Cloning and nucleotide sequence analysis of *psbD/C* operon from chloroplasts of *Populus deltoides*

M. S. SRINIVASA REDDY*, PRABODH K. TRIVEDI, RAKESH TULI and PRAFULLACHANDRA V. SANE†

Centre for Plant Molecular Biology, National Botanical Research Institute, Rana Pratap Marg, Lucknow 226 001, India

Abstract

We report the cloning and nucleotide sequence analysis of *psbD/C* operon from a dicotyledonous tree species, *Populus deltoides* (poplar). The coding regions of *psbD* and *psbC* and deduced amino acid sequences show very high homology with those from other higher plants. In pairwise alignment of the gene sequences, *P. deltoides* clustered with dicotyledonous annuals rather than with *Pinus*, the only other tree whose *psbD/C* nucleotide sequence is available. Comparison of several reported sequences showed that synonymous substitutions were distributed in both *psbD* and *psbC* uniformly, throughout the length of the genes. The frequency of nonsynonymous substitutions located in the amino-terminal end of *psbD* was distinctly higher, suggesting a lower degree of structural constraints in this region of the encoded D2 protein. The arrangement of reading frames and Northern analysis suggest that organization and expression of *psbD/C* operon in *P. deltoides* is similar to that in other higher plants.

[Reddy M.S.S., Trivedi P.K., Tuli R. and Sane P.V. 1998 Cloning and nucleotide sequence analysis of *psbD/C* operon from chloroplasts of *Populus deltoides*. *J. Genet.* **77**, 77–83]

Introduction

The *psbD* and *psbC* chloroplast genes encode D2, a reaction centre protein, and CP43, a chlorophyll-*a*-binding antenna protein, respectively. The upstream regions of the *psbD/C* operon are structurally distinct in the chloroplast genomes of cereals and dicotyledonous plants. The organization of *psbD* and *psbC* genes is conserved in higher plants, including the holoparasitic flowering plant, *Cuscuta reflexa* (Habeshausen *et al.* 1992). In contrast, the organization of *psbD* and *psbC* genes differs widely in lower forms. In *Euglena gracilis*, *psbD* and *psbC* contain introns (Orsat *et al.* 1994). The *psbD* and *psbC* genes in *Chlamydomonas reinhardtii* are monocistronic and lack introns (Rochaix *et al.* 1989). The 12 amino acid residues at the amino-terminal end of CP43 in higher plants are absent in *Chlamydomonas*. In the cyanobacteria *Synechococcus* sp. strain PCC 7942 (Golden *et al.* 1989), *Synechocystis* sp. strain PCC 6803 (Yu and Vermaas 1990) and *Anabaena* sp. strain PCC 7120 (Leonhardt and Straus 1994) two functional *psbD* genes have been reported.

Cotranscription of *psbD* with *psbC* and overlapping reading frames are characteristic features in several monocots (Hiratsuka *et al.* 1989) and dicots (Yao *et al.* 1989). Transcription of the *psbD/C* operon is highly complex in most higher plants. Transcription patterns differ between dark-grown and illuminated barley leaves (Sexton *et al.* 1990). In tobacco eight transcripts ranging in size from 1.5 to 4.4 kb were reported. Two promoters were detected in *psbD/C* operon in dicots (Meng *et al.* 1991) whereas four promoters were reported in cereals (Chen *et al.* 1994).

Most of the studies on the characterization of chloroplastic operons have been carried out on annual crop plants and algae. Unlike annuals which experience only one set of environmental conditions characteristic of their growing season, the leaves of poplar, *Populus deltoides*, experience extreme conditions of temperature, light and water stress. It was therefore of interest to find out if organization and

* Present address: Department of Agronomy, College of Agriculture, University of Kentucky, Lexington, KY 40546-0091, USA.
† For correspondence. E-mail: manager@nbri.sirnetd.ernet.in.

Keywords. chloroplast genome; nonsynonymous substitutions; *Populus deltoides; psbD/C* operon; sequence analysis; synonymous substitutions.

expression of different operons in tree species would be different and contain elements that permit greater flexibility to suit changing environmental conditions. So far the chloroplast genome of only the gymnospermic tree *Pinus*, whose leaves are needle-like, has been examined. Our laboratory has a major programme to establish the nucleotide sequence and expression of chloroplastic genes from poplar, a fast-growing tree, which grows well in alkaline soil as well as withstands high light and high temperature conditions in summer. The main objective of this study was to examine if the organization and expression of *psbD/C* in *P. deltoides* was similar to that in other crop plants. Establishment of the complete nucleotide sequence of the plastid genome of *Populus deltoides* and studies on the regulation of its expression will also help in genetic manipulation of the chloroplastic traits.

