Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies

Min Chen[a], Jer-Young Lin[a,1], Xiaomeng Wu[a,2], Nestor R. Apuya[a], Kelli F. Henry[a], Brandon H. Le[a,3], Anhthu Q. Bui[a,4], Julie M. Pelletier[b], Shawn Cokus[a], Matteo Pellegrini[a], John J. Harada[b], and Robert B. Goldberg[a,5]

[a]Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, CA 90095; and [b]Department of Plant Biology, College of Biological Sciences, University of California, Davis, CA 95616

An important question is what genes govern the differentiation of plant embryos into suspensor and embryo proper regions following fertilization and division of the zygote. We compared embryo proper and suspensor transcriptomes of four plants that vary in embryo morphology within the suspensor region. We determined that genes encoding enzymes in several metabolic pathways leading to the formation of hormones, such as gibberellic acid, and other metabolites are up-regulated in giant scarlet runner bean and common bean suspensors. Genes involved in transport and Golgi body organization are up-regulated within the suspensors of these plants as well, strengthening the view that giant specialized suspensors serve as a hormone factory and a conduit for transferring substances to the developing embryo proper. By contrast, genes controlling transcriptional regulation, development, and cell division are up-regulated primarily within the embryo proper. Transcriptomes from less specialized soybean and *Arabidopsis* suspensors demonstrated that fewer genes encoding metabolic enzymes and hormones are up-regulated. Genes active in the embryo proper, however, are functionally similar to those active in scarlet runner bean and common bean embryo proper regions. We uncovered a set of suspensor- and embryo proper–specific transcription factors (TFs) that are shared by all embryos irrespective of morphology, suggesting that they are involved in early differentiation processes common to all plants. Chromatin immunoprecipitation sequencing (ChIP-Seq) experiments with scarlet runner bean and soybean WOX9, an up-regulated suspensor TF, gained entry into a regulatory network important for suspensor development irrespective of morphology.

suspensor | embryo proper | transcriptome | laser capture microdissection | ChIP-Seq

One of the major unsolved questions in plant biology is how regulatory networks embedded in the genome choreograph processes leading to the specification and differentiation of different embryonic regions following zygote formation. In most higher plants, the zygote divides asymmetrically giving rise to an embryo consisting of two regions with distinct developmental fates—the embryo proper and the suspensor (1–3) (Fig. 1). The embryo proper differentiates into regions that enable the next plant generation to develop following seed germination. These include shoot and root meristems that generate the plant body, and cotyledons which serve as an energy source until the germinating seedling is able to survive on its own via photosynthesis (1). By contrast, the suspensor is a terminally differentiated embryonic region that anchors the embryo proper to surrounding seed tissue and degenerates by the time embryogenesis is complete (4–7). One hundred and forty years ago it was known that the suspensor varies greatly in morphology among different plant species, in contrast with the less variant embryo proper (8). Enlarged highly specialized suspensors were speculated to produce substances required for early embryo development (8, 9)—a hypothesis that has stood the test of time (4–7). A series of elegant experiments has illuminated the signaling pathways and regulators responsible for establishing zygotic polarity and directing the embryo to follow embryo proper and suspensor differentiation pathways (3, 10–12). However, most of the regulatory genes and genomic wiring that control these processes are largely unknown.

Legume embryos exhibit a wide spectrum of suspensor sizes and shapes (8, 13). For example, scarlet runner bean (SRB) (*Phaseolus coccineus*) and the common bean (CB) (*Phaseolus vulgaris*) have large multicellular suspensors (Fig. 1 *A–E*) that contain polytene chromosomes resembling those of *Drosophila* salivary glands, suggesting high metabolic activity (14, 15). These closely related legumes are separated by only 1 to 2 my (Fig. 1*J*), and have similar genome sizes and chromosome numbers (16–19). By contrast, soybean (SB) (*Glycine max*), a more distant legume (Fig. 1*J*), has smaller and less specialized suspensors (Fig. 1 *F* and *G*) resembling those of *Arabidopsis thaliana* (Fig. 1 *H*

Significance

How plant embryos are differentiated into embryo proper and suspensor regions following fertilization is a major unanswered question. The suspensor is unique because it can vary in morphology in different plant species. We hypothesized that regulatory genes controlling the specification of embryo proper and suspensor regions should be shared by all plants irrespective of embryo morphology. We compared embryo proper and suspensor transcriptomes of plants with distinct suspensor morphologies. Scarlet runner bean and common bean have highly specialized giant suspensor regions; soybean and *Arabidopsis* suspensors are smaller and less specialized. We uncovered a small set of embryo proper– and suspensor-specific transcription factors shared by all embryos irrespective of morphology, suggesting that they play an important role in early embryo differentiation.

Fig. 1. Plant species and embryos used for suspensor and embryo proper transcriptome experiments. (*A*) SRB open flower. (*B*) Plastic section (1 μm) of SRB globular-stage embryo. (*C*) Hand-dissected SRB (Pc), *Arabidopsis* (At), and SB (Gm) globular-stage embryos. (*D*) CB open flower. (*E*) Paraffin section (6 μm) of CB globular-stage embryo. (*F*) SB open flower. (*G*) Paraffin section (6 μm) of SB globular-stage embryo. (*H*) *Arabidopsis* open flower. (*I*) Paraffin section (6 μm) of *Arabidopsis* globular-stage embryo. (*J*) Evolutionary relationships between SRB, CB, SB, and *Arabidopsis* (mya, million years ago) (16, 57, 58). Embryo cartoons are not drawn to scale and were traced from paraffin sections of globular-stage embryos. Green and yellow colors indicate the embryo proper and suspensor, respectively. (*K–N*) Paraffin sections (6 μm) of SRB (*K*), CB (*L*), SB (*M*), and *Arabidopsis* (*N*) globular-stage embryos before and after capture by LCM. EP, embryo proper; SUS, suspensor.

and *I*). The large size of SRB and CB embryos permits hand dissection of embryo proper and suspensor regions for experimental studies (*SI Appendix*, Fig. S1*A*), which is not practical in SB and *Arabidopsis* (20). Almost five decades ago Ian Sussex, Mary Clutter, Ed Yeung, and their colleagues took advantage of this property to show that giant SRB and CB suspensors are highly active and supply the embryo proper with substances responsible for embryo proper growth (21–23), sustaining the ideas of pioneering plant embryologists of the late 19th and early 20th centuries (8, 9). Subsequently, experiments with isolated SRB and CB suspensors showed that they contain several hormones, including gibberellic acid (GA), auxin, cytokinin (CK), and abscisic acid (ABA) (24–27), suggesting that hormone signaling might play a critical role in suspensor differentiation and function (24).

We have been using giant SRB embryos as a system to dissect the genomic processes that control suspensor and embryo proper differentiation (28–31). We identified genes that are up-regulated in the suspensor shortly after fertilization, several of which encode enzymes in the GA biosynthetic pathway (28, 30). We dissected the regulatory regions of two suspensor-specific genes—*G564*, encoding a protein of unknown function, and *GA20-oxidase* (*GA20ox*), specifying an enzyme in the GA biosynthetic pathway (28, 31). Our experiments uncovered a *cis*-regulatory module containing five *cis*-elements that are each required for suspensor-specific transcription of the *G564* and *GA20ox* genes. Studies with transgenic plants showed that this regulatory module works in the suspensors of distantly related tobacco and *Arabidopsis* embryos (6), suggesting a conserved suspensor-specific regulatory pathway across the plant kingdom.

Here, we take advantage of the morphological differences between several plant embryos and use laser capture microdissection (LCM) and RNA sequencing (RNA-Seq) to characterize SRB, CB, SB, and *Arabidopsis* embryo proper and suspensor transcriptomes (Fig. 1). Our hypothesis is that irrespective of embryo morphology there is a shared set of embryo proper and suspensor transcription factors (TFs) that drive the differentiation of these regions and are conserved in higher plants. Our experiments uncovered 1) the spectrum of functional differences between the suspensor and embryo proper in each plant species, 2) a high degree of metabolic specialization in large SRB and CB suspensors, and 3) a small set of embryo proper– and suspensor-specific TFs common to all plant

species investigated that might play a major role in early embryo differentiation. Finally, SRB and SB chromatin immunoprecipitation sequencing (ChIP-Seq) experiments with one of the shared suspensor-specific TF mRNAs, WUSCHEL-RELATED HOMEOBOX 9 (WOX9), uncovered several TFs that are putative WOX9 targets, some of which are conserved between both plants. How embryo region–specific TFs are integrated into circuitry required for early embryo differentiation and function remain to be determined.

Results

CB Can Be Used as an SRB Reference Genome. We examined whether we could use CB as a reference genome for SRB expression data because these two legumes are separated by only 1 to 2 my (Fig. 1*J*), and an excellent CB genome draft exists (32). We generated several hundred thousand SRB expressed sequence tags (ESTs) from hand dissected globular-stage embryo proper and suspensor regions (*SI Appendix*, Fig. S1*A*) and aligned these ESTs with the predicted transcripts of CB genes. Approximately 15,000 diverse transcripts were represented in our embryo EST population, and there was >95% similarity between SRB and CB sequences (*SI Appendix*, Fig. S1 *B* and *C*), reflecting the close evolutionary relationship between SRB and CB at the gene level.

We used LCM to isolate embryo proper and suspensor regions from SRB globular-stage embryos (Fig. 1*K* and *SI Appendix*, Fig. S1*D*) and uncovered the spectrum of transcripts present in each of these embryonic regions using RNA-Seq. Approximately 95% of our EST collection was represented in the RNA-Seq reads (*SI Appendix*, Fig. S1*E*). To ensure that the CB genome can be used as a reference for SRB expression data, we sequenced the SRB genome at ~100× coverage (*SI Appendix*, Fig. S1*F*) and compared the alignments of RNA-Seq reads with both CB and SRB genome sequences (*SI Appendix*, Fig. S1*G*). The same number of mappable reads was obtained with both genomes (*SI Appendix*, Fig. S1*G*), indicating that the CB draft genome and annotated genes can be used as a reference for analyzing SRB RNA-Seq data.

