Nucleotide sequence of the maize transposable element *Mul*

Richard F.Barker*, David V.Thompson, Duncan R.Talbot, Jean Swanson¹ and Jeffrey L.Bennetzen²

Agrigenetics Advanced Research Division, 5649 East Buckeye Road, Madison, WI 53716,
¹Department of Genetics, University of California, Berkeley, CA 94720, and ²Department of
Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

Received 22 June 1984; Accepted 3 July 1984

ABSTRACT

A cloned DNA fragment from the maize allele Adhl-S3034 contains all of Mul, an insertion element involved in Robertson's Mutator activity. The element is 1367 base pairs (bp) long and is flanked by nine bp direct repeats of insertion site DNA. It has inverted terminal repeats of 215 and 213 bp showing 95% homology. Within the element are two direct repeats of 104 bp showing 96% homology. Four open reading frames (ORFs) were found, two in each DNA strand. Mul can be divided into two halves, each containing one terminal inverted repeat, an internal direct repeat, and two overlapping ORFs. The GC content of each half is high (70%), while that of a central 60 base portion of the element is low (26%). The central region contains the only sequence resembling the TAATA Goldberg and Hogness eukaryotic promoter signal. Multiple copies of DNA sequences related to Mul found in Mutator maize plants are generally similar in organization to the cloned element. A larger version containing a discrete 300 to 400 base pair insertion was found in some Mutator lines.

INTRODUCTION

Activity of the Robertson's Mutator system in maize is defined by frequent induction of mutations, many of them somatically and germinally unstable (1-4). Mutant alleles of the Adhl-S (alcohol dehydrogenase-1) gene have been isolated from Mutator plants which contain a 1400 base pair (bp) insertion in the first intron of the structural gene (4,5,6). All such mutants exhibit depressed alcohol dehydrogenase (ADH) mRNA and enzyme activity levels (4). The insertion element isolated from a genomic clone of Adhl-S3034, an allele which produces ADH enzyme and mRNA at 40% of the wild type level (5,6), possesses long terminal inverted repeats and a 9 bp direct duplication of target DNA flanking the element (4,5,7). Multiple copies (10-60 per diploid genome) of DNA sequences homologous to this insertion element, termed Mul (5,7), are present in maize lines which exhibit Mutator activity (4,7). Activity is not lost in progeny of outcrosses to lines originally lacking Mutator, and copies of sequences homologous to Mul have been maintained in these lines over several genera-

tions (1,4). Integration of elements homologous to Mul into new genomic DNA sites during outcrossing has been demonstrated (M. Allemen, cited in 4; J. L. Bennetzen, unpublished).

Characteristics of Mul, and perhaps of maize DNA sequences related to it, namely alteration of gene expression, inverted terminal repeat sequences, and target site DNA duplications, strongly indicate (see 8) that Mul is an active transposable element in the Robertson's Mutator system. In order to elucidate some of the molecular properties of Mutator elements, we have determined the nucleotide sequence of Mul. The GC content of Mul is strikingly dissimilar from that of the whole maize genome. Mul contains 4 open reading frames, each beginning with an ATG codon, two on each DNA strand. The element can be divided into two symmetrical halves, each containing a long inverted terminal repeat, a long direct internal repeat, a region of very high GC content, and a pair of open reading frames. We also describe results showing that most DNA sequences in Mutator plants homologous to Mul are the same size as Mul itself, although some Mutator plants contain a few copies of slightly larger Mul-like elements.

MATERIALS AND METHODS

Materials

A genomic clone of Adhl-S3034 (5) was previously subcloned into pBR322 by insertion of a 2.9 kbp BamH I-Hind III fragment containing all of Mul. This subclone, termed pMJ9, was used for nucleotide sequence determinations and was provided by M. J. Freeling and W. C. Taylor, University of California, Berkeley, CA and J. L. Bennetzen, Purdue University, West Lafayette, IN.

