Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 8244–8248, September 1993
Genetics

An upstream open reading frame represses expression of *Lc*, a member of the *R/B* family of maize transcriptional activators

(anthocyanin/helix–loop–helix/translational control/*Zea mays*)

RONALD D. DAMIANI, JR., AND SUSAN R. WESSLER*

Department of Botany, University of Georgia, Athens, GA 30602

Communicated by Charles S. Levings III, June 14, 1993 (received for review March 19, 1993)

ABSTRACT **The *R/B* genes of maize encode a family of basic helix–loop–helix proteins that determine where and when the anthocyanin-pigment pathway will be expressed in the plant. Previous studies showed that allelic diversity among family members reflects differences in gene expression, specifically in transcription initiation. We present evidence that the *R* gene *Lc* is under translational control. We demonstrate that the 235-nt transcript leader of *Lc* represses expression 25- to 30-fold in an *in vivo* assay. Repression is mediated by the presence *in cis* of a 38-codon upstream open reading frame. Furthermore, the coding capacity of the upstream open reading frame influences the magnitude of repression. It is proposed that translational control does not contribute to tissue specificity but prevents overexpression of the Lc protein. The diversity of promoter and 5′ untranslated leader sequences among the *R/B* genes provides an opportunity to study the coevolution of transcriptional and translational mechanisms of gene regulation.**

Anthocyanin biosynthesis in maize is controlled by two regulatory gene families: the *R/B* family and the *C1/Pl* family. One functional gene from each family is usually required to activate transcription of the structural genes in the pathway, such as *A1*, *Bz1*, and *C2* (1–7). The pigmentation pattern of the plant is determined by the genetic constitution of the *R/B* gene family, comprised of the *R* locus on chromosome 10 and the *B* locus on chromosome 2 (8).

The ease of visually distinguishing subtle differences in plant pigmentation patterns has led to the identification of nearly 100 naturally occurring variants of *R* and *B*. Members of the *R* and *B* family share >80% amino acid identity and contain the basic helix–loop–helix (HLH) DNA-binding and dimerization motif (2, 4, 9, 10). In addition to their structural similarity, transient transformation assays indicate that R and B proteins are functionally equivalent. For example, the *R* gene *Lc* and the *B* alleles *Peru* and *I* condition dramatically different patterns of plant pigmentation. However, constructs containing the Lc or B-Peru or B-I cDNA fused to the constitutive cauliflower mosaic virus (CaMV) 35S promoter induced pigmentation in most cell types after particle bombardment (3, 11). The results from these studies have led to the view that allelic diversity is from differences in gene regulation rather than from protein function. This notion is supported by two observations: (*i*) mRNAs for the *R* genes *Sn*, *R-sc*, and *S* and the *B* allele *I* are detected only in the tissues pigmented by each allele (10, 12, 13) and (*ii*) the 5′ flanking sequences of *R* and *B* genes, including the untranslated leaders, are more heterogeneous than are the coding sequences. The strongest evidence to date that allelic diversity reflects differences in the expression of members of the *R/B* gene family came from the use of transient transformation to show that elements controlling tissue specificity reside in the distinct 5′ flanking regions of the *B-I* and *B-Peru* alleles (12).

Although differences in transcriptional regulation clearly serve to distinguish *R/B* family members, a role for translational control of gene expression has not been demonstrated. Comparison of the 5′ untranslated regions of *R* and *B* genes indicates extensive sequence divergence. Of particular interest is the presence of upstream AUGs in some, but not all, members of the *R/B* family. Upstream AUGs can have dramatic effects on the efficiency of mRNA translation (for review, see refs. 14 and 15). In addition, among the *R* and *B* genes with upstream AUGs, some have upstream open reading frames (uORFs) that overlap the translation initiation site, whereas other alleles have uORFs that terminate before the start of the *R/B* coding regions.

As a first step in determining whether *R/B* allelic diversity results from differences in both transcriptional and translational mechanisms we have focused on the complex 5′ untranslated leader of the *R* gene *Lc* (16). The *Lc* gene was the first complete *R/B* family member to be cloned and sequenced (2). Mapping of the transcription start site revealed that the Lc mRNA has a 5′ leader region of 235 nt containing three upstream AUG codons that are all part of a 38-codon uORF. We have used site-directed mutagenesis in conjunction with *in vivo* particle bombardment to analyze the role of this uORF in regulating gene expression. We show that the uORF substantially represses *Lc* expression, that it functions only *in cis*, and that the coding capacity of the uORF is involved in establishing the level of repression.

