Article

# LEC2 induces somatic cell reprogramming through epigenetic activation of plant cell totipotency regulators

Received: 22 September 2024

Accepted: 16 April 2025

Published online: 06 May 2025

Check for updates

Jing Peng  $\mathbb{D}^{1,4}$ , Qi Zhang<sup>1,4</sup>, Li Ping Tang  $\mathbb{D}^{1,4}$ , Biao Jie Xu  $\mathbb{D}^1$ , Thomas Laux  $\mathbb{D}^{2,3}$ , Xian Sheng Zhang  $\mathbb{D}^1$  & Ying Hua Su  $\mathbb{D}^{1,2}$ 

Many plant species can develop embryos from somatic cells without fertilization. During this process, known as somatic embryogenesis, changes in the DNA methylation patterns are characteristic of reprogramming somatic cells into an embryogenic state. However, the underlying mechanisms connecting DNA methylation and activating totipotency-regulating genes have remained largely unknown. Here, we show that during somatic embryogenesis induced by overexpressing the totipotency-regulating transcription factor LEAFY COTYLEDON2 (LEC2) in Arabidopsis, CHH hypermethylation is deposited by the LEC2-activated RNA-directed DNA methylation (RdDM) pathway. A reader complex composed of SU(VAR)3-9 HOMOLOGS (SUVH) and its chaperone SUVH-INTERACTING DNAJ DOMAIN-CONTAINING PROTEIN (SDJ) binds to the CHH hypermethylated regions and recruits AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED (AHL) chromatin modification proteins to increase chromatin accessibility, resulting in the transcriptional activation of totipotency-regulating genes. Our work reveals a molecular framework of how epigenetic modifications mediate somatic cell reprogramming, offering a pathway toward enhancing somatic embryogenesis in agricultural regeneration biology.

Unlike animal cells, plant somatic cells exhibit a high degree of developmental plasticity, allowing them to generate new plants without fertilization<sup>1</sup>. This process, known as somatic embryogenesis, is of significant economic importance for the asexual propagation of (hybrid) crops and for eliminating viruses and viroids during agricultural breeding<sup>2</sup>.

Somatic embryos (SEs) occur spontaneously only in a few plant species but can be induced through various exogenous stimuli in many other species. As a key event, somatic cells are reprogrammed by activating regulatory genes that lead to the acquisition of cell totipotency<sup>1</sup>. This can be achieved traditionally by culturing explants on different auxin media<sup>3</sup>. However, SEs can also be induced in the absence of exogenous plant growth regulators by overexpression of totipotency-regulating transcription factors, such as LEAFY COTYLEDON2 (LEC2)<sup>4</sup>, LEC1<sup>5</sup>, or BABY BOOM (BBM)<sup>6</sup>, that have been associated with cellular totipotency<sup>7</sup>. Loss-of-function mutations in several genes, including *LEC1*, *LEC2*, *ABA INSENSITIVE3* (*ABI3*), *FUSCA3* (*FUS3*), and *BBM*, can severely reduce the efficiency of somatic embryogenesis<sup>8-10</sup>. While *LEC1*, *LEC2*, or *BBM* overexpression can directly induce somatic embryogenesis, *ABI3* and *FUS3* overexpression can only enhance somatic embryogenesis induced by other means<sup>11,12</sup>. Previous research has shown that BBM promotes somatic embryogenesis through transcriptional activation of *LEC2*<sup>13</sup>. Thus, it is reasonable to hypothesize that LEC2 plays a more direct role in the totipotency regulatory network<sup>14</sup>.

<sup>1</sup>National Key Laboratory of Wheat Improvement, College of Life Sciences, Shandong Agricultural University, Tai'an, Shandong, China. <sup>2</sup>Sino-German Joint Research Center on Agricultural Biology, Shandong Agricultural University, Tai'an, Shandong, China. <sup>3</sup>Signalling Research Centres BIOSS and CIBSS, Faculty of Biology, University of Freiburg, Freiburg, Germany. <sup>4</sup>These authors contributed equally: Jing Peng, Qi Zhang, Li Ping Tang. e-mail: suyh@sdau.edu.cn

During somatic embryogenesis, differentiated somatic cells undergo a dramatic cell fate transition and exhibit changes in cellular characteristics. For example, cotyledon cells overexpressing LEC2 accumulate auxin<sup>15</sup> due to LEC2-mediated activation of auxin biosynthesis genes such as YUCCAs (YUC1, YUC4, and YUC10) by binding to the RY motifs in their promoters<sup>15,16</sup>. Studies have shown that endogenous auxin biosynthesis and polar transport are essential for somatic embryogenesis<sup>17,18</sup>. In addition, LEC2 also promotes the accumulation of lipids by directly binding to the RY motifs in the promoters of key lipid biosynthesis genes, critical for SE formation<sup>15,19</sup>. During this process, the expression of several genes involved in lipid biosynthesis, such as WRINKLED1 (WRI1), OLEOSIN3 (OLE3), 3-KETOACYL-ACYL CAR-RIER PROTEIN SYNTHASE I (KASI), KASIII, and TRIACYLGLYCEROL BIO-SYNTHESIS DEFECT 1 (TAG1), was induced by LEC2 overexpression<sup>19,20</sup>. Moreover, loss of WRI1 function, a key regulator of lipid biosynthesis, inhibited LEC2-induced somatic embryogenesis<sup>19</sup>. However, the mechanisms of LEC2 function in the initiation of somatic embryogenesis remain elusive.

DNA methylation at cytosine residues in different contexts has been intensively studied in eukarvotes. DNA methylation is essential for regulating gene transcription through its influence on chromatin structure and protein-DNA interaction<sup>21</sup>. In plants, the genome undergoes selective methylation in CG, CHG, and CHH (where H represents A, T, or C) contexts. All these methylations are established de novo via the RNA-directed DNA methylation (RdDM) pathway<sup>22</sup>. The RdDM pathway relies on the DNA-dependent RNA polymerase IV (Pol IV) to generate RNAs, which are processed into siRNAs and loaded onto ARGONAUTE 4/6 (AGO4/6)23. AGO4/6 is then recruited by Pol V, the largest subunit NUCLEAR RNA POLYMERASE E1 (NRPE1) through the transcription elongation factor KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1)<sup>23</sup>. The resulting complex recruits the DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) by RNA-DIRECTED DNA METHYLATION 1 (RDM1) for de novo DNA methylation<sup>23</sup>. Once established, DNA methylation patterns are maintained by context-specific methyltransferases. METHYLTRANSFERASE 1 (MET1) maintains CG methylation<sup>24</sup>, while the plant-specific methyltransferase CHROMOMETHYLASE 3 (CMT3) maintains CHG methylation<sup>25</sup>, CMT2 maintains CHH methylation of heterochromatin. and the DRM2/RdDM pathway maintains CHH methylation of euchromatic regions<sup>26</sup>. CHH methylation deposited through the DRM2/RdDM pathway is recognized by SU(VAR)3-9 HOMOLOGS 1/3 (SUVH1/3), which recruit the chaperone SUVH-INTERACTING DNAJ DOMAIN-CONTAINING PROTEIN 1/2/3 (SDJ1/2/3) to form a reader complex that activates transcription<sup>27,28</sup>. By contrast, recruitment of methyl-CpG binding domain 5/6 (MBD5/6) readers with the chaperone SILENZIO (SLN) to CG sites results in a negative effect on transcription<sup>29</sup>. These examples highlight the dual role of DNA methylation in regulating gene expression.

DNA methylation is crucial during early embryogenesis. In Arabidopsis, the DNA methyltransferases DRM2, CMT3, and MET1 are highly expressed in embryos<sup>30</sup>, and the loss of function of MET1 leads to abnormal embryos<sup>31</sup>. Recent research revealed changes in DNA methylation levels during somatic embryogenesis in soybean<sup>32</sup> and cotton<sup>33</sup>. During the direct induction of somatic embryogenesis from immature zygotic embryos in Arabidopsis, the DNA methyltransferases DRM2, CMT3, and MET1 are also highly expressed<sup>34</sup>. However, our understanding of the mechanism linking DNA methylation with somatic embryogenesis remains incomplete.

DNA methylation patterns are associated with the chromatin structure<sup>35</sup>. Recent studies have shown increased chromatin accessibility and expression of totipotency-regulating genes, including *LEC1*, *LEC2*, *BBM*, *FUS3*, and *PLETHORAs* (*PLTs*), during early somatic embryogenesis induced by auxin application from Arabidopsis immature zygotic embryos, but not in seedlings<sup>14</sup>. In addition, mutations in histone modifiers, such as the histone deacetylases HISTONE

DEACETYLASE 6 (HDA6) and HDA19<sup>36</sup>, and the POLYCOMB REPRES-SIVE COMPLEX 2 (PRC2) subunits CURLY LEAF (CLF) and SWINGER (SWN)<sup>37</sup>, can induce SE formation from Arabidopsis seedlings. Mutations in these genes increase histone acetylation or reduce histone 3 lysine 27 tri-methylation (H3K27me3), activating totipotencyregulating genes<sup>36,37</sup>. Finally, overexpression of the *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED 15* (*AHL15*) triggers chromatin decondensation and somatic embryogenesis<sup>38</sup>. These findings highlight the importance of the chromatin state for acquiring totipotency.

Here, we show that LEC2 promotes CHH hypermethylation via upregulation of *DRM2* during LEC2-induced Arabidopsis somatic embryogenesis. Subsequently, the SUVH-SDJ complex binds the CHH hypermethylated regions within the promoters of totipotencyregulating genes and recruits AHLs, increasing chromatin accessibility. Furthermore, the SUVH-SDJ-AHL complex interacts with the LEC2 protein, enhancing its ability to activate totipotency-regulating genes. Our findings provide a mechanistic framework by which epigenetic reprogramming enables differentiated somatic plant cells to acquire totipotent fates.