Maize plants from four lines exhibiting Mutator activity were used as sources of genomic DNA. These lines, referred to as Mu A, B, C, and D were obtained from D. S. Robertson, Iowa State University, Ames, IA, and from M. J. Freeling.

Ultra pure urea was obtained from BRL (Gaithersburg, Maryland), polyacrylamide from BDH (Poole, England), calf intestinal alkaline phosphatase from Boehringer (Mannheim, W. Germany), polynucleotide kinase from P.L. Biochemicals Inc. (Milwaukee, Wisconsin), and [γ -32P] ATP from New England Nuclear (Boston, Massachusetts). The restriction enzymes Ava I, Bgl I, BstE II, Mst I, Nco I, and Tth I were from New England Biolabs (Beverly, Massachusetts), Hind III was purchased from Promega Biotec (Madison,

Wisconsin) and Sst II from BRL (Gaithersburg, Maryland). All were used according to suppliers' recommendations. Chemicals used for DNA sequencing reactions were from vendors recommended by Maxam and Gilbert (9). Kodak (Rochester, NY) X-Omat AR-5 film was supplied as long rolls. All other reagents were of analytical grade unless otherwise stated.

Methods

All sequencing reactions were performed using the method of Maxam and Gilbert (9) and the sequencing gels were run according to Ansorge and Barker (10) with modifications reported by Barker *et al.* (11). Computer analyses of the DNA and protein sequences were performed using computer programs (12) made available by Drs. O. Smithies, J. Devereaux, and F. Blattner (University of Wisconsin, Madison). The techniques for maize genomic DNA purification from single plants and nitrocellulose blot hybridization analysis have been reported previously (5,7).

RESULTS AND DISCUSSION

Nucleotide sequence of Mul

The strategy employed for sequencing clone pMJ9 and the restriction map derived for Mul are presented in Figure 1. The map shown here agrees with one published elsewhere (7) with regard to sites for restriction endonucleases Tth III I, Ava I, Bgl I, BstE II, and BstN I. It differs with respect to some sites such as Sst II and Ava II, which are found in dense clusters. These clusters also carried dense arrays of sites for Tac I, Fnu4H I, and Mnl I which occur in GC-rich stretches of DNA. Analy-

Figure 1.
Physical map showing restriction enzyme sites within clone pMJ9 and sequencing strategy for Mul. The dots indicate the sites of 5'-end labeling and the arrows show the distance and direction of sequencing.

sis of the nucleotide sequence showed that the clusters result from extensive reiterations of short sequences. These GC-rich regions will be further discussed below.

The complete nucleotide sequence of Mul is shown in Figure 2. The 1367 bp Mul sequence shown here is flanked by duplications of the genomic DNA sequence TTTTGGGGA (positions 5-13 and 1380-1388) as previously shown (5). The Mul element is delineated by two long terminal inverted repeats extending for 215 and 213 bp respectively. These repeats are 95% sequence homologous and match perfectly over the terminal 28 bp (Fig. 2, positions 14-42 and 1352-1380). These terminal 28 bp within the inverted repeats of Mul are not homologous to the inverted repeat termini of the Ds, Cntl, Spm-18, Tam1, or soybean lectin insertion elements characterized previously in plants (13,14,15,16,17). The complete inverted terminal repeats of Mul are longer than those of any other plant insertion element characterized to date. Only the Spm-18 element, which has terminal inverted repeats roughly 100 bp long (15), is comparable to Mul in this respect. However, there are several mismatched regions up to 29 nucleotides in length within the inverted terminal repeats of Spm-18. The inverted terminal repeats of Cin 1, Tam 1, and the soybean lectin insertion have as well several stretches containing multiple mismatched nucleotides (14,16,17), and thus differ from the long inverted terminal repeats of Mul. The organization pattern of Mul with respect to length of inverted terminal repeats resembles those of DIRS elements in Dictyostelium, FB elements in Drosophila, and the bacterial transposon Tn10 (18,19,20). However, analysis of the nucleotide sequences of the Mul inverted repeats reveals no substructure of short repeats as is found in FB elements (19). Also, no promoter activity has yet been identified residing within the long terminal repeats of Mul as occurs in DIRS elements (18). As will be discussed below, the only eukaryotic promoter-like sequences within Mul are found elsewhere in the element. The long terminal inverted repeats of

Figure 2.

