Tobacco nuclear gene for the 31 kd chloroplast ribonucleoprotein: genomic organization, sequence analysis and expression

Yuqing Li, Lizhen Ye, Mamoru Sugita and Masahiro Sugiura*

Center for Gene Research, Nagoya University, Chikusa, Nagoya 464-01, Japan

Received February 20, 1991; Revised and Accepted May 1, 1991

EMBL accession no. X57079

ABSTRACT

We have previously identified three chloroplast ribonucleoproteins and characterized their cDNAs. Here we present the genomic organization, sequence and expression of one of their genes. The 31 kd ribonucleoprotein (cp31) from tobacco (*Nicotiana sylvestris*) chloroplasts is coded for by a single-copy nuclear gene. This gene was isolated and its sequence was determined. The gene contains four exons and three introns. The position of its first intron is conserved among the genes for the maize abscisic acid-induced glycine-rich protein, the human hnRNP A1 protein and cp31. The transcription start site was determined to be 168 bp upstream from the translational initiation codon in both leaf and root tissues. No alternatively spliced transcripts was detected, suggesting that a diversity of chloroplast ribonucleoproteins is generated probably by gene amplification rather than alternative splicing.

INTRODUCTION

Chloroplast gene expression is light and developmentally regulated both at the transcriptional and post-transcriptional levels (1,2). About 20% of the chloroplast protein species are coded for by chloroplast DNA and the remaining majority are encoded in nuclear DNA (3,4). Thus both the chloroplast and nuclear genomes are involved in the regulation of chloroplast gene expression. For example, a couple of nuclear genes have been shown to affect the stability of chloroplast *psbD* transcripts in barley and *Chlamydomonas* (5,6). The characterization of nuclear-encoded chloroplast proteins involved in chloroplast gene expression is essential to understand the regulatory interaction between the nuclear and chloroplast genomes.

We have previously isolated three ribonucleoproteins of 28 kd, 31 kd and 33 kd (cp28, cp31 and cp33, respectively) from tobacco chloroplasts and characterized their cDNAs (7). The three proteins have high homology to each other and contain two ribonucleoprotein consensus sequences (RNP-CS, for review see 8). They were postulated to be involved in the splicing and/or processing of chloroplast pre-RNAs. The RNP-CS-type RNA-

binding domain, or the consensus sequence-type RNA-binding domain (CS-RBD, 9), has been shown to be the functional unit in RNA-binding (10–12). All three proteins are members of a large group of RNP-CS-type RNA-binding proteins, which are involved in splicing, polyadenylation, ribosomal biogenesis and sex determination (8,13).

Alternative splicing of pre-mRNAs has been proposed to generate a diversity of proteins containing CS-RBD. The human hnRNP protein A1 binds specifically to the 3' splice site *in vitro* and is probably involved in splicing (14,15). The A1 pre-mRNA is alternatively spliced to produce another protein A1^B (16). In addition, alternatively spliced mRNAs from *Drosophila* sex-determining gene (*tra-2*) and *Xenopus* gene for nervous system-specific RNP protein-1 (*nrp-1*) have been reported (17,18). These results led us to study the nuclear genes for tobacco chloroplast ribonucleoproteins.

Here we present the genomic organization, sequence and expression of a tobacco nuclear gene for one of the three chloroplast ribonucleoproteins, cp31. Genomic Southern analysis indicated that it is coded for by a single-copy gene. Sequence analysis revealed that the gene contains three introns. The position of the first intron is identical with those in the genes for the maize abscisic acid-induced glycine-rich protein (AAIP) (19,20) and the human hnRNP A1 protein. Ribonuclease protection assays showed that the transcription start site and the splicing of its pre-mRNA are the same in both leaf and root.

MATERIALS AND METHODS

DNA isolation and genomic Southern analysis

Nuclei were isolated according to Jofuku and Goldberg (21) from 35 g mature leaves of *Nicotiana sylvestris*. They were suspended in 10 ml TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), mixed gently with 20 ml of DNA extraction buffer (100 mM Tris-HCl, pH 9.4, 2% SDS, 10 mM EDTA, 100 mM NaCl) for 3 to 5 min at room temperature and extracted twice with 30 ml phenol:chloroform:isoamyl alcohol (25:24:1) for 20 min. The aqueous layer was mixed with 2 volumes of ethanol. Flocculated DNA fibers were rinsed with cold 70% ethanol, dried *in vacuo* briefly and dissolved in 1 ml TE buffer (440 µg/ml).

