Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 5083–5087, August 1985
Genetics

Inheritance, organization, and mapping of *rbcS* and *cab* multigene families in pea

(restriction fragment polymorphisms/linkage maps/nuclear gene clusters)

NEIL O. POLANS*, NORMAN F. WEEDEN†, AND WILLIAM F. THOMPSON*

*Department of Plant Biology, Carnegie Institution of Washington, 290 Panama Street, Stanford, CA 94305; and †Department of Horticultural Sciences, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456

Communicated by R. W. Allard, April 9, 1985

ABSTRACT **DNA restriction endonuclease fragment patterns corresponding to both the *rbcS* and *cab* multigene families of pea are each shown to segregate as single Mendelian units in the F_2 progeny of two separate crosses. All of the observed variation in each of the multigene families is thus organized on the chromosome in a tightly linked complex. Linkage relationships between both multigene families and an array of morphological and isozyme markers establish the location of the *rbcS* and *cab* gene clusters on pea chromosomes 5 and 2, respectively. Our results, which indicate a high level of DNA restriction fragment length polymorphism in pea, suggest sufficient variation to permit the construction of a highly detailed linkage map.**

Two of the most prominent and abundant chloroplast proteins, the small subunit of ribulose-1,5-bisphosphate carboxylase and the light-harvesting chlorophyll a/b binding protein, are specified by nuclear genes (*rbcS* and *cab*, respectively) (1, 2) whose transcript levels are light-regulated (3–6). The isolation of cDNA and genomic clones corresponding to *rbcS* and *cab* has facilitated investigations into their structure and expression in pea (5, 7–12), soybean (13), petunia (14, 15), and wheat (16). In particular, evidence has been accumulated that both proteins are encoded by small multigene families (7, 9–11, 13–17). In this study the inheritance of restriction endonuclease fragment length variants containing genes belonging to either the *rbcS* or *cab* multigene family of pea is described, and genes encoding both families are mapped onto the pea genome. To the best of our knowledge, the use of DNA restriction fragment length polymorphism to link plant structural genes to markers on a genetic map has not been reported previously.

MATERIALS AND METHODS

Plant Material. Seed for *Pisum sativum* L. inbred lines was obtained through the courtesy of G. A. Marx (New York State Agricultural Experiment Station). These accessions are labeled 1–11 as follows: (1) A778-26-6, (2) A73-91, (3) A1179-395, (4) B980-686, (5) C879-344, (6) A1078-236, (7) A78-237, (8) C482-236, (9) A578-238, (10) A780-388, (11) A1078-234. F_2 progeny were obtained by crossing C879-344 with A73-91 and A73-91 with A78-237. The 43 F_2 individuals resulting from the first cross and the 31 F_2 individuals resulting from the second cross were each surveyed for DNA restriction fragment length variants, isozymes, and morphological characters.

DNA Methodology. DNAs were extracted from the leaves of individual pea plants in a rapid, small-scale procedure (18) modified by R. A. Jorgensen from the procedure of Murray and Thompson (19). Approximately 5-μg DNA samples, digested with an appropriate restriction endonuclease (usually *Eco*RI, *Bam*HI, *Hin*dIII, *Bcl* I, or *Bgl* II; New England Biolabs and Bethesda Research Laboratories) according to manufacturers' specifications, were fractionated on 0.9% agarose gels with 100 mM Tris acetate/12.5 mM sodium acetate/1 mM Na_2EDTA, pH 8.1, as both the gel and tray electrophoresis buffer. After the transfer of the DNA fragments to GeneScreen membranes (New England Nuclear) (20) following procedures outlined in New England Nuclear catalog NEF-972, specific *rbcS* and *cab* sequences were localized by hybridization with nick-translated pea cDNA clones pSS15 and pAB96 (kindly provided by N.-H. Chua), which correspond to the small subunit of ribulose-1,5-bisphosphate carboxylase and polypeptide 15 of the chlorophyll a/b light-harvesting complex, respectively (12). Nick-translation and hybridization generally followed the protocols of Maniatis *et al.* (21). Hybridization buffer, added to a volume of 0.1 ml/cm^2 of GeneScreen, consisted of 5 × concentrated SSPE buffer at pH 7.4 (0.9 M NaCl/50 mM $NaH_2PO_4 \cdot H_2O$/5 mM Na_2EDTA), 10 × concentrated Denhardt's reagent [0.2% polyvinylpyrrolidone (M_r 40,000)/0.2% bovine serum albumin/0.2% Ficoll (M_r 400,000)], 100 μg of denatured calf thymus DNA per ml, 0.1% $NaDodSO_4$, 10% dextran sulfate, and 30% deionized formamide. Prehybridization in this same buffer was allowed to proceed for 6 hr at 42°C prior to adding probe at 10^6 dpm/ml of buffer. After 15–20 hr of hybridization at 42°C in the presence of the probe, blots were washed numerous times in 0.3 M sodium chloride/0.03 M sodium citrate, pH 7/0.1% $NaDodSO_4$ both at room temperature and at 65°C.