The complete nucleotide sequence of Mul. Asterisks above the sequence indicate 9 base pair duplications of the Adhl allele AdhI-S3034 DNA at the borders of the insertion. The solid line beneath the sequence shows the extent of the long terminal inverted repeats. The broken line above delineates the extent of the long internal direct repeats. The solid line within the direct repeats shows the area of strong homology with the maize Adhl gene. The box encloses the 60 base pair area of low GC content containing the possible eukaryotic promoter signals which are lined. The four open reading frames are indicated by single letter amino acids.

Figure 3.

Composite diagram of several features found in the nucleotide sequence of Mul. Map positions are taken from the sequence shown in Figure 2.

- A) Plot of the percentage GC content, within 20 base segments across Mul. Dotted line represents 49.5% GC, which is the GC content of total maize DNA (23).
- B) The positions and numbers of short stem loops found within Mul. These were summed over 100 bp regions. See text for computer analysis parameters used.
- C) The positions of the inverted (long opposing arrows) and direct (short arrows) repeats.
- D) Restriction map of Mul.
 A. Ava I, A2. Ava II, B. Bgl I, BE. BstE II, BN. BstN I, H. Hinf I, M. Mst I, ML Mlu I, N. Nco I, S. Sst II, T. Tth III I, Tq. Taq I.
 The position and structure of a possible eukaryotic promoter signal is indicated.
- E) The lengths and polarities of four open reading frames.

Mul contain recognition sites for restriction endonucleases Hinf I, Taq I, Tth III I, and Mlu I. The left-hand inverted repeat is defined by the presence of a site for Mst I (Figures 1 and 3D). Use of these sites allows separation of Mul-like sequences from other genomic DNA fragments, as will be shown in the next section. The structure of Mul is unique when compared with other characterized eukaryotic transposable elements in that it contains two 104 bp, 96% homologous, long internal direct repeats beginning at positions 468 and 870. These direct repeats contain sites for restriction endonucleases Sst II and Bgl I (Figures 1 and 3D) which allow isolation of the central portion of Mul.

Homologies found between eukaryotic transposable elements and target genes may be involved in conferring site specificity or preference to

transposition processes (17,21). No homologies between Mul and Adhl sequences were found near the borders of the insertion. However, the long internal direct repeats each contain a 25 bp region which shows direct homology (76%) to a region in the second exon of the maize Adhl gene (22):

<u>Mul</u> Direct Repeats	5'-C G C C G C C G C C G G C C C G G A G G C G A G-3'
Maize <u>Adhl</u>	5'-C G C C T C C G C A G G C C A T G G A G G T G C G-3'

This 25 bp region of Adhl DNA, which is 76% GC, occurs 526 bases 3' to the point of insertion. Within Mul, these 25 bp regions occur 513 bp from the left end (position 14) and 479 bp from the right end (position 1379). Although these regions constitute a striking feature of the Adhl-S3034 allele, a role for involvement of sequence homologies such as this in the recognition of insertion sites by Mutator elements has not yet been demonstrated. Analysis of more insertion events is needed to determine whether these repeats are functional in the transposition of Mutator elements.

Mul contains four open reading frames (ORFs), each beginning with an ATG initiation codon (Figures 2 and 3E). The open reading frames, numbered 1 to 4, extend for 402, 621, 435, and 483 bp and potentially encode polypeptides of 13.6, 21.2, 15.7, and 17.0 kilo Daltons (kD) respectively. The amino acid composition data shown in Table 1 indicate that polypeptides 2 and 3 have similar compositions owing to the fact that they are in the same reading frame with respect to the placement of the long internal direct repeat, giving them an area of high homology. Polypeptides 1 and 2 would have neutral isoelectric points and 3 and 4 would be basic, assuming that no post-translational processing occurs.

