The EMBO Journal vol.5 no.5 pp.829-833, 1986

Molecular cloning of the *c* locus of *Zea mays*: a locus regulating the anthocyanin pathway

Javier Paz-Ares, Udo Wienand, Peter A.Peterson[1] and Heinz Saedler

Max-Planck Institut für Züchtungsforschung, Egelspfad, D-5000 Köln 30, FRG, and [1]Department of Agronomy, Iowa State University, Ames, IA 50011, USA

Communicated by H.Saedler

The *c* locus of *Zea mays*, involved in the regulation of anthocyanin biosynthesis, has been cloned by transposon tagging. A clone (#18En) containing a full size *En1* element was initially isolated from the *En* element-induced mutable allele *c-m668655*. Sequences of clone #18En flanking the *En1* element were used to clone other *c* mutants, whose structure was predicted genetically. Clone #23En (isolated from *c-m668613*) contained a full size *En1* element, clone #3Ds (isolated from *c-m2*) a *Ds* element and clone #5 (isolated from *c+*) had no element on the cloned fragment. From these data we conclude that the clones obtained contain at least part of the *c* locus. Preliminary data on transcript analysis using a 1-kb DNA fragment from wild-type clone #5 showed that at least three transcripts are encoded by that part of the locus, indicating that *c* is a complex locus.

Key words: Zea mays/*c* locus/molecular cloning/transposable elements

Introduction

In plants, particularly in *Zea mays*, the pathway leading to anthocyanin pigmentation is a genetically well-defined system (Coe and Neuffer, 1977; Dooner, 1982) and therefore is especially amenable to studying gene regulation at the molecular level. At least nine loci, which tentatively can be ordered in the following gene action sequence: *vp, c, r, c2, (pr), a1, a2, bz, bz2* are required for the production of anthocyanin pigment in the aleurone tissue (Coe, 1957; Reddy and Coe, 1962; Kirby and Styles, 1970; Reddy and Reddy, 1971; Styles and Ceska, 1972, 1977; Dooner and Nelson, 1977, 1979; McCormick, 1978; Dooner, 1983, 1985). Among these, *c, vp* and *r* are known to be involved in the regulation of the anthocyanin pathway.

The *c* locus is required for anthocyanin synthesis only in the aleurone and scutellum. Its regulatory effect was suggested by the observation that in a homozygous recessive *c* mutant the expression of at least two unlinked structural genes in the anthocyanin pathway, *c2* (encoding chalcone synthase, Dooner, 1983; Wienand *et al.*, 1986) and *bz* (coding for UFTG-transferase, Larson and Coe, 1977; Dooner and Nelson, 1977) are no longer detectable on the basis of enzyme activity measurements (Dooner and Nelson, 1979; Dooner, 1983). However, nothing is known about the molecular basis of its regulatory role.

The molecular cloning of the *c* locus, and the analysis of its structure and expression, could contribute to an understanding of the nature of its regulatory effects. In addition, the molecular analysis of *c* alleles like *C-I* (a dominant color-inhibiting mutant) and *c-p* (develops color, if germinated in the presence of light; see Kirby and Styles, 1970; Chen and Coe, 1977) would further unravel the regulatory function of the *c* locus.

Since the product of the *c* locus has yet to be identified, the cloning strategy chosen was that of transposon tagging (Bingham *et al.*, 1981), in which a transposon-induced mutant allele of the desired gene is initially isolated using element-specific sequences as a molecular probe. Subsequently segments of the mutant gene can be used to clone the wild-type gene.

This strategy in plants has been successfully applied in the isolation of the *bz* (Fedoroff *et al.*, 1984), *a1* (O'Reilly *et al.*, 1985) and *c2* (Wienand *et al.*, 1986) loci of *Z. mays* and the *pallida* locus (Martin *et al.*, 1985) of *Antirrhinum majus*.

Here we report cloning of the *c* locus from two *En* element-induced mutants (*c-m668655, c-m668613*; Reddy and Peterson, 1983), from one *Ds* element-induced mutant (*c-m2*; McClintock, 1948) and from a line containing the wild-type allele (c^+). For the first time the cloning of a gene involved in the regulation of gene expression in *Z. mays* is described. Preliminary data from transcript analysis suggesting a complex structure for the *c* locus is also presented.

