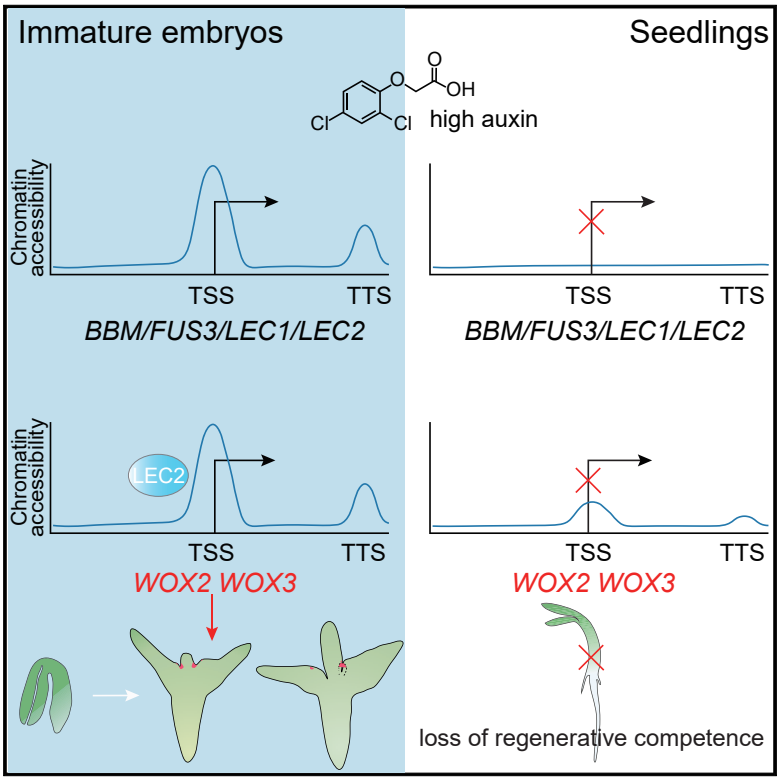
Article

Developmental Cell

Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis

Graphical Abstract



Authors

Fu-Xiang Wang, Guan-Dong Shang, Lian-Yu Wu, Zhou-Geng Xu, Xin-Yan Zhao, Jia-Wei Wang

Correspondence

jwwang@sippe.ac.cn

In Brief

Somatic embryogenesis provides a powerful system to produce genetically modified crops and to obtain artificial seeds. In combination with reverse genetics, ATAC-seq, ChIP-seq, and RNA-seq, Wang et al. reveal a hierarchical transcriptional regulatory network for somatic embryogenesis.

Highlights

- Description of the chromatin accessibility landscape for somatic embryogenesis in plants
- Auxin rapidly rewires the cell totipotency network by altering chromatin accessibility
- The embryonic nature of explants is a prerequisite for somatic cell reprogramming
- A molecular link between cell totipotent genes and early embryonic development pathway

Wang et al., 2020, Developmental Cell *54*, 1–16
September 28, 2020
https://doi.org/10.1016/j.devcel.2020.07.003

CellPress

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

Developmental Cell

CellPress

Article

Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis

Fu-Xiang Wang,[1,2,5] Guan-Dong Shang,[1,2,5] Lian-Yu Wu,[1,3] Zhou-Geng Xu,[1,2] Xin-Yan Zhao,[1,2,4] and Jia-Wei Wang[1,3,6,*]

[1]National Key Laboratory of Plant Molecular Genetics (NKLPMG), CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology (SIPPE), Chinese Academy of Sciences (CAS), Shanghai 200032, China

[2]University of Chinese Academy of Sciences (UCAS), Shanghai 200032, China

[3]ShanghaiTech University, Shanghai 200031, China

[4]Shanghai Center for Plant Stress Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences (CAS), Shanghai 200032, China

[5]These authors contributed equally

[6]Lead Contact

*Correspondence: jwwang@sippe.ac.cn

https://doi.org/10.1016/j.devcel.2020.07.003

SUMMARY

Plant somatic embryogenesis refers to a phenomenon where embryos develop from somatic cells in the absence of fertilization. Previous studies have revealed that the phytohormone auxin plays a crucial role in somatic embryogenesis by inducing a cell totipotent state, although its underlying mechanism is poorly understood. Here, we show that auxin rapidly rewires the cell totipotency network by altering chromatin accessibility. The analysis of chromatin accessibility dynamics further reveals a hierarchical gene regulatory network underlying somatic embryogenesis. Particularly, we find that the embryonic nature of explants is a prerequisite for somatic cell reprogramming. Upon cell reprogramming, the B3-type totipotent transcription factor LEC2 promotes somatic embryo formation by direct activation of the early embryonic patterning genes *WOX2* and *WOX3*. Our results thus shed light on the molecular mechanism by which auxin promotes the acquisition of plant cell totipotency and establish a direct link between cell totipotent genes and the embryonic development pathway.

INTRODUCTION

Owing to their sessile nature, plants maintain cell pluripotency or totipotency throughout their life cycles. Somatic cells are able to regenerate themselves in response to chemical or mechanical stimuli or go through somatic embryogenesis (SE) to regenerate whole plants (Birnbaum and Sánchez Alvarado, 2008; Ikeuchi et al., 2019, 2016; Sena and Birnbaum, 2010; Sugimoto et al., 2011). SE refers to the development of ectopic embryos from somatic cells independent of gamete formation and fertilization. Since its first documentation in 1950s (Reinert, 1958; Steward et al., 1958; Waris, 1957), SE has become a powerful tool in plant biotechnology for the propagation of endangered species and generation of genetically modified plants with improved traits (Lowe et al., 2016). Thus, the elucidation of the molecular and cellular basis of SE is of great importance to understand the basic principles underlying embryonic patterning and epigenetic reprogramming in plants (Birnbaum and Roudier, 2017; Méndez-Hernández et al., 2019; Palovaara et al., 2016; Radoeva et al., 2019b; Radoeva and Weijers, 2014; Smertenko and Bozhkov, 2014; Winkelmann, 2016; Wójcik et al., 2020; Xu and Huang, 2014).

Somatic embryos can be induced by exposing immature zygotic embryos or male gametophytes to the synthetic auxin 2,4-dichlorophenoxyacetic acid (2,4-D) or to abiotic stress (Custers et al., 1994; Fehér, 2015; Gaj, 2011). It has been shown that local auxin biosynthesis and polar auxin transport are essential for the establishment of auxin gradients during somatic embryo formation (Bai et al., 2013; Soriano et al., 2014; Su and Zhang, 2009). How auxin induces cell totipotency and how auxin promotes embryogenesis remains unclear.

Overexpression of certain key transcription factors (TFs) has been used to induce the differentiation of stem cells or SE in both animals and plants. In 2006, Takahashi and Yamanaka showed that a combination of four specific TFs were involved in the conversion of differentiated fibroblasts to a pluripotent state resembling embryonic stem cells derived from the blastocyst inner cell mass (Takahashi and Yamanaka, 2006). Similarly, SE in Arabidopsis can be achieved by ectopic overexpression of a single TF, including the homeodomain TF *WUSCHEL* (*WUS*), AP2-domain TF *PLETHORA4/BABY BOOM* (*PLT4/BBM*) or *PLT5/EMBRYO MAKER* (*PLT5/EMK*), MADS-box TF *AGAMOUS-LIKE15* (*AGL15*), NF-Y (nuclear factor of the Y box) TF *LEAFY COTYLEDON1* (*LEC1*), B3 TF *LEC2* and *FUSCA3* (*FUS3*), MYB TF *MYB118* and RWP-RK DOMAIN-CONTAINING4 (RKD4)/GROUNDED (GRD) (Boutilier et al., 2002; Gallois et al., 2004; Harding et al., 2003; Lotan et al., 1998; Stone et al., 2001; Thakare

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

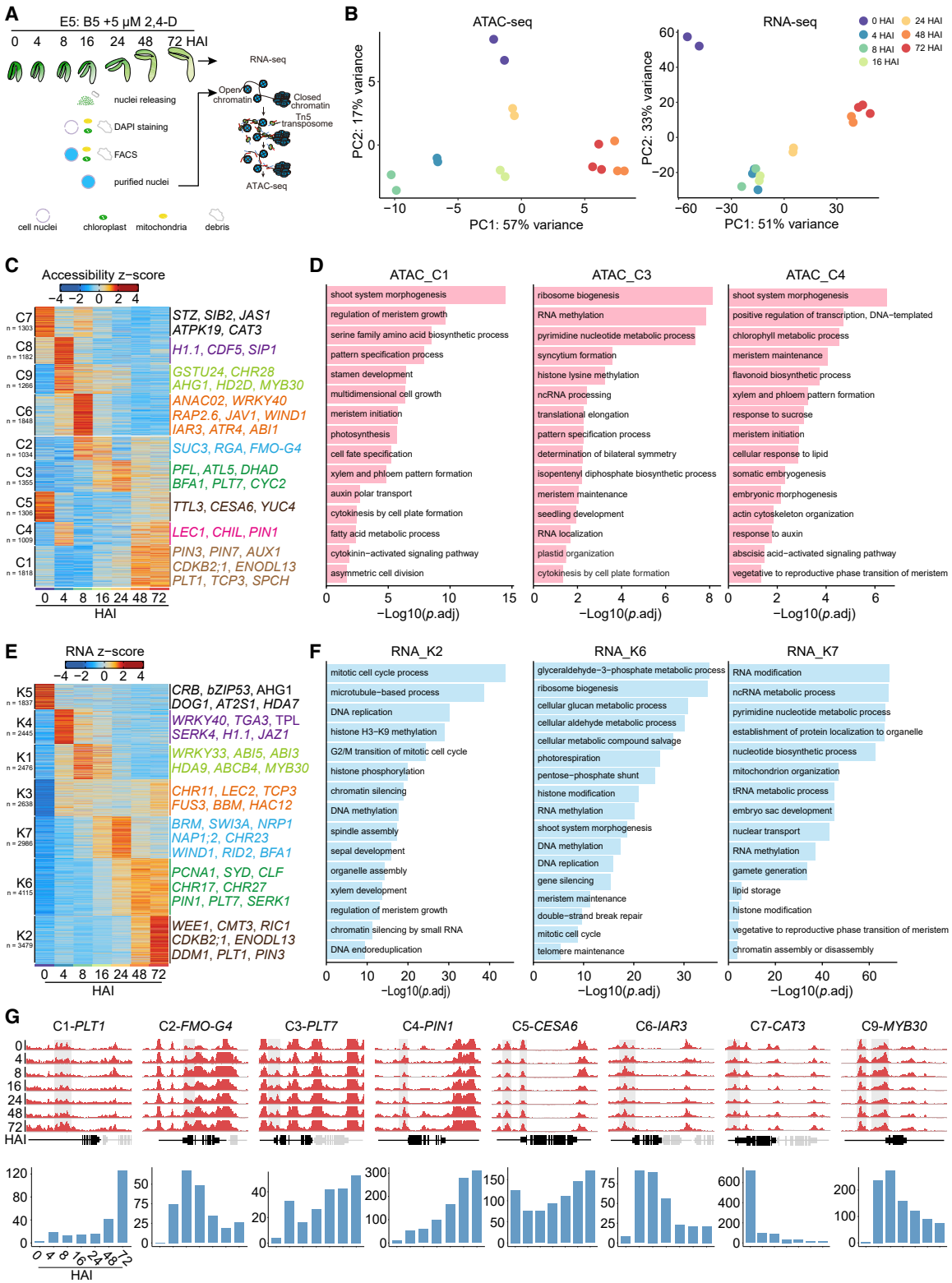


Figure 1. Overview of Chromatin Accessibility and Transcriptome Dynamics

(A) Schematic outline of genome-wide ATAC-seq and RNA-seq assays and time points of sample collection. HAI, hours after inoculation; DAPI, 4′,6-diamidino-2-phenylindole; FACS, fluorescence activated cell sorting.

(B) Principal-components plots of ATAC-seq and RNA-seq data, respectively. Color code is shown. Each dot represents one sample.

(legend continued on next page)

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

et al., 2008; Tsuwamoto et al., 2010; Waki et al., 2011; Wang et al., 2009b; Zuo et al., 2002). In addition, a large number of other TFs that are differentially expressed during SE have been identified through comprehensive time-course analyses (Gliwicka et al., 2013; Szczygieł-Sommer and Gaj, 2019; Wickramasuriya and Dunwell, 2015). The downstream events of these TFs were also analyzed. Transcriptional crosstalk among these different TFs was revealed by genome-wide target identification, and notably, BBM, LEC2, AGL15, and PLT5 were found to regulate common pathways, in particular that of auxin (Braybrook et al., 2006; Horstman et al., 2017b; Pinon et al., 2013; Stone et al., 2008; Zheng et al., 2009). However, how the overexpression of individual TFs from disparate families is able to induce SE is poorly understood, and whether these TFs promote embryogenesis through a common developmental pathway is not known.

As in animals, epigenetic reprogramming also plays an important role in the acquisition of totipotency and SE (Wójcikowska et al., 2020). POLYCOMB REPRESSIVE COMPLEX2 (PRC2), a chromatin regulator that maintains gene repression through the deposition of the histone H3 lysine 27 trimethylation (H3K27me3) marker, constitutes a major barrier to the hormone-mediated establishment of embryogenic competence in mature somatic cells in Arabidopsis (Ikeuchi et al., 2015; Mozgová et al., 2017). Accordingly, the mutation of PRC2 subunits leads to the formation of callus on the shoot apex or in some cases to disorganized cell masses and somatic embryos that develop from single root hairs (Chanvivattana et al., 2004; Ikeuchi et al., 2015). It has been proposed that chemical perturbation or genetic disruption of PRC2 may induce SE through de-repression of the TF genes, such as *AGL15*, *BBM*, *LEC1*, *LEC2*, and *PLT5* (Bouyer et al., 2011; Ikeuchi et al., 2015; Liu et al., 2016; Mozgová et al., 2017). Furthermore, treatment of Arabidopsis explants with trichostatin A (TSA), a chemical inhibitor of histone deacetylases, induces SE without the exogenous application of auxin (Wójcikowska et al., 2018).

To sum up, much progress has been made in the elucidation of the underlying mechanisms of SE. However, our understanding of how auxin, TFs, and epigenetic regulation collaboratively regulate somatic cell fate transition is still limited. Here, we report the chromatin accessibility landscape at the early stage of SE. We find that auxin rapidly induces the acquisition of cell totipotency by altering chromatin accessibility. The analysis of chromatin accessibility and transcriptome dynamics further reveal a hierarchical TF cascade underlying auxin-induced SE. In particular, our results uncover a long sought-after molecular link between cell totipotency genes and the early embryonic development pathway.

RESULTS

Overview of Chromatin Accessibility and Transcriptome Dynamics during SE

SE can be induced from explants of different origins and proceeds either directly or indirectly via an embryonic callus phase (Horstman et al., 2017a). In this study, somatic embryo formation was induced by culturing immature Arabidopsis embryos at the late-bent-cotyledon stage of development on the E5 medium (Gamborg B5 medium supplemented with 2,4-D). The embryo-like protuberances were further subcultured on auxin-free Murashige and Skoog (MS) medium, where the plantlets were formed (Figures 1A and S1) (Gaj, 2011).

Genome architecture and chromatin accessibility determine the functional state of a cell. Studying the changes in dynamic chromatin accessibility that occur during SE may provide a detailed understanding on how auxin coordinates overall genome architecture from a somatic cell to a totipotent state. To probe the reprogramming process at the chromatin level, we performed assays for transposase-accessible chromatin sequencing (ATAC-seq) on the explants at 0, 4, 8, 16, 24, 48, and 72 h after induction (HAI) on E5 media (Figure 1A) (Bajic et al., 2018; Buenrostro et al., 2013; Lu et al., 2017, 2019; Maher et al., 2018; Ricci et al., 2019). In parallel, we conducted RNA sequencing (RNA-seq) to understand the impact of chromatin accessibility on the transcriptome.

For ATAC-seq experiments, we obtained an average of 97.1% mappability and 15.7 million qualified reads per sample (Table S1). ATAC-seq data from two or three replicate samples showed high correlation (Figure 1B), indicating that ATAC-seq can reliably and reproducibly measure chromatin accessibility in these samples. In total, 25,530 high-confidence open chromatin peaks (or regions) were identified. Chromatin-accessible regions are widely distributed throughout the genome, with highest enrichment at transcriptional start sites (TSS) and mild enrichment at transcriptional termination sites (TTS) (Figures S2A and S2B). We assigned each ATAC-seq peak to the nearest gene based on its annotated TSS. It should be noted that, because enhancers often engage in physical contact with their cognate genes through long-range chromosomal interactions

(C) Heatmap of differentially accessible peaks sorted by k-means clustering across the samples collected at different time points (0 to 72 HAI). Color bar, accessibility *Z* score of differentially accessible peaks identified by ATAC-seq. The representative genes are shown on the right. The number of the peaks for each cluster is given.

(D) GO term analyses of three ATAC-seq clusters (C1, C3, and C4). The 15 selected enriched GO biological processes of differentially accessible genes are indicated below each cluster. The $-\log_{10}$ (*p*.adj) is given.

(E) Heatmap of differentially expressed genes sorted by k-means clustering across the samples collected at different time points (0 to 72 HAI). Color bar, RNA *Z* score of the differentially expressed genes identified by RNA-seq. The representative genes are shown on the right. The number of the genes for each cluster is given.

(F) GO term analyses of three RNA-seq clusters (K2, K6, and K7). The selected 15 enriched GO biological processes of differentially expressed genes are indicated below each cluster. The $-\log_{10}$(*p*.adj) is given.

(G) The ATAC-seq tracks (upper panel) and RNA-seq data (lower panel) for representative genes of each ATAC-seq cluster. The genomic regions are shown, and the selected genes are highlighted in black. The location of differential peaks is shadowed. Expression *Z* score (RNA_counts per million, RNA_CPM) for each gene at different time point is shown. The accession numbers for selected genes are *PLT1* (At3g20840), *FMO-G4* (*FLAVIN-MONOOXYGENASE GLUCOSINOLATE S-OXYGENASE4*, At1g62570), *PLT7* (At5g65510), *PIN1* (At1g73590), *CESA6* (*CELLULOSE SYNTHASE6*, At5g64740), *IAR3* (*IAA-ALANINE RESISTANT3*, At1g51760), *CAT3* (At1g20620), and *MYB30* (At3g28910).

See also Figures S1 and S2.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

(Schoenfelder and Fraser, 2019), this simple peak-gene association rule based on the gene nearest to the peak may not be accurate in all cases. On the basis of peak location, we further divided all the peaks into four sets: peaks within 2.0 kilobase pair (kb) upstream of transcription start sites (TSS-proximal), peaks within 1.0 kb downstream of TTS (TTS-proximal), peaks within coding regions, and other peaks (distal intergenic). Our results show that over 87% of the intergenic peaks belong to TSS-proximal (Figure S2B), so we used this dataset for subsequent analyses except the TF motif analysis by chromVAR (see below).

Results from principal component analysis (PCA) analyses showed the similar ATAC-seq peak profiles between 48 and 72 HAI (Figures 1B and S2C), suggesting that auxin-induced reprogramming of chromatin accessibility is largely complete by 72 HAI. K-means clustering approach based on the level of chromatin accessibility was used to sort all differential peaks resulting in nine clusters, named C1 to C9, respectively (Figure 1C; Table S1). The sites in clusters C5 and C7 were open at 0 HAI and lost accessibility after 4 HAI. Clusters C6, C8, and C9 exhibited similar accessibility dynamics. Their regulatory elements became transiently accessible at 4 or 8 HAI but gradually closed over the course of treatment. Interestingly, the wounding response genes, including *WOUND INDUCED DEDIFFERENTIATION1* (*WIND1*) and *RELATED TO AP2.6* (*RAP2.6*) and *JASMONATE-ASSOCIATED VQ MOTIF GENE1* (*JAV1*) were identified in cluster C6 (Iwase et al., 2011; Rymen et al., 2019; Yan et al., 2018). The sites in clusters C1 and C3 were closed from 0 to 8 HAI but progressively opened after 24 or 48 HAI. The genes involved in plant regeneration, auxin transport and cell cycle, such as *PLT1*, *CYCLIN-DEPENDENT KINASE B2;1* (*CDKB2;1*), *PIN-FORMED3* (*PIN3*), and *PIN7* were identified in these two clusters (Adamowski and Friml, 2015; Kareem et al., 2015; Menges et al., 2005). Gene ontology (GO) enrichment analyses revealed that the genes involved in wounding, cell death, and jasmonic acid response were enriched in clusters C2, C7, and C9 (Figure S2D; Table S1). In contrast, the genes belonging to clusters C1, C3, and C4 are known to play important roles in meristem maintenance and cell growth (Figure 1D). Thus, the above results reveal that the genome accessibility landscape of somatic cells undergoes rapid and massive transitions within 72 h.

