The Structure and Paramutagenicity of the $R$-marbled Haplotype of *Zea mays*

Tadas Panavas, Jessica Weir and Elsbeth L. Walker

Biology Department, University of Massachusetts, Amherst, Massachusetts 01003

Manuscript received April 9, 1999
Accepted for publication June 21, 1999

ABSTRACT

Paramutation is the meiotically heritable silencing of a gene that can occur in particular heterozygous combinations. The $R$-marbled ($R$-mb) haplotype is paramutagenic: it causes paramutable $r_1$ haplotypes like $R-r$ to become heritably silenced. $R$-mb was found to comprise three distinct $r_1$ genes arranged as direct repeats. The most distal gene of $R$-mb, $Scm$, contains a novel transposable element, $Sho$ ($Sho$). Excision of the $Sho$ element early in aleurone development results in the characteristic “marbled” aleurone pigmentation pattern conferred by $R$-mb. The effect of gene copy number on the paramutagenic strength of $R$-mb was tested. Paramutagenic strength of $R$-mb is directly correlated with $r_1$ gene copy number. Paramutagenic strength of $R$-mb was not affected by removal, through crossing over, of the $Sho$ transposon. Finally, $R$-mb does not appear to contain the transposable element, $Doppia$, which is associated with paramutability of $R-r$, and has been suggested to play a role in paramutagenicity of another paramutagenic haplotype, $R$-stippled.

The tissue-specific pattern of anthocyanin (red and purple pigment) deposition in maize is determined by members of the $r_1$ and structurally related $b_1$ gene families. The $r_1$ family includes a large number of single-gene and multigene haplotypes of the $r_1$ locus, plus the chromosomally displaced $r_1$ genes $Sn$ and $Lc$ (reviewed by Dooner et al. 1991). Some $r_1$ haplotypes are subject to the epigenetic silencing phenomenon, paramutation. Two key features define the phenomenon of paramutation. First, two variants of a gene or locus interact in a heterozygote such that one of the variants becomes epigenetically silenced. Second, the silencing is meiotically heritable, i.e., it is maintained in subsequent generations. Although paramutation was first described over 40 years ago (Brink 1956), a mechanism that explains how one copy of a gene can heritably affect the expression of another has remained obscure. There has been a recent resurgence of interest in paramutation, in part due to its similarity to various other epigenetic phenomena like cosuppression (Jorgensen 1995) and other forms of transgene silencing (Meyer and Saedler 1996) in plants, the methylation induced premeiotically (MIP) system of *Aspergillus immersus* (Rossignol and Faugeron 1994; Colot et al. 1996), the repeat induced point mutation (RIP) system of *Neurospora crassa* (Selker and Garrett 1988), polycistron group-mediated regulation of transcription (Pirotta 1996) and trans-sensing effects in *Drosophila* (Tartof and Henikoff 1991; Henikoff 1994), and regulation of transposable element activity (Brutnell and Dellaporta 1994; Fedoroff 1996; Martienssen 1996). Several systems of heritable epigenetic silencing (i.e., paramutation) are now known, including those at the $r_1$ (Brink 1956; Ronchi et al. 1995; Kermicle 1996), $b_1$ (Coe 1966; Patterson et al. 1993), and $p_1$ (Hollick et al. 1995) loci of maize, the transgenic A1 locus in petunia (Meyer et al. 1993), and the paramutagenic 271 locus in tobacco (Park et al. 1996).

$r_1$ genes encode nearly identical myc-homologous, helix-loop-helix proteins (Ludwig et al. 1989; Perrot and Cone 1989; Consonni et al. 1992) that are capable of activating transcription from promoters of structural genes in the anthocyanin biosynthetic pathway (Ludwig et al. 1989; Goff et al. 1990). $B$ or $Lc$ proteins are sufficient to confer pigmentation to most maize tissues when they are expressed from a constitutive viral promoter (Ludwig et al. 1989; Goff et al. 1990), indicating that the pattern of anthocyanin deposition conferred by a given $r_1$ gene is due to the tissue-specific expression of $r_1$, not to functional differences in $R$ proteins. Because the $r_1$ locus often contains multiple $r_1$ genes, particular variants of the $r_1$ locus are referred to as distinct $r_1$ haplotypes with the designation, $R$-suffix, in which the suffix is based on a single striking phenotypic characteristic of the haplotype. For example, $R$-marbled ($R$-mb) confers a marbled pattern of pigmentation to the aleurone. A particular haplotype will comprise one to several individual $r_1$ genes. These genes are not all identical; many alleles exist. These alleles are distinguished from one another without use of the $R$ prefix. For example, the $R-r$ haplotype contains four $r_1$ genes with the distinct allele combination $P$ $q$ $S1$ $S2$.

In paramutation, a paramutable haplotype, e.g., $R-r$, is made heterozygous with one of the paramutagenic $r_1$ haplotypes. In the heterozygote, simple dominance of...
R-r is observed, producing dark, solid pigmentation. However, when the heterozygote is subsequently crossed, expression from R-r is changed and confers a paler, "mottled" pattern of anthocyanin deposition. The paramutant R-r haplotype has been partially silenced. Silencing of R-r in paramutation happens in virtually 100% of R-r haplotypes transmitted, but the level of silencing is quite variable, ranging from nearly complete silence to nearly complete expression. This range of phenotypes is illustrated in Figure 1. The silenced (paramutant) phenotype can be inherited in subsequent generations, but can also be partially or completely reversed, a condition referred to as metastability. The pattern of expression of the paramutagenic (silencing) haplotype is unchanged following the interaction.

