A novel soybean mitochondrial transcript resulting from a DNA rearrangement involving the 5S rRNA gene

Peter H.Morgens, Elizabeth A.Grabau and Raymond F.Gesteland

Howard Hughes Medical Institute, Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

Received 23 April 1984; Accepted 18 June 1984

ABSTRACT

A tissue culture line of soybean cells (SB-1) is shown to have an unusual arrangement of mitochondrial genes. Fusion of the mitochondrial 5S gene to the distal end of an unidentified gene results in the abundant expression of an 800 nucleotide RNA with the 5S rRNA at its 5' end. Since only the fused 5S rRNA gene is found in this cell line, the functional 5S rRNA must arise by processing of the 800 nucleotide RNA or by intermittent termination of transcription. The 800 nucleotide transcript and its DNA arrangement are not detected at comparable levels in other soybean sources, including the parent plant of the SB-1 tissue culture line. The role of recombination in the origin of this gene fusion and in plant mitochondrial DNA in general is discussed.

INTRODUCTION

The sequence complexity of plant mitochondrial DNA is highly variable from one species to another, ranging from 170 (1) to 2500 kb (2) and is considerably greater than that found for other organisms. DNA isolated from plant mitochondria is generally a mixture of linear and circular subgenomic sized molecules whose proportions vary with species and culture conditions (for reviews see Leaver and Gray (3) and Levings (4)). These complications make it difficult to define the plant mitochondrial genome.

Lonsdale et al. (5) have examined the structure of the maize mitochondrial genome by restriction mapping and they suggest that the DNA undergoes recombination at specific sites to give a variety of circular molecules. This is strongly supported by the recent work of Palmer and Shields (6) which shows that the complete *Brassica campestris* mitochondrial genome can be represented by a single "master" circle and that two subgenomic circles could arise by recombination between 2 kb repeat sequences.

Dale and colleagues have demonstrated the presence of various sized circular molecules in plant tissue culture cells, some of which share homology with each other (7,8). The ability of plant mitochondrial DNA (mtDNA) to recombine has been demonstrated in cytoplasmic hybrids where plant cells with genetically different cytoplasms have been fused and non-parental restriction fragments are found in the mtDNA of the resulting hybrids (9,10,11).

There are numerous examples of DNA rearrangements in plant mtDNA, but what functional role, if any, these might play in plant development is far from clear. Sederoff et al. (12) have proposed that mtDNA rearrangements have been involved in the evolution of the genus Zea. Changes in mitochondrial restriction patterns have been found following selection for resistance to a pathotoxin in tissue culture (13), in association with cytoplasmic male sterility (14,15), and in plants regenerated from tissue culture cells without apparent selection (16,17). However, in a direct search for a possible developmental role for mtDNA rearrangements in several plants, tissue specific differences in mtDNA restriction fragment patterns were not found (18).

Soybean mtDNA is estimated to have a complexity in the range of 240-400 kb and includes at least some circular molecules (19,20,21). Our initial examination of the organization and expression of soybean mitochondrial genome has revealed the presence of a unique 800 nucleotide transcript resulting from a DNA rearrangement in the tissue culture line SB-1.

MATERIALS AND METHODS

Plant Material

The soybean SB-1 tissue culture line was originally isolated from the roots of Glycine max L. Merr., cultivar 'Mandarin' (22). SB-1 and other soybean tissue culture lines, fr₁, fr₂ and fr₂ x fr₁ (23) were maintained in B5C media. Cultivars 'Mandarin' and 'Wayne' were used for hypocotyl preparations.

Nucleic Acid Purification

Total cell RNA was prepared from plant or tissue culture cells by hot (65°C) phenol:chloroform:isoamyl extraction in a Waring blender (24). RNA was enriched for poly(A)⁺ sequences

by oligo-dT cellulose chromatography as described (25).

Whole SB-1 cell DNA was prepared by a modified-Marmur extraction procedure (26). MtDNA was isolated by the method of Sparks and Dale (27). DNA fragments were purified either from SeaPlaque (Marine Colloids) agarose (28) or by electro-elution from SeaChem agarose or acrylamide (29).

Gel Transfer and Hybridization

RNA was fractionated by electrophoresis in methylmercury hydroxide agarose gels (30), transferred to diazobenzylloxymethyl-paper (DBM-paper), and hybridized with ³²P labeled DNA as described (31). DNA was fractionated by electrophoresis in agarose gels, transferred to nitrocellulose (32), and hybridized with DNA labeled by nick translation (33).

Cloning and Sequencing

A partial EcoRI library from the soybean tissue culture line SB-1 was constructed in the bacteriophage lambda vector Charon 4A (24). The DNA library was screened with ³²P labeled cDNA made by reverse transcription of total RNA from the same cells using oligo-dT as a primer (34). The plasmid pSB2-2 was constructed by subcloning a 1.8 kb EcoRI-SalI fragment from the lambda mitochondrial clone SB4 into the EcoRI-SalI region of pBR322. Small fragments from pSB2-2 (see Fig. 2) were subcloned into M13 mp9 (35). DNA sequencing was performed by the method of Maxam and Gilbert (36).

