Identifying cis-*Regulatory Elements for Embryo Region-Specific Transcription*

Kelli Henry, Tomokazu Kawashima, Michael Gavino, Koen Weterings and Robert B. Goldberg
University of California, Los Angeles/Molecular, Cell, & Developmental Biology, Los Angeles, U.S.A.

Abstract

In most higher plants, the asymmetrical division of the zygote produces a small apical cell, which gives rise to most of the embryo proper, and a large basal cell, which generates the suspensor. The embryo proper becomes the next generation plant, whereas the suspensor is a terminally differentiated embryonic region that nourishes and supports the embryo proper and degenerates by the end of embryogenesis. The Scarlet Runner Bean *G564* gene is expressed specifically in the suspensor during the early stages of seed development and in the embryo proper during the later stages of seed development. In order to understand cell fate specification in the embryo, it is necessary to identify the *cis*-regulatory elements controlling transcription in the suspensor and embryo proper. I used mutagenesis to identify suspensor and embryo proper *cis*-regulatory sequences in the upstream region of the *G564* gene. The -662 to +56 upstream region of *G564* lacks suspensor activity, but is expressed in the later embryo proper. One 10-bp sequence (GAAAAGCGAA) added to *G564* -662 to +56 is able to rescue suspensor activity, but has no effect on transcription in the embryo proper. Therefore, *G564* -662 10bp-1 contains both suspensor and embryo proper *cis*-regulatory sequences. 45-bp scanning mutagenesis of the *G564* -662 10bp-1 construct was used to identify these important *cis*-regulatory sequences. The results indicate that the -662 to -618 (m1) region is important for suspensor transcription. Mutagenesis of the *G564* upstream region did not reveal any *cis*-elements important for transcription in the mature embryo proper, possibly due to redundancy of these sequences. Therefore, 5' deletions of the *G564* upstream region were examined. Embryo proper transcription was not affected, even in the *G564* -22 deletion, in which the TATA box at -24 was deleted. 5'RACE was performed on RNA isolated from *G564* -22 and -448 dry seeds and revealed two transcription start sites within the *G564* upstream region at +1 and +39, showing that the transcription start site is not affected in *G564* -22 deletion.

G564 -662 10bp-1 Contains Both Suspensor and Embryo Proper Regulatory Sequences

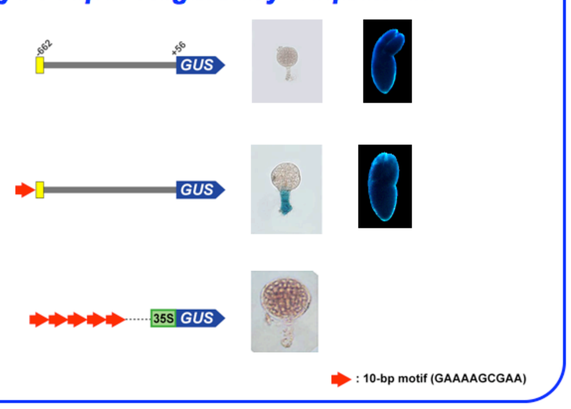
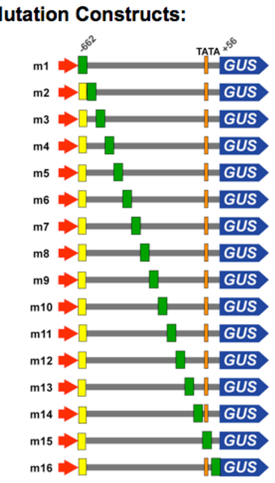


Figure 1. *G564* -662 10bp-1 Contains Both Suspensor and Embryo Proper Regulatory Sequences. When fused to the *GUS* reporter gene, the -662 to +56 upstream region of *G564* lacks suspensor activity, but is expressed in the later embryo proper. One 10-bp sequence (GAAAAGCGAA) added to *G564* -662 to +56 is able to rescue suspensor activity, but has no effect on transcription in the embryo proper. The 10-bp sequence cannot activate suspensor transcription on its own when fused to the Cauliflower Mosaic Virus 35S minimal promoter. Therefore, the 10-bp sequence may be acting with a sequence in the -662 to +56 region to activate suspensor transcription.