### Results

#### LEC2 overexpression induces the DRM2/RdDM pathway

Our previous observations revealed that SEs were induced from the cotyledons of seedlings by β-estradiol (ES)-induced ectopic overexpression of the *pER8-LEC2* transgene<sup>19</sup>. To detect differential gene expression during early steps in this process, we performed RNA-seq from pER8-LEC2-induced cotyledons at 72, 84, and 96 hours after induction (HAI; Supplementary Fig. 1a). We identified 2,721 upregulated and 3537 downregulated differentially expressed genes (DEGs) across the three-time points in ES-treated pER8-LEC2 plants (hereafter referred to as pER8-LEC2) compared to DMSO-treated pER8-LEC2 plants (hereafter referred to as Mock; Supplementary Fig. 1b, c and Supplementary Data 1). Gene ontology (GO) analysis revealed that the downregulated genes were associated with the GO-terms photosynthesis, trichoblast differentiation, or cell maturation, consistent with inhibited cell differentiation in cotyledons upon LEC2-induction (Supplementary Fig. 1d and Supplementary Data 1). By contrast, the upregulated genes were mainly associated with the GO-terms somatic embryogenesis, fatty acid biosynthesis, or DNA methylation (Supplementary Fig. 1d and Supplementary Data 1). We observed that the upregulated genes in the DNA methylation term encode proteins that primarily function in the RdDM pathway. These include the DNA methyltransferase DRM2, the essential RNA Pol V component NRPE1, and the AGO4/6 proteins, which are the key players of the RdDM pathway (Fig. 1a). We confirmed the pER8-LEC2-mediated upregulation of these genes by quantitative real-time PCR (qRT-PCR) (Supplementary Fig. 1e). Furthermore, we found that the promoters of most of these genes contain LEC2 binding peaks (Supplementary Fig. 2), suggesting that these genes involved in the RdDM pathway may be direct targets of LEC2.

Therefore, we chose to explore the role of the RdDM in LEC2induced somatic embryogenesis. We focused on DRM2 for further analysis because it is the rate-limiting factor in de novo DNA methylation, regulating the enzymatic processes of the RdDM pathway<sup>39</sup>. To explore the *DRM2* expression pattern during LEC2-induced somatic embryogenesis, we employed the functional genomic fusion construct *gDRM2-GFP*<sup>30</sup>. *gDRM2-GFP* expression was high in the epidermal cells of the cotyledons from 24 HAI in *pER8-LEC2*. In contrast, it was restricted to stomata lineage cells starting from 48 HAI in the Mock control (Fig. 1b, c). The *DRM2* promoter contains two putatively LEC2-binding RY motifs<sup>14</sup> (TGCATG, -491/-496 and -719/-724). We found that in the ChIP-Seq<sup>14</sup> analysis, LEC2 was enriched in the region containing the -491/-496 but not the -719/-724 motif (Supplementary Fig. 2). Utilizing chromatin immunoprecipitation followed by qPCR analysis (ChIP-qPCR) of *pER8-LEC2-GFP* seedlings at 24 HAI, we detected



Fig. 1 | The DRM2/RdDM pathway is essential for LEC2-induced somatic embryogenesis. a Expression heatmap illustrating the upregulation of genes involved in the DNA methylation pathway during LEC2-induced somatic embryogenesis. **b**, **c** The expression pattern of *gDRM2-GFP* in *pER8-LEC2* with DMSO (referred to as Mock, **b**) or  $\beta$ -estradiol (ES) treatment (referred to as *pER8-LEC2*, **c**); hours after induction (HAI) are indicated. The magenta signal represents the cell wall stained with 4',6-diamidino-2-phenylindole (DAPI). Scale bars, 10 µm. **d** Upper panel: schematic representation of the 800 bp region upstream of the ATG codon (+1) of the *DRM2* gene. PCR fragments labeled "a" to "c" were utilized for the ChIPqPCR analysis, with "c" serving as a negative control. Lower panel: ChIP-qPCR of LEC2-GFP showing enrichment of the "b" region of the *DRM2* promoter at 24 HAI. *P*-values are shown, two-tailed Student's *t* test. **e** EMSA of GST-LEC2 and the *DRM2* promoter region containing the LEC2 binding site. Four replicates showed similar results. The black arrowheads indicate the position of the shifted bands, nonspecific bands, and the free probe, respectively. **f** Relative ratio of Firefly LUC to Renilla LUC (REN) activity in tobacco leaves. *3SSpro::LEC2* was used as effector and *DRM2pro::LUC* as reporter. Control and LUC indicate the corresponding empty vectors. **g**–**j** Phenotypes of seedlings of indicated genotypes and treatments at 14 days after induction (DAI). White arrowheads in **h** and **j** point to somatic embryos (SEs). Scale bars, 0.5 mm. **k**, **l** The statistical analysis of the frequencies of somatic embryogenesis **k** and the average number of SEs per explant **l**. n, number of transgenic lines analyzed. The data in **d**, **f**, **k** are presented as means ± SD, with **d**, **f** representing results from three biological replicates. For box plots in **l**, the horizontal line represents the median value, the boxes represent the interquartile range (25th to 75th percentiles), and the whiskers extend to the maximum and minimum values. Dots indicate individual values. Different letters in **f**, **k**, **l** indicate significant differences (*P* < 0.05), one-way ANOVA with Tukey's multiple comparison test. Source data are provided as a Source Data file.

enrichment of fragments harboring the -491/-496 RY motif of the *DRM2* promoter (Fig. 1d). Electromobility shift assay (EMSA) confirmed a direct interaction between LEC2 and a 21-bp promoter DNA fragment containing this RY motif (Fig. 1e). In addition, LEC2 activated a luciferase reporter driven by the *DRM2* promoter in tobacco leaves

(Fig. 1f). These findings suggest that LEC2 can directly activate *DRM2* transcription.

To address whether DRM2 is an essential downstream component for *pER8-LEC2*-induced SE formation, we introduced the transgene into the homozygous drm1/2 loss-of-function double mutant

https://doi.org/10.1038/s41467-025-59335-8

(DRM1 encodes a lowly expressed paralog of DRM2) to ensure a complete loss of DRM function (Supplementary Fig. 3). Indeed, the *drm1/2* mutation strongly suppressed LEC2-induced SE formation (Fig. 1g-i, k, l and Supplementary Data 2), including excessive lipid accumulation (Supplementary Fig. 4a–d) and cell proliferation (Supplementary Fig. 4e–k), which we previously showed to be integral to LEC2-induced somatic embryogenesis<sup>19</sup>. On the contrary, somatic embryogenesis was further enhanced by increasing *DRM2* expression (Fig. 1g, h, j–l, Supplementary Figs. 3 and 4, and Supplementary Data 2). These results suggest that the DRM2/RdDM pathway is an essential downstream component of LEC2-induced somatic embryogenesis.

### DRM2/RdDM enhances totipotency-regulating gene expression

To examine whether the upregulation of DRM2 affects DNA methylation levels during somatic embryogenesis, we performed wholegenome bisulfite sequencing (WGBS) and found increased CHH methylation levels 96 hours after pER8-LEC2 induction (Fig. 2a and Supplementary Fig. 5a). In contrast, no significant differences in CG and CHG methylation levels were observed (Fig. 2a and Supplementary Fig. 5a). Accordingly, CHH differentially methylated regions (DMRs) compared to Mock in pER8-LEC2 were the most abundant among the three types of DNA methylation (Fig. 2b and Supplementary Data 3). The global CHH hypermethylation was reversed into CHH hypomethylation compared to Mock in pER8-LEC2 drm1/2 (Fig. 2a and Supplementary Fig. 5a). In line with this, the majority of CHH hyper-DMRs observed in pER8-LEC2 was reversed into hypo-DMRs by drm1/2 (Fig. 2b, Supplementary Fig. 5b, c and Supplementary Data 3). Furthermore, when comparing our data to the published ChIP-seq data for the largest POL V subunit NRPE1<sup>40</sup>, we found that NRPE1 co-localized with CHH DMRs (Supplementary Fig. 5d). These results indicate that DRM1/2 are required for CHH hypermethylation during LEC2-induced somatic embryogenesis.

We then investigated the genomic distribution of CHH hyper-DMRs after *pER8-LEC2* induction. Our results revealed that the vast majority (96.37%) of DRM1/2-dependent CHH hyper-DMRs were within 5 kb upstream of the transcription start site (TSS) of genes, with 88.15% within 3 kb, and 55.72% within 1 kb. The remaining (3.63%) DRM1/2dependent CHH hyper-DMRs were distributed in exons, introns, downstream regions, and distal intergenic regions (Fig. 2c and Supplementary Data 3). A ratio of observation/expectation (O/E) of CHH hyper-DMRs located at each genomic feature was calculated, demonstrating that the CHH hyper-DMRs were mainly located within 3 kb of the promoter regions (Supplementary Fig. 5e). Notably, we also found that LEC2 binds to genes rather than transposable elements (TEs) (Supplementary Fig. 5f). Given that most DRM1/2-dependent CHH hyper-DMRs are located in the promoter regions of genes, our subsequent association analysis focused on the DRM1/2-dependent CHH hyper-DMRs within the promoter regions associated with the up- and down-regulated genes, through RNA-seq and WGBS data. We identified 678 upregulated and 841 downregulated genes associated with DRM1/2-dependent CHH hypermethylation during LEC2-induced SE formation (Fig. 2d, Supplementary Fig. 6a, and Supplementary Data 4). GO enrichment analysis revealed that the downregulated genes were related to biological cell differentiation processes, including lateral root morphogenesis [MYB DOMAIN PROTEIN 88 (MYB88) and MICROTUBULE-ASSOCIATED PROTEIN 18 (MAP18)], stomatal closure [RESPONSIVE TO DEHYDRATION 21 (RD21A), BETA GLUCOSIDASE 37 (BGLU37), and NONPHOTOTROPIC HYPOCOTYL 1 (NPH1)], and root morphogenesis [PLASMA MEMBRANE INTRINSIC PROTEIN 2;4 (PIP2;4)]. The genes involved in these pathways showed hypermethylation and downregulation of expression levels after LEC2 induction (Supplementary Fig. 6b-d and Supplementary Data 4). Conversely, the upregulated genes were mainly associated with the GO-terms somatic

Based on LEC2 ChIP-seq<sup>14</sup> data and binding motif analyses, we discovered that LEC2 peaks are more enriched in the promoters of upregulated than downregulated genes (Supplementary Fig. 5g). Furthermore, the upregulated genes contain a significantly higher density of RY motifs in their promoter regions compared to the downregulated genes (Supplementary Fig. 5h). Because of their association with GO-terms related to somatic embryogenesis, we further focused on the 678 upregulated genes with CHH hypermethylation. Among these genes, we identified 63 genes from SE induction-related GO terms: somatic embryogenesis, auxin biosynthesis, and fatty acid biosynthesis, etc., including known totipotency-regulating genes LEC1<sup>5</sup>, ABI3<sup>12</sup>, MYB118<sup>41</sup>, AHL29<sup>38</sup>, YUC1<sup>17</sup>, YUC10<sup>17</sup>, KASI<sup>19</sup>, KASIII<sup>19</sup>, and TAG1<sup>19</sup> (Fig. 2f and Supplementary Data 4). Similar expression patterns and CHH methylation levels of the totipotency-regulating genes YUC5<sup>18</sup>, BBM<sup>13</sup>, and FUS3<sup>11</sup> genes were also shown in our RNA-seq and WGBS data (Fig. 2f). We confirmed that CHH hypermethylation is associated with the high expression of these totipotency-regulating genes by qRT-PCR and bisulfite sequencing (Supplementary Fig. 7a, b). In contrast to CHH methylation, CG and CHG methylation levels did not show significant upregulation upon *pER8-LEC2* induction for most totipotency-regulating genes (Supplementary Fig. 7c, d). In addition, we conducted both RNA-seq and qRT-PCR analyses on cotyledons from Mock, *pER8-LEC2*, and *pER8-LEC2 drm1/2* at 96 HAI, respectively. Our analysis revealed that the drm1/2 mutations suppressed the expression of genes upregulated by LEC2, which were associated with GO terms related to somatic embryogenesis, fatty acid biosynthesis, and auxin biosynthesis, including the totipotency-regulating genes (Supplementary Figs. 8 and 9 and Supplementary Data 5). These results indicate that CHH hypermethylation deposited by the DRM2/RdDM pathway is associated with the upregulation of totipotency-regulating genes during somatic embryogenesis.