* To whom correspondence should be addressed

Ten μg of the genomic DNA in 100 μl was cut with a given restriction enzyme for 3 hr, phenol extracted and precipitated with ethanol. The DNA was dissolved and digested again with the same enzyme. The resultant DNA fragments were separated by 1.0% agarose gel electrophoresis, transferred to Hybond-N nylon membrane (Amersham) by capillary blot for at least 20 hr and fixed by UV irradiation. The cp31 cDNA was labeled by the random primer method (22) to specific activities of 5×10^8 to 2×10^9 cpm/ μg . Prehybridization, hybridization and washing of the membrane were done essentially according to the Amersham's protocol. The final wash was conducted at 65°C in $0.1 \times \text{SSPE}$ containing 0.1% SDS.

Gene isolation and sequence analysis

N.sylvestris genomic DNA was further purified by a CsCl centrifugation. A genomic library was constructed in λ dash (Stratagene) according to the standard method (23) and had 6.1×10^5 independent clones before amplification. The amplified library was screened according to the standard procedure (23,24). Recombinant λ DNAs were prepared by a mini-lysate method (25). The inserts were cut with *EcoRI* or *Sall* and subcloned into Bluescript M13(+) vectors. Sequencing and computer-assisted analysis were essentially as described (26).

RNA isolation and primer extension analysis

Oligonucleotide 31k4, 5'-GCCTCTCTACGGAAGGGTTTCG-AGTAGAAG3', was labeled with ^{32}P using T4 polynucleotide kinase (23). Leaf and root total RNAs were prepared as described (7) and hybridized overnight at 30°C with the above primer (1×10^5 cpm). Primer extension was done essentially according to Sambrook *et al* (23) at 42°C in 25 μl reaction mixture containing 50 mM Tris-HCl (pH 8.3), 55 mM KCl, 3 mM MgCl_2 , 10 mM DTT, 0.5 mM each of dATP, TTP, dGTP and dCTP, 1 unit of Inhibit-ACE (5 Prime $\rightarrow$ 3 Prime, Inc.) and 400 units of Moloney murine leukemia virus reverse transcriptase (BRL). Molecular weight markers were sequencing ladders of the corresponding DNA region using the same primer.

Ribonuclease protection assay

Plasmid DNAs of the deletion clones 10, 11, 13, 23 and clone pNS31KGB were linearized with *HindIII*, *Tth111I*, *NcoI*, *HincII* and *HpaI*, respectively. Radioactive RNA probes were synthesized in 25 μl reaction mixture containing 40 mM Tris-

Figure 1. Southern blot analysis of *N.sylvestris* nuclear DNA digested with *EcoRI* (E), *EcoRV* (V), *HindIII* (H) and *XbaI* (X). Probes were a near full-length cp31 cDNA (full, positions 15–1201), its 5' part (5', positions 154–827) and its 3' part (3', positions 827–1201, positions refer to Figure 2b of ref. 7). Numbers on right are kb.

HCl (pH 8.0), 8 mM MgCl_2 , 2 mM spermidine, 50 mM NaCl, 30 mM DTT, 1 unit of Inhibit-ACE (5 Prime $\rightarrow$ 3 Prime, Inc.), 400 μM each of rATP, rGTP and rCTP, 6 μM rUTP, 2.5 μl of $\alpha\text{-}^{32}\text{P}$ -rUTP (800 Ci/mmol, 10 mCi/ml), 0.2–1.0 μg of template DNA and 10 units of T7 RNA polymerase. The RNA probes were then purified by a 5% polyacrylamide gel containing 7.5 M urea. Leaf and root RNAs were treated with RNase-free DNase to remove a trace amount of contaminating DNA. Ribonuclease protection assay (27,28) was done essentially according to the instruction manual of a ribonuclease protection assay (RPATM) kit from Ambion, Inc. USA. The hybridized RNA was digested with 0.1 unit of RNase A and 20 units of RNase T1 at 37°C for 30 min. Digestion was conducted at 16°C for probe B to reduce a background due to 'breathing' of AT-rich sequence (see Figure 3, nucleotides 2773–2794). The protected fragments were separated on a 5% polyacrylamide gel containing 7.5 M urea.