Materials and methods

Chloroplast DNA isolation and Southern hybridization: Chloroplast DNA (cpDNA) was prepared from green leaves of *P. deltoides* as described in Nath *et al.* (1993). Four micrograms of cpDNA digested with restriction endonucleases was resolved on agarose gel and transferred onto Zeta Probe membrane (Biorad). The DNA was crosslinked to the nylon membrane and used in Southern hybridization using radiolabelled probes prepared by random priming method (Sambrook *et al.* 1989). A heterologous probe carrying *psbD* from *Vigna aconitifolia* was taken out as a 1.1-kbp *Pst*I fragment from pVAII3 (a gift from Prof. A. K. Tyagi, Delhi University, India). The *Spinacia oleracea psbC* probe was obtained as a 837-bp *Eco*RI/*Xba*I fragment from pPSII44 (a gift from Prof. R. G. Herrmann, Botanisches Institüt der Universität, Munich, Germany).

Cloning of psbD and psbC genes: Two libraries of cpDNA digested with *Hin*dIII and *Pst*I were constructed in pBluescriptIIKS(+) phagemid and used to transform *E. coli* XL1 Blue (Sambrook *et al.* 1989). Recombinant clones pPHR7 and pPPR89, carrying genes *psbD* and *psbC* only, were identified by colony hybridization (Sambrook *et al.* 1989).

Subcloning and sequencing: All DNA manipulation techniques were carried out as described in Sambrook *et al.* (1989). A 1.4-kbp fragment of pPHR7 obtained after double digestion as well as 2.2-kbp and 1.0-kbp *Pst*I fragments of pPPR89 were subcloned in pBluescriptIIKS(+). Unidirectional nested deletions of pPHR7 and a subclone of pPPR89 containing the 1.0-kbp insert were constructed using the exonuclease III/mungbean nuclease system (Stratagene) according to manufacturer's instructions and procedures described by Trivedi *et al.* (1994). Sequencing reactions on double-stranded plasmid were carried out using [α-^{32}P]dATP (BRITS, Mumbai, India) and T7 sequencing kit (Pharmacia). The reaction products were resolved on urea–acrylamide gel and autoradiographed for determining the sequence. Alternatively, the automated DNA sequencing system (373, ABI) was used, with the dye termination cycle sequencing kit (Perkin Elmer).

Nucleotide substitution analysis: Pairwise multiple alignments of nucleotide sequences of the coding regions were carried out using the CLUSTAL program in PC/GENE release 14.0 (Intelligenetics Inc., Mountain View, USA) and sites of nucleotide substitutions were determined. Nucleotide substitutions were divided into two categories: (i) synonymous substitutions, i.e. nucleotide sequence changes that result in change of codon with no change of amino acid, and (ii) nonsynonymous substitutions, i.e. changes in nucleotide sequence that result in change of amino acid.

RNA isolation and Northern hybridization: Total RNA was extracted from leaves by the phenol/chloroform method (Van Slogteren *et al.* 1983). RNA (30 µg) was electrophoresed on a denaturing formaldehyde–MOPS agarose gel and transferred onto a Zeta Probe membrane (Biorad). Northern hybridization was carried out using the 2.1-kbp *Hin*dIII fragment of pPHR7 as probe. Prehybridization, hybridization and washing of blots were performed at 42°C as described (Sambrook *et al.* 1989).

Results and discussion

The hybridization patterns of *psbD* and *psbC* probes with cpDNA of *P. deltoides* are shown in figure 1. *Hin*dIII produced a single 2.1-kbp band hybridizing to both the probes indicating that the two genes are located on the same fragment. The results from digestion with other restriction enzymes, particularly *Dra*I and *Sac*I, suggest that *psbD* and *psbC* are located adjacent to each other. The results demonstrate that *psbD* and *psbC* genes in *P. deltoides* are present in single copy, as in higher plants and unlike in cyanobacteria.

Cloning and sequencing of the psbD and psbC genes

Two recombinant clones—pPHR7, positive for *psbD* as well as *psbC*, and pPPR89, positive for *psbC*—were selected after colony hybridization. Inserts in both the clones together represent a 5.1-kbp cpDNA region containing *psbD* and *psbC* genes as shown in figure 2. The original inserts, their subclones and nested deletions were sequenced from both the ends as shown in figure 2. The sequenced portion revealed the presence of reading frames for *psbD* (1062 bp) and *psbC* (1422 bp) genes (figure 3) in a bicistronic operon. Motifs representing putative ribosome binding sites are located seven nucleotides upstream of *psbD* and 20 nucleotides upstream of *psbC*. The 3′ end of *psbD* and 5′ end of *psbC* overlap by 53 nucleotides, as in other higher plants