Isozyme Electrophoresis. Isozyme phenotypes were determined by horizontal starch gel electrophoresis as described in Weeden and Marx (22).

RESULTS

Eleven pea experimental lines known to vary for a large number of morphological and isozyme characters were selected for a survey of genetic variation in *rbcS* and *cab* because of both the generally variable nature of these lines relative to one another and their background of extant genetic markers amenable to linkage analysis. Identification of the degree and pattern of *rbcS* and *cab* DNA restriction fragment length variation across the lines led to the construction of several crosses that most efficiently encompassed the range of available variation. Segregation analysis was then performed on *rbcS* and *cab* DNA restriction fragment variants to characterize the genetics of the band pattern phenotypes involved in each cross. Subsequently, the newly characterized genetic loci were examined for linkage with existing genetic markers and, based upon these linkage relationships, were mapped onto the pea genome.

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

Abbreviation: kb, kilobase(s).

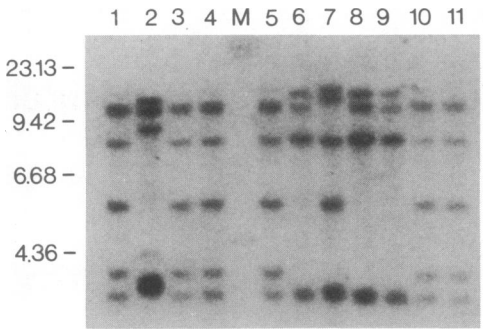

FIG. 1. Variation in pea DNA restriction fragments hybridizing with the *rbcS* probe. DNAs were extracted from the leaves of single pea plants and digested with *Eco*RI. Lanes 1–11 contain DNA from the 11 inbred lines described in the text. Lane M contains *Hin*dIII-cut λ DNA size standards (not visible) whose fragment lengths in kilobases (kb) are indicated on the left. The four distinct band patterns that can be discerned are represented by (*i*) lanes 1, 3, 4, 5, 10, and 11, (*ii*) lane 2, (*iii*) lanes 6, 8, and 9, and (*iv*) lane 7.

Substantial *rbcS* and *cab* DNA restriction fragment length polymorphism was found among the lines examined. Four different *rbcS* band patterns (Fig. 1) and three different *cab* band patterns (data not shown) were displayed among the 11 *Pisum* lines for each of the five restriction endonucleases used. Since the same grouping of lines was obtained with each of the five enzymes, it is unlikely that methylation polymorphisms contributed to the variability shown; however, the complexity of patterns within lanes, which suggests small multigene families, obscures the nature of these polymorphisms (i.e., whether they derive from insertion/deletion events or cleavage-site changes) to the extent that unambiguous interpretation is not possible by inspection alone.

Based upon these data, two crosses were selected that efficiently exploited the available variability, C879-344 × A73-91 (Fig. 1, lanes 5 and 2, respectively) and A73-91 × A78-237 (Fig. 1, lanes 2 and 7, respectively). F_2 progeny from these crosses were used to determine the genetic structure underlying the *rbcS* and *cab* band patterns.