Clues to why certain r1 haplotypes are prone to paramutation while other haplotypes are not can be obtained by examining the structure of the particular haplotypes involved. Four distinct genes are present in the paramutable R-r haplotype: Plant color (P), which colors the coleoptile, leaf tips, roots, and anthers; q, a nonfunctional gene fragment with strong sequence similarity to the promoter region of P; and two Seed color (S) genes, S1 and S2, which color the aleurone of the seed. The S genes are strongly subject to paramutation; the P gene is only very weakly susceptible (Brink and Mikula 1958; Brown 1966). The S1 and S2 genes are arranged close together in a head-to-head orientation and lack a typical r1 promoter. Instead, a 381-bp region called α constitutes the promoter for both of the S genes (Walker et al. 1995). The α region is derived from a fractured and rearranged transposable element named Doppia. Doppia sequences adjacent to S2 include 26 bp of the terminal inverted repeat and multiple copies of a 12-bp Doppia subterminal repeat element. Deletion of α severely compromises the ability of the R-r haplotype to undergo paramutation (Kermicle 1996; Walker 1998). Whether it is the loss of Doppia transposable element sequences or the lack of a functional promoter for the S genes that causes this effect is not known. However, it is clear that paramutability of R-r is not predominantly a function of the presence of multiple gene repeats or the presence of the inverted repeat that constitutes the S genes.

Structural analyses indicate that the paramutagenic R-st haplotype contains four r1 genes designated Sc (Seed color), Nc1, Nc2, and Nc3 (Near colorless; Eggleston et al. 1995). Because R-st consists of four tandemly repeated r1 gene copies, it is possible to generate deletion derivatives by unequal crossing over. Using this approach, Kermicle (Kermicle et al. 1995; Kermicle 1996) demonstrated that the ability of R-st to cause silencing of R-r is dependent on gene copy number. The R-st haplotype can be rated as strong (+++ in Kermicle's 1996 notation), while derivative haplotypes having only three r1 gene copies are moderate (++ in Kermicle's notation), derivative haplotypes with two r1 genes are weak (+ in Kermicle's notation), and those with only a single r1 gene are no longer paramutagenic. Addition of a displaced r1 gene copy (Lc) onto the same chromosome with R-st results in an extremely strong (++++ in Kermicle's notation) paramutagenic complex. Thus, paramutagenicity of R-st is incremental and appears to depend on the r1 gene copy number.

The correlation of paramutagenicity with gene number implies that copy number itself leads to paramutagenicity. However, when segmental deletions of R-st arise through unequal crossing over, both the r1 gene dosage and the dosage of the intergenic regions of the complex are altered. Thus, it is possible that loss of paramutagenicity is due to the loss of the portions of the complex that lie between the r1 genes. These intergenic regions contain Doppia elements that are located at the 5' ends of each of the Nc genes, which has caused some authors to question whether the Doppia elements pres-
ent at R-st have a role in causing paramutagenicity (Matzke et al. 1996).

Here we present studies aimed at elucidating the structure and paramutation properties of the R-mb haplotype, a paramutagenic r1 complex. We show that R-mb contains three genes and that derivatives of R-mb that have fewer r1 gene copies are less paramutagenic than the intact R-mb complex. Evidence is presented that the R-mb haplotype lacks Doppia sequences, hence paramutagenicity cannot be due solely to the presence of Doppia within the locus. We further report the identification of a novel transposable element at R-mb, Shooter (Sho), which is responsible for causing the marbled aleurone phenotype typical of the R-mb haplotype. We present evidence that the Sho transposable element does not play a significant role in paramutagenicity. The significance of the R-mb structure in r1 complex evolution and models for paramutagenicity are discussed.

**MATERIALS AND METHODS**

**Genetic stocks:** All stocks were maintained in the W22 inbred background. All stocks are homozygous dominant for the a1, a2, c1, c2, bz1, and bz2 genes necessary for anthocyanin synthesis in aleurone and homozygous recessive for the p1 and b1 genes (see Dooner et al. 1991 for descriptions). The R-mb haplotype confers strong pigmentation to the aleurone and embryo in a marbled pattern of dark blotches on a colorless background and also weakly pigments the coleoptile (Weyer's 1961). The R-r haplotype (R-r:standard) used in this study confers strong pigmentation of the aleurone of the seed, the coleoptile, the roots, and leaf tip of seedlings and to the roots and anthers of mature plants; it has been described previously (Stadler 1948; Walker et al. 1995). The colorless plant, colorless aleurone tester allele, R-g, is r-g:Sd described by Ker micle (1984). The haplotype designated here as R-nj was kindly provided by J. Ker micle. R-nj comprises a Nj gene interrupted by a Ds transposable element (n:j-m:Ds) and an S-subcomplex (consisting of q, S1 and S2) derived from R-r. Because the n:j-m:DS gene was derived (J. L. Ker micle, personal communication) by crossing over with the scm3 allele (All eman and Ker micle 1993), the Ss insertion in n:j-m:DS is likely to be the same as that in r-sc, i.e., a 2.1-kb Ds6 element in intron 5 of r1. The Δa allele confers a colorless (null) phenotype and does not hybridize to r1 probes on genomic blots.

**Paramutagenicity tests:** Each R-mb/rO derivative haplotype was tested for paramutagenicity as follows: R-mb/rO heterozygotes were crossed with homozygous R-r females to give ears that segregated R-r/rR-mb and R-r/r-rR kernels. From each ear, twenty kernels comprising both R-r/r-mb and R-r/r-r kernels were grown and crossed as mates to plants homozygous for the r-g tester allele. The level of paramutation of R-r in response to each R-mb haplotype was scored in the resulting progeny. The half of the parent plants that were of genotype r-g/R-r were used as a control for potential genetic background effects on R-r imprinting. No such background effects were actually observed (not shown), and thus all of these served as equivalent nonparamutagenic controls. The ears from control parental plants were distinguished from those from R-r/R-mb parental plants either by visual inspection, in cases in which the Nj phenotype was clearly visible, or by PCR analysis of DNA from 2–3 colorless kernels on each ear for R-mb haplotypes containing Ds at Nj, which, in the absence of Ac, are indistinguishable phenotypically from r-g. Ears bearing colorless kernels that showed PCR fragments typical of Nj were used for measurement of paramutagenicity, ears with colorless kernels that gave PCR fragments typical of r-g were designated controls, and ears that were ambiguous were discarded.