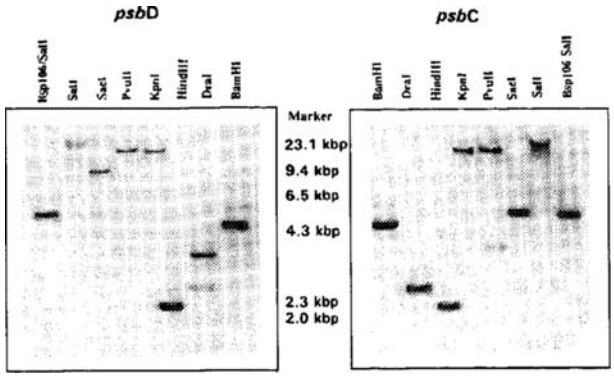

Figure 1. Southern blot analyses of cpDNA from *P. deltoides* using a *Vigna aconitifolia psbD* gene probe (left) and a *Spinacia oleracea psbC* gene probe (right). The DNA samples were digested with restriction enzymes *Bam*HI, *Dra*I, *Hin*dIII, *Kpn*I, *Pvu*II, *Sac*I, and *Bsp*106/*Sal*I. The positions of six molecular weight markers (*Hin*dIII-digested λ DNA) are shown. The faint bands of 6.3 kbp in *Bam*HI lane and 3.9 kbp in *Sac*I lane of the *psbD* Southern blot and 6.3 kbp in *Bam*HI lane of the *psbC* Southern blot are due to nonspecific hybridization.

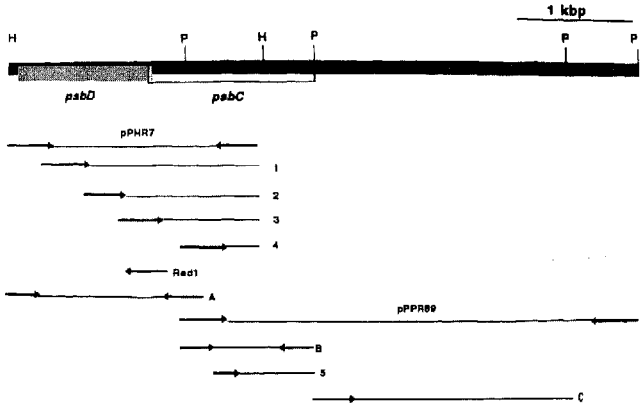

Figure 2. Diagrammatic representation of different clones, unidirectional deletions (1–5) and subclones (A–C) used for sequencing the *psbD/C* operon. Direction and length of the arrows indicate the direction and number of nucleotides read from one reaction. Red1 is an internal primer used to establish complete coding sequence.

and unlike in certain algae and cyanobacteria. In dicots, *psbD/C* is separated from the upstream *psbK/I* operon by over 23 kb. In contrast, *psbD/C* in cereals resides 1–2 kb downstream of and in the same polarity as *psbK/I*. Monocots like duckweed (DeHeij *et al.* 1983) and orchids (Chase and Palmer 1989) also have the dicot arrangement of the *psbD/C* operon. In *Euglena gracilis* the *psbD* and *psbC* genes together cover a 22-kb region, with 10 and 11 exons respectively (Orsat *et al.* 1994). In *Chlamydomonas reinhardtii* the *psbD* and *psbC* genes are transcribed as a monocistronic transcript (Rochaix *et al.* 1989). One of the two *psbD* genes in cyanobacteria, *psbDII*, is expressed as monocistronic RNA while *psbDI* is cotranscribed with *psbC* and overlaps its reading frame (Golden *et al.* 1989; Yu and Vermaas 1990; Leonhardt and Straus 1994).