Results

Isolation of the En-induced c-m668655 allele

The mutant *c-m668655* (Reddy and Peterson, 1983) was shown to contain only one functional *En* element in the genome, cosegregating with the *c* locus. For the isolation of this *c* allele, DNA was prepared from a homozygous *c-m668655* plant, partially digested with *Mbo*I and cloned into the λ phage EMBL4.

The recombinant clones ($\sim 2 \times 10^5$) were screened in parallel with three different fragments from the *En1* element (Pereira *et al.*, 1985) as probes. These three fragments (fragments 1, 2 and 3 in Figure 1a) represent almost the entire *En1* element and allowed the isolation of 14 clones which hybridized to all three probes. The DNA of these clones was analyzed by Southern hybridization for the presence of a complete *En* element. One clone, #18En, contained an *En* element with a restriction map (for the restriction enzymes used) indistinguishable from that of the *En1* element (Pereira *et al.*, 1985) (see Figure 1a).

The analysis of heteroduplexes formed between phage DNAs from clones #18En and *wx-844*-148 (this clone defines the *En1* element at the *wx* locus; Pereira *et al.*, 1985) revealed a double-stranded region of 8.5 kb (Figure 1b). This is in agreement with the size of 8.4 kb determined for *En1* (Pereira *et al.*, 1985). To examine the possibility that clone #18En also contained sequences derived from the *c* locus, a fragment flanking the *En* element was used to probe genomic libraries constructed from *Z. mays* lines with various *c* alleles.

Isolation of the wild-type c locus and other c mutants

If the *En* element present in the mutant *c-m668655* is integrated at the *c* locus one would expect to find gene-specific sequences flanking the element. Those should be present either as unique or low copy sequences in the genome. Indeed, the 1-kb *Eco*RI fragment (Figure 1a; fragment 4, wavy line) adjacent to the *En*