MATERIALS AND METHODS

Plant Material. The maize inbred line W22 (*r-g*, *A1*, *A2*, *Bz1*, *Bz2*, *C1*, *C2*, *pl*, *B-b*), provided by J. Kermicle (University of Wisconsin, Madison) was used for all transient transformation studies.

RNA Blot Analysis. Total RNA was isolated as described (2) from mature kernels 3 hr after bombardment with either pLcWT, pLcΔ, pLcm123, or mock treated (bombarded without plasmid DNA). Slot blot analysis was done according to the manufacturer by using a Minifold II apparatus (Schleicher & Schuell). Nitrocellulose filters were prehybridized, hybridized, and washed at high stringency (65°C) as described (2). DNA probes were labeled by the random primer method (17). The *Lc* probe used was an 886-bp *Sst* I fragment of the Lc cDNA (nt 884–1770) (2). The chloramphenicol acetyltransferase (CAT) probe consisted of a 563-bp *Nco* I–*Xba* I fragment of the CAT coding region isolated from pCATBasic (Promega).

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

Abbreviations: uORF, upstream open reading frame; CAT, chloramphenicol acetyltransferase; CaMV, cauliflower mosaic virus; HLH, helix–loop–helix.

*To whom reprint requests should be addressed.

Secondary Structure Analysis. Secondary structures within the Lc mRNA (nt 1–300) were analyzed with the FOLD program as part of the Genetics Computer Group sequence analysis package for the VAX (18).

Site-Directed Mutagenesis. Site-directed mutagenesis was done by using the Altered Sites kit (Promega), according to the manufacturer, using a 175-bp *Eco*RI–*Xba* I fragment containing the 5′ end of the Lc cDNA (nt 20–195) subcloned into vector pSelect-1. Transformants were screened by both restriction digestion and dideoxynucleotide sequencing (19, 20). For constructs containing more than one mutation multiple oligonucleotides were used. Mutagenic oligonucleotides used are shown below after the plasmid name with nucleotide positions in parentheses: pLcm1, 5′-ACCCTTCGCCTG-GAAGTTC-3′ (nt 51–69); pLcm2, 5′-GCATACGCAGGC-CCCTCGT-3′ (nt 118–136); pLcm3, 5′-ATCGTCTGAG-GATCAGTAG-3′ (nt 154–172); pLcmStop, 5′-GTTCTTG-CATAGCTTCGTTGC-3′ (nt 66–86); pLcmL4→I, 5′-ATGGAAGTTATTGCATTGC-3′ (nt 60–78); pLcmT30→S, 5′-GCCGGTTCTCGAGGCATCG-3′ (nt 139–157); pLcm-Syn, 5′-CACGAGGCACCGTCTGATG-3′ (nt 146–164).

Plasmid Constructions. All plasmids were constructed by using standard techniques and purified by CsCl equilibrium centrifugation (19). The Lc cDNA is numbered according to Ludwig *et al.* (2). pLcWT contains a 2.4-kb Lc cDNA fragment (nt 20–2398) cloned into the *Xba* I site of vector pDH51 (21). pLcΔ consists of a 2.2-kb *Xba* I Lc cDNA fragment (nt 195–2398) cloned into pDH51. pLcm1 is like pLcWT (see Fig. 2) except that an A → C mutation at nt 60 of the cDNA has been introduced. pLcm1 was constructed by liberating the mutagenized *Lc* fragment (see above) from pSelect-1 and directionally subcloning into pLcΔ digested with *Sma* I and *Xba* I. This placed the mutagenized fragment in its correct position within the Lc cDNA. Constructs pLcm2, pLcm3, pLcm23, pLcm123, pLcmStop, pLcmL4→R, pLcmL4→I, pLcmT30→S, and pLcmSyn were all constructed as pLcm1. pLcm2 contains a U → G substitution at nt 127. pLcm3 contains a U → G substitution at nt 161. pLcm23 contains both mutations found in pLcm2 and pLcm3. pLcm123 contains the three mutations described for pLcm1, pLcm2, and pLcm3. In pLcmStop, a U → A change was made at nt 76, creating a nonsense codon at position six of the uORF. pLcmL4→R was derived from a fortuitous mutation (UU → GG) at nt 70 and 71, altering a leucine codon to arginine. pLcmL4→I also alters codon 4 with a C → A substitution at nt 69. pLcmT30→S has an A → U substitution at nt 147, changing Thr-30 to serine. pLcmSyn has a U → C mutation at nt 155. pPEP was derived from pLcWT (above) and consists of the 5′ end of the Lc cDNA (nt 20–195) cloned into pDH51. pPEPm1 was created as pPEP but was derived from pLcm1.

