Mol Gen Genet (1989) 215:337–344

MGG

Novel *cis*-acting elements in Petunia *Cab* gene promoters

David Gidoni*, Paul Brosio, Diane Bond-Nutter, John Bedbrook, and Pamela Dunsmuir
Advanced Genetic Sciences, Inc., 6701 San Pablo Avenue, Oakland, CA 94611, USA

Summary. In order to identify specific *cis*-acting elements which regulate the expression of the divergent *Cab22R* and *Cab22L* genes of Petunia, we conducted systematic mutational studies of the 1 kb intergenic promoter region. Sequence analysis revealed three GATA box sequence repeats positioned between the TATA and CAAT box elements. These GATA elements are conserved in corresponding promoter regions of all LHCII Type I *Cab* genes in Petunia and other dicotyledonous plants we have examined. Site-specific mutations in the CAAT box and the GATA box elements of the *Cab22R* promoter resulted in 8-fold and 5-fold reductions in *Cab22R* transcript levels respectively. A deletion of 52 bp, adjacent and upstream from the CAAT box (−92 to −145) in the *Cab22R* promoter reduced transcript levels 20-fold. This deletion contains a region of 13 bp which is conserved between many Petunia *Cab* genes. These results indicate that the quantitative expression of the *Cab22* promoters is regulated by multiple *cis*-acting elements including CAAT and GATA box elements as well as sequences located between −92 and −145. The deletion of the region between −92 and −145 is partially compensated by homologous sequences present in the adjacent divergent promoter *Cab22L*.

Key words: Divergent promoters – Chlorophyll a/b binding protein genes – *cis*-acting elements – RNA expression

Introduction

The induction or suppression of gene activity during development of eukaryotic organisms often occurs at the transcriptional level. Transcription is regulated by specific interactions between *trans*-acting protein factors and *cis*-acting promoter and enhancer sequences (reviewed in Dynan and Tjian 1985; McKnight and Tjian 1986). Mammalian genes contain in their promoter regions the sequence motifs CAAT and TATA that are involved in determination of both the level and position of transcription initiation. In addition to such sequences there are other elements that modulate transcription in a promoter-specific fashion. Thus, the promoter is comprised of a complex of *cis*-acting elements which can be activated by corresponding *trans*-acting factors with the result that expression of the adjacent gene is either constitutive, induced by external factors, tissue specific or some combination of these (reviewed in Dynan and Tjian 1985; McKnight and Tjian 1986; Maniatis et al. 1987).

Many plant genes also contain CAAT and TATA sequences in their promoters (Messing et al. 1983). In addition, potential promoter-specific *cis*-acting elements have been identified as highly conserved sequences in the promoters of specific classes of plant genes such as the storage protein of legumes (Baumlein et al. 1986) and cereals (Forde et al. 1985). In maize, the identified upstream sequence of the zein gene is known to interact specifically with a nuclear protein factor which is postulated to be involved in the developmental regulation of zein gene expression (Maier et al. 1987).

Studies to identify and test specific *cis*-acting elements in plant promoter regions are possible using either direct gene transfer techniques (Fromm et al. 1985; Potrykus et al. 1985) or *Agrobacterium tumefaciens* T-DNA as a vector to reintroduce foreign DNA into the chromosomes of dicotyledonous plant cells (Hoekema et al. 1983). The transformed plant cells are regenerated into plants where the expression of the introduced genes (and mutagenized variants) can be analyzed. Such studies have defined enhancer-like elements in the promoter region of the 35S gene of cauliflower mosaic virus (Kay et al. 1987), the chlorophyll a/b binding protein (*Cab*) gene (Simpson et al. 1986), the ribulose 1,5-bisphosphate carboxylase small subunit (*rbcS*) gene from pea (Timko et al. 1985; Fluhr et al. 1986; Kuhlemeier et al. 1987), and the β-conglycinin gene of soybean (Chen et al. 1988). These elements appear necessary for light regulated and tissue-specific expression of the *Cab* and *rbcS* genes, as well as tissue-specific and temporal expression of β-conglycinin. In addition, *cis*-acting elements of the chalcone synthase gene have been shown to respond to UV light (Kaulen et al. 1986), those of the soybean heat shock gene, to heat shock (Baumann et al. 1987) and those of the maize *Adh-I* gene to anaerobiosis (Ellis 1987).

In a recent study, we characterized the expression properties of two divergent promoters derived from adjacent *Cab* genes of Petunia – *Cab22R* and *Cab22L* (Gidoni et al. 1988). Our studies indicated that these promoters are coordinately expressed and contain the *cis*-acting elements which are sufficient for directing coordinated, high level, light-regulated, and tissue-specific expression of foreign gene coding regions in transgenic tobacco plants. Here we describe a mutational analysis of these two adjacent promoters.

* *Present address:* Agricultural Research Organization, The Volcani Center, Institute of Field and Garden Crops, P.O. Box 6, Bet Dagan 50250, Israel

Offprint requests to: P. Dunsmuir

Materials and methods

DNA modifications and cloning. Molecular cloning procedures were carried out as described in Maniatis et al. (1982). Restriction endonucleases and T4 DNA ligase were from Promega, and Bal31 endonuclease was from New England Biolabs. DNA polymerase I large fragment (Klenow) was from Boehringer-Mannheim Biochemicals, T4 polynucleotide kinase was from P-L Biochemicals and [γ-^{32}P]ATP (5000 Ci/mmol) was from Amersham. Oligonucleotides were synthesized on an Applied Biosystems 381A DNA synthesizer. Oligonucleotide site-directed mutagenesis was according to Kunkel (1985); mutations were confirmed by dideoxy sequencing (Sanger et al. 1977).