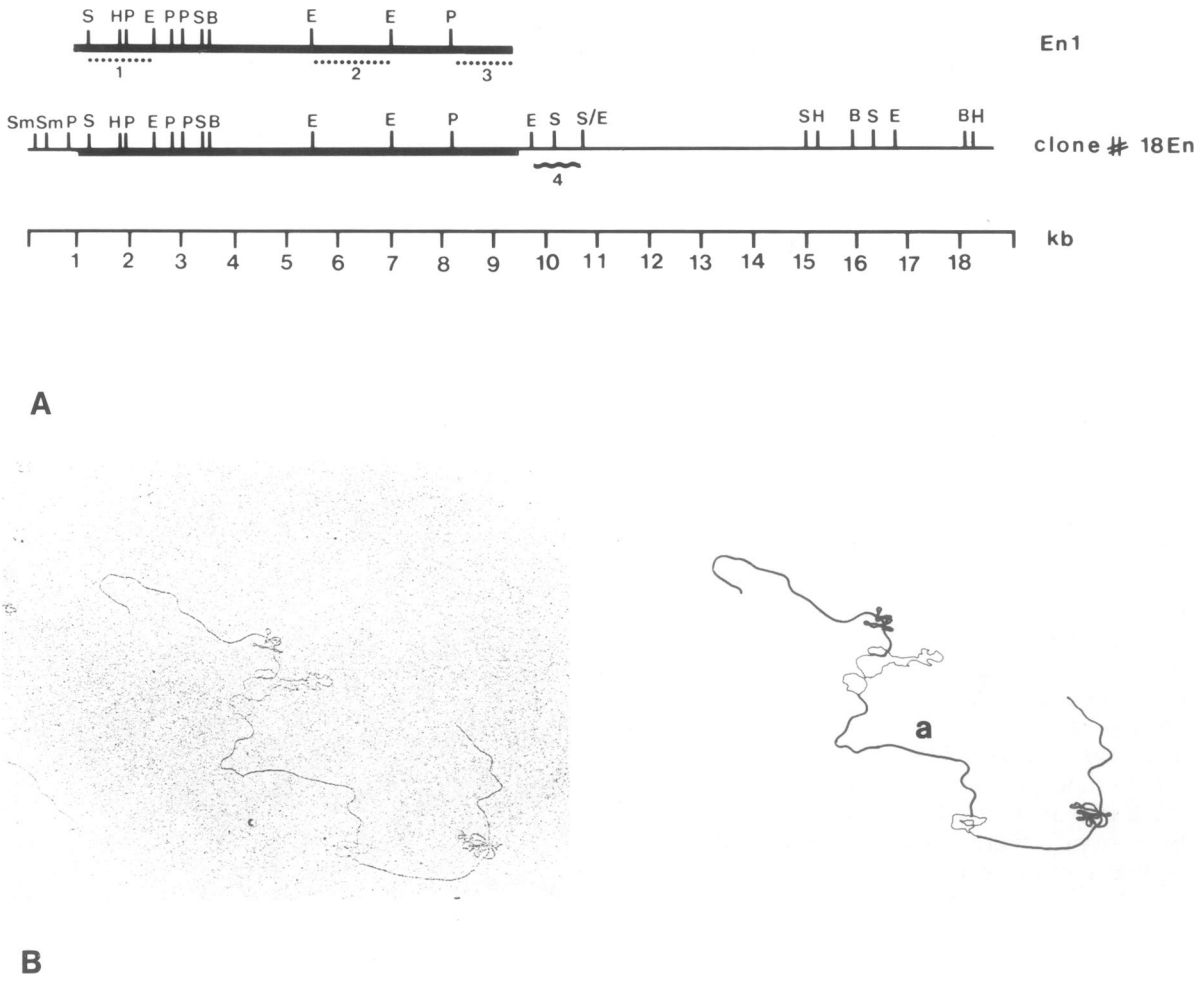

Fig. 1. Comparison between the *En1* element and the *En* element present on clone # 18En. (**A**) The physical map of the *En1* element (Pereira *et al.*, 1985) is compared with homologous sequences present on clone # 18En (isolated from *c-m668655::En*). The bold lines represent the regions corresponding to *En1* sequences. The dotted lines below the *En1* restriction map (no. 1, 2, 3 in the upper part of the figure) represent the probes used in the isolation of clone # 18En. The 1-kb *Eco*RI fragment present on clone # 18En (fragment 4, wavy line) was used as probe in the subsequent cloning experiments. For restriction enzymes the following abbreviations were used E = *Eco*RI, B = *Bam*HI, H = *Hin*dIII, S = *Sal*I, Sm = *Sma*I, P = *Pst*I. (**B**) Heteroduplex analysis between the λ clones *wx-m844*-148 (this clone carries the *En1* element at the *wx* locus, Pereira *et al.*, 1985) and # 18En. The length of homology between the inserts (a) of the two phages is 8.5 kb based on the measurements of six molecules.

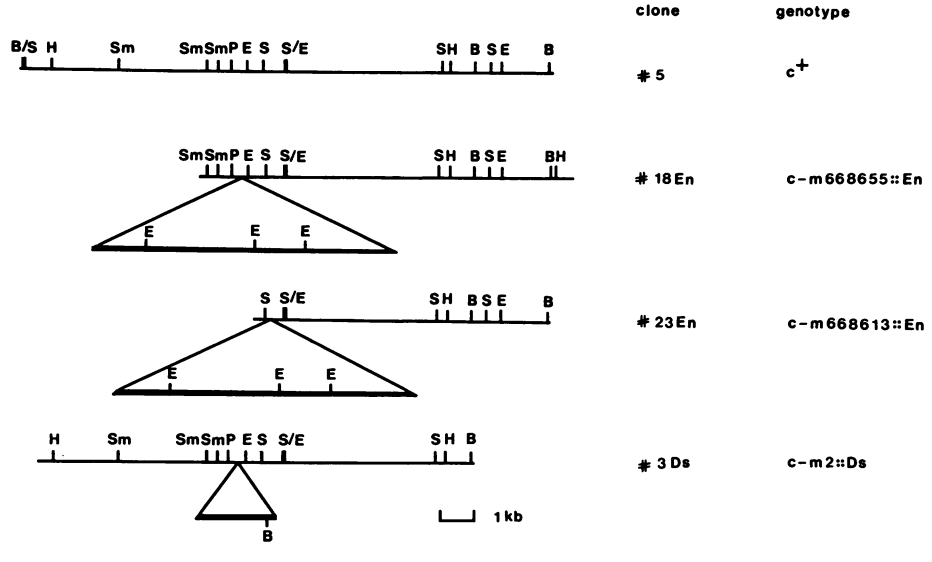

Fig. 2. Comparison of clones isolated from *En-* and *Ds*-induced *c* mutants with a homologous clone isolated from a wild-type *Z. mays* line. The physical map of clone # 5 [isolated from a wild-type *Z. mays* line (LC)] is compared with the restriction maps of clones # 18En, # 23En (isolated from *c-m668655::En* and *c-m668613::En*, respectively) and # 3Ds (isolated from *c-m2::Ds*). Symbols for restriction enzymes as in Figure 1.

