	ASN	LYS	GLU	LEU	GLU	---	---	---	TER													

Fig. 3. ORF25 predicted amino acid sequences of the four maize cytoplasms (N, T, C and S) and tobacco (TBC). Amino acid substitutions relative to N in T, C, S, and TBC are indicated. Amino acid deletions are represented by (---). TER indicates termination sites. The higher plant mitochondrial code (Fox and Leaver 1981) was used for translation

of these cytoplasms. The 5' flanking regions of ORF25 in N and T cytoplasms differ because of a recombinational event at position -16 (Fig. 2B). There is no significant sequence homology in the 5' flanking regions of N and T upstream from the recombination point except for an 18 bp sequence with perfect identity at positions -36 to -53 of N cytoplasm (Fig. 2B). Fauron et al. (1987) have determined, based on hybridization analyses, that two recombinational events in the maize *cms-T* male-fertile revertant (V3) mitochondrial genome result in a sequence arrangement identical to the region upstream of ORF25 in the male-fertile (N) mitochondrial genome. Importantly, one of these recombinations has occurred in the region between *urf13-T* and

ORF25, eliminating *urf13-T* from the *cms-T* revertant genome.

Nucleotide and predicted amino acid sequence comparisons among the ORF25 coding regions of the four maize cytoplasms reveal several variations (Fig. 3). A six base pair (bp) deletion/insertion at nucleotide positions 635-640 (Fig. 4) (amino acid positions 212-214) results in the addition of two amino acids (lys and arg) in T, C and S that are absent in N, along with the conversion of a threonine residue in N to a serine residue in T, C and S (Fig. 3). Aside from the insertion/deletion, ORF25 sequences of the N, C and S cytoplasms differ only in a single silent nucleotide substitution in S at position 603 (Table 1). ORF25 in

## ORF 25

N	1	<u>ATG</u> AGA TTT AGT GGA ATG GAT ATG AAG GGT ATA AAT ATG CTA TTT GCT GCT ATT CCA TCT AIT TGT GCA TCA AGT CCG 78
TBC		<u>ATG</u> G TCC C A C C G G T
N	79	AAG AAG ATC TCA ATC TAT AAT GAA GAA ATG ATA GTA GCT CGT TGT TTT ATA GGC TTT CTC ATA TTA AGT CGG AAG AGT 156
TBC		T T A C A C
N	157	TTA GGT AAG ACT TTC AAA GAA ACT CTC GAC GGG AGA ATC GAG TCT AIT CAG GAA TCA TTG CAG CAA TTC TTC AAT CCT 234
TBC		TG C G GA C CC
N	235	AAC GAA GTC AIT CTG GAG GAA TCC AAT GAA CAA CAA CGA TTA CTT AAT CTA CGG ATC AGC TTG CGA AIT TGC AGC ACC 312
TBC		A G CT CC --- --- A T G
N	313	GTA AAA GTA GTA GAA TCA TTA CCA GCG GCA CGC TGT GCG CCT AAG TGC GAA AAG ACA GTG CAA GCT TTG TTA TGC CGA 390
TBC		--- --- AT
N	391	AAC CTA AAT GTC AAG TCA GCA ACA CTT CTA AAT GCC ACT TCT TCC CGT CGC ATC CGT CTT CAG GAC GAT ATA GTC ACA 468
TBC		T C
N	469	GGT TTT CAC TTT TCA GTG AGT GAA AGA TTA GTA TCC GGG TCT ACA ACT TTG GTA GAA GCT TCT ACC GTA GAA CAA ATT 546
TBC		AAG A G G C T G G A T TGC C C G GC T G AAA C AGA --- --- TTC
N	547	CGA GAG GCC TTC TTA TTA GAA CCC AGA GAC CTA ATT CGA GAA GGC TTT ATA GTC CTC AGA AAG GTG AGG GTG GGG GGT 624
TBC		--- --- --- --- --- --- --- --- --- --- --- A G G G T A T T C A
N	625	ATC CCC GGG --- --- ACC TGT GGA GAC GGG GTG GGC CTG <u>TAG</u> 657
TBC		TCT TT AA --- --- AT AAA A TTA AA --- --- --- <u>TAG</u>

Fig. 4. Nucleotide sequence of ORF25 in normal (N) maize and tobacco (TBC). Nucleotide substitutions in tobacco are indicated. Codon deletions are indicated by (---). Putative initiation and termination codons are *underlined*

Table 1. ORF25 coding region. Nucleotide and amino acid substitutions among the four maize cytoplasms (N, T, C, S)

