Curr Genet (1993) 23:148–153

Current Genetics

© Springer-Verlag 1993

Characterization of the *Brassica campestris* mitochondrial gene for subunit six of NADH dehydrogenase: *nad6* is present in the mitochondrion of a wide range of flowering plants

Jacqueline M. Nugent* and Jeffrey D. Palmer

Department of Biology, Indiana University, Bloomington, IN 47405, USA

Received May 8, 1992/Accepted June 25, 1992

Abstract. We have isolated the *Brassica campestris* mitochondrial gene *nad6*, coding for subunit six of NADH dehydrogenase. The deduced amino-acid sequence of this gene shows considerable similarity to mitochondrially encoded NAD6 proteins of other organisms as well as to NAD6 proteins coded for by plant chloroplast DNAs. The *B. campestris nad6* gene appears to lack introns and produces an abundant transcript which is comparable in size to a previously described, unidentified transcript (#18) mapped to the *B. campestris* mitochondrial genome. An alignment of NAD6 proteins (deduced from DNA sequences) suggests that *B. campestris nad6* transcripts are edited. Southern-blot hybridization indicates that *nad6* is present in the mitochondrial genome of all of a wide range of flowering plant species examined.

Key words: *nad6* – Flowering plants – Mitochondrial gene

Introduction

NADH: ubiquinone reductase (NADH dehydrogenase, complex I) is a multimeric inner mitochondrial membrane enzyme that forms the link between the Krebs cycle and the electron transport chain (Moller 1986). This enzyme is very complex and contains about 30 different protein subunits. The genes coding for these subunits are partitioned between the nuclear and mitochondrial genomes of most organisms. Surprisingly, genes coding for 11 subunits of NADH dehydrogenase have also been found in land plant chloroplast genomes (Berger et al. 1991; Dupuis et al. 1991), suggesting that chloroplasts contain an NADH dehydrogenase that may be involved in a chloro-respiratory pathway.

The completely sequenced mitochondrial genomes of the land plant *Marchantia polymorpha* (Oda et al. 1992), the filamentous ascomycete *Podospora anserina* (Cummings et al. 1990) and a variety of animals (reviewed by Attardi and Schatz 1988) contain genes that specify the same seven subunits of NADH dehydrogenase (subunits 1, 2, 3, 4, 4L, 5 and 6). Mitochondrial DNAs (mtDNAs) of other organisms contain a subset of these genes or additional genes for other subunits of complex I. The *Chlamydomonas reinhardtii* mitochondrial genome contains genes for subunits 1, 2, 4, 5, and 6 (Michaelis et al. 1990); Paramecium mtDNA codes for subunits 1, 2, 3, 4, 5, 7, and 9 (Pritchard et al. 1990); and *Trypanosoma brucii* mtDNA encodes, at a minimum, subunits 1, 4, 5, 7 and 8 (Payne et al. 1985; Jasmer et al. 1985; Feagin and Stuart 1985; Koslowsky et al. 1990; Souza et al. 1992). Notable exceptions are the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, whose mitochondrial genomes do not contain any NADH dehydrogenase genes (reviewed by Attardi and Schatz 1988), and whose mitochondria may entirely lack a complex I-like NADH dehydrogenase (deVries and Grivell 1988). So far, seven genes coding for subunits of NADH dehydrogenase have been identified in flowering plant mtDNAs: *nad1* (Chapdelaine and Bonen 1991; Wissinger et al. 1991) *nad2* (Xue et al. 1990; Binder et al. 1992) *nad3* (Gualberto et al. 1988), *nad4* (Lamattina and Grienenberger 1991; Gass et al. 1992) *nad5* (Knoop et al. 1991), *nad7* (L. Bonen personal communication), and *nad9* (J. M. Grienenberger personal communication). However, since flowering plant mtDNAs are by far the largest and most complex mtDNAs known, ranging in size from 208 to 2400 kb (reviewed by Newton 1988) and being extremely prone to structural rearrangement, determination of the complete gene content and organization of these mtDNAs is still far from complete.

A detailed transcriptional map of the *Brassica campestris* mitochondrial genome revealed 24 abundant and positionally distinct transcripts longer than 500 nucleotides (Makaroff and Palmer 1987). Fifteen of these abundant transcripts could not be identified as producing any known product (protein or RNA). We have sequenced a region of the *B. campestris* mitochondrial genome that

* *Present address:* John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

Correspondence to: J. D. Palmer

produces one of these abundant but heretofore unidentified transcripts (#18; Makaroff and Palmer 1987). We identify an open reading frame (ORF) which is highly expressed and whose transcript is similar in size to the one mapped by Makaroff and Palmer (1987). This ORF codes for subunit 6 of NADH dehydrogenase by virtue of the similarity of its predicted amino-acid sequence to those of other known *nad6* genes. The *B. campestris nad6* is located just downstream from a virtually intact chloroplast-derived *rbc*L gene.