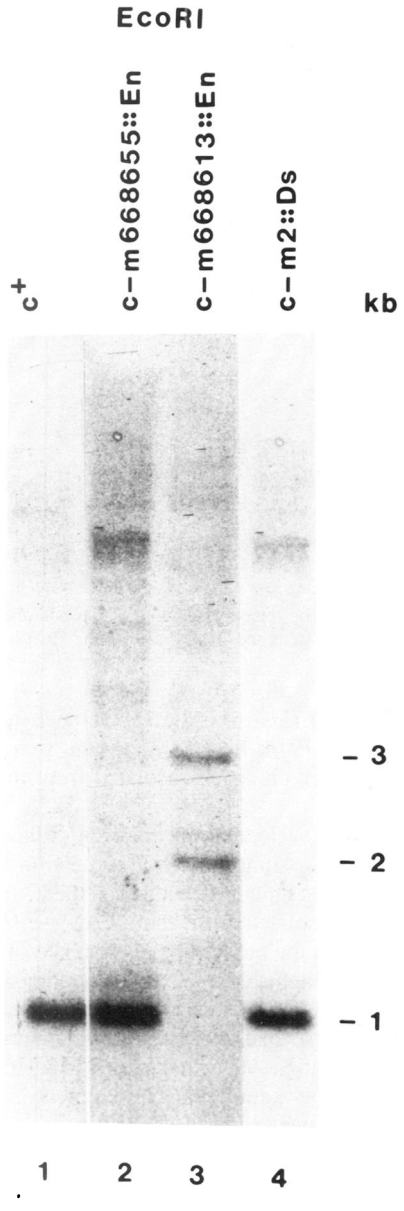

Fig. 3. Southern blot analysis of DNA isolated from wild-type and mutant *c* alleles. *Eco*RI-digested genomic DNA from plants with the genotypes shown above each lane were separated by electrophoresis in 0.8% agarose gels. After transfer to nitrocellulose, the filters were hybridized with the 1-kb *Eco*RI fragment present on wild-type clone #5 (see also fragment 4 of Figure 1). The size of the bands hybridizing is given in kilobases.

element in clone #18En contained only unique (low copy number) sequences as tested by Southern blot analysis with total *Z. mays* genomic DNA as a radioactive probe (data not shown). This fragment, if part of the *c* locus, should not be flanked by transposable element-specific sequences in a wild-type *Z. mays* line. In other transposable element-induced *c* mutants, however, it should have the corresponding element-specific sequences in its vicinity.

To test this assumption, the 1-kb *Eco*RI fragment was used as a probe to isolate clones from three different genomic libraries. These libraries were prepared using DNA from a wild-type *Z. mays* line (LC) and from two different transposable element-induced *c* mutants.

One of the mutants *c-m668613* (Reddy and Peterson, 1983) represents a second *c* allele containing an *En* element. The other mutant *c-m2* carries a stable *Ds* element at the *c* locus (McClintock, 1948). The libraries prepared from these two mutants were each probed in parallel with the 1-kb *Eco*RI fragment (Figure 1a, fragment 4) and either *En*- (Figure 1a, fragments 1 and 3) or *Ds*- (pAc7, see Materials and methods) specific probes, respectively.

The screening of the three libraries revealed three different clones isolated from wild-type (#5), *c-m668613* (#23En) and *c-m2* DNA (#3Ds). A partial restriction map of these clones is given in Figure 2. The integration sites of the elements also can be taken from these data.

The comparison of the restriction maps showed that the *En* elements present in clone #18En and #23En are identical in size. Both elements are integrated in the same orientation but at different positions. Heteroduplexes formed between the clones #23En and #18En also confirmed this finding (data not shown). The restriction map of clone #23En showed that the *En* element is integrated into the 1-kb *Eco*RI fragment previously mentioned (see Figure 1; fragment 4). This could be seen on the genomic level as well, using the 1-kb *Eco*RI fragment to probe *Eco*RI-digested DNAs isolated from the different *c* alleles. As shown in Figure 3 the probe lights up a 1-kb fragment in those alleles where no element is present in this fragment (Figure 3, lanes 1, 2 and 4). However, DNA isolated from the *c-m668613* allele, instead of a 1-kb band shows two other bands, 2 and 3 kb in size, reflecting the integration of the *En* element in that particular fragment (Figure 3, lane 3). The *Ds* element present in the clone #3Ds was compared with the *Ac* element (isolated from *wx-m7*, Müller-Neumann *et al.*, 1984) and, as a result of heteroduplex and Southern analysis, was found to be a deletion derivative of the *Ac* element (data not shown). The *Ds* element found in clone #3Ds interestingly has a stem and loop structure, as detected by heteroduplex analysis between the wild-type clone #5 and clone #3Ds (Figure 4). The stem is formed between the terminal parts of this element. Whether the sequences of the stem are element specific or the product of some other DNA rearrangement remains to be shown.