DNA constructions and deletion mutagenesis. The starting plasmid for constructing deletions in the *Cab22* divergent promoter region was pBRN/B. This plasmid was previously constructed for rapid linker-scanner mutagenesis (Haltiner et al. 1985). We introduced two modifications into this plasmid. First we inserted the M13 origin region: the 880 bp *Ball-Pvu*II fragment from M13mp18 (Messing 1983) was inserted into the *Cla*I site of pBRN/B. Second, to protect the *Nru*I/*Bgl*II (N/B) cassette from Bal31 exonuclease digestion, the 1.1 kb *Sma*I-*Pvu*II DNA fragment of the octopine synthase (*ocs*) gene (DeGreve et al. 1983) was isolated and introduced into the *Bam*HI site of pBRN/B. The resultant vector, designated as pBRN/Bm, was linearized by *Hin*dIII digestion, treated with DNA polymerase I large fragment (Klenow) to produce blunt ends, and ligated with a genomic *Eco*RI-*Xho*I fragment (also treated with DNA polymerase I), containing the intact *Cab22L* and *Cab22R* genes. This ligation generated the *Cab22*-A and *Cab22*-B plasmids shown in Fig. 3. Following the linearization of *Cab22*-A with *Hin*dIII and *Cab22*-B with BamHI, each plasmid was treated with Bal31 endonuclease, and a series of 5′ and 3′ deletion mutants was generated. Each population of Bal31-digested DNA was cleaved with *Nru*I, treated with DNA polymerase I large subunit to generate blunt ends, then recircularized and transformed into *Escherichia coli* JM101, where both double-stranded and single-stranded DNAs were generated. The approximate size of each deletion was assessed by restriction analysis and its endpoint determined by dideoxy sequencing using a synthetic primer which hybridized to the pBR322 region at a position adjacent to the N/B cassette.

To reconstruct the divergent *Cab22* region containing the internal deletions, *Cab22*-A and *Cab22*-B deletion plasmids were digested with *Bgl*II, and ligated together to form the promoter deletion mutant plasmids as shown in Fig. 3. Each of the reconstituted *Cab22* plasmids containing the promoter mutations, and the wild-type *Cab22* gene control, were inserted into the *Cla*I site of the pAGS140 binary vector (C. Dean, M. Favreau, S. Tamaki, D. Bond-Nutter, P. Dunsmuir, and J. Bedbrook, submitted). To simplify the interpretation of the results, we selected recombinant clones containing the inserted *Cab* DNA in one orientation only. Thus in all binary constructions studied, the *Cab22* DNA was oriented such that *Cab22R* gene was transcribed toward the left border.

Transformation of tobacco plants. The binary vectors were introduced into *E. coli* HB101 then mobilized to *A. tumefaciens* strain LBA4404 (Hoekema et al. 1983) by triparental mating with *E. coli* HB101/pRK2013 (Figurski and Helinski 1979). *Agrobacterium* clones were used to transform leaf cells of *Nicotiana tabacum* cv. W38 by leaf disk inoculation (Horsch et al. 1985). Kanamycin-resistant plantlets were transferred to the greenhouse, and after 6 weeks, at the 12–16 leaf stage, RNA was extracted from the top two leaves of each plant. We have found it important to standardize the harvesting time when considering the photosynthetic genes which are developmentally and light regulated and subject to diurnal fluctuation effects (Kloppstech 1985; Dunsmuir et al. 1988a).

RNA extraction and primer extension analysis. Total RNA was extracted from plant tissue as previously described (Dunsmuir et al. 1988b). The transcript levels of the *Cab22R* and *Cab22L* genes were determined by the primer extension method of McKnight et al. (1981) using gene-specific, synthetic oligonucleotide primers, with previously described modifications (Dunsmuir et al. 1988; Gidoni et al. 1988). The X-ray autoradiograms were scanned by an LKB scanning densitometer.

Results

Sequence analysis of the Cab22 *intergenic region*

The sequence of the 1.1 kb DNA region, containing the divergent *Cab22L* and *Cab22R* promoters, is shown in Fig. 1. In each promoter the translation initiation codon, the transcription initiation sites, the TATA and the CAAT boxes are shown. Sequence comparison between the *Cab22L* and *Cab22R* promoters indicates that there are four conserved regions upstream of the CAAT box, (numbered 1–4), and three GATA boxes downstream of the CAAT box (Fig. 1).

Comparison of these sequences with the promoter regions for *Cab* genes in several different plant species indicates two or three GATA sequence repeats, between the CAAT box and the TATA box, in all LHCII Type I *Cab* genes which we have examined (Fig. 2). The conservation of these GATA sequences and their spatial arrangement in the promoters of this class of *Cab* gene in different plant species suggest a possible role in the regulation of transcription of these genes. However this motif is absent in the LHCII Type II genes (Stayton et al. 1986), the LHCI genes (Pichersky et al. 1987, 1988) and the *rbcS* genes (Dean et al. 1985).

Cis-*acting control elements in the* Cab *promoter*

To test the importance of these potential promoter regulatory sequences identified by sequence analysis, we prepared a series of mutant *Cab22* promoters in transgenic W38 tobacco plants. Using oligonucleotide-specific mutagenesis, we made the following three mutations in the *Cab22R* promoter: (a) in the GATA-1 mutant, the GATA-1 box sequence was changed to CTTA; (b) in the GATA-123 mutant, the sequence of all three GATA boxes was changed to CCCA; and (c) in the CAAT mutant, the CAAT box sequence was modified from CCACCAATT to CCCGGGATT (Fig. 4A).

In addition to the site-specific promoter mutations in the *Cab22R* promoter, we used the procedure diagrammed in Fig. 3 and detailed in Materials and methods to create

THE SEQUENCE OF CAB22L PROMOTER

```
                                                    GTACATTAACTTTTCACATTT
                                          4
GAAATAAAAGTGGGGGTCTAAATAATTTTCTTGAATACGAGAGGAAGAGAGGCGAAACAT
TAAAAGAACATTCATGTTATTCAAATTAACAAAGGAATCTTGAAAGTAAACAAAAAGAAT
CCTTTTGCATCCTTCATTTTCACACGTCTCATAGACAAGTTCATGTCTTAATTTTGTGGT
GCATTTACAGATTTAAAGCATGTATTGAGACTAAATTCAGAAAGGTAAAACTGTAAAATG
                                    3
TCTATTACCAGGTAAAACCCACCAGACACATTGACAATATAATCTTGCCAGAAATATCTT
ACTGCGTATCTTATAGGTCAGTTTGCAATCGTCAAACTAAAGCCACATATTGATTGGACC
                2                                   1
CCAACAGTAAAAATATCTTAGGAAGACTGATGGATTATAGAACGACAAGTAGACAAGACA
                          III          II          I
TAACGCTAATGGAATCTTGAAACCCAATGAAATTGTAGATAGAGATATCATAAGATAAGA
ATCTTTCCATCCCTCCCCTATATATACACTTAAGACCATCTGGCCTAACAGCACCACAGC
TCATTTCTCTATTACTTCAGCCATAACAAAAGAACTCTTTTCTCTTCTTATTAAACCATG
```