Materials and methods

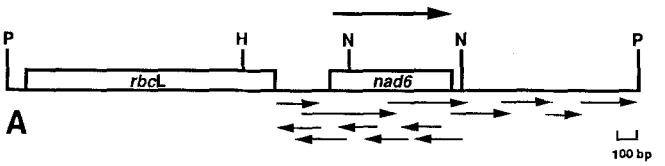
Mitochondrial DNAs were isolated from either green leaves or etiolated seedlings according to the DNase I procedure (Kolodner and Tewari 1972). *B. campestris* total RNA was prepared from etiolated shoot tissue according to a scaled-up version of the guanidinium isothiocyanate extraction procedure of Hamby et al. (1988). *B. campestris* seed (var. Purple-top white globe) was purchased from W. Atlee Burpee and Co. PA, U.S.A. The Southern-blot hybridization was carried out using standard procedures (Sambrook et al. 1989). The membrane was prehybridized and hybridized at 60 °C in Blotto solution (0.5% nonfat dry milk, 1% SDS, 4 × SSC, preheated to 60 °C). After hybridization the blot was washed in 2 × SSC, 0.5% SDS at 60 °C. RNA was electrophoresed through a 1.2% agarose gel after denaturation with glyoxal and dimethyl sulfoxide (Sambrook et al. 1989) and transferred to a Zetabind membrane. The filter was prehybridized and hybridized at 60 °C in 5% SDS, 10% dextran sulfate, and 1 *M* NaCl. Post-hybridization washes were as described above. Hybridization probes were labeled using random hexamer oligonucleotide primers (Sambrook et al. 1989). DNA sequencing was by the dideoxynucleotide chain-termination method (Sanger et al. 1977) using a series of nested deletions created by exonuclease III digestion (Erase-A-Base, Promega). All of the P3.1 clone downstream from the chloroplast-derived *rbc*L gene was sequenced on one strand; the *nad6* locus was sequenced completely on both strands. The GenBank/EMBL accession number of this sequence is X68256.

Protein and nucleotide sequences were searched using the GCG sequence-analysis software package, version 7.0, April 1991 (Devereux et al. 1984). Protein sequences were searched through the SwissProt database, release number 17.0 (2/91); nucleotide sequences were searched through the GenBank database, release 67 (3/91). The computer program CLUSTAL was used to align protein sequences.

