B-Bolivia, an Allele of the Maize *b1* Gene with Variable Expression, Contains a High Copy Retrotransposon-Related Sequence Immediately Upstream[1]

David A. Selinger and Vicki L. Chandler*

Department of Plant Sciences, University of Arizona, Tucson, Arizona 85721

The maize (*Zea mays*) *b1* gene encodes a transcription factor that regulates the anthocyanin pigment pathway. Of the *b1* alleles with distinct tissue-specific expression, *B-Peru* and *B-Bolivia* are the only alleles that confer seed pigmentation. *B-Bolivia* produces variable and weaker seed expression but darker, more regular plant expression relative to *B-Peru*. Our experiments demonstrated that *B-Bolivia* is not expressed in the seed when transmitted through the male. When transmitted through the female the proportion of kernels pigmented and the intensity of pigment varied. Molecular characterization of *B-Bolivia* demonstrated that it shares the first 530 bp of the upstream region with *B-Peru*, a region sufficient for seed expression. Immediately upstream of 530 bp, *B-Bolivia* is completely divergent from *B-Peru*. These sequences share sequence similarity to retrotransposons. Transient expression assays of various promoter constructs identified a 33-bp region in *B-Bolivia* that can account for the reduced aleurone pigment amounts (40%) observed with *B-Bolivia* relative to *B-Peru*. Transgenic plants carrying the *B-Bolivia* promoter proximal region produced pigmented seeds. Similar to native *B-Bolivia*, some transgene loci are variably expressed in seeds. In contrast to native *B-Bolivia*, the transgene loci are expressed in seeds when transmitted through both the male and female. Some transgenic lines produced pigment in vegetative tissues, but the tissue-specificity was different from *B-Bolivia*, suggesting the introduced sequences do not contain the *B-Bolivia* plant-specific regulatory sequences. We hypothesize that the chromatin context of the *B-Bolivia* allele controls its epigenetic seed expression properties, which could be influenced by the adjacent highly repeated retrotransposon sequence.

The alleles of the *b1* locus of maize (*Zea mays*) display a high degree of phenotypic diversity in terms of tissue- and developmental stage-specific expression (Styles et al., 1973; Coe, 1979; Selinger and Chandler, 1999). Studies on several alleles have served as a useful system to investigate how major changes in tissue-specific gene expression occurred (Radicella et al., 1992; Selinger et al., 1998; Selinger and Chandler, 1999). The *b1* locus encodes a transcription factor that regulates anthocyanin pigment expression, which provides an excellent visual marker for gene expression. The *B-I* and *B-Peru* alleles represent the extremes of the phenotypic diversity of *b1* alleles. *B-I* is highly expressed in most of the vegetative tissues of the plant but is not expressed in the embryo or aleurone tissues of the seed. In contrast, *B-Peru* is weakly and variably expressed in vegetative tissues of the plant, but is highly expressed in part of the embryo and in the aleurone layer of the seed (for a detailed description of these two alleles, see Radicella et al., 1992). The *B-Bolivia* allele has an intermediate phenotype between these two alleles. *B-Bolivia*, like *B-Peru,* pigments the aleurone layer of the seed and these are the only known *b1* alleles that confer aleurone-specific pigmentation. However, the consistency of pigmentation in the aleurone layer of the seed is quite different in the two alleles (Styles et al., 1973). The plant pigmentation directed by *B-Bolivia* can be as dark as that in *B-I,* but *B-Bolivia* pigments a subset of the plant vegetative tissues relative to *B-I*.

The ability of *B-Bolivia* to pigment both seed and plant tissues is reminiscent of alleles of *r1*. The *r1* gene encodes a homologous and functionally duplicate protein to that of *b1* (Ludwig et al., 1989; Goff et al., 1990; Ludwig et al., 1990) and like *b1*, the *r1* gene has many phenotypically diverse alleles, many of which color the aleurone layer of the seed (Styles et al., 1973). Several of the *r1* alleles that color both seed and plant tissues have separate coding regions that are expressed in the seed or in the plant tissues (Stadler and Neuffer, 1953; Robbins et al., 1991; Walker et al., 1995).

Previous investigations of *B-Peru* and *B-I* have demonstrated that each is a simple allele consisting of a single coding sequence. Characterization of the sequences responsible for the aleurone expression of the *B-Peru* allele and investigation of the phylogenetic relationships between several *b1* alleles have revealed that distinct phenotypes correlate with rearrangements or insertions in the upstream region

[1] This work was supported by a postdoctoral fellowship from the Jane Coffin-Childs Memorial Fund for Medical Research (to D.A.S.), by a grant from the U.S. Department of Agriculture National Research Initiative (grant no. 96–35301–3179 to V.L.C.), and by the Department of Army Research for the purchase of the Molecular Dynamics Storm 860 system used in this work (grant no. DAAG559710102).

* Corresponding author; e-mail chandler@ag.arizona.edu; fax 520–621–7186.

of several alleles (Radicella et al., 1992; Selinger et al., 1998; Selinger and Chandler, 1999).

Using genetic and molecular techniques, we have characterized the expression and structure of the *B-Bolivia* allele. Our results indicate that *B-Bolivia* contains a single coding region and that the *B-Peru* and *B-Bolivia* alleles share most of the sequences required for aleurone expression. Immediately upstream of the aleurone-specific sequences in *B-Bolivia* is a highly repeated retrotransposon-related sequence. Transient transformation assays and transgenic plants were used to characterize sequences required for seed expression. Our results suggest that the retro-element-related sequences immediately upstream of the aleurone-specific promoter contributes to some but not all of the epigenetic differences in seed expression between *B-Peru* and *B-Bolivia*.

RESULTS

The *B-Bolivia* Allele Shows Uniform Plant Expression But Variable Seed Expression

The *B-Bolivia* allele, similar to *B-I*, conditions strong anthocyanin pigmentation in several vegetative plant tissues (Coe, 1979), including culm, leaf sheath, and husk tissues (Fig. 1A). *B-Bolivia* also directs anthocyanin expression in the aleurone layer of the seeds. It is the only *b1* allele besides *B-Peru* that confers aleurone pigmentation. Whereas the *B-Peru* allele always conferred uniform and intense pigmentation of the aleurone layer (Fig. 1B), aleurone pigmentation by *B-Bolivia* was weaker and variable (Fig. 1C). As noted by previous workers (Styles et al., 1973), we observed three major differences relative to *B-Peru*. First, not every seed that inherits *B-Bolivia* was pigmented. Second, the amount of pigment in different kernels varied, even though each kernel was homozygous for *B-Bolivia*. Third, the amount of pigment even in the darkest *B-Bolivia* seeds was less than that conferred by *B-Peru*. We have further observed that the proportion of *B-Bolivia* seeds expressing pigment varied between different genetic stocks. The ear in Figure 1C is illustrative of the low number of purple kernels observed when *B-Bolivia* is in the K55 genetic background. The ear in Figure 1D illustrates that a larger number of purple kernels are observed when *B-Bolivia* is in a different background, in this case one derived from the George Sprague (GS) *B-Bolivia* line. The vegetative plant expression is equivalent in both the K55 and GS genetic backgrounds (data not shown).

B-Bolivia Is Not Expressed in the Aleurone if Transmitted through the Male Parent

During our characterization of *B-Bolivia* seed expression we discovered that the presence of colored kernels in progeny from outcrosses between *B-Bolivia* stocks with stocks containing recessive *b1* alleles depended on which stock was the male parent. Ears from *B-Bolivia* plants crossed by pollen from plants with recessive *b1* alleles displayed frequencies of pigmented seeds and intensity of pigmentation that was indistinguishable from self-pollinated ears (Fig. 1D). In contrast, ears from plants with recessive *b1* alleles crossed by pollen from homozygous or heterozygous *B-Bolivia* plants displayed no colored kernels (Fig. 1E). *B-Peru*, the other *b1* allele with aleurone pigmentation does not show female-specific expression.

To further explore *B-Bolivia* transmission, reciprocal crosses were performed between stocks with recessive *b1* alleles and the K55 and GS *B-Bolivia* stocks. The data presented in Table I (experiment 1–4) showed that seeds were pigmented on *B-Bolivia* ears that were pollinated by the *b1* tester stock. However, the ears from *b1* tester plants pollinated by *B-Bolivia*

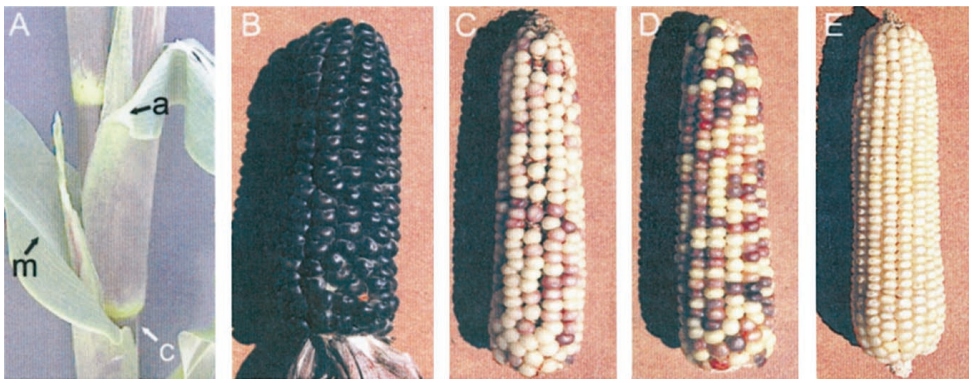

Figure 1. Phenotypes of *B-Bolivia*. A, Vegetative plant pigmentation in a *B-Bolivia* plant: a, auricle; c, culm; and m, leaf mid-vein. B, A self-pollinated ear from a plant homozygous for *B-Peru*. C, A self-pollinated ear from a homozygous *B-Bolivia* plant that shows the incomplete penetrance of seed expression typical of *B-Bolivia*. This ear is in the K55 background. D, This ear resulted from pollination of a heterozygous *B-Bolivia/b* plant by a *b1* tester line that has recessive nonfunctional alleles of *b1* and *r1*, and functional alleles of all other genes required for anthocyanin production. Self-pollinated ears from sibling plants were indistinguishable. E, An ear from a *b1* tester plant pollinated by a heterozygous *B-Bolivia/b* plant. This ear is representative of ears from the reciprocal cross that produced the ear in D.