Kernels were stripped from both control and experimental ears, and all r-g/ R-r kernels, distinguishable by their mottled phenotype, were sorted from their r-g/ R-mb or r-g/ r-g sibs. A random sample of 50 r-g/ R-r kernels was then separated and double-blinded prior to scoring by a naive observer. Scoring was performed by comparing each kernel to a seven-kernel set of standards in which 7 is full color and 1 is colorless. Weighted average scores for each ear were produced and these weighted averages were used for statistical analyses.

**Construction and screening of genomic libraries:** A primary library of 5 × 10^6 pfu was constructed from genomic DNA of plants homozygous for the R-mb complex. Genomic DNA was partially digested with restriction enzyme Sau3A1 and was cloned into the XhoI site of pJGEM11 (Promega, Madison, WI) using the partial fill-in method according to the manufacturer’s instructions. This library was screened with a mixture of probes that hybridize to promoter, coding, and 3′ noncoding portions of typical r1 genes. The probes used were as follows: pR-nj1 (Robbins et al. 1991), which detects the promoter and/or first exon of r1; SAH (Walker et al. 1995), which detects the 5′ portion of the second intron of r1; cDNA-B, a probe derived from position 812–1334 in the 5n cDNA (Consonni et al. 1992), which detects exons 4–7 of r1; and P1.8E, a probe from the 3′ noncoding portion of the P gene of the R-r complex, which hybridizes distal to some r1 genes (Robbins et al. 1991). Clones isolated from this library bear the prefix λM.

**Inverse PCR:** DNA prepared from the leaves of plants with the desired genotype was digested to completion with an appropriate restriction enzyme, and then fractionated on a 1% agarose gel. The band corresponding to the desired molecular weight was excised from the gel and purified using Qiagen gel extraction (QIAKEN, Valencia, CA) procedure. Self-ligation of the resulting DNA was performed overnight using a DNA concentration of 2 ng/μl. Self-ligated DNA (100 ng) was used as template in a 50-μl reaction with Expand DNA polymerase (Boehringer Mannheim, Indianapolis) according to the manufacturer’s instructions. The primer pairs used for I-PCR were as follows:

Scm with Sho excision: oR42 (5′ AGGAGAAGAATAGAGAG GAACGAAG 3′) and oR60 (5′ GAGGCCCATCCAGATAA CATAACTGCTT 3′)

Sho 3′ end: oScm3 (5′ TGCATGTACATGGAATCTTGG GAGGA 3′) and oScm1 (5′ TGCAAGATTTCTCTCTT CCACCTCA 3′)

Sho 5′ end: oShoot1 (5′ GGAAGCACCTCTGTGTCTGG ATGAGGG 3′) and oR59 (5′ TTGTGACAAATACG GATTCTG 3′) with nested primer oR60 (5′ GAGGCCCAT CCAGATAACATAAGCTT 3′) for the second round of amplification.

**PCR:** All PCR reactions were performed using Expand polymerase (Boehringer Mannheim) according to the manufacturer’s instructions. The primers used for PCR were as follows:

oR67 (5′ CCTTGTGTTGCCAATCTCCCCC 3′)
oRrev (5′ CATGCAGGATTAGAGAAAGCG 3′)
oLC3676 (5′ GGTGAGGCCCATCCAGATAACATAAGC 3′)
oLC6296 (5′ TTGTAGACCGCTGGCCAGCTC 3′)
oScm1 (5′ TGCAAGATTTCTCTCTCCACCTCA 3′)
oShoot20 (5′ TGGTTGCGCAGCATGTATTGTTTGGC 3′)

**Preparation of genomic DNA and Southern blotting:** DNA from aleurone sectors was prepared using the microprep pro-
The genes of the *R-mb* haplotype: Twenty-three clones that hybridized to *r1* probes were isolated from a genomic library made using homozygous *R-mb* genomic DNA. The restriction maps of the clones were determined, allowing placement of each clone into one of three groups representing the three genes of the *R-mb* complex. The restriction map of each gene is shown in Figure 2. By comparison with the maps of the genes from the *R-r* complex and with the maps for the displaced *r1* gene, *Lc*, and based on hybridization patterns for each clone (not shown), the predicted exon/intron structure for each gene from the *R-mb* complex has been superimposed on the maps. Together, the *R-mb* clones account for all but one of the fragments observed using *r1* gene probes on genomic blots of *R-mb* DNA. The fragment that is not accounted for, an 8-kb *HindIII* fragment containing the 5' end of an *r1* gene, is partially represented by clone λMB4. The remainder of this fragment was cloned using inverse PCR (I-PCR; see below for details). The positions of each λ clone and I-PCR product used in this study are shown in Figure 2.

Sequence information: Because the pattern of expression of a particular *r1* gene is determined by its upstream portions (Kermicle et al. 1988; Robbins et al. 1991; Eggleston et al. 1995), the DNA sequence of this part of each gene was determined. Two of the genes have sequences that are very similar to the corresponding portions of the displaced *r1* gene, *Lc* (S. Ludwig, L. Habera and S. Wessler, personal communication), and slightly less similar to the P gene of *R-r* (Walker et al. 1995). On the basis of both this sequence similarity and the similarity of the restriction maps of these genes and the *Lc* gene, these two genes were named *Lcm1* and *Lcm2* (*Lc* similarity at Marbled). It is important to note, however, that the *R-mb* haplotype normally confers neither the leaf color typical for *Lc* nor the root, coleoptile, and anther color typical of *P*. The overall sequence similarity of the 2166-bp *Lcm1* and 2170-bp *Lcm2* sequenced regions is 95%. The sequences of the upstream portions of *Lcm1* and *Lcm2* are available in GenBank (*Lcm1*: accession no. AF135456; *Lcm2*: accession no. AF135457).

