Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 548–552, January 1990
Botany

A sequence encoding a maturase-related protein in a group II intron of a plant mitochondrial *nad1* gene

(broad bean/reverse transcriptase/evolutionary conservation/sequence rearrangements)

JILL A. WAHLEITHNER*, JANE L. MACFARLANE, AND DAVID R. WOLSTENHOLME

Department of Biology, University of Utah, Salt Lake City, UT 84112

Communicated by Hewson Swift, October 16, 1989

ABSTRACT **We have determined from nucleotide sequence analysis that the subterminal and terminal exons of a respiratory-chain NADH dehydrogenase subunit I gene in broad bean mitochondrial DNA (mtDNA) are separated by a group II intron. Within this intron is a 687-codon open reading frame that, from considerations of similarity between amino acid sequences predicted from this open reading frame and maturase-coding sequences in group II introns of certain fungal mitochondrial genes, appears to encode a maturase-related protein. Transcripts complementary to this broad bean sequence (designated a *mat-r* gene) were detected among RNAs isolated from broad bean mitochondria. Data obtained from DNA·DNA hybridizations indicated that soybean and corn mtDNAs also contain a *mat-r* gene and suggested that only one copy of this gene occurs in each plant mtDNA. The putative protein specified by the broad bean *mat-r* gene contains amino acid sequences characteristic of reverse transcriptases. Because of this, consideration is given to the possibility that the maturase-related protein may be functional in the mechanisms by which plant mtDNA sequences are rearranged and foreign sequences are incorporated into plant mtDNAs.**

Plant mitochondrial DNAs (mtDNAs) contain the genes for a number of subunits of proteins involved in oxidative phosphorylation and genes for three ribosomal proteins (*rpS12*, *rpS13*, and *rpS14*) (1–3). Introns have been reported only in some plant *coxII* genes, in the corn and watermelon *nad1* genes, and in the *nad5* gene of *Oenothera* (4–7). Each of these introns is of the group II type found only in organelle genes (8, 9). Unlike the situation in yeast and other fungi, where some group II (and group I) introns encode enzymes (maturases) necessary for excision of the maturase-encoding intron from RNA transcripts of the gene (10–12), none of the aforementioned plant mitochondrial (mt) introns appears to encode a protein.

In this paper, we report the occurrence within a group II intron of a broad bean mt respiratory-chain NADH dehydrogenase subunit I (*nad1*) gene of a sequence that appears to encode a maturase-related protein.† Furthermore, we provide evidence that this broad bean sequence (designated a *mat-r* gene) is transcribed and that it has been evolutionarily conserved.

MATERIALS AND METHODS

The broad bean (*Vicia faba*) seed used in this work was Fava Broad Windsor (Musser Seed, Twin Falls, ID) and has been used by us previously (2, 13).

Details regarding isolation of mtDNA and mtRNA; restriction enzyme digestions; cloning into pUC8, pUC9, M13mp18, and M13mp19; preparation of M13 DNAs, and of libraries of *Eco*RI, *Hin*dIII, *Sal* I, *Bam*HI, and *Sal* I/*Eco*RI restriction fragments of total broad bean mtDNA; ^{32}P labeling of single-stranded probes derived from M13 mtDNA clones; DNA·RNA and DNA·DNA hybridizations; and colony hybridizations are given or referred to in refs. 1 and 13. DNA sequences were obtained from sets of deletion clones containing overlapping sequences representing the entire nucleotide sequences of both complementary strands of the fragment. Other sequencing details are given in ref. 13.

RESULTS

The nucleotide sequence of a 4481-nucleotide-pair (ntp) segment of broad bean mtDNA is given in Fig. 1 and our analysis of this sequence is summarized in Fig. 2. Three regions of the amino acid sequence predicted from a 2061-ntp (687 codons) open reading frame within this sequence show considerable similarity to corresponding consecutive regions of amino acid sequences of the maturase encoded by the aI1 intron of the cytochrome *c* oxidase subunit I (*coxI*) gene of *Saccharomyces cerevisiae* and by the following fungal putative maturase gene-containing mt intron sequences (Fig. 3): aI2 of the *coxI* gene of *S. cerevisiae* (14), I and IA of the *coxI* gene of *Podospora anserina* (15, 16), and bl of the cytochrome *b* (*cob*) gene of *Schizosaccharomyces pombe* (17). Because of these amino acid sequence similarities, we have designated the putative broad bean protein a maturase-related protein (MAT-R) and the sequence that encodes it a *mat-r* gene.

