Identification of *cis*-regulatory sequences that activate transcription in the suspensor of plant embryos

Tomokazu Kawashima, Xingjun Wang[1], Kelli F. Henry, Yuping Bi[1], Koen Weterings[2], and Robert B. Goldberg[3]

Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, CA 90095-1606

Contributed by Robert B. Goldberg, December 27, 2008 (sent for review December 12, 2008)

Little is known about the molecular mechanisms by which the embryo proper and suspensor of plant embryos activate specific gene sets shortly after fertilization. We analyzed the upstream region of the scarlet runner bean (*Phaseolus coccineus*) *G564* gene to understand how genes are activated specifically within the suspensor during early embryo development. Previously, we showed that the *G564* upstream region has a block of tandem repeats, which contain a conserved 10-bp motif (GAAAAG^{C}/$_{T}$GAA), and that deletion of these repeats results in a loss of suspensor transcription. Here, we use gain-of-function (GOF) experiments with transgenic globular-stage tobacco embryos to show that only 1 of the 5 tandem repeats is required to drive suspensor-specific transcription. Fine-scale deletion and scanning mutagenesis experiments with 1 tandem repeat uncovered a 54-bp region that contains all of the sequences required to activate transcription in the suspensor, including the 10-bp motif (GAAAAGCGAA) and a similar 10-bp-like motif (GAAAAACGAA). Site-directed mutagenesis and GOF experiments indicated that both the 10-bp and 10-bp-like motifs are necessary, but not sufficient to activate transcription in the suspensor, and that a sequence (TTGGT) between the 10-bp and the 10-bp-like motifs is also necessary for suspensor transcription. Together, these data identify sequences that are required to activate transcription in the suspensor of a plant embryo after fertilization.

promoter analysis | scarlet runner bean

In many higher plants, the zygote divides asymmetrically into a small apical cell and a large basal cell with distinct developmental fates (1). The apical cell differentiates into the embryo proper, containing 1 or 2 cotyledons, and an axis with shoot and root meristems that generate the next generation plant. By contrast, the basal cell differentiates into the hypophysis and suspensor. The former contributes to root meristem development, whereas the latter is a terminally differentiated structure that anchors the embryo proper to surrounding maternal tissues and serves as a conduit for nutrients and growth regulators required for embryo proper development. At later stages of embryo development, the suspensor undergoes programmed cell-death (2). Several studies have shown that different genes are expressed in the embryo proper and suspensor regions (3–6); however, the DNA sequences and transcription factors that are responsible for activating different gene sets in these 2 embryo regions are not yet known.

Previously, we investigated gene activity within scarlet runner bean (SRB) embryos to gain an understanding of the processes by which genes are activated specifically in the suspensor (4). Taking advantage of the large size of SRB seeds and embryos (Fig. 1 *A* and *B*), globular-stage embryo propers and suspensors were hand dissected, and gene activity within each embryo region was studied (3, 4). We identified an mRNA, designated as G564, that encodes a protein with an unknown function and begins to accumulate in the 2 basal cells of a 4-cell embryo (4). The SRB G564 mRNA accumulation pattern is controlled primarily at the transcriptional level, persists in the suspensor of a globular-stage embryo (Fig. 1*B*), and expands to the embryo proper at later stages (3, 4). We identified tandemly repeated sequences in the *G564* upstream region that contain a 10-bp conserved motif (5′-GAAAAG^{C}/$_{T}$GAA-3′), which is also present in the upstream region of the SRB suspensor-specific *C541* gene (Fig. 1 *H* and *I*) (3, 4); 5′-deletion experiments in transgenic tobacco plants revealed that the −921 to −663 *G564* upstream region containing the tandem repeats is required for transcription in the suspensor (Fig. 1 *C*, *D*, and *H*), but not in the embryo proper (Fig. 1 *E* and *H*) (4). We speculated that the 10-bp motif might be important for suspensor transcription (4).

In this article, we present experiments that identify *cis*-regulatory sequences required for *G564* suspensor transcription. Gain-of-function (GOF) analyses revealed that 1 repeat sequence in the *G564* upstream region is sufficient for suspensor transcription. Mutagenesis and deletion experiments identified that a 54-bp region (−880 to −827) contains all of the sequences required for suspensor transcription, and that the 10-bp motif (5′-GAAAAGCGAA-3′) and a related 10-bp-like motif (5′-GAAAAACGAA-3′) within the 54-bp region are required for transcription in the suspensor. Also, sequence comparisons between the tandem repeats and site-directed and scanning mutagenesis experiments uncovered an additional sequence (5′-TTGGT-3′) that is also required for suspensor transcription. Together, our results have identified 3 adjacent regulatory elements required to activate transcription within the suspensor shortly after fertilization.

Results

The *G564* Upstream Region Activates Suspensor Transcription in Both Tobacco and *Arabidopsis*. We introduced a *G564* (−921 to +56)/*β-glucuronidase* (*GUS*) reporter gene construct into *Arabidopsis* to test whether the SRB *G564* upstream region can activate suspensor transcription in this plant species. Transgenic *Arabidopsis* embryos were hand dissected and tested for GUS activity (see *Materials and Methods*). GUS activity was first detected in the *Arabidopsis* preglobular embryo and became stronger at the globular stage (Fig. 1 *F* and *G*) (3). A similar pattern of transcriptional activity was observed in the suspensor of transgenic tobacco embryos (Fig. 1 *C* and *D*) (3, 4), indicating that both the timing and spatial orientation of *G564* transcription is the same in divergent plant embryos after fertilization.

Author contributions: T.K., K.W., and R.B.G. designed research; T.K., X.W., K.F.H., Y.B., and K.W. performed research; T.K., X.W., and R.B.G. analyzed data; and T.K. and R.B.G. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. FJ535440 and FJ535441).

[1]Present address: High-Tech Research Center, Shandong Academy of Agricultural Sciences, Key Laboratory for Genetic Improvement of Crop, Animal, and Poultry of Shandong Province, 11 Sang Yuan Road, Jinan, 250100, People's Republic of China.

[2]Present address: Bayer BioScience N.V., Technologiepark 38, B-9052, Gent, Belgium.