THE SEQUENCE OF CAB22R PROMOTER

```
                                                    GTATGTGCTGTTGAACTTGTAGTATCAAA
CGTTCAATTGAAATCATAGTTAAAAGTTAATCATGAGAGCTTAGCTAACTGTTGGGACAC
                4
TTGGACTGAAATTTTCTTACTTACACTTTTATATTTTTCTGTTCTTTCTCTAACATTTGT
          3
TCTCATTGACAATTCACCACACATATGAGTGGTTCGCTAGTTCGATATGGCCATGAGTTG
AGATTATATATGCTTTGGCCAAGTGGATATTATATTGCAATTAATCTACTATCAGATGTG
GCAACCTTGGATTTGCTGAAAACGGAAAATCTGCATTGGGTTGGATTTCTTAAAAGTAAT
          2                                         1
GTATCTAAAAAAATATAGTCATGTTTAACGGTGCTGAATTTTGCCAACTGGACAAGAATG
                                    III         II          I
CAAATGTTACACATTGTCATCCACCAATTAGGAAATAGATAGTGATATTCAAGGATAAGG
ACTTAGGGTCTTTCGAGTCATTTAAATAAACTTGTTGGAAGATCCATGAAACTCATCAAC
TCTTCTTTCTGTGTAATAGCTGCATTCAAGAGTTTTTCAGTTACTTGTACAATG
```

Fig. 1. The sequence of the *Cab22R* and *Cab22L* divergent promoters. The bidirectional promoter region is represented in two pieces which when fused together form the entire promoter (no overlap in sequence is given). Promoter elements including the transcription start sites, the TATA boxes, the three GATA box repeats, the CAAT boxes and the four upstream sequence homology regions are indicated

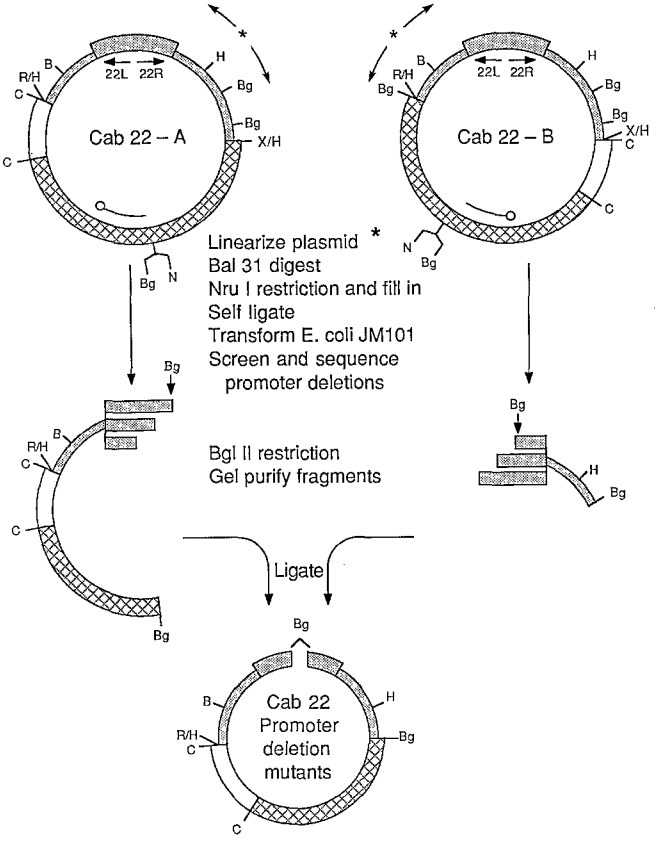

Fig. 3. Construction of internal deletions in the *Cab22* divergent promoter region. *Cab22*-A and *Cab22*-B indicate the insertions, (opposite orientations) of the *Cab22R* and *Cab22L* divergent promoters and genes (*narrow regions*), into a modified pBR322 plasmid DNA (*cross hatched*). The directions of transcription in the *Cab22* promoter region are indicated. The M13 origin of replication is represented by an *open box*, and the Bal31 deletion start points are indicated by an *asterisk*. The multiple steps to construct the internal deletions and sequence the *Cab22* promoter region by a pBR-specific primer (o-) are shown. H-*Hin*dIII, B-*Bam*H1, Bg-*Bgl*II, X-*Xho*1, C-*Cla*1, N-*Nru*1

a series of deletions throughout the *Cab22* promoter region. The deletions, located with respect to the *Cab22R* initiation site, are shown in Fig. 4B. We chose to create systematic internal deletions throughout the *Cab22* promoter as a first step towards identifying sequences which affect the activity of either or both promoters. We regenerated populations of 8 to 14 plants for each construction.

RNA from transgenic plants which carry the *Cab22* wild-type or the mutant promoter constructions were analyzed by the primer extension assay. In this assay a 5′ end-labeled gene-specific oligonucleotide is hybridized to total RNA in solution, and then extended to the 5′ end of the mRNA by reverse transcriptase. The resultant primer-extended products are analyzed on denaturing sequencing gels and in this way precise definition of the transcription start site is possible together with direct quantitation of specific transcript levels. Primer extensions were designed such that transcription from both promoters (*Cab22L* and *Cab22R*), could be assayed simultaneously using distinct primers complementary to the unique 5′ untranslated leader sequence in each RNA species.

Primer extensions were performed on RNA extracted from independent transformants. There was considerable (five- to tenfold) variability in expression between different

			III	II	I	
Petunia	22L:	CCAATGAAATTGTA	GATAGA	GATATCATAA	GATAAGA...19..."TATA"
	22R:	CCAATTAGGAAATA	GATAGT	GATATTCAAG	GATAAGG...23..."TATA"
	91R:	CCAATGAGAAAACAA	GATAAT	GATATTCAAG	GATAAGG...21..."TATA"
	25:	CCAATGAAGAAACC	GATAGT	GATATTCTAG	GATAAGG...23..."TATA"
Arabidopsis	165:	CCAATGAATGAACA	GATAAA	GATTACTTCA	GATATAA...19..."TATA"
	180:	CCAATGAGTAGAGA	GATATA	GATTACTTCATA	GATAACA...21..."TATA"
	140:	CCAATAGCAACCTC	AGAGATT	GATATTTCAA	GATAAGA...21..."TATA"
Pea	AB80:	CCAACTAGCCATAG	CTTTAT	GATACACAC	GATAAGA...10..."TATA"
Lemna	AB30:	CCAATGGCGTGCGG	CCAGTA	GATATCGGTG	GATAATG...31..."TATA"
Tomato	1B:	CCAATGAGATCATA	CATATA	GATATCACTT	GATAAGA...20..."TATA"
	3C:	CCAATGAAAAAGCA	GATAAT	GATATTCTAA	GATAAGG...22..."TATA"