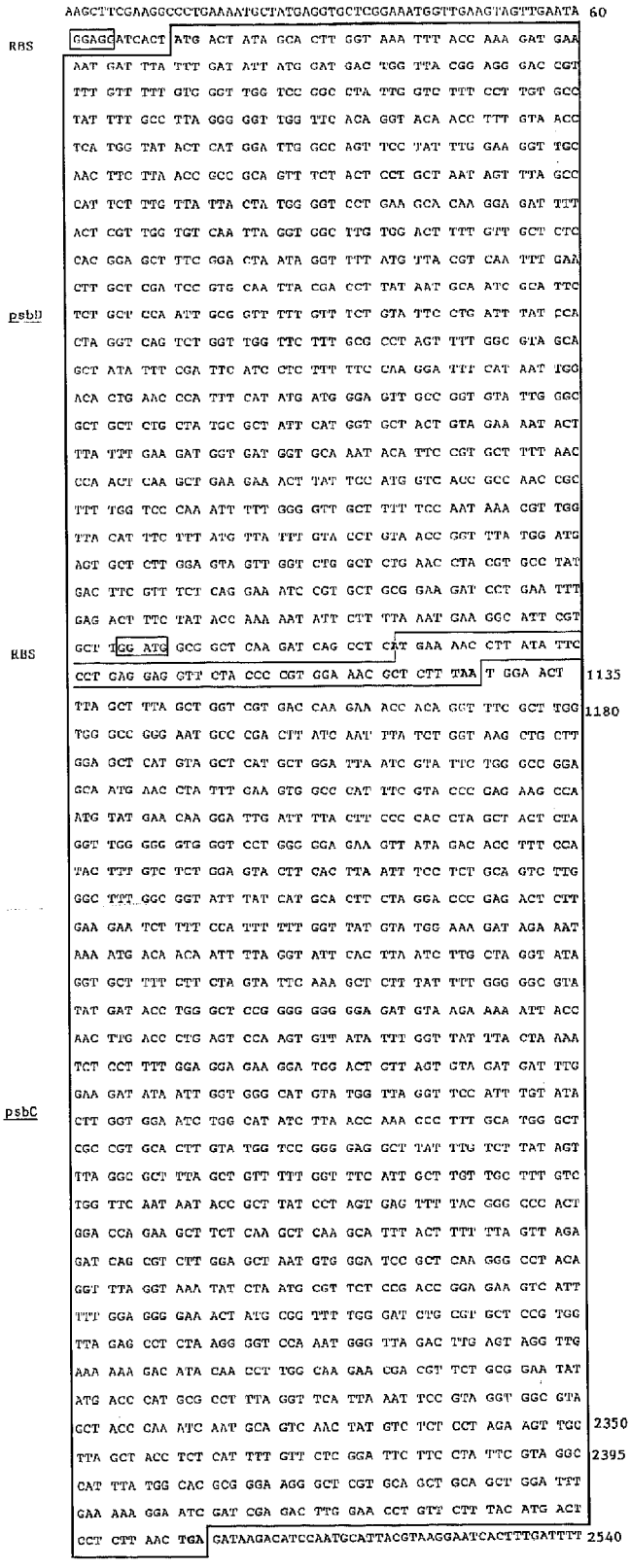

Figure 3. Complete nucleotide sequence of *psbD/C* operon of *P. deltoides* along with flanking regions. The reading frames of *psbD* (nucleotides 72–1133) and *psbC* (nucleotides 1081–2502) overlap by 53 nucleotides. The putative ribosome binding sites, the start and stop codons, and the overlapping reading frames are indicated in the sequence.

Table 1. Per cent similarity in sequences of *psbD* genes (above diagonal) and D2 proteins (below diagonal) from various plant species.

	PD	AM	NT	PS	SO	HV	OS	SC	PT	MP	CR	CC
PD		94.2	93.8	93.6	93.6	90.0	90.7	90.7	89.7	84.8	77.8	76.9
AM	99.1		96.5	93.1	94.0	90.6	91.3	91.1	89.5	85.1	78.2	77.0
NT	99.1	100		93.1	94	90.6	91.3	91.1	89.5	85.1	78.2	77.0
PS	97.7	98.5	98.5		92.7	89.5	89.5	90.0	88.7	84.6	78.0	78.3
SO	98.0	98.3	98.3	96.9		90.4	91.7	90.9	88.9	85.5	78.5	77.2
HV	96.3	96.6	96.6	95.2	96.3		96.1	98.7	87.6	84.0	77.4	76.9
OS	97.7	98.0	98.0	96.6	97.4	98.5		96.2	87.2	83.6	76.8	77.1
SC	97.2	97.4	97.4	96.0	97.1	98.8	99.1		88.1	84.7	77.7	77.6
PT	96.9	96.6	96.6	95.7	96.3	96.0	96.6	96.9		85.3	79.0	76.6
MP	96.3	96.6	96.6	95.5	96.9	95.7	96.6	96.6	97.1		79.9	81.4
CR	92.6	91.0	91.0	89.9	91.3	90.4	91.3	91.3	91.8	91.8		76.6
CC	86.1	85.4	85.4	85.1	85.1	84.8	85.4	85.4	85.4	86.8	85.1	

PD, *Populus deltoides*; AM, *Antirrhinum majus*; NT, *Nicotiana tabacum*; PS, *Pisum sativum*; SO, *Spinacea oleracea*; HV, *Hordeum vulgare*; OS, *Oryza sativa*; SC, *Secale cereale*; PT, *Pinus thunbergii*; MP, *Marchantia polymorpha*; CR, *Chlamydomonas reinhardtii*; CC, *Cyanidium caldarium*.

Table 2. Per cent similarity in sequences of *psbC* genes (above diagonal) and CP43 proteins (below diagonal) from various plant species.