Mapping and Sequence Analysis

End labeled DNA fragments were used as hybridization probes in S1 nuclease experiments performed as described (37,38). Restriction enzymes (Bethesda Research Laboratories or New England Biolabs) were used under conditions recommended by the manufacturers. Analyses of DNA sequences were performed using the SEQ program (39) on the DEC-20 computer available through the College of Science, University of Utah. The program NULIGN, a modification by Eric Kofoed of a program by Wilbur and Lipman (40) was used to generate "bestfit" comparison of sequences shown in Fig. 3.

RESULTS

Isolation of a highly expressed mitochondrial sequence

In a search for an abundantly expressed gene in soybean

FIGURE 1. Preliminary characterization of SB-1 mitochondrial sequence. (A) Northern analysis of SB-1 tissue culture cell RNAs. 5 ug RNA from SB-1 soybean tissue culture cells was loaded per lane. Lanes 1 and 2: ethidium bromide staining patterns of DNase treated total RNA and poly(A)+ RNA. Lanes 3 and 4: corresponding autoradiographs from hybridizations of those RNAs with labeled lambda SB4 DNA. Lanes 5 and 6: autoradiographs of total RNA and poly(A)+ RNA hybridized with labeled mitochondrial DNA from SB-1 cells. Lanes 7 and 8: autoradiographs of hybridization of those same RNAs with labeled pSB2-2 DNA. (B) Southern analysis of total cell DNA and mtDNA. Lane 9: ethidium bromide staining pattern of EcoRI digested SB4 DNA. Lanes 10 and 11: ethidium bromide staining patterns of 30 ug of total SB-1 DNA digested with EcoRI and corresponding autoradiograph following hybridization with labeled SB4 DNA. Lanes 12 and 13: ethidium bromide staining patterns of 1.5 ug of EcoRI digested mtDNA from SB-1 cells and the corresponding autoradiograph following hybridization with labeled SB4 DNA. Lane 14 is the autoradiograph from 2 ug of EcoRI digested SB-1 mtDNA following hybridization with labeled pSB2-2.

tissue culture cells, a partial EcoRI library was constructed in the vector Charon 4A with DNA from the SB-1 tissue culture line (see materials and methods for cloning and screening). One particular clone, designated SB4 (Fig. 2a), is mitochondrial in origin and specifies a number of transcripts including one of 800 nucleotides which is the subject of this study. The plasmid pSB2-2 is a subclone of an 1800 base pair EcoRI-Sall fragment from SB4 (Fig. 2b). Both SB4 and pSB2-2 were labeled by nick-translation (33) for use as hybridization probes for Northern analysis of soybean tissue culture cell RNA; the results

are shown in Figure 1A. Two major transcripts were observed, an 800 nucleotide transcript that is enriched by selection on oligo-dT cellulose and a 2000 nucleotide RNA which is not (see Fig. 1, lanes 4,6,8 vs. 3,5,7).

The possibility that the cloned DNA sequence was mitochondrial in origin was suggested by the high frequency of isolation of SB4 from the library and by the realization that the abundant 2000 nucleotide RNA, which apparently lacks poly(A) sequences, is the appropriate size for the 18S RNA of the small subunit of mitochondrial ribosomes. This hypothesis was supported by the observation that the same two transcripts are among those seen by hybridization using purified mitochondrial DNA as probe (Fig. 1, lanes 5 and 6). By hybridization analyses and density measurements we have ruled out the possibility that this is chloroplast DNA (data not shown). The mitochondrial origin of the insert was confirmed by hybridization of labelled SB4 DNA to the same set of fragments from both total cellular DNA and mtDNA digested with EcoRI (Fig. 1, lanes 11 and 13). Unexpectedly, however, hybridization to five EcoRI fragments was observed rather than the two fragments predicted from the restriction map of SB4 DNA (Fig. 2a). This suggests that at least part of the sequence in the SB4 insert is present in more than one location in the SB-1 mitochondrial genome. The fragment in plasmid pSB2-2 also contains at least part of the sequence found in multiple locations because it hybridizes to two EcoRI fragments from mtDNA (Fig. 1, lane 14) rather than the one predicted from its restriction map (Fig. 2b). We conclude that the cloned 1800 bp fragment in pSB2-2 is mitochondrial in origin, contains at least part of the sequence for two abundant transcripts of 800 and 2000 nucleotides and contains sequences that are found in more than one location in the SB-1 mitochondrial genome. The 800 nucleotide RNA appears to be polyadenylated, suggesting that it might be a mitochondrial mRNA. Because it is so abundant, this RNA was of particular interest and the cloned region hybridizing to this transcript was sequenced.