RESULTS AND DISCUSSION

Genomic organization

We first analyzed the genomic organization of the cp31 gene in *N.sylvestris*, the female progenitor of *N.tabacum*. *N.sylvestris* genomic DNA was digested with four restriction enzymes and probed with the entire cDNA of cp31. In each case, one or two dense bands (*EcoRI*: 4.8 kb; *EcoRV*: ~24 kb; *HindIII*: 2.8 kb

Figure 2. Restriction map of an insert in a λ Dash clone, λ NS31KG83, containing the gene for cp31. E: *EcoRI*; H: *HindIII*; S: *Sall*; V: *EcoRV*; X: *XbaI*. The map was constructed by restriction enzyme analysis and sequencing. Two additional *HindIII* sites located between the first *HinIII* and first *EcoRI* sites from left remain to be mapped. Thin lines above the map indicate subcloned fragments (plasmid names on right). The sequenced region is enlarged below the map. Open boxes represent exons and protein domains. TS: transit peptide; AD: acidic domain. Arrows indicate both the position and direction of RNA probes used for ribonuclease protection assays (Figures 5B and 6) and the regions covered are: 23, –224 to 586; 13, 207 to 847; 11, 804 to 1219; 10, 1047 to 1987; B, 1873 to 2850 (numbers correspond to those in Figure 3).

and 1.4 kb; *Xba*I: ~12 kb; Figure 1, full) were detected together with several weak bands (the weak bands produced by *Eco*RV digestion clustered below the major 24 kb band). Some of the weak bands corresponded to the bands detected by the cp28 cDNA (data not shown), which showed high homology to the cp31 cDNA (7). This indicates that the weak bands were resulted from the cross-hybridization of the cp31 probe to other genes coding for proteins containing CS-RBD. The results suggest that cp31 is coded for by a single locus.

To determine the exact copy number, we probed the genomic blot with 5' and 3' parts of the cDNA. Both probes produced single bands from *Eco*RI, *Eco*RV and *Xba*I digests (Figure 1). The 5' part hybridized to a 2.8 kb *Hind*III band while the 3' part hybridized to a 1.4 kb *Hind*III band. In addition, no weak bands were detected. Taken together, we concluded that cp31 is encoded by a single-copy gene in the nuclear genome of *N.sylvestris*.

The genomic organization of the tobacco cp31 gene is similar to those of animal U1 snRNP 70 K protein genes and a mouse

nucleolin gene (29–32), while different from those of human hnRNP A1 and C protein genes which constitute multi-gene families (33,34). Processed pseudo-genes for A1 have also been detected (33). It is also different from the organization *Xenopus* U1 snRNP 70K protein genes, which have at least two non-allelic genes (35).

Structure of the nuclear cp31 gene

Eight positive clones were obtained after screening 1.4×10^6 clones with the cp31 cDNA probe (7). Clone λ NS31KG83 was shown to contain an intact gene by hybridization with the cDNA probe (data not shown). A 4.8 kb *Eco*RI fragment was subcloned (pNS31KGB, Figure 2) and sequenced. The 3' flanking region of the gene was sequenced using the pNS31KG2 plasmid with a synthetic oligonucleotide. The sequenced region (5040 bp) consists of a 2098 bp 5' flanking region, the structural gene (2545 bp containing three introns) and a 397 bp 3' flanking region (Figure 3).

Figure 3. Portions of the cp31 gene sequence together with the predicted amino acid sequence. Uppercase letters indicate exon and flanking sequences. Lowercase letters indicate introns. A filled triangle shows the transcription start site. Putative TATA and CAAT boxes are underlined. The transit peptide sequence is italicized, the acidic domain is boxed and two CS-RBDs are stippled.

Figure 4. Comparison of the intron sites in CS-RBDs from tobacco cp31, maize AAIP (19), human hnRNP A1 protein (37), mouse nucleolin (32) and *Xenopus* U1 snRNP 70 K protein (35). Open triangles indicate intron sites interrupting codons and filled triangles indicate introns inserted between codons. Among four sequenced genes for U1 snRNP 70K proteins, murine and human have the same intron sites in CS-RBDs as *Xenopus* and *Drosophila* has no intron in CS-RBD (29–31,35).

Figure 5. Determination of the transcription start site of the cp31 gene. **A.** Primer extension analysis. Extended products (arrow) in parallel with the ladders of the corresponding DNA. L: leaf RNA; R: root RNA. **B.** Ribonuclease protection assay. Protected bands from 5 μ g leaf RNA (L) and 20 μ g root RNA (R) hybridized with RNA probe 23 (~224 bp ~ 586 bp). Size markers (denatured ϕ X174 RF-DNA HincII digest) are on right in nt. The experimental design is illustrated on right.