	PD	NT	PS	SO	OS	SC	PT	MP	CR	CC
PD		93.5	92.2	90.6	89.9	89.8	87.3	85.1	74.2	74.1
NT	98.5		92.5	91.6	90.4	90.0	87.7	85.0	74.4	73.8
PS	97.9	98.1		90.3	89.1	88.8	86.2	85.2	74.4	74.4
SO	97.7	97.8	97.0		87.5	88.8	86.0	83.8	73.6	74.2
OS	97.0	96.8	95.7	95.3		95.6	86.2	85.3	74.0	74.5
SC	96.4	96.4	94.7	95.5	98.3		85.7	84.8	73.5	74.4
PT	94.3	94.9	93.4	94.3	93.9	94.1		84.8	74.2	74.0
MP	95.8	95.7	95.1	94.1	95.5	94.1	94.7		77.4	77.8
CR	85.9	86.3	85.2	85.6	85.2	84.4	85.4	86.9		73.7
CC	82.9	82.8	82.2	83.3	82.3	82.8	83.1	82.8	81.5	

Abbreviations of plant names are as in table 1.

Nucleotide sequence homology analyis

The nucleotide sequences of the *psbD* and *psbC* genes of *P. deltoides* and the deduced amino acid sequences show a very high homology with those from other plants, as shown in tables 1 and 2. Similar results with high homology in coding regions of plastid-encoded genes from *P. deltoides* and those from other plants have been reported (Trivedi *et al.* 1994; Naithani *et al.* 1997a,b). Dendrograms constructed on the basis of pairwise similarity scores of the *psbD* (figure 4) and *psbC* (figure 5) genes place *P. deltoides* close to dicotyledonous plants and monocots in a separate cluster. The only other tree species in which chloroplast genes have been sequenced is *Pinus thunbergii* (Wakasugi *et al.* 1994). The nucleotide sequence comparison establishes that *psbD* and *psbC* of *P. deltoides* group with dicotyledonous annuals rather than with *Pinus*. To examine the phylogenetic relationship between a chloroplast gene and a nuclear gene, we constructed a dendrogram for the nuclear gene *rbcS* (figure 6) which codes for the small subunit of ribulose bisphosphate carboxylase. The relationships for this nuclear gene (figure 6) and for the chloroplast genes (figures 4 and 5) are very similar. Six sequences of *rbcS* from the family Brassicaceae grouped together and were more closely related to sequences from other dicots than to those from the monocots included in the analysis. The genes from plants of the Leguminoseae and the Solanaceae also clustered together. However, like in the case of the chloroplast genes, the monocots grouped separately.

Analysis of nucleotide substitutions in psbD and psbC genes

The qualitative and quantitative representation of synonymous and nonsynonymous substitutions in the reading frames of the *psbD* and *psbC* genes are shown in figures 7 and 8 respectively. The analysis of *psbD* coding sequences from 12 species indicates synonymous substitutions in 244 codons and nonsynonymous substitutions in 64. This gives a frequency of synonymous substitutions of 23 per 100 nucleotides and of nonsynonymous substitutions of 6 per 100 nucleotides. These numbers suggest a high degree of constraint on mutation frequency due to function-related restrictions imposed on the structure of the D2 protein. Substitution of Glu-70 and His-189 of D2 (Pakrasi and Vermaas 1992) has been reported to impair photoautotrophic growth. These positions indeed show no nonsynonymous substitutions in *psbD* throughout evolution.

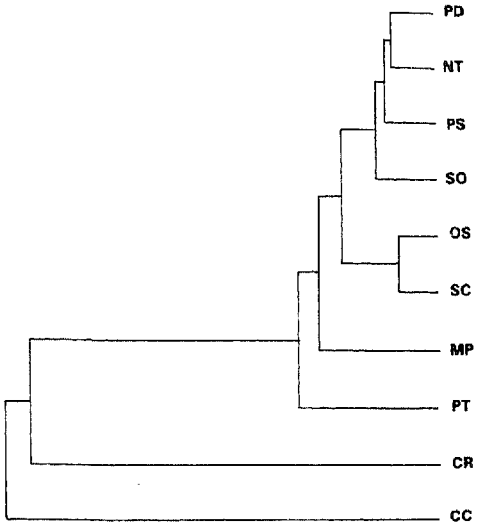

Figure 4. Dendrogram for *psbD* genes from different plants. Nucleotide sequences of *psbD* were compared using the CLUSTAL program of PC GENE. Abbreviations of plant names are as in table 1.

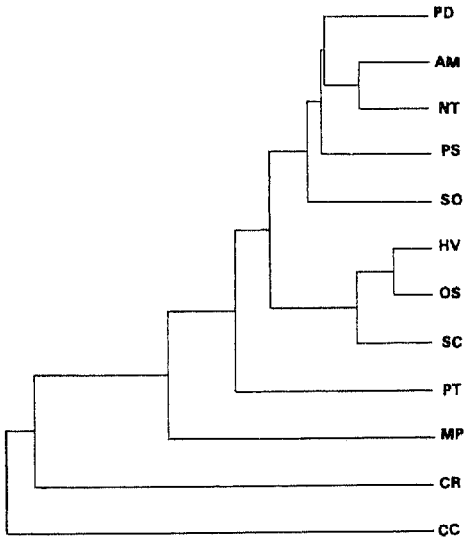

Figure 5. Dendrogram for *psbC* genes from different plants. Abbreviations are as in table 1.