Sequencing of the insert in pSB2-2

The 1800 bp EcoRI-SalI insert in pSB2-2 was sequenced (36) according to the strategy shown in Figure 2. The sequence and

FIGURE 2. Partial characterization of lambda SB4 sequence. a. Restriction map of the mitochondrial DNA insert of lambda SB4. b. Restriction map of the pSB2-2 subclone insert. c. Transcription map of the pSB2-2 mitochondrial insert based on S1 nuclease analysis as seen in figure 4. d. Probes A,B,C (which were subcloned into M13mp9) and D (which was gel purified) used in hybridizations. e. Sequencing strategy for the insert of pSB2-2. f. Scale in base pairs.

the deduced locations of transcripts (see next section for S1 nuclease experiments) are shown in Figure 3.

The hypothesis that the 2000 nucleotide RNA is the mitochondrial 18S rRNA is supported by its near identity to known sequences from wheat (41) and maize (42), and strong homology with *E. coli* 16S rRNA (43). Adjacent to the 18S rRNA gene is

the sequence for the 5S rRNA which shows a very high degree of homology with the known sequences of both wheat (44) and maize (42). Two of the three differences between soybean and the other plants would not alter the secondary structure proposed by Spencer et al. (44) because these regions are predicted to lie in loop structures that are not involved in secondary structure base pairing. The third difference is a single nucleotide substitution. The intergenic region between the 18S and 5S rRNA genes of maize (108 bp) and soybean (150 bp) have diverged considerably. The soybean sequence lacks the sequence TATGTA (at position 825) proposed by Chao et al. (42) as the promoter for the 5S gene. There is, however, a region of strong homology (26 of 30 bases identical) just 5' to the 5S rRNA gene sequence (position 866 to 895).

As will be shown in the next section the 800 nucleotide transcript shares a 5' end with the 5S rRNA making this a highly unusual arrangement. An open reading frame within the sequence for the 800 nucleotide transcript could encode a protein of 170 amino acids but there is no ATG initiation codon near the beginning of the open reading frame. Our attempts to demonstrate in vitro translation of the 800 nucleotide transcript failed. However, it would be premature to make much of this negative result until the codon identities in plant mitochondria are determined. Nevertheless, the predicted amino acid sequence was compared with a large collection of protein sequences but no convincing homologies were found (Russell Doolittle, personal communication).

Mapping of transcripts

To determine the precise locations of transcripts on the pSB2-2 insert, S1 nuclease mapping techniques were used (24,37,38). From the data in Figure 4 we conclude that the region specifying the 3' end of the 2000 nucleotide RNA extends 742 bp from the left end of the cloned fragment (Fig. 4A, lanes 2,4,6 and 4B, lines a,b,c,f and g). The "800" nucleotide RNA spans the region from 900 to 1650 (Fig. 4A, lanes 4,6 and 4B, lines b,c,h,i and j) with the same 5' to 3' orientation as the sequences for the 2000 nucleotide RNA. The sequence for the "800" nucleotide RNA comprises only 750 bp in the DNA. The size estimate of

800 nucleotides based on electrophoretic mobility of the RNA (above) could be inaccurate or more likely, a poly(A) tail could contribute the extra 50 nucleotides. Although we have not directly demonstrated the presence of a poly(A) tail on the 800 nucleotide transcript, the apparent size difference noted here and the fact that the transcript is enriched on oligo-dT cellulose and does not contain internal stretches of A's longer than 6 or 7 nucleotides are consistent with its presence. The results depicted in Figure 4 also reveal the 3' end of a transcript at position 1020 that is a 120 nucleotide 5S rRNA with its 5' end at position 900 (Fig. 4A, lanes 12 and 17-19 and 4B, lines d and e). Note that the 5' ends of the 120 nucleotide RNA and the 800 nucleotide RNA appear to be identical (Fig. 4B, lines e and h).

Multiple sequence environments in mtDNA

The soybean 5S rRNA gene identified here is clearly unusual since it is located in a region that can also be expressed to give an 800 nucleotide RNA that is apparently polyadenylated. Expression of 5S rRNA from this gene would necessitate either processing from the 800 nucleotide RNA or frequent but not total termination of transcription after the 5S sequence. An alternative is that there is another functional gene for 5S rRNA that has different downstream sequences so that it does not express an 800 nucleotide RNA. This possibility was considered in light of the fact that at least some portion of the cloned sequence is present in more than one environment in the mitochondrial genome. To determine which regions of the insert in pSB2-2 are present in multiple locations, fragments specific for 5S sequences (probe A) and for downstream sequences within the 800 nucleotide region (probe B) were subcloned for use as hybridization probes (see Fig. 2d) to SB-1 mitochondrial DNA. The

FIGURE 3. Sequence of the pSB2-2 insert. The sequence of the 1.8 kb EcoRI-SalI insert from pSB2-2 is shown and compared to published sequences for part of the *E. coli* 16S rRNA gene (50), part of the maize mitochondrial 18S and 5S rRNA genes (49) and the wheat mitochondrial 5S rRNA sequence (51). The numbering refers to base pairs in the soybean mitochondrial sequence. The dashes indicate the absence of a corresponding nucleotide in the sequence. The predicted amino acid sequence from the open reading frame is shown.