The broad bean *mat-r* gene, like the *S. cerevisiae*, *P. anserina*, and *Sc. pombe mat* genes, is contained in a group II intron of another gene—subunit 1 of a respiratory-chain NADH dehydrogenase (*nad1*). This conclusion is based on the following observations. Beginning 740 ntp upstream from the broad bean *mat-r* gene is a sequence of 78 nt (Fig. 1) that has primary and secondary sequence characteristics of domain 5 and domain 6 of group II introns (d5–d6 gpII; Fig. 4; refs. 8 and 9). Immediately following this upstream d5–d6 gpII sequence is a reading frame of 20 codons that has 50% similarity to amino acids 212–231 of mouse NAD1 (20) and 50% similarity to a corresponding region of *Drosophila yakuba* NAD1 (21).

Beginning 375 ntp downstream from the termination codon of the broad bean *mat-r* gene is a second d5–d6 gpII intron sequence but of 107 nt (Figs. 1, 2, and 4). Following this sequence is a reading frame of 85 codons that ends with a TAA and appears to be the terminal exon of a broad bean *nad1* gene: the mouse NAD1 ends 84 amino acids after the 20-amino acid sequence that corresponds to the broad bean putative *nad1* subterminal exon (Fig. 5); a sequence of 32 amino acids predicted from the broad bean 85-codon reading frame is 69% similar to a correspondingly located segment of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: nt, nucleotide(s); ntp, nucleotide pair(s); cpDNA, chloroplast DNA.

*Present address: Department of Biochemistry and Biophysics, University of California–Davis, Davis, CA 95616.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M30176).