Next, we analyzed RNA-seq data of samples collected at the same time points. PCA on our paired datasets showed high concordance between replicates (Figure 1B). Clustering analysis revealed 7 clusters, named K1 to K7, respectively (Table S2). The expression levels of the genes in cluster K5 were significantly decreased after 4 HAI (Figure 1E). GO enrichment analysis revealed that these genes were mainly involved in seed maturation and dormancy processes (Figure S2E; Table S2), indicating discontinued embryonic development upon culturing on E5. The genes in cluster K4 were transiently induced at 4 HAI and showed gene signatures for wounding or JA response (Figures 1E and S2D; Table S2). The genes in cluster K3 exhibited increased expression from 4 HAI. Among them were the known genes with critical roles in SE, such as *BBM*, *LEC2*, and *FUS3*. Clusters K2 and K6, which were highly activated at 24 or 48 HAI, were preferentially associated with the genes involved in cell cycle and ribosome biogenesis, implying that explants undergo extensive cell division at these stages.

Chromatin accessibility is necessary, but not sufficient, for enhancer or promoter activity, with gene expression also being determined by the level of DNA methylation and the type of histone modification of the promoter (Klemm et al., 2019). Therefore, the chromatin accessibility for a given gene is not always correlated with its expression level. Nevertheless, we found that peak accessibilities were positively correlated with gene expression in general (Figures 1G and S2F). For instance, the expression of *PLT1* in cluster C1 progressively increased with increasing proximal chromatin accessibility. Similarly, the expression of *CATALASE3* (*CAT3*) in cluster C7 was greatly reduced after 4 HAI, accompanied with the loss of accessibility in the proximal regulatory elements. The correlation between accessibility and gene expression level decreased with the increasing distance between the accessible peak and the TSS (Figure S2F).

Chromatin Accessibility and Transcriptome Dynamics Suggest a TF Network with Hierarchical Organization

Results from the time-course ATAC-seq has the potential to reveal the sequential action of TF binding based on characteristic chromatin footprints. Thus, we inferred chromatin-based, hierarchical gene regulatory networks for SE by integrative analysis of chromatin accessibility and transcriptome dynamics. To accurately infer TF variability and dynamics, we used HOMER and chromVAR (Heinz et al., 2010; Schep et al., 2017), two algorithms for characterization of *de novo* or known sequence motifs associated with variation in chromatin accessibility. In principle, the HOMER algorithm scores a list of motifs within differentially accessible peaks by computing the enrichment of motifs compared with a background set of peaks (Heinz et al., 2010). As shown in Figure 2A, the binding motifs for bHLH and BES1 TF were highly enriched in all the clusters (Figure 2A). Cluster C1 also showed enrichment for the B3 TF-binding motif, supporting the notion that *LEC2* is rapidly induced by auxin and plays important roles in initiation of SE (Horstman et al., 2017b; Khanday et al., 2019). The calmodulin binding transcription activator (CAMTA)1- and WRKY-binding motifs were preferentially detected in clusters C6 and C7. It has been shown that *WRKYs* encode key immune regulators in plants (Birkenbihl et al., 2017), while CAMTA1, together with CAMTA2 and CAMTA3, were recently identified as master TFs for salicylic acid biosynthesis (Kim et al., 2020; Sun et al., 2020). Consistent with these functionalities, the genes involved in wounding and defense response were overrepresented in both clusters (Figure S2D).

It is well known that auxin induces downstream signaling events by auxin response factor (ARF) TFs (Chandler, 2016; Roosjen et al., 2018; Wang and Estelle, 2014; Weijers and Wagner, 2016). Surprisingly, the ARF-binding motif was only identified in cluster C1 but not in the clusters that functioned at earlier stages (Figure 2A). This result suggests that ARFs do not contribute to chromatin remodeling until 24 HAI. In addition, although *LEC2* and *BBM* were rapidly induced upon auxin treatment, their regulatory sequences harboring putative ARF-binding motifs were constantly accessible over the time course (Figures 2D and 2E) (Chandler, 2016; Roosjen et al., 2018). Thus, high auxin levels may play two separate roles before 24 HAI: inducing totipotent genes through ARFs and evoking rapid chromatin remodeling in an ARF-independent manner.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

CellPress

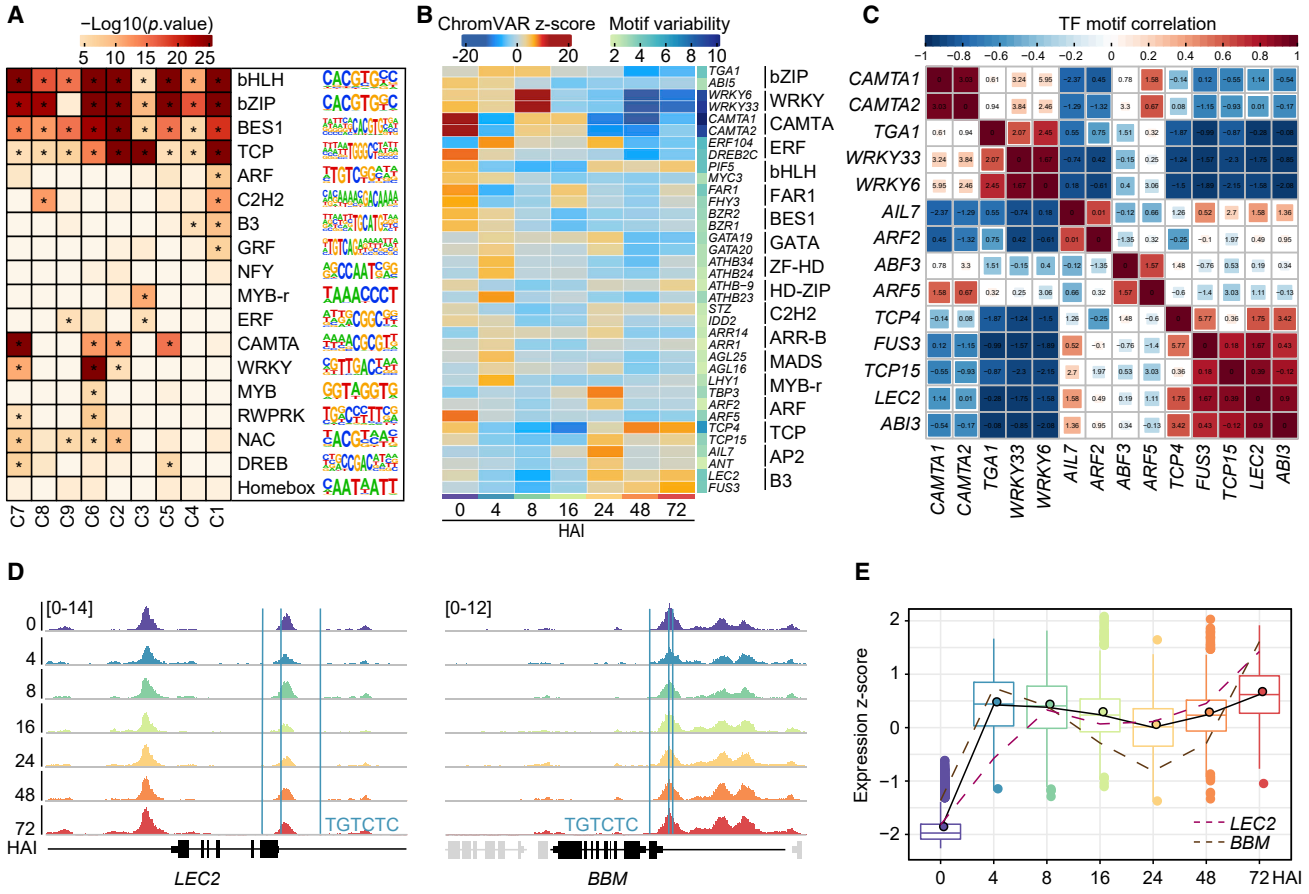


Figure 2. Chromatin-Based Hierarchical Gene Regulatory Networks for SE

(A) HOMER DNA-motif enrichment analyses of accessible peaks. The enrichment of the binding motifs of 18-TF families are shown. Asterisk stands for p value < 1e-5.

(B) Heatmap showing temporal changes in chromatin accessibility for 18-TF families (two representative TFs for each TF family) with greatest accessibility variability between samples. TF variability stands for the variability of TF binding site at different time points.

(C) Correlation and synergy between pairs of TF motifs. The color and size of each cell indicate the "the accessibility correlation level" of pairs of TF motifs. The red color indicates that the paired TF motif has the trend of co-accessibility while the blue color indicates the opposite-accessibility. The number of each cell stands for the "synergy level" between pairs of TF motifs. The positive or negative score indicates the possibility of cooperative or competitive binding in accessible peaks, respectively.

(D and E) The ATAC-seq tracks (D) and RNA-seq data (E) for the *LEC2* and *BBM* loci (cluster K3). The genomic regions are shown, and the selected genes are highlighted in black. The location of the putative ARF binding site (TGTCTC) is indicated by blue line. Expression *Z* score for *LEC2* and *BBM* at different time points are shown. Boxplot represents gene expression tendency of cluster K3 (Figure 1E).

See also Figure S2.

We next applied chromVAR to all the accessible peaks identified by our ATAC-seq datasets (Schep et al., 2017). Briefly, chromVAR first groups all accessible regions sharing the same TF-binding motif. Then it compares the observed accessibility of all peaks containing that TF-binding motif with a background set of peaks for normalization using known technical confounders. Therefore, although chromVAR does not determine the location of TF-binding motifs as is done by HOMER, it does provide a measure of the activity of TF motifs in a given sample. As an input of chromatin features, we curated a set of Arabidopsis TF position frequency matrices (PFMs, i.e., TF binding preferences) from the JASPAR (2018) database (Khan et al., 2018). Overall, chromVAR algorithm gave rise to similar results obtained by HOMER. For example, among the most variable TF motifs determined (Figure S2G; Table S3), we identified important TF-binding motifs for CAMTA1 and WRKY at 0 or 8 HAI (Figure 2B). The bZIP TF ABI5 regulates seed maturation through a subset of late embryogenesis-abundant genes (Bensmihen et al., 2005; Santos-Mendoza et al., 2008). The enrichment of bZIP TF-binding motif at 0–4 HAI implies that the gene loci involved in embryonic maturation undergo reprogramming. In support of this hypothesis, our RNA-seq survey revealed that the transcript levels of a vast number of seed maturation genes were gradually reduced upon the transfer of explants to E5 media (cluster K5, Figures 1E and S2E). At 24 HAI, the involvement of AP2, B3, and TEOSINTE BRANCHED1/CYCLOIDEA/PCF (TCP) TFs in regulating chromatin accessibility became evident.

In summary, our analyses of chromatin accessibility dynamics suggested a series of sequential TF actions involved in shaping the chromatin landscape. bHLH and BES1 TFs function

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

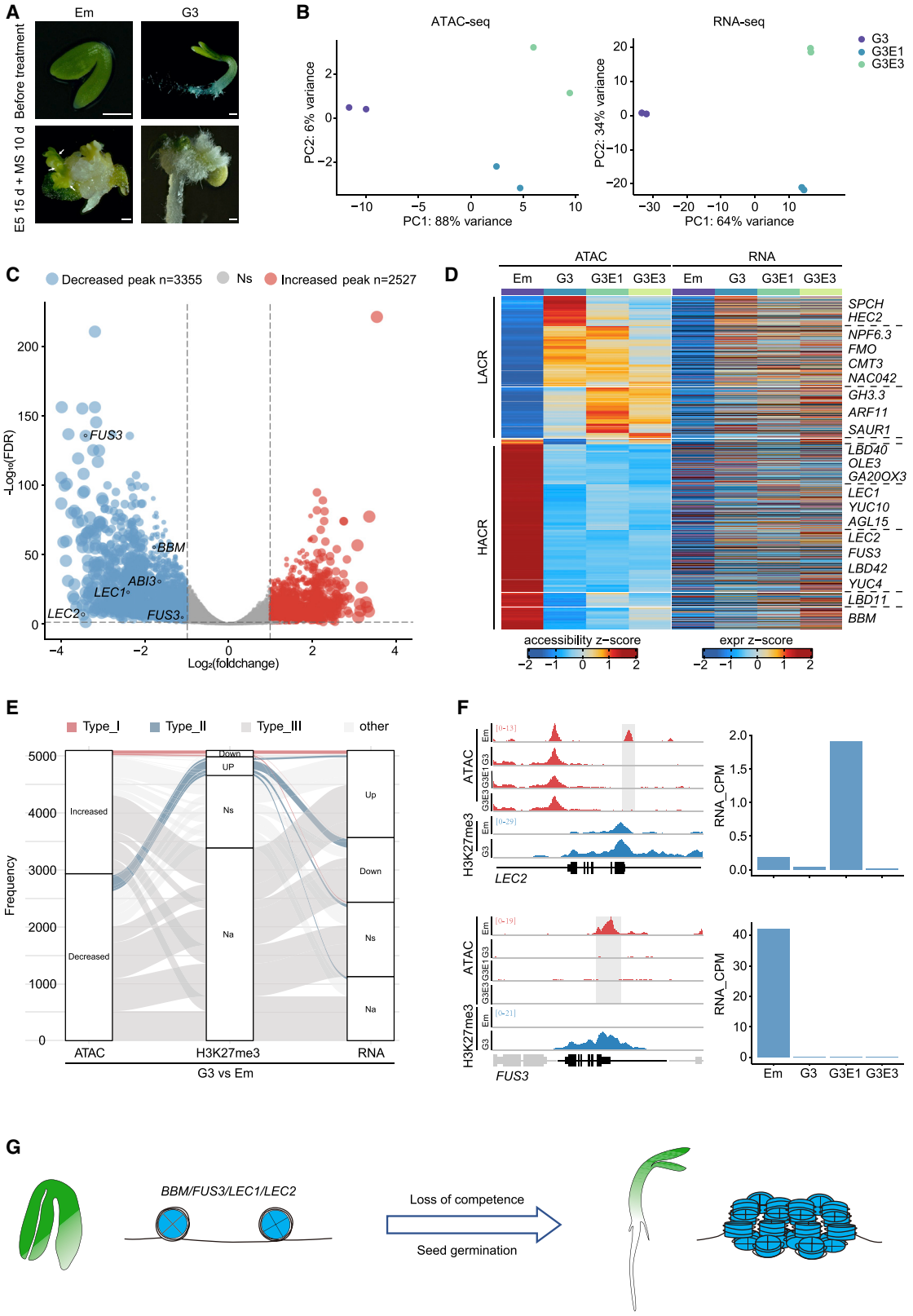


Figure 3. The Embryonic Stage Is a Prerequisite for Auxin-Induced Chromatin Accessibility and Transcriptome Alteration

(A) SE with different types of explants. Please note that G3 (3-day-old seedlings) failed to develop somatic embryos (arrows). Em, embryo. Scale bars represent 500 μm.

(legend continued on next page)

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

throughout the early stage of SE, while WRKY and CAMTA1 TFs (early-action TFs) act from 0 to 8 HAI, and ARF, AP2, B3, and TCP TFs (late-action TFs) play critical roles after 24 HAI. Importantly, the timing of the accessibility of specific TF-binding motifs coincides with altered expression of their downstream genes.

We explored the relationship between pairs of TF motifs according to the TF shared peaks and its variability in accessibility as determined by chromVAR. The "correlation" represents the co-accessibility tendency of two TF-binding motifs, whereas "synergy" stands for the frequency by which two TF-binding motifs co-localize in the same accessible peak. As shown in Figure 2C, the early-action TFs exhibited low correlation with the late-action TFs, suggesting a sequential and independent genomic remodeling process. Interestingly, we found that the B3 TF-binding motif had high correlation and synergy with the TCP TF motif, implying a collaborative action of these two TF families in the remodeling of chromatin accessibility. Similarly, the synergies of ARF-B3, ARF-TCP, and WRKY-CAMTA pairs were observed (Figure 2C).

The Embryonic Nature of Somatic Cells Is a Prerequisite for Auxin-Induced Chromatin Accessibility and Transcriptome Alteration

It is generally accepted that only immature zygotic embryos can be used as explants for SE (Custers et al., 1994; Fehér, 2015; Gaj, 2011). However, the molecular mechanism underlying this phenomenon is still unknown. To address this question, we compared the chromatin accessibility and transcriptome dynamics of immature embryos (Em), 3-day-old seedlings (G3), and 3-day-old seedlings cultured on E5 media for 1 day (G3E1) and 3 days (G3E3). Consistent with previous observations, 3-day-old seedlings were not able to regenerate (Figure 3A). All the ATAC-seq samples showed high reproducibility and similar peak distribution, with highest enrichment at TSS and TTS (Figures 3B and S3A). The comparison with G3 and Em samples revealed that 3,355 peaks closed upon seed germination whereas 2,527 peaks became more accessible (Figures 3C and S3B; Table S4). Clustering and GO term analyses revealed that closing genes were associated with the biological processes related to seed maturation (Figures 3D, S3C, and S3D; Table S4). For example, the lipid storage gene (*OLEOSIN3*, *OLE3*) and the gibberellin catabolic genes (*GIBBERELLIN 20-OXIDASE3*, *GA20ox3*) tended to be switched off. In contrast, the genes associated with opening peaks were enriched in the genes regulating auxin signaling and shoot system morphogenesis. The comparison of ATAC-seq and RNA-seq data between Em and G3 revealed that alterations in chromatin accessibility were not tightly correlated with changes in gene expression levels (Figure 3D). This discrepancy is probably due to changes in accessibility requirements for gene activation at different developmental stages. For a given gene, a decrease in accessibility in a region that is required for gene expression at the embryonic stage may not lead at to a change in expression level in the seedlings. Similarly, a gain of accessibility may represent a priming state for gene activation and so would not necessarily lead to increased transcript levels.

Of particular interest, we found that the regulatory elements of *ABI3*, *BBM*, *FUS3*, *LEC1*, and *LEC2* became closed after seed gemination (Figures 3C, 3D, and 3F). More importantly, these loci were resistant to auxin-induced chromatin accessibility alteration (Figure 3D). These results suggest that the developmental stage of explants dominates the hierarchical network of SE and that the B3 TF gene loci for totipotency are not competent for activation upon seed germination. Indeed, expression analyses found that *FUS3* was not induced in G3E1 and G3E3 samples while the *LEC2* transcripts were only transiently accumulated in G3E1 (Figure 3F). Histone modification plays an important role in gene expression; in particular, high levels of H3K27me3 are a marker for gene silencing in plants (Liu et al., 2010). However, integration of H3K27me3 chromatin immunoprecipitation assays with sequencing (ChIP-seq) datasets revealed that most peaks with a decreased level of accessibility after seed germination were not associated with high levels of H3K27me3 (Figures 3E, 3F, and S3E–S3H; Table S5).

Taken together, the above results reveal a hierarchical mechanism for cell reprogramming during SE: (1) The developmental stage of explants is at the top of the hierarchy, with the embryonic nature of explants being a prerequisite for somatic cell reprogramming. Furthermore, the loss of competence for SE upon seed germination coincides with the acute degradation of the permissive chromatin signature of AP2, B3, and NF-Y TFs (Figure 3F). (2) Auxin acts at the second tier of the hierarchical network, by inducing massive changes in chromatin accessibility and concomitant changes in gene expression. Particularly, auxin directly activates AP2 and B3 TF expression and contributes to the termination of the embryonic maturation process. (3) The AP2, B3, and NF-Y TFs, including BBM, LEC1, and LEC2 in turn serve as third-tier regulators and initiate somatic embryo formation.

(B) Principal component plots of ATAC-seq and RNA-seq data. Color code is shown. Each dot represents one sample. G3E1, 3-day-old seedling on E5 media for 1 day; G3E3, 3-day-old seedling on E5 media for 3 days.

(C) Volcano plot of the genes associated with decreased (blue) or increased (red) accessible peaks between Em and G3. The known SE genes are shown. Ns (gray in color), no difference between two samples.

(D) Heatmap of differentially accessible peaks sorted by k-means clustering approach. Color bars, accessibility *Z* score of differentially peaks identified by ATAC-seq and RNA *Z* score of the corresponding genes. HACR and LACR refer to relatively highly accessible chromatin region (Em versus G3) and relatively lowly accessible chromatin region (Em versus G3), respectively. The representative genes are shown on the right.

(E) Alluvial diagram of differentially accessible genes divided into four types (type I to III and other) according to their accessibilities, H3K27me3 deposition and transcript levels. Blue lines stand for the gene loci with decreased accessibility after gemination, accompanied with low transcript level and high level of H3K27me3. Na, not expressed or undetectable deposition of H3K27me3; Ns, no difference in transcripts or H3K27me3 levels among the samples.

(F) Representative ATAC-seq tracks and H3K27me3 ChIP-seq data for SE genes. The genomic loci are shown, and the representative genes are highlighted in black. The differentially accessible peaks at each locus are shadowed.

(G) Model. The loss of accessibility at totipotent genes loci (*BBM*, *LEC1*, *LEC2*, and *FUS3*) leads to defects in SE. Upon closure after seed germination, these loci are not competent for auxin-induced chromatin alterations.