FIGURE 4. S1 nuclease transcript mapping. Examples of S1 nuclease mapping results are presented in part A. A summary of all other S1 results are shown in parts B and C. S1 nuclease analyses were performed on neutral acrylamide (lanes 1-7, also part B lines a-c) and denaturing sequencing gels (lanes 8-19, also part B lines d and e). The labeled DNA fragment and unlabeled RNAs are: (1) 3' end-labeled EcoRI - no RNA, (2) 3' end-labeled EcoRI - 3 ug SB-1 RNA, (3) 3' end-labeled AvaI - no RNA, (4) 3' end-labeled AvaI - 3 ug SB-1 RNA, (5) 5' end-labeled AvaI - no RNA, (6) 5' end-labeled AvaI - 3 ug SB-1 RNA, (11) 5' end-labeled BglII-Sau3A - no RNA, (12) 5' end-labeled BglII-Sau3A - 3 ug SB-1 RNA, (16) 3' end-labeled HpaI-AvaI - no RNA, (17) 3' end-labeled HpaI-AvaI - 10 ug SB-1 RNA, (18) 3' end-labeled HpaI-AvaI - 2 ug SB-1 RNA, (19) 3' end-labeled HpaI-AvaI - 0.4 ug SB-1 RNA. 20 ug tRNA was added as carrier in each S1 experiment. Note that there will be a labeled end within the pBR322 sequences in addition to the labeled ends designated in the schematic in part B, lines a-c. There will also be a small labeled AvaI fragment (position 480-536) that would not have been detected in these gels (see part A, lanes 3-6, part B, lines b and c). The lanes 1 and 2, 3 and 4, and 5 and 6 are from separate gels, each with corresponding size markers. (To confirm the results observed here, the experiments were repeated

with the isolated fragments designated in lines d-j.) Lane 7: end-labeled HpaI digested T7 included as size markers, lanes 8-10: sequencing reactions for 5' end-labeled BglI-Sau3A included as size markers, lanes 13 and 14: sequencing reactions of 3' end-labeled EcoRI-HgiAI included as size markers, lane 15: end-labeled HpaII digested pBR322 included as size markers. (B) The schematic of the various S1 nuclease transcript mapping experiments. (C) Schematic of various S1 nuclease transcript experiments which did not reveal protected fragments. (D) Transcripts and corresponding restriction map of pSB2-2 insert and a scale in bp.

5S specific A probe detected only one fragment when mtDNA was digested with either EcoRI or SalI (Fig. 5, lanes 1 and 3). The B probe hybridized to these same fragments plus one additional fragment in each case (Fig. 5, lanes 2 and 4). Thus the 5S rRNA gene is present only in a sequence environment adjacent to a B homology (i.e. adjacent to the remainder of the 800 nucleotide RNA sequence) and therefore must be the source for 5S rRNA in mitochondria of SB-1 cells. The B region is clearly found in another location in the mitochondrial genome.

Differential expression of the 800 nucleotide transcript

Is the unusual 800 nucleotide transcript peculiar to the SB-1 cell line? Different sources of soybean material, including three tissue culture lines and two sources of soybean hypocotyls ('Wayne' and 'Mandarin') were surveyed for this RNA using probe D (the 700 bp AvaI-SalI fragment from pSB2-2, see Fig. 2d). The 800 nucleotide transcript is apparent only in RNA from the SB-1 tissue culture line (Fig. 6). The dilution series and the longer autoradiograph exposure shown here suggest that the 800 nucleotide RNA is at least 20 fold lower in abundance in all sources compared to SB-1. Examination of a variety of plants including mung bean, maize, tobacco, pea and cucumber failed to reveal the 800 nucleotide transcript (data not shown).

All sources of soybean tested contain a 2600 nucleotide RNA with homology to probe D. We could not detect hybridization between this 2600 nucleotide RNA and probes that were 5S specific (probe A), 18S specific (probe C) or from a region downstream from the 800 nucleotide sequence (probe pSB1-1, a clone of the 2.0 kb SalI-EcoRI fragment shown in Fig. 2a) (data not shown). Probe B, containing sequences from the 800 nucleotide RNA gene but 3' to the 5S sequence, does hybridize with the 2600 nucleotide

FIGURE 5. Southern analysis of SB-1 mitochondrial DNA with probes A and B. MtDNA from the tissue culture line SB-1 was digested with restriction endonucleases, fractionated in 1% agarose gels, transferred to nitrocellulose and hybridized by the method of Southern (35). Lanes 1 and 2 show autoradiographs from digestions with EcoRI and hybridization with probe A or B. Lanes 3 and 4 show autoradiographs of mtDNA digested with SalI and hybridized with probes A or B.