Table I. *Parent of origin differences on B-Bolivia expression*

Experiment	Stock	Cross[a]	Mean[b]	*N*[c]	High[d]	Low[d]
			%		%	
1	K55	*b1*×	0	36	0	0
2	K55	×*b1*	8.0	11	13.5	2.4
3	GS	*b1*×	0	10	0	0
4	GS	×*b1*	64.1[e]	11	83.6	41.1
5	GS	Self	71.1[e]	9	93.3	55.9
6	414	×*b1*	12.1[e]	19	31.3	1.4
7	414	Self	9.5[e]	36	27.2	1.3

[a] The direction and identity of the other parent in the cross that generated the ears that were scored is indicated as follows: ×*b1*, the *B-Bolivia* stock was crossed by *b1* tester pollen; and *b1*×, the reciprocal cross in which *B-Bolivia* pollen was placed on *b1 tester* ears. [b] The average of the proportion of colored kernels on each ear is given. Kernels were counted as pigmented if they could be clearly distinguished from colorless sibling kernels; a wide range from dark to very lightly pigmented kernels were included. [c] The no. of ears that were counted and used to determine the mean. [d] The proportion of colored kernels on the ears with the highest and lowest percentages in the sample. [e] We calculated the probability that the difference between the self cross and the outcross by *b1* pollen (*B-Bolivia* × *b1*) is due to chance variation for GS and 414 samples using the ANOVA method. This test calculates the probability that the difference in the mean of two groups of numbers is due to random chance. The test results revealed that neither of these differences is statistically significant (GS, $P = 0.24$; 414, $P = 0.21$).

stocks produced no pigmented kernels. Monitoring plant pigmentation demonstrated that *B-Bolivia* is transmitted through both the male and female gametes. Colorless seeds from the reciprocal crosses produced as darkly pigmented plants as colored seeds with no consistent differences in plant pigmentation (data not shown).

The above results indicated that *B-Bolivia* was subject to parent of origin-specific expression in the aleurone, either because of genomic imprinting or because of an effect of gene dosage on *B-Bolivia* expression. Because the aleurone layer is derived from the triploid endosperm there is only one copy present when *B-Bolivia* is transmitted through the male. In contrast, kernels on the reciprocally crossed ear, in which *B-Bolivia* is the female parent and recessive *b1* the male, have two copies of *B-Bolivia*, and kernels on a self-pollinated ear from a *B-Bolivia* homozygote have three copies of *B-Bolivia*.

The classic experiment to determine whether dosage effects or gamete transmission are responsible for expression differences in maize seed is to increase the dosage in the male or reduce the dosage in the female using genetic tools (Kermicle, 1970). Despite considerable effort, we were unable to use similar methods to alter the dosage of *B-Bolivia* through either the male or female. To examine whether we could see a dosage effect between seeds carrying two and three doses of *B-Bolivia*, we compared the proportion of pigmented seeds on ears that were self-pollinated with ears crossed by plants with recessive *b1* alleles in two different stocks, GS and 414. The results presented in Table I (experiment 4–7) showed that in both stocks there were slight differences between the proportion of colored kernels on self-pollinated and outcrossed ears, but this difference was not statistically significant. Further evidence that the differences were not biologically significant came from the observation that the differences in GS and 414 were in opposite directions. The GS outcross ears had fewer colored kernels, but the 414 outcrossed ears had more colored kernels than the self-pollinated ears. The observation that kernels carrying three copies of *B-Bolivia* were no more likely to be pigmented than those carrying two copies suggests that dosage differences are not responsible for the pigment differences.

Pigmented *B-Bolivia* Seeds Are Not Heritably Different from Colorless Seeds with Respect to the Proportion of Pigmented Seeds Produced in the Next Generation

Given the variability of seed expression, we were interested in determining the heritability of aleurone expression in *B-Bolivia*. To test for a correlation between pigmentation of the seed of a parent plant and the proportion of pigmented seeds in its progeny, we planted colored and colorless seeds from the same homozygous *B-Bolivia* ears, self-pollinated the resulting plants, and determined the percentage of colored seeds in the progeny. The experiment, summarized in Table II, was performed with two different *B-Bolivia* stocks, a stock that produced low numbers of purple seeds, 1470 (the parental ear had 9% colored kernels), and a stock that produced increased numbers of purple seeds, GS (the parental ears averaged 64% colored kernels). Plants grown from pigmented 1470 seeds produced ears that averaged 28.7% colored kernels, whereas pigmented GS seeds produced plants with ears averaging 75.5% colored kernels. The plants grown from colorless 1470 seeds

Table II. *Effect of seed color on expression in the next generation*

Colored and colorless seeds from homozygous *B-Bolivia* were planted, the resulting plants selfed, and ears scored for the proportion of purple colored kernels.

Stock	Seed Color[a]	Mean[b]	*N*[c]	High[d]	Low[d]	*P*[e]
		%		%		
1470	Purple	28.7	3	52.2	13.2	–
1470	Colorless	20.9	10	35.2	9.2	0.34
GS	Purple	75.5	6	93.3	55.9	–
GS	Colorless	62.1	3	64.4	58.2	0.19

[a] The color of the seeds that were planted is indicated. Purple seeds were clearly distinguishable from colorless seeds, but were not necessarily the darkest seeds on the ear. [b] The average of the proportion of colored kernels on each ear is given. [c] The no. of ears that were counted and used to determine the mean. [d] The proportion of colored kernels on the ears with the highest and lowest percentages in the sample. [e] The probability that the difference between the ears derived from plants grown from purple and colorless seeds is due to chance variation. The probability was determined using the ANOVA method. In both cases, the differences between the plants grown from colored and colorless kernels were not statistically significant.

averaged 20.9% colored kernels and those of GS averaged 62.1% colored kernels. In both cases, the plants grown from colored seeds produced a somewhat higher proportion of colored kernels, however this difference between the average values was not statistically significant. These results indicate that the on or off pigment expression state of a particular seed is not heritable because both types produce similar numbers of colored kernels in progeny.

The loss of pigment that occurs upon male transmission is also not heritable. When colorless kernels from such ears as shown in Figure 1E are planted and crossed by pollen from plants carrying a recessive *b1* allele, a similar number of colored kernels are observed when compared with kernels that derive from only female transmission. For example, in one experiment, an individual *B-Bolivia/b1* plant was both self-pollinated and outcrossed as male to a *b1* tester plant. Colorless seeds from both ears were planted, plants carrying the *B-Bolivia* allele were determined by plant color, and such ears were crossed by pollen from a *b1* tester line. The 10 plants, which resulted from the colorless *B-Bolivia/b1* seeds that resulted from the original cross as male to the *b1* tester plant, produced ears in which 46% of the kernels carrying *B-Bolivia* were pigmented with a range between 17% to 74%. This was equivalent to the six *B-Bolivia/b1* heterozygous plants derived from colorless kernels on the self-pollinated ear, as the proportion of *B-Bolivia* kernels that expressed color on these ears averaged 47% with a range from 19% to 77%.

B-Bolivia Allele Has Part of the *B-Peru* Aleurone-Specific Promoter Region

To investigate whether *B-Bolivia* is a simple or complex allele and to determine if its unique expression patterns result from unique promoter sequences, a molecular study of *B-Bolivia* was initiated. Initially a restriction map was generated for *B-Bolivia* using DNA gel blots probed with DNA fragments derived from the *B-I* and *B-Peru* alleles. Using the 550b probe that lies near the 5′ end of the transcribed region in *B-I* and *B-Peru* (Patterson et al., 1995), we found that there is a single *b1* coding region in *B-Bolivia* and that it had the same map as the *B-Peru* and *B-I* coding regions (Fig. 2A). Because *B-Peru* and *B-I* differ dramatically in the upstream region, we extensively mapped the region of *B-Bolivia* upstream of the transcribed sequences. We used the 550b probe and combined double digests with *Bam*HI and other enzymes. *Bam*HI cuts at the 3′ end of the 550b sequence providing an anchor site to facilitate mapping (Fig. 2A). We found that the upstream region of *B-Bolivia* contained many distinct RFLP when compared with the same regions of *B-Peru* and *B-I* (Fig. 2A; data not shown).

Based on the mapping, we identified a 2.8-kb *Bam*HI fragment from *B-Bolivia* for cloning, which contained approximately 2.1 kb of upstream sequence ("Materials and Methods"). Two λ-clones were isolated from a size selected *Bam*HI digested genomic DNA library, converted to a plasmid and the inserts were restriction mapped. Both inserts had identical restriction maps that matched the restriction fragment sizes determined from DNA-blot analysis of genomic DNA. One of the two clones was completely sequenced and this sequence was compared with that of the *B-Peru* and *B-I* alleles. These sequence comparisons revealed that part of the upstream region of the *B-Bolivia* clone was almost identical to part of the aleurone-specific promoter of *B-Peru*, differing by a single 4-bp insertion in the *B-Bolivia* sequence relative to *B-Peru*. This homology extended to 530 bp upstream of the start of transcription (for diagram, see Fig. 2B). Beyond this point the sequences completely diverged.

Divergent Sequence in *B-Bolivia* Has Homology to Retrotransposons and Is Present in the Maize Genome in Very High Copy Number

To determine the nature of the divergent sequence in *B-Bolivia*, sequence homology searches of GenBank were conducted and DNA-blot analyses were performed using the divergent sequence as a probe.

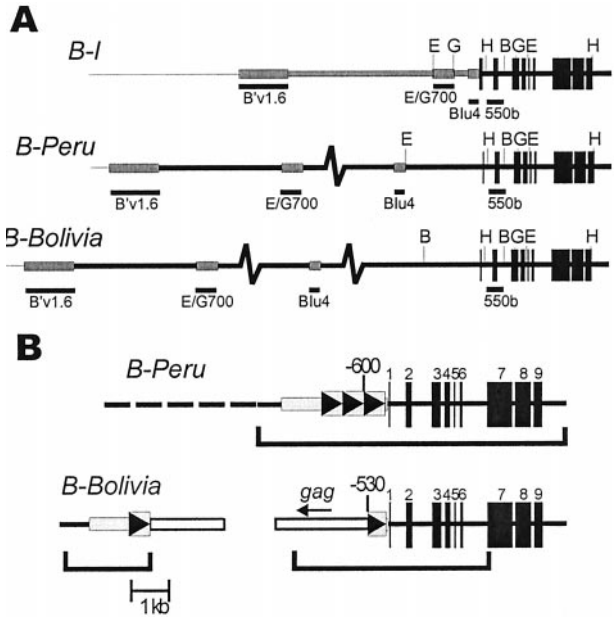

Figure 2. Structures of the *b1* alleles. A, Restriction maps comparing *B-Bolivia* with the previously characterized *B-Peru* and *B-I* alleles. The probes used in DNA-blot analyses are indicated on each of the maps, and the gray boxes indicate the regions they hybridize to. For clarity, only a subset of the mapped sites are shown. Restriction sites are *Bam*HI (B), *Bgl*II (G), *Hin*dIII (H), and *Spe*I (E). B, The regions of the two alleles that have been cloned and completely sequenced are indicated by the brackets underneath the structures. The sequences that are homologous between these two alleles and other *b1* alleles are shown in black, and the exons are indicated as large boxes and numbered. The upstream sequences of *B-Peru* that are homologous to regions in *B-Bolivia* but not shared with other *b1* alleles are shown in light gray, and the 534-bp repeats are indicated by black arrowheads. The number above the *B-Peru* sequence shows the end of the proximal 534-bp repeat sequence and the number above the *B-Bolivia* sequence shows the point of divergence between the two alleles. The putative insertion in *B-Bolivia* is represented by open boxes that are separated by a gap indicating that the intervening distance and the orientation of these two sequences has not been determined. The region labeled *gag* indicates sequences homologous to the *gag* proteins of various retrotransposons. The arrow below *gag* indicates the direction of the reading frame.