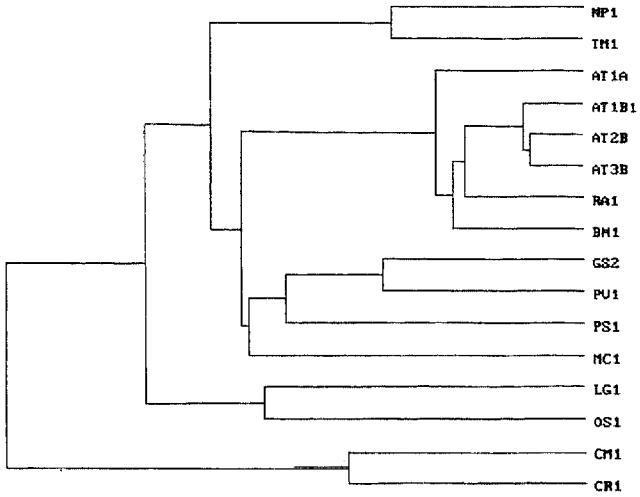

Figure 6. Dendrogram for *rbcS* genes from different plant taxa. The plants are *Nicotiana plumbaginifolia* (NP1), *Lycopersicon esculentum* (TM1), *Arabidopsis thaliana*, (AT1A, AT1B1, AT2B, AT3B), *Raphanus sativus* (RA1), *Brassica napus* (BN1), *Glycine soja* (GS2), *Phaseolus vulgaris* (PV1), *Pisum sativum* (PS1), *Mesembryanthemum crystallinum* (MC1), *Lemna gibba* (LG1), *Oryza sativa* (OS1), *Chlamydomonas moewusii* (CM1) and *Chlamydomonas reinhardtii* (CR1).

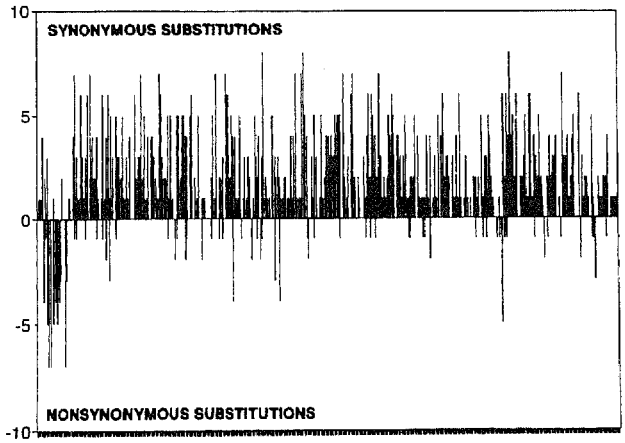

Figure 7. Distribution of synonymous and nonsynonymous substitutions in reading frames of *psbD* from the 12 plants in table 1. The horizontal axis, from 0 to 353, represents the codons and the lengths of the vertical lines at the codon positions correspond to the number of plants (the vertical axis) in which a given codon shows substitution.

The analysis of *psbC* sequences from 10 species shows synonymous substitutions in 284 codons and nonsynonymous substitutions in 124. The relatively higher frequency of nonsynonymous substitutions in *psbC* probably means less selective constraint on amino acid replacements in CP43 compared to that in D2. Synonymous substitutions in both *psbD* and *psbC* are distributed reasonably uniformly throughout the length of the genes. A similar finding was reported earlier for the *psbEFLJ* operon (Naithani *et al.* 1997a). However, nonsynonymous substitutions are more frequent in certain domains, specially the N-terminal domain of both the proteins and the central domain of CP43. Fourteen nonsynonymous substitutions are clustered within the first 25 codons of *psbD* while the rest of the gene shows no such clustering. The type of synonymous substitution in *psbD* and *psbC* in different species may reflect species-specific differences in codon usage. It has been suggested that such selection at codon level increases translational efficiency (Ikemura 1985; Sharp and Li 1987), and that variation in the rate of synonymous substitutions among different organisms results from the level of gene expression (Sharp 1991), with highly expressed genes reported to show bias for certain codons among synonymous ones (Morton 1994). Codon usage in chloroplast genes is biased towards a high representation of NNT and NNA (N