# The SUVH-SDJ reader complex is essential for somatic embryogenesis

Given that DRM1/2-dependent CHH hypermethylation is required for LEC2-induced somatic embryogenesis and that genes encoding components of the CHH methylation reader complex SUVH-SDJ were upregulated during the process, as shown by RNA-seq analysis (Supplementary Fig. 10a), we asked whether SUVH-SDJ components might have a functional role in somatic embryogenesis. We first confirmed the upregulation of SUVH1, SUVH3, SDJ1, SDJ2, and SDJ3 genes through qRT-PCR (Supplementary Fig. 10b). The promoter of the SUVH1 gene contains two putatively LEC2-binding RY motifs (TGCATG, -143/-148 and + 277/ + 282; Fig. 3a), but no putative binding motif was observed in the promoters of SUVH3, SDJ1, SDJ2, and SDJ3 genes (Supplementary Fig. 10c). Utilizing LEC2-GFP in ChIP-qPCR analysis of pER8-LEC2-GFP seedlings, we detected enrichment of fragments encompassing the -143/-148 RY motif (Fig. 3a). Furthermore, an EMSA and luciferase assay demonstrated that LEC2 could bind to the SUVH1 promoter region and activate its expression (Fig. 3b, c). These results suggest that LEC2 directly activates SUVH1 transcription.

To investigate the function of the SUVH-SDJ complex in somatic embryogenesis, we analyzed mutants with compromised SUVH and SDJ functions. We found that each of the three null mutations of *suvh1/ 3*, *sdj1/2/3*, and *suvh1/3 sdj1/3* significantly suppressed LEC2-induced SE formation (Fig. 3d–g, j, k, Supplementary Fig. 11a, and Supplementary Data 6). We further observed that overexpression of *SUVH1* (*3SSpro::-SUVH1-6Myc*) completely restored and even enhanced LEC2-induced somatic embryogenesis in *pER8-LEC2 suvh1/3* plants (Fig. 3d, e, h, j, k, Supplementary Fig. 11a, and Supplementary Data 6). This was not the case with a mutant version of SUVH1 (SUVH1<sup>Y277A</sup>) that cannot bind to CHH methylation<sup>27</sup> (Fig. 3d, e, i–k, Supplementary Fig. 11a, and



Supplementary Data 6). These results suggest that the CHH reader complex SUVH-SDJ is required for LEC2-induced somatic embryogenesis.

The binding of SUVH in a complex with SDJ to CHH methylationrich regions in gene promoters has been associated with activating transcription<sup>27,28</sup>. In agreement with this notion, we found that the upregulation of totipotency-regulating genes by LEC2 was decreased by mutations compromising SUVH-SDJ functions but increased by *SUVH1* overexpression even in the *suvh1/3* mutants (Supplementary Fig. 12). We also showed by ChIP-qPCR that during LEC2-induced SE formation, SUVH1-6Myc was enriched at the CHH hypermethylated regions of totipotency-regulating genes, including the embryonic regulators (*LEC1, ABI3, FUS3,* and *BBM*), auxin biosynthesis genes (*YUC1, YUC5,* and *YUC10*), and fatty acid biosynthesis gene (*KASI*)

**Fig. 2** | **Totipotency-regulating genes exhibit increased expression levels and CHH hypermethylation during LEC2-induced somatic embryogenesis. a** The average cytosine methylation levels across whole genes with 5 kb flanking regions in CG, CHG, and CHH contexts. TSS, transcription start site; TTS, transcription termination site. **b** The number of differentially methylated regions (DMRs) in CG, CHG, and CHH methylation contexts was determined for the indicated comparisons. **c** Genomic features overlapping with the DRM1/2-dependent CHH hyper-DMRs that are hypermethylated in *pER8-LEC2* vs. Mock and hypomethylated in *pER8-LEC2 drm1/2* vs. *pER8-LEC2*. **d** Venn diagram showing genes with DRM1/2-

(Fig. 3l and Supplementary Fig. 11b). By contrast, the non-CHH-binding SUVHI<sup>Y277A</sup>-6Myc version did not exhibit either an association with the promoters of these genes or the enhancement of their expression in the *suvh1/3* mutants (Fig. 3l and Supplementary Figs. 11b and 12). The results suggest that the SUVH-SDJ complex binds to the CHH hypermethylated regions of the totipotency-regulating genes, resulting in their transcriptional upregulation during SE formation.

### The SUVH-SDJ complex acts as a co-activator of LEC2

Given that SUVH1 was enriched at the CHH hypermethylated regions of totipotency-regulating genes (Fig. 3l), we further observed that SUVH1 peaks were co-located with CHH hypermethylated regions at the promoters of totipotency-regulating genes, including ABI3, FUS3, BBM, and YUC5 (Fig. 4a). We also discovered that many upregulated genes associated with DRM1/2-dependent CHH hyper-DMRs were LEC2 target genes, which were related to lipid storage and somatic embryogenesis (Supplementary Fig. 13a, b and Supplementary Data 7). This suggests that LEC2 may cooperate with SUVH1 to regulate somatic embryogenesis. In addition, the downregulated LEC2 target genes associated with DRM1/2-dependent CHH hyper-DMRs were primarily involved in GO terms related to secondary metabolic process and photosynthesis process (Supplementary Fig. 13c, d and Supplementary Data 7). We further observed that the peaks of SUVH1 and LEC2 were both located at the promoters of totipotency-regulating genes, including ABI3, FUS3, BBM, and YUC5 (Fig. 4a and Supplementary Fig. 13e). Based on these findings, we hypothesize that LEC2 and SUVH1 might cooperate to upregulate promoter activity of totipotencyregulating genes during somatic embryogenesis. We then asked whether LEC2 and SUVH1 might physically interact. Indeed, we observed LEC2 and SUVH1 protein-protein interaction by in vitro pull-down, Co-IP in Arabidopsis protoplasts, and bimolecular fluorescence complementation (BiFC) assays in tobacco leaves (Fig. 4b-d). We also confirmed this interaction by Co-IP in transgenic plants co-expressing pER8-LEC2-GFP and gSUVH1-6Myc (Fig. 4e). To address whether LEC2 also interacts with other components of the SUVH-SDJ complex, we utilized pull-down and BiFC assays. The analysis confirmed that LEC2 directly interacted with SUVH3 (Fig. 4f, g). Furthermore, LEC2 interacted with SDJ1, and this interaction required the presence of SUVH1 (Fig. 4h, i). Thus, LEC2 physically interacts with the SUVH1/3 component of the SUVH-SDJ complex.

We then addressed the potential role of the physical interaction between SUVH1 and LEC2. The results demonstrated that SUVH1 can enhance LEC2's binding to its target genes, whereas the mutated SUVH1<sup>Y277A</sup> cannot (Fig. 5a and Supplementary Fig. 11c). Furthermore, co-overexpressing LEC2, SUVH1, and SDJ1 significantly increased transcription from *LEC1*, *ABI3*, *FUS3*, *BBM*, and *YUC5* transcriptional reporters in luciferase assays (Fig. 5b). By contrast, expressing only LEC2, SUVH1, or LEC2 and SUVH1 was not or less effective (Fig. 5b). In addition, using a series of truncated versions of SUVH1 for pull-down assays, we identified an AT-rich domain of the SUVH1 protein that is both necessary and sufficient for the interaction with LEC2 (Fig. 4j and Supplementary Fig. 14). Importantly, we find that deleting the AT-rich domain (SUVH1-ΔAT) abolished the ability of SUVH1 to enhance the LEC2-mediated transcriptional activation of its target genes (Fig. 5c). dependent CHH DMR between the indicated treatment and genotypes. *P*-value indicates statistical significance by hypergeometric tests. **e** GO term analysis of the 678 upregulated genes with CHH hypermethylation. The selected 10 enriched GO biological processes are indicated. The -log<sub>10</sub>(*P*-value) is given. *P*-values were calculated using hypergeometric tests in the clusterProfiler package. **f** Upper panel: heatmap showing the expression levels of totipotency-regulating genes in Mock and *pER8-LEC2* at 72, 84, and 96 HAI. Lower panel: heatmap showing the CHH methylation levels of DMRs at totipotency-regulating genes in Mock, *pER8-LEC2*, and *pER8-LEC2 at m1/2* at 96 HAI.

To investigate the role of CHH hypermethylation in gene promoters for transcriptional activity, we performed luciferase assays with deletions in the CHH hypermethylated regions ( $\Delta$ mCHH) of totipotency-regulating genes, including *LEC1*, *ABI3*, *FUS3*, *BBM*, and *YUC5* (Supplementary Fig. 15). This resulted in a significantly reduced transcription upon co-expressing LEC2, SUVH1, and SDJ1, compared to the full-length promoters (Fig. 5c), indicating that CHH hypermethylation is necessary for SUVH-SDJ to enhance LEC2 activation of its target genes. Collectively, these results suggest that the SUVH1-SDJ1 complex acts as a co-activator of LEC2 during the activation of totipotency-regulating genes.

