

An *Arabidopsis* AT-hook motif nuclear protein mediates somatic embryogenesis and coinciding genome duplication

Omid Karami^{1✉}, Arezoo Rahimi¹, Patrick Mak^{1,4}, Anneke Horstman^{2,3}, Kim Boutilier², Monique Compier^{1,5}, Bert van der Zaal¹ & Remko Offringa^{1✉}

Plant somatic cells can be reprogrammed into totipotent embryonic cells that are able to form differentiated embryos in a process called somatic embryogenesis (SE), by hormone treatment or through overexpression of certain transcription factor genes, such as *BABY BOOM* (*BBM*). Here we show that overexpression of the *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED 15* (*AHL15*) gene induces formation of somatic embryos on *Arabidopsis thaliana* seedlings in the absence of hormone treatment. During zygotic embryogenesis, *AHL15* expression starts early in embryo development, and *AHL15* and other *AHL* genes are required for proper embryo patterning and development beyond the globular stage. Moreover, *AHL15* and several of its homologs are upregulated and required for SE induction upon hormone treatment, and they are required for efficient *BBM*-induced SE as downstream targets of *BBM*. A significant number of plants derived from *AHL15* overexpression-induced somatic embryos are polyploid. Polyploidisation occurs by endomitosis specifically during the initiation of SE, and is caused by strong heterochromatin decondensation induced by *AHL15* overexpression.

¹Plant Developmental Genetics, Institute of Biology Leiden, Leiden University, Leiden, Netherlands. ²Bioscience, Wageningen University and Research, Wageningen, Netherlands. ³Laboratory of Molecular Biology, Wageningen University and Research, Wageningen, Netherlands. ⁴Present address: Sanquin Plasma Products B.V., Amsterdam, Netherlands. ⁵Present address: Rijk Zwaan Netherlands B.V., De Lier, The Netherlands. ✉email: o.karami@biology.leidenuniv.nl; r.offringa@biology.leidenuniv.nl

The conversion of somatic cells into embryonic stem cells is a process that occurs in nature in only a few plant species, for example on the leaf margins of *Bryophyllum calycinum*¹ or *Malaxis paludosa*², or from the unfertilized egg cell or ovule cells of apomictic plants^{3,4}. By contrast, for many more plant species, somatic cells can be converted into embryonic cells under specific laboratory conditions^{5,6}. The process of inducing embryonic cell fate in somatic plant tissues is referred to as somatic embryogenesis (SE). Apart from being a tool to study and understand early embryo development, SE is also an important tool in plant biotechnology, where it is used for asexual propagation of (hybrid) crops or for the regeneration of genetically modified plants during transformation⁷.

SE is usually induced in *in vitro* cultured tissues by exogenous application of plant growth regulators. A synthetic analog of the plant hormone auxin, 2,4-dichlorophenoxyacetic acid (2,4-D), is the most commonly used plant growth regulator for the induction of SE^{8,9}. During the past two decades, several genes have been identified that can induce SE on cultured immature zygotic embryos or seedlings when overexpressed in the model plant *Arabidopsis thaliana*^{6,10}. Several of these genes, including *BABY BOOM* (*BBM*) and *LEAFY COTYLEDON 1* (*LEC1*) and *LEC2*, have now been recognized as key regulators of SE^{11–14}.

Here we show that overexpression of *Arabidopsis AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED 15* (*AHL15*) can also induce somatic embryos (SEs) on germinating seedlings in the absence of plant growth regulators. AT-hook motifs exist in a wide range of eukaryotic nuclear proteins, and are known to bind to the narrow minor groove of DNA at short AT-rich stretches^{15,16}. In mammals, AT-hook motif proteins are chromatin modification proteins that participate in a wide array of cellular processes, including DNA replication and repair, and gene transcription leading to cell-growth, cell-differentiation, cell-transformation, cell-proliferation, and cell-death¹⁷. The *Arabidopsis* genome encodes 29 AHL proteins that contain one or two AT-hook motifs and a plants and prokaryotes conserved (PPC) domain, that directs nuclear localization and contributes to the physical interaction of AHL proteins with other nuclear proteins, such as transcription factors^{18,19}. AHL gene families are found in angiosperms and also in early diverging land plants such as *Physcomitrella patens* and *Selaginella moellendorffii*^{20,21}. *Arabidopsis* AHL proteins have roles in several aspects of plant growth and development, including flowering time, hypocotyl growth^{20,21}, flower development²², vascular tissue differentiation²³, and gibberellin biosynthesis²⁴. Recently, we have shown that *AHL15* also enhances plant longevity by suppressing axillary meristem maturation²⁵. How plant AHL proteins regulate these underlying biological events is largely unknown. Here, we show that *AHL15* and its homologs play major roles in directing plant cell totipotency during both zygotic embryogenesis and 2,4-D-mediated and BBM-mediated SE. Furthermore, our data show that *AHL15* has a role in chromatin opening, and that its overexpression induces SE coinciding with endomitosis and polyploidy.

Results

Overexpression of AHL genes induces SE. *AHL15* was originally identified in a yeast one-hybrid screen, and functional analysis revealed that overexpression of the gene in *Arabidopsis thaliana* (*Arabidopsis*) under control of the 35S promoter (*p35S:AHL15*) results in induction of SE on germinating seedlings²⁶. *AHL15* overexpression seedlings initially remained small and pale and then developed very slowly (Fig. 1a). Three to four weeks after germination, seedlings from the majority of the transgenic lines (41 of 50 lines) recovered from this growth retardation (Fig. 1a)

and underwent relatively normal development, producing rosettes, flowers, and finally seeds. However, in the remaining *p35S:AHL15* lines (9 of 50 lines), globular structures could be observed on seedling cotyledons 1–2 weeks after germination (Fig. 1b). These structures developed into heart-shaped or torpedo-shaped somatic embryos (Fig. 1c) that could be germinated to produce fertile plants. SE was not observed in other tissues, such as roots, hypocotyls or leaves (Fig. 1b).

In *Arabidopsis*, the cotyledons of immature zygotic embryos (IZEs) are the most competent tissues for SE in response to the synthetic auxin 2,4-dichlorophenoxyacetic acid (2,4-D)⁸. Remarkably, a high percentage (85–95%) of the IZEs from the nine selected *p35S:AHL15* lines were able to produce somatic embryos when cultured on medium lacking 2,4-D. When left for a longer time on this medium, these primary *p35S:AHL15* somatic embryos produced secondary somatic embryos (Fig. 1d, e), and in two of the nine *p35S:AHL15* lines, this repetitive induction of SE resulted in the formation of embryonic masses (Fig. 1f). Overexpression of other *Arabidopsis* AHL genes encoding proteins with a single AT-hook motif (i.e., the closest paralogs *AHL19* and *AHL20*, and *AHL29* as a more distant one; Supplementary Fig. 1), did not induce SE on germinating seedlings, but did induce SE on a low percentage (20–30%) of the IZEs in the absence of 2,4-D (Supplementary Fig. 2). These results suggest that AHL proteins can enhance the embryonic competence of plant tissues, with *AHL15* being able to most efficiently induce a totipotent state without addition of 2,4-D.

AHL genes are important during zygotic embryogenesis. Given the role of AHL genes in promoting *in vitro* totipotency, we determined whether these genes also have a role during zygotic embryogenesis. Expression analysis using the *pAHL15:AHL15-GUS* and *pAHL15:AHL15-tagRFP* lines, both in the wild-type background, showed that *AHL15* is expressed in zygotic embryos (ZEs) from the four cell embryo stage onward (Fig. 2a–i). In line with the previously reported nuclear localization of AHL proteins^{22,27}, the *AHL15-tagRFP* fusion protein was detected in the nucleus (Fig. 2e–i and Supplementary Fig. 3). Single *ahl15* or *ahl19* loss-of-function mutants or *ahl15 ahl19* double mutants carrying an artificial microRNA targeting *AHL20* (*ahl15 ahl19 amiRAHL20*) showed wild-type ZE development (Supplementary Fig. 4). Also *pAHL15:AHL15-GUS* plants produced wild-type embryos (Fig. 2n) and seeds (Fig. 2j). However, in reciprocal crosses of this reporter line with the *ahl15* mutant we were unable to obtain homozygous *ahl15 pAHL15:AHL15-GUS* seedlings among 50 F2 plants that were genotyped. Irrespective of the direction in which the cross was made, F1 siliques showed a wild-type phenotype, whereas siliques of *ahl15/+ pAHL15:AHL15-GUS* F2 plants contained around 25% brown, shrunken seeds (Fig. 2k and Supplementary Fig. 5) that were unable to germinate. Embryos in these shrunken seeds showed patterning defects and did not develop past the globular stage (Fig. 2o). These results suggested that the *AHL15-GUS* fusion protein has a strong dominant negative effect on AHL function in the absence of the wild-type *AHL15* protein. This seemed specific for the *AHL15-GUS* fusion, as fertile homozygous *ahl15 pAHL15:AHL15-tagRFP* plants showing wild-type development could be obtained for three independent *pAHL15:AHL15-tagRFP* lines.

To confirm that the chimeric *AHL15-GUS* fusion protein caused the mutant phenotypes only in the mutant background, we introduced the *pAHL15:AHL15-GUS* construct in the *ahl15 pAHL15:AHL15* background. ZEs of *ahl15 pAHL15:AHL15* plants did not show any morphological defects (Fig. 2l), and the resulting *ahl15 pAHL15:AHL15-GUS pAHL15:AHL15* siliques showed normal seed development (Fig. 2m). These results

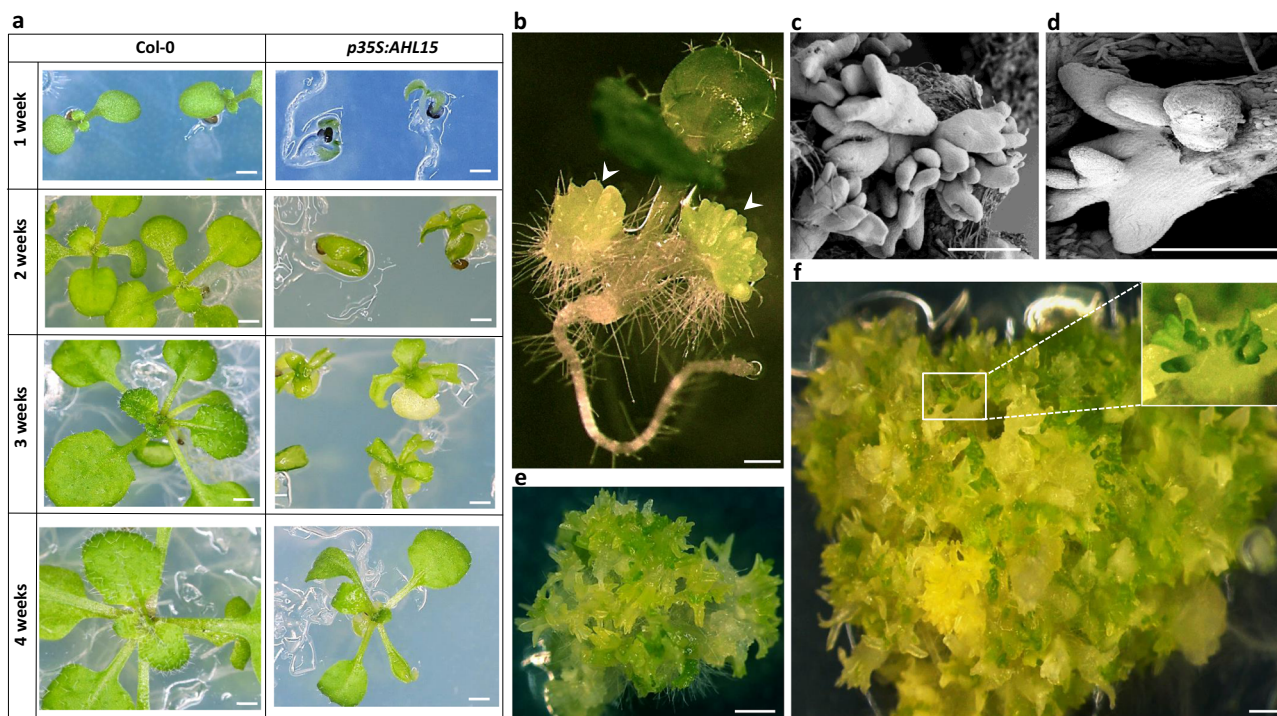


Fig. 1 Overexpression of *AHL15* delays *Arabidopsis* seedling development and induces SE. **a** The morphology of 1-week-old, 2-week-old, 3-week-old, and 4-week-old wild-type and *p35S:AHL15* seedlings and plants grown in long day conditions (16 h photoperiod). **b** Two-week-old *p35S:AHL15* *Arabidopsis* seedling with somatic embryos on the cotyledons (arrowheads). **c, d** Scanning electron micrographs showing torpedo stage somatic embryos on *p35S:AHL15* cotyledons (**c**) or the secondary somatic embryos formed on a *p35S:AHL15* primary somatic embryo (**d**). **e** The morphology of a 3-week-old *AHL15* overexpression-induced embryonic mass following secondary SE. **f** The morphology of a 2-month-old embryonic mass formed from a *p35S:AHL15* seedling. Size bars indicate 0.5 cm (**a**) and 1 mm (**b–f**). **a–f** Similar results were obtained from four independent experiments.

indicate that wild-type *AHL15*, when sufficiently expressed, is able to negate the dominant negative effect of the *AHL15*-GUS fusion.