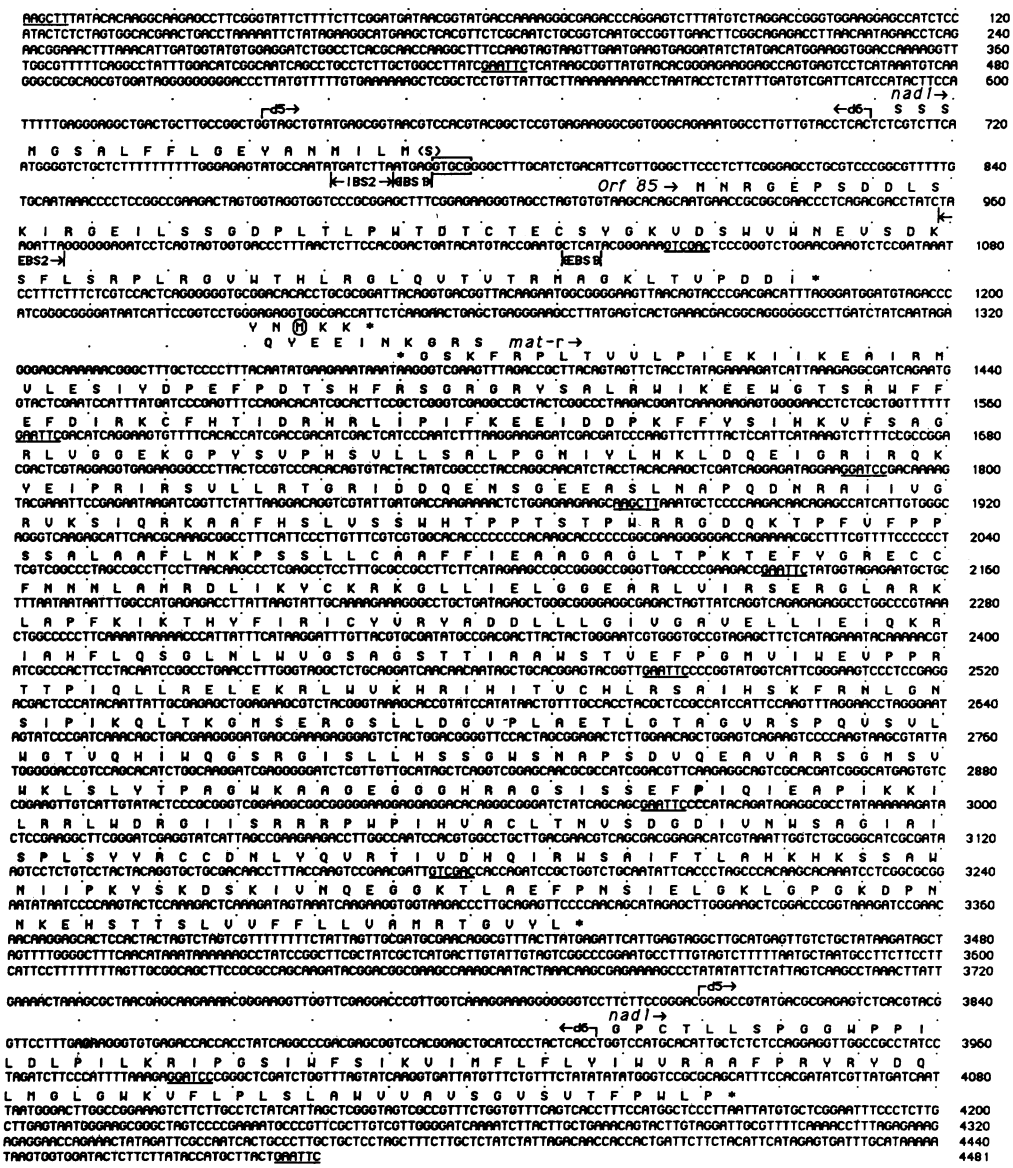

FIG. 1. The 4481-ntp *mat-r* gene-containing sequence of broad bean mtDNA. For the *nad1* exons, the *mat-r* gene, and an 85-codon open reading frame (*orf85*) the direction of transcription is indicated (arrows) and the predicted amino acid sequences (single-letter code) are shown. Partial amino acid sequences predicted for all three reading frames of nucleotides 1350–1381 are shown. The circled M identifies an ATG codon that might, as the result of frameshifting, act as a translation initiation codon for the *mat-r* gene. CGG codons are translated as tryptophan (4). An asterisk indicates a termination codon. The location of putative intron binding sequences (IBS1, IBS2), exon-binding sequences (EBS1, EBS2), and two domains 5 and 6 group II intron sequences (d5–d6) are indicated. Restriction sites (see Fig. 2) are underlined.

the aforementioned 84-amino acid sequence of mouse NAD1 (and 56% similar to the corresponding amino acid sequence of *D. yakuba* NAD1; Fig. 5).

In the broad bean sequence, following the 20-codon subterminal *nad1* exon sequence, and separated from it by two nucleotides, is the pentanucleotide GTGCG. This is the

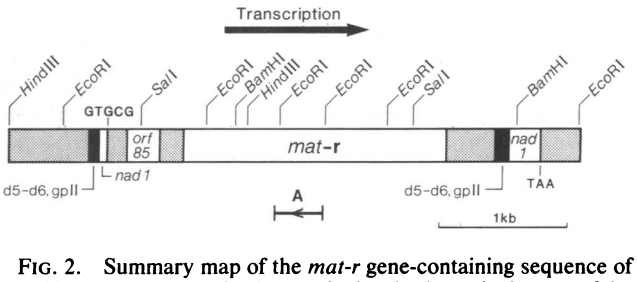

FIG. 2. Summary map of the *mat-r* gene-containing sequence of broad bean mtDNA. *nad1*, the terminal and subterminal exons of the *nad1* gene. d5–d6 gpII, domains 5 and 6 of group II intron sequences (see Fig. 4). *orf85*, an unidentified, 85-codon open reading frame. GTGCG, the putative 5′ terminus of the *mat-r* gene-containing group II intron. Stippled regions lack extensive open reading frames. The *mat-r*, *nad1*, and *orf85* sequences would be transcribed in the direction of the arrow. The location and orientation of a probe (A) used in hybridization experiments (see Figs. 6 and 7) are shown. Restriction sites used for sequencing are shown. kb, Kilobase.

sequence found at the 5′ terminus of the *S. cerevisiae* aI1 and aI2 introns, the *P. anserina* I and IA introns, and many other group II introns (9). Interpretation of this GTGCG sequence as the 5′ terminus of the broad bean *mat-r* gene-containing group II intron is further supported by the observation that cleavage of an RNA transcript immediately 5′ to the GUGCG sequence, and at the predicted location following the downstream d5–d6 gpII intron sequence (Figs. 1, 4, and 5), followed by splicing, would restore the *nad1* gene reading frame by forming an AGT codon (Fig. 5). Also, there occur in the broad bean sequence sets of complementary oligonucleotide pairs whose locations relative to each other, to the GTGCG sequence, and to sequences with secondary structure potential are consistent with them being equivalent to the EBS–IBS (exon binding site–intron binding site) sequences identified in other exon–group II introns (9): that is, immediately preceding the GTGCG sequence are two adjacent sequences of 5 and 8 nt that are complementary to sequences beginning 253 and 181 nt, respectively, downstream from the GTGCG sequence (Fig. 1).