Absence of Contamination from Surrounding Seed Tissue in Embryo Proper and Suspensor Regions Captured by LCM. We used SRB embryo medial sections to avoid potential contamination from surrounding seed tissue during embryo proper and suspensor

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies
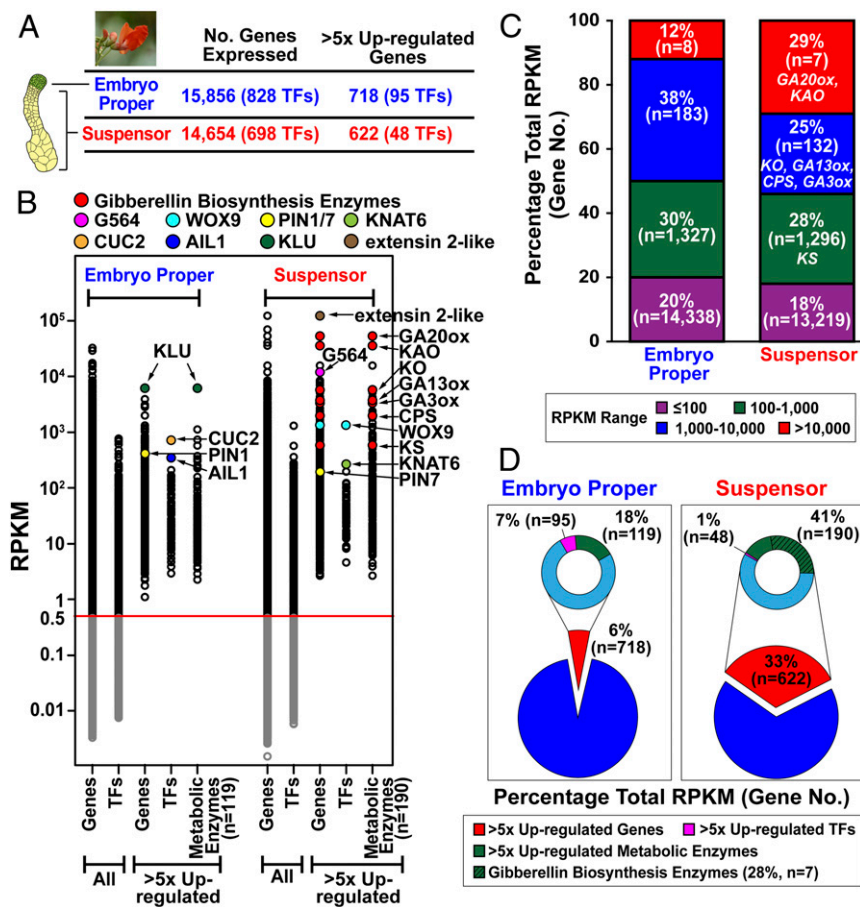
region LCM (Fig. 1K) (*SI Appendix, Materials and Methods*) (20, 33). We searched our embryo proper and suspensor RNA-Seq reads for endothelium-specific transcripts (G563) (*SI Appendix,* Fig. S2A), seed coat–specific transcripts (GA2-oxidase) (*SI Appendix,* Fig. S2B), and endosperm-specific transcripts (AGL62) that we (G563 and GA2-oxidase) (28, 30), and others (AGL62) (34), identified previously. We compared the RNA-Seq reads for these transcripts with those of ent-kaurene oxidase (KO), a suspensor-specific mRNA (Fig. 2 and *SI Appendix,* Fig. S2C) (28). There were few, if any, RNA-Seq reads for surrounding seed tissue transcripts in comparison with the KO control (*SI Appendix,* Fig. S2 *D and E*), indicating that our embryo proper and suspensor LCM sections had little, or no, contaminating transcripts from adjacent seed tissues.

SRB Embryo Proper and Suspensor Transcriptomes Contain a Spectrum of mRNAs. Approximately 15,000 diverse mRNAs were present in each embryo region, including 700 to 800 TF mRNAs (Fig. 2A and Dataset S1). The union of these mRNA sets indicated that there were 17,500 diverse mRNAs in the embryo as a whole, including 1,000 TF mRNAs and a small set of transcripts specific for embryo proper and suspensor regions. We filtered the RNA-Seq reads to include only those with reads per kilobase of transcript per million reads mapped (RPKM) values >0.5 (Fig. 2B). We estimated that this

criterion scored mRNAs at functionally meaningful mRNA levels of more than one molecule/cell (35, 36). Both the embryo proper and suspensor mRNA populations spanned a wide range of prevalences consistent with those found in plant embryos (37). TF mRNA prevalences also spanned several orders of magnitude, suggesting that a range of regulatory inputs is required by each embryo region (Fig. 2B).

A small fraction of both the embryo proper and suspensor mRNA mass (~20%) contained most of the diverse mRNAs (Fig. 2C). On average, these transcripts were represented only a few times per cell (~15) and resembled a complex class of rare mRNAs typical of plant embryo cells (37). The suspensor, however, contained a large fraction of highly prevalent mRNAs present in tens of thousands of copies per cell as compared with the embryo proper (Fig. 2 *B* and *C*). Almost 30% of the suspensor mRNA mass contained only seven different sequences, including those encoding the GA biosynthesis enzymes GA20ox and KAO (Fig. 2C). In fact, GA20ox and KAO mRNAs were among the most prevalent mRNAs in the entire embryo (Fig. 2 *B* and *C*), suggesting a high degree of metabolic specialization within the suspensor.

SRB Embryo Proper and Suspensor Regions Contain Specific mRNA Sets. We used EdgeR (false discovery rate [FDR] <0.05) to identify mRNAs that were more than fivefold more prevalent, or up-regulated,

Fig. 2. Gene activity in SRB suspensor and embryo proper regions. (A) Genes active (>0.5 RPKM) in at least one biological replicate, and more than fivefold up-regulated genes in embryo proper and suspensor (*SI Appendix, Materials and Methods*). All expressed genes and those up-regulated more than fivefold are listed in Dataset S1. TFs, transcription factor mRNAs. (B) The abundance distribution of mRNAs in different embryo proper and suspensor mRNA populations represented by RPKMs. Each circle represents a different mRNA sequence. RNAs with ≤0.5 RPKM in both biological replicates were not used in subsequent analyses, as indicated by the gray circles below the red line. Enzyme mRNAs were identified using the Plant Metabolic Pathway Database (https://plantcyc.org/). (C) Percentage of embryo proper and suspensor mRNA populations with different prevalence levels. Parentheses indicate the number of different mRNAs in each prevalence group. (D) Percentage of total RPKM for different embryo proper and suspensor mRNA populations. Parentheses represent the number of mRNAs in each population. Gene abbreviations are defined in *SI Appendix,* Table S1.

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies

PNAS | 3 of 12
https://doi.org/10.1073/pnas.2024704118

in each embryo region (*SI Appendix, Materials and Methods*). We uncovered 718 and 622 embryo proper– and suspensor-specific mRNAs, respectively (Fig. 2*A* and Dataset S1) and then identified mRNAs encoding metabolic enzymes and TFs in each up-regulated mRNA set (Fig. 2 *B* and *D* and Dataset S1). Region-specific mRNAs spanned a range of prevalences analogous to those in the unselected populations (Fig. 2*B*). Suspensor-specific mRNAs, however, occupied a greater proportion of mRNA mass than their embryo proper counterparts (33% vs. 6%), reflecting the presence of highly prevalent mRNAs (Fig. 2 *B–D*). Over 40% of the suspensor-specific mRNA mass consisted of 190 diverse metabolic enzyme mRNAs, including those encoding GA biosynthesis enzymes (Fig. 2*D*). By contrast, only 18% of the embryo proper–specific mRNA mass encoded 119 metabolic enzymes (Fig. 2*D*), indicating that a greater proportion of gene activity within the suspensor is directed toward specialized metabolic processes.

In contrast with mRNAs involved in metabolism (Fig. 2*D*), TF mRNAs represented a larger fraction of the embryo proper up-regulated mRNA set compared with the suspensor (7% vs. 1%) (Fig. 2*D*). We uncovered 95 and 48 diverse TF mRNAs specific to the embryo proper and the suspensor, respectively (Fig. 2*A* and Dataset S1), representatives of which are listed in *SI Appendix*, Fig. S3. The precise roles that the majority of these region-specific TF mRNAs perform within the SRB embryo proper and suspensor at the globular stage are unknown. However, they include TF mRNAs with known developmental and hormone signaling functions (*SI Appendix*, Fig. S3*A* and Dataset S1). For example, embryo proper–specific TF mRNAs include those involved in shoot meristem development (SHOOT MERISTEMLESS [STM]) and cotyledon separation (CUP-SHAPED COTYLEDON2 [CUC2]) (12), among others (*SI Appendix*, Fig. S3*B* and Dataset S1). By contrast, the suspensor-specific TF mRNA set contains WOX9, AUXIN RESPONSE FACTOR16 (ARF16), and HOMEODOMAIN GLABROUS11 (HDG11) TFs that are required for suspensor specification (*SI Appendix*, Fig. S3*C* and Dataset S1) (3). The most prevalent up-regulated TF mRNAs in the embryo proper and suspensor were CUC2 and WOX9, respectively (Fig. 2*B*). A higher percentage (80%) of embryo proper–specific TFs play a role in developmental and hormone response processes compared with their counterparts in the suspensor (40%) (*SI Appendix*, Fig. S3*A*), indicating significant functional differences between the embryo proper– and suspensor-specific TF mRNA populations.

SRB Embryo Proper and Suspensor Regions Differ Significantly in Biological Processes. We performed Gene Ontology (GO) analysis (FDR <0.05) on up-regulated embryo proper– and suspensor-specific mRNAs to characterize the major functions that are carried out in each embryo region (Fig. 3). A summary of up-regulated mRNAs in major GO categories is presented in *SI Appendix*, Fig. S4 and Dataset S2 which contains a list of all embryo proper and suspensor GO terms.

Embryo proper. The most significant GO terms reflected the cell division, differentiation, and developmental regulatory processes occurring in this embryo region during the globular stage. These included regulation of transcription, transcription factor activity, histone and DNA methylation, cytokinesis by cell plate formation, and cell proliferation, among others (Fig. 3 *A–C*). Developmental GO terms, such as meristem initiation and polarity specification of the adaxial/abaxial axis reflected patterning events taking place within the embryo proper region. The most significant metabolic process GO terms were carbohydrate, lipid, and cell wall macromolecule (polysaccharide) biosynthesis, although less so than regulatory events. Finally, signaling pathways were prominent among embryo proper GO terms, including polar transport and response to a spectrum of hormones, such as GA, jasmonic acid (JA), CK, ABA, and auxin, the latter hormone being transported within the embryo proper in a basal

direction toward the suspensor by the PIN FORMED1 (PIN1) transporter to form an auxin gradient that triggers root pole differentiation within the embryo proper (Fig. 2*B* and *SI Appendix*, Fig. S5) (38). GO terms for other transporters, including those for lipids and metal ions, were also overrepresented in the embryo proper–specific mRNA set (Fig. 3*A*). Specific embryo proper mRNAs encoding these transporters, and others (e.g., PIN1), are summarized in *SI Appendix*, Fig. S5.