Fig. 2. Sequence comparison of the GATA box repeat elements in *Cab* gene promoters from several plant species, including Petunia (Dunsmuir 1985), *Arabidopsis thaliana* (Leutwiler et al. 1986), pea (Simpson et al. 1985), *Lemna gibba* (Kohorn et al. 1986), and tomato (Picherski et al. 1985)

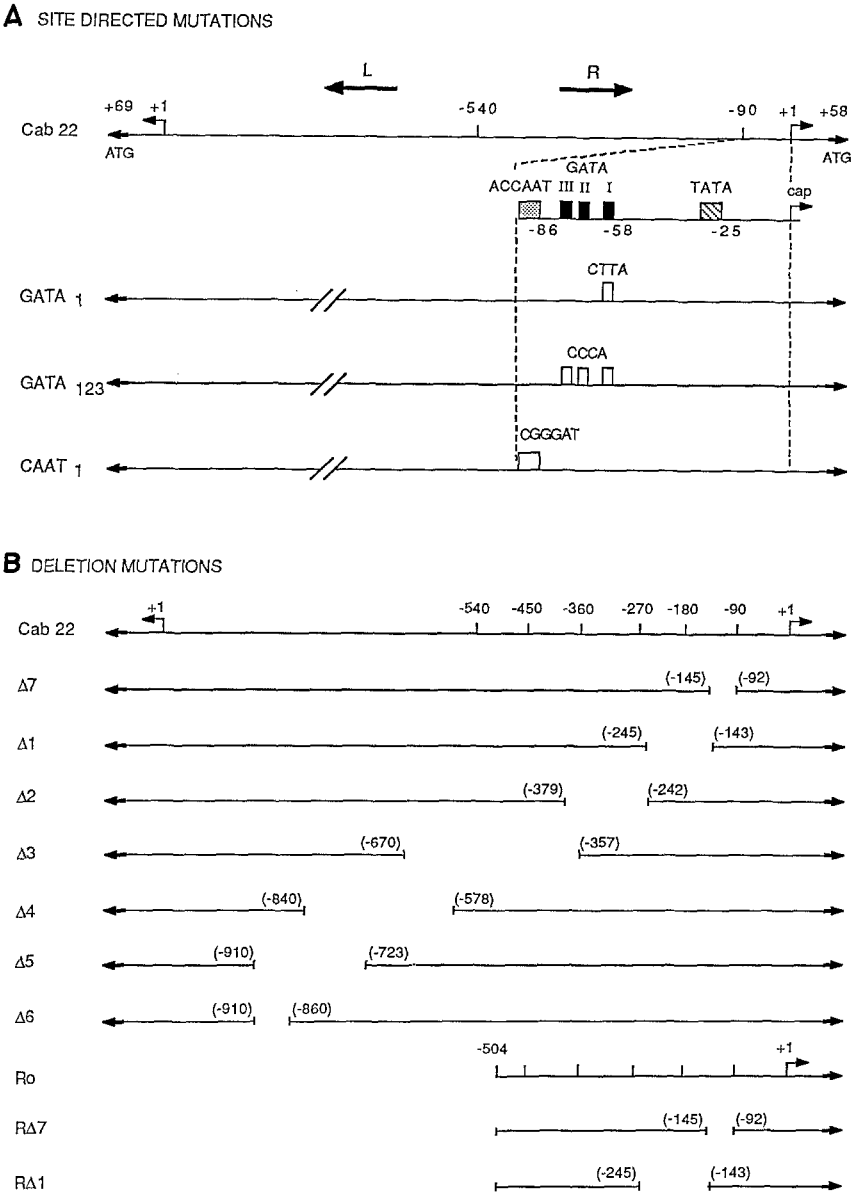

Fig. 4A and B. Schematic diagrams of *Cab22* promoter mutants.
A Site-specific mutations. The *Cab22R* and *Cab22L* intergenic region containing the *Cab22R* and *Cab22L* divergent promoters is shown. The TATA, GATA and CAAT boxes are indicated and the numbers designate their distances from the transcription start site of the *Cab22R* mRNA. The GATA1, GATA123 and the CAAT box mutations are indicated.
B Internal deletion mutations. The *Δ1–Δ7* promoter deletions, with their numbered endpoints relative to the *Cab22R* transcription start site, are illustrated. The truncated *Cab22* promoter (Ro) and its corresponding deletions *RΔ7* and *RΔ1* are shown

plants carrying the same construction. This effect may be due in part to differences in copy number for the introduced gene but also to the "position effect" resulting from the genomic integration site of the new DNA (Jones et al. 1985). Hence RNA samples for the 8–14 plants of each experimental group were pooled and the primer extension assay repeated on each pooled RNA sample. This averages the variability in expression between transformants and allows for somewhat quantitative comparisons of steady-state transcript levels in the different constructions (Fig. 5).

The site-directed mutations in the CAAT (Fig. 5, lane d) and the GATA boxes (Fig. 5, lanes b, c) of the *Cab22R* promoter caused eight- and fivefold reductions, respectively when compared with wild-type *Cab22R* transcript levels (Fig. 5, lane a). As expected, the *Cab22L* RNA levels did not vary between the transformants carrying the wild-type or mutant *Cab22R* promoter constructions. Alteration in the GATA-1 box, (most proximal to the transcription start site), had the same effect on *Cab22R* expression as when all three GATA boxes were mutated, suggesting that the GATA-1 box is necessary for ensuring high expression levels.

In the promoter deletion series mutation *Δ7* (Fig. 5, lane e) had the greatest effect on the level of *Cab22R* expression. This mutant, in which 50 bp were deleted between nucleotides −92 and −145, upstream from the *Cab22R* transcription initiation site results in a fivefold reduction in the *Cab22R* RNA level. This deletion is upstream of the CAAT box of *Cab22R* and encompasses a region of conserved sequence homology among Petunia *Cab* genes (Dunsmuir 1985). Of the other deletions in the series (*Δ1–Δ6*) (Fig. 5, lanes f–k) which are located further upstream of the deletion *Δ7*, only deletion *Δ1* has an effect (twofold reduction) on *Cab22R* expression. We believe that none of the promoter mutations had a specific effect upon *Cab22L* expression, although the *Δ5* and *Δ6* delections appeared to have lowered *Cab22L* expression; this was coupled with lowered *Cab22R* expression (so that the relative expression levels were similar to that of wild-type plants). It may be that these populations (*Δ5* and *Δ6*) have generally