High-Velocity Particle Bombardment. Plasmid DNA was delivered to sterile maize aleurone tissue by high-velocity microprojectiles using the Biolistic PDS-1000 (DuPont). Microprojectiles were prepared by precipitating a total of 5.005 μg of plasmid DNA [3 μg of pBz1LUC (1), 2 μg of pAdhCAT (previously called pA11CN in ref. 22), and 5 ng of *Lc* expression vector] onto 1.0-μm gold particles (60 mg/ml) as described (1, 3). Microprojectiles prepared for cobombardments with pPEP or pPEPm1 contained 4 μg of either of these plasmids in addition to the plasmids described above, except that 1.25 ng of the *Lc* vector (pLcm123) was used. Half kernels were incubated on MS medium (Sigma M5519) for 46–48 hr at 28°C with illumination and then harvested for enzyme assays.

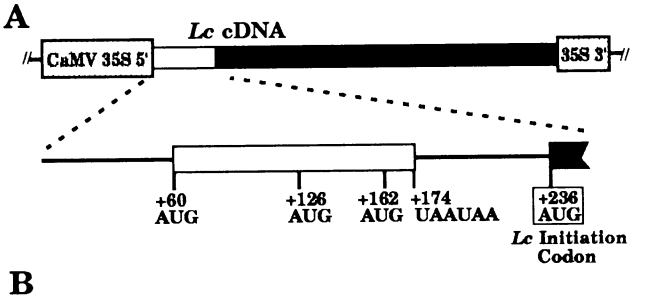
Enzyme Assays. Luciferase and CAT activities of bombarded half kernels were measured after 46–48 hr, as described (1, 22, 23). Luciferase activity was measured with a model 3010 luminometer (Analytic Scientific Instruments, Alameda, CA) and is expressed as the number of light units detected in 10 sec per 100 μl of extract at 25°C. CAT activity was measured according to Sleigh (23) and is expressed as ethyl acetate-soluble cpm for 25 μl of extract in 1 hr at 37°C.

RESULTS

The 5′ Leader Represses Lc Expression. The 235-nt *Lc* leader contains three AUGs at positions +60, +126, and +162. All three AUGs are part of an uORF that terminates with two UAA stop codons at +174 (Fig. 1). If translated, the uORF would encode a 38-amino acid peptide (Fig. 1*B*). To determine whether this region regulated *Lc* expression we used high-velocity particle bombardment to transiently transform maize aleurone tissue with plasmids expressing *Lc* with variations in the *Lc* leader. To measure *Lc* expression and correct for variation between transfection experiments, three plasmids were simultaneously delivered to aleurone cells: an *Lc* expression vector under constitutive transcriptional control of the CaMV 35S promoter; a reporter plasmid, pBz1LUC, containing the firefly luciferase coding region fused to *Bz1* 5′ and 3′ sequences (1); and a delivery control plasmid, pAdhCAT (22), consisting of the CAT coding region under control of the maize *Adh1* promoter. Unlike the *Bz1* promoter, the *Adh1* promoter is not *Lc*-responsive, so *Lc* expression is measured indirectly as a luciferase/CAT activity ratio (1).

Removal of the upstream AUGs by deletion of the terminal 175 nt of the leader results in a 22-fold increase in reporter gene activity over the wild-type control (Fig. 2*A*, compare pLcΔ and pLcWT). Elimination of all three upstream AUGs by site-directed mutations led to a 30-fold increase in expression (Fig. 2*A*, compare pLcm123 and pLcWT). That mutation of the AUGs led to more reporter gene activity than their deletion suggests that lengthening the 5′ leader region increases expression in maize.