We have been unable to identify the amino acid sequence specified by the 85-codon open reading frame (*orf85*) that is located upstream from the broad bean *mat-r* gene and within the group II intron (Figs. 1 and 2). Also, within the 631-ntp sequence 5′ to the upstream d5–d6 gpII intron sequence (Fig.

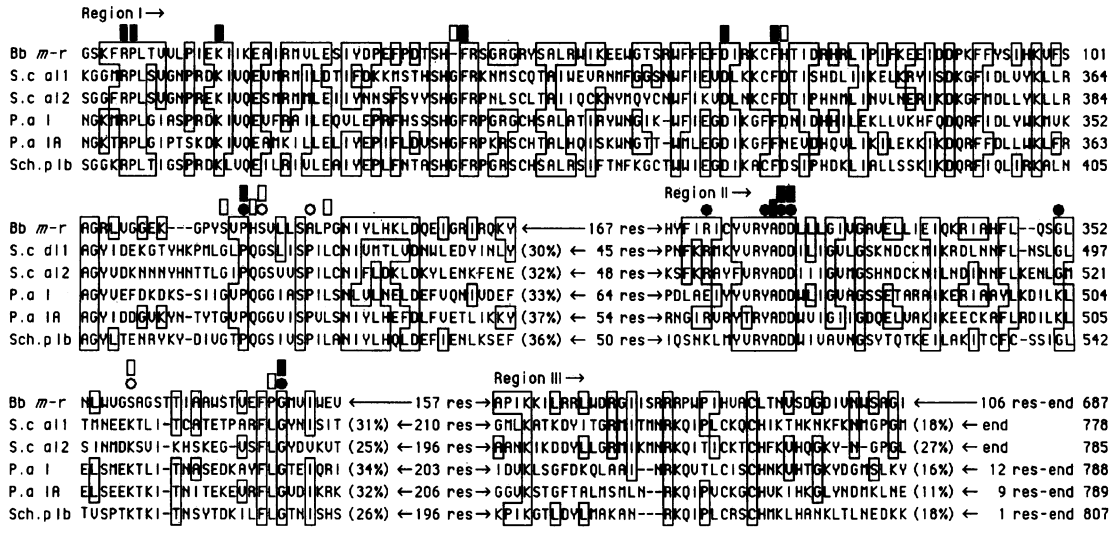

FIG. 3. Comparisons of three consecutive regions (I–III) of predicted amino acid sequences (single-letter code) of the broad bean maturase-related gene (Bb *m-r*) and the maturase genes of introns of the *S. cerevisiae coxI* gene (S.c aI1 and S.c aI2; ref. 14) (but ATA = methionine; ref. 15), the *P. anserina coxI* gene (P.a I and P.a IA; refs. 15 and 16) and the *Sc. pombe cob* gene (Sch.p Ib; ref. 17). Residues identical to those at corresponding locations in the broad bean sequence are boxed. In the broad bean and fungal amino acid sequences that separate regions I, II, and III, there is low similarity. Numbers in parentheses indicate percentage similarity to broad bean of amino acid sequences of the preceding region. The first residues in the S.c aI1, S.c aI2, P.a I, P.a IA, and Sch.p Ib sequences shown are nos. 263, 283, 252, 263, and 304 of the respective maturase. Numbers on the right identify the amino acid locations in the sequences and the relative end location of each sequence. Of the 10 amino acids that are highly conserved among all known and putative reverse transcriptases (18), the locations of those present and absent in the broad bean sequence are indicated by solid and open circles, respectively. Of 19 amino acids that are highly conserved in a subset of reverse transcriptase-like sequences designated non-long terminal repeat retrotransposons (19) that include the fungal group II maturases, the locations of those present and absent in the broad bean sequence are indicated by solid and open rectangles, respectively.

1), we have been unable to locate any other potential exon of the *nad1* gene.

In contrast to the situation in broad bean mtDNA where 598 ntp separate the subterminal *nad1* gene exon and the *mat-r* gene, the *S. cerevisiae, P. anserina*, and *Sc. pombe mat* gene sequences contained in group II introns are each an in-phase continuation of the preceding (*coxI* or *cob*) exon (10, 14–17). None of these fungal sequences begins with an ATG codon, and the location of translation initiation has not been defined in the excised introns. Also in broad bean, there is ambiguity regarding the triplet at which translation of the *mat-r* gene transcripts might be initiated. The 2061-ntp open reading frame begins with a GGG (glycine) triplet, which is preceded by TAA. An ATG is located 23 codons downstream from the beginning of the open reading frame of the broad bean *mat-r* gene. However, as this is well within the first region of amino acid sequence similarity between the broad bean *mat-r* gene and fungal *mat* genes (Fig. 3), it seems doubtful that this ATG acts as the translation initiation codon. Translation initiation of the *mat-r* gene could involve an unorthodox codon as occurs in many animal *mt* protein genes (see ref. 22). Alternatively, a MAT-R protein beginning with a methionine might be produced by frame-shifting during reading of the RNA transcript, but this would require two separate −1 nucleotide shifting events within the 13 ntp upstream from the *mat-r* gene (Fig. 1). A MAT-R protein beginning with a methionine could also be generated if in the *mat-r* gene transcript there occurred appropriate adjustments in the reading frame by ribonucleotide addition (RNA editing; ref. 23) or if the *mat-r* gene transcript acquired an AUG initiation codon by the splicing to its 5′ end of a short segment of an upstream sequence.