The **Shooter transposable element**: The third gene of the *R-mb* complex, named *Scm* (after Styles 1993), is 100% similar to the *Lcm2* gene from position +403.
Figure 3.—Comparison of Lcm1, Lcm2, and Scm upstream regions. (A) Schematic representation of 5' ends of the genes of R-mb. Similar sequences are shown in gray. The portion of Scm that differs from Lcm1 and Lcm2 is indicated in black. The position of insertion and orientation of Sho is indicated below Scm. (B) Sequence of the Sho insertion site in Scm. Underlined sequences indicate the target site duplication in the Scm gene. (C) Schematic representation of the cloned portion of the Shooter transposable element. A 700-amino acid open reading frame bearing strong amino acid sequence similarity to TnpD of the related En/Spm element (Masson et al. 1991) is indicated. The positions of multiple copies on both strands of a 9-bp motif (5'-TNTBGVCAC-3') located in Sho subterminal regions are indicated by black arrowheads above and below. Several large direct repeats in the Sho ends are indicated with triangles and arrows.

(R-mb) genetic stocks frequently give rise to “self-color” (R-scm) derivatives that confer solid pigmentation to the aleurone instead of the typical marbled pattern of large blocks of pigment on a colorless background (Styles et al. 1993). These R-scm derivatives have been presumed to arise by germinal excision of the Sho element from the3A. This element, which is named Shooter (Sho) for its role in causing the marbled phenotype, is novel. Because only 1130 bp of the left end of the Sho element was present on the Scm genomic clone, λMB4 (see Figure 2), we used a PCR-based approach to clone additional Sho sequences.

Genomic mapping indicated that an 8-kb HindIII fragment contained the 5' end of Scm. To obtain the missing portion of this 8-kb HindIII fragment, we used l-PCR. DNA from R-mb homozygotes was digested with HindIII and fractionated on an agarose gel. The 8-kb fraction was purified from the gel, self-ligated, and used as the template for PCR amplification using primers derived from r1 (oR59 and nested primer oR60) and from Sho (oShoot1). Sequence analysis of the resulting 4.1-kb product showed that no Sho left end was present, indicating that Sho is more than 4 kb in length, and that our usual r1 probes (listed in materials and methods) were not detecting R-mb fragments corresponding to the portion of Scm upstream of the Sho insertion.

(Aval) to position –395 relative to the typical r1 start of transcription (Tonelli et al. 1991). At this position, the sequence similarity abruptly ends due to insertion of a transposable element of the CACTA family (Upadhyaya et al. 1985), illustrated schematically in Figure 3A. This element, which is named Shooter (Sho) for its role in causing the marbled phenotype, is novel. Because only 1130 bp of the left end of the Sho element was present on the Scm genomic clone, λMB4 (see Figure 2), we used a PCR-based approach to clone additional Sho sequences.
An Scm-specific probe, Scm5', was made from the Scm 1-PCR product described above and used to probe genomic blots of R-mb. Sho is estimated to be at least 10 kb in length based on genomic blots of R-mb and its R-scm derivatives using this probe. Genomic blotting using the Scm5' probe also indicated that a 2.6-kbp EcoRI fragment in R-mb DNA should contain both the Scm upstream region and the left end of the Sho transposon. We made use of this fragment in a third 1-PCR using the primers oScm3 and oScm1, which amplified a 2.4-kbp fragment. Sequencing of this fragment indicated that it contained the left end of Scm and the 3' end of Sho. The entire Sho element was not cloned for this analysis.

The overall structure of the cloned portions of Sho is illustrated in Figure 3, B and C. The Sho element has a 15-bp terminal inverted repeats and created a 3-bp target site duplication (TSD) as is typical of CACTA family elements. The nucleotide sequence of Sho is not detectably similar to either En/Spm or Doppia, two other CACTA family elements found in maize. However, a 799-aa open reading frame in the 5' end of Sho bears strong (47% identity, 65% similarity) amino acid sequence similarity to the TNP2 protein of the Antirrhinum transposable element, Tam1 (Nacken et al. 1991), and the TnpD protein of the maize transposable element, En/Spm (Masson et al. 1991; 34% identity, 54% similarity). The cloned portion of the right end of Sho did not contain any apparent open reading frames. The Scm and Sho 5' and 3' end sequences are available in GenBank under the following accession nos.: Scm, AF135485; Sho 5', AF136220; Sho 3', AF136221.

Within 800 bp of the element termini, both ends of Sho contain multiple copies on both strands of a 9-bp motif ('5' TNTBGCVAC 3') that are indicated by black arrowheads above and below in Figure 3C. The arrangement of these motifs is similar to that of the subterminal repeat elements of En/Spm that are binding sites for the element-encoded protein, TNPA. Sho also contains several large direct repeats in its ends. The Sho 5' end contains two complete and one partial copy of a 189-bp sequence, three complete and one partial copy of a 61-bp sequence, two copies of a 60-bp sequence, and two copies of a 58-bp sequence. The 3' end of Sho contains a single 96-bp direct repeat. The positions of these motifs are illustrated in Figure 3C.

**Sho excision:** Fifteen R-scm derivatives, including the R-scm1, R-scm2, R-scm3, and R-scm4 derivatives described by Styles (1993), plus 11 new R-scm derivatives that arose in our stocks were tested for Sho excision. PCR primers, oScm1 and oRev, shown in Figure 4A, which flank the Sho insertion site and are specific to the Scm gene, were used to amplify a product from each R-scm derivative. For every R-scm derivative, a PCR fragment of the expected molecular weight was generated. These products were sequenced to determine whether an excision footprint was present. The results are shown in Figure 4B. Footprints found in these derivatives fell into four classes: perfect excisions (R-scm5 and R-scm6), in which the TSD is lost and the presumed original Scm sequence is precisely restored, and three types of imperfect excision. Imperfect excision footprints included a two-nucleotide deletion (R-scm4 haplotype), a two-nucleotide insertion that is part of the TSD (R-scm1, R-scm3, R-scm7, R-scm8, R-scm9, and R-scm10), and a four-nucleotide insertion consisting of the TSD plus a single A residue (R-scm2, R-scm11, R-scm12, R-scm13, R-scm14, and R-scm15).