See also Figure S3.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

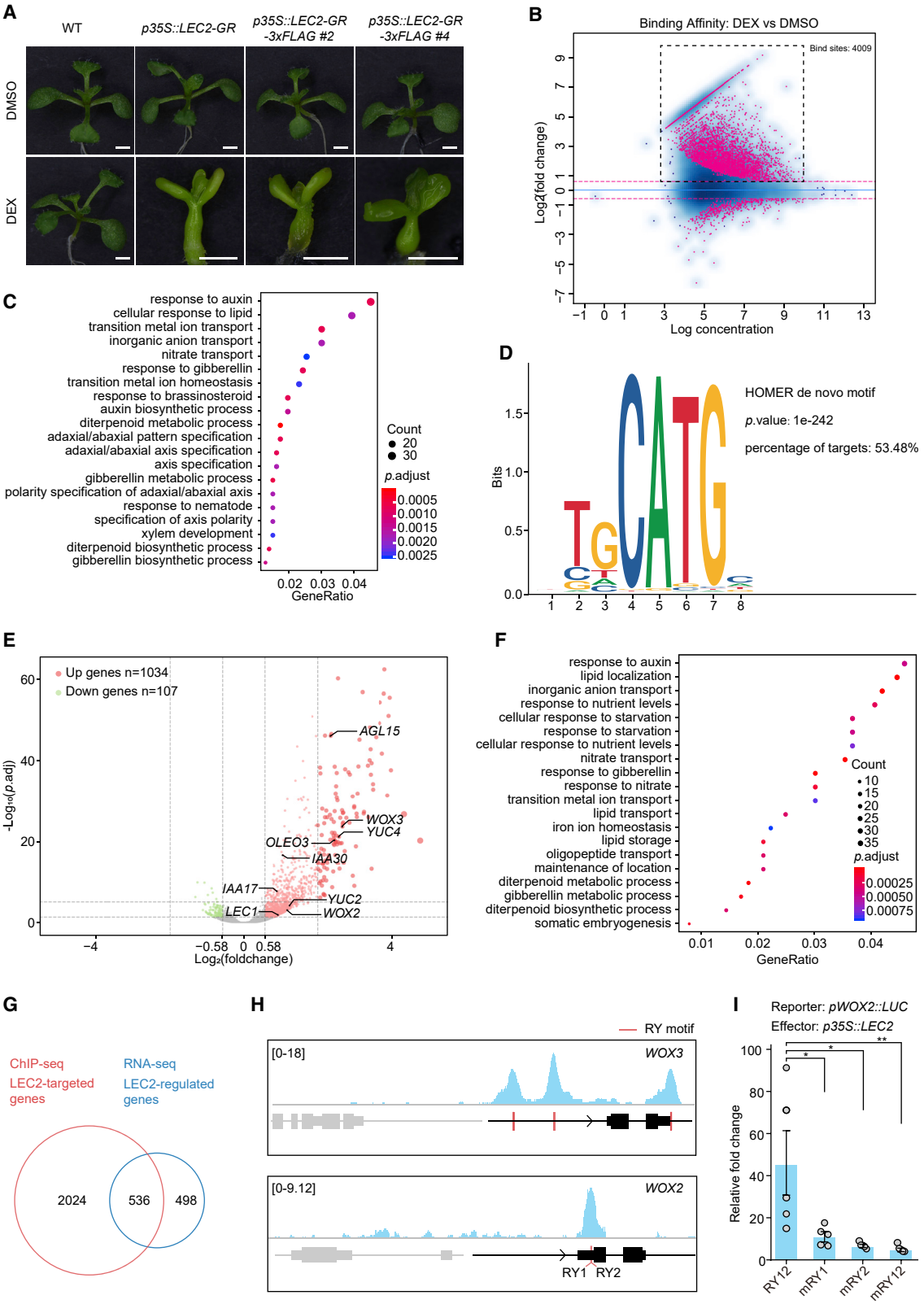


Figure 4. Identification of Downstream Targets of LEC2

(A) Induction of somatic embryos by LEC2-GR. Wild type (WT), *p35S::LEC2-GR* and two independent *p35S::LEC2-GR-3xFLAG* lines are shown. Scale bars represent 1 mm.

(legend continued on next page)

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

WOX2 and *WOX3* Are Direct Targets of LEC2

Previous work has placed BBM upstream of LEC1 and LEC2 because BBM-induced embryogenesis relies on transcriptional activation of *LEC1* or *LEC2* (Horstman et al., 2017b). Therefore, it is plausible to assume that LEC2 acts at the output node of cell totipotent genes. However, how LEC2 is linked to the early embryonic development pathway remains elusive. To answer this question, we generated two *LEC2* inducible lines (*p35S::LEC2-GR* and *p35S::LEC2-GR-3xFLAG*), in which the wild-type *LEC2* or 3xFLAG tagged *LEC2* was fused to the hormone-binding domain of the rat glucocorticoid receptor (GR) and expressed from the *35S* promoter. The chimeric *LEC2-GR/LEC2-GR-3xFLAG* gene was functional because treatment with the steroid hormone ligand dexamethasone (DEX), which leads to a translocation of the LEC2-GR/LEC2-GR-3xFLAG fusion proteins from the cytoplasm to the nucleus, induced embryo-like protuberances on explants (Figure 4A). Moreover, induction of LEC2 by DEX treatment led to increased chromatin accessibility of the gene loci which became closed upon seed germination (Figure S3I). Explants expressing *p35S::LEC2-GR-3xFLAG* were harvested 4 h after treatment with DEX or dimethyl sulfoxide (DMSO, mock) and subjected to ChIP-seq analysis. In total, we identified 4,009 LEC2 binding peaks in the genome (Figure 4B; Table S6). Consistent with earlier reports (Braybrook et al., 2006; Stone et al., 2008; Wójcikowska et al., 2013), the auxin biosynthetic genes *YUC4*, *LEC1*, and *AGL15* were identified as direct targets of LEC2 (Figures S4A and S4B). In addition, GO term analysis revealed that auxin response genes were highly enriched (Figure 4C; Table S6). It has been shown that LEC2 binds to a DNA element called the RY motif (TGCATG) (Braybrook and Harada, 2008). Indeed, the HOMER algorithm revealed that the RY motif was overrepresented in the LEC2-bound genes (Figure 4D).

To further verify the direct targets of LEC2, we performed RNA-seq using *p35S::LEC2-GR*. Upon 4 h treatment of DEX, LEC2 rapidly induced 1,034 genes and repressed 107 genes (Figures 4E and 4F). In line with GO term analysis, the expression of auxin biosynthesis and signaling genes, including *YUC2*, *YUC4*, *IAA17*, and *IAA30* was upregulated in response to the induction of LEC2 (Figures 4E and 4F). Integration of ChIP-seq data further revealed that 536 genes bound by LEC2 were induced by LEC2 (Figure 4G; Table S6). Among them, two *WUSCHEL HOMEOBOX* (*WOX*) genes, named *WOX2* and *WOX3*, were identified (Figures 4H, S4B, and S4C). The *WOX* genes form a plant-specific subclade of the eukaryotic homeobox TF superfamily (van der Graaff et al., 2009). *WOX1*, *2*, and *3* belong to the same WUS sub-clade (Breuninger et al., 2008). Expression analyses reveal that *WOX2* is initially co-expressed in the egg cell and the zygote, but becomes restricted to the apical lineages after the zygotic division (Haecker et al., 2004; Kao and Nodine, 2019). Functional study further indicates that *WOX2*, together with *WOX1*, *WOX3*, and *WOX5*, regulates embryonic shoot patterning (Breuninger et al., 2008). Because the critical roles of *WOX2* and *WOX3* in early embryonic development, we characterized their functions in SE in detail. The binding of LEC2 on the *WOX2* promoter was verified by ChIP-PCR analysis (Figure S4D). Survey of our ATAC-seq data revealed that the regulatory regions of *WOX2* and *WOX3* were accessible from 0 to 72 HAI (Figure S4E). Two RY motifs are predicted at the promoter region of *WOX2* (Figure 4H). Competitive electrophoretic mobility shift assay (EMSA) demonstrated that LEC2 bound to both *cis* elements (Figure S4F). Moreover, a transient activation assay in Arabidopsis protoplasts indicated that LEC2 was able to then activate *WOX2* reporter through binding to these two RY motifs (Figure 4I). Therefore, we conclude that *WOX2* and *WOX3* are direct targets of LEC2.

WOX2 and *WOX3* Are Essential but Not Sufficient for SE

We next performed the regeneration experiments using the immature embryo explants of *wox* single and high-order mutants. Compared with wild type, explants of the *wox2* mutant showed a reduced somatic embryo formation rate (Figures 5A–5C), while regeneration of *wox1* and *wox1 wox3* explants was indistinguishable from wild type. The *wox2 wox3* double mutation nearly abolished somatic embryo formation. It has been shown that *wox1 wox2 wox3* triple mutants displayed a strong increase in the frequency of *wox2* defects in embryonic shoot patterning (~30% embryos showing a single-cotyledon phenotype) (Breuninger et al., 2008). However, the defect of *wox2 wox3* explants in regeneration was slightly enhanced by the mutation in *WOX1* (Figures 5B and 5C), suggesting that *WOX1* might act redundantly with *WOX2* and *WOX3* in SE. It should be noted that the impairment of *wox2 wox3* and *wox1 wox2 wox3* in SE is not likely due to their developmental defects because the single-cotyledon embryos were excluded in our SE experiments (Figure S5A). SE may be induced by an indirect developmental route (Horstman et al., 2017a). The regeneration of somatic embryos was also compromised when the embryonic calli of *wox1 wox2*

(B) Identification of LEC2 targets by ChIP-seq. The *p35S::LEC2-GR* explants cultured on B5 media were treated with DEX or DMSO (mock) for 4 h and subjected to ChIP-seq analyses. In total, 4,009 binding sites were identified, fold change >1.5, FDR < 0.05.
(C) GO term analysis of putative LEC2 targets. The selected 20 categories are shown.
(D) The enrichment of the RY motif (TGCATG) in LEC2-bound genes. p value and percent of targets are given. e represents times ten raised to the power ($\times 10^n$).
(E) Volcano plot for RNA-seq analysis. 3-day-old seedlings were treated with DEX or DMSO for 4 h in liquid B5 medium and subjected to RNA-seq analyses. 1,034 genes were induced while 107 genes were repressed. FDR < 0.05; $\log_2$(fold change) > 0.58 or < −0.58.
(F) GO enrichment analysis of 1,034 upregulated genes. Twenty, GO processes with the largest gene ratios are plotted in order of gene ratio. The size of the dots represents the number of genes associated with GO term and the color of the dots represents the *p*.adj values.
(G) Venn diagrams showing overlapped genes identified by RNA-seq and ChIP-seq experiments. The LEC2-targeted genes (n = 2,560) are defined by $\log_2$(fold change) >4, FDR < 0.05; the LEC2-regulated genes (n = 1,034) are defined by $\log_2$(fold change) >0.58, FDR < 0.05.
(H) The ChIP-seq tracks for *WOX2* and *WOX3* loci. The genomic regions are shown, and the selected genes are highlighted in black. The location of the RY motif is indicated in red line. Four regions for ChIP-PCR analysis (see Figure S4E) are given.
(I) Arabidopsis protoplast transient activation assays. *p35S::LEC2* was used as effector. RY12, the wild-type *WOX2* promoter; mRY1, mRY2, and mRY12, the *WOX2* promoter with the mutation in RY motif 1, RY motif 2 or both RY1 and RY2 motifs, respectively. Relative LUCIFERASE (LUC) activity was normalized to that of Renilla (REN). Relative fold change refers to LUC/REN in *p35S::LEC2 pWOX2::LUC* versus LUC/REN in *pWOX2::LUC*. Data are represented as mean ± SEM (n = 5). Student's t test, *p < 0.05, **p < 0.01; ns, not significant.
See also Figure S4.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

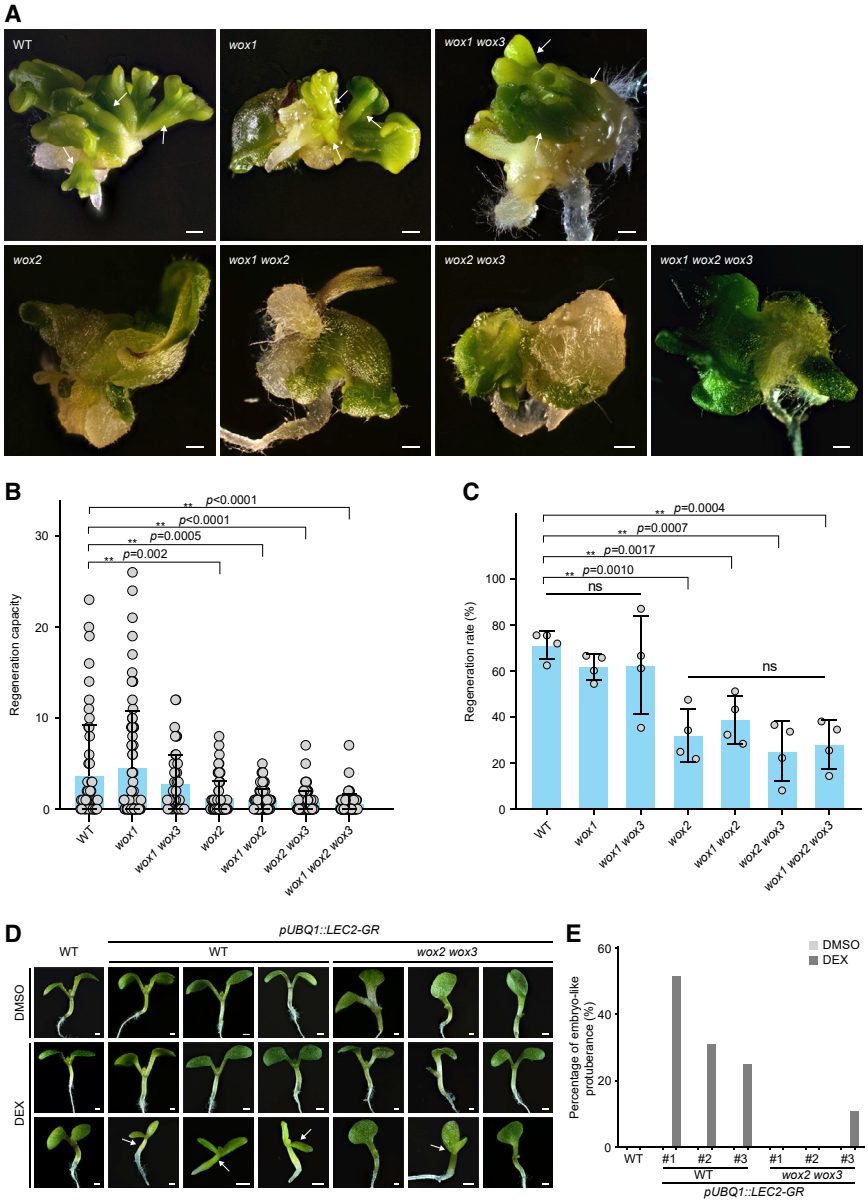


Figure 5. Essential Role of *WOX2* and *WOX3* in SE

(A) Regeneration assays using the immature embryo as explants from wild-type (WT) and *wox* mutants. Arrow indicates somatic embryo. Scale bars represent 500 μm.

(B) Regeneration capacity analysis. Wild-type, *wox* single and high-order mutants were used. The regeneration capacity was represented by the number of somatic embryos per explant (n > 60). Data are from one representative experiment of four independent experiments. Data are represented as mean ± SD; each dot represents the number of somatic embryos per explant. One-way ANOVA was performed followed by a Turkey's multiple comparisons test, **p < 0.01.

(C) Regenerative rate analysis. The regenerative rate was represented by the percentage of explants with somatic embryos. Data are from four independent experiments. Data are represented as mean ± SD. **p < 0.01.

(D and E) Regeneration assays using *pUBQ1*::*LEC2-GR* and *pUBQ1::LEC2-GR wox2 wox3* explants. The seedling cultured on the E5 medium without (mock) or with 10 μM DEX treatment. Two independent lines for each construct are shown (D). Arrow indicates embryo-like protuberance. The regeneration rate (E) was represented by the percentage of explants with embryo-like protuberance. Data from three independent experiments are shown. Scale bars (D) represent 1 mm. See also Figure S5.

wox3 triple mutants were used as explants (Figures S5B–S5D). Thus, *WOX2* and *WOX3* are essential for SE irrespective of regeneration protocols.

To understand whether *WOX2* is sufficient to trigger SE, we generated an inducible line for *WOX2*, in which WOX2-GR fusion protein was expressed from constitutive *RIBO* promoter (*pRibo::WOX2-GR*, At2g18020). However, unlike LEC2-GR, the induction of WOX2-GR by DEX treatment did not trigger somatic embryo formation on the seedlings grown on MS media (Figures S5E and S5F). Instead, we occasionally observed green rod-like structures forming on the roots (Figures S5E and S5F). Therefore, these results indicate that *WOX2* and *WOX3* are necessary but not sufficient to induce somatic embryo formation.

WOX8 expression, together with *WOX9*, is restricted to the basal daughter cell (Haecker et al., 2004). Genetic and phenotypic analyses reveal that both genes are required for programming gene expression and normal development in both the basal and apical embryo lineages at the early zygotic stage (Breuninger et al., 2008). In contrast to *wox2 wox3* and *wox1 wox2 wox3* mutants, the *wox8* mutant did not exhibit an altered somatic embryo formation rate (Figures S5G and S5H), suggesting that the apical lineage-specific *WOX* genes play a dominant role in SE.

Finally, to probe whether LEC2 promotes SE through *WOX2* and *WOX3*, we generated a transgenic plant (*pUBQ1::LEC2-GR*), in which *LEC2-GR* was expressed from the *UBIQUITIN1* (*UBQ1*, At3g52590) promoter. The usage of *UBQ1* promoter prevented gene silencing caused by the *35S* promoter in the transfer-DNA (T-DNA) insertion mutants of *WOX2* and *WOX3*. As shown in Figures 5D and 5E, upon DEX treatment, *pUBQ1::LEC2-GR* developed embryo-like protuberances on the seedlings. In contrast, these tissues were rarely observed in *pUBQ1::LEC2-GR wox2 wox3* explants (Figures 5D and 5E).

Expression Pattern of *WOX2*, *WOX3*, and *LEC2* during SE

During normal embryogenesis, *WOX2* is expressed in the apical lineage and is barely detectable in the immature embryos at late-bent-cotyledon stage of development (Breuninger et al., 2008; Haecker et al., 2004). To precisely analyze *WOX2* expression pattern during SE, we performed *in situ* hybridization assays. As shown in Figures 6A and 6C, the *WOX2* transcripts

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

CellPress

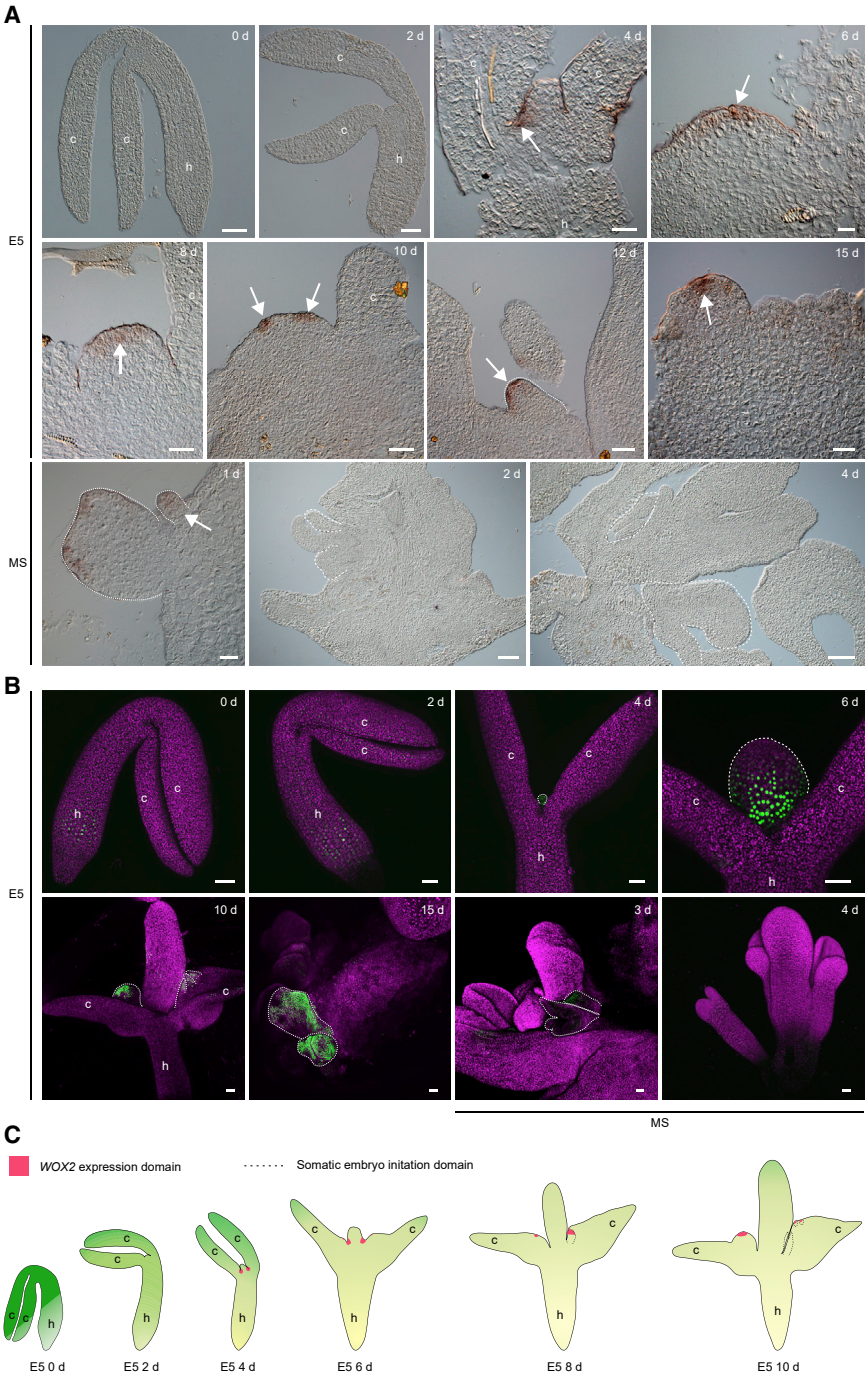


Figure 6. The Expression Pattern of *WOX2* during SE

(A) Expression of *WOX2* (arrow) in the wild-type explants. The explants were cultured on E5 media for 15 days and transferred to MS media for 4 days. The samples were collected at different time points. Dash lines, embryo-like protuberance; c, cotyledon; h, hypocotyl. Scale bars represent 50 μm.