The putative MAT-R protein, like the fungal group II maturases, contains amino acids (7 of 10; Fig. 3) that are highly conserved in all known and putative reverse transcriptases (16–18, 24). The specific locations of other amino acids in the MAT-R protein, together with the amino acid sequence similarities of this protein to the fungal group II maturases (Fig. 3), suggest that it belongs to the recently defined non-long terminal repeat retrotransposon class of reverse transcriptase-like proteins (19).

The predicted amino acid sequence of the broad bean *mat-r* gene, like those of the fungal maturases, is rich in basic residues (Arg, 53; Lys, 41; His, 22; compared to Asp, 23; Glu, 37), as might be expected for a protein whose function requires an intimate association with nucleic acids.

The G+C contents of the broad bean *mat-r* gene (51.2%) and of remaining portions of the group II intron (51.0%) are higher than those of the broad bean *nad1* gene exons (45.2%) and five other broad bean mt protein genes (*cob, coxIII, atp6,*

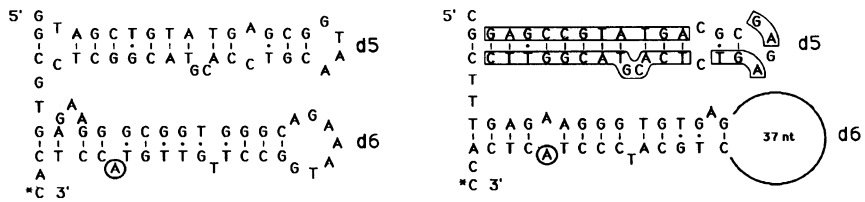

FIG. 4. The domain 5 (d5)–domain 6 (d6) group II intron sequences (9) of the broad bean *mat-r* gene-containing intron (*Right*; the boxed sequences are those conserved in the *S. cerevisiae* aI2 intron) and the intron located 5′ to the subterminal *nad1* gene exon (*Left*). By analogy to other group II introns (9), the circled A in each sequence is the nucleotide to which the 5′ end of the intron is expected to join during splicing and the *C in each sequence is the nucleotide following which intron–exon cleavage is expected to occur.

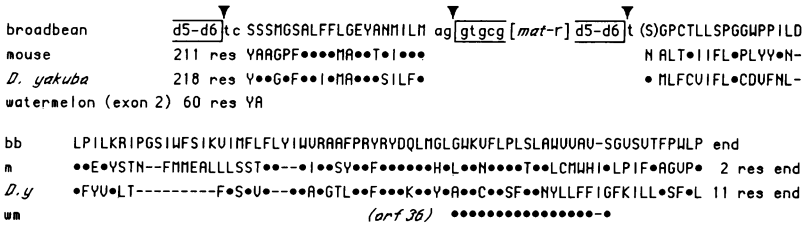

FIG. 5. Identification of the terminal and subterminal exons of the broad bean *nad1* gene. Amino acid sequences (single-letter code) of the broad bean subterminal *nad1* gene exon [predicted from the sequence between the upstream d5–d6 group II intron sequence and the 5′ end of the *mat-r* gene-containing group II intron (gtgcg, boxed)], and of the terminal exon (predicted from the sequence downstream from the d5–d6 sequence of the *mat-r* gene-containing group II intron). Nucleotides are represented by lowercase letters. The serine residue predicted from the AGT codon reconstituted by splicing (inverted arrowheads) is shown as (S). The broad bean amino acid sequences are aligned with the C-terminus proximal 104 and 106 amino acids of mouse NAD1 and *D. yakuba* NAD1, respectively (21). Also shown are the relative locations of amino acids predicted from the exon 2 and *orf36* sequences of the watermelon *nad1* gene (6).

atp9, *rpS14*; mean, 40.9%; range, 36.2–43.7%) (refs. 2 and 25; unpublished observations). Also, in the broad bean *mat-r* gene, 55.8% of codons end in G or C compared to 42.5% in the *nad1* exons and 30.7–36.7% (mean, 33.9%) in the above-mentioned five broad bean mt genes. These observations suggest that the *mat-r* gene-containing intron may have had a separate origin than other portions of the broad bean mt genome. A similar suggestion has been made regarding fungal *mat* gene-containing sequences (17) and is consistent with the idea that group II introns originated as infectious elements (26).