# SUVH-SDJ complex recruits AHLs to increase chromatin accessibility

DNA methylation has been linked to chromatin accessibility<sup>42</sup>. To examine the mechanism by which the SUVH-SDJ complex enhances the transcription of totipotency-regulating genes, we conducted ATAC-seq in *pER8-LEC2, pER8-LEC2 suvh1/3* and Mock cotyledons at 96 HAI (Fig. 6a and Supplementary Fig. 16a). We found that 507 LEC2 target genes showed increased chromatin accessibility after LEC2 induction, and this effect was significantly suppressed by *suvh1/3* mutation (Figs. 4a and 6a, Supplementary Fig. 16b, c, and Supplementary Data 8). The corresponding gene functions were enriched for the GO terms regeneration, meristem growth, somatic embryogenesis, and cell population proliferation, including the totipotency-regulating genes *LEC1, ABI3, FUS3, BBM,* and *YUC5* (Fig. 6a, Supplementary Fig. 16d, and Supplementary Data 8). These results suggest that SUVH1/3 are required for LEC2-induced chromatin opening of the totipotency-regulating genes.

DNA methylation regulates transcription through reader proteins, which recruit chromatin complexes, altering the chromatin state and causing transcriptional activation or repression<sup>43,44</sup>. In Arabidopsis, the SUVH family protein SUVH9 interacts with the DNAJ domain protein ADMETOS (ADM) and AHL10, forming a complex that regulates transcription via the deposition of H3K9me2 during reproductive isolation<sup>45</sup>. In addition, it has been reported that overexpression of the *AHL* genes (*AHL15, AHL19, AHL20,* and *AHL29*) can induce somatic embryogenesis from immature zygotic embryos without the requirement of exogenous hormones and that overexpression of *AHL15* promotes chromatin decondensation in this process<sup>38</sup>. Therefore, we addressed whether SUVH-SDJ increased chromatin accessibility through AHLs. We found that AHL15 and all other AHLs tested directly interacted with SDJ1 protein but not with SUVH1 or LEC2 (Fig. 6b–d and Supplementary Fig. 17).

To investigate whether AHLs are required for LEC2-induced SE formation, we introduced the inducible *pER8-LEC2* transgene into *ahl15, ahl15/19,* and *ahl29* mutants, respectively (Supplementary Fig. 18). We found that *ahl15, ahl15/19,* and, to a lesser extent, *ahl29* significantly reduced the efficiency of somatic embryogenesis (Fig. 6e–j and Supplementary Data 9). These results indicate that AHL15 plays a predominant role in LEC2-induced somatic embryogenesis.

To explore whether AHL15 affects chromatin accessibility during SE formation, we conducted ATAC-seq analysis between *pER8-LEC2* and *pER8-LEC2 ahl15* cotyledons at 96 HAI. Indeed, we found a



significant reduction in chromatin accessibility in *pER8-LEC2 ahl15* compared to *pER8-LEC2* (Figs. 4a and 6a and Supplementary Fig. 16b). When comparing chromatin accessibility across *pER8-LEC2*, *pER8-LEC2 suvh1/3*, and *pER8-LEC2 ahl15*, we identified 639 genes with LEC2-induced chromatin accessibility that was reduced in *suvh1/3* and *ahl15* backgrounds (Supplementary Fig. 16e and Supplementary Data 8). These genes were overrepresented in the GO-terms somatic

embryogenesis, auxin biosynthesis, and meristem maintenance (Supplementary Fig. 16f and Supplementary Data 8). Importantly, SUVH-SDJ-AHL-dependent chromatin opening was also observed around the putative LEC2-binding regions in the totipotency-regulating genes (Fig. 4a).

Open chromatin domains have been associated with increased histone 3 lysine 9 acetylation (H3K9ac) and H3K14ac<sup>46</sup>. Furthermore,

**Fig. 3** | **The SUVH-SDJ complex binding to the CHH hypermethylated regions of totipotency-regulating genes is required for LEC2-induced SE formation. a** Upper panel: schematic representation of the 400 bp region upstream of the ATG codon (+ 1) of the *SUVHI* gene. PCR fragments labeled "a" to "c" were utilized for the ChIP-qPCR analysis, with "c" serving as a negative control. Lower panel: ChIP-qPCR of LEC2-GFP showing enrichment of the "a" region of the *SUVHI* promoter at 96 HAI. *P*-values are shown, two-tailed Student's *t* test. **b** EMSA of GST-LEC2 and the *SUVHI* promoter region containing the LEC2 binding site. Four replicates showed similar results. The black arrowheads indicate the position of the shifted bands, non-specific bands, and the free probe, respectively. **c** Relative ratio of firefly LUC to REN activity in tobacco leaves. *35Spro::LEC2* was used as effector and *SUVHIpro::-LUC* as reporter. **d–i** Phenotypes of seedlings of the indicated genotypes at 14 DAI. White arrowheads in **d. e.** and **h** point to individual SEs. Scale bars. 0.5 mm. **i. k** The

the histone deacetylase inhibitor Trichostatin A (TSA) results in chromatin decondensation during somatic embryogenesis<sup>38</sup>, indicating that increased chromatin accessibility can be achieved by enhancing the level of histone acetylation. Therefore, we tested whether this was true for the LEC2-induced open chromatin. The results showed increased H3K9ac and H3K14ac levels at the LEC2 binding regions of several totipotency-regulating genes at 96 HAI after LEC2 induction (Fig. 6k–p). Importantly, this increase and the binding of LEC2 to the totipotency-regulating genes were suppressed by the *ahl15* mutation.

These results indicate that the SUVH-SDJ-AHL complex increases H3K9ac and H3K14ac levels and chromatin accessibility and enhances the activation of totipotency-regulating genes during LEC2-induced somatic embryogenesis.

### Discussion

The transcription factor LEC2 regulates various aspects of zygotic embryogenesis<sup>47</sup>. During the early morphogenesis phase, when the basic body plan of the embryo is established, loss-of-function mutations in LEC2 disrupt the maintenance of embryonic cell fate and the specification of cotyledon identity. Later in embryogenesis, lec2 mutants display cotyledon tips that fail to accumulate storage reserves or acquire desiccation tolerance<sup>47</sup>. LEC2 directly regulates several processes related to embryo maturation and auxin signaling through binding to the RY motifs in the promoters of genes, such as ENHANCED EM LEVEL (EEL) and INDOLE-3-ACETIC ACID INDUCIBLE 30 (IAA30)<sup>48</sup>. During somatic embryogenesis, LEC2 can regulate the reprogramming of differentiated plant cells to expand their developmental potential. However, the molecular mechanisms underlying these events remain poorly understood. Here, we show that LEC2-induced CHH hypermethylation is essential for the upregulation of totipotency-regulating genes, promoting somatic embryogenesis. LEC2 appears to perform this function by directly activating the transcription of totipotencyregulating genes on the one hand, and by promoting CHH hypermethylation to establish transcription-competent chromatin, resulting in the transcriptional activation of cell totipotency-regulating genes on the other hand (Fig. 7).

Plants use DNA methylation to regulate gene expression in various developmental processes, such as embryogenesis, pollen development, trichome differentiation, and flower development<sup>49,50</sup>. Studies have shown that DNA methylation plays an important role in zygotic embryogenesis. Mutations in the DNA methyltransferase MET1 result in a series of embryonic developmental defects<sup>31,49</sup>. Furthermore, mutations in DRM2 lead to early embryonic morphological abnormalities and defects in the division plane associated with the global loss of CHH methylation in the egg cell<sup>49,51</sup>. In our study, the functional loss of *DRM1/2* inhibited somatic embryo formation, indicating that DRM1/2-mediated CHH methylation plays a conserved role in both zygotic and somatic embryogenesis. In addition, studies showed that DNA methylation is increased during somatic embryogenesis<sup>32,33</sup>. SUVH1/3, as DNA methylation readers, can recognize and bind to CHH methylation and recruit the chaperone SDJ1/2/3 to enhance gene

statistical analysis of the frequencies of somatic embryogenesis calculated as the proportion of explants with at least one somatic embryo (SE) relative to all explants (**j**) and the average number of SEs per explant (**k**). n, number of transgenic lines analyzed. **I** ChIP-qPCR of SUVH1-6Myc showing enrichment of CHH methylation regions at the indicated gene promoters at 96 HAI. mCHH represents CHH hypermethylated regions. The data in **a**, **c**, **j**, **I** are presented as means  $\pm$  SD, with (**a**, **c**, **I**) representing results from three biological replicates. For box plots in **k**, the horizontal line represents the median value, the boxes represent the interquartile range (25th to 75th percentiles), and the whiskers extend to the maximum and minimum values. Dots indicate individual values. Different letters in (**c**, **j**-**I**) indicate significant differences (*P* < 0.05), one-way ANOVA with Tukey's multiple comparison test. Source data are provided as a Source Data file.

transcription through currently unknown mechanisms<sup>27,28</sup>. While recent studies have shown that DNA methylation is related to chromatin accessibility<sup>42</sup>, the underlying mechanism was unclear. In this study, we discover that the SUVH-SDJ reader complex can interact with AHLs, increasing chromatin accessibility at the promoters of totipotency-regulating genes, thereby activating these genes (Fig. 6 and Supplementary Figs. 16 and 17).

AHLs belong to the High Mobility Group A (HMGA) family proteins. They are chromatin modification proteins involved in various cellular processes, including DNA replication and repair, gene transcription of cell growth, cell differentiation, and cell transformation<sup>52</sup>. Studies have shown that AHL-mediated chromatin decondensation facilitates the transformation of differentiated cells into totipotent cells during Arabidopsis somatic embryogenesis<sup>38</sup>. Our results show that AHL enhances histone acetylation H3K9ac and H3K14ac levels and chromatin accessibility of the LEC2-activated gene regions (Fig. 6k–p). In mammals, the AHL homolog HMGA2 interacts with the histone acetyltransferase General control non-depressible 5 (GCN5) to induce chromatin conformational remodeling and promote the expression of *Matrix metalloproteinase 2 (MMP2)*<sup>53</sup>. Future studies will address whether AHL15 employs a similar mechanism during LEC2-induced somatic embryogenesis.

Our results show that the SUVH-SDJ-AHL complex, which is recruited to CHH sites, directly interacts with LEC2, enhancing LEC2mediated activation of totipotency-regulating genes (Fig. 7). This demonstrates the coordinated regulation of transcription factors and epigenetic factors during cell fate transition during somatic embryogenesis and highlights that the precise regulation of chromatin decondensation and the binding of key totipotency-regulating transcription factors at the promoters of totipotency-regulating genes are intricately intertwined.

