MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

Association for Information and Image Management
1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202

1 of 1

**MOLECULAR CHARACTERIZATION OF A
MAIZE REGULATORY GENE**

Annual Progress Report

**Dr. Susan R. Wessler
University of Georgia Research Foundation, Inc.
Athens, GA 30602**

"This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof."

November 1991 Through October 1992

**PREPARED FOR THE
U. S. DEPARTMENT OF ENERGY
UNDER GRANT
DE-FG09-86ER13621**

MASTER

Progress Report for Grant # DE-FG09-86ER13621
Report covers November 1991 through October 1992
Molecular Characterization of a Maize Regulatory Gene

Previously Stated Objectives:

Three overall goals comprised the last competitive renewal. These goals were revised for the reasons stated in my last progress report. I have summarized our progress on these revised goals and have included a brief description of the planned experiments for the coming year.

Summary:

All aspects of this years work have converged on the central theme of post-transcriptional control of *R* gene expression. Unlike transcriptional control, relatively little is known about post-transcriptional regulation, especially in plants. We believe that three levels of post-transcriptional regulation have been identified: (1) control of translation initiation as evidenced by the maize *Lc* gene, (2) control of nuclear localization as evidenced by the *Ds* allele *r-m9* of maize, and (3) control of nuclear localization through alternative splicing of the rice *R* homolog.

There are over 50 naturally occurring alleles of *R* in maize and many uncharacterized alleles in rice. These alleles condition different patterns of anthocyanin expression in both rice and maize. Current data suggests that allelic diversity reflects primarily different patterns of transcription initiation. However, comparison of sequenced *R* genes in maize and rice also reveals a variety of 5' leaders indicative of diverse patterns of translational control. It is unclear why the plant would go to such trouble to evolve these apparently elaborate mechanisms of translational control for a dispensable gene product that is regulated predominantly at the level of transcription initiation. We propose that although the *R* protein is dispensable, its overproduction would be detrimental because it may interact with other, essential helix-loop-helix transcriptional activators. It is hypothesized that the post-transcriptional mechanisms we have identified serve to down-regulate *R* expression.

I. Translational control of the *R* gene *Lc*:

The *Lc* cDNA, cloned and sequenced in my lab (Ludwig et al., 1989, 1990), has a 240 nt 5' leader region that contains three AUG codons upstream of the initiation codon of the *Lc* protein (figure 1). Within this leader region is a 38 codon minicistron that contains the three in-frame AUGs and is terminated by four in-frame stop codons. We have used site-directed mutagenesis of the leader region coupled with the particle bombardment assay to demonstrate that this leader region represses downstream *Lc* expression approximately 30 fold (R. Damiani and S. Wessler, in preparation). The mechanism underlying this repression has been and continues to be under investigation.

The data presented in figure 1 indicates that translation of the upstream minicistron is required for repression. This result has interesting evolutionary implications because not all of the *R* genes sequenced to date have this upstream ORF: some have upstream AUGs but no ORF whereas others have no upstream AUGs. Over the next year we hope to see whether these differences contribute to allelic diversity and how they have evolved.

Although our results clearly implicate translation of the upstream ORF (uORF) in repression of downstream *Lc* translation, the precise mechanism of this repression remains unresolved. Two models will be tested this year, (1) the product of uORF, a 38 amino acid peptide, is directly involved in repressing the initiation of *Lc* translation or (2) it is the translation of the uORF, not the peptide, per se, that represses downstream *Lc* expression by stalling ribosomes and reducing the number of initiation complexes that ultimately reach the downstream *Lc* ORF. We plan to use three experimental approaches to test these models, (1) immunolocalization and ELISAs in an attempt to detect and quantify the peptide, (2) *in vitro* translation of leader mutants and measurement of peptide/*Lc* ratios. We have previously demonstrated that the leader represses downstream *Lc* expression in rabbit reticulocyte lysates and (3) test whether the leader represses downstream expression in yeast. If so, yeast genetics will be invaluable in the isolation of second site suppressors that relieve or enhance repression and reveal cellular components that are necessary for repression.

II. Characterization of the leaky *Ds* allele, *r-m9*:

We initially pursued this allele because it was hypothesized that expression resulted from the splicing of *Ds* sequences from pre-mRNA. Instead, we have preliminary evidence that inefficient nuclear localization may be responsible for its leaky expression (Y.Liu, J. Kermicle, M. Alleman and S. Wessler, in preparation). This year we determined that (1) *r-m9* (a derivative of the *R-sc* allele from Jerry Kermicle) contains a 2 kb *Ds* insertion near the middle of exon 8 (Figure 2). The *Ds* element is not inserted in the orientation previously shown to be the required for splicing (Wessler et al. 1987), (2) sequencing of *r-m9* cDNAs indicate that insertion leads to the synthesis of a chimeric poly A+ transcript that terminates in the *Ds* element, (3) this transcript appears to be as abundant as the *R-sc* wild-type transcript and should direct the synthesis of a truncated protein due to the presence of an in-frame translation stop right at the beginning of *Ds* sequences. We have also characterized a leaky derivative of *r-m9* (isolated by Mary Alleman) in which *Ds* has excised and left behind a 7 bp transposon footprint. This frameshift mutation also leads to the production of a truncated protein that is missing about 30 amino acids at the carboxy terminus.