To test whether the broad bean *mat-r* gene is transcribed, a ^{32}P-labeled single-stranded DNA probe comprising a 345-nt internal segment (probe A, Fig. 2) of the *mat-r* gene was hybridized to electrophoretically fractionated whole broad bean mtRNA. In the autoradiograph of the hybridization product, two bands were discernable at positions expected for linear RNA molecules of 3800 and 2500 nt (Fig. 6). Linear molecules of this size would each be large enough to contain a transcript of the coding region of the *mat-r* gene, but only the larger molecule could contain the entire intron.

The same ^{32}P-labeled DNA probe that was used to detect RNA transcripts of the broad bean *mat-r* gene (probe A, Fig. 2) was hybridized to blots of electrophoretically separated *Eco*RI, *Hin*dIII, and *Bam*HI fragments of mtDNAs of broad bean, soybean, and corn and the products were autoradiographed (Fig. 7). A single labeled band of 345-ntp fragments was observed in the lanes containing *Eco*RI fragments of mtDNA of each of the three species. Also, only a single labeled band was observed in each of the lanes that contained *Hin*dIII or *Bam*HI digests of the broad bean, soybean, and corn mtDNAs. These data are consistent with the interpretation that mtDNA of each of the three species contains a single *mat-r* gene.

DISCUSSION

The observations that there appears to be only one copy of the *mat-r* gene in broad bean mtDNA, that this gene is transcribed, and that sequences homologous to the *mat-r* gene sequence are found in soybean and corn mtDNAs strongly support the view that the *mat-r* gene is functional and has been conserved during plant evolution.

The amino acid sequence predicted from the *mat-r* gene is considerably more similar to the amino acid sequences of known and putative fungal group II intron-encoded maturases than to any other published amino acid sequence. However, as colinearity between the amino acid sequences of the broad bean *mat-r* gene and the fungal group II maturase genes is only partial (Fig. 3), it remains an open question as to whether the broad bean putative MAT-R protein might be involved in RNA splicing.

Sequence characteristics of the putative broad bean MAT-R protein suggest that it may function as a reverse transcriptase. Reverse transcriptase activity in mitochondria has only

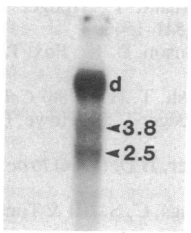

FIG. 6. Autoradiograph of a Northern blot experiment to detect transcripts of the broad bean *mat-r* gene. The lane contains whole broad bean mtRNA fractionated by electrophoresis through a 1.5% agarose gel containing 2.2 M formaldehyde. The blot of the gel was probed with a ^{32}P-labeled M13 single-stranded clone containing a 345-nt sequence complementary to an internal section of the *mat-r* gene sense strand (probe A in Fig. 2; ntp 2138–2482 in Fig. 1). The sizes (in kilobases) of fragments contained in the two bands (determined using an RNA ladder of 0.24–9.5 kilobases) are indicated. d, Contaminating mtDNA present in this RNA preparation.