Suspensor. A striking aspect of the suspensor was the large number of GO terms reflective of processes related to metabolism, hormone synthesis, and transport, in contrast with the embryo proper. For example, oxidation reduction process and catalytic activity were the most significant biological process and molecular function GO terms, respectively (Fig. 3 *A* and *C*). By contrast, there were fewer regulatory and developmental GO terms, and cell proliferation GO terms were absent, reflecting the cessation of suspensor cell division at the globular stage (22). Major metabolic pathways such as glycolysis, pentose phosphate shunt, and the synthesis of several hormones, including JA, GA, and indoleacetic acid (auxin), were significant GO terms (Fig. 3*A*). GO terms for response to these hormones were also observed (Fig. 3*A*), indicating the presence of hormone signal transduction pathways within the suspensor. The plastid cellular component GO term reflected the location of several major metabolic processes (e.g., GA biosynthesis) within this organelle (Fig. 3*D*). Finally, transmembrane, golgi organization, water, ion, amino acid, and oligopeptide transport were major GO terms (Fig. 3 *A* and *B*). These were encoded by many up-regulated suspensor-specific transporter mRNAs distinct from those present in the embryo proper, including PIN FORMED7 (PIN7) which is involved in auxin transport and suspensor development (Fig. 2*B* and *SI Appendix*, Fig. S5) (3).

SRB Suspensor mRNAs Encoding Enzymes in Several Interconnected Biosynthetic Pathways Leading to Hormone Production Are Up-Regulated.
Methyl erythritol-4-phosphate isoprenoid and GA pathways. mRNAs encoding enzymes in the methyl erythritol-4-phosphate (MEP) isoprenoid pathway were up-regulated within the suspensor (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S6*A*). MEP mRNAs, together with those required for GA biosynthesis (Figs. 2*B* and 4*C* and *SI Appendix*, Fig. S6*B*), indicated that mRNAs encoding enzymes in the metabolic pathway from pyruvate to bioactive GA1 and GA4 were up-regulated within the suspensor (Fig. 4 *B* and *C*).

We examined mRNAs specifying proteins in the GA signal transduction pathway (Fig. 4 *A* and *D*). DELLA mRNAs, such as those encoding GIBBERELLIN INSENSITIVE (GAI) and REPRESSOR OF GA-LIKE1 and 2 (RGL1 and RGL2), were up-regulated in the suspensor (Fig. 4*D*). mRNAs encoding the GA receptor, GIBBERELLIN INSENSITIVE DWARF1A (GID1A), and GA transporters, NITRATE TRANSPORTER5.4 (NPF5.4) and PEPTIDE TRANSPORTERS1 and 3 (PTR1 and PTR3), were up-regulated in the suspensor as well (Fig. 4*D*). Significantly, different paralogs of the DELLA GAI and RGA1 mRNAs and GA receptor GID1A and GID1B mRNAs were up-regulated in the embryo proper (Fig. 4*D*). These data suggest that GA is 1) synthesized in the suspensor, 2) transported to the embryo proper (Fig. 4*A*), and 3) elicits responses in both embryonic regions (Fig. 3). This is consistent with experiments carried out decades ago that detected the presence of bioactive GA in the SRB suspensor (39) and suggested that GA moves from the suspensor to the embryo proper affecting its development (40, 41).

Glycolysis and pentose phosphate shunt pathways. Plastid-localized glycolysis and the pentose phosphate shunt utilize starch as a substrate and are required to produce precursor molecules for the isoprenoid and GA biosynthetic pathways, and other hormones such as JA, CK, ABA, and auxin (*SI Appendix*, Fig. S7). We examined mRNAs encoding enzymes in glycolytic and pentose phosphate shunt pathways, and found that one or more
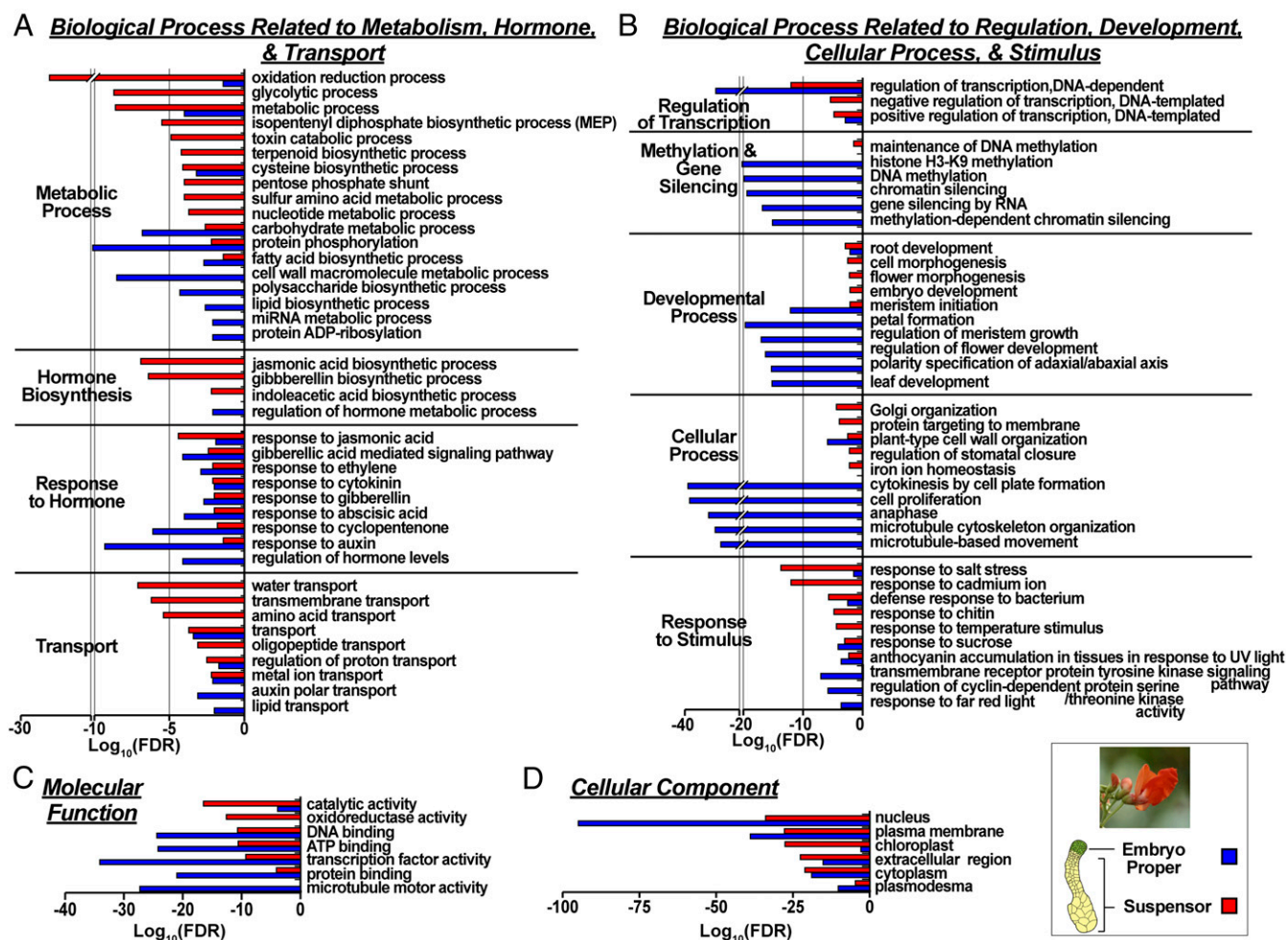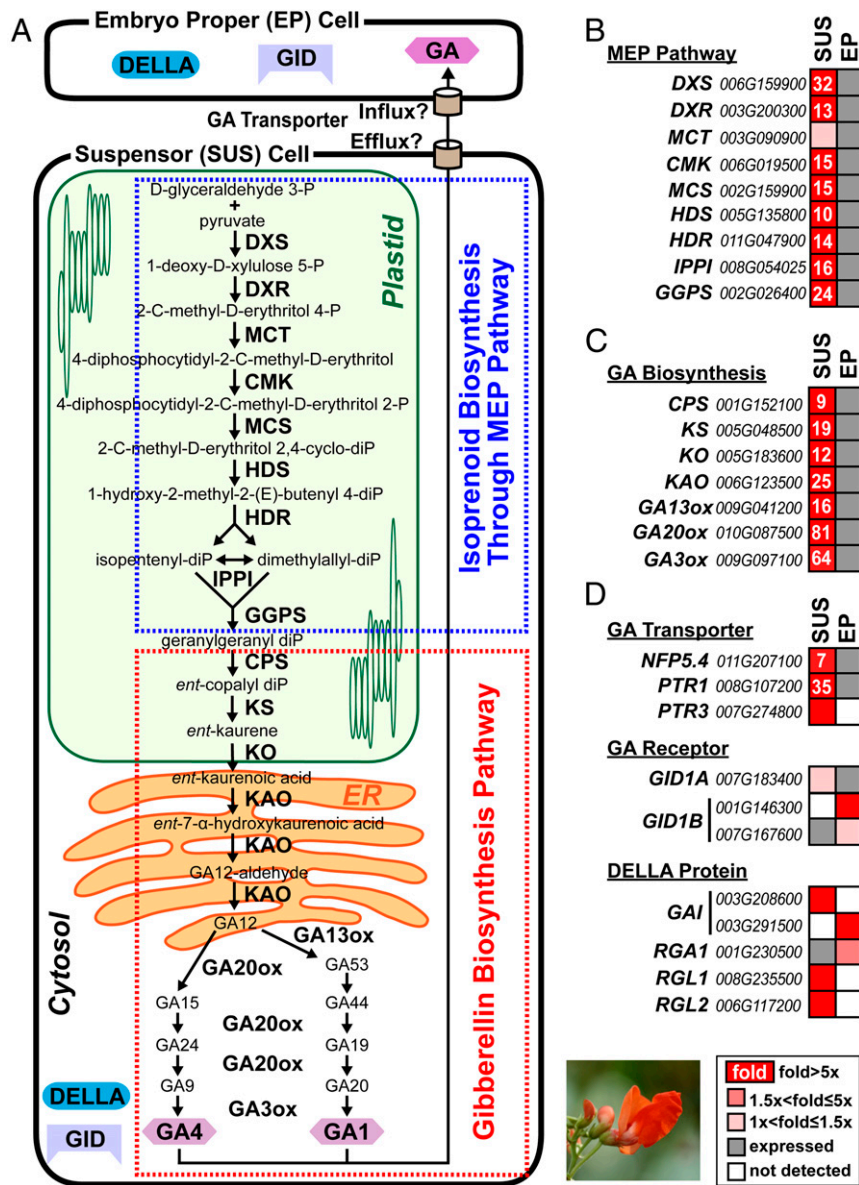
Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos
with different morphologies

Fig. 3. GO terms that are enriched in SRB globular-stage embryo proper and suspensor up-regulated mRNAs. (*A* and *B*) Enriched biological process GO terms related to metabolism, hormone, and transport (*A*) and regulation, response to regulation, development, cellular process, and stimulus (*B*). (*C* and *D*) Enriched molecular function (*C*) and cellular component (*D*) GO terms. Only the top 5 to 10 GO terms are listed. All GO terms are listed in an interactive format in Dataset S2.

paralogs of these enzyme mRNAs were either up-regulated or detected within the suspensor (*SI Appendix*, Figs. S6 *C* and *D* and S8). For example, phosphoglucose isomerase (GPI) and glucose-6-phosphate dehydrogenase (G6PD) mRNAs, encoding rate-limiting enzymes of the glycolytic and pentose phosphate shunt pathways, respectively, were up-regulated more than fivefold (*SI Appendix*, Figs. S6 *C* and *D* and S8). In addition, mRNAs for the first three enzymes in starch biosynthesis were up-regulated (e.g., adenosine diphosphate [ADP] glucose pyrophosphorylase large subunit) (Datasets S1 and S2), and starch granules are present in SRB suspensor cells (21). These data suggest a remarkable co-ordination of metabolic events within the suspensor beginning with starch biosynthesis and culminating in pyruvate, glyceraldehyde-3-phosphate, and chorismite precursors required for several hormone biosynthetic pathways (*SI Appendix*, Fig. S7).