Using BLASTN (Altschul et al., 1990) to search for DNA sequences homologous to the upstream region of *B-Bolivia* no significant homologies were found outside of the region that is nearly identical to *B-Peru*. Using BLASTX and FASTX searches (Altschul et al., 1990; Pearson et al., 1997), in which the nucleotide sequence of *B-Bolivia* is translated into all six possible polypeptide sequences and compared with the protein sequence database, we found a sequence within the *B-Bolivia* upstream region that, when translated, had highly significant sequence identity to protein sequences in GenBank. This search identified a reading frame with 35% identity over 308 amino acids (Bit score of 129, E(513612) was 3e-27) to a *gag* polyprotein from Sorghum (gb: AAD19359.1), a protein characteristic of retro-elements. This was located in the 5′-most 973 bp that is in the opposite orientation from the *b1* coding region (Fig. 2B). Although this region of *gag* homology has similarity at the amino acid level to several high copy retrotransposons in maize, the *Grande* element (accession no. X97604 and X97605), and elements in the *adh1* flanking region (accession no. AF123535; SanMiguel et al., 1996), the lack of significant identity at the nucleotide level precludes the sequences in *B-Bolivia* from belonging to any identified maize retrotransposon family.

To better place the *B-Bolivia* insertion sequence within the context of these other *gag* sequences, we used the PROTPARS program of the PHYLIP package to produce a phylogenetic tree of the *B-Bolivia* insertion sequence with the 19 *gag* proteins identified ("Materials and Methods"). The results of this analysis, shown in Figure 3, demonstrated that the *B-Bolivia* insertion sequence is related to *gag* proteins from other plant retrotransposons, but has significantly diverged from its nearest relation identified to date, the *Sorghum bicolor Retrosor* element (gi: 4378066). These results strongly suggest that the sequences in *B-Bolivia* that are absent from *B-Peru* represent a retrotransposon- or retrotransposon-related sequence.

Several different classes of retrotransposons have been found in maize. Elements like *Bs1*, *B5*, *G*, *Hop-*

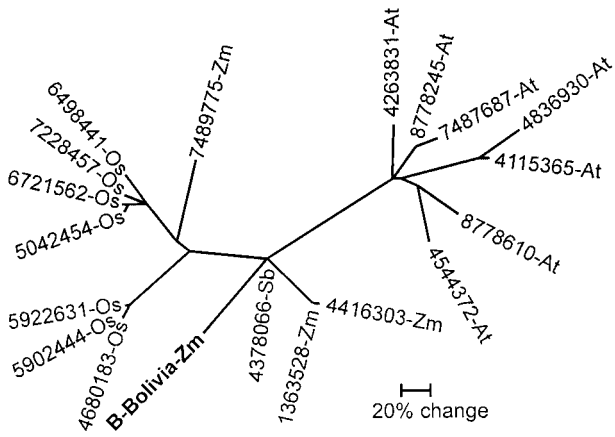

Figure 3. Parsimony phylogenetic tree of *gag* protein homologs of the *B-Bolivia* upstream sequence. This phylogenetic tree represents the consensus parsimony tree from analysis with 500 bootstrap replications. The labels indicate the GenBank gi number for that protein sequence and the two-letter abbreviation for the species from which the sequence derives. At, Arabidopsis; Os, *Oryza sativa*; Sb, *Sorghum bicolor*; Zm, *Zea mays*. Sequence 1363528 is from the *Zeon-1* element (Hu et al., 1995), all of the other putative proteins are from uncharacterized retrotransposon-like elements. The branch lengths indicate evolutionary distance and were determined from the FITCH program of the PHYLIP package (for details, see "Materials and Methods"). The bar labeled 20% change indicates the distance produced by a 20% difference in sequence identity. Labels that appear to be at nodes of the tree are sequences that have very short distances from the presumed ancestral sequence. Except for three of the nodes in the Arabidopsis clade, all branches are supported in greater than 75% of the trees.

scotch, and *Stonor* are relatively low copy elements associated with genic sequences (Johns et al., 1985; Varagona et al., 1992; White et al., 1994). In contrast, there are several families of retrotransposons in maize that have very high copy numbers and make up a significant fraction of the maize genome (Bennetzen et al., 1994; SanMiguel et al., 1996). To determine the approximate copy number of the sequences adjacent to *B-Bolivia*, we blotted known amounts of the cloned *B-Bolivia* upstream sequence and known amounts of genomic maize DNA from several distinct stocks on a slot blot. We probed the resulting slot blot with the cloned *B-Bolivia* sequence. After quantifying the blot, we found that 3 μg of maize DNA had approximately 6 times the number of counts as did 10 ng of the cloned DNA from which the probe was made (Fig. 4). This intensity difference provides an estimate of approximately 38,000 copies in the 2,500 megabase maize genome ("Materials and Methods"). Because the comparison of the genomic DNA signal was to a DNA sample that exactly matches the probe sequence, the actual number of copies in the maize genome could be higher due to sequence divergence. The 38,000-copy number is very similar to that estimated for the 9-kb *Opie* retrotransposon (SanMiguel et al., 1996) and several other retrotransposons found in maize (Bennetzen et al., 1994; SanMiguel et al., 1996). However, the sequence in *B-Bolivia* is clearly not any of these previously described high copy retrotransposons.

B-Bolivia Has a Large Insertion or DNA Rearrangement Relative to *B-Peru*

We next set out to define the size of the putative insertion in *B-Bolivia*. We used several probes to map restriction enzyme sites in the upstream region of *B-Bolivia* and *B-Peru*. By hybridizing with a probe near the 5′ end of the transcribed region (550b; Patterson et al., 1995), we mapped several sites upstream of the start of transcription in *B-Bolivia* (Fig. 5). We next probed the same blot with a probe located 2.5 kb upstream of the start of transcription in

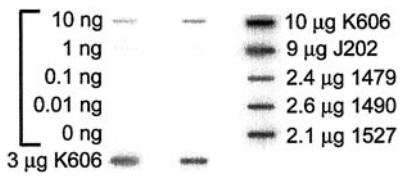

Figure 4. Slot-blot analysis of the copy number of the putative retrotransposon from *B-Bolivia*. The two columns on the left and middle are duplicates. The bracketed slots contain the indicated quantity of unlabeled cloned *B-Bolivia* upstream DNA (a 1.3-kb region) diluted into 3 μg of genomic carrier DNA (from petunia). The bottom most slots of the first two columns contain 3 μg of maize genomic DNA from the K55 *B-Bolivia* line (K606). The column on the right contains the indicated amount of maize genomic DNA from various stocks with the following *b1* alleles: K606, *B-Bolivia*; J202, a recessive *b1* allele; 1479, *B-Gua31*; 1490, *B-marker*; and 1527, *B-Peru*.

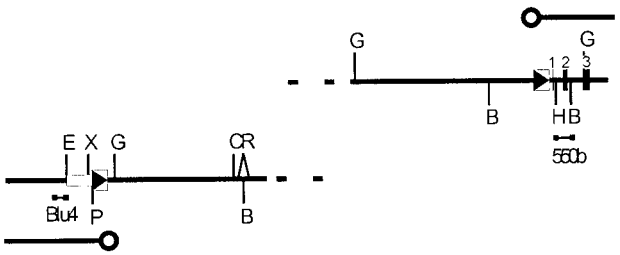

Figure 5. Restriction map of the upstream region of *B-Bolivia*. Exons 1 to 3 are numbered, and the *B-Peru* homologous sequences are indicated in the same manner as in Figure 2B. The labeled square ended lines indicate the BIu4 and 550b probes. The restriction sites are abbreviated as follows: *Bam*HI (B), *Bcl*I (C), *Bgl*II (G), *Eco*RI (R), *Hin*dIII (H), *Pst*I (P), *Spe*I (E), and *Xba*I (X). The lines above and below the *B-Bolivia* map that end in white circle represent the regions of *B-Bolivia* that have the same restriction map as *B-Peru*.

B-Peru (BIu4; Fig. 2A). This analysis did not reveal the size of the insertion as none of the restriction enzymes tested yielded fragments that hybridized to both the transcribed sequence and upstream sequence probes.

Mapping of the regions farther upstream than the BIu4 probe revealed that *B-Bolivia* and *B-Peru* share the same pattern of restriction fragments with eight different enzymes ("Materials and Methods"). Furthermore, PCR cloning of the region of *B-Bolivia* including the BIu4 probe and flanking *B-Peru*-like sequence, confirmed that this region of *B-Bolivia* matched the sequence in B-*Peru* between 1 and 3 kb upstream of the start of transcription (left bracket in Fig. 2B, underlined in Fig. 5). This region of *B-Bolivia* is almost identical to the *B-Peru* sequence and contains one of the three 534-bp direct repeats found in the *B-Peru* upstream region.

Having mapped the upstream sites, we could now determine the location in *B-Bolivia* of sites downstream of the BIu4 probe. We found that all of the restriction sites within 1.5 kb downstream of the BIu4 probe in *B-Bolivia* were identical to the sites in *B-Peru* (data not shown). In contrast, all restriction enzymes with sites farther than 1.5-kb downstream of BIu4 produced different sized fragments, indicating that the sequences further than 1.5 kb downstream of the BIu4 probe were quite different in *B-Bolivia* relative to *B-Peru* (data not shown). Comparison of restriction sites mapped upstream of the transcribed region (using the 550b probe) and downstream of the BIu4 probe, revealed that the two maps do not overlap (Fig. 5). Assuming the full 2.5 kb of *B-Peru*-like sequence is still found between these probes, there is approximately 4.5 kb of non-*B-Peru* homologous sequence between the BIu4 probe and the *Hin*dIII site. Adding this to the approximately 6 kb of non-homologous sequence between the 550b probe and the *Bgl*II site gives a minimum size of 10.5 kb for the "insertion." It is also possible that a more complex DNA rearrangement is responsible for the juxtaposition of the retrotransposon sequences next to *B-Bolivia*.

Transient Transformation Assays Reproduce the Quantitative Difference between *B-Peru* and *B-Bolivia* Aleurone Pigment

One difference between *B-Bolivia* and *B-Peru* kernel pigmentation is that even the darkest *B-Bolivia* kernels are less pigmented than *B-Peru* kernels. A quantitative transient transformation assay was used to determine if the −2,100 *B-Bolivia* upstream region produces a different level of aleurone expression relative to the −2,500 *B-Peru* upstream region that was previously studied (Selinger et al., 1998). In this assay the *B-Bolivia* or *B-Peru* upstream sequences were fused to the reporter gene, firefly luciferase, and the constructs were introduced into maize aleurone cells (Selinger et al., 1998). To normalize for transformation efficiency, the test promoter:luciferase constructs were co-transformed with either a β-glucuronidase (GUS) reporter gene construct or a Renilla luciferase construct both driven by the cauliflower mosaic virus (CaMV) 35S promoter. All luciferase values were normalized to one of these transformation controls. We found that the 2.1-kb *B-Bolivia* upstream sequence generated only 42% of the aleurone expression produced by the −2,500 *B-Peru* construct, to which we scaled all of our results (Fig. 6). This result indicated that either some part of the *B-Bolivia* sequence reduced aleurone expression or some part of the *B-Peru* sequence that is missing from *B-Bolivia* enhanced aleurone expression. Our results discussed below suggest both types of events may be operating.

