

Histone H3.3 deposition in seed is essential for the post-embryonic developmental competence in *Arabidopsis*

Received: 13 January 2022

Accepted: 7 December 2022

Published online: 13 December 2022

 Check for updatesTing Zhao^{1,4}, Jingyun Lu^{1,4}, Huairan Zhang^{1,4}, Mande Xue^{1,2}, Jie Pan^{1,2}, Lijun Ma^{1,2}, Frédéric Berger³ & Danhua Jiang^{1,2}  

The acquisition of germination and post-embryonic developmental ability during seed maturation is vital for seed vigor, an important trait for plant propagation and crop production. How seed vigor is established in seeds is still poorly understood. Here, we report the crucial function of *Arabidopsis* histone variant H3.3 in endowing seeds with post-embryonic developmental potentials. H3.3 is not essential for seed formation, but loss of H3.3 results in severely impaired germination and post-embryonic development. H3.3 exhibits a seed-specific 5' gene end distribution and facilitates chromatin opening at regulatory regions in seeds. During germination, H3.3 is essential for proper gene transcriptional regulation. Moreover, H3.3 is constantly loaded at the 3' gene end, correlating with gene body DNA methylation and the restriction of chromatin accessibility and cryptic transcription at this region. Our results suggest a fundamental role of H3.3 in initiating chromatin accessibility at regulatory regions in seed and licensing the embryonic to post-embryonic transition.

The formation of seeds is one of the most important innovations that contribute to the success of flowering plants. By seed formation, the embryo is protected in a desiccated and quiescent state, and, meanwhile, ready to swiftly initiate germination and post-embryonic development when optimal growth condition arises¹. Producing highly vigorous seeds is thus key for plant reproduction and agriculture. Embryo obtains its germination capacity during the late seed maturation stage². However, the underlying mechanisms establishing germination and post-embryonic developmental competence in seeds remain poorly understood.

The eukaryotic genome is tightly packed in the form of chromatin, which consists of DNA and histone-formed nucleosomes as basic subunits. Chromatin compaction restricts the access of transcription factors to *cis*-regulatory elements such as promoters and enhancers, and thus affecting gene transcription^{3,4}. During cellular differentiation

and developmental transitions, epigenetic reprogramming induces chromatin structure changes, thereby modifying the accessibility of regulatory elements to transcription machinery^{5,6}. The establishment of chromatin accessibility is thus a pioneer event in developmental transition and lineage-specific transcriptional regulation. Several mechanisms have been identified to control chromatin structure dynamics, including ATP-dependent chromatin remodeling, post-translational modifications on histones, and the exchange of histone variants^{7,8}. Besides their sequence divergence, histone variants often possess different expression patterns and in some cases unique modifications^{9–11}. The deposition of histones variants can modify nucleosome stability, which differentiates chromatin into specialized property and cellular function^{4,12}.

In animals, the histone H3 variant H3.3 differs from the replicative H3.1/H3.2 with only 4–5 amino acids¹³. Unlike H3.1, which is DNA

¹State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, The Innovative Academy for Seed Design, Chinese Academy of Sciences, Beijing, China. ²University of Chinese Academy of Sciences, Beijing, China. ³Gregor Mendel Institute, Austrian Academy of Sciences, Vienna BioCenter, Dr. Bohr-Gasse 3, 1030 Vienna, Austria. ⁴These authors contributed equally: Ting Zhao, Jingyun Lu, Huairan Zhang.

 e-mail: dhjiang@genetics.ac.cn

replication-dependently expressed and incorporated into the chromatin, H3.3 is expressed and incorporated in a DNA replication-independent manner^{14–18}. H3.3 is associated with transcribed regions and gene regulatory elements^{19,20}. In *Xenopus*, H3.3 is essential for gastrulation, and it is proposed that H3.3 creates a permissive chromatin state critical for embryo development^{21,22}. Knockout one of the two H3.3-coding genes in mouse causes reduced viability and infertility^{23–25}, and complete deletion of H3.3 results in early embryonic lethality²⁶. H3.3 is required to maintain a decondensed chromatin state and active histone modifications in mouse embryo²⁷. Nevertheless, loss of H3.3 seems to have a limited impact on gene transcription, rather it causes abnormal chromosome segregation and impaired genome integrity^{26,27}.

Although evolved independently in separate kingdoms, H3.1 and H3.3 in higher plants acquired comparable features as their counterparts in animals through convergent evolution^{28–30}. For instance, in *Arabidopsis* H3.3 differs from H3.1 with four amino acids³¹, and H3.3 and H3.1 are DNA replication-independently and dependently expressed respectively³². Moreover, *Arabidopsis* H3.3 is mostly enriched at the body of transcribed genes^{32,33}. Three genes *HTR4*, *HTR5* and *HTR8* in *Arabidopsis* encode H3.3. Knockdown of *Arabidopsis* H3.3 (*h3.3kd*) results in mild developmental defects, including leaf serration, early flowering and slightly reduced fertility^{34,35}. H3.3 promotes permissive histone modifications at the floral repressor *FLOWERING LOCUS C (FLC)*, leading to its activation to repress flowering³⁵. Compared with the broad distribution of H3.3 on transcribed genes, only hundreds of genes are misexpressed in *h3.3kd*, indicating that H3.3 may not regulate transcription directly³². Hence, the role of H3.3 in plants remains to be investigated.

Here, we report the critical function of *Arabidopsis* H3.3 in the establishment of chromatin accessibility and endowing seeds with post-embryonic developmental competence. Complete deletion of H3.3 shows no obvious impact on embryogenesis, but results in severe defects in germination and vegetative growth. H3.3 is enriched around both 5' and 3' gene ends in mature seeds. The 5' enrichment of H3.3 is seed-specific, and it facilitates chromatin opening, particularly at the regulatory regions of responsive and post-embryonic developmental genes, and H3.3 is essential for their transcriptional regulation during germination. On the other hand, 3' localization of H3.3 represents a canonical H3.3 deposition pattern that correlates with gene body DNA methylation and the restriction of chromatin accessibility and cryptic transcription. Our results support a fundamental role of H3.3 in the establishment of chromatin accessibility and gene transcriptional regulation important for post-embryonic development.

Results

H3.3 is necessary for plant post-embryonic development

In a previous study, we created a complete H3.3 knockout mutant by combining a CRISPR-Cas9 generated *htr4;htr5* double mutant with an *htr8* mutant carrying a T-DNA insertion at *HTR8*. At that point, we did not obtain triple homozygous plants when screening the *htr4/htr4;htr5/htr5;htr8/+* (*h3.3ko/+*) progenies, and thus we speculated that the *h3.3ko* mutant was embryonic lethal³⁴. To analyze *h3.3ko* mutant phenotypes in detail, we examined the embryo development on *h3.3ko/+* plants. Surprisingly, all seeds developed without abortion (Supplementary Fig. 1a). Around 7.5% to 11.5% of seeds collected from *h3.3ko/+* plants could not germinate after 7 days of imbibition, and thus these were likely *h3.3ko* (Fig. 1a) (Supplementary Fig. 1b). This ratio is lower than 25% because knockout of H3.3 partially impairs male gametogenesis, and when crossing *h3.3ko/+* with the wild-type (WT) Columbia (Col) mother, the transmission of *h3.3ko* was reduced to around 16% instead of the expected 50%³⁴. These ungerminated seeds were then subjected to prolonged imbibition (up to two months), a portion of these seeds eventually germinated and produced small seedlings, and genotyping analysis confirmed that these seedlings

were *h3.3ko* (Fig. 1b, c). Most of the *h3.3ko* seedlings stopped development immediately after germination without expanding cotyledons, and only a few of them could produce two true leaves (Fig. 1c, d). In rare cases, *h3.3ko* plants developed until the bolting stage and produced an inflorescence-like structure (Fig. 1c and Supplementary Fig. 1c). However, these plants ceased growth afterward and no silique was generated.

In order to efficiently obtain the *h3.3ko* mutant seeds, we transformed a construct carrying a WT *pHTR5::HTR5* and a mature seed stage-specific mCherry fluorescence selection marker into *h3.3ko/+*³⁶. The WT *HTR5* successfully rescued the *h3.3ko* phenotypes (Fig. 1d, e). *Arabidopsis* H3.3 differs from H3.1 with only four amino acids. To examine whether the *h3.3ko* phenotypes were specifically caused by the loss of H3.3 or merely due to the reduced H3 supply, we attempted the *h3.3ko* complementation by expressing *HTR13* (an H3.1-coding gene) under the *HTR5* promoter (Supplementary Fig. 1d). Expressing *HTR13* only slightly accelerated the germination of *h3.3ko* (Fig. 1e), suggesting that the *h3.3ko* defects were largely H3.3 specific. We then expressed *HTR5* using the *HTR13* promoter, which was successfully employed in driving *HTR13* expression to complement the *h3.1kd* phenotypes³⁷. Nevertheless, expressing *HTR5* under the *HTR13* promoter also failed to rescue the *h3.3ko* germination defects (Fig. 1e, Supplementary Figs. 1d and S1e). Together, we conclude that both the H3.3 protein sequences and its expression pattern are required for its function in seed germination.

We collected seeds from heterozygous *pHTR5::HTR5* transgenic plants in the *h3.3ko* background, from which the *h3.3ko* seeds could be identified by selecting mature seeds without the mCherry signal (Supplementary Fig. 1f). The *h3.3ko* seeds selected in this way behaved the same as those segregated from *h3.3ko/+* (Fig. 1a, b, d and e), and hence were used in all subsequent studies.

All germination tests were performed with seeds stored for three months after harvesting (after-ripened). Moreover, seeds were stratified for two days during germination, further excluding the possibility that the *h3.3ko* germination defects were due to its strong dormancy. Endosperm and testa (seed coat) are critical tissues in dormancy maintenance, and they also impose a physical constraint on the radicle protrusion^{38,39}. To examine which tissues are responsible for the *h3.3ko* germination defects, we isolated *h3.3ko* embryos by removing endosperm and testa. The isolated *h3.3ko* embryos showed much slower growth compared with WT (Fig. 1f), indicating strong defects in *h3.3ko* embryos.

h3.3ko shows strong defects in transcriptome

To search for a mechanism to explain the strongly impaired germination in *h3.3ko*, we profiled transcriptomes of WT and *h3.3ko* using RNA-seq. Seeds stored for three months after harvesting (hereinafter referred to as mature seeds) were subjected to imbibition and RNA was extracted from seeds collected at various time points (Fig. 2a). Principle component analysis (PCA) showed that compared with WT, the overall transcriptome changes in *h3.3ko* were less dynamic during imbibition (Fig. 2b). We combined all transcript level significantly decreased and increased genes in *h3.3ko* (fold change >2 and *P* adjust <0.05) identified from each time point during imbibition and performed gene ontology (GO) analysis, and found that responsive genes were especially enriched (Fig. 2c) (Supplementary Data 1), suggesting compromised imbibition response in *h3.3ko* compared with WT. GA and ABA are two major hormones that antagonistically regulate germination, during which GA biosynthesis and ABA catabolic genes need to be activated^{40,41}. However, the activation of key GA synthetic and ABA catabolic genes such as *GA3OX1*, *GA3OX2*, *GA20OX1*, *GA20OX2*, and *CYP707A2* was strongly abolished in *h3.3ko* (Fig. 2d and Supplementary Fig. 2a), in line with its impaired germination. Intriguingly, the variation between *h3.3ko* and Col transcriptomes was already notable at the mature seed stage and nearly 5000 genes were significantly