The Sho transposable element found in the upstream portion of the Scm gene is also likely to be responsible for the production of the 5' terminus of the element-encoded protein, TNPA.
for causing the aleurone pattern consisting of large pigmented blocks in a colorless background typical of R-mb. This pattern would be generated due to excisions of Sho occurring early in aleurone development. To test this, DNA was prepared from colored and colorless sectors from aleurones of R-mb homozygotes and used in PCR assays as described above. To test for Sho excision from Scm, the primer pair oScm1 and oRrev was used. This primer pair will produce a product only if the large Sho element has excised from Scm. The results are shown in Figure 4C. With primers oScm1 and oRrev, no fragment was amplified from any colorless sector, while the colored sectors invariably contained an excision product. As a control for integrity of the genomic DNA template, a second set of primers (oScm1 and oShoot20) were used to amplify a Scm/Sho junction fragment. These primers produce a product that includes a portion of Scm upstream of Sho plus part of the Sho 3' end. A fragment was amplified from all samples as expected; colored sectors are heterozygous for one Scm excision allele and an unexcised copy of Scm-m::Sho.

**Generation of R-mb crossover derivatives:** To establish the placement of the three r1 genes within the R-mb complex, we screened for haplotypes that arose by crossing over between genes in the R-mb complex and the Navajo (Nj) gene of a complex called R-njs. The R-njs complex is very similar in structural organization to the R-r complex (Robbins et al. 1991; Walker et al. 1995), except that a Nj allele of r1 replaces the P allele found in R-r. Nj alleles of r1 confer color in the aleurone, but only in the "crown" or top of the kernel. This phenotype is easily distinguished from the marbled pattern conferred by Scm and from full color pattern conferred by the S genes. The Nj gene in R-njs is interrupted by a Ds transposon so that it gives color in the crown of the kernel in a sectored pattern in the presence of Ac and does not confer aleurone color when Ac is not present. Thus this allele is properly designated nj-m::Ds. The phenotypic effect of the nj-m::Ds gene is, however, masked by the expression of the S genes in the R-njs complex, which appears as simply full color due to the expression of the S1 and S2 genes throughout the aleurone.

The scheme for the generation of crossover derivatives is shown in Figure 5A. When a crossover event occurs between the nj-m::Ds gene and any one of the genes of R-mb, the Nj promoter drives expression of the r1 coding sequence distal to the site of recombination and the S-subcomplex is replaced by the portion of R-mb distal to the site of the crossover. Thus, crossover events are readily detected as navajo or navajo mutable kernels on ears segregating mostly marbled (intact R-mb complex) and full color (intact R-njs complex) kernels. It should be noted that this scheme will succeed only if Scm is the most proximal member of the complex. This is likely, since if Scm were at any other position, R-mb (or R-scm derivatives from it) would be expected to give rise to crossover derivatives in which the Scm gene was lost, much as the R-r complex gives rise to derivatives that lose the S-subcomplex (Robbins et al. 1991). Such crossover derivatives are not apparent in our stocks (E. L. Walker, unpublished observations).

Plants that were heterozygous R-mb/R-njs at the r1 locus and homozygous P-WR at the p1 locus were test-crossed as females using pollen from homozygous r-g P-VV plants. The P-VV haplotype of p1 contains an Ac element that is used in this experiment to activate transposition of the Ds element at the nj-m::Ds gene of the R-njs complex. The resulting ears were screened for
To determine the identity of each 

to genomic blots with DNA from seven representa-
Li 5

is navajo, the promoter portion of the most proximal a single 

mobilized derivative haplotype, the order of 

or crossed to plants homozygous for \( r^A \) or for \( R-r \). Thirteen of these \( R-nmb \) derivative haplotypes were propagated successfully and used for further analysis.

**Physical organization of the complex:** Each \( R-nmb \) derivative haplotype generated is expected to contain either one, two, or three \( r^1 \) gene copies, depending on which gene of the \( R-nmb \) complex recombined with \( Nj \).

al. By determining which genes are retained and which are lost in each \( R-nmb \) derivative haplotype, the order of the three genes within the \( R-mb \) complex can be determined, and simultaneously, the number of genes present in each of the \( R-nmb \) derivative haplotypes is discovered. Because the phenotype of each \( R-nmb \) derivative is navajo, the promoter portion of the most proximal (left-most) gene of the complex is expected to be derived from the \( Nj \) gene. Furthermore, the \( Scm \) 5' portion is expected to be lost in all \( R-mb \) derivatives. We tested this on genomic blots with DNA from seven representati
tive derivatives plus the two parental haplotypes (\( R-mb \) and \( R-njs \)). Figure 5B shows that the \( q \), \( S1 \), and \( S2 \) alleles, represented by 3.5-kb and 5-kb HinIII fragments, respectively, that are present in the \( R-njs \) progenitor haplo-
type are lost in all \( R-nmb \) derivatives, as expected based on their phenotype. The \( Scm \) and \( Nj \) fragments that hybridize to the \( pR-nj1 \) probe are very similar in size at \(~ 8 \) kb and are not readily distinguished with the \( pR-nj1 \) probe. The \( Scm5^\prime \) probe, however, detects a 4-kb HinIII fragment containing \( Scm \) upstream portions and \( Sho \) right end, as shown in Figure 5B. This fragment is clearly absent from the \( R-mb \) derivatives, as expected from their phenotypes. This probe also detects, less strongly, a fragment corresponding to the \( q \) gene from the \( S-subcomplex \) in \( R-njs \), which is also missing in all \( R-nmb \) derivatives. Together, these data indicate that \( R-nmb \) derivatives contain the \( Nj \) upstream fragment and lack the \( Scm \) upstream fragment. In addition, the \( q \), \( S1 \), and \( S2 \) alleles present in the \( R-njs \) progenitor allele are lost in all \( R-nmb \) derivatives. The \( pR-nj1 \) probe detects both the \( Lc1 \) and \( Lc2 \) genes of \( R-mb \), which migrate as a doublet of 3.9 kb as seen in Figure 5B. One \( R-nmb \) derivative, \( R-nmb8 \), is missing both of these fragments. Derivatives \( R-nmb4 \) and \( R-nmb7 \) appear to have only a single \( Lc \) gene while derivatives \( R-nmb1 \), \( R-nmb2 \), \( R-nmb3 \), and \( R-nmb6 \) appear to retain both \( Lc \) genes, based on the intensity of hybridization with the \( pR-nj \) probe.