Wang et al. propose that the chromatin states, auxin, and totipotency-regulating transcription factors represent three tiers of a hierarchical mechanism for cell reprogramming during somatic embryogenesis<sup>14</sup>. They propose that establishing transcriptionally competent chromatin of the totipotency-regulating transcription factors such as LEC2, LEC1, and BBM is essential to endow somatic cells with regenerative competence<sup>14</sup>. And that auxin can also induce the expression of these transcription factors<sup>14</sup>. Changes in chromatin state can significantly affect gene expression and ultimately lead to changes in cell fate<sup>54,55</sup>. Compared to the dense chromatin state (heterochromatin) of somatic cells, most pluripotent stem cells exhibit a more open and relaxed chromatin state (euchromatin), which allows for more rapid transcriptional changes<sup>54,55</sup>. Somatic cell reprogramming requires overcoming the difference between these two states. In Arabidopsis, the loss of cellular embryogenic competence during seed germination is likely linked to removing permissive chromatin marks at loci associated with totipotency-regulating transcription factors, such as LEC1 and LEC2<sup>14,56</sup>. This developmentally controlled process blocks the reactivation of totipotency-regulating transcription factor expression, making it more difficult for post-embryonic explants to





Glutathione Sepharose show that His-LEC2 is pulled down. Three biological repeats were analyzed, all showing similar results. **g** BiFC assays show that SUVH3 interacts with LEC2 (green nuclear signal, YFP). The blue fluorescent signal represents the nuclei stained with DAPI. Scale bars, 40 μm. Two biological repeats were performed with similar results. **h** Co-IP assays show that SDJ1 interacts with LEC2 in the presence of SUVH1. To conduct these assays, Arabidopsis protoplasts from the *suvh1/3* and *3SSpro::SUVH1-6Myc suvh1/3* lines were transformed with *3SSpro::SUVH1-6Myc suvh1/3* lines were transformed with *3SSpro::SUVH1-6Myc suvh1/3* lines were analyzed, and both showed similar results. **i** Pull-down assays showing the interaction between SDJ1 and LEC2 in the presence of SUVH1. Two biological repeats were performed with similar results. **j** Pull-down assays show that LEC2 interacts with SUVH1, but not with the SUVH1-ΔAT variant. Two biological repeats were performed with similar results.



**Fig. 5** | **The SUVH-SDJ complex enhances the binding and activation capacity of LEC2. a** Upper panel: the diagram depicts the promoter regions of the indicated genes. Lower panel: ChIP-qPCR assays show that SUVH1 enhances the ability of LEC2 to bind to the promoter of the indicated genes at 96 HAI. The data are presented as means ± SD from three biological replicates. *P*-values are shown, twotailed Student's *t* test. **b**, **c** Relative ratio of firefly LUC to REN activity in Arabidopsis

protoplasts. The effectors and reporters are indicated. Relative LUC activity was normalized to that of the control. The data are presented as means  $\pm$  SD. n, the number of replicates. Different letters indicate significant differences (*P* < 0.05), one-way ANOVA with Tukey's multiple comparison test. Source data are provided as a Source Data file.

recover their capacity for somatic embryogenesis<sup>56</sup>. These findings provide insight into why immature zygotic embryos are often used as explants for somatic embryogenesis in Arabidopsis<sup>56</sup>. Totipotency-regulating transcription factors function independently and cooperate with other mechanisms (such as DNA methylation, histone

modifications, and TEs) to exert their effects<sup>54</sup>. Our results here add chromatin dynamics to the regulating network controlling the transcriptional output of totipotency-regulating transcription factors during somatic embryogenesis. LEC2 directly activates the expression of *DRM2*, maintaining high CHH methylation levels to recruit the



SUVH-SDJ-AHL complex, which results in chromatin decondensation and high expression levels of totipotency-regulating genes (Fig. 7).

Somatic embryogenesis enables the production of transgenic plants in various economically important species, such as maize<sup>57</sup>, soybean<sup>58</sup>, and sweet pepper<sup>59</sup>. Several observations suggest that our findings in Arabidopsis may be relevant to other species: Studies have shown that CHH hypermethylation occurs during somatic embryogenesis in several species, including globular somatic embryos of soybean<sup>32</sup> and longan<sup>60</sup> and embryogenic calli of cotton<sup>33</sup>. In addition, CHH hypermethylation has been observed during zygotic embryogenesis in Arabidopsis<sup>61</sup>, soybean<sup>62</sup>, and *Brassica rapa*<sup>63</sup>. These findings suggest that CHH hypermethylation may be a hallmark of embryogenesis. It is worth investigating whether the mechanism by which the SUVH-SDJ-AHL complex promotes totipotency-regulating Fig. 6 | Mutations in AHLs reduce chromatin accessibility and inhibit SE

induction. **a** The volcano plot showing the genes associated with decreased (blue) or increased (red) accessible-chromatin peak regions of the indicated samples. The names of known totipotency-regulating genes are shown. Ns (gray), no significant difference. Statistical significance was determined using Benjamini-Hochberg false discovery rate (FDR) correction via the DiffBind package. **b** Pull-down assays show that AHL15 interacts with SDJ1, not with SUVH1 or LEC2. Three biological repeats showed similar results. **c** Co-IP analysis of the interaction between SDJ1 and AHL15 in Arabidopsis protoplasts co-transformed with *3SSpro::AHL15-3HA-3Flag* and *3SSpro::SDJ1-GFP* or *3SSpro::GFP-6Myc*. *3SSpro::GFP-6Myc* served as a control. Three biological repeats were analyzed and showed similar results. **d** BiFC assays in tobacco leaf epidermal cells show that AHL15 interacts with SDJ1 (green nuclear signal, YFP). The blue fluorescent signal represents the nuclei stained with DAPI. Scale bars, 20 μm. Two biological repeats were performed with similar results. **e**-**h** Phenotypes of seedlings of the indicated genotypes at 14 DAI. White

gene expression is conserved in somatic embryogenesis of crop species, especially considering species where hyper-CHH methylation has been observed during zygotic embryogenesis.

Ectopic expression of the pluripotency-associated transcription factors Octamer-binding protein 4 (Oct4), Sex-determining region Y-box 2 (Sox2), Krüppel-like factor 4 (Klf4), and c-Myc (collectively termed OSKM), can reprogram differentiated cells in human or mouse, such as skin fibroblasts, to an embryonic stem cell (ESC)-like state, yielding induced pluripotent stem cells (iPSCs)<sup>64</sup>. Some additional regulators, such as epigenetic modifiers, play important roles in the induction and maintenance of the pluripotent state<sup>65</sup>. The pluripotency factor Oct4 has been found to physically associate with and regulate the expression of various chromatin regulators<sup>66</sup>. Moreover, during iPSC generation from mouse and human somatic cells, externally applied epigenome-modifying molecules can substitute for one or more OSKM reprogramming factors<sup>67</sup>. Our results suggest that the role of epigenetic modifiers may also be applied to the regeneration processes in plants. Perhaps the plant species with poor regeneration efficiency may have lost the ability to reprogram chromatin to enable totipotency-associated gene expression. Even in the presence of totipotency-regulating transcription factors, reprogramming somatic cells in these species remains challenging. Our results reported here suggest that despite fundamental differences, the role of epigenetic regulation may be central to regulating totipotency in plants and animals.

### Methods

### Plant materials and growth conditions

All plant materials were in *Arabidopsis thaliana* Columbia (Col-0) accession background. The T-DNA insertion line for *ahl29* (SALK\_151047C) was obtained from AraShare (a non-profit Arabidopsis share center, http://www.arashare.cn). The lines *drm1/2* (CS16383)<sup>68</sup>, *suvh1/3*<sup>28</sup>, *sdj1/2/3*<sup>28</sup>, *ahl15* (SALK\_040729)<sup>38</sup>, and *ahl15/19*<sup>38</sup> have been described previously. Primers used for genotyping are described in Supplementary Data 10.

Arabidopsis seeds were surface sterilized with 70% ethanol for 5 min, 3% sodium hypochlorite for 10 minutes, and sterile water for six times. These seeds were then grown on half-strength Murashige and Skoog (1/2 MS) media (Hopebio, HB8469-5) containing 0.8% (w/v) agar and 1% (w/v) sucrose for 8 days and then transplanted into soil. Plants were grown under a 16-hour light/8 h dark cycle at 22 °C.

### Plasmid construction and plant transformation

To generate the *pER8-LEC2* and *pER8-LEC2-GFP* lines, the *LEC2* coding sequence fragment or *LEC2* coding sequence and *GFP* coding sequence fragments were amplified using Phanta Max Master Mix (Vazyme, P515-01) and cloned into the  $\beta$ -estradiol-inducible XVE<sup>19</sup> binary vector by ClonExpress II One Step Cloning Kit (Vazyme, C112-01), respectively. For the *gDRM2-GFP* construct, the 5.2 kb genomic region<sup>30</sup>, including

arrowheads in **e**, **f**, and **h** point to individual SEs. Scale bars, 0.5 mm. **i**, **j** The statistical analysis of frequencies of somatic embryogenesis (**i**) and the average number of SEs per explant (**j**). n, number of transgenic lines analyzed. **k–o** Upper panel: the diagram of the promoter regions of the indicated genes. Middle panel and lower panel: ChIP-qPCR assays showing the levels of H3K9ac (middle panel) and H3K14ac (lower panel) at the LEC2-binding regions of the indicated genes at 96 HAI. **p** ChIP-qPCR assays show that AHL15 mutation reduces the binding of LEC2-GFP to the indicated promoters at 96 HAI. The data in **i**, **k–p** are presented as means ± SD, with (**k–p**) representing results from three biological replicates. For box plots in **j**, the horizontal line represents the median value, the boxes represent the interquartile range (25th to 75th percentiles), and the whiskers extend to the maximum and minimum values. Dots indicate individual values. Different letters in **i–p** indicate significant differences (*P* < 0.05), one-way ANOVA with Tukey's multiple comparison test. Source data are provided as a Source Data file.

the promoter and the gene body, was amplified from Col-0 and cloned into the *106HGFP* vector. To construct *35Spro::DRM2*, *35Spro::SUVH1-6Myc*, and *gSUVH1-6Myc*, the coding sequences of *DRM2* and *SUVH1*, and a 3.3 kb genomic region of *SUVH1*, including the promoter and gene body, were amplified from Col-0 and cloned into the *pSuper1300-Myc* plasmid, respectively. To generate the *35Spro::SUVH1*<sup>Y277A</sup>-*6Myc* construct, the mutated SUVH1 sequence was amplified using the Fast Site-Directed Mutagenesis Kit (Tiangen, KM101) and cloned into *pSuper1300-Myc*.