The boundary sequences of all three introns fit the consensus sequences (36). The predicted precursor protein is composed of four domains: a transit peptide, an acidic amino terminal domain and two CS-RBDs. The first and second introns are located in the first CS-RBD while the third intron is in the second CS-RBD (Figures 2 and 3). There is no intron in the region coding for the transit peptide and the acidic amino terminal domain. The second and third introns interrupt the two CS-RBDs at the same position (Figure 4), suggesting that the occurrence of an intron within an ancestral gene encoding CS-RBD preceded the duplication of this domain. In contrast to the human hnRNP A1 and mouse nucleolin genes (37,32), there is no intron between the first and second CS-RBDs.

The positions of three cp31 introns were compared with those of several other proteins containing CS-RBD. Surprisingly, the position of the first intron is precisely the same as those of the maize AAIP and human hnRNP A1 genes (Figure 4). Two main hypotheses have been proposed for the origin of introns. One hypothesis is that introns are as ancient as genes and the present prokaryotes have lost most of them (38–40). The second hypothesis is that introns are of more recent origin and were inserted into eukaryotic genes after the divergence of the prokaryote and eukaryote lineages (41,42). Our results support the first hypothesis.

The origin of chloroplast ribonucleoproteins is intriguing. According to the endosymbiotic theory, chloroplasts are thought to descend from cyanobacteria (43). Later a vast gene transfer to the nucleus occurred (44,45) and in some cases symbiotic genes were lost entirely and functionally replaced by pre-existing nuclear genes which had acquired transit peptides to direct their

Figure 6. Expression of the cp31 gene in leaf (L: 5 μ g RNA; L*: 10 μ g RNA) and root (R: 20 μ g RNA) tissues as revealed by ribonuclease protection assay. The size (nt) of protected fragments are on right. The experimental design is shown below. Arrows indicate the positions of RNA probes (see also Figure 2) and solid lines with nt numbers below indicate protected fragments.

products to chloroplasts (46). Therefore there are at least two possibilities for the origin of chloroplast ribonucleoproteins: nuclear origin or cyanobacterial origin. Using a mixture of synthetic oligonucleotides corresponding to the RNP-CS octapeptides as probes, we failed to isolate any positive clone from a cyanobacterial (*Anacystis nidulans*) genomic library (47) even if the number of clones screened were approximately 40-fold of that required for the cyanobacterial genome (data not shown). This indicates the absence of chloroplast ribonucleoprotein counterparts in the present cyanobacterium. Although we can not rule out the possibility that chloroplast ribonucleoproteins were originated from the ancestral cyanobacterium and the present cyanobacterium has lost the genes, we favor the possibility that they are of nuclear origin.

Transcription start site

Primer extension analysis showed that the transcription start site in both leaf and root is 168 bp upstream from the putative translation initiation codon (Figure 5A). The start site was verified by ribonuclease protection assay. An RNA probe spanning the above start site was protected by both tobacco leaf and root RNAs to yield a single protected fragment, the size of which matches the primer extension datum (Figure 5B).

Putative TATA and CAAT elements were located in the upstream region through comparison with the plant consensus sequence (48, Figure 3). A repeated sequence (TA)₂₂ is observed 1271–1314 bp upstream from the start site. Kay *et al* found a similar repeat of (TA)₂₁ in the upstream region of the rice phytochrome gene (49). The physiological role of the TA repeat remains to be investigated.

Expression of the cp31 gene

We carried out ribonuclease protection assays using both tobacco leaf and root RNAs. Leaf represents chloroplasts and root represents non-photosynthetic plastids. Four RNA probes were prepared to cover the entire sequence of the gene. Figure 6 shows

that there is no qualitative difference between leaf and root RNAs in all regions studied. The band intensity from 5 μ g of leaf RNA is roughly equal to that from 20 μ g of root RNA, indicating a 4-fold difference in the steady-state transcript level between leaf and root.

The size of protected fragments fits those calculated from the cp31 cDNA sequence (7). A weak but reproducible protected fragment of 128 nucleotides (nt) was also detected by the clone 10 probe in addition to the 282 nt fragment expected from the cDNA sequence (Figure 6). By using a sense-strand transcript from the cp28 cDNA (7), we could show that this 128 nt fragment was resulted from the protection by the cp28 mRNA (data not shown). To rule out alternative splicing events due to the possible presence of mini exons, we carried out a further protection assay using an anti-sense transcript made from the cp31 cDNA and found no alternatively spliced transcripts both in leaf and root tissues (data not shown). Unlike the human hnRNP proteins whose diversity has been mainly generated by alternative splicing (9,16,50,51), a diversity of chloroplast ribonucleoproteins might be generated by the amplification of a proto-gene.