In addition to inducing luciferase expression, bombardment of *Lc*-containing constructs activates the anthocyanin pathway and turns aleurone cells red. If red spots are counted before assaying kernels for luciferase/CAT ratios, 20- to



B

```
1   TAGGCGTTCGATCCCCTTAGCGCGGAGGAGAGCTCCTCCGGTTCTTCTCTACCCTTCGC

60  ATG GAA GTT CTT GCA TTG CTT CGT TGC TTC TCT AGT TTC TTC CTT
      M  E  V  L  A  L  L  R  C  F  S  S  F  F  L

105 CTA CGT CTT TCC AGC ATA CGC ATG CCC CTC GTC CGC CGG TTC ACG
      L  R  L  S  S  I  R  M  P  L  V  R  R  F  T

150 AGG CAT CGT CTG ATG ATC AGT AGA TAA TAA GCA ATA TAA TAC TGA
      R  H  R  L  M  I  S  R  -  -

195 TCTAGAATCGAGTTGTTGTACTCTTCGCAGATAGGCGCGTG ATG
```

FIG. 1. Features of the *Lc* 5′ leader region and diagram of the expression construct. (*A*) Wild-type 2.4-kb Lc cDNA or mutant leader constructs were inserted between the CaMV 35S promoter and terminator in vector pDH51 (21). The unfilled box represents the uORF, and the filled box represents the *Lc* coding region. (*B*) Nucleotide sequence of *Lc* 5′ leader region. *Lc* sequences within all expression vectors begin with nt 20 (2). The derived amino acid sequence of the uORF, including the three AUGs (underlined), is shown in single-letter code.

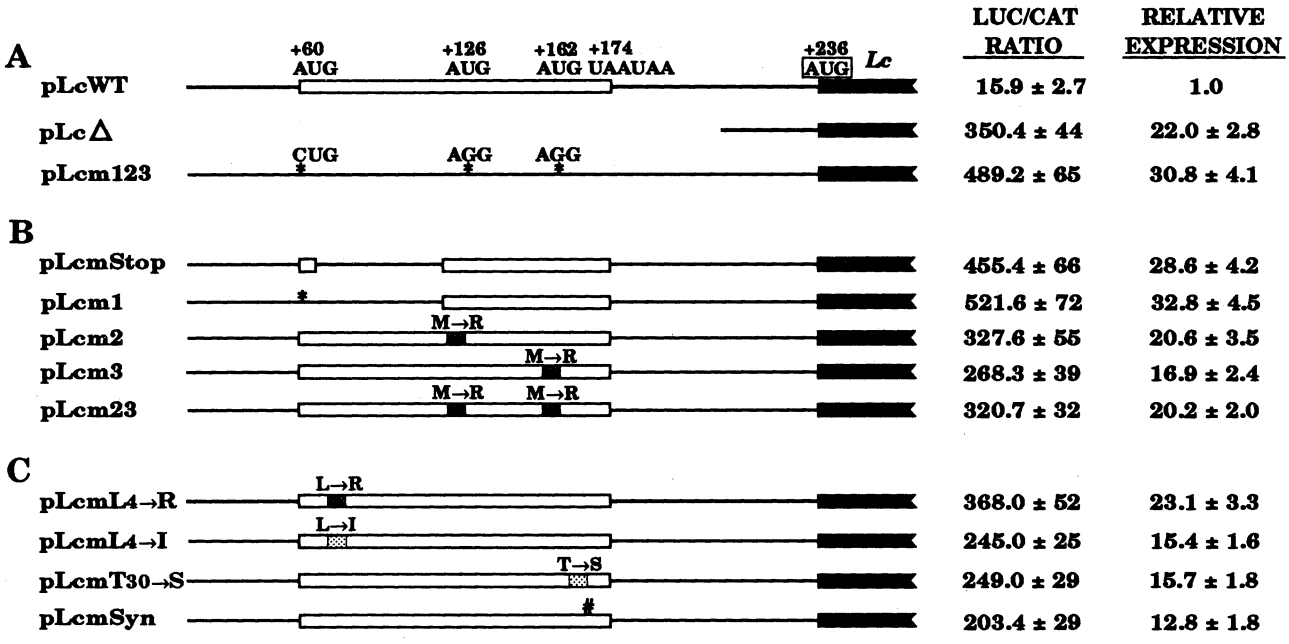


FIG. 2. Schematic representation of the 5′ ends of mRNAs encoded by *Lc* constructs and the relative expression levels obtained after particle bombardment. The uORF and *Lc* reading frames are noted as in Fig. 1. *Lc* expression is presented as a luciferase/CAT ratio (LUC/CAT) as described (1). Values are the average of 14–18 independent bombardments per construct. Relative expression was calculated by dividing the average LUC/CAT ratio for each construct by the LUC/CAT ratio obtained for pLcWT. (*A*) Effect of deletion or mutation of the three upstream AUGs. Mutations are indicated by asterisks with the nucleotide sequence change shown above. (*B*) Effect of eliminating the uORF or altering each AUG individually. The predicted small uORF created by mutating the first AUG (pLcm1) or introduction of a stop codon (pLcmStop) is probably not translated (see *Results*). The amino acid change predicted by each codon change is noted. (*C*) Effects of non-AUG codon substitutions within the uORF. Predicted nonconservative and conservative amino acid changes are indicated by dark and stippled boxes, respectively. Position of the synonymous codon change (#) is noted.