Results

Sequence analysis and identification of the nad6 gene

A transcriptional map of the 218-kb *B. campestris* mitochondrial genome revealed three transcripts (2000, 1900 and 800 nucleotides) which hybridize to a cloned *Pst*I fragment of 3.1 kb (P3.1) from this genome (Makaroff and Palmer 1987). The two larger transcripts are products of the chloroplast *rbc*L gene (encoding the large subunit of ribulose biphosphate carboxylase), which hybridize to P3.1 because a virtually intact chloroplast-derived *rbc*L gene is contained within this mtDNA fragment (Makaroff and Palmer 1987; Nugent and Palmer 1988, and unpublished). The smallest transcript (#18 in the scheme of Makaroff and Palmer 1987) was inferred to be a mitochondrial transcript but did not correspond to




```
ATACCACAAGGCTATCTGAGTTCACAGGTTTAGTTGCCTGTACCTTTTTTTGAATCGAGA -226
TTTTTGTGGTGTTCAGTCTACCGCTCCGTGGGTACAAGATCGAAAAGAATGCATTAAATG -166
GATGCATTGAGATTCCGTAAGTAACTCAGTGAGTGCTTTCTAAGAAAGAAGGGCAAGGAA -106
GAAGAAAATGAAATAGGAACAACCGCGCTGGTCGTAATAGATCGACTTTCATGCTGAAGC  -46
AAGAACTAGCATGAAAGTTCCATTTCAGGGAAGGACGACGTACCCATGATACTTTCTGTT   15
                                                  M  I  L  S  V
TTGTCGAGCCCTGCTTTGGTCTCTGGTTTGATGGTTGCACGTGCTAAAAATCCGGTACAT   75
 L  S  S  P  A  L  V  S  G  L  M  V  A  R  A  K  N  P  V  H
TCCGTTTTGTTTCCCATCCCAGTCTTTCGCGACACTTCAGGTTTACTTCTTTTGTTAGGT  135
 S  V  L  F  P  I  P  V  F  R  D  T  S  G  L  L  L  L  L  G
CTCGACTTCTTCGCTATGATCTTCCCAGTAGTTCATATAGGAGCTATAGCCGTTTCATTC  195
 L  D  F  F  A  M  I  F  P  V  V  H  I  G  A  I  A  V  S  F
CTTTTCGTTGTTATGATGTTCCATATTCAAATAGCGGAGATTCACGAAGAAGTATTGCGC  255
 L  F  V  V  M  M  F  H  I  Q  I  A  E  I  H  E  E  V  L  R
TATTTACCAGTGAGTGGGATTATTGGACTGATCTTTTGGTGGGAAATGTTTTTCATTTTA  315
 Y  L  P  V  S  G  I  I  G  L  I  F  W  W  E  M  F  F  I  L
GATAATGAAAGCATTCCATTACTACCAACCCAAAGAAATACGACCTCTCTTAGATATACG  375
 D  N  E  S  I  P  L  L  P  T  Q  R  N  T  T  S  L  R  Y  T
GTTTATGCCGGAAAGGTACGAAGTTGGACTAATTTGGAAACATTGGGAAATTTACTTTAT  435
 V  Y  A  G  K  V  R  S  W  T  N  L  E  T  L  G  N  L  L  Y
ACCTACTATTCTGTCTGGTTTTTGGTTCCTAGTCTTATTTTATTAGTAGCCATGATTGGG  495
 T  Y  Y  S  V  W  F  L  V  P  S  L  I  L  L  V  A  M  I  G
GCTATAGTACTGACTATGCATAGGACTACTAAGGTGAAAAGACAGGATGTATTCCGACGA  555
 A  I  V  L  T  M  H  R  T  T  K  V  K  R  Q  D  V  F  R  R
AATGCTATTGATTTAGGAGGACTATAATGAGGAGGACGACTGACCCACTCACGATCTAC  615
 N  A  I  D  F  R  R  T  I  M  R  R  T  T  D  P  L  T  I  Y
TAAAGGGCAAGTAGCTTGATTGGTTCGAATCCAATTCGCGA                     656
 *
```

B

Fig. 1 A, B. Organization (**A**) and sequence (**B**) of *nad6* in *B. campestris* mtDNA. **A** direction of transcription is indicated by the *upper arrow: arrows below* the drawing indicate the sequencing strategy used in the vicinity of the *nad6* gene. Restriction sites mentioned in the text are shown (*P*, *Pst*I; *H*, *Hin*dIII; *N*, *Nru*I). **B** the amino-acid translation according to the 'universal' genetic code is shown below the DNA sequence. A putative ribosome-binding site (Fig. 3) upstream of the first in-frame ATG is *boxed.* Nucleotides (250–254) showing identity to the conserved 5′ splice site of group II A1 introns (Michel et al. 1989) are *underlined.* The EMBL/GenBank accession number of this sequence is X68256

any of the known mitochondrial genes in the *B. campestris* genome (Makaroff and Palmer 1987).