Genetically *c-m668655* and *c-m668613* should have an *En* element, *c-m2* a *Ds* element and c^+ no element at the *c* locus. These predictions are fulfilled in the clones obtained from these lines. We consider this to be strong evidence in support of our premise that the four clones described contain at least portions of the *c* locus.

Transcription of the c locus

The transposable elements described above are integrated close to or in the 1-kb *Eco*RI fragment present in all clones (Figure 2) thereby possibly interfering with the activity of the locus. We therefore reasoned that at least part of the transcription unit might be located at or close to this fragment. In Northern experiments this fragment was used to probe poly(A)$^+$ mRNA isolated from kernels of a wild-type *Z. mays* line (LC) 35 days after pollination. As show in Figure 5 (lane a) three transcripts, ~1.6, 1.4 and 0.3 kb in size, hybridize with this probe. When splitting the 1-kb *Eco*RI fragment into two similar-sized fragments the 0.3-kb transcript hybridizes to one of the fragments, while the 1.4- and 1.6-kb bands hybridize to the other (Figure 5; lanes b and c).

Discussion

Sequences specific for the c gene could be identified on a clone (#18En) isolated from an *En* element containing *c* allele (*c-m668655*). Therefore at least part of the *c* locus has been cloned.

The evidence for this statement comes from clones obtained

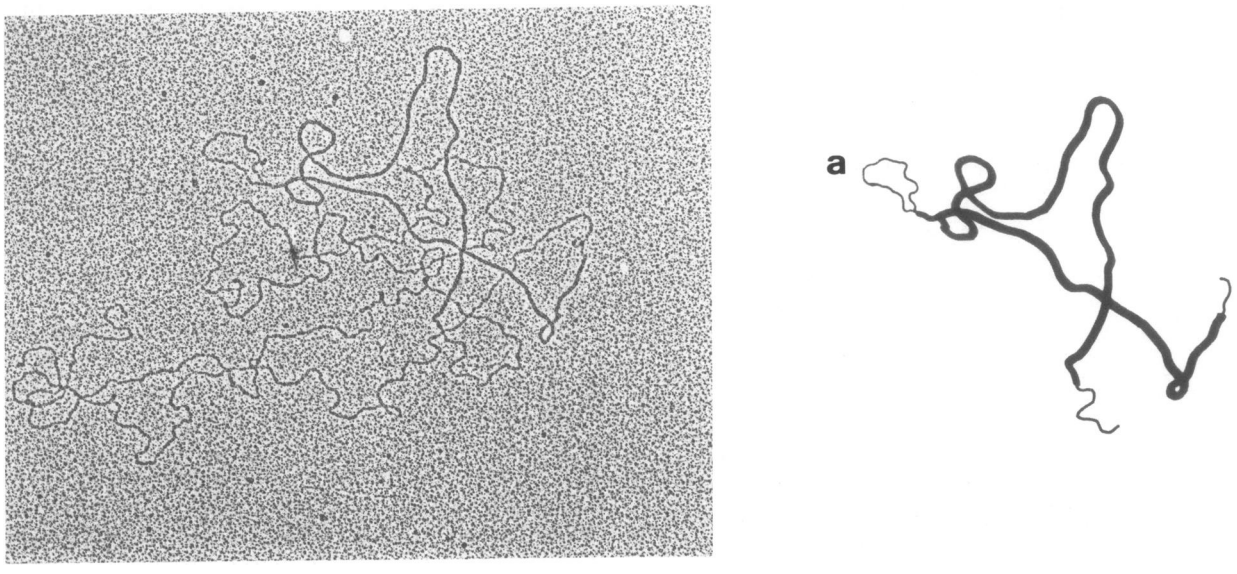

Fig. 4. Heteroduplex between the λ clone #3Ds and the 12 kb *Bam*HI insert of the wild-type clone #5. The region of homology expands over the whole 12-kb wild-type fragment. The *Ds* sequence of 2 kb (a) not hybridizing to the wild-type fragment shows a stem and loop structure (the stem is ~0.2 kb in size).