JA, auxin, ABA, and CK hormone pathways. We detected suspensor GO terms for JA and auxin biosynthesis (Fig. 3), suggesting that mRNAs encoding enzymes in these hormone pathways were also up-regulated in the suspensor. Previously, others demonstrated the presence of auxin, ABA, and CK within the SRB suspensor (24, 26, 27), although their sites of synthesis within the embryo were not known. We examined mRNAs encoding enzymes in the JA, auxin, ABA, and CK biosynthetic pathways to determine whether the suspensor had the potential for synthesizing these hormones (*SI Appendix*, Figs. S6 *E–I*, S9, and S10). In each

hormone biosynthesis pathway, all of the enzyme mRNAs required for catalytic steps from precursor to final product were detected in the suspensor (*SI Appendix*, Figs. S6 *E–I*, S9, and S10), and many were up-regulated, although not to the levels of isoprenoid and GA biosynthesis mRNAs (*SI Appendix*, Fig. S6). We also detected receptor, regulatory, and transporter mRNAs for JA, auxin, ABA, and CK in the suspensor mRNA population (*SI Appendix*, Figs. S9 and S10). For example, several mRNA paralogs of the JA receptor JASMONATE ZIM-DOMAIN PROTEIN (JAZ) were detected, and one, JAZ10, was up-regulated 10-fold compared with the embryo proper (*SI Appendix*, Fig. S9). In addition, JA regulator mRNAs, NOVEL INTERACTOR OF JAZ (NINJA) and CORONATINE IN-SENSITIVE1 (COI1), were also up-regulated in the suspensor (*SI Appendix*, Fig. S9). Similarly, the auxin receptor, TRANS-PORT INHIBITOR RESPONSE 1 (TIR1) mRNA, and the auxin response TF mRNA, ARF16, were up-regulated within the suspensor along with the PIN7 efflux carrier mRNA (*SI Appendix*, Figs. S5 and S9). These data, together with hormone response GO terms (Fig. 3), suggest that, in addition to GA, the SRB suspensor has the machinery for synthesizing, transporting, and utilizing JA, ABA, CK, and auxin in signal transduction pathways.

Gene Expression Activities in the CB Embryo Proper and Suspensor Are Similar to Those in SRB. Developmental and ultrastructure studies by others showed that CB and SRB embryos are

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies

PNAS | 5 of 12
https://doi.org/10.1073/pnas.2024704118

Fig. 4. GA biosynthesis and isoprenoid pathway mRNAs in SRB suspensor and embryo proper regions. (*A*) GA and MEP pathways, and fold-change levels (number in red squares) of pathway mRNAs. GA pathway, MEP pathway, and enzyme intracellular locations were taken from published information (59–62). Enzyme intracellular localizations were confirmed using the DeepLoc machine learning tool (63) (*SI Appendix, Materials and Methods* and Dataset S3). Suspensor and adjacent embryo proper cell model is based on representation of enzyme, receptor, and transporter mRNAs in these embryo regions (*A* and *C*). GA efflux and influx from the suspensor to the embryo proper is based on classical experiments with SRB embryos, suggesting that GA is transferred from the suspensor to the embryo proper (41, 64). (*B*) The interaction of GA, GA receptors (GID), and DELLA in GA signaling pathway (65). (*C*) Representation of GA transporter, receptor, and DELLA mRNAs in suspensor and embryo proper. (*D*) GA biosynthesis enzyme, MEP pathway enzyme, and GA signaling protein abbreviations. Enzyme abbreviations are defined in *SI Appendix*, Table S1.

indistinguishable from each other, except that the CB embryo is slightly smaller (Fig. 1 *B* and *E*) (42, 43). We used LCM to capture CB globular-stage embryo proper and suspensor regions (Fig. 1*L*), and then profiled their mRNAs using RNA-Seq (Datasets S1 and S2). The embryo proper and suspensor mRNA populations were similar to corresponding SRB mRNAs in every feature, including: 1) similar numbers of genes (16,500) and TFs (900) (*SI Appendix*, Fig. S11*A*); 2) small sets of embryo proper– and suspensor-specific mRNAs (*SI Appendix*, Fig. S11*A*); 3) abundance distributions (*SI Appendix*, Fig. S11*B*); 4) presence of highly prevalent suspensor mRNAs, including those encoding GA20ox and KAO (*SI Appendix*, Fig. S11 *B* and *D*); 5) profile of significant GO terms, including embryo proper enrichment for developmental, transcriptional, and cell proliferation processes and suspensor enrichment for metabolic and hormone biosynthesis processes (*SI Appendix*, Fig. S12 and Dataset S2); 6) up-regulation of metabolic pathway mRNAs, such as those for MEP, GA biosynthesis, pentose phosphate shunt, and glycolysis (*SI Appendix*, Figs. S11*C* and S13*A* and Dataset S1); 7) presence of GA signaling protein mRNAs in the suspensor (*SI Appendix*, Fig. S13*A*); 8) similar distribution of embryo proper– and suspensor-specific transporters (*SI Appendix*, Fig. S13 *B–D*); and 9) similar profiles of embryo proper– and suspensor-specific TF mRNAs, including the presence of STM and WOX9 mRNAs in the embryo proper and suspensor, respectively, among others (*SI Appendix*, Fig. S11 *E–G* and Dataset

S1). Together, these data indicate that SRB and CB globular-stage embryos are virtually indistinguishable in their gene expression profiles and specialized activities, as predicted by their similar morphology and close evolutionary relationship (Fig. 1).

Gene Expression Activities in SB and *Arabidopsis* Globular-Stage Embryos Differ from Those in SRB and CB. SB and *Arabidopsis* suspensor regions are much simpler than those of SRB and CB (Fig. 1). At the globular stage, the *Arabidopsis* suspensor consists of a linear file of five to seven small cells ~200 times smaller than SRB and CB suspensors (Fig. 1 *C* and *I*) (44). On the other hand, the SB suspensor consists of a small collection of cells shaped in a V-like structure at the globular stage (Fig. 1*G*) that will elongate into a narrow column with 10 tiers of small cells at the heart stage (45), and is much reduced in size and shape compared with SRB and CB suspensors (Fig. 1 *B*, *C*, and *E*) (13). We used LCM to capture SB and *Arabidopsis* embryo proper and suspensor regions (Fig. 1 *M* and *N*), sequenced each mRNA population
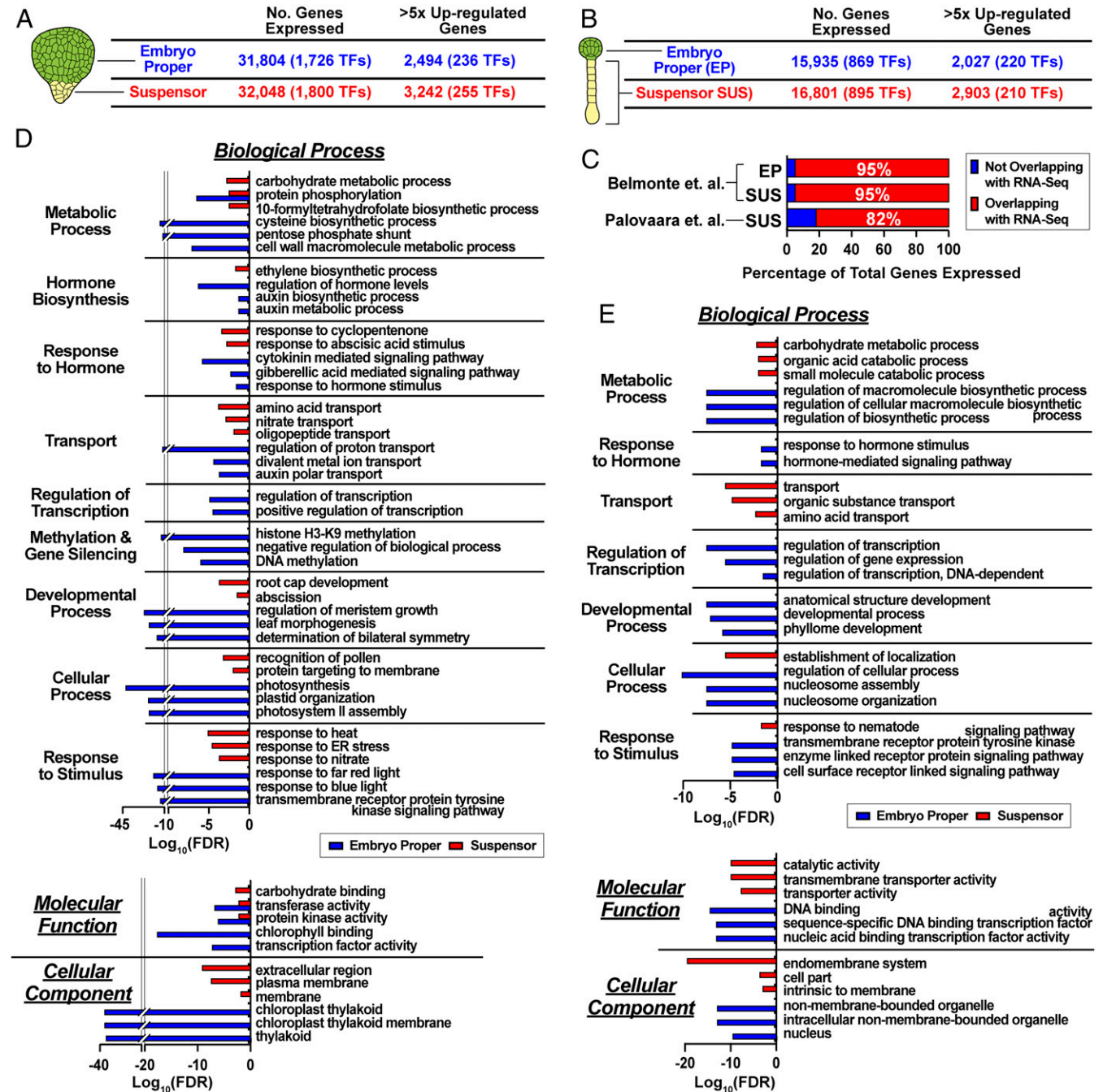
Fig. 5. Gene activity in SB and *Arabidopsis* suspensor and embryo proper regions. (*A* and *B*) SB (*A*) and *Arabidopsis* (*B*) genes active (>0.5 RPKM) in at least one biological replicate and more than fivefold up-regulated genes in embryo proper and suspensor (*SI Appendix, Materials and Methods*). (*C*) Percentage of *Arabidopsis* embryo proper and suspensor mRNAs that overlap with datasets published previously by Goldberg-Harada laboratories (33) and others (46). (*D* and *E*) Enriched biological process, cellular component, and molecular function GO terms for up-regulated SB (*D*) and *Arabidopsis* (*E*) embryo proper and suspensor mRNAs. Only the top three GO terms for SB and *Arabidopsis* are listed. All expressed genes and those up-regulated more than fivefold are listed in Dataset S1. GO terms are listed by functional categories in Dataset S2.