We had shown previously that 5′-promoter deletions of *B-Peru* to −176 had the same aleurone expression in the transient assay as the −2,500 construct (Selinger et al., 1998). To determine if the single difference in this region between the two alleles, a four base insertion in *B-Bolivia* relative to *B-Peru* just downstream of the TATA box, had any effect on expression, we compared a −176 *B-Peru* promoter construct with the comparable −180 *B-Bolivia* construct. Both constructs gave the same level of expression, which was comparable with the level of expression produced by the −2,500 *B-Peru* construct (Fig. 6). This result indicated that the four base insertion immediately downstream of the TATA box in *B-Bolivia* does not affect expression and that the *B-Peru* homologous sequences of *B-Bolivia* are fully capable of driving levels of aleurone expression equivalent to that of *B-Peru*.

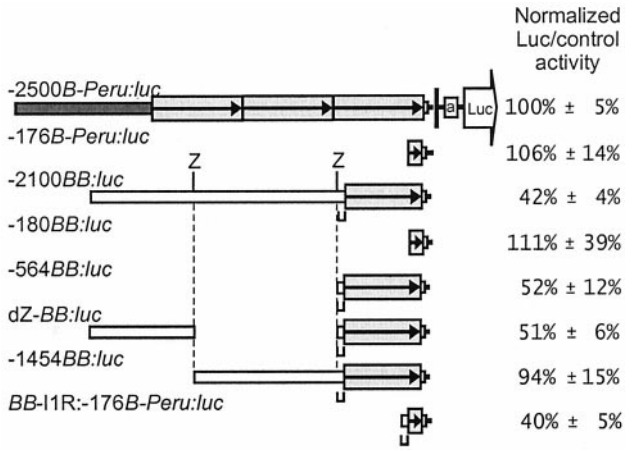

Figure 6. Deletion analysis of the *B-Bolivia* upstream promoter proximal region. The constructs shown represent promoter fusions to luciferase (shown as an open arrow labeled Luc) with the maize *Adh1* Intron1 sequence (labeled "a") as a leader sequence. Details of the reporter gene are in "Materials and Methods." Each construct was introduced into maize aleurones along with either a GUS or Renilla luciferase expressing control to normalize for transformation efficiency. The Luc/control activities were normalized to results obtained with the −2,500 *B-Peru* construct (100%; Selinger et al., 1998). The results were from at least five independent transformations and are given with the SE of measurement. The *Xho*I sites (Z) are labeled on the −2,100BB:luc construct and dashed lines indicate the same sites in the deletion constructs. The sequences found in *B-Bolivia* that are not present in *B-Peru* are indicated by the unshaded boxes. The bracket under the constructs indicates the 33 bp of *B-Bolivia*-specific sequence that is in common between five of the constructs and that appears to contain a negative regulatory element.

A 33-bp Sequence from *B-Bolivia* Reduces Aleurone Expression

The results with the −180 *B-Bolivia* construct suggested that some other part of the 2,100 bp of *B-Bolivia* upstream sequence was responsible for the reduced aleurone expression. To determine if part of the 1.5 kb of novel *B-Bolivia* sequence is responsible for the reduced expression we analyzed several deletion derivatives (Fig. 6). We used the two *Xho*I sites to produce three deletion derivatives. The smallest construct, a deletion to the *Xho*I site at −564 (−564*BB:luc*), produced a reduced expression level that was very similar to that of the −2,100 construct. The deletion of the 900 bp between the *Xho*I sites at −1,454 and −564 (d*Z-BB:luc*) also produced a level of expression that was similar to that of the −2,100 *BB:luc* construct, corroborating the results from the −564 deletion and suggesting that the 33 bp between the −564 *Xho*I site and the start of the *B-Peru* homologous sequence at position −531 was sufficient to reduce aleurone expression (Fig. 6). However, deletion to the first *Xho*I site (−1,454*BB:luc*) produced luciferase expression equivalent to the −2,500 *B-Peru* construct (Fig. 6). Because of these conflicting results, we decided to test the ability of the 33 bp between −564 and the *B-Peru* homologous sequences at −531 to reduce the expression of the −176 *B-Peru* promoter:luciferase construct. We tested a construct with a single copy of the 33-bp sequence subcloned upstream of the −176 *B-Peru:luc*, and found that this single copy of the 33-bp sequence reduced aleurone expression to 40% (Fig. 6). Thus, the effect of the 33 bp of sequence accounted for the reduction seen in the −2,100 *B-Bolivia* construct. In addition, the reduction to 40% in the transient assay correlated nicely

with the approximately 30% level of anthocyanin pigment found in the darkest *B-Bolivia* seeds relative to *B-Peru* seeds (data not shown). However, the results with this chimeric *B-Bolivia*:*B-Peru* construct do not explain why the 33-bp sequence, which is present in the −1,400 deletion construct, did not reduce expression in that context.

The Region of the *B-Peru* Aleurone-Specific Promoter That Is Missing from the Promoter Proximal Region of *B-Bolivia* Contributes to Aleurone Expression

Besides the sequence that is unique to *B-Bolivia* relative to *B-Peru*, *B-Bolivia* is missing part of the 534-bp repeated sequence that is found in *B-Peru*. The *B-Peru* promoter region contains three identical 534-bp direct repeats. All of the important aleurone regulatory sequences are found in a single 534-bp repeat sequence (Selinger et al., 1998) and a *B-Peru* deletion derivative allele with a single 534-bp repeat has the same expression and stability as the native *B-Peru* allele (Harris et al., 1994). In the analysis of the *B-Peru* promoter, the −710 and −176 promoter regions produced equivalent expression in transient transformation experiments (Selinger et al., 1998). In *B-Bolivia*, the aleurone-specific promoter consists of 464 bp of the 534-bp repeat sequence; the distal 70 bp of the 534-bp repeat is either deleted or located elsewhere due to the insertion or rearrangement. The contrast between the stable expression of *B-Peru* and the variable expression of *B-Bolivia* suggested the hypothesis that the 70 bp of the repeat that are missing in *B-Bolivia* might have a quantitative effect on aleurone expression or the stability of this expression.

To determine if the 70-bp region might contain regulatory elements important for aleurone expression, we tested whether the presence of this region could suppress mutations in two critically important regions in the first 176 bp of the promoter. We had previously identified and characterized the E1 and E2 regions of the *B-Peru* aleurone-specific promoter (Selinger et al., 1998). In the context of the −176 *B-Peru* promoter, mutation of E1, which corresponds to the −120 to −109 region, results in a loss of expression to 17%, whereas mutation of E2, which is located between positions −96 and −85 results in a reduction of expression to 7% (Fig. 7). When these mutations were made in the context of the −600 *B-Peru* promoter and tested in the aleurone transient transformation assay, expression was 40% and 60%, respectively, for E1 and E2 (Fig. 7). Importantly, we had previously shown that a −559 promoter construct carrying the E2 mutation had the same 7% level of expression seen in the −176 promoter (Fig. 7). These results suggest that the sequences in *B-Peru* between −600 and −176 contribute to aleurone expression and specifically localizes the element(s) critical for the suppression of the E2 mutation to 31 bp between −600 and −559, which is within the 70 bp missing in *B-Bolivia*.

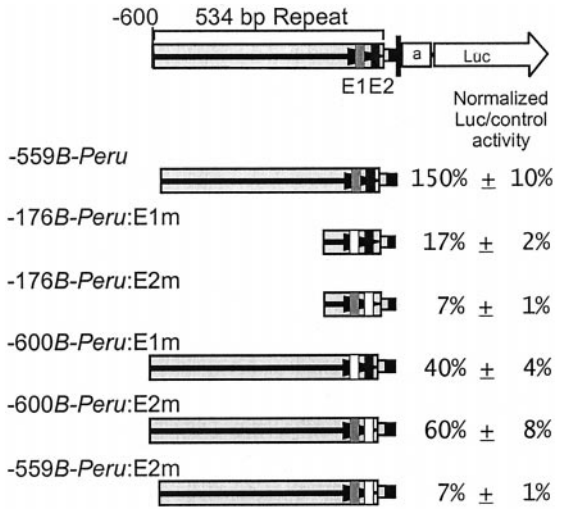

Figure 7. The distal part of the 534-bp *B-Peru* repeat contains important regulatory sequences. The structure of one of the *B-Peru* 534-bp repeats (boxed black arrow) fused to the *Adh1* intron 1 (a) and firefly luciferase (Luc) is shown at the top of the figure. The E1 and E2 regions, which were previously shown to be important for aleurone expression (Selinger et al., 1998), are indicated by a gray and black box, respectively. The mutated versions of these two elements are indicated by white boxes at the same position that E1 or E2 should be. These mutants were generated by substitution mutagenesis as previously described (Selinger at al., 1998). Each construct was introduced into maize aleurones along with either a GUS or Renilla luciferase expressing control to normalize for transformation efficiency. The expression levels of the test constructs are indicated as percentages of the −2,500 *B-Peru*:luciferase expression level set at 100%, after normalization to the internal control. The results for the −559 *B-Peru* construct are taken from Selinger et al. (1998). Each construct was assayed in five independent transformations and the mean value is given with the SE of measurement.

Transgenic Plants Containing the 2.1-kb Upstream Region of *B-Bolivia* Can Produce Aleurone-Specific Expression

Previous studies with the −2,500 bp region of *B-Peru* had demonstrated that these sequences could confer aleurone but no plant expression. To determine if the cloned upstream region of *B-Bolivia* could reproduce the aleurone and plant expression phenotypes characteristic of this allele, we produced transgenic maize plants carrying the *B-Bolivia* construct. We ligated the 2.8-kb *B-Bolivia* clone, which has 2.1 kb of upstream sequence with the cloned *B-I* coding region to produce a full length reconstruction of a genomic clone of the *B-Bolivia* upstream and *B-I* coding region (Fig. 8A). We generated transgenic plants containing this construct by cobombarding the pBB2100 plasmid with a selectable marker, the *bar* gene driven by the CaMV 35S promoter, into immature embryos. We regenerated plants from 51 independent stably transformed lines that were resistant