For the F_2 progeny from the cross A73-91 × A78-237, a comparison of observed and expected numbers of progeny in each phenotypic category showed good agreement with the 1:2:1 Mendelian ratio for segregation at a single locus, as indicated by χ^2 goodness-of-fit tests; $\chi^2_{(2)} = 0.66$, $0.50 < P < 0.75$ not significant, and $\chi^2_{(2)} = 0.70$, $0.50 < P < 0.75$ not significant, for *rbcS* and *cab*, respectively (Fig. 2). Similar analyses of the F_2 progeny from the cross C879-344 × A73-91 also demonstrated allele-like segregation at a single locus for both *rbcS* and *cab*, the A73-91 and C879-344 *Eco*RI phenotypes being designated aa and bb, respectively. The numbers of observed (and expected) aa, ab, and bb F_2 progeny from this cross were as follows: 16 (10.75), 19 (21.5), 8 (10.75), $\chi^2_{(2)}$

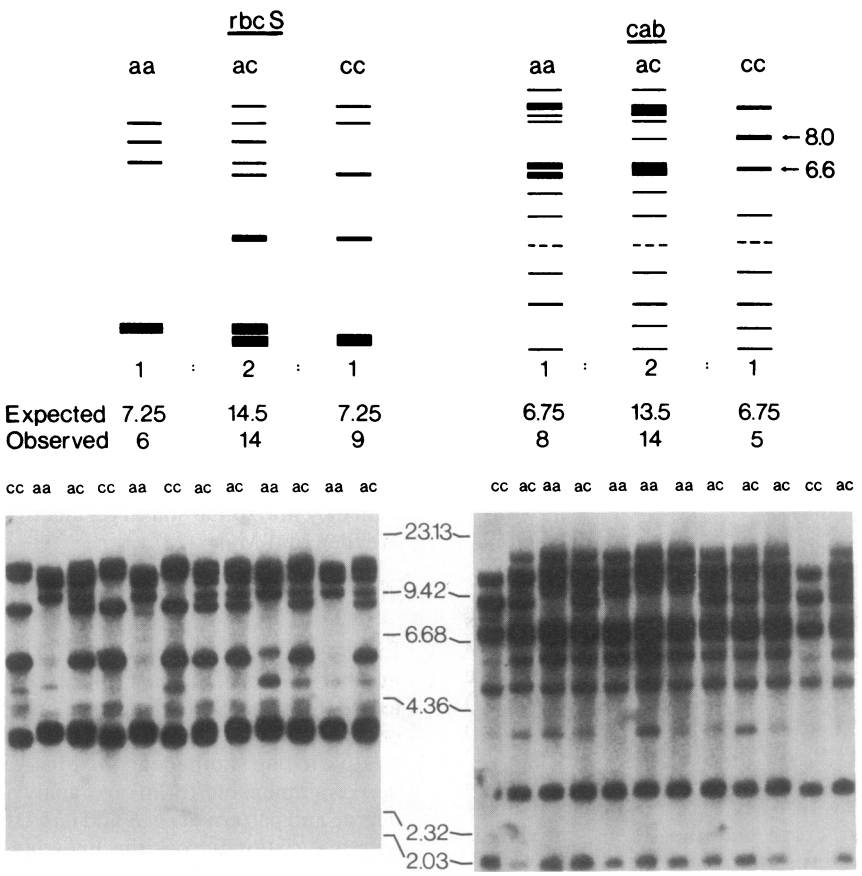

FIG. 2. Segregation in the F_2 of pea DNA restriction fragment length variants hybridizing with the *rbcS* and *cab* probes. DNAs were extracted from the F_2 progeny of a cross between pea lines A73-91 and A78-237 and digested with *Eco*RI. The resulting F_2 band patterns are either identical to the parental types (A73-91:aa, A78-237:cc) or a heterozygote-like combination of these two types (ac). Autoradiograms containing representative samples of the F_2 progeny arrays indicate that the two lowest molecular weight and most intensely hybridized *rbcS* bands are probably doublets and/or correspond to multiple *rbcS* gene sequences; the two lowest molecular weight bands pictured in the *cab* schematic diagram are not represented in the autoradiogram. Dotted lines in the *cab* schematic diagram denote lightly hybridizing, closely migrating bands that could not be consistently scored and that were excluded from all analyses. Similarly anomalous *rbcS* bands, particularly noticeable in lane 9, were also excluded from