All constructs were transduced into *Agrobacterium tumefaciens* strain GV3101, which was grown at 28 °C for 2 days and then transformed into the wild type or related mutants using the floral dip method. This study identified homozygous plants carrying the transgene (T2 generation). The primers used in this study are listed in Supplementary Data 10.

### Chemical induction of somatic embryogenesis

For in vitro plant culture, the seeds of *pER8-LEC2* in wild type or mutant background were cultured on 1/2 MS supplemented with 10  $\mu$ M  $\beta$ estradiol (Sigma, E2758) dissolved in DMSO (Sangon Biotech, A1000231) for 2 weeks. The seeds of *pER8-LEC2* cultured on 1/2 MS with an equal volume of DMSO were served as a control (Mock). After 2 weeks of culture, the capacity to induce somatic embryogenesis was scored under a stereomicroscope by scoring the frequencies of somatic embryogenesis (the explants with at least one somatic embryo (SE)/all explants) and the average number of SEs per explant. At least 50 explants were cultured for each test, and at least 8 transgenic lines were analyzed for each material for statistical analysis. One-way ANOVA with Tukey's multiple comparison test was used for statistical analysis of the data. Embryonic lipids were visualized by 30 min staining with Sudan Red 7B (Sigma, 46290).

### RNA-seq experiment and data analysis

Total RNA was extracted from the cotyledons of Mock and *pER8-LEC2* at 72, 84, and 96 HAI and then sent to Shanghai OE Biotech Company (China) for library preparation and deep sequencing. The clean reads were aligned to the TAIR10 Arabidopsis genome sequence using HISAT2<sup>69</sup> (v2.2.1) and counted using FeatureCounts<sup>70</sup> (v2.0.3) on the exons. The expression levels for TAIR10 gene models were measured and normalized as fragments per kilobase of transcript per million mapped reads (FPKM). Differentially expressed genes (DEGs) (P < 0.05 and  $\log_2$ [fold change |  $\geq$  1) were chosen by the DEseq2<sup>71</sup> (v1.40.2) package in R.

### Imaging

Live seedlings were imaged by an Olympus SZX16 dissecting microscope. Confocal microscopy images were achieved using a Zeiss LSM 880 NLO confocal microscope. A 488 nm laser was used as the excitation light for GFP signals, and emissions were collected at



Fig. 7 | CHH hypermethylation mediates LEC2-induced somatic embryogenesis through the SUVH-SDJ-AHL complex. Overexpression of LEC2 triggers the upregulation of genes associated with the DRM2-DDR complex, wherein the DDR complex comprises DRD1, DMS3, and RDM1. This activation occurs within the RdDM pathway, leading to maintaining CHH hypermethylation at the promoters of totipotency-regulating genes during somatic embryogenesis. In addition, LEC2 can stimulate the expression of genes that encode the DNA methylation reader

complex SUVH-SDJ. This complex recruits AHL chromatin modification proteins, which play a crucial role in directly or indirectly interacting with histone acetyl-transferases (HATs)<sup>53</sup>. This interaction promotes histone acetylation, thereby increasing chromatin accessibility at the promoters of totipotency-regulating genes. Furthermore, the SUVH-SDJ-AHL complex interacts with LEC2, reinforcing LEC2-mediated activation of totipotency-regulating genes and somatic embryogenesis.

505–550 nm. A 514 nm laser was used as the excitation light to detect YFP signals, and emissions were collected at 530–600 nm. 4',6-diamidino-2-phenylindole (DAPI) signals were detected using a 405 nm laser as the excitation light, and emissions were collected at 410–490 nm. The method has been described previously<sup>19</sup>.

### Total RNA isolation and qRT-PCR analysis

Total RNA was extracted using an RNA isolation kit (Cwbio, CW0581M) according to the manufacturer's protocol, and three biological replicates were performed in each experiment. Reverse transcription was performed using a cDNA synthesis kit (Transgen, AE311). The resulting cDNA sequences were used as a template for quantitative real-time PCR (qRT-PCR). qRT-PCR was performed using TransStart® Green qPCR SuperMix (Transgen, AQ101-01) on a LightCycler 96 instrument (Roche). Data were calculated by the Ct method and averaged over three biological replicates. *UBQ10* was used as an internal control. The primers used for qRT-PCR reactions are listed in Supplementary Data 10.

# Chromatin immunoprecipitation followed by qPCR analysis (ChIP-qPCR)

Seedlings of transgenic lines at 96 HAI, including DMSO-treated 35Spro::GFP, 35Spro::6Myc, pER8-LEC2-GFP 35Spro::SUVH1-6Myc suvh1/ 3, and pER8-LEC2-GFP 35Spro::SUVH1<sup>Y277A</sup>-6Myc suvh1/3, and β-estradioltreated pER8-LEC2-GFP, pER8-LEC2-GFP suvh1/3, pER8-LEC2-GFP 35Spro::SUVH1-6Myc suvh1/3, and pER8-LEC2-GFP 35Spro::SUVH1<sup>Y277A</sup>- *6Myc suvh1/3* were harvested for fixation with 1% formaldehyde. Nuclei were isolated, and chromatin DNA was sonicated to an average size of approximately 250 bp. The solubilized chromatins were immunoprecipitated (IP) by Protein A/G Magnetic Beads (Selleck, B23202) with either anti-GFP antibody (Sigma, G1544) or anti-Myc antibody (Sigma, M4439) as IP, and the elution product was recovered at 65 °C for 5 h. The recovered DNA was purified using the ChIP DNA Clean & Concentrator<sup>™</sup> (Zymoresearch, D5205) and subjected to qPCR with SYBR qPCR Mix (Vazyme, Q711-O2). The relative fold enrichment was determined by normalizing the target DNA fragment against the input DNA, followed by normalization against *35Spro::GFP* or *35Spro::6Myc* as a negative control.

### EMSA

The GST-LEC2 protein was purified as described for the pull-down assays. WT and mutant double-stranded oligonucleotides were commercially synthesized and labeled with biotin at the 5' end. A Light Shift Chemiluminescent EMSA kit (Thermofisher, 21048) was used for the binding reaction. All primers used for probes and competitors are listed in Supplementary Data 10.

### Transient dual luciferase assays (LUC)

To generate the LUC reporters, we amplified the following promoters from Col-0: 2 kb of *DRM2*, 1.3 kb of *SUVH1*, 4 kb of *ABI3*, 4.1 kb of *BBM*, 2 kb of *LEC1*, 5 kb of *FUS3*, and 4.7 kb of *YUC5*. These amplified promoter regions were then ligated into the *pGreenII 0800-LUC* vector. In

addition, the promoters of these genes, from which the CHH hypermethylated regions were deleted, were cloned into *pGreenll 0800-LUC* vector. The *LEC2*, *SUVH1*, and *SDJ1* gene coding regions were fused to a *pGreenll 62-SK* vector downstream of the *35S* promoter, respectively. Transient transactivation assays were performed using *Nicotiana benthamiana* leaves or Arabidopsis protoplasts. The tobacco leaves and Arabidopsis protoplasts were lysed and detected by the Dual-Luciferase Reporter Gene Assay Kit (Yeasen, 11402ES60). The empty vectors of *pGreenll 0800-LUC* and *pGreenll 62-SK* were used as controls. The primers are listed in Supplementary Data 10.

# Whole-genome bisulfite sequencing (WGBS), bisulfite-sanger sequencing, and data analysis

For WGBS, 1 g 96 HAI cotyledons of Mock, pER8-LEC2, and pER8-LEC2 *drm1/2*, treated with either DMSO or  $\beta$ -estradiol, were sent to Shanghai OE Biotech Company (China) for bisulfite treatment, library preparation, and deep sequencing. The reads were aligned to the TAIR10 Arabidopsis genome for data analysis using Bismark2<sup>72</sup> (v0.24.1). The methylation level of each cytosine was represented as the percentage of reads reporting methylated cytosine (mC) relative to the total number of reads reporting cytosine (C) and thymine  $(T)^{72}$ . The methylation levels at genes with 5 kb flanking regions were calculated by deepTools<sup>73</sup> (v3.5.4). Differentially methylated regions (DMRs) were identified using DMRfinder<sup>74</sup> (v0.3) with the following criteria: a minimum difference in absolute methylation level of 0.1 for CHH, 0.2 for CHG, and 0.4 for CG, a false discovery rate (FDR) < 0.05, and a minimum of three methylation sites in at least two replicates of each sample. The overlapping of DMRs was performed in R. ChIPpeakAnno<sup>75</sup>(v3.6.5). The methylation levels of DMRs in all samples were obtained using ViewBS<sup>76</sup> (v0.1.10).

For observation to expectation (O/E) ratios, the same number of regions was randomly selected from the genome, with the same distribution of the real DMRs (chromosome and length). The value of observation was defined as the real DMRs overlapping with genomic features, while the expectation value was defined as the random regions overlapping with genomic features. Then the O/E ratio of DMRs was calculated<sup>77</sup>. The *P*-value was calculated as the proportion of random DMR sets that exhibited more overlap with these regions than the actual DMR set (permutation test).

For bisulfite-sanger sequencing, the genomic DNA was converted using the DNA Bisulfite Conversion Kit (Tiangen, DP215). PCR products were cloned into the pEASY-Blunt3 vector. Then, 20–40 clones were sequenced and analyzed by the online CyMATE software (https:// cymate.org/).

### In vitro pull-down assays

For pull-down assays, the coding sequence of *LEC2* was cloned into *pGEX-4T-1* and *pET28a* vectors. The coding sequence of *SUVH1* and *SDJ1* was cloned into *pMAL-c2X* and *pGEX-4T-1* vectors, respectively. For the *His-GFP-SUVH1* construct, the coding sequences of *GFP* and *SUVH1* were amplified by overlap extension polymerase chain reaction and assembled into *pET28a*. To generate the *GST-SUVH3* construct, the coding sequence of *SUVH3* was cloned into *pGEX-4T-1*. Moreover, the *AHL15, AHL19, AHL20,* and *AHL29* coding sequences were cloned into *pET28a*. A series of truncated SUVH1 cDNA sequences were cloned into *pGEX-4T-1*. The primers are listed in Supplementary Data 10.