We have sequenced most of the P3.1 clone and present the sequence of a portion of the clone (941 bp) that lies downstream from the chloroplast-derived *rbc*L gene (Fig. 1). The *rbc*L sequence will be presented in a separate paper on the transfer and evolution of chloroplast *rbc*L genes in plant mitochondrial genomes. The region presented here contains a single ORF of significant size, with the potential to encode a 205 amino-acid polypeptide (Fig. 1 B). The best matches between this ORF and sequences in the GenBank nucleotide sequence database and the SwissProt protein sequence database are to mitochondrial and chloroplast *nad6* genes and their protein products (Fig. 2). The ends of the *nad6* proteins are moderately well conserved in length and sequence, whereas the middle is highly variable in both respects (Fig. 2 and legend to Fig. 2). For example, the *B. campestris nad6* is 35% identical to the *C. reinhardtii* protein over the former's first 78 amino acids and 37% identical over their last 38 alignable amino acids, whereas the two proteins show no significant similarity over a region corresponding to the central 61 amino acids of the *B. campestris* protein. We conclude that the *B. campestris* ORF is an *nad6* gene.

```
                                    .    .. .  .  .  .  .
Brassica                        MILSVLSSPALVSGLMVARAKNPVHS  26
Chlamy                        MFFENSAILLCALLSIA-VGYTKSPFMS  27
Podos     MNNNYPLFWINEILTNGFVEYILDIFSIMAFLTGIYVILTKNPIVS  46
Mar (cp)            MKLPESFYETIFLFLESGLILGSLGVILLTNIVYS  35
Chick                       MTYFVIFLGICFMLGVLAVASNPSPYYG  28
                                      ^    ^ ^ +    ^   ^
                          .  .  .  .  .  .  # .# .. ..#.... .
Brassica VLFPIPVFRDTSGLLLLLGLDFFAMIFPVVHIGAIAVSFLFVVMMF  72
Chlamy   LMYSVMLFINSSFVLMMLGFEFLALVNLLVYVGALAVLFLFVIMLL  73
Podos    VLFLILLFGGISSYLNIIGLNFIGLSYIIVYIGAVSILFLFILMLI  92
Mar (cp) ALFLGFVFVCISLLYLLLNADFVAAAQILIYVGAVNVLIIFAVMLI  81
Chick    VVGLVVASVMGCGWLVSLGVSFVSLALFLVYLGGMLVVFVYSVSLA  74
          ^         ^     ^^ ^+ ^    ^^^^+^^ ^ ^^^ ^ ^
                    .  .              .
Brassica HIQIAEIHEEVLRYLPVSGIIGLIFWWEMFFILDNESIPLLPTQRN 112
Chlamy   EIPATELRAYSRGWSTLGIFVFIINGVFQITPSMGPRGIITGLPGA  97
Podos    NIRTSELQSNTSNSIPLTIFIGIIFSNFLFPMLPYDIVMLSNFYNN 138
Mar (cp) NKKQ---YSNFFVYWTIGDGITLTLCTSIFLLLNNFISNTSWSKIF 117
Chick    ADPYPEAWGD---WRVVGYGLGFVLVVWMGVVLGGLVDFWKVGVVT 109
                ^        ^^  ^

Brassica TTSLRYTVYAGKVRSWTNLET------------------------- 139
Chlamy   ESITN---------------------------------------- 124
Podos    YFSEDFYTIDVNINDNNLNNLYNNVLYFMTSVIWDGSVIDFNHITA 184
Mar (cp) LMTKPNLVVKDIILINTVRH------------------------- 144
Chick    VDGGGVSFARLDF---------------------------------- 130

                .. . .  .  .. . .  .  ...
Brassica LGNLLYTYYSV-WFLVPSLILLVAMIGAIVLTMHRTTKVKRQDVFR 184
Chlamy   LGHALYLYFAD-LLILNSLVLTVALFGRFAIAPVRTTGR*       162
Podos    IGNIMYTIYNI-WLIIASFILLLAMVGSIVITIKQRKI*        221
Mar (cp) IGSELLTEFIL-PFELMSIILLVALIGAITLARREKKIELEKNDFF 189
Chick    SGVAVFYSCGVGLFLVAGWGLLLALF--VVLELVRGLSRGAIRAV* 173
          +  ^         ^ ^ ^  + ^+^    ^^^

Brassica RNAIDFRRTIMRRTTDPLTIY*                         205
Chlamy
Podos
Mar (cp) NF*                                            191
Chick
```

Fig. 2. Alignment of five NAD6 protein sequences. The alignment contains four mitochondrial NAD6 sequences, from *B. campestris* (*Brassica*; Fig. 1), *C. reinhardtii* (*Chlamy*; Boer and Gray 1989), *P. anserina* (*Podos*; Cummings and Domenico 1988), and chicken (*chick*; Desjardins and Morais 1990), and the chloroplast NAD6 sequence from *M. polymorpha* [*Mar (cp)*; Ohyama et al. 1986]. Three amino-acid positions in the *Brassica* sequence that may be altered by RNA editing are denoted by '#'. Gaps are indicated by '−'. Amino acids that are conserved across all five species are indicates by '+'. Conservative amino-acid changes are indicated by '^'. Residues that are identical between *Brassica* and *Chlamydomonas* are marked by '.'. The initial alignment produced by the pairwise, multiple sequence alignment program of CLUSTAL showed a region (corresponding to amino acids 115–184 of *Podospora*) that was notable for being rich in gaps (three in *Brassica*, two different ones in *Chlamydomonas*, four different ones in *Marchantia*, and four different ones in chicken) and extremely poorly matched in sequence. In our view, the alignment in this region is completely 'forced', i.e., we cannot discern any evidence of evolutionary homology in it. Accordingly, to reduce the 'gappiness' of the alignment shown, and thereby avoid the implication that the gaps are evolutionarily meaningful, we have deliberately altered the alignment in this region such that all of the amino acids for the four shorter sequences (i.e., all but *Podospora*) are shown as 'left-justifed', starting at *Podospora* position 115

This is the first report of a flowering plant mitochondrial gene coding for subunit six of NADH dehydrogenase.