analysis. On the basis of similar fragment sizes and banding patterns (see text), the two *cab* bands designated by arrows are thought to correspond to *cab* genes contained in the pea genomic clones AB80 and AB66 described by Timko and Cashmore (9). Size standards (in kb) were *Hin*dIII fragments of λ DNA.

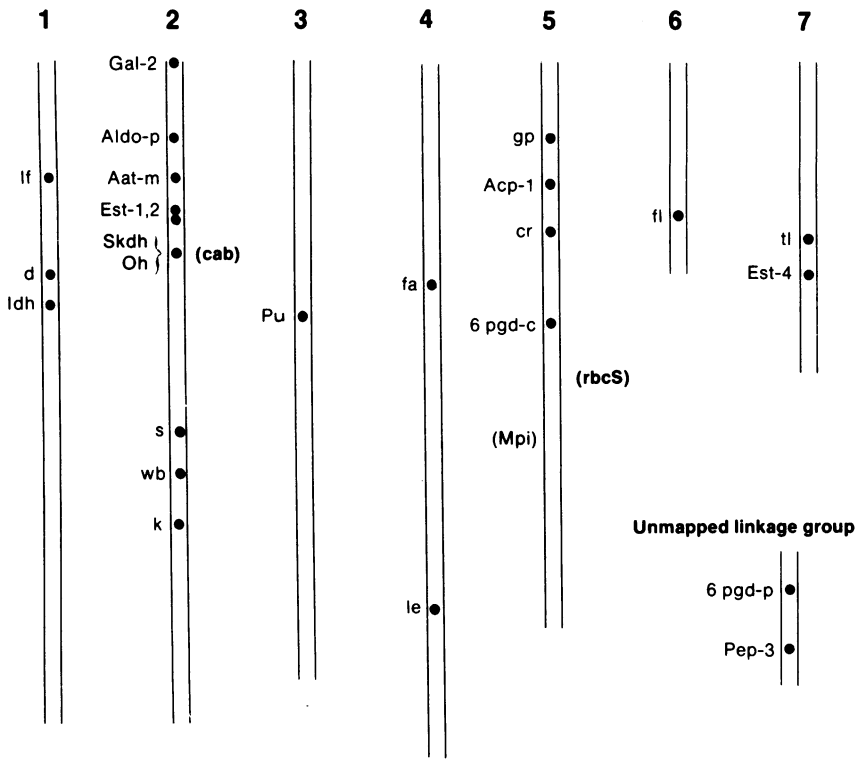

FIG. 3. Schematic diagram of the pea linkage map indicating the locations of selected markers and the *rbcS* and *cab* gene clusters. Linkage relationships were determined with the aid of the LINKAGE-1 computer program (24). The approximate chromosomal locations of morphological and isozyme markers segregating in two F_2 populations are shown to the left of the appropriate chromosome. Several unmapped markers that also segregated in these populations are not pictured.

$= 3.56, 0.10 < P < 0.25$ not significant, for *rbcS* and 5 (10.75), 22 (21.5), 16 (10.75), $\chi^2_{(2)} = 5.65$, $0.05 < P < 0.10$ not significant, for *cab*. *Bam*HI digestions of the DNAs from this second F_2 progeny array gave numerical results identical to those obtained with the *Eco*RI digestions.

In both cases, then, complex band pattern phenotypes were shown to segregate together in an allele-like fashion as would be expected for a tight cluster of genes, similar results being obtained independently from both crosses, although, in the case of the *cab* phenotypes, several of the bands were invariant and, hence, were not amenable to segregation analysis (see *Discussion*). The maximum genetic size of the clusters, measured in terms of maximum recombination frequency (r), was estimated to be 0.04 (or no more than 4 map units in length) where $\alpha = 0.05 = (1-r)^{74}$. α is the probability of obtaining the 74 nonrecombinant events observed in the two sets of F_2 progeny ($n = 43 + 31$) when the probability of each individual event is r (23).