The average GC content of Mul, 67%, is very different from that of the maize genome as a whole, which has been estimated to be 49.5% (23). Within the long terminal inverted repeats, the GC content averages 51.5%, which is comparable to maize DNA. The GC contents of the regions containing ORFs 1 and 2 and ORFs 3 and 4 average 66% and 70% respectively. Near the centers of the ORFs are several short, approximately 20 bp, regions containing 90% GC. These very highly GC-rich regions correspond with positions in the 104 bp long internal direct repeats which themselves average 79% GC and which, as stated above, contain dense clusters of GC-rich restriction sites. Subsequently, a marked codon bias is introduced at second and/or third base positions for G and C residues (Table I). However, codon composition in ORFs 1-4 is not restricted. All coding and termination triplets except TAT (tyr) were found. Comparison

Table 1. Amino acid composition of polypeptides 1 to 4.

Proteins		1	2	3	4	Proteins		1	2	3	4
Ala (A)	GCA	2	4	0	1	Leu (L)	CUU	1	2	0	0
	GCC	5	20	7	4		UUA	0	2	1	1
	GCG	5	16	7	5		UUG	4	3	2	3
	GCU	2	5	1	3		Total	13	23	14	9
	Total	14	45	15	13	Lys (K)	AAA	0	2	2	0
Arg (R)	AGA	0	2	3	0		AAG	1	0	8	2
	AGG	4	1	5	5		Total	1	2	10	2
	CGA	0	4	1	2	Met (M)	AUG	4	2	5	2
	CGC	0	6	0	14	Phe (F)	UUC	6	4	2	2
	CGG	2	2	1	6		UUU	4	1	0	0
Asn (N)	CGU	1	4	1	4		Total	10	5	2	2
	Total	7	19	11	31	Pro (P)	CCA	1	1	5	0
	AAC	1	2	3	2		CCC	0	5	6	1
	AAU	2	1	1	1		CCG	6	6	10	2
	Total	3	3	4	3		CCU	0	5	1	0
Asp (D)	GAC	5	10	8	7	Ser (S)	Total	7	17	22	3
	GAU	0	1	0	1		AGC	6	3	2	4
	Total	5	11	8	8		AGU	2	0	3	3
Cys (C)	UGC	3	1	0	4		UCA	1	0	1	0
	UGU	0	2	1	0	Thr (T)	UCC	5	6	4	3
	Total	3	3	1	4		UCG	2	2	7	3
Gln (Q)	CAA	1	3	1	0		UCU	3	1	0	0
	CAG	1	4	4	2		Total	19	12	17	13
	Total	2	7	5	2	Trp (W)	ACA	1	0	0	1
Glu (E)	GAA	0	4	3	2		ACC	1	1	1	1
	GAG	4	6	3	11		ACG	2	7	3	2
	Total	4	10	6	13		ACU	0	3	0	0
Gly (G)	GGA	3	5	0	5		Total	4	11	4	4
	GGC	12	4	6	20	Tyr (Y)	UGG	3	1	2	0
	GGG	5	2	0	4		UAC	2	4	3	1
	GGU	3	1	1	1	Val (V)	UAU	0	0	0	0
His (H)	Total	23	12	7	30		Total	2	4	3	1
	CAC	0	4	1	2		GUA	0	2	1	4
	CAU	1	0	0	0		GUC	2	7	1	8
Ile (I)	Total	1	4	1	2		GUG	2	2	2	4
	AUA	1	0	1	0	End(.)	GUU	0	3	0	1
	AUC	1	1	1	0		Total	4	14	4	17
	AUU	2	0	1	1		UAA	0	1	0	1
Leu (L)	Total	4	1	3	1		UAG	1	0	0	0
	CUA	1	0	2	0		UGA	0	0	1	0
	CUC	4	12	5	4	M. wt.	kD	13.6	21.5	15.7	17.0
	CUG	3	4	4	1						

of the codon composition in Mul with codon usages in two other maize genes, Adhl and a 19 kD zein storage protein (22,24), showed that the only codons found in Mul exclusively are arginine triplets CGA and CGT. Thus, it is likely that Mul transcripts would be compatible with maize translation systems. A consequence of the nucleotide composition of Mul is that the four ORFs may have arisen fortuitously. There is a high statistical