AHL proteins bind to each other through their PPC domain and form complexes with non-*AHL* transcription factors through a conserved six-amino-acid region in the PPC domain. In *Arabidopsis*, expression of an *AHL* protein without the conserved six-amino-acid region in the PPC domain leads to a dominant negative effect, as the truncated protein can still bind to other *AHL* proteins, but creates a non-functional complex¹⁹. Based on this finding and a similar dominant-negative approach reported for human AT-Hook proteins²⁸, we deleted these six amino acids from the PPC domain of *AHL15* (Fig. 2p) and expressed it under the *AHL15* promoter (*pAHL15:AHL15-ΔG*) in wild-type plants. All plant lines generated with a *pAHL15:AHL15-ΔG* construct ($n = 20$) were fertile and developed wild-type ZEs (Fig. 2q). However, like for the *ahl15/+ pAHL15:AHL15-GUS* lines, also *ahl15/+ pAHL15:AHL15-ΔG* plants showed defective seed development (Fig. 2r) and were unable to produce homozygous *ahl15 pAHL15:AHL15-ΔG* progeny, indicating that this genetic combination is also embryo lethal. These results provide additional support for the dominant negative effect caused by the *AHL15*-GUS fusion protein, showing that *AHL15* and homologs are important during zygotic embryogenesis.

***AHL* genes are required for 2,4-D-induced and BBM-induced SE.** Next we investigated the contribution of *AHL* genes to 2,4-D-induced SE, by culturing IZEs from *ahl* loss-of-function mutants on medium containing 2,4-D. Only a slight reduction in SE induction efficiency was observed in the single *ahl15* loss-of-function mutant (Fig. 3a), which triggered us to examine the contribution of other *AHL* genes in this process. RT-qPCR

analysis showed that *AHL15*, *AHL19*, and *AHL20* expression was significantly upregulated in IZEs following 7 days of 2,4-D treatment (Fig. 3b). Moreover, analysis of *pAHL15:AHL15-GUS* expression in IZEs showed that *AHL15* expression was specifically enhanced in the cotyledon regions where somatic embryos are initiated (Fig. 3c, d). IZEs from triple *ahl15 ahl19 amiRAHL20* mutants produced significantly less somatic embryos (Fig. 3e). Moreover, whereas *pAHL15:AHL15-GUS* IZEs showed a near wild-type capacity to form somatic embryos in the presence of 2,4-D, a strong decrease in the embryogenic capacity of the IZE explants was observed when the *pAHL15:AHL15-GUS* reporter was in the *ahl15/+* mutant background (Fig. 3a). The majority of the IZEs produced non-embryogenic calli that were not observed in the other genotypes (Fig. 3e). These results indicate that *AHL15* and several homologs are required for 2,4-D-induced somatic embryo formation starting from IZEs.

Overexpression of the AINTEGUMENTA-LIKE (*AIL*) transcription factor *BBM* efficiently induces SE in the absence of exogenous growth regulators^{13,29}. Genome-wide analysis of *BBM* binding sites using chromatin immunoprecipitation (ChIP)³⁰ in 2,4-d and *p35S:BBM*-induced somatic embryos showed significant *BBM* binding to the proximal promoter (~ -200 to -500 bp) of *AHL15*, *AHL19*, and *AHL20* (Fig. 3f–i), suggesting that these *AHL* genes are direct downstream *BBM* targets. Analysis of gene expression changes in *p35S:BBM-GR* plants 3 h after treatment with dexamethasone (DEX) and the translational inhibitor cycloheximide (CHX) showed that *BBM* activated the expression of *AHL15* and *AHL20*. A slight but statistically not significant enhancement in expression ($p = 0.1$) was observed for *AHL19* (Fig. 3j). As expected, no difference in expression ($p = 0.9$) was observed for *AHL29* (Fig. 3j), which proximal promoter was not bound by *BBM* (Fig. 3i).

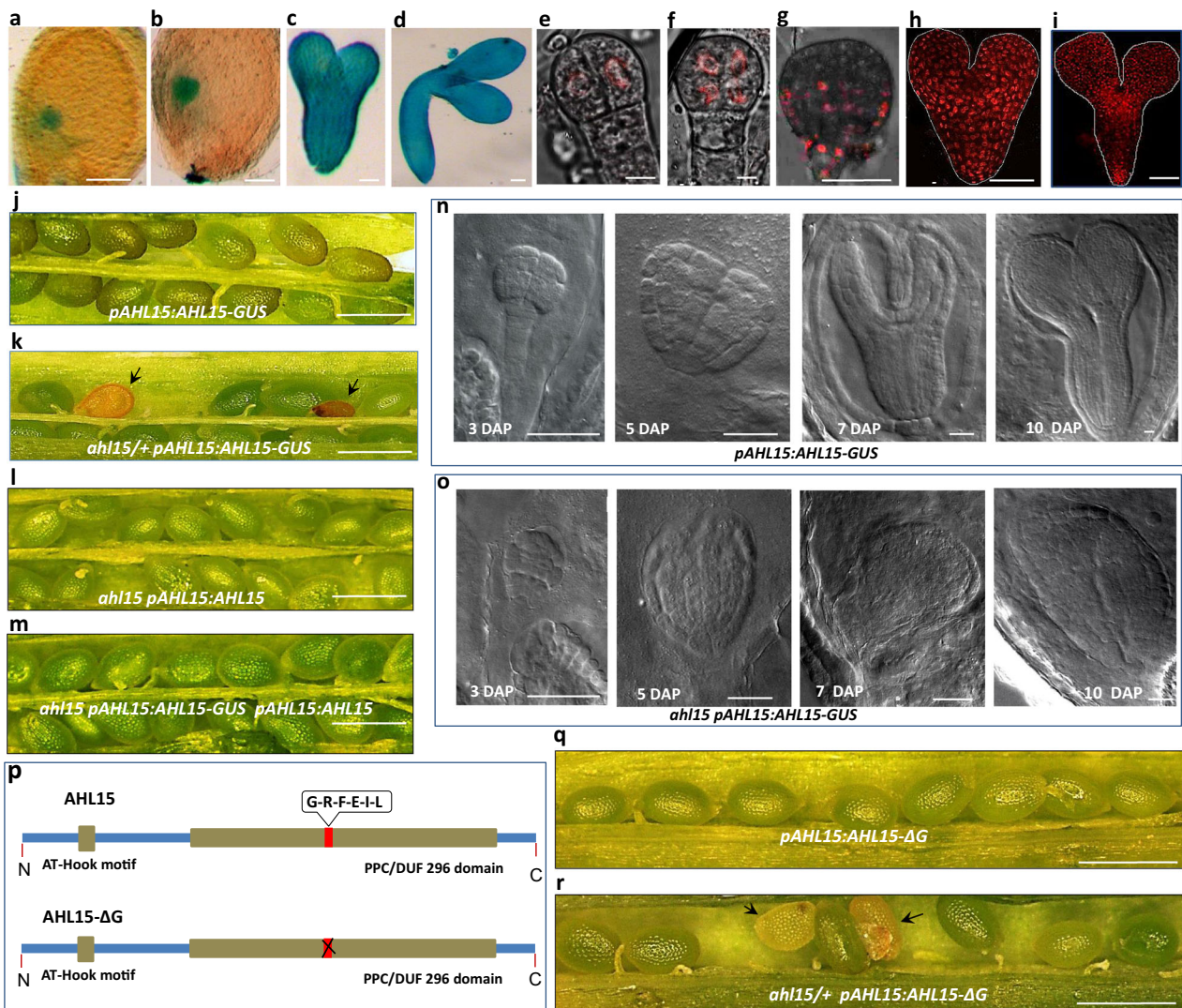


Fig. 2 *AHL15* is expressed and essential during ZE. **a–d** Expression pattern of *pAHL15:AHL15-GUS* in globular-stage (**a**), heart-stage (**b**), torpedo-stage (**c**) and bent cotyledon (**d**) stage embryos. **e–i** Confocal microscopy images of *pAHL15:AHL15-tagRFP* four cell (**e**), 8 cell-stage (**f**), globular-stage (**g**), heart-stage (**h**), and torpedo-stage (**i**) embryos. **j** Wild-type seed development in a *pAHL15:AHL15-GUS* silique. **k** Aberrant seed development (arrowheads) in a *ah115/+ pAHL15:AHL15-GUS* silique. **l, m** Wild-type seed development in *ah115 pAHL15:AHL15* (**l**) or *ah115 pAHL15:AHL15-GUS pAHL15:AHL15* (**m**) siliques. **n, o** DIC images of zygotic embryo development in siliques of *pAHL15:AHL15-GUS* (**n**) or *ah115/+ pAHL15:AHL15-GUS* (**o**) plants at 3, 5, 7, or 10 days after pollination (DAP). **p** Schematic domain structure of *AHL15* and the dominant negative *AHL15-ΔG* version, in which six-conserved amino-acids (Gly-Arg-Phe-Glu-Ile-Leu, red box) are deleted from the C-terminal PPC domain. **q** Wild-type seed development in *pAHL15:AHL15-ΔG* siliques. **r** Aberrant seed development (arrowheads) in *ah115/+ pAHL15:AHL15-ΔG* siliques (observed in three independent *pAHL15:AHL15-ΔG* lines crossed with the *ah115* mutant). **a–r** Similar results were obtained from two independent experiments. Size bar indicates 100 μm (**a–d, g–l**), 10 μm (**e, f**), 1 mm (**j–m, q, r**) and 40 μm (**n, o**).

Next, we investigated the requirement for *AHL* genes in BBM-induced SE by comparing the effectiveness of the *p35S:BBM-GR* construct in inducing SE in the wild-type or *ah115 ah119 amiRAHL20* triple mutant background. DEX treatment induced SE in 40 of the 554 primary *p35S:BBM-GR* transformants (7%) in the wild-type background, but this was completely abolished in the *ah115 ah119 amiRAHL20* background (0 of the 351 primary transformants). These results, together with the observation that *AHL15* overexpression, like *BBM* overexpression induces spontaneous SE, suggest that induction of *AHL* gene expression is a key regulatory component of the BBM signaling pathway.

***AHL15* overexpression-mediated chromatin decondensation correlates with SE induction.** Based on the observation in animal cells that AT-hook proteins are essential for the open chromatin in neural precursor cells^{31–33}, we investigated whether *AHL15*

modulates the chromatin structure during SE initiation. As a measure for global chromatin structure we quantified the amount of tightly condensed, transcriptionally-repressed regions (heterochromatin), which can be visualized using fluorescent chromatin markers or DNA staining. In somatic plant cells, large scale changes in heterochromatin have been associated with cell identity reprogramming^{34,35}.