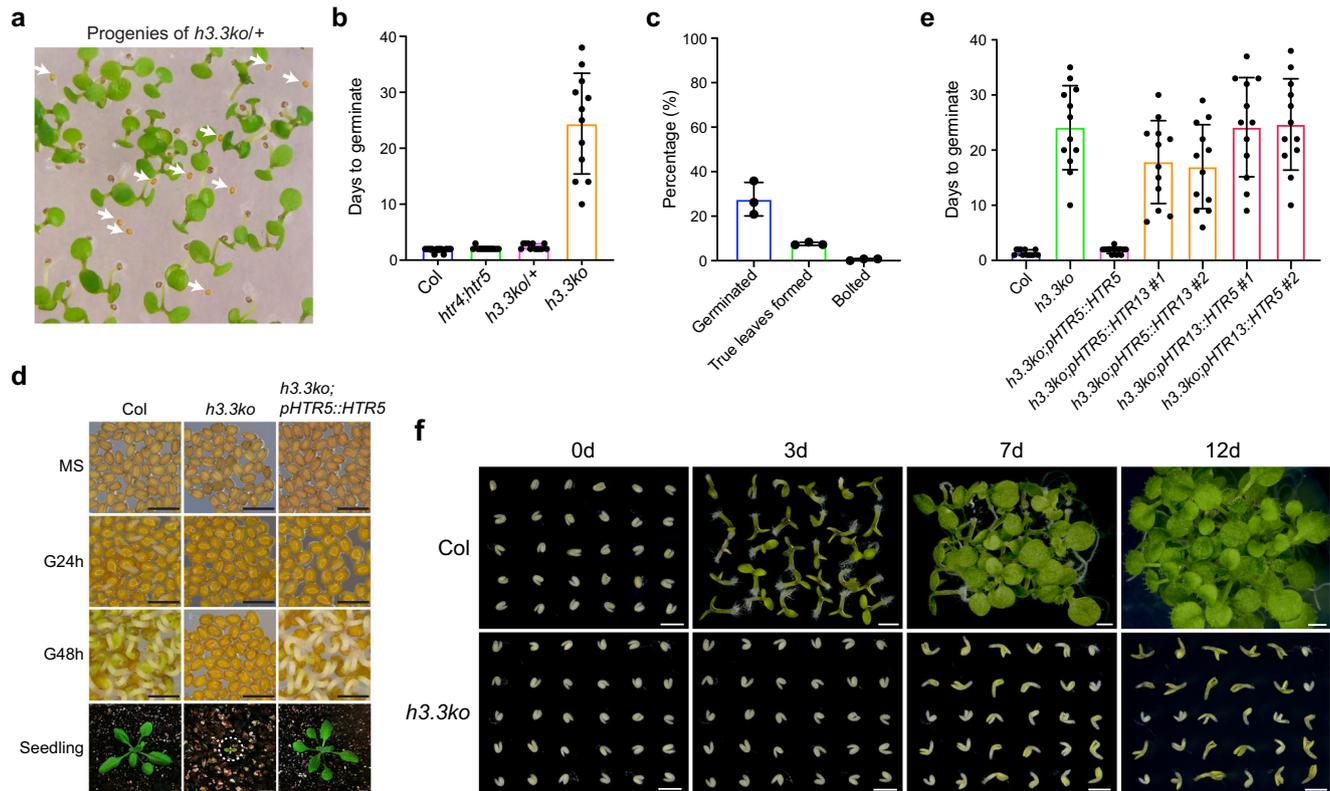


Fig. 1 | H3.3 is required for post-embryonic development. **a** The germination phenotype of *h3.3ko/+* (*htr4/htr4;htr5/htr5;htr8/+*) progenies. Arrows indicate *h3.3ko* seeds that are strongly delayed in germination. **b** Days to germinate of Col, *htr4/htr4;htr5/htr5*, *h3.3ko/+* (*htr4/htr4;htr5/htr5;htr8/+*) and *h3.3ko* (*htr4/htr4;htr5/htr5;htr8/htr8*) seeds. Seeds collected from *h3.3ko/+* plants were sown on 1/2 MS, their genotypes were determined with seedlings formed after germination, and the first 12 germinated seeds were scored for each genotype. Values are means \pm SD. ($n = 12$ biologically independent seeds). Radicle protrusion was considered as the completion of seed germination. **c** Percentages of *h3.3ko* seeds that germinated (within two months), produced true leaves and bolted. Values are means \pm SD of three biological replicates. At least 110 seeds were used for each replicate. **d** The

germination and growth phenotypes of Col, *h3.3ko* (segregated from *h3.3ko;pHTR5::HTR5/-*) and *h3.3ko;pHTR5::HTR5*. Scale bars = 1 mm (for seeds) or 1 cm (for seedlings). MS: mature seeds, G24h: subjected to germination for 24 h, G48h: subjected to germination for 48 h. **e** Days to germinate of Col, *h3.3ko* (segregated from *h3.3ko;pHTR5::HTR5/-*) and transgenic seeds expressing *HTR5* or *HTR13*. The first 12 germinated seeds were scored. More than 110 seeds were sown on 1/2 MS. Values are means \pm SD. ($n = 12$ biologically independent seeds). Radicle protrusion was considered as the completion of seed germination. For *h3.3ko;pHTR5::HTR13* and *h3.3ko;pHTR13::HTR5*, the results from two independent transgenic lines are presented. **f** The germination and growth phenotypes of isolated Col and *h3.3ko* embryos. Scale bars = 1 mm. d: days.

misexpressed in mature *h3.3ko* seeds (Fig. 2b, e). Therefore, H3.3 is already required for the establishment of proper transcriptome in mature seeds.

A distinctive pattern of H3.3 enrichment over genes in mature seeds

To test whether the impaired germination of *h3.3ko* is due to defects in its mature seeds or is caused by the lack of H3.3 during imbibition, we introduced into *h3.3ko/+* an *HTR5* expression construct based on an estradiol inducible system (*pER8::HTR5*), which induces a strong uniform gene expression pattern similar to that of *HTR5*^{32,42,43}. Seeds collected from *h3.3ko/+;pER8::HTR5* were subjected to imbibition with or without estrogen, and *HTR5* transcription was successfully induced by estrogen (Fig. 3a). Interestingly, the induction of *HTR5* expression during imbibition showed no impact on the germination of *h3.3ko* (Fig. 3b, c). The transcript levels of genes involved in H3.3 deposition were comparable in WT and *h3.3ko*, indicating that the H3.3 deposition pathway is generally intact in *h3.3ko* (Supplementary Fig. 2b). Therefore, it appears that loss of H3.3 already caused strong defects in mature seeds that affect germination, and a supply of H3.3 to *h3.3ko* during imbibition is not sufficient to complement the defects caused by the loss of H3.3 in seeds.

The above results prompted us to examine if there were any morphological defects in *h3.3ko* seeds. The seed size and seed weight

of mature *h3.3ko* seeds were comparable to that of WT (Supplementary Fig. 3a–c). In addition, mature embryos isolated from WT and *h3.3ko* seeds were morphologically similar (Fig. 1f and Supplementary Fig. 3d). Seed storage proteins (SSPs) are synthesized during the late stages of seed development and are the major nutritional sources for subsequent germination and early seedling development⁴⁴. The abundance of SSPs was examined with mature seeds, and the major SSPs including 12S globulins and 2S albumins were normally accumulated in *h3.3ko* (Supplementary Fig. 3e), suggesting that H3.3 is not required for their synthesis. Hence, loss of H3.3 did not affect embryo morphology and the accumulation of SSPs.

In WT, the transcript levels of H3.3-coding genes markedly increased towards the end of embryo development, contrasting with the decreased expression of H3.1-coding genes (Fig. 3d, Supplementary Fig. 1d, e)⁴⁵. Since H3.3 is not required for seed formation (Supplementary Figs. 1a and 3), we reasoned that H3.3 could become critical at the late seed maturation stage soon before seeds turn fully mature to prepare them for germination. To test this possibility, we expressed H3.3 in *h3.3ko* with the promoter of *at2S3* (a seed storage protein gene) or *H2B.S* (a seed-specific H2B variant gene), both of which are expressed specifically at the late stages of seed development^{44,46}. Expressing *HTR5* under either promoter increased the germination rate of *h3.3ko*, with the *H2B.S* promoter showing stronger capability (Fig. 3e, f and Supplementary Fig. 4a). However, in

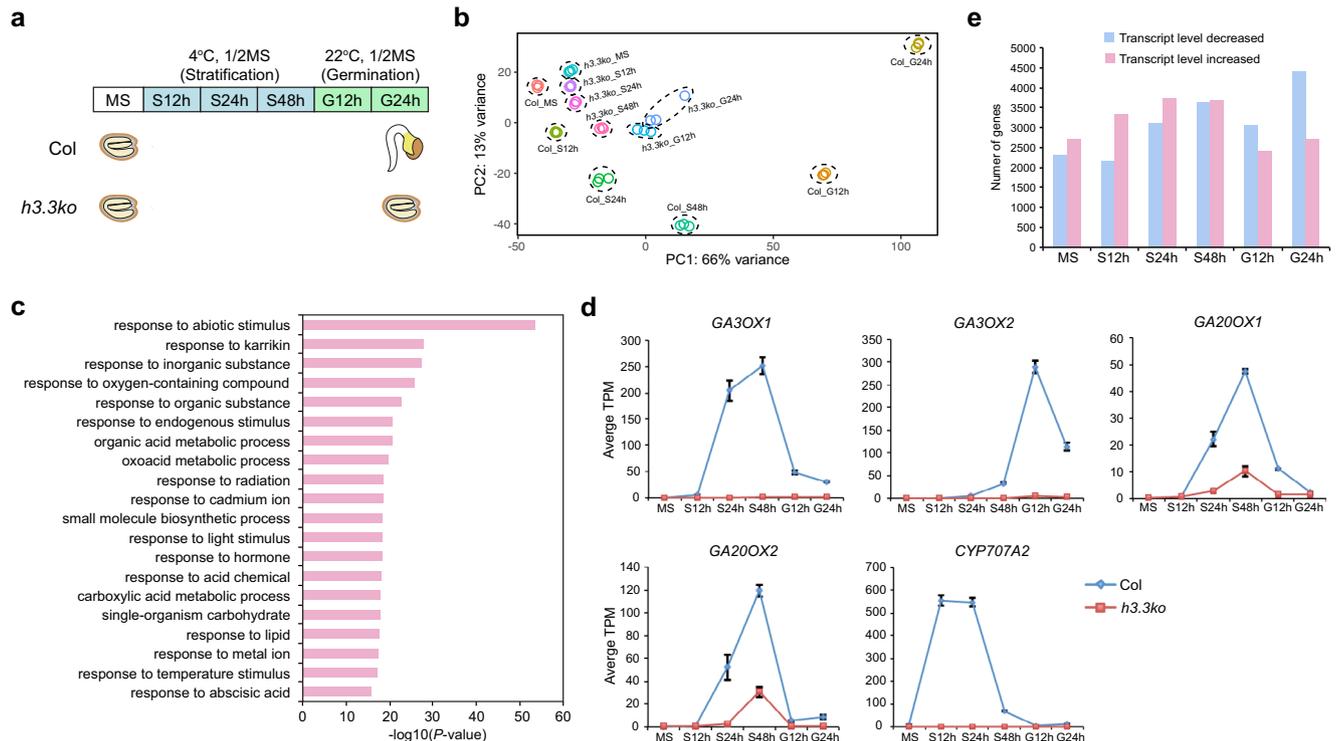


Fig. 2 | Transcriptome changes in *h3.3ko*. **a** The germination time course in this study. MS: mature seeds, S12h/S24h/S48h: stratified for 12/24/48 h, G12h/G24h: subjected to germination for 12/24 h. At G24h, the germination of Col but not *h3.3ko* has been completed. **b** PCA plot of Col and *h3.3ko* seeds imbibed for different times showing their transcriptome differences determined by RNA-seq. PC1 covers the highest amount of variance between samples; PC2 covers most of the remaining variance. Three biological replicates were performed for each line at each time point. **c** GO analysis of a combined set of all significantly misexpressed

genes in *h3.3ko* identified from each imbibition time point. Top 20 representative terms are listed and ranked by *P* value, which was calculated with hypergeometric test. **d** Transcript levels of *GA3OX1*, *GA3OX2*, *GA20OX1*, *GA20OX2* and *CYP707A2* during imbibition determined by RNA-seq (TPM) in Col and *h3.3ko*. Values are means \pm SD of three biological replicates. **e** Number of transcript levels significantly decreased and increased genes in *h3.3ko* compared with Col at different imbibition time points.

both cases, most of the seedlings ceased development after germination, similar to the *h3.3ko* seedlings. This is probably because the seedling development also requires H3.3 expressed during and after germination. Nevertheless, these results further support the notion that H3.3 expressed during seed maturation is necessary for seed germination.