50-fold more spots are seen on kernels bombarded with constructs containing leader deletions or mutations (data not shown). Correlation between the spot and luciferase assays provides additional evidence that repression has a physiological basis.

Repression Is Not Due to mRNA Stability or Secondary Structure. The 5′ leader region, specifically the upstream AUGs, appears to interfere with the efficient translation of the *Lc* reading frame. However, alterations in leader sequences could enhance *Lc* expression by decreasing the rate of Lc mRNA degradation. To address this question, RNA slot blots were used to assess the relative amounts of steady-state Lc mRNA in transfected tissue. Fig. 3 shows that the amounts of steady-state Lc mRNA isolated from kernels bombarded with either pLcWT, pLcΔ, or pLcm123 do not differ significantly when the data are normalized to the levels of CAT mRNA present. In addition, Northern blot analysis

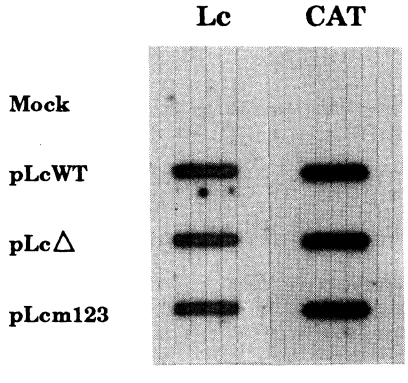


FIG. 3. Slot blots of RNA from bombarded kernels. Seven micrograms of total RNA, isolated 3 hr after bombardment with either pLcWT, pLcΔ, pLcm123, or mock treatment, was probed with Lc cDNA or with the CAT coding sequence.

indicated that no preferential degradation of *Lc* sequences occurs among these RNA samples (data not shown).

It is also possible that the leader region may repress expression of *Lc* by forming secondary structures that impair translation initiation (24). Computer-aided analysis of Lc mRNA (nt 1–300) (see *Materials and Methods*) found that the most stable secondary structure has a $\Delta G^0 = -18$ kcal/mol. On the basis of mammalian studies, such weak interactions do not present a barrier to scanning 40S ribosomal subunits (24). In addition, secondary structure effects are unlikely to be responsible for repression because point mutations in the *Lc* leader derepress *Lc* expression to the same extent as deletion of most of the leader (see Fig. 2).

The uORF Is the Negative Element in Lc Repression. From the data presented above, it is unclear whether repression is due to initiation at the first AUG and subsequent translation of the uORF or if translation initiates at each of the three upstream AUGs. Comparison of the sequence context surrounding these AUGs with either the mammalian (25, 26) or plant (27) consensus for initiation codons suggests that translation might initiate at the first AUG but is unlikely to initiate at the second or third AUG because of their poor sequence context (Fig. 4).

Analysis of additional mutant constructs indicates that repression is mediated by initiation at the first AUG and translation of the uORF (Fig. 2 *B* and *C*). Introduction of a premature stop codon (pLcmStop) led to fully derepressed levels of expression, despite the presence of all three upstream AUGs. Similar results were obtained when a frameshift mutation was introduced into the uORF (data not shown). Furthermore, elimination of the first AUG (pLcm1) led to fully derepressed levels, even though the second and third AUGs were intact.

Involvement of uORF Codons. Repression is apparently not mediated by initiation at the second or third AUG because either elimination of the first AUG or termination of the

A

```
                    -3 +1 +4
                     |  |  |
AUG 1:          CUUCGCAUGGA (nt 54-64)
AUG 2:          AUACGCAUGCC (nt 120-130)
AUG 3:          CGUCUCAUGAU (nt 156-166)
(Lc initiation codon) AUG 4: CGCGUGAUGGC (nt 230-240)
Vertebrate Consensus:  CCPCCAUGG
```

B

```
                    -3 +1 +4
                     |  |  |
AUG 1:          CUUCGCAUGGA (nt 54-64)
AUG 2:          AUACGCAUGCC (nt 120-130)
AUG 3:          CGUCUCAUGAU (nt 156-166)
(Lc initiation codon) AUG 4: CGCGUGAUGGC (nt 230-240)
Plant Consensus:  AACAAUGGC
```

FIG. 4. Sequence context surrounding the first four AUG codons of the Lc mRNA. Comparison with the vertebrate (*A*) and plant (*B*) consensus for translation initiation codons (25–27). The adenine residue of the AUG triplet is designated as +1 after Kozak (25). Matching positions are shown in boldface type.

uORF before the second or third AUG led to fully derepressed expression. For this reason it was surprising that mutation of the second (pLcm2) or third AUG (pLcm3) derepressed expression by 20.6- and 16.9-fold, respectively. No additional derepression was observed when both mutations were included in a single construct (pLcm23).