[3]To whom correspondence should be addressed. E-mail: bobg@ucla.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0813276106/DCSupplemental.

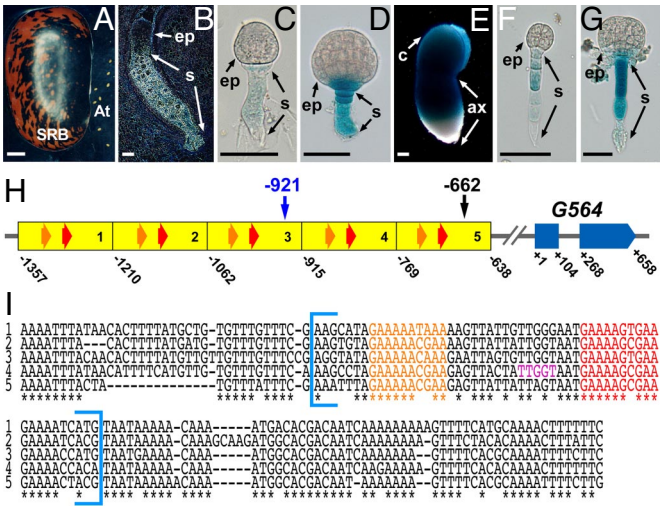

Fig. 1. SRB *G564* gene and transcriptional activity in tobacco and *Arabidopsis*. (*A*) Comparison of SRB and *Arabidopsis* seeds. (*B*) Localization of G564 mRNA in a SRB globular-stage embryo. In situ image was taken from Weterings *et al.* (4). (*C–E*) GUS activity in transgenic tobacco embryos carrying the *G564* (−921 to +56)/*GUS* (*C* and *D*) and the *G564* (−662 to +56)/*GUS* (*E*) chimeric genes. (*F* and *G*) GUS activity in transgenic *Arabidopsis* embryos carrying the *G564* (−921 to +56)/*GUS* chimeric gene. The photograph (*G*) was taken from Le *et al.* (3). Preglobular-stage (*C* and *F*), globular-stage (*D*), transition-stage (*G*), and mature (*E*) embryos were hand dissected from seeds and tested for GUS activity. GUS assay incubation time was 8 h (*C–E*) and 2 h (*F* and *G*). (*H*) Conceptual representation of the *G564* gene and upstream region. Blue and yellow boxes represent exons and 150-bp tandem repeats, respectively. Red arrows indicate a 10-bp motif (5′-GAAAAG$^{C}/_{T}$GAA-3′) identified to be conserved in the upstream regions of SRB *G564* and *C541* genes (4). Orange arrows indicate a sequence similar to the 10-bp motif (5′-GAAAAA$^{C}/_{T}{}^{G}/_{A}$AA-3′), designated as the 10-bp-like motif. Blue arrow indicates a 5′-deletion at −921 that showed GUS activity in the suspensor of transgenic tobacco embryos (*C* and *D*). Black arrow indicates a 5′-deletion at −662 that showed a weak GUS activity in the suspensor of transgenic tobacco embryos (Fig. 5*A*). The 5′-deletion results were taken from Weterings *et al.* (4) and data presented here (Fig. 5*A*). Numbers indicate positions relative to the transcription start site (+1). (*I*) Nucleotide sequence alignment of the 5 150-bp tandem repeats. Nucleotides conserved across all 5 repeats are indicated by asterisks. The 10-bp-like and 10-bp motifs, and region 2 (5′-TTGGT-3′) in the fourth repeat are shown in orange, red, and pink, respectively. Light blue brackets include the 54-bp regions analyzed further (Fig. 4*A*). The number to the left of aligned sequences indicates positions of the repeats. Gaps were introduced for optimal alignment. At, *Arabidopsis*; ax, axis; c, cotyledon; ep, embryo proper; s, suspensor. (Scale bars: *A*, 1 cm; *B–G*, 50 μm.)

The *G564* Upstream Region Contains 5 150-bp Tandem Repeats. We further analyzed the repeated sequences in the *G564* upstream region and found that the unit repeat sequence was ≈150 bp in length and occurred 5 times in tandem (−1357 to −638) (Fig. 1*H*). Sequences within these 5 150-bp repeats were similar, but not identical, and gaps were uncovered when they were aligned (Fig. 1*I*). The fifth repeat was the least conserved among the 5 repeats and had a large gap (14 bp) at its 5′-end (Fig. 1*I*). Sequences closely related to the 10-bp motif, designated as a 10-bp-like motif (5′-GAAAAA$^{T}/_{C}{}^{A}/_{G}$AA-3′), were identified 17-bp upstream from the 10-bp motif within each repeat (Fig. 1 *H* and *I*). Clustal analysis showed that (*i*) one duplication event generated either the first/second or third/fourth repeats, (*ii*) the ancestor pair duplicated, and (*iii*) the divergent fifth repeat was probably derived from the fourth repeat.

One 150-bp Tandem Repeat Can Program Suspensor Transcription. To identify suspensor *cis*-regulatory sequences in the *G564* upstream region, we generated GOF constructs in which each fragment from the upstream region was fused to a *Cauliflower mosaic virus* (*CaMV*) *35S* minimal promoter/*GUS* vector (7), and then introduced these