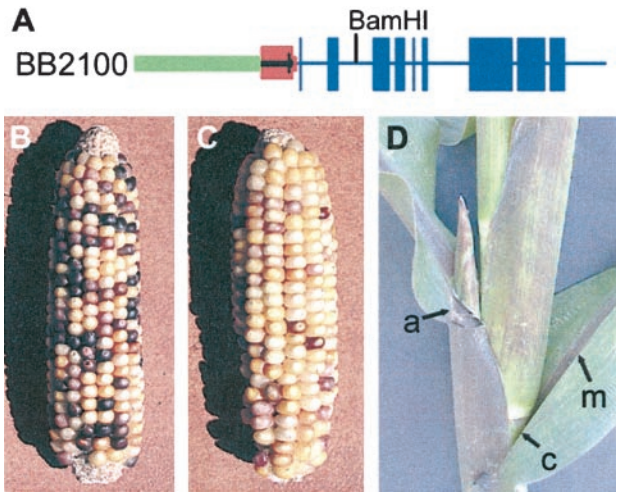

Figure 8. Phenotypes of *B-Bolivia* transgenic plants. A, Diagram of the BB2100 construct. The green block indicates *B-Bolivia*-specific upstream sequence, the red box with a black arrow represents the sequence that is homologous to the *B-Peru* aleurone-specific promoter, and the blue region indicates the exons (boxes) and introns (lines) of the transcribed region. This construct contains 2.1 kb of upstream sequence from *B-Bolivia* together with exon1, intron1, and exon2 of *B-Bolivia* fused at the *Bam*HI site in intron2 with the remaining genomic coding region of *B-I*. B and C, Two ears from hemizygous BB2100 transgenic lines crossed by a *b1 tester* line. The VLC 40-64 ear (B) has approximately 50% colored kernels, whereas the VLC 40-59 ear (C) has approximately 30% colored kernels. D, The plant phenotype of a BB2100 transgene line (VLC 40-20). Note the phenotypic differences in the auricle (a), culm (c), and leaf mid-vein (m) relative to Figure 1A.

to the herbicide Basta, indicating that they were expressing the *bar* gene. Plants from 14 of these independent lines expressed anthocyanin pigment in seeds when pollinated by a *b1* tester stock, homozygous for functional alleles of all of the anthocyanin pathway genes except *b1* and *r1*. The proportion of BB2100 lines with seed pigmentation (27.5%) was similar to the proportion of *B-Peru* transgenic lines with seed pigmentation previously isolated (36%, Selinger et al., 1998). DNA blots and/or PCR analyses on progeny derived from crossing the 14 different hemizygous transgenic lines with non-transgenic *b1* tester demonstrated that, in at least nine of the 14 lines, *B-Bolivia* transgene copies cosegregated as a single locus (data not shown). One line clearly showed segregation of at least two transgenic loci (data not shown) and for the remaining four lines, insufficient data were obtained to clearly determine that a single locus was segregating.

To determine if the BB2100 transgenic plants would show the variable seed expression characteristic of kernels carrying the native *B-Bolivia* allele, we counted the proportion of colored kernels on ears from T1 plants. For those lines with a single transgene locus, when each hemizygous transgenic plant is outcrossed, 50% of the progeny will receive the transgene. If seed expression of the BB2100 transgene was completely penetrant, like that of *B-Peru*, then we would expect all of the seeds from these nine lines carrying single locus BB2100 transgenes to be pigmented, producing 50% colored kernels. Table III contains a comparison of the proportions of colored kernels in ears from all 14 lines with the nine lines verified to have a single transgene locus indicated. Two of the nine lines produced ears in which 50% of the kernels expressed pigment (lines VLC 40-16 and 40-64, Table III). An example of one such ear is shown in Figure 8B. When we tested molecularly for the presence of the BB2100 transgene in 20 colorless seeds from one of these lines, we found that none of the kernels carried the BB2100 transgene. These results suggest that, in these two lines, the BB2100 transgene was completely penetrant, like *B-Peru* and the *B-Peru* transgenic lines (Selinger et al., 1998). In the other seven verified single transgene locus BB2100 lines, less than 50% of the seeds were pigmented. The ears were similar to that of the native *B-Bolivia* allele, in that these had reduced numbers of colored kernels and the colored kernels were generally less pigmented (Fig. 8C). When we molecularly tested for the presence of the BB2100 transgene in colorless kernels from one of these lines with fewer colored kernels (VLC40-59), we found that 25% of the colorless kernels carried the transgene (5 of 20).

Colorless kernels that received the native *B-Bolivia* allele produced plants that had colored seeds in the next generation. To test if the transgene loci behaved similarly, we grew and test-crossed transgenic BB2100 plants grown from colorless seeds from two independent lines known to have a single transgene locus. We observed that transgenic plants grown from colorless kernels produced colored kernels in the next generation in similar proportions to their siblings grown from colored seeds. These results indicated that in these lines the BB2100 transgenes in the colorless kernels were not heritably silenced. This expression pattern is reminiscent of the incomplete penetrance in seed expression seen in the native *B-Bolivia* stocks.

In addition to the incomplete penetrance of seed expression, the native *B-Bolivia* allele normally fails to produce any pigmented seeds when transmitted through the pollen. We tested 11 of the 14 BB2100 transgenic lines for seed pigment expression when transmitted through pollen. For all 11 lines, colored kernels were observed, regardless of which direction the cross was performed with recessive *b1* alleles. Data from two lines with multiple ears from crosses in both directions are presented in Table IV. These observations indicate that none of the transgenic lines display parent of origin-specific expression.

BB2100 Transgenic Lines Produce Novel Plant Pigmentation

Six of the 14 BB2100 lines, in which the transgene was expressed in the seed, displayed plant pigmen-

Table III. *B-Bolivia transgenic lines*

Plants hemizygous for each transgenic locus were outcrossed by *b1* tester and the number of colored kernels determined.

Transgenic Line	% Colored Seeds[a]	χ^2 test[b]	Plant Expression[c]	No. of Transgene Copies[d]	Single Transgene Locus[e]
VLC 38-42	33 ± 4	N/A	+	1 intact + 6	N.D.
VLC 39-2	47 ± 5	N/A	−	1 intact + 5	N.D.
VLC 40-9	48 ± 3	N/A	−	1 intact + 3	N.D.
VLC 40-12	41 ± 5	9.91×10^{-13}	+/−	1 intact + 5	Yes
VLC 40-16	48 ± 3	0.0356	−	1 intact + 3	Yes
VLC 40-20	44 ± 3	N/A	+++	0 intact + 7	No
VLC 40-32	40 ± 4	5.70×10^{-16}	−	1 intact + 1	Yes
VLC 40-37	47 ± 1	N/A	++	0 intact + 2	N.D.
VLC 40-38	44	N/A	−	? intact + 10	Yes
VLC 40-45	40 ± 7	1.59×10^{-26}	−	1 intact + 4	Yes
VLC 40-59	38 ± 7	1.90×10^{-25}	+++	0 intact + 5	Yes
VLC 40-64	49 ± 2	0.4752	++	1 intact + 13	Yes
VLC 40-68	18	N/A	+++	0 intact + 3	Yes

[a] The mean and SD of kernel counts from at least four ears (except for the VLC 40-38 and 40-68 lines for which only one ear was available for analysis). [b] The probability from a χ^2 analysis that the actual seed counts are consistent with the expectation that 50% of the seeds are colored. N/A, Not applicable. [c] Strong pigmentation of the auricle and leaf mid-vein tissues is represented by +++, moderate pigmentation by ++, light pigmentation by +, very weak pigmentation by +/−, and no anthocyanin pigment by −. [d] As determined by DNA blots probed with the 550-b probe. [e] Determined from segregating populations by either PCR- or DNA-blot analysis. N.D., Not determined. The VLC40-20 line is clearly segregating at least two transgene loci, although only one may be functional.

tation (Table III; Fig. 8D). The phenotypes of these six lines were quite similar to each other, but were different from that of the native *B-Bolivia* allele. The plant phenotype of the transgenic lines is essentially opposite of the phenotype of the native allele in three respects (compare Fig. 1A with Fig. 8D). First, the transgenic lines had strong expression in the auricle tissue that separates the sheath from the leaf blade and in the leaf mid-vein. Plants carrying the native *B-Bolivia* allele strongly pigment the sheath, but never produce pigment in the auricle and rarely produce weak pigmentation of the leaf mid-vein. Second, the transgenic lines produced relatively weak pigmentation of the culm, a tissue that is strongly and uniformly pigmented by the native *B-Bolivia* allele. Third, three of the six transgenic lines produced pigment that was strong in the margins of the sheath. In contrast, the native allele produced no pigment in sheath margins, even in plants with intense pigmentation in the rest of the sheath.

We considered several hypotheses to explain why only six of the 14 lines with seed expression showed plant expression and why this plant expression did not mimic the native *B-Bolivia* allele. One possibility was that enhancers at the integration sites were influencing the expression pattern. This seemed unlikely given that the chromosome position was different in each line while the pattern of pigment expression was similar. In addition, in other experiments we generated many other transgenic lines containing either the *B-Peru* (Selinger et al., 1998) or *B′* (K Kubo, V. Chandler, personal communication) genomic clones, in which none of the plants exhibited any plant pigmentation. Thus, it is unlikely there are a fortuitously high number of plant-specific enhancers in the genome. A second possibility was that expression of an endogenous *b1* allele in the transgenic plants was influencing transgene expression through an RNA silencing mechanism (Jorgensen, 1995). This seemed unlikely as the presence or absence of plant pigmentation was stable and did not depend on the identity of the endogenous *b1* allele or whether or not the endogenous allele was expressed in the plant (data not shown).

Another possibility was that the lines lacking plant expression have transgenes with promoter deletions and are thus missing key regulatory sequences for plant expression. We used DNA-blot analysis and restriction enzymes *Eco*RI and *Pac*I, which flank the promoter and coding region, respectively, to assess the copy number of intact and partially deleted transgene copies in the 14 lines with seed expression. These results, which are shown in Figure 9 and summarized in Table III, indicate that nine of the BB2100 lines have at least one intact copy, and one or more rearranged or partially deleted copies of the transgene. Contrary to the expectations of our hypothesis, four of the lines that have plant color have no intact copies (Fig. 9). Six of the nine lines with intact copies have no plant expression, indicating that the presence of an intact 2.1-kb upstream region fused to an intact coding region is not sufficient for plant expression.

Four of the independent lines that produced strong seed pigmentation had no bands consistent with having an intact transgene copy, suggesting that they have at least one copy with an intact coding region fused to promoter sequences sufficient for seed expression. Because the *Pac*I site is located a few bases beyond the stop codon, most deletions of this site are likely to result in the production of nonfunctional

Table IV. *Transgenic lines do not show parent of origin differences in expression*

Plants hemizygous for each transgenic locus were outcrossed with *b1* tester and the no. of colored kernels determined.

Stock	Cross[a]	Mean[b]	*N*[c]	High[d]	Low[d]	*P*[e]
		%		%		
VLC 38-42	*b1*×	32	4	39	29	–
VLC 38-42	×*b1*	25	8	44	10	0.3337
VLC 40-20	*b1*×	44	6	48	41	–
VLC 40-20	×*b1*	46	4	50	44	0.1859

[a] The direction of the cross that generated the ears that were scored is indicated as follows: ×*b1*, the BB2100 stock was crossed by *b1* stock pollen; *b1*×, the reciprocal cross in which BB2100 pollen was placed on *b1* ears. [b] The average of the proportion of colored kernels on each ear is given. [c] The no. of ears that were counted and used to determine the mean. [d] The proportion of colored kernels on the ears with the highest and lowest percentages in the sample. [e] The probability obtained by the ANOVA method that the difference between the outcross by *b1* pollen (×*b1*) and the outcross to *b1* ears (*b1*×) is due to chance variation.

proteins. The *Eco*RI site is located a few base pairs upstream of the *Bam*HI site that is the 5′-most end of the *B-Bolivia* upstream sequence (Fig. 9). Deletion of the *Eco*RI site and up to 1.5 kb of downstream sequence may have little effect on seed expression because the *B-Peru* homologous aleurone promoter sequence would remain intact.