As is the case for several plant mitochondrial genes (MacFarlane et al. 1990), *nad6* shows considerable bias towards the use of A/T (62%) over G/C (38%) in the third positions of codons. A block of nine nucleotides upstream of the first in-frame AUG of *nad6* is identical at seven positions to a region found upstream of the *B. campestris nad4* gene (Gass et al. 1992) and is complementary at eight of nine positions to the 3′-terminus of the mitochondrial 18 s rRNA from *Oenothera* (Fig. 3).

```
Oenothera mt 18S rRNA      CC UAAGUUAGG UCGGUGUCC
 (3' terminus)
                              **** ****
B. campestris nad6 5'     AAA GUUCCAUUU CAGGGAAGGACGACGUACCCAUG
                              *********
B. campestris nad4 5'     AAU GUUCGAUUC ACUCUAUG
```

Fig. 3. A putative ribosome binding site is found upstream of the first in-frame AUG (*underlined*) of *nad6*. A nononucleotide upstream of the *B. campestris nad6* and *nad4* genes is complementary to the 3-terminus of *Oenothera* 18 s mitochondrial rRNA (allowing for G-U pairing). Nucleotides upstream of *nad6* and *nad4* that are complementary to the 3′-terminus of 18*s* rRNA are indicated by *stars*. An 18 *s* rRNA-sequence (5′-UGAA) complementary to putative 'Shine-Dalgarno' type sequences (Dawson et al. 1984) of other plant mitochondrial mRNAs is *underlined*. The 18 *s* rRNA sequence is from Brennicke et al. (1985); the *nad4* sequence is from Gass et al. (1992)

Therefore, we propose that this region may serve as a ribosome binding site. The *B. campestris* sequence compares well with other putative plant mitochondrial ribosome binding sites (Dawson et al. 1984), both in sequence and in position (20 nucleotides) from the presumed start site of translation. However, it should be stressed that no experimental studies have yet been carried out to define functional ribosome binding sites in plant mitochondrial mRNAs and that some plant mitochondrial genes lack any homolog of the proposed ribosome binding site (Lonsdale 1989).

Transcription of the nad6 gene

A 564-bp *Nru*I fragment, which contains most of the *B. campestris nad6* locus (Fig. 1 A), was used to probe a Northern blot of *B. campestris* total RNA. A single abundant transcript of approximately 750 nt was detected (Fig. 4). A transcript of approximately 1 700 nt also hybridizes weakly to the *nad6* probe. Since the *Nru*I probe was gel-isolated from a digest of a 2.0-kb *Hin*dIII-*Pst*I subclone that contains part of the *rbc*l sequence (Fig. 1 A), this signal may represent hybridization to the highly abundant (Gruissem et al. 1988) chloroplast *rbc*L transcripts of this size by a small amount of contaminating *rbc*L DNA in the probe. Consistent with this hypothesis, a probe specific for the 3′ half of the tobacco chloroplast *rbc*L gene hybridizes to the same 1 700-nt transcript [data not shown; neither the *nad6* nor the 3′ *rbc*L probe hybridize to the larger *rbc*L transcript detected by Makaroff and Palmer (1987)]. However, since the ends of the 750- and 1 700-nt transcripts have not been mapped we cannot rule out other possibilities. The size of the major transcript correlates well with the abundant, mitochondrially-derived transcript (#18; 800 nt) mapped to P3.1 by Makaroff and Palmer (1987).

nad6 is present in a wide range of flowering plant mtDNAs

To determine whether *nad6* was present in other plant mitochondrial genomes, the 564-bp *Nru*I probe was used to probe mtDNAs from a wide range of flowering plants

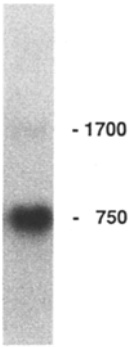


Fig. 4. The *B. campestris nad6* gene produces a single abundant transcript. Northern analysis of the *B. campestris nad6* locus using a 564 bp *Nru*I fragment (Fig. 1) as a probe to *B. campestris* total RNA. Sizes of transcripts are indicated in nucleotides