For the interaction of LEC2 with SDJ1 in the presence of SUVH1, GST-LEC2 or GST protein was incubated with Glutathione Sepharose 4 Fast Flow (GE Healthcare, USA; 17-5132-01) in PBSN buffer (PBS buffer containing 0.3% NP-40, 1 × protease inhibitor cocktail, and 5 mM DTT) at 4 °C for 2 h and then washed once with PBSN buffer. Subsequently, His-GFP-SUVH1 protein was added (or not protein added for the control) and incubated in PBSN buffer at 4 °C for 2 h, followed by another wash with PBSN buffer. Finally, MBP-His-SDJ1 was added to each tube

and incubated in PBSN buffer at 4 °C for 2 h. After five washes, the proteins were separated by SDS-PAGE and detected using the following primary antibodies: anti-GST (1:2,000 dilution; Yeasen, 30901ES-50), anti-His (1:2,000 dilution; Abbkine, ABT2050), and anti-MBP (1:2,000 dilution; Yeasen, 31201ES-20) antibodies. Goat anti-mouse IgG (1:5,000 dilution; ZSGB-BIO, ZB-5305) was used as a secondary antibody. The protein blots were visualized using a Chemiluminescence assay kit (Thermofisher, 34095) and imaged using Image Lab software (v.6.0.1, Bio-Rad).

### Co-immunoprecipitation (Co-IP) assays

For Co-IP assays of *LEC2* and *SUVH1* in Arabidopsis, the transgenic lines *pER8-LEC2-GFP gSUVH1-6Myc, pER8-LEC2-GFP*, and *gSUVH1-6Myc* were used. Samples were crushed in liquid nitrogen and suspended in 10 mL lysis buffer (100 mM Tris-HCl pH 8.8, 150 mM NaCl, 1 mM EDTA, 20% glycerol,  $1 \times$  protease inhibitor cocktail (Roche, 04693132001)). After centrifugation, 100 µL supernatant was used as input, and the remainder was incubated with anti-GFP agarose beads (Chromotek, GTA-20) for 3 h. After four washes, the bead-bound proteins and input were run on SDS-PAGE gel and detected with anti-GFP antibody (1:2,000 dilution; Transgen, HT801) and anti-Myc antibody (1:2,000 dilution; Yeasen, 30601ES60).

The coding sequences of LEC2, SUVH1, SUVH3, SDJ1, AHL15, AHL19, AHL20, and AHL29 were cloned into transient expression vectors pSuper1300-Myc, pSuper1300-GFP, or pCAMBIA1307-C-HF to generate 35Spro::LEC2-6Myc, 35Spro::SUVH1-3HA-3Flag, 35Spro::LEC2-3HA-3Flag, 35Spro::SDJ1-GFP, 35Spro::AHL15-3HA-3Flag, 35Spro::AHL19-3HA-3Flag, 35Spro::AHL2O-3HA-3Flag, and 35Spro::AHL29-3HA-3Flag constructs, respectively. For Co-IP assays in Arabidopsis protoplasts, plasmids (20 µg each) were transformed. The transfected protoplasts were harvested, and total proteins were extracted using IP buffer (50 mM pH 8.0 Tris-HCl, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, 10% glycerol, 1% NP-40, 1 mM PMSF, 1 mM DTT, and 1× protease inhibitor cocktail). Immunoprecipitation was performed using anti-Myc (Sigma, A7470) agarose or anti-GFP agarose, followed by five washes with IP buffer. The proteins were eluted from agarose by 1 × SDS loading buffer at 100 °C, followed by western blotting using the anti-HA antibody (1:2,000 dilution; Transgen, HT301), anti-GFP antibody, and anti-Myc antibody.

### Bimolecular fluorescence complementation (BiFC) assays

For the assays, we generated the following constructs: *LEC2-nYFP*, *SUVH1-CYFP*, *SUVH3-CYFP*, *LEC2-cYFP*, *SDJ1-CYFP*, *AHL15-nYFP*, *AHL19-nYFP*, *AHL20-nYFP*, and *AHL29-nYFP* constructs using the *pFGC-YC155* and *pFGC-YN173* vectors. These constructs and the co-transformation vector *35Spro::P19* were transformed into the *Agrobacterium tumefaciens* EHA105. The strains were then incubated, harvested, and resuspended in infiltration buffer (10 mM MES, 0.15 mM acetosyringone, and 10 mM MgCl<sub>2</sub>) at a final concentration of  $OD_{600} = 0.6-0.8$ . Equal volumes of different combinations of strains were mixed and injected into tobacco leaves using a needleless syringe. The leaves were incubated at 24 °C for 72 h to detect YFP fluorescence using a Zeiss 880 laser scanning microscope with 514 nm laser excitation and 530–600 nm emission spectra.

### ATAC-seq and data analysis

For the assays, 0.3 g 96 HAI cotyledons from Mock, *pER8-LEC2*, *pER8-LEC2 suvh1/3*, and *pER8-LEC2 ahl15* with DMSO or  $\beta$ -estradiol treatment were used for nucleus extraction<sup>78</sup>. The extracted nuclei were used for library preparation using TruePrep DNA Library Prep Kit V2 (Vazyme, TD50102) and TruePrep Index Kit V2 for Illumina (Vazyme, TD202) and then sent to Beijing Biomarker Technologies (China) for deep sequencing. The analysis of ATAC-seq was performed as previously described<sup>78</sup>. To identify differentially accessible chromatin regions, we employed DiffBind (v3.12.0) for peak analysis.

### **Bioinformatics analysis**

The ChIP-seq data of NRPE1<sup>40</sup>, LEC2<sup>14</sup>, and SUVH1<sup>27</sup> were obtained from previous studies. The ChIP-seq reads were aligned using Bowtie2 (v2.5.2), and duplicates were removed using SAMtools (v1.18). Peaks were called using MACS<sup>79</sup> (v2 2.2.9.1). The distribution of NRPE1 in CHH DMRs and LEC2 in genes and TE were analyzed using DeepTools<sup>73</sup> (v3.5.4). The Venn diagram was performed using Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/index.html), and Gene Ontology analysis was performed in R. clusterProfiler<sup>80</sup> (v4.8.3). The density of RY motifs was calculated by dividing the number of RY motifs within the promoter region by the length of the promoter.

### Quantification and statistical analysis

Microsoft Excel and GraphPad Prism 8 software were used for statistical analysis. A two-tailed Student's *t* test and one-way ANOVA with Tukey's multiple comparison test were used to calculate significant differences among samples or genotypes. Details regarding biological or technical replicates and statistical parameters used in various experiments are provided in the Methods section and the captions of the Figures and Supplementary Figures, wherever necessary.

### **Reporting summary**

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

# Data availability

The sequence data can be accessed at the Arabidopsis Genome Initiative (https://www.arabidopsis.org/). All analysis tools and associated code used in this study are publicly accessible and described in the Methods section. The RNA-seq, WGBS, and ATAC-seq data generated in this study have been deposited in the Beijing Institute of Genomics Data Center (BIGD) under the accession numbers BioProject PRJCA023840 and PRJCA033386. The publicly available ChIP-seq dataset used in this study can be accessed via the NCBI Gene Expression Omnibus (GEO) under the accession numbers GSE108414<sup>27</sup> and GSE100010<sup>40</sup>, as well as from BIGD under the accession number BioProject PRJCA02620<sup>14</sup>. Source data are provided in this paper.

### References

- Su, Y. H., Tang, L. P., Zhao, X. Y. & Zhang, X. S. Plant cell totipotency: Insights into cellular reprogramming. *J. Integr. Plant Biol.* 63, 228–243 (2020).
- Gambino, G. & Perrone, I. Somatic embryogenesis as a tool for studying grapevine-virus interaction. *Methods Mol. Biol.* 2536, 381–394 (2022).
- Ikeuchi, M. et al. Molecular mechanisms of plant regeneration. Annu. Rev. Plant Biol. **70**, 377–406 (2019).
- Stone, S. L. et al. *LEAFY COTYLEDON2* encodes a B3 domain transcription factor that induces embryo development. *Proc. Natl. Acad. Sci. USA* 98, 11806–11811 (2001).
- Lotan, T. et al. Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195–1205 (1998).
- Boutilier, K. et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14, 1737–1749 (2002).
- Wang, F. X., Shang, G. D. & Wang, J. W. Towards a hierarchical gene regulatory network underlying somatic embryogenesis. *Trends Plant Sci.* 27, 1209–1217 (2022).
- El Ouakfaoui, S. et al. Control of somatic embryogenesis and embryo development by AP2 transcription factors. *Plant Mol. Biol.* 74, 313–326 (2010).