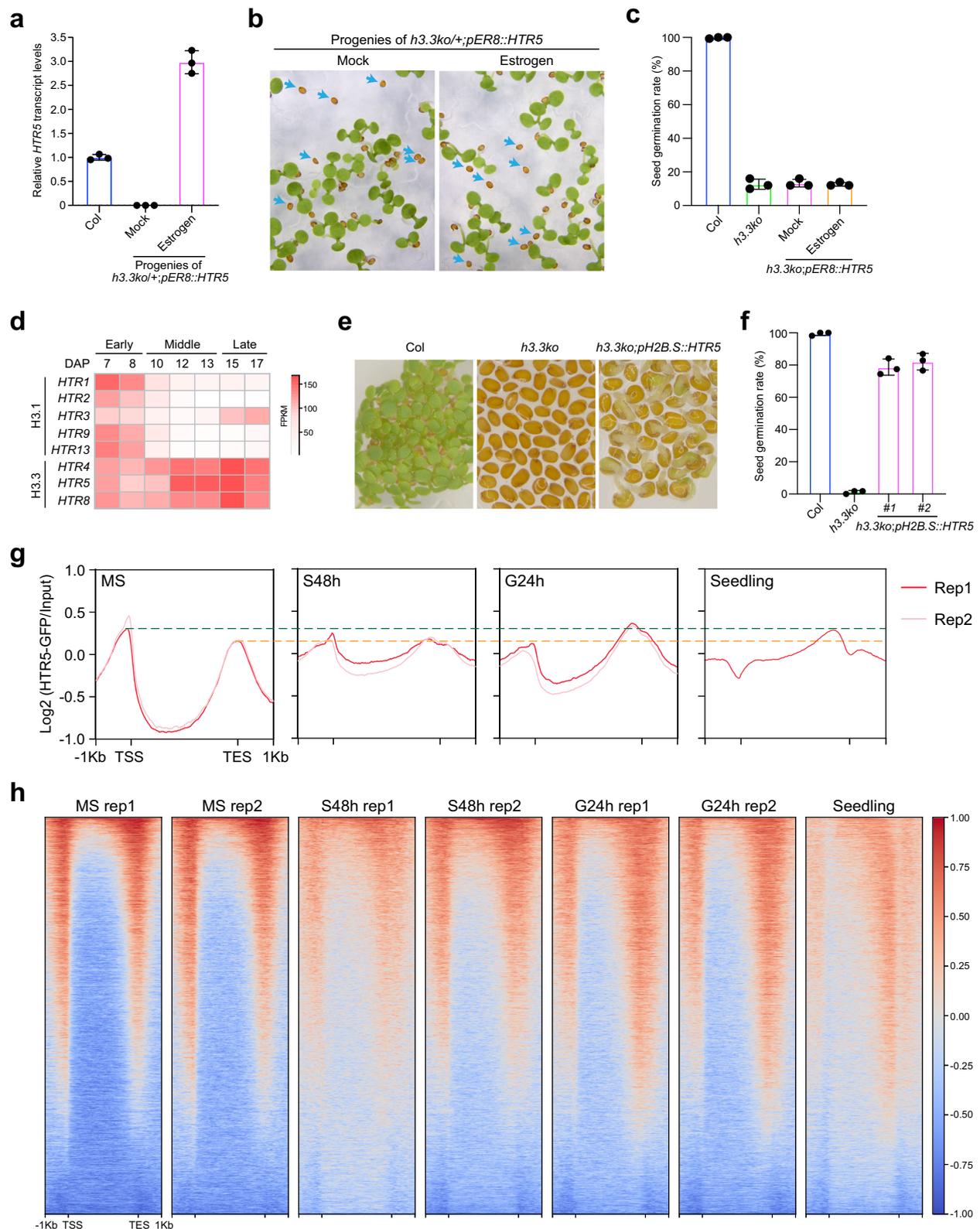
Ideally, a comparison of WT and *h3.3ko* at the molecular level (e.g. RNA-seq) during seed maturation would help the identification of the earliest stage when *h3.3ko* starts to show defects. However, the *h3.3ko* seeds were morphologically identical to WT (Supplementary Figs. 1a and 3), making them impracticable to isolate until they were fully mature (Supplementary Fig. 1f). To this end, we focused on the mature seed stage, and profiled the genome-wide accumulation of H3.3 with an HTR5-GFP reporter line using ChIP-seq³². H3.3 was localized at euchromatin regions and depleted over transposable elements (TEs) in mature seeds (Supplementary Fig. 4b, c), a pattern similar to that in vegetative tissues^{32,33,35}. Metagene analyses showed enrichment of H3.3 at both 5' gene ends/promoters and 3' gene ends (Fig. 3g, h and Supplementary Fig. 4d, e). Notably, the strong accumulation of H3.3 around the 5' gene ends was not detected previously in vegetative tissues^{32,33,35}. We thus further analyzed H3.3 enrichment during imbibition and in 10-day-old seedlings and observed a change in H3.3 enrichment patterns. In particular, H3.3 accumulation was gradually reduced at the 5' gene end but increased at the gene body and the 3' gene end (Fig. 3g, h and Supplementary Fig. 4d, e). These results suggest that H3.3 supplied during imbibition could be predominantly deposited to the gene body and the 3' gene end but not the 5' gene end, and probably explain why induction of *HTR5* expression during imbibition was not sufficient to rescue the germination defects of

h3.3ko (Fig. 3a–c). Together, we conclude that H3.3 enrichment around the 5' gene end is specific to seeds, suggesting its importance at this stage of development.

H3.3 regulates chromatin accessibility in mature seeds

The enrichment of H3.3 at the 5' gene ends coincides with the localization of chromatin accessible regions around the gene transcription start sites^{47,48}, we thus examined chromatin accessibility changes in *h3.3ko* mature seeds by Assay for Transposase Accessible Chromatin sequencing (ATAC-seq). Overall, accessible regions were mainly enriched at euchromatin in both WT and *h3.3ko* (Fig. 4a). Chromatin accessibility over TEs was not changed in *h3.3ko* (Supplementary Fig. 5a), in line with the low H3.3 abundance at TEs. At genes, the majority of the open chromatin regions localized at 5' gene ends/promoters in mature WT seeds (Fig. 4b–d and Supplementary Fig. 5b, c), similar to findings in vegetative tissues^{47,48}. However, compared with WT, the chromatin accessibility was overall lower around the 5' gene ends in *h3.3ko* and higher towards the 3' ends (Fig. 4b–e and Supplementary Fig. 5b, c). Hence, loss of H3.3 results in a global change of chromatin accessibility at genes in mature seeds.

We further identified regions that had significantly altered chromatin accessibility in *h3.3ko* (fold change >2 and *P* adjust <0.05). 3278 and 3436 regions significantly lost or gained accessibility respectively (Supplementary Fig. 6a, b) (Supplementary Data 2). Less accessible regions in *h3.3ko* were preferentially localized at promoters, but regions with increased chromatin accessibility were mainly enriched at genic regions and 3' gene ends (Supplementary Fig. 6c), parallel with the chromatin accessibility changes in *h3.3ko* over genes (Fig. 4b–e). Accessibility increased and decreased regions in *h3.3ko* were both



enriched with H3.3 in WT (Fig. 4e–g and Supplementary Fig. 6d, e), supporting the notion that H3.3 directly or indirectly regulates chromatin accessibility at these regions.

Notably, 5' and 3' localized H3.3 seems to have distinct impacts on chromatin accessibility. In animals, the coexistence of H3.3 and another histone variant H2A.Z destabilizes nucleosomes⁴⁹. We thus profiled the enrichment of H2A.Z in mature seeds. Like in vegetative

tissues, H2A.Z in mature seeds is enriched at the 5' gene ends but depleted from the 3' gene ends (Supplementary Fig. 7a)^{50,51}. Moreover, accessibility decreased regions in *h3.3ko* carried higher levels of H2A.Z than accessibility increased regions (Supplementary Fig. 7b, c). To assess the requirement of H2A.Z in chromatin accessibility control, we performed ATAC-seq using mature seeds of WT and a hypomorphic *h2a.z* mutant, in which *HTA9*, one of the three H2A.Z-coding genes in

Fig. 3 | H3.3 in seeds is essential for germination. **a** The induction of *HTR5* by estrogen during imbibition. Seeds were imbibed with or without estrogen for 48 h. Values are means \pm SD of three biological replicates. *PP2A* was used as an endogenous control for normalization. **b** The germination phenotype of *h3.3ko/+;pER8::HTR5* progenies treated with or without estrogen during imbibition. Arrows indicate *h3.3ko;pER8::HTR5* seeds that are strongly delayed in germination. **c** Percentages of *h3.3ko;pER8::HTR5* seeds that germinated within 2 month of imbibition with or without estrogen. Values are means \pm SD of three biological replicates. 50 seeds were assessed for each replicate. **d** Heatmap showing the transcript levels of H3.1 and H3.3-coding genes in dissected embryos at different

(early/middle/late) developmental stages. DAP: day after pollination. The expression data were extracted from the published RNA-seq datasets⁴⁵. **e** The germination phenotypes of Col, *h3.3ko*, and *h3.3ko;pH2B.S::HTR5* after subjected to germination for 84 h. **f** Percentages of *h3.3ko;pH2B.S::HTR5* seeds that germinated within 10 days of imbibition. Values are means \pm SD of three biological replicates. At least 38 seeds were assessed for each replicate. **g** Metaplot of HTR5-GFP ChIP-seq signals over all genes in mature seeds, during imbibition, and in the 10-day-old seedling. TSS: transcription start sites, TES: transcription end sites. **h** Heatmap of HTR5-GFP ChIP-seq signals over all genes in mature seeds, during imbibition, and in the 10-day-old seedling.