(B) Expression of *pWOX2::3xVENUS-N7* (green) during SE. Max intensity projection in the z-axis of one representative sample is shown. Scale bars represent 50 μm.

(C) Schematic drawing of the various stages of SE and expression domain of *WOX2*.

See also Figure S6.

became detectable at the boundary domain between shoot apical meristem and cotyledons of the explants at day 4 or day 6 after treatment, respectively. This domain overlapped with the region where a somatic embryo would eventually form (Figures 6A and 6C), suggesting that *WOX2* serves as an early-stage marker for the progenitor cells for SE. *WOX3* showed a similar expression pattern as *WOX2* (Figure S6A). Notably, the expression domain of both genes coincides with that of *LEC2* (Figure S6A), which is in agreement with a direct activation of *WOX2* and *WOX3* by LEC2. The *WOX2* expression was greatly reduced once the explants were transferred to MS media (Figures 6A and 6C).

To confirm the above results, we examined the expression of a nuclear yellow fluorescent protein (3xVENUS-N7) driven by *WOX2* promoter. In contrast to the expression pattern revealed by *in situ* hybridization assays, VENUS signals could be detected at the basal region of hypocotyls as well as the cells in the cotyledon cells before 4 HAI (Figures 6B and S6B). From 4 HAI onward, strong VENUS signals were observed at the boundary domain of the shoot apical meristem, were persistently detectable in the apical region of developing somatic embryos, and diminished after somatic embryos converted to plantlets (Figures 6B and S6B).

DISCUSSION

A Hierarchical Transcriptional Network for SE

On the basis of chromatin accessibility dynamics, we found that the developmental stage is at the top of the regulatory hierarchy that governs SE (Figure 7). This finding explains why post-embryonic somatic tissues are resistant to reprogramming for SE. It is very likely that the cellular status of the juvenile phase is less amenable to the reshaping of the chromatin status of the gene loci determining totipotency, albeit the precise molecular mechanism for this is still unknown. One possible candidate gene involved in this process is VAL (VP1/ABI3-LIKE), which suppresses ABI3/FUS3/LEC2 function in order to initiate germination and vegetative development through facilitating the deposition of the histone H2AK119 ubiquitination (H2Aub) marker (Suzuki and McCarty, 2008; Suzuki et al., 2007; Yang et al., 2013). In this scenario, H2Aub marking of ABI3/FUS3/LEC2 leads to their initial repression, which is further maintained by PRC2-mediated H3K27me3 and might contribute to the

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

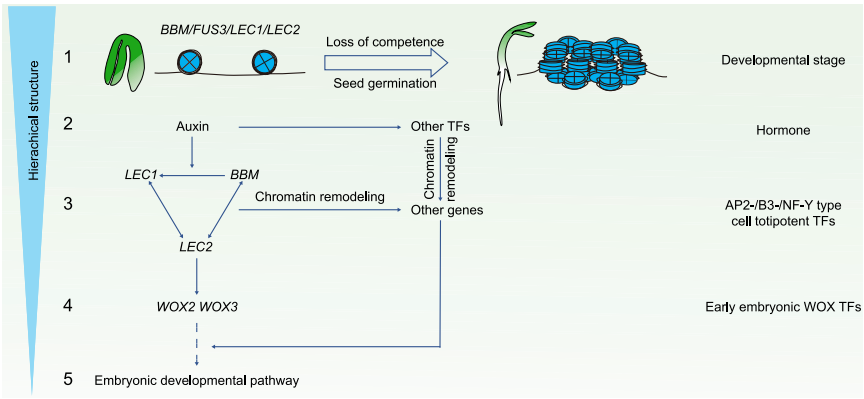


Figure 7. Hierarchical Mechanism for SE

Five hierarchical levels are shown. The cell totipotent genes and auxin form a feed-forward loop to reinforce cell fate transition. LEC2 acts at the output node of cell totipotent gene network by direct activation of early embryonic development genes, such as *WOX2* and *WOX3*. see text for detail.

loss of reprogramming competence in the somatic cells after seed germination.

We further classified TFs into different hierarchical levels depending on their abilities to induce SE in the absence of auxin (Figure 7). For example, overexpression of a single high-hierarchical TF, such as *LEC2* and *BBM* is sufficient to induce SE without auxin (Horstman et al., 2017b); In contrast, overexpression of *WOX2* alone was insufficient to trigger cell fate transition in explants in this study. These observations imply that either LEC2 or BBM triggers somatic embryo formation through multiple downstream targets.

Growing evidence indicates that high-hierarchical TFs form a feed-forward loop to reinforce cell fate transition (Figure 7). For instance, it is shown that BBM transcriptionally regulates *LEC1* and *LEC2* while BBM-induced SE is functionally dependent on LEC1 and other genes (Horstman et al., 2017b). Although the *lec2* mutant displays mild defects during early zygotic embryogenesis, it exhibits severe defects in somatic embryo induction from immature embryos (Gaj et al., 2005). This discrepancy suggests that the induction of SE might rely on LEC2 function during late embryogenesis, where it maintains embryonic fate. The compromised embryonic identity in *lec2* immature embryos leads to reduced competence for auxin-induced SE. However, the function of *LEC2* in SE might be not limited to this. Our results, together with the findings from the Harada and Boutilier labs (Horstman et al., 2017b; Stone et al., 2008), indicate that LEC2 directly induces embryogenesis by activating early embryonic genes *WOX2* and *WOX3*. At the same time, LEC2 reinforces cell fate transition through upregulation of auxin biosynthetic genes as well as other totipotent genes including *BBM* and *LEC1* (Figure 7).

The identity of the progenitor cell of somatic embryos within explants is poorly understood. Histological studies have found that somatic embryos originate from the cells in the protodermis and subprotodermis of the adaxial side of cotyledons and that they can display a single- or multicellular origin (Kurczyńska et al., 2007). Thus, profiling of the chromatin accessibility dynamics and gene expression at the single-cell level would shed light on how auxin spatiotemporally induces cell reprogramming.

Auxin Evokes Cell Reprogramming at the Chromatin Level

In the current protocol, the embryo induction and early development phase requires 15 days of explant culturing on E5 induction medium (Gaj, 2011). However, our results suggest that auxin reprograms the chromatin accessibility of explants within 3 days (Figure 1B, 48 HAI versus 72 HAI). More interestingly, our HOMER and chromVAR analyses revealed that auxin rapidly evokes chromatin remodeling at 0 to 8 HAI independent of ARFs (Figures 2A and 2B). Since WRKY and CAMTA1 TF-binding motifs were overrepresented at this stage, it is plausible to assume that, in addition to inducing totipotent gene expression through the canonical AUX/IAA-ARF signaling cascade, the high-auxin environment may trigger a general stress response in the explants. This finding is in agreement with earlier reports that abiotic stress is able to induce somatic embryo formation (Fehér, 2015; Gaj, 2011). How the transient induction of the genes involved in immunity and abiotic stress contributes to SE awaits further investigations.

Our results showed that the ARF-binding motif was enriched in the cluster C1 after 24 HAI. How do ARFs achieve this reprogramming effect at the chromatin level? The recent studies of the role of ARFs in auxin response and floral cell fate transition provide some useful hints (Szemenyei et al., 2008; Weijers and Wagner, 2016; Wu et al., 2015). Based on the model proposed by the Wagner lab, we envision that an unknown Aux/IAA-ARF complex associates with target loci and prevents their expression, either by recruiting the TOPLESS-histone deacetylase complex or by preventing recruitment of the chromatin-remodeling complex at 0 HAI. Upon auxin treatment, Aux/IAA proteins are degraded, which leads to eviction of TPL-histone deacetylase and recruitment of chromatin-remodeling complex. The chromatin-remodeling complex in turn opens up the compacted chromatin at target genomic DNA near ARF-bound sites by reducing nucleosome occupancy. Thus, the identification of the ARFs involved in SE and the elucidation of their underlying mechanism are two important research directions.

The Comparison between SE and Suspensor Embryogenesis

Recently, the Weijers lab used an auxin-dependent suspensor embryogenesis as a model to determine transcriptome changes during embryonic reprogramming (Radoeva et al., 2019a). In contrast to our method, they induced suspensor embryogenesis by blocking auxin signaling. We noted that SE and suspensor embryogenesis share some similarities at the molecular level. (1) Radoeva et al. revealed that the cell totipotent genes, such as *FUS3* and *LEC1-like* (*L1L*) are upregulated at the early stage of suspensor embryogenesis. (2) Radoeva et al. further uncovered that four auxin-regulated bHLH genes, especially *bHLH49* and *bHLH100*, are required for normal embryo development,

particularly for the embryo-suspensor junction. Moreover, overexpression of *bHLH49* is able to induce proliferative cell divisions and (subsequent) suspensor identity loss. Interestingly, we found that three of them (*bHLH49*, *bHLH60*, and *bHLH100*) are progressively induced during SE (Figure S7A). (3) Our results indicate that *WOX2* and *WOX3* play important roles in SE. Intriguingly, *WOX2* is also upregulated when auxin signaling is blocked in suspensor (Figure S7B). Altogether, these results suggest that suspensor embryogenesis may use the cell totipotent genes described here as a molecular trigger, and that *WOX2* and probably *WOX3* as well, are essential for suspensor embryogenesis. In this scenario, different artificial embryogenesis routes probably deploy the same hierarchical transcriptional regulatory network (Figure 7).

Functional Redundancy, Compensation, and Specificity among *WOX* Genes

Based on assays for explant regeneration ability, *WOX2* seems to play a more important role in somatic embryo formation than *WOX1* and *WOX3*. The phenotypic severity of the *wox2* mutant was further enhanced in the *wox1 wox3* double-mutant background, indicating that these three *WOX* genes are functionally redundant. Alternatively, the loss of *WOX2* function may lead to the activation of a compensatory network to buffer against the impairment of SE. Indeed, a similar genetic compensation induced by deleterious mutations has been found in a zebrafish model (Rossi et al., 2015).

Interestingly, *WOX8* is a basal-lineage-specific *WOX* gene, and its mutation did not result in a defect in somatic embryo induction. In agreement with this, we did not detect the activation of *WOX8* within 72 HAI (Table S2). Moreover, it was recently reported that somatic and zygotic embryo transcriptomes are distinct from each other (Hofmann et al., 2019). Thus, these observations collectively suggest that the somatic embryo is formed through a previously unknown developmental pathway. The identification of the downstream targets of *WOX2* and comparison between zygotic and somatic embryos at transcriptome and chromatin levels are important future research directions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Plant Growth Conditions
  - Constructs
  - Generation of Transgenic Plants
  - SE
  - Plant Materials Used for ATAC-seq, ChIP-seq and RNA-seq
  - ATAC-seq Experiment
  - RNA-seq, ChIP-PCR and ChIP-seq Experiments
  - Transient Protoplast Analysis
  - Expression Analyses
  - EMSA
  - RNA *In situ* Hybridization
  - Tissue Embedding and Sectioning
  - Microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data Preprocessing and Reads Alignment
  - RNA-seq Data Analyses
  - ChIP-seq Data Analyses
  - ATAC-seq Data Analyses

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.devcel.2020.07.003.

ACKNOWLEDGMENTS

We thank Arabidopsis Biological Resource Center for T-DNA insertion lines; Dr. Xingliang Hou (South China Botanical Garden, CAS) for the *lec2* seeds; Dr. Hongtao Liu (NKLPMG, SIPPE, CAS) for the dual-LUC system; Hong Zhu for technical support in FACS (Cell Biology Core Facility of Shanghai Center for Plant Stress Biology, Center for Excellence in Molecular Plant Sciences, CAS); members in J.-W.W. lab for discussion and comments on the manuscript. This work was supported by National Natural Science Foundation of China grant numbers 31430013, 31222029, 912173023, and 31525004 to J.-W.W. and Strategic Priority Research Program of the Chinese Academy of Sciences grant number XDB27030101 to J.-W.W.

AUTHOR CONTRIBUTIONS

F.-X.W. and J.-W.W. designed the research; F.-X.W. performed most of the experiments; G.-D.S. and Z.-G.X. analyzed ATAC-seq, ChIP-seq, and RNA-seq data; L.-Y.W. established ATAC-seq protocol; X.-Y.Z. purified LEC2-DBD protein; F.-X.W., G.-D.S., and J.-W.W. analyzed the rest of the data; J.-W.W. wrote the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 11, 2020
Revised: June 2, 2020
Accepted: July 8, 2020
Published: August 4, 2020

REFERENCES

Adamowski, M., and Friml, J. (2015). PIN-dependent auxin transport: action, regulation, and evolution. Plant Cell *27*, 20–32.

Bai, B., Su, Y.H., Yuan, J., and Zhang, X.S. (2013). Induction of somatic embryos in Arabidopsis requires local YUCCA expression mediated by the down-regulation of ethylene biosynthesis. Mol. Plant *6*, 1247–1260.

Bajic, M., Maher, K.A., and Deal, R.B. (2018). Identification of open chromatin regions in plant genomes using ATAC-seq. Methods Mol. Biol. *1675*, 183–201.

Bensmihen, S., Giraudat, J., and Parcy, F. (2005). Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the ABI5 family during *Arabidopsis thaliana* embryo maturation. J. Exp. Bot. *56*, 597–603.

Birkenbihl, R.P., Liu, S., and Somssich, I.E. (2017). Transcriptional events defining plant immune responses. Curr. Opin. Plant Biol. *38*, 1–9.

Birnbaum, K.D., and Roudier, F. (2017). Epigenetic memory and cell fate reprogramming in plants. Regeneration *4*, 15–20.

Birnbaum, K.D., and Sánchez Alvarado, A. (2008). Slicing across kingdoms: regeneration in plants and animals. Cell *132*, 697–710.

Boutilier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.M., van Lammeren, A.A., Miki, B.L., et al. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. Plant Cell *14*, 1737–1749.

Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.P., Grini, P.E., Colot, V., et al. (2011). Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. PLoS Genet. *7*, e1002014.

Braybrook, S.A., and Harada, J.J. (2008). LECs go crazy in embryo development. Trends Plant Sci. *13*, 624–630.

Braybrook, S.A., Stone, S.L., Park, S., Bui, A.Q., Le, B.H., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (2006). Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. Proc Natl Acad Sci USA *103*, 3468–3473.

Breuninger, H., Rikirsch, E., Hermann, M., Ueda, M., and Laux, T. (2008). Differential expression of WOX genes mediates apical-basal axis formation in the Arabidopsis embryo. Dev. Cell *14*, 867–876.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Carroll, T.S., Liang, Z., Salama, R., Stark, R., and de Santiago, I. (2014). Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. Front. Genet. *5*, 75.

Chandler, J.W. (2016). Auxin response factors. Plant Cell Environ. *39*, 1014–1028.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J. (2004). Interaction of polycomb-group proteins controlling flowering in Arabidopsis. Development *131*, 5263–5276.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890.

Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S., and Town, C.D. (2017). Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant J. *89*, 789–804.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. *16*, 735–743.

Custers, J.B., Cordewener, J.H., Nöllen, Y., Dons, H.J., and Van Lockeren Campagne, M.M. (1994). Temperature controls both gametophytic and sporophytic development in microspore cultures of *Brassica napus*. Plant Cell Rep. *13*, 267–271.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048.

Fehér, A. (2015). Somatic embryogenesis - stress-induced remodeling of plant cell fate. Biochim. Biophys. Acta *1849*, 385–402.

Gaj, M.D. (2011). Somatic embryogenesis and plant regeneration in the culture of *Arabidopsis thaliana* (L.) Heynh. immature zygotic embryos. Methods Mol. Biol. *710*, 257–265.

Gaj, M.D., Zhang, S., Harada, J.J., and Lemaux, P.G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of Arabidopsis. Planta *222*, 977–988.

Gallois, J.L., Nora, F.R., Mizukami, Y., and Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. Genes Dev. *18*, 375–380.

Gliwicka, M., Nowak, K., Balazadeh, S., Mueller-Roeber, B., and Gaj, M.D. (2013). Extensive modulation of the transcription factor transcriptome during somatic embryogenesis in *Arabidopsis thaliana*. PLoS One *8*, e69261.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849.

Haecker, A., Gross-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrmann, M., and Laux, T. (2004). Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. Development *131*, 657–668.

Harding, E.W., Tang, W., Nichols, K.W., Fernandez, D.E., and Perry, S.E. (2003). Expression and maintenance of embryogenic potential is enhanced through constitutive expression of AGAMOUS-Like 15. Plant Physiol. *133*, 653–663.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hofmann, F., Schon, M.A., and Nodine, M.D. (2019). The embryonic transcriptome of *Arabidopsis thaliana*. Plant Reprod. *32*, 77–91.

Horstman, A., Bemer, M., and Boutilier, K. (2017a). A transcriptional view on somatic embryogenesis. Regeneration *4*, 201–216.

Horstman, A., Li, M., Heidmann, I., Weemen, M., Chen, B., Muino, J.M., Angenent, G.C., and Boutilier, K. (2017b). The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. Plant Physiol. *175*, 848–857.

Ikeuchi, M., Favero, D.S., Sakamoto, Y., Iwase, A., Coleman, D., Rymen, B., and Sugimoto, K. (2019). Molecular mechanisms of plant regeneration. Annu. Rev. Plant Biol. *70*, 377–406.

Ikeuchi, M., Iwase, A., Rymen, B., Harashima, H., Shibata, M., Ohnuma, M., Breuer, C., Morao, A.K., de Lucas, M., De Veylder, L., et al. (2015). PRC2 represses dedifferentiation of mature somatic cells in Arabidopsis. Nat. Plants *1*, 15089.

Ikeuchi, M., Ogawa, Y., Iwase, A., and Sugimoto, K. (2016). Plant regeneration: cellular origins and molecular mechanisms. Development *143*, 1442–1451.

Iwase, A., Mitsuda, N., Koyama, T., Hiratsu, K., Kojima, M., Arai, T., Inoue, Y., Seki, M., Sakakibara, H., Sugimoto, K., and Ohme-Takagi, M. (2011). The AP2/ERF transcription factor WIND1 controls cell dedifferentiation in Arabidopsis. Curr. Biol. *21*, 508–514.

Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J., and Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. *45*, D1040–D1045.

Kao, P., and Nodine, M.D. (2019). Transcriptional activation of Arabidopsis zygotes is required for initial cell divisions. Sci. Rep. *9*, 17159.

Kareem, A., Durgaprasad, K., Sugimoto, K., Du, Y., Pulianmackal, A.J., Trivedi, Z.B., Abhayadev, P.V., Pinon, V., Meyerowitz, E.M., Scheres, B., and Prasad, K. (2015). PLETHORA genes control regeneration by a two-step mechanism. Curr. Biol. *25*, 1017–1030.

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Cheneby, J., Kulkarni, S.R., Tan, G., et al. (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. *46*, D260–D266.

Khan, A., and Mathelier, A. (2017). Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. BMC Bioinformatics *18*, 287.

Khanday, I., Skinner, D., Yang, B., Mercier, R., and Sundaresan, V. (2019). A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. Nature *565*, 91–95.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.

Kim, Y., Gilmour, S.J., Chao, L., Park, S., and Thomashow, M.F. (2020). Arabidopsis CAMTA transcription factors regulate pipecolic acid biosynthesis and priming of immunity genes. Mol. Plant *13*, 157–168.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. *20*, 207–220.

Kurczyńska, E.U., Gaj, M.D., Ujczak, A., and Mazur, E. (2007). Histological analysis of direct somatic embryogenesis in *Arabidopsis thaliana* (L.) Heynh. Planta *226*, 619–628.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. Annu. Rev. Plant Biol. *61*, 395–420.

Liu, J., Deng, S., Wang, H., Ye, J., Wu, H.W., Sun, H.X., and Chua, N.H. (2016). CURLY LEAF regulates gene sets coordinating seed size and lipid biosynthesis. Plant Physiol. *171*, 424–436.

Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. Cell *93*, 1195–1205.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.J., Scelonge, C., Lenderts, B., Chamberlin, M., Cushatt, J., et al. (2016). Morphogenic regulators Baby boom and Wuschel improve monocot transformation. Plant Cell *28*, 1998–2015.

Lu, Z., Hofmeister, B.T., Vollmers, C., DuBois, R.M., and Schmitz, R.J. (2017). Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Res. *45*, e41.

Lu, Z., Marand, A.P., Ricci, W.A., Ethridge, C.L., Zhang, X., and Schmitz, R.J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. Nat. Plants *5*, 1250–1259.

Ludwig, L.S., Lareau, C.A., Bao, E.L., Nandakumar, S.K., Muus, C., Ulirsch, J.C., Chowdhary, K., Buenrostro, J.D., Mohandas, N., An, X., et al. (2019). Transcriptional states and chromatin accessibility underlying human erythropoiesis. Cell Rep. *27*, 3228–3240.e7.

Maher, K.A., Bajic, M., Kajala, K., Reynoso, M., Pauluzzi, G., West, D.A., Zumstein, K., Woodhouse, M., Bubb, K., Dorrity, M.W., et al. (2018). Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. Plant Cell *30*, 15–36.

Méndez-Hernández, H.A., Ledezma-Rodríguez, M., Avilez-Montalvo, R.N., Juárez-Gómez, Y.L., Skeete, A., Avilez-Montalvo, J., De-la-Peña, C., and Loyola-Vargas, V.M. (2019). Signaling overview of plant somatic embryogenesis. Front. Plant Sci. *10*, 77.

Menges, M., de Jager, S.M., Gruissem, W., and Murray, J.A. (2005). Global analysis of the core cell cycle regulators of Arabidopsis identifies novel genes, reveals multiple and highly specific profiles of expression and provides a coherent model for plant cell cycle control. Plant J. *41*, 546–566.

Mozgová, I., Muñoz-Viana, R., and Hennig, L. (2017). PRC2 represses hormone-induced somatic embryogenesis in vegetative tissue of *Arabidopsis thaliana*. PLoS Genet. *13*, e1006562.

Palovaara, J., de Zeeuw, T., and Weijers, D. (2016). Tissue and organ initiation in the plant embryo: a first time for everything. Annu. Rev. Cell Dev. Biol. *32*, 47–75.

Pinon, V., Prasad, K., Grigg, S.P., Sanchez-Perez, G.F., and Scheres, B. (2013). Local auxin biosynthesis regulation by PLETHORA transcription factors controls phyllotaxis in Arabidopsis. Proc. Natl. Acad. Sci. USA *110*, 1107–1112.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Radoeva, T., Lokerse, A.S., Llavata-Peris, C.I., Wendrich, J.R., Xiang, D., Liao, C.Y., Vlaar, L., Boekschoten, M., Hooiveld, G., Datla, R., and Weijers, D. (2019a). A robust auxin response network controls embryo and suspensor development through a basic helix loop helix transcriptional module. Plant Cell *31*, 52–67.

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

Radoeva, T., and Weijers, D. (2014). A roadmap to embryo identity in plants. Trends Plant Sci. *19*, 709–716.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191.

Reinert, J. (1958). Morphogenese und ihre Kontrolle an Gewebekulturen aus Carotten. Naturwissenschaften *45*, 344–345.

Ricci, W.A., Lu, Z., Ji, L., Marand, A.P., Ethridge, C.L., Murphy, N.G., Noshay, J.M., Galli, M., Mejía-Guerra, M.K., Colomé-Tatché, M., et al. (2019). Widespread long-range cis-regulatory elements in the maize genome. Nat. Plants *5*, 1237–1249.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Roosjen, M., Paque, S., and Weijers, D. (2018). Auxin response factors: output control in auxin biology. J. Exp. Bot. *69*, 179–188.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., and Stainier, D.Y. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature *524*, 230–233.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389–393.

Rymen, B., Kawamura, A., Lambolez, A., Inagaki, S., Takebayashi, A., Iwase, A., Sakamoto, Y., Sako, K., Favero, D.S., Ikeuchi, M., et al. (2019). Histone acetylation orchestrates wound-induced transcriptional activation and cellular reprogramming in Arabidopsis. Commun. Biol. *2*, 404.

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., and Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. Plant J. *54*, 608–620.

Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978.

Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Genet. *20*, 437–455.

Sena, G., and Birnbaum, K.D. (2010). Built to rebuild: in search of organizing principles in plant regeneration. Curr. Opin. Genet. Dev. *20*, 460–465.

Smertenko, A., and Bozhkov, P.V. (2014). Somatic embryogenesis: life and death processes during apical-basal patterning. J. Exp. Bot. *65*, 1343–1360.

Soriano, M., Li, H., Jacquard, C., Angenent, G.C., Krochko, J., Offringa, R., and Boutilier, K. (2014). Plasticity in cell division patterns and auxin transport dependency during in vitro embryogenesis in Brassica napus. Plant Cell *26*, 2568–2581.

Stephens, M. (2017). False discovery rates: a new deal. Biostatistics *18*, 275–294.

Steward, F.C., Mapes, M.O., and Mears, K. (1958). Growth and organized development of cultured cells. II. Organization in cultures grown from freely suspended cells. Am. J. Bot. *45*, 705–708.

Stone, S.L., Braybrook, S.A., Paula, S.L., Kwong, L.W., Meuser, J., Pelletier, J., Hsieh, T.F., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (2008). Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: implications for somatic embryogenesis. Proc. Natl. Acad. Sci. USA *105*, 3151–3156.

Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. Proc. Natl. Acad. Sci. USA *98*, 11806–11811.

Su, Y.H., and Zhang, X.S. (2009). Auxin gradients trigger de novo formation of stem cells during somatic embryogenesis. Plant Signal. Behav. *4*, 574–576.

Sugimoto, K., Gordon, S.P., and Meyerowitz, E.M. (2011). Regeneration in plants and animals: dedifferentiation, transdifferentiation, or just differentiation? Trends Cell Biol. *21*, 212–218.

Sun, T., Huang, J., Xu, Y., Verma, V., Jing, B., Sun, Y., Ruiz Orduna, A., Tian, H., Huang, X., Xia, S., et al. (2020). Redundant CAMTA transcription factors negatively regulate the biosynthesis of salicylic acid and N-hydroxypipecolic acid by modulating the expression of SARD1 and CBP60g. Mol. Plant *13*, 144–156.

Suzuki, M., and McCarty, D.R. (2008). Functional symmetry of the B3 network controlling seed development. Curr. Opin. Plant Biol. *11*, 548–553.

Suzuki, M., Wang, H.H., and McCarty, D.R. (2007). Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/abscisic acid INSENSITIVE 3-LIKE B3 genes. Plant Physiol. *143*, 902–911.

Szczygieł-Sommer, A., and Gaj, M.D. (2019). The miR396-GRF regulatory module controls the embryogenic response in Arabidopsis via an auxin-related pathway. Int. J. Mol. Sci. *20*, 5221.

Szemenyei, H., Hannon, M., and Long, J.A. (2008). TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. Science *319*, 1384–1386.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Tan, G., and Lenhard, B. (2016). TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics *32*, 1555–1556.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032–2034.

Thakare, D., Tang, W., Hill, K., and Perry, S.E. (2008). The MADS-domain transcriptional regulator AGAMOUS-LIKE15 promotes somatic embryo development in Arabidopsis and soybean. Plant Physiol. *146*, 1663–1672.

Tsuwamoto, R., Yokoi, S., and Takahata, Y. (2010). Arabidopsis EMBRYOMAKER encoding an AP2 domain transcription factor plays a key role in developmental change from vegetative to embryonic phase. Plant Mol. Biol. *73*, 481–492.

van der Graaff, E., Laux, T., and Rensing, S.A. (2009). The WUS homeobox-containing (WOX) protein family. Genome Biol. *10*, 248.

Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics *33*, 3645–3647.

Waki, T., Hiki, T., Watanabe, R., Hashimoto, T., and Nakajima, K. (2011). The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. Curr. Biol. *21*, 1277–1281.

Wang, J.W., Czech, B., and Weigel, D. (2009a). miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. Cell *138*, 738–749.

Wang, R., and Estelle, M. (2014). Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway. Curr. Opin. Plant Biol. *21*, 51–58.

Wang, X., Niu, Q.W., Teng, C., Li, C., Mu, J., Chua, N.H., and Zuo, J. (2009b). Overexpression of PGA37/MYB118 and MYB115 promotes vegetative-to-embryonic transition in Arabidopsis. Cell Res. *19*, 224–235.

Waris, H. (1957). A striking morphogenetic effect of amino acid in seed plant. Suomen Kemistilehti *30b*, 121.

Weijers, D., and Wagner, D. (2016). Transcriptional responses to the auxin hormone. Annu. Rev. Plant Biol. *67*, 539–574.

Wickramasuriya, A.M., and Dunwell, J.M. (2015). Global scale transcriptome analysis of Arabidopsis embryogenesis in vitro. BMC Genomics *16*, 301.

Winkelmann, T. (2016). Somatic versus zygotic embryogenesis: learning from seeds. Methods Mol. Biol. *1359*, 25–46.

Wójcik, A.M., Wójcikowska, B., and Gaj, M.D. (2020). Current perspectives on the auxin-mediated genetic network that controls the induction of somatic embryogenesis in plants. Int. J. Mol. Sci. *21*, 1333.

Wójcikowska, B., Botor, M., Morończyk, J., Wójcik, A.M., Nodzyński, T., Karcz, J., and Gaj, M.D. (2018). Trichostatin A triggers an embryogenic transition in Arabidopsis explants via an auxin-related pathway. Front. Plant Sci. *9*, 1353.

Wójcikowska, B., Jaskóła, K., Gąsiorek, P., Meus, M., Nowak, K., and Gaj, M.D. (2013). LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of Arabidopsis, via YUCCA-mediated auxin biosynthesis. Planta *238*, 425–440.

Wójcikowska, B., Wójcik, A.M., and Gaj, M.D. (2020). Epigenetic regulation of auxin-induced somatic embryogenesis in plants. Int. J. Mol. Sci. *21*, 2307.

Wu, M.F., Yamaguchi, N., Xiao, J., Bargmann, B., Estelle, M., Sang, Y., and Wagner, D. (2015). Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. eLife *4*, e09269.

Xu, L., and Huang, H. (2014). Genetic and epigenetic controls of plant regeneration. Curr. Top. Dev. Biol. *108*, 1–33.

Yan, C., Fan, M., Yang, M., Zhao, J., Zhang, W., Su, Y., Xiao, L., Deng, H., and Xie, D. (2018). Injury activates Ca^{2+}/calmodulin-dependent phosphorylation of JAV1-JAZ8-WRKY51 complex for jasmonate biosynthesis. Mol. Cell *70*, 136–149.e7.

Yan, F., Powell, D.R., Curtis, D.J., and Wong, N.C. (2020). From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. Genome Biol. *21*, 22.

Yang, C., Bratzel, F., Hohmann, N., Koch, M., Turck, F., and Calonje, M. (2013). VAL- and AtBMI1-mediated H2Aub initiate the switch from embryonic to postgerminative growth in Arabidopsis. Curr. Biol. *23*, 1324–1329.

Yoo, S.D., Cho, Y.H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat. Protoc. *2*, 1565–1572.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS A J. Integr. Biol. *16*, 284–287.

Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382–2383.

Yu, S., Cao, L., Zhou, C.M., Zhang, T.Q., Lian, H., Sun, Y., Wu, J.Q., Huang, J.R., Wang, G.D., and Wang, J.W. (2013). Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. eLife *2*, e00269.

Zhang, T.Q., Lian, H., Zhou, C.M., Xu, L., Jiao, Y., and Wang, J.W. (2017). A two-step model for de novo activation of WUSCHEL during plant shoot regeneration. Plant Cell *29*, 1073–1087.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol *9*, R137.

Zheng, Y., Ren, N., Wang, H., Stromberg, A.J., and Perry, S.E. (2009). Global identification of targets of the Arabidopsis MADS domain protein AGAMOUS-Like15. Plant Cell *21*, 2563–2577.

Zuo, J., Niu, Q.W., Frugis, G., and Chua, N.H. (2002). The WUSCHEL gene promotes vegetative-to-embryonic transition in Arabidopsis. Plant J. *30*, 349–359.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-histone H3K27me3	Merck	Cat#07449; RRID: AB_310624
Anti-Digoxigenin-AP Fab fragments	Roche	Cat#11093274910; RRID: AB_2734716
Bacterial and Virus Strains		
Escherichia coli DH5α	N/A	N/A
Escherichia coli BL21 (DE3)	N/A	N/A
Agrobacterium tumefaciens GV3101 (pMP90)	N/A	N/A
Chemicals, Peptides, and Recombinant Proteins		
Murashige & Skoog basal salts medium	Phyto Technology	Cat#M519
Gamborg B-5 basal medium	Phyto Technology	Cat#G398
Sucrose	ABCONE	Cat#57501
2-(N-morpholino) ethanesulfonic acid (MES)	BBI Life Sciences	Cat#145224948
Phytagel	Sigma-Aldrich	Cat#7101052-1
Agarose G-10	BIOWEST	Cat#EEO015
Kanamycin sulfate	FCNCS	Cat#M211
Rifampicin solution	FCNCS	Cat#M213
Hygromycin B	YEASEN	Cat#60224ES03
Gentamycin sulfate	FCNCS	Cat#M215
Agar Bacteriological Grade	Shanghai Jiafeng	Cat#H8145
KH_2PO_4	Sigma-Aldrich	Cat#7778770
Sodium acetate buffer solution	Sigma-Aldrich	Cat#126965
Dextran Sulfate 50% solution	Millipore	Cat#S4030
Dimethyl sulfoxide (DMSO)	Amresco	Cat#67685
Triton X-100	Sigma-Aldrich	Cat#9002931
Glycerol	Sigma-Aldrich	Cat#56815
Dexamethasone (DEX)	Sigma-Aldrich	Cat#D4902
4% Paraformaldehyde	SolelyBio	Cat#AR0211
EDTA-Na_2	Sigma-Aldrich	Cat#6381926
Glycine	Sigma-Aldrich	Cat#56406
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Sigma-Aldrich	Cat#7365459
Piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES)	Sigma-Aldrich	Cat#5625376
$CH_3COOLi \cdot 2H_2O$	Sigma-Aldrich	Cat#6108174
Na_2HPO_4	Sigma-Aldrich	Cat#7558794
Na_2CO_3	Sigma-Aldrich	Cat#497198
$NaH_2PO_4 \cdot H_2O$	Sigma-Aldrich	Cat#7558807
KCl	Sigma-Aldrich	Cat#7447407
NaCl	Sigma-Aldrich	Cat#7647145
$MgCl_2 \cdot 6H_2O$	Sigma-Aldrich	Cat#7791186
Tris base	Sigma-Aldrich	Cat#77861
DAPI	AAT Bioquest	Cat#28718903
2,4-D	Sigma-Aldrich	Cat#94757

(*Continued on next page*)

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
isopropyl ß-D-1-thiogalactopyranoside	Sangon Biotech	Cat#367931
2-mercaptoethanol	Ruibio	Cat#60242
Spermidine	Sigma-Aldrich	Cat#S2626
Eva Green Dye	Biotium	Cat#31000
Phenylmethylsulfonyl fluoride (PMSF)	Sigma-Aldrich	Cat# 93482
Phusion High-Fidelity DNA Polymerase	Thermo Fisher Scientific	Cat#022021
PEG 4000	BBI Life Sciences	Cat#25322683
Complete Protease Inhibitor Cocktail	Merck	Cat#04693132001
NBT/BCIP Stock Solution	Roche	Cat#11681451001
Proteinase K	Roche	Cat#03115828001
Critical Commercial Assays		
LightShift EMSA Optimization & Control Kit	Thermo Fisher Scientific	Cat#20148X
TB Green Premix Extaq II	Takara	Cat#RR820A
CloneExpress II One Step Cloning Kit	Vazyme	Cat#C11202
pEASY-Uni Seamless Cloning and Assembly Kit	TRANS	Cat#CU10103
NEBNext Ultra II DNA Library Prep Kit for Illumina	New England Biolabs	Cat#E7645S
NEBNext Multiplex Oligos for Illumina	New England Biolabs	Cat#E7335S
RNeasy Plant Mini Kit	Qiagen	Cat#74904
Plasmid Midi Kit	Qiagen	Cat#12145
Mouse monoclonal anti-flag affinity beads	Smart	Cat#SA042001
MinElute Reaction Cleanup Kit	Qiagen	Cat#28204
Dynabeads Protein G for Immunoprecipitation	Invitrogen	Cat#10004D
TruePrep DNA Library Prep Kit v2	Vazyme Biotech	Cat#TD50102
TruePrep Index Kit v2	Vazyme Biotech	Cat#TD202
Dual-Luciferase® Reporter Assay System	Promega	Cat#E1910
TRIzol™ Reagent	Thermo Fisher Scientific	Cat#15596018
DNase I	Thermo Fisher Scientific	Cat#EN0521
T3 RNA Polymerase	Thermo Fisher Scientific	Cat#EP0101
T7 RNA Polymerase	Thermo Fisher Scientific	Cat#EP0111
RevertAid RT Reverse Transcription Kit	Thermo Fisher Scientific	Cat#K1691
LightShift Chemiluminescent EMSA kit	Thermo Fisher Scientific	Cat#20148
2x NEBNext High fidelity PCR mix	New England Biolabs	Cat#M0541L
AMPure beads	Beckman	Cat#A63880
Deposited Data		
ATAC-seq, ChIP-seq and RNA-seq experiment data	This paper	BioProject PRJCA002620, Beijing Institute of Genomics Data Center; http://bigd.big.ac.cn
Experimental Models: Organisms/Strains		
A. thaliana: Col-0	N/A	N/A
A. thaliana: *wox1*	Arabidopsis Biological Resource Center	SALK_148070
A. thaliana: *wox2*	Arabidopsis Biological Resource Center	SAIL_254_A01
A. thaliana: *wox3*	Arabidopsis Biological Resource Center	SALK_108644
A. thaliana: *wox8*	Arabidopsis Biological Resource Center	SALK_014799
A. thaliana: *wox1 wox2 wox3*	This paper	SALK_148070/ SAIL_254_A01/ SALK_108644
A. thaliana: *wox2 wox3*	This paper	SAIL_254_A01/ SALK_108644

(*Continued on next page*)

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
A. thaliana: *lec2*	Dr. Xingliang Hou (South China Botanical Garden, CAS)	SALK_015228
A. thaliana: *p35S::LEC2-GR*	This paper	N/A
A. thaliana: *p35S::LEC2-GR-3xFLAG*	This paper	N/A
A. thaliana: *pUBQ1::LEC2-GR*	This paper	N/A
A. thaliana: *pWOX2::3xVENUS-N7*	This paper	N/A
A. thaliana: *pRibo::WOX2-GR*	This paper	N/A
Oligonucleotides		
See Table S7	N/A	N/A
Recombinant DNA		
p35S::LEC2-GR	This paper	FX019
p35S::LEC2-GR-3xFLAG	This paper	FX395
pUBQ1::LEC2-GR	This paper	FX492
pWOX2::3xVENUS-N7	This paper	FX059
pRibo::WOX2-GR	This paper	FX018
p35S::LEC2	This paper	FX159
pWOX2::LUC RY12	This paper	FX474
pWOX2::LUC mRY1	This paper	FX475
pWOX2::LUC mRY2	This paper	FX476
pWOX2::LUC mRY12	This paper	FX477
AA00-LUC p35S::REN	(Zhang et al., 2017)	TQ379
AA00-35S	This paper	JW807
AA00-hyg	This paper	LZ010
AA00::3xVENUS-N7	This paper	LH297
pRibo::GR	This paper	LZ234
Software and Algorithms		
ImageJ	https://imagej.nih.gov/ij/	RRID: SCR_003070; https://imagej.nih.gov/ij/
R version 3.6	The R Foundation	RRID: SCR_001905; https://www.r-project.orgl
GraphPad Prism 8	GraphPad Software	RRID: SCR_002798; http://www.graphpad.com/
Adobe Photoshop CC 2018	Adobe Acrobat	N/A
Adobe Illustrator CC 2018	Adobe Acrobat	N/A
Fastp	(Chen et al., 2018)	RRID: SCR_016962; https://github.com/OpenGene/fastp
FastQC v0.11.7	FastQC	RRID:SCR_014583; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
MulitQC v1.6	(Ewels et al., 2016)	https://multiqc.info/
Bowtie2 v2.3.4.3	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/
hisat2 v2.1.0	(Kim et al., 2015)	RRID: SCR_015530; http://ccb.jhu.edu/software/hisat2/index.shtml
Samtools v1.9	(Li et al., 2009)	RRID: SCR_002105; http://www.htslib.org/
sambamba v0.6.7	(Tarasov et al., 2015)	https://lomereiter.github.io/sambamba/
bedtools v2.25.0	(Quinlan and Hall, 2010)	RRID: SCR_006646; https://bedtools.readthedocs.io/en/latest/
deepTools v3.1.2	(Ramírez et al., 2014)	RRID: SCR_016366; https://deeptools.readthedocs.io/en/develop/
featureCounts v1.6.2	(Liao et al., 2014)	RRID: SCR_012919; http://bioinf.wehi.edu.au/featureCounts/

(*Continued on next page*)

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DESeq2 v1.26.0	(Love et al., 2014)	RRID: SCR_015687; https://bioconductor.org/packages/release/bioc/html/DESeq2.html
ashr 2.2-39	(Stephens, 2017)	https://cran.r-project.org/web/packages/ashr/index.html
clusterProfiler v3.14.0	(Yu et al., 2012)	RRID: SCR_016884; https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
complexHeatmap	(Gu et al., 2016)	RRID: SCR_017270; https://jokergoo.github.io/ComplexHeatmap-reference/book/
MACS2 v2.1.2	(Zhang et al., 2008)	https://github.com/taoliu/MACS
DiffBind v2.14.0	(Ross-Innes et al., 2012)	RRID: SCR_012918; https://bioconductor.org/packages/release/bioc/html/DiffBind.html
ChIPseeker v1.22.0	(Yu et al., 2015)	https://bioconductor.org/packages/release/bioc/html/ChIPseeker.html
HOMER v4.10	(Heinz et al., 2010)	RRID: SCR_010881; http://homer.ucsd.edu/homer/
Intervene v0.6.1	(Khan and Mathelier, 2017)	https://intervene.readthedocs.io/en/latest/
Integrative Genomics Viewer	(Robinson et al., 2011)	RRID: SCR_011793; http://software.broadinstitute.org/software/igv/
ggseqlogo	(Wagih, 2017)	https://github.com/omarwagih/ggseqlogo
ChIPQC	(Carroll et al., 2014)	http://bioconductor.org/packages/release/bioc/html/ChIPQC.html
chromVAR v1.8.0	(Schep et al., 2017)	https://bioconductor.org/packages/release/bioc/html/chromVAR.html
JASPAR database	http://jaspar.genereg.net/	N/A
TFBStools v1.24.0	(Tan and Lenhard, 2016)	https://bioconductor.org/packages/release/bioc/html/TFBSTools.html
Other		
Eppendorf realplex2	Eppendorf	Cat#A248709R
Olympus FV3000	Olympus	N/A
GST-affinity column	GE healthcare	Cat#17075604
Falcon Cell Strainers	Corning	Cat#352340
Cell imaging dish	NEST	Cat#MH0031
Lecia Sliding Microtome 1200S	Lecia	N/A
The BD FACSAria™ II	BD	N/A
GloMax® 20/20 Luminometer	Promega	Cat#E5311
Percival chamber	Percival	N/A
Olympus BX3-CBH	Olympus	N/A
Lecia ASP200S	Lecia	N/A

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Jia-Wei Wang (jwwang@sippe.ac.cn).