Nucleotide #	N/C	S	T	Amino acid #	N/C/S	T	Substitution <sup>a</sup>
28	G	G	A	10	GLY	LYS	NON
29	G	G	A				
30	T	T	A				
32	T	T	G	11	ILE	ARG	NON
36	T	T	A	12	ASN	LYS	NON
227	T	T	G	76	PHE	CYS	NON
287	G	G	A	96	TRP	GLN	NON
348	T	T	C	116	CYS	CYS	SIL
498	A	A	T	166	LEU	PHE	CON
513	A	A	T	171	THR	THR	SIL
603	C	A	A	201	LEU	LEU	SIL

<sup>a</sup> Substitutions are conserved (CON), nonconserved (NON) or silent (SIL)

the T cytoplasm, however, contains eleven base pair substitutions; three exchanges are silent, one results in a conserved amino acid change, and the remaining nucleotide substitutions involve the exchange of polar with nonpolar residues, and ionic with nonionic residues, resulting in five nonconserved amino acid changes (Table 1). Three of these nonconserved changes are positively-charged residues clustered near the N terminus of the putative polypeptide (Fig. 3).

Recombination accounts for differences in the 3' flanking regions of ORF25 among the maize cytoplasms (Fig. 1). All four cytoplasms are homologous from the TAG termination codon until a recombination point in N at nucleotide position 723, except for a five bp deletion and two bp substitutions in N (Figs. 1, 5). The deletion in N is a five bp direct repeat (AGCTC) present in T, C and S that is a component of the chloroplast-derived pseudo tRNA-Arg gene (Dewey et al. 1986). T,

ORF 25 3' FLANKING REGION

N	651	{GGGCTGTAG	CTC	TG	AGGATTAGAG	CACGTGGCTA	CGAACACGG	712
T,C,S	657	TAG	AGCTCA					714
N	713	TGTCGGGGGT	TGGAATCC}AC	TTCTGAGGG	GCTGGGGCT	CCACGGGAG	763	
T,C,S	715	T		AA	AA TATTTG	T TGCACCC	765	
N	764	CGAGCCTACT	755					
T,C,S	766	TTC TGA	757					

**Fig. 5.** ORF25 3' flanking regions of maize. Nucleotide substitutions are shown. The ORF25 termination codon is *underlined*. Chloroplast pseudo tRNA-Arg sequences are bordered by *brackets*. *Arrow* indicates the point of possible recombination in N

C and S are identical from the TAG stop codon to a recombination site in C at position 1164 (see Fig. 1) except for: a nucleotide substitution in T at position 802, a five bp insertion in T between positions 954 and 960, and a four bp deletion in T at position 1065 (data not shown). Schuster et al. (1986) have characterized a potential stem-loop structure which can be formed immediately 5' of this recombination point. This sequence occurs 426 bp downstream from ORF25 in T, C and S (data not shown). The recombination in C occurs six nucleotides 3' from this structure. A similar stem-loop structure can be formed 60 bp downstream from maize *atp9* (data not shown). Interestingly, homology occurs between the 3' flanking regions of ORF25 in C and *atp9* in N beginning six nucleotides downstream from this potential stem-loop structure (data not shown). Similar structures occur in the 3' flanking regions of *N. tabacum atp9* (Bland et al. 1986), ORF1 of the S2 plasmid (data not shown), *Oenothera atpA*, *Oenothera coxII*, and *Zea mays coxI* and may function as transcriptional processing or termination sites (Schuster et al. 1986).

In many systems, site-specific recombinational events are involved in the regulation of gene expression. A sequence inversion regulates the phase state of the H2 gene in *Salmonella* (Zieg et al. 1978) whereas transposition of cassettes determines the mating type in *Saccharomyces* (Klar et al. 1981). The *urf13-T* and ORF25 coding and flanking regions in T nonrestored (sterile) and T restored (fertile) cytoplasm were compared to determine if the primary DNA sequence was affected by the nuclear restorer genes. T restored (*Rf1 Rf2*) *urf13-T* and ORF25 sequences, including 1,124 bp of 5' flanking region and 1,106 bp of 3' flanking region, were identical to their counterparts in the sterile T (*rf1 Rf2*) cytoplasm (data not shown). These results indicate that fertility restoration is not associated with a change in the primary DNA sequences of these ORFs.