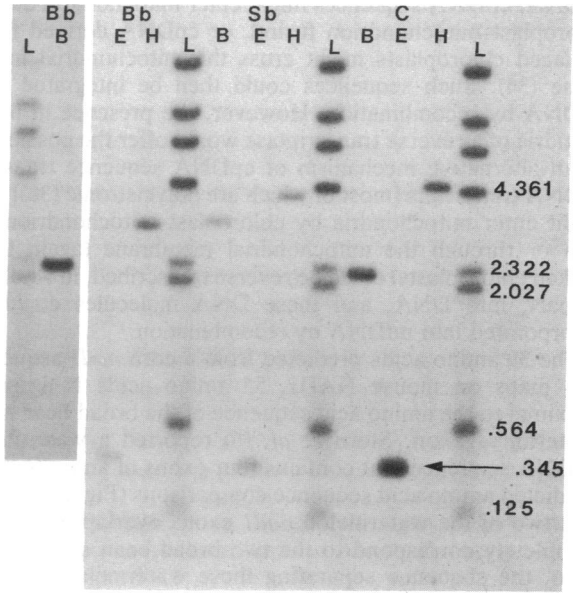

FIG. 7. Autoradiographs of Southern blot experiments to determine the distribution of *mat-r* gene-homologous sequences in broad bean (Bb), soybean (Sb), and corn (C) mtDNAs. Restriction enzyme-digested DNAs were electrophoresed through a 1% agarose gel and a blot of the gel was probed with a ^{32}P-labeled internal 345-ntp sequence of the broad bean *mat-r* gene (probe A in Fig. 1). Lanes B, E, and H contain *Bam*HI, *Eco*RI, and *Hin*dIII digestion products, respectively. The numbers indicate sizes (in kilobases) of ^{32}P-labeled, *Hin*dIII digestion products of phage λ in lanes L and the mtDNA 345-ntp *Eco*RI fragment (arrow).

been demonstrated in *Neurospora crassa*, where it is associated with the presence of one of two closely related plasmids that encode a protein with the amino acid sequence characteristics of reverse transcriptases (27). Schuster and Brennicke (28) reported a putative reverse transcriptase gene in *Oenothera* mtDNA. This gene is much smaller (142 codons) and has only the sequence Tyr-Xaa-Asp-Asp-Leu in common with the broad bean and fungal maturases. Also, the *Oenothera* gene is bounded by sequences of nuclear and chloroplast origin and it has not been detected in other plant mtDNAs. Nevertheless, the occurrence of this sequence in *Oenothera* mtDNA raises the possibility that some plant mtDNAs contain at least two distinct reverse transcriptases.

Reverse transcriptase activity may be involved in the mechanism of group II intron excision (see ref. 26). However, a reverse transcriptase could also be functional in recombination events that occur in plant mtDNAs (29, 30), including exchanges that result in both sequence reorganizations and sequence duplications (see ref. 31). Either the reverse transcriptase or a separate activity encoded by the *mat-r* gene could be involved in the breaking and rejoining of DNA strands. Relevant to this suggestion are the findings that the reverse transcriptase-like retrotransposon R2Bm of *Bombyx mori* encodes a sequence-specific endonuclease (32) and evidence that the bI4, group I intron-encoded maturase of *S. cerevisiae cob* gene is involved in recombination (12). Alternatively, in plant mitochondria, a reverse transcriptase might copy RNAs transcribed from various segments of the mt genome, and the DNA products might then be incorporated by recombination at different locations in the mtDNA.

The mt genomes of many dicotyledons and monocotyledons contain sequences of various sizes and gene content derived from chloroplast DNA (cpDNA) (33–35). Because, in contrast to plant mtDNA, chloroplast genome size has been highly conserved during evolution, it has been argued that these sequence exchanges occurred in the direction from cpDNA to mtDNA. cpDNA sequences might enter mitochondria during chloroplast–mitochondrion fusion, or cpDNA derived from damaged chloroplasts might cross the mitochondrial membrane (34). Such sequences could then be integrated into mtDNA by recombination. However, the presence in mitochondria of a reverse transcriptase would offer the possibility of an alternative mechanism of cpDNA sequence transfer. cpDNA transcripts [most of which are polycistronic (36)] that might enter mitochondria by chloroplast–mitochondrion fusion or through the mitochondrial membrane (again from broken chloroplasts) could be reverse transcribed, in whole or in part, into DNA, and these DNA molecules could be incorporated into mtDNA by recombination.

The 39 amino acids predicted from a corn *nad1* sequence (37) maps on mouse NAD1, 55 amino acids N-terminus proximal to the amino acid sequence of the broad bean *nad1* subterminal exon. Stern *et al.* (6) reported a watermelon mtDNA sequence that contains four exons of an *nad1* gene. Predicted amino acid sequence comparisons (Fig. 5) indicate that two of the watermelon *nad1* exons overlap but do not completely correspond to the two broad bean *nad1* exons. Also, the sequence separating these watermelon exons is only 1357 ntp (3374 ntp in broad bean) and lacks a *mat-r* gene. These observations suggest either that each of the plant mtDNAs examined contains two *nad1* genes that differ in regard to exon–intron structure or that each plant mtDNA contains a single *nad1* gene, but exon–intron structure of this gene differs between species.

We thank Kirk Thomas for oligonucleotides. This work was supported by National Institutes of Health Grants GM18375 and RR07092 and Genetics Training Program GM07464 (J.A.W.).