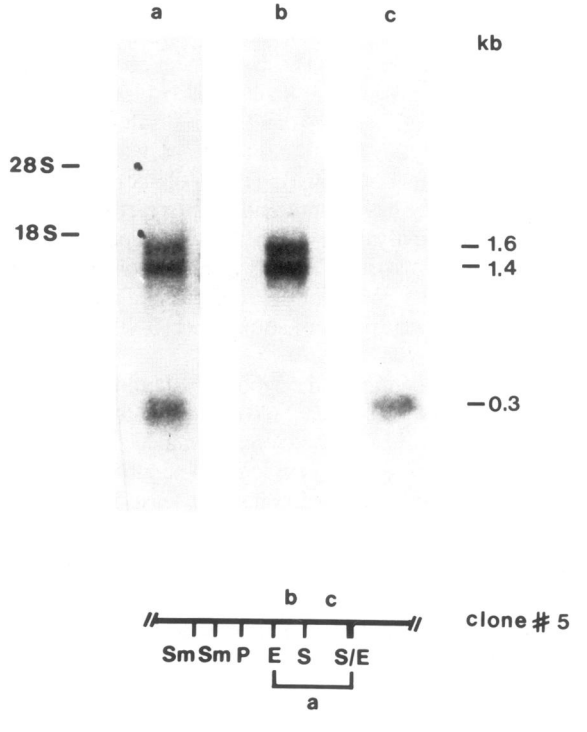

Fig. 5. Northern blot analysis with poly(A)$^+$ RNA. Poly(A)$^+$ RNA was isolated from kernels (35 days after pollination) separated on 1.5% agarose and transferred to nitrocellulose. **Lane a** shows the hybridization pattern using the 1-kb *Eco*RI fragment of clone #5 as a probe (fragment a of the partial restriction map of clone #5 shown below the lanes). Subfragments of this *Eco*RI fragment have been used as probes on the same RNA and light up either the two larger transcripts (**lane b**, using the *Eco/Sal* fragment b as probe) or the small transcript (**lane c**, using the *Sal/Sal* (*Eco*RI) fragment c as probe).

from three different *c* alleles, which on the molecular level showed the constitution that was predicted from genetic experiments. Clone #23En isolated from the *En* element-induced mutant *c-m668613*) (Reddy and Peterson, 1983) contained a full size *En* element and clone #3Ds from the *Ds* element-induced mutant *c-m2* (McClintock, 1948) had a *Ds* element present, while no element was found in clone #5, isolated from the wild-type *c* allele.

The analysis of *En* elements present on clones #13En and #23En supports the assumption that the functional *En* elements are very conserved as unique structures in the *Z. mays* genome. The two elements isolated from the alleles *c-m668655* and *c-m668613* indeed, on the basis of restriction sites and size determination, are indistinguishable from the previously isolated *En* elements at the *waxy* and *a1* locus (Pereira *et al.*, 1985; O'Reilly *et al.*, 1985). These characteristics prove the utility of the *En* element in the isolation of virtually any gene in the *Z. mays* genome in which the autonomous *En* element is integrated.

Using a small segment of the cloned *c* locus as a probe in Northern experiments it became apparent that this locus has a complex structure. The three transcripts observed differed considerably in size. There are two mRNAs, 1.6 and 1.4 kb in size which hybridize to the same genomic fragment (see Figure 5b). Whether this reflects products of different splicing and/or 5′ and 3′ heterogeneity remains to be shown. The third transcript, 0.3 kb in size, hybridized to a different genomic sequence. These results have to be confirmed, for example by cDNA cloning of the different transcripts and/or transcript analysis from various *c* mutants. However, compoundness of the *c* locus has already been predicted by McClintock on the basis of studies on the complementarity among different *c-m2* derivative alleles (McClintock, 1949). The failure to prove this in recombination tests (Coe, 1964) could be explained by the close proximity of the three transcripts.

The data currently available do not allow a definition of the function of the *c* locus. However, the size of one of its transcripts (0.3 kb) which might be too small to encode a protein, leads one to speculate on the *c* locus being, at least in part, a true regulatory gene, where an RNA molecule could be involved directly in its regulatory mechanism. Further detailed analysis of structure and expression of the *c* locus using *c* mutants like *C-I* or *c-p* hopefully will help to clarify the molecular nature of its regulatory function. Of these mutants, the *C-I* allele has already been cloned

and its characterization is currently underway (Paz-Ares, unpublished).