To determine the identity of each \( Lc \) gene in the

### TABLE 1

<table>
<thead>
<tr>
<th>Number of genes</th>
<th>Phenotype</th>
<th>Mean aleurone score</th>
<th>Count (number of ears scored)</th>
<th>Minimum individual ear aleurone score</th>
<th>Maximum individual ear aleurone score</th>
</tr>
</thead>
<tbody>
<tr>
<td>r-g</td>
<td>1</td>
<td>colorless</td>
<td>6.245</td>
<td>19</td>
<td>5.880</td>
</tr>
<tr>
<td>R-nmb1</td>
<td>3</td>
<td>navajo</td>
<td>4.226</td>
<td>11</td>
<td>3.240</td>
</tr>
<tr>
<td>R-nmb2</td>
<td>3</td>
<td>navajo</td>
<td>5.167</td>
<td>4</td>
<td>4.792</td>
</tr>
<tr>
<td>R-nmb3</td>
<td>3</td>
<td>navajo mutable</td>
<td>NS</td>
<td>2</td>
<td>4.080</td>
</tr>
<tr>
<td>R-nmb4</td>
<td>2</td>
<td>navajo</td>
<td>5.410</td>
<td>6</td>
<td>6.040</td>
</tr>
<tr>
<td>R-nmb5</td>
<td>3</td>
<td>navajo mutable</td>
<td>4.065</td>
<td>9</td>
<td>3.480</td>
</tr>
<tr>
<td>R-nmb6</td>
<td>3</td>
<td>navajo mutable</td>
<td>4.985</td>
<td>12</td>
<td>4.340</td>
</tr>
<tr>
<td>R-nmb7</td>
<td>2</td>
<td>navajo mutable</td>
<td>5.966</td>
<td>10</td>
<td>5.640</td>
</tr>
<tr>
<td>R-nmb8</td>
<td>1</td>
<td>navajo mutable</td>
<td>6.655</td>
<td>5</td>
<td>6.236</td>
</tr>
<tr>
<td>R-nmb9</td>
<td>2</td>
<td>navajo mutable</td>
<td>6.142</td>
<td>12</td>
<td>5.561</td>
</tr>
<tr>
<td>R-nmb10</td>
<td>3</td>
<td>navajo mutable</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
</tr>
<tr>
<td>R-nmb11</td>
<td>3</td>
<td>navajo</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
</tr>
<tr>
<td>R-nmb12</td>
<td>3</td>
<td>navajo</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
</tr>
<tr>
<td>R-nmb13</td>
<td>3</td>
<td>navajo</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
</tr>
<tr>
<td>R-mb</td>
<td>3</td>
<td>marbled</td>
<td>4.275</td>
<td>6</td>
<td>3.571</td>
</tr>
</tbody>
</table>

Gene number refers to the number of \( r^1 \) gene copies present within each haplotype. Phenotype of \( R-nmb \) derivative haplotypes can be navajo if no \( Ds \) element is present in the \( Nj \) gene, or navajo mutable if \( Ds \) is present. Mean, minimum, and maximum aleurone scores refer to the ear weighted average score for \( R-r \) transmittted following one generation of heterozygosity with the haplotype indicated. Number of ears refers to the number of individual ears that were scored for paramutation. NS, not significant; an insufficient number of ears were obtained for scoring. ND, no data.
R-nmb derivatives, we tested for restriction site polymorphisms in PCR-generated fragments. To do this, we designed a primer pair that specifically amplifies the Lcm1 and Lcm2 upstream portions, but does not amplify Nj. The Lcm1 and Lcm2 genes are then distinguished from each other by virtue of polymorphisms at four restriction sites, as shown in Figure 6A. PCR reactions using DNA from each R-nmb derivative as template were performed, and the resulting fragments were digested with NdeI, XmnI, and BslI. The results of this experiment are shown in Figure 6B. Derivative R-nmb1, R-nmb2, R-nmb3, R-nmb5, R-nmb6, R-nmb10, R-nmb11, R-nmb12, and R-nmb13 contain both Lcm1 and Lcm2 (as well as Nj), and thus have the same number of genes as the R-mb complex but with a Nj promoter substituted for the Scm promoter. Derivatives R-nmb4, R-nmb7, and R-nmb9 contain only two genes, Nj and Lcm2. These derivatives formed by crossing over between Lcm1 and Nj to result in the loss of the Scm gene. These derivatives also establish the gene order for R-mb as Scm-Lcm1-Lcm2 in order from proximal to distal. Only one derivative was identified that had a single gene. This derivative, R-nmb8, arose from crossing over between Nj and Lcm2.