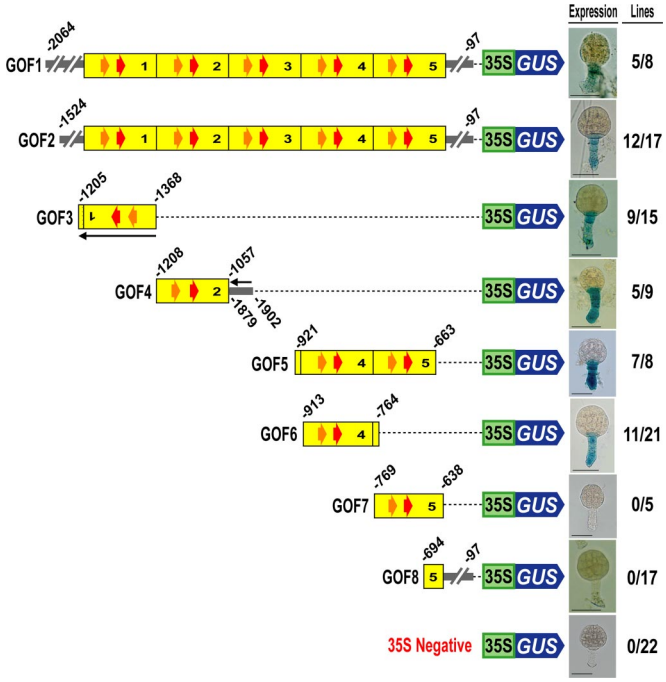

Fig. 2. One 150-bp tandem repeat is sufficient for suspensor transcription. Suspensor GUS activity in transgenic tobacco embryos containing different GOF constructs. Names and conceptual representations of GOF constructs to the left of each embryo. Numbers indicate positions relative to the *G564* transcription start site (+1). Yellow boxes indicate 150-bp tandem repeats. Orange and red arrows indicate the 10-bp-like and 10-bp motifs, respectively; *35S/GUS* indicates the *CaMV 35S* minimal promoter/*GUS* gene (7). Black arrows indicate fragments inserted in the opposite direction to the transcription start site. The GOF4 construct contains 2 fragments (−1208 to −1057, and −1902 to −1879). Numbers in the Lines column indicate the number of individual transformants displaying suspensor GUS activity over total number of transformants analyzed. Photographs of GOF1, GOF2, and GOF4 were taken after 16-h GUS incubation. Photographs of GOF3, GOF5 to GOF8, and 35S Negative were taken after 24-h GUS incubation. (Scale bar: 50 μm.)

constructs into tobacco plants (Fig. 2). GOF constructs that contained all 5 repeats showed strong suspensor GUS activity (GOF1 and GOF2). By contrast, neither the GOF construct without the tandem repeats (GOF8) nor the *CaMV 35S* minimal promoter/*GUS* construct (35S Negative) showed suspensor GUS activity (Fig. 2). These results indicate that suspensor transcription can be driven by a heterologous *CaMV 35S* minimal promoter, and that sequences required for transcription in the suspensor are within the −1524 to −97 fragment (GOF2).

Previously, we identified that the region between −921 and −663 contains sequences important for suspensor transcription (4). To investigate whether the −921 to −663 fragment (GOF5) was sufficient for transcription in the suspensor, transgenic tobacco plants containing the GOF5 construct were tested for GUS activity. Strong GUS activity was observed in the suspensor of GOF5 embryos (Fig. 2), demonstrating that all of the sequences required for suspensor transcription were within the −921 to −663 *G564* upstream region.

The GOF5 fragment contains the fourth repeat and most of the fifth repeat (Fig. 2). To determine whether 1 repeat was sufficient for suspensor transcription, transgenic tobacco embryos containing GOF constructs with the first (GOF3), second (GOF4), fourth (GOF6), and fifth (GOF7) repeat regions were examined individually for GUS activity. The first, second, and fourth repeat constructs (GOF3, GOF4, and GOF6) showed strong GUS activity in the suspensor (Fig. 2). The first repeat fragment was inserted into the GOF vector in the opposite

direction, indicating that suspensor *cis*-regulatory sequences functioned in either orientation. By contrast, the fifth divergent repeat construct (GOF7) did not show GUS activity in the suspensor (Fig. 2). Together, these results show that 1 150-bp repeat sequence contains sequences required for suspensor transcription, and sequence differences in the fifth repeat might be within important suspensor *cis*-regulatory sequences (Fig. 1*I*).

A 54-bp Region Contains All Sequences Required for Suspensor Transcription. To identify suspensor *cis*-regulatory sequences within the repeats, we mutagenized the fourth repeat sequence (GOF6) by dividing the sequence into 10 15-bp segments, and replacing each segment with a 15-bp mutation sequence (5′-GGC-CGCGGGGGGCCC-3′). This sequence did not contain any known plant *cis*-elements or sequences similar to the original GOF6 fragment [see supporting information (SI) *Materials and Methods*]. Transgenic tobacco embryos containing the 15-bp mutation fused to the *CaMV 35S* minimal promoter/*GUS* construct (15-bp Mutation Negative) did not show GUS activity (Fig. 3*A*).

We examined 10 mutagenesis constructs (M1 to M10) for GUS activity in the suspensor. Mutations M1, M2, and M6 to M10 did not affect suspensor GUS activity (Fig. 3*A*). By contrast, GUS activity in the suspensor was significantly decreased when the mutation sequence replaced the 10-bp-like motif region (M3 and M4), and abolished when the mutation sequence replaced the 10-bp motif region (M5) (Fig. 3*A* and Fig. S1). Ectopic GUS activity was not observed in the embryo proper in any of the M1 to M10 constructs, suggesting the absence of negative elements. Together, these results show that the M3 to M5 region, spanning 45 bp of *G564* upstream sequence (−883 to −839), contains positive elements that are necessary for transcription in the suspensor.

Because short sequence redundancies were observed within the 150-bp tandem repeats (Fig. 1*I*), it was possible that sequences flanking the 45-bp region (−883 to −839) were also important for suspensor transcription and would not have been uncovered in our mutagenesis experiment (Fig. 3*A*). To test this possibility, we generated GOF constructs containing a core 54-bp fragment (−880 to −827) that included the 10-bp-like and 10-bp motifs and different lengths of 5′- or 3′- flanking sequences (GOF9 to GOF12) (Fig. 3*B*). Each of these constructs showed strong GUS activity in the suspensor at the same level as the control GOF6 construct (Fig. 3 *A* and *B*). Together, these results indicate that the 54-bp fragment (GOF12) contains sequences required for suspensor transcription.

The 10-bp-Like and 10-bp Motifs Are Required for Suspensor Transcription. To determine whether the 10-bp-like and 10-bp motifs within the 54-bp GOF12 fragment were required for suspensor transcription, we replaced the 10-bp-like motif (M11) and the 10-bp motif (M12) with a 10-bp mutagenesis sequence (5′-CGGGGGGCCC-3′) (Fig. 3*C*; see *SI Materials and Methods*). Mutation of the 10-bp-like motif (M11) caused a significant decrease in suspensor GUS activity (Fig. 3*C*; Fig. S2). By contrast, no GUS activity was detected when the 10-bp motif was mutated (M12) (Fig. 3*C*; Fig. S2). These results show that the 10-bp-like and 10-bp motifs within the 54-bp fragment are required for transcription in the suspensor.