Substantial nucleotide and predicted amino acid variations were observed between the ORF25 sequences of maize and tobacco. These differences are distributed asymmetrically, primarily located near the carboxy

terminus (Figs. 3, 4). 91.4% nucleotide sequence conservation is observed between nucleotides 1 and 468. The majority of the substitutions are either silent or result in conserved amino acid changes. Interestingly, some of the altered residues in T more closely resemble the analogous residues in tobacco. In the T cytoplasm, the arginine and lysine residues at positions 11 and 12 and the phenylalanine at position 166 are homologous with their counterparts at these positions in tobacco (Fig. 3). Other residues altered in T are also non-conserved replacements in the tobacco ORF25. Although homology rapidly degenerates after position 468, the tobacco reading frame is maintained for another 138 bp. The entire tobacco ORF25 coding region is 594 nucleotides: 63 nucleotides shorter than the N maize sequence. Part of the heterogeneity at the 3' end of ORF25 in tobacco is due to the absence of a chloroplast derived tRNA-Arg pseudogene that terminates ORF25 in maize (Dewey et al. 1986). This insert in maize is 103 bp in length and contains extensive homology with a tobacco chloroplast tRNA-Arg gene, including 37 bp of 5' flanking sequence (Dewey et al. 1986). ORF25 initiates at the same position in maize and tobacco and both utilize a TAG termination codon. No significant homology is detected in the flanking regions of ORF25 between maize and tobacco (data not shown), which includes the absence of ORF13 in tobacco.

The preference for a thymine or adenine residue at the third codon position is common to higher plant mitochondrial genes (Dawson et al. 1984). For example, the *atpA* gene shows a third position bias for adenine (Braun et al. 1985). ORF25, like the *atpA* gene, displays a third position preference for adenine. In the maize cytoplasm and tobacco, ORF25 contains 34% and 32% adenine, respectively, as third position residues.

The hydrophobicity patterns of ORF25 in maize and tobacco are strikingly similar despite the extensive amino acid variation at the carboxy termini of the predicted polypeptides. Computer generated hydrophobicity plots (data not shown) of the maize and tobacco ORF25s reveal that the predicted polypeptide is mostly hydrophilic and, therefore, may be a soluble matrix protein.

Northern blot analyses, using a maize ORF25-specific fragment as a hybridization probe (positions 83 to 633), indicate that ORF25 is transcribed in all four maize cytoplasm and *N. tabacum* (Fig. 6). The ORF25 steady state transcriptional patterns of C and S are similar, each having transcripts of approximately 3.2, 2.1 and 1.35 kb (Fig. 6A). ORF25 in N displays an altered pattern; it contains transcripts of approximately 3.4, 2.3 and 1.6 kb (Fig. 6A). The transcriptional pattern in N may be due to the recombination in the 3' flanking region which eliminates or translocates the stem-loop structure. T exhibits a complex transcriptional pattern.

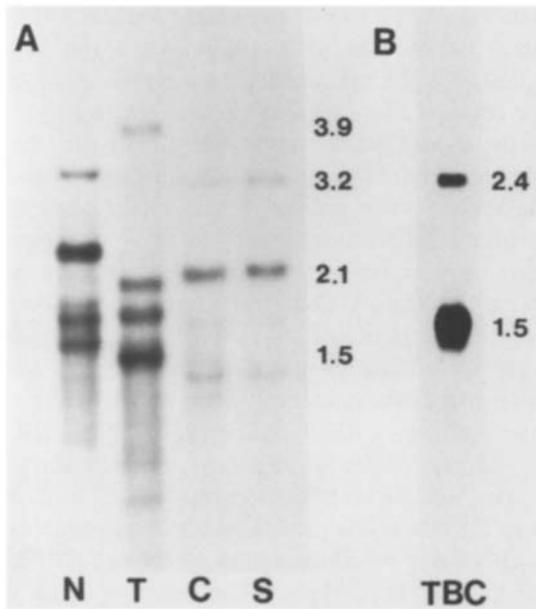


Fig. 6A, B. Northern blot analysis of ORF25 in the four cytoplasms of maize (A) and *Nicotiana tabacum* (B). Approximate RNA sizes, determined from ribosomal RNA size markers, are expressed in kilobases

The analysis detects many RNA species ranging in size from 3.9 to 1.1 kb (Fig. 6A; Dewey et al. 1986). The unique T pattern is due to the novel chimeric nature of this sequence (Dewey et al. 1986). ORF25 hybridizes to two RNA species in *N. tabacum*: a minor band approximately 2.4 kb in length and an abundant transcript of 1.5 kb, possibly representing the mature message (Fig. 6B).