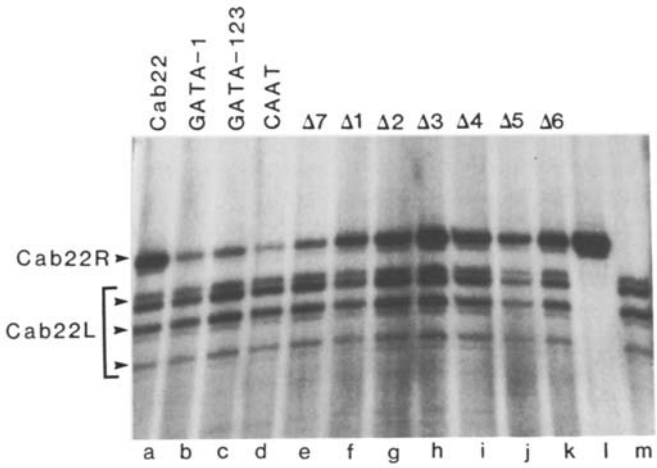

Fig. 5. Mutational analysis of the *Cab22* divergent promoter. Primer extension of pooled RNA from populations of independently transformed plants. Lanes l and m, analysis with *Cab22R* primer and *Cab22L* primer, respectively. Lane a, analysis of RNA from *Cab22*-transformed plants with *Cab22R* and *22L* primers together. Lanes b–k, analysis of RNAs from plants transformed with site-directed mutants (lanes b–d) and internal deletion mutants (lanes l–k; see Fig. 4). The primer extension products indicating accurate transcripts of the *Cab22R* and *22L* genes are designated by *arrowheads*

lower numbers of loci for the transferred genes which would result in overall lowered expression levels.

Light induction and tissue specificity

Previous gene transfer experiments have suggested that transcriptional regulation of *Cab* genes is important in determining light-inducible and tissue-specific expression patterns (Simpson et al. 1985, 1986; Nagy et al. 1986; Gidoni et al. 1988). To determine whether these properties are affected in the *Cab22* promoter, mutant RNA samples were extracted from five plants of each experimental group at four different stages: (1) leaves of greenhouse-grown plants (L_1); (2) leaves of the same plants after 5 days in darkness (D); (3) leaves (L_2) and roots (R) from these plants grown in the greenhouse for 7 days after the dark period.

Primer extension analysis of pooled RNA samples from each experimental group is shown in Fig. 6. The wild-type *Cab22* promoter directs expression in light-grown leaves but not in dark-grown leaves or roots (Fig. 6, lanes 1–4). Similarly, in every mutant promoter which we analyzed (site-specific mutations and deletions) the mutant *Cab22* promoter DNA directs the expression of RNA in a light-regulated and tissue-specific manner (Fig. 6, lanes 5–44). As a control in these experiments we used transformants which carry the cauliflower mosaic virus 35S constitutive promoter fused to the bacterial chitinase (*chiA*) gene. In these

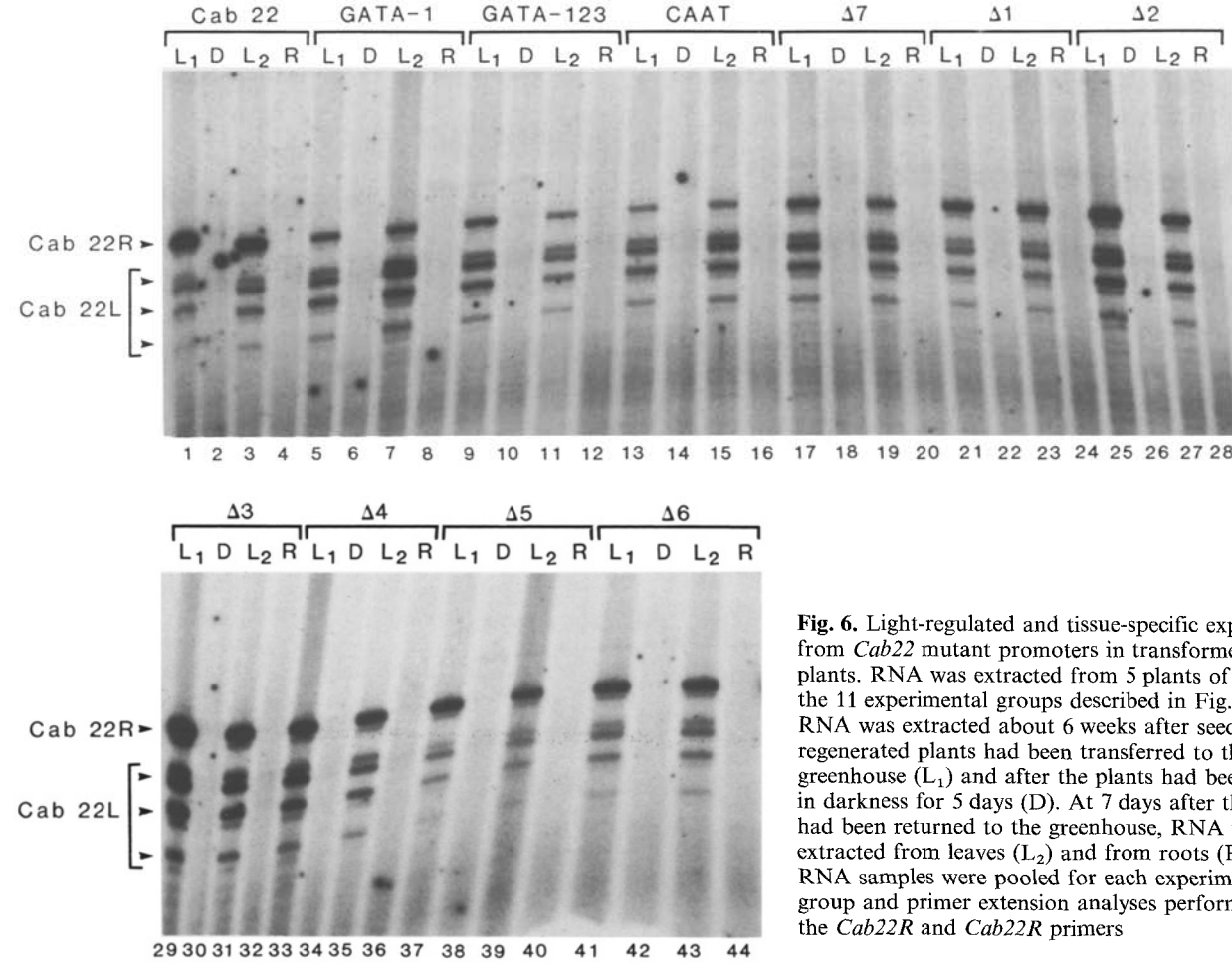

Fig. 6. Light-regulated and tissue-specific expression from *Cab22* mutant promoters in transformed tobacco plants. RNA was extracted from 5 plants of each of the 11 experimental groups described in Fig. 4. Leaf RNA was extracted about 6 weeks after seedlings or regenerated plants had been transferred to the greenhouse (L_1) and after the plants had been placed in darkness for 5 days (D). At 7 days after the plants had been returned to the greenhouse, RNA was again extracted from leaves (L_2) and from roots (R). The RNA samples were pooled for each experimental group and primer extension analyses performed with the *Cab22R* and *Cab22R* primers

plants chitinase mRNA was expressed in both dark-grown leaves and roots (data not shown).