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies

PNAS | 7 of 12
https://doi.org/10.1073/pnas.2024704118

using RNA-Seq, and selected for more than fivefold up-regulated embryo proper– and suspensor-specific mRNA sets, including those encoding TFs (Fig. 5 *A* and *B* and Dataset S1). We obtained 95% overlap with *Arabidopsis* embryo proper and suspensor mRNA sequences identified by us previously using GeneChip technology and the same initial cDNAs (Fig. 5*C*) (33) (*SI Appendix, Materials and Methods*). In addition, there was 82% overlap with suspensor nuclear RNA sequences generated by others using RNA-Seq, indicating that we have a good representation of *Arabidopsis* globular embryo region mRNAs (46).

Each mRNA population had a wide range of prevalences, including TF mRNAs and more than fivefold region-specific sets (*SI Appendix*, Fig. S14), analogous to what was observed with SRB and CB embryo mRNA populations (Fig. 2*B* and *SI Appendix*, Fig. S11*B*). Significantly, highly prevalent GA biosynthesis mRNAs were not present in the *Arabidopsis* and SB suspensor up-regulated mRNA sets (*SI Appendix*, Fig. S14), in marked contrast with SRB and CB mRNA suspensor populations (Fig. 2*B* and *SI Appendix*, Fig. S11*B*). We generated GO terms using the up-regulated SB and *Arabidopsis* embryo proper and suspensor mRNAs (Fig. 5 *D* and *E* and Dataset S2). The spectrum of GO terms for the SB and *Arabidopsis* embryo proper regions was similar to those obtained with SRB and CB (Fig. 3 and *SI Appendix*, Fig. S12). For example, the most significant embryo proper GO terms were those involved in regulation of transcription, developmental processes, and response to hormone stimulus, among others (Fig. 5 *D* and *E*). By contrast, GO terms obtained with SB and *Arabidopsis* suspensor regions differed significantly from those obtained with SRB and CB (Fig. 3 and *SI Appendix*, Fig. S12). The most significant SB and *Arabidopsis* GO terms, such as those for carbohydrate metabolic process and transport (Fig. 5 *D* and *E*), were similar to those obtained with SRB and CB suspensor mRNAs (Fig. 3 and *SI Appendix*, Fig. S12). Missing, however, were GO terms for specialized metabolic processes and pathways, such as oxidation reduction process, glycolysis, pentose phosphate shunt, MEP pathway, and several hormone biosynthesis pathways (e.g., GA and JA), among others, which were hallmarks for SRB and CB suspensors (Fig. 3 and *SI Appendix*, Fig. S12). Together these data indicate that, on a functional level, the spectrum of processes carried out by the embryo proper region of all plants investigated was similar irrespective of embryo size and morphology. However, major differences occurred between the small and relatively simple *Arabidopsis* and SB suspensor regions on the one hand, and the giant, specialized SRB and CB suspensors on the other.

Identification of a Set of Shared Embryo Proper and Suspensor TF mRNAs. We searched the up-regulated embryo proper and suspensor populations for TF mRNAs that were specific to the embryo regions of all plants investigated (Fig. 6 and *SI Appendix*, Table S2). For this comparison we used a stringent criterion that included 1) more than fivefold up-regulation and 2) concordance between all biological replicates (*SI Appendix, Materials and Methods*). We obtained two small sets of globular-stage embryo proper– and suspensor-specific TF mRNAs that were up-regulated in SRB, CB, *Arabidopsis*, and SB embryo regions (Fig. 6 *A* and *B*, and *SI Appendix*, Table S2). Suspensor TF mRNAs included the known regulators WOX8/9, HDG11, and ARF16 (Fig. 6*A*) (12). On the other hand, embryo proper TF mRNAs included those encoding CUC2, HANABA TANARU (HAN), and TARGET OF MONOPTEROS5-LIKE1 (TMO5L1), among others (Fig. 6*B*). These TF mRNAs are involved in shoot meristem, root pole, and vascular tissue differentiation processes, within the embryo proper, respectively (12). STM TF mRNA was specific for SRB, CB, and SB embryo proper regions, but was absent for unknown reasons in the *Arabidopsis* embryo proper–specific mRNA set, although it was up-regulated within our *Arabidopsis* GeneChip population (33).

Several WOX9 TF Targets Are Shared by SRB and SB. *Arabidopsis WOX8*, and its close relative *WOX9*, play important roles in suspensor differentiation (12). *WOX8/9* gene relatives in SRB, CB, and SB most closely resembled *Arabidopsis WOX9* by phylogenetic analysis (Fig. 6*C*), but *WOX8* by their suspensor-specific expression patterns (Fig. 6 *D–G*) (12). We named the SRB, CB, and SB relatives as *WOX9*, and considered these genes as functionally equivalent to *Arabidopsis WOX8*.

We carried out ChIP-Seq experiments with SRB and SB WOX9 peptide antibodies to 1) uncover WOX9 downstream target genes, 2) characterize target functions, and 3) ascertain whether any targets were shared between SRB and SB (Fig. 7) (*SI Appendix, Materials and Methods*). We used whole globular-stage seeds for our experiments, because in both plants 1) WOX9 was the most prevalent suspensor-specific TF mRNA (Fig. 2*B* and *SI Appendix*, Fig. S14*A*) and 2) only present within the suspensor region of the seed (Fig. 6*D* and Harada-Goldberg LCM datasets [seedgenenetwork.net]). We obtained 660 and 178 potential WOX9 target genes in SB and SRB, respectively, including 88 and 21 TF gene targets (Fig. 7*A* and Dataset S4). We defined target genes as those that were at the intersection of WOX9-bound genes and genes up-regulated more than fivefold within the suspensor (*SI Appendix, Materials and Methods*) (47). We carried out GO term analysis on the SB and SRB WOX9 target genes (Fig. 7*B*). The most significant GO term in both plants was regulation of transcription, reflecting the large representation of TF genes in SB and SRB target gene sets (Fig. 7*A*). Significantly, the range of GO terms for SRB WOX9 targets was greater than those for SB, and included GO terms that mirrored many functional activities unique to the SRB suspensor region (Fig. 3). These included oxidation reduction processes and JA biosynthesis, among others (Fig. 7*B* and Dataset S4), but did not include GA biosynthesis.
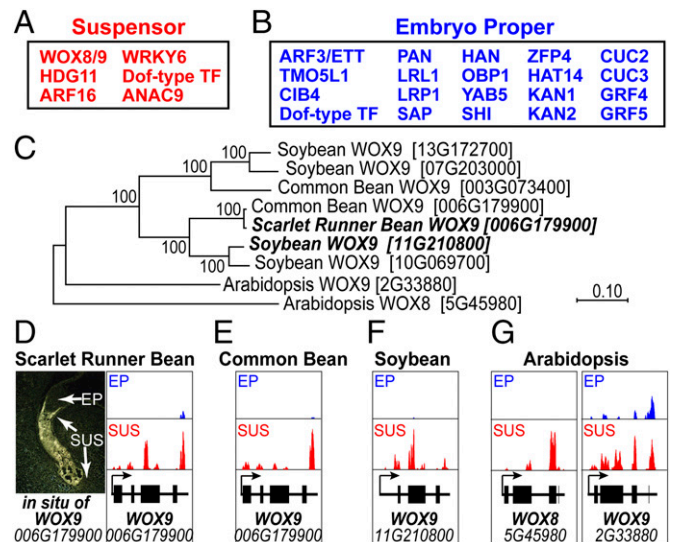
Fig. 6. Transcription factor mRNAs that are up-regulated in the suspensor and embryo proper of all four plant species investigated. (*A* and *B*) Shared up-regulated transcription factor mRNAs in the suspensor (*A*) and embryo proper (*B*). Gene abbreviations and IDs are listed in *SI Appendix*, Tables S1 and S2, respectively. (*C*) Relationships between SRB, CB, SB, and *Arabidopsis WOX9* genes. The phylogenetic trees were constructed using MEGA X based on the protein sequence alignments with the maximum likelihood method and default parameters (66). Bootstrap confidence values are shown as percentages on each branch. SB (Glyma.11G210800) and SRB (Phvul.006G179900) *WOX9* genes used for ChIP-Seq experiments (Fig. 7) are in italics. (*D–F*) Representation of WOX9 mRNAs in SRB (*D*), CB (*E*), and SB (*F*). SRB WOX9 mRNA in situ hybridization image in *D* from ref. 29, copyright American Society of Plant Biologists. (*G*) Representation of *Arabidopsis* WOX8 and WOX9 mRNAs.

We searched for TF target genes that were shared by SB and SRB, and found a small number that included *HDG11* and *ARF16* genes that were present in the suspensor-specific gene set that was shared by all plants investigated (Figs. 6*A* and 7*C*). RNA-Seq and ChIP-Seq genome browser views show 1) the suspensor-specific expression pattern of SB (Fig. 7*D*) and SRB (Fig. 7*E*) *HDG11* and *ARF16* genes and 2) indicate that WOX9 binding peaks were near their transcription start sites (Fig. 7 *D* and *E*). We carried out motif analysis on the SB and SRB WOX9 targets and uncovered two distinct binding motifs for SB and SRB (Fig. 7*F*). Approximately 30% of SB and SRB target genes had one, or both, of these motifs in their upstream regions. One SB motif recognized a Dof-type zinc finger TF (motif 1), while the other a beta helix–loop–helix TF (motif 2) (Fig. 7*F*). By

contrast, both SRB motifs recognized Dof-type zinc finger TFs and were similar to the SB Dof-type zinc finger motif (Fig. 7*F*). One candidate binding to this motif might be the Dof-type zinc finger TF that was up-regulated in all suspensor mRNA populations we investigated (Fig. 6 and *SI Appendix*, Table S2). Together, these data suggest that 1) WOX9 binds to a large number of potential target genes with distinct overall functions in SB and SRB suspensors, 2) several TF gene targets are shared between SB and SRB, and 3) WOX9 appears to form a complex with other TFs in order to bind to a subset of target genes.