Arabidopsis, is still expressed at low levels^{50,52}. In *h2a.z*, a slight reduction of chromatin accessibility was observed at the 5' gene end, while chromatin accessibility at the 3' gene end was not changed compared with WT (Supplementary Fig. 8a). Importantly, regions with reduced accessibility in *h3.3ko* partially lost accessibility in *h2a.z* (Supplementary Fig. 8b), but loss of H2A.Z did not induce the chromatin accessibility at regions with increased accessibility in *h3.3ko* (Supplementary Fig. 8c). In addition, we observed moderately delayed germination of *h2a.z* compared with WT (Supplementary Fig. 8d). These results indicate that H2A.Z probably works together with H3.3 at the 5' gene end to facilitate chromatin opening in mature seeds, while without H2A.Z at the 3' gene end H3.3 somehow restricts chromatin accessibility. It is of note that the chromatin accessibility changes in the *h2a.z* mutant were less pronounced compared with that in *h3.3ko*. This could be because there is still residual H2A.Z in this *h2a.z* mutant, or other factors such as histone modifications enriched specifically at the 5' or 3' gene end may also contribute to the accessibility control together with H3.3.

Altered chromatin accessibility in *h3.3ko* correlates with changes in transcription and gene body DNA methylation

As mentioned above, the mature seed stage is the earliest time point when we could distinguish *h3.3ko* from WT (Supplementary Figs. 1a, f and 3). At this stage, the transcriptome of *h3.3ko* was already drastically different from that of WT (Fig. 2b, e). We then analyzed the correlations between changes in chromatin accessibility and transcriptome in mature *h3.3ko* seeds. In general, the distribution patterns of transcripts on genes were slightly altered in *h3.3ko* compared with WT, that the transcript levels were reduced around the 5' gene end, and were increased towards the 3' gene end (Fig. 5a, b, Supplementary Fig. 9a, b).

Next, we examined the transcripts of genes with significantly changed accessibility in *h3.3ko*. 4043 and 4388 genes were associated with accessibility decreased and increased regions respectively; with 920 genes associating with both (Supplementary Fig. 9c) (Supplementary Data 3). In *h3.3ko*, genes with accessibility decreased regions showed a strong reduction in chromatin accessibility around the 5' end, and genes with accessibility increased regions mainly displayed a gain of accessibility towards the 3' end (Supplementary Fig. 9d, e). These changes are consistent with the localization of accessibility decreased and increased regions in *h3.3ko* (Supplementary Fig. 6c). Overall, around accessibility decreased regions, fewer transcripts were detected in *h3.3ko* compared with WT (Fig. 5c), and genes associated with accessibility decreased regions showed an overall reduction in transcript levels (Fig. 5d). A detailed analysis revealed that a significant amount of them were indeed transcript level strongly decreased genes in mature *h3.3ko* seeds (Figs. 2e and 5e). Notably, some accessibility decreased region-associated genes showed increased transcript levels in *h3.3ko* compared with WT (Figs. 2e and 5e). Chromatin accessibility has also been found to link with gene repression^{53,54}, which may involve chromatin binding proteins such as transcription factors acting as repressors^{55,56}. 2928 accessibility decreased region-associated genes did not show strong transcript level changes in mature *h3.3ko* seeds

compared with WT (Fig. 5e). Hence, these genes may bear regulatory capacity mainly necessary in later developmental stages.

At accessibility increased regions, transcript levels were overall higher in *h3.3ko* (Fig. 5c), and accessibility increased region-associated genes had more transcripts (Fig. 5d). Nevertheless, these transcripts produced in *h3.3ko* were unevenly distributed, with more being generated towards the 3' gene end (Fig. 5d), indicating cryptic transcription. If this is the case, RNA Polymerase II (RNA Pol II) may randomly transcribe from both sense and antisense directions within these regions. We thus performed strand-specific RNA-seq to measure antisense transcripts in WT and *h3.3ko* mature seeds. The overall levels of antisense transcripts on genes were not increased in *h3.3ko* compared with WT (Supplementary Fig. 10a). However, at accessibility increased regions in *h3.3ko*, antisense transcripts were increased as well as sense transcripts (Supplementary Fig. 10b). Similarly, accessibility increased region-associated genes showed increased transcript levels from both sense and antisense, especially towards the 3' gene end (Supplementary Fig. 10c). In mature seeds, RNA Pol II may remain at gene loci or still transcribe at a low rate⁵⁷. We further selected a few genes showing increased chromatin accessibility and transcript levels towards the 3' gene end and measured RNA Pol II enrichment levels at their loci by ChIP-qPCR in mature seeds (Fig. 5f). The results showed that the RNA Pol II enrichment levels were increased around the 3' gene ends (Fig. 5g). Hence, the induction of transcripts at the 3' gene end in *h3.3ko* is likely due to spurious transcription at both sense and antisense directions by RNA Pol II.

Knockdown of H3.3 induces the loss of CG DNA methylation at the gene body, especially towards the 3' gene end³⁴. Using genome-wide bisulfite sequencing (BS-seq), we confirmed a strong reduction of gene body CG methylation in mature *h3.3ko* seeds. In addition, CHG and CHH methylation levels at the gene body were also slightly reduced (Supplementary Fig. 11a). At TE regions, only a reduction of CHH methylation levels was observed in *h3.3ko* (Supplementary Fig. 11b). H3.3 is largely absent from TEs, and therefore the loss of CHH methylation at TEs could be caused by an indirect effect. Accessibility increased but not decreased regions in *h3.3ko* showed reduced CG methylation levels (Fig. 5g, h). Consistently, hypo CG differentially methylated regions (DMR) in *h3.3ko* displayed increased chromatin accessibility (Supplementary Fig. 11c). It is unlikely that loss of gene body CG methylation in *h3.3ko* causes increased chromatin accessibility since a complete loss of CG methylation in *met1* fails to induce chromatin accessibility at the gene body⁵⁸. Together, these results suggest that H3.3-mediated chromatin accessibility control is probably required for proper transcriptional regulation and gene body DNA methylation.

Loss of chromatin accessibility in *h3.3ko* is associated with gene misregulation during germination

The above results indicate that H3.3 in seeds regulates chromatin accessibility and impacts on transcription. The enrichment of H3.3 around the 5' gene end was gradually diminished along germination and vegetative growth (Fig. 3f, g). We thus examined chromatin accessibility in WT during germination and in seedlings, and found that the overall chromatin accessibility at the 5' gene end was not strongly

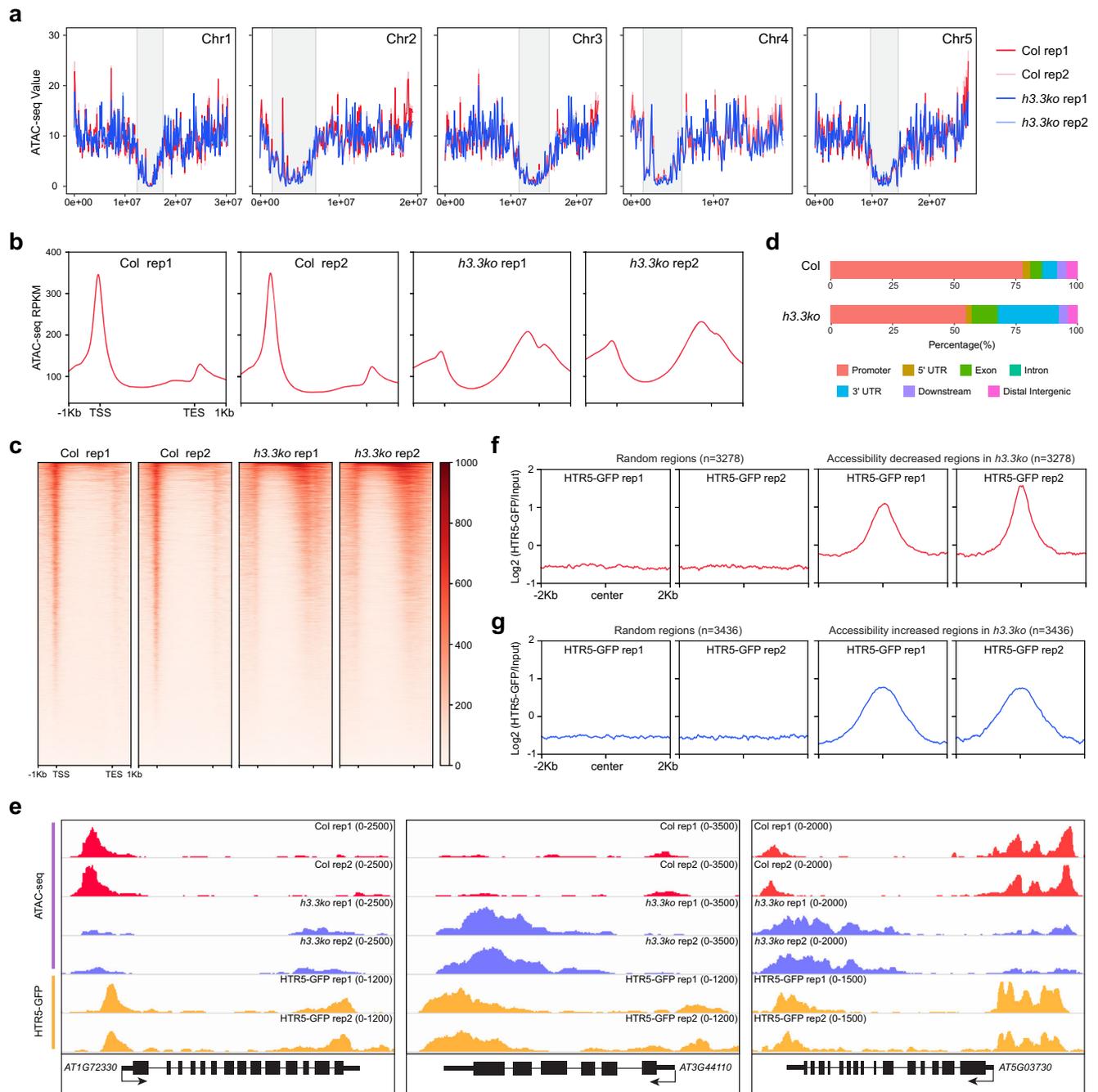


Fig. 4 | Chromatin accessibility changes in mature *h3.3ko* seeds. **a** ATAC-seq signals in Col and *h3.3ko* mature seeds over *Arabidopsis* chromosomes. Signals were calculated in 100 kb bins. Pericentromeric heterochromatin regions are indicated with gray shading. **b** Metaplot of ATAC-seq signals in Col and *h3.3ko* mature seeds over all genes. **c** Heatmap of ATAC-seq signals in Col and *h3.3ko* mature seeds over all genes. **d** Genomic distribution of accessible regions in Col and *h3.3ko* mature seeds. **e** Genome browser view of ATAC-seq and HTR5-GFP ChIP-seq signals in Col and *h3.3ko*

mature seeds over a few representative loci. *AT1G72330*: a gene mainly lost accessibility at the 5' end/promoter region in *h3.3ko*, *AT3G44110*: a gene mainly gained accessibility around the 3' end/genic region in *h3.3ko*, *AT5G03730*: a gene with both accessibility increased and decreased regions in *h3.3ko*. **f**, **g** Metaplot of HTR5-GFP ChIP-seq signals in mature seeds over accessibility significantly decreased (**f**) and increased (**g**) regions in mature *h3.3ko* seeds compared with Col. HTR5-GFP ChIP-seq signals over randomly selected regions are served as control.

changed (Supplementary Fig. 12a). We then focused on the H3.3-dependent open chromatin regions in mature seeds (accessibility significantly decreased regions in mature *h3.3ko* seeds that preferentially enriched around the 5' gene end). At these regions, H3.3 enrichment levels were drastically reduced during germination and in seedlings, yet this was only coupled with a moderate reduction in chromatin accessibility (Fig. 6a, b). In *h3.3ko*, the accessibility at these regions remained low after imbibition (Fig. 6c and Supplementary Fig. 12b). These results indicate that the H3.3-dependent chromatin

accessibility in seeds is largely maintained during germination and subsequent development.