Because they behave as single loci, both of the gene clusters were mapped onto the pea genome as discrete units (Fig. 3). Linkage relationships and recombinational values (i.e., map distances) between *rbcS* and *cab* genes and an array of other genetic markers were determined with the aid of the LINKAGE-1 computer program (24) and were based on the method of maximum likelihood (25). Seventeen markers segregated in the F_2 progeny from the cross A73-91 × A78-237 and 21 markers segregated in the F_2 progeny from the cross C879-344 × A73-91.

In cross A73-91 × A78-237, no recombinants were observed between *cab* and either *Skdh*, the locus encoding shikimate dehydrogenase, or *Oh*, a chromosome 2 marker (26). These data are summarized in Table 1. Tight linkage between *Skdh* and *Oh* has been reported (22). Our results indicate then that the *cab* multigene family is located on the upper portion of chromosome 2 as depicted in Fig. 3. Further evidence of this linkage is the association observed between *cab* and the esterase 1 and esterase 2 loci (*Est-1,2*), also shown to be closely linked to *Skdh* and *Oh* on chromosome 2 (22). *Est-1,2* and *cab* are separated by ≈11 map units (Table 1).

In cross C879-344 × A73-91, segregation at *rbcS* exhibited nonrandom assortment with markers on chromosome 5. The linkage group *Gp–Acp-1–6pgd-c* has been firmly established on chromosome 5 (22, 26), and *rbcS* showed linkage with the

Table 1. Joint segregation analysis involving *cab*, *rbcS*, and marker loci on chromosomes 2 and 5

Loci	F_2 progeny in each phenotypic class,* no.									n	χ^2	P	Recombinant fraction
	11, 11	11, 12	11, 22	12, 11	12, 12	12, 22	22, 11	22, 12	22, 22				
cab, Skdh	8	0	0	0	14	0	0	0	5	27	54.0	<0.01	0.01 ± 0.01
cab, Oh†	8	0	14	0	0	5				27	27.0	<0.01	0.01 ± 0.02
cab, Est-1,2	5	0	0	1	10	3	0	1	4	24	25.5	<0.01	0.11 ± 0.05
6pgd-c, Acp-1	8	5	1	1	16	2	0	2	3	38	21.2	<0.01	0.18 ± 0.05
rbcS, 6pgd-c	11	5	0	3	14	2	2	3	3	43	16.8	<0.01	0.22 ± 0.05
rbcS, Acp-1	6	6	1	2	11	4	1	6	1	38	6.2	0.1–0.25	0.36 ± 0.07

*Phenotypic designations: 1 = A73-91 phenotype, 2 = either A78-237 or C879-344 phenotype.
†11 and 12 cannot be differentiated due to dominance.

latter two loci in this group (Table 1). The recombination frequencies calculated between *rbcS* and each of the other loci indicate that this multigene family is positioned about 22 map units from the gene encoding the cytosolic isozyme of 6-phosphogluconate dehydrogenase (*6pgd-c*) on the side opposite *Acp-1* (Fig. 3). Another locus, *Mpi*, which encodes mannose phosphate isomerase, is also situated on this side of *6pgd-c*, although its precise location has yet to be determined. Our data on segregation at *Mpi* were incomplete due to the difficulty of scoring certain phenotypes; however, linkage was observed between *rbcS* and *Mpi*, again confirming that the former locus is on chromosome 5.

The rest of the segregating markers assorted independently of *rbcS* and *cab* and showed expected linkage relationships among themselves. These results further support the organization of these multigene families as tightly linked clusters and reduce the possibility that such linkages were caused by karyotypic differences between the parents.