To confirm that the lack of amplification from R-nmb8 did not result from a problem with the R-nmb8 template, we amplified a fragment spanning from just upstream of the HindIII site in the large r1 second intron through the r1 third exon using primer pair oLc3676, and oLc6296. Fragments of 2.6 kb were amplified from the Lcm1 clone λMB2 and the Lcm2 clone λMB23. A
smaller fragment of 2.4 kb was generated from genomic DNA of each R-nmb derivative, indicating that the Nj gene has a deletion in this region relative to the Lcm genes. Both the 2.4- and 2.6-kb products were amplified from each two-gene R-nmb derivative. Amplification from the single gene R-nmb8 derivative produced only the smaller, Nj-derived product, demonstrating the integrity of the R-nmb8 template and confirming the presence of a single gene in this derivative. Furthermore, the presence of the smaller Nj-derived fragment in all two- and one-gene derivatives (Figure 6C) indicates that the position of crossing over in each occurred distal to this size polymorphism in Lcm1 (for derivatives R-nmb4, R-nmb7, and R-nmb9) or Lcm2 (for derivative R-nmb8).

Paramutagenicity of R-nmb derivative haplotypes: Previous studies of the paramutagenic R-st haplotype indicated that paramutagenicity was dependent on the gene copy number within the haplotype (Kermicle et al. 1995). To test whether this correlation holds for the R-mb haplotype as well, nine of the R-nmb derivatives described above were tested for paramutagenicity. Each R-nmb derivative was made heterozygous with the paramutable haplotype R-r by crossing R-nmb/r-g heterozygotes with homozygous R-r females. The resulting R-nmb/R-r and r-g/R-r heterozygotes were then crossed to a recessive tester strain homozygous for r-g. The progeny of the r-g/R-r heterozygous parents serve as nonparamutagenic controls for any possible genetic background effects that might arise in these experiments. As a control for the normal level of paramutagenicity of R-mb, R-mb/R-r heterozygotes were similarly tested.

The average aleurone color score was determined (by a naive observer following double-blinding of each sample) for each ear, and these scores were then used in the statistical analyses below. The average aleurone score for each derivative tested is given in Table 1. The aleurone scores were then grouped according to the number of genes in the haplotype as represented graphically in Figure 7. An unpaired t-test indicates that the three-gene derivatives are not different in paramutagenic strength from the R-mb progenitor haplotype (P = 0.35). The single one-gene derivative that we obtained was nonparamutagenic, giving average aleurone scores that were even higher than the r-g, nonparamutagenic, controls. The behavior of the two-gene R-nmb derivatives was more complex: they were obviously less paramutagenic than three-gene haplotypes (P < 0.0001), but were not significantly different from r-g controls (P = 0.20) by nonparametric Mann-Whitney U-test. They were, however, significantly different (P = 0.01) from the one-gene R-nmb derivative that, as noted above, gave aleurone scores that were even higher than those from the r-g controls. In terms of qualitative scores, R-r segregating from two-gene R-nmb derivatives conferred lighter colored aleurone (average score of 6.136 ± 0.057) than R-r segregating from either the one-gene R-nmb derivative (average score of 6.700 ± 0.123) or from r-g (average score of 6.245 ± 0.071), as shown in Table 1 and Figure 7. Taking all the data into account, we conclude that there is a gradient of paramutagenicity that is directly proportional to gene copy number and that the single gene R-nmb haplotype is not paramutagenic.

Absence of Doppia sequences within R-mb: Matzke et al. (1996) have questioned whether the gene copy number effect in paramutagenicity of R-st could be the result of incremental loss of the Doppia elements found in R-st intergenic regions. Sequencing of the relevant portions of Scm, Lcm1, and Lcm2 genes of R-mb indicated that Doppia sequences are not found at corresponding positions in any gene of the R-mb haplotype. To determine whether Doppia sequences exist elsewhere within the R-mb complex, each R-mb clone was hybridized to probes from the Doppia-containing regions of the S (PCR-A) and q (Doppia-L) genes of R-r. No hybridization was found (not shown). The absence of Doppia elements within the cloned portions of R-mb does not rule out the presence of Doppia within the complex, since Doppia could be present in intergenic regions not represented in the set of clones. Genomic blot analysis was used to determine whether any Doppia-containing fragments associated with R-mb could be detected. Genomic DNA from R-r and R-mb was analyzed using Doppia probes from R-r (PCR-A and Doppia-L). To determine whether any Doppia-hybridizing fragment cosegregated with R-mb, 14 R-r/rΔ and 14 R-mb/rΔ kernels from an ear resulting from the cross R-r/R-mb × rΔ were analyzed. No Doppia fragments that cosegregated with R-mb were identified (not shown).
DISCUSSION

The R-mb haplotype comprises three r1 genes. Because crossover events were observed between Nj and each of the three genes of R-mb, the R-mb genes are arranged as direct repeats. The most proximal gene of the complex, Scm, is responsible for the marbled aleurone pattern characteristic of the haplotype. The colored aleurone sectors arise when a transposable element, Sho, excises from its position at -700 in the Scm promoter, thus restoring Scm gene function. Sho also transposes generically to give the stable self-color (R-scm) derivative haplotypes that frequently arise in R-mb stocks (Weyers et al. 1961; Styles 1993). That the Sc gene of R-mb also colors the embryo is indicated by occasional colored embryo sectors visible in R-mb seeds and by the fully colored embryos conditioned by R-scm derivatives. It is interesting that R-scm derivatives also condition anthocyanin pigmentation of seedling leaves and of tassel branches (Styles 1993). The other two genes of R-mb, Lcm1 and Lcm2, do not appear to be expressed, based on the lack of pigmentation in any tissues other than aleurone and scutellum in R-mb plants. In combination with particular alleles of the anthocyanin regulatory genes P11 and a3, the pattern of anthocyanin expression conferred by R-mb and its R-scm derivatives is extended to include weak pigmentation of the anthers (Styles 1993), possibly indicating that Lcm genes can be active in certain genetic backgrounds. However, it is also possible that anther expression in a3 and P11 plants reflects activity of the Scm promoter that is not affected by the presence of the Sho element. The leaf color conferred by R-scm derivatives is increased to include all leaves of the plant in a3 or P11 backgrounds (Styles 1993).