In mature seeds, many genes associated with the H3.3-dependent open chromatin regions did not show strong transcript level changes in *h3.3ko* compared with WT (Fig. 5e). To understand the importance of these H3.3-dependent open chromatin regions in transcriptional regulation during germination, we compared the transcript levels of their associated genes in WT and *h3.3ko* during imbibition. These genes were grouped into five clusters based on their expression

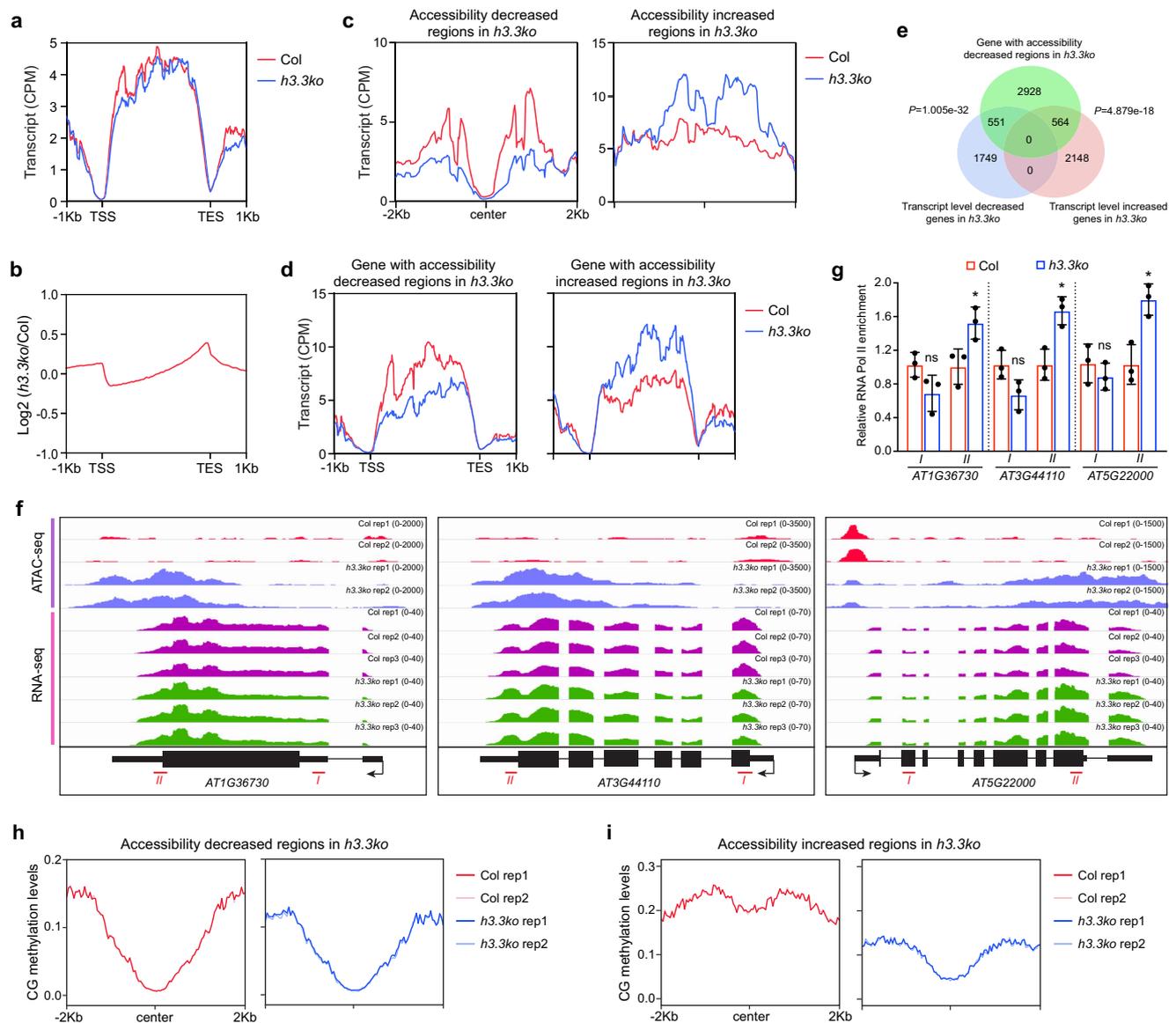


Fig. 5 | Altered chromatin accessibility in *h3.3ko* correlates with changes in transcription and gene body DNA methylation. **a** Metaplot of transcripts in Col and *h3.3ko* mature seeds over all genes. The profiles were generated after merging three biological replicates. **b** Metaplot of *h3.3ko* transcripts vs Col transcripts in mature seeds over all genes. The profiles were generated after merging three biological replicates. **c** Metaplot of transcripts in Col and *h3.3ko* mature seeds over accessibility significantly decreased and increased regions in mature *h3.3ko* seeds compared with Col. The profiles were generated after merging three biological replicates. **d** Metaplot of transcripts in Col and *h3.3ko* mature seeds over genes with accessibility significantly decreased and increased regions in mature *h3.3ko* seeds compared with Col. The profiles were generated after merging three biological replicates. **e** Venn diagrams of accessibility significantly decreased region-associated genes in mature *h3.3ko* seeds and transcript level significantly decreased

and increased genes in mature *h3.3ko* seeds. *P* values are based on the hypergeometric test. **f** Genome browser view of ATAC-seq and RNA-seq signals in Col and *h3.3ko* mature seeds over a few representative loci showing increased chromatin accessibility and transcript levels around the 3' gene end in *h3.3ko* compared with Col. Red lines indicate regions examined by ChIP-qPCR. **g** Relative RNA Pol II enrichment levels determined by ChIP-qPCR at *AT1G36730*, *AT3G44110*, and *AT5G22000* in mature seeds. RNA Pol II enrichment levels were first normalized to input and then to that of Col. Values are means \pm SD of three biological replicates. Statistical significance was determined by two-tailed Student's *t* test (**P* < 0.05; ns, *P* > 0.05). **h**, **i** Metaplot of CG methylation levels in Col and *h3.3ko* mature seeds over accessibility significantly decreased (**h**) and increased (**i**) regions in mature *h3.3ko* seeds compared with Col.

patterns during imbibition in WT (Fig. 6d). Importantly, in *h3.3ko*, most of the genes in clusters 1–4 failed to fully activate during imbibition, while those in cluster 5 were not sufficiently repressed (Fig. 6d). GO analysis of these genes revealed mainly responsive processes and post-embryonic development (Fig. 6e). Hence, H3.3-dependent chromatin accessibility at these responsive and developmental genes is likely necessary for their transcriptional control during germination.

At regions where the accessibility was repressed by H3.3 in mature seeds (mainly localized at gene body and 3' gene end), H3.3 remained deposited in WT during imbibition and in seedlings and is required to

maintain their low chromatin accessibility and to repress cryptic transcription at these regions (Fig. 6f and Supplementary Fig. 12c–e), indicating a canonical H3.3 distribution pattern around the 3' gene end that constantly prevents chromatin opening and cryptic transcription at genic regions.

Discussion

In *Arabidopsis*, both paternal and maternal-derived H3.3 is removed from the zygote genome a few hours after fertilization, followed by de novo synthesis of H3.3 before the zygote divides. This process was

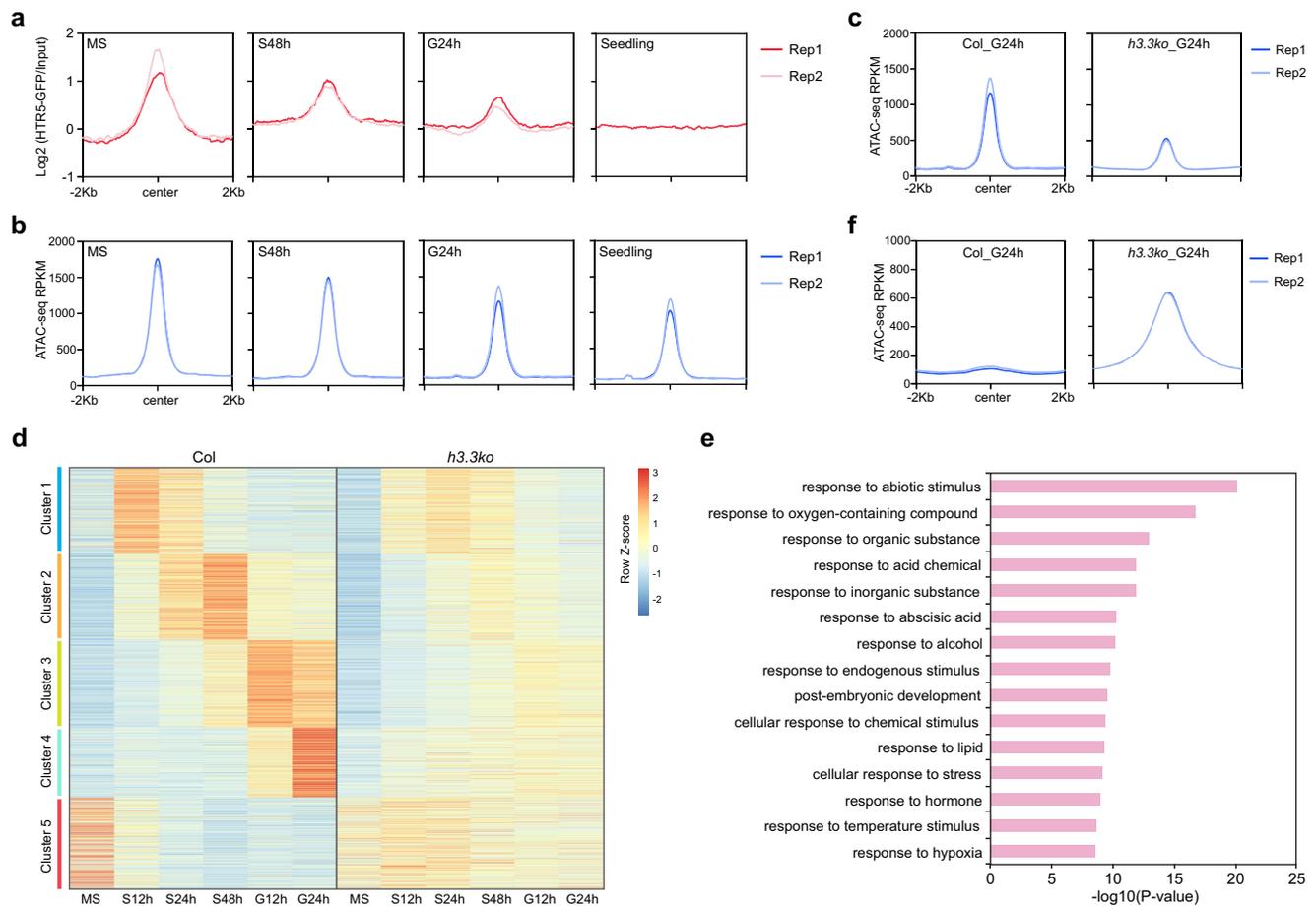


Fig. 6 | Loss of chromatin accessibility in *h3.3ko* is associated with gene misregulation during germination. **a, b** Metaplot of HTR5-GFP ChIP-seq (a) and ATAC-seq (b) signals in mature seeds, during imbibition and in the 10-day-old seedlings over accessibility decreased regions in mature *h3.3ko* seeds compared with Col. **c** Metaplot of ATAC-seq signals in imbibed Col and *h3.3ko* seeds over accessibility significantly decreased regions in mature *h3.3ko* seeds compared with Col. **d** Heatmap showing expression profiles (z-normalized) of genes associated with

accessibility significantly decreased regions in mature *h3.3ko* seeds compared with Col. Genes were grouped into five clusters based on their expression patterns during imbibition in Col. **e** GO analysis of genes associated with accessibility significantly decreased regions in mature *h3.3ko* seeds compared with Col. Top 15 representative terms are listed and ranked by *P* value, which was calculated with hypergeometric test. **f** Metaplot of ATAC-seq signals in imbibed Col and *h3.3ko* seeds over accessibility significantly increased regions in mature *h3.3ko* seeds compared with Col.