A Sequence Between the 10-bp-Like and 10-bp Motifs Is also Required for Suspensor Transcription. The fifth repeat (GOF7) did not activate transcription in the suspensor, as compared with the first, second, and fourth repeats (GOF 3, 4, and 6), even though the 10-bp-like and 10-bp motifs were present in the fifth repeat (Figs. 1*H* and 2). This result suggested that additional sequences in the 54-bp fragment were required for suspensor transcription. We aligned and compared the 54-bp repeat regions to identify sequence differences between suspensor GUS-positive repeats (first, second, and fourth) and the GUS-negative fifth repeat (Fig. 4*A*). We uncovered 3 nucleotides that differed between GUS-positive and

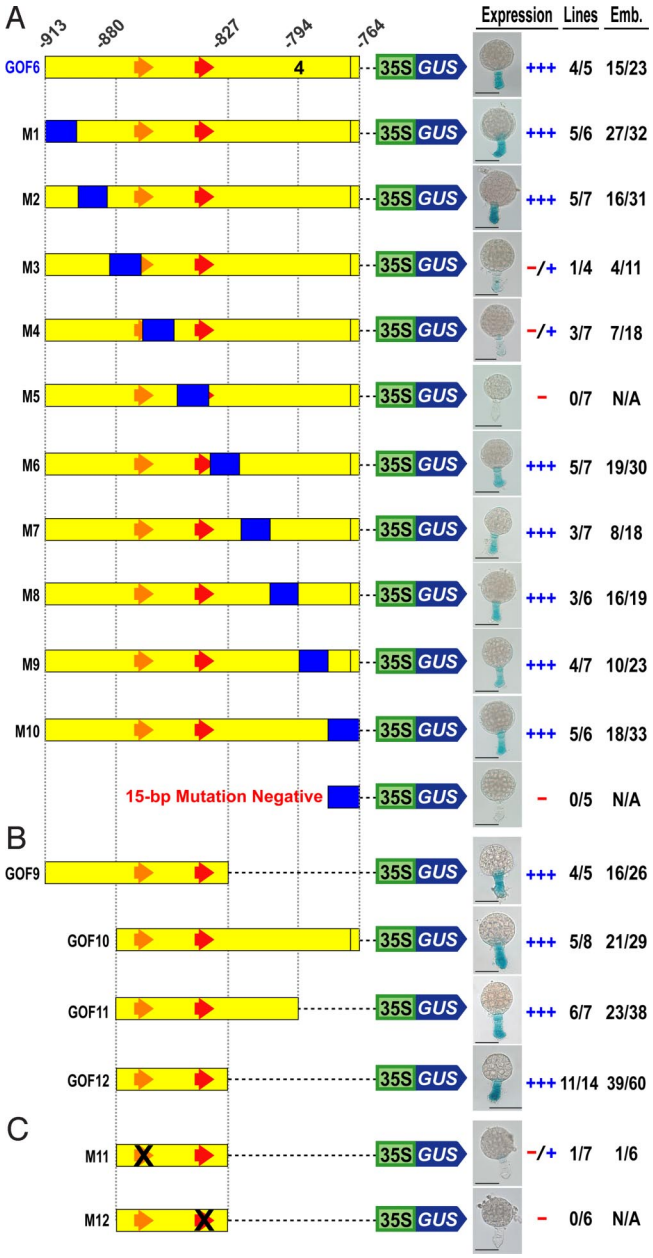

Fig. 3. The 10-bp-like and 10-bp motifs are required for suspensor transcription. Suspensor GUS activity in transgenic tobacco embryos containing 15-bp scanning mutatgenesis constructs (*A*), constructs with GOF fragments focusing on the −880 to −827 region (*B*), and the GOF12 constructs with mutagenesis in the 10-bp-like and 10-bp motifs (*C*). Blue blocks indicate the 15-bp mutation sequence (5′-GGCCGCGGGGGGCCC-3′). Black crosses on the 10-bp-like and 10-bp motifs indicate mutagenesis of these motifs (see *SI Materials and Methods*); +++ in the Expression column indicates that suspensor GUS activity was strong and detected by 2-h incubation in the majority of GUS-positive lines; −/+ in the Expression column indicates that suspensor GUS activity was weak and not detected by 2-h incubation in the majority of GUS-positive lines; − in the Expression column indicates no detectable suspensor GUS activity. Numbers in the Emb. column indicate the number of embryos displaying suspensor GUS activity by 24-h incubation over total number of analyzed embryos of GUS-positive lines. Other figure details are the same as those outlined in Fig. 2. Higher magnification photos of these results are shown in Figs. S1 and S2. Emb., embryos; N/A, not applicable. Photographs were taken after 24-h GUS incubation. (Scale bar: 50 μm.)

GUS-negative repeats (Fig. 4*A*). Two nucleotide differences were located in regions (designated as regions 1 and 2) within the 45-bp sequence (−883 to −839) (Figs. 3*A* and 4*B*). To test whether these