Analysis of a truncated Cab22 *promoter*

The analysis of deletions throughout the *Cab22* promoter region did not reveal regions with dramatic effects on the expression of the adjacent *Cab22L* and *Cab22R* genes. The two most likely explanations are (a) there are multiple interchangeable *cis*-acting elements in each promoter (Maniatis et al. 1987; Kuhlemeier et al. 1987), or (b) there are promoter-enhancer elements in the *Cab22L* promoter region which have compensatory effects on internal deletions in the *Cab22R* promoter region and vice versa. The blocks of sequence homology conserved between the two promoters could correspond to such elements. To address this latter possibility, we dissected the *Cab22* promoter and evaluated the effect of deletions *Δ7* and *Δ1* in the *Cab22R* promoter only, thus comparing the levels of *Cab22R* gene expression in the Ro, R*Δ7* and R*Δ1* transformed plants with those of *Cab22*, *Δ7* and *Δ1* transformed plants, (Fig. 4B). In order to construct the Ro, R*Δ7* and R*Δ1* DNAs, the appropriate plasmids, *Cab22* -*Δ7* and -*Δ1* were digested with *Sna*bI which cleaves the *Cab22* DNA at position −504 from the *Cab22R* transcription start site. Subsequently, *Cla*I linkers were added and the deletions were cloned into pAGS140 then introduced into plants by the process described in Materials and methods. For each deletion 15 plants were analyzed.

The autoradiographs showing primer extension reactions for pooled RNA samples are shown in Fig. 7A. THe comparison of lanes a and b with d and e indicates clearly that in the R*Δ7* promoter deletion in elimination of the *Cab22L* promoter region reduces the steady-state mRNA level of the *Cab22R* 20-fold. However a comparison of lanes c and f in Fig. 7 A shows little difference between *Δ1* and R*Δ1*. These data suggest that there are elements in the *Cab22L* promoter (−504 to −1124 upstream from the *Cab22R* transcription start site) which are able to complement sequences of the *Δ7* deletion (−92 to −145) and partially restore activity to the deleted *Cab22R* promoter.

Light induction and tissue specificity of the truncated Cab22 *promoter*

To test whether elements in the *Cab22L* promoter compensate for those deleted in the *Cab22R* promoter mutations with respect to qualitative aspects of expression, selected transformants (five for each construction) carrying the truncated *Cab22* promoters, Ro, R*Δ7* or R*Δ1*, were incubated under the light/dark regime described previously. RNA was extracted and samples were pooled for primer extension analyses. Figure 7B shows the autoradiograms from these experiments. The qualitative aspects of the *Cab22R* expression are identical for the truncated *Cab22R* promoter and the R*Δ7* and R*Δ1* deletions; in all plants the introduced *Cab22R* gene is expressed only in light-grown leaves. Furthermore this pattern of light-dependent expression in leaves parallels that measured for the intact *Cab22* promoter. Thus, while we have been able to identify a region which functions in the high level expression directed by the *Cab22R* promoter, we have not identified an element which functions in the regulation of the tissue-specific and light-dependent properties of this promoter.

Discussion

The *Cab22* promoter region directs the divergent transcription of two distinct *Cab* genes, *Cab22L* and *Cab22R* in a light-inducible and tissue-specific fashion. The transcription of these two genes is tightly coordinated in Petunia. After transfer of this genomic region to tobacco the two genes are still coordinately expressed, and furthermore this coordination in expression persists when foreign coding regions are fused to the divergent promoter region (Gidoni et al. 1988).

There are several possible models for this type of coordinated expression. In one, the two promoters are distinct in sequence and each responds (separately) to different *trans*-acting factors which are in turn coordinately synthesized; as a consequence the genes respond together. An alternative possibility is that there are similar *cis*-acting ele-

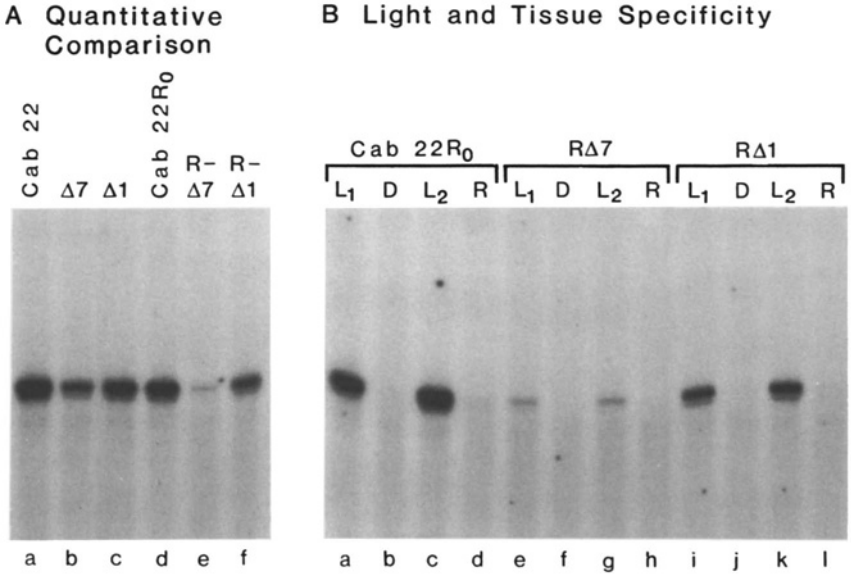

Fig. 7A and B. Mutational analysis of a truncated *Cab22R* promoter. Primer extension of pooled RNA from populations of independently transformed plants.
A Comparison of steady-state mRNA levels in leaf tissue of light-grown plants carrying divergent (lanes a–c) or truncated (lanes d–f) *Cab22* promoter mutants.
B Light-regulated and tissue-specific expression from *Cab22R* truncated promoter mutants. The experiment was conducted as described for Fig. 6

ments in both of the promoter regions which simultaneously respond to a single *trans*-acting factor. Potentially such elements could also function to regulate transcription from the adjacent promoter. A combination of these two mechanisms may also occur.