First, by tracking SE induction on *p35S:AHL15* IZEs, we observed that protodermal cells at the adaxial side of cotyledons started to divide around 6 days after culture (Supplementary Fig. 6A), leading to the formation of *pWOX2:NLS-GFP* expressing pro-somatic embryos (Supplementary Fig. 6B). Propidium iodide (PI) staining of chromosomal DNA in cotyledon protodermal cells of *35S:AHL15* IZEs showed a remarkable dispersion of heterochromatin coinciding with the appearance of *pWOX2:NLS-GFP* expressing pro-somatic embryos at 7 days after culture

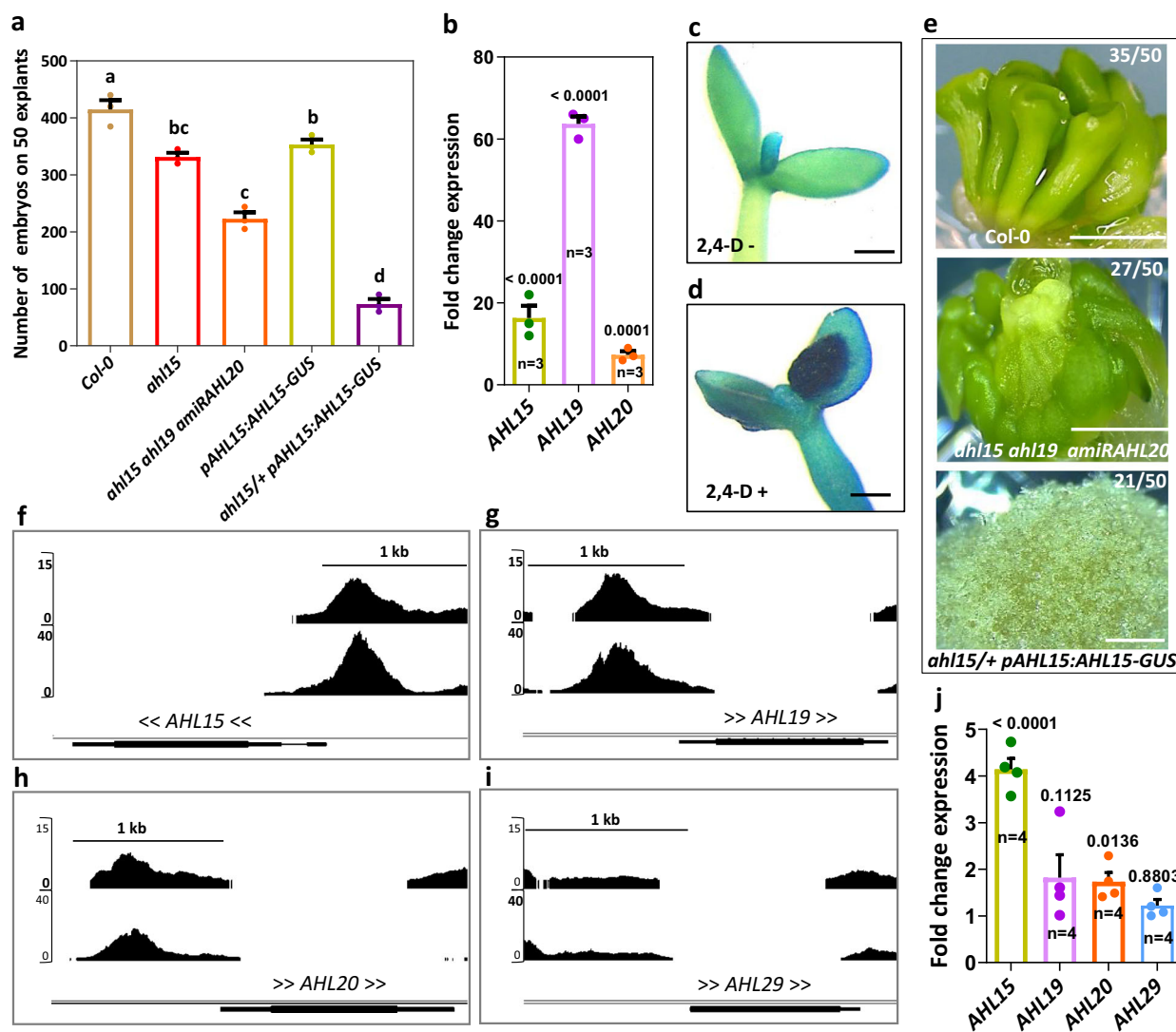


Fig. 3 *AHL* genes are essential for 2,4-D-induced and BBM-induced SE. **a** The effect of *ahl* loss-of-function on the capacity to induce somatic embryos on IZEs by 2,4-D. Dots indicate the values of three biological replicates per plant line with 50 IZEs per replicate, bar indicates the mean, and error bars the s.e.m. For the *ahl15/+ pAHL15:AHL15-GUS* combination IZEs were harvested from *ahl15/+ pAHL15:AHL15-GUS* plants and scored and genotyped after culture. Only scores of *ahl15/+ pAHL15:AHL15-GUS* explants were used. The letters indicate statistically different values, as determined by a one-way ANOVA with Tukey's post hoc test. The *p*-values are provided in the source data file. **b** RT-qPCR analysis of the fold change in expression of *AHL15*, *AHL19* and *AHL20* in IZEs cultured for 7 days on medium with 5 μ M 2,4-D relative to medium without 2,4-D. **c, d** *pAHL15:AHL15-GUS* IZEs cultured for 7 days in the absence (**c**) or presence (**d**) of 5 μ M 2,4-D and histochemically stained for GUS activity. Similar results were obtained from two independent experiments. Size bar indicates 1 mm. **e** Predominant phenotype of wild-type (upper panel), *ahl15 ahl19 amiRAHL20* (middle panel), or *ahl15/+ pAHL15:AHL15-GUS* (lower panel) IZEs cultured for 2 weeks on 2,4-D medium. Numbers indicate the frequency with which similar phenotypes were observed in two independent experiments. Size bar indicates 1 mm. **f-i** ChIP-seq BBM binding profiles for *AHL15* (**f**), *AHL19* (**g**), *AHL20* (**h**), and *AHL29* (**i**). The binding profiles from the *pBBM:BBM-YFP* (upper profile) and *p35S:BBM-GFP* (lower profile) ChIP-seq experiments are shown. The x-axis shows the position of DNA binding relative to the location of the selected gene (TAIR 10 annotation), the y-axis shows the ChIP-seq score (fold enrichment of the BBM-GFP/YFP ChIP to the control ChIP), and the arrow brackets around the gene name indicate the direction of gene transcription. **j** RT-qPCR analysis of the fold change in expression of *AHL* genes in DEX + CHX treated *p35S:BBM-GR* seedlings relative to that in DEX + CHX treated Col-0 wild-type seedlings. **b, j** Dots indicate the values of three or four biological replicates, bar indicates the mean, and error bars the s.e.m. The *p*-values were determined by a two-sided Student's *t*-test.

(Fig. 4a and Supplementary Fig. 7). In contrast, cotyledon protodermal cells of wild-type IZEs did not show a clear change in heterochromatin state at this time point (Fig. 4a and Supplementary Fig. 7). Based on the fraction of compacted chromatin per nucleus, we categorized nuclei into either the condensed (Supplementary Fig. 8a) or the dispersed phenotype (Supplementary Fig. 8b). Quantification showed that the percentage of dispersed nuclei highly increased in cotyledon

protodermal cells of *35S:AHL15* IZEs after 7 days of culture relative to wild-type IZEs (Fig. 4b).

The *Arabidopsis* HISTONE 1.1-GFP³⁶ and HISTONE 2B-GFP³⁵ proteins are incorporated into nucleosomes, providing markers for the chromatin state in living cells. H1.1-GFP and H2B-GFP fluorescence observations confirmed that the chromocenters in cotyledon protodermal cells of 7-day-cultured *35S:AHL15* IZEs (Fig. 4c, d) were much more diffuse compared to

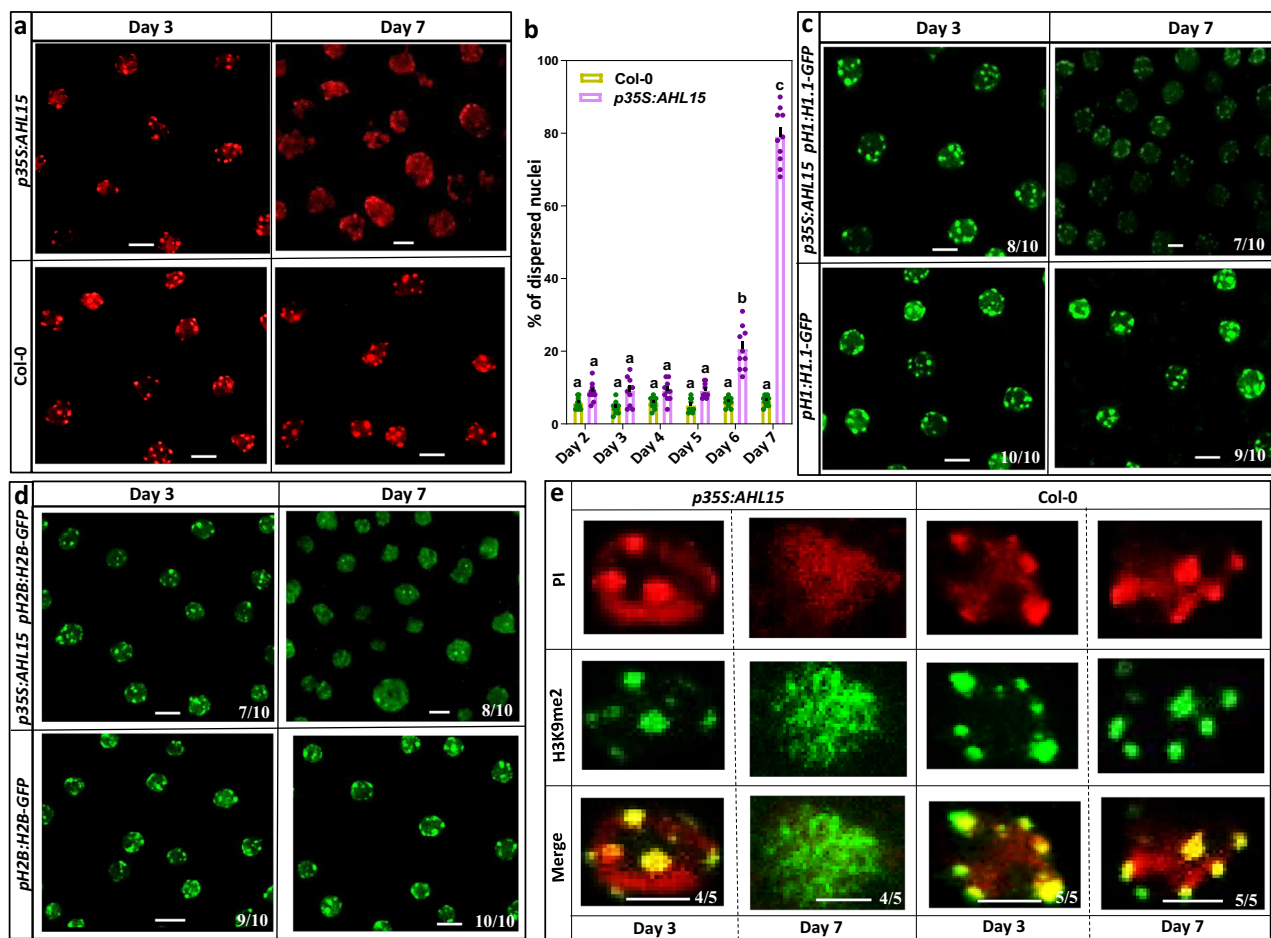


Fig. 4 *AHL15* overexpression reduces heterochromatin condensation during SE induction. **a** Visualization of DNA compaction using propidium iodide (PI) staining in cotyledon protodermal nuclei of wild-type and *p35S::AHL15* IZEs 3 or 7 days after culture. **b** Quantification of the percentage of dispersed PI-labeled nuclei (**a**) according to the classification shown in Supplementary Fig. 7. Dots indicate the values of ten biological replicates per plant line with about 200 nuclei analyzed per replicate, bar indicates the mean, and error bars the s.e.m. Different letters indicate statistically significant differences as determined by one-way ANOVA with Tukey's honest significant difference post-hoc test. The *p*-values are provided in the source data file. **c, d** Visualization of DNA compaction in H1.1-GFP (**c**) or H2B-GFP (**d**) labeled nuclei in cotyledon protodermal cells of wild-type and *p35S::AHL15* IZEs 3 or 7 days after culture. **e** Visualization of heterochromatin compaction in nuclei of cotyledon protodermal cells stained with PI or immunostained against the heterochromatin marker H3K9me2 in wild-type and *p35S::AHL15* IZEs 3 or 7 days after culture. **a, c–e** Size bar indicates 6 μ m and numbers in images indicate the frequency with which similar images from different samples were obtained in two independent experiments.

cells of 3-day-cultured IZEs (Fig. 4c, d). No significant differences in H1.1-GFP and H2B-GFP signals were detected between cotyledon protodermal cells of three and seven day-incubated wild-type IZEs (Fig. 4c, d).