RNA and is also known to be found in at least one other sequence environment in the genome. This led us to suspect that the 2600 nucleotide RNA might be transcribed from the second location of B-specific sequences in the mitochondrial genome. Indeed

FIGURE 6. Analysis of the 800 nucleotide RNA in various soybean sources. Total cell RNA from the various soybean sources was fractionated by electrophoresis in agarose gels, blotted onto DBM paper as described in materials and methods and hybridized with probe D. Lane 1: 5 ug hypocotyl RNA from cultivar 'Wayne', Lane 2: unlabeled marker, Lane 3: 5 ug RNA from fr2 x fr1 tissue culture cells, Lane 4: 5 ug RNA from fr2 tissue culture cells, Lane 5: 5 ug RNA from fr1 tissue culture cells, Lane 6: RNA from SB-1 tissue culture cells, Lane 7: labeled size markers (AvaI-SalI digested pBR322 and AvaII digested pBR322). Lanes 8-11: 0.04, 0.2, 1 and 5 ug RNA from SB-1 tissue culture cells, respectively. Autoradiograph is shown with 2 exposures.

FIGURE 7. Analysis of the 3' end of the 2600 nucleotide transcript. Lanes 1-7: The 3' end labeled *Ava*I-*Hgi*AI fragment (see fig 4B, line i) was used in S1 nuclease analyses without soybean RNA (lane 1), with 10 ug, 2 ug, 0.4 ug of total SB-1 cell RNA, respectively (lanes 2-4). The same fragment was also used in S1 nuclease experiments with 10 ug, 2 ug or 0.4 ug of total *fr*₂ RNA (lanes 5-7). Lane 8 shows the same fragment with no treatment. Lanes 9 and 10 show the G+A and C+T sequencing reactions for the 3' end labeled 856 bp *Eco*RI-*Hgi*AI fragment (see Fig 4B, line f) used here as size markers. Lane 11 is end labeled *Hpa*II digested pBR322 DNA used as size markers. Because of the prominent "smiling" effect seen on this gel, it is difficult to accurately determine size estimates from these results but the 550 bp size is consistent with the results in Figure 4. These results were confirmed using a smaller fragment, the *Sau*3A-*Sal*I fragment (Fig 4B, line j) and again showed that the 3' ends of the transcripts present in both SB-1 and "normal" soybean mitochondria are identical. The results are presented to indicate that the transcripts share 3' ends for at least 550 nucleotides. The 450 bp fragment appears to be protected at levels inversely proportional to the amount of RNA present. Presumably it is due to a favored DNA-DNA hybrid configuration that is outcompeted by an RNA-DNA hybrid and does not represent a transcript.

S1 nuclease protection experiments using RNA from cell line *fr*₂, which contains the 2600 nucleotide RNA but no detectable 800 nucleotide RNA, showed that an RNA 3' end is present that is identical to that from the 800 nucleotide RNA (Fig. 7). These data indicate that the 2600 and 800 nucleotide RNAs may have the same 3' terminal 550 nucleotides but that their 5' sequences are different. Thus it appears that all soybean mitochondria have a 2600 nucleotide RNA, and in addition the SB-1 cells have an 800 nucleotide RNA with the same 3' sequences as the 2600 nucleotide RNA. The differential expression of the 800 nucleotide transcript suggested a possible difference in genome organization in the SB-1 cell line.

Gene organization

To investigate the organization of DNA sequences in the vicinity of these genes, mitochondrial DNAs from the soybean sources used above were compared by hybridization. As was shown

FIGURE 8. Southern analysis of mitochondrial DNAs from either SB-1 cells or other soybean sources using probe B. MtDNAs from 4 different sources were digested with EcoRI, fractionated in 1% agarose gels, transferred to nitrocellulose and hybridized to probe B by the method of Southern (35). Lane 1: SB-1 tissue culture line mtDNA, lane 2: fr₂ tissue culture line mtDNA, lane 3: 'Mandarin' hypocotyl mtDNA, lane 4: 'Wayne' hypocotyl mtDNA.