## Discussion

The ORF25 region of maize *cms-T* mitochondria is closely linked to *urf13-T*: a gene of chimeric origin that is associated with cytoplasmic male sterility and susceptibility to the T toxin of *Helminthosporium maydis* (Dewey et al. 1986). Sequence analyses suggest that *urf13-T* has arisen by recombinations among the coding and/or flanking regions of the maize mitochondrial 26S ribosomal gene and *atp6* gene. The open reading frame of this chimeric gene is unique to the T cytoplasm of maize and is actively transcribed. A 13 kD protein product has been identified which corresponds to *urf13-T* (Dewey et al. 1987). *Urf13-T* is absent in the maize N, C and S cytoplasms (Fig. 1) and tobacco. The transcription of *urf13-T* and ORF25 in the T cytoplasm is presumably under the control of the *atp6*

promoter. The absence of *urf13-T* and its 5' flanking region in N, C and S suggests that the transcription of ORF25 in these cytoplasms is under the control of a different promoter. Moreover, the tobacco ORF25 5' flanking region contains no homology with this region in maize and, therefore, a different promoter must be utilized. Homology between T and the other maize cytoplasms ends in the intergenic region of *urf13-T* and ORF25: 16 bp, 12 bp and 17 bp from the ORF25 ATG initiator codon in N, C and S, respectively. Further upstream, T displays 18 bp of perfect homology with the other maize cytoplasms, suggesting that more than one recombination has occurred (Fig. 2B). The ORF25 coding sequences of N, C and S cytoplasms of maize are nearly identical. In contrast, the ORF25 in T contains several substitutions, most of which could result in nonconserved amino acid substitutions (Fig. 3).

The ORF25 of tobacco shows substantial variation compared with maize at both the nucleotide and predicted amino acid levels. Sequence homology begins in the proximity of the putative initiation codon and ends abruptly near the 3' terminus. This sequence heterogeneity in tobacco involves the absence of a chloroplast-derived pseudo tRNA-Arg sequence. The chloroplast sequence insertion may have occurred in maize after the two species diverged or may have been present in a common ancestor and subsequently lost from the tobacco genome. This sequence is identical in the four maize cytoplasms except for the absence of a five bp repeat in N.

Computer searches indicate that sequences similar to ORF25 do not occur in animal or fungal mitochondrial genomes. Conversely, the presence of sequences homologous with ORF25 in diverse plant species has led to speculation that ORF25 may be a gene common to higher plant mitochondria. Although a protein product has not yet been isolated, our results suggest that ORF25 encodes a functional plant mitochondrial gene because: 1) Homologous sequences are present in maize, tobacco, bean, wheat, pea and rice. 2) This gene is transcribed in four maize cytoplasms and tobacco (transcription studies of bean, wheat, pea and rice have not been performed). 3) The open reading frame is conserved among the four major maize cytoplasms and tobacco.

*Acknowledgements.* We gratefully acknowledge the technical assistance of Jane Suddith, Carol Griffin and Wilma Hu. Computer resources were furnished by BIONET whose funding is provided by the Biomedical Research Technology Program, Division of Research Resources, National Institutes of Health, Grant # U41 RR-01685-02. This work was supported by research grants from Agrigenetics, Inc. and the National Science Foundation.

**References**

- Beckett JB (1971) *Crop Sci* 11:724–726
- Bland MM, Matzinger DF, Levings CS III (1985) *Theor Appl Genet* 69:535–541
- Bland MM, Levings CS III, Matzinger DF (1986) *Mol Gen Genet* 204:8–16
- Braun CJ, Levings CS III (1985) *Plant Physiol* 79:571–577
- Dawson, AJ, Jones VP, Leaver CJ (1984) *EMBO J* 3:2107–2113
- Dewey RE, Levings CS III, Timothy DH (1986) *Cell* 44:439–449
- Dewey RE, Levings CS III, Timothy DH (1987) *Proc Natl Acad Sci USA* 84:5374–5378
- Fauron CM-R, Abbott AG, Brettell RIS, Gesteland RF (1987) *Curr Genet* 11:339–346
- Fox TD, Leaver CJ (1981) *Cell* 26:315–323
- Klar AJS, Strathern JN, Broach JR, Hicks JB (1981) *Nature* 289:239–244
- Maniatis T, Fritsch EF, Sambrook J (eds) (1982) *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Pring DR, Levings CS III (1978) *Genetics* 89:121–136
- Sanger F, Nicklen S, Coulson AR (1977) *Proc Natl Acad Sci USA* 74:5463–5467
- Schuster AM, Sisco PH, Levings CS III (1983) Two unique RNAs in cms-S and RU maize mitochondria. In: Goldberg RB (ed) *Plant molecular biology*. Liss, New York, pp 437–444
- Schuster W, Hiesel R, Isaac PG, Leaver CJ, Brennicke A (1986) *Nucleic Acids Res* 14:5943–5954
- Thomas PS (1980) *Proc Natl Acad Sci USA* 77:5201–5205
- Vieira J, Messing J (1982) *Gene* 19:259–268
- Yanisch-Perron C, Vieira J, Messing J (1985) *Gene* 33:103–119
- Zieg J, Hilmen M, Simon M (1978) *Cell* 15:237–244

Communicated by C. W. Birky

Received May 29, 1987