DISCUSSION

In the case of the *cab* multigene family, several of the *Eco*RI bands observed are shared by both of the parents of a cross (C879-344 × A73-91 and A73-91 × A78-237 have in common 4 of 17 and 4 of 15 bands, respectively); they are generally of lower molecular weight than their variable counterparts. Since invariant bands are not amenable to segregation analysis, we cannot exclude the possibility that they may arise from genes or pseudogenes located outside the genetic locus described in Fig. 2.

We think it is likely that the variant bands mapped to chromosome 2 correspond to actual *cab* genes rather than pseudogenes. This conclusion is suggested by an indirect comparison between these restriction fragments and similar fragments described in a study by Timko and Cashmore (9) using pea variety Progress No. 9.

In their study, a *cab* probe hybridized to an *Eco*RI digest of nuclear DNA produced a profile very similar to the profile we observe for pea accession C879-344. This profile is quite common among the pea genotypes we have sampled. Given the similarities of the profiles, we feel it is reasonable to assume that corresponding sequences are present in the corresponding bands from different varieties. Thus, the *Eco*RI fragments of approximately 6.6 and 8.0 kb that we see in our profile probably contain the *cab* genes sequenced by Timko and Cashmore (9) from their clones of 6.6- and 8.0-kb fragments of Progress No. 9 DNA. These bands in our profile are distinct from bands found in pea accession A73-91 and, hence, were among those bands determined to segregate as a unit in the F_2 progeny of the cross C879-344 × A73-91. Bands of about 6.6 and 8.0 kb were also found in pea accession A78-237. As such, they were among those bands directly mapped to chromosome 2 on the basis of the cross A73-91 × A78-237 (see Fig. 2). These results suggest that at least some actual *cab* genes, and not just pseudogenes, are part of the gene cluster described in Fig. 2.

Our results indicate that the entire *rbcS* multigene family (for which all of the observed bands segregated) is clustered on pea chromosome 5 within a region of <4 map units. At least most members of the *cab* multigene family (those for which segregation was observed) appear to be in a similar-sized cluster on chromosome 2. Studies on electrophoretic variants of the holoenzyme of ribulose-1,5-bisphosphate carboxylase in *Phaseolus vulgaris* indicate that the *rbcS* gene cluster also segregates as a single Mendelian unit in this species (27). Evidence of tight linkage between two *rbcS* genes in pea (10) as well as between two *rbcS* genes and between two *cab* genes in petunia (17) has also been reported, each pair of genes having been localized to a single genomic clone.

This type of genetic organization (i.e., clustering) has interesting implications regarding the origin, maintenance, and regulation of these gene families. It has been argued that a tightly linked family of genes would most likely derive from a rare gene duplication followed by (relatively rapid and continuous) unequal crossing-over events. One advantage of maintaining such a configuration, once it had evolved, might be the facility with which coordinate expression of the genes could be regulated (28). In pea and other systems, both *rbcS* and *cab* mRNA levels are regulated by light acting through the phytochrome system (3–4, 6). It is possible that the presence of the pea *rbcS* and *cab* gene families in tightly linked assemblages reflects a regulatory strategy common to both systems. Alternatively, the tight linkage may only reflect relatively recent duplication events that have not yet undergone rearrangement.

Studies on the nature and mechanisms of light control on *rbcS*, *cab*, and a number of other genes in pea have been reported recently (29, 30) and complementary studies on the inheritance, localization, and organization of these other light-regulated gene systems are continuing.

It is noteworthy that high levels of restriction fragment polymorphism similar to those observed for the *rbcS* and *cab* loci were also observed with several other cDNA probes (unpublished). The frequency of such variation thus appears to be quite high even for closely related pea lines, indicating that this approach to genetic mapping might quite rapidly be extended by using large numbers of random clones to produce a highly detailed pea linkage map.

The authors gratefully acknowledge the assistance of G. A. Marx for providing the inbred lines used in this study and for helping to score plants for certain morphological characters. We are also grateful for the assistance and cooperation of R. A. Jorgensen, J. D. Palmer, G. A. Ford, L. S. Kaufman, D. B. Stern, and J. C. Watson and for the comments of several anonymous reviewers. This work was supported in part by U.S. Department of Agriculture Grant 78-59-2115-0-1-009-1 and in part by Department of Agriculture Hatch Project 33491 and is CIW-DPB publication 869.