Materials and methods

Plant strains

The *c-m668655 (c-m668655::En)* and *c-m668613 (c-m668613::En)* strains containing an autonomous *En* element at the *c* locus were isolated independently (Reddy and Peterson, 1983).

The *c-m2 (c-m2::Ds)* containing a *Ds* element at the *c* locus was first described by McClintock (1948).

LC used as a source of the wild-type *c* allele is a color-converted W22 maize line developed by Dr R.A.Brinck, University of Wisconsin.

Isolation of plant, bacteriophage and plasmid DNA

Plant DNA was obtained from leaf material as described by Schwarz-Sommer *et al.* (1984). Isolation of plasmid DNA was done by the boiling method described by Holmes and Quigley (1981). After lysis, the plasmid DNA was precipitated with polyethylene glycol and purified twice on CsCl ethidium bromide gradients. Purification of phages and phage DNA was as described by Maniatis *et al.* (1982).

Cloning of plant DNA in λ EMBL4

λ EMBL4 vector arms were prepared by digestion of phage DNA with *Bam*HI and *Sal*I (Frischauf *et al.*, 1983) followed by purification on potassium acetate gradients (Maniatis *et al.*, 1982). *Mbo*I partial digestion of plant DNA was carried out as described by Maniatis *et al.* (1982). After electrophoresis on a 0.6% low melting agarose gel the DNA fragments in the size range of 15–20 kb were cut out of the gel and purified as described by Langride *et al.* (1980). Ligation and *in vitro* packaging were performed according to Wienand *et al.* (1982). The *Escherichia coli* strain K803 was used as host (Fedoroff, 1983). The selection of recombinant phages by plaque hybridization was done by the method of Benton and Davis (1977) using the probes mentioned in Results. Fragments to be subcloned were ligated in the plasmid pUC18 (Yanisch-Perron *et al.*, 1985).

Recombinant plasmids

The recombinant plasmids containing *En1* fragments (fragments 1, 2 and 3, see Figure 1a) were a gift from Andy Pereira and Zsuzsanna Schwarz-Sommer. The recombinant plasmid pAc7 (kindly provided by Markus Müller-Neumann) contained the entire *Ac* sequence present in the *wx* mutant *wx-m7* and was used as a probe to screen the library derived from the *Ds*-induced mutant *c-m2*.

Radioactive labelling

The nick translation of DNA was done as described by Maniatis *et al.* (1982).

Southern blot hybridization

Southern hybridizations (Southern, 1975; Wahl *et al.*, 1979) were carried out as described by Schwarz-Sommer *et al.* (1984) with minor modifications. The filters were washed several times in 0.2 × SSPE at 72°C (1 × SSPE = 10 mM NaH_2PO_4, 150 mM NaCl, 1 mM EDTA; pH 7.4).

Isolation of poly(A)$^+$ RNA

Total RNA from 35-day-old developing kernels was prepared by the method of Kloppstech and Schweiger (1976). Poly(A)$^+$ RNA was purified further by adsorption on oligo(dT)-cellulose (Aviv and Leder, 1972).

Northern hybridization of RNA

Poly(A)$^+$ RNA was fractionated in 1.5% agarose gels containing formaldehyde (Maniatis *et al.*, 1982). After electrophoresis, the gel was soaked in 20 × SSPE for 30 min and the RNA then transferred to nitrocellulose. The conditions for hybridization were the same as those used for Southern hybridization except that the washings were done in 2 × SSPE at 65°C.

Heteroduplex analysis

Heteroduplex analysis were performed essentially according to Davis *et al.* (1971) and Davis and Hyman (1971).

Acknowledgements

We thank Bill Martin and Arjula R.Reddy for critical reading of the manuscript. We wish to thank Andy Pereira and Zsuzsanna Schwarz-Sommer for providing the *En* element-specific plasmids and Markus Müller-Neumann for providing the pAc7 plasmid. The technical assistance of Irmgard Weisskirchen is gratefully acknowledged. We also thank Isolde Bertram and Margot Küsters for DNA heteroduplex analysis in the electron microscope. J.Paz-Ares was supported by a long-term EMBO post-doctoral fellowship.