Because repression appears to be mediated by the uORF and not initiation at AUG2 or AUG3, mutation of these AUGs may derepress expression because of changes in the coding capacity of the uORF. When viewed in this light, these changes introduced nonconservative mutations of methionine to arginine codons. Another nonconservative mutation of position 4 from leucine to arginine (pLcmL4→R) resulted in 23-fold derepression. Interestingly, changing this same codon to isoleucine (pLcmL4→I) creates a conservative substitution that partially restored repression. Another conservative substitution at residue 30 (Thr → Ser in pLcmT30 →S) also resulted in partial repression.

These data indicate that the magnitude of repression is influenced by the uORF codons. However, all mutations presented thus far alter both nucleotide sequence of the uORF and amino acid sequence of the putative uORF peptide. To change the uORF sequence without changing the uORF-peptide sequence, a synonymous codon change (CAU → CAC) was introduced at position 32 (pLcmSyn). Although pLcWT and pLcmSyn encode the same uORF peptides, expression of the latter was 12.8-fold higher. Thus, a synonymous change still derepresses expression relative to wild type but derepresses it to a lesser extent than all other constructs tested.

The uORF Is Required *in cis*. If the uORF encodes a peptide that directly represses *Lc* translation, then it should function *in trans* as well as *in cis*. Alternatively, if the process of translating the uORF represses *Lc* expression, rather than the peptide encoded by uORF, then the uORF will only repress *in cis*. To provide the uORF *in trans*, vectors pPEP and pPEPm1 were constructed to express only wild-type and mutant leader RNA, respectively (Fig. 5*A*). Reporter gene activity was assayed after cobombardment of a large excess of pPEP or pPEPm1 with pLcm123, a construct shown previously to be fully derepressed for *Lc* expression because of point mutations in the leader AUGs. Fig. 5*B* shows that reporter gene activity does not significantly differ between cobombardments with pPEP or pPEPm1, implying that repression does not work when the uORF is provided *in trans*. Similar results were obtained in a series of experiments varying both the concentration of pLcm123 and the concentration of either pPEP, pPEPm1, or pDH51 (data not shown).

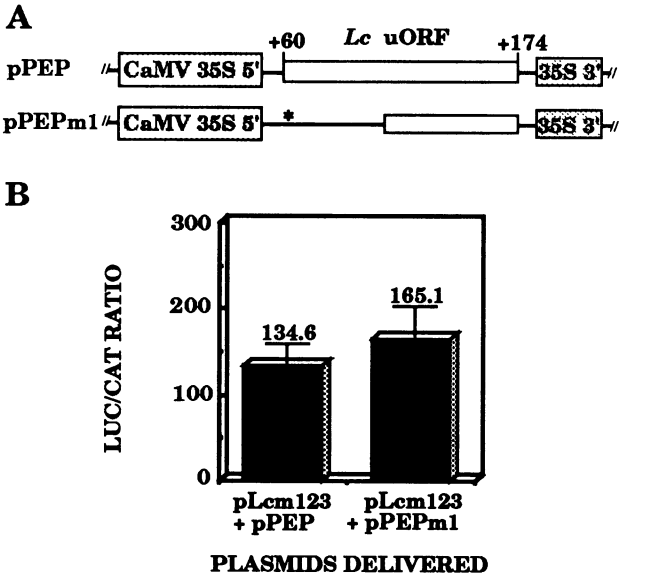


FIG. 5. Effect of uORF provided *in trans*. (*A*) Constructs designed to express wild-type (pPEP) or mutant (pPEPm1, derived from pLcm1) *Lc* leader from a separate mRNA. (*B*) Results of cobombardment of a large excess of each construct with pLcm123. The average of 12 bombardments is displayed graphically.