Our results indicate that the quantitative expression of the *Cab22R* and *Cab22L* promoters is regulated by multiple *cis*-acting control elements. The importance of the CAAT box sequence (at −86 from the *Cab22R* transcription start site +1) was tested by introducing the sequence CGGGAT to replace the ACCAAT sequence. The eightfold reduction in *Cab22R* transcript levels in this mutant is consistent with the importance of the CAAT box element in mammalian promoters (Dynan and Tjian 1985; Maniatis et al. 1987).

Between the CAAT and the TATA box in each gene we identified three 5′-GATA-3′ sequence (GATA box) repeats. Furthermore, we found that these GATA box elements are also repeated two or three times, at the same location, in *Cab* gene promoters of several other plant species. The conserved spacing between these GATA elements and the CAAT box suggests that they may play a role as recognition site(s) for *Cab trans*-activating transcription factor(s), and there might be linkage in activity between the CAAT and the GATA elements.

The fivefold reduction in *Cab22R* transcript level that resulted from a 2 bp substitution in GATA box 1 (mutant GATA-1), or a 3 bp substitution in all three of the GATA boxes (mutant GATA-123) suggests that the GATA box elements modulate transcription in a positive rather than negative fashion. Furthermore, GATA-1, the most proximal to the transcription start site, appears to be necessary for maximum promoter function. This result is consistent with previous observations with repeated promoter elements in mammalian viruses (Gidoni et al. 1985; Jones et al. 1986).

Further upstream from the CAAT box, included within the *Δ*7 deletion mutant, we detected an additional region which has an important role in *Cab* gene expression. The *Δ*7 deletion (ranging from −92 to −145) and the *Δ*1 deletion (ranging from −143 to −245) were tested using two approaches. In the initial experiment the *Cab22L* promoter region was linked to the *Cab22R* promoter; the internal *Δ*7 and *Δ*1 deletions in the *Cab22R* promoter region showed approximately 5-fold and 2-fold reductions, respectively, in RNA levels measured for the *Cab22R* gene. In the second experiment, the *Cab22L* promoter region was deleted hence the *Cab22R* promoter alone was tested. In this case the R*Δ*1 deletion reduced the *Cab22R* mRNA levels 2-fold compared with Ro, while the RNA levels from the R*Δ*7 deletion were reduced approximately 20-fold.

These results suggest that regions in the *Cab22L* promoter function, from a distance, to compensate partially for the loss of promoter activity caused by the *Δ*7 deletion in *Cab22R*. This hypothesis is supported by the occurrence of an extensive region of sequence homology (13 nucleotides allowing 2 bp mismatch) shared between the *Cab22R* and *Cab22L* promoters at positions −115 and −110 respectively (see homology box 1 in Fig. 1). This sequence is also found in other Petunia LHCII Type I *Cab* genes (Dunsmuir 1985). An enhancer-like element which was previously characterized in a pea *Cab* gene promoter also occurs in this same region relative to the CAAT box (Simpson et al. 1986). A direct test of the importance of this region will be possible by preparing a mutant *Cab22* promoter which combines the *Δ*7 deletion with a corresponding deletion in the *Cab22L* promoter.

The mutational analysis conducted here indicates that the GATA repeat sequences, the CAAT sequence, and the upstream boxes of sequence homology (1 and possibly 2, see Fig. 1) are important for the quantitative expression of the *Cab22R* gene. Our data for the deletions further upstream in the *Cab22R* promoter (that include homology boxes 3 and 4) are less conclusive. The sequences beyond the *Δ*1 deletion region (at −245) may have no effect on expression level, or it may be that the effects of these upstream deletions are fully compensated by elements remaining in the adjacent *Cab22L* promoter region. We have not examined the effects of these deletions in a truncated *Cab22R* promoter construction.

The qualitative aspects of expression were not perturbed in any of the mutant promoters which we tested. The most likely explanation for this result is that the qualitative expression properties of these divergent promoters are controlled by a complex of multiple *cis*-acting elements, and the effect of deletion of any one of these elements can be fully compensated by the remaining elements. Since the qualitative expression from the deletions in the truncated promoter is also identical to that from the wild-type promoter, there must be multiple interchangeable elements within the *Cab22R* promoter which regulate qualitative expression.

We are now pursuing an alternative and complementary approach to evaluate further *Cab* promoter elements. Promoter fusions are being used to test how each element, or combinations of these different elements, affect quantitative as well as qualitative expression of heterologous promoters. Such studies will identify elements that confer tissue- and light-induced expression in the *Cab22* promoters, and provide more understanding of the nature of their action.

Acknowledgements. We thank C. Dean, P. Lund and E. Crump for critical reading of this manuscript and C. Stephens, R. Narberes and L. Kruger for assistance in manuscript preparation. We also thank E. Penzes for tissue culture assistance, and T. Lemieux and C. Robinson for plant propagation and maintenance.

References

Baumann G, Raschke E, Beven M, Schoffl F (1987) Functional analysis of sequences required for transcriptional activation of a soybean heat shock gene in transgenic tobacco plants. EMBO J 6:1161–1166

Baumlein H, Wobus U, Pustell J, Kafatos FC (1986) The legumin gene family: structure of a B-type gene of *Vicia faba* and a possible legumin gene specific regulatory element. Nucleic Acids Res 14:2707–2720

Chen ZL, Pan NZ, Beachy RN (1988) A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. EMBO J 7:297–302

Dean C, van den Elzen P, Tamaki S, Dunsmuir P, Bedbrook J (1985) Differential expression of the eight genes of the Petunia *rbcS* gene family. EMBO J 4:3055–3061

DeGreve H, Dhaese P, Seurinck J, Lemmers M, van Montagu M, Schell J (1983) Nucleotide sequence and transcript map of the Agrobacterium tumefaciens Ti plasmid-encoded octopine synthase gene. J Mol Appl Gen 1:499–511

Dunsmuir P (1985) The petunia chlorophyll *a*/*b* binding protein genes: a comparison of *Cab* genes from different gene families. Nucleic Acids Res 13:2503–2518

Dunsmuir P, Bond D, Lee K, Gidoni D, Townsend J (1988a) Fate of introduced genes in transformed plants. In: Gelvin S, Schilperoot R (eds) Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht C1:1–17