Discussion

We compared the embryo proper and suspensor transcriptomes of SRB, CB, SB, and *Arabidopsis* globular-stage embryos. At this
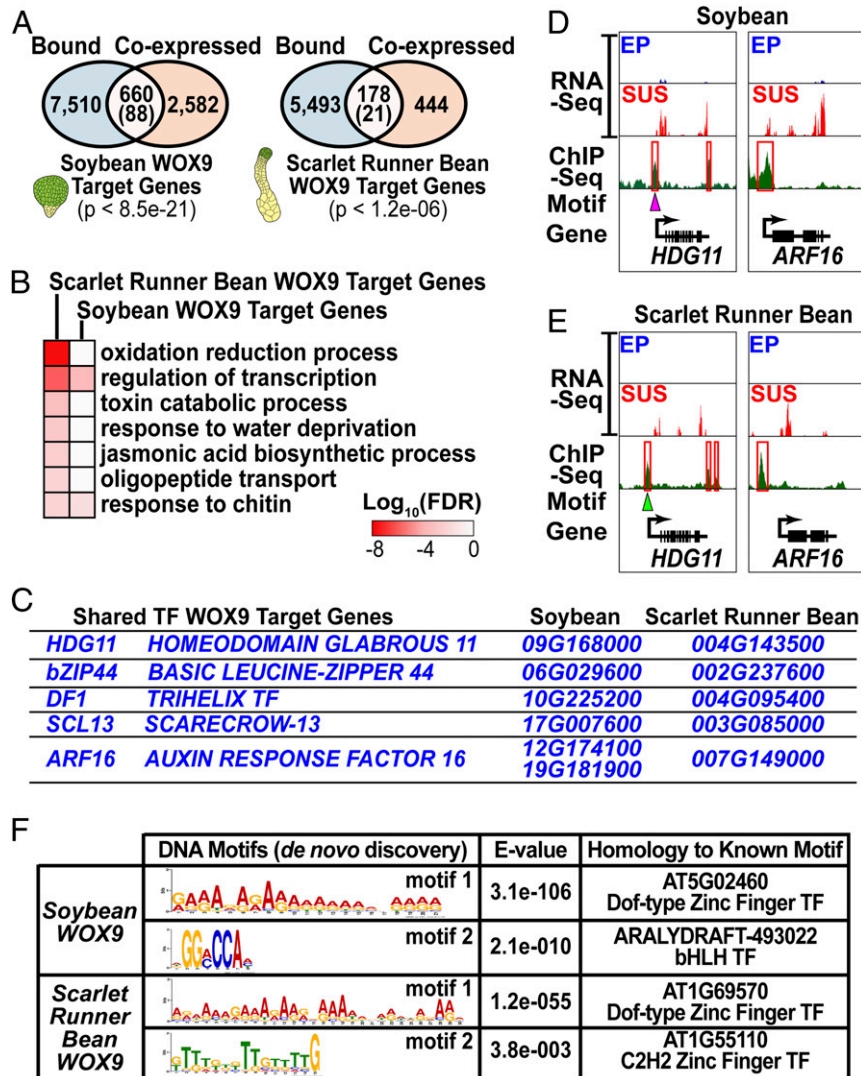
Fig. 7. WOX9 transcription factor targets in SRB and SB globular-stage seeds. (*A*) Venn diagrams between WOX9-bound genes (blue) and coexpressed genes (orange). Bound genes are those with peaks within 1 kb upstream of the transcription start site (47). Coexpressed genes are the more than fivefold up-regulated genes shown in Fig. 2 (SRB) and Fig. 5 (SB). WOX9 target genes are assumed to be those in the intersection between bound and coexpressed genes (pink) (47, 56). TF genes are in parentheses. *P* values were obtained using a hypergeometric distribution test (56). WOX9-bound regions and target genes are listed in Dataset S4. (*B*) The most significant GO terms for SRB and SB WOX9 target genes. All GO terms are listed in Dataset S4. (*C*) WOX9 TF target genes that are shared by SRB and SB (*SI Appendix, Materials and Methods*). (*D* and *E*) Genome browser view of shared SB (*D*) and SRB (*E*) WOX9 target gene activity (RNA-Seq) and bound gene regions (ChIP-Seq). Arrows point to motif 1 in both SB and SRB *HDG11* genes. *ARF16* did not have an enriched motif. (*F*) Enriched motifs in target genes bound by WOX9. The MEME-ChIP suite (67) was used for de novo motif discovery as previously described (47, 56). The E value is the probability of obtaining a specific motif compared with a randomly generated set of sequences (67). The Tomtom tool and plant TF databases within the MEME-ChIP suite were used to identify TFs with binding sites similar to the discovered motifs (68–70). The zinc finger TFs were present in the *Arabidopsis* DNA affinity purification sequencing (DAP-Seq) database (69), whereas the bHLH TF was identified from the *Arabidopsis* protein-binding microarray database (70). Target genes associated with each enriched motif are listed in Dataset S4.

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies

PNAS | 9 of 12
https://doi.org/10.1073/pnas.2024704118

stage of development, major regulatory decisions are being made, particularly within the embryo proper region (12), and embryos of these species differ in size and morphology, reflecting their final seed sizes and differences in suspensor morphology (Fig. 1).

Giant SRB and CB Suspensors Are Specialized to Express Specific Metabolic Pathways. A comparison of major GO terms derived from up-regulated suspensor mRNAs shows the dramatic functional differences between SRB and CB suspensors with those of SB and *Arabidopsis* (Figs. 3 and 5, *SI Appendix*, Fig. S12, and Dataset S2) and are summarized conceptually in Fig. 8. Significantly, a large number of metabolic pathways leading to the synthesis of several important hormones are overrepresented in SRB and CB suspensor-specific mRNAs. Up-regulated mRNAs are also enriched in plastid organelle processes (Fig. 8), as well as other plastid-associated GO terms such as starch biosynthesis, plastid stroma, and plastid membrane, because many of the major pathways are localized within specialized plastids (Dataset S3). Remarkably, SRB and CB suspensor-specific mRNAs are overrepresented in a continuum of biosynthetic pathways starting with glycolysis, pentose phosphate shunt, and the MEP isoprenoid pathway on one hand, and ending with major metabolic pathways leading to the production of GA, JA, and auxin, among others (Fig. 3, *SI Appendix*, Figs. S6 and S12, and Dataset S2). The machinery driving these pathways is active, because their end-product hormones are present in SRB suspensors (24, 26, 27, 39). SRB and CB up-regulated mRNAs are also overrepresented in the transport and signaling processes, making it possible for hormones (e.g., GA, JA, and auxin) synthesized in the suspensor to be either utilized within the suspensor region or exported to the embryo proper to facilitate developmental and physiological events (Figs. 3 and 8, *SI Appendix*, Fig. S12, and Dataset S2).

With the exception of auxin, suspensor-specific mRNAs are not overrepresented in GO terms for the majority of pathways leading to hormone production (e.g., MEP pathway, glycolysis, and pentose phosphate shunt) in simpler SB and *Arabidopsis* suspensors (Figs. 5 and 8 and Dataset S2). Nor are up-regulated mRNAs enriched for plastid-associated GO terms in these plants (Fig. 8). Where then are hormone synthesizing processes being carried out in SB and *Arabidopsis* seeds? We previously showed that *Arabidopsis* chalazal-endosperm–specific mRNAs are over-represented for GA, ABA, CK, and auxin GO terms at the globular stage, suggesting that these hormones are produced within this specialized endosperm subregion and not in the suspensor (33). We searched the Harada-Goldberg SB seed LCM datasets and found that these hormones are most likely produced in globular-stage endosperm and seed coat regions (http://seedgenenetwork.net) (47). For example, the only the seed region that has all of the mRNAs required for GA biosynthesis, including GA3ox, is the outer integument seed coat layer, and none of the rate limiting enzymes for JA, ABA, CK, and auxin are up-regulated within the suspensor. Thus, highly specialized SRB and CB suspensors may have co-opted regulatory events that occur within the endosperm and seed coat layers of plants that have simpler suspensors, such as SB and *Arabidopsis*—a hypothesis that was proposed over 100 y ago (5, 9, 27). Whether pathways leading to similar hormones in distinct seed parts utilize the same or different gene regulatory pathways remains to be determined.

The high degree of metabolic functional specialization within giant SRB and CB suspensors reported here is in remarkable agreement with elegant histological studies carried out by Yeung and Clutter with SRB suspensors over four decades ago (22). Their experiments revealed that the SRB suspensor has a large number of specialized starch-containing plastids, extensive networks of wall ingrowths, and substantial amounts of smooth membranes and dictyosomes consistent with 1) an embryo region that is synthesizing and transporting essential materials to the embryo proper and 2) the GO terms we uncovered that are generated by up-regulated suspensor mRNAs. Finally, are the giant suspensor regions of other plants specialized for hormone production as well? *Cytisus laburnum*, or Golden Chain, and *Tropaeolum majus*, the garden nasturtium, both have giant suspensors and have been shown to synthesize bioactive GAs (48). Thus, the development of giant specialized suspensors appears to be associated with specialized metabolic processes such as those leading to hormone production. How plants such as SRB and CB evolved morphologically unique suspensor regions as well as the specialized regulatory processes that drive and coordinate the expression of large numbers of specific metabolic pathway genes remains an important unanswered question.

The Embryo Proper Region Carries Out Similar Processes in Plants with Different Suspensor Morphologies. In contrast with the suspensor, the constellation of functions carried out by the embryo proper regions of SRB, CB, SB, and *Arabidopsis* are similar, irrespective of suspensor morphology (Figs. 3 and 5, *SI Appendix*, Fig. S12, and Dataset S2). Up-regulated embryo proper mRNAs are enriched for GO terms reflecting gene regulation, pattern formation, hormone responses, cell proliferation, and DNA replication, among many others (Fig. 8). These GO terms mirror the cell division and differentiation processes that occur within SRB, CB, SB, and *Arabidopsis* globular-stage embryo proper regions as they form cells, tissues, and subregions that will constitute the embryo when it matures. Thus, the evolutionary events that give rise to morphologically diverse suspensor regions that degenerate during seed development are uncoupled from those that maintain continuity of embryo proper form and function in order to guarantee plant survival from generation to generation.