We have estimated that *r-m9* and its derivative may have less than 10% of wild-type *R-sc* activity. We will attempt to firm up this value over the next year. Why should deletion of the terminal 30 amino acids reduce *R* expression so dramatically? A clue to this puzzle comes from a collaborative effort with Natashe Raikhel's lab (M. Shieh, S. Wessler and N. Raikhel, submitted). They have shown that sequences deleted in *r-m9* contain a nuclear localization domain (Figure 2). The possibility that the truncated *r-*

m9 protein may be inefficiently transported to the nucleus will be tested this year.

III. Alternative splicing produces a truncated R protein in rice:

The rice *R* homolog has been cloned from a red rice variety (B. Johnston and S. Wessler, in preparation). When fused to the 35S promoter and bombarded into maize aleurones, the rice cDNA complements *R* function. Due to the use of a 5' splice site for intron 8 that is 20 nt downstream from the corresponding site in maize, an additional 20 nt are added to exon 8. This addition results in a frameshift that, like that described for the *r-m9* derivative above, leads to truncation of the protein and deletion of exon 9 sequences. However, the possibility that this represents an alternatively spliced transcript, that is, it is not the only transcript encoded by the rice gene, is suggested by conservation of exon 9 amino acids in rice and maize. This year we will test whether the sequenced cDNA is the only transcript from the rice *R* gene or whether alternative splicing generates an additional transcript that includes exon 9 sequences in the ORF. We will also look to see whether deletion of exon 9 sequences from the rice *R* protein leads to inefficient nuclear localization.

Putting it all together: These three, seemingly disparate paths of investigation, may have a common underlying theme. The uORF of *Lc* is clearly involved in repressing downstream expression of *Lc* in maize. We believe that translational control serves to prevent overexpression of the *R(Lc)* protein as high levels of this protein may be toxic to the cell. Unfortunately, data generated from analysis of expression of the rice *R* homolog seems to nullify this hypothesis. Unlike maize *R* genes, we find that the rice gene transcript is very abundant and lacks upstream AUGs. These findings suggest that the rice protein is probably abundant in rice cells. If rice can tolerate high levels of *R* protein then it is probably not toxic to the plant. However, the rice protein appears to resemble the *r-m9* mutant protein in that it is truncated for sequences that have been implicated in efficient nuclear transport. Perhaps rice has evolved this novel mechanism of post-translational control to prevent efficient nuclear localization and limit the toxic effect(s) of *R* overexpression?

1A

Reporter plasmid:

Delivery control plasmid:

Lc plasmid:

Fig. 1A. Schematic representation of plasmids delivered to maize tissue by particle bombardment. The *Bz1*LUC construct serves as a reporter construct since it is *trans*-activated by *Lc*. Luciferase activity was measured as described (6). The *Adh1*CAT construct is not *Lc*-responsive and will be expressed in maize aluerone. CAT activity, measured by the method of Sleight (40), serves to normalize luciferase values to adjust for variations between bombardments. Various *Lc* plasmids were used containing alterations in the 5' UTR (see blowup) to measure the effects on expression.

1B

		LUC/CAT Ratio	Relative Expression
pLcWT		15.9 ± 2.7	1.0
pLcΔ		350.4 ± 44	22.0 ± 2.8
pLcmut123		489.2 ± 65	30.8 ± 4.1
pLcmut1		521.6 ± 72	32.8 ± 4.5
pLcmut2		327.6 ± 55	20.6 ± 3.5
pLcmut3		268.3 ± 39	16.9 ± 2.4
pLcmut23		320.7 ± 32	20.2 ± 2.0

1C

pLcmutStop		455.4 ± 66	28.6 ± 4.2
pLcmutL4->R		368.0 ± 52	23.1 ± 3.3
pLcmutL4->I		245.0 ± 25	15.4 ± 1.6
pLcmutT30->S		249.0 ± 29	15.7 ± 1.8
pLcmutH32Syn		203.4 ± 29	12.8 ± 1.8

Fig. 1B & 1C. Schematic representation of mRNA 5' ends produced from *Lc* expression plasmids and results obtained. **b**, Site-directed mutations of AUG codons in the 5' leader of *Lc* and their effects in transient transfection experiments. The uORF is shown as a shaded rectangle with the positions of each upstream AUG codon marked. The *Lc* reading frame is indicated by a truncated black box with an arrow at the initiation codon. * indicates mutations at AUG codons, the nucleotide sequence changes are given above the * in pLcmut123. Amino acid codon changes are noted above constructs where relevant. **c**, Site-directed mutations altering the coding capacity of the uORF. The nonconservative change in pLcmutL₄->R is indicated by a lightly colored box. Conservative changes made in pLcmutL₄->I and pLcmutT₃₀->S are indicated by a shaded box. # marks the position of mutations made where applicable. The results obtained with each construct is shown at the right. Values represent the average of 12-18 samples $\pm$ SE. Relative expression values were calculated by dividing the LUC/CAT ratio of a given construct by the LUC/CAT ratio obtained for pLcWT.

The *R-sc* Gene:

The *r-m9* (*Ds*) mutation:

Figure 2: The intron/exon structure of the *R* gene (*R-sc* and *Lc* are identical). The 2 kb *Ds* element in the *r-m9* allele is inserted in exon 8 near the carboxy terminus. Insertion of *Ds* leads to polyadenylation of *r-m9* mRNA after transcription of only 150 nt of *Ds* sequences. This chimeric transcript contains an in-frame stop and results in the deletion of exon 8 and 9 sequences (shaded). The position of the nuclear localization sequence within the shaded area is indicated by a black box.

DATE

FILMED

6/15/94

END