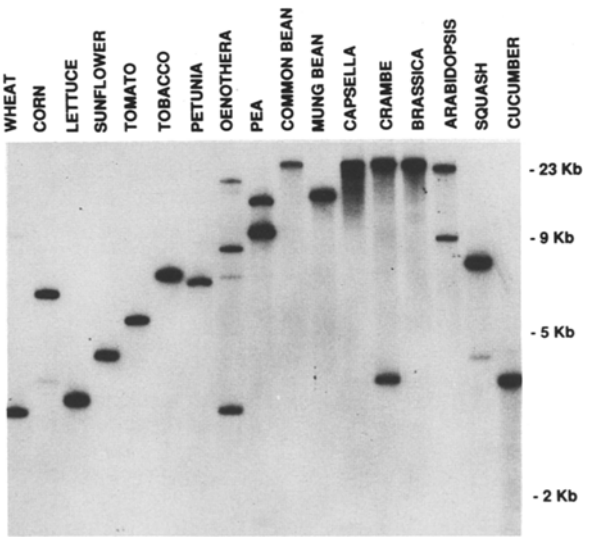


Fig. 5. *Nad6* is present in the mtDNA of a wide range of flowering plant species. Hybridization to a filter-blot of mtDNAs with a 564-bp *Nru*I fragment (Fig. 1) specific to the *B. campestris* mitochondrial *nad6* gene. MtDNAs were cut with restriction enzymes, electrophoresed through a 1.0% agarose gel and blotted onto a nylon membrane. Wheat (*Sal*I fragments); corn (*Bam*HI); lettuce (*Sac*I); sunflower (*Sac*I), tomato (*Sal*I); tobacco (*Sal*I); *Petunia* (*Sal*I); *Oenothera* (*Sal*I); pea (*Sal*I); common bean (*Sal*I); mungbean (*Sal*I); *Capsella* (*Sal*I); *Crambe* (*Sal*I); *Brassica* (*Sal*I); *Arabidopsis* (*Sal*I); squash (*Sac*I); cucumber (*Sac*I). Other mtDNAs included in the survey but not shown here are watermelon, spinach and radish

(two monocotyledons and 18 dicotyledons). Positive hybridization signals were obtained for all species examined (Fig. 5), demonstrating that the *nad6* gene is a conserved feature of flowering plant mitochondrial genomes. Many of the mtDNAs exhibit a single strong hybridizing fragment, indicating that *nad6* is generally a single-copy mitochondrial gene. Several species (*Oenothera*, pea, *Crambe, Arabidopsis*) show relatively strong hybridization to more than one fragment both in this experiment and when the mtDNAs are cut with a second restriction enzyme (data not shown), suggesting that part or all of the gene is repeated in these mitochondrial genomes. However, other possibilities, such as disruption of *nad6* by *cis*- or *trans*-spliced introns in these organisms, cannot be ruled out. At least three flowering plant *nad* genes contain exons that are scattered widely in the mitochondrial genome (Chapdelaine and Bonen 1991; Wissinger et al. 1991; Conklin et al. 1991; Knoop et al. 1991; Binder et al. 1992), transcripts of these genes require *trans*-splicing to produce a functional mRNA.

Discussion

We have isolated the first gene coding for subunit six of NADH dehydrogenase (*nad6*) from a flowering plant mitochondrial genome. In *B. campestris* this gene is located on a 3.1-kb *Pst*I fragment just downstream from a chloroplast-derived *rbc*L sequence. *Nad6* produces a single abundant transcript similar in size to a transcript (#18) mapped to P3.1 by Makaroff and Palmer (1987).

Southern-blot hybridization indicates that an *nad6* gene is present in the mitochondrial genome of all of a wide range of flowering plants examined. Although plant mitochondrial protein gene content is generally well conserved, a few genes (three ribosomal protein genes and *coxII*) are thought to have been lost from the mitochondrion during the evolution of flowering plants (Bland et al. 1986; Wahleithner and Wolstenholme 1988; Schuster et al. 1990b; Nugent and Palmer 1991). In the case of *coxII*, the gene loss is tolerated because a functional copy of the gene is present in the nucleus of those legumes lacking the mitochondrial gene (Nugent and Palmer, 1991); presumably this is also the case for the putative ribosomal protein gene losses.