TF mRNA Sets Have Been Identified that Are Shared By Embryo Proper and Suspensor Regions Irrespective of Embryo Morphology.
Suspensor-specific TF mRNAs. We uncovered a small number of mRNAs that are up-regulated in SRB, CB, SB, and *Arabidopsis* suspensor regions (Fig. 6A). The precise role that most of these TF mRNAs play in suspensor differentiation and function is not yet known. However, most likely they function within all suspensor cells as localization experiments carried out with suspensor-specific mRNAs [e.g., WOX9 (Fig. 6D) and GA enzymes], by us (28, 30)
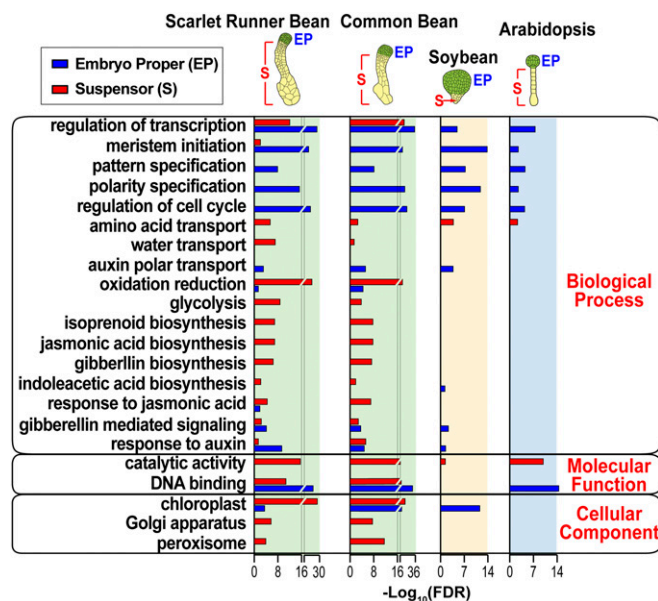

Fig. 8. Comparison of SRB, CB, SB, and *Arabidopsis* GO terms from up-regulated suspensor and embryo proper mRNAs. Summary of the most significant shared and specific GO terms. Data are taken from Dataset S2, Figs. 3 and 5, and *SI Appendix*, Fig. S12.

and by others (11, 49), showed that transcripts are distributed relatively evenly across the suspensor. The absence of cellular diversity suggests that genetic regulatory networks responsible for controlling suspensor form and function are probably simpler than those that operate in the embryo proper which undergoes a more complex set of developmental events required to establish the diverse cell types and tissues of this embryonic region.

WOX8/9, HDG11, and WRKY2 TF mRNAs have been shown to play essential roles in *Arabidopsis* suspensor differentiation (11, 50). WRKY2 mRNA is up-regulated more than fivefold in our *Arabidopsis* suspensor mRNA population (Dataset S1). Close relatives in SRB (Phvul.005G005800; Phvul.008G054100), CB (Phvul.005G005800; Phvul.008G054100), and SB (Glyma.09250500) suspensor mRNAs are up-regulated three- to sixfold depending upon the gene (Dataset S1), but collectively failed to meet our more than fivefold criterion for shared mRNAs, in contrast with WOX8/9 and HDG11 TF mRNAs (Fig. 6A and *SI Appendix*, Figs. S3C and S11G). Both WRKY2 and HDG11 TFs are essential for *WOX8/9* gene activation within the suspensor and are part of the SHORT SUSPENSOR (SSP)/YODA signaling cascade required for suspensor differentiation (11, 50). HDG11 mRNA is present in the egg cell, suggesting that maternal factors play a role in suspensor specification and are important for *WOX8/9* gene activation (11). We proposed over two decades ago that localized maternal factors in the egg cell might be distributed asymmetrically to the basal cell after zygote division promoting suspensor differentiation similar to maternal localization processes that occur in animal embryos such as the sea urchin (30). HDG11 TF mRNA localization within the egg cell and its role in suspensor development is consistent with this hypothesis. Nevertheless, our results suggest that WRKY2, HDG11, and WOX8/9 TF mRNAs perform similar roles in SRB, CB, and SB, and that regulatory events giving rise to the suspensor early in embryogenesis are conserved in plants, irrespective of suspensor morphology.

Embryo proper–specific TF mRNAs. We uncovered a larger number of globular-stage embryo proper–specific TF mRNAs that are shared between SRB, CB, SB, and *Arabidopsis* (Fig. 6B), most likely a result of the greater complexity of the embryo proper region. Many of these TF mRNAs have known roles in early embryo proper development as a result of the elegant genetic experiments carried out with *Arabidopsis* (12, 51). What emerges from these studies is that the globular embryo is divided into specific territories marking specification events leading to shoot and root meristems, vasculature regions, cotyledons, and other subregions and tissues of the mature embryo. Many of these territories are generated by auxin gradient signaling that sets off a cascade of events leading to the differentiation of specific embryo parts and subregions (12). In this respect, the embryo proper region resembles conceptually an early sea urchin embryo that is divided into specific territories that require different regulatory inputs to specify unique differentiation events leading to the mature embryo (52). The up-regulation of shared TF mRNAs in the embryo proper of SRB, CB, SB, and *Arabidopsis* suggests that common regulatory pathways operate within the embryo proper across the plant kingdom to define analogous specification domains. In contrast with the suspensor, it will be a significant challenge to unravel the mosaic of genetic regulatory networks that govern differentiation events within each embryo proper territory. The precise architecture of these networks and how they program the differentiation of a mature plant embryo remain to be determined.

Gaining Entry into WOX9 Suspensor Regulatory Networks. One of the interesting aspects of the WOX9 ChIP-Seq experiments in both SRB and SB is that there is a significant enrichment in TF target genes. This is consistent with the essential role that *WOX9* plays in suspensor differentiation (11) and suggests that *WOX9* is upstream in the hierarchy of suspensor regulatory networks. It is surprising that the *HDG11* TF gene appears to be a WOX9 target. This suggests that there is a feedback loop in the regulatory circuit containing *WOX9*, and that after HDG11 activation of *WOX9*,

WOX9 plays a role in reinforcing *HDG11* transcription. What is interesting is that both HDG11 and WOX8/9 mRNAs are localized within the *Arabidopsis* egg cell (11) and that the maternally derived *HDG11* allele is expressed early in suspensor development (53). This suggests that the *WOX9* regulatory network might be activated during egg cell development and sets off a cascade of events leading to suspensor differentiation following fertilization.

What is also significant is the divergence of WOX9 target gene functions in SRB and SB (Fig. 7). These functional differences correlate with morphological differences between SRB and SB suspensor regions (Fig. 1). For example, GA biosynthesis is a key marker of giant specialized suspensors (Fig. 8), but GA genes are not direct targets of WOX9 (Fig. 7B and Dataset S4). How are these genes activated in SRB and CB suspensors? Previously, we uncovered a *cis*-regulatory module (CRM) that is required for *GA20ox* and *G564* transcription within the suspensor (28, 31). Both of these genes are up-regulated in SRB and CB suspensor regions (Fig. 2B and *SI Appendix*, Fig. S11B). This module contains three distinct *cis*-elements, designated as the 10-bp motif, region 2 motif, and the fifth motif, which have sequences that recognize Dof-type zinc finger, myb, and C2H2 zinc finger TFs, respectively (28). We searched SRB WOX9 targets for mRNAs encoding these TFs and found candidates for all three that are up-regulated in the SRB suspensor (Datasets S1 and S4). Thus, one model is that activation of *WOX9* leads to the activation of TFs which switch on GA biosynthesis and other specialized suspensor-specific genes that are downstream of *WOX9* in the regulatory circuitry (*SI Appendix*, Fig. S15).

It would appear, therefore, that WOX9 plays a dual role—one that is responsible for basic developmental events required for suspensor differentiation across the plant kingdom and another that has been adapted for highly specialized functions in giant suspensors such as those in SRB and CB. The regulatory circuitry controlling these specialized suspensor functions has evolved over the 19 my since the separation of SRB and SB from their common ancestor (Fig. 1J). Clearly, the major tasks before us are to 1) unravel the basic regulatory circuitry required for the differentiation of all suspensors irrespective of morphology and 2) determine how regulatory genes in that network (e.g., *WOX9*) are utilized differently in circuits that control specialized downstream suspensor functions in plants with morphologically distinct suspensor regions. What these circuits are and how they are organized within plant genomes remain to be discovered.

Materials and Methods

All of the methods used for the experiments reported in this paper were published recently by our laboratories (20, 33, 47, 54–56). Briefly, methods related to growth of SRB and CB plants, and procedures used for LCM of embryo proper and suspensor regions are detailed in Chen et al. (20). Specific RNA-Seq methods, including RNA isolation, sequencing library construction, and bioinformatic analyses, are presented in Belmonte et al. (33) and Pelletier et al. (47). DNA sequencing methodology and analysis of DNA sequences are outlined in detail by Lin et al. (55) and Chen et al. (54). Specific ChIP-Seq procedures, including peptide antibody synthesis and bioinformatic analyses, are described in detail by Pelletier et al. (47) and Jo et al. (56). Information relevant to each experiment reported here is contained within figure legends and complete details are contained within *SI Appendix, Materials and Methods*.

Data Availability. The transcriptome and ChIP-Seq data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database (scarlet runner bean and common bean RNA-Seq [accession no. GSE57537]; soybean RNA-Seq [accession no. GSE57349]; *Arabidopsis* RNA-Seq [accession no. GSE135393]; scarlet runner bean ChIP-Seq [accession no. GSE153644]; and soybean ChIP-Seq [accession no. GSE135267]). The genome sequences were deposited into the GenBank database (scarlet runner bean [contigs QBDZ01000001 to QBDZ01192921]). All study data are included in the article and/or supporting information.

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos with different morphologies

PNAS | 11 of 12
https://doi.org/10.1073/pnas.2024704118

1. R. B. Goldberg, G. de Paiva, R. Yadegari, Plant embryogenesis: Zygote to seed. *Science* **266**, 605–614 (1994).

2. T. Radoeva, P. Vaddepalli, Z. Zhang, D. Weijers, Evolution, initiation, and diversity in early plant embryogenesis. *Dev. Cell* **50**, 533–543 (2019).

3. K. Wang, H. Chen, Y. Miao, M. Bayer, Square one: Zygote polarity and early embryogenesis in flowering plants. *Curr. Opin. Plant Biol.* **53**, 128–133 (2020).

4. D. Jacob, J. Brian, The short and intricate life of the suspensor. *Physiol. Plant.* **169**, 110–121 (2020).

5. K. F. Henry, R. B. Goldberg, Using giant scarlet runner bean embryos to uncover regulatory networks controlling suspensor gene activity. *Front. Plant Sci.* **6**, 44 (2015).

6. T. Kawashima, R. B. Goldberg, The suspensor: Not just suspending the embryo. *Trends Plant Sci.* **15**, 23–30 (2010).

7. E. C. Yeung, D. W. Meinke, Embryogenesis in angiosperms: Development of the suspensor. *Plant Cell* **5**, 1371–1381 (1993).

8. L. Guignard, Recherches d 'embryogenie vegetale comparee. Premier memoire: Legumineuses. *Ann. Sci. Nat. Bot. VI.* **12**, 5–166 (1881).

9. M. A. Tison, Sur le suspenseur du *Trapa natans* L. Revue générale de botanique. **31**, 219–228 (1919).

10. K. Wabnik, H. S. Robert, R. S. Smith, J. Friml, Modeling framework for the establishment of the apical-basal embryonic axis in plants. *Curr. Biol.* **23**, 2513–2518 (2013).