To further confirm the changes in the heterochromatin state in cotyledon cells, we analyzed the dynamic localization of the heterochromatin mark H3K9me2 by immunostaining^{37,38}. Dispersal of the PI stain in nuclei of cotyledon protodermal cells of *35S::AHL15* IZEs 7 days after culture correlated very well with the dispersed H3K9me2 signal (Fig. 4e and Supplementary Fig. 9). In contrast, both the PI staining and the H3K9me2 signal remained condensed in wild-type nuclei. Together these data indicate that *AHL15* promotes heterochromatin decondensation. Surprisingly, in cells expressing both *pAHL15::AHL15-tagRFP* and *pH2B::H2B-GFP* reporters, *AHL15-tagRFP* did not co-localize with the chromocenters (Supplementary Fig. 10), but showed a more diffuse nuclear distribution, suggesting that *AHL15* action is not limited to heterochromatin, but that the protein rather regulates global chromatin decondensation.

To explore whether the observed heterochromatin decondensation in *35S::AHL15* IZEs also occurs in other SE systems, we

investigated the heterochromatin state in embryonic cells induced by 2,4-D. The H1.1-GFP signals in cotyledon protodermal cells of *pH1.1::H1.1-GFP* IZEs cultured for 7 days on medium containing 2,4-D displayed only moderate decondensation of heterochromatin, with chromocenters becoming significantly smaller in cotyledon cells after 2,4-D treatment (Supplementary Fig. 11a–c), but not showing the diffuse signal observed in *p35S::AHL15* IZEs (Fig. 4c). This suggests that the *AHL15*-induced chromatin decondensation is not a general trait of cells undergoing SE, but rather, is specific for *AHL15*-overexpressing cells. Although 2,4-D strongly upregulates *AHL15* expression during somatic embryogenesis (Fig. 3b–d), *35S* promoter-driven *AHL15* expression resulted in approximately 3-fold higher mRNA levels than observed with 2,4-D treatment (Supplementary Fig. 11d). This difference in expression levels might explain the strong chromatin decondensation observed in *35S::AHL15* IZEs compared to 2,4-D-treated IZEs.

To obtain insight into the role of *AHL15* in chromatin decondensation during zygotic embryogenesis, we introduced the *pH2B::H2B-GFP* reporter into the *ahl15/+ pAHL15::AHL15-GUS* background. In defective *ahl15 pAHL15::AHL15-GUS* embryos, we

Table 1 Ploidy level of plants derived from SEs induced by *AHL15* overexpression, *BBM* overexpression or by 2,4-D treatment.

Genotype	SE-derived plants	Ploidy level of plants ^a			Ploidy percentage
		2n	4n	8n	
<i>p35S:AHL15-2</i>	16	5	11	–	69
<i>p35S:AHL15-4</i>	6	4	1	1	33
<i>p35S:AHL15-13</i>	11	7	4	–	36
<i>p35S:AHL15-14</i>	17	14	2	1	18
<i>p35S:AHL15-15</i>	15	11	4	–	27
Wild type, 2-4-D	30	30	–	–	0
<i>p35S:BBM</i>	20	20	–	–	0

^aThe ploidy level was analyzed by counting the chloroplast number in guard cells. For plants derived from *p35S:AHL15*-induced SEs, the ploidy level was confirmed using flow cytometry (Supplementary Fig. 15).

we observed irregular shaped chromocenters that were much larger than those in wild-type cells (Supplementary Fig. 12). This result, together with the reduced heterochromatin condensation observed in cotyledon cells of *p35S:AHL15* IZEs, suggests that *AHL15* plays a role in regulating the chromatin architecture during embryogenesis.

***AHL15* overexpression induces polyploidy during SE initiation.**

Plants regenerated from somatic embryos derived from *p35S:AHL15* IZEs (without 2,4-D treatment) developed large rosettes with dark green (Supplementary Fig. 13) leaves and large flowers (Supplementary Fig. 14a). These phenotypes were not observed in *p35S:AHL15* progeny obtained through zygotic embryogenesis. As these phenotypes are typical for polyploid plants, we determined the ploidy level of the plants. The number of chloroplasts in guard cells³⁹ of plants showing large flowers was two times higher (8–12) than that of diploid wild-type plants (4–6) (Supplementary Fig. 14b). Polyploidisation is also correlated with an increase in cell size and nuclear size in *Arabidopsis* and many other organisms⁴⁰. Indeed root cells of tetraploid *35S::AHL15* seedlings showed a larger nucleus than diploid control plants (Supplementary Fig. 14c). Using the centromere-specific HIS-TONE3-GFP fusion protein (CENH3-GFP)^{41,42}, seven to eight CENH3-GFP-marked centromeres could be detected in root cells of wild-type *Arabidopsis* plants and diploid *35S:AHL15* SE-derived plants (Supplementary Fig. 12d). By contrast, around 12–16 centromeres were observed in the larger nuclei in root cells of tetraploid *p35S:AHL15* plants (Supplementary Fig. 14d). This confirmed that the plants with large organs that were regenerated from *AHL15* overexpression-induced somatic embryos are polyploid. Additional flow cytometry analysis on SE-derived plant lines confirmed that most of these plants were tetraploid, and two were octoploid (Table 1 and Supplementary Fig. 15). The frequency of SE-derived polyploidy varied per *p35S:AHL15* line, ranging from 18 to 69% (Table 1). This variety among lines most likely relates to the level of *AHL15* overexpression in the different *p35S:AHL15* lines. No polyploid plants were obtained from somatic embryos induced by 2,4-D on wild-type IZEs, or by *BBM* overexpression in seedlings (Table 1), indicating that polyploidisation is specifically induced by *AHL15* overexpression.

The considerable frequency of polyploid plants regenerated from *p35S:AHL15* somatic embryos posed the question as to when polyploidisation is induced, and whether it is correlated with, or even promoted by SE induction. We observed a variable number of CENH3-GFP labeled centromeres (6–8, 12–15, and 25–30) in nuclei of protodermal cells at the adaxial side of of *p35S:AHL15* IZE cotyledons 7–8 days after of the start of culture, coinciding with SE induction and reflecting the presence of diploid, tetraploid, and octoploid cells (Fig. 5a). Quantitative analysis showed that nuclei with more than 11 detectable

centromeres were only observed in cotyledon protodermal cells of *p35S:AHL15* IZEs after 7 days of culture (Fig. 5b), demonstrating that polyploidisation coincides with SE induction. No evidence was obtained for polyploidy in root meristems (Supplementary Fig. 16a) or young leaves (Supplementary Fig. 16b) of ZE-derived *p35S:AHL15* plants, nor was polyploidy observed in the 2,4-D-induced non-embryogenic calli found on leaf and root tissues of *p35S:AHL15* plants (Supplementary Fig. 16c, d). When we followed the *pH2B:H2B-GFP* reporter in cotyledons of *p35S:AHL15* IZEs in time, we did not observe any cells with an increased number of chromocenters during the first week of culture (Fig. 5c). At 7 days of IZE culture, however, an increase in chromocenter number could be detected in proliferating *p35S:AHL15* cotyledon cells (Fig. 5d). This result showed that polyploidisation in *p35S:AHL15* cotyledon cells is tightly associated with the induction of SE. Based on these results, and in line with the observation that *p35S:AHL15* polyploid plants were only obtained from *p35S:AHL15* somatic embryos, we conclude that the *AHL15*-induced polyploidisation occurs specifically during somatic embryo initiation.

***AHL15* overexpression-induced polyploidy occurs by endomitosis due to chromosome mis-segregation.** Endoreduplication normally occurs in expanding cells to facilitate cell growth. During endoreduplication, duplicated chromosomes do not enter into mitosis and the number of chromocenters does not increase^{43–45}. We observed an increase in H2B-GFP-marked chromocenters in *p35S:AHL15* cotyledon cells that coincided with polyploidisation events (Fig. 5c, d), but not in endoreduplicated nuclei of wild-type root hair, leaf, or root cells (Fig. 5e–g), suggesting that these polyploid *p35S:AHL15* cells are not derived from endoreduplication. Thus, duplication of segregated chromosomes in *35S:AHL15* cotyledons cells must be caused by endomitosis, during which mitosis is initiated and chromosomes are separated, but cytokinesis fails to occur. Although ectopic overexpression of *AHL15* resulted in a high percentage of polyploid plants (Table 1), polyploidy of the embryo itself was not a prerequisite for further development of somatic embryos into plants, as most of the SE-derived *AHL15* overexpressing plants were still diploid.

Disruption of heterochromatin in human mitotic cells leads to mis-segregation of chromosomes^{46–49} and cellular polyploidization⁵⁰. We hypothesized that heterochromatin disruption and more global chromatin decondensation in dividing *p35S:AHL15* cotyledon cells might contribute to endomitosis resulting in polyploid somatic embryo progenitor cells. Compared to chromosome segregation in wild-type dividing cotyledon cells (Fig. 5h), chromosome segregation in dividing *p35S:AHL15* cotyledon cells lagged behind (Fig. 5i) and binucleate cells (Fig. 5j, k) could be detected in cotyledons of *p35S:AHL15* IZE 7 days