- Gaj, M. D., Zhang, S., Harada, J. J. & Lemaux, P. G. Leafy cotyledon genes are essential for induction of somatic embryogenesis of Arabidopsis. *Planta* 222, 977–988 (2005).
- Gaj, M. D. et al. Hormone-response mutants of Arabidopsis thaliana (L.) Heynh. impaired in somatic embryogenesis. *Plant Growth Regul.* 49, 183–197 (2006).
- Liu, Z. et al. Overexpression of the CsFUS3 gene encoding a B3 transcription factor promotes somatic embryogenesis in Citrus. Plant Sci. 277, 121–131 (2018).
- 12. Chen, B. et al. ABA signalling promotes cell totipotency in the shoot apex of germinating embryos. *J. Exp. Bot.* **72**, 6418–6436 (2021).
- Horstman, A. et al. The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. *Plant Physiol.* **175**, 848–857 (2017).
- Wang, F. X. et al. Chromatin accessibility dynamics and a hierarchical transcriptional regulatory network structure for plant somatic embryogenesis. *Dev. Cell* 54, 742–757 (2020).
- Stone, S. L. et al. Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. Proc. Natl. Acad. Sci. USA 105, 3151–3156 (2008).
- Wojcikowska, B. et al. *LEAFY COTYLEDON2 (LEC2)* promotes embryogenic induction in somatic tissues of Arabidopsis, via *YUCCA*-mediated auxin biosynthesis. *Planta* 238, 425–440 (2013).
- Bai, B., Su, Y. H., Yuan, J. & Zhang, X. S. Induction of somatic embryos in Arabidopsis requires local YUCCA expression mediated by the down-regulation of ethylene biosynthesis. *Mol. Plant* 6, 1247–1260 (2013).
- Karami, O. et al. Endogenous auxin maintains embryonic cell identity and promotes somatic embryo development in Arabidopsis. *Plant J.* **113**, 7–22 (2022).
- 19. Zhang, W. J. et al. A WRI1-dependent module is essential for the accumulation of auxin and lipid in somatic embryogenesis of *Arabidopsis thaliana*. *N. Phytol.* **242**, 1098–1112 (2024).
- Kim, H. U. et al. Senescence-inducible LEC2 enhances triacylglycerol accumulation in leaves without negatively affecting plant growth. *Plant Biotechnol. J.* 13, 1346–1359 (2015).
- 21. Peng, J. et al. The dynamics of chromatin states mediated by epigenetic modifications during somatic cell reprogramming. *Front. Cell Dev. Biol.* **11**, 1097780 (2023).
- 22. Wendte, J. M. & Pikaard, C. S. The RNAs of RNA-directed DNA methylation. *Biochim. Biophys. Acta Gene Regul. Mech.* **1860**, 140–148 (2017).
- Matzke, M. A. & Mosher, R. A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408 (2014).
- 24. Zubko, E., Gentry, M., Kunova, A. & Meyer, P. De novo DNA methylation activity of METHYLTRANSFERASE 1 (MET1) partially restores body methylation in *Arabidopsis thaliana*. *Plant J.* **71**, 1029–1037 (2012).
- Lindroth, A. M. et al. Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* 292, 2077–2080 (2001).
- 26. He, L. et al. Pathway conversion enables a double-lock mechanism to maintain DNA methylation and genome stability. *Proc. Natl. Acad. Sci. USA* **118**, e2107320118 (2021).
- 27. Harris, C. J. et al. A DNA methylation reader complex that enhances gene transcription. *Science* **362**, 1182–1196 (2018).
- Zhao, Q. Q. et al. A methylated-DNA-binding complex required for plant development mediates transcriptional activation of promoter methylated genes. J. Integr. Plant Biol. 61, 120–139 (2019).
- 29. Ichino, L. et al. MBD5 and MBD6 couple DNA methylation to gene silencing through the J-domain protein SILENZIO. *Science* **372**, eabg6130 (2021).

### Article

- 30. Jullien, P. E. et al. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr. Biol.* **22**, 1825–1830 (2012).
- 31. Xiao, W. et al. DNA methylation is critical for Arabidopsis embryogenesis and seed viability. *Plant Cell* **18**, 805–814 (2006).
- Ji, L. et al. Genome-wide reinforcement of DNA methylation occurs during somatic embryogenesis in soybean. *Plant Cell* **31**, 2315–2331 (2019).
- Li, J. et al. Multi-omics analyses reveal epigenomics basis for cotton somatic embryogenesis through successive regeneration acclimation process. *Plant Biotechnol. J.* **17**, 435–450 (2019).
- Wickramasuriya, A. M. & Dunwell, J. M. Global scale transcriptome analysis of Arabidopsis embryogenesis in vitro. *BMC Genom.* 16, 301–301 (2015).
- 35. Keshet, I., Lieman-Hurwitz, J. & Cedar, H. DNA methylation affects the formation of active chromatin. *Cell* **44**, 535–543 (1986).
- Tanaka, M., Kikuchi, A. & Kamada, H. The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiol.* 146, 149–161 (2008).
- Ikeuchi, M. et al. PRC2 represses dedifferentiation of mature somatic cells in Arabidopsis. *Nat. Plants* 1, 15089 (2015).
- Karami, O. et al. An Arabidopsis AT-hook motif nuclear protein mediates somatic embryogenesis and coinciding genome duplication. *Nat. Commun.* 12, 2508 (2021).
- Mosher, R. A. Pinpointing a puzzling polymerase. Nat. Struct. Mol. Biol. 19, 865–866 (2012).
- Liu, W. et al. RNA-directed DNA methylation involves cotranscriptional small-RNA-guided slicing of polymerase V transcripts in Arabidopsis. *Nat. Plants* 4, 181–188 (2018).
- Wang, X. et al. Overexpression of *PGA37/MYB118* and *MYB115* promotes vegetative-to-embryonic transition in Arabidopsis. *Cell Res.* 19, 224–235 (2009).
- 42. Zhong, Z. et al. DNA methylation-linked chromatin accessibility affects genomic architecture in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **118**, e2023347118 (2021).
- 43. Schübeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–326 (2015).
- Grimanelli, D. & Ingouff, M. DNA methylation readers in plants. J. Mol. Biol. 432, 1706–1717 (2020).
- Jiang, H. et al. Ectopic application of the repressive histone modification H3K9me2 establishes post-zygotic reproductive isolation in Arabidopsis thaliana. Genes Dev. **31**, 1272–1287 (2017).
- Chen, Y. C., Koutelou, E. & Dent, S. Y. R. Now open: Evolving insights to the roles of lysine acetylation in chromatin organization and function. *Mol. Cell* 82, 716–727 (2022).
- Liu, B., Sun, G., Liu, C. & Liu, S. LEAFY COTYLEDON 2: A regulatory factor of plant growth and seed development. *Genes* 12, 1896 (2021).
- Braybrook, S. A. et al. Genes directly regulated by LEAFY COTYLE-DON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc. Natl. Acad. Sci. USA* **103**, 3468–3473 (2006).
- Markulin, L. et al. Taking the wheel de novo DNA methylation as a driving force of plant embryonic development. *Front. Plant Sci.* 12, 764999 (2021).
- Kumar, S. & Mohapatra, T. Dynamics of DNA methylation and its functions in plant growth and development. *Front. Plant Sci.* 12, 596236 (2021).
- Ingouff, M. et al. Live-cell analysis of DNA methylation during sexual reproduction in Arabidopsis reveals context and sex-specific dynamics controlled by noncanonical RdDM. *Genes Dev.* **31**, 72–83 (2017).
- 52. Sgarra, R. et al. HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta* **1799**, 37–47 (2010).

- Zhang, S., Zhang, H. & Yu, L. HMGA2 promotes glioma invasion and poor prognosis via a long-range chromatin interaction. *Cancer Med.* 7, 3226–3239 (2018).
- 54. Wang, Y., Bi, Y. & Gao, S. Epigenetic regulation of somatic cell reprogramming. *Curr. Opin. Genet. Dev.* **46**, 156–163 (2017).
- 55. Xu, L. & Huang, H. Genetic and epigenetic controls of plant regeneration. *Curr. Top. Dev. Biol.* **108**, 1–33 (2014).
- Chen, C. et al. Plant regeneration in the new era: from molecular mechanisms to biotechnology applications. *Sci. China Life Sci.* 67, 1338–1367 (2024).
- 57. Lowe, K. et al. Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. *Plant Cell* **28**, 1998–2015 (2016).
- Kita, Y. et al. Genetic improvement of the somatic embryogenesis and regeneration in soybean and transformation of the improved breeding lines. *Plant Cell Rep.* 26, 439–447 (2007).
- Heidmann, I. et al. Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant Cell Rep.* **30**, 1107–1115 (2011).
- 60. Chen, X. et al. Genome-wide investigation of DNA methylation dynamics reveals a critical role of DNA demethylation during the early somatic embryogenesis of *Dimocarpus longan* Lour. *Tree Physiol.* **40**, 1807–1826 (2020).
- Kawakatsu, T., Nery, J. R., Castanon, R. & Ecker, J. R. Dynamic DNA methylation reconfiguration during seed development and germination. *Genome Biol.* 18, 171 (2017).
- 62. Lin, J. Y. et al. Similarity between soybean and Arabidopsis seed methylomes and loss of non-CG methylation does not affect seed development. *Proc. Natl. Acad. Sci. USA* **114**, E9730–E9739 (2017).
- Chakraborty, T., Kendall, T., Grover, J. W. & Mosher, R. A. Embryo CHH hypermethylation is mediated by RdDM and is autonomously directed in *Brassica rapa*. *Genome Biol.* 22, 140 (2021).
- 64. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
- van den Hurk, M. et al. Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency. *Epigenomics* 8, 1131–1149 (2016).
- 66. Wang, J. et al. A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364–368 (2006).
- 67. Huangfu, D. et al. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* **26**, 795–797 (2008).
- 68. Qiu, Q. et al. DNA methylation repels targeting of Arabidopsis REF6. *Nat. Commun.* **10**, 2063 (2019).
- Kim, D., Landmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360 (2015).
- Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
- Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-seq applications. *Bioinformatics* 27, 1571–1572 (2011).
- 73. Ramírez, F. et al. deepTools: a flexible platform for exploring deepsequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
- Gaspar, J. M. & Hart, R. P. DMRfinder: efficiently identifying differentially methylated regions from MethylC-seq data. *BMC Bioinform*. 18, 528 (2017).
- 75. Zhu, L. J. et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinform.* **11**, 237 (2010).
- Huang, X. et al. ViewBS: a powerful toolkit for visualization of highthroughput bisulfite sequencing data. *Bioinformatics* 34, 708–709 (2018).

- Zhao, H. et al. Genome-wide DNA methylome reveals the dysfunction of intronic microRNAs in major psychosis. *BMC Med. Genomics* 8, 62 (2015).
- Wang, F. X. et al. Protocol for assaying chromatin accessibility using ATAC-seq in plants. STAR Protoc. 2, 100289 (2021).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
- Yu, G. C., Wang, L. G., Han, Y. Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16, 284–287 (2012).

## Acknowledgements

We are grateful to Dr. Xiaofeng Cao (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences) for providing the seeds of *drm1/2* mutant, Dr. Xinjian He (National Institute of Biological Sciences, China) for providing the seeds of *suvh1/3* and *sdj1/2/3* mutants, Prof. Remko Offringa (Leiden University) for providing the seeds of *ahl15* and *ahl15/19* mutants. The *106HGFP* construct was provided by Prof. Jian Xu (Radboud University, The Netherlands). This work was supported by the National Key Research and Development Program of China (2024YFF1000700 and 2022YFF1002902 to Y.H.S.) and the National Natural Science Foundation of China (32270377 to Y.H.S.).

### Author contributions

Y.H.S. designed the experiments. J.P., Q.Z., and B.J.X. performed the experiments. J.P., Q.Z., and L.P.T. performed the data analysis. Y.H.S., X.S.Z., and T.L. supervised the project and wrote the manuscript.

# **Competing interests**

The authors declare no competing interests.

# Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-59335-8.

**Correspondence** and requests for materials should be addressed to Ying Hua Su.

**Peer review information** *Nature Communications* thanks Daniel Bouyer and the other anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025