Materials Availability

Plasmids and transgenic plant lines generated in this study will be made available on request to the lead contact. This study did not generate new unique reagents.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

Data and Code Availability

The ATAC-seq, ChIP-seq and RNA-seq data (BioProject PRJCA002620) were deposited in Beijing Institute of Genomics Data Center (http://bigd.big.ac.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The *Arabidopsis thaliana* plants were used as the experimental model in the study. The Columbia-0 (Col-0) ecotype was used as wild type. The mutants and transgenic lines were generated in this background as detailed in the Key Resources Table.

METHOD DETAILS

Plant Growth Conditions

Wild type and mutants were grown in the growth chambers at 21°C under a 16-h light/8-h dark condition. For SE, the explants were grown in the Percival chamber at 21°C under a 16-h light/8-h dark condition. The light intensity is 45 μmol/m^2/s (white fluorescent light).

Constructs

For *p35S::LEC2-GR* (FX019), *LEC2* was fused with GR coding sequence by fusion PCR method. The resultant *LEC2-GR* fragment was then inserted into JW819 behind the *35S* promoter. For *p35S::LEC2-GR-3xFLAG* (FX395), *3xFLAG* was fused to *LEC2-GR* by fusion PCR. The *LEC2-GR-3xFLAG* fragment was then replaced the *LEC2-GR* fragment in FX019 by homologous recombination using ClonExpress II one Step Cloning kit (Vazyme Biotech, Cat No./ID: C112).

For *pUBQ1::LEC2-GR* (FX492), *LEC2-GR* was cloned into LZ010 by homologous recombination. *pWOX2::3xVENUS-N7* (FX059) was constructed by inserting the regulatory sequence of *WOX2* (4.0 kb upstream and 1.4 kb downstream of coding region) in front of *3xVENUS-N7* coding gene in LH297. For *pRibo::WOX2-GR* (FX018), *WOX2-GR* was cloned into LZ234.

For transient protoplast analysis, the 2.0 kb wild-type or mutated promoter of *WOX2* (mRY1, mRY2 and mRY12) was cloned in front of *LUCIFERASE* (*LUC*) in TQ379, which harbors the *p35S::RENILLA* (*REN*) cassette. The resultant plasmids were named as FX474, FX475, FX476 and FX477, respectively. *p35S::LEC2* (FX159) was generated by cloning the cDNA of *LEC2* into JW807 behind the *35S* promoter.

The oligonucleotide primers for all constructs are given in Table S7.

Generation of Transgenic Plants

For transgenic Arabidopsis plants, the binary constructs were delivered into *Agrobacterium tumefaciens* GV3101 (pMP90) by the freeze-thaw method. Transgenic plants were generated by the floral dipping method (Clough and Bent, 1998) and screened with 0.05% glufosinate (Basta) on soil, 40 mg/mL hygromycin, or 50 mg/mL kanamycin on half-strength MS plates.

SE

For direct SE, immature embryos were manually dissected and cultured on E5 media (B5 media supplemented with 5 μM 2,4-D) at 21°C under a 16-h light/8-h dark condition for 15 days. The explants were then transferred to the MS media supplemented with 20 g sucrose/L for another 10 days (Gaj, 2011). The SE rate was represented by the percentage of explants with regenerated embryos in a given number of explants at 10 DAM (days after transfer to MS media). The regenerative capacity was represented by the regenerated somatic embryos per explant at 10 DAM. For each experiment, 60-70 explants were used. For DEX induction experiment, 10 μM DEX was added into E5 media. Equal volume of DMSO was used as mock. For indirect SE (embryonic calli route) (Su and Zhang, 2009), immature zygotic embryos were manually harvested and used as explants to induce primary somatic embryos (PSEs). To increase the number of PSEs, casein hydrolysate (500 mg/L) was added to agar-solidified B5 media. The PSEs were then transferred to liquid B5 media containing 9.0 μM 2,4-D and pre-cultured for 14 days. The resultant embryonic calli were cultured in auxin-free liquid B5 media to induce SEs.

Plant Materials Used for ATAC-seq, ChIP-seq and RNA-seq

For each biological replicate of RNA-seq and ATAC-seq, approximately 200 immature embryo explants cultured on E5 media were collected at 0, 4, 8, 16, 24, 48 and 72 HAI. The wild-type (Col-0) seeds were geminated on B5 media. After three days, the seedlings were split into two halves: one was directly harvested for RNA-seq and ATAC-seq (G3); the other was transferred to E5 media and harvested after 1 day (G3E1) or 3 days (G3E3). The seed coats of all the seedlings were manually removed. For RNA-seq of *p35S::LEC2-GR*, *p35S::LEC2-GR* seeds were germinated in B5 liquid media and cultured for 3 days. The seedlings were harvested after 4 h treatment of 10 μM DEX or DMSO and 10 μM translation inhibitor cycloheximide. For ChIP-seq of *p35S::LEC2-GR-3xFLAG*, *p35S::LEC2-GR-3xFLAG* seeds were germinated in B5 liquid media and cultured for 3 days. The seedlings were collected after 4 h treatment of 30 μM DEX or DMSO.

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

ATAC-seq Experiment

For each biological replicate, the collected plant tissue was cut into small pieces with blade in 500 μL lysis buffer (15 mM Tris-HCl pH 7.5, 20 mM NaCl, 80 mM KCl, 0.5 mM spermidine, 5 mM 2-mercaptoethanol and 0.2% Triton X-100). The slurry was filtered with 40 μm filter into collection tube (BD biosciences, Cat No./ID: 352340) and resuspended in 2.0 mL lysis buffer. The crude nuclei were stained with 4,6-diamidino-2-phenylindole and loaded into a flow cytometer (BD biosciences, FACSAria III). The nuclei were sorted into 500 μL lysis buffer (50,000 nuclei per reaction), and centrifuged at 1,000 g for 10 min at 4 °C, and washed once with Tris-Mg buffer (10 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$). The purified nuclei were then incubated with Tn5 transposome and tagmentation buffer at 37 °C for 30 min (Vazyme Biotech, Cat No./ID: 501-02). After the tagmentation, the DNA was purified using a Qiagen MinElute PCR Purification Kit (Qiagen, Cat No./ID: 28004) and then amplified using 2x NEBNext High fidelity PCR mix (New England Biolabs, Cat No./ID: M0541L) for 8-12 cycles. PCR cycle number was determined as described previously (Buenrostro et al., 2013). Amplified libraries were purified with AMPure beads (Beckman Coulter, Cat No./ID: A63880). Two or three biological replicates were performed.

RNA-seq, ChIP-PCR and ChIP-seq Experiments

For RNA-seq, total RNAs were extracted using RNeasy Plant Mini Kit (Qiagen, Cat No./ID: 74904). Library construction and deep sequencing were performed using the Illumina HiSeq 4000 Platform according to manufacturer's protocols (Novogene, Beijing, China). Two or three biological replicates were performed.

For ChIP-seq or ChIP-qPCR, seeds were germinated and cultured in B5 liquid media with shaking (140 rpm) for 3 days. The seedlings were harvested and fixed according to our published protocol (Yu et al., 2013). Briefly, the samples were fixed twice in MC-buffer (10 mM potassium phosphate, pH 7.0, 50 mM NaCl, and 0.1 M sucrose) with 1% formaldehyde and 1 mM phenylmethylsulfonyl fluoride (PMSF) using vacuum infiltration (10 min each at room temperature). After quenching by 100 mM glycine, the samples were extensively washed with H_2O and ground into fine powder by liquid nitrogen. The powder was resuspended with lysis buffer, and the nuclei were extracted by sucrose density gradient centrifugation. The nuclei were sheared and the resultant chromatin extract was immunoprecipitated with anti-FLAG beads (Smart, Cat No./ID: SA042001/SA042005). ChIP DNAs were reverse crosslinked and purified with PCR purification kit (Qiagen, Cat No./ID: 28206). 1 μL DNA was used for real-time quantitative PCR (RT-qPCR) analyses. The relative enrichment of LEC2-GR-3xFLAG at specific locus was calculated by normalizing the amount of each immunoprecipitated fragment to input DNA, and then by normalizing the value for DEX samples against the value for DMSO samples.

For H3K27me3 ChIP-seq experiment, about 1,000 embryos at late bent cotyledon stage or 0.2 g 3-day-old seedlings were collected for each biological replicate. The samples were fixed as described above. The chromatin extract was immunoprecipitated with anti-H3K27me3 antibody (Merck, Cat No./ID:07449) overnight. The immunoprecipitated DNA was incubate with protein G Dynabeads (Thermo, Cat No./ID: 10004D). ChIP DNAs were reverse crosslinked and purified with PCR purification kit (Qiagen, Cat No./ID: 28206).

For ChIP-seq library, 5 ng input or IP DNA was used. End repair, adaptor ligation, PCR amplification and purification were performed according to the manual of NEBNext Ultra II DNA library prep kit (NEB, Cat No./ID: E7645). Two biological replicates were performed.

Transient Protoplast Analysis

For the dual-LUC reporter assay, Arabidopsis protoplasts were prepared from the leaves of 3-week-old wild type plants according to a published protocol (Yoo et al., 2007). Briefly, the fully expanded leaves were cut into leaf strips with sharp razor blade. The leaf strips were cultured in enzyme solution for 2 h at room temperature with shaking. The protoplasts were then collected, washed and transfected with plasmids. After transfection, the protoplasts were cultured at 22 °C for ~13 h. The protoplasts were lysed with Passive Lysis Buffer (Promega, Cat No./ID: E1910). LUC and REN activities were quantified and measured with a luminometer (Promega 20/20). LUC activity was calculated by normalizing to that of REN.

Expression Analyses

Total RNA was extracted with Trizol reagent (ThermoFisher, Cat No./ID: 15596018). Total RNA (1 μg) was treated with DNase I (1 unit/mL; ThermoFisher, Cat No./ID: EN0521) and used for cDNA synthesis with oligo (dT) primer by RevertAid RT Reverse Transcription Kit (ThermoFisher, Cat No./ID: K1691). The average expression levels and SD values were calculated from $2^{-\Delta\Delta Ct}$ values. *UBQ10* was used for normalization. The oligonucleotide primers for all genes are given in Table S7.

EMSA

To construct the plasmid for the expression of recombinant LEC2 protein in *Escherichia coli*, the DNA binding domain (DBD, 169-273 amino acids) of LEC2 was cloned into the pGEX vector (Novagen). The resultant construct was then transformed into *E. coli* BL21 (DE3). To induce GST-LEC2-DBD expression, the cells were grown at 37°C to OD_{600} = 0.6-0.8 and induced with 0.3 mM isopropyl ß-D-1-thiogalactopyranoside for 8-10 h at 16 °C. The cells were harvested by centrifugation for 15 min at 4,000g, resuspended in the GST-lysis buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl) and lysed with a homogenizer. The lysate was centrifuged for 1 h at 17,000g, and the supernatant was passed over a GST-affinity column (GE healthcare, Cat No./ID:17-0756-04). The GST-LEC2-DBD was recovered by the GST-elution buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 20 mM glutathione). The protein was concentrated

Please cite this article in press as: Wang et al., Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis, Developmental Cell (2020), https://doi.org/10.1016/j.devcel.2020.07.003

to 2 mg/mL. Double-stranded wild-type or mutated oligonucleotide probes were synthesized and labeled with biotin at the 5' end of positive strand. EMSA was performed using LightShift Chemiluminescent EMSA kit (ThermoFisher, Cat No./ID: 20148). Briefly, biotin-labeled probes were incubated in 1x binding buffer, 2.5% glycerol, 5 mM $MgCl_2$, 0.05% NP-40, and 50 ng/μL Poly (dI·dC) with or without purified LEC2-DBD proteins at room temperature for 20 min. For unlabeled probe competition, unlabeled probes were added to the reactions.

RNA *In situ* Hybridization

RNA *in situ* hybridization was performed as described (Wang et al., 2009; Zhang et al., 2017). Briefly, immature embryos or regenerated samples were fixed with formaldehyde. Paraffin-embedded samples were sectioned (7-9 μm) with a Lecia sliding microtome. The slides were dewaxed, digested with Proteinase K (Roche, Cat No./ID: 03115828001), dehydrated with gradient ethanol, hybridized with corresponding probes and incubated with anti-digoxigenin-AP Fab fragments (Roche, Cat No./ID: 11093274910). After washing, the signals were detected with NBT/BCIP stock solution (Roche, Cat No./ID: 11681451001). The cDNA fragments of *WOX2*, *WOX3*, *LEC1* and *LEC2* were PCR amplified and cloned into pGEM T-vector (Promega, Cat No./ID: A3600), respectively. *In vitro* transcription was performed with T3 or T7 RNA polymerase (ThermoFisher, Cat No./ID: EP0101/EP0111), in which linearized vectors were used as templates.

Tissue Embedding and Sectioning

Samples were collected and immediately placed in the vials with ice-cold phosphate buffered saline (PBS) containing 2.5% paraformaldehyde (PFA, pH 7.0). The samples were infiltrated for 30 mins by vacuum, and stored at 4 °C overnight. Tissues were then washed with sucrose gradient PBS-PFA solution, embedded with 6% low melting agarose, and sliced with a Lecia Sliding Microtome 1200S at the thickness of 50 to 70 μm.

Microscopy

For RNA *in situ* imaging, slides were mounted with water and observed under an Olympus BX63 microscope equipped with DP73 digital camera and differential interference contrast (DIC) modules. For confocal imaging, we used OLYMPUS FV3000 confocal microscopic system (Olympus, Japan). The explants were placed in a 35 mm cell imaging dish (NEST, Cat No./ID: MH0031). Proper filter sets and lasers were selected for fluorescence signal scanning. VENUS was stimulated with an argon laser at 514 nm, with emission filtered at 530 to 600 nm. Autofluorescence were used to delineate the tissue, with emission at 650 to 750 nm. The images were processed with ImageJ.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data Preprocessing and Reads Alignment

The libraries were sequenced on Illumina Hiseq-PE150. For each library, raw.fastq was trimmed by fastp v0.20.0 with default parameters (for ATAC-seq, set adapter sequence with parameter "-a CTGTCTCTTATACACATCT")(Chen et al., 2018). After trimming, FastQC v0.11.7 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MulitQC v1.6 were performed as quality control to obtain clean fq files (Ewels et al., 2016).

Reads were aligned using either Bowtie2 v2.3.4.3 (for ATAC-seq and ChIP-seq) or hisat2 v2.1.0 (for RNA-seq) to the Arabidopsis genome (TAIR10) (Kim et al., 2015; Langmead and Salzberg, 2012). The resulting SAM file containing mapped reads were converted to BAM format, sorted, and indexed using Samtools v1.9 (Li et al., 2009). The biological replicates were merged by Samtools v1.9. The sorted BAM were processed to remove duplicated and organellar reads by sambamba v0.6.7 and bedtools v2.25.0 (for ATAC-seq and ChIP-seq) (Quinlan and Hall, 2010; Tarasov et al., 2015). To normalize and visualize the individual and merged replicate datasets, the BAM files were converted to bigwig using bamCoverage provided by deepTools v3.1.2 with a bin size of 10 bp and normalized by Reads Per Kilobase per Million mapped reads (RPKM, for RNA-seq) or Bin Per Million mapped reads (BPM, for ATAC-seq and ChIP-seq) (Ramírez et al., 2014).

RNA-seq Data Analyses

The number of paired mapping reads that overlap each annotated gene (Araport11) was counted using featureCounts v1.6.2 with the parameter "-p" (Cheng et al., 2017; Liao et al., 2014). The counts files were then used as inputs for differential gene expression analysis by DESeq2 v3.10 (Love et al., 2014). The Log fold change shrinkage was estimated by the ashr package (Stephens, 2017). The threshold of differential gene expression is "p.adj < 0.05 and abs(log2FoldChange) >1". To analyze all the samples over the course of time at once, we used Likelihood Ratio Test (LRT) function provided by DESeq2. The PCA plot was generated by "PlotPCA function" provided by DESeq2.

For subsequent clustering and visualization, we obtained mean counts by merging two biological replicates. To normalize the depth effect of library, we used CPM (counts per million) normalization methods. After removing the genes, which showed small changes over the course of time, with low stringent threshold "p.adj >= 0.2", we scaled RNA-seq CPM normalized expression matrix into the Z-scale matrix and used k-means function in R to cluster genes with following parameters "set.seed(19960203), centers = 7, iter.max = 50". The GO and Gene Set Enrichment Analysis of differentially expressed genes or the clusters sorted by k-means approach were performed using clusterProfiler v3.14.0 and org.At.tair.db v3.10.0 (Yu et al., 2012) (https://bioconductor.org/

packages/release/data/annotation/html/org.At.tair.db.html). We generated the plots with following packages: ggplot2 (https://ggplot2.tidyverse.org/), tidyr (https://tidyr.tidyverse.org/), dplyr (https://dplyr.tidyverse.org), BuenColors (https://github.com/caleblareau/BuenColors) and complexHeatmap (Gu et al., 2016).

ChIP-seq Data Analyses

MACS2 v2.1.2 was used to call peaks with default parameters "-t IP -c input -f BAMPE -g 1.1e8" (Zhang et al., 2008). For H3K27me3 ChIP-seq datasets, we turned the "broad" flag on. To identify the differential binding regions, DiffBind v2.14.0 was applied with following parameters "minOverlap = 2, method = DBA_DESEQ2" (Ross-Innes et al., 2012). For H3K27me3 ChIP-seq datasets, we turned on the "summits = 250" in the step of "dba.count" to resize peaks.

The peaks called by MACS2 and differential peaks identified by DiffBind v2.14.0 were annotated by TxDb.Athaliana.BioMart.plantsmart28 and ChIPseeker v1.22.0 package with "annotatePeak" function (https://bioconductor.org/packages/release/data/annotation/html/TxDb.Athaliana.BioMart.plantsmart28.html) (Yu et al., 2015). The gene promoters are defined as ± 3.0 kb from TSS. The differential regions were scanned for the enrichment of motifs using "findMotifsGenome.pl" function provided by HOMER v4.10 with default parameters (Heinz et al., 2010).

The GO analysis for ChIP-seq datasets was performed according to the methods used for RNA-seq analysis. In addition to using the plot packages as for RNA-seq analyses, we used "dba.plotHeatmap" function in DiffBind v2.14.0 to check the correlation of biological replicates; "PlotProfile" and "PlotHeatmap" in deeptools to check the quality of samples and score associated genomic regions; Intervene to intersect and visualize the intersection results (Khan and Mathelier, 2017); Integrative Genomics Viewer to visualize the signals (Robinson et al., 2011) and ggseqlogo for TF motif graph (Wagih, 2017).