Although the functional homology to hnRNP proteins has not been demonstrated, our results may imply that a diversity of chloroplast ribonucleoproteins is important for their functions. Relevant to this, Buvoli *et al* (16) have pointed out that polymorphism in hnRNP structures could reflect a functional specialization in hnRNA metabolism. Kay *et al* have suggested that the synthesis of *Xenopus* A1 protein isoforms are developmentally regulated and/or that the isoforms are selectively incorporated into hnRNPs to alter RNA stability or splicing efficiency (51). A new member of the A and B hnRNP protein group has recently been identified (52). In addition, two highly homologous but different AAIP-like protein cDNAs have recently been found from sorghum, each contains a CS-RBD (53). These predicted proteins are probably ribonucleoproteins because at least maize AAIP binds to poly(G) at high salt concentrations (54). All these observations suggest a diversity of ribonucleoproteins.

ACKNOWLEDGMENTS

We thank Drs.T.Wakasugi, M.Gō, Dr.K.Fukami-Kobayashi, B.Y.Meng, M.Ohme for their help and stimulating discussion. Y.Li and L.Ye were supported by Monbusho pre-doctoral fellowships. This research was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture (Japan).

REFERENCES

- Mullet, J.E. (1988) *Annu. Rev. Plant Physiol. Plant Molec. Biol.*, **39**, 475–502.
- Gruissem, W. (1989) *Cell*, **56**, 161–170.
- Sugiura, M. (1989) *Annu. Rev. Cell Biol.*, **5**, 51–70.
- Palmer, J.D. (1990) *Trends Genet.*, **6**, 115–120.
- Gamble, P.E. and Mullet, J.E. (1989) *EMBO J.*, **8**, 2785–2794.
- Kuchka, M.R., Goldschmidt-Clermont, M., van Dillewijn, J. and Rochaix, J.-D. (1989) *Cell*, **58**, 869–876.
- Li, Y. and Sugiura, M. (1990) *EMBO J.*, **9**, 3059–3066.
- Bandziulis, R.J., Swanson, M.S. and Dreyfuss, G. (1989) *Genes Develop.*, **3**, 431–437.
- Burd, C.G., Swanson, M.S., Grolach, M. and Dreyfuss, G. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9788–9792.
- Query, C.C., Bentley, R.C. and Keene, J.D. (1989) *Cell*, **57**, 89–101.
- Scherly, D., Boelens, W., van Venrooij, W.J., Dathan, N.A., Hamm, J. and Mattaj, J.W. (1989) *EMBO J.*, **8**, 4163–4170.
- Surowy, C.S., van Santen, V.L., Scheib-wixted, S.M. and Spritz, R.A. (1989) *Mol. Cell. Biol.*, **9**, 4179–4186.
- Mattaj, J.W. (1989) *Cell*, **57**, 1–3.
- Swanson, M.S. and Dreyfuss, G. (1988) *EMBO J.*, **7**, 3519–3529.
- Buvoli, M., Cobianchi, F., Biamonti, G. and Riva, S. (1990) *Nucl. Acids Res.*, **18**, 6595–6600.
- Buvoli, M., Cobianchi, F., Bestagno, M.G., Mangiarotti, A., Bassi, M.T., Biamonti, G. and Riva, S. (1990) *EMBO J.*, **9**, 1229–1235.
- Amrein, H., Maniatis, T. and Nothiger, R. (1990) *EMBO J.*, **9**, 3619–3629.
- Richter, K., Good, P.J. and Dawid, I.B. (1990) *The New Biologist*, **2**, 556–565.
- Gomez, J., Sanchez-Martinez, D., Stiefel, V., Rigau, J., Puigdomenech, P. and Pages, M. (1988) *Nature*, **334**, 262–264.
- Mortenson, E. and Dreyfuss, G. (1989) *Nature*, **337**, 312.
- Jofuku, K.D. and Goldberg, R.B. (1988) In Shaw, C.H. (ed.) *Plant Molecular Biology—A Practical Approach*. IRL Press, Oxford, pp.37–66.
- Feinberg, A.P. and Vogelstein, B. (1984) *Anal. Biochem.*, **137**, 266–267.