The structure of R-mb confirms certain aspects of r1 complex formation during evolution. Both of the other well-characterized complex r1 haplotypes, R-r and R-st, contain a Doppia transposable element in the distal genes of the complex (Walker et al. 1995; Matzke et al. 1996). On the basis of the lack of a Doppia excision footprint within the proximal P gene of R-r, we previously proposed that r1 gene duplication event occurred prior to the arrival of Doppia during evolution of the R-r complex (Walker et al. 1995). The finding that no Doppia elements are present in any of the genes of R-mb strengthens the argument that r1 gene duplication occurred before Doppia inserted. It is, of course, possible that Doppia visited the progenitor allele and subsequently excised without leaving a footprint. The lack of Doppia at R-mb as well as the lack of a Doppia footprint in either of the Lcm genes also suggests that R-mb is less closely related to R-r and R-st than they are to each other.

The Sho transposable element found in Scm is a novel member of the CACTA superfamily of elements. Sho represents the third CACTA superfamily member reported in maize; En/Spm and Doppia are the other two. The Sho element is very likely to be autonomous given that the marbled phenotype of R-mb is observed in all genetic backgrounds tested. The Sho element is unlikely to be involved in paramutagenicity of the R-mb haplotype, since when a Nj allele that lacks Sho is substituted for Scm in the R-mb derivatives presented here, the same level of paramutagenicity is retained.

The R-mb derivatives also make it possible to address whether the Scm promoter per se is involved in paramutagenicity of R-mb. We find that substitution of Nj for Scm has no effect on paramutagenicity. This finding differs from the results reported by Kermicle (Satyanarayana 1970; Williams 1972; cited by Kermicle 1996) who found that substituting a Nj or P promoter for the Sc promoter of R-st did diminish paramutagenic strength. The difference may be attributable to differences in the Sc promoter of R-st and the Scm promoter of R-mb. Although the aleurone phenotypes conferred by Sc and Scm are indistinguishable, the Scm gene confers stronger color to the scutellum. This difference in expression may reflect sequence differences in the two promoters that could also affect their behavior in paramutation.

The R-mb derivatives that we have generated in which one or more genes of the R-mb complex are deleted were also tested for paramutagenicity. Consistent with previous findings for the R-st haplotype (Kermicle et al. 1995), gene copy number strongly affects paramutagenic strength of R-mb. Derivatives with a single gene copy are neutral with respect to paramutation. A model that has been suggested as an explanation for the gene copy number effect in R-st paramutagenicity is that loss of the Doppia sequences within the promoters of its distal genes contributes to paramutagenicity (Matzke et al. 1996). Because no Doppia sequences are found in the genetic portions of the R-mb complex, yet gene copy number in this haplotype has a similar effect on paramutagenicity, models that rely exclusively on the presence of Doppia at the paramutagenic haplotype are unlikely to be correct.

Models for paramutagenicity must focus on gene copy number effects. Paramutagenicity is very similar to the phenomenon dubbed RIGS (repeat induced gene silencing; Assad et al. 1993). In the RIGS system a single transgenic construct was made that contained three genes with each driven by the 3SS promoter. Crossover/gene conversion derivatives that differed in gene/promoter number were then derived from a line having a single T-DNA insert. In this system, gene copy number strongly influences the epigenetic state of genes within the complex, with multigene haplotypes being more strongly silenced than single-gene haplotypes. Furthermore, single-gene or two-gene haplotypes that were heterozygous with three-gene haplotypes were found to be more strongly silenced than they were when homozygous, a situation that is strikingly similar to the gene copy number effect on paramutagenicity.

Recent characterization (Jakowitsch et al. 1999) of
the paramutagenic H2 transgene locus revealed that, like the paramutagenic r1 haplotypes, it contains multiple copies (four complete and two partial) of the Nos promoter sequence (NosP) that appears to direct homologous gene silencing. Cloned portions of H2, each containing only three NosP sequences were found to be no more paramutagenic than a transgene carrying two NosP sequences when reintroduced into plants. Thus, in this system, too, paramutagenicity may be dependent on the presence of multiple homologous gene or promoter copies.

One of the more intriguing aspects of the r1 paramutation system is that not all genes are equally subject to paramutation, and gene copy number does not automatically lead to silencing activity: the paramutable R-r haplotype contains three highly similar functional gene copies, yet none of these copies is ordinarily strongly silenced, even in R-r homozygotes, which have a total of six r1 gene copies in the genome. In combination with a paramutagenic haplotype, certain genes of R-r are strongly silenced (the S genes), while another gene, P, is only weakly affected (Brin k and Mikula 1958; Brown 1966). Paramutagenicity itself depends on gene copy number, but even so, there is some indication that not every gene is equally able to contribute to paramutagen- icty since the Sc gene of the R-r haplotype is stronger in its ability to contribute to paramutagenicity than are two other r1 genes that have been used to replace it (Kermicle 1996). Complete understanding of the epigenetic phenomenon of paramutation will require a more detailed understanding of the specific nature of the individual genes that participate and the precise nature of the interaction of the paramutable and paramutagenic haplotypes.

We are most grateful to Jerry Kermicle, who provided us with the R-nJs stock used for this work. We thank Ronald Beckwith for his expert assistance in the greenhouse and Alison Delong and Randy Phillis for their comments on this manuscript. This work was supported by the National Science Foundation (NSF9796088).

LITERATURE CITED


Brown, D. F., 1966 Paramutability of R-g and r-mutant genes derived from an R allele in maize. Genetics 54: 899-910.


Park, Y. D., L. Papp, E. A. Moscone, V. A. Iglesias, H. Vauchert et


Williams, W. M., 1972 Variability of the R-stippled gene in maize, Ph.D Thesis, University of Wisconsin, Madison, WI.

Communicating editor: J. A. Birchler