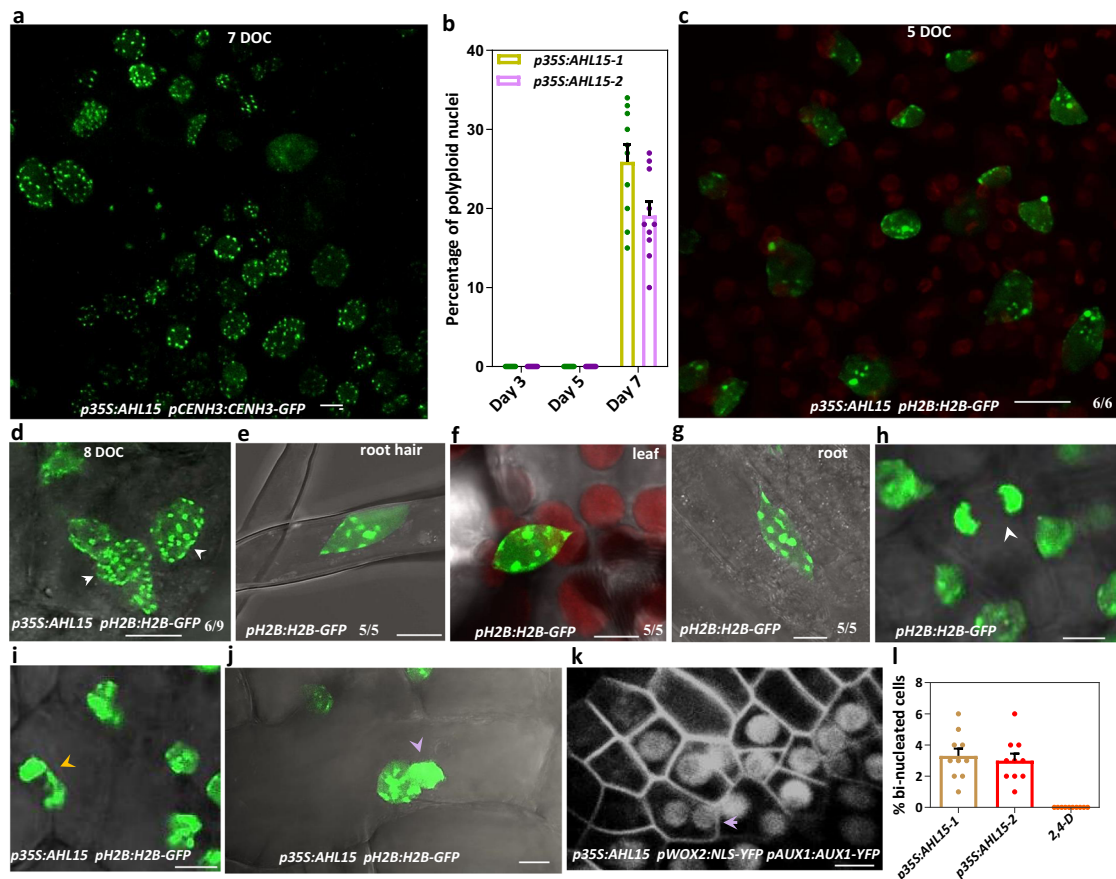


Fig. 5 *AHL15* overexpression-induced SE results in polyploidy by endomitosis due to chromosome mis-segregation. **a** Confocal image of polyloid cells detected by CENH3-GFP-mediated centromere labeling in an embryonic structure developing on a cotyledon of a *35S:AHL15* IZE after 7 days of culture (DOC). **b** Percentage of polyploid nuclei (more than 11 detectable CENH3-GFP-labeled centromeres) in cotyledon cells of *p35S:AHL15* IZEs after 5 or 7 DOC. Dots indicate the values of ten biological replicates per plant line, with about 200 nuclei analyzed per replicate, bar indicates the mean, and error bars the s.e.m. **c, d** H2B-GFP-labeled chromocenters in nuclei of cotyledon cells of *p35S:AHL15* IZEs after 5 (**c**) or 8 (**d**) DOC. White arrowheads (**d**) indicate cells with a duplicated number of chromocenters. **e–g** Confocal images of H2B-GFP labeled chromocenters in endoreduplicated nuclei of wild-type root hair (**e**), leaf (**f**), or root epidermis (**g**) cells. **h–j** Confocal microscopy analysis of chromosome segregation in dividing protodermal cells in cotyledons of *p35S:AHL15* IZEs after 8 DOC using the *pH2B:H2B-GFP* reporter. The white arrowhead indicates normal chromosome segregation during anaphase (**h**), the yellow arrowhead indicates mis-segregation of chromosomes during anaphase (**i**), and the magenta arrowhead indicates a bi-nucleated cell (**j**). **k** Confocal microscopy image of a cotyledon of a *p35S:AHL15 pWOX2:NLS-YFP pAUX1:AUX1-YFP* IZE. *pWOX2:NLS-YFP* and *pAUX1:AUX1-YFP* reporters were used to mark embryonic nuclei and plasma membranes, respectively. The magenta arrowhead indicates a bi-nucleated cell in an area of cells with *WOX2-YFP*-marked embryo cell fate. **l** Percentage of bi-nucleated cells, as detected by H2B-GFP labeling, in protodermal cotyledon cells of wild-type or *p35S:AHL15* IZEs after 7 DOC on medium with or without 2,4-D, respectively. Dots indicate the values of ten biological replicates per plant line (two *35S:AHL15* lines were included) with about 200 nuclei analyzed per replicate, bar indicates the mean, and error bars the s.e.m. **a–k** Similar results were obtained from two independent experiments. Size bar indicates 6 μ m. **d–k** Images show a merge of the transmitted light and the GFP channel (**d–j**), or the GFP (**a, c**) or YFP (**k**) channel alone. **c–g** Numbers in images indicate the frequency with which similar images from different samples were obtained.

after culture. The observation that binucleate cotyledon cells expressed the *pWOX2:NLS-YFP* embryo marker (Fig. 5k, i) confirmed that such cells can adopt embryo identity and develop into polyploid somatic embryos. By contrast, we did not observe any binucleate cells in *pWOX2:NLS-YFP* expressing cotyledon tissues of 2,4-D treated IZEs (Fig. 5I and Supplementary Fig. 17). Taken together, we conclude that heterochromatin disruption in *p35S:AHL15*-induced embryonic cotyledon cells leads to chromosome mis-segregation, the formation of binucleate cells and finally to cellular polyploidization coinciding with the development of polyploid somatic embryos.

TSA treatment or long heat stress exposure induce polyploidy during SE initiation. Inhibition of histone deacetylases by trichostatin A (TSA), a chemical that induces chromatin

decondensation⁵¹, has been reported to promote the transition of plant somatic cells and microspores toward an embryonic state^{52,53}. We therefore asked whether the decondensation of chromatin by TSA treatment induces polyploidy during SE. To test this, *pCENH3:CENH3-GFP* IZEs cultured for 7 days on medium containing 2,4-D were transferred to new medium containing 2,4-D and TSA. We observed a variable number of CENH3-GFP-labeled centromeres (6–8, 12–15, and 25–30) per cell in embryonic cells 2 days after TSA treatment, indicating that polyploidization occurred (Fig. 6a). By contrast, we did not observe such a variable number of centromeres in embryonic cells on medium containing only 2,4-D (Fig. 6a). Quantification of the number of nuclei with more than 11 detectable centromeres confirmed that TSA treatment resulted in polyploidisation in 2,4-D-induced embryonic cells (Fig. 6b). Also long heat stress (LHS) exposure has been shown to induce heterochromatin

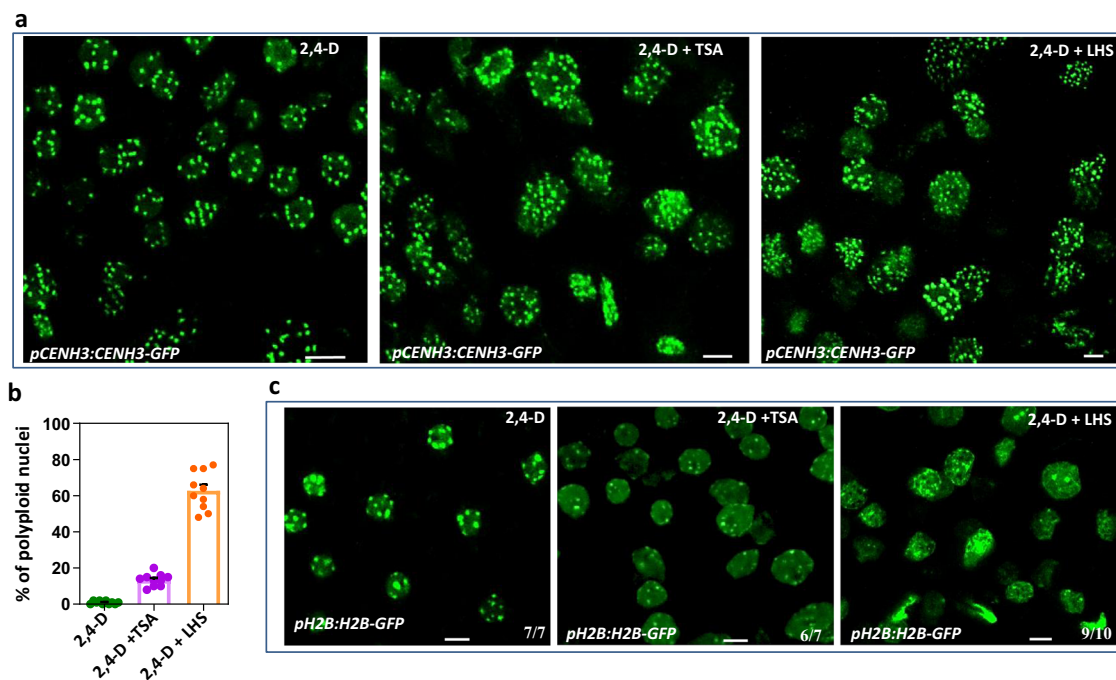


Fig. 6 TSA treatment or heat stress result in polyploidy in 2,4-D induced embryonic cells by promoting heterochromatin decondensation. **a** Confocal image of CENH3-GFP-labeled nuclei in protodermal cells at the adaxial side of IZE cotyledons after 9 DOC. IZEs were cultured for 9 days on medium with 2,4-D (left), or after 7 days on 2,4-D medium transferred for two days to medium with 2,4-D and 1 μ M TSA (middle), or exposed for 2 days to long heat stress (LHS) at 38 $^{\circ}$ C (right). **b** Percentage of polyploid nuclei (based on more than ten detectable CENH3-GFP-labeled centromeres) in cotyledon cells (shown in **a**). Dots indicate the values of ten biological replicates per treatment with about 200 nuclei analyzed per replicate, bar indicates the mean, and error bars the s.e.m. **c** Visualization of heterochromatin decondensation in pH2B:H2B-GFP IZEs after 9 DOC on medium containing 2,4-D (left), or after 7 DOC on 2,4-D medium followed by 2 days on 2,4-D and 1 μ M TSA (2,4-D + TSA, middle), or by 2 days incubation at 38 $^{\circ}$ C (2,4-D + LHS, right). Size bars indicate 6 μ m. **a–c** Similar results were obtained from two independent experiments. **c** Numbers in images indicate the frequency with which similar images from different samples were obtained.

decondensation in Arabidopsis leaves⁵⁴. Incubation of *pCENH3::CENH3-GFP* IZEs, which were precultured for 7 days on medium containing 2,4-D, for 48 h at 38 $^{\circ}$ C also resulted in cellular polyploidization in the 2,4-D-induced embryonic cells, even at a higher rate than with TSA treatment (Fig. 6a, b). No polyploidy was observed in root cells treated with TSA or LHS (Supplementary Fig. 19), suggesting that TSA treatment or LHS exposure only results in polyploidisation in 2,4-D induced embryonic cells.

About 10 or 40% of plants regenerated from the embryonic cells treated with respectively TSA or LHS developed large flowers (Supplementary Fig. 18a). Visualization of the number of CENH3-GFP-labeled centromeres in root cells showed that the plants with large flowers were indeed polyploid (Supplementary Fig. 18b).

Comparison of the heterochromatin state in IZE cotyledon cells using the H2B-GFP reporter revealed a significant decrease of H2B-GFP signal in nuclei of cells receiving 2,4-D + TSA treatment or 2,4-D + LHS exposure compared to the 2,4-D control (Fig. 6c). Together our data show that the induction of embryonic identity by either *AHL15* overexpression or 2,4-D + TSA or LHS treatment in Arabidopsis IZE cotyledon cells coincides with strong heterochromatin decondensation, which frequently leads to cellular polyploidization and thus to the production of polyploid somatic embryos.

Discussion

The herbicide 2,4-D is extensively used for SE induction in a wide range of plant species, including *Arabidopsis*. In *Arabidopsis*, SE can also be induced on IZEs or seedlings in the absence of 2,4-D

treatment by the overexpression of specific transcription factors, such as the AIL transcription factor BBM^{13,29}. In this study, we showed that *AHL15* adds to the list of nuclear proteins whose overexpression induces somatic embryos on IZEs and seedlings in the absence of 2,4-D. In line with this observation, *AHL15* and several of its homologs are upregulated and required for SE induction upon 2,4-D treatment. Furthermore, they are required for efficient *BBM*-induced SE as downstream targets of *BBM*. Like *BBM* and other members of the AIL family¹⁰, *AHL15* is expressed in the early zygotic embryo and acts redundantly with *AHL* family members to maintain embryo development.