PLANT BIOLOGY

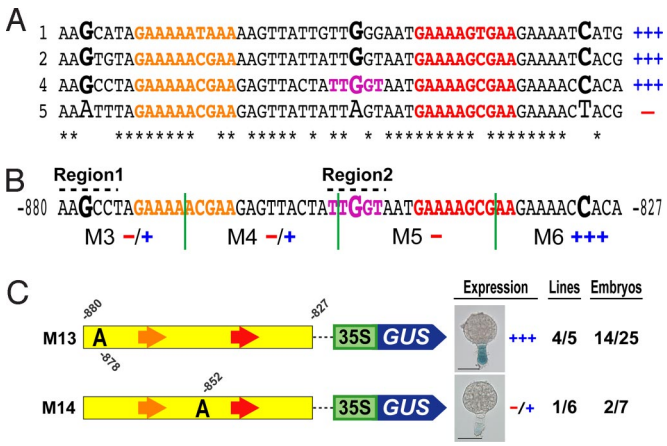

Fig. 4. A region between the 10-bp-like and 10-bp motifs is also required for suspensor transcription. (*A*) Alignment of 54-bp tandem repeat regions that were tested for GUS activity. Asterisks indicate conserved nucleotides across 4 repeats. Nucleotides shown in orange, red, and pink indicate the 10-bp-like, 10-bp, and region 2 motifs, respectively. Nucleotides shown in larger and bold fonts indicate nucleotides that are conserved in GUS-positive repeats. (*B*) Nucleotide sequences of the 54-bp fragment and summary of the 15-bp scanning mutagenesis experiment (Fig. 3*A*). Green lines display borders of the 15-bp mutation constructs (M3 to M6), as shown in Fig. 3*A*. Regions 1 and 2 indicate 5-bp sequences containing a nucleotide G at −878 and −852, respectively. (*C*) Suspensor GUS activity in embryos containing the 54-bp sequence (−880 to −827) with G to A mutations at −878 (M13) and −852 (M14). Other figure details are the same as those outlined in Figs. 2 and 3. Photographs were taken after 24-h GUS incubation. (Scale bar: 50 μm.)

2 regions were important for suspensor transcription, we mutated the guanines at positions −878 and −852 in regions 1 and 2 to adenines that were found in the fifth repeat (M13 and M14) (Fig. 4 *A*–*C*). Mutation of G to A in region 1 did not affect GUS activity in the suspensor. By contrast, mutation of G to A in region 2 significantly decreased suspensor GUS activity (Fig. 4*C*; Fig. S2). Together, these results show that region 2 between the 10-bp-like and 10-bp motifs is also required for suspensor transcription.

The 10-bp-Like and 10-bp Motifs and Region 2 Sequence Are Discrete Suspensor *cis*-Regulatory Elements. Previously, we showed that suspensor GUS activity was abolished when the *G564* upstream region was deleted to −662 (D-662) (Fig. 1*H*) (4). By contrast, the D-662 construct produced strong GUS activity in the embryo proper at later stages of development, indicating that regulatory sequences programming *G564* transcription in the suspensor and embryo proper reside in distinct *G564* upstream regions (Fig. 1*E*). Analysis of additional D-662 construct lines showed that this construct could direct weak transcription in the suspensor (Fig. 5*A*; Fig. S3). A sequence comparison of the −662 *G564* upstream region and the 54-bp GOF12 region (−880 to −827) uncovered a conserved 5′-TTGGT-3′ motif (−640 to −636) identical to the region 2 sequence (Fig. 5*A*). Deletion of the 5′-TTGGT-3′ sequence from the −662 *G564* upstream region abolished suspensor GUS activity (D-629) (Fig. 5*A*; Fig. S3), consistent with the hypothesis that this motif is important for transcription in the suspensor.

To test whether the addition of the 10-bp motif to the D-662 construct could restore GUS activity in the suspensor, 1 and 2 copies of the 10-bp motif were fused to the 5′-end of the D-662 construct (GOF14 and GOF13, respectively), and transgenic tobacco embryos were examined for GUS activity. Both GOF13 and GOF14 showed strong GUS activity in the suspensor (Fig. 5*A*). By contrast, constructs containing 2 and 5 copies of the 10-bp motif fused to the *CaMV 35S* minimal promoter/*GUS* vector did not show any GUS activity in the suspensor (Fig. 5*B*), confirming that the