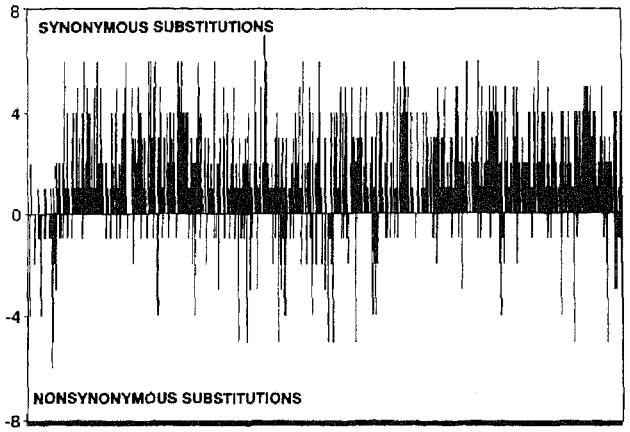

Figure 8. Distribution of synonymous and nonsynonymous substitutions in reading frames of *psbC* from 10 plants in table 1. The horizontal axis here represents codons 0 to 473.

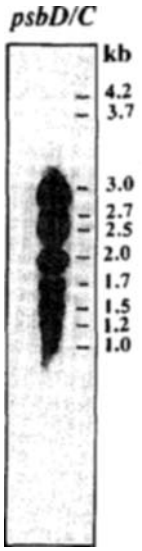

Figure 9. Northern blot with a *psbD/C* probe to detect transcripts of *psbD/C* operon in *P. deltoides*. Two faint transcripts of 3.7 and 4.2 kb are not visible in the photograph.

= any nucleotide) and it has earlier been suggested that this reflects mutational bias (Wolfe and Sharp 1988; Wolfe *et al.* 1992; Morton 1993).

Transcripts from psbD/C operon

The *psbD/C* probe hybridized to transcripts of 4.2, 3.7, 3.0, 2.7, 2.5, 2.0, 1.7, 1.5, 1.2 and 1.0 kb (figure 9). Multiple transcripts can arise either by processing of a polycistronic mRNA (Barkan *et al.* 1995) or because of generation, from multiple promoters, of several independent transcripts which may be processed further (Christopher *et al.* 1992). Twelve different *psbD/C* transcripts, driven by four different promoters, are produced in rice (Chen *et al.* 1994). Multiple promoters and transcripts with different 5′ and 3′ ends have been reported in barley (Sexton *et al.* 1990) and tobacco (Yao *et al.* 1989). Some transcripts start upstream of *psbD* and cover both the genes while others start within *psbD* and code for CP43 only. Since overall organization of *psbD* and *psbC* genes in *P. deltoides* is similar to that in other higher plants, the multiple transcripts may possibly be due to multiple initiation sites as well as processing of a bicistronic mRNA.

Acknowledgements

We thank Prof. R. G. Herrmann, Botanisches Institüt der Universität, Munich, Germany, and Prof. A. K. Tyagi, Delhi University, Delhi, India, for providing clones containing *psbC* from *Spinacia oleracea* and *psbD* from *Vigna aconitifolia* respectively; Department of Biotechnology, New Delhi, for financial support to carry out this work; and Council of Scientific and Industrial Research, New Delhi, for a senior research fellowship to M.S.S.R.

References

Barkan A., Voelker R., Mendel-Hartvig J., Johnson D. and Walker M. 1995 Genetic analysis of chloroplast biogenesis in higher plants. *Physiol. Plantarum* **93**, 163–170.

Chase M. W. and Palmer J. D. 1989 Chloroplast DNA systematics of lilioid monocots: Resources, feasibility and an example from the Orchidaceae. *Am. J. Bot.* **76**, 1720–1730.

Chen S.-C. G., Lu J.-H., Cheng M.-C., Chen L.-F. O. and Lo C. P.-K. 1994 Four promoters in the rice plastid *psbK–psbI–psbD–psbC* operon. *Plant Sci.* **99**, 171–182.

Christopher D. A., Kim M. and Mullet J. E. 1992 A novel light-regulated promoter is conserved in cereal and dicot chloroplasts. *The Plant Cell* **4**, 785–798.

De Heij, H. T., Lustig H., Moeskops D.-J. M., Bovenberg W. A., Bisanz C. and Grott G. S. P. 1983 Chloroplast DNAs of *Spinacia, Petunia* and *Spirodela* have a similar gene organization. *Curr. Genet.* **7**, 1–6.

Golden S. S., Cho D.-S. C. and Nalty M. S. 1989 Two functional *psbD* genes in the cyanobacterium *Synechococcus* sp. strain PCC7942. *J. Bacteriol.* **171**, 4707–4713.

Habeshausen G., Valentin K. and Zetsche K. 1992 Organization and sequence of photosynthetic genes from the plastid genome of the holoparasitic flowering plant *Cuscuta reflexa. Mol. Gen. Genet.* **232**, 154–161.