Although the complexity of the transgene arrays in most lines made detailed characterization of their promoter structure difficult, PCR analysis was used to determine the extent of the promoter deletions in the two BB2100 lines with the fewest number of transgene copies. These two lines, VLC 40-37 and VLC 40-68, had strong plant expression, but no intact copies of the promoter region and only two to three partially deleted copies (Table III; Fig. 9). We used a series of upstream primers starting with the 531 primer located at the upstream end of the *B-Peru* homologous aleurone promoter region with additional primers located further upstream and spaced approximately every 300 bp (Fig. 9). Amplifications were done using one of the upstream primers with either of two downstream primers, the Sac primer located at the 5′ end of intron 1 or the 532u primer located in the same region as the 531 primer but oriented in the opposite direction. Using the 531 and Sac primers, we amplified a 550-bp fragment from both transgenic lines and from the native *B-Bolivia* allele indicating that both lines contained all of the 531 bp of upstream sequence that is homologous to the aleurone-specific promoter region of *B-Peru* (data not shown). Additional reactions using upstream primers at −700, −1,092, and −1,399 with one of the two downstream primers resulted in the amplification of the appropriately sized fragments from both the VLC 40-37 and VLC 40-68 transgenic lines (data not shown). PCR with an upstream primer at −1,682 paired with the −532 primer generated the expected product with DNA from the VLC 40-68 line, but not the VLC 40-37 line. Thus, the deletion in VLC 40-37 begins somewhere between −1,399 and −1,682, whereas the deletion in VLC 40-68 begins between −1,682 and the *Eco*RI site at −2,100.

In summary, there was no clear correlation between the presence of specific promoter sequences and plant expression in the transgenic lines. A final possibility we considered is that another sequence within the transgene array may be contributing to plant pigmentation. This hypothesis is further developed in the discussion.

DISCUSSION

Our results demonstrate that a DNA rearrangement that places sequences from a high copy number retrotransposon adjacent to aleurone-specific promoter elements is associated with altered patterns of expression in *B-Bolivia*. Our finding that *B-Bolivia* shares the same aleurone-specific promoter sequences with *B-Peru* explains why both alleles are expressed in the aleurone. However, it does not explain the variability in the seed expression phenotype of *B-Bolivia* relative to *B-Peru*. We hypothesized that the presence of a highly repetitive element adjacent to regulatory sequences required for aleurone expression could be sufficient to produce the differences in seed expression between *B-Peru* and *B-Bolivia*. To examine this possibility, we produced transgenic plants containing 2.1 kb of the *B-Bolivia* upstream sequence proximal to the start of transcription and performed transient expression assays. Our results discussed below indicate that the sequences adjacent to the promoter can contribute to the variable penetrance of seed expression and to the reduced amounts of pigment, but they do not contribute to female-specific expression characteristic of the native allele.

In seven independent transgenic lines verified to have a single transgene locus, the proportion of colored kernels was significantly less than 50%, indicating that in these lines, the penetrance of expression was incomplete. In at least two of these lines, color-

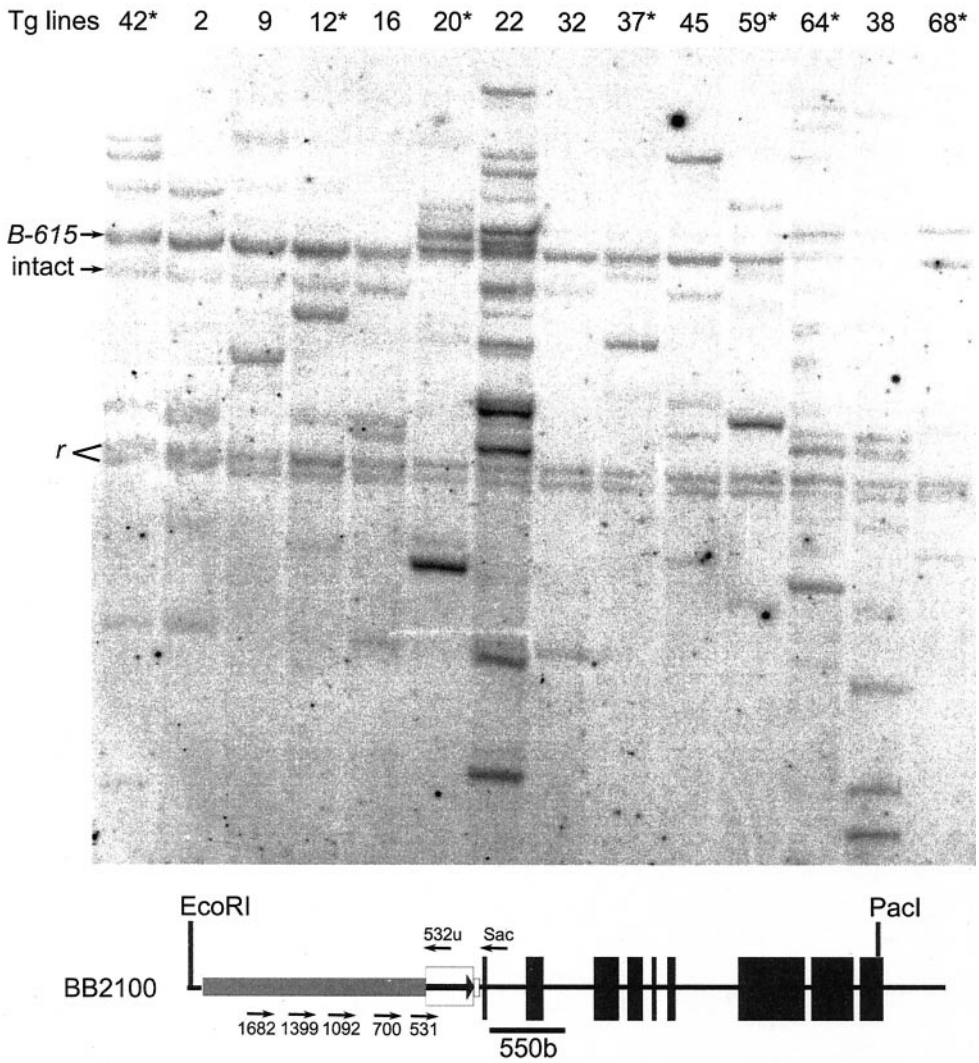

Figure 9. DNA gel-blot analysis of the 14 BB2100 transgenic lines that produced seed pigmentation. Genomic DNA from T_1 transgenic plants was prepared and digested with *Pac*I and *Eco*RI. The map below the blot indicates where these sites are within the construct that was introduced. After electrophoresis and blotting, the resulting blot was hybridized to labeled 550b probe, washed, and hybridization detected using a Molecular Dynamics Storm 2000 system. The endogenous homozygous *B-615* allele produces a characteristic band labeled "B-615" in all lanes. Two bands representing the two *r1* alleles in the T_1 heterozygotes are indicated by arrows labeled "r." The expected size for an intact *Eco*RI/*Pac*I digested BB2100 transgene insertion is indicated by the arrow labeled "intact." The lanes are labeled according to the "VLC" number of the independent transgenic lines. The asterisks indicate the lines with vegetative plant pigmentation. The primers used to determine the amount of *B-Bolivia* upstream DNA in two transgenic lines are indicated on the map.

less seeds that carry the transgene produce plants that express the transgene in progeny seed. This result is similar to the behavior of the native *B-Bolivia* allele, in that colorless seeds that carry the *B-Bolivia* allele produce plants with indistinguishable expression patterns and intensities in the next generation when compared with plants from colored seeds. However, the proportion of pigmented seeds in two of the nine independent transgenic lines verified to have a single transgene locus was approximately 50%, indicating that in these lines 100% of the seeds that received a transgene copy from the hemizygous parent expressed pigment. This proportion was significantly higher than the proportion of colored kernels produced by either of the stocks that carry the native allele. Furthermore, the copy number of the transgene showed no correlation with the penetrance of seed expression. One possibility is that there is a specific region within the *B-Bolivia* upstream sequences that confers variability and the presence or absence of this sequence is different in lines showing 100% penetrance versus lines showing reduced penetrance. It is unfortunate that the number of copies and the complex nature of the transgene loci make this difficult to rigorously test.

Results from transient expression experiments, suggest that sequences within the putative retrotransposon insertion directly affect the amount of aleurone pigment. Expression is reduced to 40% when a sequence normally located in the 33 bp of the insertion proximal to the *B-Peru* homologous sequences is placed immediately upstream of the −176 *B-Peru* promoter, a similar reduction to that seen in the comparison of the 2.1 kb *B-Bolivia* upstream region with the 176-bp *B-Peru* upstream region. However, the promoter proximal region of *B-Bolivia* also lacks sequences homologous to those between −600 and −530 in *B-Peru*. The observation that the presence of the −600 to −559 region reduces the negative effect of the two mutations (E1 and E2), suggests that the −600 to −559 element(s) and the −120 to −84 elements interact to regulate the expression of the aleurone-specific promoter. Thus, we hypothesize that the retrotransposon insertion in *B-Bolivia* at −530 displaces important elements in the aleurone-specific promoter. Transient expression studies and transgenic analyses indicate these elements are not essential, but we suspect their absence in the transgenes or displacement in the native allele contributes to the variability of expression. An intriguing idea is that these sequences serve as a boundary element preventing other regulatory sequences from influencing the promoter. The absence of the −600 to −559 sequences in *B-Bolivia* may allow the retrotransposon insertion in the native allele, and possibly other sequences in the transgenic lines, to have a greater influence on expression.

In addition to the incomplete penetrance and reduced level of aleurone pigmentation in *B-Bolivia*, this allele also shows a parent of origin effect on seed expression. All of the transgenic lines that expressed pigment in the seed, did so when transmitted through the male or female gametes. Thus the sequences immediately upstream of the aleurone-specific sequences are not sufficient to impart the parent of origin effect on expression when the gene is in ectopic locations.

There are two possible ways to explain the lack of pigment in kernels that inherit *B-Bolivia* from the male parent. The first explanation is that *B-Bolivia* expression is sensitive to dosage, as the female contributes two doses and the male one dose to the triploid endosperm, which gives rise to the aleurone. The second explanation is that *B-Bolivia* is epigenetically imprinted such that it is completely silenced when transmitted through the male gametes, whereas, when transmitted through the female gamete, *B-Bolivia* is not silenced in all kernels. Although, due to technical problems, we have not been able to directly test the dosage model, there are two observations that suggest that *B-Bolivia* is epigenetically imprinted. The first is that there is no difference in the proportion of pigmented kernels or the amount of pigment between seeds with two and three copies of *B-Bolivia* in the endosperm. The second is that we have isolated a variant of *B-Bolivia* that does show seed pigment when transmitted through the male parent (DA Selinger, VL Chandler, article in preparation). The implication that *B-Bolivia* is an imprinted allele is particularly interesting in that there are only a few well characterized imprinting systems in plants (for review, see Alleman and Doctor, 2000). In one of these systems, at the *r1* locus in maize (Kermicle, 1978), certain alleles, such as the *R-r:std* allele induce uniform pigmentation of kernels when passed through the female gametes, but induce a weaker, mottled expression when passed through the male gametes. At the Arabidopsis *medea* locus, expression in certain ecotypes is solely from the maternally transmitted allele in the endosperm and embryo tissues of the developing seed (Kinoshita et al., 1999; Vielle-Calzada et al., 1999).