The G564 -662 10bp-1 Construct Can Be Mutated to Reveal Both Embryo Proper and Suspensor Regulatory Sequences

A) Mutation Constructs:



m	Mutation
1	-662 to -618
2	-617 to -573
3	-572 to -528
4	-527 to -483
5	-482 to -438
6	-437 to -393
7	-392 to -348
8	-347 to -303
9	-302 to -258
10	-257 to -213
11	-212 to -168
12	-167 to -123
13	-122 to -78
14	-77 to -33
15	-32 to +13
16	+14 to +56

B) Mutation sequence:

5' GGCCGCGGCGCTGGACTGGCATGAACTTCGGTCCGCGGGGGGCCC 3'
3' CCGGCGCCGCGACCTGACCGTACTTGAACGGAGGCGCCCCCCGGG 5'

: 10-bp motif (GAAAAGCGAA) ; : 45-bp mutation sequence ; : TATA box at -24 (TACATAA)

Figure 2. The *G564* -662 10bp-1 Construct Can Be Mutated to Reveal Both Embryo Proper and Suspensor Regulatory Sequences. A) Every 45-bp of the *G564* –662 to +56 upstream region was replaced with a mutation sequence using Fusion PCR. B) The 45-bp mutation sequence was designed from the *GUS* coding region, contains no known plant *cis*-elements and has minimal homology to the *G564* -662 to +56 region. However, the mutation sequence does contain an ATG, which may cause a frameshift in *GUS* in the m16 construct.

ACKNOWLEDGEMENTS: We are grateful to the members of our laboratory. Kelli Henry was supported by the NIH Genetics Training Grant. This work was supported by DOE and HHMI grants to Robert B. Goldberg.

The Mutation Sequence Is Unable to Activate Transcription in the Embryo

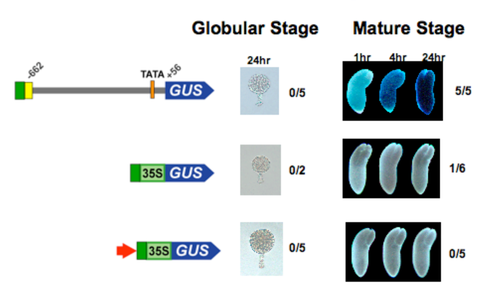


Figure 3. The Mutation Sequence Is Unable to Activate Transcription in the Embryo. The numbers to the right of each embryo indicate the number of individual transformants displaying GUS activity in the embryo per total number of transformed individuals analyzed. GUS assay incubation time is indicated above each embryo.

What Sequences in the G564 -662 Region Are Important for Suspensor Transcription?

Figure 4. Mutagenesis Analysis of the *G564* Upstream Region in the Globular Stage Suspensor. The numbers to the right of each embryo indicate the number of individual transformants displaying GUS activity in the globular stage suspensor per total number of transformed individuals analyzed, and the number of GUS positive embryos per total embryos analyzed from GUS positive lines. GUS assay incubation time is 24 hr. Chi-squared test was performed to test whether embryos from GUS positive lines are segregating 3:1 (GUS positive:GUS negative). Null hypothesis was rejected for p-value < 0.05.

TTGGT Is Present in the m1 Mutation Region and Important for Suspensor Transcription