Hiratsuka J., Shimada H., Whittier R., Ishibashi T., Sakamoto M., Mori M., Kondo C., Honji Y., Sun C. R., Meng B. Y., Li Y. Q., Kanno A., Nishikawa Y., Hirai A., Shinozaki K. and Sugiura M. 1989 The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217**, 185–194.

Ikemura T. 1985 Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* **2**, 13–34.

Leonhardt K. and Straus N. A. 1994 Photosystem II genes *isiA, psbDI* and *psbC* in *Anabaena* sp. PCC 7120: cloning, sequencing and the transcriptional regulation in iron-stressed and iron-repleted cells. *Plant Mol. Biol.* **24**, 63–73.

Meng B. Y., Wakasugi T. and Sugiura M. 1991 Two promoters within the *psbK–psbI–trnG* gene cluster in tobacco chloroplast DNA. *Curr. Genet.* **20**, 259–264.

Morton B. R. 1993 Chloroplast DNA codon use: evidence for selection at the *psbA* locus based on tRNA availability. *J. Mol. Evol.* **37**, 273–280.

Morton B. R. 1994 Codon use and the rate of divergence of land plant chloroplast genes. *Mol. Biol. Evol.* **11**, 231–238.

Naithani S., Trivedi P. K., Tuli R. and Sane P. V. 1997a The *psbE-F-L-J* operon from chloroplast genome of *Populus deltoides*: cloning, nucleotide sequence and transcript analysis. *J. Genet.* **76**, 61–72.

Naithani S., Trivedi P. K. and Sane P. V. 1997b Characterization of *orf31–petG* gene cluster from plastid genome of *Populus deltoides*. *Biochem. Mol. Biol. Int.* **43**, 433–442.

Nath P., Sane A. P., Bijola V., Trivedi P. K., Arora J. and Sane P. V. 1993 A simple method for purification of organelle DNA of plants. *J. Plant Biochem. Biotechnol.* **2**, 117–120.

Orsat B., Spielmann A., Marc-Martin S., Lemberger T. and Stutz E. 1994 Analysis of the 22 kbp long *psbD–psbC* gene cluster of *Euglena gracilis* chloroplast DNA: evidence for overlapping transcription units undergoing differential processing. *Biochim. Biophys. Acta* **1218**, 75–81.

Pakrasi H. B. and Vermaas W. F. J. 1992 In *The photosystems: structure, function and molecular biology* (ed. J. Barber), pp. 231– 257. Elsevier, Amsterdam.

Rochaix J.-D., Kuchka M., Mayfield S., Schirmer-Rahire M., Girard-Bascou J. and Bennoun P. 1989 Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast *psbC* gene product in *Chlamydomonas reinhardtii*. *EMBO J.* **84**, 1013–1021.

Sambrook T., Fritsch E. F. and Maniatis T. 1989 *Molecular cloning—a laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sexton T. B., Christopher D. A. and Mullet J. E. 1990 Light induced switch in barley *psbD–psbC* promoter utilization: A novel mechanism regulating chloroplast gene expression. *EMBO J.* **9**, 4485–4494.

Sharp P. M. 1991 Determinants of DNA sequence divergence between *Escherichia coli* and *Salmonella typhimurium*: codon usage, map position and concerted evolution. *J. Mol. Evol.* **33**, 23–33.

Sharp P. M. and Li W. H. 1987 The codon adaptation index — a measure of directional synonymous codon usage bias and its potential applications. *Nucl. Acids Res.* **15**, 1281–1295.

Trivedi P. K., Nath P. and Sane P. V. 1994 Characterization of the *psbA* gene from chloroplasts of *Populus deltoides*. *J. Plant Biochem. Biotechnol.* **3**, 97–102.

Van Slogteren C. M. S., Hoge J. H. C., Hooykaas P. J. J. and Schilperoot R. A. 1983 Clonal analysis of heterogeneous crown gall tumor tissues induced by wild-type and shooter mutant strains of *Agrobacterium tumefaciens* expression of T-DNA genes. *Plant Mol. Biol.* **2**, 321–333.

Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T. and Sugiura M. 1994 Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. USA* **91**, 9794–9798.

Wolfe K. H. and Sharp P. M. 1988 Identification of functional open reading frames in chloroplast genomes. *Gene* **66**, 215–222.

Wolfe K. H., Morden C. W. and Palmer J. D. 1992 Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc. Natl. Acad. Sci. USA* **89**, 10648–10652.

Yao W. B., Meng B. Y., Tanaka M. and Sugiura M. 1989 An additional promoter within the protein-coding region of the *psbD–psbC* gene cluster in tobacco chloroplast DNA. *Nucl. Acids Res.* **17**, 9583–9591.

Yu J. and Vermaas W. F. J. 1990 Transcript levels and synthesis of photosystem II components in cyanobacterial mutants with inactivated photosystem II genes. *The Plant Cell* **2**, 315–322.

Received 7 July 1997; in revised form 14 May 1998