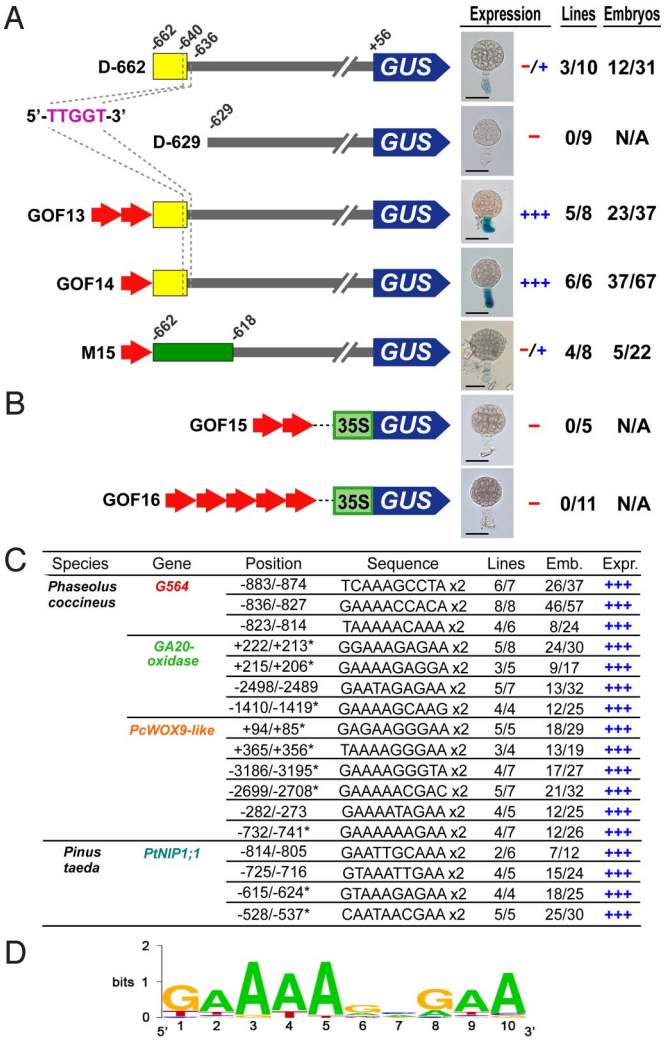

Fig. 5. The 10-bp motif and degenerate 10-bp motifs combined with the TTGGT sequence can program suspensor transcription. (*A*) Suspensor GUS activity in transgenic tobacco embryos containing the −662 *G564* deletion (D-662), the −629 *G564* deletion (D-629), D-662 constructs with 1 (GOF14) and 2 (GOF13) copies of the 10-bp motif, and the GOF14 construct with a 45-bp mutation in the region at −662 to −618 (M15). Higher magnifications are shown in Fig. S3. (*B*) Suspensor GUS activity in transgenic tobacco embryos containing the *CaMV 35S* minimal promoter/*GUS* constructs with 2 (GOF15) and 5 (GOF16) copies of the 10-bp motif. (*C*) Sequences and results of suspensor GUS activity in transgenic tobacco embryos containing the D-662 constructs with 2 copies of sequences similar to the 10-bp motif present in 5′-flanking regions of suspensor-active genes. (*D*) Consensus 10-bp motif sequence generated from 17 sequences similar to the 10-bp motif shown in *C* and the 10-bp motif (5′-GAAAAGCGAA-3′) using WebLogo (28). Green boxes indicate 45-bp replacement mutations. Numbers in the Position column indicate positions of 10-bp sequences relative to their transcription start sites (+1). Asterisks indicate 10-bp sequences found in the opposite direction to the transcription start site. Other figure details are the same as those outlined in Figs. 2 and 3. Emb., Embryos; Expr., Expression; N/A, not applicable. Photographs were taken after 24-h GUS incubation. (Scale bar: 50 μm.)

10-bp motif by itself cannot activate suspensor transcription (GOF 7) (Fig. 2).

To determine what sequences in the −662 *G564* upstream region together with the 10-bp motif can activate suspensor transcription, we replaced the −662 *G564* upstream region of the GOF14 construct with a 45-bp mutation sequence (see *SI Materials and Methods*). Transgenic tobacco embryos that contained the GOF14 construct without the 5′-TTGGT-3′ sequence (M15) showed significantly weaker GUS activity in the suspensor, similar to the

D-662 construct (Fig. 5*A*; Fig. S3). By contrast, 45-bp replacements at all other −662 *G564* upstream positions did not affect GUS activity in the suspensor (e.g., −617 to −573, M16) (Fig. S3). Together, these results demonstrate that (*i*) both the −662 *G564* upstream region containing the 5′-TTGGT-3′ sequence and the 10-bp motif are required for suspensor transcription, and (*ii*) the 10-bp motif is a discrete regulatory element distinct from the 5′-TTGGT-3′ sequence.

Divergent 10-bp Motif Sequences Can Program Transcription in the Suspensor. We identified divergent 10-bp motif sequences in the *G564* upstream region as well as in 5′-flanking sequences of several genes known to be active in the suspensor (Fig. 5*C*) (3). To determine whether these divergent 10-bp motifs could program transcription in the suspensor, we fused 2 copies of each sequence to the −662 *G564* upstream region, generating GOF constructs similar to GOF13 (Fig. 5*A*). All of the constructs showed strong suspensor GUS activity (Fig. 5*C*), suggesting that the transcription factor binding to the 10-bp motif can recognize divergent sequences. A comparison of all divergent 10-bp motif sequences suggested the 5′-GAAAA-3′ core sequence shown in Fig. 5*D*.

Discussion

At Least 3 Positive *cis*-Elements Are Required to Activate *G564* Suspensor Transcription. We used the SRB *G564* gene to identify *cis*-regulatory sequences important for suspensor transcription. The *G564* upstream region contains 5 150-bp tandem repeats (Fig. 1*H*), and 1 repeat is sufficient for suspensor transcription (Fig. 2). This unique tandem repeat organization does not occur in the 5′-flanking regions of *G564*-related soybean genes (*Glycine max*), which diverged from SRB ≈19 million years ago (8). Nor does it occur in 5′-flanking regions of *G564-related* genes in other legumes such as *Medicago truncatula* and *Lotus japonicas*, suggesting that the SRB *G564* tandem repeats originated relatively recently (i.e., within 19 million years).

GOF experiments identified a 54-bp GOF12 region within the fourth repeat (−880 to −827) that contains all sequences required for spatial and temporal control of suspensor transcription (Fig. 3*B*). Site-directed mutagenesis showed that both the 10-bp-like and 10-bp motifs within the 54-bp region are essential for suspensor transcription (Fig. 3*C*; Fig. S2). However, these motifs are not sufficient to activate suspensor transcription, because they are present in the 54-bp region of the fifth repeat that cannot drive suspensor GUS activity (GOF7) (Fig. 2). Comparison of 54-bp regions within the tandem repeats, and additional site-directed mutagenesis experiments, identified a G nucleotide at −852 in the region 2 sequence (5′-TTGGT-3′) that is also required for suspensor transcription (Fig. 4 *A*–*C* and Fig. S2). Detailed analyses of 5′-deletion data uncovered a sequence (−640 to −636) identical to that in region 2 (Fig. 5*A*). Consistent with this finding, GOF and mutagenesis experiments with the D-662 construct showed that the region between −662 and −618 contains sequences that activate suspensor transcription with the 10-bp motif (Fig. 5*A*; Fig. S3). These results indicate that the 10-bp-like, 10-bp, and region 2 motifs constitute individual *cis*-elements, and that the 54-bp region can be considered as a positive suspensor *cis*-regulatory module (9).

Addition of the 10-bp motif to D-662 (GOF13 and GOF14) restored strong GUS activity in the suspensor (Fig. 5*A*). How can this occur in the absence of a 10-bp-like motif? Perhaps this is explained by the high degeneracy of the 10-bp motif (Fig. 5 *C* and *D*), which might allow sequences similar to the 10-bp and 10-bp-like motifs near the 5′-TTGGT-3′ sequence in D-662 to function as a 10-bp-like motif (e.g., −654 to −658 and −641 to −645). Collectively, these sequences plus the 10-bp and 5′-TTGGT-3′ motifs are able to allow GOF13 and GOF14 to program transcription in the suspensor. Which sequences in D-662 act as a 10-bp-like motif, and whether these sequences function in the endogenous promoter context, remain to be determined.