In addition to the differences in aleurone expression between *B-Peru* and *B-Bolivia*, the two alleles have strikingly different patterns and levels of plant pigmentation. It is interesting that another allele of *b1*, *B-Gua31*, isolated from an exotic land race (Negro de Chimaltenango) collected in Guatemala, has a strikingly similar pattern of plant pigmentation to that of the native *B-Bolivia* allele. *B-Gua31* lacks both the aleurone-specific promoter sequences of *B-Peru* and *B-Bolivia* and the putative retrotransposon sequences of *B-Bolivia* (Selinger and Chandler, 1999). Phylogenetic analyses of upstream sequences that are shared by all maize *b1* alleles indicate that *B-Peru*, *B-Bolivia*, and *B-Gua31* are closely related. However, the *B-Gua31* and *B-Bolivia* alleles are the only members of the clade with strong vegetative plant pigmentation phenotypes (Selinger and Chandler, 1999). One inference from this observation is that the putative retrotransposon insertion does not carry the plant-specific regulatory elements that produce the plant expression seen in *B-Bolivia*, but rather, plant expression is due to sequences outside of the retrotransposon region that differentiates *B-Bolivia* from *B-Peru*. These sequences are presumably shared by the phenotypically similar *B-Bolivia* and *B-Gua31* alleles.

Although the plant expression produced by the six BB2100 transgenic lines is very similar between the independent transgenic lines, it is quite different from the plant expression of the native *B-Bolivia* allele. These differences in plant phenotype suggest that if the sequences in the *B-Bolivia* upstream region are producing plant expression, they are not behaving as they normally do in the native allele. Alternatively, another sequence in the transgene array may be contributing to plant pigmentation.

Because these transgenic lines were produced by particle gun bombardment, all of the lines have multiple copies of the BB2100 construct along with the CaMV 35S promoter:*bar* gene construct, which serves as the selectable marker. An intriguing hypothesis is

that the enhancers that are part of the CaMV 35S promoter of the selectable marker construct are interacting with the BB2100 transgene to induce plant expression. The same selectable marker was used in the generation of nine *B-Peru* transgenic lines, all of which showed no plant expression. However, in the *B-Peru* transgenic lines, the potential insulator function of the −600 to −559 sequences in the aleurone-specific promoter may have prevented this sort of interaction with the CaMV 35S enhancers. An alternative explanation is that the highly repetitive nature of the *B-Bolivia* retrotransposon sequence in BB2100 construct may produce very different interactions between some of the transgene loci and the maize genome, some of which may result in plant expression.

As a model for the evolution of novel expression patterns, *B-Bolivia* reveals that the insertion in the upstream regulatory region of a high copy number element can change the expression pattern of a gene. Unlike the insertion in *B-Peru* that produces aleurone pigmentation, the insertion in *B-Bolivia* does not appear to be carrying promoter elements that have been translocated from another gene. Instead it appears that insertion of this large, extremely high copy sequence has altered the expression of the aleurone-specific sequences. Although these retrotransposons have achieved extremely high copy numbers in an evolutionarily short time (SanMiguel et al., 1998), there is no evidence that any of them are still active, and they are very rarely found immediately next to genes (SanMiguel et al., 1996). In contributing to the reduced and unstable seed expression at *B-Bolivia*, this insertion may illustrate the consequences of having a large highly repetitive element near the promoter proximal and coding regions. Continued study of the native *B-Bolivia* allele and various transgenic lines promises to define the roles played by different sequence and chromatin structures in the control of gene expression in plants.

Changes in the spatial and temporal expression of genes, especially genes encoding regulatory proteins, are likely to contribute to the evolution of new species and morphologies. A few genes have been identified as major factors in conferring the morphological differences between maize and teosinte, its wild relative (Beadle, 1939; Doebley and Stec, 1991, 1993). Recent work on one of these genes, *teosinte branched 1*, *tb1*, suggests that a change in expression is responsible for the difference in the morphological phenotypes produced by the maize and teosinte alleles of this gene (Doebley et al., 1997; Wang et al., 1999). Similar types of changes in the cis-acting regulatory regions of genes have been hypothesized to be responsible for many instances of morphological change during evolution (Doebley and Lukens, 1998). However, what the changes are and the molecular mechanisms that created the phenotypic variation found at *tb1* and many other genes are not well characterized. It will be interesting to determine if DNA sequence polymorphisms or DNA arrangements such as those observed at *b1* are operating.

MATERIALS AND METHODS

Plant Materials

B-Bolivia in the K55 background was obtained from G. Neuffer (University of Missouri, Columbia) and *B-Bolivia* in the GS background from George Sprague, Sr. (the University of Illinois, Champaign/Urbana). Several *b1* testers were used, and all the testers carried recessive, nonfunctional alleles of *b1* and *r1*, and functional, dominant alleles of the *C1* regulatory gene, and the anthocyanin biosynthetic genes. The 414 stock resulted from a cross between the K55 *B-Bolivia* line and a *b1*, *r1*, *pl1-sr* tester line. The 1,470 stock is the result of a cross between the K55 and GS *B-Bolivia* that was then outcrossed to a *b1*, *r1*, *Pl-Rhodes* tester.

Cloning *B-Bolivia*

Genomic DNA was extracted from an immature cob that was homozygous for the *B-Bolivia* allele, digested with *Bam*HI, and fragments of approximately 2.8 kb, based on size markers, were recovered by phenol extraction of the melted gel slices. DNA from the fraction showing the strongest *B-Bolivia*-specific hybridization to the 550b probe (Patterson et al., 1995) was ligated into *Bam*HI digested lambda ZAP-Express arms and packaged using a Stratagene XL-Gold packaging extract and plated (Stratagene, La Jolla, CA). Plaque lifts and hybridizations were performed using standard techniques (Sambrook et al., 1989) and the 550b probe. Positively hybridizing plaques were picked, purified through a second round of plating, and hybridization, and the pBK phagemid containing the insert was excised from the lambda ZAP vector according to the manufacturer's instructions. The insert was subsequently subcloned into pTZ 18U for sequencing. This sequence along with additional downstream sequence derived from PCR experiments was deposited in GenBank (accession no. AF326577). Additional upstream *B-Bolivia* sequences were obtained by PCR using oligos specific to *B-Peru* sequences (GenBank accession no. AF205801).

Sequence Analysis and Phylogenetic Analysis

The complete sequence of the 2.8-kb *B-Bolivia* clone was used as a query for several searches using the BLASTN, BLASTX, and FASTX programs (Altschul et al., 1990; Pearson et al., 1997). The potential protein sequence identified by these searches was then used for FASTA and TBLASTN searches to identify homologous protein sequences and coding sequences. The FASTA search returned 19 predicted proteins with an E-value less than 0.01. A TBLASTN search with the same polypeptide sequence against the nucleotide database yielded 80 hits to nucleotide sequences with E-values less than 0.01. The 19 protein sequences found in the FASTA search with E() values less than 0.01

were aligned using the ClustalW program at the BCM server (http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html). The alignment was imported into the Genedoc program (Nicholas et al., 1997, available at http://www.psc.edu/biomed/genedoc/). The N and C termini of the protein sequences were trimmed to remove sequences that did not align well. The core region that aligned well for all sequences was then formatted for input into the PHYLIP package (Felsenstein, 1993) for phylogenetic analysis. We used the PROTPARS program to produce a tree based on the maximum parsimony method. The tree was bootstrapped to generate confidence levels for all the branches (Felsenstein, 1985). The same alignment that was used for the parsimony run was also put through the PROTDIST program using the "Categories" distance matrix to generate a table of evolutionary distances between the sequences. The distance table and the consensus tree were then put into the FITCH program to generate a tree with branch lengths that reflect evolutionary distance.

Copy Number Determination

We prepared dilutions of unlabeled DNA, corresponding to 1.3 kb of the *B-Bolivia* upstream sequence from the upstream *Bam*HI site to the point of divergence with *B-Peru* at −532, into genomic Petunia DNA, which served as a carrier. Twenty-microliter samples containing 3 μg of Petunia genomic DNA and 10, 1, 0.1, or 0.01 ng of the unlabeled *B-Bolivia* upstream sequence were diluted into 200 μL of 0.5 M NaOH, 1.5 M NaCl denaturation solution and vacuum blotted onto a Hybond N^+ membrane (Amersham-Pharmacia Biotech, Uppsala). In addition, 3-μg samples of maize (*Zea mays*) genomic DNA from a *B-Bolivia* line and 2- to 10-μg samples of genomic DNA from other maize lines were similarly treated and blotted. The blot was then hybridized with a [^{32}P]-labeled probe that was the same as the unlabeled cloned *B-Bolivia* upstream sequence used to spike the petunia DNA samples. After washing the blot, hybridization was quantified using a Storm 2600 PhosphorImager. We used the following calculations to determine copy number based on hybridization intensity. Ten nanograms of 1.3-kb DNA equals 7.02 × 10E-9 copies of the sequence. Multiplying this number by 6, which is the difference in intensity between the 10-ng probe signal and the 3 μg of genomic DNA, and then dividing by the number of maize genomes in 3 μg of DNA (1.09 × 10E6) gives an estimate of approximately 38,000 copies in a single 2,500-megabase genome.

RFLP Mapping of *B-Bolivia* and *B-Peru*

Restriction mapping of the *B-Peru* and *B-Bolivia* alleles was conducted using the same series of probes used to map *B-I* and *B-Peru* (Patterson et al., 1995). The 550b probe hybridizes to the region including intron1, exon2, and part of intron2 (Fig. 2A). The BIu4 probe hybridizes to the region immediately upstream of the start of transcription in *B-I*, the E/G700 probe hybridizes to the region between −1,600 and −950 in *B-I* and the B'v1.6 probe hybridizes to a region that is approximately 8 kb upstream in *B-I* (Fig. 2A). All of these probes hybridize to regions tightly linked to the *B-Peru* and *B-Bolivia* alleles. In the case of *B-Peru*, the isolation of recombinants has demonstrated that these sequences are located much further upstream relative to their position in *B-I* (Patterson et al., 1995; M Stam, V Chandler, unpublished data). Band sizes were estimated by performing linear regression on the log of the migration distance of the hybridizing bands compared with those of the known size standards run on the same gel. The E/G700 and B'v1.6 probes, which are located much farther upstream of the BIu4 probe in *B-Peru* than in *B-I* (Fig. 2A; Patterson et al., 1995), had no polymorphic restriction fragments in *B-Bolivia* relative to *B-Peru* when tested with eight restriction enzymes that produce polymorphic fragments in *B-Peru* relative to *B-I*.