ATAC-seq Data Analyses

ATAC-seq data analyses were performed according to published methods with some modifications (Ludwig et al., 2019; Yan et al., 2020). Briefly, "PlotPCA" in DESeq2 was applied to plot the correlation of biological replicates. "PlotProfile" in deeptools and "ChIPQC" in ChIPQC were used for quality control (Carroll et al., 2014). For differentially accessible region analyses, we used "DiffBind" with the parameter "minOverlap = 1". "DiffBind" was also used to calculate merged peak locations, based on the outer boundaries of overlapping peaks from all the samples analyzed. The peak accessibility scores for each biological replicate were calculated by "DiffBind" with the setting "DBA_SCORE_TMM_READS_EFFECTIVE_CPM" or "DBA_SCORE_READS". The peak annotation, Motif and GO analyses were performed according to the methods used for ChIP-seq data analyses. The location of annotated merge peaks was defined by "ChIPSeeker" with the parameter "tssRegion = c(-2000,0)". Subsequently, we divided the peaks into four sets on the basis of peak location: P_U (TSS-proximal, < 2.0 kb from TSS), P_D (TTS-proximal, < 1.0 kb from TTS), GB (gene body) and D_I (distal intergenic).

The raw peak counts (DBA_SCORE_READS) were used as inputs to identify differentially accessible peaks over the course of time by DESeq2 LRT. We defined the peaks with changes in accessibility by the threshold "p.adj < 0.05". To cluster differentially accessible peaks, we used the peaks in the P_U dataset with the setting "p.adj < 0.05". The peak accessibility scores (DBA_SCORE_TMM_READS_EFFECTIVE_CPM) were Z-scaled and sorted by k-means with the parameters "set.seed(19960203), centers = 9, iter.max = 50". ClusterProfile and HOMER were then applied to call GO and Motif information of corresponding clusters. We also used the peak accessibility scores as the inputs of chromVAR suite to calculate TF activity scores, synergy, and correlation (Schep et al., 2017). Since chromVAR was designed for single-cell ATAC-seq, we used "set.seed(19960203)" as the parameter to obtain relatively stable results for synergy and correlation analyses. The TF motif defined in chromVAR was downloaded from the JASPAR2018 package (https://bioconductor.org/packages/release/data/annotation/html/JASPAR2018.html). Because of the absence of WOX, BBM and LEC1 TF motifs in JASPAR2018, we imported the corresponding motifs from plantTFDB by TFBStools (Jin et al., 2017; Tan and Lenhard, 2016). In addition to the packages described above, we used corrplot (https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html) to plot the relationship of synergy and correlation between TF motifs.

To test the correlation between gene expression level and accessibility of associated peak, we generated the peak-gene associations. GenomicFeatures and GenomicRanges packages were first used to obtain gene flanking regions (3, 6, 9, 12 and 15 kb from TSS) and identify associated ATAC-seq peaks (Lawrence et al., 2013). The permutation of peak-gene associations were then produced by sample function in R (all permutation seed is 19960203). Pearson correlation method was subsequently applied to calculate the correlation score. DBA_SCORE_TMM_READS_EFFECTIVE_CPM and CPM were used as peak accessibility scores and gene expression values, respectively.

We applied ggalluvial (https://corybrunson.github.io/ggalluvial/) package to produce the sankey plot with gene as flow unit (Figure 3E). Briefly, we annotated ATAC-seq or H3K27me3 peaks to nearest genes and intersect them according to following definitions: p.adj < 0.05 and log2FoldChange > 1 as increased peak/gene, p.adj < 0.05 and log2FoldChange < -1 as decreased peak/gene, p.adj > 0.05 or abs(log2FoldChang2) <=1 as non-significant peak/gene, mean count < 1 as ‘‘not-expressed’’ gene, and no H3K27me3 peaks as the genes without H3K27me3 modification.

Developmental Cell, Volume *54*

Supplemental Information

Chromatin Accessibility Dynamics and a Hierarchical Transcriptional Regulatory Network Structure for Plant Somatic Embryogenesis

Fu-Xiang Wang, Guan-Dong Shang, Lian-Yu Wu, Zhou-Geng Xu, Xin-Yan Zhao, and Jia-Wei Wang

Figure S1. Regeneration assays using immature wild type embryo as explant. related to Figure 1.

Open and solid arrows indicate embryo-like protuberance on the explants cultured on E5 and somatic embryos on the explants after transfer to MS media, respectively. c, cotyledon; h, hypocotyl. Scale bars represent 500 μm.

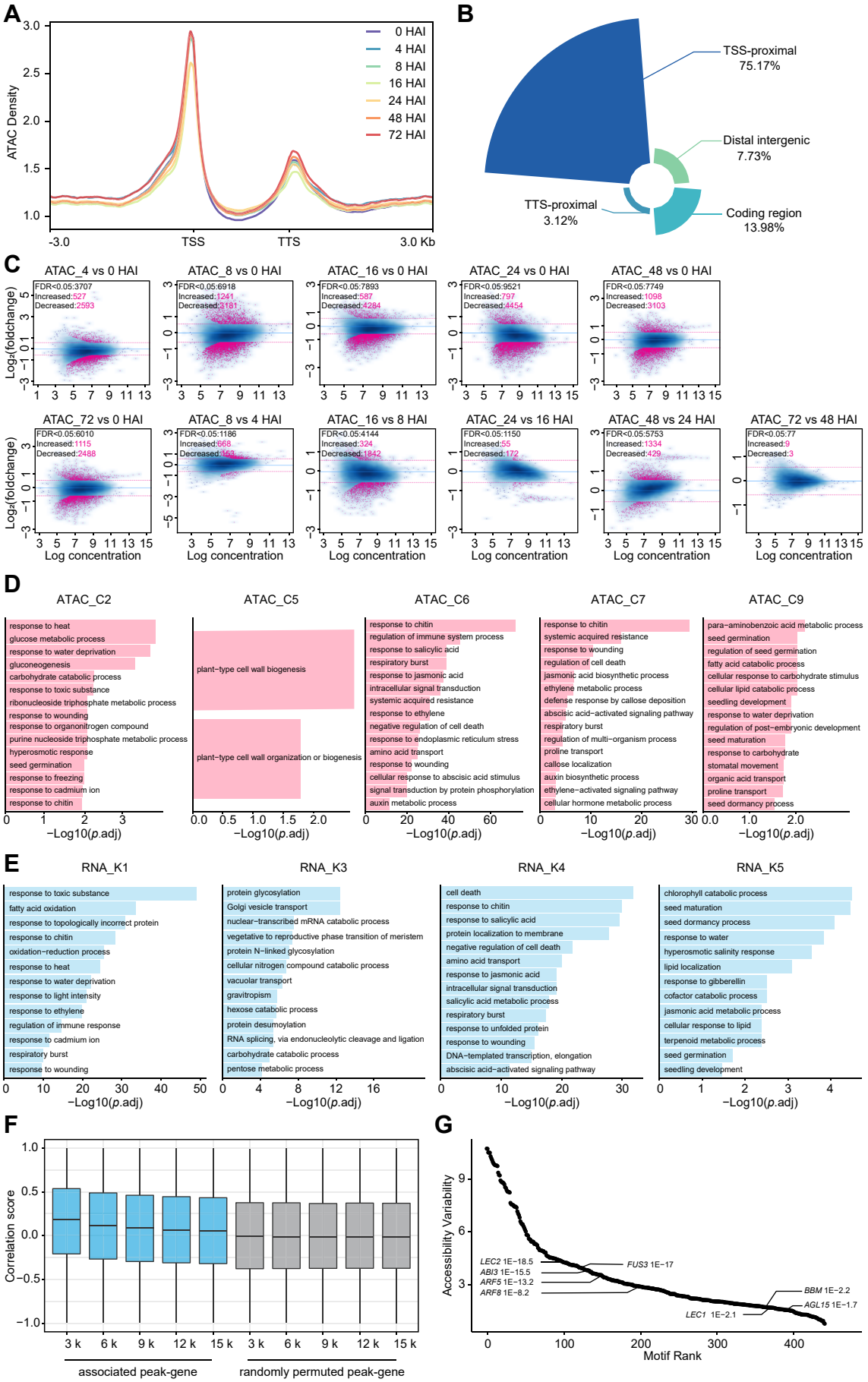

A
ATAC Density
0 HAI
4 HAI
8 HAI
16 HAI
24 HAI
48 HAI
72 HAI
-3.0
TSS
TTS
3.0 Kb
B
TSS-proximal
75.17%
Distal intergenic
7.73%
TTS-proximal
3.12%
Coding region
13.98%
C
ATAC_4 vs 0 HAI
ATAC_8 vs 0 HAI
ATAC_16 vs 0 HAI
ATAC_24 vs 0 HAI
ATAC_48 vs 0 HAI
ATAC_72 vs 0 HAI
ATAC_8 vs 4 HAI
ATAC_16 vs 8 HAI
ATAC_24 vs 16 HAI
ATAC_48 vs 24 HAI
ATAC_72 vs 48 HAI
Log2(foldchange)
Log concentration
D
ATAC_C2
response to heat
glucose metabolic process
response to water deprivation
gluconeogenesis
carbohydrate catabolic process
response to toxic substance
ribonucleoside triphosphate metabolic process
response to wounding
response to organonitrogen compound
purine nucleoside triphosphate metabolic process
hyperosmotic response
seed germination
response to freezing
response to cadmium ion
response to chitin
ATAC_C5
plant−type cell wall biogenesis
plant−type cell wall organization or biogenesis
ATAC_C6
response to chitin
regulation of immune system process
response to salicylic acid
respiratory burst
response to jasmonic acid
intracellular signal transduction
systemic acquired resistance
response to ethylene
negative regulation of cell death
response to endoplasmic reticulum stress
amino acid transport
response to wounding
cellular response to abscisic acid stimulus
signal transduction by protein phosphorylation
auxin metabolic process
ATAC_C7
response to chitin
systemic acquired resistance
response to wounding
regulation of cell death
jasmonic acid biosynthetic process
ethylene metabolic process
defense response by callose deposition
abscisic acid−activated signaling pathway
respiratory burst
regulation of multi−organism process
proline transport
callose localization
auxin biosynthetic process
ethylene−activated signaling pathway
cellular hormone metabolic process
ATAC_C9
para−aminobenzoic acid metabolic process
seed germination
regulation of seed germination
fatty acid catabolic process
cellular response to carbohydrate stimulus
cellular lipid catabolic process
seedling development
response to water deprivation
regulation of post−embryonic development
seed maturation
response to carbohydrate
stomatal movement
organic acid transport
proline transport
seed dormancy process
−Log10(p.adj)
E
RNA_K1
response to toxic substance
fatty acid oxidation
response to topologically incorrect protein
response to chitin
oxidation−reduction process
response to heat
response to water deprivation
response to light intensity
response to ethylene
regulation of immune response
response to cadmium ion
respiratory burst
response to wounding
RNA_K3
protein glycosylation
Golgi vesicle transport
nuclear−transcribed mRNA catabolic process
vegetative to reproductive phase transition of meristem
protein N−linked glycosylation
cellular nitrogen compound catabolic process
vacuolar transport
gravitropism
hexose catabolic process
protein desumoylation
RNA splicing, via endonucleolytic cleavage and ligation
carbohydrate catabolic process
pentose metabolic process
RNA_K4
cell death
response to chitin
response to salicylic acid
protein localization to membrane
negative regulation of cell death
amino acid transport
response to jasmonic acid
intracellular signal transduction
salicylic acid metabolic process
respiratory burst
response to unfolded protein
response to wounding
DNA−templated transcription, elongation
abscisic acid−activated signaling pathway
RNA_K5
chlorophyll catabolic process
seed maturation
seed dormancy process
response to water
hyperosmotic salinity response
lipid localization
response to gibberellin
cofactor catabolic process
jasmonic acid metabolic process
cellular response to lipid
terpenoid metabolic process
seed germination
seedling development
−Log10(p.adj)
F
Correlation score
associated peak-gene
randomly permuted peak-gene
G
Accessibility Variability
LEC2 1E−18.5
ABI3 1E−15.5
ARF5 1E−13.2
ARF8 1E−8.2
FUS3 1E−17
BBM 1E−2.2
AGL15 1E−1.7
LEC1 1E−2.1
Motif Rank

Figure S2. Integrative analysis of ATAC-seq and RNA-seq data. related to Figure 1 and Figure 2.

(A) The genome-wide distribution of ATAC-seq peaks. Window size: gene body ± 3.0 kb.

(B) The location of ATAC-seq peaks are classified into four groups: peaks within 2.0 kb upstream of TSS (TSS-proximal), peaks within 1.0 kb downstream of TTS (TTS-proximal), peaks within the coding region, and other intergenic peaks (distal intergenic).

(C) MA plots showing fold-change of differentially accessible peaks. Blue, distribution of constitutive peaks; pink dots, individual differential peaks. The numbers of differentially accessible peaks (opening or closing) according to false discovery rates (FDR) alone or FDR with Fold change cut off are indicated at the lower right corner of each plot. FDR cut off is < 0.05 while Fold change cut off is >1.5 or <-1.5.

(D) GO term analyses showing distinct gene ontologies of target genes linked to differentially accessible peaks. The selected 15 enriched GO biological processes are indicated. The $-Log_{10}(p.adj)$ is given.

(E) GO term analyses showing distinct gene ontologies of differentially expressed genes revealed by RNA-seq. The selected 15 enriched GO biological processes are indicated. The $-Log_{10}(p.adj)$ is given.

(F) The correlation between gene expression and nearby peak accessibility. The mean gene expression scores for each gene and mean peak accessibility scores for associated peaks at different distance (±3.0 to 15.0 kb from TSS) were calculated. The values were used to compute Pearson correlation scores between each gene and associated peaks. The distribution of observed correlations between gene expression and associated peak accessibility are shown in blue, compared to randomly permuted associations between genes and peaks, shown in gray. Boxplots show medians (middle bar), quartiles (boxes), and 1.5 × interquartile ranges (whiskers). Kolmogorov–Smirnov tests between any observed correlation and matching randomly-permuted correlation set all had p.adj < 1e-07, while any comparison between two randomly-permuted correlation sets had a p.adj > 0.01.

(G) Rank order plot of about 400 TF binding motifs with greatest variability in chromatin accessibility across samples at the different time points.

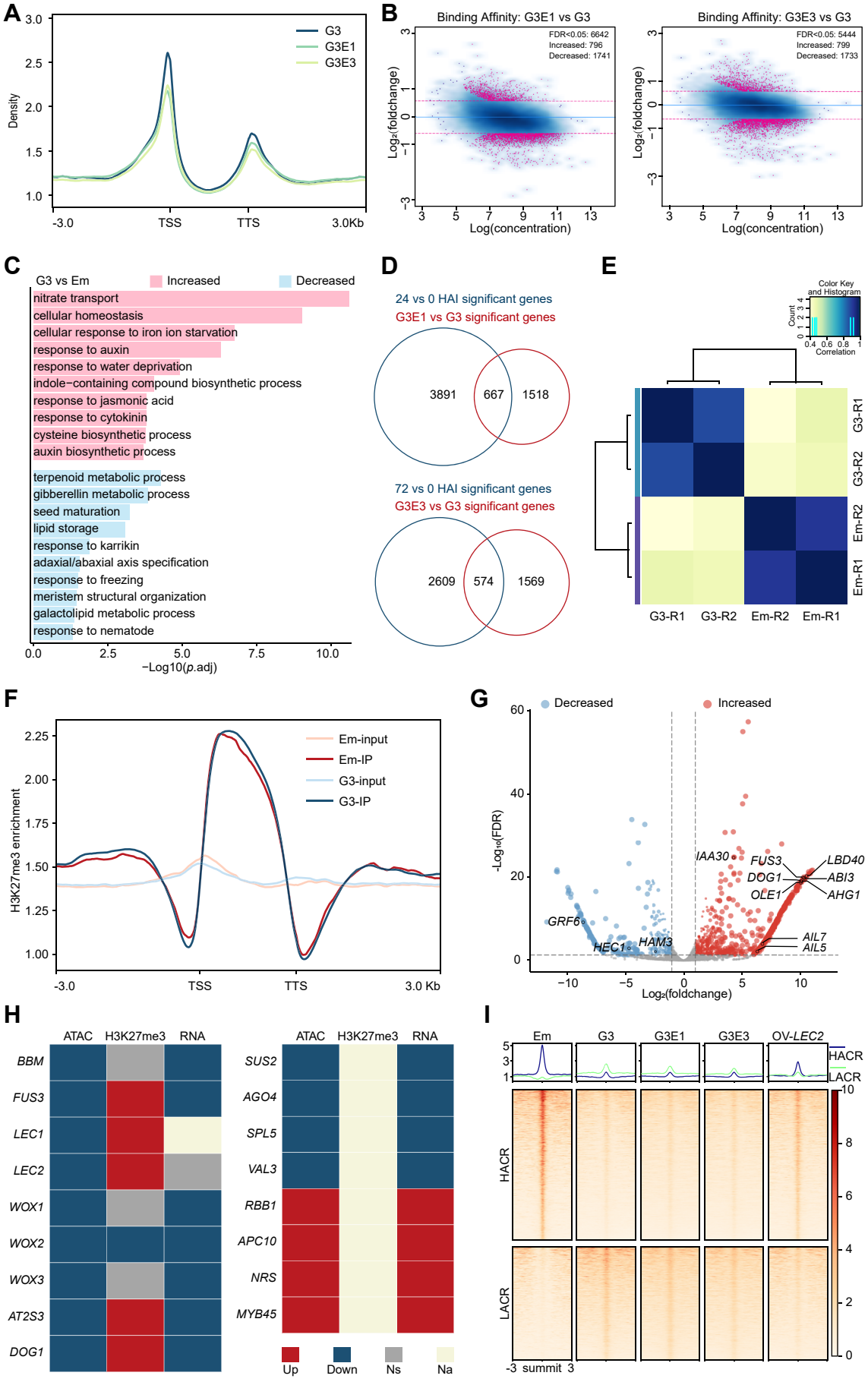

A
G3
G3E1
G3E3
Density
TSS
TTS
3.0Kb
B
Binding Affinity: G3E1 vs G3
FDR<0.05: 6642
Increased: 796
Decreased: 1741
Binding Affinity: G3E3 vs G3
FDR<0.05: 5444
Increased: 799
Decreased: 1733
Log2(foldchange)
Log(concentration)
C
G3 vs Em
Increased
Decreased
nitrate transport
cellular homeostasis
cellular response to iron ion starvation
response to auxin
response to water deprivation
indole−containing compound biosynthetic process
response to jasmonic acid
response to cytokinin
cysteine biosynthetic process
auxin biosynthetic process
terpenoid metabolic process
gibberellin metabolic process
seed maturation
lipid storage
response to karrikin
adaxial/abaxial axis specification
response to freezing
meristem structural organization
galactolipid metabolic process
response to nematode
−Log10(p.adj)
D
24 vs 0 HAI significant genes
G3E1 vs G3 significant genes
3891
667
1518
72 vs 0 HAI significant genes
G3E3 vs G3 significant genes
2609
574
1569
E
Color Key and Histogram
Count
Correlation
G3-R1
G3-R2
Em-R2
Em-R1
F
Em-input
Em-IP
G3-input
G3-IP
H3K27me3 enrichment
TSS
TTS
3.0 Kb
G
Decreased
Increased
-Log10(FDR)
Log2(foldchange)
IAA30
FUS3
LBD40
DOG1
ABI3
OLE1
AHG1
GRF6
HEC1
HAM3
AIL7
AIL5
H
ATAC
H3K27me3
RNA
BBM
FUS3
LEC1
LEC2
WOX1
WOX2
WOX3
AT2S3
DOG1
SUS2
AGO4
SPL5
VAL3
RBB1
APC10
NRS
MYB45
Up
Down
Ns
Na
I
Em
G3
G3E1
G3E3
OV-LEC2
HACR
LACR
summit

Figure S3. The comparisons among the explants of different origins. related to Figure 3 and Figure 4.

(A) The genome-wide distribution of ATAC-seq peaks.

(B) MA plots showing fold-change of differentially accessible peaks. Blue, distribution of constitutive peaks; pink dots, individual differential peaks. The numbers of differentially accessible peaks (opening or closing) according to FDR alone or FDR with Fold change cut off are indicated at the lower right corner of each plot. FDR cut off is < 0.05 while Fold change cut off is >1.5 or <-1.5.

(C) GO term analyses showing distinct gene ontologies of target genes linked to differentially accessible peaks (G3 vs Em). The selected 20 enriched GO biological processes are indicated. The -$\log_{10}$(*p*.adj) is given.

(D) Venn plot showing differentially accessible peaks between the samples. Please note that, compared to Em samples, G3 samples showed clear reduction in their responsiveness to high auxin environment.

(E) Reproducibility of H3K27me3 ChIP-seq experiments. Two biological replicates (R1 and R2) for G3 and Em are shown.

(F) The genome-wide distribution of H3K27me3 in Em and G3 samples. The immunoprecipitation (IP) and input data are shown.

(G) Volcano plot of the genes associated with decreased (blue) or increased (red) level of H3K27me3. The representative genes are shown. Grey, no difference between two samples.