1. Ellis, R. J. (1981) *Annu. Rev. Plant Physiol.* **32,** 111–137.
2. Kirk, J. T. O. & Tilney-Basset, R. A. E. (1978) *The Plastids: Their Chemistry, Structure, Growth and Inheritance* (Elsevier, Amsterdam), 2nd Ed.
3. Thompson, W. F., Everett, M., Polans, N. O., Jorgensen, R. A. & Palmer, J. D. (1983) *Planta* **158,** 487–500.
4. Tobin, E. M. (1981) *Plant Mol. Biol.* **1,** 35–51.
5. Bedbrook, J., Smith, S. & Ellis, R. (1980) *Nature (London)* **287,** 692–697.
6. Apel, K. (1979) *Eur. J. Biochem.* **97,** 183–188.
7. Coruzzi, G., Broglie, R., Edwards, C. & Chua, N.-H. (1984) *EMBO J.* **3,** 1671–1679.
8. Cashmore, A. R. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2960–2964.
9. Timko, M. P. & Cashmore, A. R. (1983) in *Plant Molecular Biology*, ICN–UCLA Symposium on Molecular and Cellular Biology, ed. Goldberg, R. B. (Liss, New York), pp. 403–412.
10. Cashmore, A. R. (1983) in *Genetic Engineering in Plants: An Agricultural Perspective*, eds. Kosuge, T., Meredith, C. P. & Hollaender, A. (Plenum, New York), pp. 29–38.
11. Coruzzi, G., Broglie, R., Cashmore, A. & Chua, N.-H. (1983) *J. Biol. Chem.* **258,** 1399–1402.
12. Broglie, R., Bellemare, G., Bartlett, S. G., Chua, N.-H. & Cashmore, A. R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 7304–7308.
13. Berry-Lowe, S. L., McKnight, T. D., Shah, D. M. & Meagher, R. B. (1982) *J. Mol. Appl. Genet.* **1,** 483–498.
14. Dunsmuir, P., Smith, S. & Bedbrook, J. (1983a) *Nucleic Acids Res.* **1,** 4177–4183.
15. Dunsmuir, P., Smith, S. M. & Bedbrook, J. (1983b) *J. Mol. Appl. Genet.* **2,** 285–300.
16. Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. & Chua, N.-H. (1983) *Biotechnology* **1,** 55–61.

17. Dunsmuir, P. & Bedbrook, J. (1983) in *Structure and Function of Plant Genomes*, eds. Ciferri, O. & Dure, L. (Plenum, New York), pp. 221–230.
18. Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. & Allard, R. W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 8014–8018.
19. Murray, M. G. & Thompson, W. F. (1980) *Nucleic Acids Res.* **8,** 4321–4326.
20. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
22. Weeden, N. F. & Marx, G. A. (1984) *J. Hered.* **75,** 365–370.
23. Bailey, N. T. J. (1961) *Introduction to the Mathematical Theory of Genetic Linkage* (Clarendon, Oxford).
24. Suiter, K. A., Wendel, J. F. & Case, J. S. (1983) *J. Hered.* **74,** 203–204.
25. Allard, R. W. (1956) *Hilgardia* **24,** 235–278.
26. Blixt, S. (1974) in *Handbook of Genetics*, ed. King, R. C. (Plenum, New York), Vol. 2, pp. 181–221.
27. Weeden, N. F. (1984) *Plant Physiol.* **75,** s159.
28. McCarthy, B. J. (1980) in *Gene Families of Collagen and Other Proteins*, eds. Prockop, D. J. & Champe, P. C. (Elsevier/North–Holland, New York), pp. 37–46.
29. Kaufman, L. K., Thompson, W. F. & Briggs, W. R. (1984) *Science* **226,** 1447–1449.
30. Kaufman, L. K., Briggs, W. R. & Thompson, W. F. (1985) *Plant Physiol.* **78,** in press.