Figure 4

Nitrocellulose blot hybridization analysis of genomic DNAs from individual Robertson's Mutator maize plants. DNA samples were digested with the restriction endonuclease(s) identified above each lane, fractionated on agarose gels, and transferred to nitrocellulose. The nitrocellulose blots were hybridized with Mul probe prepared by digesting clone pMJ9 DNA with Tth IIII I, followed by nick translation of the isolated insert. Hybridizations were performed in a buffer system containing 0.98 M Na⁺ at 70°C (7). Following hybridization, blots were washed extensively at 65°C in a buffer containing 0.04 M Na⁺ (7). Mu A, B, C and D represent individuals from different maize lines containing Mutator. Molecular weight markers in kbp are from Eco RI + Hind III-digested bacteriophage lambda DNA. **Panel A:** Digestion of DNAs from Mutator plants using enzymes with recognition sites in the long terminal inverted repeats of Mul. Molecular weights indicated are approximate sizes in kbp of bands obtained in a Tth IIII I digest of Mutator plant A DNA. **Panel B:** Digestion of Mutator plant D DNA with Tth IIII I singly (lane 8) or in combination with BstE II or BstNI I (lanes 9 and 10). Digests with BstNI I alone or with Bgl II, which does not cut within Mul, are shown for comparison (lanes 11 and 12). Lane 13 is a single haploid genomic copy equivalent of clone pMJ9 DNA digested with Tth IIII I.

probability of long ORFs occurring randomly in exceptionally GC-rich DNA regions (E. J. Gubbins, I. M. Larrinua, and L. Bogorad, unpublished). Proof that ORFs 1-4 in Mul are functional requires further investigation, including determination whether Mul sequences are transcribed in maize.

In a small central portion of Mul, between bases 659-709, the GC content drops to 26%. The four ORFs do not cross this region (Fig. 3E). Thus, Mul can be divided into halves which range from 26% to 52% GC at their ends, reach 90% GC near their centers, and which contain one long inverted terminal repeat copy, one long direct internal repeat copy, and one pair of ORFs.

The low GC central region of Mul contains the sequence TATTATTTT which is the only sequence in the element resembling the 'TAATA' Goldberg-Hogness eukaryotic promoter signal (25). Short intercistronic regions of low GC content having regulatory functions are found in other transposable elements (26). The nucleotide sequence of Mul was searched for other eukaryotic regulatory signals. Sequences with perfect homology to the eukaryotic promoter CCAAT sequence (26) were not found, but several with partial homology occur. For example the sequence GCAAT, which may be involved in transcription of ORF1, occurs at position 747. One sequence, AATAA (position 1244), which is homologous to the poly(A) addition sequence AATAAA found in animal systems (27) was observed +74 bp from the termination codon of ORF 3 (Fig. 2). In maize, the sequences AATAAG and/or AATAAA appear to be used for polyadenylation of a 19 kD zein mRNA (24). Because short inverted repeats might constitute recognition and regulatory sites on DNA (11,28), the secondary structure of Mul was analyzed by determining the positions of stem loops along its length. The stem loop parameters chosen were a minimum stem length of 10 bp with a bond value of 20 arbitrary units where paired G-C=3, A-T=2, G-T=1, and a maximum loop size of 20 bases (12). The numbers and positions of stem loops are shown in Figure 3B. A comparison with Figure 3E shows that there is a correlation between the numbers of short stem loops and the positions of open reading frames. This result appears to be related to the GC-richness of the central portions of the open reading frames rather than to the placement of regulatory signals in the DNA.