DISCUSSION

We have demonstrated that the 235-nt leader of the *Lc* gene represses *Lc* expression 25- to 30-fold in an *in vivo* assay. Our data indicate that the leader does not destabilize *Lc* mRNA nor does it assume a secondary structure that inhibits translation initiation of Lc. Rather, repression is apparently mediated by the presence *in cis* of a 38-codon uORF. Furthermore, the coding capacity of the uORF influences the magnitude of repression.

Although uORFs have been associated with modulating downstream gene expression, it is unusual for the codons of an uORF to be involved in the repression mechanism (28–30). The effect of uORFs on gene expression, without concern for the coding capacity of the uORF, has been examined in mammalian cells (31–33). Kozak (31) demonstrated that increasing the intercistronic distance enhanced the level of translation reinitiation at the downstream reading frame. An intercistronic distance of 79 nt permitted 50–100% of the expression obtained with constructs lacking an uORF. Decreasing the intercistronic distance to 45 nt and 11 nt still allowed 20% and 10%, respectively, of control expression. In comparison, wild-type Lc mRNA has an intercistronic distance of 62 nt, which should allow, based on comparison with the mammalian study, 20–50% of the expression levels seen with no uORF. Instead, constructs with the wild-type uORF express only 3% of the activity of constructs that lack the uORF. We propose that it is the coding capacity of the *Lc* uORF that is responsible for the additional repression seen.

There are at least two ways that the coding capacity of the uORF could repress translation of the downstream *Lc* reading frame. The uORF may encode a peptide that decreases translation of *Lc* (the peptide repressor model). Alternatively, the process of translating the uORF may retard the flow of ribosomes to the *Lc* reading frame (the ribosome stall model) (33).

The peptide repressor model alone appears insufficient to describe all data for the following reasons. Cobombardment experiments suggest that the uORF only represses *in cis*. More significantly, a synonymous codon change within the uORF (pLcmSyn) results in 13-fold more expression than in wild type. That this mutation alleviates much of the repres-

sion indicates that the codon usage within the uORF may be important. pLcmSyn contains the most common histidine codon used in maize, CAC, whereas the wild-type uORF contains the least preferred histidine codon, CAU. Ribosomes that may normally stall at the rare, wild-type CAU codon might not stall at the preferred, mutant codon CAC. It should be noted that only 10 of the 38 codons of the uORF are the most preferred maize codons, whereas 6 are the least preferred (34). In addition, multiple codons are used for most amino acids: five of the six codons for arginine and leucine are present, and four of the six serine codons are used. Although such codon usage seems more consistent with an uORF that serves to stall ribosomes rather than one that encodes a functional peptide, the peptide model cannot be dismissed because (*i*) derepression is observed with several constructs (Fig. 2 *B* and *C*) that do not exchange rare codons for preferred codons but, rather, change the composition of the putative peptide and (*ii*) the cobombardment experiments (Fig. 5) do not address the possibility that the putative uORF peptide acts locally to repress translation.

Why should such a sophisticated mechanism of translational control regulate the expression of a dispensable gene? It is unlikely that the uORF contributes to the distinct spatial pattern of *Lc* expression because another *R* gene, *R-sc*, which conditions a dramatically different pigmentation pattern, encodes a virtually identical 5′ leader (M. Alleman and J. Kermicle, personal communication). However, because the 5′ leader regions of other *R* and *B* genes are diverse, a role for translational control in determining spatial patterns of gene expression cannot be ruled out. A more likely explanation for translational control is that repression of *Lc* prevents overexpression of a basic-HLH transcriptional activator. HLH proteins can form heterodimers with other HLH proteins (35). Although *Lc* is dispensable, its overexpression might lead to deleterious interactions with essential HLH proteins involved in growth and development. Experiments with transgenic plants containing *Lc* constructs suggest that overexpression of *Lc* can alter developmental patterns. *Arabidopsis* plants transformed with the maize Lc cDNA lacking the uORF (like pLcΔ) displayed distinctive patterns of trichome development when compared with transgenic plants containing the full-length Lc cDNA (like pLcWT) (36). Moreover, mature Petunia plants containing pLcΔ could not be isolated because all transformants died as purple seedlings (F. Quattrocchio and J. Mol, personal communication).

The 38-codon uORF does not have a role in preventing the overexpression of all members of the *R/B* family because some genes, such as *B-Peru*, have no upstream AUGs and others, including *S* and *B-I*, have uORFs distinct from the *Lc* uORF. It is conceivable that for each allele, a balance has evolved between control of transcription initiation and the efficiency of mRNA translation. In this regard, the diversity of promoter and leader sequences associated with the numerous members of the *R/B* gene family provides an excellent opportunity to examine the coevolution of transcriptional and posttranscriptional mechanisms of gene expression.