AT-hook motif-containing proteins are generally considered to be chromatin architecture factors^{17,31–33,55}. Studies in animals have shown that chromatin decondensation precedes the induction of pluripotent stem cells and their subsequent differentiation⁵⁶. In the *Arabidopsis* zygote, predominant decondensation of the heterochromatin configuration is likely to contribute to the totipotency of this cell⁵⁷. Our data indicate that *AHL15* overexpression induces a global reduction of the amount of heterochromatin in induced somatic embryonic cells, whereas *ahl* loss-of-function mutants show enhanced heterochromatin formation in vitro cultured explants, correlating with a reduced embryonic competence of their explants. Based on our results, we suggest a model in which chromatin opening is required for the acquisition of embryonic competence in somatic plant cells (Fig. 7). In this model, chromatin opening is mediated by upregulation of *AHL* gene expression, which can be achieved by 35S promoter-driven overexpression.

During the S phase of the cell cycle, eukaryotic cells duplicate their chromosomes after which the mitosis machinery ensures

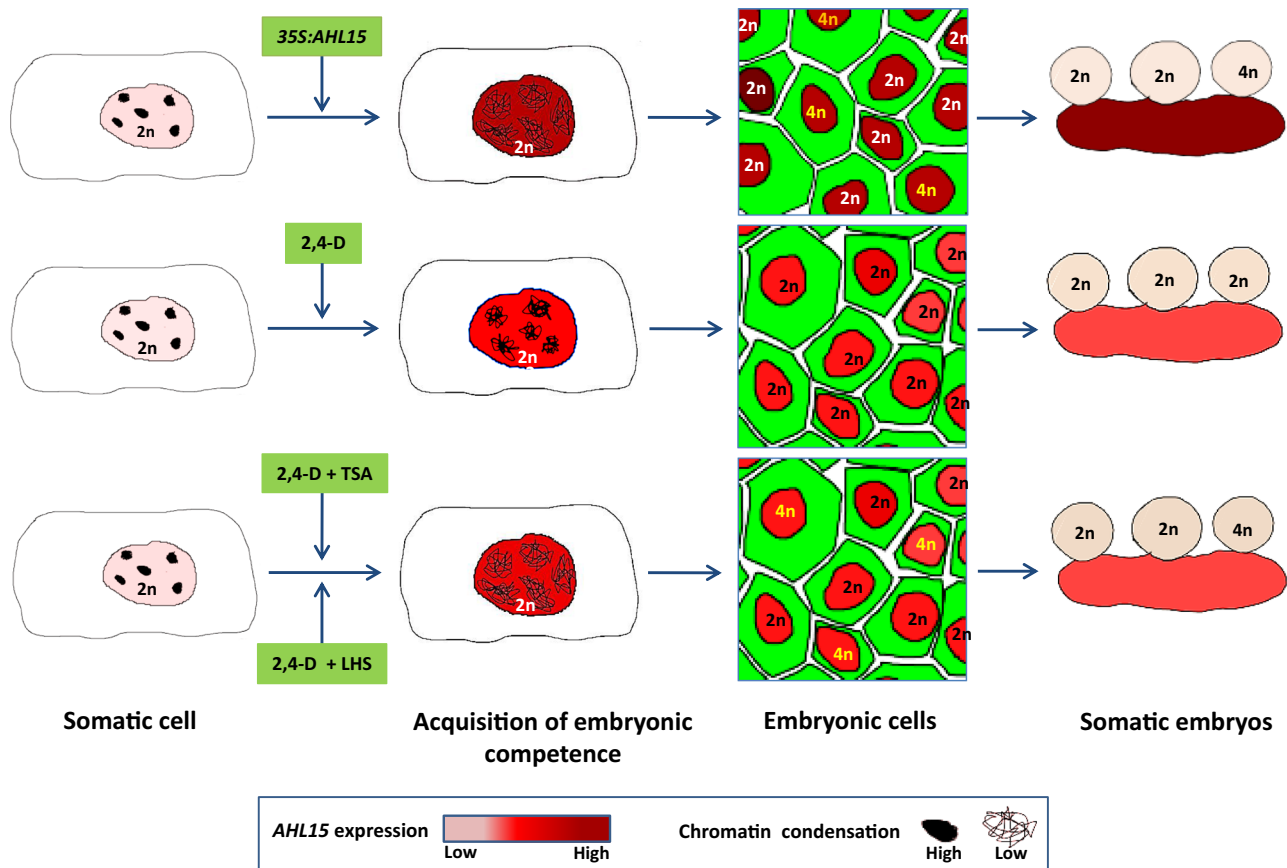


Fig. 7 Cellular polyloidization occurs during somatic-to-embryonic reprogramming due to strong chromatin decondensation. *AHL15* overexpression induces strong chromatin decondensation, coinciding with the induction of embryonic competence in somatic cells. This prevents chromosome segregation in some cells, leading to endomitosis events that give rise to polyloid embryonic cells and subsequently to polyloid somatic embryos. By contrast, 2,4-D treatment leads to enhanced *AHL15* expression that is sufficient to induce somatic-to-embryonic reprogramming, but insufficient to induce the level of chromatin decondensation needed for endomitosis. However, when 2,4-D treatment is combined with a treatment that induces strong chromatin decondensation (TSA or heat stress), endomitosis does occur, giving rise to polyloid somatic embryos.

that the sister chromatids segregate equally over the two daughter cells. However, some cell types do not separate the duplicated chromosomes, leading to a polyloid state known as endopolyploidy⁵⁸. In plants, endopolyploidy is commonly classified either as endomitosis or endoreduplication⁵⁸. Endoreduplication occurs during cellular differentiation, where chromosomes are duplicated but do not segregate, leading to the formation of polytene chromosomes⁴⁴. By contrast, during endomitosis sister chromatids are separated, but the last steps of mitosis including nuclear division and cytokinesis are skipped, generally leading to a duplication of the chromosome number. In this work, we show that polyloid cells resulting from endomitosis can be specifically detected during *p35S:AHL15* induced SE.

Previous studies have shown that defects in heterochromatin condensation in animal cells lead to mis-separation of chromosomes during mitosis^{46–49}, and that mis-segregation of chromosomes subsequently leads to cellular polyploidisation⁵⁰. In our experiments, we found a high reduction of heterochromatin coinciding with mis-segregation of chromosomes in in vitro-cultured *p35S:AHL15* cotyledon cells. Consistent with the strong conservation of chromosome segregation mechanisms between animal and plant cells⁵⁹, we propose that cellular polyploidisation in *p35S:AHL15* embryonic cells is caused by an ectopic *AHL15*-mediated reduction in chromosome condensation during mitosis, which results in chromosome mis-segregation. The observation that polyloid embryos and plants are not obtained after 2,4-D treatment or by *BBM* overexpression suggests that in these somatic

embryos *AHL15* expression levels are finely balanced to prevent strong heterochromatin decondensation and subsequent chromosome mis-segregation (Fig. 7). The fact that 2,4-D culture can lead to polyloid somatic embryos when combined with treatments that enhance chromatin decondensation (TSA or LHS, Fig. 7) provides further evidence that polyloid cells are formed when strong chromatin decondensation coincides with cell proliferation.

A low frequency of polyloid plants derived from somatic embryo cultures has been reported^{8,60–63}, but the molecular and cytological basis for this genetic instability in relation to in vitro embryogenesis has not been described. Since SE requires cell division, these genome duplications most likely have been caused by endomitosis and not by endoreduplication. Until now, endomitosis in plants has been described as a result of defective cytokinesis due to aberrant spindle-plate or cell-plate formation^{42,64,65}. Our results show that chromosome mis-segregation by *AHL15* overexpression-induced chromatin decondensation provides an alternative mechanism of endomitosis in plants, possibly acting not only during SE but also during gametogenesis or zygotic embryogenesis as infrequent environmentally-induced events, that could lead to genome duplication-enabled speciation during evolution. Environmentally-induced or chemically-induced *AHL15* overexpression, combined with TSA treatment or LHS exposure when needed, might also provide an alternative method for chromosome doubling in cultured embryos derived from haploid explants, such as egg cells or microspores⁶⁶ and for the production of polyloid crops.

Methods

Plant material and growth conditions. T-DNA insertion mutants *ahl15* (SALK_040729) and *ahl19* (SALK_070123) were obtained from the European Arabidopsis Stock Center (<http://arabidopsis.info/>). Primers used for genotyping are described in Supplementary Table 1. The reporter lines *pHL1.1:H1.1-GFP³⁶*, *pCENH3:CENH3-GFP⁴¹* and *pH2B:H2B-GFP⁴¹*, *pWOX2:NLS-YFP⁶⁷*, and *pAUX1:AUX1-YFP⁶⁸* have been described previously. For in vitro plant culture, seeds were sterilized in 10% (v/v) sodium hypochlorite for 12 min and then washed four times in sterile water. Sterilized seeds were plated on MA medium⁶⁹ containing 1% (w/v) sucrose and 0.7% agar. Seedlings, plants, and explants were grown at 21 °C, 70% relative humidity, and 16 h photoperiod.

Plasmid construction and plant transformation. The 35S:*AHL15* construct was generated by PCR amplification of the full-length *AHL15* (At3g55560) cDNA from ecotype Columbia (Col-0) using primers 35S:*AHL15-F* and 35S:*AHL15-R* (Supplementary Table 1). The resulting PCR product was cloned as a *Sma*I/*Bgl*II fragment into the *p35S:3'OC*S expression cassette of plasmid pART7, which was subsequently cloned as *Not*I fragment into the binary vector pART27⁷⁰. To generate the other overexpression constructs, the full-length cDNA clones of *AHL19* (At3g04570), *AHL20* (At4g14465), and *AHL29* (At1g76500) from *Arabidopsis* Col-0 were used to amplify the open reading frames (ORFs) using primers indicated in Supplementary Table 1. The ORFs were cloned into plasmid pJET1/blunt (GeneJET™ PCR Cloning Kit, #K1221), and next transferred as *Not*I fragments to binary vector *pGPTV 35S:FLAG⁷¹*. To generate the *pAHL15:AHL15-GUS* and *pAHL15:AHL15-tagRFP* translational fusions, a 4 kb fragment containing the promoter and the full coding region of *AHL15* was amplified using PCR primers *AHL15-GUS-F* and *AHL15-GUS-R* (Supplementary Table 1), and inserted into pDONR207 using a BP reaction (Gateway, Invitrogen). LR reactions were carried out to fuse the 4 kb fragment upstream of *GUS* and *tagRFP* in respectively destination vectors pMDC163⁷² and pGD121⁷³. The artificial microRNA (amiR) targeting *AHL20* was generated as described by Schwab and colleagues⁷⁴ using oligonucleotides I-IV miR-a/s *AHL20* (Supplementary Table 1). The fragment of the *amiRAHL20* precursor was amplified using PCR primers *amiRNA AHL20-F* and *AHL20-R* (Supplementary Table 1), and subsequently introduced into the entry vector pDONR207 via a BP reaction (Gateway, Invitrogen). The *amiRAHL20* precursor was recombined into destination vectors pMDC32⁷² downstream of the 35S promoter via an LR reaction (Gateway, Invitrogen). To generate the *pAHL15:AHL15-ΔG* construct, a synthetic *Kpn*I-*Spe*I fragment (BaseClear) containing the *AHL15* coding region lacking the sequence encoding the Gly-Arg-Phe-Glu-Ile-Leu amino acids in the C-terminal region was used to replace the corresponding coding region in the *pAHL15:AHL15* construct. The *p35S:BBM-GR* construct has been described previously⁷⁵. All binary vectors were introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation⁷⁶ and transgenic *Arabidopsis* Col-0 lines were obtained by the floral dip method⁷⁷.

Somatic embryogenesis. Immature zygotic embryos (IZEs) at the bent cotyledon stage of development (10–12 days after pollination) or germinating dry seeds were used as explants to induce SE using a previously described protocol⁸. In short, seeds and IZEs were cultured on solid B5 medium⁷⁸ supplemented with 5 μM 2,4-D, 2% (w/v) sucrose and 0.7% agar (Sigma) for 2 weeks. Control seeds or IZEs were cultured on solid B5 medium without 2,4-D. To allow further embryo development, explants were transferred to medium without 2,4-D. One week after subculture, the capacity to induce SE was scored under a stereomicroscope as the number of somatic embryos produced from 50 explants cultured on a plate. Three plates were scored for each line. The Student's *t*-test was used for statistical analysis of the data.