The results of the D-662 GOF experiment with divergent 10-bp motifs indicate that there is a core 5′-GAAAA-3′ sequence at the 5′-end, followed by 2 degenerate middle nucleotides and a 3′-end 5′-GAA-3′ sequence in the consensus motif (Fig. 5*D*). The 15-bp M3 mutation replacing region 1 and the 10-bp-like motif of 5′-GAAAA-3′ sequence caused a drastic decrease in the suspensor GUS activity (Figs. 3*A* and 4*B*; Fig. S1), whereas mutation of the G nucleotide in region 1 had no effect on GUS activity (Fig. 4*C*; Fig. S2). This result suggests that the 5′-GAAAA-3′ is critical for the function of the 10-bp-like and 10-bp motifs. Similarly, the 15-bp M16 mutation replacing the last 2 nucleotides (5′-AA-3′) of the 10-bp motif showed no effect on GUS activity (Figs. 3*A* and 4*B*). Together, these results suggest that the 8 5′-nucleotides of the 10-bp-like and 10-bp motifs, including the 5 conserved 5′-GAAAA-3′ nucleotides, may be essential to activate suspensor transcription.

The *G564* Suspensor Regulatory Apparatus Is Relatively Simple. The suspensor is derived from the basal cell of a 2-cell embryo, whereas the apical cell gives rise to the embryo proper and the next generation plant (1). Previously, we hypothesized that transcription factors localized asymmetrically in the basal cell are responsible for activating suspensor-specific transcription (4). Here, we did not find negative elements involved in *G564* suspensor transcription. This result suggests that the 54-bp *cis*-regulatory module interacts with positive transcription factors either localized or active specifically in the basal cell lineage to initiate *G564* suspensor transcription after fertilization. In contrast with the relatively simple organization of the *G564* suspensor regulatory module, genes that are active in the embryo proper appear to have more complex regulatory architecture. For example, storage protein genes, such as *β-Phaseolin*, have several positive and negative *cis*-elements, scattered throughout the promoter that are responsible for programming gene activity with respect to both time and space during embryo proper development (10). Similarly, the *Arabidopsis ATML1* gene has different positive *cis*-regulatory modules that are responsible for programming transcription in embryo proper epidermal cells depending on their positions along the embryonic axis (11). Perhaps, the difference between the *G564* suspensor regulatory architecture and that of genes active in the embryo proper can be explained by the cellular complexity and developmental fate of each embryo region. That is, the embryo proper differentiates into multiple cell types, tissues, and organs with precise orientations during development, whereas the suspensor is a terminally differentiated structure consisting of relatively uniform cell types that are direct descendants of the basal cell (2).

Recently, *Arabidopsis* WOX8/WOX9 transcription factor mRNAs were found to be localized specifically in the 2-cell embryo basal cell (5) and to be important for suspensor development (12). *Zea mays* ZmWOX9 mRNAs are suspensor-specific (13), and we uncovered a SRB PcWOX9-like mRNA that accumulates to a high level in the suspensor (3). Transcription factors like WOX8/9 that are asymmetrically localized in the basal cell might have a major role directly or indirectly in activating *G564* transcription in the suspensor.

We used PLACE (14) to check whether any known plant *cis*-elements are present in the 54-bp *G564* suspensor regulatory module. Although no known regulatory elements were found in region 2, sequences closely related to a GT element are present in the 10-bp-like and 10-bp motif regions. GT elements contain a core sequence of 1 or 2 G nucleotides followed by 4 to 5 T or A nucleotides (5′-G$^{G}/_{A}$A$^{A}/_{T}$AA$^{A}/_{T}$-3′) (15), which is similar to the 5′-GAAAA-3′ consensus sequence at the 5′-end of the 10-bp-like and 10-bp motifs (Fig. 5*D*). There are several common features between the 10-bp-like/10-bp motif and GT elements. First, GT elements are usually present in tandem repeats (15). Second, GT elements in genes such as *RBCS-3A* are necessary, but not sufficient for transcription (16). Last, 2 GT elements are required for full

PLANT BIOLOGY

transcription (16, 17). Transcription factors that bind to GT elements are trihelix DNA-binding factors (GT factors) (18–20). Recent GeneChip and EST sequencing studies in our laboratory showed that GT factor mRNAs are present in SRB, *Arabidopsis*, and soybean suspensors (http://estdb.biology.ucla.edu/seed/). Because the SRB *G564* suspensor regulatory module can activate transcription in tobacco and *Arabidopsis* suspensors (Fig. 1 *C*, *D*, *F*, and *G*), transcription factors binding to this module are present in the suspensor of these embryos. Whether GT factors bind to the 10-bp-like and 10-bp motifs and have a major role in suspensor transcription remain to be determined.

A Model for *G564* Suspensor Transcription. The 10-bp-like, 10-bp, and region 2 5′-TTGGT-3′ motifs constitute discrete regulatory elements within the 54-bp module that are essential for suspensor transcription (Figs. 3 *B* and *C*, 4*C*, and 5*A*). How do these 3 elements interact to activate transcription in the suspensor? Fig. S4 illustrates a heuristic model for *G564* suspensor transcription by the 54-bp module. This model assumes that the 10-bp-like and 10-bp motifs bind the same transcription factor X. However, region 2 binds a different transcription factor Y. Transcription factors X and Y presumably form a complex that interacts with the basal transcriptional machinery to activate transcription in the suspensor (Fig. S4). Within the 54-bp module, there is 9-bp between the 10-bp-like motif and region 2 sequence (Fig. 4*B*), providing enough space for another element that would have gone undetected in our experiments. Further analysis will be required to determine whether there is an additional positive *cis*-element important for *G564* suspensor transcription.

In addition to the SRB G564 mRNA, there are several other SRB mRNAs that accumulate specifically in the globular-stage suspensor (e.g., C541, GA 20-oxidase, PcWOX9-like) (3, 4). How are genes encoding these mRNAs activated in the suspensor? A simple hypothesis is that the transcriptional machinery activating *G564* in the suspensor also regulates these suspensor-active genes. Indeed, 10-bp motif and region 2 sequences are present in the *C541* upstream region (4). Also, 10-bp sequences similar to the 10-bp motif are present in the upstream regions of the *GA 20-oxidase* (GenBank accession no. FJ535441), *PcWOX9-like* (GenBank accession no. FJ535440), and Pine *PtNIP1;1* (21) genes, and function similarly to the 10-bp motif in the context of the *G564* D-662 promoter (Fig. 5*C*). The region 2 5′-TTGGT-3′ sequence is also present in these upstream regions. These observations suggest that the *G564* suspensor regulatory apparatus uncovered in our experiments is conserved among seed plants, and originated before the divergence of angiosperms and gymnosperms >300 million years ago (22). Clearly, the presence of sequences identical to specific *cis*-elements does not necessarily mean that these elements function to activate transcription in their endogenous promoters (23, 24). Further analysis will be required to determine whether genes active in suspensors of divergent plants contain regulatory modules similar to that found in *G564*. The precise nature of regulatory networks required for activating genes in the suspensor after fertilization remains to be determined.