Construction of the BB2100 Transgenic Lines

The 2.8-kb clone of *B-Bolivia* containing 2.1 kb of upstream sequence was subcloned between the *Bam*HI sites in a 6-kb *Sal*I clone of *B-I* that contained the complete transcribed region and 3′ UTR, replacing the upstream region of the *B-I* clone and the 5′ most part of the transcribed sequence with *B-Bolivia* sequences (Fig. 8A). This construct was introduced into maize plants using biolistic bombardment of immature embryos as previously described (Koziel et al., 1993; Selinger et al., 1998).

Characterization of BB2100 Transgenic Lines

The primary transgenic plants (T_0) regenerated from stably transformed callus were crossed to a *b1* tester line that carries functional alleles of all of the anthocyanin structural genes and the regulatory genes *c1* and *pl1*, which are co-required with a functional *b1* or *r1* allele to generate anthocyanin pigment. The tester line carried nonfunctional alleles of the *b1* and *r1* genes. The T_0 plants, which were in the Ciba-Geigy CG00526 inbred line, contained nonfunctional, recessive alleles of the regulatory genes *c1*, *pl1,* and *r1*, and a *b1* allele that is weakly expressed in the plant, but that does not color the seed. Because of the nonfunctional *c1* and *pl1* alleles in the line used for transformation, we could not assess pigmentation until the T_1 generation. Transgenic lines that displayed seed color were further characterized. DNA samples collected from leaves or immature cobs were used to characterize the copy number and structure of the BB2100 transgene loci in those lines with seed color. We used *Eco*RI, which cuts in the polylinker of the pTZ 18U plasmid backbone and *Pac*I, which cuts 10 bp beyond the stop codon. DNA blots probed with the 550b probe indicated the number of transgene copies, and by comparison with the size of the *Eco*RI/*Pac*I digested BB2100 plasmid, the presence or absence of intact transgene copies. To determine the number of transgene loci, colored and colorless seeds from each line were planted, DNA prepared and the presence of transgene copies determined either by PCR using primers specific for the transgene promoter and first exon region or by DNA

blots probed with the 550b probe. Lines were considered single locus if there was no segregation of the transgene bands in purple seeds, and/or no colorless seeds carried transgene bands (minimum 7 samples, which equals a 99% confidence that a 50% probability event has not occurred). In some lines that had less than 50% colored kernels, we did see transgenic bands in the colorless seeds, and these were scored as single locus if the individual bands did not segregate when compared with purple seeds and if the frequency of transgene positive, colorless kernels was close to that expected from the frequency of colored kernels.

Transient Transformation Assays for Aleurone Expression

Transient transformation of aleurone cells and luciferase expression assays were performed as previously described (Selinger et al., 1998). Briefly, expression constructs with *B-Bolivia* promoter fragments driving firefly luciferase expression were introduced by biolistic methods into aleurone tissue along with a CaMV 35S promoter driven transformation control. The control plasmid was expressing either GUS (Sainz et al., 1997) or Renilla luciferase (Lorenz et al., 1991; Selinger et al., 1998). Expression was determined by normalizing the luciferase values to the transformation control value and then dividing this number by the normalized luciferase value of the 2.5-kb *B-Peru* promoter construct to generate a percent expression value. The *B-Bolivia*:luciferase expression constructs were produced by subcloning the *Bam*HI to *Sna*BI fragment (−2,100 to +2, relative to the start of transcription) from the 2.8-kb *Bam*HI *B-Bolivia* clone into *Bam*HI/*Sna*BI digested pABPluc plasmid, which was previously described (Selinger et al., 1998). This subcloning step replaced the *B-Peru* promoter sequences in pABPluc with the *B-Bolivia* upstream sequence, which was located just upstream of the *adh1* intron 1 sequence and the firefly luciferase cDNA. Internal *Xho*I sites were used to generate deletion derivatives pBB1400luc, pBBdZluc, and pBB564luc. To create the fusions with the −176 *B-Peru* promoter, oligonucleotides containing the sequence of *B-Bolivia* between −564 and −530 were synthesized with engineered *Bam*HI compatible 5′ overhangs (Marshall University DNA core facility, Huntington, WV). The oligonucleotides were phosphorylated, annealed, and ligated into the *Bam*HI site of the BP176luc plasmid (Selinger et al., 1998). BP600 constructs were constructed by using a −600 *B-Peru*-specific oligo with an engineered 5′-*Bam*HI site together with the BP120A or BP96A oligos (Selinger et al., 1998) to PCR from −600 to the mutant site. The engineered *Xho*I site of the linker-scan mutants together with the *Bam*HI site in the polylinker were used to subclone the −600 to −120 or −96 fragment into the corresponding −176 *B-Peru* LS120 or LS96 luciferase construct. Other than the large deletions that were confirmed by restriction analysis, all the constructs were confirmed by sequencing.

ACKNOWLEGEMENTS

We are grateful to G. Neuffer and G.S. Sprague, Sr. for providing *B-Bolivia* stocks. We thank Susan Belcher for help producing and maintaining the transgenic lines, Catherine Clay for help with DNA preparations and Southern blots, and Lyudmila Sidorenko and Teresa Lavin for comments on the manuscript.

Received November 20, 2000; accepted December 21, 2000.

LITERATURE CITED

Alleman M, Doctor J (2000) Genomic imprinting in plants: observations and evolutionary implications. Plant Mol Biol **43:** 147–161

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Beadle GW (1939) Teosinte and the origin of maize. J Heredity **30:** 245–247

Bennetzen JL, Schrick K, Springer PS, Brown WE, SanMiguel P (1994) Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. Genome **37:** 565–576

Coe EH Jr (1979) Specification of the anthocyanin biosynthetic function by *B* and *R* in maize. Maydica **XXIV:** 49–58

Doebley J, Lukens L (1998) Transcriptional regulators and the evolution of plant form. Plant Cell **10:** 1075–1082

Doebley J, Stec A (1991) Genetic analysis of the morphological differences between maize and teosinte. Genetics **129:** 285–295

Doebley J, Stec A (1993) Inheritance of the morphological differences between maize and teosinte: comparison of results for two F_2 populations. Genetics **134:** 559–570

Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature **386:** 485–488

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783–791

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle

Goff SA, Klein TM, Roth BA, Fromm ME, Cone KC, Radicella JP, Chandler VL (1990) Transactivation of anthocyanin biosynthetic genes following transfer of *b* regulatory genes into maize tissues. EMBO J **9:** 2517–2522

Harris LJ, Currie K, Chandler VL (1994) Large tandem duplication associated with a *Mu2* insertion in *Zea mays B-Peru* gene. Plant Mol Biol **25:** 817–828

Hu W, Das OP, Messing J (1995) *Zeon-1*, a member of a new maize retrotransposon family. Mol Gen Genet **248:** 471–480

Johns MA, Mottinger J, Freeling M (1985) A low copy number copia-like transposon in maize. EMBO J **4:** 1093–1102

Jorgensen RA (1995) Cosuppression, flower color patterns, and metastable gene expression states. Science **268:** 686–691

Kermicle JL (1970) Dependence of the R-mottled aleurone phenotype in maize on mode of sexual transmission. Genetics **66:** 69–85

Kermicle JL (1978) Imprinting of gene action in maize endosperm. *In* DB Walden, ed, Maize Breeding and Genetics. John Wiley & Sons, New York, pp 357–371

Kinoshita T, Yadegari R, Harada JJ, Goldberg RB, Fischer RL (1999) Imprinting of the *MEDEA* polycomb gene in the *Arabidopsis* endosperm. Plant Cell **11:** 1945–1952

Koziel MG, Beland GL, Bowman C, Carozzi NB, Crenshaw R, Crossland L, Dawson J, Desai N, Hill M, Kadwell S et al. (1993) Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. Biotechnology **11:** 194–200

Lorenz WW, McCann RO, Longiaru M, Cormier MJ (1991) Isolation and expression of a cDNA encoding *Renilla reniformis* luciferase. Proc Natl Acad Sci USA **88:** 4438–4442

Ludwig SR, Bowen B, Beach L, Wessler SR (1990) A regulatory gene as a novel visible marker for maize transformation. Science **247:** 449–450

Ludwig SR, Habera SL, Dellaporta SL, Wessler SR (1989) *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region. Proc Natl Acad Sci USA **86:** 7092–7096

Nicholas KB, Nicholas HB Jr, Deerfield DW II (1997) GeneDoc: analysis and visualization of genetic variation. EMBNEW.NEWS **4:** 14

Patterson GI, Kubo KM, Shroyer T, Chandler VL (1995) Sequences required for paramutation of the maize *b* gene map to a region containing the promoter and upstream sequences. Genetics **140:** 1389–1406

Pearson WR, Wood T, Zhang Z, Miller W (1997) Comparison of DNA sequences with protein sequences. Genomics **46:** 24–36

Radicella JP, Brown D, Tolar LA, Chandler VL (1992) Allelic diversity of the maize *b* regulatory gene: different leader and promoter sequences of two *b* alleles determine distinct tissue specificities of anthocyanin production. Genes Dev **6:** 2152–2164

Robbins TP, Walker EL, Kermicle JL, Alleman M, Dellaporta SL (1991) Meiotic instability of the *R-r* complex arising from displaced intragenic exchange and intrachromosomal rearrangement. Genetics **129:** 271–283

Sainz MB, Goff SA, Chandler VL (1997) Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation *in vivo*. Mol Cell Biol **17:** 115–122

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 2.60–2.112

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet **20:** 43–45

SanMiguel P, Tikhonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z et al. (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science **274:** 765–768

Selinger DA, Chandler VL (1999) Major recent and independent changes in the levels and patterns of expression have occurred at the *b* gene, a regulatory locus in maize. Proc Natl Acad Sci USA **96:** 15007–15012

Selinger DA, Lisch D, Chandler VL (1998) The maize regulatory gene, *B-Peru*, contains a DNA rearrangement that specifies tissue-specific expression through both positive and negative promoter elements. Genetics **149:** 1125–1138

Stadler LJ, Neuffer M (1953) Problems of *R* gene structure: II. Separation of *R-r* elements (P) and (S) by unequal crossing over. Science **117:** 471–472

Styles ED, Ceska O, Seah KT (1973) Developmental differences in action of *r* and *b* alleles in maize. Can J Genet Cytol **15:** 59–72

Varagona MJ, Purugganan M, Wessler SR (1992) Alternative splicing induced by insertion of retrotransposons into the maize *waxy* gene. Plant Cell **4:** 811–820

Vielle-Calzada J-P, Thomas J, Spillane C, Coluccio A, Hoeppner MA, Grossniklaus U (1999) Maintenance of genomic imprinting at the *Arabidopsis medea* locus requires zygotic DDM1 activity. Genes Dev **13:** 2971–2982

Walker EL, Robbins TP, Bureau TE, Kermicle J, Dellaporta SL (1995) Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize *R-r* complex. EMBO J **14:** 2350–2363

Wang R-L, Stec A, Hey J, Lukens L, Doebley J (1999) The limits of selection during maize domestication. Nature **398:** 236–239

White SE, Habera LF, Wessler SR (1994) Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA **91:** 11792–11796