11. M. Ueda *et al.*, Transcriptional integration of paternal and maternal factors in the *Arabidopsis* zygote. *Genes Dev.* **31**, 617–627 (2017).

12. S. Lau, D. Slane, O. Herud, J. Kong, G. Jürgens, Early embryogenesis in flowering plants: Setting up the basic body pattern. *Annu. Rev. Plant Biol.* **63**, 483–506 (2012).

13. N. R. Lersten, Suspensors in leguminoseae. *Bot. Rev.* **49**, 234–257 (1983).

14. W. Nagl, The Phaseolus suspensor and its polytene chromosomes. *Z. Pflanzenphysiol.* **73**, 1–44 (1974).

15. T. Brady, M. Clutter, Structure and replication of *Phaseolus* polytene chromosomes. *Chromosoma* **45**, 63–79 (1974).

16. A. Delgado-Salinas, R. Bibler, M. Lavin, Phylogeny of the genus *Phaseolus* (leguminoseae): A recent diversificatio in an ancient landscape. *Syst. Bot.* **31**, 779–791 (2006).

17. A. R. Schwember, B. Carrasco, P. Gepts, Unraveling agronomic and genetic aspects of runner bean (Phaseolus coccineus L.). *Field Crops Res.* **206**, 86–94 (2017).

18. A. Guerra-García, M. Suárez-Atilano, A. Mastretta-Yanes, A. Delgado-Salinas, D. Piñero, Domestication genomics of the open-Pollinated scarlet runner bean (*Phaseolus coccineus* L.). *Front. Plant Sci.* **8**, 1891 (2017).

19. P. E. McClean, M. Lavin, P. Gepts, S. A. Jackson, "Phaseolus vulgaris: A diploid model for soybean" in *Genetics and Genomics of Soybean. Plant Genetics and Genomics: Crops and Models*, G. Stacey, Ed. (Springer, New York, NY, 2008), 2, pp. 1–22.

20. M. Chen, A. Q. Bui, R. B. Goldberg, Using giant scarlet runner bean (Phaseolus coccineus) embryos to dissect the early events in plant embryogenesis. *Methods Mol. Biol.* **2122**, 205–222 (2020).

21. E. C. Yeung, M. E. Clutter, Embryology of *Phaseolus coccineus:* The ultrastructure and development of the suspensor. *Can. J. Bot.* **57**, 120–136 (1979).

22. E. C. Yeung, M. E. Clutter, Embryogeny of *Phaseolus coccineus*: Growth and microanatomy. *Protoplasma* **94**, 19–40 (1978).

23. I. Sussex, M. Clutter, V. Walbot, T. Brady, Biosynthetic activity of the suspensor of phaseolus coccineus. *Caryologia* **25**, 261–272 (1973).

24. P. Picciarelli, N. Ceccarelli, F. Paolicchi, G. Calistri, Exogenous auxins and embryogeneisis in *Phaseolus coccineus*. *Aust. J. Plant Physiol.* **28**, 73–78 (2001).

25. A. Piaggesi, N. Ceccarelli, P. Picciarelli, A. Alpi, Biosynthesis of kaurene in Phaseolus coccineus embryos at a very early developmental stage. *J. Plant Physiol.* **137**, 120–122 (1990).

26. P. Perata, P. Picciarelli, A. Alpi, Pattern variations in abscisic acid cntent in suspensors, embryos, and integuments of developing *Phaseolus coccineus* seeds. *Plant Physiol.* **94**, 1776–1780 (1990).

27. R. Lorenzi, A. Bennici, P. G. Cionini, A. Alpi, F. D'Amato, Embryo-suspensor relations in *Phaseolus coccineus*: Cytokinins during seed development. *Planta* **143**, 59–62 (1978).

28. K. F. Henry, A. Q. Bui, T. Kawashima, R. B. Goldberg, A shared *cis*-regulatory module activates transcription in the suspensor of plant embryos. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5824–E5833 (2018).

29. B. H. Le *et al.*, Using genomics to study legume seed development. *Plant Physiol.* **144**, 562–574 (2007).

30. K. Weterings *et al.*, Regional localization of suspensor mRNAs during early embryo development. *Plant Cell* **13**, 2409–2425 (2001).

31. T. Kawashima *et al.*, Identification of cis-regulatory sequences that activate transcription in the suspensor of plant embryos. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3627–3632 (2009).

32. J. Schmutz *et al.*, A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* **46**, 707–713 (2014).

33. M. F. Belmonte *et al.*, Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E435–E444 (2013).

34. I. H. Kang, J. G. Steffen, M. F. Portereiko, A. Lloyd, G. N. Drews, The AGL62 MADS domain protein regulates cellularization during endosperm development in Arabidopsis. *Plant Cell* **20**, 635–647 (2008).

35. A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).

36. Q. Tu, R. A. Cameron, E. H. Davidson, Quantitative developmental transcriptomes of the sea urchin Strongylocentrotus purpuratus. *Dev. Biol.* **385**, 160–167 (2014).

37. R. B. Goldberg, G. Hoschek, S. H. Tam, G. S. Ditta, R. W. Breidenbach, Abundance, diversity, and regulation of mRNA sequence sets in soybean embryogenesis. *Dev. Biol.* **83**, 201–217 (1981).

38. C. A. ten Hove, K. J. Lu, D. Weijers, Building a plant: Cell fate specification in the early Arabidopsis embryo. *Development* **142**, 420–430 (2015).

39. N. Ceccarelli, R. Lorenzi, A. Alpi, Gibberellin biosynthesis in Phaseolus coccineus suspensor. *Z. Pflanzenphysiol.* **102**, 37–44 (1981).

40. P. G. Cionini, A. Bennici, A. Alpi, F. D' Amato, Suspensor, gibberellin and in vitro development of Phaseolus coccineus embryos. *Planta* 131, 115–117 (1976).

41. T. Brady, E. D. Walthall, The effect of the suspensor and gibberellic acid on Phaseolus vulgaris embryo protein content. *Dev. Biol.* **107**, 531–536 (1985).

42. V. Walbot, M. Clutter, I. Sussex, Reprodcutive development and embryogeny in *Phaseolus*. *Phytomorphology* **22**, 59–68 (1972).

43. T. Brady, "Cytological and biochelical studies on the suspensor and suspesnor polytene chromosomes of *Phaseolus*" in *Biology* (Yale University, New Haven, CT, 1972).

44. S. G. Mansfield, L. G. Briarty, Early embryogenesis in *Arabidopsis thaliana.* II. The developing embryo. *Can. J. Bot.* **69**, 461–476 (1991).

45. M. A. Chamberlin, H. A. Horner, R. G. Palmer, Nutrition of ovule, embryo sac, and young embryo in soybean: An anatomical and atoradiographic study. *Can. J. Bot.* **71**, 1153–1168 (1993).

46. J. Palovaara *et al.*, Transcriptome dynamics revealed by a gene expression atlas of the early Arabidopsis embryo. *Nat. Plants* **3**, 894–904 (2017).

47. J. M. Pelletier *et al.*, LEC1 sequentially regulates the transcription of genes involved in diverse developmental processes during seed development. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6710–E6719 (2017).

48. P. Picciarelli, A. Alpi, L. Pistelli, M. Scalet, Gibberellin-like activity in suspensors of Tropaeolum majus L. and Cytisus laburnum L. *Planta* **162**, 566–568 (1984).

49. A. Haecker *et al.*, Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in Arabidopsis thaliana. *Development* **131**, 657–668 (2004).

50. M. Ueda, Z. Zhang, T. Laux, Transcriptional activation of Arabidopsis axis patterning genes WOX8/9 links zygote polarity to embryo development. *Dev. Cell* **20**, 264–270 (2011).

51. A. Capron, S. Chatfield, N. Provart, T. Berleth, "Embryogenesis: Pattern formation from a sincle cell" in *The Arabidopsis Book* (American Society of Plant Biologists, 2009), vol. 2009, pp. 7.

52. E. H. Davidson, R. A. Cameron, A. Ransick, Specification of cell fate in the sea urchin embryo: Summary and some proposed mechanisms. *Development* **125**, 3269–3290 (1998).

53. P. Zhao, X. Zhou, Y. Zheng, Y. Ren, M. X. Sun, Equal parental contribution to the transcriptome is not equal control of embryogenesis. *Nat. Plants* **6**, 1354–1364 (2020).

54. M. Chen *et al.*, Seed genome hypomethylated regions are enriched in transcription factor genes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8315–E8322 (2018).

55. J. Y. Lin *et al.*, Similarity between soybean and *Arabidopsis* seed methylomes and loss of non-CG methylation does not affect seed development. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9730–E9739 (2017).

56. L. Jo *et al.*, Combinatorial interactions of the LEC1 transcription factor specify diverse developmental programs during soybean seed development. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1223–1232 (2020).

57. M. Lavin, P. S. Herendeen, M. F. Wojciechowski, Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst. Biol.* **54**, 575–594 (2005).

58. J. E. Bowers, B. A. Chapman, J. Rong, A. H. Paterson, Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433–438 (2003).

59. B. B. Buchanan, W. Gruissem, R. L. Jones, *Biochemistry & Molecular Biology of Plants* (John Wiley & Sons, United Kingdom, ed. 2, 2015), pp. 1264.

60. P. Hedden, S. G. Thomas, Gibberellin biosynthesis and its regulation. *Biochem. J.* **444**, 11–25 (2012).

61. H. Magome *et al.*, CYP714B1 and CYP714B2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1947–1952 (2013).

62. J. Binenbaum, R. Weinstain, E. Shani, Gibberellin localization and transport in plants. *Trends Plant Sci.* **23**, 410–421 (2018).

63. J. J. Almagro Armenteros, C. K. Sønderby, S. K. Sønderby, H. Nielsen, O. Winther, DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics* **33**, 3387–3395 (2017).

64. A. Alpi, F. Tognoni, F. D'Amato, Growth regulator levels in embryo and suspensor of *Phaseolus coccineus* at two stages of development. *Planta* **127**, 153–162 (1975).

65. T. P. Sun, The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants. *Curr. Biol.* **21**, R338–R345 (2011).

66. K. Tamura *et al.*, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).

67. P. Machanick, T. L. Bailey, MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* **27**, 1696–1697 (2011).

68. S. Gupta, J. A. Stamatoyannopoulos, T. L. Bailey, W. S. Noble, Quantifying similarity between motifs. *Genome Biol.* **8**, R24 (2007).

69. R. C. O'Malley *et al.*, Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**, 1280–1292 (2016).

70. J. M. Franco-Zorrilla *et al.*, DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2367–2372 (2014).

12 of 12 | PNAS
https://doi.org/10.1073/pnas.2024704118

Chen et al.
Comparative analysis of embryo proper and suspensor transcriptomes in plant embryos
with different morphologies