References

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.
Bingham,P.M., Lewis,R. and Rubin,G.M. (1981) *Cell*, **25**, 693–704.
Chen,S.M. and Cole,E.M., Jr. (1977) *Biochem. Genet.*, **13**, 333–346.
Coe,E.H.,Jr. (1957) *Am. Nat.*, **91**, 381–385.
Coe,E.H.,Jr. (1964) *Genetics*, **50**, 581–588.
Coe,E.H.,Jr. and Neufer,G.M. (1977) In Sprague G.F. (ed.), *Corn and Corn Improvement*. The American Society of Agronomy Inc., Madison, WI, pp. 111–223.
Davis,R.W. and Hyman,R.W. (1971) *J. Mol. Biol.*, **62**, 287–301.
Davis,R.W., Simon,M. and Davidson,N. (1971) *Methods Enzymol.*, **21**, 413–428.
Dooner,H.K. (1982) In Sheridan,W.F. (ed.), *Maize for Biological Research*. Plant Molecular Biology Association, University Press, pp. 123–128.
Dooner,H.K. (1983) *Mol. Gen .Genet.*, **189**, 136–141.
Dooner,H.K. (1985) *Plant Physiol.*, **77**, 486–488.
Dooner,H.K. and Nelson,O.E. (1977) *Biochem. Genet*, **15**, 509–515.
Dooner,H.K. and Nelson,O.E. (1979) *Genetics*, **91**, 309–315.
Fedoroff,N.V. (1983) *Plant Mol. Biol. Rep.*, **1**, 27–29.
Fedoroff,N.V., Furtek,D.B. and Nelson,O.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3825–3829.
Frischauf,A.M., Lehrach,H., Poutska,A. and Murray,N. (1983) *J. Mol. Biol.*, **170**, 827–842.
Holmes,D.S. and Quigley,M. (1981) *Anal. Biochem.*, **114**, 193–197.
Kirby,L.T and Styles,E.D. (1970) *Can. J. Genet. Cytol.*, **12**, 934–940.
Kloppstech,K. and Schweiger,H.G. (1976) *Cytobiologie*, **13**, 394–400.
Langridge,J., Langridge,P. and Berquist,P.L. (1980) *Anal. Biochem.*, **103**, 264–271.
Larson,R. and Coe,E.H.,Jr. (1977) *Biochem. Genet.*, **15**, 153–156.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY.
Martin,C., Carpenter,R., Sommer,H., Saedler,H. and Coen,E.S. (1985) *EMBO J.*, **4**, 1625–1630.
McClintock,B. (1948) *Carnegie Inst., Washington Yearboook*, **47**, 155–169.
McClintock,B. (1949) *Carnegie Inst., Washington Yearbook*, **48**, 142–154.
McCormick,S. (1978) *Biochem. Genet.*, **16**, 777–785.
Müller-Neumann,M., Yoder,J.I. and Starlinger,P. (1984) *Mol. Gen. Genet.*, **198**, 19–24.
O'Reilly,C., Shepherd,N.S., Pereira,A., Schwarz-Sommer,Zs., Bertram,I., Robertson,D.S., Peterson,P.A. and Saedler,H. (1985) *EMBO J.*, **4**, 877–882.
Pereira,A., Schwarz-Sommer,Zs., Gierl,A., Bertram,I., Peterson,P.A. and Saedler,H. (1985) *EMBO J.*, **4**, 17–23.
Reddy,G.M. and Coe,E.H.,Jr. (1962) *Science*, **138**, 149–150.
Reddy,A.R. and Reddy,G.M. (1971) *Curr. Sci.*, **40**, 335–337.
Reddy,A.R. and Peterson,P.A. (1983) *Mol. Gen. Genet.*, **192**, 21–31.
Schwarz-Sommer,Zs., Gierl,A., Klösgen,R.B., Wienand,U., Peterson,P.A. and Saedler,H. (1984) *EMBO J.*, **3**, 1021–1028.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
Styles,E.D. and Ceska,O. (1972) *Phytochemistry*, **11**, 3019–3021.
Styles,E.D. and Ceska,O. (1977) *Can. J. Genet. Cytol.*, **19**, 289–302.
Wahl,G.M., Stern,M., Stark,G.R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3683-3687.
Wienand,U., Sommer,H., Schwarz,Zs., Shepherd,N.S., Saedler,H., Kreuzaler,F., Ragg,H., Fautz,E., Hahlbrock,K., Harrison,B. and Peterson,P.A. (1982) *Mol. Gen. Genet.*, **187**, 195–201.
Wienand,U., Weydemann,U., Niesbach-Klösgen,U., Peterson,P.A. and Saedler,H. (1986) *Mol. Gen. Genet.*, in press.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received on 3 February 1986