Most plant mitochondrial mRNAs are edited via a cytidine to uridine conversion at various positions along the mRNA molecule (Covello and Gray 1989; Gualberto et al. 1989; Hiesel et al. 1989); more rarely, uridine to cytidine alterations are observed (Gualberto et al. 1990; Schuster et al. 1990a). We do not have a sequence for *B. campestris* mitochondrial *nad6* transcripts or cDNAs and thus do not know whether, or to what extent, the *nad6* mRNAs are edited. One way of indirectly inferring editing sites in *B. campestris nad6* transcripts is to produce an alignment of the *B. campestris* NAD6 protein (deduced from the DNA sequence) and other known NAD6 proteins. Amino-acid positions that are invariant in non-plant mitochondrial proteins but have undergone radical substitutions in the plant mitochondrial protein are likely candidates to be changed by RNA editing (Covello and Gray 1989). Within the most highly conserved region of the NAD6 protein there are at least three amino acids that might be altered as a result of cytidine to uridine editing on *B. campestris nad6* transcripts to produce a more structurally conserved protein (Fig. 2). Two of these proposed edits involve radical amino-acid changes, Pro(CCA) to Leu(CUA) at position 54, and Ser(UCA) to Leu(UUA) at position 64. The third possible edited site, at position 57, does not involve a radical amino-acid change but would alter His(CAU) to

Tyr(UAU), which is found at this position in all non-plant mitochondrial *nad6* genes.

Seven flowering plant mitochondrial genes are known to contain introns, all 21 of which are group II introns. Three of these split genes (*nad1, nad2, nad5*) are fragmented into exons that are located far apart from each other on the mitochondrial genome and whose transcripts are joined by *trans*-splicing (Chapdelaine and Bonen 1991; Binder et al. 1992; Conklin et al. 1991; Wissinger et al. 1991; Knoop et al. 1991). In the case of *nad5*, *trans*-splicing integrates an exon of only 22 nucleotides into the mature transcript (Knoop et al. 1991). Direct sequencing of *nad6* mRNA or cDNA is needed to determine conclusively whether the gene contains any *cis*- or *trans*-spliced introns. However, we think it unlikely that the *B. campestris nad6* gene contains introns for several reasons. First, we have searched the *B. campestris nad6* sequence for conserved features of group II introns (Michel et al. 1989). The *nad6* sequence shows no similarity to the highly conserved domain V of group II introns. A single region (nucleotides 250–254; Fig. 1) is identical to a conserved sequence found at the 5′ splice site of the A1 subgroup of group II introns (Michel et al. 1989). However, the deduced protein sequence downstream from this region continues to show similarity to the other NAD6 proteins examined (Fig. 2), making it unlikely that splicing occurs at this region. Second, the deduced *B. campestris* protein sequence is reasonably coextensive with *Chlamydomonas* NAD6; there are no internal gaps of sufficient length to possibly constitute a *cis*-spliced group II (or group I) intron. Third, both proteins show considerable similarity at their C- and N-termini (Fig. 2); it is, therefore, highly unlikely that the *nad6* mRNA contains a short, terminal, *trans*-spliced exon that we would have missed. Fourth, the *B. campestris nad6* gene exhibits a very simple transcriptional pattern. In general, *trans*-spliced genes produce complex transcriptional patterns that reflect the complex organization of the genes (Chapdelaine and Bonen 1991; Wissinger et al. 1991). Fifth, no other region of the *B. campestris* mitochondrial genome hybridizes to a transcript (the presumptive mRNA) of the size produced at this *nad6* locus (Makaroff and Palmer 1987).

Since flowering plant mtDNAs are still far from being completely characterized with respect to gene content it remains to be seen whether additional subunits of complex I are coded for by these DNAs. The completely sequenced mitochondrial genome of *M. polymorpha* (186 kb in size; Oda et al. 1992) contains at least seven genes coding for subunits of NADH dehydrogenase. This genome also contains 32 open reading frames whose protein products are so far unidentified. It is possible that some of these ORFs may code for additional subunits of NADH dehydrogenase and that flowering plants may also contain additional, as yet undescribed, *nad* genes.

Acknowledgements. We thank M. Hanson, K. Newton and B. Sears for providing several of the mtDNAs used in Fig. 5 and L. Bonen and J. M. Grienenberger for providing unpublished data. This work was supported by NIH grant GM-35087 to J.D.P. and a Midwest Biotechnology Consortium Fellowship to J.M.N.