1. Lonsdale, D. M. (1988) *Plant Mol. Biol. Rep.* **6,** 266–273.
2. Wahleithner, J. A. & Wolstenholme, D. R. (1988) *Nucleic Acids Res.* **16,** 6897–6913.
3. Gualberto, J. M., Wintz, J. H. & Grienenberger, J.-M. (1988) *Mol. Gen. Genet.* **215,** 118–127.
4. Fox, T. D. & Leaver, C. J. (1981) *Cell* **26,** 315–323.
5. Hiesel, R. & Brennicke, A. (1983) *EMBO J.* **2,** 2173–2178.
6. Stern, D. B., Bang, A. G. & Thompson, W. F. (1986) *Curr. Genet.* **10,** 857–869.
7. Wissinger, B., Hiesel, R., Schuster, W. & Brennicke, A. (1988) *Mol. Gen. Genet.* **202,** 56–65.
8. Michel, F. & Dujon, B. (1983) *EMBO J.* **2,** 33–38.
9. Michel, F. & Jacquier, A. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** 201–212.
10. Carignani, G., Netter, P., Bergantino, E. & Robineau, S. (1986) *Curr. Genet.* **11,** 55–63.
11. Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E. & Slonimski, P. P. (1983) *Cell* **35,** 733–742.
12. Kotylak, Z., Lazowska, J., Hawthorne, D. C. & Slonimski, P. P. (1985) *Achievements and Perspectives of Mitochondrial Research: Biogenesis*, eds. Quagliariello, E., Slater, E. C., Palmieri, F., Saccone, C. & Kroon, A. M. (Elsevier Science, Amsterdam), Vol. 2, pp. 1–20.
13. Wahleithner, J. A. & Wolstenholme, D. R. (1987) *Curr. Genet.* **12,** 55–67.
14. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E. & Tzagoloff, A. (1980) *J. Biol. Chem.* **255,** 11927–11941.
15. Osiewacz, H. D. & Esser, K. (1984) *Curr. Genet.* **8,** 299–305.
16. Matsuura, E. T., Domenico, J. M. & Cummings, D. J. (1986) *Curr. Genet.* **10,** 915–922.
17. Lang, B. F., Ahne, F. & Bonen, L. (1985) *J. Mol. Biol.* **184,** 353–366.
18. Toh, H., Hayashida, H. & Miyata, T. (1983) *Nature (London)* **305,** 827–829.
19. Xiong, Y. & Eickbush, T. H. (1988) *Mol. Biol. Evol.* **5,** 675–690.
20. Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. (1981) *Cell* **26,** 167–180.
21. Clary, D. O., Wahleithner, J. A. & Wolstenholme, D. R. (1984) *Nucleic Acids Res.* **12,** 3747–3762.
22. Wolstenholme, D. R., Macfarlane, J. L., Okimoto, R., Clary, D. O. & Wahleithner, J. A. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1324–1328.
23. Benne, R., Van den Burg, J., Brakenhoff, J., Sloof, P., Van Broom, J. A. & Tromp, M. C. (1986) *Cell* **46,** 819–826.
24. Michel, F. & Lang, B. F. (1985) *Nature (London)* **316,** 641–643.
25. Macfarlane, J. L., Wahleithner, J. A. & Wolstenholme, D. R. (1989) *Curr. Genet.*, in press.
26. Lambowitz, A. M. (1989) *Cell* **56,** 323–326.
27. Kuiper, M. T. N. & Lambowitz, A. M. (1988) *Cell* **55,** 693–704.
28. Schuster, W. & Brennicke, A. (1987) *EMBO J.* **6,** 2857–2863.
29. Rothenberg, M., Boeshore, M. L., Hanson, M. R. & Izhar, S. (1985) *Curr. Genet.* **9,** 615–618.
30. Rottman, W. H., Brears, T., Hodge, T. P. & Lonsdale, D. M. (1987) *EMBO J.* **6,** 1541–1546.
31. Bailey-Serres, J., Hanson, D. K., Fox, T. D. & Leaver, C. J. (1986) *Cell* **47,** 567–576.
32. Xiong, Y. & Eickbush, T. H. (1988) *Cell* **55,** 235–246.
33. Lonsdale, D. M., Hodge, T. P., Howe, C. J. & Stern, D. B. (1983) *Cell* **34,** 1007–1014.
34. Stern, D. B. & Palmer, J. D. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1946–1950.
35. Dewey, R. E., Levings, C. S., III & Timothy, D. H. (1986) *Cell* **44,** 439–449.
36. Woodbury, N. W., Roberts, L. L., Palmer, J. D. & Thompson, W. F. (1988) *Curr. Genet.* **14,** 75–89.
37. Bland, M. M., Levings, C. S., III & Matzinger, D. F. (1986) *Mol. Gen. Genet.* **204,** 8–16.