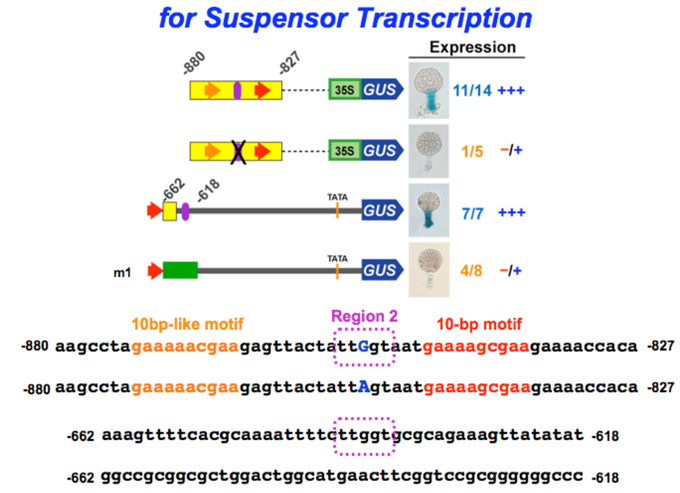


Figure 5. TTGGT and the 10-bp Sequence Are Important for Suspensor Transcription. When Region 2 in *G564* -880 to -827 is mutated, GUS expression decreases significantly. *G564* -662 to -618 contains an identical sequence (TTGGT), which is not present in -617 to +56. When -662 to -618 is mutated (m1), GUS expression also decreases significantly.

What Sequences in the G564 -662 Region Are Important for Embryo Proper Transcription?

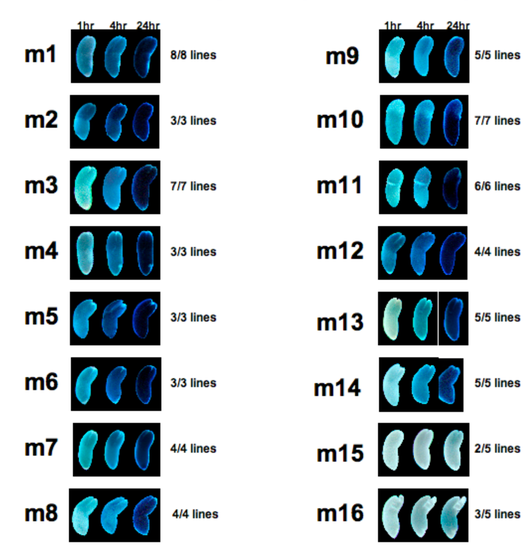


Figure 6. Mutagenesis Analysis of the *G564* Upstream Region in the Mature Stage Embryo Proper. The numbers to the right of each embryo indicate the number of individual transformants displaying GUS activity in the mature stage embryo per total number of transformed individuals analyzed. GUS assay incubation time is shown above each embryo.

5' Deletion Analysis of the *G564* Upstream Region

Figure 7. 5' Deletion Analysis of the *G564* Upstream Region. The numbers to the left indicate the deletion position relative to the transcription start site for each of the 5' truncated *G564/GUS* constructs. The numbers to the right of each embryo indicate the number of individual transformants displaying GUS activity in the mature stage embryo per total number of transformed individuals analyzed, and the number of individual embryos displaying GUS activity per total number of embryos analyzed from the GUS positive lines. GUS assay incubation time is shown above each embryo.

5' RACE Reveals Two Transcription Start Sites for G564 -22 Deletion

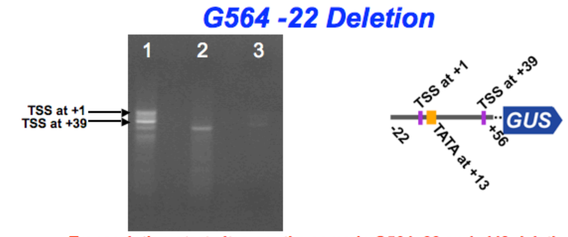


Figure 8. 5' RACE Reveals Two Transcription Start Sites for *G564* -22 Deletion. RNA was isolated from *G564* -22 and -448 dry seeds and 5' RACE was performed using the Clontech SMART RACE cDNA amplification kit. Lanes: 1. *G564* -22 cDNA amplified with SMART and *GUS* primers. 2. *G564* -22 cDNA amplified with SMART primer only. 3. *G564* -22 cDNA amplified with *GUS* primer only. Two bands in lane 1 were sequenced to reveal transcription start sites at +1 and +39. A possible TATA box (TAAAGAA) at +13 was also identified.