speculated to prevent the transmission of epigenetic memory from parents and meanwhile establish the zygotic epigenome critical for embryogenesis^{42,59,60}. Interestingly, our genetic evidence suggests that H3.3 is not essential for early plant embryonic development. It is likely that H3.1 or other H3 proteins largely compensate for the loss of H3.3 at this stage, especially considering that H3.1 is abundantly produced during cell division, a dominant event in early embryogenesis⁶¹. After pattern formation at the torpedo stage, the embryo gradually ceases cell division and initiates maturation². This is coupled with a strong reduction in H3.1 expression and an increase in H3.3 expression (Fig. 3d), which coincident with the extensive H3.1 replacement by H3.3 during embryo maturation⁶². Thus, H3.3 likely starts to be critical at embryo maturation stages, and reprograms the embryo epigenome for the establishment of post-embryonic developmental potentials. Consistently, embryos before torpedo stages are not capable to germinate when excised from seeds, while later stage embryos could germinate². This role of H3.3 is comparable to the function of a specialized form of non-replicative H3 variant H3.10 in sperm that establishes chromatin accessibility at promoters of genes expressed in early embryogenesis^{9,59}. It is unknown by which mechanism H3.3 is deposited to the 5' gene ends in seeds but removed during germination and seedling development. This possibly involves the H3.3 deposition chaperone HIRA⁶³ that plays a moderate function in seed germination⁶⁴, or ATRX involved in the deposition of H3.3^{65,66}. In

addition, specific chromatin remodelers might recruit H3.3 deposition machinery. Alternatively, removal of H3.1 from this region during embryo maturation may allow H3.3 to accumulate.

Several pieces of evidence suggest that seed germination is largely programmed during seed maturation⁶⁷. Protein synthesis from stored mRNA in mature seeds is required for the completion of germination⁶⁸⁻⁷⁰. Moreover, mRNA abundance changes are observed within 1–2 h after imbibition without de novo protein synthesis required, highlighting the function of seed-stored transcription machinery in quickly adjusting transcription⁷¹. Our results suggest that H3.3 deposited in seeds is vital for their acquisition of germination capacity. Loss of H3.3 causes reduced and increased chromatin accessibility at 5' and 3' gene ends respectively, and results in impaired regulation of responsive and post-embryonic development genes and increased cryptic transcription. Since H3.3 is part of the chromatin core (nucleosome), it is unlikely that H3.3 impacts on other physiological processes that would in turn affect chromatin accessibility. Therefore, we propose the following model (Fig. 7): H3.3 is loaded at the 5' gene ends/promoters in mature seeds, facilitating chromatin opening (Figs. 3f, g and 4). During germination and subsequent seedling development, 5' enriched H3.3 is gradually lost. However, the chromatin accessibility is still maintained (Fig. 6). A plausible explanation is that the 5' enriched H3.3 may initiate chromatin opening to enable transcription factor binding in seeds and during germination.

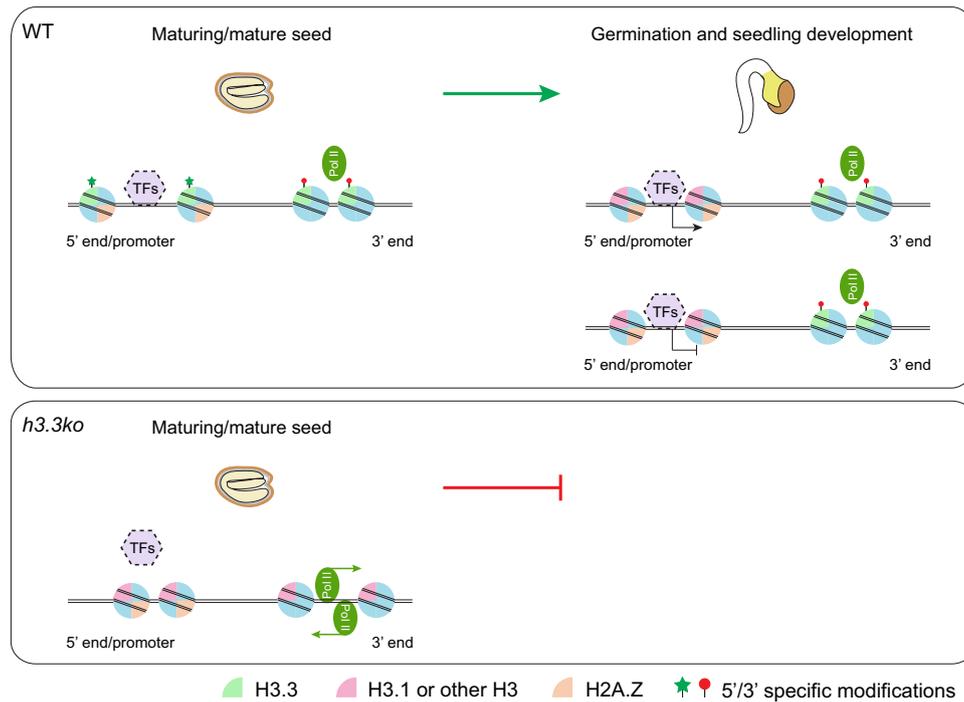


Fig. 7 | A proposed function of H3.3 in chromatin accessibility regulation. In maturing and mature seeds, H3.3 is loaded at the 5' gene ends/promoters and establishes chromatin accessibility. During germination and seedling development, H3.3 is gradually removed from the 5' gene ends, but the accessibility established by H3.3 is largely maintained probably by the binding of transcription factors (as indicated by the dashed border), and this may allow proper gene transcriptional regulation (activation or repression) during post-

embryonic development. On the other hand, H3.3 is constantly deposited around the 3' gene ends and prevents chromatin opening and cryptic transcription. Loss of H3.3 causes reduced and increased chromatin accessibility at 5' and 3' gene ends respectively, and results in compromised transcriptional regulation and severely impaired post-embryonic development. H2A.Z and the 5' and 3' specific histone modifications may also contribute to the H3.3-mediated chromatin accessibility control.

Afterward, the binding of transcription factors could sustain chromatin accessibility without H3.3. We found that the H3.3-established open chromatin regions are enriched with the transcription factor binding motifs especially those of bZIP, bHLH and BZR/BES family members (Supplementary Fig. 13 and Supplementary Data 4). On the other hand, H3.3 is constantly deposited around the 3' gene ends and prevents chromatin accessibility and cryptic transcription. Nevertheless, it is still unclear whether the altered chromatin accessibility in *h3.3ko* is the direct cause of transcription changes and the impaired post-embryonic development, and the exact role of H3.3 in chromatin regulation and post-embryonic development remains to be investigated.

As a replacement histone, H3.3 deposition has been proposed to alter the nucleosome property^{4,12}. In human cells, the presence of H3.3 destabilizes nucleosome, especially when co-existed with H2A.Z, leading to a nucleosome-free region at the 5' gene end for transcription factor binding^{49,72}. In mouse embryonic stem cell, H3.3 and H2A.Z collaborate around gene transcription start sites to set up a poised chromatin state for gene activation during cell differentiation⁷³. Despite a predominant localization of H3.3 at the 3' gene end in *Arabidopsis* vegetative tissues^{32,33,35}, our findings show that at certain developmental stages it is highly enriched around the 5' gene end and facilitates chromatin opening. With fully developed leaves that have ceased cell division and expansion, a small fraction of H3.3 is also detected at the 5' gene ends/promoters⁷⁴. Besides in mature embryo, a considerable amount of H3.3 is deposited to replace H3.1 in plant cells that have exited proliferation and become terminal differentiated⁶². This implies a broad role of H3.3 in reprogramming the epigenome for cell fate determination. Further study of detailed mechanisms might provide a potential to manipulate chromatin accessibility and cellular differentiation by engineered H3.3 deposition.

Methods

Plant materials and growth conditions

Arabidopsis h3.3ko and HTR5-GFP lines were described previously^{32,34}. Plants were grown in long-day conditions (16 h light/8 h dark) at 22 °C. For seed germination, after-ripened mature seeds (seeds harvested and stored at room temperature for three months) were sown on 1/2 MS medium. After being kept at 4 °C for two days (stratification), seeds were subjected for germination in long days at -22 °C. Three replicates were performed to assess the germination and growth of *h3.3ko*, at least 110 seeds were used for each replicate. *h3.3ko* seeds germinated within two months after sowing were considered as germinated, and the rest seeds were discarded. The germination date of the first 12 germinated seeds was scored.

For dissected embryo growth analysis, embryos from after-ripened Col and *h3.3ko* seeds were isolated by manually removing endosperm and testa. The isolated embryos were cultivated on 1/2 MS medium under long-day conditions.

Plasmid construction and *h3.3ko* complementation

For *h3.3ko* complementation constructs, *HTR5* or *HTR13* coding sequences were placed under the control of *HTR5* or *HTR13* promoter in the binary vector pBlligatorR43, which contains a mature seed stage-specific mCherry fluorescence selection marker³⁶. For the estradiol induction, *HTR5* coding sequences were placed under the *pER8* promoter in the binary vector pMDC7^{43,75}. To express *HTR5* with *at2S3* or *H2B.S* promoter, the *at2S3* or *H2B.S* promoter sequences were first cloned into pGWBS10⁷⁶, followed by the insertion of *HTR5* coding sequences.

For the germination test of *h3.3ko;pER8::HTR5* seeds with or without estrogen (10 μM) treatment during imbibition, 50 seeds were assessed in each replicate and seeds were transferred onto a new 1/2

MS plate supplemented with or without estrogen every 5 days in case of estrogen degradation. The germination rates were scored after 2 month of imbibition. For the germination test of *h3.3ko;pat2S3::HTR5* and *h3.3ko;pH2B.S::HTR5* seeds, germination rates were scored after 1 month and 10 days imbibition respectively. At least 38 seeds were assessed in each replicate.

Seed size measurement

Complete dried mature seeds were spread on a white background. After the image taken, seed size was measured with ImageJ. At least 47 seeds were scored for each genotype.

Seed storage protein analysis

One hundred and fifty Col or *h3.3ko* mature seeds were homogenized in protein extraction buffer (100 mM Tris-HCl pH 8.0, 0.5% SDS, 10% glycerol, 2% β -ME). Extracted proteins were resolved on 15% SDS-PAGE gel and stained with Coomassie Blue.

RNA-seq

Total RNA was extracted from mature seeds (seeds stored for three months after harvesting) or imbibed seeds using Minibest plant RNA extraction kit (Takara). Three independent biological replicates were performed. Sequencing libraries were prepared with BGISEQ-500 RNA-seq library preparation kit according to the manufacturer's instruction. For strand-specific RNA-seq, total RNA extracted from Col and *h3.3ko* mature seeds were subjected to library preparation with MGIEasy RNA directional library preparation kit. Prepared libraries were sequenced on a DNBSEQ-T7 platform and paired-end 150 bp reads were generated.