Analysis of sequences related to Mul in maize genomic DNA

Nitrocellulose blot hybridization analysis of genomic DNAs from individuals of Mutator lines A,B,C, and D shows that multiple copies of sequences homologous to Mul are present in their genomes (Ref. 7; Fig. 4). Digestions with Bgl II, which does not cut Mul, or with BstE II, which cuts Mul in one position (position 751), show multiple copies of the element dispersed in the genome of Mutator plant D (Fig. 4B, lanes 11 and 12). Digestion with Tth III I, Hinf I, Taq I, and Mlu I, which have

recognition sites in Mul only within the long terminal inverted repeats, yielded intensely hybridizing bands of approximate sizes ranging from 1000 to 1300 bp, as predicted from the nucleotide sequence of Mul (Fig. 1; 4A, lanes 1,5,6, and 7). These bands are present in several copies per genome, as seen from a comparison with a single copy reconstruction sample prepared by digestion of pMJ9 with Tth 111 I (Fig. 4B, lane 13). Most Mul-related sequences gave the fragment sizes predicted from the nucleotide sequence, and are thus similar in organization to Mul (Fig. 4A, lanes 2 and 3). Minor classes of larger fragments were also observed in Mutator maize lines A and D (Fig. 4A, lanes 1 and 4). The large fragments are approximately 350 bp greater in size than the major classes of Mul-related elements found in Mutator plants A,B,C, and D. Large variants are present in multiple copies, approximately two to five depending on the line examined (Ref. 7; Fig. 4, compare lanes 1,4,8, and 13). These larger fragments are not partial digestion products because they are of conserved size in all lines, and they appear whether Tth 111 I, Mlu I, or Hinf I are used as digesting enzymes. These restriction enzyme sites map within the larger elements in the order expected from the nucleotide sequences of the long terminal inverted repeats in Mul (Figs. 3 and 4A, lanes 4,5, and 7), suggesting that the large variants have inverted repeat structures homologous to the ones in Mul. The fact that all of the variants in a given line are the same size suggests that they have incorporated discrete insertions of DNA into the basic Mul structure.

Figure 4B depicts nitrocellulose blot hybridization analysis of genomic DNA from Mutator plant D which contains approximately five copies of a large variant (compare lanes 8 and 13). Digestion with Tth 111 I together with BstN I (which has one recognition site in Mul, at position 1026) gave two pairs of bands, the larger bands of which are BstN I partial digestion products differing from the smaller by approximately 150-200 bp. A band at approximately 1175 bp (arrow, Fig. 4B, lane 10), differs by approximately 350 bp from the major BstN I/Tth 111 I band (asterisk, Fig. 4B, lane 10), which indicates the insertions in the large variants are to the left of the BstN I sites. A 175 bp band produced from DNA between the BstN I sites and the Tth 111 I sites within the right-hand long terminal inverted repeats is not shown in the figure. Similar analysis of genomic DNA digested by BstE II in combination with Tth 111 I gave results consistent with the insertions being positioned to the right of the BstE II sites (Fig. 4B, lane 9). Our data also show that the insertions have

internal Taq I sites (Fig. 4A, lane 6). The insertions within the larger Mul-related variants in Mutator plant D are thus localized to a region between the BstE II and BstN I sites. These insertions may interrupt DNA sequences corresponding to open reading frames 3 and 4 within Mul. The exact organization of these variants awaits confirmation at the nucleotide sequence level.

ACKNOWLEDGEMENTS

The authors thank M. Freeling for providing a sample of pMJ9, Mutator maize lines, unpublished materials, and for many helpful discussions, D. Robertson for Mutator lines, L. Blomquist, K. McIntosh, and R. P. Fracasso for technical assistance, and J. Strommer for helpful discussions.

Agrigenetics Advanced Research Division Manuscript No. 28.