- Sambrook, J., Maniatis, T. and Fritsch, E.F. (1989) *Molecular Cloning*. 2nd ed. Cold Spring Harbor Laboratories, New York.
- Slightom, J.L. and Drong, R.F. (1988) In Gelvin, S.B., Schilperoort, R.A. and Verma, D.P.S. (ed.), *Plant Molecular Biology Manual*. Kluwer Academic Publishers, Dordrecht, A8:1–42.
- Manfioletti, G. and Schneider, C. (1988) *Nucl. Acids Res.*, **16**, 2873–2884.
- Hiratsuka, J., Shimada, H., Whittier, R.F., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, M., Honji, Y., Sun, C.R., Meng, B.Y., Li, Y.Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989) *Mol. Gen. Genet.*, **217**, 185–194.
- Zinn, K., DiMaio, D. and Maniatis, T. (1983) *Cell*, **34**, 865–879.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucl. Acids Res.*, **12**, 7035–7056.
- Spritz, R.A., Strunk, K., Surowy, C.S. and Mohrenweiser, H.W. (1990) *Genomics*, **8**, 371–379.
- Mancebo, R., Lo, P.C.H. and Mount, S.M. (1990) *Mol. Cell. Biol.*, **10**, 2492–2502.
- Hornig, H., Fischer, U., Costas, M., Rauh, A. and Luhrmann, R. (1989) *Eur. J. Biochem.*, **182**, 45–50.
- Bourbon, H., Lapeyre, B. and Amalric, F. (1988) *J. Mol. Biol.*, **200**, 627–638.
- Buvoli, M., Biamonti, G., Ghetti, A., Riva, S., Bassi, M.T. and Morandi, C. (1988) *Nucl. Acids Res.*, **16**, 3751–3770.
- Nakagawa, T.Y., Swanson, M.S., Wold, B. and Dreyfuss, J. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2007–2011.
- Etzerodt, M., Vignali, R., Ciliberto, G., Scherly, D., Mattaj, J.W. and Philipson, L. (1988) *EMBO J.*, **7**, 4311–4321.
- Brown, J.W.S., Feix, G. and Frendewey, D. (1986) *EMBO J.*, **5**, 2749–2758.
- Biamonti, G., Buvoli, M., Bassi, M.T., Morandi, C., Cobianchi, F. and Riva, S. (1989) *J. Mol. Biol.*, **207**, 491–503.
- Darnell, J.E. and Doolittle, W.F. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1271–1275.
- Gilbert, W., Marchionni, M. and McKnight, G. (1986) *Cell*, **46**, 151–154.
- Tomioka, N. and Sugiura, M. (1984) *Mol. Gen. Genet.*, **193**, 427–430.
- Cavalier-Smith, T. (1985) *Nature*, **315**, 283–284.
- Dibb, N.J. and Newman, A.J. (1989) *EMBO J.*, **8**, 2015–2021.
- Margulis, L. (1981) *Symbiosis in Cell Evolution*. W.H. Freeman and Co., San Francisco.
- Shih, M.-C., Lazar, G. and Goodman, H.M. (1986) *Cell*, **47**, 73–80.
- Baldauf, S.L. and Palmer, J.D. (1990) *Nature*, **344**, 262–265.
- Tingey, S.V., Tsai, F.-Y., Edwards, J.W., Walker, E.L. and Coruzzi, G.M. (1988) *J. Biol. Chem.*, **263**, 9651–9657.
- Meng, B.Y., Shinozaki, K. and Sugiura, M. (1989) *Mol. Gen. Genet.*, **216**, 25–30.
- Grierson, D. and Covey, S.N. (1988) *Plant Molecular Biology*. 2nd ed., Blackie and Son Ltd, Glasgow, pp.35.
- Kay, S.A., Keith, B., Shinozaki, K., Chye, M.-L. and Chua, N.-H. (1989) *Plant Cell*, **1**, 351–360.
- Merrill, B.M., Barnett, S.F., LeStourgeon, W.M. and Williams, K.R. (1989) *Nucl. Acids Res.*, **17**, 8441–8449.
- Kay, B.K., Sawhney, R.K. and Wilson, S.H. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1367–1371.
- Haynes, S.R., Johnson, D., Raychaudhuri, G. and Beyer, A.L. (1991) *Nucl. Acids Res.*, **19**, 25–31.
- Cretin, C. and Puigdomenech, P. (1990) *Plant Mol. Biol.*, **15**, 783–785.
- Skriver, K. and Mundy, J. (1990) *Plant Cell*, **2**, 503–512.