RNA-seq data analysis

Adapter trimming was performed and low-quality reads were filtered with fastp version 0.20.1⁷⁷. Reads were mapped to the *Arabidopsis* genome (TAIR10) using Hisat2 version 2.1.10⁷⁸. The numbers of mapped reads were determined with SAMtools version 1.9⁷⁹, and are listed in Supplementary Table 1. Reads per gene were counted by HTseq version 0.11.2⁸⁰. Transcripts per million (TPM) values were generated using R. Differential gene expression analysis was performed using DESeq2 version 1.26.0⁸¹. Genes that displayed a more than two-fold expression change and had a *P* adjust value <0.05 were considered as differentially expressed. For data visualization, bigwig coverage files were generated using deepTools utility bamCoverage⁸² with a bin size of 10 bp and normalized with CPM. The PCA plot was generated with plotPCA in DESeq2 version 1.26.0. The expression of *HTR4*, *HTR5* and *HTR8* during embryogenesis was extracted from the published datasets⁴⁵. Gene ontology analysis was performed with DAVID (<https://david.ncifcrf.gov>)⁸³. Clustering of genes associated with accessibility decreased regions in mature *h3.3ko* seeds compared with Col was performed based on their expression changes during germination in Col using K-Means clustering in R.

RT-qPCR

Total RNA was extracted from mature seeds (seeds stored for three months after harvesting), imbibed seeds or 7-day-old seedlings using Minibest plant RNA extraction kit (Takara). Reverse transcription was performed using HiScript III 1st Strand cDNA Synthesis Kit (Vazyme, R312-02). Real-time quantitative PCR was conducted on an Applied Biosystems QuantStudio 6 Flex Real-Time PCR System using ChamQ Universal SYBR qPCR Master Mix (Vazyme, Q711-02). *PP2A* was used as an endogenous control for normalization. Three independent biological replicates were performed for each line and condition. Primers used to amplify *HTR5*, *HTR13*, *GA3OX1*, *GA3OX2*, *GA20OX1*, *GA20OX2*, *CYP707A2*, and *PP2A* are listed in Supplementary Table 2. The *HTR5* amplification primers were designed to only amplify the WT but not Crispr-mutated *HTR5*. Primers amplifying *HTR5* and *HTR13* were tested to have similar amplification efficiency.

ChIP-seq

ChIP was performed with mature HTR5-GFP seeds (seeds stored for three months after harvesting) or seeds imbibed for different times. Materials were ground with liquid nitrogen into fine powder and immediately fixed with 1% formaldehyde^{32,33,35,50–52}. Nuclei were extracted and mononucleosomes were generated with micrococcal nuclease (MNase) (Sigma, N5386) digestion as previously described⁵². Immunoprecipitation was conducted with anti-GFP (Thermo Fisher Scientific, A-11122) or anti-H2A.Z⁵¹ antibody. After antibody incubation, Protein A Dynabeads (Thermo Fisher Scientific, 10002D) were added to collect the immunocomplexes. After washing, elution, reverse cross-linking and DNA purification, recovered DNA was subjected to library preparation with VAHTS universal DNA library prep kit for Illumina (Vazyme, ND607) according to the manufacturer's instruction and sequenced with Illumina NovaSeq 6000 to generate paired-end 150 bp reads. ChIP-seq experiments were performed with two independent biological replicates.

ChIP-seq data analysis

Adapter trimming was performed and low-quality reads were filtered with fastp version 0.20.1⁷⁷. Reads were mapped to the *Arabidopsis* genome (TAIR10) with Bowtie2 version 2.4.2⁸⁴, and filtered for duplicated reads by using Picard version 2.24.0 MarkDuplicates (<https://github.com/broadinstitute/picard>). The numbers of mapped reads were determined with SAMtools version 1.9⁷⁹, and are listed in Supplementary Table 1. For data visualization, bigwig coverage files were generated using deepTools utility bamCoverage⁸² with a bin size of 10 bp and normalized to input or sequencing depth using RPKM. Visualization was performed with IGV version 2.7.2⁸⁵. Heatmaps and average normalized ChIP-seq profiles were generated using deepTools utilities plotHeatmap and plotProfile. Published HTR5-GFP ChIP-seq data in seedlings were processed in the same manner³⁵.

ChIP-qPCR

Mature seeds were ground with liquid nitrogen into fine powder and immediately fixed with 1% formaldehyde. Chromatin was sheared by sonication and immunoprecipitation was performed with anti-RNA Pol II antibody (abcam, ab26721). The amounts of immunoprecipitated DNA were quantified by real-time quantitative PCR. Three independent biological replicates were performed. Primers used for PCR amplification are listed in Supplementary Table 2.

ATAC-seq

Mature seeds (seeds stored for three months after harvesting), seeds imbibed for different times or 10-day-old seedlings were chopped in lysis buffer (15 mM Tris-HCl pH 7.5, 20 mM NaCl, 80 mM KCl, 0.5 mM Spermine, 0.2% Triton X-100, 5 mM β -ME, protease inhibitor cocktail). After being filtered with a 30 μ m filter (CellTrics), nuclei were stained with DAPI and subjected to fluorescence-activated cell sorting (FACS). 50,000 nuclei per sample were collected and mixed with Tn5 transposase (Mei5bio, MF650-01), after 20 min of incubation at 37 °C, tagged DNA was recovered with ChIP DNA Clean & Concentrator kit (ZYMO, D5205). After PCR amplification, sequencing libraries were purified with VAHTS DNA Clean Beads (Vazyme N411) and sequenced with Illumina NovaSeq 6000 to generate paired-end 150 bp reads. ATAC-seq experiments were performed with two independent biological replicates.

ATAC-seq data analysis

Adapter trimming was performed and low-quality reads were filtered with fastp version 0.20.1⁷⁷. Reads were mapped to the *Arabidopsis* genome (TAIR10) with Bowtie2 version 2.4.2⁸⁴. Reads mapped to chloroplast and mitochondria genome were removed and duplicated reads were filtered with Picard version 2.24.0 MarkDuplicates (<https://github.com/broadinstitute/picard>). The numbers of mapped reads were

determined with SAMtools version 1.9⁷⁹, and are listed in Supplementary Table 1. ATAC-seq open chromatin peaks were identified using MACS2 version 2.1.2 with default parameters⁸⁶. The q-value cutoff for peak calling was 0.05. Only peaks identified from both biological replicates were kept. Peak distributions were analyzed with ChIPseeker version 1.22.1⁸⁷. Promoter refers to accessible regions located <2 kb upstream of the transcription start sites (TSS), downstream refers to accessible regions located <1 kb downstream of the transcription end sites (TES), Intergenic region refers to accessible regions located >2 kb upstream of TSS and more than 1 kb downstream of TES. To identify differentially enriched peaks, a common peak set was created by merging peaks in Col and *h3.3ko*, scores for open chromatin peak regions were calculated with deepTools utility multiBigwigSummary⁸², and differential peaks were called by requiring more than two-fold difference and *P* adjust <0.05. Genes were assigned to the peaks when the genic and 1 kb upstream regions overlapped with the peak regions for at least one base pair. For data visualization, bigwig coverage files were generated using deepTools utility bamCoverage⁸² with a bin size of 10 bp and normalized with RPKM. DNA motif analysis was performed with HOMER (version 4.11) using “findMotifs.pl”⁸⁸.

BS-seq

Genomic DNA was extracted from mature seeds (seeds stored for three months after harvesting) with Quick-DNA Plant/Seed Miniprep Kit (ZYMO, D6020). 200 ng fragmented genomic DNA (200 bp–500 bp) was used for library preparation with VAHTS universal Pro DNA library prep kit for Illumina (Vazyme, ND608). NEBNext Multiplex Oligos for Illumina (NEB, E7535) were used for adapter ligation. After purification with VAHTS DNA Clean Beads (Vazyme N411), bisulfite conversion was performed with EZ DNA Methylation-Gold Kit (ZYMO, D5005), followed by DNA purification and PCR amplification. Libraries were sequenced with Illumina NovaSeq 6000 to generate paired-end 150 bp reads. BS-seq experiments were performed with two independent biological replicates.

BS-seq data analysis

Adapter trimming was performed and low-quality reads were filtered with fastp version 0.20.1⁷⁷. Reads were mapped to the *Arabidopsis* genome (TAIR10) with BS-Seeker2 version 2.1.8 using default parameters⁸⁹. Duplicated reads were filtered with Picard version 2.24.0 MarkDuplicates. The numbers of mapped reads are listed in Supplementary Table 1. CG, CHG, and CHH methylation levels were calculated with CGmappro (version 0.1.2)⁹⁰. Methylation data were visualized with ggplot2 in R (version 1.1.1106). CG DMRs were called with CGmappro using dmr function with default parameters.

Statistical analysis

Statistical tests performed are indicated in the figure captions.

Reporting summary

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

Data availability

The datasets generated during the current study are available in the GEO repository [GSE209645](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE209645). The ChIP-seq data of HTR5-GFP in seedlings were downloaded from [GSE167384](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE167384). Source data are provided with this paper.

References

- Leprince, O., Pellizzaro, A., Berriri, S. & Buitink, J. Late seed maturation: drying without dying. *J. Exp. Bot.* **68**, 827–841 (2017).
- Raz, V., Bergervoet, J. H. & Koornneef, M. Sequential steps for developmental arrest in *Arabidopsis* seeds. *Development* **128**, 243–252 (2001).
- Lai, W. K. M. & Pugh, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* **18**, 548–562 (2017).
- Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 178–189 (2015).
- Krishnakumar, R. & Blesch, R. H. Epigenetics of cellular reprogramming. *Curr. Opin. Genet. Dev.* **23**, 548–555 (2013).
- Li, D., Shu, X., Zhu, P. & Pei, D. Chromatin accessibility dynamics during cell fate reprogramming. *EMBO Rep.* **22**, e51644 (2021).
- Swygert, S. G. & Peterson, C. L. Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochim. Biophys. Acta* **1839**, 728–736 (2014).
- Talbert, P. B. & Henikoff, S. Histone variants on the move: substrates for chromatin dynamics. *Nat. Rev. Mol. Cell Biol.* **18**, 115–126 (2017).
- Borg, M., Jiang, D. & Berger, F. Histone variants take center stage in shaping the epigenome. *Curr. Opin. Plant Biol.* **61**, 101991 (2021).
- Martire, S. & Banaszynski, L. A. The roles of histone variants in fine-tuning chromatin organization and function. *Nat. Rev. Mol. Cell Biol.* **21**, 522–541 (2020).
- Probst, A. V., Desvoyes, B. & Gutierrez, C. Similar yet critically different: the distribution, dynamics and function of histone variants. *J. Exp. Bot.* **71**, 5191–5204 (2020).
- Henikoff, S. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* **9**, 15–26 (2008).
- Filipescu, D., Muller, S. & Almouzni, G. Histone H3 variants and their chaperones during development and disease: contributing to epigenetic control. *Annu. Rev. Cell Dev. Biol.* **30**, 615–646 (2014).
- Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9**, 1191–1200 (2002).
- Gaillard, P. H. L. et al. Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. *Cell* **86**, 887–896 (1996).
- Ray-Gallet, D. et al. HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol. Cell* **9**, 1091–1100 (2002).
- Smith, S. & Stillman, B. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* **58**, 15–25 (1989).
- Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51–61 (2004).
- Goldberg, A. D. et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678–691 (2010).
- Wirbelauer, C., Bell, O. & Schubeler, D. Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes Dev.* **19**, 1761–1766 (2005).
- Sitbon, D., Boyarchuk, E., Dingli, F., Loew, D. & Almouzni, G. Histone variant H3.3 residue S31 is essential for *Xenopus* gastrulation regardless of the deposition pathway. *Nat. Commun.* **11**, 1256 (2020).
- Szenker, E., Lacoste, N. & Almouzni, G. A developmental requirement for HIRA-dependent H3.3 deposition revealed at gastrulation in *Xenopus*. *Cell Rep.* **1**, 730–740 (2012).
- Bush, K. M. et al. Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. *Epigenet. Chromatin* **6**, 7 (2013).
- Couldrey, C., Carlton, M. B., Nolan, P. M., Colledge, W. H. & Evans, M. J. A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. *Hum. Mol. Genet.* **8**, 2489–2495 (1999).
- Tang, M. C. et al. Contribution of the two genes encoding histone variant h3.3 to viability and fertility in mice. *PLoS Genet.* **11**, e1004964 (2015).