Quantitative real-time PCR (RT-qPCR) and ChIP seq analysis. To determine the expression of *AHL* genes during SE induction, RNA was isolated from 25 IZEs cultured for 7 days on B5 medium with or without 2,4-D in four biological replicates using a Qiagen RNeasy Plant Mini Kit. The RNA samples were treated with Ambion® TURBO DNA-free™ DNase. To determine the expression of *AHL* genes in 2,4-D treated Col-0 IZEs by qRT-PCR, 1 μg of total RNA was used for cDNA synthesis with the iScript™ cDNA Synthesis Kit (BioRad). PCR was performed using the SYBR-Green PCR Master mix (Biorad) and a CHOROMO 4 Peltier Thermal Cycler (MJ RESEARCH). The relative expression level of *AHL* genes was calculated according to the 2^{-ΔΔCt} method⁷⁹, using the without 2,4-D value to normalize and the β-*TUBULIN-6* (At5g12250) gene as a reference. The gene-specific PCR primers used are described in Supplementary Table 1.

The effect of *BBM* overexpression on *AHL* gene expression was examined by inducing five-day-old *Arabidopsis thaliana* Col-0 and 35S:*BBM-GR* seedlings (four biological replicates for each line) for 3 h with 10 μM dexamethasone (DEX) plus 10 μM cycloheximide (CHX). RNA was isolated using the Invitex kit, treated with DNase I (Invitrogen) and then used for cDNA synthesis with the Taqman cDNA synthesis kit (Applied Biosystems). qRT-PCR was performed as described above. The relative expression level of *AHL* genes was calculated according to the 2^{-ΔΔCt} method⁷⁹, using the wild-type Col-0 value to normalize and the *SAND* (At2g28390) gene⁸⁰ as a reference. The gene-specific PCR primers are listed in Supplementary Table 1.

The ChIP-seq data and analysis was downloaded from GEO (GSE52400). Briefly, the experiments were performed using somatic embryos from either 2,4-D-induced *BBM:BBM-YFP* cultures (with *BBM:NLS-GFP* as a control) or a 35S:*BBM-GFP* overexpression line (with 35S:*BBM* as a control), as described in the ref. ³⁰.

Ploidy analysis. The ploidy level of plants derived from *p35S:AHL15*-induced somatic embryos was determined by flow cytometry (Plant Cytometry Services, Schijndel, Netherlands), and confirmed by counting the total number of chloroplasts in stomatal guard cells and by comparing flower size and/or the size of the nucleus in root epidermal cells. The number of chloroplasts in stomatal guard cells was counted for plants derived from 2,4-D-induced and *BBM*-induced somatic embryos.

Histological staining and microscopy. Histochemical β-glucuronidase (*GUS*) staining of *pAHL15:AHL15-GUS* IZEs or ovules was performed as described previously⁸¹ for 4 h at 37 °C, followed by rehydration in a graded ethanol series (75, 50, and 25%) for 10 min each. *GUS* stained tissues were observed and photographed using a LEICA MZ12 microscopy (Switzerland) equipped with a LEICA DC500 camera.

DNA staining of wild-type and *p35S:AHL15* seedlings was performed using propidium iodide (PI) according to the protocol described by Baroux et al.⁸². The IZEs were fixed in 1% formaldehyde, 10% DMSO, 2 mM EGTA, 0.1% Tween in 1× phosphate-buffered saline (PBS) 20 for 30 min, followed by three washes in PBT (0.1% Tween 20 in 1× PBS), incubation twice in methanol and twice in ethanol for 2 min each, xylene:ethanol (1:1, v/v) for 30 min, and twice in ethanol for 10 min. The sample was subsequently rehydrated in a series of ethanol (90, 70, 50, and 30%, 5 min each). Following rehydration, tissues were washed twice with PBT for 5 min and incubated in enzyme digestion solution (0.5% driselase [Sigma-Aldrich], 0.5% cellulase [Duchefa Biochemie], 0.5% pectolyase [Duchefa Biochemie] in 50 mM PIPES, 5 mM MgCl₂, and 5 mM EGTA) at 37 °C for 120 min. This was followed by two washes in PBT for 10 min each and RNase treatment at 37 °C (100 μg/ml RNase A [Roche] in 1× PBS with 1% Tween 20) for 120 min. The samples were then washed 15 min with PBT containing 1% formaldehyde, and were washed twice in PBS for 15 min each and stained with 10 μg/ml PI in PBS for 30 min. Stained samples were placed on a microscope slide mounted with a cover slip.

To stain nuclei, the samples were incubated for 30 min in 4',6-diamidino-2-phenylindole (DAPI) staining solution (1 μg/ml DAPI in 1× PBS just before observation).

Immunostaining of cultured IZEs was performed as described previously⁸³. The IZEs were fixed in 1% formaldehyde, 10% DMSO, 2 mM EGTA, 0.1% Tween in 1× PBS for 30 min, and subsequently embedded in 5% acrylamide on microscope slides. Samples were then processed by incubating them in ethanol and methanol for 5 min each, xylene:ethanol (1:1, v/v) for 30 min, 5 min in methanol, 5 min in ethanol, and 15 min in methanol:PBT (1:1, v/v), complemented with 2.5% formaldehyde. The samples were then washed with PBT and were incubated in enzyme digestion solution (0.5% driselase [Sigma-Aldrich], 0.5% cellulase [Duchefa Biochemie], 0.5% pectolyase [Duchefa Biochemie] in 50 mM PIPES, 5 mM MgCl₂, and 5 mM EGTA) at 37 °C for 120 min. This was followed by two washes in PBT, RNase treatment at 37 °C (100 μg/ml RNase A [Roche] in 1× PBS with 1% Tween 20) for 120 min. After that the samples were permeabilized in PBS with 2% Tween-20 at 4 °C for 120 min. For immunostaining, samples were incubated with the primary antibody against H3K9me2 (1:500 dilution, 07-212, Sigma-Aldrich) for 12 h and subsequently with the goat anti-rabbit IgG (H+L) secondary antibody (1:200 dilution, SAB4600234-125UL, Sigma-Aldrich) for 24 h at 4 °C. The samples were mounted in anti-fading liquid mountant supplemented with 10 μg/ml PI.

For scanning electron microscopy (SEM), seedlings were fixed in 0.1 M sodium cacodylate buffer (pH 7.2) containing 2.5% glutaraldehyde and 2% formaldehyde. After fixation, samples were dehydrated by a successive ethanol series (25, 50, 70, 95, and 100%), and subsequently critical-point dried in liquid CO₂. Dried specimens were gold-coated and examined using a JEOL SEM-6400 (Japan).

For morphological studies of embryos, fertilized ovules were mounted in a clearing solution (glycerol:water:chloral hydrate = 1:3:8 v/v) and then incubated at 65 °C for 30 min and observed using a LEICA DC500 microscope (Switzerland) equipped with differential interference contrast (DIC) optics.

The number of chloroplasts in leaf guard cells, the size of the DAPI stained nuclear area in root cells and the number of conspicuous heterochromatin regions of the PI stained nuclei of cotyledon cells were recorded using a confocal laser scanning microscope (ZEISS-003-18533), using a 633 laser and 488 nm LP excitation and a 650–700 nm BP emission filters for chlorophyll signals in guard cells, a 405 laser and 350 nm LP excitation and 425–475 nm BP emission filters for DAPI signals in cotyledon cells, and a 633 laser and 488 nm LP excitation and 600–670 nm emission BP filters for PI signals in cotyledon cells. The relative nuclei area in root cells were determined by measuring the DAPI stained nuclei on confocal images using ImageJ software (Rasband).

Cellular and subcellular localization of *AHL15-tagRFP* and *H2B-GFP* or *CENH3-GFP* protein fusions were visualized using the same laser scanning microscope with a 633 laser and 532 nm LP excitation and 580–600 nm BP emission filters for tagRFP signals, or a 534 laser and 488 nm LP excitation and 500–525 nm BP emission filters for GFP or Anti-rabbit IgG (H+L) signals.

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

Data availability

All processed data are contained in the manuscript or in the Supplementary Information. Materials generated in this study are available upon reasonable request. Source data are provided with this paper.

Received: 18 December 2019; Accepted: 30 March 2021;