in Figure 5, the probe B hybridizes to 4.4 kb and 3.8 kb EcoRI fragments from the SB-1 tissue culture cell line. In the other sources of mitochondrial DNA the 4.4 kb fragment is detected but the 3.8 kb fragment is missing and a 1.45 kb fragment is found instead (Fig. 8). The 5S specific A probe hybridizes to the 3.8 kb fragment from SB-1 tissue culture mtDNA and also to the 1.45 kb fragments found in the other mtDNAs but not to the 4.4 kb fragment (data not shown). The mtDNA from 'Mandarin' hypocotyls, the cultivar from which the SB-1 tissue culture line was derived, shows the same arrangement as the other "normal" soybean mitochondrial genomes. In these experiments we have repeatedly observed minor bands including a 3.8 kb fragment. These bands may represent regions with only partial homology to the probe or actually indicate the presence of the SB-1 cell line DNA organization (3.8 kb EcoRI fragment) at some low level in the normal mitochondrial DNA population. Since mitochondrial DNA is prepared from a large number of cells, we would not be able to determine whether this organization might be present as a minor component in the mitochondrial genome or as a major component in a subpopulation of cells. In similar experiments to the Southern analyses shown in Figure 8, mtDNAs from SB-1 and fr₂ cells were digested with a series of different restriction endonucleases and probed with A and B DNA. The sequence of the pSB2-2 insert was used to predict the expected SB-1 mtDNA fragment sizes which were compared to the observed fragments (Table 1) to deduce the partial restriction maps of the various

Table 1. Southern analysis with probes A and B

Restriction enzyme	Predicted fragment	Fragments detected in SB-1 DNA	Fragments detected in fr ₂ DNA
DNA fragments in bp with probe A			
AvaI	550	550	550
EcoRI	3800	3800	1450
SalI	>1800	>10,000	>10,000(doublet)
AvaI & EcoRI	550	550	550
AvaI & SalI	550	550	550
SalI & EcoRI	1800	1800	1450
AluI	233	233	233
HaeIII	513	513	513
Sau3A	334, 519	334, 519	334, 650
DNA fragments in bp with probe B			
AvaI	>712	2800	1300, 2800
EcoRI	3800	3800, 4400	1450, 4400
SalI	>1800	2250, >10,000	2250, >10,000(doublet)
AvaI & EcoRI	2712	2712	370, 2712
AvaI & SalI	712	712	712, 1300
SalI & EcoRI	1800	1800, 2250	1450, 2250
AluI	246	246	246, 450
HaeIII	525	450, 525	300, 525
Sau3A	519	519	450, 650

Restriction mapping data from two alternate sources of soybean mitochondrial DNA. MtDNA from SB-1 and "normal" (fr₂) soybean was digested with a series of restriction endonucleases, fractionated on agarose gels, transferred to nitrocellulose and hybridized with probes A and B (36). The sequence of the 1.8 kb pSB2-2 insert was used to predict fragment sizes which are indicated along with the observed fragments. Differences from the predicted sizes were used to deduce the points at which the maps differed among the various environments.

sequence environments (Fig. 9). Line c shows the 1.8 kb EcoRI-SalI insert that has been sequenced and lines a and b show the deduced maps of the "normal" plant mtDNA regions containing homology to B. The solid lines indicate regions of identity between the different environments. From the restriction mapping data we conclude that the sequence identity between a and c breaks down in a region delimited by the Sau3A and HaeIII restriction sites and between b and c in the region delimited by AluI and HaeIII restriction sites which is in a region 80 to 350 base pairs beyond the 5S rRNA gene (see stippled regions on the maps). Since there is homology to the B probe beyond the AluI site

FIGURE 9. Restriction mapping. The maps were deduced from the data in Table 1. All three regions contain homology to probe B. Line a shows the map of the region in common to all soybean sources that presumably encodes the 2600 nucleotide transcript. Line b shows the map of the region encoding the 5S rRNA in "normal" soybean mitochondria. Line c shows the 1.8 kb EcoRI-SalI insert from pSB2-2 that has been sequenced. The probes A and B are noted at the bottom. The solid lines indicate regions of sequence identity between the environments shown here.

in both sequence environments, the putative point of change must lie far enough to the right to include the majority of the B probe. Sequencing of these regions will allow exact determination of the point of non-identity. The possibility that the SB-1 organization arose from the "normal" sequences via homologous recombination is suggested by the presence of regions homologous to sequence B in the different environments. Recombination in this region would fuse the 3' end of the gene for the 2600 nucleotide transcript to the 5S rRNA gene to give rise to the gene for the 800 nucleotide RNA. This rearrangement must have become predominant during growth in tissue culture since it is not abundant in mtDNA from 'Mandarin' hypocotyls.

The simplest overall organization of these regions in both soybean sources is shown by the schematic in Figure 10. The

FIGURE 10. The simplest organization of the regions containing homology to B in both "normal" and SB-1 soybean mitochondrial genomes. The DNA regions shown are the corresponding EcoRI fragment from each environment. The region B is designated by the black line. Other patterns represent other regions of identity. The positions of the transcripts have been included.

picture presented here for the mtDNA organization suggests that the ribosomal genes are present only in a single environment in the soybean mitochondrial genome. However, if we examine the sequences further to the right of these regions the situation appears to be more complicated. Preliminary evidence suggests either multiple copies of the rRNA genes or genome heterogeneity.