References

Attardi G, Schatz G (1988) Annu Rev Cell Biol 4:289–333
Berger S, Ellersiek U, Steinmuller K (1991) FEBS Lett 286:129–132
Binder S, Marchfelder A, Brennicke A, Wissinger B (1992) J Biol Chem 267:7615–7623
Bland MM, Levings III CS, Matzinger, DF (1986) Mol Gen Genet 204:8–16
Boer PH, Gray MW (1989) Nucleic Acids Res 17:3993
Brennicke A, Moller S, Blanz PA (1985) Mol Gen Genet 198:404–410
Chapdelaine Y, Bonen L (1991) Cell 65:465–472
Conklin PL, Wilson RK, Hanson MR (1991) Genes Dev 5:1407–1415
Covello PS, Gray MW (1989) Nature 341:662–666
Cummings DJ, Domenico JM (1988) J Mol Biol 204:815–839
Cummings DJ, McNally KL, Domenico JM, Matsuura ET (1990) Curr Genet 17:375–402
Dawson AJ, Jones VP, Leaver CJ (1984) EMBO J 3:2107–2113
Devereux J, Haeberli P, Smithies O (1984) Nucleic Acids Res 12:387–395
Desjardins P, Morais R (1990) J Mol Biol 212:599–634
deVries S, Grivell LA (1988) Eur J Biochem 176:377–384
Dupuis A, Skehel M, Walker JE (1991) Biochemistry 30:2954–2960
Feagin JE, Stuart K (1985) Proc Natl Acad Sci USA 82:3380–3384
Gass DA, Makaroff CA, Palmer JD (1992) Curr Genet 21:423–430
Gruissem W, Barkan A, Deng XW, Stern D (1988) Trends Genet 4:258–263
Gualberto JM, Wintz H, Weil JH, Grienenberger JM (1988) Mol Gen Genet 215:118–127
Gualberto J, Lamattina L, Bonnard G, Weil JH, Grienenberger JM (1989) Nature 341:660–662
Gualberto JM, Weil JH, Grienenberger JM (1990) Nucleic Acids Res 18:3771–3776
Hamby RK, Sims LE, Issel LE, Zimmer EA (1988) Plant Mol Biol Rep 6:175–192
Hiesel R, Wissinger B, Schuster W, Brennicke A (1989) Science 246:1632–1634
Jasmer DP, Feagin JE, Stuart K (1985) Mol Cell Biol 5:3041–3047
Knoop V, Schuster W, Wissinger B, Brennicke A (1991) EMBO J 10:3483–3493
Kolodner R, Tewari KK (1972) Proc Natl Acad Sci USA 69:1830–1834
Koslowsky DJ, Bhat GJ, Perrollaz AL, Feagin JE, Stuart K (1990) Cell 62:901–911
Lamattina L, Grienenberger JM (1991) Nucleic Acids Res 19:3275–3282
Lonsdale DM (1989) The plant mitochondrial genome. In: Marcus A (ed) The biochemistry of plants, vol 15, Academic Press, New York, pp 230–295
Macfarlane JL, Wahleithner JA, Wolstenholme DR (1990) Curr Genet 17:33–40
Makaroff CA, Palmer JD (1987) Nucleic Acids Res 15:5141–5156
Michaelis G, Vahrenholz C, Pratje E (1990) Mol Gen Genet 223:211–216
Michel F, Umesono F, Ozeki H (1989) Gene 82:5–30
Moller IM (1986) Physiol Plant 67:517–520
Newton KJ (1988) Annu Rev Plant Physiol Plant Mol Biol 39:503–532
Nugent JM, Palmer JD (1988) Curr Genet 14:501–509
Nugent JM, Palmer JD (1991) Cell 66:473–481
Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogur Y, Kohchi T, Ohyama K (1992) J Mol Biol 223:1–7
Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota SI, Inokuchi H, Ozeki H (1986) Nature 322:572–574
Payne M, Rothwell V, Jasmer DP, Feagin JE, Stuart K (1985) Mol Bio Parasitol 15:159–170
Pritchard AE, Seilhamer JJ, Mahalingam R, Sable CL, Venuti SE, Cummings DJ (1990) Nucleic Acids Res 18:173–180

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci USA 74:5463–5467
Schuster W, Hiesel R, Wissinger B, Brennicke A (1990a) Mol Cell Biol 10:2428–2431
Schuster W, Wissinger B, Unseld M, Brennicke A (1990b) EMBO J 9:263–269
Souza AE, Myler PJ, Stuart K (1992) Mol Cell Biol 12:2100–2107
Wahleithner JA, Wolstenholme DR (1988) Nucleic Acids Res 16:6897–6913
Wissinger B, Schuster W, Brennicke A (1991) Cell 65:473–482
Xue Y, Davies DR, Thomas CM (1990) Mol Gen Genet 221:195–198

Communicated by D. R. Wolstenholme