*To whom reprint requests may be addressed

REFERENCES

1. Robertson, D. S. (1978) *Mutat. Res.* 51, 21-28.
2. Robertson, D. S. and Mascia, P. N. (1981) *Mutat. Res.* 54, 283-289.
3. Freeling, M., Cheng, D. S.-K., and Alleman, M. (1982) *Develop. Genet.* 3, 179-196.
4. Freeling, M. (1984) *Ann. Rev. Plant Physiol.* 35, 277-298.
5. Bennetzen, J. L., Swanson, J., Taylor, W. C., and Freeling, M. (1984) *Proc. Natl. Acad. Sci. USA* 81, in press.
6. Strommer, J. N., Hake, S., Bennetzen, J. L., Taylor, W. C., and Freeling, M. (1982) *Nature* 300, 542-544.
7. Bennetzen, J. L. (1984) *J. Mol. Appl. Genet.* 2, in press.
8. Shapiro, J. A., Ed. (1983) "Mobile Genetic Elements", p. 688, Academic Press, New York.
9. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
10. Ansoorge, W., and Barker, R. F. (1984) *J. Biochem. Biophys. Methods*, in press.
11. Barker, R. F., Idler, K. B., Thompson, D. V., and Kemp, J. D. (1983) *J. Plant Mol. Biol.* 2, 335-350.
12. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucl. Acids Res.* 12, 387-395.
13. Döring, H. P., Tillman, P., and Starlinger, P. (1984) *Nature* 307, 127-130.
14. Saedler, H., Bonas, J., Deumling, B., Gupta, M., Hahlbrock, K., Harrison, B. J., Kreuzaler, F., Peterson, P. A., Reif, J., Schwarz-Sommer, Z., Shepard, N., Sommer, H., Ubben, D., and Wienand, U. (1983) in "Genetic Rearrangement", Chater, K. F., Cullis, C. A., Hopwood, D. A., Johnston, A. W. B., and Woolhouse, H. W., Eds., pp. 107-115, Sinauer Associates, Sunderland, MA, U.S.A.
15. Schwarz-Sommer, Zs., Gierl, A., Klösgen, R. B., Wienand, U., Peterson, P. A., and Saedler, H. (1984) *EMBO J.* 3, 1021-1028.
16. Bonas, U., Sommer, H., and Saedler, H. (1984) *EMBO J.* 3, 1015-1019.
17. Vodkin, L. O., Rhodes, P. R., and Goldberg, R. B. (1983) *Cell* 34,

- 1023-1031.
18. Zuker, C., Cappello, J., Lodish, H. F., George, P., and Chung, S. (1984) *Proc. Natl. Acad. Sci. USA* 81, 2660-2664.
 19. Truett, M. A., Jones, R. S., and Potter, S. S. (1981) *Cell* 24, 753-763.
 20. Kleckner, N. (1983) in "Mobile Genetic Elements", Shapiro, J. A., Ed., pp. 261-298, Academic Press, New York.
 21. Farabaugh, P. J. and Fink, G. R. (1980) *Nature* 286, 352-356.
 22. Dennis, E. S., Gerlach, W. L., Pryor, A. J., Bennetzen, J. L., Inglis, A., Llewellyn, D., Sachs, M. M., Ferl, R. J., and Peacock, W. J. (1984) *Nucl. Acids Res.*, in press.
 23. Hake, S., and Walbot, V. (1980) *Chromosoma* 79, 251-270.
 24. Geraghty, D., Peifer, M. A., Rubenstein, I., and Messing, J. (1981) *Nucl. Acids Res.* 9, 5163-5174.
 25. Breathnach, R., and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349-383.
 26. Heffron, F., McCarthy, B. J., Ohtsubo, H., and Ohtsubo, E. (1979) *Cell* 18, 1153-1163.
 27. Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) *Nucl. Acids Res.* 8, 127-142.
 28. Panayotatos, N., and Wells, R. D. (1981) *Nature* 289, 466-470.