DISCUSSION

The simplest interpretation of the data presented here is that in the tissue culture line SB-1 the mitochondrial 5S rRNA gene has been fused to the downstream end of an unknown gene that normally encodes a 2600 nucleotide RNA. This results in abundant expression of a novel 800 nucleotide transcript whose function is not known. Since a normal, unfused copy of the 5S rRNA gene is not maintained in the SB-1 tissue culture line, the fused copy of the 5S gene is apparently sufficient for the mitochondrial 5S rRNA requirement. It is likely that the 5S rRNA in SB-1 tissue culture cells is produced by the same mechanism as in normal plant cells since the site of fusion appears to be in a region at least 80 and up to 350 base pairs downstream from the 5S rRNA gene. The fused arrangement requires either processing of the large 800 nucleotide transcript or incomplete termination of transcription after the 5S gene.

The reciprocal recombinant form that would link the 5' end of the gene for the 2600 RNA to the sequences normally found

downstream from the 5S gene has not been detected and the normal 2600 nucleotide gene is still present in SB-1 cells in amounts about equimolar to the fused gene. This might implicate a gene conversion-like mechanism or transposition event in the generation of the altered DNA organization. However, since most mitochondria are likely to have more than one copy of the genome, an unaltered copy of the 2600 gene could readily have survived. Since both DNA arrangements are maintained in SB-1 cells, the fused arrangement cannot be unduly deleterious.

Lonsdale et al. (5) have suggested that the molecular heterogeneity of the maize mitochondrial genome arises by recombination at repeated regions in the mtDNA. A very similar picture is emerging in Brassica (6) in which a 2 kb repeat is thought to provide a specific recombination site for the interconversion of three circular molecules. Various size classes of circular molecules have also been reported for soybean (19). The apparent repeated sequences located after the 5S gene and within the 2600 gene may be another example of potential recombination sites. Unlike the maize and Brassica cases, however, we find that the two forms of the 5S rRNA gene (fused and unfused) do not approach equimolar amounts in soybean. There are suggestions that the altered DNA rearrangement may be present in normal (non SB-1) mitochondria at low levels. The minor bands observed in hybridization to normal soybean mtDNA appear identical to the major 3.8 kb SB-1 species (Fig. 8). Since the altered arrangement, if indeed present in normal cells, is present at very low levels it may represent a minor recombinant class such as ones suggested by both Lonsdale et al. (5) and Palmer and Shields (6).

The novel RNA described here is the first direct evidence of altered transcription of the plant mitochondrial genome associated with a specific DNA rearrangement. Unfortunately, the role of the unusual 800 nucleotide transcript is unknown and cannot be deduced from its sequence organization. Perhaps a similar differential expression exists in maize and Brassica but at this point it is not known whether the repeated regions in maize and Brassica are transcribed.

The rearrangement that we have characterized here and some of the other examples cited above may be merely consequences

of tissue culture manipulations. Nonetheless studying them may help unravel the mechanisms involved in mitochondrial DNA rearrangement. More importantly, these events may have resulted in fixing the mitochondrial DNA in particular arrangements that are functionally important but normally present only transiently or in a minority of molecules and may be particularly useful in the study of differential gene expression in plant mitochondria.

ACKNOWLEDGMENTS

We would like to thank the following people for materials, technical assistance and/or laboratory space: Jill Roth, Paul Jackson, Gordon Lark, Reid Palmer, William Thompson, Russell Doolittle, and Eric Kofoed. P.H.M. was supported in part by N.I.H. Training Grant GM07531.