Dunsmuir P, Dean C, Gidoni D, Stayton M (1988b) Elements which regulate the expression of photosynthetic genes. In: Goldberg R (ed) The molecular basis of plant development. UCLA symposia on molecular and cellular biology, new series 92. Alan R. Liss, New York

Dynan WS, Tjian R (1985) Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature 316:774–778

Ellis JG, Lewellyn DJ, Dennis ES, Peacock WJ (1987) Maize Adh-1 promoter sequences control anaerobiotic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. EMBO J 6:11–16

Figurski DH, Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc Natl Acad Sci USA 76:1648–1652

Fluhr R, Kuhlemeier C, Nagy F, Chua H (1986) Organ-specific and light-induced expression of plant genes. Science 232:1106–1112

Forde BG, Heyworth A, Pywell J, Kreis M (1985) Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize. Nucleic Acids Res 13:7327–7339

Fromm M, Taylor LP, Walbot V (1985) Expression of genes transferred into monocot and dicot plant cells by electroporation. Proc Natl Acad Sci USA 82:5824–5328

Gidoni D, Kadonaga JT, Barrera-Saldana H, Takahashi K, Chambon P, Tjian R (1985) Bidirectional SV40 transcription mediated by tandem Spl binding interactions. Science 230:511–517

Gidoni D, Bond-Nutter D, Brosio P, Jones JDG, Bedbrook J, Dunsmuir P (1988) Coordinated expression between two photosynthetic petunia genes in transgenic plants. Mol Gen Genet 211:507–514

Haltiner M, Kempe T, Tjian R (1985) A novel strategy for constructing clustered point mutations. Nucleic Acids Res 13:1015–1025

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoot RA (1983) A binary plant vector strategy based on separation of vir and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303:179–180

Horsch RB, Fry JE, Hoffmann NL, Eichholz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Science 227:1229–1231

Jones JDG, Dunsmuir P, Bedbrook J (1985) High level expression of introduced chimaeric genes in regenerated transformed plants. EMBO J 4:2411–2418

Jones KA, Kadonaga JT, Luciw PA, Tjian R (1986) Activation of the AIDS retrovirus promoter by the cellular transcription factor Spl. Science 232:755–759

Kaulen H, Schell J, Kreuzaler F (1986) Light-induced expression of the chimeric chalcone synthase-NPT-II gene in tobacco cells. EMBO J 5:1–8

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236:1299–1302

Kloppstech K (1985) Diurnal and circadian rhythmicity in the expression of light-induced plant nuclear messenger RNAs. Plant 165:502–506

Kohorn BD, Harel E, Chitnis PR, Thornber JP, Tobin EM (1986) Functional and mutational analysis of the light-harvesting chlorophyll a/b – protein of thylakoid membranes. J Cell Biol 102:972–981

Kuhlemeier C, Fluhr R, Green PJ, Chua N-H (1987) Sequences in the pea *rbcS-3A* gene have homology to constitutive mammalian enhancers but function as negative regulatory elements. Genes Devel 1:247–255

Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc Natl Acad Sci USA 82:488–492

Leutwiler LS, Meyerowitz EM, Tobin EM (1986) Structure and expression of three light-harvesting chlorophyll a/b-binding protein genes in *Arabidopsis thaliana*. Nucleic Acids Res 14:4051–4064

Maier U-G, Brown JWS, Toloczyki C, Feix G (1987) Binding of a nuclear factor to a consensus sequence in the 5′ flanking region of zein genes from maize. EMBO J 6:17–22

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Maniatis T, Goodbourn S, Fischer JA (1987) Regulation of inducible and tissue-specific gene expression. Science 236:1237–1245

McKnight SL, Tjian R (1986) Transcription selectivity of viral genes in mammalian cells. Cell 46:795–805

McKnight SL, Gavis ER, Kingsbury R (1981) Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell 25:385–398

Messing J (1983) New m13 vectors for cloning. Methods Enzymol 101:20–78

Messing J, Geraghty D, Heidecker G, Hu N-T, Kridl J, Rubenstein I (1983) In: Kosuge T, Meredith CP, Hollander A (eds) Genetic engineering of plants. Plenum Press, New York, pp 211–227

Nagy F, Kay SA, Boutry M, Hsu MY, Chua N-H (1986) Phytochrome-controlled expression of a wheat Cab gene in transgenic tobacco seedlings. EMBO J 5:1119–1124

Pichersky E, Bernatzky R, Tanksley SD, Breidenbach R, Kansch AP, Cashmore AR (1985) Molecular characterization and genetic mapping of two clusters of genes encoding chlorophyll *a/b*-binding proteins in *Lycopersicon esculentum* (tomato). Gene 40:247–258

Pichersky E, Hoffman NE, Bernatzky R, Piechulla B, Tanksley SD, Cashmore AR (1987) Molecular characterization and genetic mapping of DNA sequences encoding the Type I chlorophyll a/b-binding polypeptide of photosystem I in *Lycopersicon esculentum* (tomato). Plant Mol Biol 9:205–216

Pichersky E, Tanksley SD, Piechulla B, Stayton MM, Dunsmuir P (1988) Nucleotide sequence and chromosomal location of *Cab-7*, the tomato gene which encodes the type II chlorophyll a/b-binding polypeptide of photosystem I. Plant Mol Biol

Potrykus I, Shillito R, Saul M (1985) Direct gene transfer to cells of a graminaceous monocot. Mol Gen Genet 199:183–188

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Simpson J, Timko MP, Cashmore AR, Schell J, Van Montagu M, Herrera-Estrella L (1985) Light-inducible and tissue-specific expression of a chimaeric gene under control of the 5′-flanking sequence of a pea chlorophyll *a/b*-binding protein gene. EMBO J 4:2723–2729

Simpson J, Schell J, Van Montagu M, Herrera-Estrella L (1986) Light-inducible and tissue-specific pea *lhcp* gene expression involves an upstream element combining enhancer- and silencer-like properties. Nature 323:551–554

Stayton M, Black M, Bedbrook J, Dunsmuir P (1986) A novel chlorophyll a/b-binding (*Cab*) protein gene from petunia which encodes the lower molecular weight *Cab* precursor protein. Nucleic Acids Res 14:9781–9796

Timko MP, Kausch AP, Castresana C, Fassler J, Herrera-Estrella L, Van den Broeck G, Van Montague M, Schell J, Cashmore AR (1985) Light regulation of plant gene expression by an upstream enhancer-like element. Nature 318:579–582

Communicated by J. Schell

Received June 16, 1988