(H) Heatmap showing the accessibilities, H3K27me3, and transcript levels of selected genes. Na, not expressed or undetectable deposition of H3K27me3; Ns, no difference in transcripts or H3K27me3 levels among the samples. Up, increased accessibility or transcription levels; Down, decreased accessibility or transcription levels. See also Figure 3E.

(I) Heatmap and pileup of ATAC-seq signals. The peaks with increased or decreased accessibility (Log_2(foldchange) >1 or <-1) are shown. OV-LEC2, *p35S::LEC2-GR* seedlings treated with DEX for 7 days. Heat maps are ranked in decreasing order of ATAC-seq signal. Window size: peak summit ± 3.0 kb. HACR and LACR refer to relatively highly accessible chromatin region (Em vs G3) and

relatively lowly accessible chromatin region (Em vs G3), respectively.

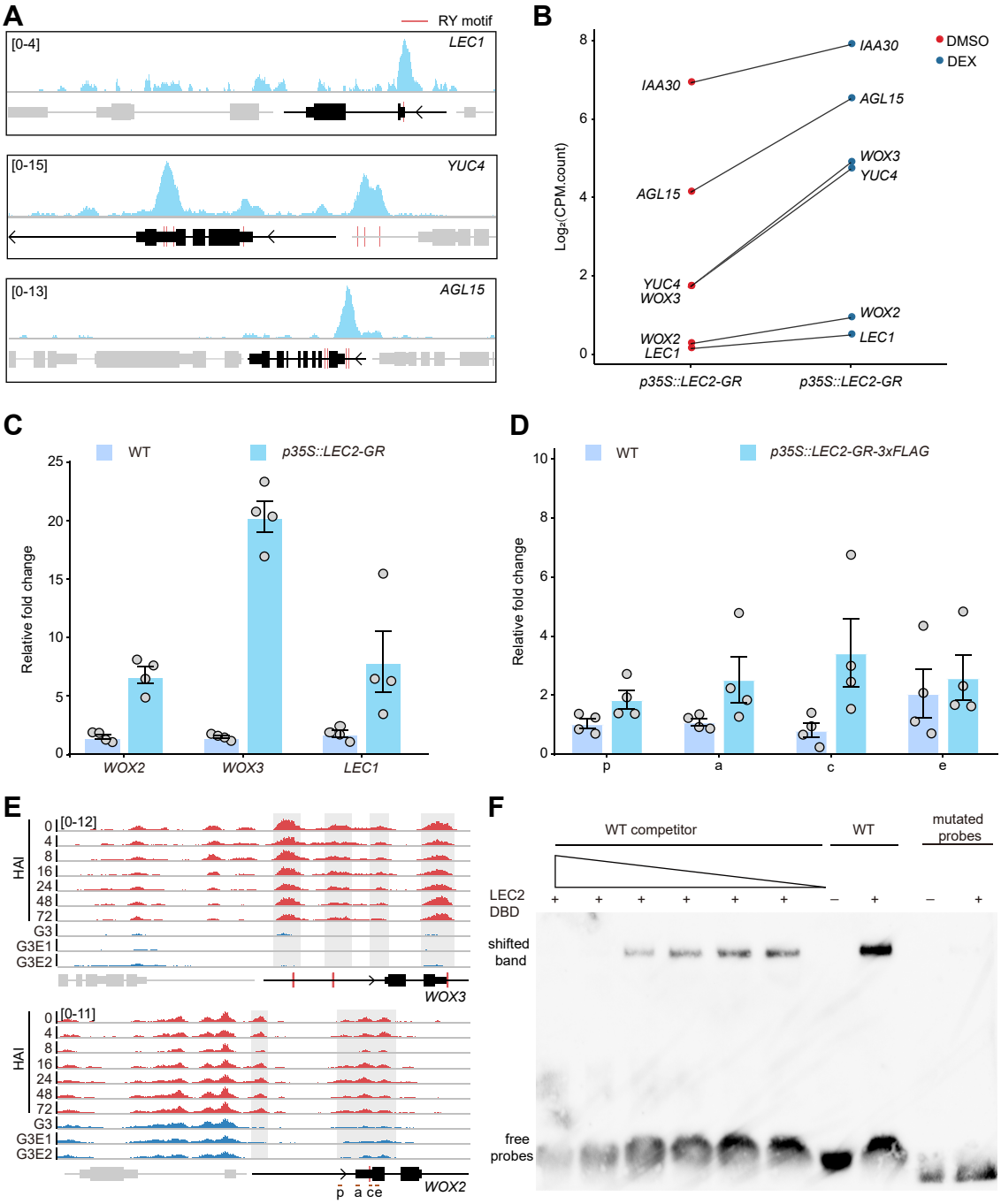


Figure S4. *WOX2* and *WOX3* are direct targets of LEC2. related to Figure 4.

(A) Identification of known LEC2 targets. The genomic regions are shown and the selected genes are highlighted in black. The location of RY motif is indicated in red line.

(B) Expression of LEC2 targets. Log_2CPM (RNA_counts per million, RNA_CPM) for each gene is shown.

(C) Validation of RNA-seq data by RT-qPCR. Wild type and *p35S:LEC2-GR* were used. Data are represented as mean ± SD (*n*=4).

(D) ChIP-qPCR analysis. The *p35S::LEC2-GR-3xFLAG* seedlings cultured in liquid B5 media were treated with DEX or DMSO (mock) for 4 hours and subjected to ChIP analyses. The relative enrichment of ChIP DNA was determined by qPCR. Data are represented as mean ± SD (*n*=4).

(E) The ATAC-seq tracks for the *WOX2* and *WOX3* loci. The genomic regions are shown and the selected genes are highlighted in black. The location of RY motif is indicated in red line. Accessible peaks of each gene are shadowed.

(F) Competitive EMSA showing binding of LEC2-DBD (DNA-binding domain of LEC2) to the RY motif in *WOX2* promoter. The relative amount (labeled oligonucleotide was set to 1.0) of unlabeled competitive oligonucleotide is indicated on the top. Schematic of *WOX2* genomic region and the positions of the RY motif I and II are shown in Figure 4H. The assays without LEC2-BD protein (-) or with mutated probes were performed as controls. The probe sequences are listed in Table S7.

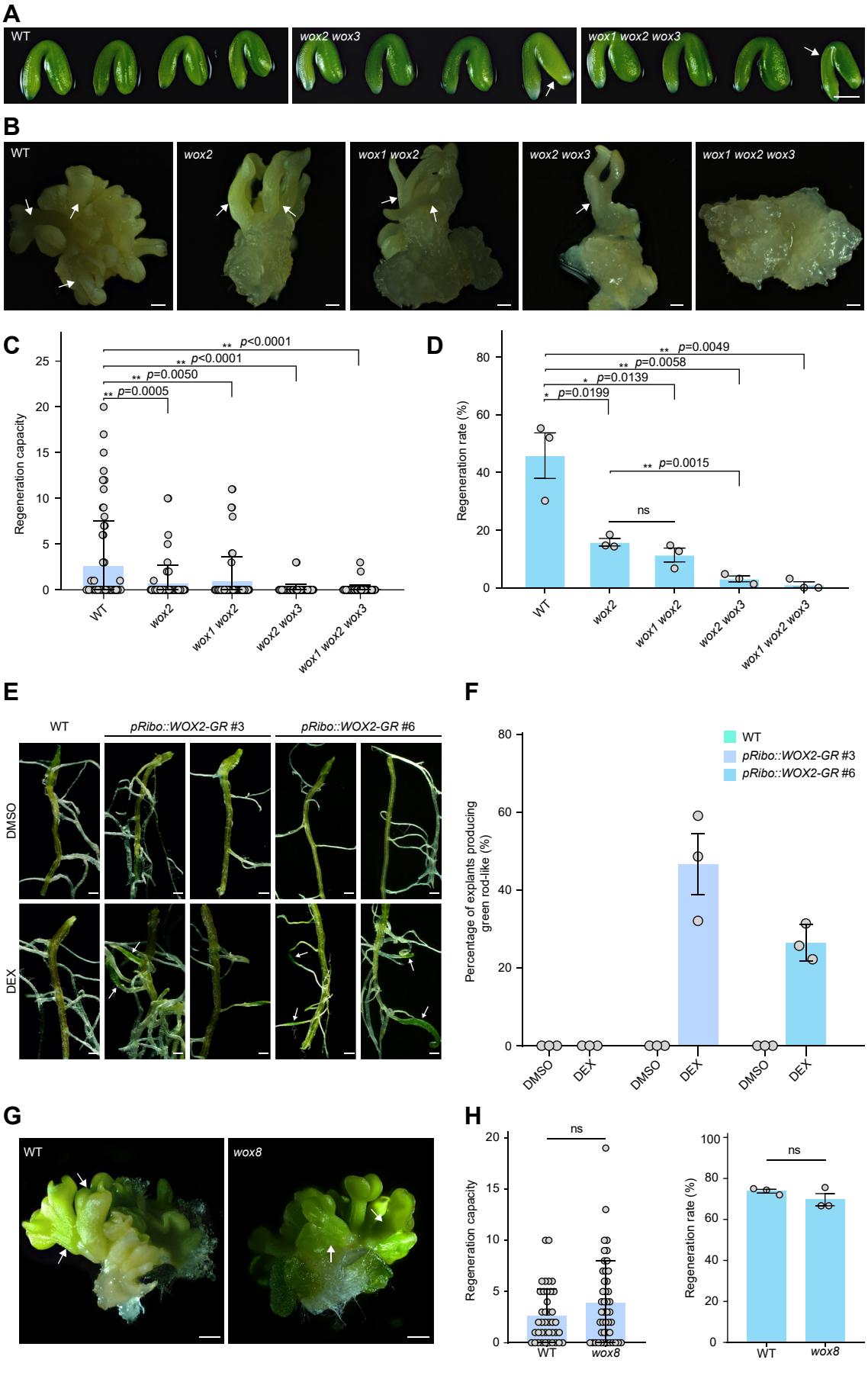

A
WT
wox2 wox3
wox1 wox2 wox3
B
WT
wox2
wox1 wox2
wox2 wox3
wox1 wox2 wox3
C
Regeneration capacity
** p<0.0001
** p<0.0001
** p=0.0050
** p=0.0005
WT
wox2
wox1 wox2
wox2 wox3
wox1 wox2 wox3
D
Regeneration rate (%)
** p=0.0049
** p=0.0058
* p=0.0139
* p=0.0199
** p=0.0015
ns
WT
wox2
wox1 wox2
wox2 wox3
wox1 wox2 wox3
E
WT
pRibo::WOX2-GR #3
pRibo::WOX2-GR #6
DMSO
DEX
F
WT
pRibo::WOX2-GR #3
pRibo::WOX2-GR #6
Percentage of explants producing green rod-like (%)
DMSO
DEX
G
WT
wox8
H
Regeneration capacity
ns
WT
wox8
Regeneration rate (%)
ns
WT
wox8

Figure S5. *WOX2* and *WOX3* are essential for SE. related to Figure 5.

(A) Phenotype of wild type, *wox2 wox3* and *wox1 wox2 wox3* explants. Arrow indicates mono-cotyledon embryo which was not used for SE experiments. Scale bars represent 500 μm.

(B) SE experiments with embryonic calli. Wild type, *wox2*, *wox1 wox2*, *wox2 wox3* and *wox1 wox2 wox3* mutants were used as explants. Arrow indicates somatic embryo. Scale bars represent 200 μm.

(C) Regeneration capacity assays using embryonic calli. The regeneration capacity was represented by the number of somatic embryos per explant ($n > 60$). Data is from represent single experiment which was repeated three times independently. Data are represented as mean ± SD; each dot represent the number of somatic embryos per explant. One-way ANOVA was performed followed by a Turkey's multiple comparisons test, **P < 0.01.

(D) Regeneration rate assays using embryonic calli. Wild type, *wox* single and high-order mutants were used. The regeneration rate was represented by the percentage of explants with somatic embryos. Data was from three independent experiments. Data are represented as mean ± SD; Unpaired t test was performed; *P < 0.05, **P < 0.01. ns, not significant.

(E) *WOX2* could not induce SE alone. *pRibo::WOX2-GR* explants were treated with DEX. Arrow indicates green rod-like structures. Please note that no somatic embryo was formed on these explants. Scale bars represent 200 μm.

(F) Quantification of green rod-like structures. Two independent *pRibo::WOX2-GR* lines are shown.

(G) Regeneration assays with wild type and *wox8* mutant. Arrow indicates somatic embryo. Scale bars represent 200 μm.

(H) Quantification of regeneration capacity of wild type and *wox8* mutant. Please note that *wox8* single mutant did not exhibit altered SE regeneration rate. Unpaired *t*-test was performed. ns, not significant.

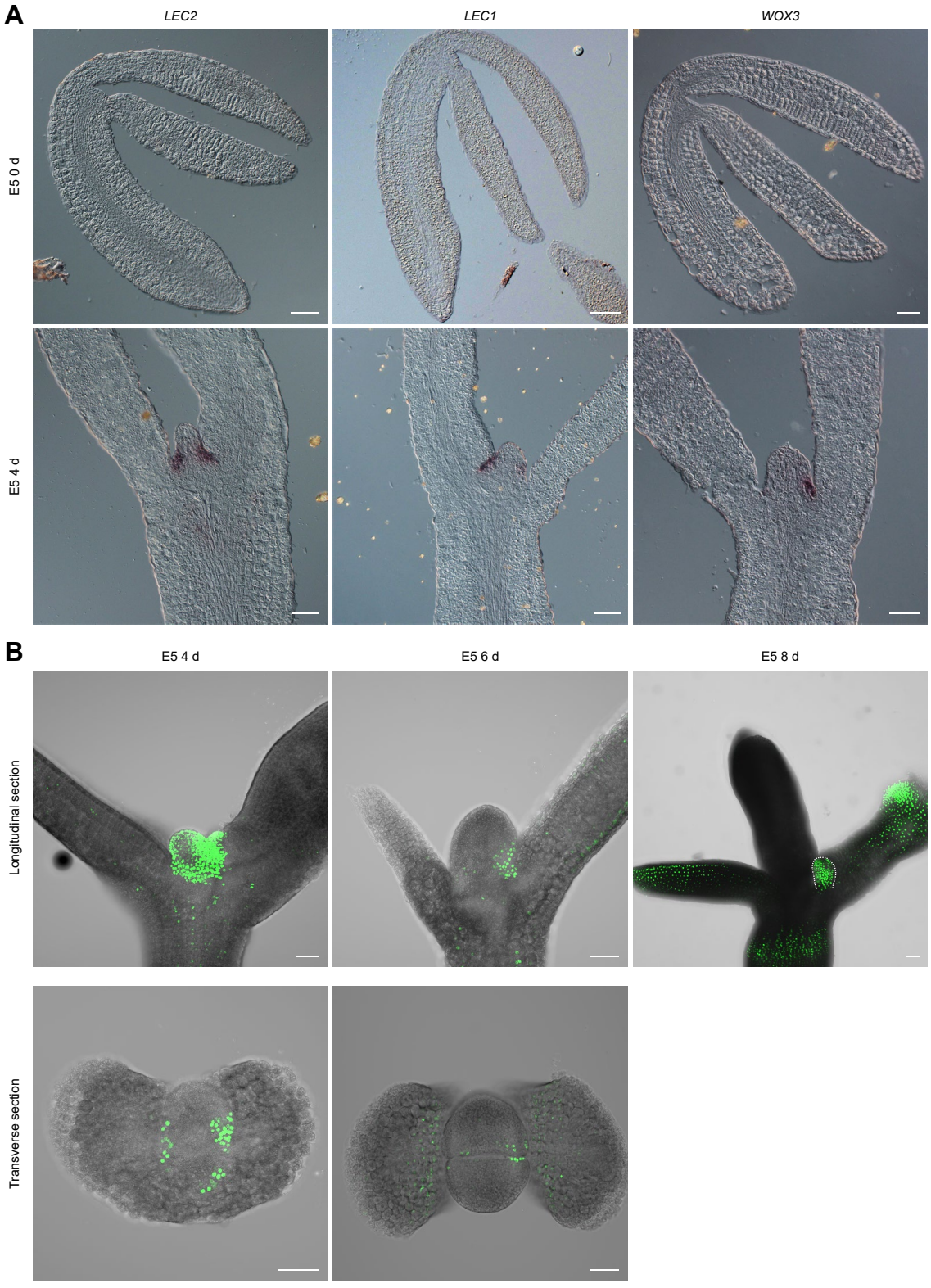


Figure S6. Expression pattern of *LEC1*, *LEC2*, *WOX2* and *WOX3* during SE. related to Figure 6.

(A) The expression pattern of *LEC1*, *LEC2* and *WOX3* revealed by *in situ* hybridization. The longitudinal sections are shown. Please note that all these genes were expressed at the boundary domain of shoot apical meristem after 4 days of treatment on E5 media. Scale bars represent 50 μm.

(B) The analysis of *WOX2* reporter. The explants at different stages were embedded and sectioned. The longitudinal and transverse sections are shown. Please note that strong *WOX2* expression (green) was observed at the boundary domain of shoot apical meristem. Dash lines, embryo-like protuberance at E5 8 d. Max intensity projection in the z-axis of one representative sample is shown. Scale bars represent 50 μm.

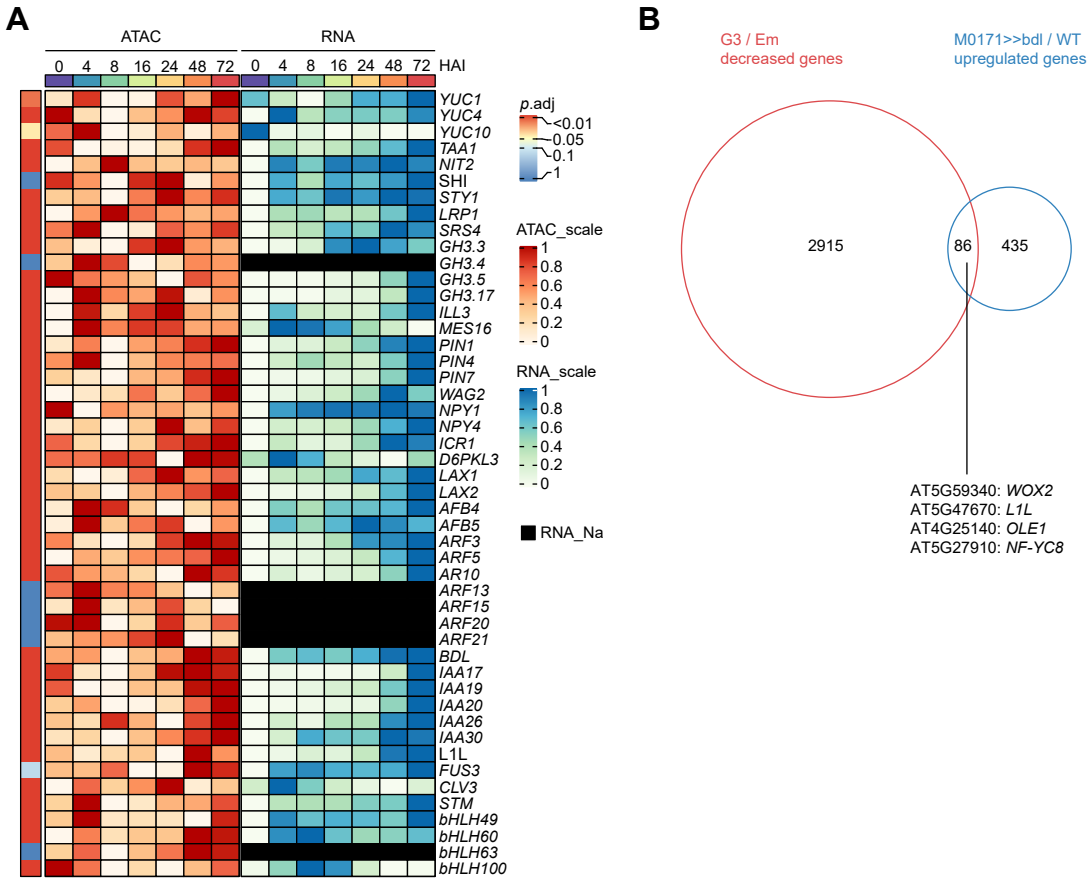


Figure S7. The comparison between SE and suspensor embryogenesis. related to Figure 7.

(A) Chromatin accessibility and expression pattern of differentially expressed genes during suspensor embryogenesis (Radoeva et al., 2019a). Heatmap showing accessibility and transcript levels of selected genes. The accessibility of the genomic region (-3.0 to 0.5 kb from TSS) of selected gene was analyzed. The CPM values are scaled using min-max normalization method. Left bar, the LRT *p*.adj of gene expression level; Black cell, gene not expressed (RNA_Na).

(B) The comparison between the genes up-regulated during suspensor embryogenesis (M0171-bdl/WT, fold change>=2.0, Radoeva et al., 2019a) and the genes with decreased chromatin accessibility during seed germination (G3 vs Em). In total, 86 genes are shared by two datasets. Among them, the key regulators for SE such as *WOX2* and *L1L* (*LEC1-like*) are identified.