REFERENCES

1. Lebacqz, P. and Vedel, F. (1981) *Plant Science Letts.* 23, 1-9.
2. Ward, B.L., Anderson, R.S. and Bendich, A.J. (1981) *Cell* 25, 793-803.
3. Leaver, C.J. and Gray, M.W. (1982) *Ann. Rev. Plant Physiol.* 33, 373-402.
4. Levings, C.S., III (1983) *Cell* 32, 659-661.
5. Lonsdale, D.M., Hodge, T.P., Fauron, C.M.R. and Flavell, R.B. (1983) In *ICN-UCLA Symposia on Molecular and Cellular Biol.*, R.B. Goldberg, Ed., new ser. 12, pp. 445-456, Alan R. Liss, Inc., New York.
6. Palmer, J.D. and Shields, C.R. (1984) *Nature* 307, 437-440.
7. Dale, R.M.K. (1981) *Proc. Nat. Acad. Sci. USA* 78, 4453-4457.
8. Dale, R.M.K., Wu, M. and Kiernan, M.C.C. (1983) *Nucl. Acids Res.* 11, 1673-1685.
9. Belliard, G., Vedel, F. and Pelletier, G. (1979) *Nature* 281, 401-403.
10. Nagy, F., Török, I. and Maliga, P. (1981) *Mol. Gen. Genet.* 183, 437-439.
11. Boeshore, M.L., Lifshitz, I., Hanson, M.R. and Izhar, S. (1983) *Mol. Gen. Genet.* 190, 459-467.
12. Sederoff, R.R., Levings, C.S., III, Timothy, D.H. and Hu, W.W.L. (1981) *Proc. Nat. Acad. Sci. USA* 78, 5953-5957.
13. Gengenbach, B.G., Connelly, J.A., Pring, D.R. and Conde, M.F. (1981) *Theor. Appl. Genet.* 59, 161-167.
14. Levings, C.S., III, Kim, B.D., Pring, D.R., Conde, M.F., Mans, R.J., Laughnan, J.R. and Gabay-Laughnan, S.J. (1980) *Science* 209, 1021-1023.
15. Kemble, R.J. and Mans, R. J. (1983) *J. Mol. Appl. Genet.* 2, 161-171.
16. Kemble, R.J., Flavell, R.B. and Brettell, R.I.S. (1982) *Theor. Appl. Gen.* 62, 213-217.
17. Brettell, R.I.S., Thomas, E. and Ingram, D.S. (1980) *Theor.*

- Appl. Genet. 58, 55-58.
18. Quetier, F. and Vedel, F. (1980) In Genome Organization and Expression in Plants, C.J. Leaver, Ed., pp. 401-406, Plenum Press, New York.
19. Synenki, R.M., Levings, C.S., III and Shah, D.M. (1978) Plant Physiol. 61, 460-464.
20. Levings, C.S., III and Pring, D.R. (1979) In Genetic Engineering, J.K. Setlow and A. Hollaender, Eds., vol. 1, pp. 205-222, Plenum Press, New York.
21. Wallace, D.C. (1982) Microbiological Reviews 46, 208-240.
22. Gamburg, O.L., Miller, R.A. and Ojima, K. (1968) Exp. Cell Res. 50, 151-158.
23. Roth, E.J., Weber, G. and Lark, K.G. (1982) Plant Cell Reports 1, 205-208.
24. Morgens, P.H. (1984) Ph.D. thesis, University of Utah.
25. Aviv, H. and Leder, P. (1972) Proc. Nat. Acad. Sci. USA 69, 1408-1412.
26. Walbot, V. and Goldberg, R. (1979) In Nucleic Acids in Plants, T.C. Hall and J.W. Davies, Eds., vol. 1, pp. 3-40, CRC Press, Boca Raton, Florida.
27. Sparks, R.B., Jr. and Dale, R.M.K. (1980) Molec. Gen. Genet. 180, 351-355.
28. Weislander, L. (1979) Anal. Biochem. 98, 305-309.
29. Smith, H.O. (1980) In Methods in Enzymology, L. Grossman and K. Moldave, Eds., vol. 65, pp. 371-380, Academic Press, New York.
30. Bailey, J.M. and Davidson, N. (1976) Anal. Biochem. 70, 75-85.
31. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Proc. Nat. Acad. Sci. USA 74, 5350-5354.
32. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
33. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Proc. Nat. Acad. Sci. USA 72, 1184-1188.
34. Young, M.W. and Hogness, D. S. (1977) In Molecular Approaches to Eucaryotic Genetic Systems, ICN-UCLA Symposia on Molecular and Cellular Biology, G. Wilcox, J. Abelson and C.F. Fox, Eds., vol. VIII, pp. 315-331, Academic Press, New York.
35. Messing, J., Gronenborn, B., Müller-Hill, B. and Hofschneider, P.H. (1977) Proc. Nat. Acad. Sci. USA 74, 3642-3646.
36. Maxam, A.M. and Gilbert, W. (1980) In Methods in Enzymology, L. Grossman and K. Moldave, Eds., vol. 65, pp. 499-560, Academic Press, New York.
37. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
38. Weaver, R.F. and Weismann, C. (1979) Nucl. Acids Res. 7, 1175-1193.
39. Brutlag, D.L., Clayton, J., Friedland, P. and Kedes, L.H. (1982) Nucl. Acids Res. 10, 279-294.
40. Wilbur, W.J. and Lipman, D.J. (1983) Proc. Nat. Acad. Sci. USA 80, 726-730.
41. Schnare, M.N. and Gray, M.W. (1982) Nucl. Acids Res. 10, 3921-3932.
42. Chao, S., Sederoff, R.R. and Levings, C.S., III (1983) Plant Physiol. 71, 190-193.
43. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Nat. Acad. Sci. USA 75, 4801-4805.
44. Spencer, D.F., Bonen, L. and Gray, M.W. (1981) Biochemistry 20, 4022-4029.