We thank Steve Goff for providing plasmid pAdhCAT and for helpful comments on using the luciferase/CAT reporter system. We also thank Ingo Potrykus and Brad Roth for plasmids pDH51 and pBz1LUC, respectively. We are grateful to S. Ludwig for plasmids pLcWT and pLcΔ and to M. Purugganan and T. Bureau for critical reading of the manuscript. This study was supported by a U.S. Department of Energy grant to S.R.W.

1. Klein, T. M., Roth, B. A. & Fromm, M. E. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6681–6685.
2. Ludwig, S. R., Habera, L. F., Dellaporta, S. L. & Wessler, S. R. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7092–7096.
3. Goff, S. A., Klein, T. M., Roth, B. A., Fromm, M. E., Cone, K. C., Radicella, J. P. & Chandler, V. L. (1990) *EMBO J.* **9,** 2517–2522.
4. Goff, S. A., Cone, K. C. & Chandler, V. L. (1992) *Genes Dev.* **6,** 864–875.
5. Dooner, H. K. (1983) *Mol. Gen. Genet.* **189,** 136–141.
6. Cone, K. C., Burr, F. A. & Burr, B. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9631–9635.
7. Coe, E. H., Jr., Neuffer, M. G. & Hoisington, D. A. (1988) in *Corn and Corn Improvement*, eds. Sprague, G. F. & Dudley, J. W. (Am. Soc. Agronomy, Madison, WI), pp. 81–258.
8. Styles, E. D., Ceska, O. & Seah, K. T. (1973) *Can. J. Genet. Cytol.* **15,** 59–72.
9. Perrot, G. H. & Cone, K. C. (1989) *Nucleic Acids Res.* **17,** 8003.
10. Tonneli, C., Consonni, G., Dolfini, S. F., Dellaporta, S. L., Viotti, A. & Gavazzi, G. (1991) *Mol. Gen. Genet.* **225,** 401–410.
11. Ludwig, S. R., Bowen, B., Beach, L. & Wessler, S. R. (1990) *Science* **247,** 449–450.
12. Radicella, J. P., Brown, D., Tolar, L. A. & Chandler, V. L. (1992) *Genes Dev.* **6,** 2152–2164.
13. Consonni, G., Geuna, F., Gavazzi, G. & Tonelli, C. (1993) *Plant J.* **3,** 335–346.
14. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241.
15. Hinnebusch, A. G. & Klausner, R. D. (1991) in *Translation in Eukaryotes*, ed. Trachsel, H. (CRC, Boca Raton, FL), pp. 243–272.
16. Dooner, H. K. & Kermicle, J. L. (1976) *Genetics* **82,** 309–322.
17. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
18. Devereaux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
21. Pietrzak, M., Shillito, R. D., Hohn, T. & Potrykus, I. (1986) *Nucleic Acids Res.* **14,** 5857–5868.
22. Callis, J., Fromm, M. & Walbot, V. (1987) *Genes Dev.* **1,** 1183–1200.
23. Sleigh, M. J. (1986) *Anal. Biochem.* **156,** 251–256.
24. Kozak, M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2850–2854.
25. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148.
26. Kozak, M. (1986) *Cell* **44,** 283–292.
27. Lutcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. & Scheele, G. A. (1987) *EMBO J.* **6,** 43–48.
28. Werner, M., Feller, A., Messenguy, F. & Pierard, A. (1987) *Cell* **49,** 805–813.
29. Schleiss, M. R., Degnin, C. R. & Geballe, A. (1991) *J. Virol.* **65,** 6782–6789.
30. Hill, J. R. & Morris, D. R. (1993) *J. Biol. Chem.* **268,** 726–731.
31. Kozak, M. (1987) *Mol. Cell. Biol.* **7,** 3438–3445.
32. Peabody, D. S. & Berg, P. (1986) *Mol. Cell. Biol.* **6,** 2695–2703.
33. Sedman, S. A., Good, P. J. & Mertz, J. E. (1989) *J. Virol.* **63,** 3884–3893.
34. Murray, E. E., Lotzer, J. & Eberle, M. (1989) *Nucleic Acids Res.* **17,** 477–498.
35. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lasser, A. B., Weintraub, H. & Baltimore, D. (1989) *Cell* **58,** 537–544.
36. Lloyd, A. M., Walbot, V. & Davis, R. W. (1992) *Science* **258,** 1773–1775.