26. Jang, C. W., Shibata, Y., Starmer, J., Yee, D. & Magnuson, T. Histone H3.3 maintains genome integrity during mammalian development. *Genes Dev.* **29**, 1377–1392 (2015).
27. Lin, C. J., Conti, M. & Ramalho-Santos, M. Histone variant H3.3 maintains a decondensed chromatin state essential for mouse pre-implantation development. *Development* **140**, 3624–3634 (2013).
28. Loppin, B. & Berger, F. Histone variants: the nexus of developmental decisions and epigenetic memory. *Annu. Rev. Genet.* **54**, 121–149 (2020).
29. Lu, L., Chen, X., Qian, S. & Zhong, X. The plant-specific histone residue Phe41 is important for genome-wide H3.1 distribution. *Nat. Commun.* **9**, 630 (2018).
30. Talbert, P. B. et al. A unified phylogeny-based nomenclature for histone variants. *Epigenet. Chromatin* **5**, 7 (2012).
31. Jiang, D. & Berger, F. Histone variants in plant transcriptional regulation. *Biochim. Biophys. Acta Gene Regul. Mech.* **1860**, 123–130 (2017).
32. Wollmann, H. et al. Dynamic deposition of histone variant H3.3 accompanies developmental remodeling of the Arabidopsis transcriptome. *PLoS Genet.* **8**, e1002658 (2012).
33. Stroud, H. et al. Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana. *Proc. Natl Acad. Sci. USA* **109**, 5370–5375 (2012).
34. Wollmann, H. et al. The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. *Genome Biol.* **18**, 94 (2017).
35. Zhao, F., Zhang, H., Zhao, T., Li, Z. & Jiang, D. The histone variant H3.3 promotes the active chromatin state to repress flowering in Arabidopsis. *Plant Physiol.* **186**, 2051–2063 (2021).
36. Kawashima, T. et al. Dynamic F-actin movement is essential for fertilization in Arabidopsis thaliana. *eLife* <https://doi.org/10.7554/eLife.04501> (2014).
37. Jiang, D. & Berger, F. DNA replication-coupled histone modification maintains Polycomb gene silencing in plants. *Science* **357**, 1146–1149 (2017).
38. Debeaujon, I., Leon-Kloosterziel, K. M. & Koornneef, M. Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. *Plant Physiol.* **122**, 403–414 (2000).
39. Holdsworth, M. J., Bentsink, L. & Soppe, W. J. J. Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *N. Phytologist* **179**, 33–54 (2008).
40. Ogawa, M. et al. Gibberellin biosynthesis and response during Arabidopsis seed germination. *Plant Cell* **15**, 1591–1604 (2003).
41. Okamoto, M. et al. CYP707A1 and CYP707A2, which encode abscisic acid 8'-hydroxylases, are indispensable for proper control of seed dormancy and germination in Arabidopsis. *Plant Physiol.* **141**, 97–107 (2006).
42. Ingouff, M. et al. Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in Arabidopsis. *Curr. Biol.* **20**, 2137–2143 (2010).
43. Zuo, J., Niu, Q. W. & Chua, N. H. Technical advance: an estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J.: Cell Mol. Biol.* **24**, 265–273 (2000).
44. Fujiwara, T., Nambara, E., Yamagishi, K., Goto, D. B. & Naito, S. Storage proteins. *Arabidopsis book / Am. Soc. Plant Biologists* **1**, e0020 (2002).
45. Schneider, A. et al. Potential targets of VIVIPAROUS1/ABI3-LIKE1 (VAL1) repression in developing Arabidopsis thaliana embryos. *Plant J.* **85**, 305–319 (2016).
46. Jiang, D. et al. The evolution and functional divergence of the histone H2B family in plants. *PLoS Genet.* **16**, e1008964 (2020).
47. Lu, Z., Hofmeister, B. T., Vollmers, C., DuBois, R. M. & Schmitz, R. J. Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Res.* **45**, e41 (2017).
48. Maher, K. A. et al. Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. *Plant Cell* **30**, 15–36 (2018).
49. Jin, C. Y. & Felsenfeld, G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* **21**, 1519–1529 (2007).
50. Coleman-Derr, D. & Zilberman, D. Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. *PLoS Genet.* **8**, e1002988 (2012).
51. Yelagandula, R. et al. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell* **158**, 98–109 (2014).
52. Xue, M. et al. The INO80 chromatin remodeling complex promotes thermomorphogenesis by connecting H2A.Z eviction and active transcription in Arabidopsis. *Mol. Plant* **14**, 1799–1813 (2021).
53. Maezawa, S., Yukawa, M., Alavattam, K. G., Barski, A. & Namekawa, S. H. Dynamic reorganization of open chromatin underlies diverse transcriptomes during spermatogenesis. *Nucleic Acids Res.* **46**, 593–608 (2018).
54. Starks, R. R., Biswas, A., Jain, A. & Tuteja, G. Combined analysis of dissimilar promoter accessibility and gene expression profiles identifies tissue-specific genes and actively repressed networks. *Epigenet. Chromatin* **12**, 16 (2019).
55. Gaston, K. & Jayaraman, P. S. Transcriptional repression in eukaryotes: repressors and repression mechanisms. *Cell Mol. Life Sci.* **60**, 721–741 (2003).
56. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).
57. Comai, L. & Harada, J. J. Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. *Proc. Natl Acad. Sci. USA* **87**, 2671–2674 (1990).
58. Zhong, Z. et al. DNA methylation-linked chromatin accessibility affects genomic architecture in Arabidopsis. *Proc. Natl Acad. Sci. USA* <https://doi.org/10.1073/pnas.2023347118> (2021).
59. Borg, M. et al. Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nat. Cell Biol.* <https://doi.org/10.1038/s41556-020-0515-y> (2020).
60. Ingouff, M., Hamamura, Y., Gourgues, M., Higashiyama, T. & Berger, F. Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. *Curr. Biol.* **17**, 1032–1037 (2007).
61. ten Hove, C. A., Lu, K. J. & Weijers, D. Building a plant: cell fate specification in the early Arabidopsis embryo. *Development* **142**, 420–430 (2015).
62. Otero, S., Desvoves, B., Peiro, R. & Gutierrez, C. Histone H3 dynamics reveal domains with distinct proliferation potential in the Arabidopsis root. *Plant Cell* **28**, 1361–1371 (2016).
63. Nie, X., Wang, H., Li, J., Holec, S. & Berger, F. The HIRA complex that deposits the histone H3.3 is conserved in Arabidopsis and facilitates transcriptional dynamics. *Biol. open* **3**, 794–802 (2014).
64. Layat, E. et al. The Histone chaperone HIRA is a positive regulator of seed germination. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms22084031> (2021).
65. Duc, C. et al. Arabidopsis ATRX modulates H3.3 occupancy and fine-tunes gene expression. *Plant Cell* **29**, 1773–1793 (2017).
66. Wang, H. et al. LHP1 Interacts with ATRX through plant-specific domains at specific loci targeted by PRC2. *Mol. Plant* **11**, 1038–1052 (2018).
67. Rajjou, L. et al. Seed germination and vigor. *Annu. Rev. Plant Biol.* **63**, 507–533 (2012).
68. Dure, L. & Waters, L. Long-lived messenger RNA: evidence from cotton seed germination. *Science* **147**, 410–412 (1965).

69. Rajjou, L. et al. The effect of alpha-amanitin on the Arabidopsis seed proteome highlights the distinct roles of stored and neosynthesized mRNAs during germination. *Plant Physiol.* **134**, 1598–1613 (2004).
70. Nakabayashi, K., Okamoto, M., Koshihara, T., Kamiya, Y. & Nambara, E. Genome-wide profiling of stored mRNA in Arabidopsis thaliana seed germination: epigenetic and genetic regulation of transcription in seed. *Plant J.: Cell Mol. Biol.* **41**, 697–709 (2005).
71. Kimura, M. & Nambara, E. Stored and neosynthesized mRNA in Arabidopsis seeds: effects of cycloheximide and controlled deterioration treatment on the resumption of transcription during imbibition. *Plant Mol. Biol.* **73**, 119–129 (2010).
72. Jin, C. et al. H3.3/H2A.Z double variant-containing nucleosomes mark ‘nucleosome-free regions’ of active promoters and other regulatory regions. *Nat. Genet.* **41**, 941–945 (2009).
73. Yang, Y. et al. HIRA complex presets transcriptional potential through coordinating depositions of the histone variants H3.3 and H2A.Z on the poised genes in mESCs. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkab1221> (2021).
74. Shu, H. et al. Arabidopsis replacement histone variant H3.3 occupies promoters of regulated genes. *Genome Biol.* **15**, R62 (2014).
75. Curtis, M. D. & Grossniklaus, U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **133**, 462–469 (2003).
76. Nakagawa, T. et al. Improved gateway binary vectors: high-performance vectors for creation of fusion constructs in Transgenic analysis of plants. *Biosci. Biotech. Bioch.* **71**, 2095–2100 (2007).
77. Chen, S. F., Zhou, Y. Q., Chen, Y. R. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, 884–890 (2018).
78. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907+ (2019).
79. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
80. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
81. 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).
82. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
83. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
84. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–U354 (2012).
85. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
86. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
87. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
88. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
89. Guo, W. et al. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics* **14**, 774 (2013).
90. Guo, W. et al. CGmapTools improves the precision of heterozygous SNV calls and supports allele-specific methylation detection and visualization in bisulfite-sequencing data. *Bioinformatics* **34**, 381–387 (2018).

Acknowledgements

We thank Dr. Hidenori Takeuchi for providing CRISPR-Cas9 generated *htr4;htr5* double mutant. Work in D.J. laboratory was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Precision Seed Design and Breeding, XDA24020303), the National Natural Science Foundation of China (31970527, 32170545, and 32150610472) and the National Key R&D Program of China (2019YFA0903903). Work in F.B. laboratory was supported by FWF grant (P30802).

Author contributions

F.B. and D.J. conceived the research. T.Z., J.L., M.X., J.P., L.M., and D.J. performed experiments. H.Z. performed bioinformatics analysis. F.B. and D.J. wrote the paper with inputs from all co-authors.

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/s41467-022-35509-6>.

Correspondence and requests for materials should be addressed to Danhua Jiang.

Peer review information *Nature Communications* thanks Yuda Fang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are 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 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) 2022