Published online: 04 May 2021

References

- Yarborough, J. A. Anatomical and developmental studies of the foliar embryos of *Bryophyllum calycinum*. *Am. J. Bot.* **19**, 443–453 (1932).
- Taylor, R. L. The foliar embryos of *Malaxis paludosa*. *Can. J. Bot.* **45**, 1553–1556 (1967).
- Ozias-Akins, P. & van Dijk, P. J. Mendelian genetics of apomixis in plants. *Annu. Rev. Genet.* **41**, 509–537 (2007).
- Hand, M. L. & Koltunow, A. M. G. The genetic control of apomixis: asexual seed formation. *Genetics* **197**, 441–450 (2014).
- Birnbaum, K. D. & Sánchez Alvarado, A. Slicing across kingdoms: regeneration in plants and animals. *Cell* **132**, 697–710 (2008).
- Smertenko, A. & Bozhkov, P. V. Somatic embryogenesis: life and death processes during apical-basal patterning. *J. Exp. Bot.* **65**, 1343–1360 (2014).
- Bhojwani, S. S. *Plant Tissue Culture: Applications and Limitations* (Elsevier B. V., 2012).
- Gaj, M. D. Direct somatic embryogenesis as a rapid and efficient system for in vitro regeneration of *Arabidopsis thaliana*. *Plant Cell Tissue Organ Cult.* **64**, 39–46 (2001).
- Jiménez, V. M. Involvement of plant hormones and plant growth regulators on in vitro somatic embryogenesis. *Plant Growth Regul.* **47**, 91–110 (2005).
- Horstman, A., Bemer, M. & Boutilier, K. A transcriptional view on somatic embryogenesis. *Regeneration* **4**, 201–206 (2017).
- Lotan, T. et al. *Arabidopsis* *LEAFY COTYLEDON1* is sufficient to induce embryo development in vegetative cells. *Cell* **93**, 1195–1205 (1998).
- Stone, S. L. et al. *LEAFY COTYLEDON2* encodes a B3 domain transcription factor that induces embryo development. *Proc. Natl Acad. Sci.* **98**, 11806–11811 (2001).
- Boutilier, K. et al. Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. *Plant Cell* **14**, 1737–1749 (2002).
- Zuo, J., Niu, Q.-W., Frugis, G. & Chua, N.-H. The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis*. *Plant J.* **30**, 349–359 (2002).
- Aravind, L. & Landsman, D. AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.* **26**, 4413–4421 (1998).
- Reeves, R. Nuclear functions of the HMG proteins. *Biochim. Biophys. Acta* **1799**, 3–14 (2010).
- Sgarra, R. et al. HMG A molecular network: from transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta* **1799**, 37–47 (2010).
- Fujimoto, S. et al. Identification of a novel plant MAR DNA binding protein localized on chromosomal surfaces. *Plant Mol. Biol.* **56**, 225–239 (2004).
- Zhao, J., Favero, D. S., Peng, H. & Neff, M. M. *Arabidopsis thaliana* AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain. *Proc. Natl Acad. Sci.* **110**, E4688–E4697 (2013).
- Street, I. H., Shah, P. K., Smith, A. M., Avery, N. & Neff, M. M. The AT-hook-containing proteins SOB3/AHL29 and ESC/AHL27 are negative modulators of hypocotyl growth in *Arabidopsis*. *Plant J.* **54**, 1–14 (2008).
- Xiao, C., Chen, F., Yu, X., Lin, C. & Fu, Y.-F. Over-expression of an *AT-hook* gene, *AHL22*, delays flowering and inhibits the elongation of the hypocotyl in *Arabidopsis thaliana*. *Plant Mol. Biol.* **71**, 39–50 (2009).
- Ng, K.-H., Yu, H. & Ito, T. AGAMOUS controls *GIANT KILLER*, a multifunctional chromatin modifier in reproductive organ patterning and differentiation. *PLoS Biol.* **7**, e1000251 (2009).
- Zhou, J., Wang, X., Lee, J.-Y. & Lee, J.-Y. Cell-to-cell movement of two interacting AT-hook factors in *Arabidopsis* root vascular tissue patterning. *Plant Cell* **25**, 187–201 (2013).
- Matsushita, A., Furumoto, T., Ishida, S. & Takahashi, Y. AGF1, an AT-hook protein, is necessary for the negative feedback of *AtGA3ox1* encoding *GA 3-oxidase*. *Plant Physiol.* **143**, 1152–1162 (2007).
- Karami, O. et al. A suppressor of axillary meristem maturation promotes longevity in flowering plants. *Nat. Plants* **6**, 368–376 (2020).
- van der Zaal, E. J. & Hooykaas, P. J. J. Control of plant growth and developmental processes. WO2004/066985, <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2004069865> (2004).
- Lim, P. O. et al. Overexpression of a chromatin architecture-controlling AT-hook protein extends leaf longevity and increases the post-harvest storage life of plants. *Plant J.* **52**, 1140–1153 (2007).
- Himes, S. R. et al. The role of high-mobility group I(Y) proteins in expression of IL-2 and T cell proliferation. *J. Immunol.* **164**, 3157–3168 (2000).
- Khanday, I., Skinner, D., Yang, B., Mercier, R. & Sundaresan, V. A male-expressed rice embryonic trigger redirected for asexual propagation through seeds. *Nature* **565**, 91–95 (2019).
- Horstman, A. et al. AIL and HDG proteins act antagonistically to control cell proliferation. *Development* **142**, 454–464 (2015).
- Catez, F. et al. Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. *Mol. Cell. Biol.* <https://doi.org/10.1128/MCB.24.10.4321>. (2004).
- Kishi, Y., Fujii, Y., Hirabayashi, Y. & Gotoh, Y. HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells. *Nat. Neurosci.* **15**, 1127–1133 (2012).
- Jagannathan, M., Cummings, R. & Yamashita, Y. M. A conserved function for pericentromeric satellite DNA. *Elife* **7**, 1–19 (2018).
- Meister, P., Mango, S. E. & Gasser, S. M. Locking the genome: nuclear organization and cell fate. *Curr. Opin. Genet. Dev.* **21**, 167174 (2011).
- Bourbousse, C. et al. Light signaling controls nuclear architecture reorganization during seedling establishment. *Proc. Natl Acad. Sci.* **112**, E2836–E2844 (2015).
- She, W. et al. Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. *Development* **140**, 4008–4019 (2013).
- Jasencakova, Z. et al. Histone modifications in *Arabidopsis*—high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. *Plant J.* **33**, 471–480 (2003).
- Yelagandula, R. et al. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell* **158**, 98–109 (2014).
- Finn, T. E. et al. Transgene expression and transgene-induced silencing in diploid and autotetraploid *Arabidopsis*. *Genetics* **187**, 409–423 (2011).
- Tsukaya, H. Does ploidy level directly control cell size? Counterevidence from *Arabidopsis* genetics. *PLoS ONE* **8**, e83729 (2013).
- Fang, Y. & Spector, D. L. Centromere positioning and dynamics in living *Arabidopsis* plants. *Genetics* **16**, 5710–5718 (2005).
- De Storme, N. et al. Glucan synthase-like8 and sterol methyltransferase2 are required for ploidy consistency of the sexual reproduction system in *Arabidopsis*. *Plant Cell* **25**, 387–403 (2013).
- Edgar, B. A. & Orr-Weaver, T. L. Endoreplication cell cycles: more for less. *Cell* **105**, 297–306 (2001).
- Lermontova, I. et al. Loading of *Arabidopsis* centromeric histone CENH3 occurs mainly during G2 and requires the presence of the histone fold domain. *Plant Cell* **18**, 2443–2451 (2006).
- Lee, H. O., Davidson, J. M. & Duronio, R. J. Endoreplication : polyploidy with purpose. *Genes Dev.* <https://doi.org/10.1101/gad.1829209.results>. (2009).
- Maison, C. & Almouzni, G. HPI and the dynamics of heterochromatin maintenance. *Nat. Rev. Mol. Cell Biol.* **5**, 296–304 (2004).
- Kondo, Y. et al. Downregulation of histone H3 lysine 9 methyltransferase G9a induces centrosome disruption and chromosome instability in cancer cells. *PLoS ONE* **3**, e2037 (2008).
- Carone, D. M. & Lawrence, J. B. Heterochromatin instability in cancer: from the Barr body to satellites and the nuclear periphery. *Semin. Cancer Biol.* **23**, 99–108 (2013).
- Hahn, M. et al. Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev.* **27**, 859–872 (2013).
- Shi, Q. & King, R. W. Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines. *Nature* **437**, 1038–1042 (2005).
- Yang, F. et al. Trichostatin A and 5-azacytidine both cause an increase in global histone H4 acetylation and a decrease in global DNA and H3K9 methylation during mitosis in maize. *BMC Plant Biol.* **10**, 178 (2010).
- 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).
- Li, H. et al. The histone deacetylase inhibitor trichostatin A promotes totipotency in the male gametophyte. *Plant Cell* **26**, 195–209 (2014).
- Pecinka, A. et al. Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell* **22**, 3118–3129 (2010).
- Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat. Rev. Cancer* **7**, 899–910 (2007).
- Gaspar-Maia, A., Alajem, A., Meshorer, E. & Ramalho-Santos, M. Open chromatin in pluripotency and reprogramming. *Nat. Rev. Mol. Cell Biol.* **12**, 36–47 (2011).
- Pillot, M. et al. Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in *Arabidopsis*. *Plant Cell* **22**, 307–320 (2010).

58. Breuer, C., Braidwood, L. & Sugimoto, K. Endocycling in the path of plant development. *Curr. Opin. Plant Biol.* **17C**, 78–85 (2014).
59. Yanagida, M. Basic mechanism of eukaryotic chromosome segregation. *Philos. Trans. R. Soc. Lond. B* **360**, 609–621 (2005).
60. Winkelmann, T., Sangwan, R. S. & Schwenkel, H.-G. Flow cytometric analyses in embryogenic and non-embryogenic callus lines of *Cyclamen persicum* Mill.: relation between ploidy level and competence for somatic embryogenesis. *Plant Cell Rep.* **17**, 400–404 (1998).
61. Borchert, T., Fuchs, J., Winkelmann, T. & Hohe, A. Variable DNA content of *Cyclamen persicum* regenerated via somatic embryogenesis: rethinking the concept of long-term callus and suspension cultures. *Plant Cell Tissue Organ Cult.* **90**, 255–263 (2007).
62. Orbović, V. et al. Analysis of genetic variability in various tissue culture-derived lemon plant populations using RAPD and flow cytometry. *Euphytica* **161**, 329–335 (2007).
63. Prado, M. J. et al. Detection of somaclonal variants in somatic embryogenesis-regenerated plants of *Vitis vinifera* by flow cytometry and microsatellite markers. *Plant Cell Tissue Organ Cult.* **103**, 49–59 (2010).
64. Weingartner, M. et al. Expression of a nondegradable *cyclin B1* affects plant development and leads to endomitosis by inhibiting the formation of a phragmoplast. *Plant Cell* **16**, 643–657 (2004).
65. Iwata, E. et al. GIGAS CELL1, a novel negative regulator of the anaphase-promoting complex/cyclosome, is required for proper mitotic progression and cell fate determination in *Arabidopsis*. *Plant Cell* **23**, 4382–4393 (2011).
66. Soriano, M., Li, H. & Boutilier, K. Microspore embryogenesis: establishment of embryo identity and pattern in culture. *Plant Reprod.* **26**, 181–196 (2013).
67. Breuning, H., Rikirsch, E., Hermann, M., Ueda, M. & Laux, T. Differential expression of *WOX* genes mediates apical-basal axis formation in the *Arabidopsis* embryo. *Dev. Cell* **14**, 867–876 (2008).
68. R, S. et al. Root gravitropism requires lateral root cap and epidermal cells for transport and response to a mobile auxin signal. *Nat. Cell Biol.* **7**, 1057–1065 (2005).
69. Masson, J. & Paszkowski, J. The culture response of *Arabidopsis thaliana* protoplasts is determined by the growth conditions of donor plants. *Plant J.* **2**, 829–833 (1992).
70. Gleave, A. P. A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.* **20**, 1203–1207 (1992).
71. Becker, D., Kemper, E., Schell, J. & Masterson, R. New plant binary vectors with selectable markers located proximal to the left T-DNA border. *Plant Mol. Biol.* **20**, 1195–1197 (1992).
72. Karimi, M., Depicker, A. & Hilson, P. Recombinational cloning with plant gateway vectors. *Plant Physiol.* **145**, 1144–1154 (2007).
73. Immink, R. G. H. et al. Characterization of SOC1's central role in flowering by the identification of its upstream and downstream regulators. *Plant Physiol.* **160**, 433–449 (2012).
74. Schwab, R., Ossowski, S., Riester, M., Warthmann, N. & Weigel, D. Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell* **18**, 1121–1133 (2006).
75. Passarinho, P. et al. BABY BOOM target genes provide diverse entry points into cell proliferation and cell growth pathways. *Plant Mol. Biol.* **68**, 225–237 (2008).
76. Den Dulk-Ras, A. & Hooykaas, J. P. Electroporation of *Agrobacterium tumefaciens*. *Plant Cell Electroporation Electroporation Protoc.* **55**, 63–72 (1995).
77. Clough, S. J. & Bent, F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
78. Gamborg, O. L., Miller, R. & Ojima, K. Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **50**, 151–158 (1968).
79. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method. *Methods* **25**, 402–408 (2001).
80. Czechowski, T., Stitt, M., Altmann, T. & Udvardi, M. K. Genome-wide identification and testing of superior reference genes for transcript normalization. *Plant Physiol.* **139**, 5–17 (2005).
81. Anandalakshmi, R. et al. A viral suppressor of gene silencing in plants. *Proc. Natl Acad. Sci.* **95**, 13079–13084 (1998).
82. Baroux, C., Pecinka, A., Fuchs, J., Schubert, I. & Grossniklaus, U. The triploid endosperm genome of *Arabidopsis* adopts a peculiar, parental-dosage-dependent chromatin organization. *Plant Cell* **19**, 1782–1794 (2007).
83. She, W., Grimanelli, D. & Baroux, C. An efficient method for quantitative, single-cell analysis of chromatin modification and nuclear architecture in whole-mount ovules in *Arabidopsis*. *J. Vis. Exp.* (2014) <https://doi.org/10.3791/51530>. (2014).

Acknowledgements

We thank Gerda Lamers for help with microscopy and are grateful to Johan Pinas (†), Nick Surtel, Ward de Winter, and Jan Vink for their help with plant growth and media preparation. O.K. was supported in part by Grant G14_006.02 to R.O. from Generade, and by the Building Blocks of Life grant 737.016.013 to R.O. from the Netherlands Organisation for Scientific Research. M.C. and P.M. were partially supported through grant IS054064 from the Ministry of Economic Affairs (SenterNovem).

Author contributions

B.v.d.Z. and R.O. conceived the project, all authors designed experiments, and analyzed and interpreted results, O.K. performed the majority of the experiments, A.R. contributed to polyploidisation analysis and generated *pAHL15:AHLL15-ΔG* transgenic lines, P.M. generated and analysed *Arabidopsis* lines overexpressing the *AHL* genes, A.H. performed the BBM-related experiments, O.K. and R.O. wrote the manuscript, K.B., B.v.d.Z. and A.H. assisted in finalizing 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/s41467-021-22815-8>.

Correspondence and requests for materials should be addressed to O.K. or R.O.

Peer review information *Nature Communications* thanks Danny Geelen, Sachihoro Matsunaga, and Iva Mozgova for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permission 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 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021