Materials and Methods

Generating GOF and Mutagenesis Constructs. Details are described in *SI Materials and Methods*. Primer and oligonucleotide sequences are listed in Table S1.

Plant Transformation. Tobacco (*Nicotiana tabacum* cv SR1) and *Arabidopsis* (Col-0 ecotype) plants were used for generating transformants. Tobacco plants were transformed and regenerated by using the leaf disk procedure (25). *Arabidopsis* plants were transformed by using the floral dip method (26). Each individual transformant was checked for T-DNA insertion by PCR analysis. Promoter/*GUS* regions of 1 or 2 tobacco lines of suspensor GUS negative constructs were sequenced to ensure no rearrangements.

GUS Histochemical Assay. Transgenic tobacco and *Arabidopsis* seeds were harvested at different stages of development. Embryos were dissected from seeds and assayed for GUS activity for 2 to 24 h at 37 °C as described previously (27). Embryos were photographed under bright- or dark-field illumination by using compound (LEICA 5000 B; Leica) and dissecting (OLYMPUS SZH; Olympus) microscopes.

ACKNOWLEDGMENTS. We thank the members of our laboratory for generous help with different aspects of this project, and Professors Eric Davidson and John Harada for insightful advice. This work was supported by grants from the Department of Energy and Ceres, Inc. (to R.B.G.). T.K. was supported by a Nakajima Foundation Predoctoral Fellowship, and K.F.H. was supported by a National Institutes of Health Predoctoral Traineeship.

1. Goldberg RB, de Paiva G, Yadegari R (1994) Plant Embryogenesis: Zygote to Seed. *Science* 266:605–614.
2. Yeung EC, Meinke DW (1993) Embryogenesis in Angiosperms: Development of the Suspensor. *Plant Cell* 5:1371–1381.
3. Le BH, *et al.* (2007) Using genomics to study legume seed development. *Plant Physiol* 144:562–574.
4. Weterings K, *et al.* (2001) Regional localization of suspensor mRNAs during early embryo development. *Plant Cell* 13:2409–2425.
5. Haecker A, *et al.* (2004) Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in Arabidopsis thaliana. *Development* 131:657–668.
6. Okamoto T, Scholten S, Lorz H, Kranz E (2005) Identification of genes that are up- or down-regulated in the apical or basal cell of maize two-celled embryos and monitoring their expression during zygote development by a cell manipulation- and PCR-based approach. *Plant Cell Physiol* 46:332–338.
7. Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RB (1990) Different Temporal and Spatial Gene Expression Patterns Occur during Anther Development. *Plant Cell* 2:1201–1224.
8. Lavin M, Herendeen PS, Wojciechowski MF (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst Biol* 54:575–594.
9. Davidson EH, *et al.* (2002) A genomic regulatory network for development. *Science* 295:1669–1678.
10. Bustos MM, Begum D, Kalkan FA, Battraw MJ, Hall TC (1991) Positive and negative cis-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. *EMBO J* 10:1469–1479.
11. Takada S, Jurgens G (2007) Transcriptional regulation of epidermal cell fate in the Arabidopsis embryo. *Development* 134:1141–1150.
12. Breuninger H, Rikirsch E, Hermann M, Ueda M, Laux T (2008) Differential expression of WOX genes mediates apical-basal axis formation in the Arabidopsis embryo. *Dev Cell* 14:867–876.
13. Nardmann J, Zimmermann R, Durantini D, Kranz E, Werr W (2007) WOX gene phylogeny in Poaceae: A comparative approach addressing leaf and embryo development. *Mol Biol Evol* 24:2474–2484.
14. Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27:297–300.
15. Zhou DX (1999) Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plants Sci* 4:210–214.
16. Sarokin LP, Chua NH (1992) Binding sites for two novel phosphoproteins, 3AF5 and 3AF3, are required for rbcS-3A expression. *Plant Cell* 4:473–483.
17. Kuhlemeier C, *et al.* (1988) Localization and conditional redundancy of regulatory elements in rbcS-3A, a pea gene encoding the small subunit of ribulose-bisphosphate carboxylase. *Proc Natl Acad Sci USA* 85:4662–4666.
18. Dehesh K, Hung H, Tepperman JM, Quail PH (1992) GT-2: A transcription factor with twin autonomous DNA-binding domains of closely related but different target sequence specificity. *EMBO J* 11:4131–4144.
19. Gilmartin PM, Memelink J, Hiratsuka K, Kay SA, Chua NH (1992) Characterization of a gene encoding a DNA binding protein with specificity for a light-responsive element. *Plant Cell* 4:839–849.
20. Perisic O, Lam E (1992) A tobacco DNA binding protein that interacts with a light-responsive box II element. *Plant Cell* 4:831–838.
21. Ciavatta VT, Egertsdotter U, Clapham D, von Arnold S, Cairney J (2002) A promoter from the loblolly pine PtNIP1;1 gene directs expression in an early-embryogenesis and suspensor-specific fashion. *Planta* 215:694–698.
22. Bowe LM, Coat G, dePamphilis CW (2000) Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proc Natl Acad Sci USA* 97:4092–4097.
23. Punwani JA, Rabiger DS, Lloyd A, Drews GN (2008) The MYB98 subcircuit of the synergid gene regulatory network includes genes directly and indirectly regulated by MYB98. *Plant J.* 55:406–414.
24. Etchberger JF, *et al.* (2007) The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. *Genes Dev* 21:1653–1674.
25. Horsch RB, *et al.* (1985) A Simple and General Method for Transferring Genes into Plants. *Science* 227:1229–1231.
26. Clough SJ, Bent AF (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 16:735–743.
27. Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: Beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* 6:3901–3907.
28. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188–1190.